



Харківський національний університет радіоелектроніки
Державне підприємство
"Південний державний
проектно-конструкторський та науково-дослідний
інститут авіаційної промисловості"



Kharkiv National University of Radio Electronics
State Enterprise
"Southern National Design & Research Institute of Aerospace Industries"

Сучасний стан наукових досліджень та технологій в промисловості

Innovative technologies and scientific solutions for industries

Щоквартальний науковий журнал
№ 2 (32), 2025

Затверджений до друку
Науково-технічною Радою
Харківського національного університету радіоелектроніки
(Протокол № 5 від 19 червня 2025 р.)

ЗАСНОВНИКИ

Харківський національний університет радіоелектроніки,
Державне підприємство "Південний державний
проектно-конструкторський та науково-дослідний інститут
авіаційної промисловості"

АДРЕСА РЕДАКЦІЇ:

Україна, 61166, м. Харків, проспект Науки, 14
Інформаційний сайт: <http://itssi-journal.com>
<https://journals.uran.ua/itssi>
E-mail редколегії: journal.itssi@gmail.com

Quarterly scientific journal
No. 2 (32), 2025

Approved for publication
by the Scientific and Technical Council
of the Kharkiv National University of Radio Electronics
(Protocol No. 5 d.d. 18 June 2025)

ESTABLISHERS

Kharkiv National University of Radio Electronics,
State Enterprise
"National Design & Research Institute
of Aerospace Industries"

EDITORIAL OFFICE ADDRESS:

14 Nauky Ave., Kharkiv, Ukraine, 61166
Information site: <http://itssi-journal.com>
<https://journals.uran.ua/itssi>
E-mail of the editorial board: journal.itssi@gmail.com

Ідентифікатор медіа R30-03878
Витяг з реєстру суб'єктів у сфері медіа – реєстрантів від 25.04.2024 № 1410

*Журнал включено до Переліку наукових фахових видань України, в яких можуть публікуватися
результати дисертаційних робіт на здобуття наукових ступенів доктора і кандидата наук
наказом Міністерства освіти і науки України від 16.07.2018 №775 (додаток 7)*

Харків – 2025

© Харківський національний університет радіоелектроніки,
Державне підприємство "Південний державний проектно-конструкторський
та науково-дослідний інститут авіаційної промисловості"

РЕДАКЦІЙНА КОЛЕГІЯ

Головний редактор

Бодянський Євгеній Володимирович,
д-р техн. наук, професор

Заступник головного редактора

Айзенберг Ігор Наумович,
канд. техн. наук, професор (США);
Шекер Серхат, д-р техн. наук, професор (Туреччина)

Члени редколегії:

Артюх Роман Володимирович, канд. техн. наук;
Бабенко Віталіна Олексіївна,
д-р екон. наук, канд. техн. наук, професор;
Безкоровайний Володимир Валентинович,
д-р техн. наук, професор;
Гасімов Юсіф, д-р мат. наук, професор (Азербайджан);
Гопєєнко Віктор, д-р техн. наук, професор (Латвія);
Го Цян, д-р техн. наук, професор (КНР);
Зайцева Єлена, д-р техн. наук, професор (Словаччина);
Зачко Олег Богданович, д-р техн. наук, професор;
Коваленко Андрій Анатолійович, д-р техн. наук, професор;
Косенко Віктор Васильович, д-р техн. наук, професор;
Костін Юрій Дмитрович, д-р екон. наук, професор;
Левашенко Віталій, д-р техн. наук, професор (Словаччина);
Лемешко Олександр Віталійович, д-р техн. наук, професор;
Малєєва Ольга Володимирівна, д-р техн. наук, професор;
Момот Тетяна, д-р екон. наук, професор (США);
Музика Катерина Миколаївна, д-р техн. наук, професор;
Назарова Галина Валентинівна, д-р екон. наук, професор;
Невлюдов Ігор Шакирович, д-р техн. наук, професор;
Опанасюк Анатолій Сергійович,
д-р фіз.-мат. наук, професор;
Павлов Сергій Володимирович, д-р техн. наук, професор;
Паржнн Юрій, д-р техн. наук, професор (США);
Перова Ірина Геннадіївна, д-р техн. наук, професор;
Петленков Едуард, канд. техн. наук (Естонія);
Петришин Любомир, д-р техн. наук, професор (Польща);
Рубан Ігор Вікторович, д-р техн. наук, професор;
Семенець Валерій Васильович, д-р техн. наук, професор;
Семенов Сергій, д-р техн. наук, професор (Польща);
Сетлак Галина, д-р. техн. наук, професор (Польща);
Терзіян Ваган, д-р техн. наук, професор (Фінляндія);
Тєлєтов Олександр Сергійович, д-р екон. наук, професор;
Тімофєєв Володимир Олександрович,
д-р техн. наук, професор;
Філатов Валентин Олександрович, д-р техн. наук, професор;
Чумаченко Ігор Володимирович, д-р техн. наук, професор;
Чухрай Наталія Іванівна, д-р екон. наук, професор;
Юн Джин,
канд. фіз.-мат. наук, професор (КНР);
Ястремська Олена Миколаївна, д-р екон. наук, професор.

EDITORIAL BOARD

Editor in Chief

Bodyanskiy Yevgeniy,
Dr. Sc. (Engineering), Professor, Ukraine

Deputy Chief Editor

Igor Aizenberg,
PhD (Computer Science), Professor (United States)
Serhat Seker, Dr. Sc. (Engineering), Professor (Turkey)

Editorial Board Members:

Artiukh Roman, PhD (Engineering Sciences) (Ukraine);
Babenko Vitalina,
Dr. Sc. (Economics); PhD (Engineering Sciences), Professor (Ukraine);
Bezkorovainyi Volodymyr,
Dr. Sc. (Engineering), Professor (Ukraine);
Gasimov Yusif, Dr. Sc. (Mathematical), Professor (Azerbaijan);
Gopeyenko Vectors, Dr. Sc. (Engineering), Professor (Latvia);
Guo Qiang, Dr. Sc. (Engineering), Professor (P.R. of China);
Zaitseva Elena, Dr. Sc. (Engineering), Professor (Slovak Republic);
Zachko Oleh, Dr. Sc. (Engineering), Professor (Ukraine);
Kovalenko Andrey, Dr. Sc. (Engineering), Professor (Ukraine);
Kosenko Viktor, Dr. Sc. (Engineering), Professor, (Ukraine);
Kostin Yuri, Dr. Sc. (Economics), Professor (Ukraine);
Levashenko Vitaly, Dr. Sc. (Engineering), Professor (Slovakia);
Lemeshko Oleksandr, Dr. Sc. (Engineering), Professor (Ukraine);
Malyeyeva Olga, Dr. Sc. (Engineering), Professor (Ukraine);
Momot Tatiana, Dr. Sc. (Economics), Professor (USA);
Muzyka Kateryna, Dr. Sc. (Engineering), Professor (Ukraine);
Nazarova Galina, Dr. Sc. (Economics), Professor (Ukraine);
Nevliudov Igor, Dr. Sc. (Engineering), Professor (Ukraine);
Opanasyuk Anatoliy,
Dr. Sc. (Physical and Mathematical), Professor (Ukraine);
Pavlov Sergii, Dr. Sc. (Engineering), Professor (Ukraine);
Parzhyn Yurii, Dr. Sc. (Engineering), Professor (USA);
Perova Iryna, Dr. Sc. (Engineering), Professor (Ukraine);
Petlenkov Eduard, PhD (Engineering Sciences) (Estonia);
Petryshyn Lubomyr, Dr. Sc. (Engineering), Professor (Poland);
Ruban Igor, Dr. Sc. (Engineering), Professor, (Ukraine);
Semenets Valery, Dr. Sc. (Engineering), Professor, (Ukraine);
Semenov Serhii, Dr. Sc. (Engineering), Professor (Poland);
Setlak Galina, Dr. Sc. (Engineering), Professor (Poland);
Terziyan Vagan, Dr. Sc. (Engineering), Professor, (Finland);
Teletov Aleksandr, Dr. Sc. (Economics), Professor (Ukraine);
Timofeyev Volodymyr,
Dr. Sc. (Engineering), Professor (Ukraine);
Filatov Valentin, Dr. Sc. (Engineering), Professor (Ukraine);
Chumachenko Igor, Dr. Sc. (Engineering), Professor (Ukraine);
Chukhray Nataliya, Dr. Sc. (Economics), Professor (Ukraine);
Yu Zheng,
PhD (Physico-Mathematical Sciences), Professor (P.R. of China);
Iastremska Olena, Dr. Sc. (Economics), Professor (Ukraine)

ЗМІСТ

Інформаційні технології

- 5 **Барковська О. Ю., Головченко О. С., Сторчай Д. Г., Костін А. О., Лєгезін Н. В.**
Дослідження методів штучного зору в системах внутрішньої навігації (en)
- 16 **Данишина С. Ю., Андрєєв С. М.**
Формування моделі розумного міста з огляду на динаміку змін міської забудови (ua)
- 33 **Косаревський Б. В., Тецький А. Г.**
Сучасні підходи до розгортання інфраструктури мобільних інтелектуальних систем (ua)
- 49 **Мухін В. Є., Хабло Я. О.**
Порівняльний аналіз алгоритмів узгодження модальностей у мультимодальних трансформерах для синтезу звуку (en)
- 58 **Невлудов І. Ш., Євсєєв В. В., Максимова С. С., Артюх Р. В.**
Математична модель адаптивного ієрархічного високорівневого керування триланкового колаборативного робота-маніпулятора (ua)
- 69 **Поддубний В. О., Сєвєрінов О. В.**
Метод автентифікації користувачів в інформаційно-комунікаційних системах із застосуванням нульового водяного знака (ua)
- 79 **Прокопенко О. С., Смяляков С. В.**
Розроблення об'єктно-орієнтованого алгоритму порівняння зображень для їх ефективного пошуку (ua)
- 102 **Сироткін К. В.**
Оцінювання точності та аналіз помилок розбору залежностей для текстів українською мовою (en)
- 111 **Сопов Є. О., Крицький Д. М., Артьомова А. В., Артьомов І. В.**
Методика побудови довірчої маршрутизації в ройових мережах БПЛА на основі аналізу трафіку та виявлення аномалій (ua)
- 129 **Чухран І. Д., Удовенко С. Г., Шергін В. Л., Чала Л. Е.**
Метод доповнення 3D-моделей точкових хмар з використанням графових нейронних мереж (ua)
- 151 **Шкурко В. В., Поляков А. О.**
Організація програмних і нейромережових алгоритмів машинного аналізу текстових повідомлень, поданих природною мовою (ua)

Сучасні технології управління підприємством

- 168 **Бушуєв С. Д., Бушуєва Н. С., Лобок Є. А., Мурованський Г. А.**
Менеджмент інноваційних проєктів на основі застосувань штучного інтелекту у турбулентному оточенні (ua)

Інженерія та промислові технології

- 177 **Бизкровний О. М., Смяляков К. С.**
Моделі класифікатора передумов виникнення ДТП для передбачення небезпечних ситуацій на перехрестях (ua)
- 188 **Зенкін М. А., Іванко А. І., Черниш М. Л.**
Вплив вібрацій різального інструмента на якість поверхні обрізаних листових матеріалів та методи їх мінімізації (ua)
- 199 **Алфавітний покажчик**

CONTENTS

Information Technology

- 5 **Barkovska O., Holovchenko O., Storchai D., Kostin A., Lehezin N.**
Investigation of computer vision techniques for indoor navigation systems (en)
- 16 **Danshyna S., Andrieiev S.**
Formation of a smart city model based on the dynamics of changes in urban building (ua)
- 33 **Kosarevskiy B., Tetskiy A.**
Modern approaches to deploying the infrastructure of mobile intelligent systems (ua)
- 49 **Mukhin V., Khablo Ya.**
Comparative analysis of modality alignment algorithms
in multimodal transformers for sound synthesis (en)
- 58 **Nevliudov I., Yevsieiev V., Maksymova S., Artiukh R..**
Mathematical model of adaptive hierarchical high-level control
of a three-link collaborative robot-manipulator (ua)
- 69 **Poddubnyi V., Sievierinov O.**
Method of user authentication in information and communication systems
using zero watermark (ua)
- 79 **Prokopenko O., Smelyakov S.**
Development of an object-oriented image comparison algorithm for efficient search (ua)
- 102 **Syrotkin K.**
Accuracy evaluation and error analysis of dependency parsing of texts in Ukrainian (en)
- 111 **Sopov I., Krytskiy D., Artomova A., Artomov I.**
Trust-based routing methodology in uav swarm networks
based on traffic analysis and anomaly detection (ua)
- 129 **Chukhran I., Udovenko S., Shergin V., Chala L.**
Method for augmenting 3D point cloud models using graph neural networks (ua)
- 151 **Shkurko V., Poliako, A.**
Organization of software and neural network algorithms
for machine analysis of textual data presented in natural language (ua)

Modern Enterprise Management Technology

- 168 **Bushuyev S., Bushuyeva N., Lobok Y., Murovansky H.**
Management of innovative projects based on artificial intelligence applications
in a turbulent environment (ua)

Engineering & Industry Technologies

- 177 **Byzkrovnyi O., Smelyakov K.**
Models of the classifier of prerequisites for the occurrence
of road accidents to predict dangerous situations at intersections (ua)
- 188 **Zenkin M., Ivanko A., Chernysh M.**
Influence of cutting tool vibrations on the surface quality
of cut sheet materials and methods for their minimization (ua)
- 199 **Alphabetical index**

O. BARKOVSKA, O. HOLOVCHENKO, D. STORCHAI, A. KOSTIN, N. LEHEZIN

INVESTIGATION OF COMPUTER VISION TECHNIQUES FOR INDOOR NAVIGATION SYSTEMS

The subject of this article is the development and implementation of computer vision methods that can be integrated into an indoor navigation system designed for individuals with visual impairments. **The goal** of the study is to enhance such a system with advanced object recognition capabilities in enclosed environments by combining modern technologies, including artificial intelligence, spatial analysis, voice control, and Bluetooth-based localization. To achieve this, a number of **tasks** were carried out. These included an analysis of the problem domain and justification of the study's relevance, a comparison of existing solutions, and the development of a generalized model of the navigation system with a voice interface, enabling real-time search for locations and items. A specialized dataset was prepared, containing images of key obstacle classes typically encountered in indoor environments – such as shopping carts, barrier tape, forklifts, and people. A new two-stage object recognition method was proposed to detect these classes in complex scenes. Additionally, a comparative analysis of deep learning architectures for object detection was conducted, followed by experimental studies to assess training quality and system robustness. The research employed various image preprocessing **methods** – bilateral filtering, Gaussian blurring, enhancement of specific color channels, motion blur removal, and noise reduction using averaging filters – as well as neural network-based methods for data analysis and statistical evaluation approaches. The **results** demonstrate that the proposed method significantly improves object detection performance on real-world images, achieving an average intersection-over-union (IoU) of 68% and a confidence level of 69%, which is 79% and 89% higher, respectively, compared to baseline recognition results on noisy inputs. However, the findings also revealed the necessity of integrating additional sensors, such as LiDAR, to reliably detect low-contrast or reflective obstacles like glass storefronts, which are difficult to identify using computer vision alone. **Conclusions.** The study confirms that the proposed two-stage preprocessing, and recognition pipeline significantly enhances navigation system performance for users with visual impairments, while also highlighting the importance of combining vision-based methods with complementary sensing technologies to ensure safe and reliable operation in complex indoor environments.

Keywords: system; localization; navigation; blindness; recognition; computer vision; classification.

Introduction

Vision impairments lead to a lower quality of life, especially in large cities with complex navigation environments. They also reduce independence, autonomy, and overall safety. Therefore, the task of creating robotic assistants based on artificial intelligence or developing technological solutions capable of ensuring comfort, safety, mobility, and independence remains highly relevant [1–3].

In Ukraine, there is a growing number of people with visual impairments. According to the National Health Service of Ukraine (NHSU), in 2021, 17,478 people were diagnosed with blindness or vision impairment. In 2022, this number increased to 19,551 people. Moreover, in the first seven months of 2023, the number of registered cases has already reached 19,336 patients, which is close to the total number for the previous year.

These data and the growing number of people with visual impairments highlight the social need for and importance of developing indoor and outdoor navigation systems. Such systems are important for the accessibility

and independence of people with visual impairments. The implementation of these navigation systems is in line with the objectives of the European Disability Strategy, ensuring social integration and equal access to infrastructure, public services, and employment opportunities for people with disabilities.

Analysis of last achievements and publications

Computer vision (CV) is a field of artificial intelligence that teaches computers to interpret and understand the visual world. Using digital images from cameras and videos, as well as deep learning models, computers can accurately identify and classify objects and then respond to what they "see". At the heart of CV is image processing, which involves improving image data (removing noise, sharpening or brightening the image) and preparing it for further analysis.

Computer vision is actively being implemented in various industries, improving accuracy and speed of work, as well as automating complex processes. These examples confirm that computer vision (CV) already plays a key role in the future of technology (Fig. 1) [4–8].

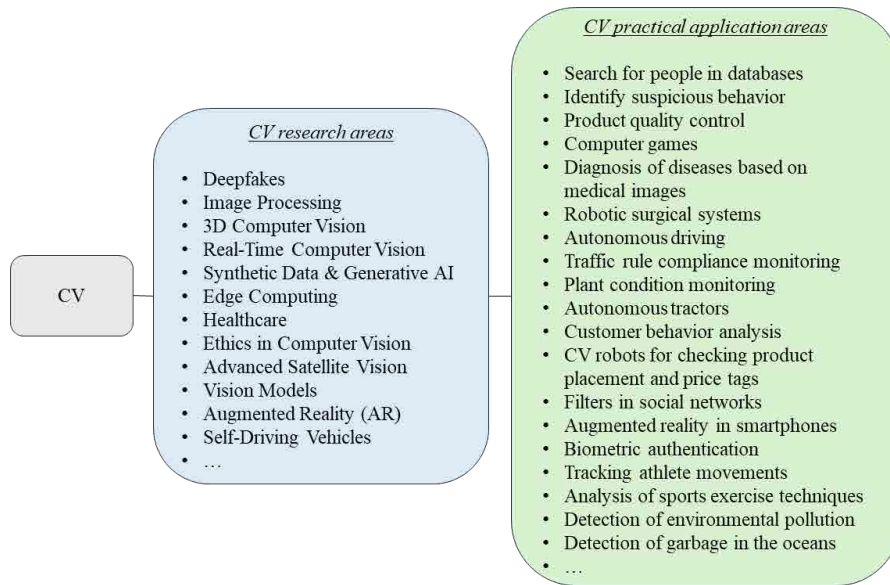


Fig. 1. Some of computer vision research and practical application areas overview

Detecting obstacles for a blind person's shopping cart in a supermarket can be considered a special case of real-time computer vision (real-time CV) and autonomous navigation. The system can use cameras, LiDAR, depth sensors, and modern CV models to analyze the path and warn the user about obstacles using voice commands or tactile feedback.

Modern deep learning architectures for computer vision tasks can be divided into several categories:

- convolutional neural networks (CNNs);
- transformers;

- hybrid models;
- generative networks.

Below is a comparison table of modern deep learning architectures used for various computer vision tasks (Table 1). The table presents the key characteristics of the models, including the year of development, the number of parameters, accuracy on known datasets (ImageNet, COCO), image output speed, and real-world application areas. The architectures are grouped by type: convolutional neural networks (CNNs), transformers, hybrid models, and generative networks.

Table 1. A comparative table highlighting key characteristics of various deep learning architecture classes used in computer vision

Model Name	Year	Parameters	Top-1 Accuracy (ImageNet)	Top-1 Accuracy (COCO)	Inference Speed (ms/image)	Real-world Applications	Datasets Used
1	2	3	4	5	6	7	8
Architecture Type: Convolutional Neural Networks							
ResNet-50	2015	25.6M	79%	–	4.6	Image classification, Feature extraction	ImageNet [9]
DenseNet-121	2016	8.0M	74.9%	–	6.9	Medical imaging, Anomaly detection	ImageNet [10]
EfficientNet-B0	2019	5.3M	77.1%	–	3.2	Mobile vision applications	ImageNet [11]
Faster R-CNN	2015	41.0M	–	42.0%	100+	Object detection, Autonomous driving, Medical imaging	COCO, Pascal VOC [12]
Architecture Type: Transformers							
Vision Transformer (ViT-B/16)	2020	86.4M	77.9%	–	7.9	Large-scale image classification	ImageNet [13]
Swin Transformer (Swin-B)	2021	88.0M	83.5%	53.5%	8.7	Object detection, Segmentation	ImageNet [14]
DETR	2020	41.5M	–	44.0%	50-60	Object detection, Autonomous systems, Retail AI	COCO, ADE20K [15]

Continuation of the table 1

1	2	3	4	5	6	7	8
Architecture Type: Hybrid Models							
ConvNeXt-B	2022	88.6M	83.8%	–	8.0	General-purpose vision tasks	ImageNet [16]
YOLOv8	2023	~25.0M	–	51.5%	5–10	Object detection, Surveillance, Robotics	COCO, OpenImages [17]
Architecture Type: Generative Networks							
BigGAN	2018	112.0M	–	–	–	Image synthesis, DeepFakes	ImageNet [18]

Analysis of the table showed that for the task of recognizing visual obstacles in the path of a shopping cart between aisles, it is important to focus on the following aspects:

- Speed of logical inference. YOLOv8 demonstrates the highest image processing speed (5–10 ms per image). This is critical for a task where obstacles must be detected in real time so that the cart can quickly respond to changes in the environment.

- Training on COCO and OpenImages. The YOLOv8 model was trained on COCO and OpenImages, which include many classes of objects found in real life (e.g., bottles, boxes, shopping carts). This allows the model to adapt to typical obstacles in a supermarket;

- Versatility. YOLOv8 supports detection, classification, and segmentation, allowing you to extend the functionality of the system. For example, you can use YOLOv8 not only for obstacle avoidance but also for environment analysis (product identification or navigation).

- Compact model (parameters). YOLOv8 has a relatively small number of parameters (~25 million) compared to other detectors (e.g., Faster R-CNN – 41 million). This makes YOLOv8 more suitable for use in devices with limited computing resources, such as carts with integrated cameras.

- Real-world use cases. YOLOv8 is already being used in video surveillance and robotics, where real-time object detection is also required. These examples confirm that YOLOv8 is well suited for dynamic environments such as supermarket traffic.

With its high social relevance, the project focuses on achieving scientific and engineering goals.

The aim of this research is to develop and implement artificial vision methods focused on indoor obstacle recognition in a proposed indoor navigation system for visually impaired people. The system integrates modern technologies such as artificial intelligence, voice control, spatial analysis, and Bluetooth-based navigation.

The system aims to promote social integration and increase the independence of visually impaired people

when navigating complex indoor environments such as shopping malls, warehouses, or public facilities.

The implementation of the project includes the following tasks:

- analysis of the problem area and justification of the relevance of the research topic;

- development of a model of an indoor navigation system with a voice interface for real-time search of places and objects by users with visual impairments;

- preparation of a working dataset containing images of specific classes – shopping carts, barrier tape, people, forklifts;

- creation of an improved method for object detection in complex scenes;

- conducting experimental studies to evaluate the quality of training and identify limitations and shortcomings in the use of visual detection tools.

Materials and methods

Based on the analysis of the subject area, the key components of the proposed system are as follows:

- voice message interpretation technology;
- interactive maps of premises with spatial localization of key objects;

- a system for positioning moving objects within a room;

- a subsystem for determining the route from a moving object to a destination point, provided in the form of a voice query.

A functional model of the system for planning a route to a specific place or product is shown in Figure 2. The diagram is used to determine the requirements for the system and to clarify the functions necessary for further development of the system.

To start working with the proposed route planning system for a specific location or product, it is necessary to obtain input data. The input data includes the user's local coordinates and voice command, the analysis of which determines the desired location or product for which a route needs to be built within the premises (indoor localization).

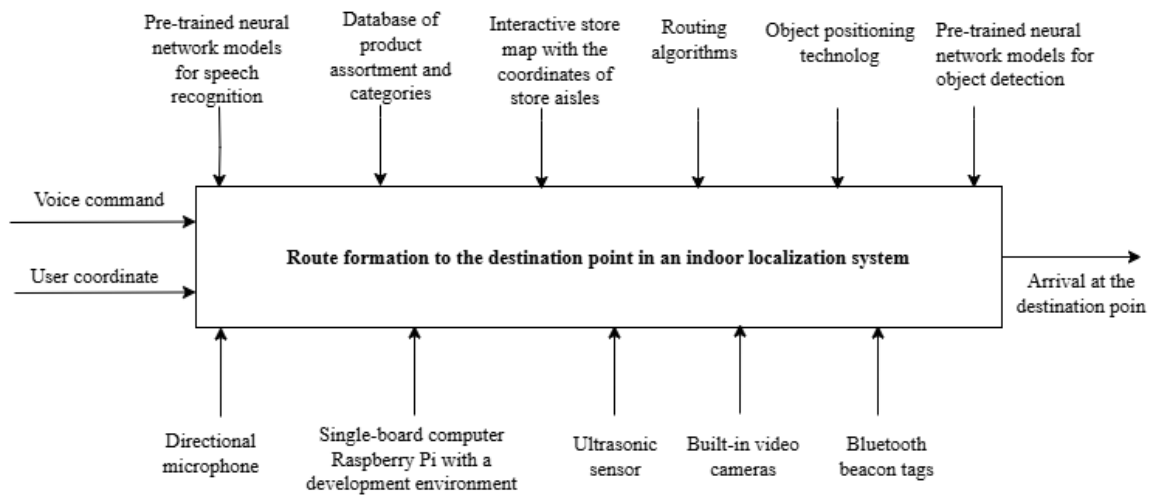


Fig.2. Generalized Functional Model of the System for Route Planning to a Specific Location or Product

As a result of decomposing the generalized functional model of the route planning system for a specific location or product, five structural blocks were identified: interpretation of a voice request into a command, determination of the destination coordinates, route construction, analysis of obstacles in the user's path, and reactive route updating (Figure 3).

To interpret a voice request into a command, the system uses the mathematical apparatus of artificial neural networks (ANN) and methods of preliminary processing of voice information, including normalization, reverberation reduction, compression, spectrogram equalization, and frequency filtering (using time and frequency masks).

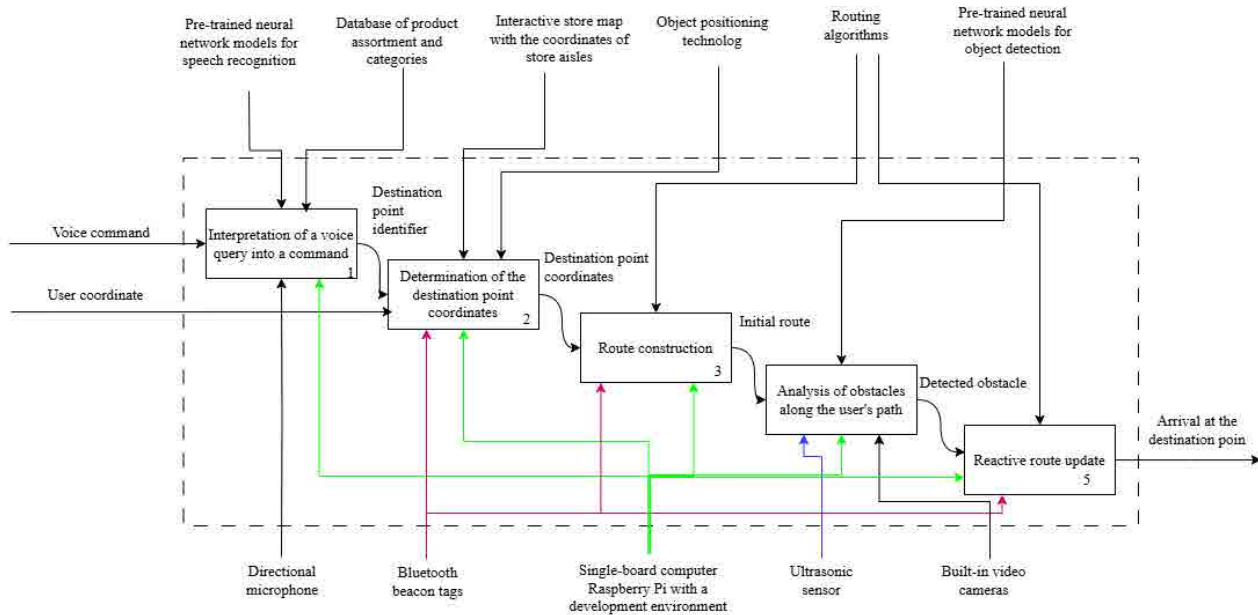


Fig.3. Decomposition of the Process into Interconnected and Interdependent Functional Blocks

The use of directional microphones allows the system to focus on sounds from a specific direction, ignoring noise from other sources. This approach significantly improves the signal-to-noise ratio (SNR), which is critical for accurate speech recognition.

To determine the location of the requested item, the system analyzes the product database, where all entries

are considered a training dictionary for the neural network during recognition.

The coordinates of the store aisles are stored on an interactive map of the premises, which links information about specific aisles (e.g., product assortment) to their physical location. This data is synchronized with positioning technology, allowing the system to accurately locate objects in space.

Object positioning in the system is implemented using Bluetooth Low Energy (BLE) beacons, which transmit signals with unique identifiers at fixed intervals (Figure 4). Compared to other positioning technologies (UWB, WiFi, RFID, optical systems, etc.), BLE technology offers several key advantages:

- high accuracy (up to 1-2 meters depending on the number of beacons and environmental conditions);
- energy efficiency;
- ease of deployment;
- scalability.

These characteristics make BLE a suitable choice for large, complex indoor environments such as shopping malls or warehouses.

Obstacles in the user's path are detected by analyzing the surrounding space. This information is processed by a Raspberry Pi single-board computer, which analyzes the input data and detects obstacles in the route. The system provides real-time obstacle detection, which is important for user safety and navigation accuracy.

Routing algorithms build the optimal and safest path to the requested product, taking into account the detected obstacles.

The system uses the user's current coordinates, and the coordinates of the destination point as input data.

The resulting route is presented to the user in the form of a series of voice instructions for smooth control.

The first stage of implementing the proposed system is to develop a model of a local navigation system with a voice interface, which serves as the initial step necessary for implementing the proposed system. This stage includes the following tasks:

- creation of a digital map of selected premises with the location of key objects;
- selection of positioning technology (Wi-Fi, Bluetooth beacons, UWB, ultrasound, etc.);
- development of a wireless personal area network (WPAN) model based on Bluetooth Low Energy (BLE) technology;
- integration of the BLE network into the RTLS system infrastructure (with the ability to add new objects and assign unique identifiers) to create an effective positioning and tracking mechanism in real time;
- integration of an algorithm for determining the current location of the user using an interactive map of the premises to display the current position of the user.

The second stage involves the development of a routing subsystem based on interactive floor plans for indoor navigation, which is another key step in the implementation of the proposed system.

This stage includes the following tasks:

- review of route planning algorithms, including triangulation algorithms, Dijkstra's algorithm, and SLAM;
- implementation of a method for detecting obstacles/moving objects;
- implementation of a context-dependent routing algorithm in the RTLS system, taking into account detected obstacles;
- development of methods for reactive route updating.

The most research-intensive stage necessary for the implementation of the proposed system is the study and development of methods for interpreting voice queries in an indoor navigation system. This stage is the most science-intensive, since the accuracy of voice query interpretation directly affects the accuracy of the generated route. This stage includes the following tasks:

- systematization of knowledge about hardware for voice capture in noisy environments;
- research of noise suppression methods and reduction of reverberation effects to minimize the impact of reflected sound on voice command recognition;
- evaluating the impact of using directional microphones or microphone arrays to improve voice command recognition accuracy by isolating the speaker's voice;
- developing methods for analyzing voice commands with adaptability to lexical ambiguity.

Further research in this work focuses on analyzing computer vision methods, which are the basis of the user interference analysis module. The YOLO8nano neural network model was trained on two prepared datasets. The key characteristics of the first dataset are as follows:

- the total volume is 300 images, divided into training (80%, 240 images) and test (20%, 60 images) subsets;
- includes four classes – faces (highlighted on a portion of images of other classes) and carts, forklifts, and barrier tape (100 images for each class). Image size varies from 141x149 to 2010x2010 pixels. The images were normalized and reduced to 640x640 for accurate neural network training.

The images were collected from random sources, so the quality varies:

- lighting – from bright to dim, both daylight and artificial;
- noise – JPEG compression artifacts, blur, some images have a "grainy" texture;
- viewing angle – images show objects from different angles and perspectives. However, preference was given to images where the viewing point is no lower

than one meter above the ground and no higher than human eye level.

The second dataset was a test dataset, and the model was trained on it separately, as training on this dataset could distort the weights of other classes. The key characteristics of the dataset are the same as for the first one, with the following exceptions:

- total volume is 100 images;
- it includes two classes: people (not labeled separately; pre-trained weights were left) and shop windows/glass walls (100 images (80 for the training set and 20 for the test set). The image size varies from 339x295 to 715x474.

Research results

This study focuses on the ability of a pre-trained neural network to accurately recognize four main classes of obstacles. The fifth class (glass walls/showcases) is a test class and is placed in a separate network, as it is predicted that the selected neural network is not capable of learning to recognize poorly visible objects, which

could ultimately reduce the detection accuracy of the other classes.

Since images are often accompanied by noise in real conditions, a two-stage detection algorithm is proposed, the essence of which is as follows:

- the input image is fed to the input of the neural network;
- after detecting an object in the image, if the confidence that this object belongs to a certain class is less than 85%, then preprocessing methods are applied to this image, depending on which class the neural network assigned the object to for the first time. After that, the image is fed back into the neural network;
- the neural network moves on to the next image or completes its work.

Figure 4 shows a simplified algorithm for analyzing obstacles in the path of the module user.

Table 2 shows the dependence of the IoU and Confidence metrics on the use of preprocessing methods on images with significant noise (from 21 to 55%). For simplicity, Table 3 shows only one experiment for each class.

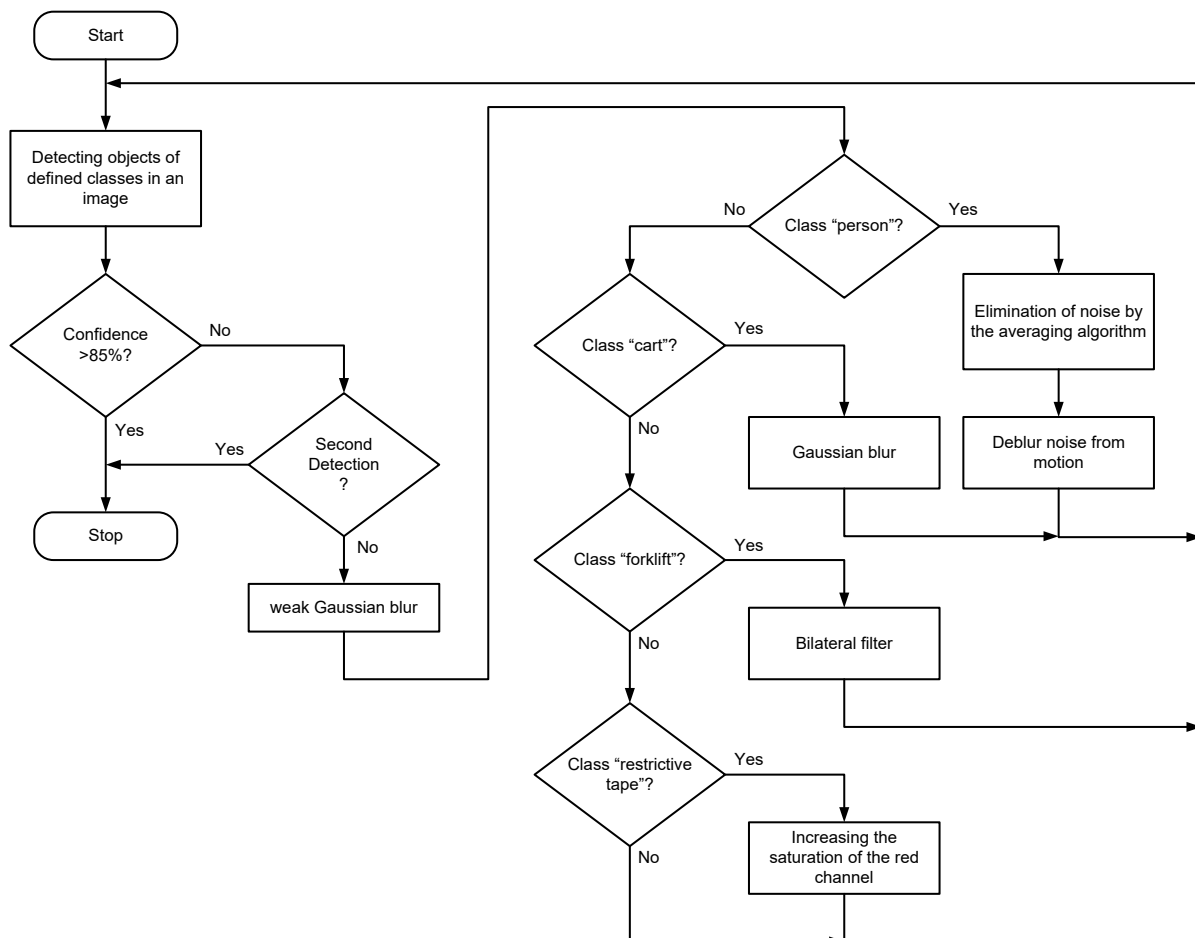


Fig. 4. Simplified algorithm of the analysis of obstacles along the user's path module

Table 2. Results of experimental studies of the impact of using preprocessing methods on IoU and Confidence indicators for noisy input images

Cart recognition demonstration	Forklift recognition demonstration	People recognition demonstration	Restrictive tape recognition demonstration
Input noisy image			
			
Reference marked-up image			
			
Detected object in an image without preprocess-sing			
			
Average IoU and confidence for an image without preprocessing, %			
IoU = 65 conf. = 32	IoU = 57 conf. = 46	IoU = 36 conf. = 55	IoU = 25 conf. = 51
Detected object in the image after preprocessing			
			
Average IoU and confidence for pre-processed images, %			
IoU = 69 conf.= 81,5	IoU = 67 conf. = 72	IoU = 70 conf. = 61	IoU = 65 conf. = 63
Average increase in IoU and confidence when using preprocessing, %			
IoU = 15 conf. = 154	IoU = 47 conf. = 56	IoU = 94 conf. = 110	IoU = 160 conf. = 24

Analysis of the results presented in Table 3 shows that the use of image preprocessing methods improves the ability of the neural network (NN) to recognize objects.

On average, the IoU (intersection over union) metric can be increased by 79% and the Confidence metric by 86%. For better visualization, a summary graph of the averaged

IoU and Confidence metrics is provided for images without additional noise, noisy images, and images that have undergone preprocessing.

Table 3 shows the results of determining the training quality of NN YOLO8nano on a selected training dataset using averaged IoU (intersection over union) and confidence metrics using high-quality input images (noise level up to 20%).

The graph in Figure 5 clearly shows a trend toward improvement in the average IoU and Confidence metrics for images with noise when using preprocessing methods. In some places, when using preprocessing methods, the selected metrics exceed the IoU and Confidence for images without additional noise.

Table 3. Results of experimental studies of IoU and Confidence metrics for high-quality input images

	Cart	Forklift	People	Restrictive tape
Average class confidence, %	91	64	59	65
Average class IoU, %	80	58	62	71

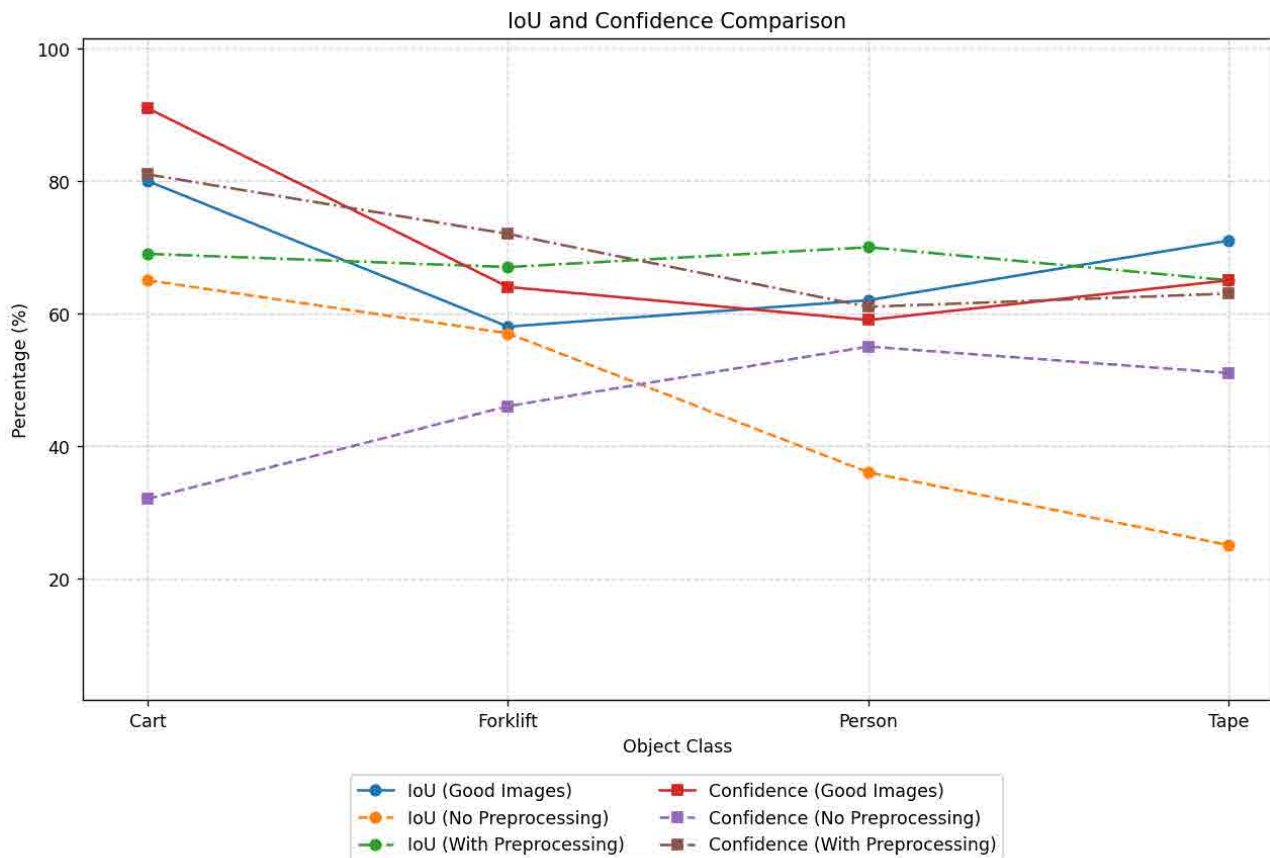


Fig. 5. Summary graph of averaged metrics

The last experiment analyzed the capabilities of computer vision methods for detecting poorly visible obstacles, such as glass windows/walls. Table 4 shows a selection of four experiments that reflect general trends in the object recognition results of a neural network trained on the second dataset. Table 4 shows that more

than half of the images of shop windows/glass walls are free of these obstacles. Thus, the use of photo and video sensors is insufficient to detect this class of obstacles. Therefore, it is advisable to equip the system with ultrasonic sensors or lidars to confirm the creation of such an object based on a significant distance to it.

Table 4. Evaluation of the ability of the selected neural network to detect faint objects after training on the selected dataset

input images	reference marked-up image	image with detected object
shopfront1.JPG (confidence=0.85)		
		
shopfront2.JPG (confidence=0.32)		
		
shopfront3.JPG (confidence=0.84)		
		
shopfront4.JPG (confidence=0.88)		
		

Conclusions

The relevance of the research is due to the growing number of people with visual impairments and the need to develop technologies that ensure their independence and safety in urban environments.

The practical significance lies in the creation of a technology that can be used in shopping centers, warehouse complexes, and public buildings to ensure the convenience and safety of people with visual impairments.

The scientific novelty lies in the development of an integrated system that combines computer vision for obstacle detection and their implementation in an adaptive routing subsystem, a voice interface for user interaction, and BLE navigation for indoor positioning.

The research aims to improve object recognition methods in commercial premises.

The results showed that the proposed two-stage recognition method based on YOLOv8 increases the confidence in detecting objects in the user's path (people, carts, forklifts, barrier tapes) by 89%, reaching an average confidence of 69%, while IoU increased by 79%, reaching an average IoU = 68%. A two-stage preprocessing algorithm is proposed, according to which the sequence of preprocessing methods for input images depends on a predefined class. It was also found that for some classes of obstacles (e.g., mirrors or glass showcases), the system requires additional tools to process additional data obtained from sensors that compensate for the limitations of photo and video cameras.

Prospects for further research. Future research aims to improve the system by adding specialized sensors for detecting hard-to-recognize objects and developing context-dependent routing methods in the proposed system.

References

1. Khan, S., Nazir, S., & Khan, H. U. (2021), "Analysis of navigation assistants for blind and visually impaired people: A systematic review". *IEEE access* 9 (2021), P. 26712–26734. DOI:10.1109/ACCESS.2021.3052415
2. Барковська, О., Сердечний, В. (2024), "Intelligent assistance system for people with visual impairments". *Innovative technologies and scientific solutions for industries*, (2 (28)), P. 6–16. DOI:10.30837/2522-9818.2024.28.006
3. Ashmafee, M. H., & Sabab, S. A. (2016), "Blind Reader: An intelligent assistant for blind". In 2016 19th *International Conference on Computer and Information Technology*. DOI: 10.1109/ICCITECHN.2016.7860200
4. Wu, M., Li, C., & Yao, Z. (2022), "Deep active learning for computer vision tasks: methodologies, applications, and challenges". *Applied Sciences*, 12(16), 8103 p. DOI: <https://doi.org/10.3390/app12168103>
5. Paneru, S., Jeelani, I. (2021), "Computer vision applications in construction: Current state, opportunities & challenges". *Automation in Construction*, 132, 103940 p. DOI: 10.1016/j.autcon.2021.103940
6. Elyan, E., Vuttipittayamongkol, P., Johnston, P., Martin, K., McPherson, K., Moreno-García, C. F., Sarker, M. M. K. (2022), "Computer vision and machine learning for medical image analysis: recent advances, challenges, and way forward". *Artificial Intelligence Surgery*, 2(1), P. 24–45. DOI: 10.20517/ais.2021.15
7. Naik, B. T., Hashmi, M. F., Bokde, N. D. (2022), "A comprehensive review of computer vision in sports: Open issues, future trends and research directions". *Applied Sciences*, 12(9), 4429 p. DOI: <https://doi.org/10.3390/app12094429>
8. Zablocki, É., Ben-Younes, H., Pérez, P., & Cord, M. (2022), "Explainability of deep vision-based autonomous driving systems: Review and challenges". *International Journal of Computer Vision*, 130(10), P. 2425–2452. DOI: <https://doi.org/10.1007/s11263-022-01657-x>
9. He, K., Zhang, X., Ren, S., & Sun, J. (2016), "Deep residual learning for image recognition". In *Proceedings of the IEEE conference on computer vision and pattern recognition*. P. 770–778. DOI: 10.1109/cvpr.2016.90
10. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K. Q. (2017), "Densely connected convolutional networks". In *Proceedings of the IEEE conference on computer vision and pattern recognition*. P. 4700–4708. DOI: 10.1109/cvpr.2017.243
11. Tan, M., & Le, Q. (2019), "Efficientnet: Rethinking model scaling for convolutional neural networks". In *International conference on machine learning*. P. 6105–6114. DOI: <https://doi.org/10.48550/arXiv.1905.11946>
12. Ren, S., He, K., Girshick, R., & Sun, J. (2016), "Faster R-CNN: Towards real-time object detection with region proposal networks". *IEEE transactions on pattern analysis and machine intelligence*, 39(6), P. 1137–1149. DOI:10.1109/tpami.2016.2577031
13. Alexey, D. (2020), "An image is worth 16x16 words: Transformers for image recognition at scale". *Computer Vision and Pattern Recognition*.
14. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Guo, B. (2021), "Swin transformer: Hierarchical vision transformer using shifted windows". In *Proceedings of the IEEE/CVF international conference on computer vision*. P. 10012–10022. DOI: <https://doi.org/10.1109/ICCV48922.2021.00986>
15. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020), "End-to-end object detection with transformers". In *European conference on computer vision*. Cham: Springer International Publishing. P. 213–229. DOI: https://doi.org/10.1007/978-3-030-58452-8_13
16. Liu, Z., Mao, H., Wu, C. Y., Feichtenhofer, C., Darrell, T., Xie, S. (2022), "A convnet for the 2020s". In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. P. 11976–11986. DOI: 10.1109/CVPR52688.2022.01167
17. Redmon, J. (2016), "You only look once: Unified, real-time object detection". In *Proceedings of the IEEE conference on computer vision and pattern recognition*. DOI:10.1109/CVPR.2016.91
18. Brock, A. (2018), "Large Scale GAN Training for High Fidelity Natural Image Synthesis", DOI:10.48550/arXiv.1809.11096

Надійшла (Received) 13.04.2025

Відомості про авторів / About the Authors

Barkovska Olesia – Ph.D (Engineering Sciences), Associate Professor, Kharkiv National University of Radio Electronics, Associate Professor of the Department of Electronic Computers, Kharkiv, Ukraine; e-mail: olesia.barkovska@nure.ua; ORCID ID: <https://orcid.org/0000-0001-7496-4353>

Holovchenko Oleksandr – Kharkiv National University of Radio Electronics, Master's student of Department of Electronic Computers, Kharkiv, Ukraine; e-mail: oleksandr.holovchenko@nure.ua; ORCID ID: <https://orcid.org/0009-0002-7582-1746>

Storchai Denis – Kharkiv National University of Radio Electronics, Master's student of the Department of Electronic Computers, Kharkiv, Ukraine; e-mail: denys.storchai@nure.ua; ORCID ID: <https://orcid.org/0009-0006-0527-0095>

Kostin Anton – Kharkiv National University of Radio Electronics, Master's student of Department of Electronic Computers, Kharkiv, Ukraine; e-mail: anton.kostin@nure.ua; ORCID ID: <https://orcid.org/0009-0000-1517-0035>

Lehezin Nikita – Kharkiv National University of Radio Electronics, Master's student of Department of Electronic Computers, Kharkiv, Ukraine; e-mail: nikita.lehezin@nure.ua; ORCID ID: <https://orcid.org/0009-0005-6349-0603>

Барковська Олеся Юрївна – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри Електронних обчислювальних машин, Харків, Україна.

Головченко Олександр Сергійович – Харківський національний університет радіоелектроніки, магістрант кафедри Електронних обчислювальних машин, Харків, Україна.

Сторчай Денис Геннадійович – Харківський національний університет радіоелектроніки, магістрант кафедри Електронних обчислювальних машин, Харків, Україна.

Костін Антон Олексійович – Харківський національний університет радіоелектроніки, магістрант кафедри Електронних обчислювальних машин, Харків, Україна.

Легезін Нікіта Вікторович – Харківський національний університет радіоелектроніки, магістрант кафедри Електронних обчислювальних машин, Харків, Україна.

ДОСЛІДЖЕННЯ МЕТОДІВ ШТУЧНОГО ЗОРУ В СИСТЕМАХ ВНУТРІШНЬОЇ НАВІГАЦІЇ

Предметом статті є методи комп'ютерного зору, які можуть бути імплементовані у систему навігації в приміщенні для людей з вадами зору. **Метою цього дослідження** є розробка та впровадження методів штучного зору, орієнтованих на розпізнавання перешкод у закритому просторі, у пропоновану систему внутрішньої навігації для людей з вадами зору, яка інтегрує сучасні технології штучного інтелекту, голосового керування, просторового аналізу та Bluetooth-навігації. Для досягнення мети були вирішені такі **завдання**: виконано аналіз проблемної області, включаючи обґрунтування актуальності теми, порівняння існуючих рішень; запропоновано узагальнену модель системи з описом передбачених модулів; запропоновано новий метод розпізнавання основних класів перешкод, які можуть зустрітися в торгових залах (люди, візки, навантажувачі, обмежувальна стрічка) за допомогою вдосконаленого методу двоетапного розпізнавання об'єктів; проведено порівняльний аналіз архітектур глибокого навчання для задач розпізнавання об'єктів; виконано експериментальні дослідження для оцінки якості навчання. Використані **методи**: попередньої обробки зображення (білатеральна фільтрація, гаусівське розміття, підвищення насиченості певного каналу зображення, видалення розміття рухом, видалення шумів алгоритмом усереднення), неймережеві методи аналізу вхідних даних, методи статистичних досліджень. **Результат**: запропонований метод показав хороші результати на реальних тестових зображеннях. (досягнуто: IoU = 68% і достовірність = 69% в середньому, що в середньому на 79% і 89% більше, ніж вихідні зображення з шумом). Була виявлена необхідність розширення системи додатковими інструментами (наприклад, лідачами) для виявлення важко помітних перешкод типу дзеркальних вітрин. **Висновки**. Спираючись на проведений аналіз, запропонований двохетапний метод препроцесингу значно покращує якість розпізнавання. Пропонованій системі потрібні додаткові датчики, оскільки не всі об'єкти можуть бути розпізнані за допомогою методів комп'ютерного зору.

Ключові слова: система; локалізація; навігація; сліпота; розпізнавання; комп'ютерний зір; класифікація.

Бібліографічні описи / Bibliographic descriptions

Барковська О. Ю., Головченко О. С., Сторчай Д. Г., Костін А. О., Легезін Н. В. Дослідження методів штучного зору в системах внутрішньої навігації. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 2 (32). С. 5–15. DOI: <https://doi.org/10.30837/2522-9818.2025.2.005>

Barkovska, O., Holovchenko, O., Storchai, D., Kostin, A., Lehezin, N. (2025), "Investigation of computer vision techniques for indoor navigation systems", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (32), P. 5–15. DOI: <https://doi.org/10.30837/2522-9818.2025.2.005>

С. ДАНШИНА, С. АНДРЕЄВ

ФОРМУВАННЯ МОДЕЛІ РОЗУМНОГО МІСТА З ОГЛЯДУ НА ДИНАМІКУ ЗМІН МІСЬКОЇ ЗАБУДОВИ

Мета роботи – підвищення ґрунтовності оцінок для формування рішень з інтелектуалізації міста на основі аналізу змін міської забудови. У статті розв’язано такі **завдання**: розглянуто тенденції розвитку концепцій *Smart City*; проаналізовано вимоги до розумних міст і запропоновано теоретико-множинну модель розумного міста; подальшого розвитку набув метод аналізу міської забудови за дистанційними даними, експериментально підтверджено можливість його використання. Упроваджено **методи** системного аналізу, теорії множин і глибокого навчання. **Досягнуті результати**. Створення зручного міського середовища сприяє економічному зростанню, покращує якість життя та забезпечує сталий розвиток, що можливо завдяки впровадженню концепції розумного міста. Для уніфікації кроків зі створення розумних міст запропоновано теоретико-множинну модель розумного міста, яка відповідно до вимог стандартів серії ISO 37100 поєднує показники, що забезпечують реалізацію певних напрямів інтелектуалізації. Вибір напрямку інтелектуалізації має ґрунтуватися на вивченні проблем урбанізації. Складність цього питання потребує вдосконалення процесів аналізу міської забудови, наприклад, з використанням даних дистанційного зондування Землі (ДЗЗ). Для отримання оцінок, придатних для подальшого аналізу та формування рішень, розроблено метод аналізу міської забудови за даними ДЗЗ, що використовує глибоке навчання в обробленні структурованої інформації за допомогою інструменту *Image Analyst ArcGIS Pro 3.4*. Розроблено схему алгоритму отримання моделі *Deep Learning* для дешифрування об’єктів міської забудови, запропоновано її програмну реалізацію. Дослідження можливостей розробленого методу відбувалося в процесі розв’язання завдань актуалізації інформації про міську забудову Києва. Модель *Deep Learning* добре розпізнає об’єкти міської забудови та має добру пристосованість до об’єктів, що різняться від об’єктів навчальної вибірки. **Висновки**. Вибір напрямку інтелектуалізації має ґрунтуватися на ретельному вивченні ситуації, що виникла внаслідок урбанізації міських районів. Експериментальний аналіз динаміки змін міської забудови району м. Києва за період 2005–2021 рр. показав: збільшення кількості об’єктів житлової (на 5,4%) та комерційної забудови (на 11%), нехтування розвитком соціальної інфраструктури, нерозбудову транспортної інфраструктури, ігнорування питань озеленення території району тощо. Досягнуті результати можуть бути корисними у формуванні пріоритетних способів упровадження інтелектуальних рішень у повсякденне життя містян.

Ключові слова: теоретико-множинна модель; дистанційне зондування; глибоке навчання; актуалізація даних; напрямки інтелектуалізації міста.

Вступ

Феномен урбанізації впливає на спосіб життя людей, стимулює їхню міграцію до міських районів із сільської місцевості. Ці зміни активізують комерційний розвиток, кооперацію та економічне зростання. Водночас швидке збільшення міст загострює наявні проблеми та призводить до нових, створюючи значні складності науково-технічної та соціально-економічної спрямованості, вимагаючи від влади впровадження нових способів забезпечення безпеки громадян і покращення якості їхнього життя [1]. У цих умовах упровадження концепції *Smart City* (розумного міста) суттєво впливає на міський розвиток і планування: під тиском збільшення кількості населення за наявних обмежених ресурсів її спрямовано на підвищення стійкості міст, розвиток якісних громадських послуг і задоволення потреб мешканців [2, 3].

Розумне місто – це концепція, що використовує інформаційно-комунікаційні технології з метою підвищення ефективності функціонування та використання міських ресурсів відповідно до потреб його мешканців у забезпеченні належної якості життя в умовах прагнення до міської стійкості. Моделі *Smart City* розробляють для виконання найактуальніших завдань, що постають перед міськими районами: покращення інфраструктури та боротьба із заторами, забезпечення мобільності мешканців і підвищення адекватності громадських послуг, розвиток системи поводження з відходами та протидія засміченню територій, керування енергоспоживанням і розв’язання енергетичних питань тощо [3, 4].

Результативність подібних моделей та ефективність подолання зазначених проблем у міському середовищі щорічно оцінює *World Competitiveness Center* – Центр світової

конкурентоспроможності (<https://www.imd.org/smart-city-observatory/home/>), який формує рейтинг міст за *Smart City Index*, зосереджуючи увагу на економічних, технологічних аспектах і "гуманних вимірах", реалізованих у моделях (якість життя, навколишнє середовище та інклюзивність), узгоджуючи їх зі стратегіями керування містами всього світу. Цей індекс поєднує результати аналізу статистичних показників щодо рівня охорони здоров'я, безпеки та мобільності, перспектив розвитку й зростання міст із результатами опитування мешканців, які оцінюють упровадження розумних технологій у місті. У підсумку отримують щорічний рейтинг, який станом на 2024 р. містив відомості щодо 142 найбільш технологічних міст усього світу [5]. Тут лідерами визнано столиці країн Північної Європи та Австралії, також рейтинг *Smart City* очолили азіатські мегаполіси.

Поєднання *Smart City Index* з *Global Competitiveness Index* – рейтингом країн світу за показником економічної конкурентоспроможності,

що традиційно визначає здатність країн світу забезпечувати стабільні темпи економічного зростання (за показниками Всесвітнього економічного форуму) [6], демонструє, що майже 22% найбільш технологічних міст світу розташовані в країнах з найбільшою конкурентоспроможністю національних економік (серед країн з населенням понад 20 млн). У цьому разі сильна додатна кореляція (з коефіцієнтом кореляції $r \approx 0,87$) між світовим рейтингом конкурентоспроможності країни та кількістю міст у країні з найбільшою інтеграцією в повсякденне життя високих технологій і смарт-рішень (рис. 1) свідчить про те, що сучасну конкурентоспроможність економік країни не можна оцінювати лише за економічними показниками (ВВП або продуктивністю). Політичні, соціальні та культурні показники набувають суттєвого значення, і, отже, уряди країн мають створювати зручне міське середовище з ефективною інфраструктурою, соціальними інститутами, екологією та інклюзивною політикою, які зможуть забезпечити стійкий розвиток міст і країн загалом.

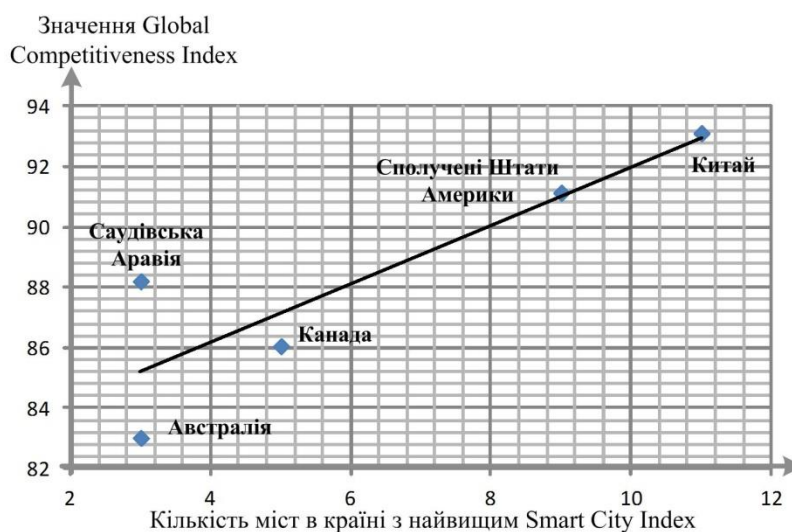


Рис. 1. Діаграма розсіювання, що демонструє статистичний взаємозв'язок між світовим рейтингом конкурентоспроможності країни та кількістю міст у країні з найбільшою інтеграцією в повсякденне життя громадян високих технологій і смарт-рішень

Дослідники зазначають, що 90% міст у всьому світі не мають офіційної стратегії розумного розвитку, віддаючи перевагу впровадженню корпоративних технологій без осмислення соціальних цілей, яким мають слугувати моделі *Smart City*. Як наслідок, питання розумної урбанізації стають найбільш актуальною темою наукової спільноти [5, 7, 8].

Аналіз проблеми та наявних методів

Стійкість міст невід'ємно пов'язана з цілями Стратегії сталого розвитку. Урбанізація змінює навколишнє середовище, значно покращує умови життя громадян, забезпечуючи незалежність та економічне зростання. Однак вона ж стає причиною багатьох проблем (крім погіршення стану довкілля, збільшується кількість викидів парникових газів,

забруднюються ґрунти та вода, зростають нерівність і міська бідність тощо), які внаслідок прискорення урбанізації, ймовірно, не лише погіршать якість життя мешканців, а й спричинять значні складнощі для майбутньої стійкості міст. Крім цього, сучасна практика міського розвитку й управління не спроможна загальмувати ці загрозові наслідки нестійкої урбанізації: розширення багатьох міст спрямовано на досягнення короткострокової, а не тривалої вигоди, що унеможливорює підвищення якості життя міського населення в довгостроковій перспективі [9, 10].

Стійкий розвиток залежить від балансу економічного, соціального й екологічного складників. Міста – це складні системи, на які впливають зростання населення, транспортна інфраструктура, ринкова поведінка, земельна політика уряду, екологічні процеси. Це призводить до нелінійного розвитку міст: зазначені фактори взаємопов'язані, а причинно-наслідкові зв'язки стають засобом інтеграції економічних, соціальних і екологічних підсистем. Коли кінцевою метою є соціальна стійкість, економічна підсистема має забезпечувати необхідну підтримку, а екологічна – стати її основою. Крім цього, всі три підсистеми мають розвиватися одночасно. Саме розуміння окреслених проблем сприяло стрімкому розвитку технологій розумного міста, пік якого спостерігався з 2015 до 2021 рр., коли кількість міст, що ефективно впроваджували в муніципальну систему управління ці технології, збільшилася майже вдвічі – з 70 до 138 [5, 8, 9]. Але вже зараз спостерігається сповільнення темпів поширення моделей *Smart City*, хоча, за оцінками

фахівців, нині та в найближчому майбутньому альтернативи їм немає: їх визнано результативним інструментом покращення життя, розв'язання проблем з ресурсами, оптимізації інфраструктури для забезпечення стійкого та якісного проживання в умовах урбанізації [7–9]. Ці обставини вимагають вивчення суперечливого характеру розвитку розумних міст, їх всебічного дослідження та систематизації властивостей [8, 10].

Бібліографічний пошук за ключовою фразою "*Smart City*" у базі даних *Sciencedirect* (<https://www.sciencedirect.com>) показує, що близько сотні нових публікацій за 2024 р. (рис. 2) спрямовано на дослідження:

- взаємозв'язків між впровадженням моделей розумного міста й ефективністю досягнення цілей сталого розвитку (наприклад, роботи [8, 11], всього 23% від проаналізованих досліджень);
- можливостей *Smart City* та їх подальшої інтеграції до систем управління міським середовищем (наприклад, роботи [12, 13], всього 21% від проаналізованих досліджень);
- наявного досвіду реалізацій моделей розумного міста, його узагальнення, типізацію та тиражування (наприклад, роботи [3, 7], всього 16% від проаналізованих досліджень);
- питань залучення мешканців і громад у проекти розумного міста та поширення ідей *Smart City* серед населення (наприклад, роботи [1, 14], всього 15% від проаналізованих досліджень);
- проблем конфіденційності та безпеки даних у смарт-рішеннях (наприклад, роботи [2, 10], всього 16% від проаналізованих досліджень) тощо.



Рис. 2. Розподіл публікацій 2024 р. за проблемними питаннями *Smart City*

Зауважимо, що значна кількість розглянутих публікацій безпосередньо є результатом узагальнення досвіду впровадження моделей *Smart City*, містить тематичний огляд літератури, поєднаний з опитуванням експертів і зацікавлених сторін (наприклад, роботи [3, 7, 8 та ін.]), що доводить ґрунтовність отриманих висновків і пильну увагу до цього питання.

Визначення не розв'язаних раніше частин загальної проблеми

Отже, досвід упровадження моделей *Smart City* показує поки локальний вплив на умови життя містян [4, 7], їх спрямовано на розв'язання окремих проблем [15], вони зазвичай не зважають на динаміку зміни міського середовища [9], суттєво залежать від прийнятої практики керування містом і політичної культури управління не лише містом, але й країною [13], потребують соціальної підтримки, роз'яснення та залучення громадськості [16], а також постійного застосування передових інноваційних технологій і методів [11, 17]. Також у значній кількості моделей не береться до уваги просторовий складник міського середовища, структура розселення, доступна інфраструктура або наявний потенціал міських ресурсів [3, 10]. Усе це призводить до ускладнень, коли порівнюють моделі *Smart City* (особливо їх технічний складник), а також коли оцінюють ці моделі, визначають рівень інтелектуальності та ефективності функціонування. А отже, виникає необхідність дослідження наявної нормативної бази для формування критеріїв ефективності розумного міста, визначення його основних принципів і ознак для забезпечення доцільності функціонування.

Міське середовище – це розподілена сукупність інфраструктур, зокрема транспортної (дороги, мости, тунелі, місця паркування, пішохідні зони, велосипедна інфраструктура), водної (водопостачання, водовідведення, санітарії та зливової каналізації), енергетичної (системи постачання електроенергії, газу, централізованого теплопостачання, громадського освітлення), інфраструктура для збирання та перероблення відходів; громадські будівлі, які використовують для спорту, мистецтва, культури, освіти, охорони здоров'я, адміністрації тощо [10, 18].

Зазначимо, що найпоширенішою концепцією реалізації ідей *Smart City* є *Internet of Things* (IoT), що підтримує мережу передачі даних між фізичними

об'єктами ("речами"), оснащеними вбудованими засобами взаємодії один з одним і зовнішнім середовищем, що забезпечують збір інформації, необхідної для належного управління міськими активами, ресурсами та послугами [1, 8]. Але така мережа не дає змогу системно аналізувати міські проблеми, комплексно досліджувати об'єкти міської інфраструктури та розробляти можливі способи її вдосконалення та інтелектуалізації. У зв'язку з цим географічні інформаційні системи відкривають нові перспективи, пов'язані зі збором даних про об'єкти, їх ідентифікацією, перевіркою на відповідність законодавству тощо [18]. Розуміння динаміки змін міського середовища є важливим джерелом інформації для прийняття управлінських рішень. Але тут потенціал геоінформаційних систем (ГІС) не використовують у повному обсязі: їх не розглядають як ефективний засіб, що допомагає приймати ґрунтовні рішення з інтелектуалізації міст з огляду на вимоги до містобудування, землеустрою, охорони довкілля та ін. Це потребує розроблення методів, оснований на використанні сучасних досягнень ГІС для підтримки прийняття рішень щодо інтелектуального розвитку міст, які можуть ґрунтуватися на актуальній цифровій моделі міста з географічною прив'язкою його об'єктів за допомогою застосування дистанційних показників зондування Землі (ДЗЗ) [18].

Предмет дослідження – ГІС-методи аналізу міської забудови для формування рішень з інтелектуалізації міста.

Метою роботи є підвищення ґрунтовності оцінок для формування рішень з інтелектуалізації міста на основі аналізу змін міської забудови.

Завдання дослідження:

- вивчити тенденції розвитку концепцій *Smart City*;
- проаналізувати сучасні вимоги до розумних міст і запропонувати модель розумного міста;
- розробити метод аналізу змін міської забудови за дистанційними даними та експериментально підтвердити можливість його впровадження у визначенні напрямів інтелектуалізації міста..

Теоретико-множинна модель розумного міста

Процес трансформації міст у розумні суттєво різняться по країнах і навіть по різних містах однієї країни, що пов'язано з різними наборами контекстних

факторів (показників, що визначають середовище, в якому живуть місцеві мешканці), що беруть до уваги в моделях *Smart City* [7]. За цих обставин виникають питання, що вважати розумним містом та як оцінити ефективність його функціонування, адже наявні локальні показники найчастіше не стандартизовані, не узгоджені або не піддаються порівняльному аналізу за часом.

Аналізуючи серію міжнародних стандартів ISO 37100, спрямовану на уніфікацію кроків щодо створення стратегій з досягнення стійкості міст і громад, зазначимо, що значну частку тут становлять документи, які містять набір індикаторів для вимірювання ефективності сфер діяльності з метою отримання кращих практик з усього світу й знаходження інноваційних рішень міських проблем. Зокрема ISO 37122 "Сталі міста та громади – індикатори розумних міст" [19] містить групи показників, сформованих відповідно до цілей Сталого розвитку, а також методи їх застосування для оцінювання динаміки руху до розумного міста економічної, екологічної та соціальної сфер його діяльності. Цей стандарт використовує статистичні показники, що надходять від відповідних служб міської адміністрації та компаній, які досліджують ринок відповідних послуг.

Нехай модель розумного міста повністю визначається сукупністю показників, сформованих у ISO 37122. У теоретико-множинному вигляді цю модель подамо як

$$SC = \bigcup_{i \in I} X_i, \quad (1)$$

де X_i – множина показників розумного міста i -ї групи, визначеної ISO 37122;

$I = 1, \dots, 19$ – максимальна кількість груп показників розумного міста.

У цьому разі справедливим є таке:

– групи показників можуть поєднувати різну кількість чинників (табл. 1), необхідних для формування комплексних властивостей розумного міста, тобто

$$\exists i, j \in I, \text{ якщо } i \neq j \mid |X_i| \neq |X_j|; \quad (2)$$

– кількість показників групи може бути розширена для дотримання більш комплексного

та прикладного підходу до аналізу сталості та інтелектуальності [19], але з дотриманням умови

$$X_i \cap X_j = \emptyset, \text{ якщо } i \neq j \text{ у разі } i, j \in I; \quad (3)$$

– відповідно до стандарту ISO 37122 максимальна кількість показників, що беруть до уваги в моделі (1), сягає 80, тобто $|SC|_{\max} = 80$, але висновок щодо "розумності" міста формують за результатами аналізу їх мінімально необхідної кількості:

$$|SC|_{\min} \leq |SC|_{\max} / 2. \quad (4)$$

Зазначимо, що частина показників не залежить від упровадження інноваційних технологій, а є, ймовірно, наслідком їх поширення та популяризації. Так, наприклад, коефіцієнт виживання нових компаній в місті здебільшого залежить від сприятливого економічного клімату в країні, ніж від "розумних" ініціатив міського керівництва; частка міського населення, що професійно володіє більш ніж однією мовою спілкування, є наслідком впровадження нових стандартів освіти або поширенням нових вимог ринку праці, а частка активних користувачів публічних бібліотек пов'язана із загальними традиціями освіти, притаманними для певної країни, та є результатом залучення до неї традиційних методів навчання тощо.

Виділимо в ISO 37122 сукупність показників моделі розумного міста, які, на думку авторів, безпосередньо визначають ефективність упровадження в повсякденне життя містян високих технологій і смарт-рішень. Для зручності подамо їх у вигляді таблиці (табл. 1), де вказано назву показників, збережено їх поділ на групи за ISO 37122, а також використано позначки множин відповідно до виразу (1).

Аналізуючи табл. 1, зазначимо, що для ефективного впровадження концепції *Smart City* та розвитку міста необхідно ретельно досліджувати наявну міську інфраструктуру [20, 21], адже розвиток технологій розумних міст спрямовано на використання передових технологій та інновацій. Це неможливо без сучасної доступної інфраструктури, яка охоплює всі райони міста. Вимоги до її стану для розумного міста більш жорсткі, ніж для звичайних міст і регіонів [20]. А отже, для їх забезпечення необхідно мати актуальну інформацію про міську забудову й тенденції її розвитку, адже щільність забудови постійно збільшується внаслідок спорудження нових об'єктів.

Таблиця 1. Показники моделі розумного міста, що визначають ефективність упровадження інтелектуальних рішень у повсякденне життя містян (джерело: сформовано авторами на основі [19])

Відповідно до ISO 37122			Напря́м інтелектуаліза́ції
познака множини, її назва (назва групи), кількість показників у групі	показник, значення якого залежить від упровадження інноваційних технологій		
	познака	назва	
1	2	3	4
X ₁ – економіка, 4	x ₁ ¹	частка контрактів на надання міських послуг, основаних на політиці відкритих даних	управління цільовою інформацією, що гарантує відкрите та підзвітне керування
	x ₁ ³	частка мешканців, зайнятих у сфері інформаційно-комунікаційних технологій (ІКТ)	упровадження інновацій у промислову сферу міста та сферу послуг
X ₂ – освіта, 3	x ₂ ²	кількість персональних комп’ютерів, ноутбуків, планшетів та інших навчальних пристроїв на 1 000 учнів	розвиток комп’ютерної грамотності для досягнення особистих, професійних, соціальних та освітніх цілей
X ₃ – енергетика, 10	x ₃ ⁶	частка вуличного освітлення, що регулюється системою управління робочими характеристиками	підвищення енергоефективності міста
	x ₃ ⁹	частка міських будівель, обладнаних інтелектуальними датчиками електро– споживання	підвищення енергоефективності міста
X ₄ – довкілля, 3	x ₄ ²	кількість інтерактивних дистанційних станцій атмосферного моніторингу на 1 км2	отримання оперативної інформації щодо стану міського середовища
	x ₄ ³	частка міських будівель, обладнаних системами моніторингу повітря в приміщеннях	підвищення комфорту користувачів будівель, покращення їхнього здоров’я
X ₅ – фінанси, 2	x ₅ ²	кількість платежів, що надходять місту за допомогою електронних систем оплати	збільшення росту продуктивності, зменшення витрат міста, бізнесу, мешканців міста
X ₆ – керівництво, 4	x ₆ ¹	щорічна кількість online-відвідувань муніципального порталу відкритих даних	поширення доступу мешканців до інформації від керівництва міста
	x ₆ ²	частка міських послуг, що доступні в online-режимі	поширення взаємодії керівництва міста з мешканцями та бізнесом
	x ₆ ³	середній термін відповіді на неекстрені запити (у днях)	поширення взаємодії керівництва міста з мешканцями та бізнесом
	x ₆ ⁴	середня тривалість збоїв у роботі інформаційно-технічної структури міста	підвищення ефективності міської сфери послуг та інфраструктури міста
X ₇ – охорона здоров’я, 3	x ₇ ¹	частка мешканців, які мають уніфікований online-профіль стану здоров’я	підвищення ефективності постачальників медичних послуг
	x ₇ ²	щорічна кількість дистанційних консультацій медичних фахівців	підвищення доступності первинної медичної допомоги
	x ₇ ³	частка мешканців, що мають доступ до інтерактивних систем екстреного оповіщення про стан довкілля	мінімізація шкідливого впливу забруднень на здоров’я населення
X ₈ – житло, 2	x ₈ ¹	частка домогосподарств, обладнаних інтелектуальними датчиками енергоспоживання	підвищення енергоефективності міста
	x ₈ ²	частка домогосподарств, обладнаних інтелектуальними лічильниками енергоспоживання	підвищення енергоефективності міста
X ₉ – соціальні умови, 4	x ₉ ⁴	частка міського бюджету, яка спрямовується на програми з подолання цифрового бар’єра	розвиток комп’ютерної грамотності серед населення міста
X ₁₀ – відпочинок, 1	x ₁₀ ¹	частка послуг з організації відпочинку, що надаються online	підвищення доступності послуг для мешканців міста
X ₁₁ – безпека, 1	x ₁₁ ¹	частка території міста, що покривається камерами online-спостереження	підвищення безпеки міста

Продовження таблиці 1

1	2	3	4
X_{12} – тверді відходи, 6	X_{12}^1	частка пунктів збору сміття, обладнаних системами телеметричного контролю	оптимізація та підвищення ефективності системи збирання сміття
	X_{12}^2	частка мешканців, які мають доступ до системи збирання відходів з індивідуальним обліком обсягів	оптимізація процесу збирання сміття та трафіку спеціалізованих транспортних засобів
	X_{12}^5	частка муніципальних контейнерів для сміття, обладнаних сенсорними вимірювальними датчиками	удосконалення системи моніторингу та збирання відходів, оптимізація маршрутів і графіків вивезення сміття
X_{13} – спорт та культура, 6	X_{13}^1	кількість online-бронювання відвідування об'єктів культури	підвищення доступності установ культури та спорту
	X_{13}^3	кількість найменувань електронних книг і книг з фондів міських публічних бібліотек	підвищення рівня культури й освіти мешканців міста, спрощення доступу до книг
X_{14} – телекомунікації, 3	X_{14}^2	частка міської території, яка розташована в "проблемних" зонах покриття телекомунікаційних мереж	покращення доступу до базових послуг міста
	X_{14}^3	частка міської території, яка розташована в зонах покриття муніципальних інтернет-мереж загального користування	покращення процесу спілкування, розв'язання завдань перспективного планування
X_{15} – транспорт, 14	X_{15}^1	частка вулиць міста, під'єднаних до online-системи попередження про ситуації на дорогах у режимі реального часу	підвищення ефективності планування розвитку транспортної інфраструктури міста
	X_{15}^5	частка ліній громадського транспорту, обладнаних загальнодоступними інформаційними системами реального часу	підвищення ефективності функціонування транспортної системи міста
	X_{15}^6	частка послуг громадського транспорту, за які сплачують за допомогою єдиної системи платежів	підвищення ефективності функціонування транспортної системи міста
	X_{15}^7	частка громадських паркувань, що оплачуються за допомогою системи електронних платежів	створення системи розумного ціноутворення, підвищення ефективності оплати міських послуг
	X_{15}^8	частка громадських паркувань, обладнаних інтерактивною системою індикації вільних місць	підвищення ефективності міських послуг, покращення стану довкілля
	X_{15}^9	частка інтелектуальних світлофорних постів	покращення стану довкілля, оптимізація маршрутів екстрених служб, підтримання мобільності
	X_{15}^{10}	відсоток території міста, охопленої інтерактивними дорожніми мапами	підтримання мобільності, оптимізація маршрутів
	X_{15}^{12}	частка маршрутів громадського транспорту, де надають послугу під'єднання до муніципальної мережі Інтернет	покращення процесу спілкування, розв'язання завдань перспективного планування
	X_{15}^{13}	частка міських доріг, які підтримують систему автономного керування	покращення транспортної інфраструктури й комунікаційної мережі міста
X_{16} – продовольча безпека, 3	X_{16}^3	частка міської території, охопленої інтерактивною системою керування постачання продуктів харчування	підвищення доступності продовольчих ресурсів міста, оптимізація запасів продуктів харчування
X_{17} – містобудування, 4	X_{17}^1	частка мешканців, яких щорічно долучають до процесів планування	підвищення ефективності планування та розроблення містобудівної політики
	X_{17}^2	частка дозволів на будівництво, отриманих за допомогою електронних систем розгляду документів	підвищення рентабельності будівництва, зменшення корупційного складника
X_{18} – стічні води, 5	X_{18}^5	частка трубопроводів каналізаційної мережі, які контролюють в реальному часі системою датчиків спостереження	підвищення ефективності каналізаційної мережі, впровадження системи моніторингу стоків
X_{19} – вода, 4	X_{19}^2	кількість станцій моніторингу якості природної води, що працюють в режимі реального часу	підвищення ефективності моніторингу стану природної води
	X_{19}^3	частка міської системи водопостачання, яку контролюють інтелектуальною системою моніторингу	підвищення ефективності керування системою водопостачання міста
	X_{19}^4	частка міських будівель, обладнаних інтелектуальними лічильниками води	підвищення ефективності системи водопостачання міста

Аналіз змін міської забудови

Зазвичай розвиток міської забудови відповідно до генерального плану регламентується національними нормативними документами, наприклад Державними будівельними нормами, які визначають організаційно-планувальну структуру міста й мають забезпечувати населення місцем для мешкання, роботи та відпочинку. На жаль, усе частіше розвиток мегаполісів України пов'язують з корупційними скандалами, він має широке коло й масштаб містобудівних проблем стратегічного генерального планування, що суттєво зменшує можливість інтелектуалізації населених пунктів. Збільшення кількості об'єктів капітального будівництва є лише метою отримання швидкого комерційного зиску, ігнорування інновацій під час створення комфортного середовища за одночасного обмеження засобів контролю в просторовому охопленні призводять до перевантажених автошляхів, шумового забруднення, погіршення екологічного стану міста тощо. Усе це потребує поєднання класичних і нових способів керування забудовою міських територій, регулювання містобудівного процесу для забезпечення комфортних умов життя та нормального функціонування інфраструктури [21]. У цьому контексті актуальним стає вдосконалення аналізу міської забудови, що склалася внаслідок урбанізації районів міст і формування ґрунтовних напрямів для смарт-рішень. Використання даних ДЗЗ дає змогу розв'язати частину проблем, пов'язаних із аналізом наявної міської забудови. Крім того, застосування цих даних має низку переваг, зокрема [18, 20, 22]:

- охоплює велику територію, що особливо актуально для дослідження забудови мегаполісів;
- немає необхідності в залученні до процесу аналізу міської забудови значної кількості виконавців;
- не створює перешкод у використанні міської інфраструктури під час аналізу;
- методи отримання даних, їх кількість та якість можуть бути визначені для конкретних завдань з огляду на особливі вимоги до збирання та подальшого аналізу наборів даних тощо.

Водночас зауважимо, що дані ДЗЗ мають значний обсяг, їх якість залежить від сукупності

факторів, які не завжди можуть бути ідеальними, а їх оброблення займає тривалий час [20, 22]. Тому виникає необхідність у створенні методу, що максимально спростить процедуру оброблення даних ДЗЗ і надасть оцінки, придатні для подальшого аналізу та формування рішень.

Нині перспективним варіантом розв'язання складних просторових завдань стає машинне навчання, зокрема глибоке навчання, або *Deep Learning*. Саме тому останнім часом ці інструменти стають основними компонентами просторового аналізу в геоінформаційних системах [23].

Спираючись на типовий алгоритм машинного навчання в процесі оброблення структурованої інформації [24], сформуємо метод аналізу міської забудови за даними ДЗЗ. Він передбачає кілька етапів.

Етап 1. Завантаження супутникового знімку міської забудови для створення навчальної вибірки.

Етап 2. Формування навчальної вибірки з типових просторових об'єктів міської забудови (рис. 3, а):

- виділення полігонів типових об'єктів;
- вилучення просторових фрагментів зображення з типовими об'єктами;
- збереження одержаних фрагментів у базу даних.

На цьому етапі створюють навчальні вибірки, обираючи з просторових даних міської забудови репрезентативні зразки об'єктів інтересу, які точно відтворюють просторові та спектральні властивості забудови, її орієнтацію, розмір і стан. У процесі виконання цієї операції необхідно пам'ятати, що чим повніше виділені об'єкти відбивають об'єкт інтересу, то кращим буде навчання моделі й точнішими будуть класифікація та виявлення логічних зв'язків на зображенні [24].

Етап 3. Отримання моделі *Deep Learning* (рис. 3, а):

- циклічний пошук ознак у зображеннях навчальної вибірки;
- класифікація ознак за типовими об'єктами зображення;
- збереження моделі *Deep Learning*.

Схема алгоритму реалізації другого та третього етапів методу дослідження міської забудови зображена на рис. 3, а. Тут передбачено, що операції третього етапу виконують за допомогою інструменту *Image Analyst ArcGIS Pro 3.4*.

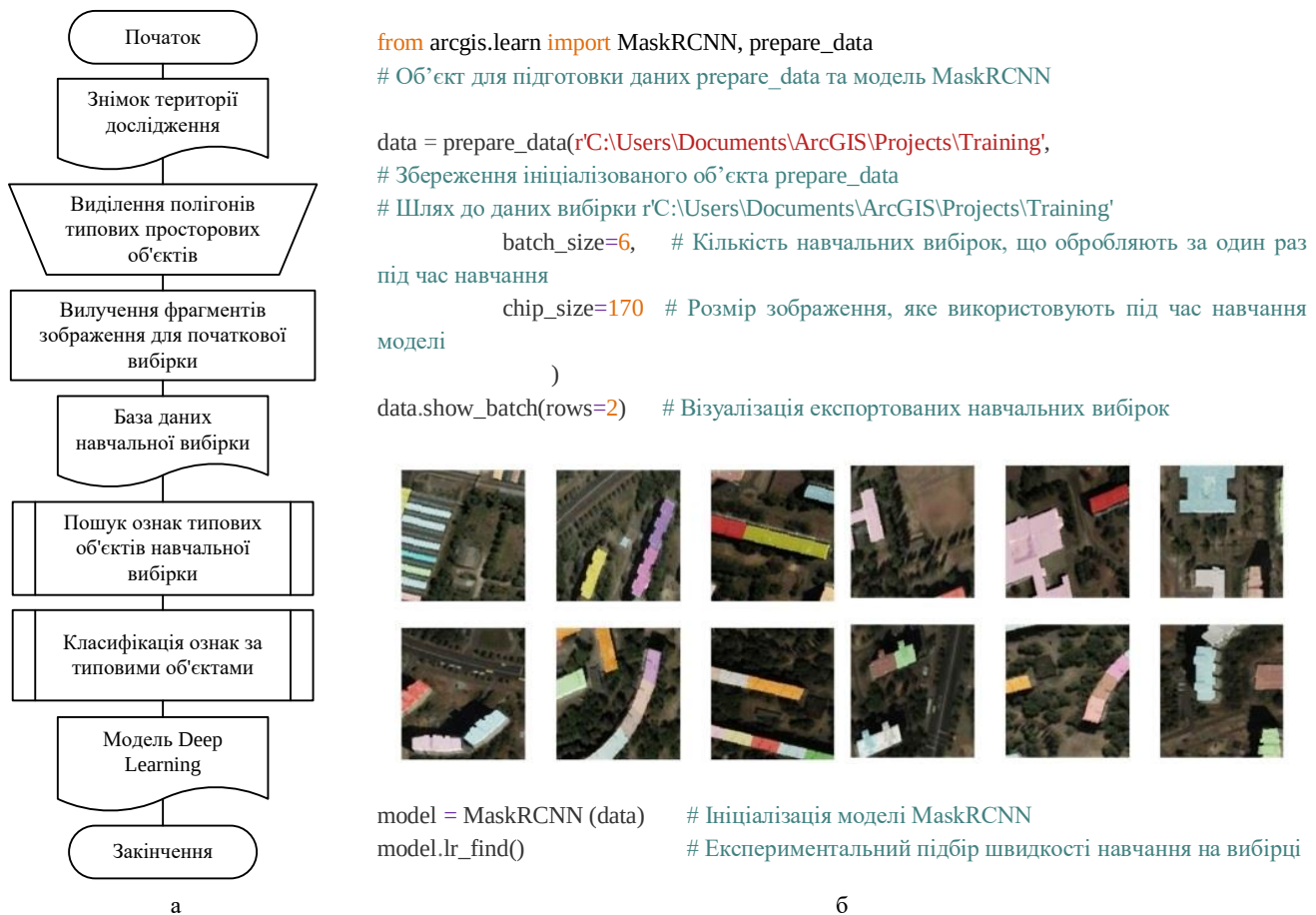


Рис. 3. Програмно-алгоритмічна реалізація методу дослідження міської забудови за супутниковими знімками:

- а – схема алгоритму отримання моделі *Deep Learning* для дешифрування об'єктів міської забудови;
б – лістинг програми циклічного аналізу ознак об'єктів навчальної вибірки

В *Image Analyst* модель *Deep Learning* формується як результат навчання на зображеннях навчальної вибірки, отриманої на другому етапі.

Стандартний уніфікований код для навчання моделі *Deep Learning*, реалізований на *Python*, зазвичай поєднує дані такого типу [24]:

```

TrainDeepLearningModel(in_folder, out_folder, {max_epochs}, {model_type}, {batch_size},
{arguments}, {learning_rate}, {backbone_model}, {pretrained_model}, {validation_percentage},
{stop_training}, {freeze}, {augmentation}, {augmentation_parameters}, {chip_size}, {resize_to},
{weight_init_scheme}, {monitor}),
  
```

(5)

де *in_folder* та *out_folder* – шляхи до місць, де зберігають елементи навчальної вибірки та куди після навчання завантажують навчену модель;

max_epochs – максимальна кількість періодів навчання моделі;

model_type та *arguments* – тип моделі, яку використовують для навчання, та її аргументи (за необхідністю);

batch_size – кількість навчальних вибірок, які обробляють за один крок навчання;

learning_rate – темп навчання;

backbone_model – тип нейронної мережі, що застосовують як архітектуру для навчання нової моделі;

validation_percentage – відсоток навчальних вибірок, що використовують для перевірки моделі;

stop_training та *monitor* – ознака ранньої зупинки та метрика для перевірки контрольних точок у разі її застосування;

augmentation та *augmentation_parameters* – тип збільшення навчальних даних (за необхідністю) та його значення для кожного перетворення в параметрі збільшення;

chip_size – розмір зображення, що використовують для навчання моделі (за необхідністю);

weight_init_scheme – ознака схеми, за якою визначають вагу для шару нейронної мережі (у разі застосування в навчальній вибірці мультиспектральних, а не RGB-даних).

Зазначимо, що зміст виразу (5) може суттєво розрізнятися залежно від типу моделі, обраної для навчання.

На рис. 3, б наведено лістинг програми циклічного аналізу ознак об'єктів вибірки під час навчання моделі з використанням моделі *Mask R-CNN*, яка найбільш пристосована для сегментації об'єктів з однаковими просторовими ознаками [25], властивими для контурів будівель міської забудови. Параметри *max_epochs*, *learning_rate*, *stop_training* прийнято за замовчуванням, тобто кількість періодів навчання моделі становить 20, оптимальну швидкість навчання отримують за кривою навчання в процесі, навчання моделі зупиняється в разі відсутності покращення моделі [24, 25].

Етап 4. Перевірка навченої моделі *Deep Learning*:

– на вихідному супутниковому знімку міської забудови для порівняння результатів розпізнавання з огляду на елементи навчальної вибірки;

– на довільних знімках для визначення можливості адаптації моделі до інших вихідних даних.

Етап 5. Використання моделі *Deep Learning* для виявлення об'єктів міської забудови відповідно до знімків.

Результати застосування методу дають змогу отримати кількісні та якісні властивості об'єктів будівництва, оцінити динаміку реальних змін міського середовища, актуалізувати інформацію про міську

забудову тощо. Це є необхідним кроком для формування напрямів інтелектуалізації міста [23].

Аналіз забудови міста Києва за супутниковими знімками

Фахівці зазначають, що Київ будують швидко, неконтрольовано й найчастіше непередбачувано. Чинний генеральний план міста вважають застарілим через наявність у ньому нормативів, які є наслідком радянської планувальної традиції, вони часто суперечать один одному, а рішення щодо міського планування приймаються непрозоро. Ці проблеми не може усунути новий генеральний план, адже жоден з планів Києва ніколи не виконувався через корупцію, а також незацікавленість судової влади та правоохоронних органів у розв'язанні проблем. Кількість незаконних забудов, споруджених без відповідних дозвільних документів або всупереч їм, постійно збільшується. За не завжди законно зведеними житловими масивами не встигають будувати школи, а установи соціальної інфраструктури, транспортна інфраструктура й необхідні комунікації не відповідають Державним будівельним нормам тощо [4, 21]. За цих обставин дані ДЗЗ стають майже єдиним джерелом актуальної об'єктивної інформації щодо міської забудови. Тому для дослідження можливостей запропонованого методу розглянемо супутникові знімки одного з районів м. Києва.

На першому етапі для актуалізації інформації про міську забудову Києва та оцінювання динаміки її зміни з використанням сервісу *Google Earth Pro* завантажимо два супутникових знімки району столиці за різні часові періоди – 2005 р. (рис. 4, а) і 2022 р.



а



б

Рис. 4. Об'єкт дослідження в Подільському районі м. Києва: а – супутниковий знімок території дослідження (2005 р.); б – виділення фрагментів зображення об'єктів забудови для навчальної вибірки

У програмному забезпеченні *ArcGIS Pro* визначимо полігони, створимо набір класів об'єктів навчальної вибірки (рис. 4, б), за допомогою інструменту *Export_Training_Data_for_Deep_Learning* отримаємо фрагменти зображення об'єктів для встановлених класів у форматі .TIFF. У такий спосіб сформуємо базу даних навчальної вибірки.

Модель навчалася з використанням нейронної мережі *Mask R-CNN* за допомогою циклічного аналізу ознак об'єктів вибірки. Зазначимо, що вже на п'ятій ітерації було досягнуто мінімальних втрат під час навчання ($loss_train \approx 0,483$), у цьому разі швидкість навчання в процесі руху до цього мінімуму становила $learning_rate \approx 8,32 \cdot 10^{-6}$.

На рис. 5 наведено динаміку зміни втрат ($loss_train$), які є властивістю похибки навчання на навчальних даних для кожної ітерації, та втрат валідації ($loss_validation$), що вимірюють похибку в окремому перевірчому наборі даних, які модель не бачила під час її навчання. Процес змін $loss_train$ і $loss_validation$ демонструє добре навчання моделі та її рух до мінімуму втрат не лише на навчальному наборі даних. У цьому разі забезпечується узагальнення моделі на новій інформації та підтримується задовільна продуктивність на перевірчому наборі даних.

Отже, досягнення мінімального значення похибок свідчить про те, що отримано добре навчену модель *Deep Learning*.

Переходимо до етапу тестування та перевіримо її:

- на вихідному супутниковому знімку міської забудови для порівняння результатів розпізнавання на навчальній вибірці (рис. 6, а);
- на довільному супутниковому зображенні міської забудови, яке майже не має однакових із навчальною вибіркою об'єктів (рис. 6, б).

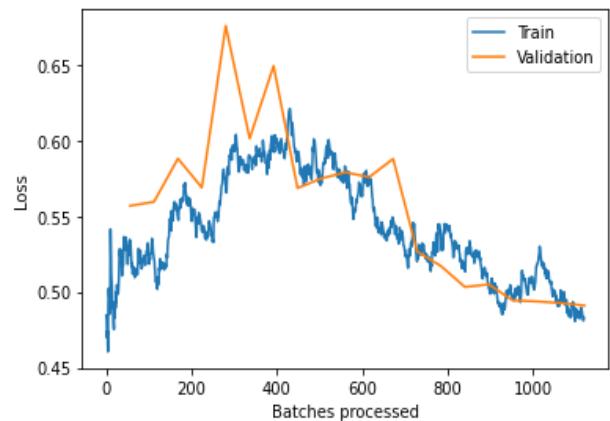


Рис. 5. Динаміка зміни похибок $loss_train$ і $loss_validation$ для кожної ітерації навчання моделі

Аналізуючи рис. 6, зазначимо, що модель дає добрі результати, розпізнає об'єкти міської забудови за межами визначеної навчальної ділянки та гарно пристосовується до інших об'єктів, на яких модель *Deep Learning* не тренувалася.

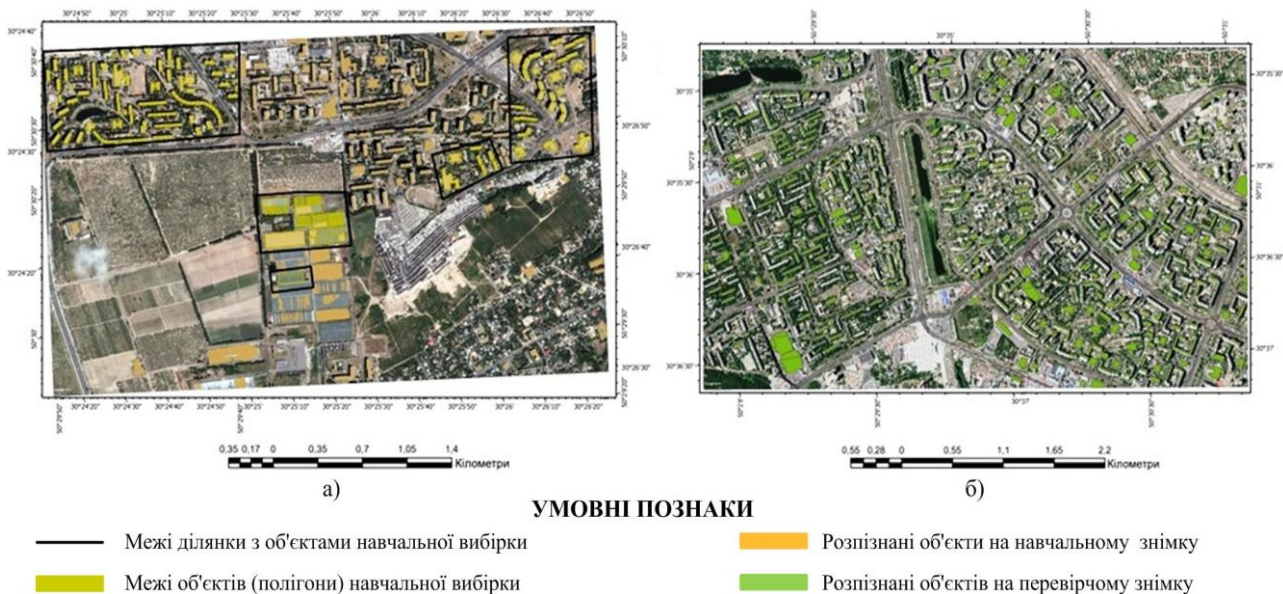


Рис. 6. Перевірка навченої моделі *Deep Learning*: а – порівняння результатів розпізнавання на навчальній вибірці; б – результати розпізнавання на довільному супутниковому знімку

Результати дослідження та їх обговорення

Аналіз міської забудови району м. Києва та її динаміки за допомогою моделі *Deep Learning* відбувався за супутниковими знімками, зробленими 2005 р. і 2022 р.

На знімку 2005 р. (рис. 7) знайдено та класифіковано 380 об'єктів міської забудови, зокрема житлову забудову (доля якої становить 68,5%), школи, лікарні, тепличний комплекс та інші об'єкти (супермаркети, аптеки, кінотеатр, автозаправні станції тощо). Біля тепличного комплексу також ідентифіковано землі сільськогосподарського призначення та сади, які займають значну площу району (ліва нижня частина знімку).

На знімку 2021 р. (рис. 8) знайдено та класифіковано 509 об'єктів міської забудови. Крім визначених раніше об'єктів, дешифровано стадіон і споруди, що будуються. Тут майже повністю зник тепличний комплекс, на його місці виявлено нові завершені розбудови (п'ять нових житлових комплексів) і житлові забудови на стадії будівництва, з'явився торговий центр. На території земель сільськогосподарського призначення ідентифіковано житловий комплекс і великий торговий комплекс.

Отже, спостерігаємо значну кількість багатоповерхової житлової забудови, частка якої стабільно збільшується (сягає приблизно 73,9%),

водночас кількість закладів освіти та лікарень, якщо порівнювати з 2005 р., не змінилася. Не розвиваються садово-паркові зони, обмежуються території озеленення та прибудинковий простір. Зростає кількість комерційної забудови – торгових центрів, супермаркетів – приблизно на 11% порівняно з 2005 р. Тому необхідно більшу увагу приділити конкретним заходам модернізації транспортної та інженерної інфраструктури, оптимізації вуличної мережі, будівництву мостів і багаторівневих транспортних розв'язок тощо.

За цих обставин можливими напрямками інтелектуалізації відповідно до моделі (1) можуть стати:

- підвищення ефективності постачальників медичних послуг, доступності первинної лікарської допомоги не лише завдяки будівництву нових закладів, але внаслідок збільшення кількості дистанційних консультацій у медичних фахівців, впровадження електронних систем ведення медичного обліку, створення online-профілів пацієнтів тощо;
- підвищення ефективності планування розвитку транспортної інфраструктури міста з пріоритизацією на розвиток електротранспортної мережі, на організацію інтелектуальних світлофорних постів тощо;
- підвищення ефективності функціонування транспортної системи міста завдяки впровадженню системи диспетчеризації на основі GPS-навігації, дотримання розкладу руху транспорту, застосування єдиної системи платежів тощо.



УМОВНІ ПОЗНАКИ

 Житлова забудова	 Тепличний комплекс
 Школи	 Інші об'єкти
 Лікарні	

Рис. 7. Класифіковані об'єкти міської забудови в Подільському районі м. Києва (2005 р.)



УМОВНІ ПОЗНАКИ

Житлова забудова	Стадіон	Інші об'єкти
Будинки, що будуються	Школи	
Тепличний комплекс	Лікарні	

Рис. 8. Класифіковані об'єкти міської забудови в Подільському районі м. Києва (2021 р.)

Висновки та перспективи подальшого розвитку

Тепер майже половина населення планети живе в містах, а в подальшому кількість містян може збільшитися до 5 млрд [1]. Узагальнення тенденцій розвитку мегаполісів підтверджує, що створення зручного міського середовища сприяє економічному зростанню та розв'язанню проблем з ресурсами, покращує якість життя, забезпечує сталий розвиток. Але, як свідчать результати бібліографічного пошуку, це можливо завдяки впровадженню концепції розумного міста, яку спрямовано на підтримку балансу його економічного, соціального та екологічного складників [3].

Для уніфікації кроків зі створення розумних міст і формування способів розв'язання міських проблем розроблено теоретико-множинну модель розумного міста, яка, ґрунтуючись на вимогах міжнародних стандартів серії ISO 37100 [19], поєднує сукупність показників, що забезпечують реалізацію певних напрямів інтелектуалізації. У цьому разі визначена група чинників, що безпосередньо характеризують ефективність упровадження інноваційних технологій у життя містян (табл. 1).

Вибір напрямку інтелектуалізації має ґрунтуватися на ретельному вивченні ситуації, що виникла внаслідок урбанізації міських районів. Складність цього питання потребує вдосконалення аналізу

міської забудови та залучення нових підходів та інформації, зокрема даних ДЗЗ, адже інтеграція цих технологій забезпечує можливості з багатовимірного моніторингу, підвищує ґрунтовність прогнозних оцінок і пом'якшує наслідки нестійкої урбанізації [20]. Для спрощення процедури оброблення цих даних та отримання оцінок, придатних для подальшого аналізу й формування рішень, запропоновано метод аналізу змін міської забудови за даними ДЗЗ. Він ґрунтується на використанні глибокого навчання в обробленні структурованої інформації за допомогою інструменту *Image Analyst ArcGIS Pro 3.4* [24]. Розроблено схему алгоритму отримання моделі *Deep Learning* для дешифрування об'єктів міської забудови, наведено її програмну реалізацію.

Експериментальне оцінювання можливостей розробленого методу відбувалося під час виконання завдань актуалізації інформації про міську забудову Києва.

Отже, отримана в процесі реалізації методу модель *Deep Learning* добре розпізнає об'єкти міської забудови та добре пристосована до об'єктів, що різняться від об'єктів навчальної вибірки. У цьому разі внаслідок аналізу динаміки змін міської забудови району м. Києва за період 2005–2021 рр. встановлено таке:

– збільшення кількості об'єктів житлової (на 5,4%) і комерційної забудови (на 11%), що підтверджує

загальносвітову тенденцію нестійкої урбанізації, коли забудова відбувається лише з метою одержання швидкого комерційного зиску [2, 8];

– нехтування розвитком соціальної інфраструктури, зокрема не зафіксовано збільшення кількості освітніх і медичних закладів;

– відсутність розбудови транспортної інфраструктури, зокрема на знімках не розпізнано нових транспортних розв'язок або мостів;

– актуальним стає питання щодо озеленення території району та створення садово-паркових зон тощо.

Отже, ґрунтуючись на отриманих оцінках міської забудови району м. Києва, формують пріоритетні напрями інтелектуалізації та визначають шляхи впровадження інтелектуальних рішень у повсякденне життя містян відповідно до табл. 1.

Зазначимо, що запропоновані підходи можуть успішно використовуватися під час воєнного стану (наприклад, для визначення обсягів руйнувань або пошкоджень внаслідок обстрілів) і в умовах післявоєнного відновлення країни (наприклад, для встановлення обсягів відшкодувань), адже модель *Deep Learning* добре пристосована до інших об'єктів, на яких вона не тренувалася.

Роботу виконано за підтримки Міністерства освіти і науки України (державні реєстраційні номери проєктів 0125U000573, 0125U001680) і Регіонального центру космічного моніторингу Землі "Слобожанщина".

Внесок авторів. Огляд та аналіз літератури – **С. Даншина, С. Андрєєв**; розроблення концептуальних положень дослідження – **С. Даншина**; практична апробація результатів статті – **С. Андрєєв**; аналіз результатів дослідження **С. Даншина, С. Андрєєв**; адміністрування проєкту – **С. Даншина**. Усі автори прочитали та погодилися з опублікованою версією рукопису.

Висловлюємо подяку **Тарасу Ісаєнку** за практичну перевірку ефективності розробленого програмно-алгоритмічного забезпечення.

Список літератури

1. Tricomi G., CV POp-CoRN: The (smart) city-vehicle participatory-opportunistic cooperative route navigation system / G. Tricomi, C. Scaffidi, A. Puliafito, S. Distefano. *Ad Hoc Networks*. 2024. Article 103604. DOI: <https://doi.org/10.1016/j.adhoc.2024.103604>
2. Teng Q., Bai X., Apuke O. D. Modelling the factors that affect the intention to adopt emerging digital technologies for a sustainable smart world city. *Technology in Society*. 2024. Vol. 78. Article 102603. DOI: <https://doi.org/10.1016/j.techsoc.2024.102603>
3. Yang R., Zhen F. Smart city development Models: A cross-cultural regional analysis from theory to practice. *Research in Globalization*. 2024. Vol. 8. Article 100221. DOI: <https://doi.org/10.1016/j.resglo.2024.100221>
4. Чичкало-Кондрацька І. Б., Буряк А. А., Кондрацька Д. С. Особливості створення та перспективи розвитку Smart Cities у країнах світу. *Ефективна економіка*. 2020. № 8. DOI: <https://doi.org/10.32702/2307-2105-2020.8.9>
5. IMD Smart City Index 2024. URL: https://www.imd.org/wp-content/uploads/2024/04/20240412-SmartCityIndex-2024-Full-Report_4.pdf. (дата звернення 1.07.2024).
6. IMD World Competitiveness Booklet. URL: https://www.imd.org/wp-content/uploads/2023/06/WCY_Booklet_2023-FINAL.pdf. (дата звернення 1.07.2024).
7. Zhu J. How different can smart cities be? A typology of smart cities in China / J. Zhu et al. *Cities*. 2024. Vol. 149. Article 104992. DOI: <https://doi.org/10.1016/j.cities.2024.104992>
8. Okonta D. E., Vukovic V. Smart cities software applications for sustainability and resilience. *Heliyon*. 2024. Vol. 10, Issue 2. Article e32654. DOI: <https://doi.org/10.1016/j.heliyon.2024.e32654>
9. Liu Y., Wen X. Sustainability assessment of cities using multicriteria decision-making combined with deep learning methods *Sustainable Cities and Society*. 2024. Vol. 115. Article 105571. DOI: <https://doi.org/10.1016/j.scs.2024.105571>
10. Ішпак О., Федорка П., Пригара М. Розумні міста та Інтернет речей: вплив розробок у сфері ІТ на розвиток міст і покращення якості життя. *Сучасний стан наукових досліджень та технологій в промисловості*. 2023. № 3 (25). С. 114 – 128. DOI: <https://doi.org/10.30837/ITSSI.2023.25.114>
11. Kim J. S., Feng Y. Understanding complex viewpoints in smart sustainable cities: The experience of Suzhou, China. *Cities*. 2024, Vol. 147, Article 104832. DOI: <https://doi.org/10.1016/j.cities.2024.104832>

12. Han M.J.N., Kim M. J. A systematic literature review of the smart city transformation process: The role and interaction of stakeholders and technology. *Cities*. 2024. Vol. 150. Article 105027. DOI: <https://doi.org/10.1016/j.cities.2024.105027>
13. Abdullah Kaiser Z. R. M. Smart Governance for Smart Cities and Nations: The Case of Smart Bangladesh. *Journal of Economy and Technology*. 2024, Article 003. DOI: <https://doi.org/10.1016/j.ject.2024.07.003>
14. Makkonen T., Inkinen T. Inclusive smart cities? Technology-driven urban development and disabilities. *Cities*. 2024. Vol. 154. Article 105334. DOI: <https://doi.org/10.1016/j.cities.2024.105334>
15. Ulya A. Major Dimensions of Smart City: A Systematic Literature Review / A. Ulya, T. D. Susanto, Y. S. Dharmawan, A. P. Subriadi. *Procedia Computer Science*. 2024. Vol. 234. P. 996-1006. DOI: <https://doi.org/10.1016/j.procs.2024.03.089>
16. Florentin K. M., Onuki M., Yarime M. Facilitating citizen participation in Greenfield smart city development: The case of a human-centered approach in Kashiwanoha international campus town. *Telematics and Informatics Reports*. 2024. Vol. 15. Article 100154. DOI: <https://doi.org/10.1016/j.teler.2024.100154>
17. Trencher G. Towards the smart city 2.0: Empirical evidence of using smartness as a tool for tackling social challenges. *Technological Forecasting and Social Change*. 2019. Vol. 142. P. 117-128. DOI: <https://doi.org/10.1016/j.techfore.2018.07.033>
18. Даншина С. Ю., Андреев С. М. Дистанційні дані у процесі реставрації пам'яток архітектури. *Управління розвитком складних систем*. 2024. No. 59. С. 138-147. DOI: <https://doi.org/10.32347/2412-9933.2024.59.138-147>
19. ISO 37122:2019. Sustainable cities and communities – Indicators for smart cities. URL: <https://www.iso.org/standard/69050.html>. (дата звернення 5.07.2024).
20. Zheng Y. Wavelet Transform Cluster Analysis of UAV Images for Sustainable Development of Smart Regions Due to Inspecting Transport Infrastructure / Y. Zheng et al. *Sustainability*. 2025. Vol. 17(3), Article 927. DOI: <https://doi.org/10.3390/su17030927>
21. Дюжев С. Генеральне стратегічне містобудівне планування та проблеми планувального управління розселенням (частина друга: проблеми, перешкоди щодо їх розв'язання, актуальні завдання та технологічні вимоги до змісту містобудівної документації). *Містобудування та територіальне планування*. 2023. № 84. С. 64-131. DOI: <https://doi.org/10.32347/2076-815x.2023.84.64-131>
22. Danshyna S., Nechausov A. Andrieiev S. Information technology of transport infrastructure monitoring based on remote sensing data. *Radio Electronics, Computer Science, Control*. 2022. No. 4 (63). P. 86–97. DOI: <https://doi.org/10.15588/1607-3274-2022-4-7>
23. Zhou H. Identifying influencing factors and characterizing key issues in urban sustainable development capacity through machine learning / H. Zhou et al. *Chinese Journal of Population, Resources and Environment*. 2024. Vol. 22, issue 3. P. 291-304. DOI: <https://doi.org/10.1016/j.cjpre.2024.09.008>
24. Deep learning using the ArcGIS Image Analyst extension. URL: <https://pro.arcgis.com/en/pro-app/3.3/help/analysis/image-analyst/deep-learning-in-arcgis-pro.htm>.
25. Shaoqing R. University of Science and Technology of China, Hefei, Anhui, ChinaFaster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks / Sh. Ren, Kaiming He, R. Girshick, J. Sun. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017. Vol. 39, Issue 6. P. 1137-1149. DOI: <https://doi.org/10.1109/TPAMI.2016.2577031>

References

1. Tricomi, G., Scaffidi, C., Puliafito, A., Distefano, S. (2024), "CV POp-CoRN: The (smart) city-vehicle participatory-opportunistic cooperative route navigation system", *Ad Hoc Networks*, Article 103604. DOI: <https://doi.org/10.1016/j.adhoc.2024.103604>
2. Teng, Q., Bai, X., Apuke, O. D. (2024), "Modelling the factors that affect the intention to adopt emerging digital technologies for a sustainable smart world city", *Technology in Society*, Vol. 78, Article 102603. DOI: <https://doi.org/10.1016/j.techsoc.2024.102603>
3. Yang, R., Zhen, F. (2024), "Smart city development Models: A cross-cultural regional analysis from theory to practice", *Research in Globalization*, Vol. 8, Article 100221. DOI: <https://doi.org/10.1016/j.resglo.2024.100221>
4. Chychkalo-Kondratska, I., Buriak, A., Kondratska, D. (2020), "Features of Creation and Prospects of Development of Smart Cities in the Countries of the World" ["Osoblyvosti stvorennia ta perspektyvy rozvytku Smart Cities u krainakh svitu"], *Effective Economy*, No. 8, DOI: <https://doi.org/10.32702/2307-2105-2020.8.9>
5. "IMD Smart City Index 2024". available at: https://www.imd.org/wp-content/uploads/2024/04/20240412-SmartCityIndex-2024-Full-Report_4.pdf. (last accessed 1.07.2024).
6. "IMD World Competitiveness Booklet". available at: https://www.imd.org/wp-content/uploads/2023/06/WCY_Booklet_2023-FINAL.pdf. (last accessed 1.07.2024).

7. Zhu, J., Gianoli, A., Noori, N., de Jong, M., Edelenbos, J. (2024), "How different can smart cities be? A typology of smart cities in China", *Cities*, Vol. 149, Article 104992. DOI: <https://doi.org/10.1016/j.cities.2024.104992>
8. Okonta, D. E., Vukovic, V. (2024), "Smart cities software applications for sustainability and resilience", *Heliyon*, Vol. 10, issue 2, Article e32654. DOI: <https://doi.org/10.1016/j.heliyon.2024.e32654>
9. Liu, Y., Wen, X. (2024), "Sustainability assessment of cities using multicriteria decision-making combined with deep learning methods", *Sustainable Cities and Society*, Vol. 115, Article 105571. DOI: <https://doi.org/10.1016/j.scs.2024.105571>
10. Shpak, O., Fedorka, P., Prygara, M. (2023), "Smart cities and the internet of things: the impact of its developments on the development of cities and improving the quality of life" ["Rozumni mista ta Internet rechei: vplyv rozrobok u sferi IT na rozvytok mist i pokrashchennia yakosti zhyttia"], *Innovative Technologies and Scientific Solutions for Industries*, No. 3 (25), P. 114 – 128. DOI: <https://doi.org/10.30837/ITSSI.2023.25.114>
11. Kim, J. S., Feng, Y. (2024), "Understanding complex viewpoints in smart sustainable cities: The experience of Suzhou, China", *Cities*, Vol. 147, Article 104832. DOI: <https://doi.org/10.1016/j.cities.2024.104832>
12. Han, M.J.N., Kim, M. J. (2024), "A systematic literature review of the smart city transformation process: The role and interaction of stakeholders and technology", *Cities*, Vol. 150, Article 105027. DOI: <https://doi.org/10.1016/j.cities.2024.105027>
13. Abdullah Kaiser, Z. R. M. (2024), "Smart Governance for Smart Cities and Nations: The Case of Smart Bangladesh", *Journal of Economy and Technology*, Article 003. DOI: <https://doi.org/10.1016/j.ject.2024.07.003>
14. Makkonen, T., Inkinen, T. (2024), "Inclusive smart cities? Technology-driven urban development and disabilities", *Cities*, Vol. 154, Article 105334. DOI: <https://doi.org/10.1016/j.cities.2024.105334>
15. Ulya, A., Susanto, T. D., Dharmawan, Y. S., Subriadi, A. P. (2024), "Major Dimensions of Smart City: A Systematic Literature Review", *Procedia Computer Science*, Vol. 234, P. 996 – 1006. DOI: <https://doi.org/10.1016/j.procs.2024.03.089>
16. Florentin, K. M., Onuki, M., Yarime, M. (2024), "Facilitating citizen participation in Greenfield smart city development: The case of a human-centered approach in Kashiwanoha international campus town", *Telematics and Informatics Reports*, Vol. 15, Article 100154. DOI: <https://doi.org/10.1016/j.teler.2024.100154>
17. Trencher, G. (2019), "Towards the smart city 2.0: Empirical evidence of using smartness as a tool for tackling social challenges", *Technological Forecasting and Social Change*, Vol. 142, P. 117 – 128. DOI: <https://doi.org/10.1016/j.techfore.2018.07.033>
18. Danshyna, S., Andrieiev, S. (2024), "Remote data in the process of architectural monuments restoration", ["Dystantsiini dani u protsesi restavratsii pamiatok arkhitektury"], *Management of Development of Complex Systems*, No. 59. pp. 138 – 147. DOI: <https://doi.org/10.32347/2412-9933.2024.59.138-147>
19. "ISO 37122:2019. Sustainable cities and communities – Indicators for smart cities". available at: <https://www.iso.org/standard/69050.html>. (last accessed 5.07.2024).
20. Zheng, Y., Shcherbakova, G., Rusyn, B., Sachenko, A., Volkova, N., Kliushnikov, I., & Antoshchuk, S. (2025), "Wavelet Transform Cluster Analysis of UAV Images for Sustainable Development of Smart Regions Due to Inspecting Transport Infrastructure", *Sustainability*, Vol. 17(3), Article 927. DOI: <https://doi.org/10.3390/su17030927>
21. Dyuzhev, S. (2023), "General strategic city planning designing and problems of planning management for settling (the second part: problems, hindrances concerning their solution, actual targets and technological demands to the content of city planning documentations)" ["Heneralne stratehichne mistobudivne planuvannia ta problemy planuvalnoho upravlinnia rozselenniam (chastyna druha: problemy, pereshkody shchodo yikh rozviazannia, aktualni zavdannia ta tekhnolohichni vymohy do zmistu mistobudivnoi dokumentatsii) "], *Urban Development and Spatial Planning*, No. 84, P. 64 – 131. DOI: <https://doi.org/10.32347/2076-815x.2023.84.64-131>
22. Danshyna, S., Nechausov, A. Andrieiev, S. (2022), "Information technology of transport infrastructure monitoring based on remote sensing data", *Radio Electronics, Computer Science, Control*, No. 4 (63). P. 86 – 97. DOI: <https://doi.org/10.15588/1607-3274-2022-4-7>
23. Zhou, H., Gao, L., Xu, Z., Shi, L., Lv Q. (2024), "Identifying influencing factors and characterizing key issues in urban sustainable development capacity through machine learning", *Chinese Journal of Population, Resources and Environment*, Vol. 22, Issue 3, P. 291 – 304. DOI: <https://doi.org/10.1016/j.cjpre.2024.09.008>.
24. "Deep learning using the ArcGIS Image Analyst extension". available at: <https://pro.arcgis.com/en/pro-app/3.3/help/analysis/image-analyst/deep-learning-in-arcgis-pro.htm>. (last accessed 15.01.2025).
25. Shaoqing, R., Kaiping, He., Girshick, R., Sun, J. (2017), "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, Issue 6, P. 1137 – 1149. DOI: <https://doi.org/10.1109/TPAMI.2016.2577031>

Відомості про авторів / About the Authors

Даншина Світлана Юрївна – доктор технічних наук, Національний аерокосмічний університет "Харківський авіаційний інститут", професор кафедри геоінформаційних технологій та космічного моніторингу Землі, Харків, Україна; e-mail: s.danshyna@khai.edu; ORCID ID: <https://orcid.org/0000-0001-7354-4146>

Андрєєв Сергій Михайлович – кандидат технічних наук, Національний аерокосмічний університет "Харківський авіаційний інститут", доцент кафедри геоінформаційних технологій та космічного моніторингу Землі, Харків, Україна; e-mail: s.andreev@khai.edu; ORCID ID: <https://orcid.org/0000-0003-4256-2637>

Danshyna Svitlana – Doctor of Sciences (Engineering), National Aerospace University "Kharkiv aviation institute", Professor at the Department of Geo-information Technologies and Space Monitoring of the Earth, Kharkiv, Ukraine.

Andriev Sergey – PhD (Engineering Sciences), National Aerospace University "Kharkiv aviation institute", Associate Professor at the Department of Geo-information Technologies and Space Monitoring of the Earth, Kharkiv, Ukraine.

FORMATION OF A SMART CITY MODEL BASED ON THE DYNAMICS OF CHANGES IN URBAN BUILDING

The **purpose** of the article is improving the validity of assessments for the formation of decisions on the intellectualization of the city based on analysis of urban building. **Objectives.** The trends in the development of Smart City concepts were analyzed; requirements for Smart Cities are analyzed and a set-theoretical model of a Smart City is proposed; the method of analyzing urban building using remote sensing data has been improved and the possibility of its use has been experimentally confirmed. The following **methods** used are – methods of system analysis, set theory, and deep learning. The following **results** were obtained. Creating a comfortable urban environment promotes economic growth, improves the quality of life and ensures sustainable development. This is possible through the implementation of the smart city concept. To unify the steps to create Smart Cities, a set-theoretic model of a Smart City is proposed. It generalizes the requirements of standards of the ISO 37100 series and combines a set of indicators, ensuring the implementation of certain areas of intellectualization. The choice of the intellectualization direction should be based on the study of urbanization problems. The complexity of this issue requires improvement of urban building analysis processes, for example, using remote sensing data. To obtain estimates suitable for further analysis and decision-making, a method for studying urban development based on remote sensing data is proposed. It is uses of deep learning in processing structured data using the Image Analyst ArcGIS Pro 3.4. A scheme of the algorithm for obtaining a Deep Learning model for decrypting urban building objects has been developed, and its software implementation has been proposed. The study of the capabilities of the developed method was carried out when solving the problems of updating information about the urban development of Kyiv. The resulting Deep Learning model recognizes urban building objects well and has good adaptability to objects other than the training sample. **Conclusions:** The choice of the direction of intellectualization should be based on a thorough study of the situation that has arisen as a result of the urbanization. An experimental analysis of the dynamics of changes in the urban development of the Kyiv district for the period 2005 – 2021 showed: an increase in the number of residential (by 5.4%) and commercial development (by 11%), neglect of the development of social infrastructure, non-development of transport infrastructure, ignoring issues of landscaping the territory of the district, etc. The results obtained can be useful in forming priority ways of introducing intelligent solutions into the everyday life of city residents.

Keywords: set-theoretic model; remote sensing data; deep learning, updating data; directions of city intellectualization.

Бібліографічні описи / Bibliographic descriptions

Даншина С. Ю., Андрєєв С. М. Формування моделі розумного міста з огляду на динаміку змін міської забудови. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 2 (32). С. 16–32. DOI: <https://doi.org/10.30837/2522-9818.2025.2.016>

Danshyna, S., Andriev, S. (2025), "Formation of a smart city model based on the dynamics of changes in urban building", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (32), P. 16–32. DOI: <https://doi.org/10.30837/2522-9818.2025.2.016>

Б. КОСАРЕВСЬКИЙ, А. ТЕЦЬКИЙ

СУЧАСНІ ПІДХОДИ ДО РОЗГОРТАННЯ ІНФРАСТРУКТУРИ МОБІЛЬНИХ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ

Предмет дослідження – інфраструктура мобільних інтелектуальних систем (МІС) для моніторингу критичних об'єктів із застосуванням груп безпілотних літальних апаратів (БПЛА), що поєднує крайові та хмарні обчислення, методи балансування навантаження та кібербезпеки. **Мета роботи** – аналіз методів розгортання інфраструктури МІС моніторингу критичних об'єктів, орієнтованих на використання груп БПЛА. Основну увагу зосереджено на підвищенні ефективності, надійності та адаптивності таких систем унаслідок інтеграції хмарних і крайових обчислень, розроблення механізмів автоматизованого керування ресурсами, кіберзахисту та адаптації до умов експлуатації. **Завдання:** розробити архітектуру МІС з гібридним розподілом обчислень між *Edge*- та *Cloud*-компонентами; дослідити вплив механізмів балансування ресурсів на продуктивність за умови змінного навантаження; оцінити ефективність мультиканальних технологій зв'язку в разі втрати основного каналу; впровадити методи кіберзахисту для забезпечення безперебійної роботи системи. **Методи:** теоретичний аналіз наявних підходів, моделювання та симуляція для оцінювання їх продуктивності й надійності. Дослідження передбачає використання генетичних алгоритмів, ройових алгоритмів і штучних потенційних полів для управління траєкторіями ботів. **Результати дослідження.** Експериментальні випробування в середовищах *Microsoft Azure* та *CoppeliaSim EDU* підтвердили ефективність запропонованих підходів. Досягнуто зменшення середньої затримки на 35%, покращено енергоспоживання, забезпечено безперервність передачі даних у 92% сценаріїв втрати зв'язку. Аналіз надійності показав доцільність резервування компонентів і впровадження методів прогнозування, що дало змогу знизити ймовірність критичних відмов на 22% та скоротити середній час відновлення на 42%. **Висновки.** Використання автоматизованого управління ресурсами в МІС забезпечує стабільність і безперервність роботи груп БПЛА навіть у змінних умовах експлуатації. Оптимізація обчислень і адаптивне масштабування сприяють підвищенню продуктивності системи, зменшенню затримки передачі інформації та покращенню енергоспоживання. Розроблені підходи до кібербезпеки гарантують захист даних і надійність функціонування інфраструктури в разі зовнішніх загроз та мережових атак.

Ключові слова: точкова хмара; глибинне навчання; графова нейронна мережа; автоматичне доповнення точкових хмар; кодування графів; 3D-реконструкція графових моделей.

Вступ

Сучасні технології відіграють ключову роль у забезпеченні моніторингу критичних об'єктів, що вимагає впровадження мобільних інтелектуальних систем (МІС), зокрема безпілотних літальних апаратів (БПЛА), наземних роботизованих платформ і хмарних обчислювальних технологій. Використання таких систем дає змогу значно підвищити ефективність контролю за станом стратегічно важливих інфраструктурних об'єктів, транспортних мереж і екологічно небезпечних зон, що потребують постійного спостереження та аналізу.

Розгортання інфраструктури таких систем передбачає інтеграцію передових технологій оброблення інформації, децентралізованих мереж, машинного навчання та механізмів кіберзахисту. Зокрема актуальним є впровадження *Mobile Edge Computing* (MEC) для оброблення даних безпосередньо в місці їх збору, що дає змогу мінімізувати затримки та зменшити навантаження на центральні сервери [1]. У цьому разі

важливу роль відіграє ієрархічна архітектура периферійних обчислень, яка сприяє оптимізації процесів прийняття рішень та покращенню масштабованості систем.

Використання БПЛА для моніторингу критичних об'єктів допомагає підвищити автономність систем спостереження, забезпечуючи швидке розгортання мережі сенсорів і збирання інформації в режимі реального часу. Однак зростання складності таких систем спричиняє низку викликів, пов'язаних із забезпеченням надійного зв'язку, з енергоспоживанням, управлінням роєм БПЛА та протидією потенційним кібератакам [2].

Окремої уваги заслуговує безпека, оскільки системи моніторингу, що базуються на БПЛА, можуть стати об'єктами кібератак, спрямованих на злам каналів зв'язку та компрометацію переданих даних. Запропоновані в останніх дослідженнях методи захисту інформаційних потоків, зокрема комбіновані підходи, що поєднують сигнатурний аналіз та виявлення аномалій, є важливим напрямом у забезпеченні безпечного функціонування таких систем.

Крім того, автоматизація розгортання інфраструктури МІС та керування нею є ключовим чинником підвищення ефективності їх застосування в умовах мінливого середовища. Використання алгоритмів оптимізації та нейронних мереж дає змогу значно покращити точність прийняття рішень, мінімізувати вплив людського фактора та скоротити час реакції на позаштатні ситуації [3].

Отже, розгортання інфраструктури мобільних інтелектуальних систем моніторингу критичних об'єктів є складним процесом, що вимагає поєднання різних технологій і уваги до низки чинників, таких як надійність, масштабованість, безпека й ефективність управління. З огляду на сучасні виклики та тенденції подальші дослідження в цьому напрямі мають бути спрямовані на розроблення нових підходів до інтеграції МЕС, оптимізацію алгоритмів керування групами БПЛА та впровадження вдосконалених механізмів кіберзахисту [4].

Аналіз останніх досліджень і публікацій

Робота С. Лютфі та інших [5] демонструє інтеграцію МЕС у 6G-мережі, що дає змогу оптимізувати передачу інформації між мобільними агентами. Т. Хуан [6] досліджує оптимізацію якості обслуговування в МЕС для IoT, що є актуальним для моніторингу критичних об'єктів. А. Рахмані та інші [7] запропонували використання блокчейн-технологій у МЕС, що підвищує безпеку оброблення даних. Крім того, автори статті [8] досліджують флюїдні антени, які покращують адаптивність МЕС-систем.

А. Фірузі [9] розглядає використання штучного інтелекту (AI) та Інтернету речей (IoT) для дистанційного моніторингу критичної інфраструктури. Автор аналізує найкращі практики розгортання та експлуатації таких систем, що має важливе значення для мобільних інтелектуальних систем. С. Го та інші [10] досліджують використання МЕС у системах реального часу для виявлення об'єктів на відео в промислових середовищах. Запропонований метод трекінгу покращує продуктивність і зменшує потребу в обчислювальних ресурсах. Автори праці [11] описали методи екологічного моніторингу на основі IoT, зокрема сенсорні системи та інфраструктурні рішення для оцінювання навколишнього середовища в режимі реального часу. А. Марданшагі та інші [12] проаналізували сучасні методи структурного моніторингу та використання цифрових двійників

для оцінювання стану критичних об'єктів, що безпосередньо пов'язано з ефективністю мобільних інтелектуальних систем моніторингу.

Забезпечення кіберстійкості комп'ютерних систем критичної інфраструктури є ключовим аспектом захисту мобільних інтелектуальних систем від можливих загроз. У своїй роботі М. Комаров [4] пропонує комбінований метод виявлення загроз, що поєднує сигнатурний аналіз і методи виявлення аномалій. Це сприяє покращенню механізмів реагування на кібератаки та підвищенню безпеки інформаційних потоків у критичних системах.

О. Погудіна та інші [13] розглядають підходи до управління роєм БПЛА в автономних системах моніторингу. Запропоновані алгоритми дають змогу забезпечити ефективний обмін інформацією між апаратами, підвищуючи координацію та адаптивність групових операцій.

Л. Артюшин, О. Кононов, А. Кир'янов [14] у дослідженні запропонували програмний модуль для керування групою БПЛА, що допомагає покращити точність і швидкість виконання завдань моніторингу критичних об'єктів. Застосування такого модуля сприяє підвищенню автономності та ефективності розгортання груп БПЛА.

А. Позднякова [15] аналізує впровадження інтелектуальних технологій у міській інфраструктурі, наголошуючи на ролі IoT та AI в моніторингу критичних об'єктів. Авторка аналізує перспективи для розвитку автоматизованих систем управління міськими та промисловими об'єктами.

Нерозв'язані проблеми й мета дослідження

Попри значний прогрес у використанні БПЛА, залишається низка нерозв'язаних проблем, пов'язаних із розгортанням та управлінням інфраструктурою [3].

Однією з ключових проблем є адаптивне управління ресурсами хмарних і крайових обчислень, необхідних для оброблення значного обсягу інформації, що надходить від БПЛА. Оскільки такі системи часто функціонують у розподіленому середовищі, необхідно розробити ефективні методи балансування навантаження, що дадуть змогу оптимально використовувати обчислювальні ресурси без затримок у передачі та аналізі інформації. Крім того, динамічний характер розташування БПЛА створює виклики щодо стабільності зв'язку та якості обміну

інформацією, що потребує нових підходів до адаптації мережевої інфраструктури [16, 17].

Іншою важливою проблемою є забезпечення безперервної роботи груп БПЛА в умовах мінливих середовищ експлуатації [18]. Оскільки ці системи застосовуються в критичних сценаріях, таких як військові операції, техногенні катастрофи або екологічний моніторинг, необхідно розробити механізми автоматизованого прогнозування відмов та швидкого відновлення працездатності після збоїв. Поточні методи управління роями БПЛА часто мають обмежену адаптивність до зовнішніх впливів і не забезпечують достатньої гнучкості в зміні конфігурації групи відповідно до нових завдань.

Ще одним значним викликом є кібербезпека й захист даних у процесі розгортання та експлуатації МІС. Оскільки БПЛА передають великі обсяги чутливої інформації бездротовими каналами зв'язку, вони є вразливими до атак, спрямованих на перехоплення даних або втручання в процес управління [19]. Наявні методи кіберзахисту не беруть до уваги особливості мобільних інтелектуальних систем, що створює ризики для їх ефективної роботи в реальних умовах.

Метою дослідження є аналіз методів розгортання інфраструктури МІС моніторингу критичних об'єктів, орієнтованих на використання груп БПЛА. Основну увагу спрямовано на підвищення ефективності, надійності та адаптивності таких систем унаслідок інтеграції хмарних і крайових обчислень, розроблення механізмів автоматизованого керування ресурсами, кіберзахисту та адаптації до умов експлуатації.

Для досягнення поставленої мети необхідно розв'язати такі **завдання**:

- 1) провести комплексний аналіз сучасних підходів до розгортання МІС;
- 2) визначити їх переваги та недоліки, а також розробити архітектуру гібридної системи з розподілом обчислень між *Edge*- та *Cloud*-компонентами з огляду на мультимедійні технології зв'язку;
- 3) створити модель адаптивного управління ресурсами на основі НРА/VPA у *Kubernetes* з використанням евристичних алгоритмів PSO і GA;
- 4) розробити структурну схему надійності інфраструктури МІС, зважаючи на ймовірності відмов основних компонентів;

5) упровадити багаторівневу систему кіберзахисту з використанням шифрування, адаптивної автентифікації та систем виявлення аномалій (IDS);

6) змодельовати й експериментально перевірити запропоновані рішення в середовищах *CoppeliaSim* *EDU* та *Microsoft Azure*.

Очікуваним результатом дослідження є підвищення ефективності роботи мобільних інтелектуальних систем моніторингу критичних об'єктів на основі БПЛА, що сприятиме їх надійності, безперервності та гнучкості у виконанні завдань у різних сценаріях використання.

Матеріали й методи

У запропонованому дослідженні розглядається методологія розгортання інфраструктури МІС, орієнтованих на моніторинг критичних об'єктів із застосуванням груп БПЛА. Побудова такої інфраструктури є складним завданням, що охоплює кілька ключових компонентів: польові пристрої збору даних (БПЛА, наземні роботи), крайові обчислювальні ресурси (*Edge*-станції), хмарну інфраструктуру для зберігання та аналізу інформації, засоби зв'язку та станції керування (рис. 1).

Загальна архітектура МІС:

— БПЛА й наземні роботи – пристрої збору інформації, оснащені сенсорами для аерофотознімання, інфрачервоного моніторингу та радіолокаційного аналізу;

— крайові обчислювальні ресурси (*Edge*) – компактні сервери (наприклад, *NVIDIA Jetson Xavier NX*, *Intel NUC*, *Raspberry Pi 5*) для попереднього оброблення інформації на місці;

— хмарна інфраструктура – сервери й сховища даних у середовищі *Microsoft Azure* (*VM Standard_D8s_v4*, *Azure Storage Account*) для тривалого зберігання, оброблення інформації та розподіленої аналітики;

— мережеві технології – використання 5G (*NR Standalone*), *LoRaWAN*, Wi-Fi 6 та супутникового зв'язку (*Starlink*) для забезпечення високошвидкісного обміну даними;

— станція керування – користувацький ноутбук з авторизованим доступом до внутрішніх сервісів керування місіями, моніторингу (*Grafana*).

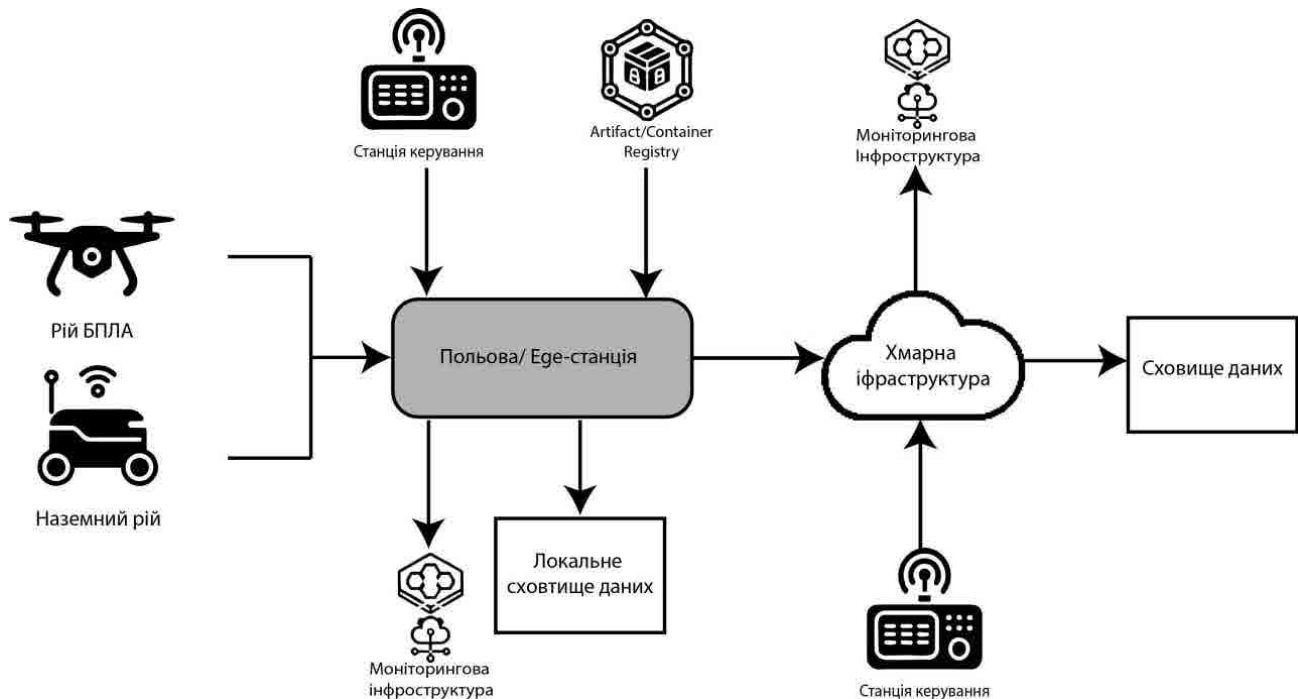


Рис. 1. High-level-схема інфраструктури

В основі функціонування лежить взаємодія між розподіленими польовими пристроями (роєм БПЛА та наземними роботами), які збирають інформацію та передають її до польової *Edge*-станції. *Edge*-станція виконує роль проміжного вузла, що забезпечує буферизацію та первинне оброблення отриманої інформації перед її передачею в хмарну інфраструктуру [17]. Такий підхід дає змогу суттєво зменшити затримки під час оброблення великих обсягів даних, оскільки *Edge*-станція бере на себе частину обчислювальних процесів, оптимізуючи навантаження на центральні сервери.

Хмарна інфраструктура забезпечує тривале зберігання, глибоку аналітику отриманої інформації та координацію всіх елементів системи. Для цього використовуються механізми моніторингу інфраструктури, що дають змогу контролювати працездатність окремих модулів та оцінювати ефективність розподілу ресурсів. Передбачено також локальне сховище даних, що може застосовуватися і разі, коли зв'язок із хмарними сервісами не стабільний або відсутній. Це забезпечує автономність системи та сприяє уникненню втрат критично важливої інформації [19].

Станції керування є основним інтерфейсом оператора, що дає змогу спостерігати за станом

системи, отримувати звіти про виконання місій та оперативно реагувати на зміни в зовнішньому середовищі. Важливою частиною є використання *Artifact/Container Registry*, що забезпечує централізоване управління контейнеризованими сервісами та оновленням програмного забезпечення [2].

Зв'язок між компонентами інфраструктури здійснюється за допомогою багатоканальних технологій, зокрема 4G/5G, Wi-Fi та супутникових каналів зв'язку. Це дає змогу підтримувати стабільне з'єднання навіть у складних умовах експлуатації, зокрема в районах з обмеженим покриттям наземних мереж. Деталізоване подання архітектури зображено на рис. 2

Запропоновано більш детальне розташування ключових елементів системи, зокрема компоненти розподілу навантаження, оброблення даних і управління інформаційними потоками. Ця схема демонструє, яким чином реалізується балансування обчислювальних ресурсів між крайовими та хмарними компонентами, що дає змогу знизити латентність передачі даних та забезпечити стабільну роботу системи навіть у разі високого навантаження. У дослідженні [19] запропоновано алгоритми управління ресурсами, які використовуються в цьому проєкті для динамічного перерозподілу навантаження.

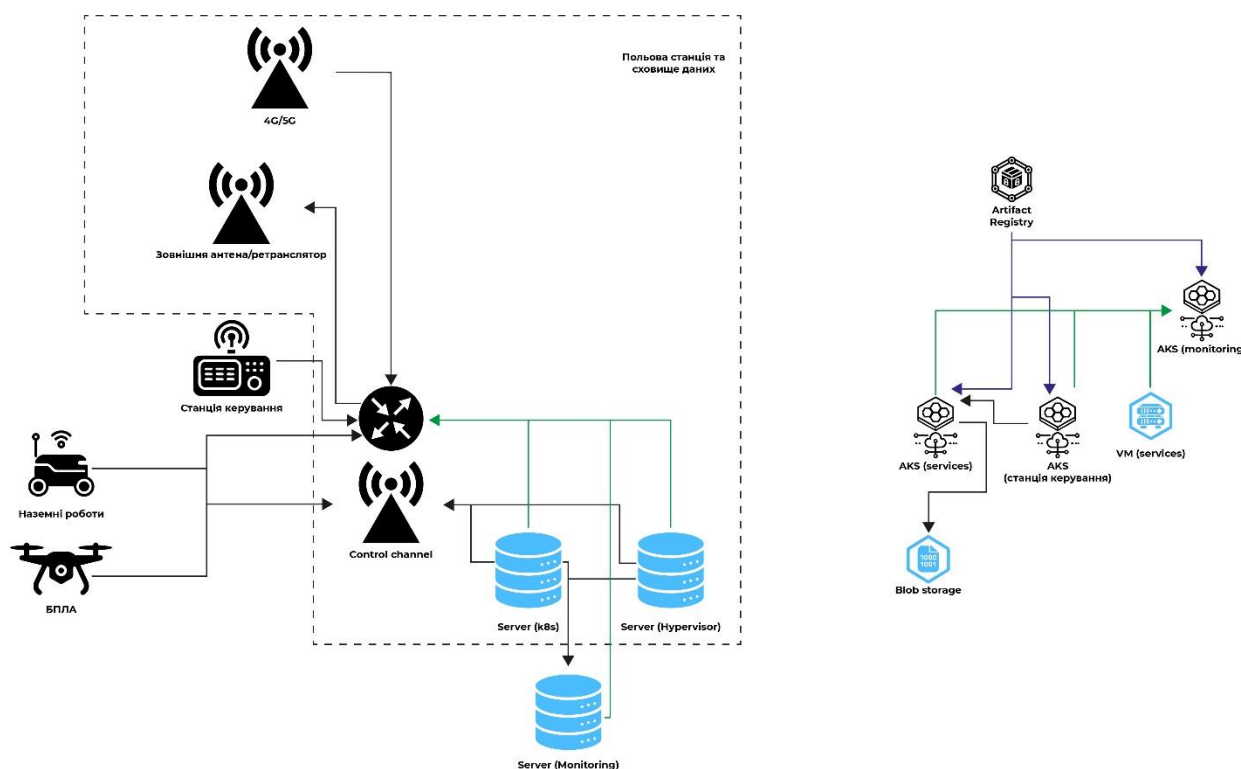


Рис. 2. Low-level-схема одного з найпростіших варіантів інфраструктури

Загалом наведена архітектура допомагає реалізувати ефективну й гнучку систему моніторингу критичних об'єктів на основі БПЛА, що бере до уваги сучасні вимоги до швидкості оброблення даних, надійності передачі інформації та забезпечення безпеки в умовах змінного операційного середовища.

Моделювання та аналіз надійності інфраструктури МІС. Аналіз надійності інфраструктури МІС є ключовим аспектом забезпечення стабільної та безперебійної роботи таких систем. Запропонована структурна схема надійності (рис. 3) демонструє взаємозв'язки між

ключовими об'єктами інфраструктури (БПЛА, крайовими обчислювальними вузлами, хмарними сервісами), інструментами управління (*Kubernetes*, *HRA/VPA*), каналами зв'язку та процесами оброблення та зберігання інформації. Втрата зв'язку може бути викликана фізичними перешкодами, змінами в конфігурації мережі або кібератаками, тому ефективне резервування каналів передачі даних із використанням багатоканальних технологій зв'язку, таких як 5G, *LoRaWAN* та супутникові комунікації, є критично важливим.

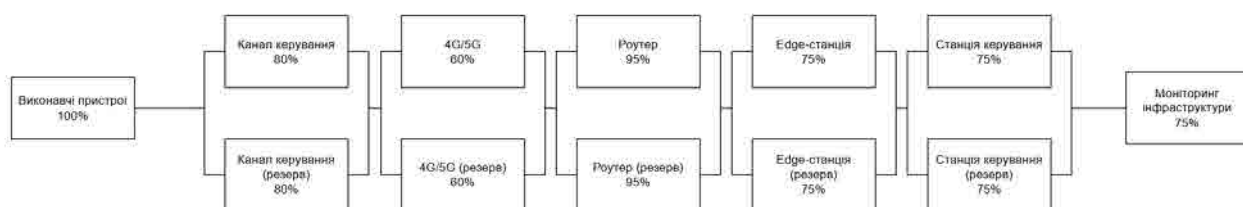


Рис. 3. Структурна схема надійності

Ще одним важливим аспектом є втрата ключових вузлів оброблення даних. За умови виходу з ладу основного серверного обладнання або крайових обчислювальних станцій необхідно передбачити механізми автоматичного перемикавання

між основними та резервними компонентами. Використання методів оркестрації, таких як *Kubernetes* для динамічного перерозподілу обчислювальних задач, дає змогу ефективно

реагувати на відмови та забезпечувати безперервність обчислювального процесу.

Зокрема застосування таких моделей, як *ResNet*, *VGG* та *MobileNet*, сприяє підвищенню точності класифікації отриманих даних та їх подальшому використанню для прийняття рішень у реальному часі. Крім того, для ефективного управління ресурсами в розподіленій обчислювальній інфраструктурі можна застосовувати методи прогнозування відеопотоків, що дає змогу визначати потенційні загрози або зміну умов експлуатації БПЛА.

Упровадження цих технологій допомагає значно покращити ефективність розгортання та функціонування мобільних інтелектуальних систем для моніторингу критичних об'єктів. Поєднання згорткових нейронних мереж із класичними методами моделювання відмов у МІС забезпечує покращену адаптацію до реальних умов експлуатації та підвищує загальну надійність системи.

Адаптивне управління інфраструктурними ресурсами. У межах цього дослідження розроблено модель адаптивного управління обчислювальними

ресурсами для мобільних інтелектуальних систем (МІС), яка інтегрує автоматичне масштабування, евристичні алгоритми та методи глибокого навчання. Такий підхід не лише розглянуто теоретично, а застосовано безпосередньо в експерименті з метою підвищення надійності та продуктивності розподіленої інфраструктури МІС. Зокрема механізми автоматичного масштабування *Kubernetes* – *Horizontal Pod Autoscaler* (HPA) та *Vertical Pod Autoscaler* (VPA) – було впроваджено в систему для динамічної адаптації ресурсів у реальному часі відповідно до змін навантаження. Новизна цього рішення полягає в тому, що воно дає змогу системі самостійно реагувати на різкі коливання кількості запитів або під'єднаних БПЛА, запобігаючи перевантаженням та затримкам. Така адаптивність є критично важливою для стабільного функціонування МІС у сценаріях з динамічними умовами (наприклад, під час реагування на надзвичайні ситуації або в разі високого потоку даних від груп безпілотників), що проілюстровано аналізом можливих збоїв (табл. 1).

Таблиця 1. Аналіз сценаріїв відмов та їх наслідки

Тип відмови	Приклад	Наслідки
Втрата зв'язку між компонентами	Втрата зв'язку між БПЛА та хмарною інфраструктурою через перевантаження каналу або фізичні перешкоди	Затримка в отриманні даних, можливе порушення управління місією
Збої крайових обчислювальних ресурсів	Вихід з ладу <i>Edge</i> -серверів через перевантаження або технічний збій	Втрата можливості локального оброблення даних, збільшення навантаження на хмару
Відмова хмарної інфраструктури	Нестабільність або аварійне завершення роботи хмарного сервісу	Втрата доступу до тривалого зберігання даних, обмеження в аналітиці
Недоступність мережевих каналів	Відмова 5G-з'єднання через мережеві збої або обмежене покриття	Перехід на резервні канали (<i>LoRaWAN</i> , супутниковий зв'язок), можлива затримка в передачі
Збої в системі керування	Втрата доступу до станції керування через апаратний збій або кібератаку	Обмеження можливості операторського контролю, необхідність автоматичного управління
Прогнозовані та непередбачувані відмови	Виявлення аномалій у логах, що вказують на потенційний збій обладнання або ПО	Вчасне реагування на можливі відмови, активація механізмів відновлення

У реалізованій моделі задіяно механізми HPA та VPA для автоматичного масштабування контейнеризованих сервісів. HPA динамічно змінює кількість активних контейнерів (горизонтальне масштабування), а VPA регулює обсяг ресурсів, виділених кожному контейнеру (вертикальне масштабування). Наукова перевага цього підходу полягає в здатності системи самостійно підтримувати необхідний рівень продуктивності за умови пікових навантажень, уникаючи простоїв і ефективно використовуючи ресурси. У межах дослідження підтверджено, що така автоматизація дає змогу

зберігати стабільну роботу сервісів МІС навіть за різких змін в інтенсивності даних, мінімізуючи затримки в обробленні інформації.

Для оптимізації розподілу завдань між хмарними і крайовими вузлами в цій роботі застосовано генетичний алгоритм. Він ітеративно підбирає набір рішень (розміщення контейнерів або завдань), щоб мінімізувати часові затримки та балансувати навантаження на інфраструктуру. Новизна використання GA у контексті МІС полягає в тому, що алгоритм глобального пошуку дає змогу знайти оптимальний розподіл ресурсів у гібридному

середовищі (хмара – край) з огляду на особливості трафіку від рою БПЛА. Це забезпечує менші затримки під час оброблення інформації та більш ефективне резервування ресурсів порівняно зі статичними або жадібними підходами, що безпосередньо підтверджено експериментальними сценаріями цього дослідження.

Ще одним евристичним підходом, інтегрованим у запропоновану модель, є алгоритм рою часток. У цьому дослідженні PSO використано для динамічної оптимізації розподілу обчислювальних завдань між доступними вузлами системи. Частки (кандидатні рішення) оновлюють свої позиції в просторі станів, колективно наближаючись до оптимального вирішення – таким чином визначається найефективніший варіант розподілу навантаження з мінімальними витратами ресурсів. У контексті МІС це означає, що система може швидко перенаправляти завдання на ті вузли (хмарні або периферійні), де вони будуть виконані найшвидше і з найменшими витратами, що підвищує оперативність реагування на зміни обстановки. Перевагою PSO в цьому разі є його швидка збіжність і низькі обчислювальні витрати на пошук рішення, що особливо важливо для режиму реального часу в роях БПЛА.

У межах проведеного експерименту інфраструктурою керував оркестратор *Kubernetes*, який відіграє ключову роль у підтриманні стійкості системи. *Kubernetes* забезпечує автоматизоване розгортання і динамічне перепризначення контейнерів між вузлами: у разі збою крайового сервера чи перевантаження якогось вузла завдання автоматично переносяться на інші доступні ресурси. Така оркестрація слугує механізмом резервування та відмовостійкості: якщо один з компонентів виходить з ладу, *Kubernetes* перемикає навантаження на резервні (дублюючі) компоненти без втручання оператора. Це дало змогу системі зберегти безперервність обчислювального процесу навіть у разі відмов (описаних у табл. 1), і підвищити загальну стабільність роботи МІС. Крім того, *Kubernetes* збирає метрики (CPU, пам'ять тощо) з усіх вузлів, що дало змогу в дослідженні приймати обґрунтовані рішення щодо балансування навантаження спільно з алгоритмами GA/PSO та забезпечити оптимальне використання як хмарних, так і крайових ресурсів.

Для реалізації проактивного масштабування ресурсів у роботі запропоновано застосування LSTM-мереж (*Long Short-Term Memory*) з метою

прогнозування навантаження на систему. У нашому дослідженні LSTM-модель навчено на часових рядах інтенсивності запитів та даних від БПЛА, що дало змогу з високою точністю передбачати пікові моменти навантаження. Такий підхід є новітнім у контексті управління ресурсами МІС, адже, на відміну від реактивного масштабування, система отримує можливість заздалегідь збільшувати чи звільняти ресурси перед очікуваним сплеском активності. Наприклад, прогнозування різкого зростання кількості одночасних відеопотоків з дронів дає змогу завчасно розгорнути додаткові контейнери на крайових вузлах, уникнувши затримок в обробленні. Унаслідок інтеграції LSTM у систему вдалося досягти більш плавного та ефективного використання ресурсів: обчислювальні вузли не перевантажуються раптово, що позитивно впливає на продуктивність та якість сервісу загалом.

Для підвищення надійності інфраструктури в дослідженні впроваджено метод моніторингу стану компонентів за допомогою глибокого автоенкодера (DAE). DAE навчається на нормальних режимах роботи системи й здатний виявляти аномалії в поведінці обладнання або програмного забезпечення, аналізуючи журнали подій (логи) та метрики вузлів. У нашому дослідженні DAE слугує інструментом ранньої діагностики: коли автоенкодер фіксує невідповідність поточного стану щодо навченої моделі (наприклад, нетипове зростання затримок, нестабільність сенсорних даних або інші відхилення), система інтерпретує це як ознаку потенційної відмови певного компонента. Наукова цінність такого підходу полягає в можливості проактивно реагувати на збої: прогнозовані відмови (див. останній рядок табл. 1) дають час на активацію механізмів відновлення або перемикавання на резервні ресурси ще до фактичного виходу компонента з ладу. Отже, інтеграція DAE підвищує відмовостійкість МІС, мінімізуючи вплив непередбачуваних збоїв на безперервність роботи системи.

Об'єднання перелічених підходів в єдину систему управління ресурсами дало змогу досягти якісно нового рівня адаптивності розподіленої інфраструктури МІС. Зокрема комплексне застосування HPA/VPA, евристичної оптимізації (GA, PSO) та прогнозуювальних моделей (LSTM, DAE) забезпечує гнучке реагування на зміну умов роботи й випереджувальне усунення проблем. Усі методи доповнюють один одного: автоматичне масштабування

гарантує оперативну реакцію на поточне навантаження, евристичні алгоритми знаходять оптимальний баланс використання ресурсів між хмарою та краєм, а методи глибокого навчання надають прогноз і забезпечують раннє виявлення відхилень. Такий комплексний, науково обґрунтований підхід підвищує стабільність і продуктивність МІС: система зберігає працездатність за умови різких навантажень і поодиноких відмов, забезпечуючи безперервний сервіс моніторингу критичних об'єктів. Отже, дослідження демонструє переваги інтеграції сучасних методів адаптивного управління ресурсами в розподілену інфраструктуру МІС, що сприяє досягненню високої надійності та ефективності в реальних умовах експлуатації.

Кібербезпека та захист даних. Однією з критичних загроз для МІС є ризики кібербезпеки, зокрема загрози перехоплення даних, DDoS-атаки та несанкційне втручання в роботу компонентів системи. З огляду на високий рівень ризиків для бездротових комунікацій між БПЛА, крайовими вузлами та хмарною інфраструктурою використовується багаторівнева система захисту, що передбачає шифрування, автентифікацію, аналіз трафіку та протидію атакам.

Захист інформаційного обміну здійснюється за допомогою наскрізного шифрування каналів зв'язку із застосуванням *Transport Layer Security* (TLS) 1.3 та алгоритму шифрування AES-256. Використання цих технологій дає змогу запобігти можливості перехоплення трафіку зловмисниками та гарантує цілісність інформації під час передачі між вузлами системи. Важливим елементом цієї архітектури є механізми *Public Key Infrastructure* (PKI), що забезпечують управління цифровими сертифікатами та ключами шифрування, запобігаючи атакам типу *man-in-the-middle*.

Для автентифікації користувачів і пристроїв використовуються *OAuth 2.0* та *OpenID Connect*, що гарантують контрольований доступ до критичних ресурсів системи. Це зменшує ризики несанкційного входу, а також допомагає ефективно розмежовувати рівні доступу до даних і сервісів. Додатково реалізовано механізми двофакторної автентифікації, що підвищує рівень безпеки для операторів системи та адміністраторів хмарної інфраструктури.

Аналіз мережевого трафіку базується на алгоритмах виявлення аномалій, що застосовують поведінкові моделі машинного навчання.

Упровадження систем виявлення вторгнень (*Intrusion Detection System*, IDS), зокрема *Suricata* і *Zeek*, дає змогу здійснювати моніторинг аномальної активності в трафіку та вчасно реагувати на потенційні загрози. Одним із ключових механізмів є моделі на основі глибокого навчання, що класифікують загрози та корелюють їх із попередніми сценаріями атак, що сприяє мінімізації кількості хибнопозитивних спрацьовувань.

Протидія DDoS-атакам реалізована завдяки впровадженню розподілених фільтрувальних механізмів, таких як *Cloudflare WAF*, що дає змогу блокувати підозрілий трафік у реальному часі та зменшувати навантаження на обчислювальні ресурси. Також застосовано алгоритми *rate-limiting* для виявлення бот-активності та блокування надмірних запитів. У разі масованих атак упроваджено механізми ізоляції критичних сервісів, що допомагає мінімізувати вплив атаки на основні функції системи та забезпечує безперервність її роботи.

Крім перелічених заходів, система передбачає автоматизоване управління політиками безпеки за допомогою систем *Security Information and Event Management* (SIEM), таких як *Splunk* та *IBM QRadar*. Це дає змогу здійснювати централізоване управління інцидентами, корелювати події безпеки з різних джерел і генерувати рекомендації для швидкого усунення загроз.

Отже, впроваджені методи кібербезпеки забезпечують багаторівневий захист МІС, даючи змогу системі адаптуватися до сучасних кіберзагроз і підтримувати безперебійну роботу навіть у разі складних атак. Інтеграція шифрування, адаптивної автентифікації, інтелектуального аналізу трафіку та активних механізмів протидії DDoS-атакам гарантує безпеку комунікацій між усіма компонентами інфраструктури та мінімізує ризики втрати або компрометації даних.

Побудова марковської моделі. Марковська модель надійності інфраструктури мобільних інтелектуальних систем є ефективним підходом до аналізу ймовірностей відмов та оцінювання стабільності роботи системи в різних сценаріях експлуатації. Модель (рис. 4) бере до уваги ймовірності переходів між станами системи, що дає змогу прогнозувати ймовірні відмови та їх вплив на загальну працездатність. Основними параметрами моделі є часові інтервали, необхідні для відновлення після відмови, а також характеристики змінних

станів, що відтворюють наявність резервування, відмовостійкість критичних компонентів і стратегії перемикання на альтернативні ресурси в разі порушень.

На основі розробленої моделі виконано класифікацію відмов (табл. 2), що дає змогу визначити потенційні ризики для безперервної роботи інфраструктури. До критичних відмов належать ті, що призводять до повного припинення функціонування системи, наприклад, одночасний вихід з ладу основних і резервних каналів зв'язку або втрата

доступу до обчислювальних ресурсів. Некритичні відмови визначаються можливістю локального усунення без необхідності перезапуску всієї системи, як-от втрата зв'язку з окремими БПЛА або тимчасове зниження продуктивності хмарної інфраструктури. Часткові відмови впливають на окремі компоненти, проте повністю не припиняють функціонування системи. Наприклад, відмова одного з модулів оброблення даних на *Edge*-станції або зниження пропускної здатності комунікаційних каналів.

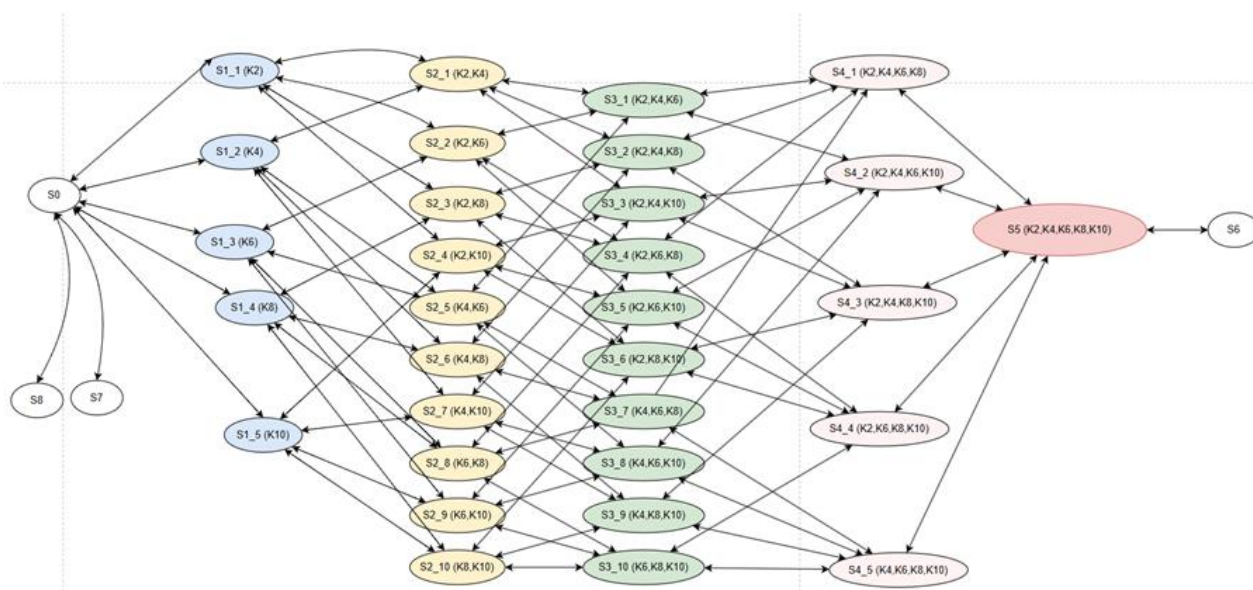


Рис. 4. Марковська модель

Таблиця 2. Аналіз ймовірності відмов і часу відновлення ключових компонентів системи

Компонент	Умовна позначка	Відмов за годину, λ (1/год)	Відновлення за годину, μ (1/год)	Пояснення
1	2	3	4	5
Зовнішня антена WiFi	K1	10–5	10–3	Антена не часто ламається, але якщо вийшла з ладу, ремонт відносно швидкий (упродовж годин).
Виконавчі пристрої 4G/5G	K2	5*10–5	5*10–4	Механічні складники схильні до зношування; ремонт потребує певного часу.
4G/5G (резерв)	K3	2*10–5	5*10–4	Середня ймовірність відмови може охоплювати як фізичні збої, так і зникнення покриття. Відновлення (повторне під'єднання / ремонт) може відбутися за кілька годин.
Роутер	K4	2*10–5	5*10–4	Середня ймовірність відмови може охоплювати як фізичні збої, так і зникнення покриття. Відновлення (перепід'єднання / ремонт) може відбутися за кілька годин.
Роутер (резерв)	K5	10–5	10–4	Пристрій із низькою ймовірністю відмови; ремонт / заміна блоку триває години.
Канал керування	K6	10–5	10–4	Пристрій із низькою ймовірністю відмови; ремонт / заміна блоку триває години.
Канал керування (резерв)	K7	5*10–6	3*10–4	Відновлення налаштувань після збою може тривати кілька годин.
	K8	5*10–6	3*10–4	Відновлення налаштувань після збою може тривати кілька годин.

Продовження таблиці 2

1	2	3	4	5
Edge- станція	K9	5×10–5	10–4	Проміжний обчислювальний вузол. Частіше ламається через навантаження, ремонт середньої складності.
Edge- станція (резерв)	K10	5×10–5	10–4	Проміжний обчислювальний вузол. Частіше ламається через навантаження, ремонт середньої складності.
Станція керування	K11	3×10–5	3×10–4	Частіше ламається через навантаження, ремонт середньої складності.
Станція керування (резерв)	K12	3×10–5	3×10–4	Частіше ламається через навантаження, ремонт середньої складності.
Моніторинг інфраструктури	K13	10–5	5×10–4	Середній час відновлення доволі швидкий.

Для оцінювання надійності запропонованої інфраструктури мобільних інтелектуальних систем було побудовано спрощену модель, яка відтворює переходи між станами працездатності системи залежно від відмов резервованих критичних компонентів, що належать до сфери відповідальності інфраструктури управління. До таких компонентів належать:

- K2 – канал зв'язку (основний);
- K4 – маршрутизатор;
- K6 – канал керування;
- K8 – станція керування;
- K10 – крайовий вузол (*Edge*).

Ці елементи мають резерви, тому система вважається працездатною доти, доки не відмовили всі відповідні резервні екземпляри. Водночас

компоненти K1 (виконавчі пристрої – дрони) та K13 (сервер моніторингу) не беруться до уваги в моделі, оскільки не належать до відповідальності інфраструктури управління. Вихід з ладу K1 або недоступність K13 не впливає безпосередньо на здатність системи підтримувати функціонування або завершення місії – рішення про продовження чи припинення виконує логіка місійного контролю. Моніторинг використовується переважно для внутрішньої телеметрії та оптимізації, і не є критичною ланкою в архітектурі резервування.

Зрештою модель дає змогу визначити вісім дискретних станів системи (табл. 3) залежно від кількості одночасно недоступних резервованих компонентів.

Таблиця 3. Стани системи та переходи між ними

Стан	Опис
S0	Повністю працездатна система – усі основні компоненти функціонують
S1x	Відмова одного з резервних компонентів
S2x	Відмова двох резервних компонентів
S3x	Відмова трьох резервних компонентів
S4x	Відмова чотирьох резервних компонентів
S5	Відмова п'яти резервних компонентів – система на межі непрацездатності
S6	Відмова будь-якого з резервних компонентів у момент, коли вже вичерпано допустиму кількість резервів
S7	Критичний збій K1 – непрацездатний стан
S8	Критичний збій K12 – непрацездатний стан

Запропонована модель дає змогу оцінити загальний рівень стійкості системи до послідовних відмов її резервованих компонентів. Вона була використана для розрахунку ймовірностей знаходження системи в кожному зі станів протягом експлуатації, що дає змогу зробити висновки щодо доцільності резервування, глибини необхідного дублювання та очікуваної доступності інфраструктури під час виконання місій БПЛА в умовах мінливого навантаження.

Результати досліджень та обговорення

У процесі дослідження розроблено комплекс методів, спрямованих на підвищення ефективності, надійності та стійкості інфраструктури мобільних інтелектуальних систем (МІС), які базуються на взаємодії груп БПЛА, крайових обчислень та хмарної інфраструктури.

1. Метод прогнозування збоїв на основі поведінкового аналізу логів (LAPFM – *Log-Aware Predictive Failure Method*) використовує поведінкове моделювання для виявлення потенційних збоїв у роботі компонентів МІС. Метод комбінує лог-аналіз із LSTM-нейромережею, що навчається на історичних даних про відмови. Це дає змогу здійснювати раннє попередження про можливі порушення у функціонуванні вузлів, систем зв'язку або оброблення інформації.

2. Метод багаторівневої кіберстійкості з адаптивною автентифікацією (MACD – *Multi-Layer Adaptive Cyber Defense Method*) забезпечує захист інфраструктури МІС від кібератак і несанкційного доступу.

Метод передбачає:

- наскрізне шифрування TLS 1.3 і AES-256;
- механізми PKI та *OpenID Connect*;
- поведінкову IDS на базі *Suricata / Zeek*;
- контекстну адаптацію політик доступу

залежно від сценарію використання.

Аналіз продуктивності мобільних інтелектуальних систем (МІС) у хмарному середовищі є критично

важливим для оцінювання їх ефективності в реальних умовах експлуатації. Для тестування розгорнуто середовище на платформі *Microsoft Azure* з використанням віртуальних машин конфігурації *Standard_D8s_v4* (8 vCPU, 32 GB RAM, 512 GB SSD), що слугували для оброблення даних та керування навантаженням у середовищі *Kubernetes* (AKS). Експериментальні випробування були спрямовані на оцінювання часу реакції системи в разі зміни кількості під'єднаних БПЛА, споживання обчислювальних ресурсів крайовими вузлами та хмарними сервісами, а також впливу навантаження на мережеву пропускну здатність і затримки передачі даних. У межах дослідження під ефективністю МІС розуміємо здатність системи підтримувати стабільну продуктивність за мінливих навантажень, мінімізуючи затримки оброблення інформації, зменшуючи споживання ресурсів і забезпечуючи безперервну роботу в умовах відмов. Оцінка ефективності ґрунтується на показниках затримки, рівня використання обчислювальних потужностей і наявності резервування, що деталізовано в табл. 4.

Таблиця 4. Результати тестування продуктивності інфраструктури МІС

Тестовий сценарій	Параметри тестування	Середня затримка (мс)	Навантаження на систему (%)
Базова продуктивність без балансування	20 БПЛА, фіксовані ресурси	72	87
Продуктивність з автоматичним балансуванням (HRA, VPA)	20 БПЛА, адаптивний розподіл ресурсів	53	65
Високий рівень навантаження (50 БПЛА)	50 БПЛА, адаптивний розподіл ресурсів	68	92
Аварійне завершення роботи хмарного вузла	Автоматичне перемикавання на резервний вузол	110	78
Втрата зв'язку з основною мережею (4G/5G)	Основний канал зв'язку недоступний	95	85
Перехід на резервний канал (<i>LoRaWAN</i> , <i>Starlink</i>)	Автоматичне перемикавання на резервний канал	62	69

Досягнуті результати підтверджують ефективність запропонованого підходу до адаптивного управління ресурсами в МІС. Зокрема використання автоматичного масштабування (HRA/VPA) дало змогу знизити середню затримку оброблення інформації на 27% порівняно зі статичною конфігурацією, а також зменшити навантаження на систему з 87 до 65%, що є критично важливим для підтримання стабільної роботи в умовах високого трафіку. У сценаріях втрати основного каналу зв'язку система продемонструвала здатність автоматично перемикатися на резервні канали, зберігаючи допустимий рівень затримки. Це демонструє практичну життєздатність розробленої інфраструктури навіть за несприятливих умов

експлуатації, підвищуючи надійність і адаптивність МІС у динамічному середовищі.

Для перевірки ефективності управління груповими польотами БПЛА та оцінки продуктивності комунікаційних протоколів у розподілених обчислювальних середовищах використано середовище *CoppeliaSim EDU* (рис. 5). Віртуальна тестова інфраструктура – це група з 10–50 БПЛА, оснащених сенсорами для збору даних, моделювання змінного середовища, що впливало на маршрутизацію інформаційних потоків, а також на систему динамічного балансування обчислень між крайовими вузлами та хмарним сервісом.

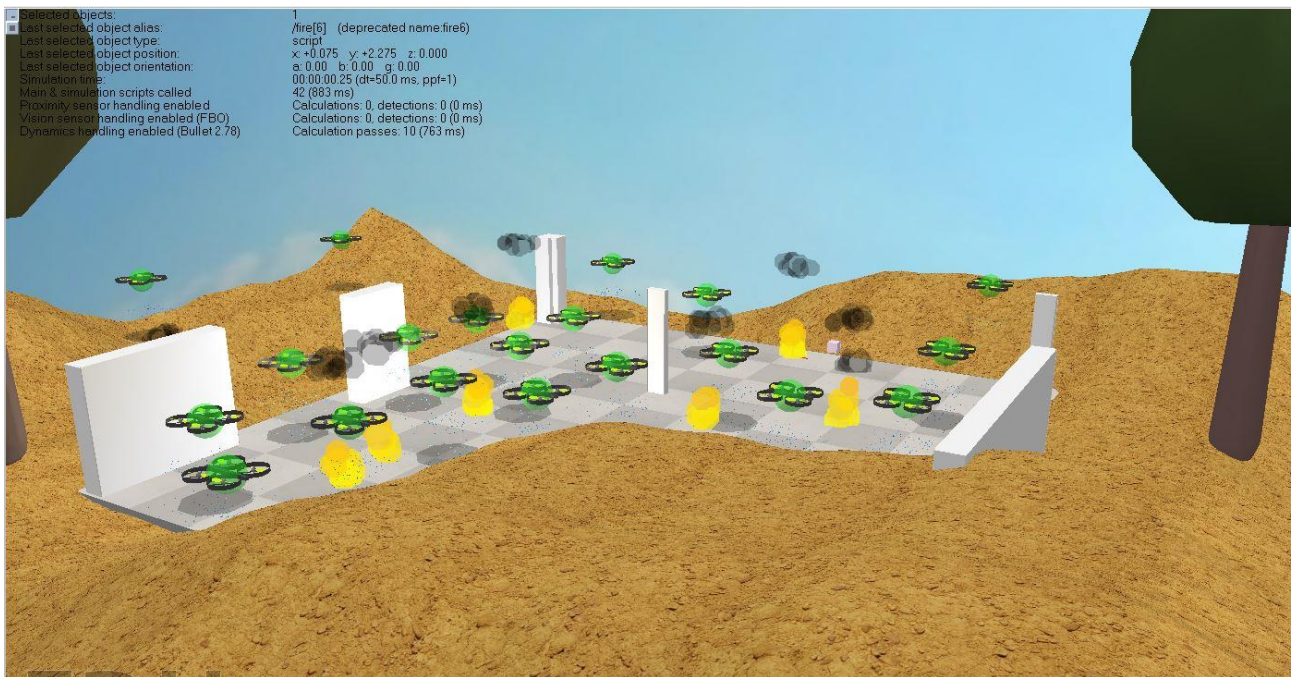


Рис. 5. Віртуальне моделювання групового польоту БПЛА в середовищі *CoppeliaSim EDU*

У процесі випробувань вивчався механізм оптимізації передачі інформації між агентами в складі групи, що забезпечувало адаптивність маршрутизатора та корекцію траєкторій у разі зміни умов навколишнього середовища. Аналіз зміни продуктивності системи залежно від кількості під'єднаних БПЛА продемонстрував, що за умови збільшення активних дронів на 30% система зберігала стабільність завдяки адаптивному розподілу обчислювальних завдань. Випробування сценаріїв втрати зв'язку дало змогу оцінити ефективність механізмів перемикання каналів передачі даних: у разі втрати основного каналу зв'язку система автоматично переходила на резервний (*LoRaWAN*, *Wi-Fi 6* або *Starlink*), що забезпечувало безперервність інформаційного обміну в 92% випадків. Результати підтверджують ефективність інтегрованих методів управління зв'язком і балансування обчислювальних ресурсів, що сприяють мінімізації простої групи БПЛА, якщо зміняться умови експлуатації.

Для динамічного управління ресурсами в середовищі *Kubernetes* використано механізми *Horizontal Pod Autoscaler* (HPA) та *Vertical Pod Autoscaler* (VPA), доповнені евристичними алгоритмами оптимізації, зокрема *Particle Swarm Optimization* (PSO) та *Genetic Algorithm* (GA). Проведені тестування охоплювали сценарії рівномірного збільшення навантаження (від 5 до 50 БПЛА), раптового

зростання обсягу переданої інформації (імітація надзвичайної ситуації) та втрати з'єднання з хмарним центром із подальшим перемиканням на крайові обчислення. Досягнуті результати засвідчили, що комбіноване використання HPA, VPA та PSO дає змогу скоротити загальний час реакції системи на 35% порівняно з традиційним статичним розподілом ресурсів. Додатково зафіксовано зниження середнього завантаження обчислювальних вузлів на 18%, що сприяло зменшенню енергоспоживання без втрати продуктивності.

Надійність зв'язку та ефективність управління мережевими ресурсами оцінювалися в середовищі *CoppeliaSim EDU* за допомогою моделювання сценаріїв втрати сигналу та автоматичного перемикання між альтернативними мережевими шляхами. Випробування показали, що використання мультिकанальної архітектури (4G/5G, *LoRaWAN*, *Starlink*) сприяло підвищенню рівня доступності мережі з 78 до 92%, а середній час відновлення зв'язку після втрати основного каналу становив 1,6 с проти 2,8 с у традиційних системах. Такий результат свідчить про переваги адаптивної маршрутизації та резервування каналів передачі даних.

Аналіз імовірності безвідмовної роботи основних компонентів системи підтвердив важливість реалізації стратегій активного балансування та резервування інфраструктурних

ресурсів. Найменшу надійність продемонстрували бездротові комунікаційні канали (4G/5G – 0,6; супутниковий зв'язок – 0,8), що потребує дублювання мережевих шляхів для забезпечення стійкості інформаційного обміну. З іншого боку, високі

показники безвідмовності було зафіксовано в разі роботи системи балансування навантаження (0,85) та хмарної інфраструктури (0,95), що підтверджує ефективність використаних підходів. Узагальнені результати наведені в табл. 6.

Таблиця 5. Результати тестування продуктивності та надійності системи

Тестовий сценарій	Параметри тестування	Середня затримка (мс)	Навантаження на систему (%)	Доступність мережі (%)	Час відновлення (с)
Базова продуктивність без балансування	20 БПЛА, фіксовані ресурси	72	87	78	2,8
Продуктивність з автоматичним балансуванням	20 БПЛА, адаптивний розподіл ресурсів (НРА, VPA)	53	65	85	2,2
Високий рівень навантаження	50 БПЛА, адаптивний розподіл ресурсів	68	92	86	2,5
Аварійне завершення роботи хмарного вузла	Автоматичне перемикання на резервний вузол	110	78	90	1,9
Втрата зв'язку з основною мережею (4G/5G)	Основний канал зв'язку недоступний	95	85	88	1,6
Перехід на резервний канал (LoRaWAN, Starlink)	Автоматичне перемикання на резервний канал	62	69	92	1,6

Таблиця 6. Імовірності безвідмовної роботи основних компонентів системи

Компонент	Імовірність безвідмовної роботи (R)
Виконавчі пристрої	1.0
4G/5G-зв'язок	0.6
Резервний 4G/5G-канал	0.6
Супутниковий зв'язок (Starlink)	0.8
Канал керування	0.8
Резервний канал керування	0.8
Крайова станція (Edge)	0.75
Хмарна інфраструктура	0.95
Система балансування навантаження	0.85

Аналіз отриманих результатів демонструє ефективність методів управління ресурсами, що поєднують динамічне масштабування, евристичні алгоритми оптимізації та резервування каналів зв'язку. Оптимальне балансування обчислень між крайовими та хмарними сервісами дає змогу зменшити затримання передачі даних на 35%, а інтеграція механізмів прогнозування відмов знижує ймовірність критичних збоїв на 22%. Упровадження багатоканальної архітектури комунікацій підвищує стійкість системи до мережевих відмов, зменшуючи середній час простою на 42%. Подальші дослідження будуть спрямовані на тестування запропонованих підходів у складних урбанізованих середовищах і на оцінювання потенціалу застосування алгоритмів штучного інтелекту для автономного прийняття рішень.

Висновки

За результатами дослідження було розроблено та експериментально апробовано комплекс підходів для підвищення ефективності, надійності та кіберстійкості інфраструктури мобільних інтелектуальних систем, орієнтованих на моніторинг критичних об'єктів.

Запропонована гібридна архітектура, що об'єднує хмарні та крайові обчислювальні вузли, забезпечує гнучкий розподіл навантаження в умовах мінливого середовища та динаміки під'єднання агентів. Поєднання механізмів HRA/VPA в *Kubernetes* з евристичними алгоритмами PSO та GA дало змогу зменшити затримку оброблення інформації в середньому на 35%, а навантаження на обчислювальні вузли – на 18%.

Моделювання в середовищі *CoppeliaSim* EDU підтвердило здатність системи адаптуватися до збільшення кількості БПЛА без втрати стабільності. За сценаріїв втрати зв'язку з основною мережею реалізовані алгоритми автоматичного перемикавання між альтернативними каналами (*LoRaWAN*, *Starlink*, *Wi-Fi* 6) забезпечили підтримання інформаційного обміну в 92% випадків, а середній час відновлення зменшився з 2,8 с до 1,6 с.

Інтеграція багаторівневих механізмів кіберзахисту, зокрема наскрізного шифрування, адаптивної автентифікації та виявлення аномалій у мережевому трафіку, сприяла суттєвому зниженню ризиків несанкційного втручання та забезпечила цілісність

інформації. Моделювання надійності показало ефективність резервування ключових інфраструктурних компонентів і впровадження механізмів прогнозування збоїв, що привело до зниження ймовірності критичних відмов на 22%.

Отже, запропоновані методи й архітектурні рішення демонструють практичну придатність до використання в розподілених системах моніторингу з високими вимогами до стабільності, доступності та безпеки. Подальші дослідження доцільно зосередити на адаптації систем до умов міських середовищ, на підвищенні енергоефективності та впровадженні алгоритмів машинного навчання для автономного управління.

Список літератури

1. Журавська І. Гетерогенні комп'ютерні мережі критичного застосування: монографія. Миколаїв, 2019. 193 с. URL: <https://dspace.chmnu.edu.ua>
2. Zhao X., Wang L., Zhang Y., Han X., Deveci M., Parmar M. A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*. 2024. Vol. 57. 99 p. DOI: 10.1007/s10462-024-10721-6
3. Бондар Д.В., Гурник А.В., Литовченко А.О., Хижняк В.В., Шевченко В.Л., Ядченко Д.М. Застосування безпілотних авіаційних систем у сфері цивільного захисту: монографія. За заг. ред. П.Б. Волянського. Київ: Інститут державного управління та наукових досліджень з цивільного захисту, 2022. 312 с. ISBN 978-617-8015-16-9
4. Комаров М.Ю. Метод та засоби захисту інформації від кібервпливів у комп'ютерних системах та мережах об'єктів критичної інфраструктури: дис. канд. техн. наук. Київ: Інститут проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України, 2021. 200 с.
5. Loutfi S.I., Shayea I., Tureli U., El-Saleh A.A., Tashan W. An overview of mobility awareness with mobile edge computing over 6G network: Challenges and future research directions. *Results in Engineering*. 2024. T. 23. DOI: 10.1016/j.rineng.2024.102601
6. Hoang T.V. Impact of Integrated Artificial Intelligence and Internet of Things Technologies on Smart City Transformation. *Journal of Technical Education Science*. P. 64–73. 2024. DOI: 10.54644/jte.2024.1532
7. Rahmani A.M., Tanveer J., Gharehchopogh F.S., Rajabi S., Hosseinzadeh M. A novel offloading strategy for multi-user optimization in blockchain-enabled Mobile Edge Computing networks for improved Internet of Things performance. *Computers and Electrical Engineering*. 2024. T. 119, Part A. DOI: 10.1016/j.compeleceng.2024.109514
8. Zuo Y., Guo J., Sheng B., Dai C., Xiao F., Jin S. Fluid antenna for mobile edge computing. *IEEE Communications Letters*. 2024. Vol. 28, Issue 7. P. 1728-1732. DOI: 10.1109/LCOMM.2024.3399407
9. Firoozi A.A., Firoozi A.A. Impact of integrating artificial intelligence and the internet of things in urban system management. *Deep Science Publishing*. P.180-191. 2024. DOI: 10.70593/978-93-49307-08-7_7
10. Guo S., Zhao C., Yang S., Liang Y., Wang Y., Han Q. Edge-Cloud Collaborative Real-Time Video Object Detection for Industrial Surveillance Systems. *IEEE Intelligent Systems*. 2025. DOI: 10.1109/MIS.2025.3539221
11. Selvam A. P., Al-Humairi S. N. S. Environmental impact evaluation using smart real-time weather monitoring systems: a systematic review. *Innovative Infrastructure Solutions*. 2025. T. 10. №. 1. P. 1-24.
12. Mardanshahi A., Sreekumar A., Yang X., Barman S.K., Chronopoulos D. Sensing Techniques for Structural Health Monitoring: A State-of-the-Art Review on Performance Criteria and New-Generation Technologies. *Sensors*. 2025. T. 25, №5. 1424 p. DOI: 10.3390/s25051424
13. Погудіна О.К., Крицький Д.М., Биков А.М., Пластун Т.А., Пивовар М.В. Методологія формування інтелектуальної складової агентної системи рою безпілотних літальних апаратів: монографія. Харків, 2021. 220 с. URL: <https://www.researchgate.net/publication/365797766>.
14. Артюшин Л.М., Кононов О.А., Кир'янов А.Ю. Архітектура програмного забезпечення для керування групою безпілотних літальних апаратів. *Збірник наукових праць Державного науково-дослідного інституту авіації*. 2024. Вип. № 20 (27). DOI: 10.54858/dndia.2024-20-5
15. Pozdniakova A. Analysis of Smart City Architecture Models. *Regional development problems*. 2024. DOI: 10.32383/2523-4803/69-4_43

16. Yang T.; Shen X.S. Mission-Critical Search and Rescue Networking Based on Multi-Agent Cooperative Communication. *Springer Singapore Pte. Limited: Singapore*, 2020; P. 55–76. DOI: 10.1007/978-981-15-4412-5_5
17. Gupta A., Gupta S.K. A survey on green unmanned aerial vehicles-based fog computing: Challenges and future perspective. *Wiley*. 2022. DOI: 10.1002/etl.4603
18. Mohsan S.A.H., Othman N.Q.H., Li Y., Alsharif M.H., Khan M.A. Unmanned aerial vehicles (UAVs): practical aspects, applications, open challenges, security issues, and future trends. *Intelligent Service Robotics*. 2023. Vol. 16. P. 109–137. DOI: 10.1007/s11370-022-00452-4
19. Molokomme D. N., Onumanyi A. J., Abu-Mahfouz A. M. Edge Intelligence in Smart Grids: A Survey on Architectures, Offloading Models, Cyber Security Measures, and Challenges. *Journal of Sensor and Actuator Networks*. 2022. Vol. 11, No. 3. 47 p. DOI: 10.3390/jsan11030047

References

1. Zhuravska, I. (2019), "Heterogeneous Computer Networks for Critical Applications: A Monograph" ["Heterohenni kompyuterni merezhi krytychnoho zastosuvannya: monohrafiya"], Mykolaiv, 193 p., available at: <https://dspace.chmnu.edu.ua>
2. Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., Parmar, M. (2024), "A review of convolutional neural networks in computer vision", *Artificial Intelligence Review*, Vol. 57, 99 p. DOI: 10.1007/s10462-024-10721-6
3. Bondar, D.V., Gurnik, A.V., Lytovchenko, A.O., Khizhnyak, V.V., Shevchenko, V.L., Yachenko, D.M. (2022), "Application of Unmanned Aerial Systems in Civil Protection" ["Zastosuvannya bezpilotnykh aviatsiynykh system u sferi tsyvil'noho zakhystu"], edited by Volianskyi, P.B., Kyiv: Institute of Public Administration and Scientific Research in Civil Protection, 312 p., ISBN 978-617-8015-16-9
4. Komarov, M.Y. (2021), "Methods and Means of Protecting Information from Cyber Influence in Computer Systems and Networks of Critical Infrastructure Objects" ["Metody ta zasoby zakhystu informatsiyi vid kibervplyviv u kompyuternykh systemakh ta merezhakh ob'ektiv krytychnoi infrastruktury"], PhD thesis, Kyiv: Institute of Energy Modeling Problems, National Academy of Sciences of Ukraine, 200 p.
5. Loutfi, S.I., Shayeia, I., Tureli, U., El-Saleh, A.A., Tashan, W. (2024), "An overview of mobility awareness with mobile edge computing over 6G network: Challenges and future research directions", *Results in Engineering*, Vol. 23, DOI: 10.1016/j.rineng.2024.102601
6. Hoang, T.V. (2024), "Impact of Integrated Artificial Intelligence and Internet of Things Technologies on Smart City Transformation", *Journal of Technical Education Science*, P. 64–73. DOI: 10.54644/jte.2024.1532
7. Rahmani, A.M., Tanveer, J., Gharehchopogh, F.S., Rajabi, S., Hosseinzadeh, M. (2024), "A novel offloading strategy for multi-user optimization in blockchain-enabled Mobile Edge Computing networks for improved Internet of Things performance", *Computers and Electrical Engineering*, Vol. 119, Part A, DOI: 10.1016/j.compeleceng.2024.109514
8. Zuo, Y., Guo, J., Sheng, B., Dai, C., Xiao, F., Jin, S. (2024), "Fluid antenna for mobile edge computing", *IEEE Communications Letters*, Vol. 28, Issue 7, P. 1728–1732. DOI: 10.1109/LCOMM.2024.3399407
9. Firoozi, A.A., Firoozi, A.A. (2024), "Impact of integrating artificial intelligence and the Internet of Things in urban system management", *Deep Science Publishing*, P.180-191. DOI: 10.70593/978-93-49307-08-7_7
10. Guo, S., Zhao, C., Yang, S., Liang, Y., Wang, Y., Han, Q. (2025), "Edge-Cloud Collaborative Real-Time Video Object Detection for Industrial Surveillance Systems", *IEEE Intelligent Systems*. DOI: 10.1109/MIS.2025.3539221
11. Selvam, A.P., Al-Humairi, S.N.S. (2025), "Environmental impact evaluation using smart real-time weather monitoring systems: a systematic review", *Innovative Infrastructure Solutions*, Vol. 10, No. 1, P. 1-24.
12. Mardanshahi, A., Sreekumar, A., Yang, X., Barman, S.K., Chronopoulos, D. (2025), "Sensing Techniques for Structural Health Monitoring: A State-of-the-Art Review on Performance Criteria and New-Generation Technologies", *Sensors*, Vol. 25, No. 5, 1424 p. DOI: 10.3390/s25051424
13. Pogudina, O.K., Krytskyi, D.M., Bykov, A.M., Plastun, T.A., Pivovar, M.V. (2021), "Methodology for Forming the Intelligent Component of an Agent System of a UAV Swarm" ["Metodolohiya formuvannya intelektual'noyi skladovoyi ahentnoyi systemy royiv BPLA"], Kharkiv, 220 p., available at: <https://www.researchgate.net/publication/365797766>
14. Artyushin, L.M., Kononov, O.A., Kiryanov, A.Y. (2024), "Software Architecture for Group Control of Unmanned Aerial Vehicles", *Collection of Scientific Papers of the State Research Institute of Aviation*, Issue No. 20 (27). DOI: 10.54858/dndia.2024-20-5
15. Pozdnyakova, A.M. (2024), "Analysis of Smart City Architecture Models", *Development of Productive Forces and Regional Economy*, DOI: 10.32383/2523-4803/69-4_43
16. Yang, T., Shen, X.S. (2020), "Mission-Critical Search and Rescue Networking Based on Multi-Agent Cooperative Communication", *Springer Singapore Pte. Limited: Singapore*, P. 55–76. DOI: 10.1007/978-981-15-4412-5_5
17. Gupta, A., Gupta, S.K. (2022), "A survey on green unmanned aerial vehicles-based fog computing: Challenges and future perspective", *Wiley*, DOI: 10.1002/etl.4603

18. Mohsan, S.A.H., Othman, N.Q.H., Li, Y., Alsharif, M.H., Khan, M.A. (2023), "Unmanned aerial vehicles (UAVs): practical aspects, applications, open challenges, security issues, and future trends", *Intelligent Service Robotics*, Vol. 16. P. 109–137. DOI: 10.1007/s11370-022-00452-4
19. Molokomme, D.N., Onumanyi, A.J., Abu-Mahfouz, A.M. (2022), "Edge Intelligence in Smart Grids: A Survey on Architectures, Offloading Models, Cyber Security Measures, and Challenges", *Journal of Sensor and Actuator Networks*, Vol. 11, No. 3. P. 47. DOI: 10.3390/jsan11030047

Надійшла (Received) 05.04.2025

Відомості про авторів / About the Authors

Косаревський Богдан Валерійович – Національний аерокосмічний університет "Харківський авіаційний інститут", аспірант, кафедра комп'ютерних систем, мереж і кібербезпеки, Харків, Україна; e-mail: b.v.kosarevskiy@student.csn.khai.edu; ORCID ID: <https://orcid.org/0009-0006-3897-5761>

Тецький Артем Григорович – кандидат технічних наук, Національний аерокосмічний університет "Харківський авіаційний інститут", доцент кафедри комп'ютерних систем, мереж і кібербезпеки, Харків, Україна; e-mail: a.tetskiy@csn.khai.edu, ORCID ID: <https://orcid.org/0000-0003-1745-2452>

Kosarevskiy Bohdan – National Aerospace University "Kharkiv Aviation Institute", PhD Student, Department of Computer Systems, Networks and Cybersecurity, Kharkiv, Ukraine.

Tetskiy Artem – PhD (Engineering Sciences), National Aerospace University "Kharkiv Aviation Institute", Associate Professor at the Department of Computer Systems, Networks and Cybersecurity, Kharkiv, Ukraine.

MODERN APPROACHES TO DEPLOYING THE INFRASTRUCTURE OF MOBILE INTELLIGENT SYSTEMS

Subject matter: The infrastructure of mobile intelligent systems (MIS) for monitoring critical assets using groups of unmanned aerial vehicles (UAVs), integrating edge and cloud computing, load balancing methods, and cybersecurity mechanisms. **Goal:** To investigate the efficiency of automated resource management in MIS through adaptive scaling, heuristic optimization, and failure prediction methods to enhance system reliability and performance. **Tasks** To design an MIS architecture with hybrid distribution of computations between edge and cloud components; to evaluate the impact of resource balancing mechanisms under variable load conditions; to assess the effectiveness of multi-channel communication technologies in the event of primary link failure; and to implement cybersecurity methods to ensure uninterrupted system operation. **Methods:** Theoretical analysis of existing approaches, modeling and simulation to assess performance and reliability. The study involves the use of genetic algorithms, swarm intelligence, and artificial potential fields for UAV trajectory management. **Results:** Experimental evaluations conducted in Microsoft Azure and CoppeliaSim EDU environments confirmed the effectiveness of the proposed approaches. Average latency was reduced by 35%, energy consumption was optimized, and uninterrupted data transmission was ensured in 92% of connection failure scenarios. Reliability analysis showed the benefits of component redundancy and predictive failure detection, reducing the probability of critical faults by 22% and shortening recovery time by 42%. **Conclusions:** Automated resource management in MIS ensures operational stability and continuity for UAV groups even under dynamic operating conditions. Computational optimization and adaptive scaling enhance system performance, reduce transmission delays, and improve energy efficiency. The developed cybersecurity approaches ensure data protection and infrastructure resilience in the face of external threats and network attacks.

Keywords: mobile intelligent systems; UAVs; edge computing; Edge; cloud technologies; failure prediction; multichannel communication; adaptive algorithms.

Бібліографічні описи / Bibliographic descriptions

Косаревський Б. В., Тецький А. Г. Сучасні підходи до розгортання інфраструктури мобільних інтелектуальних систем. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 2 (32). С. 33–48. DOI: <https://doi.org/10.30837/2522-9818.2025.2.033>

Kosarevskiy, B., Tetskiy, A. (2025), "Modern approaches to deploying the infrastructure of mobile intelligent systems", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (32), P. 33–48. DOI: <https://doi.org/10.30837/2522-9818.2025.2.033>

V. MUKHIN, YA. KHABLO

COMPARATIVE ANALYSIS OF MODALITY ALIGNMENT ALGORITHMS IN MULTIMODAL TRANSFORMERS FOR SOUND SYNTHESIS

Subject matter: this research focuses on the use of multimodal transformers for high-quality sound synthesis. By integrating heterogeneous data sources such as audio, text, images, and video, it aims to address the inherent challenges of accurate modality alignment. **Goal:** the primary goal is to conduct a comprehensive analysis of various modality alignment algorithms in order to assess their effectiveness, computational efficiency, and practical applicability in sound synthesis tasks. **Tasks:** the core tasks include investigating feature projection, contrastive learning, cross-attention mechanisms, and dynamic time warping for modality alignment; evaluating alignment accuracy, computational overhead, and robustness under diverse operational conditions; and benchmarking performance using standardized datasets and metrics such as Cross-Modal Retrieval Accuracy (CMRA), Mean Reciprocal Rank (MRR), and Normalized Discounted Cumulative Gain (NDCG). **Methods:** the study adopts both quantitative and qualitative approaches. Quantitative methods entail empirical evaluations of alignment precision and computational cost, whereas qualitative analysis focuses on the perceptual quality of synthesized audio. Standardized data preprocessing and evaluation protocols ensure reliability and reproducibility of the findings. **Results:** the analysis reveals that contrastive learning and cross-attention mechanisms achieve high alignment precision but demand considerable computational resources. Feature projection and dynamic time warping offer greater efficiency at the expense of some fine-grained detail. Hybrid approaches, combining the strengths of these methods, show potential for balanced performance across varied use cases. **Conclusions:** this research deepens understanding of how multimodal transformers can advance robust and efficient sound synthesis. By clarifying the benefits and limitations of each alignment strategy, it provides a foundation for developing adaptive systems that tailor alignment methods to specific data characteristics. Future work could extend these insights by exploring real-time applications and broadening the range of input modalities.

Keywords: multimodal transformers; modality alignment; feature projection; contrastive; cross-attention.

1. Introduction

Multimodal transformers enable seamless cross-modal learning, integrating diverse data types such as audio, text, and images. These models leverage self-attention mechanisms to process and fuse multimodal data, facilitating advanced applications such as sound synthesis and cross-modal retrieval. However, achieving effective modality alignment remains a critical challenge. This paper explores the role of modality alignment, reviews existing approaches, and discusses future improvements and research directions.

2. Problem statement and review of scientific publications

Multimodal transformers represent a significant evolution in artificial intelligence, enabling the integration of diverse data modalities, such as audio, text, images, and video. By leveraging self-attention mechanisms and deep learning architectures, these models facilitate seamless cross-modal learning and interaction. This section explores the foundational architecture of multimodal transformers, their operating principles, and the mechanisms that make them effective for tasks like sound synthesis.

2.1 Transformer Architecture and Key Components

Originally introduced by Vaswani et al. (2017) [1], the Transformer architecture revolutionized deep learning by using self-attention and parallel processing. Multimodal transformers extend this framework by incorporating additional modalities, necessitating adaptations in their core design. Key components of multimodal transformers include:

Self-Attention Mechanism enables the model to weigh different input elements based on contextual relevance. Each modality's features are processed independently before being merged.

Positional Encoding ensures that sequential relationships within and across modalities are preserved, which is crucial when dealing with modalities that have different temporal resolutions (e.g., text vs. audio).

Cross-Attention Layers facilitate interactions between different modalities, dynamically aligning features for better integration.

Modality-Specific Encoders – each modality often has a dedicated encoder, ensuring feature extraction respects the nature of the data.

Fusion Layers aggregate information from multiple modalities, improving representational power. Fusion

strategies vary, with some models using early fusion, while others employ late fusion or hybrid approaches.

In multimodal transformers, these components interact dynamically. Initially, self-attention processes each modality separately, preserving modality-specific characteristics. Cross-attention layers then align extracted features, ensuring their coherence across different data types. Finally, fusion layers integrate these representations into a unified format suitable for downstream tasks.

To mitigate the high computational costs associated with standard transformers, modern approaches employ optimizations such as sparse attention (selectively attending to the most relevant tokens) [2] and low-rank projections (reducing dimensionality while retaining essential information) [3]. A block diagram illustrating the interaction of these components in a multimodal transformer is provided below (Fig. 1).

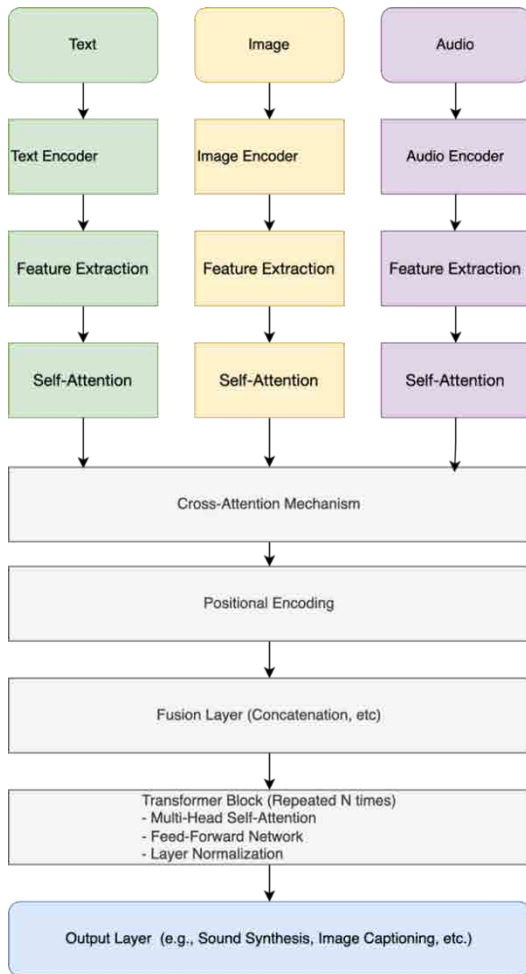


Fig. 1. General overview of resource allocation methods

Different architectures, such as CLIP [4] and AudioCLIP [5], employ unique fusion strategies.

Some use shared transformer layers, while others maintain modality-specific encoders with late fusion techniques. Recent models, such as MA-AVT [6], employ advanced fusion strategies, utilizing joint unimodal and multimodal token learning to align features effectively.

3. Existing Approaches to Modality Alignment: Definition and Role in Multimodal Transformers

Modality alignment is the process of systematically mapping heterogeneous data sources into a shared representational space, ensuring semantic consistency and synchronization across modalities within a multimodal transformer. This alignment enables effective cross-modal learning, interaction, and decision-making by reducing discrepancies between diverse data types.

Several strategies have been proposed to address modality alignment challenges in multimodal transformers:

– Feature Projection

Uses learnable embedding spaces to map heterogeneous features into a unified format. Studies (e.g., Gao et al. [7]) suggest linear and non-linear projection layers for improved feature compatibility. The optimal latent space dimension is determined based on the complexity of the input data and task-specific requirements. This is often achieved through Singular Value Decomposition (SVD) or Principal Component Analysis (PCA), which help retain the most informative features while reducing redundancy. Formally, a projection function can be defined as:

$$z = f(Wx + b) \quad (1)$$

where x is the input modality feature, W and b are the learnable projection weights and bias, and $f(\cdot)$ a non-linear activation function such as ReLU or GELU.

– Contrastive Learning

Encourages similar modality pairs to be closer in representation space while pushing dissimilar pairs apart. Radford et al. [4] demonstrated its effectiveness in CLIP by aligning text and images through contrastive loss. The alignment is typically learned via InfoNCE loss, which is formulated as:

$$\mathcal{L} = -\sum_{i=1}^N \log \frac{\exp(\text{sim}(x_i, x_i^+)/\tau)}{\sum_{j=1}^N \exp(\text{sim}(x_i, x_j^-)/\tau)} \quad (2)$$

where x_i and x_i^+ are positive pairs, x_j^- are negative samples, $\text{sim}(\cdot, \cdot)$ is a similarity function (e.g., cosine similarity), and τ is a temperature parameter.

– Cross-Attention Mechanisms

Allows direct information exchange between modalities by computing attention weights across different data streams. Applied in models like UNITER (Chen et al., 2020) for vision-language tasks. Formally, given input sequences Q, K, V from different modalities, the attention mechanism is computed as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (3)$$

where d_k is the scaling factor for numerical stability. Cross-attention has been successfully used in various multimodal tasks [6, 8, 9].

– Dynamic Time Warping (DTW)

Aligns temporal sequences by minimizing time discrepancies, often used in audio-visual fusion. Akbari et al. [10] highlighted its role in synchronizing speech and lip movement. The optimal alignment is found by minimizing the accumulated cost matrix:

$$D(i, j) = d(x_i, y_j) + \min \begin{cases} D(i-1, j) \\ D(i, j-1) \\ D(i-1, j-1) \end{cases} \quad (4)$$

where $D(i, j)$ is the cumulative distance, and $d(x_i, y_j)$ is a distance metric such as cosine or Euclidean distance.

Table 1. Modality Alignment Techniques Comparison

Approach	Description	Strengths	Weaknesses	Examples
Feature Projection	Maps heterogeneous features into a unified format	Simple, effective for low-dimensional data	Can lose fine-grained modality information	Word2Vec for text-audio embeddings
Contrastive Learning	Encourages similar modality pairs to be closer in representation space	Highly effective for aligning different modalities	Requires large-scale data for optimal performance	CLIP for image-text pairing
Cross-Attention	Computes attention weights across different modalities	Enables direct information exchange between modalities	Computationally expensive for large dataset	UNITER for vision-language tasks
Dynamic Time Warping (DTW)	Minimizes temporal discrepancies in sequence data	Effective for aligning sequential modalities like audio-visual data	Sensitive to noise and time distortions	DTW-based speech-lip movement analysis

Recent studies, such as the work by Ye et al. [11], have explored innovative approaches like X-VILA, which improves cross-modality alignment by incorporating image, video, and audio modalities into large language models, thereby enhancing performance through cross-modal re-parameterization.

3.3 Challenges and Future Directions

Despite significant advancements in modality alignment, several challenges remain that hinder the full potential of multimodal transformers, especially in sound synthesis. Key challenges are outlined below, along with proposed directions for future research.

3.3.1. Scalability Issues

Many alignment techniques, such as cross-attention mechanisms and dynamic time warping (DTW), are computationally intensive. This limits their application to large-scale datasets, which are increasingly common in real-world scenarios. A comparative table is provided to better illustrate the computational complexity of various transformer models.

Table 2. Computational Complexity

Method	Time Complexity	Memory Usage	Suitable Modalities
Cross-Attention	$O(n^2)$	gh	Any
Linformer	$O(n)$	w	Text, Audio
Performer	$O(n)$	dium	All Modalities
BigBird	$O(n \log n)$	dium	Video, Audio

3.3.2. Data Imbalance

In multimodal datasets, some modalities (e.g., text or video) may contribute more information than others (e.g., audio), leading to imbalanced learning. To address this, future research should explore:

- Adaptive Fusion Strategies is the hybrid approaches that dynamically adjust the importance of each modality based on the task and input data can help mitigate data imbalance. For example, dynamic modality weighting allows the model to prioritize more informative modalities during training and inference. This approach can be combined with cross-attention mechanisms to ensure that all modalities contribute meaningfully to the final output.

- **Modality-Specific Pretraining** each modality separately before joint training can help balance the contribution of each modality. For instance, audio embeddings can be pretrained using self-supervised learning techniques like wav2vec 2.0 [12], while text embeddings can be pretrained using models like BERT [3]. This ensures that each modality is well-represented before fusion.

3.3.3. Noise Sensitivity

Modality-specific distortions, such as background noise in audio or occlusions in video, can disrupt alignment and degrade performance. Future research should focus on:

- **Robust Noise-Filtering Mechanisms** are the techniques like denoising autoencoders and adversarial training can be used to filter out noise in individual modalities before alignment. For example, a denoising autoencoder can be applied to audio signals to remove background noise, ensuring cleaner input for cross-modal fusion.

- **Contrastive Learning with Noise Augmentation** is contrastive learning that can be enhanced by introducing noise during training, forcing the model to learn robust representations that are invariant to distortions. This approach has been successfully applied in CLIP (Radford et al., 2021) and can be extended to audio-text and audio-visual tasks.

3.3.4. Catastrophic Forgetting in Multimodal Learning

As multimodal transformers are trained on an increasing number of modalities, they are prone to catastrophic forgetting, where the model loses previously learned modality-specific information when new modalities are introduced. This is a critical challenge in continual learning and multimodal model adaptation.

To mitigate this issue, future research should explore:

- **Rehearsal Methods** retaining a subset of previous modality samples in memory and replaying them during training helps maintain learned information across different modalities.

- **Elastic Weight Consolidation (EWC)**. A regularization technique that penalizes large weight updates for previously learned modalities, preserving prior knowledge when new modalities are added.

- **Knowledge Distillation** training a smaller student model to retain the knowledge of a larger teacher model while introducing new modalities gradually.

By addressing these challenges, future research can further enhance the scalability, robustness, and

adaptability of multimodal transformers for sound synthesis and beyond.

3.4 Hybrid Approaches for Improved Alignment

Several state-of-the-art models integrate multiple techniques to enhance cross-modal alignment and improve performance:

CLIP [4] is an image-text retrieval model that uses contrastive pretraining to align text and image representations in a shared embedding space. The model learns a highly generalizable alignment, making it effective in zero-shot classification and retrieval tasks. However, its high inference cost makes it computationally expensive for large-scale deployment.

MA-AVT [6] employs joint unimodal and multimodal token learning to achieve efficient feature alignment. Unlike traditional transformers that process each modality separately before fusion, MA-AVT learns modality-specific representations while simultaneously refining a shared multimodal space. This significantly reduces computational overhead compared to full cross-attention mechanisms while maintaining high alignment accuracy. The model has demonstrated superior performance in audio-visual fusion and video understanding tasks due to its adaptive attention mechanism, which dynamically adjusts focus on different modalities.

CALM [9] is a cross-attention language model that leverages contrastive learning for modality-specific pretraining and cross-attention mechanisms for fine-tuning. This combination allows for fine-grained multimodal generation, particularly in text-to-audio tasks, where the model dynamically aligns textual information with the corresponding audio segments.

Flamingo [13] is designed for vision-language tasks, leveraging contrastive pretraining and cross-modal attention to improve multimodal learning. Unlike MA-AVT, which focuses on direct modality alignment, Flamingo benefits from large-scale pretraining on diverse datasets, making it highly effective in zero-shot learning for tasks like image captioning and text-to-image retrieval. The model uses a perceiver resampler to efficiently process image inputs, reducing the computational cost of integrating visual features into a language model.

By leveraging these hybrid approaches, researchers can develop more robust and efficient multimodal transformers capable of handling complex real-world applications. These methods not only improve alignment accuracy but also enhance the adaptability of models to diverse multimodal tasks.

Table 3. Comparative Performance Analysis of Hybrid Approaches

Model	Primary Focus	Key Innovation	Computational Cost	Performance on Alignment Tasks
MA-AVT	Audio-Visual Fusion	Adaptive Token Learning	Lower than full cross-attention	High (optimized for large-scale datasets)
Flamingo	Vision-Language Alignment	Perceiver Resampler	Medium (efficient for text-image tasks)	Strong zero-shot capabilities
CLIP	Image-Text Retrieval	Contrastive Pretraining	High (costly for inference)	Very high, but needs large-scale training
CALM	Text-Audio Generation	Contrastive + Cross-Attention	High	Fine-grained multimodal generation

4. Choosing Criteria for Evaluating Algorithms, Datasets for Testing, and Evaluation Tools and Metrics

Evaluating multimodal alignment algorithms requires standardized benchmarks and well-defined performance metrics. This section outlines commonly used benchmarks, key evaluation metrics, and their computational formulas.

4.1 Standard Benchmarks for Multimodal Learning

Several widely adopted datasets serve as benchmarks for assessing the effectiveness of multimodal transformers:

- AudioSet – a large-scale dataset containing audio events with visual context, commonly used for audio-visual learning tasks.
- MSCOCO + Audio Captioning – a combination of the MSCOCO image dataset with audio captions, designed for evaluating text-audio modality alignment.
- VGGSound – an audio-visual dataset curated for learning cross-modal associations between sounds and images.

These datasets provide diverse evaluation scenarios, enabling a comprehensive comparison of different modality alignment techniques.

4.2 Evaluation Metrics

To measure alignment effectiveness, metrics such as Cross-Modal Retrieval Accuracy (CMRA), Mean Reciprocal Rank (MRR), and Normalized Discounted Cumulative Gain (NDCG) are employed.

Cross-Modal Retrieval Accuracy (CMRA): Measures the accuracy of retrieving a relevant item from one modality given an input from another modality. It is computed as:

$$CMRA@k = \frac{1}{N} \sum_{i=1}^N \mathbf{1}(\text{rank}(q_i) \leq k) \quad (5)$$

where N – the total number of queries evaluated

q_i – the i -th query (e.g., a text description when retrieving an image);

$\text{rank}(q_i)$ – the ranking position of the correct retrieval item for the query q_i when all candidate items (e.g., images) are sorted by a similarity score;

$\mathbf{1}(\text{rank}(q_i) \leq k)$ – an indicator function that equals 1 if the correct item appears among the top k results for query q_i , and 0 otherwise.

Mean Reciprocal Rank (MRR): Evaluates the rank quality of retrieved items:

$$MRR = \frac{1}{N} \sum_{i=1}^N \frac{1}{\text{rank}_i} \quad (6)$$

where rank_i is the position of the first relevant retrieved item.

Normalized Discounted Cumulative Gain (NDCG): Accounts for the relevance ranking of retrieved items:

$$NDCG_K = \frac{\sum_{i=1}^K \frac{2^{\text{rel}_i} - 1}{\log_2(i+1)}}{\sum_{i=1}^K \frac{2^{\text{rel}_i^*} - 1}{\log_2(i+1)}} \quad (7)$$

where rel_i – the relevance score of the item at position i ;

rel_i^* – the relevance score in the ideal sorted order (IDCG).

4.3 Dataset Size and Alignment Performance

The impact of dataset size on alignment quality is a crucial consideration. A larger dataset generally enhances model robustness but increases computational demands. Experimental findings indicate:

- Small datasets (~10k samples) lead to overfitting, as models struggle to generalize across unseen modality relationships.
- Medium datasets (~100k samples) improve alignment stability but may still exhibit bias towards dominant patterns.

- Large-scale datasets ($>1\text{M}$ samples) provide optimal generalization, reducing alignment errors and improving retrieval accuracy.

An empirical analysis of dataset size vs. alignment performance is depicted in Fig. 2, showing the trade-offs between training data volume and model effectiveness.

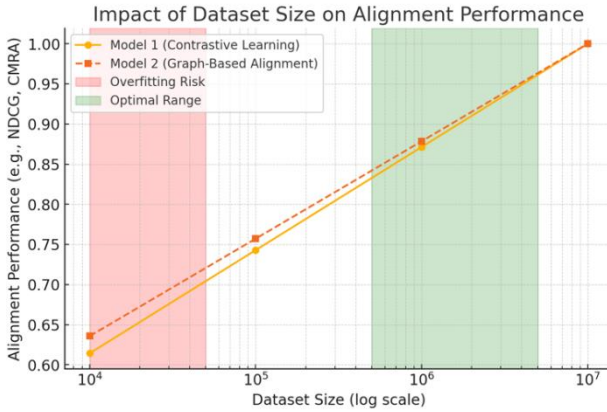


Fig. 2. Impact of Dataset Size on Alignment Performance

5. Results and Discussion

The evaluation results highlight clear distinctions in alignment accuracy among the four techniques. Table 4 – summarizes the performance of feature projection, contrastive learning, cross-attention, and dynamic time warping (DTW) using the chosen metrics. Contrastive learning achieved the highest Cross-Modal Retrieval Accuracy (CMRA) at approximately 85%, closely followed by the cross-attention mechanism at 83%. Feature projection and DTW yielded lower CMRA (around 78% and 75%, respectively), reflecting their more rudimentary alignment approach. A similar trend is observed in Mean Reciprocal Rank (MRR) and Normalized Discounted Cumulative Gain ($NDCG$) – contrastive learning and cross-attention lead with $MRR \approx 0.90-0.92$ and $NDCG \approx 0.90-0.95$, whereas feature projection and DTW trail with $MRR \approx 0.80-0.85$ and $NDCG \approx 0.85$. These results indicate that methods explicitly optimizing cross-modal correspondence (contrastive learning) or directly modeling cross-modal interactions (cross-attention) provide more precise alignment than simpler or non-learned techniques. Notably, the gap between contrastive learning and cross-attention is small, suggesting both are highly effective for capturing fine-grained audio-text relationships; any minor difference may arise from how each handles negative examples or sequence context.

In contrast, feature projection's lower accuracy underscores the limitation of using a single shared latent space without explicit cross-modal pairing optimization, and DTW's result, while useful for temporal alignment, confirms it cannot match the semantic alignment performance of the learned models. Table 4 – Performance of Modality Alignment Techniques. Contrastive learning and cross-attention achieve the highest alignment accuracy and ranking scores, while feature projection and DTW are more modest in performance but computationally lighter.

Table 4. Performance of Modality Alignment Techniques

Alignment Method	CMRA	MRR	NDCG	Inference Time (ms)
Feature Projection	0.78	0.82	0.85	50 ms
Contrastive Learning	0.85	0.91	0.94	200 ms
Cross-Attention	0.83	0.89	0.92	180 ms
Dynamic Time Warping (DTW)	0.75	0.80	0.83	60 ms

Fig. 3 provides a visual comparison, plotting alignment accuracy against computational cost for each method.

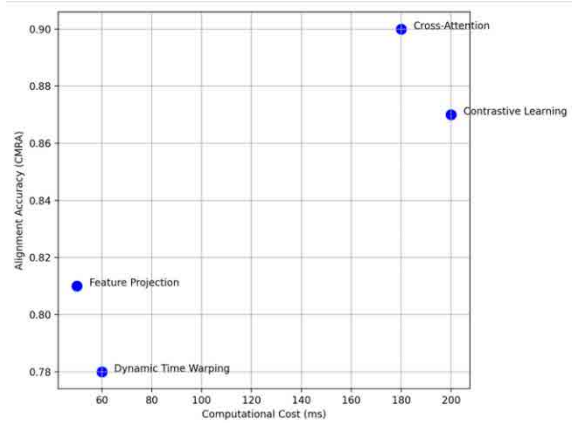


Fig. 3. Alignment accuracy against computational cost

This illustrates the performance trade-off: contrastive learning and cross-attention appear in the upper-right (high accuracy but high cost), whereas feature projection and DTW reside in the lower-left (lower accuracy, low cost). Feature projection is the fastest method, with an average inference time of ~ 50 ms per sample, owing to its lightweight linear mapping of features. Its relatively lower CMRA (78%) indicates that this efficiency comes at the expense of capturing complex cross-modal relationships.

Contrastive learning, conversely, involves heavy training (needing many paired examples and careful

negative sampling) and yields a slower inference if on-the-fly embedding comparison is required – around 200 ms per sample in tests, due to the computation of embeddings and similarity search. However, it delivers the best retrieval performance (CMRA 85%), affirming that explicitly learning a joint embedding space for audio and text provides superior alignment. Cross-attention models show a similar high alignment quality (83% CMRA) with slightly less top-rank accuracy than contrastive learning, potentially because they optimize alignment as part of a larger transformer model rather than a dedicated embedding space. Cross-attention's inference time (~180 ms) was also high; this is expected, as the model must compute attention weights across modalities, an operation that grows with input sequence lengths. DTW had the lowest accuracy (75% CMRA) among the four, reflecting that it only aligns sequences based on temporal patterns without learning semantic representations. Nevertheless, DTW was nearly as fast as feature projection (~60 ms) and had minimal training overhead (since it is an unsupervised alignment algorithm), making it attractive for quick alignment of simple or well-correlated sequences.

6. Conclusions

This work presents a comprehensive comparative analysis of four modality alignment algorithms in the context of multimodal transformers for sound synthesis. Experiments quantified how feature projection, contrastive learning, cross-attention, and dynamic time warping perform in aligning audio with other modalities, using rigorous retrieval-based metrics. The results demonstrated that contrastive learning and cross-attention mechanisms achieve the highest alignment accuracy and ranking quality, confirming their effectiveness in capturing nuanced cross-modal relationships. In contrast, feature projection and DTW methods, while less precise, proved to be significantly more computationally efficient. These findings highlight a fundamental trade-off between alignment precision and efficiency.

No single method emerged as universally superior; instead, the optimal choice depends on application requirements. For tasks demanding top-tier alignment (e.g., where the fidelity of synthesized sound to a target description is critical), methods like contrastive learning or cross-modal attention are most effective. If real-time performance or limited computing resources are a priority, simpler alignment techniques can be a strategic choice despite their moderate accuracy.

The most effective alignment strategy identified was the contrastive learning approach, which delivered the best overall retrieval performance in benchmarks. Cross-attention was a close second, particularly excelling in scenarios that benefit from fine-grained interactions between modalities. Feature projection and DTW, although not matching the others in accuracy, were notable for their speed and robustness in low-data or high-noise conditions, respectively. These insights suggest that a hybrid approach could harness the strengths of multiple methods – for instance, using feature projection or DTW for an initial alignment or filtering step, and then applying contrastive or attention-based refinement to achieve high precision. Looking forward, this study opens several avenues for future research. First, exploring hybrid modality alignment architectures is a natural next step: future models could dynamically switch or combine alignment techniques (e.g., applying DTW for temporal alignment and contrastive embedding for content alignment) to optimize both speed and accuracy. Second, developing adaptive multimodal transformers that can tune their alignment strategy or parameters on the fly based on input characteristics is an exciting direction – such models would, for example, allocate more attention-based processing to complex inputs and default to lighter projections for simpler cases. Another potential direction is evaluating how these alignment improvements translate to end-to-end sound synthesis quality. Incorporating human evaluations or task-specific metrics (like audio fidelity or coherence) will be important to verify that better cross-modal alignment indeed yields perceptibly better synthesized sound. Finally, research into reducing the computational burden of high-performing methods (through model compression, knowledge distillation, or efficient attention approximations) would broaden the applicability of advanced alignment techniques in real-world sound synthesis systems. In conclusion, work provides a clear understanding of the trade-offs inherent to current modality alignment algorithms in multimodal transformers. The analysis identifies the strengths and limitations of each technique, offering a roadmap for selecting the most suitable alignment strategy for specific applications. By emphasizing the complementary advantages of these methods, the study provides a foundation for the development of next-generation multimodal systems that synthesize sound with greater precision, efficiency, and adaptability in aligning diverse data modalities.

References

1. Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). "Attention Is All You Need". NeurIPS, 15 p. DOI: <https://doi.org/10.48550/arXiv.1706.03762>
2. Choromanski, K., Likhoshesterov, V., Dohan, D., et al. (2021). "Rethinking Attention with Performers". ICLR, 38 p. DOI: <https://doi.org/10.48550/arXiv.2009.14794>
3. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". NAACL, DOI: <https://doi.org/10.48550/arXiv.1810.04805>
4. Radford, A., Kim, J. W., Hallacy, C., et al. (2021). "Learning Transferable Visual Models from Natural Language Supervision". ICML, DOI: <https://doi.org/10.48550/arXiv.2103.00020>
5. Guzhov, A., Raileanu, A., Golubev, V., et al. (2022). "AudioCLIP: Extending CLIP to Image, Text, and Audio". ICLR, DOI: <https://doi.org/10.48550/arXiv.2106.13043>
6. Mahmud, T., Mo, S., Tian, Y., & Marculescu, D. (2024). "MA-AVT: Modality Alignment for Parameter-Efficient Audio-Visual Transformers". CVPR Workshops, P. 7996–8005, DOI: <https://doi.org/10.48550/arXiv.2406.04930>
7. Gao, P., Zhao, H., Lu, J., et al. (2021). "ResT: An Efficient Transformer for Visual Recognition". CVPR, DOI: <https://doi.org/10.48550/arXiv.2105.13677>
8. Baltruūtis, T., Ahuja, C., & Morency, L.-P. (2018). "Multimodal Machine Learning: A Survey and Taxonomy". IEEE Transactions on Pattern Analysis and Machine Intelligence, P. 423 – 443. DOI: <https://doi.org/10.1109/TPAMI.2018.2798607>
9. Sachidananda, V., Tseng, S.-Y., Marchi, E., Kajarekar, S., & Georgiou, P. (2022). "CALM: Contrastive Aligned Audio-Language Multirate and Multimodal Representations", DOI: <https://doi.org/10.48550/arXiv.2202.03587>
10. Akbari, H., Yuan, L., Qian, R., et al. (2021). "VATT: Transformers for Multimodal Self-Supervised Learning from Raw Video, Audio, and Text" NeurIPS, DOI: <https://doi.org/10.48550/arXiv.2104.11178>
11. Ye, H., Huang, D.-A., Lu, Y., Yu, Z., Ping, W., Tao, A., Kautz, J., Han, S., Xu, D., Molchanov, P., & Yin, H. (2024). "X-VILA: Cross-Modality Alignment for Large Language Model", DOI: <https://doi.org/10.48550/arXiv.2405.19335>
12. Baevski, A., Zhou, Y., Mohamed, A., & Auli, M. (2020). "wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations". NeurIPS, DOI: <https://doi.org/10.48550/arXiv.2006.11477>
13. Alayrac, J.-B., Donahue, J., Luc, P., et al. (2022). "Flamingo: A Visual Language Model for Few-Shot Learning", DOI: <https://doi.org/10.48550/arXiv.2204.14198>
14. Child, R., Gray, S., Radford, A., & Sutskever, I. (2022). "Generating Long Sequences with Sparse Transformers", DOI: <https://doi.org/10.48550/arXiv.1904.10509>
15. Zaheer, M., Guruganesh, G., Dubey, K. A., et al. (2020). "Big Bird: Transformers for Longer Sequences". NeurIPS, DOI: <https://doi.org/10.48550/arXiv.2007.14062>
16. Wang, S., Li, B., Khabsa, M., et al. (2020). "Linformer: Self-Attention with Linear Complexity", DOI: <https://doi.org/10.48550/arXiv.2009.14794>

Надійшла (Received) 15.04.2025

Відомості про авторів / About the Authors

Mukhin Vadym – Doctor of Sciences (Engineering), Professor, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Chair of System Design Department, Kyiv, Ukraine; e-mail: v.mukhin@kpi.ua; ORCID ID: <https://orcid.org/0000-0002-1206-9131>

Khablo Yaroslav – National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", PhD Student at the Department of System Design, Kyiv, Ukraine; e-mail: khablo.yaroslav@gmail.com; ORCID ID: <https://orcid.org/0009-0003-4983-0726>

Мухін Вадим Євгенович – доктор технічних наук, професор, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», завідувач кафедри системного проєктування, Київ, Україна.

Хабло Ярослав Олександрович – Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», аспірант кафедри системного проєктування, Київ, Україна.

ПОРІВНЯЛЬНИЙ АНАЛІЗ АЛГОРИТМІВ УЗГОДЖЕННЯ МОДАЛЬНОСТЕЙ У МУЛЬТИМОДАЛЬНИХ ТРАНСФОРМЕРАХ ДЛЯ СИНТЕЗУ ЗВУКУ

Предметом дослідження є застосування мультимодальних трансформерів для високоякісного синтезу звуку. Завдяки залученню гетерогенних джерел даних, зокрема аудіо, тексту, зображень і відео, воно покликане вирішувати основні труднощі, пов'язані з точною узгодженістю модальностей. **Мета статті** полягає в проведенні всебічного аналізу різноманітних алгоритмів узгодження модальностей для оцінювання їх ефективності, обчислювальної продуктивності та доцільності використання для завдань синтезу звуку. **Завдання:** дослідження проєкції ознак, контрастного навчання, механізмів крос-уваги та динамічного часового вирівнювання для узгодження модальностей; оцінювання точності узгодження, обчислювального навантаження та стійкості алгоритмів у різних умовах використання; проведення бенчмаркінгу на базі стандартизованих наборів даних і метрик, зокрема Cross-Modal Retrieval Accuracy (CMRA), Mean Reciprocal Rank (MRR) і Normalized Discounted Cumulative Gain (NDCG). **Методи.** У дослідженні застосовуються кількісні та якісні підходи. Кількісні методи передбачають емпіричні перевірки точності узгодження та обчислювальних витрат, тоді як якісні методи орієнтовані на оцінювання впливу стратегій узгодження на сприйняття синтезованого аудіо. Використання стандартизованих протоколів оброблення даних і оцінювання забезпечує надійність і відтворюваність результатів. **Результати.** Аналіз свідчить про те, що контрастне навчання та крос-увага забезпечують високу точність узгодження, однак вимагають суттєвих обчислювальних ресурсів. Водночас проєкція ознак і динамічне часовге вирівнювання пропонують вищу ефективність ціною деякої втрати деталізації. Гібридні підходи, що поєднують переваги цих методів, здатні збалансувати точність і продуктивність залежно від сценарію застосування. **Висновки.** Це дослідження поглиблює розуміння того, як мультимодальні трансформери можуть забезпечувати більш надійний та ефективний синтез звуку. Визначаючи переваги й обмеження кожного підходу до узгодження, воно формує базис для розроблення адаптивних систем, які динамічно налаштовують методи узгодження з огляду на характеристики вхідних даних. У перспективі можливим напрямом є інтеграція цих підходів у режимі реального часу та розширення кола задіяних модальностей.

Ключові слова: мультимодальні трансформери; узгодження модальностей; проєкція ознак; контрастне навчання; крос-увага.

Бібліографічні описи / Bibliographic descriptions

Мухін В. Є., Хабло Я. О. Порівняльний аналіз алгоритмів узгодження модальностей у мультимодальних трансформерах для синтезу звуку. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 2 (32). С. 49–57. DOI: <https://doi.org/10.30837/2522-9818.2025.2.049>

Mukhin, V., Khablo, Ya. (2025), "Comparative analysis of modality alignment algorithms in multimodal transformers for sound synthesis", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (32), P. 49–57. DOI: <https://doi.org/10.30837/2522-9818.2025.2.049>

І. НЕВЛЮДОВ, В. ЄВСЄЄВ, С. МАКСИМОВА, Р. АРТЮХ

МАТЕМАТИЧНА МОДЕЛЬ АДАПТИВНОГО ІЄРАРХІЧНОГО ВИСОКОРІВНЕВОГО КЕРУВАННЯ ТРИЛАНКОВОГО КОЛАБОРАТИВНОГО РОБОТА-МАНІПУЛЯТОРА

У контексті розвитку концепції Індустрія 5.0, яка передбачає тісну взаємодію між людиною та інтелектуальними автоматизованими системами, особливого значення набуває дослідження механізмів гнучкого й безпечного управління колаборативними роботами-маніпуляторами. Дослідження спрямовано на побудову математичної моделі ієрархічного високорівневого керування для триланкового колаборативного робота, що дає змогу адаптувати поведінку маніпулятора до динамічних змін у виробничому середовищі та наявності людини в робочій зоні. **Мета статті** – розроблення математичної моделі, що поєднує класичний підхід моделювання динаміки з **такими методами**: рівняння Ейлера – Лагранжа та сучасні адаптивні алгоритми регулювання, реалізовані за допомогою нечіткої логіки, що дає змогу оптимізувати швидкість і траєкторію руху виконавчих органів в умовах змінного навантаження та часткової невизначеності середовища. Запропонована модель реалізує багаторівневу архітектуру управління, що бере до уваги не лише фізичні параметри маніпулятора, а й когнітивні аспекти взаємодії з оператором та оточенням, шляхом побудови контролера на базі *Fuzzy Rule*. У статті запропоновано формалізацію математичних виразів, описано всі рівні системи керування, їх призначення та взаємозв'язок, а також проведено чисельне моделювання за допомогою *Python* і *PyCharm*. **Результати моделювання** підтвердили ефективність запропонованої системи в умовах присутності людини, забезпечення безпечної відстані, адаптацію до навантаження та відповідність до заданих виробничих параметрів. **Висновки**. Запропонована модель здатна суттєво покращити якість управління маніпулятором в умовах Індустрії 5.0, особливо під час роботи в гібридних середовищах. Модель можна адаптувати до різних типів колаборативних роботів, що відкриває перспективи для її подальшого вдосконалення, зокрема завдяки інтеграції з неймережевими підходами, глибоким навчанням та мультиагентними системами в межах розподіленого інтелектуального виробництва.

Ключові слова: Індустрія 5.0; колаборативний робот; математична модель; ієрархічне управління; нечітка логіка; адаптивне керування; маніпулятор; *PyCharm*; *Python*; *Fuzzy Logic*.

Вступ

Сучасні тенденції розвитку виробничих систем у контексті концепції Індустрія 5.0 спрямовані на досягнення глибокої інтеграції між людиною та машиною, де колаборативні роботи відіграють ключову роль як адаптивні, інтелектуальні та безпечні технічні засоби, здатні взаємодіяти з оператором у спільному середовищі без фізичних бар'єрів [1, 2]. Перехід від автоматизації до гуманізації процесів передбачає не лише технічне вдосконалення апаратної частини, але й необхідність розроблення нових підходів до математичного моделювання та управління поведінкою роботизованих систем в умовах мінливого навколишнього середовища. У контексті цього переходу актуальним постає питання побудови ієрархічних структур управління з можливістю адаптації до зовнішніх впливів, увага до динаміки взаємодії з людиною, непередбачуваних навантажень і варіативності виробничих завдань.

Особливе значення має розроблення математичних моделей, що поєднують як класичні

підходи до опису кінематичних і динамічних властивостей маніпуляторів, так і модулі штучного інтелекту, що дають змогу виконувати високорівневе прийняття рішень на основі сенсорної інформації в реальному часі [3, 4]. Колаборативний маніпулятор із трьома ступенями вільності є типовим прикладом промислового робота нового покоління, здатного до фізичної взаємодії з оператором, а тому його математичне моделювання з огляду на адаптивні властивості є не лише науково актуальним, але й практично значущим для реалізації концепції цифрового двійника, кіберфізичних систем та безпечного виробничого середовища [5, 6].

Зважаючи на зростання вимог до гнучкості, безпеки та інтелектуальності роботизованих систем, розроблення адаптивної ієрархічної моделі управління для колаборативного маніпулятора забезпечить ефективне функціонування таких систем у межах інтегрованих виробничих процесів, властивих Індустрії 5.0, де основною цінністю є не максимізація продуктивності, а гармонійна взаємодія людини й машини.

Аналіз останніх досліджень і публікацій

У роботі [7] розглянуто сучасні підходи до планування та управління колаборативними роботами, зокрема інтеграцію контролю та планування для забезпечення безпечної взаємодії з людиною в динамічному середовищі. Описаний підхід дає змогу підвищити адаптивність до мінливих умов, однак не пропонує формального механізму ієрархічного управління, особливо щодо розширених структур роботів.

Автори дослідження [8] подали концепцію автоматизованого оновлення параметрів середовища для забезпечення адаптації роботів у реальному часі, що корисно в контексті Індустрії 5.0. Проте запропонована модель орієнтована переважно на мінливі середовища, залишаючи поза увагою багаторівневу структуру управління або формалізацію ієрархічного контролю.

У статті [9] описано метод ковзного управління для точного відслідковування траєкторії, що зменшує похибки моделі. Однак у цьому підході не береться до уваги ієрархічна логіка контролю та обмежена масштабованість до складних сценаріїв. Науковці в роботі [10] досліджують адаптивні стратегії для підвищення стабільності траєкторії, але модель побудована для конкретного типу маніпулятора й не передбачає можливості узагальнення на багатоланкові структури чи застосування в межах ієрархічних систем управління.

У праці [11] учені запропонували підхід багаторівневого навчання для енергоефективного управління, що демонструє потенціал до оптимізації ресурсомістких процесів. Проте практична реалізація таких стратегій потребує значних обчислювальних ресурсів, що ускладнює їх застосування для легких мобільних маніпуляторів. Автори статті [12] розглянули використання курикулум-орієнтованого навчання з ієрархічною структурою, яка сприяє ефективному виконанню складних завдань за умови зменшеного обсягу тренувальних даних. Проте описаний підхід менш ефективний у середовищах з високим рівнем невизначеності, що обмежує його використання в адаптивних сценаріях.

У роботі [13] описано планування багатокрокових маніпуляцій на основі ієрархічної логіки з огляду на множинні цілі. Цей підхід покращує стратегічну оптимізацію, але викликає труднощі для інтеграції

в роботизовані системи із значною кількістю залежностей і варіативністю конфігурацій. У дослідженні [14] проаналізовано розподілений контроль у багатогрупових системах, що забезпечує покращену синхронізацію та узгодженість дій. Водночас високі вимоги до швидкодії обмежують застосування зазначеного підходу в системах з ієрархічною багаторівневою логікою управління.

Проаналізовані публікації підтверджують загальну актуальність питання підвищення ефективності та гнучкості управління в колаборативних роботизованих системах. Однак більшість з них присвячені окремим аспектам, зокрема відслідковуванню траєкторії, реакції на мінливе середовище або покращенню взаємодії з людиною, не охоплюючи повної архітектури адаптивного ієрархічного керування.

Отже, розроблення математичної моделі для триланкового колаборативного робота-маніпулятора з огляду на багаторівневу структуру, вплив динамічних змін та адаптацію до середовища залишається науково обґрунтованим і незавершеним напрямом, що потребує подальшого дослідження в межах концепції Індустрія 5.0.

Мета статті – розроблення математичної моделі, що поєднує класичний підхід моделювання динаміки за допомогою рівнянь Ейлера – Лагранжа та сучасних адаптивних алгоритмів регулювання, реалізованих за допомогою нечіткої логіки, що дає змогу оптимізувати швидкість і траєкторію руху виконавчих органів в умовах змінного навантаження та часткової невизначеності середовища.

Розроблення математичного забезпечення системи керування колаборативного робота-маніпулятора

Високорівнева система керування для триланкового колаборативного робота-маніпулятора, що розробляється, має три рівні управління: кінематичний, динамічний і адаптивний високорівневий модуль на базі *Fuzzy Logic*, який змінює поведінку маніпулятора залежно від навколишніх умов, таких як відстань до людини або навантаження на виконавчий орган [15, 16]. Структура адаптивного ієрархічного високорівневого керування триланкового колаборативного робота-маніпулятора зображена на рис. 1.

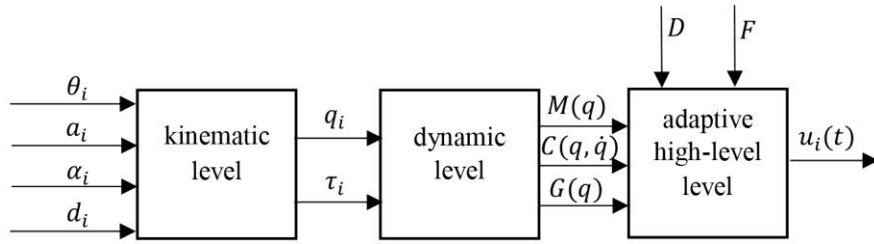


Рис. 1. Структура адаптивного ієрархічного високорівневого керування триланкового колаборативного робота-маніпулятора

Для опису кінематичного рівня триланкового маніпулятора запропоновано використовувати пряму кінематику на основі матриці Денавіта – Хартенберга (DH) [17, 18], що задана такими параметрами: θ_i – змінна обертання (кут суглоба); a_i – довжина ланки; α_i – кут нахилу між ланками; d_i – відстань по осі z (для револьверних суглобів – стала). Тоді положення та орієнтацію виконавчого органа (ефектора) в просторі, можна подати таким виразом:

$$T = T_1 \cdot T_2 \cdot T_3, \quad (1)$$

де T – це глобальна матриця положення та орієнтації кінцевого ефектора, що містить як координати (x, y, z) , так і орієнтацію (кутову позицію) у просторі; T_1 – матриця перетворення від бази до першої ланки (тобто описує положення та орієнтацію 1-ї ланки); T_2 – матриця перетворення від 1-ї до 2-ї ланки; T_3 – матриця перетворення від 2-ї до 3-ї ланки (кінцевого ефектора).

Матриці перетворення T_1, T_2, T_3 мають такий вигляд:

$$T_i = \begin{bmatrix} \cos \theta_i - \sin \theta_i \cos \alpha_i & \sin \theta_i \sin \alpha_i & \alpha_i \cos \theta_i \\ \sin \theta_i & \cos \theta_i \cos \alpha_i & -\cos \theta_i \sin \alpha_i \\ 0 & \sin \alpha_i & \cos \alpha_i \\ 0 & 0 & 1 \end{bmatrix}. \quad (2)$$

Для опису динамічного рівня запропоновано використовувати модель Ейлера – Лагранжа [19, 20], оскільки вона дає змогу точно описати динаміку руху триланкового колаборативного робота-маніпулятора, беручи до уваги кінетичну та потенційну енергії кожної ланки, а також вплив зовнішніх сил. Вона є основою для побудови адаптивної системи управління, що зважає на зміни навантаження та взаємодію з людиною в умовах невизначеності. Модель Ейлера – Лагранжа для цього дослідження має вигляд

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = \tau_i, \quad (3)$$

де $L = K - P$ – лагранжіан (різниця кінетичної та потенційної енергії); q_i – кутові положення i -го суглоба; τ_i – момент сили, що прикладається до суглоба.

Кінетичну енергію (K) знаходимо так:

$$K = \frac{1}{2} \dot{q}^T M(q) \dot{q}, \quad (4)$$

де $\dot{q}$ – вектор кутових швидкостей обертання суглобів робота, тобто швидкість зміни положення кожної ланки; $M(q)$ – матриця інерції (масово-інерціальна матриця), що залежить від положення q кожної ланки; вона описує, як маса та інерція розподілені по всій структурі маніпулятора; $\dot{q}^T$ – транспонований вектор швидкостей, що дає змогу обчислити скалярне значення кінетичної енергії в множенні на матрицю інерції та вектор швидкостей.

Потенціальну енергію (P) знаходимо так:

$$P = G(q), \quad (5)$$

де $G(q)$ – вектор гравітаційних сил.

Динамічне рівняння руху необхідне для опису взаємозв'язку між силами, моментами та рухом ланок маніпулятора, що дає змогу точно моделювати його поведінку у відповідь на зовнішні впливи. Воно лежить в основі побудови системи керування, оскільки допомагає формувати відповідні керуючі дії для досягнення бажаних траєкторій з огляду на масово-інерційні властивості та змінну динаміку системи, та має такий вигляд:

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) = \tau, \quad (6)$$

де $M(q)$ – матриця інерції; $C(q, \dot{q})$ – матриця коріолісових і відцентрових сил; $G(q)$ – вектор гравітаційних сил; τ – вектор моментів керування.

Високорівневий адаптивний модуль виконує функцію адаптації швидкості руху маніпулятора залежно від таких параметрів: D – відстан до людини (у метрах); F – навантаження на ефектор (у Ньютонах). Використаємо *Fuzzy Logic*, що формує

коефіцієнт масштабування швидкості $\mu \in [0.1, 1.0]$, який потім використовується в команді швидкості:

$$\dot{q}_{desired} = \mu(D, F) \cdot \dot{q}_{ref}, \quad (7)$$

де $\dot{q}_{desired}$ – адаптивно скоригована швидкість, яку має реалізувати робот-маніпулятор; $\mu(D, F)$ – вихід нечіткої логіки, що визначає ступінь модифікації швидкості залежно від поточної відстані до людини D та рівня навантаження F на виконавчий орган; $\dot{q}_{ref}$ – еталонна швидкість, задана на основі запланованої траєкторії, незважаючи на адаптацію.

На базі виразів (1)–(6) модель комбінованого керування з огляду на *Fuzzy Logic* до суглоба матиме такий вигляд:

$$u(t) = M(q)\ddot{q}_d + C(q, \dot{q})\dot{q}_d + G(q), \quad (9)$$

де $M(q)$ – матриця інерції, що описує залежність динаміки системи від конфігурації суглобів робота; $\ddot{q}_d$ – бажане прискорення суглобів, яке визначається планувальником руху, $\ddot{q}_d = \frac{d}{dt}\dot{q}_d$; $C(q, \dot{q})$ – матриця коріолісових і відцентрових сил, що бере до уваги нелінійні динамічні ефекти під час руху; $\dot{q}_d$ – бажана швидкість суглобів, $\dot{q}_d = \mu(D, F) \cdot \dot{q}_{ref}$; $G(q)$ – вектор гравітаційних моментів, який компенсує силу тяжіння для збереження рівноваги маніпулятора.

Контролер для адаптації швидкості на базі *Fuzzy Logic* використовується для гнучкого регулювання швидкості руху маніпулятора залежно від поточної ситуації в робочому середовищі, зокрема відстані до людини та навантаження на виконавчий орган. Такий підхід є важливим у контексті безпечної та ефективної взаємодії людини з колаборативним роботом. Модель дає змогу уникати жорстких порогових рішень, натомість створює плавні адаптивні переходи між режимами руху та має такий вигляд:

$$v^* = Defuzz\left(\bigcup_{i=1}^N \min(\mu_{D_i}(d), \mu_{L_i}(l), \mu_{V_i}(v))\right), \quad (10)$$

де v^* – адаптивна швидкість руху маніпулятора, що розраховується на основі аналізу нечітких правил; це остаточне числове значення, яке буде використано для керування виконавчим органом; *Defuzz* – операція дефазифікації, тобто перетворення результату об'єднання правил у конкретне числове значення; пропонується реалізація методом центру тяжіння; $\mu_{D_i}(d)$ – функція належності до нечіткої

множини відстані (наприклад: близько, середньо, далеко) в i -му правилі; показує, наскільки значення d (відстань до людини) відповідає лінгвістичному опису; $\mu_{L_i}(l)$ – функція належності до нечіткої множини навантаження (наприклад: низьке, середнє, високе) в i -му правилі; визначає ступінь відповідності поточного навантаження на маніпулятор; $\mu_{V_i}(v)$ – функція належності до нечіткої множини швидкості, що описує бажану швидкість (наприклад: повільна, нормальна, швидка) в i -му правилі; d – відстань до людини, що вимірюється в сантиметрах або метрах; що менша відстань, то більша небезпека, яка береться до уваги в логіці; l – рівень навантаження на маніпулятор у %, що може впливати на стабільність або точність рухів; v – поточне значення швидкості в умовних одиницях або у відсотках від максимальної.

Подана модель (10) дає змогу на основі множини нечітких правил динамічно й контекстно регулювати поведінку маніпулятора, забезпечуючи безпечну співпрацю з людиною без втрати продуктивності.

Для опису реалізації моделі високорівневого керування на основі ієрархічної логіки модифікуємо модель (6):

$$\tau = M(q)\ddot{q}_d + C(q, \dot{q})\dot{q}_d + G(q), \quad \dot{q}_d = \mu(d, l) \cdot \dot{q}_{ref}, \quad (11)$$

де $\mu(d, l) \in [0, 1]$ – вихід нечіткого логічного регулятора, що адаптує бажану швидкість до умов середовища.

Модель адаптації керування (12) на основі спостережень потрібна для того, щоб забезпечити гнучке, ситуаційно залежне та обґрунтоване керування колаборативним роботом-маніпулятором у реальних умовах, де доступ до повної інформації про стан системи часто є обмеженим або неточним.

$$\dot{b}(\dot{s}) = \eta \cdot O(o|\dot{s}, a) \sum_{s \in S} T(\dot{s}|s, a) b(s), \quad (12)$$

де $\dot{b}(\dot{s})$ – оновлене вірування (*belief*), оцінка ймовірності того, що система тепер перебуває у стані $\dot{s}$ після виконання дії a , та отримання спостереження o ; $b(s)$ – поточна оцінка ймовірності того, що система перебуває в стані s до виконання дії; $O(o|\dot{s}, a)$ – функція спостереження, тобто ймовірність отримати спостереження o , якщо система фактично перебуває в стані $\dot{s}$ після виконання дії a ; $T(\dot{s}|s, a)$ – функція переходу, тобто ймовірність того, що система перейде зі стану s у новий стан $\dot{s}$ після виконання дії a ; моделює динаміку навколишнього середовища

та реакцію робота; η – нормалізувальний коефіцієнт, який гарантує, що всі ймовірності в новому $\text{belief } \hat{b}(\dot{s})$ сумуються до 1; використовується для збереження коректної ймовірнісної інтерпретації.

Модель (12) дає змогу роботу, який не має повного доступу до фактичного стану системи, адаптувати керування на основі обмежених спостережень. Так він формує нову стратегію дій, що відповідає поточній ситуації в робочому середовищі.

Запропонована ієрархічна модель має певні переваги: використовує кінематику для обчислення положення в просторі; за допомогою динаміки формує точні моменти управління; інтегрує адаптивний модуль на основі *Fuzzy Logic*, який бере до уваги відстань до людини та рівень навантаження для адаптивного регулювання швидкості руху; підвищує безпеку й гнучкість колаборативного маніпулятора в умовах мінливого довкілля.

Розроблення програми для моделювання адаптивного ієрархічного високорівневого керування триланкового колаборативного робота-маніпулятора

Вибір мови програмування *Python* обґрунтований її високим рівнем абстракції, широкою підтримкою наукових та інженерних бібліотек, зокрема *NumPy*, *SciPy*, *Matplotlib* та *scikit-fuzzy*, що дає змогу ефективно реалізовувати математичні моделі, логіку адаптивного керування та візуалізацію результатів моделювання [21, 22]. *Python* має гнучкий синтаксис, який спрощує інтеграцію модулів штучного інтелекту, таких як *Fuzzy Logic*, нейронні мережі або методи

машинного навчання для розширення функціональності систем керування роботизованими комплексами. Середовище розробки *PyCharm* 2022.2.3 обрано завдяки своїй стабільності, підтримці інтеграції з віртуальними середовищами, зручній навігації по коду, підсвітці синтаксису та відлагодженню, а також можливості роботи з проектами, що містять складні структури даних та графічні бібліотеки. Завдяки цим перевагам поєднання *Python* і *PyCharm* створює оптимальне середовище для розроблення програмного забезпечення в галузі інтелектуального керування роботами.

Для перевірки адекватності розроблених математичних моделей і виразів на базі розробленої програми моделювання адаптивного ієрархічного високорівневого керування триланкового колаборативного робота-маніпулятора застосовуватимемо певні показники. Зокрема D – набір значень відстані до людини в сантиметрах, який рівномірно розподілений у межах від 0 до 200 см. Цей параметр моделює безпечну дистанцію між роботом і людиною, яка впливає на адаптацію швидкості та поведінки робота. Менша відстань вказує на потребу зменшення швидкості руху маніпулятора для забезпечення безпеки. Показник F – набір значень навантаження на виконавчий орган робота у відсотках, який рівномірно варіюється від 0% до 100%. Цей параметр впливає на динаміку руху та керування, оскільки збільшення навантаження потребує адаптації крутного моменту, сили та швидкості, щоб уникнути перенавантаження або втрати точності. Отримані результати моделювання за вказаних вище вхідних параметрів зображені на рис. 2.

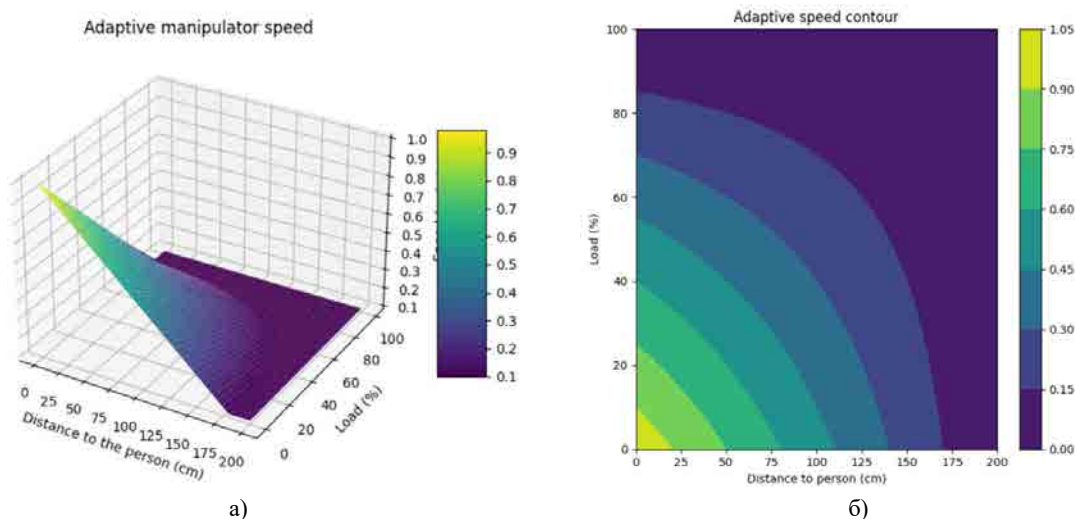


Рис. 2. Отримані результати моделювання: а) 3D-графік адаптивний швидкості маніпулятора;
б) 2D-графік контуру адаптивний швидкості

На 3D-графіку (рис. 2, а) подано залежність адаптивної швидкості маніпулятора від двох ключових параметрів – відстані до людини та навантаження на виконавчий орган. Поверхня, що утворюється в просторі параметрів, демонструє, як система управління знижує швидкість маніпулятора в умовах потенційного ризику для людини або перевантаження конструкції. Якісний аналіз показує, що швидкість руху ефектора є максимальною, коли навантаження є мінімальним, а відстань до людини – максимальною. Це свідчить про те, що робот працює на повну потужність тільки тоді, коли умови є безпечними й не створюють загрози для оточення. Із зменшенням відстані до людини, навіть за невеликого навантаження, швидкість плавно знижується, що відповідає принципам колаборативної безпеки відповідно до концепції Індустрія 5.0. У разі максимального навантаження, навіть за великої відстані, система суттєво знижує швидкість. Це свідчить про те, що взято до уваги механічні обмеження конструкції.

Отже, адаптивна стратегія забезпечує подвійний захист – як від зіткнення з людиною, так і від механічного перенавантаження. Чисельний аналіз показує, що швидкість нормалізується в межах від 0.1 до 1.0, де 1.0 відповідає оптимальному режиму роботи, а значення близькі до 0.1 – захищеним або критичним умовам. Найбільші швидкості спостерігаються в точках, де відстань перевищує 150 см, а навантаження не перевищує 20%. Тоді як за умови відстані меншій ніж 50 см і навантаження понад 80% швидкість суттєво падає. Це узгоджується з логікою нечіткого регулювання, яке гнучко адаптує поведінку робота до зовнішніх умов.

На 2D-графіку контуру адаптивної швидкості маніпулятора (рис. 2, б) візуалізовано зміну швидкості руху виконавчого органа залежно від відстані до людини та навантаження. Кожна лінія контуру відповідає певному рівню швидкості, що дає змогу чітко бачити градієнт змін. Якісний аналіз показує, що максимальна швидкість спостерігається в ділянці, де навантаження є мінімальним, а відстань до людини не значна. Цей результат демонструє, що за відсутності небезпеки зіткнення та механічного перенавантаження система дає змогу маніпулятору працювати на повну швидкість.

Залежно від зростання навантаження або зменшення відстані до людини контури зміщуються до нижчих значень швидкості, що свідчить про

впровадження адаптивної стратегії регулювання. Найбільш насичена зона з мінімальними значеннями швидкості розташована в правому верхньому куті графіка, де поєднуються високе навантаження та незначна відстань – умови, що вимагають максимальної обережності. Чисельно можна оцінити, що швидкість у цій зоні прямує до нуля або приймає мінімальні допустимі значення, що забезпечує безпеку людини та захист механізмів. Візуально видно плавний перехід між різними режимами роботи, що вказує на ефективну роботу нечіткого регулятора, реалізованого на основі моделі високорівневого адаптивного керування. Такий підхід повністю відповідає вимогам стандартів безпеки роботизованих систем у концепції Індустрія 5.0. Система здатна не лише реагувати на зміну умов, а й забезпечувати оптимальний баланс між продуктивністю та безпекою, адаптуючи поведінку маніпулятора в реальному часі.

Heatmap оптимальної дії, зображений на рис. 3, демонструє, як система залежно від відстані між виконавчим елементом робота та людиною обирає безпечну й водночас ефективну дію – зупинку, уповільнення або продовження руху.

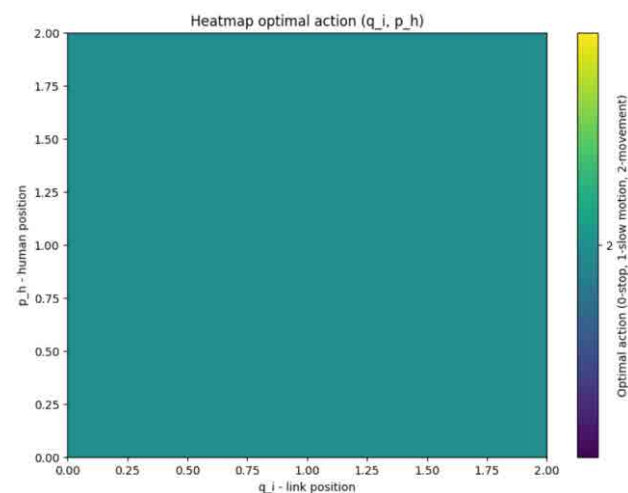


Рис. 3. *Heatmap* оптимальної дії

На графіку *Heatmap* оптимальної дії (рис. 3) видно, що система в усіх точках простору станів обирає одну й ту саму дію, що візуалізовано єдиним кольором по всій площині. Це свідчить про домінування одного рішення – рух (2), що може означати недостатню чутливість моделі до змін позиції маніпулятора та людини. Така ситуація ймовірна, якщо функція винагороди надто сильно стимулює ефективність, ігноруючи безпекові аспекти

або якщо коефіцієнти у функції винагороди некоректно налаштовані. Отриманий результат свідчить про необхідність тоншого балансування вагових коефіцієнтів α та β або модифікації функції переходу чи винагороди для забезпечення адаптивної поведінки, особливо в критичних точках взаємодії з людиною.

3D-графік функції винагороди, зображений на рис. 4, показує, що максимальна винагорода досягається за умови більшої дистанції до людини та дії "рух", але з огляду на вагові коефіцієнти безпеки та ефективності.

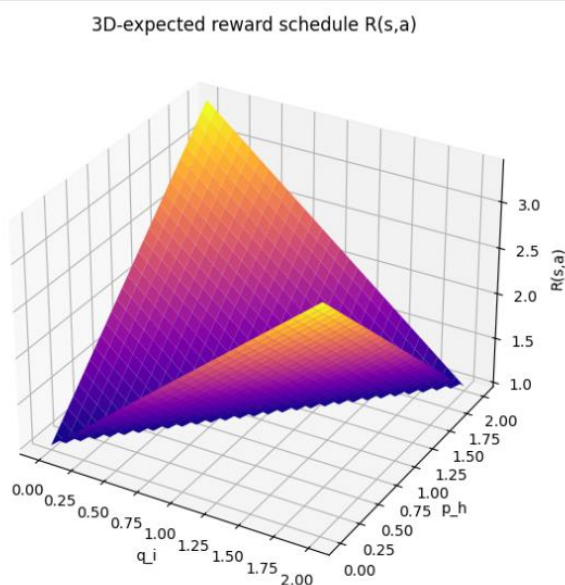


Рис. 4. 3D-графік винагороди

На поданому 3D-графіку очікуваної функції винагороди $R(s,a)$ спостерігається лінійне зростання винагороди зі збільшенням різниці між положенням маніпулятора q_i та положенням людини p_h , що свідчить про прямий вплив безпечної відстані на підвищення загального коефіцієнта винагороди. Максимальні значення $R(s,a)$ досягаються в умовах максимальної віддаленості між виконавчим органом робота та оператором, що відповідає принципу пріоритету безпеки в середовищах співпраці. Мінімальні значення фіксуються в разі збігу $q_i \approx p_i$, де ймовірність ризику є найвищою, тому адаптивна система віддає перевагу консервативній поведінці. Така залежність підтверджує адекватність побудованої функції винагороди до вимог безпечного керування. Результати експерименту також демонструють, що модель чутлива до зміни просторових параметрів, що дає змогу використовувати її в реальних сценаріях із присутністю людини.

Порівняльний аналіз швидкостей (рис. 5, а) демонструє адаптивну поведінку робота: порівняно з фіксованою швидкістю адаптивна дає змогу гнучко реагувати на зміну зовнішніх факторів. Графік накопиченого енергоспоживання (рис. 5, б) чітко ілюструє, що адаптивне керування сприяє заощадженню енергії в довгостроковій перспективі. Це підтверджує доцільність використання розроблених моделей у динамічному середовищі спільної взаємодії людини та робота.

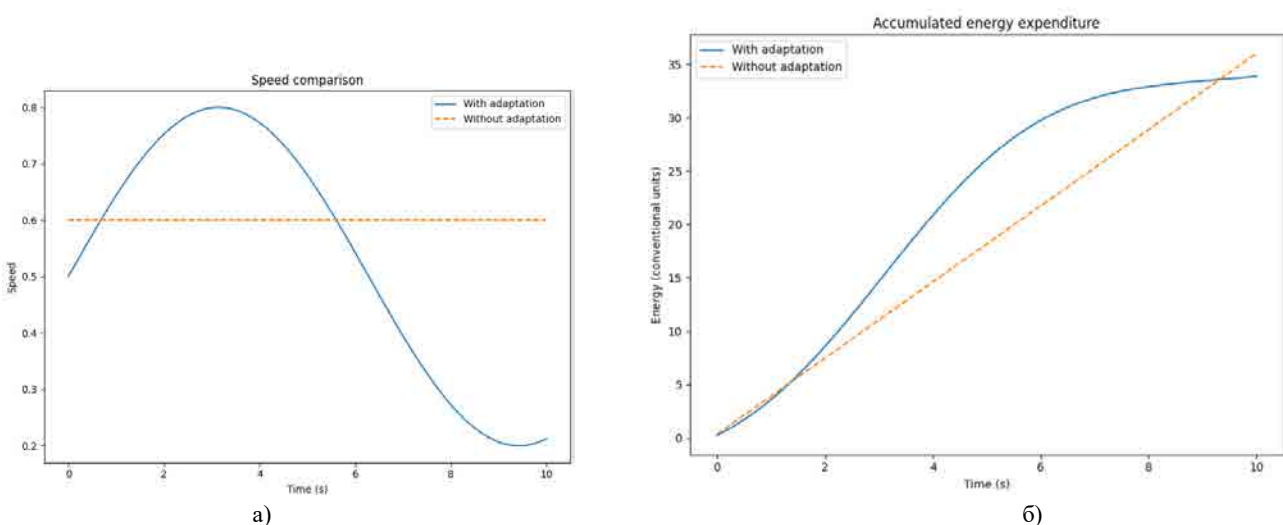


Рис. 5. Графіки порівняльного аналізу швидкостей та енергоспоживання:
а) порівняльна динаміка швидкості та б) накопиченого енергоспоживання

На графіку порівняльної динаміки швидкості (рис. 5, а), чітко видно переваги адаптивного керування над фіксованою стратегією. Крива *With adaptation* демонструє гнучке реагування на зміни умов середовища, забезпечуючи плавне регулювання швидкості в діапазоні приблизно від 0.2 до 0.8 одиниць, що дає змогу підвищити безпеку та ефективність взаємодії з людиною. Тоді як *Without adaptation* є постійним значенням близько 0.6, яке ігнорує зміни зовнішніх факторів, таких як відстань до людини чи навантаження.

Отже, адаптивне керування забезпечує більш збалансований і безпечний режим роботи маніпулятора, що особливо важливо для середовищ співпраці з людиною. Експериментальні результати підтверджують доцільність використання адаптивних моделей у реальних умовах.

На графіку накопиченого енергоспоживання (рис. 5, б) видно, що система з адаптивним керуванням спочатку демонструє більші витрати енергії, проте після 7 с спостерігається уповільнення зростання споживаної енергії. Порівняно з неадаптивним підходом, для якого властиве лінійне й стабільне зростання, адаптивна система поступово оптимізує витрати, реагуючи на зміну умов навколишнього середовища. Це свідчить про те, що, попри вищі енергетичні витрати на початковому етапі, адаптивне керування знижує сумарне навантаження на приводи у фазі стабілізації. Експеримент підтверджує, що адаптивний підхід забезпечує більш раціональне використання енергії в динамічних умовах, що є важливим для тривалої та ефективної роботи колаборативного маніпулятора.

Висновки

Під час дослідження розроблено математичну модель адаптивного ієрархічного високорівневого керування для триланкового колаборативного робота-маніпулятора, що поєднує методи класичної динаміки, нечіткої логіки та принципи безпечної взаємодії людини з роботом. Результати підтверджують ефективність запропонованого підходу в умовах мінливого навантаження на виконавчий орган та варіативної відстані до оператора, що критично важливо в контексті впровадження концепції Індустрія 5.0, де пріоритетом є гнучка взаємодія між людиною та інтелектуальними системами. Комбінована модель керування дала змогу досягти адаптації швидкісного профілю руху в реальному часі, забезпечивши баланс між безпекою, точністю та енергоефективністю. Графічні результати демонструють передбачувану поведінку системи в різних ситуаціях, що підтверджує практичну реалізованість моделі. Реалізація математичного апарату у вигляді програмного модуля мовою *Python* із використанням середовища *PyCharm* забезпечує зручність для подальшої модифікації, інтеграції у віртуальні симулятори або фізичні робототехнічні комплекси. Рекомендовано впроваджувати розроблену модель у системи симуляційного навчання, робототехнічні тренажери та цифрові двійники для колаборативних роботів, а також використовувати її як базу для досліджень у сфері предиктивного та когнітивного керування в індустріальному середовищі.

Список літератури

1. Rijwani T., Kumari S., Srinivas R., Abhishek K., Iyer G., Vara H., Gupta M. Industry 5.0: A review of emerging trends and transformative technologies in the next industrial revolution. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 19(2), P. 667–679. 2025. DOI: <https://doi.org/10.1007/s12008-024-01943-7>
2. Sun X., Song Y. Unlocking the Synergy: Increasing productivity through Human-AI collaboration in the industry 5.0 Era. *Computers & Industrial Engineering*, 200, 2025. 110657 p. DOI: <https://doi.org/10.1016/j.cie.2024.110657>
3. Gonçalves A., Pereira T., Lopes D., Cunha F., Lopes F., Coutinho F., Ferreira N. M. F. Enhancing Nut-Tightening Processes in the Automotive Industry: Integration of 3D Vision Systems with Collaborative Robots. *Automation*, 6(1), 8. 2025. DOI: <https://doi.org/10.3390/automation6010008>
4. Nevliudov I., Yevsieiev V., Baker J. H., Ahmad M. A., Lyashenko V. Development of a cyber design modeling declarative Language for cyber physical production systems. *Journal of Mathematical and Computational Science*, 11(1), 2020. P. 520–542. DOI: <https://doi.org/10.28919/jmcs/5152>
5. Yuan F., Ren G., Deng Q., Wang X. Steam Generator Maintenance Robot Design and Obstacle Avoidance Path Planning. *Sensors*, 25(2), 2025. 514 p. DOI: <https://doi.org/10.3390/s25020514>
6. Recker T., Raatz A. Design and control of flexible handling systems based on mobile cooperative multi-robot-systems. *CIRP Annals*. 2025. DOI: <https://doi.org/10.1016/j.cirp.2025.04.059>

7. Pan Y. J., Buchanan S., Chen Q., Wan L., Chen N., Forbrigger S., Smith S. Survey on recent advances in planning and control for collaborative robotics. *IEEE Journal of Industry Applications*, 14(2), 2025. P. 139–151. DOI: <https://doi.org/10.1541/ieejia.24005652>
8. Невлюдов І. ІІ., Євсєєв В. В., Гурін Д. В. Model development of dynamic representation a model description parameters for the environment of a collaborative robot manipulator within the industry 5.0 framework. *Системи управління, навігації та зв'язку. Збірник наукових праць*, 1(79), 2025. P. 42–48. DOI: <https://doi.org/10.26906/SUNZ.2025.1.42-48>
9. Hu S., Wan Y., Liang X. Adaptive nonsingular fast terminal sliding mode trajectory tracking control for robotic manipulators with model feedforward compensation. *Nonlinear Dynamics*, 2025. P. 1–19. DOI: <https://doi.org/10.3390/electronics11223672>
10. Ben Hazem Z., Guler N., Altaif A. H. A study of advanced mathematical modeling and adaptive control strategies for trajectory tracking in the Mitsubishi RV-2AJ 5-DOF Robotic Arm. *Discover Robotics*, 1(1), 2. 2025. DOI: <https://doi.org/10.1007/s44430-025-00001-5>
11. Zhou H., Lin X. Intelligent redundant manipulation for long-horizon operations with multiple goal-conditioned hierarchical learning. *Advanced Robotics*, 39(6), 2025. P. 291–304. DOI: <https://doi.org/10.1080/01691864.2025.2475295>
12. Sun Z., Pang B., Yuan X., Xu X., Song Y., Song R., Li Y. Hierarchical reinforcement learning with curriculum demonstrations and goal-guided policies for sequential robotic manipulation. *Engineering Applications of Artificial Intelligence*, №153, 2025. 110866 p. DOI: <https://doi.org/10.1016/j.engappai.2025.110866>
13. Wei Z., Luo X., Liu C. Hierarchical Temporal Logic Task and Motion Planning for Multi-Robot Systems. *Accepted by Robotics: Science and Systems*. 2025. DOI: <https://doi.org/10.48550/arXiv.2504.18899>
14. Fan S., Meng M., Fu Y., Deng C. Distributed adaptive tracking control for fuzzy nonlinear MASs under round-robin protocol. *IEEE Transactions on Fuzzy Systems*. 2025. DOI: <https://doi.org/10.1109/TFUZZ.2025.3525989>
15. Liu H., Xu S., Song J., Ma B. Fuzzy control of uncertain dual-arm cooperative robots with quantitative performance and input deadzone. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, № 47(3), 126 p. 2025. DOI: <https://doi.org/10.1007/s40430-024-05377-w>
16. Urrea C. Hybrid Fault-Tolerant Control in Cooperative Robotics: Advances in Resilience and Scalability. In *Actuators*. Vol. 14, No. 4, 177 p. 2025. DOI: <https://doi.org/10.3390/act14040177>
17. Babiola A. Recurrent Neural Networks for Real-Time Position Estimation of Robotic Arm End-Effector Coordinates. In *2025 International Conference on Artificial Intelligence and Data Engineering (AIDE)*. P. 34–40. 2025. DOI: <https://doi.org/10.1109/AIDE64228.2025.10986815>
18. Alwardat M. Y., Alwan H. Geometric jacobians derivation and kinematic singularity analysis for 6-dof robotic manipulator. *International Journal of Advanced Research in Computer Science*, №16(1). 2025. DOI: <https://doi.org/10.26483/ijarcs.v16i1.7178>
19. Wang Z., Xu L., Zhu X., Quan L., Chen W. H., Xu L., Ding S. Euler-Lagrange-Model-Based Torque Assignment Control for Dual In-Wheel PM Motors With Voltage Vectors Integrated Modulation. *IEEE Transactions on Industrial Electronics*. 2025. P. 1–14. DOI: <https://doi.org/10.1109/TIE.2025.3555041>
20. Lavaill M., Chen X., Heinrich S., Pivonka P., Leyendecker S. Muscle path predictions using a discrete geodesic Euler-Lagrange model in constrained optimisation: comparison with OpenSim and experimental data. *Multibody System Dynamics*, P. 1–23. 2025. DOI: <https://doi.org/10.1007/s11044-025-10055-3>
21. Tran H. C., Tran A. D., Le K. H. DetectVul: A statement-level code vulnerability detection for Python. *Future Generation Computer Systems*, № 163, 107504 p. 2025. DOI: <https://doi.org/10.1016/j.future.2024.107504>
22. Mechri A., Ferrag M. A., Debbah M. Secureqwen: Leveraging llms for vulnerability detection in python codebases. *Computers & Security*, № 148, 104151 p. 2025. DOI: <https://doi.org/10.1016/j.cose.2024.104151>

References

1. Rijwani, T., Kumari, S., Srinivas, R., Abhishek, K., Iyer, G., Vara, H., Gupta, M. (2025), "Industry 5.0: A review of emerging trends and transformative technologies in the next industrial revolution". *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 19(2), P. 667–679. DOI: <https://doi.org/10.1007/s12008-024-01943-7>
2. Sun, X., Song, Y. (2025), "Unlocking the Synergy: Increasing productivity through Human-AI collaboration in the industry 5.0 Era". *Computers & Industrial Engineering*, 200, 110657 p. DOI: <https://doi.org/10.1016/j.cie.2024.110657>
3. Gonçalves, A., Pereira, T., Lopes, D., Cunha, F., Lopes, F., Coutinho, F., Ferreira, N. M. F. (2025), "Enhancing Nut-Tightening Processes in the Automotive Industry: Integration of 3D Vision Systems with Collaborative Robots". *Automation*, 6(1), 8. DOI: <https://doi.org/10.3390/automation6010008>

4. Nevliudov, I., Yevsieiev, V., Baker, J. H., Ahmad, M. A., Lyashenko, V. (2020), "Development of a cyber design modeling declarative Language for cyber physical production systems". *Journal of Mathematical and Computational Science*, 11(1), P. 520–542. DOI: <https://doi.org/10.28919/jmcs/5152>
5. Yuan, F., Ren, G., Deng, Q., Wang, X. (2025), "Steam Generator Maintenance Robot Design and Obstacle Avoidance Path Planning". *Sensors*, 25(2), 514 p. DOI: <https://doi.org/10.3390/s25020514>
6. Recker, T., Raatz, A. (2025), "Design and control of flexible handling systems based on mobile cooperative multi-robot-systems". *CIRP Annals*. DOI: <https://doi.org/10.1016/j.cirp.2025.04.059>
7. Pan, Y. J., Buchanan, S., Chen, Q., Wan, L., Chen, N., Forbrigger, S., Smith, S. (2025), "Survey on recent advances in planning and control for collaborative robotics". *IEEEJ Journal of Industry Applications*, 14(2), P. 139–151. DOI: <https://doi.org/10.1541/ieejia.24005652>
8. Nevlyudov, I., Yevsieiev, V., Gurin, D. (2025), "Model development of dynamic representation a model description parameters for the environment of a collaborative robot manipulator within the industry 5.0 framework". *Control, Navigation and Communication System*. 1(79), P. 42–48. DOI: <https://doi.org/10.26906/SUNZ.2025.1.42-48>
9. Hu, S., Wan, Y., Liang, X. (2025), "Adaptive nonsingular fast terminal sliding mode trajectory tracking control for robotic manipulators with model feedforward compensation". *Nonlinear Dynamics*, P. 1–19. DOI: <https://doi.org/10.3390/electronics11223672>
10. Ben, Hazem Z., Guler, N., Altaif, A. H. (2025), "A study of advanced mathematical modeling and adaptive control strategies for trajectory tracking in the Mitsubishi RV-2AJ 5-DOF Robotic Arm". *Discover Robotics*, 1(1), 2. DOI: <https://doi.org/10.1007/s44430-025-00001-5>
11. Zhou, H., Lin, X. (2025), "Intelligent redundant manipulation for long-horizon operations with multiple goal-conditioned hierarchical learning". *Advanced Robotics*, 39(6), P. 291–304. DOI: <https://doi.org/10.1080/01691864.2025.2475295>
12. Sun, Z., Pang, B., Yuan, X., Xu, X., Song, Y., Song, R., Li, Y. (2025), "Hierarchical reinforcement learning with curriculum demonstrations and goal-guided policies for sequential robotic manipulation". *Engineering Applications of Artificial Intelligence*, №153, 110866 p. DOI: <https://doi.org/10.1016/j.engappai.2025.110866>
13. Wei, Z., Luo, X., Liu, C. (2025), "Hierarchical Temporal Logic Task and Motion Planning for Multi-Robot Systems". *Accepted by Robotics: Science and Systems*. 2025. DOI: <https://doi.org/10.48550/arXiv.2504.18899>
14. Fan, S., Meng, M., Fu, Y., Deng, C. (2025), "Distributed adaptive tracking control for fuzzy nonlinear MASs under round-robin protocol". *IEEE Transactions on Fuzzy Systems*. 2025. DOI: <https://doi.org/10.1109/TFUZZ.2025.3525989>
15. Liu, H., Xu, S., Song, J., Ma, B. (2025), "Fuzzy control of uncertain dual-arm cooperative robots with quantitative performance and input deadzone". *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, № 47(3), 126 p. DOI: <https://doi.org/10.1007/s40430-024-05377-w>
16. Urrea, C. (2025), "Hybrid Fault-Tolerant Control in Cooperative Robotics: Advances in Resilience and Scalability". *In Actuators*. Vol. 14, No. 4, 177 p. DOI: <https://doi.org/10.3390/act14040177>
17. Babiola, A. (2025), "Recurrent Neural Networks for Real-Time Position Estimation of Robotic Arm End-Effector Coordinates". *In 2025 International Conference on Artificial Intelligence and Data Engineering (AIDE)*. P. 34–40. DOI: <https://doi.org/10.1109/AIDE64228.2025.10986815>
18. Alwardat, M. Y., Alwan, H. (2025), "Geometric jacobians derivation and kinematic singularity analysis for 6-dof robotic manipulator". *International Journal of Advanced Research in Computer Science*, №16(1). DOI: <https://doi.org/10.26483/ijarcs.v16i1.7178>
19. Wang, Z., Xu, L., Zhu, X., Quan, L., Chen, W. H., Xu, L., Ding, S. (2025), "Euler-Lagrange-Model-Based Torque Assignment Control for Dual In-Wheel PM Motors With Voltage Vectors Integrated Modulation". *IEEE Transactions on Industrial Electronics*. P. 1–14. DOI: <https://doi.org/10.1109/TIE.2025.3555041>
20. Lavaill, M., Chen, X., Heinrich, S., Pivonka, P., Leyendecker, S. (2025), "Muscle path predictions using a discrete geodesic Euler-Lagrange model in constrained optimisation: comparison with OpenSim and experimental data". *Multibody System Dynamics*, P. 1–23. DOI: <https://doi.org/10.1007/s11044-025-10055-3>
21. Tran, H. C., Tran, A. D., Le, K. H. (2025), "DetectVul: A statement-level code vulnerability detection for Python". *Future Generation Computer Systems*, № 163, 107504 p. DOI: <https://doi.org/10.1016/j.future.2024.107504>
22. Mechri, A., Ferrag, M. A., Debbah, M. (2025), "Secureqwen: Leveraging llms for vulnerability detection in python codebases". *Computers & Security*, № 148, 104151 p. 2025. DOI: <https://doi.org/10.1016/j.cose.2024.104151>

Відомості про авторів / About the Authors

Невлюдов Ігор Шакирович – доктор технічних наук, професор, Харківський національний університет радіоелектроніки, завідувач кафедри комп'ютерно-інтегрованих технологій, автоматизації та робототехніки, м. Харків, Україна; e-mail: igor.nevliudov@nure.ua; ORCID ID: <https://orcid.org/0000-0002-9837-2309>.

Євсєєв Владислав В'ячеславович – доктор технічних наук, професор, Харківський національний університет радіоелектроніки, професор кафедри комп'ютерно-інтегрованих технологій, автоматизації та робототехніки, м. Харків, Україна; e-mail: vladyslav.yevsieiev@nure.ua; ORCID ID: <https://orcid.org/0000-0002-2590-7085>.

Максимова Світлана Святославівна – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри комп'ютерно-інтегрованих технологій, автоматизації та робототехніки, м. Харків, Україна; e-mail: svitlana.milyutina@nure.ua; ORCID ID: <https://orcid.org/0000-0002-1375-9337>

Артиух Роман Володимирович – кандидат технічних наук, Державне підприємство "Південний державний проектно-конструкторський і науково-дослідний інститут авіаційної промисловості", директор, Харківський національний університет радіоелектроніки, доцент кафедри комп'ютерно-інтегрованих технологій, автоматизації та робототехніки, м. Харків, Україна; e-mail: roman.artiuh77@gmail.com; ORCID ID: <https://orcid.org/0000-0002-5129-2221>

Nevliudov Igor – Dr. Sc. (Engineering), Professor, Kharkiv National University of Radioelectronics, Head the Department of Computer-Integrated Technologies, Automation and Robotics, Kharkiv, Ukraine.

Yevsieiev Vladyslav – Dr. Sc. (Engineering), Professor, Kharkiv National University of Radioelectronics, Professor the Department of Computer-Integrated Technologies, Automation and Robotics, Kharkiv, Ukraine.

Maksymova Svitlana – Candidate of Technical Sciences, Associate Professor, Kharkiv National University of Radio Electronics, Associate Professor of the Department of Computer Integrated Technologies, Automation and Robotics, Kharkiv, Ukraine.

Artiukh Roman – Candidate of Technical Sciences, National Design and Research Institute of Aerospace Industries, Kharkiv National University of Radio Electronics, Associate Professor of the Department of Computer Integrated Technologies, Automation and Robotics, Kharkiv, Ukraine.

MATHEMATICAL MODEL OF ADAPTIVE HIERARCHICAL HIGH-LEVEL CONTROL FOR A 3-LINK COLLABORATIVE ROBOT-MANIPULATOR

In the context of the development of the Industry 5.0 concept, which envisages close interaction between humans and intelligent automated systems, research into mechanisms for flexible and safe control of collaborative robot manipulators is becoming particularly important. **The objective of this study is** to construct a mathematical model of hierarchical high-level control for a three-link collaborative robot that allows the manipulator's behavior to be adapted to dynamic changes in the production environment and the presence of humans in the work area. The purpose of this article is to develop a mathematical model that combines the classical approach to modeling dynamics with the following methods: using Euler-Lagrange equations and modern adaptive control algorithms implemented through fuzzy logic, which allows optimizing the speed and trajectory of the executive bodies under variable load and partial uncertainty of the environment. The proposed model implements a multi-level control architecture that takes into account not only the physical parameters of the manipulator, but also the cognitive aspects of interaction with the operator and the environment by constructing a Fuzzy-Rule-based controller. The article presents the formalization of mathematical expressions, describes all levels of the control system, their purpose and interconnection, and performs numerical modeling using Python and PyCharm. **The results** of the modeling confirmed the effectiveness of the proposed system in the presence of humans, ensuring a safe distance, adaptation to the load, and compliance with the specified production parameters. **Conclusions** indicate that the proposed model can significantly improve the quality of manipulator control in Industry 5.0 conditions, especially when working in hybrid environments. The developed model can be adapted to different types of collaborative robots, which opens up prospects for its further improvement, in particular through integration with neural network approaches, deep learning, and multi-agent systems within the framework of distributed intelligent manufacturing.

Keywords: Industry 5.0; collaborative robot; mathematical model; hierarchical control; fuzzy logic; adaptive control; manipulator; PyCharm; Python; Fuzzy Logic.

Бібліографічні описи / Bibliographic descriptions

Невлюдов І. Ш., Євсєєв В. В., Максимова С. С., Артиух Р. В. Математична модель адаптивного ієрархічного високорівневого керування триланкового колаборативного робота-маніпулятора. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 2 (32). С. 58–68. DOI: <https://doi.org/10.30837/2522-9818.2025.2.058>

Nevliudov, I., Yevsieiev, V., Maksymova, S., Artiukh, R. (2025), "Mathematical model of adaptive hierarchical high-level control of a three-link collaborative robot-manipulator", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (32), P. 58–68. DOI: <https://doi.org/10.30837/2522-9818.2025.2.058>

В. Поддубний, О. Сєверінов

МЕТОД АВТЕНТИФІКАЦІЇ КОРИСТУВАЧІВ В ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ СИСТЕМАХ ІЗ ЗАСТОСУВАННЯМ НУЛЬОВОГО ВОДЯНОГО ЗНАКА

Предметом дослідження є системи контролю доступом в інформаційно-комунікаційних мережах. **Мета роботи** – розроблення методу автентифікації користувачів в інформаційно-комунікаційних мережах як системи контролю доступом з використанням нульового водяного знака. **Завдання:** розроблення методу автентифікації користувачів із застосуванням нульового водяного знака; тестування запропонованого методу; висування пропозицій та рекомендацій щодо подальшого впровадження методу автентифікації. Для виконання окреслених завдань використовувалися такі **методи:** комп’ютерне моделювання запропонованого методу для дослідження його характеристик; емпіричні методи – упровадження методу та спостереження за результатами перетворень; статистичні методи – обчислення показників нормальної кореляції та пікового значення сигнал / шум, коефіцієнтів Спірмена. **Досягнуті результати:** уперше запропоновано метод автентифікації користувачів в інформаційно-комунікаційній системі, оснований на унікальному алгоритмі нульового водяного знака; сформовано пропозиції щодо подальшого впровадження методу; протестовано розроблений метод. **Висновки.** У межах роботи запропоновано метод автентифікації, оснований на алгоритмі нульового водяного знака, для використання як системи контролю доступом на промислових підприємствах. Метод було протестовано, результати тестування доводять його роботоспроможність та ефективність. Запропонований метод потенційно може зменшити вартість двофакторної автентифікації в умовах розгортання систем автентифікації у великих підприємствах.

Ключові слова: системи контролю доступу; нульовий водяний знак; метод; автентифікація; зображення.

Вступ

Нульові водяні знаки – це технологія, що використовується для забезпечення авторства цифрових зображень. Проте новизна цієї технології зумовлює безліч проблем за умови їх практичного застосування. Так, наприклад, нульові водяні знаки наразі досить погано працюють у використанні в подібних зображеннях (зокрема в медицині) [1].

Одним із можливих варіантів застосування нульових водяних знаків є завдання автентифікації користувачів. Ця робота спрямована на демонстрацію методу нульового водяного знака для використання в схемі двофакторної автентифікації. Така схема будуватиметься на двох факторах, а саме: знання чогось (пароль), володіння чимось (зображення). Додатковим фактором є знання того, чи є фото основною інформацією, оскільки зловмисник не може чітко визначити ключову інформацію. Ця схема є експериментальною через досить обмежене вивчення нульових водяних знаків для автентифікації та потребує доопрацювання й подальшого дослідження алгоритмів і методів.

Запропонований метод може бути використаний у схемах автентифікації на промислових підприємствах для автентифікації користувачів як

системи контролю доступом. Метод має деякі переваги порівняно з поширеними методами багатофакторної автентифікації, зокрема: відсутність персональної інформації (на відміну від біометрії), відсутність спеціальних токенів чи карток, використання як мінімум двох факторів, висока мобільність та зручність для користувача. Перелічені переваги роблять його більш дешевою альтернативою для автентифікації користувачів в інформаційно-комунікаційних системах.

Аналіз проблеми

У багатьох вимогах і практиках використання багатофакторної автентифікації (як мінімум двофакторної) необхідне для забезпечення надійного підтвердження особистостей користувачів. Так, наприклад, відповідно до наказу Адміністрації Держспецзв’язку від 24.06.2024 № 317 "Про визначення Базового профілю безпеки інформації", використання двофакторної автентифікації призначене для ідентифікації та автентифікації користувачів, які мають привілейовані облікові записи [2].

Стандарт ІЕС 62443, що встановлює вимоги промислової інформаційної безпеки, також рекомендує впровадження багатофакторної

автентифікації для підвищення рівня безпеки доступу користувачів [3].

Системи контролю доступом – необхідний елемент, що забезпечує автентифікацію користувачів для доступу до приміщень чи засобів. Системи контролю доступом необхідні, оскільки традиційні системи фізичного доступу вже не задовольняють вимоги з безпеки та контролю [4]. Великі підприємства зазвичай потребують більш складних і коштовних систем контролю доступу через значну кількість приміщень та персоналу. Особливо це актуально для впровадження двофакторної автентифікації [5]. Методи та схеми двофакторної автентифікації під час розгортання в промислових підприємствах зазвичай потребують виділення значної кількості фінансів та людських годин, оскільки постає необхідність додаткових засобів (сканерів біометрії, токенів, карток, адміністрування).

У роботі [6] запропоновано загальну схему автентифікації, основу на алгоритмі нульового водяного знака. Більш деталізований опис цієї схеми подано в праці [7]. Крім того, в ній проаналізовано перетворення цифрових зображень з метою їх використання в алгоритмах нульового водяного знака в процесі автентифікації. Одними з найкращих алгоритмів є, зокрема, *DWT* та *K-means*, тому було прийнято рішення використати *DWT* як перетворення першого рівня, а *K-means* – другого рівня.

Метод застосовує як контейнер RGB-зображення, що описано в роботі [8]. У цій праці описано метод цифрового знака та його використання.

Розроблення методу автентифікації із застосуванням нульового водяного знака є перспективним завданням у зв'язку з можливістю зниження вартості автентифікації користувачів на виробничих підприємствах, адже водяний знак не потребує спеціалізованих носіїв інформації чи спеціальних сканерів для її оброблення. Запропонований метод інтегрується в системи контролю доступом та призначений для автентифікації користувачів за умови їх доступу до контрольованих об'єктів.

Алгоритми оброблення цифрових зображень

Дискретне вейвлет-перетворення (*DWT*) – це вейвлет-перетворення, для якого вейвлети дискретно відбираються. Як і в інших вейвлет-перетвореннях,

ключовою перевагою, на відміну від перетворення Фур'є, є часова роздільна здатність: воно фіксує як частоту, так і інформацію про місце розташування (місце розташування в часі).

DWT розділяє зображення на високочастотні та низькочастотні компоненти.

LL – низькочастотний піддіапазон, що використовується для подальших перетворень.

LH – піддіапазон, що виділяє горизонтальні особливості оригінального зображення.

HL – піддіапазон, що виділяє вертикальні особливості оригінального зображення.

HH – піддіапазон, що виділяє діагональні особливості оригінального зображення [7].

На рис. 1 показаний приклад перетворення *DWT* (перетворення Хаара) для монохромного зображення.



Рис. 1. Приклад перетворення Хаара

У статистиці та машинному навчанні кластеризація *k*-середніх – це метод кластерного аналізу, метою якого є розподілення *n* спостережень на *k* кластерів, у яких кожне спостереження належить до кластера з найближчим середнім.

Він подібний до алгоритму очікування-максимізації для сумішей Гаусса тим, що обидва вони намагаються знайти центри природних кластерів у даних, а також у підході ітераційного уточнення, який використовують ці алгоритми.

Алгоритм передбачає кілька кроків.

1. *K* точок розміщуються в просторі, що є об'єктами, які групуються. Ці точки – початкові центроїди групи.

2. Кожен об'єкт призначається до групи, яка має найближчий центроїд.

3. Коли всі об'єкти будуть призначені, перераховується положення *K* центроїдів.

4. Повторюються кроки 2 і 3, доки центроїди не перестануть рухатися. Це призводить до поділу об'єктів на групи, з яких можна обчислити показник, що потрібно мінімізувати.

Приклад використання перетворення зображено на рис. 2 [10].



Рис. 2. Приклад перетворення *K-means*

Опис методу

Наведемо порядок формування водяного знака (W).

1. Користувач має в розпорядженні зображення (O) та пароль ($PASS$).

2. Алгоритм нульового знака отримує зображення в цифровому вигляді та пароль.

3. Зображення (O) розкладається на спектри, отримуючи зображення (O_r , O_g , O_b). Ці зображення є матрицями значень пікселів від 0 до 255.

4. Після розкладання зображення (O_r , O_g , O_b) підлягають перетворенню *DWT* (як представник *DWT*-перетворень використовується перетворення Хаара). Перетворення здійснюється поблоково для блоків 8×8 . Перетворення *DWT* формує набір значень: LL , LH , HL , HH , з яких використовується лише LL (формуючи для кожного спектра LL_r , LL_g , LL_b).

5. Для кожного LL_r , LL_g , LL_b виконується перетворення *K-means*, формуючи, відповідно, K_r , K_g , K_b . Параметри перетворення *K-means*: кількість кластерів – 5, толерантність – 0.05.

6. Паралельно відбувається формування матриці з паролем, для цього пароль $PASS$ перетворюється за допомогою функції гешування $SHA512$, формуючи геш H_{pass} .

7. Геш H_{pass} записується у вигляді вектора з 64 значень V_h .

8. Вектор V_h записується 64 рази, формуючи матрицю $Matrix_{pass}$. Отримана матриця має розмір 64×64 .

9. Матриця $Matrix_{pass}$ перемножується поблоково (розмір блоку становить) на матриці K_r , K_g , K_b , формуючи M_r , M_g , M_b .

10. Формується параметр b для *Swish*-перетворення, значення формується з байтів гешу H_{pass} . Для цього обирається 10, 20, 30 та 40-й байт, формуючи цим значення int , $((int/int_{max})+1) \cdot 10$.

11. Для кожного значення з матриць M_r , M_g , M_b виконується *Swish*-перетворення, унаслідок чого формуються S_r , S_g , S_b .

12. S_r , S_g , S_b додаються, формуючи кінцеву матрицю W .

Цей водяний знак зберігається в базі даних системи контролю доступом як "еталонне значення", з яким відбуваються порівняння під час автентифікації. Тобто формування водяного знака – це реєстрація користувача в системі. У подальшому під час повторної авторизації користувач проходить процедуру формування водяного знака з наданими параметрами (пароль і зображення), а система контролю доступом порівнює цей водяний знак з еталонним.

Для підвищення швидкості перевірки можна окремо перевіряти геш-значення паролю, оскільки швидкість формування функції гешу значно вища, ніж швидкість роботи алгоритму нульового водяного знака. Проте в такому разі необхідно використовувати інший алгоритм гешування, що відрізняється від *SHA-512*, оскільки компрометація гешу зменшить стійкість методу (тому що злоумисник буде знати коефіцієнти *Swish* -перетворення).

На рис. 3 зображена загальна схема роботи методу автентифікації з використанням алгоритму нульового водяного знака.

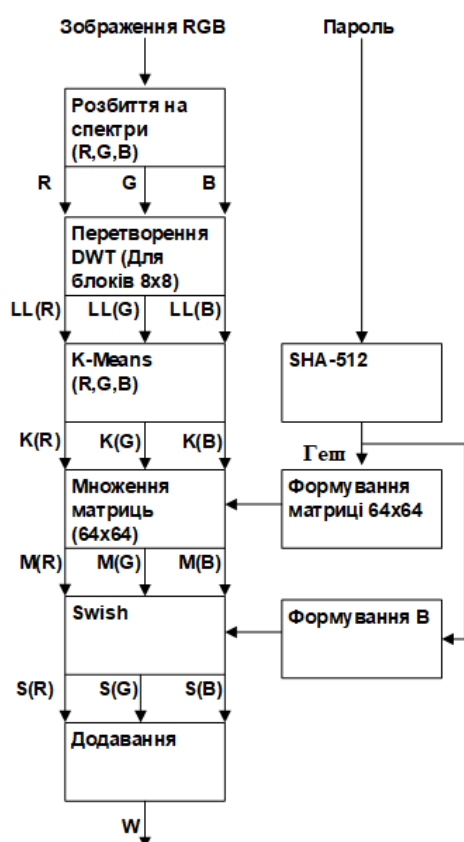


Рис. 3. Схема роботи методу автентифікації з використанням нульового водяного знака

У методі застосовано -функцію, що описується виразом

$$F(x) = \frac{x}{1 + e^{-bx}}, \quad (1)$$

де b – константа або параметр, який залежить від типу моделі. У запропонованому методі цей параметр здається частиною геш-значення.

У методі автентифікації це перетворення виконує функцію згладжування різких змін параметрів зображення. Також *Swish* виконує роль ускладнення оберненості методу внаслідок того, що пошук оберненої функції $(F(x))'$ є в декілька разів складнішим завданням, оскільки потребує знаходження похідної або обчислення значень для кожного x та b . Ці значення формуються з гешу паролю, роблячи метод більш залежним до його змін [11].

Тестування

Для тестування методу використано такі метрики:

- нормальна кореляція (коефіцієнт Пірсона);
- коефіцієнт Спірмена.

Кореляція Пірсона (або *NCC*) – це статистичний метод, що застосовується для вимірювання ступеня зв'язку (нормальної кореляції) між двома наборами змінних. Кореляція Пірсона описується такою формулою [12]:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (2)$$

де x , y – елементи наборів, що порівнюються;

$\bar{x}$, $\bar{y}$ – середні значення наборів даних.

Значення r перебуває в діапазоні від -1 до 1 . Що ближче значення до 0 , то менша подібність наборів даних (відсутність лінійної кореляції). За умови $r = 1$ – ідеальна позитивна кореляція, якщо $r = -1$ – ідеальна негативна кореляція.

Коефіцієнт Спірмена – непараметрична міра статистичної залежності між двома змінними, що оцінює, наскільки добре взаємозв'язок між ними може бути описаний монотонною функцією. Коефіцієнт Спірмена для набору даних визначається як [11]

$$p = 1 - \frac{6 \sum_{i=1}^n (x_i - y_i)^2}{n(n^2 - 1)}, \quad (3)$$

де n – кількість пар спостережень;

x , y – ранги змінних, що порівнюються.

Значення p перебувають у діапазоні від -1 до 1 . Що ближче значення p до 1 , то більша залежність між наборами даних. Якщо $p = 1$ – ідеальна монотонна залежність.

Результатом роботи методу є матриця значень. Для оцінювання схожості двох матриць використовується така формула:

$$p_w = p_{row} \cdot p_{col}, \quad (4)$$

де p_{row} – коефіцієнт Спірмена для рядків;

p_{col} – коефіцієнт Спірмена для стовпців.

Коефіцієнт p_{col} обчислюється як

$$p_{col} = \sum_1^l \left(1 - \frac{6 \sum_{i=1}^n (x_i - y_i)^2}{n(n^2 - 1)} \right), \quad (5)$$

де l – кількість стовпців.

Тобто обчислюється сума коефіцієнтів для кожного рядка матриці x та y . Аналогічним чином обчислюється p_{row} :

$$p_{row} = \sum_1^m \left(1 - \frac{6 \sum_{i=1}^n (x_i - y_i)^2}{n(n^2 - 1)} \right), \quad (6)$$

де m – кількість рядків.

Для тестування методу автентифікації за допомогою нульового знака було обрано 50 зображень 256×256 пікселів, 50 зображень 512×512 пікселів, 50 зображень 1280×1280 пікселів [14–16].

Зображення для тестування деформувалися програмним чином за допомогою

- кристалізації (розмір кристала = 4);

- гауссового розмиття ($\sigma = 3$);

- додавання шуму (випадкове значення в діапазоні яскравості пікселя $+25/-25$);

- додавання тексту (текст різного кольору, масштабований залежно від розміру зображення).

Приклад спотворення зображення показано на рис. 4.

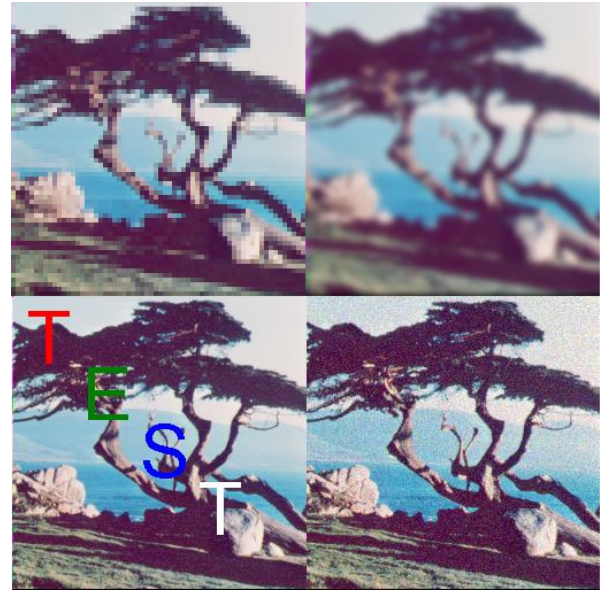


Рис. 4. Приклад спотворення зображення

Між оригінальним і спотвореним зображенням були виміряні показники NCC та $PSTR$ [17]. На рис. 5 наведений графік показників $PSTR$ для зображень 512×512 пікселів.

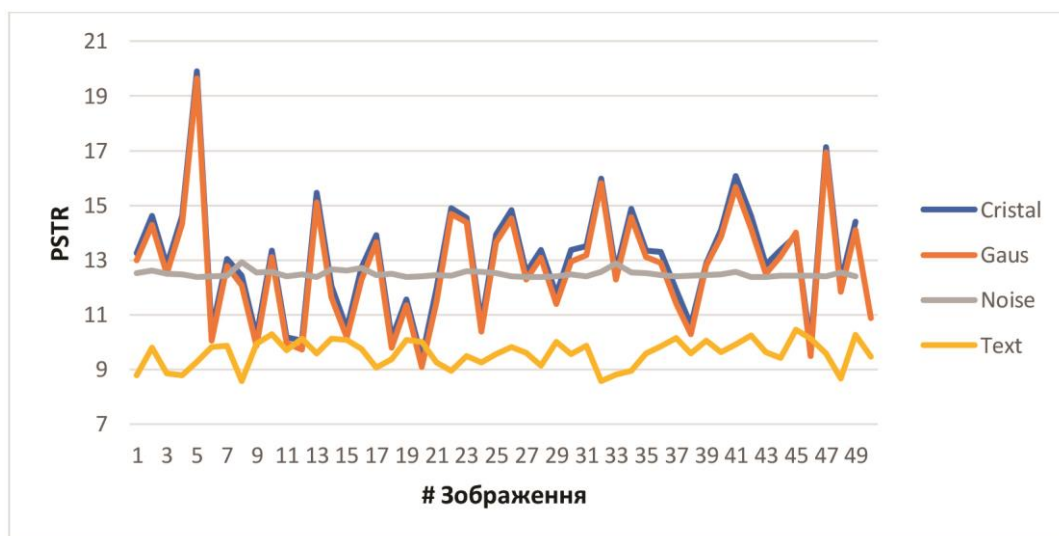


Рис. 5. Графік $PSTR$ для спотворених зображень

Як можна побачити, вимірювання проводилися на зображеннях, спотворених в одному діапазоні значень. Це було здійснено з метою переконатися в роботоспроможності методу на різних типах зображень і на різних типах спотворень, адже навіть за умови однакового рівня та методу словотворення різні зображення спотворюються по-різному. Найбільш рівномірний графік для шуму, оскільки шум додається до кожного пікселя залежно від сусідніх пікселів та яскравості самого пікселя. Найбільш нестабільним (таким, чії показники варіюються в більшому діапазоні) видом спотворення є кристалізація та гауссове розмиття, оскільки за цих видів спотворення відбуваються зміни на основі груп пікселів на зображенні.

Тестування відбувалося способом порівняння результату формування водяного знака для оригінального й спотвореного зображення. Як пароль використовувалося значення "1111".

Тестування відбувалося в такому порядку:

- формувалися водяні знаки для оригінального зображення $[W_o]$;
- формувалися водяні знаки для спотворених зображень $[W_d]$;
- перевірялися водяні знаки W_o та W_d за допомогою коефіцієнта Спірмена та Пірсона.

Загалом було сформовано 150 оригінальних водяних знаків та 600 спотворених.

Після цього випадково обрані зображення Wx_o з $[W_o]$ було порівняно з іншими $[W_o]$ з метою перевірки спроможності водяного знака розрізняти неоднакові зображення.

В описаному методі автентифікація успішна, якщо водяний знак з певною достовірністю ідентичний еталонному водяному знаку, що зберігається в системі контролю доступом.

Таблиця 1. Результати тестувань

Розмір зображення	Кристалізація	Гауссове розмиття	Шум	Додавання тексту	Порівняння випадкового Wx_o з набором $[W_o]$
256×256	100%	100%	100%	92%	0%
512×512	98%	100%	98%	92%	0%
1280×1280	100%	100%	100%	100%	0%

Як можна побачити, метод має досить високу стійкість до пошкоджень, до того ж не втрачає здібності розрізняти зображення (навіть схожі) один

Цей коефіцієнт встановлюється залежно від наявних шумів в інформаційно-комунікаційній системі та ступеня довіри, необхідному власникам.

Після проведення попередніх тестів визначено "коефіцієнт достовірності" – суму значень коефіцієнтів Пірсона та Спірмена, коли прийнято вважати, що водяний знак достовірний. Обране значення коефіцієнта становило 1.4, процедура вибору полягала в порівнянні результатів перевірок зображень і знаходженні параметра, за якого зберігається оптимальна відповідність між помилками першого та другого роду. Тобто для системи із середнім шумом $PSTR=13$ всі значення, менші за 1.4, відкидаються як недостовірні (зображення надто пошкоджені, або подано інше зображення). За такого сценарію автентифікація є невдалою, користувач не пройшов перевірку. Якщо значення вище ніж 1.4 – автентифікація успішна, користувач пройшов перевірку.

Цей коефіцієнт може встановлюватися залежно від системи. Що він вищий, то меншу кількість шумів може ігнорувати водяний знак, то менша кількість помилок другого роду та більша першого роду. Що вищий коефіцієнт, то більшу кількість шумів може ігнорувати метод, то менша кількість помилок першого роду та більша другого роду.

У системах без виникнення шумів цей коефіцієнт може дорівнювати сумі максимального значення коефіцієнтів Пірсона та Спірмена – 2.0.

У табл. 2 наведені результати тестувань, а саме скільки зображень пройшли поріг, що встановлює "коефіцієнт достовірності". Останній стовпець таблиці подає відомості про перевірку методу між зображенням Wx_o та набором $[W_o]$ з метою виявлення здатності методу розрізняти неоднакові зображення (автентифікація з хибними параметрами).

від одного. Найменший показник проходження тесту під час модифікації "додавання тексту". Це пов'язано з типом шумів, що генерує спотворення.

На рис. 6 подано графік "коефіцієнта достовірності" для зображення розміром 512×512 із спотворенням "кристалізація". Як можна побачити, одне зображення не пройшло перевірку, тобто його структура була надто пошкоджена.

На рис. 7 продемонстровано графік "коефіцієнта достовірності" для перевірки між різними зображеннями для зображення розміром 512×512 . Як можна побачити, жодне зображення не пройшло перевірку.



Рис. 6. Результати тестування для зображення розміром 512×512 з типом пошкодження "кристалізація"



Рис. 7. Результати тестування для зображення розміром 512×512 для перевірки зображень

Висновки

У роботі запропоновано метод автентифікації користувачів з використанням нульового водяного знака, оснований на перетворенні Хаара та *K-means*. У цьому методі також застосовується *Switch*-функція та функція гешування *SHA-512*.

Метод протестовано на неоднакових зображеннях з різними типами шуму, характеристиками та роздільною здатністю. Запропонований метод автентифікації на основі нульового водяного знака має певну стійкість до пошкоджень і водночас може розрізнити неоднакові зображення або схожі одне на одного (оскільки ми можемо встановити порогове значення

достовірності). За умови встановлення значення коефіцієнта достовірності 1.4 в процесі тестувань 600 зображень з рівнем шуму від 9 до 20 PSTR метод успішно розпізнав 98.3% зображень (лише 1.7% зображень пошкоджено настільки, що користувач не пройшов автентифікацію з правильними атрибутами доступу). Під час тестування подібності між зображеннями (150 перевірок) метод успішно розпізнав 0% зображень (тобто жоден користувач не автентифікувався з хибними атрибутами доступу).

Метод призначений для застосування в мережах для двофакторної автентифікації, найкраще його впроваджувати в мережах промисловості (*Industry 4.0*) для автентифікації користувачів (у системах контролю доступу), оскільки цей метод стійкий до шумів та спотворень. Застосування запропонованого

методу є більш дешевим, ніж упровадження методів автентифікації за допомогою біометричних даних чи автентифікації за допомогою токенів або смарт-карток, оскільки не потребує спеціалізованого обладнання. Порівняно з автентифікацією з використанням біометричних даних, перевагою є також те, що ключова інформація модифікована й не є чутливою, що містить персональні дані. Додатковою перевагою методу вважаємо те, що злоумисник не може напряду визначити, чи є зображення X ключовими даними, чи ні. Це робить метод більш стійким, на відміну від використанням QR-кодів.

Оскільки запропонований метод є досить новим, то потребує більшого тестування на значній кількості статистичної інформації та детального криптоаналізу.

Список літератури

1. Roček A., Javorník M., Slavíček K. et al. Zero Watermarking: Critical Analysis of Its Role in Current Medical Imaging. *Journal of Digital Imaging*. 2021. Vol. 34. P. 204–211. DOI: 10.1007/s10278-020-00396-0
2. Адміністрація Держспецзв'язку. Про визначення Базового профілю безпеки інформації: Наказ від 24.06.2024 № 317. URL: <https://cip.gov.ua/ua/docs/nakaz-administraciyi-derzhspeczv-yazku-vid-24-06-2024-317-pro-viznachennya-bazovogo-profilu-bezpeki-informaciyi>
3. Міжнародна електротехнічна комісія. EN IEC 62443. Security for industrial automation and control systems URL: <https://www.isa.org/standards-and-publications/isa-standards/isa-iec-62443-series-of-standards>
4. Chapple M., Implementing Access Control Systems, Indiana, USA, 2020, P. 110-135. URL: <https://eforum.lntu.edu.ua/index.php/jurnal/article/download/1506/1433/>
5. Айко О. В., Василенко О. Д., Степаненко В. М. Системи автоматизації контролю доступу на об'єкти. *Всеукраїнська науково-практична конференція студентів, аспірантів та молодих вчених* URL: <https://ela.kpi.ua/server/api/core/bitstreams/e4de99d9-a572-452d-8194-3ba4a39f5bb4/content>
6. Поддубний В. О., Северінов О. В., Гвоздьов Р. Ю. Використання нульових водяних знаків для підтвердження авторства зображень та багатофакторної автентифікації. *Радіотехніка*. 2024. Вип. 218. С. 35–43. DOI: <https://doi.org/10.30837/rt.2024.3.218.02>
7. Поддубний В., Северінов О., Непокритов Д. Дослідження ефективності алгоритмів оброблення зображень у схемах нульового водяного знака. Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 1(31). С. 102–114. DOI: 10.30837/2522-9818.2025.1.102
8. Poddubnyi V., Gvozдов R., Sievierinov O., Sukhoteplyi V., Bulba S., Lysytsia D. A Zero-Watermarking Algorithm for Use in RGB and Monochrome Images. 2024 *IEEE 5th KhPI Week on Advanced Technology (KhPIWeek)*. P. 1–5. DOI: <https://doi.org/10.1109/KhPIWeek61434.2024.10878081>.
9. Sethi N., Krishna R., Arora R.P. Image Compression Using Haar Wavelet Transform. *Computer Engineering and Intelligent Systems*. Vol. 2, No. 3. URL: <https://scispace.com/pdf/image-compression-using-haar-wavelet-transform-1q7fxnv973.pdf>
10. Accord.NET Framework. KMeans Class URL: http://accord-framework.net/docs/html/T_Accord_MachineLearning_KMeans.htm
11. Ramachandran P., Zoph B., Le Q.V. Swish: A Self-Gated Activation Function. URL: <https://arxiv.org/pdf/1710.05941v1.pdf>.
12. Berman J. J. Data Simplification. Chapter 4 - Understanding Your Data. 2016. P. 135–187. ISBN 9780128037812
13. McClenaghan E. Spearman Rank Correlation. *Technology Networks*. 2024. URL: <https://www.technologynetworks.com/tn/articles/spearman-rank-correlation-385744>
14. The USC-SIPI Image Database URL: <https://sipi.usc.edu/database/>

15. Nasa Image and Video Library URL: <https://images.nasa.gov/>
16. Lorem Picsum URL: <https://picsum.photos/>
17. Nadipally M. Optimization of Methods for Image-Texture Segmentation Using Ant Colony Optimization. *Intelligent Data-Centric Systems: Intelligent Data Analysis for Biomedical Applications*. Academic Press, 2019. Ch. 2. 2019. P. 21–47. DOI:10.1016/B978-0-12-815553-0.00002-1

References

1. Roček, A., Javorník, M., Slaviček, K. et al. (2021), "Zero Watermarking: Critical Analysis of Its Role in Current Medical Imaging", *Journal of Digital Imaging*, Vol. 34, P. 204–211. DOI: 10.1007/s10278-020-00396-0
2. "State Special Communications Service Administration" (2024), *Order on the definition of the Basic information security profile: Order No. 317 dated 24.06.2024* ["Administratsiia Derzhspetsvriazku. Pro vyznachennia Bazovoho profilu bezpeky informatsii: Nakaz vid 24.06.2024 № 317"]. available at: <https://cip.gov.ua/ua/docs/nakaz-administraciyi-derzhspetsvriazku-vid-24-06-2024-317-pro-vyznachennya-bazovogo-profilu-bezpeki-informaciyi>
3. "International Electrotechnical Commission (n.d.)", *EN IEC 62443. Security for industrial automation and control systems*. [Online]. available at: <https://www.isa.org/standards-and-publications/isa-standards/isa-iec-62443-series-of-standards>
4. "ChappleM., Implementing Access Control Systems", Indiana, USA, 2020, P. 110-135. available at: <https://eforum.lntu.edu.ua/index.php/jurnal/article/download/1506/1433/>
5. Aiko, O.V., Vasylenko, O.D., Stepanenko, V.M. "Systems for Access Control Automation at Facilities", *All-Ukrainian Scientific and Practical Conference of Students, Postgraduates and Young Scientists* ["Systemy avtomatyzatsii kontroliu dostupu na ob'ekty. Vseukrainska naukovo-praktychna konferentsiia studentiv, aspirantiv ta molodykh vchenykh"]. available at: <https://ela.kpi.ua/server/api/core/bitstreams/e4de99d9-a572-452d-8194-3ba4a39f5bb4/content>
6. Poddubnyi, V.O., Sievierinov, O.V., Hvozdiyov, R.Y. (2024), "Using Zero Watermarks for Image Authorship and Multifactor Authentication", *Radiotekhnika*, No. 218, P. 35–43. DOI: <https://doi.org/10.30837/rt.2024.3.218.02>
7. Poddubnyi, V., Sievierinov, O., Nepokrytyov, D. (2025), "Efficiency Study of Image Processing Algorithms in Zero Watermarking Schemes", *Modern State of Scientific Research and Technologies in Industry* ["Suchasnyi stan naukovykh doslidzhen ta tekhnolohii v promyslovosti"], No. 1(31), P. 102–114. DOI: 10.30837/2522-9818.2025.1.102
8. Poddubnyi, V., Gvozdiyov, R., Sievierinov, O., Sukhoteplyi, V., Bulba, S., Lysytsia, D. (2024), "A Zero-Watermarking Algorithm for Use in RGB and Monochrome Images", *2024 IEEE 5th KhPI Week on Advanced Technology (KhPIWeek)*, P. 1–5. DOI: <https://doi.org/10.1109/KhPIWeek61434.2024.10878081>
9. Sethi, N., Krishna, R., Arora, R.P. "Image Compression Using Haar Wavelet Transform", *Computer Engineering and Intelligent Systems*, Vol. 2, No. 3. available at: <https://scispace.com/pdf/image-compression-using-haar-wavelet-transform-1q7fxnv973.pdf>
10. "Accord.NET Framework", *KMeans Class*. available at: http://accord-framework.net/docs/html/T_Accord_MachineLearning_KMeans.htm
11. Ramachandran, P., Zoph, B., Le, Q.V. "Swish: A Self-Gated Activation Function". 2017. available at: <https://arxiv.org/pdf/1710.05941v1.pdf>
12. Berman, J.J. (2016), *Data Simplification. Chapter 4 Understanding Your Data*, P. 135–187. ISBN 9780128037812
13. McClenaghan, E. (2024), "Spearman Rank Correlation", *Technology Networks*. available at: <https://www.technologynetworks.com/tn/articles/spearman-rank-correlation-385744>
14. "University of Southern California, The USC-SIPI Image Database". available at: <https://sipi.usc.edu/database/>
15. NASA (n.d.), *NASA Image and Video Library*. [Online]. Available at: <https://images.nasa.gov/>
16. Lorem Picsum, "Free Placeholder Image Service". available at: <https://picsum.photos/>
17. Nadipally, M. (2019), "Optimization of Methods for Image-Texture Segmentation Using Ant Colony Optimization", *Intelligent Data-Centric Systems: Intelligent Data Analysis for Biomedical Applications*, Academic Press, Ch. 2, P. 21–47.

Відомості про авторів / About the Authors

Поддубний Вадим Олександрович – Харківський національний університет радіоелектроніки, аспірант кафедри безпеки інформаційних технологій, Харків, Україна; e-mail: vadym.poddubnyi@nure.ua; ORCID ID: <https://orcid.org/0000-0002-4380-491X>

Сєверінов Олександр Васильович – кандидат технічних наук, Харківський національний університет радіоелектроніки, доцент кафедри безпеки інформаційних технологій, Харків, Україна; e-mail: oleksandr.sievierinov@nure.ua; ORCID ID: <https://orcid.org/0000-0002-6327-6405>

Poddubnyi Vadym – Kharkiv National University of Radio Electronic, PhD Student at the Department of Information Technology Security, Kharkiv, Ukraine.

Sievierinov Oleksandr – PhD (Engineering Sciences), Kharkiv National University of Radio Electronic, Associate Professor at the Department of Information Technology Security, Kharkiv, Ukraine.

METHOD OF USER AUTHENTICATION IN INFORMATION AND COMMUNICATION SYSTEMS USING ZERO WATERMARK

The subject access control systems is information and communication networks. **The purpose** is to develop a user authentication method for information and communication systems as an access control system using a zero watermarking algorithm. **Objectives:** develop a user authentication method based on zero watermarking, test the proposed method, provide suggestions and recommendations for the further application of the authentication method. To accomplish the defined tasks, the following **methods** computer simulation of the proposed method to study its characteristics; empirical methods – applying the method and observing the transformation results; tatistical methods – calculating indicators such as Pearson correlation coefficients, peak signal-to-noise ratio (PSNR), and Spearman's rank correlation coefficients. **Results achieved:** novel user authentication method based on a unique zero watermarking algorithm was proposed for the first time within an information and communication system. Suggestions for further application of the method were formulated, and the method was successfully tested. **Conclusions.** During the study, an authentication method based on a zero watermarking algorithm was proposed, designed to be used as an access control system at industrial enterprises. The method was tested, and the test results confirm its operability and effectiveness. This method has the potential to reduce the cost of two-factor authentication when deploying authentication systems in large enterprises.

Keywords: ccess control systems; zero watermarking; method; authentication; image.

Бібліографічні описи / Bibliographic descriptions

Поддубний В. О., Сєверінов О. В., Метод автентифікації користувачів в інформаційно-комунікаційних системах із застосуванням нульового водяного знака. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 2 (32). С. 69–78. DOI: <https://doi.org/10.30837/2522-9818.2025.2.069>

Poddubnyi, V., Sievierinov, O. (2025), "Method of user authentication in information and communication systems using zero watermark", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (32), P. 69–78. DOI: <https://doi.org/10.30837/2522-9818.2025.2.069>

О. ПРОКОПЕНКО, С. СМЕЛЯКОВ

РОЗРОБЛЕННЯ ОБ'ЄКТНО-ОРІЄНТОВАНОГО АЛГОРИТМУ ПОРІВНЯННЯ ЗОБРАЖЕНЬ ДЛЯ ЇХ ЕФЕКТИВНОГО ПОШУКУ

Об'єктом у статті є пошук зображень на основі змісту. **Предмет дослідження** – моделі та методи пошуку зображень на основі змісту (CBIR) і управління значними обсягами медіаконтенту у великих системах збереження зображень. **Мета статті** – розроблення алгоритму порівняння об'єктно-орієнтованих дескрипторів зображень, що передбачає використання передових моделей комп'ютерного зору для виявлення об'єктів і побудови ефективних методів порівняння й пошуку цих дескрипторів. Запропонований дескриптор і алгоритм порівняння мають на меті підвищити ефективність і точність процесів пошуку зображень і управління ними. **Завдання:** аналіз сучасних підходів і рішень для створення та порівняння дескрипторів зображень та їх використання в пошуку зображень на основі змісту (CBIR); розроблення метрик і алгоритмів порівняння дескрипторів зображень, що ефективно використовують інформацію про виявлені об'єкти, такі як їх типи, розміри та місце розташування для пошуку зображень у великих сховищах даних; проведення експериментів для оцінювання запропонованого алгоритму пошуку за зображенням і порівняння ефективності з наявними рішеннями. **Методологія** передбачає всебічний огляд передових методів створення дескрипторів зображень, зокрема: геш-дескрипторів, створених вручну дескрипторів, дескрипторів на основі глибокого навчання; аналіз використання наявних дескрипторів у системах CBIR, зважаючи на їх переваги й обмеження; аналіз найкращих алгоритмів пошуку за зображенням, зокрема з підходами, що використовують глибоке навчання; розроблення алгоритму порівняння об'єктних дескрипторів для завдань пошуку за тегами, зображенням тощо. **Досягнуті результати:** розроблено дескриптор зображення, оснований на об'єктах, виявлених за допомогою сучасних моделей машинного навчання; розроблено метрики й алгоритми порівняння запропонованих дескрипторів, що дають змогу використовувати їх для пошуку зображень на основі змісту у великих сховищах даних; проведено серію експериментів для оцінювання ефективності та якості пошуку у великих системах збереження зображень за допомогою запропонованого дескриптора та алгоритмів. Експерименти дали змогу порівняти їх ефективність з наявними методами, виявивши їх переваги й обмеження, а саме: більш швидке створення дескриптора, більш швидке порівняння дескрипторів, ніж гешовані, створені вручну, та дескриптори на основі глибокого навчання, ефективне фільтрування зображень у сховищі, вища якість та швидкість пошуку зображень, але ефективність дескриптора залежить від якості моделі та даних, що використовуються для виявлення об'єктів, оскільки зображення без виявлених об'єктів не з'являються внаслідок пошуку, що може обмежувати повноту пошуку. **Висновки.** Розроблений алгоритм порівняння об'єктно-орієнтованих дескрипторів зображень є ефективним інструментом для розв'язання низки завдань пошуку зображень на основі змісту. Досягнуті результати є задовільними, оскільки розроблений алгоритм пошуку зображень перевершує більшість аналогів за швидкістю та якістю пошуку. Перспективним напрямом цього дослідження є побудова системи пошуку зображень на основі змісту з використанням розробленого дескриптора та алгоритмів, посилене застосування паралельних і розподілених обчислень, доопрацювання під конкретні потреби, що дасть змогу використовувати його не тільки в контексті зображень загального призначення, а й для більш точних наукових напрямів.

Ключові слова: оброблення зображень; детекція об'єктів; глибоке навчання; дескриптор зображень; пошук зображень; великі дані; зберігання зображень; оптимізація пошуку; інформаційні технології.

1. Аналіз сучасного стану розвитку

1.1. Мотивація

Люди отримують величезну кількість інформації з навколишнього середовища за допомогою зору. Як наслідок, з появою комп'ютерних технологій значні ресурси були спрямовані на вилучення корисної інформації із зображень, її формалізацію та використання для аналізу ситуацій і прийняття рішень у системах комп'ютерного зору, моніторингу, підтримки рішень і штучного інтелекту [1].

Спочатку через обмеження ранніх комп'ютерних ресурсів, зображення могли бути подані лише в низькій якості, займаючи всього кілька сотень байтів або кілобайтів.

Зі швидким розвитком технологій інформаційні ресурси також зазнали вибухового зростання [2]. Тепер зображення можуть займати до 32 гігабайтів за секунду в разі використання високошвидкісного відео для спеціальних фільмів. Навіть не беручи до уваги високоякісні зображення, повсякденні зображення зазвичай займають від кількох до десятків мегабайтів. Хоча це дає змогу отримати

майже реалістичне візуальне подання, це також створює значні виклики для розроблення ефективних алгоритмів, здатних обробляти такі великі обсяги інформації в режимі реального часу. Цей виклик ще більше ускладнюється швидким зростанням розмірів і кількості сучасних систем збереження інформації. Крім того, зображення часто зазнають дублювання й трансформації протягом свого життєвого циклу, чи то внаслідок покращень, ініційованих користувачами, чи через автоматичні сервіси оброблення, що стискають зображення за допомогою втратних форматів, таких як JPEG.

Багатство візуальної інформації передбачає неабиякі вимоги до пам'яті та обчислювальних ресурсів для збереження та оброблення сцен. У завданнях навігації в реальному часі така кількість інформації може стати непосильною. Тому необхідно подавати зображення за допомогою дескрипторів, що зводять інформацію до вектора ознак, але зберігають здатність розпізнавати зображення з-поміж інших у базі даних [3]. Оскільки час і увага людини є принципово обмеженими ресурсами, лише незначна частина завантажених зображень буде реально переглянута згодом [4].

Пошук зображень на основі контенту (CBIR) є дуже важливою та складною проблемою, що висвітлюється в багатьох наукових і бізнесових сферах діяльності людини, зокрема медичній промисловості [5], віртуальній реальності [6], іграх, роздрібній торгівлі, безпеці [7], соціальних медіаплатформах [8], візуальних пошукових системах [9], автомобілебудівництві тощо. Кожна із згаданих сфер потребує різного підходу до аналізу вмісту зображень. Залежно від цілей оброблення можемо детально аналізувати об'єкти на зображенні, а також оточення, текстури фону, елементи переднього плану сцени тощо. Тому CBIR є дуже широким напрямом досліджень у комп'ютерних науках. Проте, незалежно від цілей застосування та обраного методу, ефективність завжди відіграє найважливішу роль [10]. Ефективність системи вилучення визначається тим, наскільки добре вона може відновлювати ознаки за допомогою дескриптора ознак [11].

Основною ідеєю є використання зображення, предмета, теми або інтересу як вхідних даних, а відповідні зображення можуть бути отримані з бази даних за допомогою CBIR. Швидке зростання баз даних зображень для системи CBIR ускладнило отримання схожих зображень із вищою точністю [12].

Незважаючи на значні дослідження, CBIR залишається складним завданням, особливо зі зростанням розмірів наборів даних. У CBIR статистичні характеристики є надзвичайно важливими, особливо коли вони поєднуються з класифікаторами машинного навчання. Точність пошуку зображень залежить від атрибутів, що містять кольори, текстури та форми, які вилучаються із зображень [13].

Технологія пошуку контенту є обчислювальним методом, оснований на поведінці розпізнавання, що імітує взаємодію та співпрацю між індивідуумами для досягнення мети глобальної оптимізації [14].

Сучасні тенденції в пошуку зображень на основі контенту (CBIR) спрямовані на потребу в передових моделях і алгоритмах для вилучення, порівняння та пошуку інформації на основі вмісту зображень. Щоб забезпечити спеціалізованим сервісам можливість пошуку, аналізу та прийняття рішень у реальному часі, зазвичай створюється база даних дескрипторів. У цьому підході дескриптори генеруються та зберігаються в допоміжному сховищі. Коли зображення потрапляє в пошукову систему, його дескриптор створюється під час операції пошуку. Фактичний пошук і порівняння зображень у сховищі виконуються за допомогою цих дескрипторів, що значно зменшує час пошуку та зберігає ресурси системи. Використання дескрипторів є набагато ефективнішим рішенням порівняно з виконанням операцій безпосередньо на повних файлах зображень. Однак вимоги до дескрипторів постійно зростають, оскільки якість результатів пошуку й адекватність подання зображень залежать від них. Це зі свого боку впливає на складність операцій пошуку.

Постановка проблеми

Наявні дескриптори або пріоритизують незначний розмір і високу швидкість оброблення (але здебільшого спираються на кольорові параметри, ігноруючи об'єкти на зображеннях), або використовують високовимірні вектори ознак, вилучені за допомогою глибокого навчання, які часто гарантують високу точність пошуку й глибоке розуміння контексту (але вони відстають у швидкості обчислень, що робить це рішення дорогим для великих сховищ зображень і високонавантажених потоків даних).

1.2. Аналіз наявних аналогів

Будь-які параметри, статистичні дані або трансформації зображень, що дають змогу ефективно

ідентифікувати й розрізняти зображення у сховищі, можуть бути використані як дескриптори зображень. Однак одним з найпоширеніших типів дескрипторів для пошуку схожих зображень є геші, згенеровані алгоритмами перцептивного гешування. Гешування є широко вживаним методом, оснований на пошуку найближчих сусідів, і використовується у завданнях пошуку зображень у великих масштабах [15]. Ці алгоритми є функціями, розробленими для генерації порівняних гешів. Приклади містять Average Hash, DCT-Based Hash, Radial Variance-Based Hash та Marr-Hildreth Operator-Based Hash. Узагальнений процес роботи цих алгоритмів передбачає:

- застосування особливих трансформацій до зображення для вилучення шуму;
- нормалізацію зображення до фіксованого розміру;
- виконання математичних трансформацій, визначених алгоритмом;
- обчислення гешу.

Для порівняння геш-дескрипторів використовують такі метрики, як відстань Хеммінга, нормалізована відстань Хеммінга та пікова функція взаємної кореляції. Деякі алгоритми гешування стійкі до певних трансформацій зображень, зокрема розмиття, масштабування, стиснення, затемнення та незначні повороти. Важливою перевагою геш-дескрипторів є швидкість їх порівняння, оскільки геші зазвичай займають до 128 байтів. Однак суттєвим недоліком є те, що вони не беруть до уваги об'єкти на зображенні. Це може призводити до ситуацій, коли зовсім різні об'єкти реального світу з подібними кольоровими схемами неправильно ідентифікуються як схожі, а колізії є досить поширеними.

Іншим широко використовуваним типом дескрипторів зображень є ті, що основані на ключових точках. Традиційно підхід до проектування локальних дескрипторів зображень передбачав застосування створених вручну ознак, серед яких SIFT [16] та його варіанти (SURF [17], GLOH [18], PCA-SIFT [19]), є, мабуть, найпопулярнішими. Деякі створені вручну дескриптори генерують кодування в просторі Хеммінга – їх найчастіше називають бінарними дескрипторами. Прикладами таких є BRIEF [20], ORB [21], BRISK [22], FREAK [23] та LDASH [24].

Ці алгоритми зазвичай передбачають два етапи:

- виявлення ключових точок – ділянок, що значно відрізняються від свого оточення, наприклад контури або кути об'єктів;

– генерація дескрипторів для цих точок, що дає змогу порівнювати точки на різних зображеннях, щоб визначити, чи належать вони до одного об'єкта.

Сукупність дескрипторів ключових точок утворює дескриптор зображення. Порівняння цих дескрипторів дає змогу ідентифікувати однакові об'єкти на різних зображеннях, навіть за різних кутів, масштабів, умов освітлення та інших трансформацій.

Порівнюючи дескриптори на основі ключових точок із геш-дескрипторами, перші мають значні переваги. На відміну від геш-дескрипторів, вони допомагають виявляти об'єкти, реконструювати 3D-сцени, зшивати зображення та виконувати інші складні завдання комп'ютерного зору. Однак створення та використання дескрипторів ключових точок є обчислювально витратним. Це робить їх менш придатними для порівняння великих обсягів зображень, оскільки час порівняння значно залежить від кількості ключових точок, виявлених на кожному зображенні.

Зіставлення ознак зображень є невід'ємним завданням для багатьох застосунків комп'ютерного зору, таких як відстеження об'єктів, пошук зображень тощо. Зображення можна зіставляти незалежно від того, як вони змінюються внаслідок геометричних трансформацій (наприклад, повороти та зсуви), освітлення тощо. Також завдяки успішному застосуванню глибокого навчання в обробленні зображень метод глибокого навчання має перевагу у вилученні ознак із зображень [25].

Дескриптори зображень на основі глибокого навчання стали потужною альтернативою традиційним методам водночас із використанням можливості згорткових нейронних мереж (CNN) та інших нейронних архітектур для вилучення високодискримінантних ознак із зображень. На відміну від створених вручну дескрипторів, основаних на попередньо визначених алгоритмах, моделі глибокого навчання навчаються подавати ознаки безпосередньо з даних, захоплюючи як низькорівневі деталі (наприклад, краї, текстури), так і високорівневі семантичні дані (зокрема об'єкти, сцени). Ці дескриптори генеруються способом пропускання зображення крізь навчену мережу та вилучення карт ознак або вбудовувань із проміжних шарів. Зокрема для пошуку зображень низка підходів безпосередньо використовує активації мережі як ознаки зображень і успішно виконує пошук зображень [26].

Популярні підходи передбачають застосування попередньо навчених моделей, зокрема VGG, ResNet

та EfficientNet, як екстракторів ознак, або навчання спеціалізованих мереж для конкретних завдань. Дескриптори на основі глибокого навчання мають переваги в таких завданнях, як пошук зображень, розпізнавання об'єктів і семантична сегментація, оскільки вони можуть узагальнювати інформацію з різних наборів даних і адаптуватися до складних візуальних шаблонів. Однак вони часто вимагають значних обчислювальних ресурсів для навчання та виведення, а їх ефективність значно залежить від якості та різноманітності навчальної інформації. Незважаючи на окреслені виклики, здатність захоплювати багаті, ієрархічні ознаки робить їх

основним елементом сучасних систем комп'ютерного зору. Вони застосовуються в багатьох сферах, особливо в медицині [27].

Останні дослідження показують зростання інтересу до вилучення деталізованих ознак і крос-модальних підходів. Ці методи зосереджуються на навчанні деталізованої семантичної інформації для збереження високорівневої семантичної релевантності за умови усунення модальних ознак. Це приводить до більш точної та адекватної репрезентації оригінального зображення [28].

Узагальнене подання моделі CBIR зображено на рис. 1.

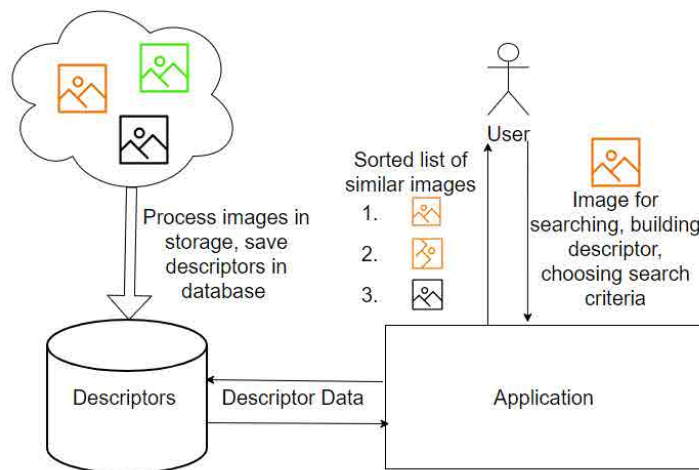


Рис. 1. Подання моделі CBIR

Пошук зображень на основі контенту (CBIR) набув значного розвитку за останні кілька десятиліть, перейшовши від традиційних методів вилучення створених у ручний спосіб ознак до сучасних підходів, основаних на глибокому навчанні. Власне, CBIR має на меті отримати зображення з бази даних, які візуально схожі на запитоване зображення, на основі їх контенту, а не текстових метаданих. Ранні системи CBIR поклалися на низькорівневі ознаки, такі як кольорові гістограми, текстури та дескриптори країв. Хоча ці методи були обчислювально ефективними, вони часто не могли охопити високорівневі семантичні дані, що призводило до обмеженої точності в реальних застосунках. Наприклад, два зображення з подібними кольоровими розподілами, але зовсім різними об'єктами могли бути неправильно зіставлені. Незважаючи на зазначені обмеження, базові техніки заклали основу для більш складних систем пошуку.

Поява глибокого навчання революціонізувала CBIR, даючи змогу системам навчатися багатого, ієрархічного подання ознак безпосередньо з даних. Сучасні системи CBIR використовують згорткові нейронні мережі (CNN) та інші архітектури глибокого навчання для вилучення високимірних векторів ознак або вбудовувань, що охоплюють як низькорівневі, так і високорівневі візуальні дані. Попередньо навчені моделі, зокрема VGG, ResNet та EfficientNet, часто застосовуються як екстрактори ознак, тоді як більш спеціалізовані підходи, а саме сіамські мережі та мережі з трійками, упроваджуються для навчання подання ознак, оптимізованих для порівняння схожості. Ці досягнення значно покращили точність пошуку, даючи змогу системам CBIR обробляти складні запити, що містять об'єкти, сцени й навіть деталізовані деталі. Наприклад, запитоване зображення певного виду птаха тепер може отримати зображення того самого виду з різних ракурсів та умов освітлення

завдяки семантичній багатогранності дескрипторів на основі глибокого навчання.

Наявні дескриптори для порівняння та пошуку зображень часто пріоритизують незначний розмір та високу швидкість оброблення, але вони здебільшого спираються на кольорові параметри, ігноруючи об'єкти на зображеннях. Хоча цей підхід добре працює в застосунках з унікальними зображеннями, він страждає від колізій, коли зображення мають подібні інтегральні властивості. Це призводить до неадекватних результатів порівняння та пошуку, оскільки дескриптори базуються винятково на інтегральних властивостях яскравості та кольору.

Іншою помітною тенденцією в сучасному CBIR є інтеграція крос-модальних технік пошуку, що уможливають пошук серед різних типів даних, наприклад, отримувати зображення за текстовими запитамі або навпаки. Це досягається за допомогою моделей, зокрема CLIP (*Contrastive Language-Image Pretraining*), які навчають спільні вбудовування для зображень і тексту в єдиному просторі ознак. Такі системи особливо ефективні в таких сферах, як електронна комерція, де користувачі можуть шукати товари за описовим текстом, або в медичній візуалізації, де діагностичні звіти використовуються для пошуку відповідних медичних зображень. Крім того, механізми уваги й трансформери все частіше впроваджуються в CBIR для зосередження на найбільш релевантних частинах зображення, що значно підвищує точність пошуку.

Незважаючи на ці досягнення, у сфері CBIR залишаються виклики. Однією з основних проблем є масштабованість систем пошуку, оскільки обчислювальна вартість порівняння високимірних векторів ознак зростає зі збільшенням розміру бази даних. Техніки, такі як гешування та векторне квантування, часто використовуються для стиснення векторів ознак і прискорення пошуку, але ці методи іноді можуть погіршувати точність. Іншим викликом є інтерпретованість систем CBIR на основі глибокого навчання, оскільки їх процеси прийняття рішень часто не прозорі. Дослідники активно вивчають способи зробити ці системи більш прозорими та зручними для користувачів. Крім того, продуктивність систем CBIR значно залежить від якості та різноманітності навчальних даних, що може призводити до упереджень та обмежувати узагальнюваність.

1.3. Мета й підходи

Метою дослідження є розроблення алгоритму порівняння об'єктно-орієнтованих дескрипторів зображень, що передбачає використання передових моделей комп'ютерного зору для виявлення об'єктів і побудови ефективних методів порівняння й пошуку цих дескрипторів. Запропонований дескриптор і алгоритм порівняння мають на меті підвищити ефективність і точність пошуку й управління зображеннями. Цей підхід забезпечує швидкість обчислень завдяки обмеженій кількості інформації, що використовується, а також надає глибокі контекстуальні дані для порівняння. Об'єкти, як складні сутності, несуть значну частину інформаційного вмісту зображення, що забезпечує більш змістовні й точні результати пошуку. Цей підхід пропонує збалансоване рішення для пошуку зображень на основі контенту, значно покращуючи точність традиційних дескрипторів, і водночас є значно швидшим, ніж високимірні вектори ознак, що застосовуються в дескрипторах на основі глибокого навчання. Топові результати, відібрані запропонованим методом, можуть бути додатково проаналізовані за допомогою більш обчислювально витратних технік, що забезпечує як ефективність, так і точність у процесі пошуку.

Основна ідея розробленого дескриптора полягає в поєднанні низьких обчислювальних витрат дескрипторів, оснований на низькорівневих ознаках зображень, із високорівневими даними про об'єкти, які надаються згортковими нейронними мережами (CNN), що зазвичай вимагають значних обчислювальних ресурсів. Цей підхід має на меті збалансувати ефективність і точність, використовуючи переваги обох методів для створення надійного та масштабованого рішення для завдань пошуку зображень.

Розроблений дескриптор має відповідати кільком критеріям / вимогам.

- Ефективність: використовувати мінімальну кількість даних про об'єкти для забезпечення обчислювальної можливості та ефективності зберігання.

- Адаптивність: бути сумісним з різними моделями детекції об'єктів, що забезпечує гнучкість у застосуванні.

- Налаштовуваність: уможлилювати налаштування для задоволення особливих потреб різних наборів даних і сценаріїв використання.

- Баланс: забезпечувати оптимальний баланс між швидкістю пошуку та якістю результатів, щоб гарантувати як швидкий пошук, так і точність.

- Простота побудови: бути простим у створенні, що дає змогу ефективно будувати базу дескрипторів і легко інтегрувати їх у пошукові системи.

Для досягнення окресленої мети необхідно виконати такі завдання:

- визначити поля даних для збереження об'єктів, щоб забезпечити легкість дескриптора й достатню інформацію для ефективного пошуку;

- розробити дизайн збереження дескриптора, що буде адаптивним до різних моделей детекції об'єктів і налаштовуваним для задоволення особливих вимог застосунків;

- розробити метрики для порівняння запропонованих дескрипторів, алгоритми для порівняння дескрипторів, які використовуються для розв'язання завдань пошуку зображень на основі змісту;

- обрати відповідну модель детекції об'єктів для експериментів, що дасть змогу справедливо порівнювати з іншими дескрипторами;

- обрати набір даних для проведення серії експериментів, які окреслять переваги й недоліки розробленого дескриптора, метрик і алгоритмів;

- створити програмне забезпечення для реалізації побудови бази дескрипторів для зберігання зображень і забезпечення ефективного порівняння дескрипторів.

З розробленим алгоритмом і реалізованим програмним забезпеченням можемо застосовувати обраний набір даних і модель детекції об'єктів для проведення точного й швидкого пошуку у великих сховищах зображень.

2. Матеріали та методи дослідження

2.1. Вибір моделі для детекції об'єктів

Детекція об'єктів є важливим складником систем комп'ютерного зору. Вона може застосовуватися в багатьох сферах, таких як відеоспостереження [29], медична візуалізація [30] та навігація роботів [31]. Для розв'язання цього завдання можна використовувати чимало алгоритмів, зокрема вилучення фону, тимчасове диференціювання, оптичний потік, фільтрація Калмана, метод опорних векторів і зіставлення контурів [32].

Ми обрали нейронну мережу, спеціалізовану на детекції об'єктів, для вилучення інформації із зображень. Обрана модель має відповідати таким критеріям:

- бути широко використовуваною та часто застосовуватися для завдань детекції об'єктів, що гарантує її ефективність у різних умовах;

- бути відносно простою у використанні, здатною виявляти значну кількість об'єктів і уможливлувати легке перетренування для адаптації до нових класів об'єктів або спеціалізації на певних типах об'єктів;

- мати кілька версій, які розрізняються кількістю параметрів, швидкістю й точністю, що дає змогу використовувати одну й ту саму архітектуру як в умовах обмежених ресурсів (наприклад, на мобільних пристроях), так і в сценаріях із значною обчислювальною потужністю.

Важливо наголосити, що різні модифікації нейронної мережі можуть бути оптимізовані для зображень певних розмірів. Це є критично важливим, оскільки швидкість оброблення залежить від цього. Нейронні мережі зазвичай добре обробляють масштабування великих зображень, але не завжди можуть адекватно масштабувати малі зображення.

Моделі детекції об'єктів можуть підвищити зрозумілість, розпізнаючи ділянки інтересу на запитуваних і знайдених зображеннях. Детекція об'єктів у глибокому навчанні охоплює дві основні парадигми: двохетапні алгоритми детекції, основані на якорних прямокутниках, та одноетапні алгоритми детекції, основані на без'якорних прямокутниках. До першої категорії належать відомі моделі, такі як R-CNN, Fast R-CNN, Faster R-CNN та R-FCN, що зазвичай досягають високої точності, але потребують більшого часу на детекцію. На противагу цьому, до другої категорії належать алгоритми SSD та YOLO, які використовують пряму мережу виведення для ефективного визначення місця розташування цілей та генерації результатів класифікації [33].

З огляду на ці вимоги до моделі детекції об'єктів, яка буде інтегрована в систему управління великими сховищами зображень, ми обрали модель YOLOv8 для цієї роботи. На момент написання YOLOv8 є одним із найбільш передових рішень у сфері комп'ютерного зору для завдань детекції об'єктів, пропонуючи баланс швидкості, точності та гнучкості.

YOLOv8 є найновішою моделлю, побудованою на успіху попередніх версій YOLO, і містить нові функції та покращення для подальшого підвищення продуктивності та гнучкості. YOLOv8 розроблена як швидкий, точний і зручний інструмент, що робить її ефективним вибором для широкого спектра завдань, зокрема детекції та відстеження об'єктів, сегментації, класифікації зображень і оцінювання постав.

Основною перевагою моделі YOLO над конкурентами в розв'язанні поставленого завдання є швидкість, що відіграє вирішальну роль. Вона у 2–2,5 рази швидша за інші популярні моделі, такі як Faster R-CNN, SSD, RetinaNet, маючи водночас порівняну (а іноді навіть вищу) точність, на відміну від своїх конкурентів. YOLOv8 має одноетапну архітектуру детекції, якщо порівнювати із сімейством нейронних мереж R-CNN, які виявляють об'єкти за два проходи.

Ще одним важливим аргументом на користь вибору YOLOv8 є можливість легко використовувати графічний процесор для прискорення детекції об'єктів способом виклику функції детекції об'єктів з одним додатковим параметром, що пришвидшує оброблення зображень у 5–10 разів.

Мережа масштабує оригінальне зображення до кількох карт ознак за допомогою пропускних зв'язків та інших архітектурних технік. Отримані карти ознак зменшуються до однієї роздільної здатності за допомогою апсемплінгу й конкатенації. Потім прогноуються класи та обмежувальні прямокутники для об'єктів, і за допомогою придушення неповних максимумів обирається найбільш імовірний обмежувальний прямокутник для кожного об'єкта. Інформація про кожен обмежувальний прямокутник подана шістьма значеннями: номер класу, координати центру X та Y , ширина, висота й рівень довіри.

2.2. Загальний алгоритм пошуку зображення

На початковому етапі, щоб забезпечити функціональність пошуку, ми виявляємо об'єкти на кожному зображенні в сховищі за допомогою нейронної мережі для детекції об'єктів. На основі цих виявлень створюємо та заповнюємо базу даних дескрипторів зображень, оснований на об'єктах. Ця база даних дескрипторів слугує основою для ефективних операцій пошуку зображень у майбутньому. Детальний опис моделей нейронних мереж, дескрипторів і методів їх побудови наведені нижче.

Далі використовуються розроблені метрики й алгоритми порівняння для оцінювання схожості між зображеннями (точніше, їх дескрипторами), а також алгоритм пошуку, який використовує ці дескриптори та метрики для оцінювання схожості зображень.

Розглянемо загальний процес пошуку.

1. Побудова дескриптора: генерація дескриптора для запитуваного зображення.

2. Налаштування пошуку: визначення критеріїв пошуку та встановлення параметрів (наприклад, пороги схожості, типи об'єктів).

3. Виконання пошуку:

- ітерація по дескрипторах у сховищі;
- порівняння кожного дескриптора з дескриптором запитуваного зображення за допомогою метрики схожості;
- сортування зображень за оцінкою їх схожості;
- відтворення результатів за схожістю.

В ідеальному сценарії відсортовані результати пошуку відтворюватимуть безпосередньо запитуване зображення (якщо воно є в сховищі) на першому місці, потім трансформації або варіації запитуваного зображення (наприклад, змінені розміри, обрізані або відредаговані версії), а потім інші зображення зі сховища, ранжовані за їх схожістю до запиту.

Для перевірки алгоритмів і програмного забезпечення розроблено допоміжні моделі, методи, критерії, метрики та інструменти програмного забезпечення, зокрема метрики для оцінювання якості пошуку. Експериментальна установка, результати й детальний аналіз подані у відповідних розділах нижче.

2.3. Дані дескрипторів

Дескриптор зображення є компактным поданням, що зберігає основну інформацію про зображення в сховищі. Це передбачає: розташування зображення та його формальні параметри (наприклад, роздільна здатність, формат), а також інформацію про об'єкти, виявлені на зображеннях, зокрема їх розташування (координати обмежувального прямокутника), розміри та площі, рівень довіри моделі комп'ютерного зору, що обмежувальний прямокутник містить об'єкт певного типу (наприклад, тип i). Завдяки цьому мінімальному, але всеосяжному набору інформації можна ефективно будувати методи пошуку, фільтрації та порівняння зображень на основі їх вмісту, зокрема об'єктів, які вони містять.

2.3.1. Дескриптор об'єкта

Кожен об'єкт подано вектором із семи чисел, як показано в табл. 1. Дескриптор об'єкта не містить номера класу, до якого належить об'єкт. Для неявного зберігання номера класу кожного об'єкта використовується структура дескриптора класу.

Таблиця 1. Дані дескриптора об'єкта

Назва	Опис
<i>x</i>	Відносна координата за віссю <i>x</i> центра обмежувальної рамки об'єкта (діапазон 0–1)
<i>y</i>	Відносна координата за віссю <i>y</i> центра обмежувальної рамки об'єкта (діапазон 0–1)
<i>w</i>	Відносна ширина об'єкта (діапазон 0–1)
<i>h</i>	Відносна висота об'єкта (діапазон 0–1)
<i>area</i>	Відносна площа об'єкта
<i>ratio</i>	Відношення меншої до більшої сторони обмежувальної рамки об'єкта
<i>conf</i>	Впевненість моделі в тому, що рамка містить об'єкт класу <i>i</i> (діапазон 0–1)

2.3.2. Дескриптор класу

Дескриптор класу описує набір об'єктів, що належать до певного класу в межах зображення. Він містить колекцію дескрипторів для кожного об'єкта цього класу, виявленого на зображенні, разом із додатковою інформацією про групу об'єктів.

Ця додаткова інформація передбачає кількість об'єктів класу, загальну площу, яку вони займають, та арифметичний центр (центроїда) групи на зображенні (див. табл. 2). Хоча ці дані можна обчислити безпосередньо з дескрипторів об'єктів, їх попереднє обчислення усуває зайві розрахунки під час порівнянь, значно скорочуючи час порівняння та підвищуючи загальну ефективність.

Таблиця 2. Дані дескриптора класу

Назва	Опис
<i>class</i>	Ідентифікатор класу
<i>number</i>	Кількість об'єктів цього класу на зображенні
<i>area</i>	Сума відносних площ, які займають об'єкти цього класу на зображенні
<i>center</i>	Арифметичний центр групи об'єктів цього класу на зображенні
<i>objects</i>	Колекція дескрипторів об'єктів цього класу на зображенні

2.3.3. Дескриптор класів

Дані дескриптора класів агрегують дескриптори класів для всіх можливих типів об'єктів. Завдання полягає в забезпеченні зручного стандартизованого способу зберігання дескрипторів класів, що дає змогу проводити майбутні обчислення на графічному процесорі (табл. 3). Кількість класів відповідає

кількості класів, використаних з метою навчання нейронної мережі для детекції об'єктів.

Таблиця 3. Дані дескриптора класів

#	Клас	Кількість	Площа	
1	<i>person</i>	2	0.012	...
2	<i>car</i>	0	0	...
3	<i>dog</i>	0	0	...
4	<i>cat</i>	0	0	...
5	<i>tie</i>	1	0.074	...

2.3.4. Дескриптор зображення

Дескриптор зображення інкапсулює основні дані, необхідні для локалізації та опису оригінального зображення. Він містить:

- розташування зображення: шлях до зображення, наприклад, локальний шлях до файлу, URI або інший ідентифікатор, щоб отримати повний файл зображення за потреби;

- розміри зображення: абсолютна ширина й висота зображення, необхідні для обчислення абсолютних координат і розмірів обмежувальних прямокутників, оскільки вся інформація про об'єкти зберігається у відносних величинах;

- дескриптори класів: колекція дескрипторів для кожного класу об'єктів, виявлених на зображенні.

Структура дескриптора зображення є гнучкою та може бути адаптована для збереження додаткової інформації про зображення та його об'єкти за потреби. Хоча дескриптор може бути налаштований для конкретних завдань, основною вимогою є збереження його компактності для швидкого оброблення під час роботи з великими сховищами зображень. Дані, пов'язані з моделлю дескриптора зображення, узагальнені в табл. 4.

Таблиця 4. Дані дескриптора зображення

Назва	Опис
<i>path</i>	Шлях до файлу із зображенням
<i>width</i>	Ширина зображення
<i>height</i>	Висота зображення
<i>classes</i>	Дескриптор класів, знайдених на зображенні

2.4 Моделі пошуку, алгоритми та метрики

Дані, отримані з нейронної мережі та використані для побудови дескрипторів, можуть застосовуватися в різних моделях пошуку.

Фільтрація за класами об'єктів

Дескриптори дають змогу фільтрувати зображення в сховищі, ідентифікуючи ті, що містять

певну кількість об'єктів із зазначених класів. Залежно від налаштувань, модель може шукати зображення, що містять або точний набір об'єктів, або принаймні зазначений набір об'єктів. Результати такого пошуку можуть бути корисними самі по собі або як попередній фільтр для звуження ділянки пошуку. Це значно зменшує обсяг для більш складних і часозатратних процедур пошуку, підвищуючи загальну ефективність.

Пошук на основі зображень

Дескриптори також полегшують пошук на основі зображень, даючи змогу користувачам знаходити конкретні зображення або ідентифікувати схожі на основі об'єктів, зображених на них, та їх розташування в межах зображення. Ця модель пошуку може знаходити точні відповідники в сховищі та ранжувати схожі зображення за їх візуальним вмістом. За замовчуванням це вважається основною моделлю пошуку.

Пошук зображень за вказаним складом об'єктів

Вимога полягає в тому, щоб повернути всі зображення в наборі даних, які містять об'єкти, зазначені користувачем. Запит містить класи й кількість об'єктів цих класів, які мають бути присутні на зображенні. Запит перетворюється на вектор, де індекси, що відповідають класам, містять кількість таких об'єктів, утворюючи вектор, аналогічний вектору *Numbers*, що зберігається в кожному дескрипторі. Відповідь на запитання – чи необхідно повертати зображення? – зрештою зводиться до порівняння двох векторів однакової довжини.

Це завдання може бути виконано в строгій та нестрогій формах, де строга форма означає, що зображення, яке шукається, містить лише об'єкти із запиту у вказаній кількості, тобто вектор запиту та вектор дескриптора ідентичні. У разі нестрогого пошуку зображення має містити принаймні типи й кількості об'єктів із запиту, тобто значення кожного елемента вектора зображення має бути більшим або рівним відповідному елементу вектора запиту. Рівняння нижче пояснюють різницю між строгою та нестрогою фільтрацією за кількістю об'єктів.

a_{ij} – кількість об'єктів

j -го класу на i -му фото

$\{Vec_1 = (a_{11}, a_{12}, \dots, a_{1n})\}$ – вектор запиту

$\{Vec_2 = (a_{21}, a_{22}, \dots, a_{2n})\}$ – вектор дескриптора

$Vec_2 = Vec_1$, if all $a_{1i} = a_{2i}$ (строгий)

$Vec_2 = Vec_1$, if all $a_{1i} \leq a_{2i}$ (нестрогий)

Аналогічно можемо фільтрувати за параметрами об'єктів, записаними в дескриптор. За потреби можна використовувати двосторонні обмеження та логічні схеми для формулювання критеріїв фільтрації.

Розв'язання цього завдання корисне не лише для надання відповідей на такі запити, але й для використання подібного підходу для виконання складніших завдань попередньої фільтрації всього набору даних. Складні методи порівняння зображень, які передбачають порівняння ділянок локалізації та площ, займаних об'єктами певних класів, зазвичай є більш обчислювально витратними. Має сенс зменшити кількість порівнянь на порядок, оскільки немає сенсу порівнювати зображення, що не містять жодної пари об'єктів одного класу для такого запиту.

2.5. Метрика ступеня схожості дескрипторів

Як метрику пропонуємо використовувати відстань між зображеннями. Очікується, що ідентичні зображення матимуть відстань, рівну нулю, і що більша відстань, то менш схожим є зображення на бажане. Нагадаємо, що метод, який розробляється, має бути стійким до типових трансформацій, що часто трапляються в сховищах зображень, таких як помірні графічні зміни (наприклад, зміни яскравості, насиченості, кольорового охоплення, стиснення, масштабування тощо). Тому необхідно додатково брати до уваги те, що зображення може бути збільшене або зменшене, і працювати здебільшого з відносними, а не абсолютними розмірами зображення та об'єктів на ньому.

Для досягнення цього використано пропорції всього зображення, оскільки такі трансформації, як збільшення або зменшення, не впливають на пропорції зображення, що беруться як відношення меншої сторони до більшої сторони зображення. Отже, перший компонент метрики – це різниця в пропорціях фотографії Δp . У формулі меншу сторону позначено як h , а більшу – як w , з яких береться модуль для інваріантності:

$$\Delta p = \left| \frac{h_1}{w_1} - \frac{h_2}{w_2} \right|, \quad (1)$$

де Δp – різниця у пропорціях двох зображень; h_1 – менша сторона першого зображення; w_1 – більша

сторона першого зображення; h_2 – менша сторона другого зображення; w_2 – більша сторона другого зображення.

Різниця в пропорціях, менша за порогове значення, вважається нульовою, щоб усунути вплив похибок масштабування, які зазвичай виникають за умови округлення довжини та ширини масштабованого зображення.

Наступна частина метрики – це порівняння параметрів зображених об'єктів, а саме співвідношення їх площ і центроїдів. Необхідно зауважити, що в цій роботі центроїд не є геометричним центром точок, які описують центри об'єктів, і не бере до уваги різні розміри об'єктів. Для обчислення центроїда об'єктів розраховано його координати як середнє арифметичне координат усіх об'єктів i -го класу. Об'єкти обробляються за типами, кожне зображення містить нуль або більше об'єктів i -го типу, і для порівняння групи об'єктів на першому зображенні з тими, що на другому, використано такий метод.

Обчислюється відстань між центроїдами групи об'єктів i -го класу d_i на одному зображенні та на іншому за допомогою евклідової відстані. Обчислюється другий компонент метрики, який позначений Δo_i і показує відстань між зображеннями за об'єктами i -го класу. Для його обчислення множимо відстань між центроїдами на різницю площ, займаних усіма об'єктами i -го типу на двох зображеннях, і додаємо штраф, якщо зображення містять різну кількість об'єктів i -го класу. Варто зазначити, що існують два особливих випадки: по-перше, коли обидва зображення не мають об'єктів i -го типу, тоді вважаємо внесок цього типу нульовим, і, по-друге, якщо одне зображення містить об'єкти, а інше – ні, тоді аналогічно множимо різницю займаних площ на максимально можливу відстань між центроїдами, додаючи штраф за різну кількість об'єктів на зображеннях.

$$d_i = \sqrt{(x_{i1} - x_{i2})^2 + (y_{i1} - y_{i2})^2}, \quad (2)$$

де d_i – відносна відстань між центроїдами групи об'єктів i -го класу на двох зображеннях; x_{i1} – координата x центроїда i -го класу на першому зображенні; x_{i2} – координата x центроїда i -го класу на другому зображенні; y_{i1} – координата y центроїда i -го класу на першому зображенні; y_{i2} – координата y центроїда i -го класу на другому зображенні.

$$\Delta o_i = d_i * |area_{i1} - area_{i2}| * (|num_{i1} - num_{i2}| + 1), \quad (3)$$

де Δo_i – відстань між зображеннями за об'єктами i -го класу; d_i – відносна відстань між центроїдами групи об'єктів i -го класу на двох зображеннях; $area_{i1}$ – сума віносних площ об'єктів i -го класу на першому зображенні; $area_{i2}$ – сума віносних площ об'єктів i -го класу на другому зображенні; num_{i1} – кількість об'єктів i -го класу на першому зображенні; num_{i2} – кількість об'єктів i -го класу на другому зображенні.

Формула (2) використовується для обчислення відстані між центрами групи об'єктів i -го класу на двох зображеннях. Формула (3) застосовується для обчислення відстані між двома зображеннями за об'єктами i -го класу. Для досягнення кращих результатів можна запропонувати кілька модифікацій.

По-перше, як згадувалося раніше, немає сенсу порівнювати пари зображень і обчислювати відстань для зображень, які не містять жодної пари об'єктів спільного класу, тому ми попередньо фільтруємо всі такі зображення з порівняння.

По-друге, порівняння загальних площ не бере до уваги кількість об'єктів i -го типу на зображеннях, тому пропонуємо поняття штрафу для зображень, які мають різну кількість об'єктів i -го типу, що збільшує різницю в площах пропорційно до різниці в кількості об'єктів на зображеннях. Поєднуючи всі компоненти, отримуємо метрику, що дає змогу знаходити зображення в наборі даних, отримані деякими трансформаціями з оригіналу, а також ранжувати інші зображення в наборі даних за їх схожістю з оригінальним зображенням для пошуку схожих зображень у наборі даних, якщо таке завдання поставлено перед користувачем.

Кожен компонент формули (4) змінюється від 0 до 1, тому вони мають однаковий відносний вплив на результат. Якщо досягнуті результати вказують на надмірний або, навпаки, незначний вплив певного компонента на кінцеве значення, то є можливість компенсувати це, додавши або помноживши на константу. Остаточна формула має такий вигляд:

$$diff(image_1, image_2) = \Delta p + \frac{\sum_{i=0}^n \Delta o_i}{n}, \quad (4)$$

де Δp – різниця в пропорціях двох зображень; Δo_i – відстань між зображеннями за об'єктами i -го класу; n – кількість унікальних класів об'єктів на двох зображеннях.

Отже, розроблена формула, що дає змогу швидко обчислювати міру схожості двох зображень, маючи їх дескриптори, отримані за допомогою моделі комп'ютерного зору.

Алгоритм пошуку зображення

1. На зображенні шукаються об'єкти за допомогою моделі YOLOv8.

2. За даними про знайдені об'єкти будується описаний дескриптор зображення.

3. Зі сховища даних відфільтровуються всі зображення, дескриптори яких не містять жодного об'єкта класів, знайдених на шуканому зображенні.

4. Для всіх дескрипторів із відфільтрованої вибірки розраховується відстань до шуканого зображення за формулою (4).

5. Результати сортуються та повертаються зображення з найменшими відстанями до шуканого.

Розроблений алгоритм використовує деякі параметри зображень і властивості, притаманні всій групі об'єктів, зокрема пропорції зображення, центроїди груп об'єктів, загальна площа, зайнята групою об'єктів на зображенні, та кількість виявлених об'єктів певного типу. Теоретично розроблені методи дають змогу повертати результати пошуку в сховищах, що містять сотні тисяч зображень, за час до 3 с.

Розроблений алгоритм використовує лише загальні параметри зображень і властивості, притаманні всій групі об'єктів, замість властивостей окремих об'єктів, оскільки порівняння двох зображень об'єкт за об'єктом не може бути дозволено заради мінімізації часу виконання пошуку у великих сховищах зображень. Розроблений метод має час порівняння двох дескрипторів, який лінійно залежить від кількості класів об'єктів, на яких навчена нейронна мережа, і не залежить від фактичної кількості виявлених об'єктів на зображеннях.

Для більш точного пошуку можна обрати знайдені зображення за певним порогом відстані або, наприклад, топ-20, топ-100 зображень, і для них провести порівняння, що бере до уваги розміри та положення окремих об'єктів замість груп. Використання таких обмежень є дуже ефективним з практичного боку. Однак це питання не розглядається детально в цій роботі. Ця тема є актуальною для подальших досліджень.

Розроблений алгоритм призначений для управління великими сховищами зображень за умови, що зображення здатні зазнавати різних

трансформацій: можуть мати різні формати файлів, скориговану яскравість, контрастність, підвищену різкість, виконане квантування, зображення може бути стиснуте або масштабоване. Водночас, це дослідження спрямоване на найпоширеніші трансформації у великих сховищах зображень; повороти, дзеркальне відтворення, додавання, модифікація та видалення окремих об'єктів, а також вбудовування іншого зображення в оригінальне на цьому етапі не розглядаються.

2.6. Метрика якості пошуку

Якість пошуку (q) визначається насамперед порядком, у якому зображення, відсортовані за зростанням відстані від оригіналу, повертаються зі сховища.

Припустимо, що сховище містить $n1$ схожих зображень (зокрема оригінальне зображення та його трансформації). Нехай m – порядковий номер останнього зображення цієї групи у відсортованому результаті пошукового запиту.

Далі знайдемо кількість $n2$ зображень у сховищі, що не належать до розглянутої групи зображень, але передують зображенню з номером m у порядку. Наявність таких зображень вказує на те, що трансформовані зображення зрештою не утворюють суцільного списку. Серед них також є нетипові, що свідчить про зниження якості пошуку. Тому оцінюємо якість пошуку таким чином:

$$q = 1 - \frac{n2}{n1 + n2}, \quad (5)$$

де q – якість пошуку; $n1$ – кількість трансформацій шуканого зображення; $n2$ – кількість зображень, що не належать до трансформацій шуканого зображення, проте мають меншу відстань під час пошуку.

Коли $n2 = 0$, перші $n1$ зображень, як наслідок, належать до бажаної групи зображень та йдуть одне за одним без проміжків. Ця ситуація вважається ідеальною ($q = 1$).

Після опису розроблених моделей дескрипторів зображень, метрик, моделей пошуку та нейронної мережі, яка використовується для побудови дескрипторів із вихідних зображень, тепер опишемо програмне забезпечення, використане для проведення експериментів.

2.7. Опис програмного забезпечення

Для реалізації запропонованих методів пошуку зображень за допомогою дескрипторів розроблено застосунки з простим вебінтерфейсом. Ці застосунки полегшують проведення експериментів і наочно демонструють результати запитів. Розроблене програмне забезпечення виконує кілька ключових функцій.

1. Побудова дескрипторів. Створює дескриптори для окремих зображень або цілих колекцій. Буде та керує сховищем зображень.

2. Інтерфейс сховища. Надає зручний інтерфейс для перегляду всіх зображень у сховищі. Дає змогу фільтрувати зображення за тегами й шукати схожі зображення в межах сховища.

3. Детекція об'єктів. Виявляє об'єкти на заданому зображенні та візуально подає виявлені об'єкти (наприклад, обмежувальні рамки).

4. Перегляд зображень. Дає змогу користувачам переглядати будь-яке зображення в сховищі. Відтворює колекцію виявлених об'єктів та інтерактивно виділяє їх на зображенні.

5. Трансформація зображень. Генерує дев'ять заданих трансформацій оригінального зображення, таких як редагування кольору, стиснення, масштабування тощо. Ці трансформації використовуються в експериментах щодо стійкості й точності пошуку зображень для подальших досліджень.

На рис. 2 показано інтерфейс розробленого застосунку, який використовується для проведення експериментів.

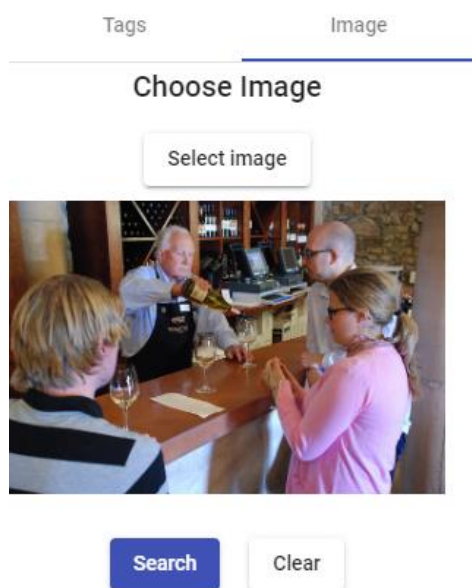


Рис. 2. Інтерфейс застосунку для експериментів

Для тестування запропонованого алгоритму пошуку за зображенням та програмного забезпечення було розроблено деякі компоненти. Опишемо їх.

Сервіс YOLO. Реалізований на Python, цей сервіс абстрагує взаємодію з моделлю YOLO. Приймає шлях до зображення або набору зображень і генерує дескриптори, повертаючи їх безпосередньо або зберігаючи в базі даних. Використовує прискорення за допомогою GPU (якщо доступне) для прискорення обчислень.

База даних. В експериментальній установці використовується MongoDB, обрана за свою гнучкість як NoSQL база даних. Під час досліджень і розроблення моделі дескрипторів часто розширюються та модифікуються, і MongoDB дає змогу вносити зміни без постійної міграції даних. MongoDB також підтримує масштабованість за допомогою таких технік, як шардування, що робить її придатною для роботи з великими сховищами даних зображень. Побудова гнучкого рішення для бази даних є критично важливою для передбачення та розв'язання проблем, що виникають у роботі зі сховищами зображень великого масштабу.

Сервіс Backend. Постійно оновлюється для підтримки актуальності моделей дескрипторів. Відповідає за виконання операцій пошуку та порівняння, обробляє запити користувачів на детекцію об'єктів у зображеннях та додавання фотоархівів до бази даних дескрипторів. Реалізований на .NET, він зберігає дескриптори в пам'яті для підвищення продуктивності та використовує ImageMagick для трансформації зображень.

Сервіс UI. Легкий фронтенд, розроблений на Angular для демонстраційних цілей. Дає змогу користувачам візуалізувати можливості системи, додавати фотографії до сховища, будувати дескриптори для зображень у певних директоріях і виконувати пошукові запити (наприклад, за тегом або схожістю до запитуваного зображення).

Крім того, був створений простий набір інструментів для генерації трансформованих версій вхідних зображень (наприклад, зміни масштабу, стиснення, гами тощо). Ці трансформації використовуються для тестування стійкості й точності методу пошуку зображень у великих сховищах.

2.8. Опис даних

Для експериментів використовуємо зображення з набору даних COCO 2017. Він містить понад

163 000 зображень із анотаціями для 80 класів об'єктів. Така значна кількість зображень дає змогу оцінити швидкість пошуку в сховищі із сотнями тисяч екземплярів.

Повна назва – *Common Objects in Context* (Поширені об'єкти в контексті) – вказує на характер зображених об'єктів, переважно повсякденних фотографій, зокрема людей, природи, міських пейзажів, спортивних подій тощо. Серед розпізнаваних класів об'єктів – люди, автомобілі, тварини, пристрої, предмети й аксесуари. Окрім оригінальних зображень з набору даних COCO 2017, додаємо трансформовані копії тестових зображень, які будуть використовуватися для пошуку.

Повний список трансформацій оригінальних зображень передбачає: зменшення яскравості на 20%, стиснення до 70%, підвищення контрастності, збільшення розміру фотографії на 20%, зменшення розміру фотографії на 20%, перетворення в градації сірого, квантування до 128 кольорів, збільшення насиченості на 30% і підвищення різкості зображення. Разом з оригінальним зображенням очікуємо, що оригінал і трансформації з'являться в топових результатах пошуку для тестового зображення.

2.9. Опис обладнання

Експерименти проводилися на робочій станції з процесором Intel(R) Core(TM) i5-9300H, що працює на частоті 2,4 ГГц, 24 ГБ оперативної пам'яті, 128 ГБ SSD та 1 ТБ HDD. Робоча станція також має графічну карту NVIDIA GeForce GTX 1650 з 4 ГБ відеопам'яті GDDR5, побудовану на архітектурі Turing. Ця графічна карта має 896 CUDA ядер, які використовуються для машинного навчання. Примітиви CUDA підтримують роботу з даними на GPU і використовуються для прискорення детекції об'єктів моделлю YOLOv8.

2.10. Побудова дескрипторів та їх репозиторію

Сховище дескрипторів було побудовано способом зберігання набору даних на робочій станції та оброблення їх створеним програмним забезпеченням. Програмне забезпечення отримувало назви директорій, де зберігалися зображення, потім обробляло всі зображення в цих директоріях, виконувало детекцію об'єктів на кожному зображенні, перетворювало дані про об'єкти та зображення в

описану модель дескриптора та зберігало отримані дескриптори в базу даних. Набір даних містить приблизно 164 000 зображень, і побудова дескрипторів для всіх зображень зайняла близько п'яти з половиною годин. Це передбачало детекцію об'єктів, побудову дескрипторів та їх збереження в базу даних.

2.11. Підготовка експериментів

Експерименти зосереджені на швидкості виконання, використанні ресурсів та якості пошуку зображень. Зокрема нас цікавить наявність трансформованих зображень у топових результатах, повернутих нашою системою.

Тест пошуку зображень на основі дескрипторів

Виконуємо пошук для 10 тестових зображень, вимірюємо час виконання та якість пошуку результатів, а також відзначаємо кількість оригінальних і трансформованих зображень із цільової групи, які безперервно з'являються у верхній частині вибірки. Це надає практичні результати для розроблених дескрипторів, моделей пошуку та метрик.

Порівняння з перцептивними гешами

Проводиться пошук зображень за допомогою відомого алгоритму для порівняння з тим, що подано в цій роботі. Перцептивні геші використовуються для цього порівняння, оскільки методи, що застосовують ключові точки, є значно більш часозатратними. Геші попередньо обчислюються для всіх зображень у сховищі, і ми обчислюємо відстань як різницю суми квадратів між гешами зображень, сортуємо результати за зростанням і аналогічно вимірюємо час виконання та якість пошуку.

Нарешті, порівнюються результати, досягнуті за допомогою вилучення ознак із зображень набору даних, з тими, що згенеровані сучасною моделлю Vision Transformer, яку запропонували наприкінці 2020 р. А. Досовіцький, Л. Бейєр, О. Колесніков та ін. [34]. Для цього обчислюємо вектори ознак для всіх зображень у наборі даних і зберігаємо їх у базі даних. Під час пошуку зображень обчислюватимемо вектори ознак для цільового зображення та порівнюватимемо їх із тими, що зберігаються в базі даних, використовуючи пошук за косинусною схожістю. Це порівняння дасть змогу оцінити запропонований метод порівняно із сучасними моделями в аналогічних умовах. Наш основний фокус буде спрямований на якість пошуку порівняно із запропонованим методом, а також на час виконання.

Обирається 10 тестових зображень, що містять різні конфігурації об'єктів (один об'єкт, група об'єктів одного класу, група об'єктів різних класів, малі та великі об'єкти тощо), для яких додаємо дев'ять трансформацій до сховища. Ці трансформації передбачають: зменшення яскравості на 20%, стиснення до 70%, підвищення контрастності, збільшення розміру фотографії на 20%, зменшення розміру фотографії на 20%, перетворення в градації сірого, квантування до 128 кольорів, збільшення насиченості на 30% та підвищення різкості зображення. Обираються зображення, що містять різні класи та кількості об'єктів, а також різні конфігурації та розташування об'єктів у межах зображень. Вимірюється час виконання для пошукових запитів із тестовими зображеннями та якість пошуку. Для тестів із гешами також фіксуємо кількість зображень, які випадають з топ-100 результатів пошуку. Мета цього експерименту – визначити ефективність розробленого методу, порівняти його результативність з наявними аналогами та виявити крайні ситуації, коли цей метод

може бути неефективним для пошуку трансформованих зображень.

3. Результати та обговорення

3.1. Пошук за зображенням

У наступному експерименті тестові зображення разом з їх трансформаціями шукаються, в ідеалі повертаючи топ-10 результатів пошуку. Пошук передбачає виявлення об'єктів на зображенні та побудову дескрипторів; потім зображення без жодного перетину в складі об'єктів фільтруються, а решта зображень порівнюються за допомогою розробленого алгоритму. Кілька скриншотів результатів пошуку наведено на рис. 3–5. Усі скриншоти досягнутих результатів пошуку за допомогою дескрипторів та перцептивних гешів, а також скриншоти розробленого програмного забезпечення можна знайти в нашому репозиторії: <https://github.com/alex-prokopenko-nure/image-search-results>



Рис. 3. Результати пошуку для тестового зображення № 3



Рис. 4. Результати пошуку для тестового зображення № 6



Рис. 5. Результати пошуку для тестового зображення № 10

У табл. 4 узагальнено результати експерименту. Як і очікувалося, для всіх 10 тестових зображень оригінальне зображення отримано як перший результат, і всі трансформації містяться в межах топ-15 результатів для всіх протестованих зображень. Для 7 із 10 тестових зображень досягнуто ідеальну якість пошуку, крім цього, оригінальне зображення та всі

трансформації з'являються в топ-10 результатах пошуку. Найгірший результат спостерігається для зображення № 6, з якістю пошуку 76,92%. Три сторонніх зображення знайдено перед деякими трансформаціями, до того ж сторонні зображення знайдено на індексах 9, 11 та 12. Середня якість пошуку серед 10 тестових зображень становить 95,12 %.

Таблиця 4. Результати пошуку за зображенням з розробленими дескрипторами

	Час пошуку	Максимальний індекс трансформованого зображення, max	Якість пошуку, q	Кількість трансформованих зображень підряд у топі вибірки
1	144 мс	10	1	10
2	122 мс	12	0.8333	9
3	123 мс	10	1	10
4	130 мс	10	1	10
5	154 мс	10	1	10
6	132 мс	13	0.7692	8
7	132 мс	10	1	10
8	131 мс	11	0.9091	9
9	155 мс	10	1	10
10	142 мс	10	1	10

Під час експерименту з пошуку зображень спостерігається, що обчислення метрики відстані для всіх зображень у сховищі за заданих умов є швидшим, ніж попередній метод фільтрації за складом об'єктів з подальшим обчисленням відстані для меншого набору зображень, оскільки копіювання даних у цьому разі займає більше часу, ніж обчислення відстаней для решти зображень. Нетипові зображення іноді заважають результатам пошуку, випереджаючи деякі трансформації, через те, що модель виявляє інший склад об'єктів порівняно з оригінальним зображенням. Найбільша різниця спостерігається, коли в трансформованому зображенні не виявлено жодного об'єкта класу, який був виявлений в оригінальному зображенні.

Іноді відстань до нетипового зображення з подібним складом об'єктів є меншою, ніж відстань до трансформованого зображення з неповним набором об'єктів у таких випадках.

3.2. Пошук за зображенням із використанням перцептивних гешів

У цьому експерименті ті самі тестові зображення з їх трансформаціями, що й у другому експерименті, шукаються, але з використанням перцептивних гешів зображень, і для порівняння обчислюється сума квадратів різниць між гешами. Досягнуті результати для порівняння із застосуванням гешів подано в табл. 5.

Таблиця 5. Результати пошуку за зображенням із використанням перцептивних гешів

	Час пошуку	Максимальний індекс трансформованого зображення, max	Якість пошуку, q	Кількість трансформованих зображень підряд у топі вибірки
1	214 мс	>1000	<0.01	8
2	221 мс	>1000	<0.01	7
3	230 мс	>1000	<0.01	7
4	242 мс	>1000	<0.01	7
5	215 мс	>1000	<0.01	6
6	226 мс	>1000	<0.01	6
7	223 мс	>1000	<0.01	7
8	252 мс	>1000	<0.01	7
9	235 мс	>1000	<0.01	8
10	228 мс	>1000	<0.01	7

Пошук на основі гешів відбувається за порівняний час. Однак, попри те, що 6–8 трансформацій містяться в топових результатах пошуку, трансформація в градації сірого для всіх тестових зображень опиняється за межами топ-1000 результатів. Крім того, для 8 із 10 тестових зображень ще 1–2 зображення опиняються за межами топ-1000 вибірки через маніпуляції з кольоровим охопленням. Це робить метод на основі гешів непрактичним для пошуку змін кольорового охоплення в зображеннях, навіть за умови перетворення зображень у градації сірого перед отриманням гешу. Це контрастує з методом на основі дескрипторів, де ті

самі трансформації успішно з'являлися в топових результатах.

3.3. Пошук за зображенням з використанням дескрипторів отриманих з Vision Transformer

У фінальному експерименті порівнюємо тестові зображення разом з їх попередньо створеними трансформаціями за допомогою вилучених ознак із моделі Vision Transformer. Ці ознаки порівнюються за допомогою косинусної схожості між векторами. У табл. 6 подано результати пошуку зображень за допомогою Vision Transformer.

Таблиця 6. Результати пошуку за зображенням із використанням дескрипторів, отриманих з Vision Transformer

	Час пошуку	Максимальний індекс трансформованого зображення, max	Якість пошуку, q	Кількість трансформованих зображень підряд у топі вибірки
1	69.7 с	10	1	10
2	62.6 с	>20	<0.5	9
3	63.2 с	18	0.5556	9
4	66.4 с	>20	<0.5	9
5	58.1 с	15	0.6667	9
6	68 с	12	0.8333	8
7	61.7 с	10	1	10
8	64.9 с	11	0.9091	9
9	62.7 с	10	1	10
10	62.8 с	>20	<0.5	9

Пошук за допомогою косинусної схожості векторів ознак, отриманих за допомогою моделі Vision Transformer, є значно довшим порівняно з іншими використаними методами. Подібно до експерименту з перцептивними гешами, трансформація у градації сірого часто випадає з топових результатів пошуку, опиняючись за межами топ-20 результатів у трьох випадках. Однак інші трансформації переважно містяться в межах першої десятки. Якщо вимкнути трансформацію в градації сірого, якість пошуку для деяких зображень є вищою, ніж у запропонованого методу, але вимагає більш тривалого часу.

Пошук способом порівняння векторів ознак зображень з подальшим уточненням або поєднанням з іншими методами може бути використаний для більш точного порівняння зображень у попередньо відфільтрованій вибірці із сотень або тисяч зображень, спочатку відфільтрованих за допомогою нашого запропонованого методу.

3.4. Обговорення

Досягнуті результати свідчать про те, що розроблений алгоритм успішно досягає поставлених

цілей, а саме пошуку схожих зображень у великих сховищах даних. Несподівано розроблені моделі та алгоритми дають змогу порівнювати зображення набагато швидше, ніж очікувалося, використовуючи базу даних дескрипторів. Простота обчислень, закладених у алгоритм, призводить до швидкого розрахунку відстаней між зображеннями, коли накладні витрати на копіювання даних займають більше часу, ніж фактичні обчислення відстаней. Якість пошуку для всіх тестових зображень перевищує 75% у разі пошуку зображень за допомогою дескрипторів, і ми обговоримо причини та можливі способи подолання цього далі.

По-перше, обговоримо час виконання пошуку, який у середньому становить близько 150 мс для понад 164 тисяч зображень. Це свідчить про те, що розроблений алгоритм може комфортно застосовуватися для великих сховищ із мільйонами зображень і забезпечувати результати пошуку в межах прийнятного часу для користувача. Паралельні та розподілені обчислення є перспективним напрямом для подальшого скорочення часу виконання.

По-друге, обговоримо якість пошуку зображень для всіх алгоритмів. Розроблений алгоритм повертає всі 10 трансформованих зображень як найкращі 15 результатів для всіх 10 тестових зображень. Сім із 10 тестових зображень повертають всі 10 трансформацій як найкращі результати без будь-якої домішки зображень, не пов'язаних із оригіналом. Навіть у разі, коли не всі трансформовані зображення потрапляють у топ-10 результатів, вони всі наявні серед 15 найближчих зображень, що дає змогу порівнювати з іншими методами з-поміж відфільтрованих топових зображень для досягнення більш точних кінцевих результатів.

Нижча якість пошуку для деяких зображень пов'язана з чутливістю алгоритму до розбіжностей у складі об'єктів зображень. Якщо зображення містить кілька об'єктів певного класу, і один з них втрачається, це зазвичай не впливає на пошук. Ця ситуація спостерігається, зокрема, у зображенні № 9, що має високу якість пошуку, незважаючи на це. Однак проблема виникає, коли втрачаються кілька об'єктів, що належать до різних класів, порівняно з оригіналом. Тоді зображення отримує високий штраф за відстань за кількома класами об'єктів, і іноді воно отримує вищий бал відстані, ніж ті зображення, що не пов'язані з оригіналом, але мають подібний склад об'єктів. Зокрема на деяких трансформаціях

зображення № 8 і № 10 втрачають класи під час детекції, а зображення № 2 містить клас об'єкта, що не був знайдений в оригінальному зображенні.

Найпоширенішою причиною того, що трансформоване зображення не потрапляє до топового списку під час експерименту, є проблема, пов'язана з роботою моделі детекції об'єктів. Зміни, виконані для трансформованих зображень, переважно незначні, але навіть вони можуть призвести до того, що малі об'єкти в оригінальному зображенні пропускаються моделлю в трансформованому зображенні, або, навпаки, об'єкти, не знайдені в оригіналі, з'являються в трансформованому.

Порівнюючи результати розробленого алгоритму пошуку зображень за допомогою дескрипторів з пошуком зображень за допомогою перцептивних гешів, а також із порівнянням векторів ознак, отриманих за допомогою моделі Vision Transformer, спостерігаємо, що метод, поданий у цій роботі, є порівняним за швидкістю (на 30–40% швидшим за перцептивні геші в проведених тестах і на порядок швидшим за порівняння векторів ознак). Крім того, він більш ефективний для пошуку зображень, що можуть містити поширені трансформації, зокрема зміни кольорового охоплення. Якість пошуку на основі запропонованої метрики в середньому перевищує 95%, тоді як якість пошуку на основі порівняння векторів ознак зазвичай нижча за 75%, а за умови використання відстані перцептивних гешів без додаткових модифікацій – менше ніж 1%. Це пов'язано насамперед з тим, що за певних трансформацій кольорового охоплення нейронна мережа може не виявити конкретні об'єкти з оригінального зображення. Однак ці об'єкти зазвичай малі й мають незначний вплив на кінцеву відстань, тому навіть у разі втрати об'єктів трансформації все ще розташовані близько до топових результатів пошуку з найгіршою продуктивністю під час тестів понад 75%.

Порівнюючи із сучасною моделлю, часто спостерігаємо, що внаслідок обчислення косинусної схожості між отриманими векторами ознак трансформації в градації сірого часто випадають з бажаних топ-10 результатів, іноді навіть за межі топ-20, вимагаючи в сотні разів більше часу на обчислення під час експериментів. Це свідчить про те, що запропонований метод може мати значні переваги в практичному використанні. Водночас, якщо зміни кольорового охоплення є суттєвими,

перцептивний геш дуже змінюється, і трансформоване зображення навіть не потрапляє до топ-1000 результатів пошуку. Це еквівалентно тому, що трансформація взагалі не виявлена.

Третій пункт – обговорення потенційних напрямів досліджень. Одним з них є поєднання з іншими методами та алгоритмами пошуку. Оскільки швидкість алгоритму висока, інші методи порівняння можна застосувати до топ-100 групи зображень, де їх можна порівнювати, наприклад, на основі спеціальних точок або гешів, згаданих на початку статті, для уточнення об'єктно-орієнтованого пошуку. Також у відфільтрованій групі можна більш точно порівняти розташування об'єктів або навіть інші дескриптори цих об'єктів, щоб отримати кінцевий результат.

Серед модифікацій алгоритму можна розглянути зміну розрахункового центру групи об'єктів із середнього арифметичного центрів на певний "центр мас", де центр буде прагнути до більших об'єктів, що робить втрату менших об'єктів менш помітною за умови трансформацій. Також серед параметрів, що присутні зараз у дескрипторі зображення, але фактично не використовуються для обчислення відстані між зображеннями, є пропорції окремих об'єктів і впевненість моделі в тому, що прямокутник належить об'єкту певного класу.

На нашу думку, обчислення деяких усереднених показників пропорцій об'єктів у групі та середньої впевненості в об'єктах цього класу не є сильним показником і може бути приблизно однаковим для абсолютно різних наборів об'єктів на основі інших параметрів. Більш того, спостережувана ситуація, коли об'єкти "випадають" під час трансформацій, може суттєво вплинути на показники та збільшити відстань між трансформованими зображеннями, що є небажаним для нас. Вважаємо згадані невикористані показники перспективними, особливо для ситуацій порівняння зображень на основі об'єктів з певного невеликого топу, коли вони є сильним показником того, чи є два зображення близькими чи далекими одне від одного.

Хотіли б наголосити, що якість розробленого дескриптора зображення та методу порівняння повністю залежить від якості даних, отриманих із нейронної мережі. У разі, коли об'єкти на зображенні не знайдені, це зображення ніколи не обирається. Тому дуже важливо ретельно обирати модель для пошуку зображень, можливо, налаштовуючи її за

допомогою трансферного навчання, знаючи особливості сховища зображень, з яким потрібно працювати в межах практичного завдання.

Логічним напрямом подальшого розвитку побудованої системи є використання гібридних моделей з одночасним використанням різних дескрипторів, що сприяє покращенню якості пошуку та усуненню недоліків конкретних моделей.

Нарешті, на основі результатів експериментів варто зазначити, що розроблений дескриптор зображення та метод порівняння успішно виконують завдання пошуку схожих зображень, маючи високу швидкість порівняння та спираючись не на формальні параметри зображень, а на зображені об'єкти. Крім упровадження розробленого методу пошуку як автономного алгоритму, його також можна використовувати для попередньої фільтрації, звуження ділянки пошуку та формування вибірки зображень, до яких можна застосувати більш складні та обчислювально витратні методи порівняння.

4. Висновки

У цій роботі досліджено проблему компромісу між швидкістю та точністю в дескрипторах зображень, які використовуються для пошуку зображень на основі контенту у великих сховищах даних.

Для розв'язання окресленої проблеми запропоновано об'єктно-орієнтований алгоритм порівняння дескрипторів зображень, що використовує дані, отримані моделями детекції об'єктів, для створення ефективного та збалансованого рішення для пошуку зображень на основі контенту, значно покращуючи точність порівнянь, отриманих за допомогою традиційних дескрипторів, і залишаючись у цьому разі більш швидким за високовимірні вектори ознак, що використовуються в дескрипторах на основі глибокого навчання. Для розроблення об'єктно-орієнтованого алгоритму було виконано кілька завдань, зокрема: визначення полів даних для зберігання об'єктів, розроблення дизайну сховища для дескриптора, вибір відповідної моделі детекції об'єктів, вибір набору даних, розроблення метрик і алгоритмів для порівняння запропонованих дескрипторів та створення програмного забезпечення з метою реалізації порівняння дескрипторів у сховищі зображень.

Для перевірки розробленого алгоритму порівняння зображень і застосунків, призначених

для управління великими сховищами зображень, було заплановано та проведено серію експериментів. Вони тестували різні аспекти розробленої системи за допомогою набору даних, призначеного для детекції об'єктів, що містить понад 160 000 зображень. Було побудовано базу даних із дескрипторами для всіх зображень у наборі даних. Результати розробленого алгоритму пошуку потім було порівняно з результатами, досягнутими за допомогою сучасної моделі Vision Transformer та відомого алгоритму пошуку зображень за допомогою перцептивних гешів.

Результати експериментів показали, що розроблений дескриптор і алгоритм порівняння чудово впоралися з поставленими завданнями. Основні критерії, що тестувалися: витрачений час, застосування пам'яті й точність, що необхідно для висновків про доцільність використання дескриптора й алгоритмів для виконання завдань пошуку зображень на основі змісту.

Розроблений алгоритм виявився достатньо точним для пошуку схожих трансформованих зображень, оскільки майже всі копії розміщені серед топових результатів пошуку для тестового зображення. Навіть коли сторонні зображення були ближчими до оригіналу, всі 10 трансформованих зображень усе ще були в межах топ-15 результатів. Експерименти продемонстрували переваги розробленого алгоритму пошуку порівняно з аналогічним пошуком за допомогою моделі Vision Transformer, особливо щодо якості пошуку й значно вищої швидкості виконання. Це наголошує

на перевагах розробленого алгоритму, на відміну від сучасної моделі, оскільки він пропонує достатню швидкість пошуку у великих сховищах із збереженням високої якості пошуку. Порівняно з пошуком на основі перцептивних гешів, розроблений алгоритм також пропонує значно кращу якість пошуку з порівняною швидкістю виконання.

Окрім позитивних результатів, досягнутих унаслідок експериментів, було зібрано достатньо велику кількість інформації для роздумів і виявлено значну кількість перспективних напрямів досліджень, які плануємо розвивати в найближчому майбутньому.

Деякі з цих напрямів передбачають застосування розробленого алгоритму для пошуку зображень із проведенням порівняння об'єктів для більш точного ранжування близьких результатів, посилене використання паралельних і розподілених обчислень для ще кращого масштабування запропонованої системи до значно більших обсягів даних, а також ідеї щодо модифікацій та покращень, якими необхідно доповнити розроблені методи й модулі системи для тонкого налаштування під конкретні потреби окремих сховищ і користувачів, що дасть змогу застосовувати її не лише в контексті зображень загального призначення, але й для більш точних наукових сфер.

Об'єктно-орієнтований дескриптор балансує ефективність і точність, використовуючи переваги наявних методів для створення надійного й масштабованого рішення для завдань пошуку зображень.

Список літератури

1. Gonzalez R. C., Woods R. E. Digital Image Processing: monograph. 4th ed., New York, 2018. 1168 p. ISBN: 9780133356724
2. Yang W., Zhao H., Wang M. Design of Intelligent Search Engine Service Performance Evaluation System. 2020 5th Asia-Pacific Conference on Intelligent Robot Systems (ACIRS). Singapore, 2020. P. 86–91. DOI: <https://doi.org/10.1109/ACIRS49895.2020.9162611>
3. Amorós F., Payá L., Mayol-Cuevas W. Holistic Descriptors of Omnidirectional Color Images and Their Performance in Estimation of Position and Orientation. *IEEE Access*. 2020. Vol. 8. P. 81822–81848. DOI: <https://doi.org/10.1109/ACCESS.2020.2990996>
4. Liu X., Cheung G., Lin C.-W. Prior-Based Quantization Bin Matching for Cloud Storage of JPEG Images. *IEEE Transactions on Image Processing*. 2018. Vol. 27. № 7. H. 3222–3235. DOI: <https://doi.org/10.1109/TIP.2018.2799704>
5. Carvalho E. D., Filho A. O. C., Silva R. R. V. Breast Cancer Diagnosis from Histopathological Images Using Textural Features and CBIR. *Artificial Intelligence in Medicine*. 2020. Vol. 105. P. 101845. DOI: <https://doi.org/10.1016/j.artmed.2020.101845>
6. Nakazato M., Huang T. S. 3D MARS: Immersive Virtual Reality for Content-Based Image Retrieval. *IEEE International Conference on Multimedia and Expo (ICME)*. Tokyo, Japan, 2001. P. 44–47. DOI: <https://doi.org/10.1109/ICME.2001.1237651>
7. Iqbal K., Odetayo M. O., James A. Content-Based Image Retrieval Approach for Biometric Security Using Colour, Texture and Shape Features Controlled by Fuzzy Heuristics. *Journal of Computer and System Sciences*. 2012. Vol. 78. № 4. P. 1258–1277. DOI: <https://doi.org/10.1016/j.jcss.2011.10.013>

8. Popescu A., Grefenstette G. Social Media Driven Image Retrieval. *Proceedings of the 1st ACM International Conference on Multimedia Retrieval (ICMR '11)*. New York, NY, USA, 2011. P. 1–8. DOI: <https://doi.org/10.1145/1991996.1992029>
9. Liu Y., Mei T., Hua X.-S. CrowdReranking: Exploring Multiple Search Engines for Visual Search Reranking. *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '09)*. New York, NY, USA, 2009. P. 500–507. DOI: <https://doi.org/10.1145/1571941.1572027>
10. Staszewski P., Jaworski M., Cao J. A New Approach to Descriptors Generation for Image Retrieval by Analyzing Activations of Deep Neural Network Layers. *IEEE Transactions on Neural Networks and Learning Systems*. 2022. Vol. 33. № 12. P. 7913–7920. DOI: <https://doi.org/10.1109/TNNLS.2021.3084633>
11. Sahmoudi Y., El-Ogri O., El-Mekkaoui J. An Efficient Biomedical Color Image Retrieval System Based on Continuous Orthogonal Legendre Fourier Quaternion. 2024 *Sixth International Conference on Intelligent Computing in Data Sciences (ICDS)*. Marrakech, Morocco, 2024. P. 1–6. DOI: <https://doi.org/10.1109/ICDS62089.2024.10756406>
12. Bano M., Matta P., Chandel S. Content Based Image Retrieval: A Study of Approaches and Techniques. 2024 *4th International Conference on Technological Advancements in Computational Sciences (ICTACS)*. Tashkent, Uzbekistan, 2024. P. 16–22. DOI: <https://doi.org/10.1109/ICTACS62700.2024.10840489>
13. Anand A., Saxena A., Singh K. Statistical Features Based Content Based Image Retrieval Using Machine Learning Classifiers. 2024 *IEEE 3rd World Conference on Applied Intelligence and Computing (AIC)*. Gwalior, India, 2024. P. 1102–1109. DOI: <https://doi.org/10.1109/AIC61668.2024.10731120>
14. Debin H., Yue Z., Shuai J. Application of Content-Based Retrieval Technology in Image Archive Management. 2024 *Global Conference on Communications and Information Technologies (GCCIT)*. Bangalore, India, 2024. P. 1–6. DOI: <https://doi.org/10.1109/GCCIT63234.2024.10862277>
15. Bai J., Ni B., Wang M. Deep Progressive Hashing for Image Retrieval. *IEEE Transactions on Multimedia*. 2019. Vol. 21. № 12. P. 3178–3193. DOI: <https://doi.org/10.1109/TMM.2019.2920601>
16. Lowe D. G. Object Recognition from Local Scale-Invariant Features. *Proceedings of the Seventh IEEE International Conference on Computer Vision*. Kerkyra, Greece, 1999. Vol. 2. P. 1150–1157. DOI: <https://doi.org/10.1109/ICCV.1999.790410>
17. Bay H., Ess A., Tuytelaars T. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*. 2008. Vol. 110. № 3. P. 346–359. DOI: <https://doi.org/10.1016/j.cviu.2007.09.014>
18. Mikolajczyk K., Schmid C. A Performance Evaluation of Local Descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2005. Vol. 27. № 10. P. 1615–1630. DOI: <https://doi.org/10.1109/TPAMI.2005.188>
19. Ke Y., Sukthankar R. PCA-SIFT: A More Distinctive Representation for Local Image Descriptors. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*. Washington, DC, USA, 2004. Vol. 2. P. II–II. DOI: <https://doi.org/10.1109/CVPR.2004.1315206>
20. Calonder M., Lepetit V., Strecha C. BRIEF: Binary Robust Independent Elementary Features. *Computer Vision – ECCV 2010. Lecture Notes in Computer Science*. 2010. Vol. 6314. P. 778–792. DOI: https://doi.org/10.1007/978-3-642-15561-1_56
21. Rublee E., Rabaud V., Konolige K. ORB: An Efficient Alternative to SIFT or SURF. 2011 *International Conference on Computer Vision*. Barcelona, Spain, 2011. P. 2564–2571. DOI: <https://doi.org/10.1109/ICCV.2011.6126544>
22. Leutenegger S., Chli M., Siegwart R. Y. BRISK: Binary Robust Invariant Scalable Keypoints. 2011 *International Conference on Computer Vision*. Barcelona, Spain, 2011. P. 2548–2555. DOI: <https://doi.org/10.1109/ICCV.2011.6126542>
23. Alahi A., Ortiz R., Vanderghenst P. FREAK: Fast Retina Keypoint. 2012 *IEEE Conference on Computer Vision and Pattern Recognition*. Providence, RI, USA, 2012. P. 510–517. DOI: <https://doi.org/10.1109/CVPR.2012.6247715>
24. Žižakić N., Pižurica A. Efficient Local Image Descriptors Learned with Autoencoders. *IEEE Access*. 2022. Vol. 10. P. 221–235. DOI: <https://doi.org/10.1109/ACCESS.2021.3138168>
25. Liu Y., Xu X., Li F. Image Feature Matching Based on Deep Learning. 2018 *IEEE 4th International Conference on Computer and Communications (ICCC)*. Chengdu, China, 2018. P. 1752–1756. DOI: <https://doi.org/10.1109/CompComm.2018.8780936>
26. Radenović F., Tolias G., Chum O. Fine-Tuning CNN Image Retrieval with No Human Annotation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2019. Vol. 41. № 7. P. 1655–1668. DOI: <https://doi.org/10.1109/TPAMI.2018.2846566>
27. Song L., Lin J., Wang Z. J. An End-to-End Multi-Task Deep Learning Framework for Skin Lesion Analysis. *IEEE Journal of Biomedical and Health Informatics*. 2020. Vol. 24. № 10. P. 2912–2921. DOI: <https://doi.org/10.1109/JBHI.2020.2973614>
28. Wang B., Zhang H., Zhu L. Multi-Level Adversarial Attention Cross-Modal Hashing. *Signal Processing: Image Communication*. 2023. Vol. 117. 117017 p. DOI: <https://doi.org/10.1016/j.image.2023.117017>
29. Gajjar V., Khandhediya Y., Gurnani A. Human Detection and Tracking for Video Surveillance: A Cognitive Science Approach. 2017 *IEEE International Conference on Computer Vision Workshops (ICCVW)*. Venice, Italy, 2017. P. 2805–2809. DOI: <https://doi.org/10.1109/ICCVW.2017.330>
30. Adel M., Moussaoui A., Rasigni M. Statistical-Based Tracking Technique for Linear Structures Detection: Application to Vessel Segmentation in Medical Images. *IEEE Signal Processing Letters*. 2010. Vol. 17. № 6. P. 555–558. DOI: <https://doi.org/10.1109/LSP.2010.2046697>

31. Truong X.-T., Yoong V. N., Ngo T.-D. RGB-D and Laser Data Fusion-Based Human Detection and Tracking for Socially Aware Robot Navigation Framework. 2015 *IEEE International Conference on Robotics and Biomimetics (ROBIO)*. Zhuhai, China, 2015. P. 608–613. DOI: <https://doi.org/10.1109/ROBIO.2015.7418835>
32. Galvez R. L., Bandala A. A., Dadios E. P. Object Detection Using Convolutional Neural Networks. *TENCON 2018 – 2018 IEEE Region 10 Conference*. Jeju, Korea (South), 2018. P. 2023–2027. DOI: <https://doi.org/10.1109/TENCON.2018.8650517>
33. Wehbe A., Hotiet H., Minetti I. Integrating YOLO for Advanced Content-Based Image Retrieval in Lung Cancer Imaging. 2024 31st *IEEE International Conference on Electronics, Circuits and Systems (ICECS)*. Nancy, France, 2024. P. 1–4. DOI: <https://doi.org/10.1109/ICECS61496.2024.10848862>
34. Dosovitskiy A., Beyer L., Kolesnikov A. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale / A. Dosovitskiy та ін. *International Conference on Learning Representations*. 2021. DOI: <https://doi.org/10.48550/arXiv.2010.11929>

References

1. Gonzalez, R.C. and Woods, R.E. (2018), *Digital Image Processing*, 4th ed., Pearson/Prentice Hall, 1168 p. DOI/ISBN: 9780133356724
2. Yang, W., Zhao, H., Wang, M. and Ji, J. (2020), "Design of Intelligent Search Engine Service Performance Evaluation System", 2020 5th *Asia-Pacific Conference on Intelligent Robot Systems (ACIRS)*, Singapore, P. 86–91. DOI: <https://doi.org/10.1109/ACIRS49895.2020.9162611>
3. Amorós, F., Payá, L., Mayol-Cuevas, W., Jiménez, L.M. and Reinoso, O. (2020), "Holistic Descriptors of Omnidirectional Color Images and Their Performance in Estimation of Position and Orientation", *IEEE Access*, vol. 8, P. 81822–81848. DOI: <https://doi.org/10.1109/ACCESS.2020.2990996>
4. Liu, X., Cheung, G., Lin, C.-W., Zhao, D. and Gao, W. (2018), "Prior-Based Quantization Bin Matching for Cloud Storage of JPEG Images", *IEEE Transactions on Image Processing*, Vol. 27, No. 7, P. 3222–3235. DOI: <https://doi.org/10.1109/TIP.2018.2799704>
5. Carvalho, E.D., Filho, A.O.C., Silva, R.R.V., Araújo, F.H.D., Diniz, J.O.B., Silva, A.C., Paiva, A.C. and Gattass, M. (2020), "Breast Cancer Diagnosis from Histopathological Images Using Textural Features and CBIR", *Artificial Intelligence in Medicine*, Vol. 105, 101845 p. DOI: <https://doi.org/10.1016/j.artmed.2020.101845>
6. Nakazato, M. and Huang, T.S. (2001), "3D MARS: Immersive Virtual Reality for Content-Based Image Retrieval", *IEEE International Conference on Multimedia and Expo (ICME)*, Tokyo, Japan, P. 44–47. DOI: <https://doi.org/10.1109/ICME.2001.1237651>
7. Iqbal, K., Odetayo, M.O. and James, A. (2012), "Content-Based Image Retrieval Approach for Biometric Security Using Colour, Texture and Shape Features Controlled by Fuzzy Heuristics", *Journal of Computer and System Sciences*, Vol. 78, No. 4, P. 1258–1277. DOI: <https://doi.org/10.1016/j.jcss.2011.10.013>
8. Popescu, A. and Grefenstette, G. (2011), "Social Media Driven Image Retrieval", *Proceedings of the 1st ACM International Conference on Multimedia Retrieval (ICMR '11)*, New York, NY, USA, Article 33, P. 1–8. DOI: <https://doi.org/10.1145/1991996.1992029>
9. Liu, Y., Mei, T. and Hua, X.-S. (2009), "CrowdReranking: Exploring Multiple Search Engines for Visual Search Reranking", *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '09)*, New York, NY, USA, P. 500–507. DOI: <https://doi.org/10.1145/1571941.1572027>
10. Staszewski, P., Jaworski, M., Cao, J. and Rutkowski, L. (2022), "A New Approach to Descriptors Generation for Image Retrieval by Analyzing Activations of Deep Neural Network Layers", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, No. 12, P. 7913–7920. DOI: <https://doi.org/10.1109/TNNLS.2021.3084633>
11. Sahmoudi, Y., El-Ogri, O., El-Mekkaoui, J. and Hjouji, A. (2024), "An Efficient Biomedical Color Image Retrieval System Based on Continuous Orthogonal Legendre Fourier Quaternion", 2024 *Sixth International Conference on Intelligent Computing in Data Sciences (ICDS)*, Marrakech, Morocco, P. 1–6. DOI: <https://doi.org/10.1109/ICDS62089.2024.10756406>
12. Bano, M., Matta, P. and Chandel, S. (2024), "Content Based Image Retrieval: A Study of Approaches and Techniques", 2024 *4th International Conference on Technological Advancements in Computational Sciences (ICTACS)*, Tashkent, Uzbekistan, P. 16–22. DOI: <https://doi.org/10.1109/ICTACS62700.2024.10840489>
13. Anand, A., Saxena, A. and Singh, K. (2024), "Statistical Features Based Content Based Image Retrieval Using Machine Learning Classifiers", 2024 *IEEE 3rd World Conference on Applied Intelligence and Computing (AIC)*, Gwalior, India, P. 1102–1109. DOI: <https://doi.org/10.1109/AIC61668.2024.10731120>
14. Debin, H., Yue, Z. and Shuai, J. (2024), "Application of Content-Based Retrieval Technology in Image Archive Management", 2024 *Global Conference on Communications and Information Technologies (GCCIT)*, Bangalore, India, P. 1–6. DOI: <https://doi.org/10.1109/GCCIT63234.2024.10862277>

15. Bai, J., Ni, B., Wang, M., Li, Z., Cheng, S. and Yang, X. (2019), "Deep Progressive Hashing for Image Retrieval", *IEEE Transactions on Multimedia*, Vol. 21, No. 12, P. 3178–3193. DOI: <https://doi.org/10.1109/TMM.2019.2920601>
16. Lowe, D.G. (1999), "Object Recognition from Local Scale-Invariant Features", *Proceedings of the Seventh IEEE International Conference on Computer Vision*, Kerkyra, Greece, vol. 2, P. 1150–1157. DOI: <https://doi.org/10.1109/ICCV.1999.790410>
17. Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L. (2008), "Speeded-Up Robust Features (SURF)", *Computer Vision and Image Understanding*, vol. 110, No. 3, P. 346–359. DOI: <https://doi.org/10.1016/j.cviu.2007.09.014>
18. Mikolajczyk, K. and Schmid, C. (2005), "A Performance Evaluation of Local Descriptors", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 10, P. 1615–1630. DOI: <https://doi.org/10.1109/TPAMI.2005.188>
19. Ke, Y. and Sukthankar, R. (2004), "PCA-SIFT: A More Distinctive Representation for Local Image Descriptors", *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Washington, DC, USA, Vol. 2, P. II–II. DOI: <https://doi.org/10.1109/CVPR.2004.1315206>
20. Calonder, M., Lepetit, V., Strecha, C. and Fua, P. (2010), "BRIEF: Binary Robust Independent Elementary Features", *Computer Vision – ECCV 2010. Lecture Notes in Computer Science*, Vol. 6314, P. 778–792. DOI: https://doi.org/10.1007/978-3-642-15561-1_56
21. Rublee, E., Rabaud, V., Konolige, K. and Bradski, G. (2011), "ORB: An Efficient Alternative to SIFT or SURF", *2011 International Conference on Computer Vision*, Barcelona, Spain, P. 2564–2571. DOI: <https://doi.org/10.1109/ICCV.2011.6126544>
22. Leutenegger, S., Chli, M. and Siegwart, R.Y. (2011), "BRISK: Binary Robust Invariant Scalable Keypoints", *2011 International Conference on Computer Vision*, Barcelona, Spain, P. 2548–2555. DOI: <https://doi.org/10.1109/ICCV.2011.6126542>
23. Alahi, A., Ortiz, R. and Vandergheynst, P. (2012), "FREAK: Fast Retina Keypoint", *2012 IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, USA, P. 510–517. DOI: <https://doi.org/10.1109/CVPR.2012.6247715>
24. Žižakić, N. and Pižurica, A. (2022), "Efficient Local Image Descriptors Learned with Autoencoders", *IEEE Access*, Vol. 10, P. 221–235. DOI: <https://doi.org/10.1109/ACCESS.2021.3138168>
25. Liu, Y., Xu, X. and Li, F. (2018), "Image Feature Matching Based on Deep Learning", *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*, Chengdu, China, P. 1752–1756. DOI: <https://doi.org/10.1109/CompComm.2018.8780936>
26. Radenović, F., Tolias, G. and Chum, O. (2019), "Fine-Tuning CNN Image Retrieval with No Human Annotation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, No. 7, P. 1655–1668. DOI: <https://doi.org/10.1109/TPAMI.2018.2846566>
27. Song, L., Lin, J., Wang, Z.J. and Wang, H. (2020), "An End-to-End Multi-Task Deep Learning Framework for Skin Lesion Analysis", *IEEE Journal of Biomedical and Health Informatics*, vol. 24, No. 10, P. 2912–2921. DOI: <https://doi.org/10.1109/JBHI.2020.2973614>
28. Wang, B., Zhang, H., Zhu, L., Nie, L. and Liu, L. (2023), "Multi-Level Adversarial Attention Cross-Modal Hashing", *Signal Processing: Image Communication*, Vol. 117, 117017 p. DOI: <https://doi.org/10.1016/j.image.2023.117017>
29. Gajjar, V., Khandhediya, Y. and Gurnani, A. (2017), "Human Detection and Tracking for Video Surveillance: A Cognitive Science Approach", *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, Venice, Italy, P. 2805–2809. DOI: <https://doi.org/10.1109/ICCVW.2017.330>
30. Adel, M., Moussaoui, A., Rasigni, M., Bourennane, S. and Hamami, L. (2010), "Statistical-Based Tracking Technique for Linear Structures Detection: Application to Vessel Segmentation in Medical Images", *IEEE Signal Processing Letters*, Vol. 17, No. 6, P. 555–558. DOI: <https://doi.org/10.1109/LSP.2010.2046697>
31. Truong, X.-T., Yoong, V.N. and Ngo, T.-D. (2015), "RGB-D and Laser Data Fusion-Based Human Detection and Tracking for Socially Aware Robot Navigation Framework", *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Zhuhai, China, P. 608–613. DOI: <https://doi.org/10.1109/ROBIO.2015.7418835>
32. Galvez, R.L., Bandala, A.A., Dadios, E.P., Vicerra, R.R.P. and Maningo, J.M.Z. (2018), "Object Detection Using Convolutional Neural Networks", *TENCON 2018 – 2018 IEEE Region 10 Conference*, Jeju, Korea (South), P. 2023–2027. DOI: <https://doi.org/10.1109/TENCON.2018.8650517>
33. Wehbe, A., Hotiet, H., Minetti, I. and Dellapiane, S. (2024), "Integrating YOLO for Advanced Content-Based Image Retrieval in Lung Cancer Imaging", *2024 31st IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, Nancy, France, P. 1–4. DOI: <https://doi.org/10.1109/ICECS61496.2024.10848862>
34. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J. and Houslsby, N. (2021), "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale", *International Conference on Learning Representations*. DOI: <https://doi.org/10.48550/arXiv.2010.11929>

Надійшла (Received) 15.04.2025

Відомості про авторів / About the Authors

Прокопенко Олександр Сергійович – Харківський національний університет радіоелектроніки, аспірант кафедри Програмної інженерії, Харків, Україна. e-mail: oleksandr.prokopenko1@nure.ua, ORCID ID: <https://orcid.org/0000-0003-0489-6820>

Смеляков Сергій Вячеславович – доктор фізико-математичних наук, професор, Харківський національний університет радіоелектроніки, професор кафедри Програмної інженерії, Харків, Україна. e-mail: serhii.smeliakov@nure.ua, ORCID ID: <https://orcid.org/0000-0002-5791-2479>, Scopus Author ID: 24527617600

Prokopenko Oleksandr – Kharkiv National University of Radio Electronics, PhD Student, Software Engineering Department, Kharkiv, Ukraine.

Smelyakov Serhii – Kharkiv National University of Radio Electronics, Doctor of Science, Professor at the Software Engineering Department, Kharkiv, Ukraine.

DEVELOPMENT OF AN OBJECT-ORIENTED IMAGE COMPARISON ALGORITHM FOR EFFICIENT SEARCH

The **object** of research is content-based image retrieval (CBIR). The **subject** of this study is models and methods for content-based image retrieval (CBIR) and managing large volumes of media content in extensive image storage systems. The **goal** of the research is to develop an algorithm for comparing object-oriented image descriptors, which involves using advanced computer vision models for object detection and constructing efficient methods for comparing and searching these descriptors. The proposed descriptor and comparison algorithm aim to enhance the efficiency and accuracy of image search and management processes. The **tasks** include: analyzing modern approaches and solutions for creating and comparing image descriptors and their use in CBIR; developing metrics and algorithms for comparing image descriptors that effectively utilize information about detected objects – such as their types, sizes, and locations – for image search in large data repositories; conducting experiments to evaluate the proposed image search algorithm and comparing its efficiency with existing solutions. The **methodology** includes: conducting a comprehensive review of advanced image descriptor generation methods, including hash-based descriptors, handcrafted descriptors, and deep learning-based descriptors; analyzing the use of existing descriptors in CBIR systems, focusing on their advantages and limitations; evaluating the best image search algorithms, including deep learning-based approaches; developing an object descriptor comparison algorithm for tag-based search, image-based search, and other tasks. The **results** obtained are as follows: an object-based image descriptor was developed using state-of-the-art machine learning models for object detection; metrics and comparison algorithms for the proposed descriptors were developed, enabling their use for CBIR in large data repositories; a series of experiments were conducted to assess the efficiency and search quality of the proposed descriptor and algorithms in large-scale image storage systems. These experiments compared their performance with existing methods, revealing their advantages and limitations, namely: faster descriptor generation; faster descriptor comparison than hashed, handcrafted, and deep learning-based descriptors; efficient image filtering in storage; higher search quality and speed for image-based queries. However, the descriptor's effectiveness depends on the quality of the model and data used for object detection, as images without detected objects do not appear in search results, which may limit search completeness. **Conclusions:** The developed algorithm for comparing object-oriented image descriptors is an effective tool for solving various CBIR tasks. The obtained results are satisfactory, as the proposed image search algorithm outperforms most alternatives in terms of speed and search quality. A promising direction for future research is the development of a CBIR system using the proposed descriptor and algorithms, enhanced by parallel and distributed computing, and further refinement for specific applications. This would allow its use not only for general-purpose images but also for more precise scientific domains.

Keywords: image processing; object detection; deep learning; image descriptor; image retrieval; big data; image storage; search optimization; information technology.

Бібліографічні описи / Bibliographic descriptions

Прокопенко О. С., Смеляков С. В. Розроблення об'єктно-орієнтованого алгоритму порівняння зображень для їх ефективного пошуку. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 2 (32). С. 79–101. DOI: <https://doi.org/10.30837/2522-9818.2025.2.079>

Prokopenko, O., Smelyakov, S. (2025), "Development of an object-oriented image comparison algorithm for efficient search", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (32), P. 79–101. DOI: <https://doi.org/10.30837/2522-9818.2025.2.079>

K. SYROTKIN

ACCURACY EVALUATION AND ERROR ANALYSIS OF DEPENDENCY PARSING FOR TEXTS IN UKRAINIAN

The **subject** of our research is the dependency parsing of sentences in the Ukrainian language using the Universal Dependencies framework. The **goal** of the work is to evaluate the accuracy of existing transition-based and graph-based parsing architectures with and without deep word embeddings on the Ukrainian dataset, and to analyze the error profiles of such parsers. The article addresses two **tasks**. One is to evaluate the accuracy of several modern dependency parsing approaches applied to a hand-annotated gold standard dataset, using labeled and unlabeled attachment scores as the metric to evaluate the parsing accuracy. The other task is to analyze and categorize the errors made by standard parsers. Resolving these errors could potentially allow us to build a more accurate parser in the future. Error rate for different categories is compared to the baseline error rate, and statistical significance of such comparison is validated using the chi-square **method**. The key **results** are as follows. For the Ukrainian language, parsing accuracy is greatly increased with the use of deep word embeddings. Transition-based parser with deep word embeddings provides the highest labeled attachment score of 84.66% for the test dataset. For the same parser, higher error rates are associated with non-projectivity of dependencies, higher sentence length and higher distance to head. Also, for pronouns and numerals the error rate for labeled attachment is significantly higher than the baseline, while the unlabeled error rate is at the baseline. **Conclusions:** parsing accuracy for the Ukrainian dataset is sub-par in comparison with other languages, but the overall trend of accuracy improvement with the use of deep word embeddings is consistent with existing research. To improve overall parsing accuracy, we must focus on such problem areas as non-projective dependencies, longer sentences, and greater distance between the head and the dependent. In future work we intend to explore ways to improve parsing accuracy by supplementing neural parsing with other approaches, like formal rules or pre- and post-processing.

Keywords: natural language processing; syntactic parsing; dependency parsing; Universal Dependencies.

1. Introduction

1.1 Relevance of dependency parsing

Dependency parsing is a method of analysis in natural language processing. It involves analyzing the grammatical structure of a sentence by identifying the relationships between words and representing these relationships as a directed graph.

Syntactic analysis is usually considered an important building block used to improve the performance of subsequent tasks.

For example, since named entities are often expressed as noun phrases, the analyzed input can greatly benefit named entity recognition (NER). Part-of-speech tagging (POS) is closely related to dependency parsing. Correct identification of the subject and object of a sentence is extremely useful for dialogue systems, sentiment analysis, and machine translation. Last but not least, syntactic analysis is often used in advanced writing systems and educational systems to check written text for compliance with a set of formal rules.

Given the importance of parsing dependencies, our focus in this work is on the accuracy of parsing.

1.2 Problems in analyzing dependencies in Ukrainian texts

1.2.1 Morphological richness and free word order

One of the main problems inherent in the Ukrainian language is its morphological richness. Morphologically rich languages are those in which grammatical information about the organization of words into syntactic units and the characteristics of syntactic relations are expressed at the word level. Since information about the relationships between syntactic elements is indicated in the form of words, these words can freely change their positions in a sentence. This is called free word order. Information about the grouping of elements can be expressed by reference to their morphological form. Such logical groups of non-contiguous elements are often called discontinuous constituents. In dependency structures, such breaks lead to non-projectivity [1].

As a language with a fairly flexible word order, Ukrainian has a higher proportion of non-projective sentences than, for example, English. The proportion of non-projective sentences in the gold standard corpus of universal dependencies for the Ukrainian language [2] used in this evaluation is 7.6%. For comparison, the Gold Standard Dependency Corpus for English [3] contains

only 2.0% non-projective sentences. An example of non-projective dependency analysis is shown in Figure 1. The arc from *сльози* to *мої* intersects the arc between *пролились* and *там*.

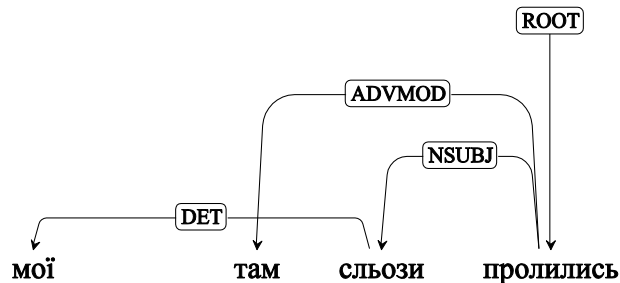


Fig. 1. Example of a non-projective dependency tree

The arc from the head word to the dependent word is called projective if there is a path from the head word to each of the dependent words that are located between the head word and the dependent word in the sentence. A dependency tree is called projective if all the arcs that form it are projective. However, there are many correct constructions that result in non-projective trees, especially in languages with relatively flexible word order [4].

Since basic transition-based parsers do not inherently produce non-projective trees, special measures must be taken to solve this problem. The parser we chose for our evaluation uses a special SWAP transition to eliminate non-projectivity.

1.2.2 Lack of resources

In terms of available tree banks and corpora, Ukrainian is a rather low-resource language. At the time of writing, there is only one available "silver standard" corpus with manually annotated dependencies. Other corpora are available, but they are either specialized, lack dependency annotations, or have machine-annotated dependencies ("silver standard"). The smaller size of the dataset may explain why the parser's accuracy for Ukrainian texts is lower than for other languages (see Table 2).

1.3 Research objectives

1.3.1 Accuracy assessment

In this study, the first objective is to evaluate the accuracy of several modern dependency analysis approaches applied to a manually annotated gold standard dataset. This allows us to establish a baseline for future work, as the ultimate goal of our research is to develop an improved dependency analyzer. To evaluate the

accuracy of syntactic analysis, we use the accuracy measures for labeled and unlabeled link analysis (LAS and UAS).

1.3.2 Error analysis

Our second goal is to analyze and classify errors made by standard syntax analyzers, as we believe that tuning the parser to handle language-specific errors is a viable way to improve the accuracy of syntax analysis. Errors that consistently appear in unrelated parser models may indicate a typical error specific to a particular parser architecture or language. Eliminating these errors could potentially allow us to create a more accurate parser in the future.

2. Other works on the topic

Most existing dependency analysis studies are not specific to the Ukrainian language. Instead, universal approaches to syntactic analysis are being developed, studied, and tested in many languages, including Ukrainian.

A brief overview is provided below, but we believe that the arc-hybrid transition-based approach, which uses deep word embeddings for encoding, is the most promising, and we should focus on improving it.

2.1 Brief overview of dependency analysis approaches

Two common approaches to dependency analysis are transition-based and graph-based architectures

The classic transition-based syntax analysis architecture is based on shift-reduce syntax analysis, a paradigm originally developed for analyzing programming languages. In transition-based parsing, we have a stack on which we build the parsing tree, a token buffer for parsing, and a parser that performs actions through a predictor called an oracle. Such parsers are generally very efficient, often having linear time complexity, but their accuracy can suffer due to the accumulation of search errors.

Graph-based dependency parsers search the space of possible trees for a given sentence for a tree (or trees) that maximizes a certain score. These methods encode the search space as directed graphs and use methods from graph theory to search the space for optimal solutions. These analyzers do not suffer from search error accumulation, but their analysis algorithms are typically more complex [4].

Although both architectures can be implemented using formal algorithms, the modern neural approach is

now almost exclusively used. Regardless of whether a neural classifier is used for the oracle in transition-based parsing, or a neural algorithm is used for scoring in graph-based parsing, the representation of tokens is a significant factor in the accuracy of parsing. The latest systems use an encoder, typically in the form of a BiLSTM, which provides contextualized representations of input words as input to transition estimation – in transition-based analyzers – or dependency arcs – in graph-based analyzers.

Neural syntactic analyzers rely on vector representations of words as their main input, often in the form of pre-trained embeddings such as word2vec [5]. These methods assign a single static representation to each word and therefore cannot capture context-dependent changes in meaning and syntactic behavior.

In contrast, deep contextualized word representations encode words with respect to the context of the sentence in which they appear. Similar to word embeddings, such models are typically trained for language modeling but produce sentence-level tensors as representations instead of individual vectors. The advantage of such models is that they not only create contextualized representations, but also do so at multiple levels of abstraction, as captured by different layers of the model, and are pre-trained on corpora much larger than typical tree banks. The use of embeddings with information about the global structure of the sentence in the representation of local functions makes transition-based and graph-based parsers practically equivalent in terms of both accuracy and error profile [6].

In addition to better computational complexity, transition-based analyzers offer another significant advantage. Both dependency analysis and sentence segmentation can be performed simultaneously using the same oracle by including an additional "end of sentence" transition. This approach by Honnibal and Johnson [7] is used in the spaCy NLP library for Python.

3. Methodology

Our work on evaluating syntactic analyzers is carried out within the Universal Dependencies framework. Universal Dependencies (UD) is a project that develops a cross-linguistically aligned annotation of a tree bank for many languages with the aim of facilitating the development of a multilingual parser, cross-lingual learning, and typological analysis research [8]. UD defines a data format for the analyzed

sentences and provides a variety of annotated datasets for different languages.

3.1 Dataset selection

To train and evaluate different parsers, we need a dataset that meets certain requirements. The data should come from various sources, such as news articles, fiction, or social media posts. The data should meet the "gold standard", i.e., be annotated manually by experts. The size of the dataset should be of the same order as the data used by Kulmiziev et al. [6], as we intend to compare our findings with the results of their work.

We have researched the available resources and found several notable Ukrainian corpora that can be downloaded and used for NLP research:

- Uber Text 2.0 [9];
- Ukrainian Brown Corpus, or BRUK [10];
- Ukrainian parliamentary stenograms annotated in Universal Dependencies [11];
- Universal Dependencies gold standard corpus for the Ukrainian language [2].

Although UberText and BRUK have a lot of diverse and balanced data, they lack the dependency annotations we need. Both the parliamentary transcripts and the gold standard corpus have dependency annotations, but the parliamentary transcripts do not have enough diverse data for our use case. The UD gold standard corpus, however, contains 122,000 tokens in 7,000 sentences from fiction, news, articles, Wikipedia, legal documents, letters, publications, and comments.

Thus, model training and evaluation should be performed on the UD gold standard corpus for the Ukrainian language.

3.2 Parser implementation

In this work, we conduct experiments described by Kulmiziev et al. [6] on our Ukrainian dataset. Therefore, we use the same system, UUParser, originally developed by de Lhoneux, Stymne, and Nivre [12]. The UUParser project includes a transition-based and a graph-based parser, both of which can be trained with or without deep word embeddings.

For an input sentence $S = w_1, \dots, w_N$, the parser creates a sequence of vectors $W_{1:N}$, where a vector $W_k = x_k \circ BiLSTM(c_{1:M})$ represents an input word w_k as a concatenation of a pre-trained word x_k embedding and a character $BiLSTM(c_{1:M})$ embedding obtained

by processing the BiLSTM sequence of the vector's w_k characters $c_{1:M}$. Next, each input element is represented as a BiLSTM vector, $h_k = \text{BiLSTM}(W_{1:N,k})$.

During transition-based syntactic analysis, BiLSTM vectors are fed into a multilayer perceptron (MLP) to count transitions using an arc-hybrid transition system extended with SWAP transitions to allow the construction of non-projective dependency trees. The evaluation is based on the top three words in the stack and the first word in the buffer, and the input to the MLP includes BiLSTM vectors for these words as well as their leftmost and rightmost dependencies (up to 12 words in total).

In graph-based analysis, BiLSTM vectors are fed into the MLP to evaluate all possible dependency relationships in the arc model, meaning that only vectors corresponding to the header and dependent are part of the input (2 words in total). Then, the syntactic analyzer extracts the maximum spanning tree over the feature matrix using the Chu-Liu-Edmonds algorithm (CLE), which allows building non-projective trees [6].

To evaluate the effectiveness of the parser using deep word embeddings, we use ELMo representations. These differ from traditional word embeddings in that each lexeme is assigned a representation that is a function of the entire input sentence. ELMo uses vectors obtained from a bidirectional LSTM trained with an object-related language model (LM) on a large text corpus. ELMo representations are deep in the sense that they are a function of all internal layers of biLM [13].

These embeddings were chosen because of the availability of a pre-trained ELMo model for the Ukrainian language [14]. More importantly, using this model allows us to make a direct comparison with the data from [6], as it includes accuracy values for both transition-based parsers and graph-based parsers with ELMo embeddings.

In this work, we conduct the experiments described by Kulmiziev et al. [6] on our Ukrainian dataset. Therefore, we use the same system, UUParser, originally developed by de Lhoneux, Stymne, and Nivre [12]. The UUParser project includes a transition-based parser and a graph-based parser, both of which can be trained with or without deep word embeddings.

For an input sentence, the parser creates a sequence of vectors, where a vector represents an input word as a concatenation of a pre-trained word embedding and a character embedding obtained by processing the

BiLSTM sequence of the vector's characters. Next, each input element is represented as a BiLSTM vector.

During transition-based syntactic analysis, BiLSTM vectors are fed into a multilayer perceptron (MLP) to count transitions using an arc-hybrid transition system extended with SWAP transitions to allow the construction of non-projective dependency trees. The evaluation is based on the top three words in the stack and the first word in the buffer, and the input to the MLP includes BiLSTM vectors for these words as well as their leftmost and rightmost dependencies (up to 12 words in total).

In graph-based analysis, BiLSTM vectors are fed into the MLP to evaluate all possible dependency relationships in the arc model, meaning that only vectors corresponding to the header and dependent are part of the input (2 words in total). Then, the syntactic analyzer extracts the maximum spanning tree over the feature matrix using the Chu-Liu-Edmonds algorithm (CLE), which allows building non-projective trees [6].

To evaluate the effectiveness of the parser using deep word embeddings, we use ELMo representations. These differ from traditional word embeddings in that each lexeme is assigned a representation that is a function of the entire input sentence. ELMo uses vectors obtained from a bidirectional LSTM trained with an object-related language model (LM) on a large text corpus. ELMo representations are deep in the sense that they are a function of all internal layers of biLM [13].

These embeddings were chosen because of the availability of a pre-trained ELMo model for the Ukrainian language [14]. More importantly, using this model allows us to make a direct comparison with the data from [6], as it includes accuracy scores for both transition-based parsers and graph-based parsers with ELMo embeddings.

3.3 Assessment metrics

To evaluate a dependency analysis system, it is necessary to measure how well it performs on a test set. The simplest approach would be to evaluate how many of the analyzed sentences exactly match the ground truth, but such a metric is too coarse to guide the development process. More detailed metrics are preferable, namely the unlabeled association score (UAS), as described by Eisner [15], and the labeled association score (LAS), the main dependency analysis metric introduced by Neuville, Hall, and Nilsson [16].

For each word, the construction of a labeled relation means the identification of the head word and the relation type, for example, NSUBJ or AMOD. Then, the accuracy of labeled relation construction, LAS, is the percentage of tokens for which both the head word and the relation type are correctly identified. The accuracy of unlabeled relations, UAS, is the percentage of tokens for which at least the headword has been correctly identified (see Fig. 2). Of the five relations, the parser incorrectly labeled *Дніпр* as NSUBJ for the word *стогне* and incorrectly identified the relation type between *широкий* and *Дніпр*. Thus, UAS is $4/5=80\%$, and LAS is $3/5=60\%$.

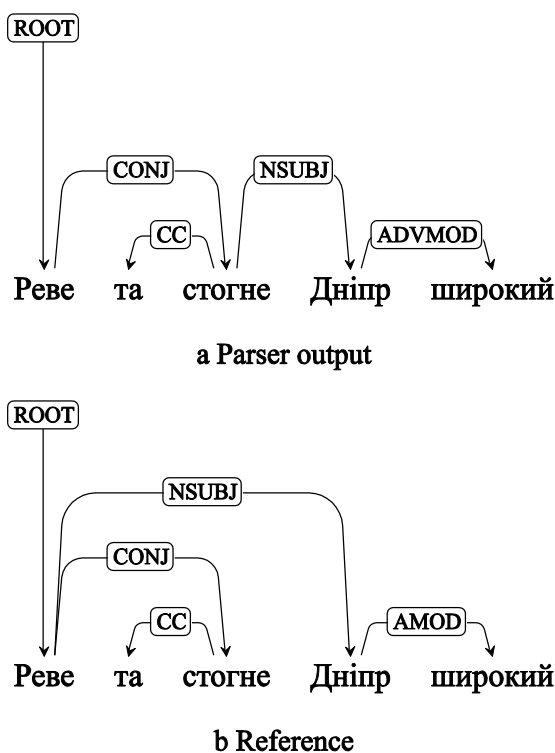


Fig. 2. Example of LAS and UAS calculation

There are other metrics that aim to achieve success where LAS falls short, namely for comparing the performance of syntactic analyzers in different languages [17]. However, since Ukrainian is the focus of our work, we will use LAS as our main metric.

3.4 Experiment plan

The dataset is divided into three parts: training (5521 sentences / 93161 dependencies), development (673 sentences / 12622 dependencies), and testing (898 sentences / 17249 dependencies). We train several models using the same dataset but different random initial weights and evaluate their performance.

3.4.1 Accuracy evaluation

We evaluate four different parser architectures: transition-based, graph-based, transition-based with deep embeddings, and graph-based with deep embeddings. For each architecture, the parser is trained on the training dataset for 30 epochs with default parser settings. After each epoch, the model is applied to the development dataset to determine the LAS. The epoch model with the highest LAS value is selected. The training process is repeated three times with different random initial weights, creating three models. Each model is applied to the test dataset to determine the LAS, and the average LAS is calculated.

3.4.2 Error analysis

To analyze the error pattern, we train 100 models based on transitions with ELMo embeddings with different random initial weights.

Error analysis is performed by comparing the error frequency for dependencies grouped by certain features with the baseline error frequency (for the entire test dataset). This analysis takes into account the error frequency in both labeled and unlabeled connections. We investigate the following features: projectivity of the relation, sentence length, distance between the head and dependent words, and part of speech of the dependent word.

For each group, the difference in error rates from the baseline is evaluated for statistical significance using the chi-square test ($p < 0.05$).

4. Experiment results

4.1 Overall accuracy and comparison with other languages

The accuracy assessment results are summarized in **Tables 1 and 2**. The assessments for languages other than Ukrainian were obtained by Kulmiziev et al. [6].

Table 1. LAS by architecture (individual experiments and average)

Architecture	#1	#2	#3	Avg.
Transition	77.57	78.03	77.80	77.80
Graph	79.34	78.81	78.97	79.04
Transition + ELMo	84.71	84.79	84.47	84.66
Graph + ELMo	83.84	83.37	83.78	83.66

Four parsing systems were analyzed: transition-based (Transition), graph-based (Graph), transition-based

with deep ELMo embeddings (Transition + ELMo), and graph-based with deep ELMo embeddings (Graph + ELMo). The LAS value was obtained for three independently trained models with different random initial weights (#1, #2, and #3), as well as the average LAS value (Avg.).

Table 2. Size of the training dataset (sentences) and average LAS for Ukrainian and other languages

Language	Train	TR	GR	TR+E	GR+E
Ukrainian	5.5k	77.8	79.0	84.7	83.7
English	12.5k	82.7	83.3	87.0	86.5
Hindi	13.3k	88.4	89.6	91.0	91.2
Italian	13.1k	88.0	88.2	90.9	90.6
Russian	48.8k	88.3	88.0	90.7	90.6
Swedish	4.3k	80.5	81.6	86.9	86.2

For each language, the number of sentences in the training dataset (Train) is shown, as well as the average LAS score for each of the systems: transition-based (TR), graph-based (GR), transition-based with deep ELMo embeddings (TR+E), and graph-based with deep ELMo embeddings (GR+E).

Our results show that without deep embeddings, the graph-based parser offers better accuracy than the transition-based parser. We believe this is due to the high frequency of non-projective parse trees in the Ukrainian language.

The use of ELMo embeddings significantly improves the accuracy of the syntactic analyzer in both transition-based parsers and graph-based analyzers. The accuracy of transition-based and graph-based parsers on the Ukrainian dataset is similar.

Our findings confirm the general trend described in [6]: the use of deep word embeddings significantly improves analysis accuracy and eliminates the gap between graph-based and transition-based parsers.

It should be noted that the accuracy of Ukrainian sentence parsing is low compared to other Indo-European languages, as shown in Table 2.

This is probably due to the fact that the dataset for Ukrainian is relatively small compared to other languages. Nevertheless, the general trend of improvement from the use of embeddings is consistent with other languages.

4.2 Error profile analysis for a transition-based parser with deep word embeddings

4.2.1 Non-projectivity

Conclusions regarding the frequency of errors grouped by projectivity are presented in Table 3. Error levels above the baseline are highlighted in bold.

Coefficients with statistically insignificant differences from the baseline are indicated in square brackets.

Table 3. Error rates for labeled and unlabeled relations grouped by projectivity at the level of a single dependency and at the sentence level

	Prevalence of feature, %	Unlabeled errors, %	Labeled errors, %
Basic level	–	12.29	15.47
Non-projective dependencies	2.57	29.88	33.71
Projective dependencies	97.43	11.83	14.99
Projective dependencies in non-representative sentences	9.76	[12.31]	[15.44]
Projective dependencies in fully projective sentences	87.67	11.77	14.94

There is a strong correlation between the projectivity of dependencies and the accuracy of parsing. The error rate among the analyzed non-projective dependencies is significantly higher than among projective dependencies: 29.88% vs. 11.83 error rates for unlabeled relations and 33.71% vs. 14.99 for labeled relations.

If a sentence contains a non-projective relation, other projective relations are not associated with higher error rates than relations in fully projective sentences. The error rates for both labeled and unlabeled relations do not differ significantly from the baseline level.

4.2.2 Sentence length

The results for error frequencies grouped by sentence length are shown in Table 4. For this analysis, we grouped sentences by length into 8 quantiles. Error levels above the baseline level are highlighted in bold. Coefficients with a statistically insignificant difference from the baseline level are indicated in square brackets.

Table 4. Error rate for labeled and unlabeled relations grouped by sentence length

	Prevalence of feature, %	Unlabeled errors, %	Labeled errors, %
Basic level	–	12.29	15.47
2 – 11 tokens	12.51	7.86	11.44
12 – 17 tokens	14.63	10.46	14.37
18 – 21 tokens	11.61	11.99	14.78
22 – 27 tokens	13.74	11.87	[15.36]
28 – 33 tokens	10.89	13.39	16.51
34 – 41 tokens	11.69	14.34	17.31
42 – 58 tokens	12.52	14.18	17.34
59 – 140 tokens	12.41	14.86	17.10

Longer sentences are clearly associated with higher error rates. The error rate for both labeled and unlabeled

connections exceeds the baseline level for sentences with a length of 28 tokens or more.

From a linguistic point of view, these longer sentences are often either "sentences within sentences" (e.g., direct speech) or several different sentence structures joined together. Since the syntactic analyzer works with greater accuracy on shorter sentences, we must explore different segmentation methods that would allow us to split sentences, analyze the parts, and then merge them into a single tree.

4.2.3 Distance to the head word

The conclusions for the frequency of errors grouped by the distance between the dependent word and the head word are shown in **Table 5**. In this grouping, a negative distance means that the head word precedes the dependent word, and a positive distance means that the head word follows the dependent word. Root nodes are highlighted in a separate group. Error rates above the baseline are highlighted in bold. All differences from the baseline are statistically significant.

Table 5. Error rate for labeled and unlabeled relations grouped by distance between dependent and head words

	Prevalence of feature, %	Unlabeled errors, %	Labeled errors, %
Basic level	–	12.29	15.47
– 4 and further	15.09	33.00	37.82
– 3	5.50	19.51	23.58
– 2	10.45	9.99	13.65
– 1	12.60	7.38	11.19
0 (root node)	6.44	8.07	8.62
1	28.25	3.42	5.54
2	10.05	8.16	10.81
3	4.53	11.23	14.84
4 and further	7.09	20.43	24.76

The error rate increases with greater distances between the dependent and main words. It should be noted that accuracy in the "forward" direction is better than in the "backward" direction, i.e., for the same absolute distance, the error rate is lower when the dependent word precedes the corresponding main word.

The base error rate is exceeded when the main word is 4 or more words away from the dependent word in the forward direction and 3 or more words away in the reverse direction.

4.2.4 Part of speech of the dependent word

Conclusions regarding the frequency of errors grouped by the part of speech of the dependent word are given in **Table 6**. Error frequencies above the baseline are highlighted in bold. Coefficients with a statistically

insignificant difference from the baseline are indicated in square brackets.

Table 6. Error rate for labeled and unlabeled links grouped by part of speech of the dependent word

	Prevalence of feature, %	Unlabeled errors, %	Labeled errors, %
Basic level	–	12.29	15.47
NOUN	26.30	15.95	19.78
PUNCT	18.15	12.47	12.47
ADJ	11.46	7.34	8.72
VERB	9.21	15.22	20.09
ADP	9.21	2.38	2.56
ADV	4.16	17.79	24.44
PROPN	3.67	15.96	20.70
DET	3.65	8.88	13.37
CCONJ	3.65	8.91	10.85
PRON	3.01	10.74	18.25
NUM	2.26	7.62	19.40
PART	2.17	19.35	26.95
SCONJ	1.51	6.17	10.10
X	0.72	37.33	52.58
AUX	0.70	13.38	[15.68]
SYM	0.10	22.35	43.06
INTJ	0.06	1.40	6.80

It should be noted that for punctuation (PUNC), the frequency of marked and unmarked errors is the same, since all punctuation tokens have the "punct" relation type.

The frequency of errors in pronouns (PRON) and numerals (NUM) deserves special attention. For these categories, the frequency of unlabeled errors is lower than the baseline, while the frequency of labeled errors exceeds it. This means that the analyzer often correctly identifies the head word but not the relation type.

Methods could be developed to improve recognition accuracy in this case. We believe that formal rules could be applied to check and possibly correct the labeling.

5. Conclusions

We trained several models for syntactic analysis with different architectures and evaluated their performance on a manually annotated Ukrainian dataset.

We found that the parsing accuracy for the Ukrainian dataset is lower than for many other Indo-European languages, but the overall trend of accuracy improvement with deep word embeddings is consistent with existing research.

We analyzed the error profile of the syntactic analyzer based on transitions with deep word embeddings and demonstrated that the decrease in syntactic analysis accuracy is related to non-projectivity of dependencies,

longer sentences, and greater distance between the head word and the dependent.

We found that pronouns and numerals meet the baseline metrics for unmarked relations but have significantly higher error rates for marked relations.

In future work, we intend to explore ways to improve syntactic analysis accuracy by supplementing neural analysis with other approaches, such as formal rules or pre- and post-processing.

References

1. Tsarfaty, R.; Seddah, D.; Goldberg, Y.; Kuebler, S.; Versley, Y.; Candito, M.; Foster, J.; Rehbein, I.; Tounsi, L. (2010), "Statistical Parsing of Morphologically Rich Languages (SPMRL) What, How and Whither". Proceedings of the NAACL HLT 2010 First Workshop on Statistical Parsing of Morphologically-Rich Languages. P. 1–12. URL: <https://aclanthology.org/W10-1401/>
2. Kotsyba, N.; Moskalevskiy, B.; Romanenko, M.; Samoridna, H.; Kosovska, I.; Lytvyn, O.; Orlenko, O.; Brovko, H.; Matushko, B.; Onyshchuk, N.; Pareviazko, V.; Rychyk, Y.; Stetsenko, A.; Umanets, S.; Masenko, L. (2021), "Gold standard Universal Dependencies corpus for Ukrainian (UD_Ukrainian-IU) v2.8". URL: https://github.com/UniversalDependencies/UD_Ukrainian-IU.
3. Silveira, N.; Dozat, T.; de Marneffe M.-C.; Bowman, S.; Connor M.; Bauer, J.; Manning C. (2014), "A Gold Standard Dependency Corpus for English". Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014). P. 2897–2904. URL: http://www.lrec-conf.org/proceedings/lrec2014/pdf/1089_Paper.pdf
4. Jurafsky D., Martin J. "Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models, 3rd ed". USA. 2025. 599 p. URL: <https://web.stanford.edu/~jurafsky/slp3/>
5. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. (2013), "Efficient Estimation of Word Representations in Vector Space". arXiv: 1301.3781 [cs.CL]. DOI: 10.48550/arXiv.1301.3781
6. Kulmizev, A.; de Lhoneux, M.; Gontrum, J.; Fano, E.; Nivre, J. (2019), "Deep Contextualized Word Embeddings in Transition-Based and Graph-Based Dependency Parsing - A Tale of Two Parsers Revisited". *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. P. 2755–2768. DOI: 10.18653/v1/D19-1277
7. Honnibal M.; Johnson M. (2015), "An Improved Non-monotonic Transition System for Dependency Parsing". Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. P. 1373–1378. DOI: 10.18653/v1/D15-1162
8. De Marneffe, M.-C.; Manning, C.; Nivre, J.; Zeman, D. (2021), "Universal Dependencies". *Computational Linguistics*. Vol. 47. No. 2. P. 255–308. DOI: 10.1162/coli_a_00402
9. Chaplynskyi D. (2023), "Introducing UberText 2.0: A Corpus of Modern Ukrainian at Scale". *Proceedings of the Second Ukrainian Natural Language Processing Workshop (UNLP)*. P. 1–10. DOI: 10.18653/v1/2023.unlp-1.1
10. Starko V.; Rysin A. (2023), "Creating a POS Gold Standard Corpus of Modern Ukrainian". *Proceedings of the Second Ukrainian Natural Language Processing Workshop (UNLP)*. P. 91–95. DOI: 10.18653/v1/2023.unlp-1.11.
11. Shvedova M.; Lukashevskiy A. (2024), "UD_Ukrainian-ParlaMint". URL: https://github.com/UniversalDependencies/UD_Ukrainian-ParlaMint.
12. De Lhoneux, M.; Szymne S.; Nivre J. (2017), "Arc-Hybrid Non-Projective Dependency Parsing with a Static-Dynamic Oracle". Proceedings of the 15th International Conference on Parsing Technologies. P. 99–104. URL: <https://aclanthology.org/W17-6314/>
13. Peters, M.; Neumann, M.; Iyyer, M.; Gardner, M.; Clark, C.; Lee, K.; Zettlemoyer, L. (2018), "Deep Contextualized Word Representations". *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. P. 2227–2237. DOI: 10.18653/v1/N18-1202
14. Che, W.; Liu, Y.; Wang, Y.; Zheng, B.; Liu, T. (2018), "Towards Better UD Parsing: Deep Contextualized Word Embeddings, Ensemble, and Treebank Concatenation". *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. P. 55–64. DOI: 10.18653/v1/K18-2005
15. Eisner J. (1996), "Three New Probabilistic Models for Dependency Parsing: An Exploration". COLING 1996 Volume 1: The 16th International Conference on Computational Linguistics. P. 340–345. URL: <https://aclanthology.org/C96-1058/>
16. Nivre J.; Hall J.; Nilsson J. (2004), "Memory-Based Dependency Parsing". Proceedings of the Eighth Conference on Computational Natural Language Learning (CoNLL-2004) at HLT-NAACL 2004. P. 49–56. URL: <https://aclanthology.org/W04-2407/>
17. Nivre J.; Fang C.-T. (2017), "Universal Dependency Evaluation". Proceedings of the NoDaLiDa 2017 Workshop on Universal Dependencies (UDW 2017). P. 86–95. URL: <https://aclanthology.org/W17-0411/>

Syrotkin Kostiantyn – National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Master, Kyiv, Ukraine; e-mail: k.syrotkin@kpi.ua; ORCID ID: <https://orcid.org/0009-0008-4150-8325>

Сироткін Костянтин Вікторович – Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", магістр, Київ, Україна.

ОЦІНЮВАННЯ ТОЧНОСТІ ТА АНАЛІЗ ПОМИЛОК РОЗБОРУ ЗАЛЕЖНОСТЕЙ ДЛЯ ТЕКСТІВ УКРАЇНСЬКОЮ МОВОЮ

Предметом дослідження є розбір залежностей в межах фреймворку *Universal Dependencies* для речень українською мовою. **Метою роботи** є оцінювання і порівняння точності розбору залежностей, яка досягається декількома сучасними системами на стандартному наборі даних українською мовою, а також аналіз профілю помилок таких систем. У статті виконано два **завдання**. Перше – оцінювання точності декількох сучасних систем розбору залежностей із використанням анотованого вручну стандартного набору даних. Метрикою точності обрано відсоток правильно побудованих позначених та непозначених зв'язків. Друге завдання – аналіз та категоризація помилок, яких припускаються стандартні системи. Виявлення та усунення цих помилок потенційно дасть змогу в подальшому створити більш точну систему розбору. У досліджуваних системах використовуються **методи** машинного навчання і нейронних мереж разом з **методами** теорії автоматів і теорії графів, а також глибокі векторні подання слів. **Основні результати**. Для української мови точність синтаксичного розбору значно підвищується з використанням глибокого векторного подання слів. Система на основі переходів із глибоким векторним поданням слів забезпечує найвищий показник точності побудови позначених зв'язків на рівні 84,66% для тестового набору даних. Для цієї самої системи підвищення рівня помилок спостерігається для непроективних залежностей, довгих речень і більшої відстані до основного слова. Крім того, для займенників і числівників рівень помилок для позначених зв'язків значно перевищує базовий рівень, тоді як помилки для непозначених зв'язків залишаються на базовому рівні. **Висновки**: точність розбору залежностей для української мови поступається аналогічним показникам для інших мов, однак загальна тенденція підвищення точності з використанням глибокого векторного подання слів узгоджується з результатами попередніх досліджень. Для покращення загальної точності розбору важливо зосередитися на таких проблемних аспектах, як непроективні залежності, довгі речення та велика відстань між основним і залежним словами. У подальших дослідженнях заплановано дослідити можливості підвищення точності синтаксичного розбору способом доповнення підходу на основі машинного навчання іншими методами, зокрема використанням формальних правил або етапів попереднього та подальшого оброблення.

Ключові слова: оброблення природної мови; синтаксичний розбір; розбір залежностей; *Universal Dependencies*.

Бібліографічні описи / Bibliographic descriptions

Сироткін К. В. Оцінювання точності та аналіз помилок розбору залежностей для текстів українською мовою. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 2 (32). С. 102–110. DOI: <https://doi.org/10.30837/2522-9818.2025.2.102>

Syrotkin, K. (2025), "Accuracy evaluation and error analysis of dependency parsing of texts in Ukrainian", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (32), P. 102–110. DOI: <https://doi.org/10.30837/2522-9818.2025.2.102>

Є. СОПОВ, Д. КРИЦЬКИЙ, А. АРТЬОМОВА, І. АРТЬОМОВ

МЕТОДИКА ПОБУДОВИ ДОВІРЧОЇ МАРШРУТИЗАЦІЇ В РОЙОВИХ МЕРЕЖАХ БПЛА НА ОСНОВІ АНАЛІЗУ ТРАФІКУ ТА ВИЯВЛЕННЯ АНОМАЛІЙ

Предметом дослідження є процеси забезпечення безпечної маршрутизації та обміну даними між безпілотними літальними апаратами в складі ройових мереж в умовах дії кіберзагроз, зокрема атак типу *Black Hole*. **Мета роботи** – розроблення та моделювання захищеного механізму передачі інформації в ройових мережах БПЛА, що зважає на рівень довіри до вузлів та можливість виявлення шкідливих учасників мережі на основі аналізу їх поведінки. **Дослідження** також спрямовано на розроблення методики захищеної та енергоефективної маршрутизації в ройових мережах безпілотних літальних апаратів на основі блокчейн-технологій і моделей оцінювання довіри. Запропонована методика забезпечує стійкість до кіберзагроз, збереження цілісності інформації та мінімізацію витрат ресурсів. **Завдання**, які передбачено розв'язати в дослідженні: здійснити комплексний аналіз вразливостей традиційних протоколів маршрутизації, що застосовуються в ройових мережах БПЛА (FANET), з метою виявлення потенційних загроз інформаційній безпеці; обґрунтувати доцільність використання механізмів оцінювання довіри (*trust-based mechanisms*) як засобу підвищення надійності та стійкості маршрутизації до внутрішніх атак; реалізувати модель атаки типу *Black Hole* в симуляційному середовищі NS-3 для дослідження її впливу на функціонування ройової мережі; розробити механізм обліку кількості переданих пакетів кожним вузлом як базовий інструмент для побудови системи оцінювання довіри між агентами мережі; здійснити візуалізацію результатів моделювання для забезпечення наочності, подальшого аналізу й порівняння ефективності запропонованих рішень. **Методи**, використані в дослідженні, передбачають моделювання мережі FANET у симуляторі NS-3.36, застосування мобільної моделі *RandomWaypoint*, протоколу маршрутизації AODV, статистичного аналізу показників передачі пакетів вузлами та побудову графічної інтерпретації довірчих метрик. Розроблений код реалізує атаку *Black Hole* внаслідок утручання на рівні *NetDevice* та забезпечує логування всіх переданих пакетів у вигляді CSV-файлу для подальшого аналізу. У дослідженні використано комплекс методів, що поєднують інструменти симуляційного моделювання, аналітичного аналізу та візуалізації. Основу експериментального середовища становить моделювання ройової мережі БПЛА (FANET) у симуляторі NS-3.36 із застосуванням мобільної моделі *RandomWaypoint*, яка відтворює характерну поведінку вузлів у динамічному середовищі. Для організації маршрутизації використано протокол *Ad hoc On-Demand Distance Vector* (AODV), що дає змогу дослідити вплив внутрішніх атак на процес побудови маршрутів у мережі. Модель атаки *Black Hole* реалізовано способом втручання у процес обміну маршрутною інформацією, що допомагає симулювати шкідливу поведінку вузлів. **Результати**. Дослідження довело ефективність запропонованого підходу до встановлення зловмисників у ройовій мережі. Метрики виявили нетипову активність вузла, який атакує, відсутність передачі пакетів, що відрізняється від поведінки звичайних вузлів. Це сприяє оперативному виявленню загроз і вилученню зловмисника з маршрутизації. Графічна візуалізація чітко демонструє розподіл активності між вузлами, що спрощує інтерпретацію результатів без потреби в глибокому аналізі логів. **Висновки**. Запропонований *trust*-механізм у поєднанні з аналізом активності вузлів ефективно захищає FANET-мережі від атак типу *Black Hole*. Перспективним є розширення моделі внаслідок упровадження більш складних методів оцінювання довіри, таких як багатофакторний аналіз, блокчейн чи машинне навчання. Також доцільно розробити адаптивні алгоритми маршрутизації з автоматичним вимкненням або ізоляцією підозрілих вузлів.

Ключові слова: ройова мережа БПЛА; FANET; AODV; атака *Black Hole*; довіра (*trust*); симуляція в NS-3; безпека маршрутизації; захист від атак; *NetDevice*; виявлення зловмисників.

Вступ

Швидкий розвиток технологій мобільного зв'язку п'ятого покоління (5G New Radio, 5G NR) відкриває нові можливості для застосування в різних сферах, зокрема в промисловому виробництві, телемедицині, хмарній освіті та автономних системах управління безпілотними літальними апаратами [1]. Завдяки таким властивостям, як надвисока пропускна

здатність, мінімальні затримки передачі та висока щільність з'єднань, 5G NR створює сприятливі умови для широкомасштабного розгортання ройових систем БПЛА. Це забезпечує ефективне виконання складних місій з багатьма завданнями в динамічному середовищі з високим рівнем автономності та кооперації між апаратами. Комплексний підхід до управління групами роботів [16] наголошує на значенні адаптивних стратегій для динамічних середовищ.

Основою взаємодії в таких системах є ройові мережі БПЛА (FANET – *Flying Ad-Hoc Networks*), що функціують як самоконфігуровані бездротові мережі без централізованої інфраструктури. Водночас відкритий характер цих мереж, висока мобільність вузлів і часта зміна топології створюють сприятливі умови для реалізації низки кіберзагроз. Найбільш поширеними серед них є атаки типу *Black Hole*, *Replay*, *Sybil* та *Man-in-the-Middle* (MITM), що дають змогу зловмисникам перехоплювати, спотворювати або повністю блокувати передачу інформації, що критично впливає на стабільність і безпечність функціонування системи загалом.

Попри значну кількість наявних рішень у сфері маршрутизації, зокрема протоколів AODV і GSPSR, а також впровадження стандартних криптографічних інструментів на основі DTLS, наявні підходи демонструють обмежену ефективність у контексті високої мобільності, змінної топології та обмежених обчислювальних ресурсів, притаманних ройовим мережам БПЛА. Застосування традиційного шифрування та централізованих механізмів виявляється недостатнім для забезпечення надійного захисту маршрутної інформації в умовах децентралізованої динамічної мережі.

У зв'язку з цим актуальним постає завдання розроблення нового підходу до захищеної маршрутизації, який поєднує механізм оцінювання довіри до мережевих вузлів, використання легкого блокчейну для збереження маршрутної інформації та консенсусного алгоритму на основі *Proof-of-Traffic* (PoT). Така архітектура має забезпечити розподілений характер прийняття рішень, знизити енергоспоживання внаслідок пасивної синхронізації блоків і водночас зберігати стійкість до типових атак, що спостерігаються у FANET. Важливим також є увага до вимог масштабованості, енергоефективності та автономності, які є ключовими для практичного застосування систем такого класу.

Аналіз останніх досліджень і публікацій

Проблематика захищеної маршрутизації в ройових мережах безпілотних літальних апаратів активно розробляється в сучасній науковій літературі. Ванг та інші [2] запропонували алгоритм маршрутизації, що базується на легкому блокчейні з використанням консенсусного механізму *Proof-of-Traffic* (PoT). Такий підхід дає змогу зменшити навантаження

на мережу та знизити енергоспоживання, що є критичним для обмежених за ресурсами мультиагентних систем.

У дослідженні Альшаммарі та інших [3] розглянуто підхід до побудови безпечної *ad-hoc*-мережі БПЛА з використанням блокчейн-технологій, який не знижує пропускну здатність системи. Запропонована мережева архітектура забезпечує підвищений рівень захисту інформації та стійкість до атак. Гупта Р. та інші [4] зосередили увагу на створенні адаптивної мережевої архітектури з використанням унікальних ідентифікаторів, механізмів участі в консенсусі та контролю безпеки маршрутів. Такий підхід є особливо актуальним для реалізації динамічних ройових систем, що функціують у змінному середовищі.

У праці [10] запропоновано архітектуру безпечного обслуговування мережі БПЛА, що поєднує методи глибокого навчання з блокчейн-технологією. Досягнуті результати демонструють високу ефективність протидії кіберзагрозам, навіть в умовах обмежених обчислювальних можливостей.

К. Мершад [11] розробив протокол PROACT, який базується на паралельній генерації блоків кількома вузлами на основі моделі накопиченої довіри. Такий механізм сприяє зниженню затримок і покращенню масштабованості системи.

Останнє з розглянутих досліджень – система BETA-UAV [12], яка забезпечує автентифікацію вузлів у ройових мережах із використанням смартконтрактів. Система гарантує актуальність сесій зв'язку та захищає від атак типу *replay*, *spoofing* та *impersonation*. Її реалізація на платформі *Ethereum* підтвердила високу ефективність і низьке споживання ресурсів.

Розглянуті дослідження створюють наукове підґрунтя для формування комплексного підходу, що інтегрує блокчейн-технології, моделі довіри та енергоефективні стратегії маршрутизації в ройових мережах БПЛА.

Визначення не розв'язаних раніше частин загальної проблеми. Мета роботи й завдання

Ройові мережі безпілотних літальних апаратів розвиваються паралельно зі зростанням складності завдань, які ці системи покликані виконувати у сферах моніторингу, безпеки, логістики та надзвичайного реагування. Тоді як мобільність, автономність і самоорганізація забезпечують таким

системам унікальні переваги, саме ці характеристики викликають значні труднощі для підтримки надійного та безпечного обміну інформацією всередині рою.

У реальних умовах ройові мережі БПЛА функціують у середовищі з високою динамікою топології, нестабільними каналами зв'язку та обмеженим енергозабезпеченням. Це робить їх особливо вразливими до атак з боку внутрішніх або зовнішніх суб'єктів, здатних впливати на процес маршрутизації, викривляти дані, порушувати логіку маршруту або навіть повністю блокувати інформаційний обмін. Найбільш поширеними є атаки типу *Black Hole*, MITM (*Man-in-the-Middle*), *Replay* та *Sybil*, спрямовані на порушення цілісності, автентичності або вчасності передачі маршрутної інформації.

Аналіз наявних рішень показує, що поширені протоколи маршрутизації, зокрема AODV та GPSR, не зважають на зміну довіри до вузлів

та не мають вбудованих засобів виявлення аномалій у поведінці учасників мережі. Застосування класичних криптографічних схем (наприклад, DTLS) є обмеженим через високу енерговитратність і невідповідність вимогам децентралізованих, масштабованих систем з обмеженими ресурсами. Крім того, відсутність об'єднаної структури, яка б поєднувала довірчі механізми із засобами захисту маршрутизації на структурному рівні, залишає такі системи вразливими до складних багаторівневих атак. Як описано в роботі [10], поєднання блокчейну з глибоким навчанням може значно підвищити ефективність виявлення атак.

З метою узагальнення основних недоліків традиційних підходів та визначення напрямів їх усунення доцільно подати відповідні спостереження в табличній формі (табл. 1).

Таблиця 1. Недоліки чинних підходів і шляхи вдосконалення

Ключова проблема	Недоліки наявних рішень	Можливі способи розв'язання
Виявлення зловмисних вузлів у режимі реального часу	Відсутність оцінки довіри у AODV, GPSR	Побудова динамічного довірчого рівня з аналізом поведінки вузлів
Захист від внутрішніх атак (<i>Black Hole</i> , <i>Sybil</i>)	DTLS не працює без централізованої інфраструктури	Інтеграція блокчейну з локальною автентифікацією вузлів
Енергоспоживання та масштабованість	PoW/PoS надто ресурсомісткі для малопотужних БПЛА	<i>Proof-of-Traffic</i> (PoT) як легкий консенсус для FANET
Неспроможність адаптації до динамічної топології	Статичні таблиці маршрутів швидко застарівають	Пасивна синхронізація маршрутної інформації через блокчейн
Роз'єднаність засобів безпеки та довіри	Системи довіри не інтегровані в логіку маршрутизації	Об'єднана архітектура маршрутизації, довіри та захисту даних

Аналіз показує, що забезпечення надійної маршрутизації в ройових мережах БПЛА потребує комплексного підходу, який передбачає оцінювання довіри, захист маршрутної інформації та адаптивність до змін у структурі мережі. Водночас такий підхід має зважати на обмеження, пов'язані з ресурсами та відсутністю централізованого керування.

Відповідно, **метою дослідження** є розроблення методики захищеної та енергоефективної маршрутизації в ройових мережах безпілотних літальних апаратів на основі блокчейн-технологій і моделей оцінювання довіри, яка забезпечує стійкість до кіберзагроз, збереження цілісності інформації та мінімізацію витрат ресурсів.

Для досягнення поставленої мети передбачається виконати низку **завдань**:

1) здійснити комплексний аналіз вразливостей традиційних протоколів маршрутизації, що застосовуються в ройових мережах БПЛА (FANET),

з метою виявлення потенційних загроз інформаційній безпеці;

2) обґрунтувати доцільність використання механізмів оцінювання довіри (*trust-based mechanisms*) як засобу підвищення надійності та стійкості маршрутизації до внутрішніх атак;

3) реалізувати модель атаки типу *Black Hole* в симуляційному середовищі NS-3 для дослідження її впливу на функціонування ройової мережі;

4) розробити механізм обліку кількості переданих пакетів кожним вузлом як базовий інструмент для побудови системи оцінювання довіри між агентами мережі;

5) візуалізувати результати моделювання для забезпечення наочності, подальшого аналізу та порівняння ефективності запропонованих рішень.

Розроблення адаптивної, енергоощадної та стійкої до атак архітектури маршрутизації на основі довіри та блокчейну має потенціал суттєво підвищити надійність ройових систем БПЛА та забезпечити

безперервність їх функціонування в агресивному кіберсередовищі.

Матеріали й методи

Для забезпечення захищеної маршрутизації в ройових мережах безпілотних літальних апаратів (UAS) було застосовано комплексний досліджувальний підхід, що охоплює побудову тривірневої архітектури з елементами довірчого аналізу, блокчейн-реєстру

та феромонної маршрутизації. Загальний фреймворк реалізовано з огляду на особливості мереж 5G NR, які забезпечують низьку затримку, високу щільність сполучень і покращену пропускну спроможність для високодинамічних сценаріїв повітряної мобільності.

Необхідність упровадження захищеної маршрутизації в ройових мережах UAS зумовлена численними загрозами, зокрема атакою типу *Black Hole*, приклад якої подано на рис. 1.

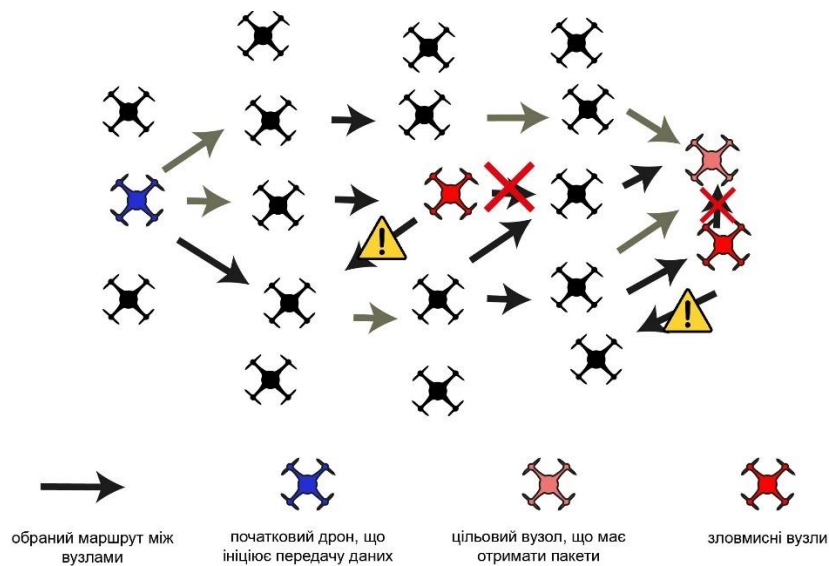


Рис. 1. Сценарій атаки типу *Black Hole* в ройовій мережі БПЛА

На рис. 1 зображено типову ситуацію, що може виникнути в ройовій мережі безпілотних літальних апаратів, коли відбувається внутрішнє проникнення зловмисних вузлів. Вони вбудовуються в маршрут між легітимними дронами – початковим дроном та цільовим вузлом – і перехоплюють або блокують передані пакети даних.

Стрілки ілюструють шляхи маршрутизації, якими рухаються пакети. Зловмисні вузли можуть виконувати атаки типу *Black Hole* (перехоплення з подальшим знищенням пакета), *Man-in-the-Middle* (втручання в процес передачі) або *Replay* (повторне надсилання раніше перехоплених повідомлень).

Через такі атаки окремі ділянки маршруту стають недоступними або небезпечними для передачі, що призводить до порушення комунікації, затримок у доставці та зниження ефективності всієї місії.

Безпечна маршрутизація в ройових мережах БПЛА на основі довіри та PoT-блокчейну. Ройові мережі БПЛА, відомі також як FANET (*Flying Ad-hoc*

Network), дають змогу координувати дії рою дронів для виконання складних місій. Водночас така мережа є вразливою до зловмисних вузлів: компрометований дрон може перехоплювати чи модифікувати пакети, відмовлятися їх ретранслювати або впроваджувати атаки "людина посередині". Традиційні підходи захисту (шифрування каналів) не гарантують виявлення зрадницьких вузлів і можуть потребувати значних ресурсів. Необхідна надійна архітектура маршрутизації, що забезпечує відсіювання зловмисників, довірче обрання маршрутів і стійкість до атак з огляду на обмежені обчислювальні та енергетичні ресурси дронів (рис. 2).

Принцип функціонування маршрутизації в рої БПЛА, основаної на поєднанні географічного аналізу, довірчих показників і блокчейн-автентифікації. Початковий дрон ініціює передачу інформації до цільового вузла, прокладаючи маршрут крізь низку проміжних вузлів, кожен з яких оцінюється з погляду надійності та участі в трафіку.

Запропонована архітектура узгоджується з підходами, використовуваними Х. Чжоу [9], де блокчейн використовується для підвищення стійкості мережі рою БПЛА.

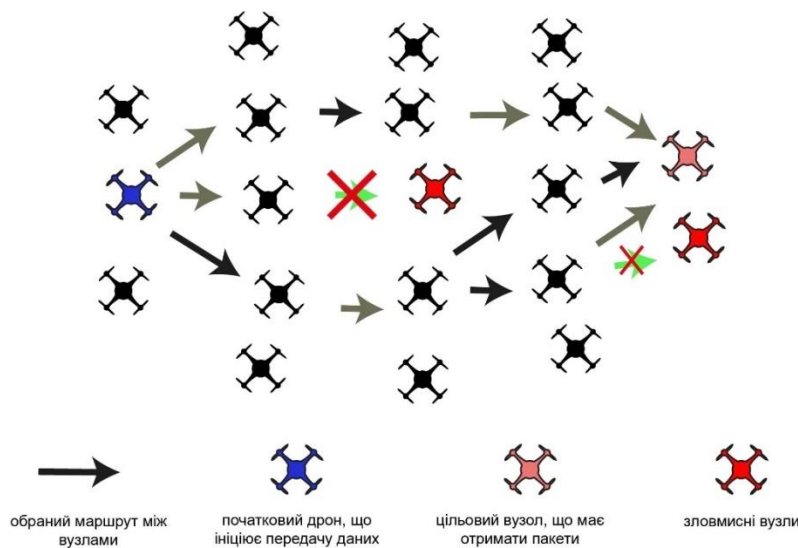


Рис. 2. Архітектура побудови довірчих маршрутів у рої БПЛА

Ключовим елементом механізму є вилучення з маршруту дронів, ідентифікованих як потенційно зловмисних, що можуть ігнорувати, блокувати або модифікувати передані пакети. Таких учасників виявляють за допомогою локальних перевірок автентичності з використанням дайджестів блоків (HN) у розподіленому реєстрі, що зберігається в кожному дроні до старту місії.

Побудова маршруту здійснюється з огляду на напрямок до цільового вузла (вектор $v_{s,d}$), що дає змогу скоротити кількість перевірок та обмежити коло потенційних хопів. Після попереднього фільтрування береться до уваги інтенсивність феромонної активності, яка відтворює ступінь залученості вузла до успішної маршрутизації в минулому. Вузол із найвищим рівнем феромону обирається як пріоритетний для подальшої передачі пакетів, оскільки ймовірність його компрометації зменшується зі зростанням навантаження.

Завдяки такому підходу система забезпечує стабільне функціонування, навіть у присутності зловмисних вузлів, значно зменшує загальне енергоспоживання (внаслідок уникнення широкомовного розсилання запитів) і підвищує загальну стійкість до атак типу *Black Hole*, *MITM* або *Replay*.

Архітектурна модель ґрунтується на тривірневій структурі, що поєднує:

1) шар довіри (*Trust Layer*) – для оцінювання надійності вузлів на основі поведінкових метрик;

2) PoT-блокчейн-рівень – для розподіленого збереження довірчих даних і автентифікації;

3) маршрутизаційне ядро (*Routing Core*) – для побудови маршрутів з огляду на довіру та підтвердження легітимності вузлів.

Такий підхід дає змогу інтегрувати блокчейн і систему оцінювання довіри для забезпечення стійкої та захищеної маршрутизації в децентралізованих ройових мережах БПЛА. Особливу увагу приділено консенсусному механізму PoT, що є альтернативою енерговитратним алгоритмам на кшталт *Proof-of-Work* і спеціально адаптований до умов обмежених обчислювальних ресурсів у бездротових мережах. Структурну взаємодію між рівнями зображено на рис. 3.

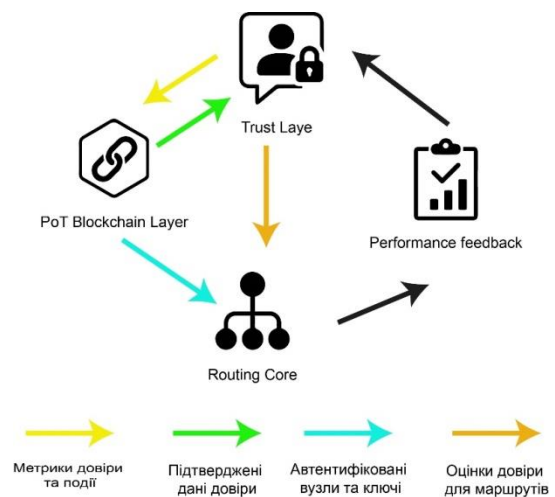


Рис. 3. Тривірнева архітектурна модель системи безпечної маршрутизації для ройової мережі БПЛА

Кожен дрон в мережі містить модуль управління довірою (верхній рівень), під'єднаний до локального блокчейн-модуля (рівень PoT) і до модуля маршрутизації (нижній рівень). Шар довіри збирає інформацію про поведінку сусідніх БПЛА (успішність передачі пакетів, затримки, відмови тощо) і розраховує метрики довіри. Ці дані можуть передаватися в блокчейн-рівень, що діє як розподілений реєстр авторитетності вузлів. Блокчейн за допомогою консенсусу PoT узгоджує між дронами спільне бачення надійних і зловмисних учасників та зберігає криптографічні "сліди" (гіші) їх поведінки. Маршрутизаційне ядро виконує пошук маршрутів у децентралізованій мережі (на основі протоколу типу AODV), але, зважаючи на додаткову інформацію, воно отримує від шару довіри оцінки надійності потенційних наступних вузлів маршруту й перевіряє за блокчейном автентичність і відсутність компрометації цих вузлів. Теоретичні основи відмовостійкої маршрутизації, розроблені Євдокименком [13], підтверджують важливість динамічної адаптації до змін топології. Як наслідок, обираються тільки безпечні проміжні вузли: дрони з достатнім рівнем довіри та присутні в реєстрі блокчейну як легітимні учасники (не вилучені за зловмисність).

1. Шар довіри (Trust Layer)

Шар довіри відповідає за оцінювання надійності кожного вузла та каналу зв'язку на основі досвіду їх поведінки. Кожен БПЛА локально веде таблицю довіри, оновлюючи її після обміну повідомленнями із сусідами. Основний принцип: якщо сусідній вузол успішно доставляє пакети й підтримує стабільний зв'язок, рівень довіри до нього зростає; якщо ж пакети губляться без причин або вузол відмовляється ретранслювати дані, рівень довіри падає. Такий підхід дає змогу виявляти зловмисних чи ненадійних учасників мережі на основі їх поведінкових властивостей.

Для кількісного оцінювання використовується метрика, аналогічна феромону в алгоритмах мурашиної колонії. Після успішної доставки пакета через сусіда n вузол-джерело підсилює мітку довіри (феромон) для цього сусіда на величину, обернено пропорційну затримці доставки або кількості стрибків. Формально приріст довіри $\Delta P_{n,t}$ обчислюється за формулою

$$\Delta P_{n,t} = 1/\ln \ell_{n,t} \quad (1)$$

де $\ell_{n,t}$ – затримка доставки пакета через вузол n у момент часу t .

Загальне оновлення феромонної мітки довіри до вузла n на момент часу $t+1$ відбувається відповідно до рівняння

$$P_{n,t+1} = (1-\rho)P_{n,t} + \Delta P_{n,t}, \quad (2)$$

де $\rho \in (0,1)$ – коефіцієнт згасання, що імітує "випаровування" довіри з часом; $P_{n,t}$ – поточне значення довіри до вузла n у момент часу t ; $\Delta P_{n,t}$ – приріст довіри на основі нової доставки.

Якщо ж передача через n (сусіда) не вдалася або виявлено факт підміни / пошкодження пакета, рівень довіри до нього зменшується. Ця метрика виконує роль локального рейтингу надійності та з часом згасає без підтвердження, зважаючи на динаміку мережі.

На основі метрик довіри кожен дрон самостійно вирішує, кого вважати підозрілим. Якщо сусід накопичує негативну історію (високу частку невдалих передач), модуль довіри може позначити його як потенційно зловмисний та повідомити про це іншим рівням. Зокрема, коли рівень довіри падає нижче певного порогу, генерується подія про зловмисний вузол, що передається в блокчейн-рівень для колективного внесення цього вузла до "чорного списку" (списку U_m – множини розпізнаних зловмисних БПЛА).

$$T_n < T_{\text{threshold}} \Rightarrow u_n \in U_m \quad (3)$$

де T_n – поточний рівень довіри до вузла u_n ; $T_{\text{threshold}}$ – встановлений пороговий рівень довіри; U_m – множина вузлів, класифікованих як зловмисні.

Отже, шар довіри є першим рубежем безпеки: він локально фільтрує сумнівних сусідів та ініціює глобальне поширення інформації про них через блокчейн.

Шар довіри тісно інтегрується з блокчейном і маршрутизатором. З одного боку, він постачає блокчейн-рівню інформацію про надійність, наприклад повідомлення про виявленого зловмисника або періодичні узагальнені показники довіри (на рис. 3 цей напрям позначено жовтою стрілкою "Метрики довіри та подій"). З іншого боку, сам отримує від блокчейну підтверджену глобальну інформацію про довіру: наприклад, рішення більшості про вимкнення певного вузла з мережі або про присвоєння йому статусу надійного після успішної автентифікації (на рис. 3 стрілка "Підтверджені дані довіри"). Крім того, маршрутизатор запитує у шару довіри рейтинги сусідів у процесі вибору наступного хопу (стрілка "Оцінки довіри для маршрутів"), а після

побудови маршруту маршрутизатор повертає до шару довіри інформацію про успішність чи неуспішність доставки *Performance feedback*. У такий спосіб формується зворотний зв'язок: маршрутизаційне ядро повідомляє, які вузли виправдали довіру або підвели, і *Trust Layer* коригує метрики відповідно.

Шар довіри виконує роль системи локальної репутації для сусідів, що слугує основою для прийняття рішень і на рівні блокчейну (глобальна репутація), і на рівні маршрутизації (вибір безпечних маршрутів). Такий підхід відповідає рекомендаціям сучасних досліджень, де зазначено, що відсутність механізмів довіри в *ad-hoc*-мережах призводить до ненадійної роботи, і пропонується відстежувати поведінку вузлів для оцінювання їх доброчесності. Збереження історії поведінки в незмінному вигляді (наприклад, на блокчейні) дає змогу виявляти зловмисників і запобігати їх повторному приєднанню. Саме це й забезпечує наступний рівень архітектури – блокчейн з консенсусом PoT.

2. PoT-блокчейн

Блокчейн-рівень інтегрує розподілений реєстр (блокчейн) у ройову мережу БПЛА, щоб глобально узгоджувати та зберігати критичні відомості про мережу: ідентифікатори дронів, їх статус автентифікації, сертифікати, а також довірчі показники (наприклад, записи про зловмисну поведінку). Кожен вузол мережі розглядається як учасник блокчейну (в уявленні авторів – як "контейнер блоку" з власними даними). Блокчейн функціонує в консорціумному режимі (без центрального сервера, але для фіксованої групи вузлів-дронів), де всі чесні вузли спільно підтримують цілісність ланцюжка блоків.

Нові записи (транзакції) можуть стосуватися приєднання або виходу дрона з ройової мережі, зміни його статусу довіри, оновлення групового ключа тощо. Основна вимога – алгоритм консенсусу має бути легким (не вимагати великих обчислень чи частих обмінів) і адаптованим до динаміки топології БПЛА.

Запропонований алгоритм досягнення згоди – PoT – це інноваційний підхід, що використовує показники мережевого трафіку як доказ, замість традиційних доказів роботи чи частки. Ідея полягає в тому, що "авторитет" вузла визначається інтенсивністю його участі в маршрутизації. Показник трафіку, позначений як феромон (від *Trust Layer*), відбиває активність і надійність дрона в передачі інформації. Вузли з вищим рівнем феромону (тобто ті,

які часто успішно передають пакети) отримують більшу вагу внаслідок консенсусу.

Формально оновлення феромонного рівня виконується за співвідношенням

$$P_n \leftarrow (1 - \rho) P_n + \Delta P_n, \\ \Delta P_n = \frac{1}{m} \sum_{k=1}^m \ln(1/l_k) \quad (4)$$

де m – кількість переданих пакетів через вузол за певний інтервал часу u_i ; P_n – поточний рівень довіри до вузла n ; $\rho \in (0,1)$ – коефіцієнт згасання, що моделює "випаровування" феромону (зменшення довіри з часом); ΔP_n – зміна рівня довіри на основі спостережених метрик трафіку; l_k – затримка доставки k -го пакета (у секундах).

Отже, із зростанням затримки пакетів (l_k) відповідний складник $1/l_k$ зменшується, що призводить до зниження рівня довіри до вузла.

Результатом формули є оновлений показник довіри до кожного вузла, що динамічно змінюється відповідно до його поведінки в мережі. Якщо довіра падає нижче порогового значення $T_{threshold}$, тоді такий вузол вважається зловмисним і додається до множини вилучених:

$$T_i < T_{threshold} \Rightarrow u_n \in U, \quad (5)$$

де T_i – рівень довіри до i -го вузла; $T_{threshold}$ – граничне значення довіри; U – множина вилучених вузлів.

Це дає змогу колективно обирати надійного лідера для генерування чергового блоку або погоджувати наступний блок. На відміну від PoW, де всі вузли змагаються в розв'язанні складних криптозавдань, PoT не потребує додаткових обчислень – достатньо метрики трафіку, яку мережа вже природно збирає.

Кожен вузол u_n зберігає

- свій ідентифікатор I_n ;
- дайджест H_n геш-блоку, що містить інформацію про вузол;
- глобальний геш I_N об'єднання всіх легітимних I_n , які є елементами множини I_N :

$$H_N = \{H_i | u_i \in I_N\}, \quad (6)$$

де I_n – свій ідентифікатор.

Блокчейн містить блоки-автентифікації вузлів. Перед зльотом дрона в його пам'ять записується

актуальний знімок блокчейну H_N , що дає змогу кожному дрону швидко перевіряти, чи має потенційний сусід запис у блокчейні та не позначений як зловмисник. У разі спроби нового вузла приєднатися генерується блок з його ідентифікатором і ключем, який додається до ланцюжка після погодження.

Якщо вузол вважається скомпрометованим, інші учасники за допомогою консенсусу вилучають його з H_N , і оновлений геш ланцюжка стає відомим усім. Блоки також можуть містити дайджести маршрутів чи подій, що впливають на безпеку, наприклад виявлення атаки або зміна ключа шифрування.

Однією з ключових інновацій є пасивна синхронізація блоків. Дрони передають блоки опортуністично, долучаючи їх фрагменти до службових повідомлень (RREQ, ADS-B тощо), або лише за запитом сусіда. Це значно знижує трафік синхронізації, що критично важливо для масштабування ройової мережі.

Отже, RoT-блокчейн виконує роль розподіленого центру автентифікації та репутації без центрального вузла. Завдяки легкому консенсусу на основі трафіку та відсутності надлишкової сигналізації система ефективно функціонує, навіть у середовищі з обмеженими обчислювальними ресурсами.

Зокрема кількість повідомлень на побудову консенсусу й розповсюдження блоків зростає повільніше, ніж у разі інших алгоритмів, що критично для масштабування ройової мережі.

Блокчейн-рівень слугує проміжною ланкою між шаром довіри та маршрутизацією. Він приймає від модуля довіри транзакції про зміну стану довіри вузлів (події) і після досягнення консенсусу зберігає цю інформацію. У такий спосіб суб'єктивні оцінки довіри з окремих дронів консоліднуються в глобально довірену інформацію. Маршрутизаційне ядро зі свого боку звертається до блокчейну для перевірки легітимності кандидатів на маршрут: у процесі RREQ/RREP під час вибору наступного хопу кожен дрон перевіряє у своєму локальному ланцюжку H_N , чи присутній кандидат в переліку зареєстрованих і не заблокованих вузлів. Якщо запису про сусіда немає або він позначений як зловмисник, такий сусід ігнорується в процесі маршрутизації. Отримання актуальної копії H_N (основного геша блокчейну) перед стартом або через пасивну синхронізацію гарантує, що всі чесні дрони мають узгоджений

список довірених учасників. Блокчейн також може зберігати спільний груповий ключ шифрування трафіку для захисту від сторонніх, як пропонується в інших рішеннях, і оновлювати його для нових учасників за допомогою транзакцій.

Отже, RoT-блокчейн виконує роль розподіленого центру автентифікації та репутації в мережі, але без виділеного вузла: усі дрони колективно підтримують цю функцію. Завдяки легкому консенсусу на основі трафіку й відсутності зайвої сигналізації ця схема є придатною для БПЛА з обмеженими ресурсами. Далі розглянемо, як маршрутизатор використовує інформацію довіри та блокчейну.

3. Маршрутизаційне ядро (Routing Core)

Основи маршрутизації в рої БПЛА: ройова мережа БПЛА зазвичай будується за принципами *ad-hoc*-мережі, де дрони взаємодіють напряму або за допомогою мультихоп-з'єднання. Наше рішення ґрунтується на вдосконаленому протоколі типу AODV, адаптованому під особливості 5G та *trust/blockchain*-інфраструктури. Маршрутизація запускається на вимогу: коли один дрон хоче відправити інформацію іншому, він розсилає запит маршруту RREQ. Завдяки можливостям 5G NR припускається директивна передача RREQ у напрямку на призначення (*drone-source* знає приблизне положення *drone-destination* через ADS-B координати). Це зменшує широкомовлення – запит спрямовується лише на тих сусідів, які лежать у секторі напрямку до цільового вузла.

$$U_{source,hops} = \{u_n \mid \angle(v_{s,d}, v_{s,n}) > \alpha\}, \quad (7)$$

де u_n – сусідній вузол (кандидат у наступники); $v_{s,d}$ – вектор напрямку від джерела s до пункту призначення d ; $v_{s,n}$ – вектор від джерела s до сусіда n ; α – порогове значення кута відхилення (у градусах або радіанах), що обмежує допустиме відхилення напрямку; $U_{source,hops}$ – підмножина сусідів, які лежать у секторі напрямку до цільового вузла та прийнятні за кутом маршрутизації.

Ця фільтрація дає змогу уникати маршрутизації через вузли, які географічно віддалені від цільового напрямку (більш ніж на α), забезпечуючи значну ефективність та передбачуваність траєкторії.

Кожен проміжний дрон, отримавши RREQ, додає себе до маршруту й пересилає далі (так формується шлях), доки RREQ не досягне призначення, яке потім відправить відповідь RREP назад до початкового

дрона по знайденому шляху. Ключова відмінність маршрутизатора – критерій вибору наступного хопу. Стандартний AODV обирає маршрут за мінімальною кількістю стрибків, не беручи до уваги надійність вузлів.

Далі виконується фільтрація з огляду на довіру:

$$U_{n,hops} = \{u_n \mid u_n \in U_{cycloid} \wedge u_n \notin I_t\}, \quad (8)$$

де I_t – множина зловмисних вузлів; u_n – вузол-кандидат у наступники; $U_{cycloid}$ – множина всіх наявних сусідів вузла u_t ; I_t – множина вузлів, визнаних зловмисними на момент часу t ; $U_{n,hops}$ – підмножина сусідів, прийнятних для маршрутизації.

Отже, обмеження (7) й (8) разом формують геометрично-фільтраційну модель вибору наступного вузла, зважаючи як на просторові, так і поведінкові (довірчі) властивості сусідів.

Algorithm 1. NEXT_HOP_SELECTION – Вибір наступного безпечного хопу

Input: $U_{cycloid}$; $v_{s,d}$; H_N ; I_t ; α .

Output:

Обраний наступний вузол u_{next} .

1. 1. $U_hops \leftarrow \emptyset$
2. 2. for each $u_n \in U_{cycloid}$ do
3. 2.1. Обчислити кут $\angle(v_{s,d}, v_{s,n})$;
4. 2.2. if $\angle(v_{s,d}, v_{s,n}) \leq \alpha$ and $u_n \in H_N$ and $u_n \notin I_t$ then
5. 2.2.1. Додати u_n до U_hops ;
6. end if
7. end for
8. 3. if $U_hops = \emptyset$ then
9. 3.1. Завершити процедуру без вибору вузла;
10. else
11. 3.2. Обрати $u_{next} \in U_hops$ з мінімальною кількістю хопів або найбільшим рівнем довіри;
12. end if
13. 4. Повернути u_{next} .

Якщо сусід x не має відповідного запису (або є в списку U_m зловмисників), u_n не розглядає його як наступний шлях маршрутизації. Це гарантує, що в маршрут не потраплять невідомі системі або заблоковані вузли. Друга умова полягає в тому, що з-поміж автентичних сусідів обирається той, який має найвищу метрику довіри P_{hn} (найбільший накопичений феромон).

$$P_{hn} \leftarrow (1 - \rho)P_{hn} + \Delta P_{hn}, \quad (9)$$

де P_{hn} – актуальне значення довіри (феромонної метрики) до сусіднього вузла n , що накопичується

дроном h на основі історії успішної маршрутизації через цей вузол; ΔP_{hn} – приріст довіри на основі нових успішних передач пакетів через вузол n , наприклад, після маршрутування RREQ/RREP або доставлення даних; $(1 - \rho)P_{hn}$ – частка поточної довіри, яка зберігається після згасання.

Перевага надається тому вузлу, який історично демонстрував кращу якість маршрутизації (низькі затримки, відсутність втрат). Цей механізм фактично будує маршрут по "найбільш довірених" сусідах від джерела до призначення. Для ініціалізації, коли історії ще немає, можуть використовуватися базові значення довіри або вибір за метричними показниками каналу (рівень сигналу, відстань). З часом же маршрути стають більш оптимізованими за безпечністю завдяки навчанню метрик довіри. Захист від атак під час маршрутизації: комбінування довіри та блокчейну в маршрутизаторі забезпечує кілька рівнів захисту. По-перше, запобігання проникненню зловмисників: якщо дрон, який атакує, не зміг пройти автентифікацію та додати себе до блокчейну, інші не посилатимуть через нього пакети (він ігнорується на етапі RREQ). Це узгоджується з принципом, запропонованим в роботі [8], де неавторизовані дрони не допускаються до участі в мережі. По-друге, ізоляція викритих зловмисників: якщо вузол починає неправомірно діяти (наприклад, відмовляє в ретрансляції), то після декількох таких випадків його сусіди знизять до нього довіру та ініціюють блокування в блокчейні. Як тільки по блокчейну зафіксовано, що вузол перейшов до U_m (*malicious*), жоден маршрут більше не буде включати його. Це підвищує стійкість маршрутизації: навіть якщо зловмисник змінив своє місце, він не зможе знову вклинитися в трафік. По-третє, захист від перехоплення і підслуховування: завдяки використанню *5G beamforming* RREQ і наступний трафік передаються вузькоспрямовано, тому сторонній дрон поза маршрутом не може легко перехопити сигнал. Крім того, відомості про присутність зловмисника поблизу можуть бути сигналізовані за допомогою блокчейну, і маршрутизатори коригують траєкторії маршруту, обходячи небезпечний район.

Припустимо, дрон A хоче надіслати інформацію дрону D через ройову мережу. Дрон A знає координати D (наприклад, унаслідок періодичних широкомовних ADS-B повідомлення від D) [15]. A формує RREQ і спрямовано передає його сусідам у бік D .

Нехай сусіди A – дрони B і X . A перевіряє їх по блокчейну: припустимо, обидва зареєстровані (тобто є в списку H_N , де зберігаються дайджести легітимних вузлів). Далі A перевіряє їх метрики довіри:

$$P_B = 0,9, \quad P_X = 0,4 \quad (11)$$

(тобто B вважається надійнішим за X відповідно до накопиченого феромону).

Отже, A обирає вузол B , оскільки

$$\arg \max_{u \in U_{\text{hops}}} P_u = B, \quad (12)$$

і передає через нього RREQ.

Дрон B отримує RREQ, додає свій ідентифікатор і передає далі сусідам, також керуючись своїм списком довіри й перевіркою автентичності через локальну копію H_N . Якщо на якомусь кроці сусід u не має запису в H_N або належить до множини зловмисників U_m , тоді

$$u \notin H_N \text{ або } u \in U_m \Rightarrow u \quad (13)$$

не обирається як наступний.

У підсумку RREQ досягає вузол D по шляху, що містить лише довірені вузли. D відправляє відповідь RREP назад тим самим маршрутом, підтверджуючи встановлення маршруту. Після завершення передачі даних дрон A та всі проміжні вузли оновлюють метрики довіри для своїх сусідів. Для кожного вузла n , через який успішно пройшов трафік, застосовується формула (2) оновлення феромонної довіри.

Це підвищує шанси використання цього маршруту в майбутньому, оскільки система запам'ятала його ефективність.

Якщо трапиться відхилення або атака, наприклад дрон X , отримавши RREQ, не пересилає його далі, тоді A не отримає RREP у відповідь. У цьому разі A зменшує довіру до X :

$$P_X \leftarrow (1 - \rho) P_X, \quad (14)$$

а також формує транзакцію в блокчейн із позначкою про підозрілу поведінку вузла X .

Інші вузли, отримавши оновлення про зниження рейтингу X через пасивну синхронізацію, також уникатимуть його у своїх маршрутах.

Отже, маршрутизаційне ядро у взаємодії з *Trust Layer* адаптивно навчається будувати все більш безпечні маршрути, зокрема проблемні ланки.

Порівняльний аналіз

алгоритму PoT та інших консенсусів

В основі розглянутої архітектури лежить консенсус PoT, покликаний узгоджувати спільний

блокчейн ройової мережі БПЛА. Важливо оцінити, чим PoT вигідно відрізняється від класичних алгоритмів консенсусу, таких як *Proof-of-Work* (PoW), *Proof-of-Stake* (PoS) та класичних алгоритмів візантійської відмовостійкості (BFT – *Byzantine Fault Tolerance*, наприклад, PBFT). На відміну від PROACT [11], де використовується паралельна генерація блоків, запропонований у дослідженні підхід орієнтований на мінімізацію затримок. Порівняння проведемо за ключовими критеріями: обчислювальна складність, енергоефективність та придатність до використання в FANET. Результати зведено в табл. 2.

Як видно з табл. 2, PoT перевершує PoW та PoS у ресурсно обмеженому середовищі дронів. Відповідно до результатів моделювання, наведених у роботі [2], PoT генерує значно менше службового трафіку та витрат енергії порівняно з PoW/PoS, особливо зі зростанням кількості вузлів. Звичайний PoW для дронів абсолютно непрактичний – він вимагав би від кожного БПЛА майнити блок, що є неприйнятним щодо і енергії, і затримок (наприклад, ІТУ відзначає необхідність відмови від PoW для Інтернету речей на користь легших підходів). PoS є легшим, проте в умовах відсутності єдиної валюти чи ставки в рої його застосування не інтуїтивне – в літературі запропоновано хіба що реалізації з довірою як "ставкою" для VANET, але це все одно не бере до уваги динамічний трафік.

BFT-алгоритми, як-от PBFT, використано в деяких роботах для дронів, але їх масштабованість обмежена: класичний PBFT працює лише для десятків вузлів через квадратичне зростання обмінів повідомленнями. Нещодавні вдосконалення BFT (наприклад, ієрархічні схеми) можуть розширити цей діапазон, проте вони все ще передбачають фіксований пул валідаторів і не розраховані на часті зміни складу.

Зазначимо, що PBFT є лише одним із представників класу алгоритмів BFT (*Byzantine Fault Tolerance*). Існують також інші варіанти, зокрема SBFT (*Scalable BFT*), *HotStuff*, *Tendermint* та ієрархічні BFT-схеми, які мають на меті покращити масштабованість та ефективність у динамічних умовах. Однак більшість з них все ще вимагають високого рівня взаємодії між вузлами або використовують фіксований пул валідаторів, що обмежує їх придатність для мобільних і ресурсно обмежених середовищ, зокрема FANET.

Таблиця 2. Порівняльний аналіз алгоритму PoT та інших консенсусів

Алгоритм консенсусу	Обчислювальна складність	Енергоефективність	Придатність для FANET
Proof-of-Traffic (PoT)	Низька – не потребує складних обчислень (використовує вже наявні дані трафіку); масштабується приблизно $O(n)$ за кількіст. учасників (тільки легкі перевірки).	Висока – мінімальні додаткові витрати енергії, відсутні ресурсомісткі завдання; економія на сигналізації внаслідок пасивної синхронізації блоків.	Дуже висока – розроблений спеціально для ройових <i>ad-hoc</i> -мереж; стійкий до динамічних змін, не потребує від дронів нічого, крім звичайної маршрутизації. Забезпечує ефективну роботу в умовах обмежених ресурсів.
Proof-of-Work (PoW)	Дуже висока – потребує розв'язання криптографічних головоломок (експоненційно складних); не масштабується (значна частина ресурсів витрачається марно).	Низька – надзвичайно енерговитратний (дрони не мають ресурсів для майнінгу, швидке розрядження батарей); неприйнятний для мобільних вузлів.	Вкрай низька – практично непридатний для FANET: висока затримка на знаходження блока, вимога потужних процесорів, відсутність прив'язки до мережевої активності. Не забезпечує швидкого консенсусу в умовах швидкої зміни топології.
Proof-of-Stake (PoS)	Невисока – обчислення відносно прості (перевірки підписів, вибір валідатора за часткою); добре масштабується до сотень вузлів.	Висока – не потребує майнінгу, енергоспоживання здебільшого на комунікації; для дронів значно кращий, ніж PoW.	Середня – може бути використаний в приватному рої, якщо кожному дрону призначити "частку" (наприклад, ваговий коефіцієнт). Проте PoS не бере до уваги реальну активність у мережі, а також вимагає відносно стабільного складу учасників. У разі частозмінюваній топології визначення часток та лідера може ускладнюватися.
BFT (PBFT і подібні)	Помірна – алгоритми BFT потребують багато повідомлень між всіма парами вузлів (комунікаційна складність $O(n^2)$; за умови великого n обчислювальне та комунікаційне навантаження швидко зростає.	Середня / низька – відсутні важкі обчислення, але численні повідомлення узгодження навантажують мережу (радіообмін витрачає енергію). Для кількох десятків дронів прийнятно, але для сотень ні.	Обмежена – підходить лише для малих стабільних груп дронів (до ~10–20 вузлів). У великих ройових мережах класичний BFT не масштабується (занадто багато трафіку узгодження). До того ж за динамічного входу / виходу вузлів необхідні додаткові протоколи перегрупування консенсусу.

Натомість PoT природно вписується в парадигму ройової мережі. Він використовує ту саму інформацію, що й протокол маршрутизації (стан мережевого трафіку), тож не створює зайвого навантаження. Завдяки пасивній синхронізації додатковий трафік консенсусу практично відсутній. Як показали експерименти, PoT забезпечує швидке узгодження: затримки на оновлення блокчейну мінімальні, оскільки оновлення передаються разом з наявним трафіком. Це критично для безпечної маршрутизації – інформація про нового зловмисника або нового учасника швидко стає відомою всьому рою. Узагальнюючи, PoT досягає балансу між безпекою та ефективністю, який потрібен у ройових мережах БПЛА, що підтверджено кількісними показниками продуктивності у відповідних наукових публікаціях

Розглянута архітектура поєднує довірчі обчислення та блокчейн для забезпечення безпечної маршрутизації в ройовій мережі безпілотників. Трирівнева модель – *Trust Layer*, PoT-блокчейн, *Routing Core* – дає змогу досягти стійкості до атак і ефективності одночасно. Шар довіри локально

відстежує поведінку сусідів, блокчейн глобально фіксує та узгоджує інформацію про довіру, а маршрутизатор використовує ці дані для побудови оптимальних маршрутів, уникаючи зловмисників. Завдяки впровадженню легкого консенсусу PoT система позбавлена основних недоліків PoW/PoS: не витрачає зайві ресурси на майнінг чи підтримання валідаторів, а натомість використовує властивості самої мережі (трафік) для узгодження стану. Наведений порівняльний аналіз підтверджує переваги PoT в контексті FANET: він найбільш енергоефективний та масштабований для роїв дронів. Сформована таким чином архітектура спирається на сучасні наукові розробки в галузі *trust*-менеджменту та блокчейн-систем для *ad-hoc*-мереж, що підтверджують автори публікацій IEEE та інших джерел.

Подальші дослідження можуть бути присвячені детальному прототипуванню цієї системи, аналізу її стійкості до різноманітних атак і сумісності з наявними стандартами зв'язку (наприклад, інтеграція з 5G MEC або використання стандартних

ідентифікаторів UAV з ASTM/ETSI для автентифікації). Підходи до забезпечення надійності рою БПЛА в умовах руйнувань, які описав Д. Тереник та інші [18], наголошують на важливості резервування комунікаційних каналів. Проте вже зараз запропонована модель демонструє життєздатність підходу "довіра + блокчейн" для забезпечення ефективної та надійної роботи ройових мереж безпілотників у реальних умовах.

Результати досліджень та обговорення

У межах дослідження реалізовано симуляційне середовище у NS-3.36, *CoppeliaSim* для моделювання рою БПЛА з використанням мобільної *ad-hoc*-мережі (FANET) на основі протоколу маршрутизації AODV (рис. 5). Було змодельовано атаку типу *Black Hole*, за

якої один із вузлів мережі (ідентифікатор 3) перехоплює пакети, але не пересилає їх далі й цим порушує цілісність маршрутизації.

На першому етапі експерименту сформовано початкову топологію рою БПЛА (рис. 4), що передбачало стартове розміщення дронів на площині з різноманітним рельєфом місцевості.

На рис. 5 зображено початкову конфігурацію рою безпілотних літальних апаратів у моделювальному середовищі *CoppeliaSim*. Дрони розташовані на заданій ділянці, що містить як рівні, так і горбисті ділянки рельєфу, що дає змогу симулювати умови складної топології місцевості. БПЛА рівномірно розміщені на стартових позиціях перед початком експериментальної сесії з аналізу поведінки мережі в присутності атак типу *Black Hole* та реалізації механізмів захищеної маршрутизації.

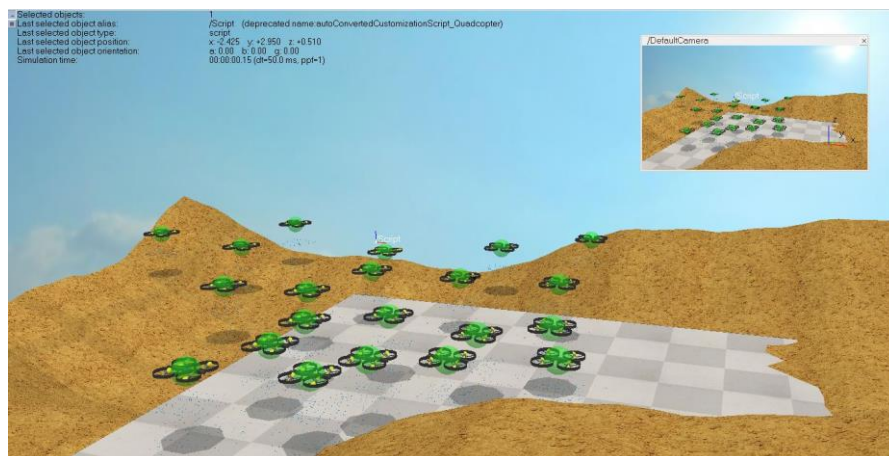


Рис. 4. Початкове розгортання роєвої мережі БПЛА в симуляційному середовищі *CoppeliaSim*

```
Black Hole Attack: Packet dropped by malicious node!
Black Hole Attack: Packet dropped by malicious node!
Black Hole Attack: Packet dropped by malicious node!

Trust-механізм: кількість переданих пакетів:
Вузол 0 передав 100 пакетів
Вузол 1 передав 100 пакетів
Вузол 2 передав 100 пакетів
Вузол 3 передав 100 пакетів
Вузол 4 передав 100 пакетів
Вузол 5 передав 100 пакетів
Вузол 6 передав 100 пакетів
Вузол 7 передав 100 пакетів
Вузол 8 передав 100 пакетів
Вузол 9 передав 100 пакетів
Вузол 10 передав 100 пакетів
Вузол 11 передав 100 пакетів
Вузол 12 передав 100 пакетів
Вузол 13 передав 100 пакетів
Вузол 14 передав 100 пакетів
Вузол 15 передав 100 пакетів
Вузол 16 передав 100 пакетів
Вузол 17 передав 100 пакетів
Вузол 18 передав 100 пакетів
Вузол 19 передав 100 пакетів
ns@ns-VirtualBox:~/ns-3-allinone/ns-3.36$ nano scratch/fanet-simulation.cc
ns@ns-VirtualBox:~/ns-3-allinone/ns-3.36$
```

Рис. 5. Результати симуляції атаки типу *Black Hole* у FANET: виведення повідомлень про втрату пакетів і кількість переданих пакетів кожним вузлом

Ця сцена використовується для подальшого моделювання сценаріїв обміну даними між дронами з огляду на динамічне зміщення та можливу появу шкідливих вузлів. Симуляція допомагає оцінити, як запропонована архітектура захисту впливає на стабільність маршрутів, затримки та загальну ефективність роботи рою.

Для виявлення аномальної поведінки та формування оцінки довіри використано механізм підрахунку кількості переданих пакетів кожним вузлом. Цей підхід дав змогу кількісно визначити відхилення в поведінці окремих учасників мережі.

За результатами моделювання було зібрано статистику у форматі CSV, що містить ідентифікатори вузлів і відповідну кількість переданих пакетів. Проведено візуалізацію даних із застосуванням *Python*, що уможливило ідентифікацію вузла-зловмисника – він мав нульову або аномально незначну кількість переданих пакетів порівняно з іншими.

На рис. 6 зображено гістограму активності вузлів. Вузол № 3, що виконував роль атакуючого, виділено червоним кольором. Графічний аналіз підтвердив ефективність використаного *trust*-механізму для виявлення зловмисної поведінки в умовах обмеженого ресурсу FANET-мереж.

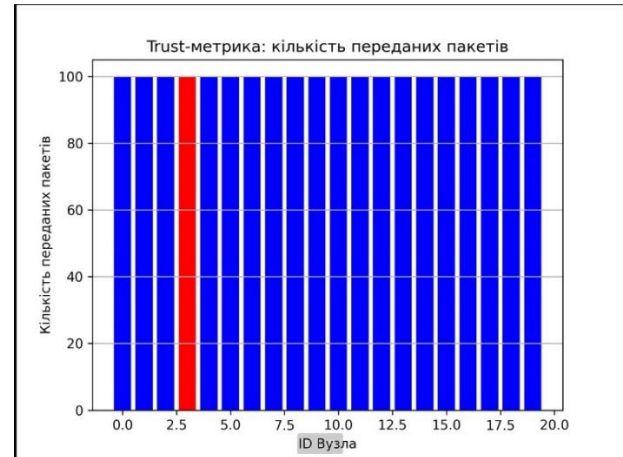


Рис. 6. Графічне подання активності вузлів у FANET: кількість переданих пакетів з виділенням зловмисного вузла

Для оцінювання ефективності запропонованого методу було порівняно ключові метрики – PDR (коефіцієнт доставки пакетів), затримки доставки (*Delay*), частоти виявлення зловмисників (*Detection Rate*) та енергоспоживання (*Energy Consumption*) – у різних сценаріях: без атак, з атакою типу *Black Hole*, та з активованим механізмом довіри та блокчейну (рис. 7).

```
GNU nano 4.8                                fanet-simulation.cc
NetDeviceContainer devices = wifi.Install(phy, mac, nodes);

// Інсталюємо стек 4. ANDV
InternetStackHelper stack;
AodvHelper aodv;
stack.SetRoutingHelper(aodv);
stack.Install(nodes);

Ipv4AddressHelper address;
address.SetBase("10.0.0.0", "255.255.255.0");
Ipv4InterfaceContainer interfaces = address.Assign(devices);

// Атака типу Black Hole та Trust-механізм безпеки
Simulator::Stop(Seconds(100.0));
Simulator::Run();
Simulator::Destroy();
return 0;

// Симуляція Black Hole атакує зловмисний вузол на мережі/на мережі
uint32_t attackerNodeId = 3; // наприклад, вузол №3 = атакуючий
Ptr<Node> attackerNode = nodes.Get(attackerNodeId);
Ptr<Ipv4> ipv4 = attackerNode->GetObject<Ipv4>();
ipv4->SetAttribute("ReceiveCallback", Callback<bool, Ptr<NetDevice>, Ptr<Internet Packet>, Ipv4::Address &&{
    [(Ptr<NetDevice> dev, Ptr<Internet Packet> pkt, uint32_t proto, const Address &from) -> bool {
        // Зловмисник просто приймає, але нічого не передає далі
        std::cout << "Black Hole Attack: Packet dropped by malicious node!" << std::endl;
        return true; // "Black" name
    }
});

// Простий Trust-механізм: перевіряємо, скільки пакетів передано кожному вузлу
std::map<uint32_t, uint32_t> packetForwardCount;

Config::ConnectWithoutContext(
    "/NodeList/*/$s3::Ipv4L3Protocol/Ix",
    MakeCallback([&packetForwardCount](Ptr<Internet Packet> p) {
        // Простий облік переданих пакетів (додаток = кількість переданих)
        uint32_t nodeId = Simulator::GetContext() / 4;
        packetForwardCount[nodeId]++;
    })
);

File Name to Write: fanet-simulation.cc
Get Help.      DOS Format      Append      Backup File
Cancel         Mac Format      Prepend     To Files
```

Рис. 7. Сценарій атаки типу *Black Hole* в ройовій мережі БПЛА

У зловмисному середовищі PDR суттєво знижується, оскільки вузли-порушники перехоплюють і

не пересилають пакети. Це особливо помітно в умовах відсутності будь-якого механізму захисту.

Застосування довірчої маршрутизації дає змогу вилучити підозрілі вузли, а блокчейну – узагальнити ці підозри по всій мережі, завдяки чому PDR зростає до майже базових значень.

Легкий блокчейн (наприклад, із використанням структур типу DAG) створює додаткову затримку, пов'язану з обробленням транзакцій та верифікацією статусу вузлів. Проте зростання загальної затримки не перевищує 8–10%, що є прийнятним для критичних FANET-застосунків із вимогами до безпеки.

Атака Black Hole

Вузол № 3 модифікує поведінку, відповідаючи на всі запити RREQ, але блокує подальшу передачу пакетів.

Реалізація у NS-3

```
Ptr<Node> attackerNode = nodes.Get(3);
Ptr<NetDevice> dev = attackerNode->GetDevice(0);
dev->SetReceiveCallback(MakeCallback(&BlackHoleDrop));
```

Захист Trust + Blockchain

Усі вузли зберігають локальні "рівні довіри" своїх сусідів.

Після досягнення порогового значення недовіри вузол додається до списку *It* (підозрілих).

Легка блокчейн-підсистема реєструє зміни в довірчих значеннях для забезпечення незмінності історії.

Фрагмент підрахунку пакетів для Trust-механізму

```
Config::ConnectWithoutContext(
    "/NodeList/*/Ns3::Ipv4L3Protocol/Tx",
    MakeCallback(&PacketTrace)
);
```

На графіках (рис. 8) зловмисний вузол (ідентифікатор 3) позначено червоним кольором. Видно, що його метрика довіри швидко знижується, і він вилучається з маршрутизації.

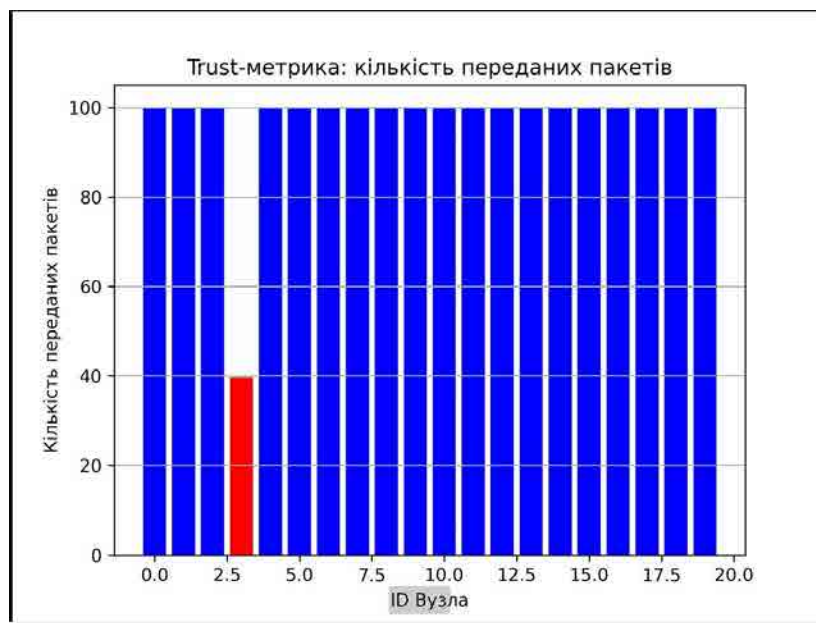


Рис. 8. Рівень передачі пакетів вузлами в сценарії атаки типу *Black Hole* (зловмисний вузол виділений червоним кольором)

Метрика довіри, основана на принципах мурашиних колоній [14], дає змогу ефективно оцінювати надійність вузлів. У табл. 3 наведено

порівняння ключових метрик (PDR, *Delay*, *Detection Rate*, *Energy*) для різних сценаріїв функціонування ройової мережі БПЛА.

Таблиця 3. Порівняльний аналіз метрик ефективності для різних сценаріїв ройової мережі БПЛА

Сценарій	PDR (%)	Delay (мс)	Detection Rate (%)	Energy (мДж)
Без атак	97.8	33	–	100
<i>Black Hole</i> (без захисту)	56.3	31	0	99
<i>Trust + Blockchain</i> (захист)	92.1	36	93.4	107

Застосування довірчої маршрутизації з блокчейн-підтримкою сприяє збереженню високого PDR і виявленню зловмисних вузлів з ефективністю понад 90%. Незначне збільшення затримки та енергоспоживання є прийнятною платою за підвищеною безпеку.

Отже, результати моделювання підтверджують доцільність використання довірчої моделі та обґрунтовують її долучення до легкого протоколу захищеної маршрутизації для роїв БПЛА, особливо в умовах кіберзагроз типу *Black Hole*. Запропонований підхід не потребує значних обчислювальних ресурсів, що є критичним для енергообмежених пристроїв, і може бути інтегрований з іншими механізмами, зокрема PoT або полегшеними варіантами блокчейну.

Висновки

У межах виконаного дослідження реалізовано симуляційне середовище з метою аналізу захищеності мобільних мереж типу FANET в умовах наявності внутрішніх загроз, зокрема атак типу *Black Hole*. Для моделювання використано середовище NS-3.36, протокол маршрутизації AODV та топологія з 20 безпілотними літальними апаратами, які переміщувалися відповідно до моделі мобільності *RandomWaypoint*.

Під час проведеного експерименту встановлено, що функціонування зловмисного вузла в межах мережі чинить суттєвий вплив на якість маршрутизації.

Такий вузол перехоплює мережевий трафік і блокує його подальшу передачу, що зумовлює порушення цілісності комунікаційних маршрутів. Для виявлення зловмисної активності застосовано вбудований механізм оцінювання довіри (*trust*-механізм), що реалізовано способом обліку кількості пакетів, переданих кожним вузлом. Аналіз мережевої активності дав змогу ідентифікувати зловмисний вузол як такий, що демонструє аномально низьку активність. Досягнуті результати було візуалізовано, що підтвердило ефективність запропонованого методу детекції.

Отже, результати моделювання засвідчили доцільність застосування простих поведінкових метрик, зокрема кількості ретрансляцій, для виявлення внутрішніх загроз у ройових мережах БПЛА. Запропонований підхід може слугувати основою для створення легковагових захисних механізмів, сумісних з обмеженими ресурсами FANET-середовищ. Як зазначено в роботах [16, 17], інтеграція елементів ройового інтелекту має перспективи для покращення систем уникнення зіткнень, особливо в умовах щільної урбанізованої інфраструктури.

Подальші дослідження буде присвячено на розроблення адаптивних *trust*-механізмів із застосуванням технологій самонавчання, а також на моделювання складніших сценаріїв атак, таких як *Collaborative Black Hole* та *Grey Hole*, з метою всебічного оцінювання стійкості протоколів маршрутизації в середовищах типу FANET.

Список літератури

1. Liu E., Effio E., Hitchcock J. Survey on health care applications in 5g networks. *IET Communications*, P. 1073–1080. 2020. DOI: doi.org/10.1049/iet-com.2019.0813
2. Wang J., Liu Y., Niu S., Song H. Lightweight blockchain-assisted secure routing of swarm UAS networking. *Computer Communications*. 2020. Vol. 157. P. 66–75. DOI: https://doi.org/10.1016/j.comcom.2020.09.035
3. Alshammari H., Niazi M. Mitigating Black Hole and Sybil Attacks in UAV Swarm Networks using Blockchain and Fuzzy Logic. *Sensors*. 2023. Vol. 23, Issue 3. Article 625. DOI: https://doi.org/10.3390/s23030625
4. Gupta R. et al. Blockchain-assisted secure UAV communication in 6G environment: Architecture, opportunities, and challenges. *IET communications*. 2021. Vol. 15. №. 10. P. 1352–1367. DOI: 10.1049/cmu2.12113
5. Rosati S., Kruzelecki K., Heitz G., Floreano D., Rimoldi B. Dynamic Routing for Flying Ad Hoc Networks. *IEEE Transactions on Vehicular Technology*. 2016. Vol. 65, Issue 3. P. 1690–1700. DOI: https://doi.org/10.1109/TVT.2015.2415417
6. Akkaya K., Guvenc I., Aygun R., Pala N., Kadri A. Routing in unmanned aerial ad hoc networks: A survey. *Ad Hoc Networks*. 2020. Vol. 92. Article 101778. DOI: https://doi.org/10.1016/j.adhoc.2019.101778
7. Wang W.; Lv M.; Ru L.; Lu B.; Hu S.; Chang X. (2022). Multi-UAV Unbalanced Targets Coordinated Dynamic Task Allocation in Phases. *Aerospace*, 9, 491 p. 2022. DOI: https://doi.org/10.3390/aerospace9090491

8. Krytskyi D., Karatanov O., Pohudina O., Shevel V., Bykov A., Pyvovar M., Plastun T. Information Technology for Determining the Flight Performance of a Paraglider Wing. *In Information Technologies in the Design of Aerospace Engineering* P. 1-42. Cham: Springer Nature Switzerland. 2023. DOI: https://doi.org/10.1007/978-3-031-43579-9_1
9. Zhou X. et al. Towards secure and resilient unmanned aerial vehicles swarm network based on blockchain. *IET Blockchain*. 2024. Vol. 4. P. 483-493. DOI: <https://doi.org/10.1049/blc2.12050>
10. Li Z. et al. A secure and efficient UAV network defense strategy: Convergence of blockchain and deep learning // *Computer Standards & Interfaces*. 2024. Vol. 90. 103844 p. DOI: <https://doi.org/10.1016/j.csi.2024.103844>
11. Mershad K. PROACT: Parallel multi-miner proof of accumulated trust protocol for Internet of Drones. *Vehicular Communications*. 2022. Vol. 36. 100495 p. DOI: <https://doi.org/10.1016/j.vehcom.2022.100495>
12. Hafeez S. et al. Beta-UAV: blockchain-based efficient authentication for secure UAV communication. *Cryptography and Security*. 2024. DOI: <https://doi.org/10.48550/arXiv.2402.15817>
13. Євдокименко М. О. Теоретичні основи відмовостійкої маршрутизації в телекомунікаційних мережах: дис. д-ра техн. наук: 05.12.02. М. О. Євдокименко. Харк. нац. ун-т радіоелектроніки, 2020.
14. Ясінчук В. І. Багатошляхова маршрутизація на основі алгоритмів мурашкових колоній 2020. URL: <https://dspace.wunu.edu.ua/bitstream/316497/1510/1/Yasinchuk%20V.I.%20C%20KSMzm-51.pdf>
15. Krytskyi, D., Karatanov, O., Pohudina, O., Shevel, V., Bykov, A., Pyvovar, M., Plastun, T. Information Technology for Determining the Flight Performance of a Paraglider Wing. *In Information Technologies in the Design of Aerospace Engineering*. Cham: Springer Nature Switzerland. P. 1-42. 2024. DOI: https://doi.org/10.1007/978-3-031-43579-9_1
16. Бінько І. В., Шевель В. В., Биков А. М., Крицький Д. М. Аналіз децентралізованої моделі управління дронів і розрахунок траєкторії перехоплення. *Сучасний стан наукових досліджень та технологій в промисловості*. 2024. № 2 (28). С. 33–47. DOI: <https://doi.org/10.30837/2522-9818.2024.2.033>
17. Єна М. Контроль міської мобільності БПЛА: ройовий інтелект і уникнення зіткнень. *Сучасний стан наукових досліджень і технологій в промисловості*. 2024. № 4. С. 210–218. DOI: [10.30837/2522-9818.2024.4.059](https://doi.org/10.30837/2522-9818.2024.4.059)
18. Тереник Д., Харченко В. Вибір стратегій розгортання і забезпечення надійності рою БПЛА для підтримки комунікацій в умовах руйнувань. *Сучасний стан наукових досліджень і технологій в промисловості*. 2024. № 3. С. 155–162. DOI: [10.30837/2522-9818.2024.3.091](https://doi.org/10.30837/2522-9818.2024.3.091)

References

1. Liu, E., Effio, E., Hitchcock, J. (2020), "Survey on health care applications in 5G networks", *IET Communications*, Vol. 14, P. 1073–1080. DOI: <https://doi.org/10.1049/iet-com.2019.0813>
2. Wang, J., Liu, Y., Niu, S., Song, H. (2020), "Lightweight blockchain-assisted secure routing of swarm UAS networking", *Computer Communications*, Vol. 157, P. 66–75. DOI: <https://doi.org/10.1016/j.comcom.2020.09.035>
3. Alshammari, H., Niazi, M. (2023), "Mitigating Black Hole and Sybil Attacks in UAV Swarm Networks using Blockchain and Fuzzy Logic", *Sensors*, Vol. 23, No. 3, Article 625. DOI: <https://doi.org/10.3390/s23030625>
4. Gupta R., Nair A., Tanwar S., Kumar N. (2021), Blockchain-assisted secure UAV communication in 6G environment: Architecture, opportunities, and challenges. *IET Communications*, Vol. 15. №. 10. P. 1352-1367. DOI: [10.1049/cmu2.12113](https://doi.org/10.1049/cmu2.12113)
5. Rosati, S., Kruzelecki, K., Heitz, G., Floreano, D., Rimoldi, B. (2016), "Dynamic Routing for Flying Ad Hoc Networks", *IEEE Transactions on Vehicular Technology*, Vol. 65, No. 3, P. 1690–1700. DOI: <https://doi.org/10.1109/TVT.2015.2415417>
6. Akkaya, K., Guvenc, I., Aygun, R., Pala, N., Kadri, A. (2020), "Routing in unmanned aerial ad hoc networks: A survey". *Ad Hoc Networks*. Vol. 92, Article 101778. DOI: <https://doi.org/10.1016/j.adhoc.2019.101778>.
7. Wang, W., Lv, M., Ru, L., Lu, B., Hu, S., Chang, X. (2022), "Multi-UAV Unbalanced Targets Coordinated Dynamic Task Allocation in Phases", *Aerospace*, Vol. 9, Article 491. DOI: <https://doi.org/10.3390/aerospace9090491>
8. Krytskyi, D., Karatanov, O., Pohudina, O., Shevel, V., Bykov, A., Pyvovar, M., Plastun, T. (2024), "Information Technology for Determining the Flight Performance of a Paraglider Wing". *Information Technologies in the Design of Aerospace Engineering*, Springer Nature Switzerland, P. 1–42. DOI: https://doi.org/10.1007/978-3-031-43579-9_1
9. Zhou, X., Chen, H., Liu, J., He, Y., Wu, H. (2024), "Towards secure and resilient unmanned aerial vehicles swarm network based on blockchain", *IET Blockchain*, Vol. 4, P. 483–493. DOI: <https://doi.org/10.1049/blc2.12050>

10. Li, Z., Chen, H., Sun, L., Wu, J., Li, Q. (2024), "A secure and efficient UAV network defense strategy: Convergence of blockchain and deep learning", *Computer Standards & Interfaces*, Vol. 90, Article 103844. DOI: <https://doi.org/10.1016/j.csi.2024.103844>
11. Mershad, K. (2022), "PROACT: Parallel multi-miner proof of accumulated trust protocol for Internet of Drones", *Vehicular Communications*, Vol. 36, Article 100495. DOI: <https://doi.org/10.1016/j.vehcom.2022.100495>
12. Hafeez, S., et al. (2024), "Beta-UAV: blockchain-based efficient authentication for secure UAV communication", *Cryptography and Security*. 2024. DOI: <https://doi.org/10.48550/arXiv.2402.15817>
13. Yevdokymenko, M. O. (2020), "Theoretical foundations of fault-tolerant routing in telecommunication networks" ["Teoretychni osnovy vidmovostiikoji marshrutyzatsii v telekomunikatsiynykh merezhakh"], Dissertation, Kharkiv National University of Radioelectronics.
14. Yasinchuk, V. I. (2020), "Multipath routing based on ant colony algorithms" ["Bahatoshliakhova marshrutyzatsiia na osnovi alhorytmiv murashkovykh kolonii"], available at: <https://dspace.wunu.edu.ua/bitstream/316497/1510/1/Yasinchuk%20V.I.%20C%20KSMzm-51.pdf> (last accessed: 08.05.2025).
15. Krytskyi, D., Karatanov, O., Pohudina, O., Shevel, V., Bykov, A., Pyvovar, M., Plastun, T. (2024), "Information Technology for Determining the Flight Performance of a Paraglider Wing". *Information Technologies in the Design of Aerospace Engineering*, Springer Nature Switzerland, P. 1–42. DOI: https://doi.org/10.1007/978-3-031-43579-9_1
16. Binko, I. V., Shevel, V. V., Bykov, A. M., Krytskyi, D. M. (2024), "Analysis of decentralized drone control model and interception trajectory calculation". *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (28), P. 33–47. DOI: <https://doi.org/10.30837/2522-9818.2024.2.033>
17. Yena, M. (2024), "Urban UAV mobility control: Swarm intelligence and collision avoidance", *Innovative Technologies and Scientific Solutions for Industries*, No. 4, P. 210–218. DOI: <https://doi.org/10.30837/2522-9818.2024.4.059>
18. Terenyk, D., Kharchenko, V. (2024), "Deployment strategy selection and swarm UAV reliability support for communication in destruction conditions", *Innovative Technologies and Scientific Solutions for Industries*, No. 3, P. 155–162. DOI: <https://doi.org/10.30837/2522-9818.2024.3.091>

Надійшла (Received) 08.05.2025

Відомості про авторів / About the Authors

Сопов Єгор Олександрович – Національний аерокосмічний університет "Харківський авіаційний інститут", аспірант кафедри інформаційних технологій проєктування, Харків, Україна; e-mail: y.o.sopov@khai.edu; ORCID ID: <https://orcid.org/0009-0004-7243-3021>

Крицький Дмитро Миколайович – кандидат технічних наук, доцент, Національний аерокосмічний університет "Харківський авіаційний інститут", доцент кафедри інформаційних технологій проєктування, Харків, Україна; e-mail: d.krickiy@khai.edu; ORCID ID: <https://orcid.org/0000-0003-4919-0194>

Артюмова Аліна Вадимівна – кандидат технічних наук, доцент, Національний аерокосмічний університет "Харківський авіаційний інститут", доцент кафедри інформаційних технологій проєктування, Харків, Україна; e-mail: a.artymova@khai.edu; ORCID ID: <https://orcid.org/0000-0002-6761-2066>

Артюмов Ігор Володимирович – Національний аерокосмічний університет "Харківський авіаційний інститут", асистент кафедри інформаційних технологій проєктування, Харків, Україна; e-mail: i.artomov@khai.edu; ORCID ID: <https://orcid.org/0009-0001-5368-2469>

Sopov Iegor – National Aerospace University "Kharkiv Aviation Institute", Postgraduate Student at the Department of Information Technology Design, Kharkiv, Ukraine.

Krytskyi Dmytro – PhD (Engineering Sciences), Associate Professor, National Aerospace University "Kharkiv Aviation Institute", Associate Professor at the Department of Information Technology Design, Kharkiv, Ukraine.

Artomova Alina – PhD (Engineering Sciences), Associate Professor, National Aerospace University "Kharkiv Aviation Institute", Associate Professor at the Department of Information Technology Design, Kharkiv, Ukraine.

Artomov Ihor – National Aerospace University "Kharkiv Aviation Institute", Assistant at the Department of Information Technology Design, Kharkiv, Ukraine.

TRUST-BASED ROUTING METHODOLOGY IN UAV SWARM NETWORKS BASED ON TRAFFIC ANALYSIS AND ANOMALY DETECTION

Subject matter. the subject of the research is the process of ensuring secure routing and data exchange among unmanned aerial vehicles (UAVs) within swarm networks under cyber threat conditions, particularly Black Hole-type attacks. **Goal.** The purpose of this study is to develop and simulate a secure information transmission mechanism for UAV swarm networks that takes into account node trust levels and enables the identification of malicious participants based on behavioral analysis. The study also aims to establish a methodology for secure and energy-efficient routing in FANETs based on blockchain technologies and trust evaluation models, ensuring cyber resilience, data integrity, and minimal resource usage. **Tasks** the following objectives were addressed during the research. Conduct a comprehensive analysis of vulnerabilities in traditional routing protocols used in FANETs to identify potential threats to information security; justify the use of trust-based mechanisms to improve routing resilience against internal attacks; implement a Black Hole attack model within the NS-3 simulation environment to analyze its impact on swarm network performance; develop a mechanism for counting forwarded packets per node as a foundation for a trust evaluation system among agents; visualize simulation results to support analysis and comparison of proposed methods. **Methods:** the research employs simulation modeling of FANETs in NS-3.36, using the RandomWaypoint mobility model and the AODV routing protocol. Methods include statistical analysis of packet forwarding metrics and graphical representation of trust metrics. The developed code simulates the Black Hole attack by manipulating the NetDevice layer and logs all transmitted packets in CSV format for post-processing. A combination of simulation tools, analytical analysis, and visualization techniques was applied to evaluate system performance under dynamic conditions. **Results.** The results demonstrate the effectiveness of the proposed approach in detecting malicious nodes within the swarm network. Trust metrics revealed anomalous attacker behavior, such as the absence of packet forwarding, distinguishing them from normal nodes. This allows for timely identification and exclusion of threats from the routing process. Graphical visualization clearly displays node activity distribution, facilitating result interpretation without the need for in-depth log analysis. **Conclusions.** The proposed trust-based mechanism, combined with node activity analysis, effectively protects FANET networks against Black Hole attacks. Future improvements may include integrating more advanced trust assessment methods, such as multifactor analysis, blockchain, or machine learning. Developing adaptive routing algorithms capable of autonomously isolating or excluding suspicious nodes is also recommended

Keywords: UAV swarm network, FANET, AODV, Black Hole attack, trust, NS-3 simulation, routing security, attack mitigation, NetDevice, malicious node detection.

Бібліографічні описи / Bibliographic descriptions

Сопов Є. О., Крицький Д. М., Артѣмова А. В., Артѣмов І. В. Методика побудови довірчої маршрутизації в ройових мережах БПЛА на основі аналізу трафіку та виявлення аномалій. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 2 (32). С. 111–128. DOI: <https://doi.org/10.30837/2522-9818.2025.2.111>

Sopov, I., Krytskyi, D., Artomova, A., Artomov, I. (2025), "Trust-based routing methodology in uav swarm networks based on traffic analysis and anomaly detection", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (32), P. 111–128. DOI: <https://doi.org/10.30837/2522-9818.2025.2.111>

І. ЧУХРАН, С. УДОВЕНКО, В. ШЕРГІН, Л. ЧАЛА

МЕТОД ДОПОВНЕННЯ 3D-МОДЕЛЕЙ ТОЧКОВИХ ХМАР З ВИКОРИСТАННЯМ ГРАФОВИХ НЕЙРОННИХ МЕРЕЖ

Побудова моделей подання тривимірних об'єктів у вигляді просторової сукупності слабкоструктурованих точок (3D-моделей точкових хмар) набуває поширення в різних сферах, зокрема в завданнях створення віртуальної реальності, маршрутизації автономних роботів і 3D-реконструкції. Необхідність формувати та обробляти дані 3D-моделей хмар точок є суттєвою для застосунків, що потребують ретельного аналізу особливостей навколишнього середовища, пов'язаних з подоланням перешкод, визначенням маршрутів транспортних засобів та моделюванням реальних сцен. Утім, з різних причин точкові хмари часто містять пропущені ділянки, що створює значні проблеми для подальшого оброблення інформації. Неповні дані точкових хмар можуть мати серйозні наслідки, наприклад, у системах автономної навігації, де помилки призводять до зіткнень або інших небезпечних ситуацій. Розв'язання цієї проблеми є ключовим для надійного оброблення 3D-даних. **Метою статті** є розроблення та дослідження методу автоматичного доповнення й реконструкції точкових хмар з використанням графових нейронних мереж. Основними **завданнями дослідження** є аналіз наявних підходів до побудови та відновлення тривимірних графових моделей, розроблення та програмна реалізація методу автоматичного доповнення точкових хмар з використанням графових нейронних мереж, а також моделювання запропонованого методу для завдань доповнення та 3D-реконструкції моделей точкових хмар. У роботі розглянуто загальну модель доповнення точкових хмар із застосуванням графових нейронних мереж, що дають змогу кодувати неповні 3D-моделі хмар точок із використанням графів та прогнозуванням положення відсутніх фрагментів моделей. Запропоноване рішення для доповнення неповних 3D-хмар точок має наукову новизну та поєднує потужність графових нейронних мереж (GNN) з архітектурою мережі *Point Completion Network* (PCN). Окреслений підхід сприяє якісному відновленню неповних 3D-даних у практичному застосуванні. Значущість роботи підтверджують результати моделювання запропонованого методу для класичних датасетів та їх порівняння з деякими наявними підходами до розв'язання досліджуваної проблеми. Перспективним напрямом продовження досліджень з окресленої теми є випробування різних архітектур нейронних графових мереж, налаштування гіперпараметрів, застосування альтернативних функцій втрат і більш потужних обчислювальних ресурсів для навчання побудованих нейромережних моделей.

Ключові слова: точкова хмара; глибоке навчання; графова нейронна мережа; автоматичне доповнення точкових хмар; кодування графів; 3D-реконструкція графових моделей.

Вступ

Застосування подання тривимірних об'єктів або середовищ у формі неструктурованих наборів точок у просторі, відомих як 3D-хмари точок, набуває дедалі більшого поширення в таких галузях, як автономна навігація, робототехніка, технології віртуальної та доповненої реальності, а також тривимірна реконструкція.

Відповідно до усталених наукових визначень 3D-хмара точок є сукупністю даних, що містить точки, координати яких задані в тривимірній системі та які відтворюють зовнішню поверхню об'єктів або просторових середовищ. Такі дані зазвичай отримують за допомогою сучасних методів лазерного сканування [1] або способом фотограмметричного аналізу, коли вимірювання відстаней трансформуються в множину точок, що характеризують поверхню досліджуваних об'єктів.

Зберігання даних 3D-хмар точок можливе в різних сучасних форматах, які мають унікальні властивості, адаптовані до особливих потреб оброблення та прикладного використання. Серед найбільш поширених форматів можна виокремити:

- формат PLY (*Polygon File Format*), що дає змогу зберігати тривимірні скановані дані;
- формат OBJ (*Object File Format*), що використовується, зокрема, для зберігання 3D-хмар точок та містить інформацію про топологію сіток та вершин;
- формати PTS або PTX, що застосовуються в геодезичних програмних застосунках для зберігання та оброблення відсканованих тривимірних наборів даних;
- формат LAS (*LiDAR Aerial Surveying Format*), що впроваджується для зберігання та оброблення великих обсягів геопросторових даних;
- формат XYZ, що зберігає 3D-координати хмар точок разом з додатковою інформацією про їх колір.

Здатність до точного захоплення та оброблення даних тривимірних (3D) хмар точок є критично важливою для застосунків, що потребують глибокого розуміння просторового середовища, зокрема для завдань уникнення перешкод, планування траєкторій та моделювання сценічного простору. Проте внаслідок таких факторів, як оклюзії, обмеження сенсорних систем або недоліки процедур збору даних, 3D-хмари точок часто визначаються наявністю пропущених ділянок або неповним відтворенням об'єктів чи сцен. Ця неповнота суттєво перешкоджає подальшим етапам оброблення, які залежать від точної та вичерпної тривимірної інформації, що може спричинити помилки або зниження ефективності в критично важливих системах.

Неповнота даних у хмарах точок має особливо серйозні наслідки для застосунків, де безпека є пріоритетом, наприклад, у системах автономної навігації безпілотних транспортних засобів або роботизованих операціях у динамічних середовищах. Відсутність чи неточність інформації про перешкоди, особливості рельєфу або інші елементи оточення може спричинити зіткнення чи навігаційні збої. Крім того, неповні хмари точок ускладнюють адекватне сприйняття та моделювання сцен, що обмежує можливості таких технологій, як віртуальна та доповнена реальність і тривимірна реконструкція. Отже, розроблення методів доповнення хмар точок є ключовим для ефективного оброблення 3D-даних і максимального розкриття потенціалу відповідних застосунків.

Традиційні методи, зокрема геометричні підходи чи техніки реконструкції поверхонь, часто виявляються недостатньо ефективними в разі наявності складних структур об'єктів моделювання або значних (у відсотках) відсутніх ділянок 3D-моделей, а також за умов неоднорідної щільності хмар точок і присутності шумів. Останні досягнення у сфері глибокого навчання, зокрема в геометричному глибокому навчанні, відкривають нові перспективи для розв'язання окресленої проблеми. Графові нейронні мережі (GNN), що належать до моделей глибокого навчання, адаптованих до роботи з нерегулярними й неєвклідовими структурами даних, продемонстрували перспективні результати в завданнях, пов'язаних із класифікацією, кластеризацією, сегментацією та генерацією даних 3D-хмар точок.

У цьому дослідженні запропоновано інноваційний підхід до доповнення хмар точок, який

оснований на використанні графових нейронних мереж. Цей метод передбачає необхідність кодування вхідної хмари точок із застосуванням графової моделі, точки якої є вузлами, а ребра відтворюють геометричні зв'язки між точками. Таке графове подання забезпечує ефективне моделювання нерегулярної структури даних хмари точок і є основою для побудови моделі кодера, що має навчитися виділяти ключові ознаки та формувати подання на основі структури графа.

Упровадження графових моделей дає змогу певною мірою вдосконалити технології оброблення хмари точок із застосуванням геометричного глибокого навчання. Автоматичне доповнення відсутніх фрагментів хмар сприяє підвищенню безпеки функціонування критично важливих систем, що мають використовувати достовірну 3D-інформацію щодо особливостей навколишнього середовища.

Мета статті – розроблення та дослідження методу автоматичного доповнення та реконструкції точкових хмар з використанням графових нейронних мереж.

Завдання дослідження:

- аналіз наявних підходів до автоматичного доповнення хмар точок;
- порівняння метрик, що можуть бути застосовані для побудови графічних моделей хмар точок;
- розроблення методу автоматичного доповнення точкових хмар з використанням графових нейронних мереж;
- моделювання запропонованого методу для завдань доповнення та 3D-реконструкції моделей точкових хмар.

Аналіз наявних підходів до автоматичного доповнення хмар точок

Побудова та доповнення тривимірних моделей базується на використанні традиційних методів і методів глибокого навчання. До традиційних насамперед належать геометричні методи та методи вирівнювання.

Традиційними методами доповнення хмар точок є методи на основі геометрії хмар і методи на основі вирівнювання. Методи, що ґрунтуються на геометричному підході [2], використовують геометричні властивості, такі як симетрія форми та її дисперсія, отримані з наявної бази даних форм, для реконструкції повноцінної моделі на основі

часткових вхідних даних. Деякі із зазначених методів [3–5] локально заповнюють пропущені ділянки внаслідок інтерполяції гладких поверхонь на основі сусідніх структур. Проте такі підходи виявляються недостатньо ефективними в разі, коли вхідні моделі мають значні пропущені ділянки, оскільки їм бракує семантичної інформації, наприклад, про структуру чи топологію. Водночас, зважаючи на симетричність багатьох об'єктів, окремі методи [6, 7] здатні виявляти симетрію у вхідних тривимірних моделях і заповнювати відсутні частини, спираючись на їх симетричну геометрію.

Методи, основані на вирівнюванні [8, 9], використовують великомасштабні бази даних хмар точок для пошуку відповідних фрагментів, які згодом застосовуються для заповнення пропущених ділянок. Ключовим етапом у цих підходах є формування бази даних, яка охоплює тривимірні форми однієї категорії. Деякі методи [10] обирають окремі частини моделей як базові елементи, тоді як інші [11, 12] використовують деформацію тривимірних фігур для створення таких баз. Зазначений підхід дає змогу успішно доповнювати 3D-моделі зі схожою структурою, однак його ефективність знижується, коли вхідні моделі виходять за межі заздалегідь визначених категорій.

Отже, традиційні методи значною мірою покладаються на геометричні властивості, сформовані вручну, для виведення структури відсутніх ділянок на основі часткової форми. Їх результативність суттєво обмежена в разі, коли тривимірні моделі не відповідають попередньо встановленим припущенням чи знанням, що ускладнює досягнення задовільних результатів.

Методи доповнення тривимірних моделей з використанням глибокого контрольованого або неконтрольованого навчання передбачають необхідність побудови спеціалізованих нейромережних моделей.

Нейромережний підхід до вирівнювання хмар точок, оснований на навчанні нейронних моделей, запропонований у роботі [13], де використано рекурентні нейронні мережі для кодування двовимірних хмар точок із подальшим декодуванням у контексті розв'язання геометричних задач, зокрема двовимірної задачі комівояжера. Запропонований у цій статті метод також розв'язує проблему генерації хмар точок змінного розміру з використанням

механізму уваги, однак він не адаптований до тривимірного (3D) простору.

У дослідженні [14] запропоновано підхід, який базується на поданні 3D-сітки об'єкта у вигляді набору фрагментів атласу частин об'єкта. Первинні форми фрагментів кодуються за допомогою багат шарового перцептрона в латентний простір, після чого точки декодуються в латентне подання щодо закодованої форми-зразка, що дає змогу генерувати 3D-сітку замість хмар точок.

У праці [15] запропоновано доповнювати хмари точок на основі груп точок, а не окремих одиниць, що дає змогу кожну модель подавати у вигляді тривимірної воксельної сітки. У статті [16] розглянуто використання воксельних мереж для доповнення хмар точок способом генерації країв. На початковому етапі хмари точок інтегруються в регулярні воксельні сітки, а потім за допомогою ребер уявної форми створюються цілісні об'єкти.

У дослідженнях [17, 18] проаналізовано підхід, що передбачає застосування одновимірних згорток до точок для розв'язання задач класифікації та сегментації на наборах хмар точок із використанням простого та ієрархічного методів. У згаданих студіях спочатку формуються локальні та глобальні ознаки, які разом із обробленими даними передаються до кодувальної мережі. На подальших етапах ці дані трансформуються в хмари точок із грубим і тонким виходом, де грубий вихід зберігає лише ключові точки, що сприяють формуванню більш точного результату на наступних етапах генерації.

Аналогічний підхід застосовано в роботі [19], де використано кодувальник PCN для вибірки різних поверхонь. У цьому разі попередня інформація збирається подібно до глобальних ознак з набору поверхонь. Мережа декодувальника навчається відбирати потрібні точки на основі аналізу узагальненого вектора кодованих ознак неповного вихідного об'єкта. Для подолання обмежень автокодера, таких як ігнорування структури, злиття та уточнення, у дослідженні [20] запропоновано метод вибірки за критерієм мінімальної щільності. Цей підхід визначає відновлювану точку з найменшою щільністю, що підвищує ефективність оброблення.

У праці [21] для заповнення хмар точок на основі реальних даних застосовано змагальне навчання. У роботі [22] запропоновано використання генеративної змагальної мережі (GAN), у якій генератор, що є адаптаційною мережею, трансформув

вхідну інформацію в латентне подання так, щоб дискримінатор не розрізняв результати перетворення від результатів, отриманих з навчальних наборів даних. Генератор, власне, виконує основну функцію відтворення необроблених часткових наборів точок у завершені та цілісні набори, що реалізується внаслідок упорядкованої роботи у двох окремих латентних просторах, сформованих на основі відсканованих і синтезованих даних об'єктів.

Альтернативний підхід, подібний до GAN, розглянуто в роботі [23], де генерація загальної форми хмари точок здійснюється з неповного набору точок із неминучими змінами наявних точок, а також з огляду на шум і геометричні втрати. У цьому методі мережа ієрархічно оцінює відсутні частини хмари точок, застосовуючи великомасштабну генеративну мережу, що базується на характерних точках.

Окремі дослідження спрямовані на вдосконалення процесу доповнення хмар точок. Так, у роботі [24] запропоновано нейронні мережі з архітектурою кодувальник-декодувальник для безпосереднього оцінювання всієї хмари точок на основі неповних даних. Цей підхід реалізовано через наскрізну нейронну архітектуру, що зосереджується на прогнозуванні відсутньої геометрії та інтеграції відомої вхідної інформації з передбаченою хмарою точок. Метод містить дві нейронні мережі: перша відповідає за прогнозування відсутньої частини, витягуючи інформацію з неповних даних, а друга забезпечує об'єднання та уточнення розподілу точок у результуючій хмарі.

У статті [25] запропоновано нейромережу автокодера, що аналізує латентні подання та формує кілька зразків доповнених 3D-хмар точок. У цьому разі інформація, що обробляється, розподіляється на два окремі потоки, а заповнення прогалів реалізується через парадигму гіпермереж. Гіпермережі, як зазначено в роботі [26], є нейронними мережами, що генерують ваги для цільової мережі, оптимізуючи її функціональність.

Інші підходи пропонують ітеративну генерацію відсутніх частин хмари точок способом поступового додавання нових елементів до наявних. Наприклад, у праці [27] завершення хмар точок у 3D-просторі здійснюється за принципом генерування дочірніх точок унаслідок послідовного розщеплення батьківських точок із використанням оператора *Snowflake Point Deconvolution* (SPD), запропонованого авторами.

У низці досліджень описані рішення, які, на нашу думку, є найбільш перспективними з погляду уваги до геометричних особливостей із застосуванням графових нейронних мереж. Зокрема в роботі [28] подано синтез мережі ECG (*Edge-aware Completion Graph*) для завершення хмар точок, що інтегрує ребра з графічною згортою. Цей підхід сприяє генерації деталізованої тривимірної (3D) хмари точок із різномасштабними крайовими властивостями. Щоб зважати на локальні геометричні деталі в разі розширеної вибірки, автори запропонували модуль *Edge-aware Feature Expansion* (EFE), який забезпечує плавне розширення та вдосконалення вибірки точкових об'єктів з огляду на їх локальні краї.

Відповідно до мети цього дослідження, далі буде розглянуто завдання розроблення методу доповнення 3D-хмар точок із використанням графових нейронних мереж.

Метрики, що можуть бути застосовані для побудови графічних моделей хмар точок

У межах аналізу можливих метрик для оцінювання відстані між наборами даних тривимірних точок розглядається проблема оцінювання якості реконструкції 3D-моделей, що передбачає порівняння двох поверхонь. Такими можуть бути, наприклад, поверхня, побудована на основі хмари 3D-точок, та еталонна поверхня. Відомі метрики в цьому разі базуються на оцінці похибок відстані між поверхнями. Їх сутність полягає в обчисленні відстані між всіма точками однієї поверхні до найближчих точок іншої поверхні з подальшим узагальненням результатів у вигляді статистичних показників. Найпоширенішими метриками цього типу є середньоквадратична помилка (RMSE) та середня абсолютна помилка (MAE). Для зазначених метрик точність реконструкції є часткою правильно реконструйованих точок у загальному наборі реконструйованих даних, тобто тих точок, відстань від яких до найближчої точки еталонної поверхні не перевищує заданого порогу.

Окремі метрики продемонстрували свою ефективність в оцінюванні результатів сегментації, передбаченої моделями глибокого навчання. Їх основна ідея полягає в одночасному вимірюванні точності класифікації та коректності локалізації. До таких метрик належать точність (*precision*),

відгук (*recall*), індекс перетину над об'єднанням (*IoU*, *Jaccard Index*) та показник *F1*. Однак застосування зазначених метрик може потребувати вокселізації вибірки даних, що дещо знижує довіру до результатів оцінювання. Далі розглядаються найбільш поширені метрики цього типу.

$$HD(P_A, P_B) = \frac{1}{2} \max_{x \in P_A} |x - NN(x, P_B)| + \frac{1}{2} \max_{x \in P_B} |x - NN(x, P_A)|, \quad (1)$$

де $NN(x, P) = \arg \min_{x' \in P} \|x - x'\|$.

Якщо розглядати розподіли ймовірностей P_A та P_B для наборів хмар точок, то відстань між цими хмарами можна визначати як дивергенцію Кульбака – Лейблера (KL):

$$D_{KL}(P_A \parallel P_B) = \sum_{x \in \chi} P_A(x) \log \left(\frac{P_A(x)}{P_B(x)} \right), \quad (2)$$

де χ – простір вибірки для точок хмар, що порівнюються.

Показник KL дає змогу, зокрема, оцінити різницю між синтезованою та реальною вибіркою в

$$CD(P_A, P_B) = \frac{1}{2n} \sum_{i=1}^n |x_i - NN(x_i, P_B)| + \frac{1}{2m} \sum_{j=1}^m |x_j - NN(x_j, P_A)|. \quad (4)$$

За припущенням, що хмари точок є вирівняними по осях, а навколо них існує канонічна сітка вокселів, доцільно використовувати метрику розбіжності Дженсена – Шеннона (*Jensen-Shannon Divergence – JSD*) між граничними розподілами, визначеними в евклідовому 3D-просторі. Ця метрика дає змогу оцінити, наскільки хмари точок A мають тенденцію збігатися в середньому з хмарами точок B . Для цього підраховуємо кількість точок, що лежать у межах кожного вокселя для всіх хмар точок A та, відповідно, для B , й оцінюємо розбіжність *JSD* між отриманими емпіричними розподілами (P_A, P_B) .

$$JSD(P_A \parallel P_B) = \frac{1}{2} D(P_A \parallel M) + \frac{1}{2} D(P_B \parallel M), \quad (5)$$

де $M = \frac{1}{2}(P_A + P_B)$ та $D(\cdot \parallel \cdot)$ – KL-дивергенція між двома розподілами [29].

Для кожної хмари точок в A спочатку знаходимо її найближчого сусіда в B . Покриття вимірюється як частка хмар точок у B , які були зіставлені з хмарами точок в A . Близькість може бути обчислена за допомогою *CD*- або

Метрика просторової відстані Хаусдорфа (HD) інколи використовується для оцінювання результатів 3D-моделювання. Вона дає змогу оцінювати відстань між точками P_A та P_B графової моделі з використанням функції найбільш близьких сусідів $NN(x, P)$:

процесі 3D-реконструкції. Альтернативним показником, що для визначення різниці між 3D-об'єктами також використовує розподіли ймовірностей, є відстань *Earth-Mover's Distance (EMD)*:

$$EMD(P_A, P_B) = \min_{\pi \in \Pi(P_A, P_B)} \left[\sum_{i=1}^n \sum_{j=1}^m \pi_{ij} |p_{Ai} - p_{Bj}| \right]. \quad (3)$$

Іншим підходом, що бере до уваги геометрію, є обчислення відстані фаски (*Chamfer distance*), яка визначається як середня відстань між парами найближчих сусідів між двома наборами хмари точок P_A та P_B відповідно:

EMD-відстані між точками з використанням метрики *COV-CD* або *COV-EMD*:

$$COV(P_A, P_B) = \frac{\left| \arg \min_{Y \in P_B} D(X, Y) \mid X \in P_A \right|}{|P_B|}, \quad (6)$$

де D може бути *CD* або *EMD*.

Високий показник покриття свідчить про те, що більша частина B подана в межах A . Покриття не вказує на те, наскільки добре хмари точок містяться в множині A ; приклади, що збігаються, не обов'язково мають бути близькими. Для оцінювання точності A щодо B кожна хмара точок B зіставляється з хмарою точок A з мінімальною відстанню (*MMD*) і визначається середнє значення відстаней у зіставленні. Тут може бути використана будь-яка відстань між наборами точок, що дає метрики *MMD-CD* і *MMD-EMD*:

$$MMD(P_A, P_B) = \frac{1}{|P_B|} \sum_{Y \in P_B} \min_{X \in P_A} D(X, Y), \quad (7)$$

де D відповідає значенням *CD* або *EMD*.

Показник *MMD* безпосередньо залежить від відстаней між порівнюваними наборами точок.

Загалом метрики MMD , $COV-CD$ та $COV-EMD$ добре корелюють між собою. Метрика JSD відрізняється від них тим, що оцінює подібність між A і B лише за допомогою граничної статистики й вимагає попередньо вирівняних даних, але водночас є більш зручною для обчислень. Проте ця метрика є ефективною, коли необхідно здійснювати послідовність порівнянь між 3D-моделями хмар точок.

Метод автоматичного доповнення точкових хмар з використанням графових нейронних мереж

Розглянемо особливості структури графових нейронних мереж (GNN) та їх функціональних компонентів. Графові мережі є методами глибокого навчання, які працюють на основі використання графів. Ці мережі дають змогу аналізувати та обробляти неевклідові дані, такі як графи, що містять вузли та ребра [30]. Спочатку такі мережі були вперше використані на спрямованих ациклічних графах. Пізніше графові моделі були застосовані до рекурентних нейронних мереж та нейронних мереж прямого поширення, що вимагають ітераційного визначення переходу станів на графах.

Останні дослідження у сфері глибоких нейронних мереж, зокрема згорткових нейронних мереж (CNN) [31], привели до розширення можливостей GNN. Мережі CNN оперують з різномасштабними просторовими ознаками та компонують їх для побудови якісних зображень [32]. Для CNN властиве локальне з'єднання використання декількох шарів. У цих мережах обробляються евклідові дані для опису зображень і текстів, що дає змогу застосовувати графове подання. Однак узагальнення CNN на графові моделі не завжди сприяє визначенню локалізованих фільтрів згортки та операторів об'єднання, що перешкоджає поширенню моделей CNN з евклідової ділянки до неевклідової (рис. 1).

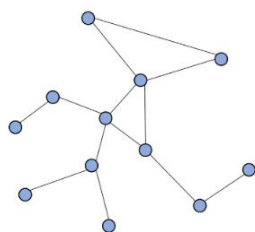


Рис. 1. Приклад зображення в евклідовому просторі (ліворуч) та його графічного спрощеного подання в неевклідовому просторі (праворуч)

Застосування геометричного глибокого навчання для побудови графічних моделей в неевклідовому просторі є новим напрямом досліджень у сфері машинного навчання. Ефективним є навчання графів [33], де здійснюється подання вершин, ребер або підграфів графів у вигляді векторів невеликої розмірності. Але традиційні підходи до машинного навчання графових моделей здебільшого використовують функції, які розробляються вручну, що зумовлює їх негнучкість і високу обчислювальну вартість. У роботах [34, 35] розглянуто метод автоматичного вбудовування графів на основі згенерованих випадкових блукань. Однак у цьому методі параметри не розподіляються між вузлами кодувальної частини моделі, що спричиняє лінійне зростання обчислюваних витрат зі зростанням кількості вузлів. Крім того, використання прямого вбудовування обмежує його здатність до узагальнення (це означає, що він не може працювати з новими типами графів).

Розглянемо деякі термінологічні особливості застосування графових моделей.

Графи цих моделей можна розподілити на структуровані та неструктуровані. Структуровані графи мають явно визначену будову, наприклад, молекули, фізичні системи та графи знань. Неструктуровані графи, навпаки, мають неявну структуру, яка має бути побудована на основі конкретного завдання, як-от графи слів або графі-сценарії для зображень. Також графи можуть бути орієнтованими (із заданими напрямками ребер між вершинами) та неорієнтованими. Графи, що можуть змінюватися з часом, називаються динамічними, а ті, що не залежать від часу, вважаються статичними.

Можна виокремити три основні типи завдань, пов'язаних з аналізом графів:

- завдання на рівні вузлів (зокрема класифікація, кластеризація та регресія вузлів);
- завдання на рівні ребер (а саме класифікація типів ребер та прогнозування зв'язків між вершинами графа);
- завдання на рівні графів (зокрема класифікація, регресія та порівняння графів).

В алгоритмах навчання мереж на основі графів можливе використання трьох різних режимів навчання:

- контрольоване, що надає позначену для навчання інформацію;
- напівконтрольоване навчання, що передбачає наявність незначної кількості позначених вузлів

і великої кількості непозначених вузлів для навчання (в більшості задач класифікації вершин і ребер упроваджується цей тип навчання);

– неконтрольоване навчання, що пропонує лише немарковані дані для налаштування моделі (цей тип навчання використовується, зазвичай, для кластеризації вузлів).

Для побудови графових нейронних мереж застосовують такі основні обчислювальні модулі:

– модуль поширення (для поширення характеристичної та топологічної інформації між вузлами мережі); у цих модулях оператор згортки

застосовується для агрегування інформації від сусідніх вузлів;

– модуль вибірки (для проведення поширення на графах великої розмірності); цей модуль зазвичай поєднується з модулем поширення;

– модуль об'єднання або субдискретизації (використовується для отримання інформації з вузлів у разі необхідності подання високорівневих підграфів або графів).

Загальну структуру конвеєра GNN зображено на рис. 2.

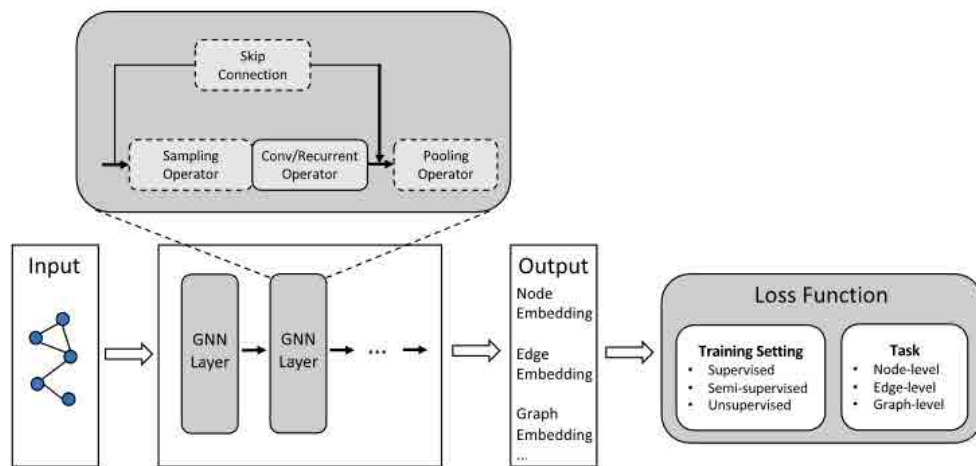


Рис. 2. Загальна структура конвеєра GNN

Розглянемо більш детально функціональні особливості окремих компонентів GNN.

Модуль поширення в моделях GNN реалізує операції згортки, основна ідея яких полягає в узагальненні згорток з різних ділянок графа. У модулі поширення використовуються як спектральні, так і просторові підходи. Спектральні базуються на обробленні графових сигналів з визначенням операторів згортки в спектральній ділянці. У цьому разі графічний сигнал x спочатку подається в спектральній ділянці за допомогою графового перетворення Фур'є, а потім здійснюється операція згортки.

Просторові підходи ґрунтуються на реалізації згорток на графі для просторово близьких сусідів. Вони основані на теоремі згортки та спрямовані на використання локалізованих операторів, що можуть бути розпаралелені та масштабовані. Відповідні перетворення формально визначаються як

$$\mathcal{F}(x) = U^T x, \quad \mathcal{F}^{-1}(x) = Ux, \quad (8)$$

де U – матриця власних векторів нормалізованого лапласіана графа

$$L = I_N - D^{-1/2} A D^{-1/2}, \quad (9)$$

де D – матриця степенів; A – матриця суміжності графа.

Лапласіан нормалізованого графа є дійсною симетричною додатною напівскінченністю, тому його можна розкласти на множники у вигляді

$$L = U \Lambda U^T, \quad (10)$$

де Λ – діагональна матриця власних значень.

З огляду на теорему згортки операція згортки здійснюється таким чином [36]:

$$g * x = \mathcal{F}^{-1}(\mathcal{F}(g) \odot \mathcal{F}(x)) = U(U^T g \odot U^T x). \quad (11)$$

Формулу (11) можна спростити, якщо замість $U^T g$ використовувати навчальну матрицю g_w :

$$g_w * x = U g_w U^T x. \quad (12)$$

У цьому разі можуть бути застосовані різні типи навчальних матриць. Зокрема в спектральній мережі

може бути використана діагональна навчальна матриця-фільтр:

$$g_w = \text{diag}(w); \quad w \in \mathbb{R}^N. \quad (13)$$

Однак фільтр (13) не є просторово локалізованим, що ускладнює його практичну реалізацію. У роботі [37] g_w апроксимується з використанням усиченого розкладання за допомогою поліномів Чебишева $T_k(x)$ до K^{th} порядку:

$$g_w * x \approx \sum_{k=0}^K w_k T_k(\bar{L}), \quad (14)$$

де $\bar{L} = \frac{2}{\lambda_{\max}} L - I_N$; $\lambda_{\max}$ – максимальне власне значення матриці L ; w – вектор коефіцієнтів Чебишева.

У рівнянні (14) поліноми Чебишева визначаються в такий спосіб:

$$T_k(x) = 2xT_{k-1}(x), \quad (15)$$

де $T_0(x) = 1$ та $T_1(x) = x$.

У праці [38] показано, що для $\lambda_{\max} \approx 2$ і для двох вільних параметрів w_0 та w_1 можна подати рівняння (13) в такому вигляді:

$$g_w * x \approx w_0 x + w_1 (L - I_N) x = w_0 x - w_1 D^{-1/2} A D^{-1/2} x, \quad (16)$$

або (за умови $w = w_0 = -w_1$)

$$g_w * x \approx w (I_N - D^{-1/2} A D^{-1/2}) x. \quad (17)$$

Практичне застосування модуля поширення в моделях GNN пов'язане з використанням інтеграції механізму уваги (GAT), що дає змогу обчислювати приховані стани кожного вузла на основі стратегії самоуваги. До того ж механізм GAT використовує багатоголову увагу для покращення процесу навчання мережі, формуючи K незалежних матриць уваги для обчислення прихованих станів, які потім об'єднуються. Іноді доцільно регулювати кількість шарів k графової нейронної мережі для покращення результатів. Збільшення кількості шарів дає змогу кожному вузлу отримувати більше інформації від сусідніх вузлів. Однак експериментально встановлено, що більш глибокі моделі не завжди покращують роботу мережі, а занадто глибокі моделі можуть навіть працювати гірше. Це пов'язано з тим, що занадто велика кількість шарів призводить до поширення зашумленої інформації від сусідів, унаслідок чого виникає ефект надмірного згладжування, тобто вузли починають мати схожі подання після агрегування.

Ефективним підходом для розв'язання окресленої проблеми є використання модулів вибірки на великих графах. Існують такі основні різновиди модулів вибірки графів: вибірка вершин (вузлів), вибірка шарів та вибірка підграфів. Одним зі способів зменшення кількості сусідніх вершин є вибірка скороченої множини вузлів з околів кожної вершини (операція *GraphSAGE*) [39]. У цьому разі обирається фіксована обмежена кількість сусідів (до 40 для однієї вершини).

Модуль пошарової вибірки (операція *FastGCN*) використовує обмежений набір вузлів для агрегації в кожному шарі з контрольованим коефіцієнтом розширення. У цьому разі важливі вузли із значною ймовірністю потрапляють до вибірки.

Модуль вибірки підграфів (операція *ClusterGCN*) способом кластеризації графів здійснює вибірку декількох підграфів графової моделі та пошук околів у межах обраних підграфів.

У мережах GNN для формування узагальнених характеристик використовуються модулі прямого об'єднання (субдискретизації) та модулі ієрархічного об'єднання. Модулі прямого об'єднання безпосередньо застосовують графову модель, упрощаючи різні стратегії вибору вузлів. Модулі ієрархічного об'єднання реалізують багаторівневу стратегію для GNN, що дає змогу захоплювати локальну й глобальну інформацію з вхідного графа. Ця стратегія передбачає рекурсивне застосування операції об'єднання для поступового формування сукупності підграфів з ієрархічною структурою. На кожному рівні ієрархії із вхідного графа вилучаються вкладені вершини. Далі вилучені вершини об'єднуються та передаються на наступний структурний рівень. Ця операція триває до досягнення заданої кількості рівнів з остаточним поданням вузлів мережі.

У статті запропоновано метод доповнення хмари точок, що дає змогу відновлювати відсутні фрагменти 3D-зображень. Метод використовує графові нейронні мережі глибокого навчання, які обробляють просторові дані вхідних графів з частково викривленою структурою.

Запропонований метод передбачає кодування хмари точок або її частини у форматі графа, кожна точка якого є вузлом, а ребра визначають локальні геометричні зв'язки між вузлами. Таке графове подання є вхідною інформацією для графової моделі кодера доповненої мережі, наприклад, графової

згорткової мережі (GCN) або *GraphSAGE*, яка навчається витягувати важливі ознаки та подання з графа хмари точок.

Цей підхід орієнтований на розв'язання таких завдань формування та доповнення хмар точок:

- оброблення інформації, що не має регулярної сіткової структури (на відміну від традиційних згорткових нейронних мереж графові нейронні мережі ефективно справляються з нерегулярними даними, зокрема і з викривленими хмарами точок);
- фіксація геометричних зв'язків між вершинами графової моделі (графове подання хмари точок дає змогу виявляти як локальні геометричні зв'язки через ребра, так і глобальну структурну інформацію через загальне графове подання моделі, що сприяє підвищенню якості відновлення відсутніх фрагментів);
- масштабування графових моделей (мережі GNN здатні обробляти розріджені графові подання структури моделі, що дає змогу ефективно працювати з великомасштабними хмарами точок різної щільності);
- забезпечення надійності та узагальнення (мережі GNN використовують закодоване графічне подання структурної інформації, закодованої в графі,

що забезпечує можливість моделі навчатися на надійних зразках та узагальнювати її за наявності неповних точкових даних).

Ефективність запропонованого методу оцінюватиметься за якістю доповнених хмар точок, що вимірюється такими показниками: відстань між точками, повнота та візуальна перевірка. Розглянемо загальний опис цього методу. У попередніх дослідженнях з реконструкції неповних хмар точок графові нейронні мережі застосовувалися здебільшого для оброблення закодованих особливостей моделі.

У цій роботі запропоновано безпосереднє використання графової структури точкових хмар. Синтез графової моделі здійснюється за допомогою алгоритму найближчих сусідів KNN (*K-nearest neighbors*). Зазначимо, що інші методи, зокрема побудова графа за радіусами точок, є менш стабільними та ефективними через ризик захоплення надмірної кількості зв'язків, якщо не передбачено додаткової адаптації радіуса. Натомість KNN завжди повертає чітко визначену кількість сусідів для кожної точки (рис. 3).

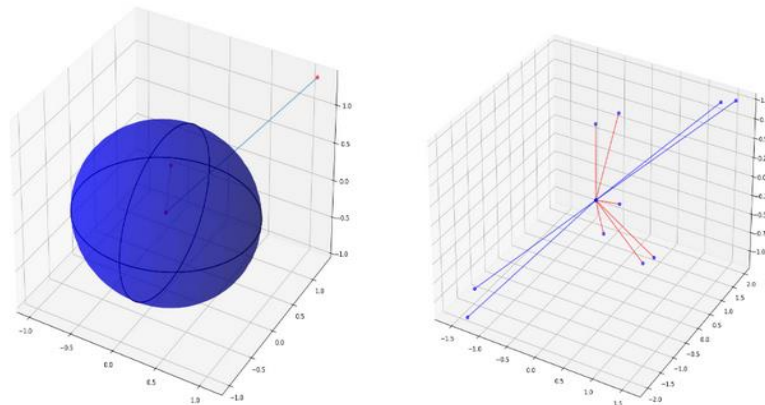


Рис. 3. Побудова графа за радіусами точок (праворуч) та відповідно до алгоритму KNN (ліворуч)

Модель кодера мережі GNN здатна в процесі навчання уточнювати структуру й локальні зв'язки в навчальних даних хмари точок. Навчена модель дає змогу використовувати фрагменти хмари точок як вхідну інформацію та генерувати відсутні точки (доповнювати модель) на основі раніше отриманих закономірностей.

Архітектура мережі, що реалізує запропонований метод, базується на мережі *Point Completion Network* (PCN), запропонованій у роботі [17], де розглянуто завдання реконструкції точкових хмар (рис. 4).

Мережа PCN містить три основні частини: кодувальну (кодувальник) *Encoder*, що здійснює глобальне подання вхідної неповної точкової хмари; мережу загального відтворення *Mapping Network*, що генерує грубу хмару точок за результатами глобального подання; декодувальну частину (декодувальник) *Decoder*, що уточнює та підвищує роздільну здатність грубої хмари точок для отримання доповненої графової моделі.

В архітектурі запропонованої графової мережі збережено загальну структуру PCN

з адаптацією її для оброблення графовими даними за допомогою кодувальника *Graph Encoder* (замість стандартного кодувальника (рис. 5).

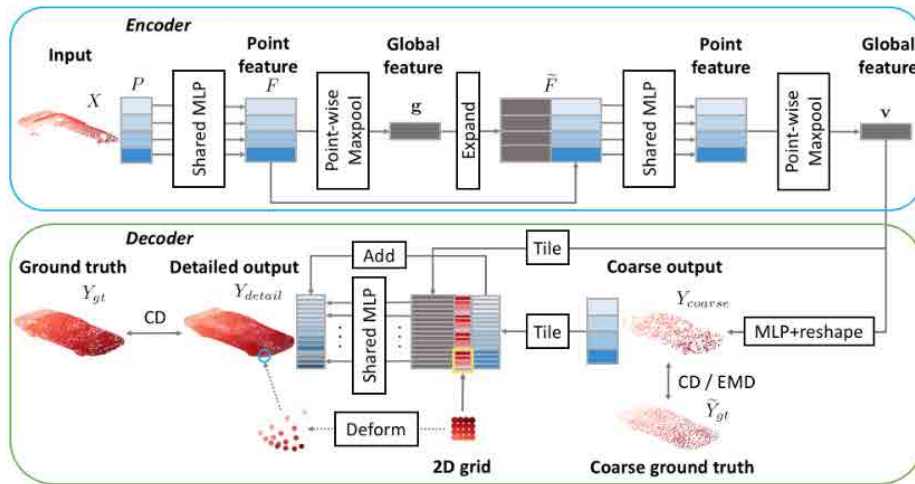


Рис. 4. Архітектура мережі доповнення точок PCN

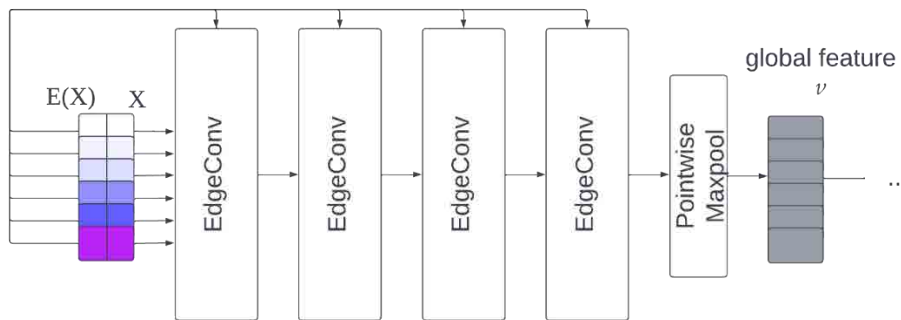


Рис. 5. Структура кодувальної частини запропонованої мережі

Основна розбіжність між кодувальником *Graph Encoder* і стандартною кодувальною частиною в мережі PCN полягає в тому, що:

- крім звичайного набору точок, кодувальник *Graph Encoder* приймає також значення індексів у побудованому графі цих точок;
- замість одновимірних операцій згортки в кодувальнику *Graph Encoder* виконуються операції графової згортки *EdgeConv*, запропоновані в роботі [40].

На відміну від операцій згортки в базових конфігураціях GCN в операціях графової згортки *EdgeConv* до зваженої агрегації даних щодо вершин графа додаються значення ребер, які з'єднують ці вершини:

$$x'_i = \sum_{j \in NN_i} h_{\Theta}(x_i \parallel x_i - x_j), \quad (18)$$

де NN_i – найближчі сусіди вершини i ; h_{Θ} – нелінійна функція з параметрами Θ .

На рис. 6 наведено приклад візуальної інтерпретації операції згортки *EdgeConv*.

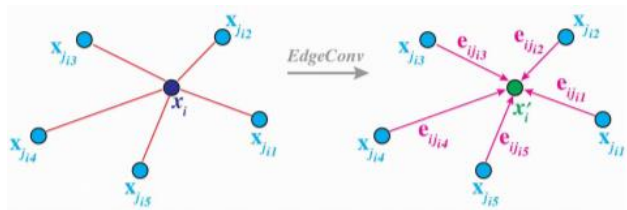


Рис. 6. Візуальна інтерпретація операції згортки *EdgeConv*

Розроблена неймережа містить кілька основних компонентів: кодувальник *Graph Encoder*, мережа загального відтворення *Mapping Network* і декодувальник *Decoder*.

Кодувальник *Graph Encoder* є ключовим елементом запропонованої модифікації нейронної мережі, відповідальним за формування глобальної репрезентації вхідного графа. Цей компонент містить шари згортки *EdgeConv*, що застосовуються

до вхідних даних для поступового розширення рецептивного поля та отримання інформації із сусідніх вершин. Під час такого розширення ознаки вершин та індекси графових ребер передаються на кожен шар для здійснення операції згортки та оновлення ознак вершин. Оновлені ознаки вершин використовуються для глобальної максимальної субдискретизації, що формує глобальне подання обробленого графа, зважаючи на дані з окремих його частин.

Глобальне подання графа отримують за допомогою *Graph Encoder*, що бере до уваги як локальні ознаки сусідів кожної вершини, так і загальну структуру графа. Аналогічне подання використовують інші компоненти мережі, зокрема *Mapping Network* і *Decoder*, для генерування точкової хмари з огляду на доповнення.

Застосування кодувальника *Graph Encoder* із шарами згортки *EdgeConvLayer* забезпечує ефективно оброблення неструктурованих даних, поданих у вигляді графа, та дає змогу отримувати найбільш інформативні ознаки для реалізації завдань графової реконструкції зображень.

Mapping Network є компонентом нейронної мережі для доповнення неповних 3D точкових хмар, що відповідає за перетворення глобального подання графа, отриманого з *Graph Encoder*, у грубу точкову хмару (*coarse point cloud*). Груба точкова хмара є проміжним поданням, що містить меншу кількість точок порівняно з остаточною доповненою точковою хмарию, але зберігає загальну структуру та форму об'єкта.

Послідовність повнозв'язних шарів *Mapping Network* виконує лінійні перетворення із застосуванням стандартних функцій активації (зазвичай *ReLU*). Груба точкова хмара, сформована *Mapping Network*, містить меншу кількість точок, якщо порівнювати з остаточною доповненою хмарию, що дає змогу зменшити обчислювальні витрати та вимоги до пам'яті для подальшого оброблення. Втім, глобальна репрезентація графа, здійснена за допомогою кодувальника *Graph Encoder*, дає змогу грубій точковій хмарі зберігати основні структурні властивості об'єкта.

Після формування грубої точкової хмари вона передається до фінального компонента *Decoder*, що збільшує кількість точок для остаточного формування нейронною мережею уточненої хмари з високою деталізацією. Згорткові шари цього декодера

виконують операції згортки вздовж ознак для кожної точки окремо (не беручи уваги інформацію від сусідніх точок). Після кожного шару згортки в декодері здійснюються операції пакетної нормалізації та активації, а також оновлюються ознаки точок із збереженням інформації з попередніх шарів.

Процес генерації уточненої точкової хмари в модулі *Decoder* ґрунтується на методі *folding*, запропонованому в дослідженні [41]. Основна ідея цього методу полягає у створенні регулярної сітки опорних точок (*folding seed*) з подальшою її деформацією за допомогою згорткових шарів для отримання остаточної точкової хмари. Такий підхід забезпечує генерацію точок із високою роздільною здатністю, зберігаючи водночас топологічні властивості об'єкта. Модуль *Decoder* відіграє центральну роль у відновленні деталізованої та повної точкової хмари на основі її початкової грубої версії. Використання згорткових шарів у поєднанні з методом *folding* дає змогу ефективно уточнювати та розширювати набір точок, зберігаючи в цьому разі геометричні особливості зображення.

Для навчання та оцінювання продуктивності нейронної мережі *GraphPointCompletionNetwork* доцільно використовувати набір даних *ShapeNet* [42]. *ShapeNet* є великомасштабним репозиторієм тривимірних моделей об'єктів, що охоплює різні категорії, зокрема літаки й автомобілі. Моделі цього набору подані сукупностями дискретних точок у тривимірному просторі, які описують геометричну структуру зображень.

Для моделювання запропонованого методу набір даних *ShapeNet* було спеціально підготовлено та поділено на два складники:

- неповні версії точкових хмар з набору *ShapeNet*, що були вхідною інформацією для навчання нейронної мережі;
- точки, що були штучно вилучені з оригінальних точкових хмар з метою створення неповних версій. Ці дані застосовуються для оцінювання якості подальшої реконструкції точкових хмар.

Для формування неповних точкових хмар з набору *ShapeNet* був розроблений спеціалізований скрипт, що реалізує такий алгоритм:

- завантаження оригінальних точкових хмар із файлів у форматі PLY;
- вибірка фіксованої кількості точок із кожної хмари для уніфікації розміру вибірки;

- формування неповних точкових хмар із застосуванням поділу вихідної хмари на дві частини за допомогою роздільної площини: точки з одного боку площини класифікуються як наявні, тоді як точки з другого боку – як відсутні;

- збереження отриманих неповних версій точкових хмар і вилучених точок у форматі *PLY*.

Для розділення точкової хмари використовується випадково згенерована гіперплощина в тривимірному просторі, що визначається нормальним вектором та зміщенням. Це дає змогу чітко визначити, по який бік площини розташована кожна точка. Крім того, було реалізовано функцію, що виконує розподіл отриманих неповних точкових хмар на тренувальну та валідаційну вибірки. Розподіл здійснюється випадковим чином відповідно до заданих пропорцій, а результати зберігаються у вигляді списків шляхів до відповідних файлів, що забезпечує можливість відтворення експерименту. Для підвищення варіативності навчальної вибірки було прийнято рішення розширити обсяг згенерованої інформації. Зокрема для кожного об'єкта створюється чотири різні гіперплощини, що дає змогу отримати різні варіанти його розділення. Такий підхід сприяє збільшенню різноманітності навчальних даних і покращенню узагальнювальної здатності моделі.

На початковому етапі моделювання запропонованого методу досліджувалися лише набори даних, що належать до класу літаків. Це допомогло ретельно опрацювати всі необхідні екземпляри та сформувати набір даних, що містить об'єкти єдиної категорії. Отже, підготовлений набір *ShapeNet* разом зі спеціалізованим скриптом для генерації неповних версій точкових хмар забезпечує ефективну підготовку інформації для навчання та оцінювання працездатності нейронної мережі *Graph Point Completion Network*. Використання випадкового поділу точкових хмар та розпаралелювання операції генерації уможливають репрезентативну множину варіантів неповних точкових хмар, що сприяє підвищенню якості навчання моделі.

Розглянемо метод навчання запропонованої графової нейронної мережі для завдання реконструкції 3D-моделей точкових хмар. Процес навчання передбачає послідовне виконання таких етапів:

- ініціалізація моделі нейронної мережі (а також параметрів оптимізатора, функції втрат і планувальника швидкості навчання);

- завантаження наборів навчальних і валідаційних даних, поданих у вигляді точкових хмар;

- аналіз наявних результатів навчання аналогічної моделі в минулому (якщо збережені результати існують, відбувається відновлення найкращих значень функції втрат і відповідної епохи навчання; завантажуються історичні значення функцій втрат для навчального й валідаційного наборів даних; відновлюється стан моделі, оптимізатора та планувальника навчання способом завантаження збережених ваг).

Навчання моделі здійснюється протягом визначеної кількості епох, кожна з яких передбачає такі основні етапи:

- прохід по навчальному набору даних, згідно з яким визначаються грубі прогнози з подальшою деталізацією;

- обчислення функцій втрат для грубого й деталізованого прогнозів, а також загальної функції втрат;

- обчислення градієнтів (на основі значення загальної функції втрат) з подальшим оновленням коефіцієнтів моделі та параметрів оптимізатора;

- прохід по валідаційному набору даних, згідно з яким (за аналогією з проходом по навчальному набору даних) обчислюються прогнозні значення та значення загальної функції втрат; на цьому етапі оновлення ваг не відбувається, оскільки метою є лише оцінювання якості моделі;

- обчислення та логування навчальних метрик для поточної епохи;

- збереження поточного стану навчання через задані проміжки часу (зокрема станів моделі, оптимізатора, планувальника швидкості навчання, а також історії значень функцій втрат);

- збереження прикладів реконструйованих точкових хмар для подальшої візуалізації навчання;

- передача збережених результатів (реконструйованих хмар точок і навчальних метрик) до відповідної бази даних;

- завершення навчання: після досягнення заданої кількості епох зберігається найкращий стан моделі для подальшого використання у завдання доповнення точкових хмар.

Такий підхід забезпечує ефективне налаштування параметрів графової нейронної мережі для доповнення точкових хмар, даючи змогу відстежувати динаміку навчання, зберігати проміжні результати та візуалізувати реконструйовані хмари. Використання

планувальника швидкості навчання (наприклад, *StepLR*) додатково підвищує ефективність навчального процесу: на початкових етапах застосовується більш висока швидкість навчання для швидкої оптимізації параметрів; на подальших етапах швидкість навчання поступово знижується, що дає змогу моделі точніше налаштуватися поблизу глобального оптимуму та уникнути небажаних коливань.

Такий механізм сприяє кращій збіжності та допомагає уникнути зупинок алгоритму навчання в локальних оптимумах. Окрім цього, можливість використання попередніх результатів навчального процесу забезпечує економію обчислювальних ресурсів, оскільки дає змогу продовжити навчання з найкращого проміжного стану, уникаючи необхідності повторювати з нуля. Це значно зменшує витрати часу та ресурсів на досягнення оптимальних результатів.

Програмна реалізація запропонованого методу

Для програмної реалізації запропонованого методу використано такі основні технології: *Python*, *PyTorch* та *PyTorch Geometric*. Застосування цих інструментів зумовлено їх ефективністю, гнучкістю та поширенням у сфері глибокого навчання та оброблення інформації. *Python* – одна з найбільш популярних мов програмування, активно впроваджена в наукових обчисленнях, аналізі даних та завданнях машинного навчання завдяки простому та зрозумілому синтаксису, а також наявності значної кількості спеціалізованих бібліотек.

PyTorch – це відкрита бібліотека, яка забезпечує гнучкий та інтуїтивно зрозумілий інтерфейс для побудови й навчання нейронних мереж, підтримує автоматичне диференціювання та ефективно використовує обчислювальні ресурси, зокрема GPU та TPU. Завдяки можливості динамічного визначення обчислювального графа *PyTorch* є зручним інструментом для дослідницьких і експериментальних проєктів, а його велика спільнота користувачів сприяє активному розвитку та підтримці бібліотеки.

PyTorch Geometric – це спеціалізована бібліотека для роботи з графовими структурами в середовищі *PyTorch*, яка містить набір оптимізованих інструментів для завантаження, оброблення й трансляції графових даних у нейронні мережі, що дає змогу реалізовувати різні архітектури графових нейронних мереж (GNNs). *PyTorch*

Geometric широко застосовується в завданнях оброблення графових структур, зокрема в 3D комп'ютерному зорі, моделюванні фізичних процесів, аналізі соціальних і біологічних мереж. Використання зазначених технологій допомагає ефективно реалізувати запропонований метод, забезпечуючи високу продуктивність, гнучкість у налаштуванні моделей та можливість масштабування обчислень.

Розглянемо особливості програмної реалізації основних компонентів графової мережі, що дають змогу виконувати окремі етапи запропонованого методу.

Одним з основних будівельних блоків кодувальної частини мережі для реконструкції неповних 3D точкових хмар є шар *EdgeConvLayer*, який реалізує операцію *EdgeConv*. Ця операція, запропонована в роботі [1], є локальним перетворенням на графових моделях, що дає змогу агрегувати інформацію з сусідніх вершин для оновлення ознак поточної вершини. Шар *EdgeConvLayer* приймає два параметри: *in_channels*, що визначає кількість ознак вершин, та *out_channels*, що задає розмірність ознак після застосування шару *EdgeConv*. У межах цього шару створюється екземпляр класу *EdgeConv* з бібліотеки *torch_geometric*. Для нелінійного перетворення h_θ використовується лінійне відтворення, реалізоване за допомогою шару *nn.Linear*, що застосовується до ознак кожної вершини та її сусідів.

Упровадження методу *forward* дає змогу реалізувати операцію *EdgeConv*, що на основі вхідних даних формує новий тензор ознак вершин розміру $(num_nodes, out_channels)$, у якому кожна вершина містить оновлену інформацію, отриману від її сусідів.

Шар *EdgeConvLayer* є ключовим компонентом кодувальної частини графової нейронної мережі, оскільки забезпечує агрегацію локальної інформації із сусідніх вершин і оновлення їх ознак з огляду на структуру графа. Використання кількох послідовних шарів *EdgeConvLayer* дає змогу збільшувати рецептивне поле й отримувати більш виразні ознаки для кожної вершини, що підвищує ефективність навчання моделі.

Кодувальник *Graph Encoder* використовується для доповнення неповних 3D точкових хмар і отримання глобального подання вхідного графа. Його структура передбачає послідовності шарів *EdgeConvLayer*, що поетапно обробляють вхідні дані, збільшуючи рецептивне поле й агрегуючи інформацію із сусідніх вершин.

Graph Encoder приймає один параметр *blocks*, що є списком цілих чисел, які визначають кількість каналів (розмірність ознак) на кожному рівні *EdgeConvLayer*.

Під час ініціалізації кодувальника *Graph Encoder* зберігається список *blocks* як атрибут класу, а також викликається метод *_build_encoder*, що створює послідовність шарів *EdgeConvLayer* згідно із заданою архітектурою. Метод *_build_encoder* приймає список *blocks*, де формується екземпляр класу *EdgeConvLayer* з відповідною кількістю вхідних і вихідних каналів, забезпечуючи ефективну побудову глибокої графової нейронної мережі.

Метод *forward* призначений для реалізації проходження крізь кодувальну мережу графа. Він використовує такі аргументи: вхідний тензор ознак вершин розміру; тензор індексів ребер розміру; тензор, що містить інформацію про належність вершин до різних графів у разі оброблення пакетів графів. Відповідно до цього методу вхідний тензор послідовно проходить крізь шари *EdgeConvLayer* кодувальника. У цьому разі шари оновлюють ознаки вершин, зважаючи на інформацію від їх сусідів. Потім застосовується операція об'єднання інформації з усіх вершин графа й формування глобального подання графа.

Подання вихідного графа, отримане за допомогою кодувальника *Graph Encoder*, є компактним і репрезентативним. Надалі це подання використовується для генерації доповненої точкової хмари іншими компонентами графової нейронної мережі, зокрема мережею відтворення *Mapping Network* та декодувальником *Decoder*. Кодувальник *Graph Encoder* з набором шарів *EdgeConvLayer* дає змогу ефективно обробляти неструктуровані дані графової моделі, що робить його зручним інструментом для аналізу та оброблення 3D точкових хмар.

У компоненті *Mapping Network*, призначеному для побудови грубих точкових хмар, використовуються такі аргументи: розмірність вхідного латентного подання графа; кількість повнозв'язних шарів у мережі; кількість точок у грубій точковій хмарі.

Ініціалізація *Mapping Network* передбачає виклик методу *_build_mappings* для створення послідовності повнозв'язних шарів на основі значень розмірності вхідного подання графа й кількості повнозв'язних шарів у мережі. Метод *_build_mappings* працює з послідовністю повнозв'язних шарів та функціями

активації *ReLU*. Останній із шарів дає змогу генерувати координати точок грубої точкової хмари.

Груба точкова хмара, згенерована з використанням *Mapping Network*, має суттєво редуковану кількість точок порівняно з повною точковою хмариною, що сприяє зменшенню обчислювальної складності подальшого оброблення. Водночас глобальне подання графа, отримане згідно з кодувальником *Graph Encoder*, дає змогу грубій точковій хмарі зберігати загальну структуру та форму 3D-об'єкта. Сформована груба точкова хмара передається до компонента *Decoder*, що збільшує кількість точок для формування остаточної доповненої точкової хмари. Операція *Mapping Network* дає змогу перейти від початкового подання графа до просторового подання у вигляді точкової хмари. Це забезпечує ефективність і якість доповнення точкових хмар на подальших етапах оброблення інформації. Функції декодувальника для нейронної мережі реалізує компонент *Decoder*, що безпосередньо доповнює неповну 3D точкову хмару. Він використовує такі атрибути: список каналів у кожному блоці; розмір сітки *grid_size* (за замовчуванням 4) та масштаб сітки *grid_scale* (за замовчуванням 0.05). Ці атрибути ініціалізуються конструктором класу *init*, що також викликає метод *_build_decoder*, який створює необхідний список блокових модулів. Кожен блок списку містить згортковий шар, шар нормалізації батча й активаційної функції *ReLU*. Процес проходження даних крізь декодувальник реалізується з використанням методу *forward*, що приймає вхідну точкову хмару, а потім формує фрагменти доповненого подання точкової хмари, які застосовуються для остаточного доповнення неповних 3D точкових хмар. Тобто він відповідає за перетворення грубої точкової хмари на більш детальну точкову хмару.

Ініціалізація повної версії запропонованої графової мережі *Graph Point Completion Network*, що використовується для доповнення неповних 3D точкових хмар. Ключовими компонентами цієї мережі є кодувальна мережа *Graph Encoder*, мережа відтворення *Mapping Network* і декодувальна мережа *Decoder*.

Ініціалізація передбачає необхідність налаштування компонентів *Graph Encoder*, *Mapping Network*, *Decoder* та обчислення тензору *folding_seed*, що надалі використовується для згортання грубої точкової хмари до остаточної точкової хмари. Програмний код ініціалізації повної версії запропонованої мережі наведено на рис. 7.

```

class GraphPointCompletionNetwork(nn.Module):
    def __init__(self, cfg: dict):
        super(GraphPointCompletionNetwork, self).__init__()
        self.num_dense: int = cfg.get('num_dense', 16384)
        self.grid_size: int = cfg.get('grid_size', 4)
        self.grid_scale: int = cfg.get('grid_scale', 0.05)
        self.num_coarse = self.num_dense // (self.grid_size ** 2)
        self.encoder = GraphEncoder(**cfg['encoder'])
        self.mapping_network = MappingNetwork(**cfg['mapping_network'], num_coarse = self.num_coarse)
        self.decoder = Decoder(**cfg['decoder'], grid_size = self.grid_size, grid_scale = self.grid_scale)
        a = (torch.linspace(-self.grid_scale, self.grid_scale, steps = self.grid_size, dtype=torch.float).view(1, self.grid_size).expand(self.grid_size, self.grid_size).reshape(1, -1))
        b = (torch.linspace(-self.grid_scale, self.grid_scale, steps=self.grid_size, dtype=torch.float).view(self.grid_size, 1).expand(self.grid_size, self.grid_size).reshape(1, -1))
        self.folding_seed = torch.cat([a, b], dim=0).view(1, 2, self.grid_size ** 2) # (1, 2, S)

```

Рис. 7. Програмний код ініціалізації повної версії запропонованої мережі

Операція ініціалізації забезпечує налаштування функціональності ключових компонентів *Graph Encoder*, *Mapping Network* і *Decoder* для отримання остаточної доповненої точкової хмари на основі вхідної неповної точкової хмари. Для формування тестувального набору даних з метою налаштування розробленої мережі застосовуватимемо клас *GraphShapeNet*, який є підкласом *Dataset* і містить набір даних *ShapeNet* у графовому поданні. Клас *GraphShapeNet* призначений для тестування завдань реконструкції зображень, пов'язаних з обробленням неповних 3D точкових хмар з використанням графових нейронних мереж.

Клас *GraphShapeNet*, що є сумісним із функціональністю класу *ShapeNet*, додатково буде за методом KNN граф k найближчих сусідів ($k=15$) для наявної точкової хмари, що зберігається

в атрибуті об'єкта *Data*. Повна точкова хмара цього об'єкта містить відповідний атрибут *y*. До основних фрагментів коду, який реалізує запропонований підхід, належить механізм навчання побудованої нейромережі. Цей механізм передбачає реалізацію таких кроків: ініціалізація даних із використанням відповідних обчислювальних засобів; формування фрагментів відтворень 3D-моделей; оброблення нейромережею поданих фрагментів відтворень 3D-моделей; порівняння грубих і повних відтворень 3D-моделей; поєднання відтворень 3D-моделей (за результатами порівняння в єдину доповнену модель). Програмний код механізму навчання графової мережі наведено на рис. 8.

Цей механізм навчання оптимізується із застосуванням оптимізатора *Adam*.

```

ground_truth = data.y.to(device)
ground_truth = ground_truth.view(batch_size, -1, 3)
coarse_predicted, dense_predicted = model(data.to(device))
num_coarse_points = coarse_predicted.shape[1]
coarse_predicted = coarse_predicted.view(batch_size, -1, 3)
dense_predicted = dense_predicted.view(batch_size, -1, 3)
coarse_loss = loss_fn(coarse_predicted, ground_truth[:, :num_coarse_points, :])
dense_loss = loss_fn(dense_predicted, ground_truth)
loss = coarse_loss + loss_multiplier * dense_loss

```

Рис. 8. Програмний код механізму навчання графової мережі

Моделювання запропонованого методу для завдань доповнення та 3D-реконструкції моделей точкових хмар

Моделювання запропонованого методу здійснювалося для реконструкції хмар точок на прикладі класу об'єктів "літак" (*airplane*) з набору даних *ShapeNet* [42]. Навчання моделі здійснювалось протягом 340 епох із використанням таких налаштувань:

- оптимізатор *Adam* з параметрами $l_r = 0.0001$, $weight_decay = 0.0001$, $betas = [0.9, 0.999]$;
- планувальник швидкості навчання *StepLR* з параметрами $step_size = 40$, $gamma = 0.8$;
- функція втрат *Chamfer Loss*.

Для тренування було використано хмарне середовище з графічним процесором *Nvidia RTX 3090 Ti* (24 GB) та 48 GB RAM. Планований час тренування становив 116 год, загальний час тренування – 30 год. На рис. 9 наведено логарифмічно масштабовані значення функцій втрат для навчальної та валідаційної вибірок залежно від кількості епох.

Результати навчання продемонстрували значення функції втрат 25.11 на навчальній вибірці та 30.21 на валідаційній вибірці. Різниця між значеннями функції втрат на навчальній та валідаційній вибірках не є дуже великою, що свідчить про відсутність

значного перенавчання моделі. Відповідно до графіка (рис. 9) можна стверджувати, що, починаючи з 200-ї епохи, нейромережна мережа досягла своїх можливостей навчання на застосованому наборі даних.

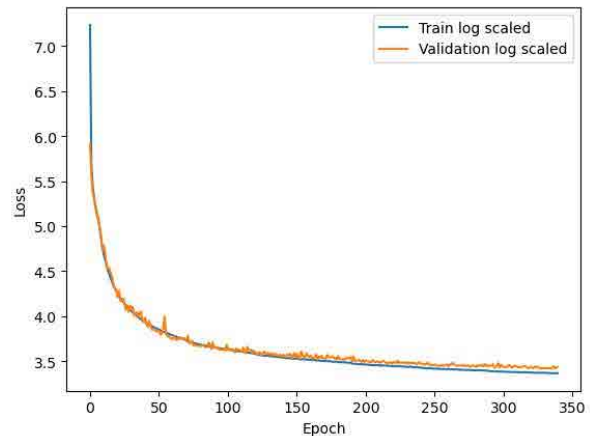


Рис. 9. Логарифмічно масштабовані значення функцій втрат навчальної та валідаційної вибірок

Стабільному процесу навчання сприяло використання оптимізатора *Adam*. Крім того, застосування планувальника швидкості навчання *StepLR* дало змогу регулювати швидкість навчання для покращення збіжності моделі. На рис. 10 наведено приклади реконструйованих хмар точок, вхідних даних та еталонних хмар точок.

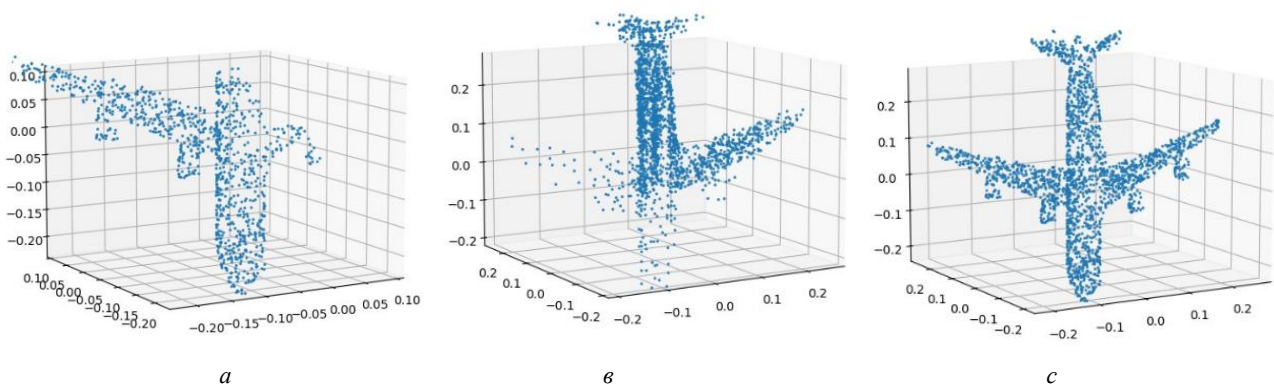


Рис. 10. Візуалізація вхідних даних (а), відновлених фрагментів моделі (в) та повної реконструкції моделі (с)

Відновлена 3D-модель літака демонструє задовільну якість доповнення порівняно з початковою неповною моделлю. Основні структурні елементи, зокрема крила та хвостова частина, були відновлені. Водночас відновлена модель має певні недоліки в деталях. Наприклад, бажаною є більш висока деталізація невеликих або дрібних елементів об'єкта. Крім того, спроба відновити менш поширені

моделі літальних апаратів (з відомих датасетів) інколи призводила до неякісних результатів доповнення (рис. 11).

Якщо порівняти отримані результати доповнення з результатами деяких інших робіт [41], [42], можна помітити відсутність чіткої деталізації дрібних елементів за рівних умов навчання та формування даних (рис. 12).

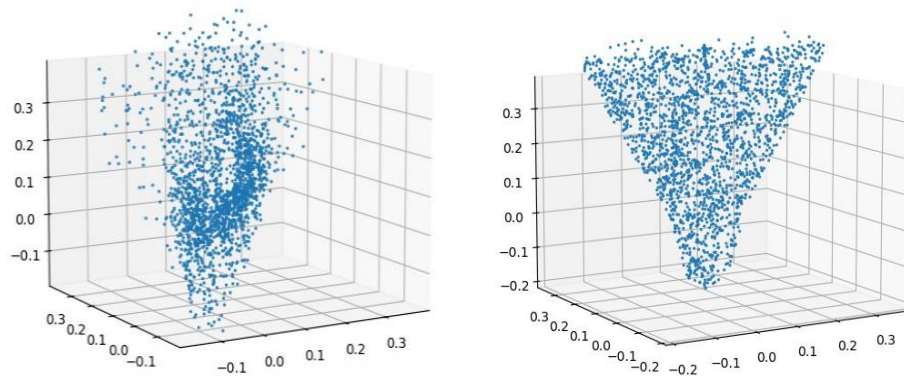


Рис. 11. Візуалізація ефекту недостатньої деталізації на малопоширеній 3D-моделі літака

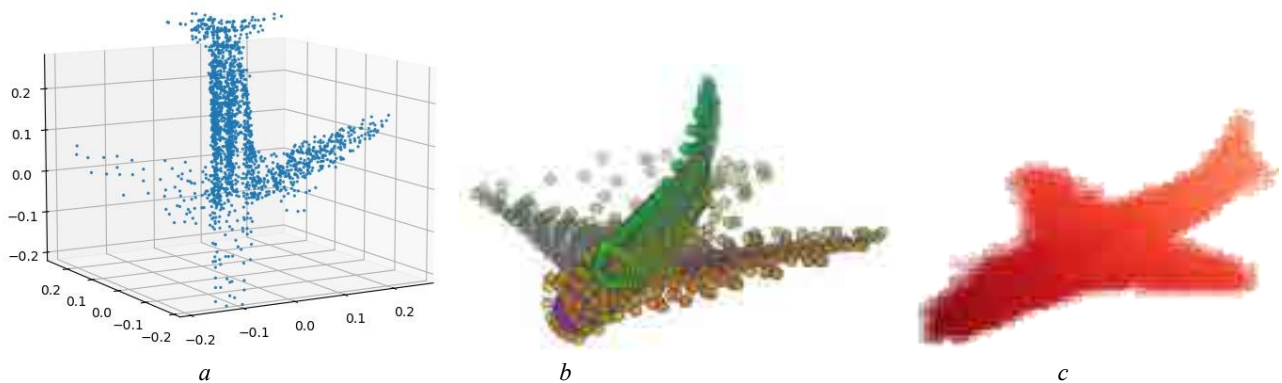


Рис. 12. Візуальне порівняння отриманих результатів для запропонованої мережі (a), мережі *FoldingNet* (b) та мережі *Point Completion Network* (c)

Важливо наголосити, що деякі сучасні модифікації мереж (зокрема *SnowflakeNet*) дають змогу інколи відновлювати об'єкти з поліпшеною деталізацією графічного подання об'єкта (рис. 13).

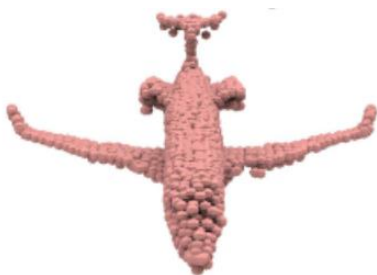


Рис. 13. Відновлення моделі літака за допомогою мережі *SnowflakeNet*

Утім, розроблений метод є першим, що призначений для розв'язання проблеми доповнення неповних 3D точкових хмар з безпосереднім використанням графових нейронних мереж без попереднього оброблення іншими засобами. Зважаючи на це, отримані результати можна вважати перспективними для досягнення поставленої в роботі мети. Зазначимо, що графове подання точкової хмари

може покращити здатність моделі виявляти та вивчати структурні залежності між точками (це може бути корисним для завдання доповнення неповних точкових хмар).

З метою покращення подальших результатів доцільним є дослідження можливості інших архітектур кодувальних мереж, які дали б змогу ефективніше частково доповнити запроповану архітектуру графової моделі. До того ж перспективним є поліпшення алгоритмів генерації самих точкових хмар (наприклад, алгоритмів подвійного спуску [43]) для підвищення якості відтвореної сцени й реконструкції дрібних деталізованих компонентів.

Більш об'єктивно оцінити узагальнювальну здатність запропонованого методу можна за результатами проведення додаткових експериментів з іншими класами об'єктів і наборів даних. Крім того, використання потужних обчислювальних ресурсів (зокрема GPU з більшим обсягом пам'яті) може сприяти навчанню більш глибоких графових нейронних моделей, що потенційно може поліпшити якість доповнення 3D-моделей хмар точок.

Висновки

У роботі запропоновано метод застосування графових нейронних мереж для доповнення 3D-моделі хмар точок.

Подане рішення для відновлення неповних 3D точкових хмар визначається науковою новизною та поєднує потужність графових нейронних мереж (GNN) з архітектурою мережі *Point Completion Network* (PCN). Це є важливим для багатьох застосувань, таких як 3D-реконструкція, навігація роботів тощо.

Автоматичне доповнення даних хмари точок надає потенціал для підвищення безпеки, надійності та продуктивності критично важливих систем, що покладаються на достовірну 3D-інформацію про навколишнє середовище.

Практичну значущість роботи підтверджують результати моделювання розробленого методу для класичних датасетів та їх порівняння з деякими наявними підходами до розв'язання досліджуваної проблеми.

Результати моделювання продемонстрували прийнятні значення функції втрат на навчальній та валідаційній вибірках, що свідчить про відсутність суттєвого ефекту перенавчання моделі.

Аналіз графіків функцій втрат і візуальне оцінювання реконструйованих хмар точок показали, що модель змогла відновити відсутні структурні елементи. Водночас спостерігалась недостатньо якісна деталізація дрібних елементів, а також низька спроможність доповнення для менш поширених моделей літаків.

Порівняння з іншими роботами в галузі доповнення хмар точок виявило подібні недоліки за аналогічних умов навчання та генерації даних. Це визначає необхідність подальшого вдосконалення архітектур графових нейронних мереж і методів навчання для підвищення якості відновлення 3D-моделей хмар точок.

Подальші дослідження доцільно зосередити на аналізі можливостей інших архітектур графових нейронних мереж, налаштуванні гіперпараметрів і застосуванні більш потужних обчислювальних ресурсів для навчання аналізованих моделей. Додаткові експерименти з іншими класами об'єктів та наборів даних також можуть допомогти більш об'єктивно оцінити узагальнювальну здатність запропонованого методу.

Загалом отримані результати демонструють потенціал графових нейронних мереж для ефективного розв'язання завдання доповнення та відновлення 3D-моделей хмар точок.

Список літератури

1. Taylor T. S. Introduction to Laser Science and Engineering. First edition. Boca Raton, FL: CRC Press/Taylor & Francis Group, 2020: CRC Press, 2019. DOI: <https://doi.org/10.1201/b22159>
2. Data-driven structural priors for shape completion / M. Sung et al. *ACM Transactions on Graphics*. 2015. Vol. 34, No. 6. P. 1–11. DOI: <https://doi.org/10.1145/2816795.2818094>
3. "State of the Art in Surface Reconstruction from Point Clouds / M. Berger et al. DSpace Repository". URL: <https://diglib.eg.org/items/d9bfd0be-9d3e-4ced-8d25-ef522dc454f3> (дата звернення: 14.05.2024).
4. A Field Model for Repairing 3D Shapes / D. T. Nguyen et al. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 27–30 June 2016. DOI: <https://doi.org/10.1109/cvpr.2016.612>
5. Fu Z., Hu W., Guo Z. Point Cloud Inpainting on Graphs from Non-Local Self-Similarity. 2018 25th *IEEE International Conference on Image Processing (ICIP)*, Athens, 7–10 October 2018. DOI: <https://doi.org/10.1109/icip.2018.8451550>
6. Mitra N. J., Guibas L. J., Pauly M. Partial and approximate symmetry detection for 3D geometry. *ACM Transactions on Graphics*. 2006. Vol. 25, No. 3. P. 560–568. DOI: <https://doi.org/10.1145/1141911.1141924>
7. Discovering structural regularity in 3D geometry / M. Pauly et al. *ACM Transactions on Graphics*. 2008. Vol. 27, No. 3. P. 1–11. DOI: <https://doi.org/10.1145/1360612.1360642>
8. An interactive approach to semantic modeling of indoor scenes with an RGBD camera / T. Shao et al. *ACM Transactions on Graphics*. 2012. Vol. 31, No. 6. P. 1–11. DOI: <https://doi.org/10.1145/2366145.2366155>
9. Structure recovery by part assembly / C.-H. Shen et al. *ACM Transactions on Graphics*. 2012. Vol. 31, No. 6. P. 1–11. DOI: <https://doi.org/10.1145/2366145.2366199>
10. A probabilistic model for component-based shape synthesis / E. Kalogerakis et al. *ACM Transactions on Graphics*. 2012. Vol. 31, No. 4. P. 1–11. DOI: <https://doi.org/10.1145/2185520.2185551>
11. Chauve A.-L., Labatut P., Pons J.-P. Robust piecewise-planar 3D reconstruction and completion from large-scale unstructured point data. 2010 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Francisco, CA, USA, 13–18 June 2010. DOI: <https://doi.org/10.1109/cvpr.2010.5539824>

12. Completing 3D object shape from one depth image / J. Rock et al. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 7–12 June 2015. DOI: <https://doi.org/10.1109/cvpr.2015.7298863>
13. Vinyals O., Fortunato M., Jaitly N. Pointer Networks. *Advances in Neural Information Processing Systems*. 2015. DOI: <https://doi.org/10.48550/arXiv.1506.03134>
14. A Papier-Mache Approach to Learning 3D Surface Generation / T. Groueix et al. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, 18–23 June 2018. DOI: <https://doi.org/10.1109/cvpr.2018.00030>
15. Dai A., Qi C. R., NieBner M. Shape Completion Using 3D-Encoder-Predictor CNNs and Shape Synthesis. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 21–26 July 2017. DOI: <https://doi.org/10.1109/cvpr.2017.693>
16. Wang X., Ang M. H., Hee Lee G. Voxel-based Network for Shape Completion by Leveraging Edge Generation. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, 10–17 October 2021. 2021. DOI: <https://doi.org/10.1109/iccv48922.2021.01294>
17. PCN: Point Completion Network / W. Yuan et al. 2018 *International Conference on 3D Vision (3DV)*, Verona, 5–8 September 2018. DOI: <https://doi.org/10.1109/3dv.2018.00088>
18. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation / R. Q. Charles et al. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 21–26 July 2017. DOI: <https://doi.org/10.1109/cvpr.2017.16>
19. Morphing and Sampling Network for Dense Point Cloud Completion / M. Liu et al. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2020. Vol. 34, No. 07. P. 11596–11603. DOI: <https://doi.org/10.1609/aaai.v34i07.6827>
20. Lawrance N. R. J., Chung J. J., Hollinger G. A. Fast Marching Adaptive Sampling. *IEEE Robotics and Automation Letters*. 2017. Vol. 2, No. 2. P. 696–703. DOI: <https://doi.org/10.1109/lra.2017.2651148>
21. Chen X., Chen B., Mitra N. J. Unpaired Point Cloud Completion on Real Scans using Adversarial Training. *International Conference On Learning Representations (ICLR 2020)*. DOI: <https://doi.org/10.48550/arXiv.1904.00069>
22. "Generative Adversarial Networks / I. J. Goodfellow et al. *Advances in Neural Information Processing Systems*". 2014. URL: https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf. (дата звернення: 15.05.2024).
23. PF-Net: Point Fractal Network for 3D Point Cloud Completion / Z. Huang et al. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, 13–19 June 2020. DOI: <https://doi.org/10.1109/cvpr42600.2020.00768>
24. Refinement of Predicted Missing Parts Enhance Point Cloud Completion / A. Mendoza et al. *Computer Vision and Pattern Recognition*. 11 p. 2020. DOI: <https://doi.org/10.48550/arXiv.2010.04278>
25. HyperPocket: Generative Point Cloud Completion / P. Spurek et al. 2022 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Kyoto, Japan, 23–27 October 2022. 2022. DOI: <https://doi.org/10.1109/iro47612.2022.9981829>
26. Ha D., Dai A., Le Q. V. Hypernetworks. *International Conference on Learning Representations*. *Machine Learning*. 2022. DOI: <https://doi.org/10.48550/arXiv.1609.09106>
27. SnowflakeNet: Point Cloud Completion by Snowflake Point Deconvolution with Skip-Transformer / P. Xiang et al. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, 10–17 October 2021. 2021. DOI: <https://doi.org/10.1109/iccv48922.2021.00545>
28. Pan L. ECG: Edge-aware Point Cloud Completion with Graph Convolution. *IEEE Robotics and Automation Letters*. 2020. Vol. 5, No. 3. P. 4392–4398. DOI: <https://doi.org/10.1109/lra.2020.2994483>
29. Kullback S., Leibler R. A. On Information and Sufficiency. *The Annals of Mathematical Statistics*. 1951. Vol. 22, No. 1. P. 79–86. DOI: <https://doi.org/10.1214/aoms/1177729694>
30. Graph Convolutional Networks with Markov Random Field Reasoning for Social Spammer Detection / Y. Wu et al. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2020. Vol. 34, No. 01. P. 1054–1061. DOI: <https://doi.org/10.1609/aaai.v34i01.5455>
31. Gradient-based learning applied to document recognition / Y. Lecun et al. *Proceedings of the IEEE*. 1998. Vol. 86, No. 11. P. 2278–2324. DOI: <https://doi.org/10.1109/5.726791>
32. LeCun Y., Bengio Y., Hinton G. Deep learning. *Nature*. 2015. Vol. 521, No. 7553. *PubMed*. P. 436–444. DOI: <https://doi.org/10.1038/nature14539>
33. What is a good medical decision? A research agenda guided by perspectives from multiple stakeholders / J. G. Hamilton et al. *Journal of Behavioral Medicine*. 2016. Vol. 40, No. 1. P. 52–68. DOI: <https://doi.org/10.1007/s10865-016-9785-z>
34. Efficient Estimation of Word Representations in Vector Space / T. Mikolov et al. *International Conference on Learning Representations*. 2013. DOI: <https://doi.org/10.48550/arXiv.1301.3781>
35. Perozzi B., Al-Rfou R., Skiena S. S. DeepWalk: online learning of social representations. *KDD '14: Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. 20th ed. 2014. DOI: <https://doi.org/10.1145/2623330.2623732>
36. Mallat S. Wavelet Tour of Signal Processing. Elsevier Science & Technology Books, 1999.
37. Hammond D. K., Vandergheynst P., Gribonval R. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*. 2011. Vol. 30, no. 2. P. 129–150. DOI: <https://doi.org/10.1016/j.acha.2010.04.005>

38. Kipf T. N., Welling M. Semi-Supervised Classification with Graph Convolutional Networks. *Machine Learning*. 2016. DOI: <https://doi.org/10.48550/arXiv.1609.02907>
39. Hamilton W. L., Ying R., Leskovec J. Inductive Representation Learning on Large Graphs. *Social and Information Networks*. 2017. DOI: <https://doi.org/10.48550/arXiv.1706.02216>
40. Dynamic Graph CNN for Learning on Point Clouds / Y. Wang et al. *ACM Transactions on Graphics*. 2019. Vol. 38, No. 5. P. 1–12. DOI: <https://doi.org/10.1145/3326362>
41. FoldingNet: Point Cloud Auto-Encoder via Deep Grid Deformation / Y. Yang et al. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, 18–23 June 2018. DOI: <https://doi.org/10.1109/cvpr.2018.00029>
42. Orthogonal Dictionary Guided Shape Completion Network for Point Cloud / P. Cai et al. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2024. Vol. 38, No. 2. P. 864–872. DOI: <https://doi.org/10.1609/aaai.v38i2.27845>
43. Nakkiran, P., Kaplun, G., Bansal, Y., Yang, T., Barak, B., and Sutskever, I. Deep double descent: Where bigger models and more data hurt. *Machine Learning*. 2019. DOI: <https://doi.org/10.48550/arXiv.1912.02292>

References

1. Taylor, T. S. (2020), "Introduction to Laser Science and Engineering". First edition. Boca Raton, FL: CRC Press/Taylor & Francis Group. DOI: <https://doi.org/10.1201/b22159>
2. Sung, M. (2015), "Data-driven structural priors for shape completion" / M. Sung et al. *ACM Transactions on Graphics*. 2015. Vol. 34, No. 6. P. 1–11. DOI: <https://doi.org/10.1145/2816795.2818094>
3. "State of the Art in Surface Reconstruction from Point Clouds" / M. Berger et al. DSpace Repository". URL: <https://diglib.eg.org/items/d9bfd0be-9d3e-4ced-8d25-ef522dc454f3> (last accessed: 14.05.2024).
4. Nguyen, D. T. (2016), "A Field Model for Repairing 3D Shapes" / D. T. Nguyen et al. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 27–30 June 2016. DOI: <https://doi.org/10.1109/cvpr.2016.612>
5. Fu, Z., Hu, W., Guo, Z. (2018), "Point Cloud Inpainting on Graphs from Non-Local Self-Similarity". 2018 25th *IEEE International Conference on Image Processing (ICIP)*, Athens, 7–10 October 2018. DOI: <https://doi.org/10.1109/icip.2018.8451550>
6. Mitra, N. J., Guibas, L. J., Pauly, M. (2006), "Partial and approximate symmetry detection for 3D geometry". *ACM Transactions on Graphics*. 2006. Vol. 25, No. 3. P. 560–568. DOI: <https://doi.org/10.1145/1141911.1141924>
7. Pauly, M. (2008), "Discovering structural regularity in 3D geometry" / M. Pauly et al. *ACM Transactions on Graphics*. 2008. Vol. 27, No. 3. P. 1–11. DOI: <https://doi.org/10.1145/1360612.1360642>
8. Shao, T. (2012), "An interactive approach to semantic modeling of indoor scenes with an RGBD camera" / T. Shao et al. *ACM Transactions on Graphics*. 2012. Vol. 31, No. 6. P. 1–11. DOI: <https://doi.org/10.1145/2366145.2366155>
9. Shen, C.H. (2012), "Structure recovery by part assembly" / C.-H. Shen et al. *ACM Transactions on Graphics*. 2012. Vol. 31, No. 6. P. 1–11. DOI: <https://doi.org/10.1145/2366145.2366199>
10. Kalogerakis, E. (2012), "A probabilistic model for component-based shape synthesis" / E. Kalogerakis et al. *ACM Transactions on Graphics*. 2012. Vol. 31, No. 4. P. 1–11. DOI: <https://doi.org/10.1145/2185520.2185551>
11. Chauve, A.L., Labatut, P., Pons, J.P. (2010), "Robust piecewise-planar 3D reconstruction and completion from large-scale unstructured point data". 2010 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Francisco, CA, USA, 13–18 June 2010. DOI: <https://doi.org/10.1109/cvpr.2010.5539824>
12. Rock, J. (2015), "Completing 3D object shape from one depth image" / J. Rock et al. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 7–12 June 2015. DOI: <https://doi.org/10.1109/cvpr.2015.7298863>
13. Vinyals, O., Fortunato, M., Jaitly, N. (2015), "Pointer Networks". *Advances in Neural Information Processing Systems*. 2015. DOI: <https://doi.org/10.48550/arXiv.1506.03134>
14. Groueix, T. (2018), "A Papier-Mache Approach to Learning 3D Surface Generation" / T. Groueix et al. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, 18–23 June 2018. DOI: <https://doi.org/10.1109/cvpr.2018.00030>
15. Dai, A., Qi, C. R., NieBner, M. (2017), "Shape Completion Using 3D-Encoder-Predictor CNNs and Shape Synthesis". 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 21–26 July 2017. DOI: <https://doi.org/10.1109/cvpr.2017.693>
16. Wang, X., Ang, M. H., Hee Lee, G. (2021), "Voxel-based Network for Shape Completion by Leveraging Edge Generation". 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, 10–17 October 2021. 2021. DOI: <https://doi.org/10.1109/iccv48922.2021.01294>
17. Yuan, W. (2018), "PCN: Point Completion Network" / W. Yuan et al. 2018 *International Conference on 3D Vision (3DV)*, Verona, 5–8 September 2018. DOI: <https://doi.org/10.1109/3dv.2018.00088>
18. Charles, R. Q. (2017), "PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation" / R. Q. Charles et al. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 21–26 July 2017. DOI: <https://doi.org/10.1109/cvpr.2017.16>
19. Liu, M. (2020), "Morphing and Sampling Network for Dense Point Cloud Completion" / M. Liu et al. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2020. Vol. 34, No. 07. P. 11596–11603. DOI: <https://doi.org/10.1609/aaai.v34i07.6827>

20. Lawrance, N. R. J., Chung J. J., Hollinger G. A. (2017), "Fast Marching Adaptive Sampling". *IEEE Robotics and Automation Letters*. 2017. Vol. 2, No. 2. P. 696–703. DOI: <https://doi.org/10.1109/lra.2017.2651148>
21. Chen, X., Chen, B., Mitra, N. J. (2020), "Unpaired Point Cloud Completion on Real Scans using Adversarial Training". *International Conference On Learning Representations (ICLR 2020)*. DOI: <https://doi.org/10.48550/arXiv.1904.00069>
22. "Generative Adversarial Networks / I. J. Goodfellow et al. *Advances in Neural Information Processing Systems*". 2014. URL: https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf. (last accessed: 15.05.2024).
23. Huang, Z. (2020), "PF-Net: Point Fractal Network for 3D Point Cloud Completion" / Z. Huang et al. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, 13–19 June 2020. DOI: <https://doi.org/10.1109/cvpr42600.2020.00768>
24. Mendoza, A. (2020), "Refinement of Predicted Missing Parts Enhance Point Cloud Completion" / A. Mendoza et al. *Computer Vision and Pattern Recognition*. 11 p. DOI: <https://doi.org/10.48550/arXiv.2010.04278>
25. Spurek, P. (2022), "HyperPocket: Generative Point Cloud Completion" / P. Spurek et al. 2022 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Kyoto, Japan, 23–27 October 2022. DOI: <https://doi.org/10.1109/iros47612.2022.9981829>
26. Ha, D., Dai, A., Le, Q. V. (2022), "Hypernetworks. International Conference on Learning Representations". *Machine Learning*. 2022. DOI: <https://doi.org/10.48550/arXiv.1609.09106>
27. Xiang, P. (2021), "SnowflakeNet: Point Cloud Completion by Snowflake Point Deconvolution with Skip-Transformer" / P. Xiang et al. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, 10–17 October 2021. DOI: <https://doi.org/10.1109/iccv48922.2021.00545>
28. Pan, L. (2020), "ECG: Edge-aware Point Cloud Completion with Graph Convolution". *IEEE Robotics and Automation Letters*. 2020. Vol. 5, No. 3. P. 4392–4398. DOI: <https://doi.org/10.1109/lra.2020.2994483>
29. Kullback, S., Leibler, R. A. (1951), "On Information and Sufficiency". *The Annals of Mathematical Statistics*. 1951. Vol. 22, No. 1. P. 79–86. DOI: <https://doi.org/10.1214/aoms/1177729694>
30. Wu, Y. (2020), "Graph Convolutional Networks with Markov Random Field Reasoning for Social Spammer Detection" / Y. Wu et al. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2020. Vol. 34, No. 01. P. 1054–1061. DOI: <https://doi.org/10.1609/aaai.v34i01.5455>
31. Lecun, Y. (1998), "Gradient-based learning applied to document recognition" / Y. Lecun et al. *Proceedings of the IEEE*. 1998. Vol. 86, No. 11. P. 2278–2324. DOI: <https://doi.org/10.1109/5.726791>
32. LeCun, Y., Bengio, Y., Hinton, G. (2015), "Deep learning. *Nature*". Vol. 521, No. 7553. *PubMed*. P. 436–444. DOI: <https://doi.org/10.1038/nature14539>
33. Hamilton, J. G. (2016), "What is a good medical decision? A research agenda guided by perspectives from multiple stakeholders" / J. G. Hamilton et al. *Journal of Behavioral Medicine*. 2016. Vol. 40, No. 1. P. 52–68. DOI: <https://doi.org/10.1007/s10865-016-9785-z>
34. Mikolov, T. (2013), "Efficient Estimation of Word Representations in Vector Space" / T. Mikolov et al. *International Conference on Learning Representations*. 2013. DOI: <https://doi.org/10.48550/arXiv.1301.3781>
35. Perozzi, B., Al-Rfou, R., Skiena, S. S. (2014), "DeepWalk: online learning of social representations". *KDD '14: Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. 20th ed. 2014. DOI: <https://doi.org/10.1145/2623330.2623732>
36. Mallat, S. *Wavelet Tour of Signal Processing*. Elsevier Science & Technology Books, 1999.
37. Hammond, D. K., Vandergheynst, P., Gribonval, R. (2011), "Wavelets on graphs via spectral graph theory". *Applied and Computational Harmonic Analysis*. 2011. Vol. 30, No. 2. P. 129–150. DOI: <https://doi.org/10.1016/j.acha.2010.04.005>
38. Kipf, T. N., Welling, M. (2016), "Semi-Supervised Classification with Graph Convolutional Networks". *Machine Learning*. 2016. DOI: <https://doi.org/10.48550/arXiv.1609.02907>
39. Hamilton, W. L., Ying, R., Leskovec, J. (2017), "Inductive Representation Learning on Large Graphs". *Social and Information Networks*. 2017. DOI: <https://doi.org/10.48550/arXiv.1706.02216>
40. Wang, Y. (2019), "Dynamic Graph CNN for Learning on Point Clouds" / Y. Wang et al. *ACM Transactions on Graphics*. 2019. Vol. 38, No. 5. P. 1–12. DOI: <https://doi.org/10.1145/3326362>
41. Yang, Y. (2018), "FoldingNet: Point Cloud Auto-Encoder via Deep Grid Deformation" / Y. Yang et al. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, 18–23 June 2018. DOI: <https://doi.org/10.1109/cvpr.2018.00029>
42. Cai, P. (2024), "Orthogonal Dictionary Guided Shape Completion Network for Point Cloud" / P. Cai et al. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2024. Vol. 38, No. 2. P. 864–872. DOI: <https://doi.org/10.1609/aaai.v38i2.27845>
43. Nakkiran, P., Kaplun, G., Bansal, Y., Yang, T., Barak, B., and Sutskever, I. (2019), "Deep double descent: Where bigger models and more data hurt". *Machine Learning*. DOI: <https://doi.org/10.48550/arXiv.1912.02292>

Відомості про авторів / About the Authors

Чухран Іван Дмитрович – Харківський національний університет радіоелектроніки, студент кафедри штучного інтелекту, Харків, Україна; e-mail: ivan.chukhran@nure.ua; ORCID ID: <https://orcid.org/0009-0004-7720-6979>

Удовенко Сергій Григорович – доктор технічних наук, професор, Харківський національний економічний університет ім. С. Кузнеця, завідувач кафедри інформатики та обчислювальної техніки, e-mail: serhiy.udovenko@hneu.net; ORCID ID: <https://orcid.org/0000-0001-5945-8647>

Шергін Вадим Леонідович – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри штучного інтелекту, Харків, Україна; e-mail: vadim.shergin@nure.ua; ORCID ID: <https://orcid.org/0000-0002-4388-8180>

Чала Лариса Ернестівна – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри штучного інтелекту, Харків, Україна; e-mail: larysa.chala@nure.ua; ORCID ID: <https://orcid.org/0000-0002-9890-4790>

Chukhran Ivan – Kharkiv National University of Radio Electronics, Student of the Artificial Department, Kharkiv, Ukraine.

Udovenko Serhii – Doctor of Sciences (Engineering), Professor, Head at the Department of Informatics and Computer Technics of S. Kuznets Kharkiv National University of Economics, Kharkiv, Ukraine.

Shergin Vadim – PhD (Engineering Sciences), Associate Professor, Kharkiv National University of Radio Electronics, Associate Professor of the Artificial Department, Kharkiv, Ukraine.

Chala Larysa – PhD (Engineering Sciences), Associate Professor, Kharkiv National University of Radio Electronics, Associate Professor of the Artificial Department, Kharkiv, Ukraine.

METHOD FOR AUGMENTING 3D POINT CLOUD MODELS USING GRAPH NEURAL NETWORKS

Constructing models representing three-dimensional objects as a set of unstructured points (3D point cloud models) in space is becoming increasingly widespread across various domains, including autonomous navigation, robotics, virtual/augmented reality, and 3D reconstruction. Accurately capturing and processing 3D point cloud data is critical for applications requiring a comprehensive understanding of the surrounding environment, such as obstacle avoidance, path planning, and scene modelling. However, due to various reasons, point clouds often contain missing regions, posing significant challenges for subsequent data processing. Incomplete point cloud data can have serious consequences, for instance, in autonomous navigation systems, where errors may lead to collisions or other hazardous situations. Addressing this issue is crucial for the reliable processing of 3D data. This work aims to develop and investigate a method for automatically completing and reconstructing point clouds using graph neural networks. The study's primary objectives include analysing existing approaches to constructing and restoring three-dimensional graph models, developing and implementing a method for automatically completing point clouds using graph neural networks and modelling the proposed method for tasks related to the completion and 3D reconstruction of point cloud models. In this work, a conceptual model for point cloud completion was developed using graph neural networks, enabling the efficient encoding of incomplete point clouds as graphs and the prediction of missing points. The proposed solution for completing incomplete 3D point clouds offers scientific novelty and combines the power of graph neural networks (GNN) with the Point Completion Network (PCN) architecture. The suggested approach allows for high-quality restoration of incomplete 3D data, essential for numerous applications, such as 3D reconstruction, robot navigation, and more. The practical significance of the work's results is validated by the modelling outcomes of the developed method on classical datasets and their comparison with existing approaches to solving the studied problem. A promising direction for further research on this topic includes testing various architectures of graph neural networks, tuning hyperparameters, applying alternative loss functions, and leveraging more powerful computational resources to train the constructed neural network models.

Keywords: point cloud; deep learning; graph neural network; automatic point cloud completion; graph encoding; 3D reconstruction of graph models.

Бібліографічні описи / Bibliographic descriptions

Чухран І. Д., Удовенко С. Г., Шергін В. Л., Чала Л. Е. Метод доповнення 3D-моделей точкових хмар з використанням графових нейронних мереж. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 2 (32). С. 129–150. DOI: <https://doi.org/10.30837/2522-9818.2025.2.129>

Chukhran, I., Udovenko, S., Shergin, V., Chala, L. (2025), "Method for augmenting 3D point cloud models using graph neural networks", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (32), P. 129–150. DOI: <https://doi.org/10.30837/2522-9818.2025.2.129>

В. ШКУРКО А. ПОЛЯКОВ

ОРГАНІЗАЦІЯ ПРОГРАМНИХ І НЕЙРОМЕРЕЖЕВИХ АЛГОРИТМІВ МАШИННОГО АНАЛІЗУ ТЕКСТОВИХ ПОВІДОМЛЕНЬ, ПОДАНИХ ПРИРОДНОЮ МОВОЮ

У статті розглянуто питання організації програмних і нейромережових алгоритмів машинного аналізу текстових даних, поданих природною мовою. Обґрунтовано актуальність завдання оброблення як стислих текстових повідомлень і відгуків, що потребують швидкого оброблення з мінімальними ресурсними витратами, так і складних структурованих документів, які вимагають збереження структурних характеристик і глибокого контекстного аналізу. Проведено комплексний аналіз сучасних методів машинного оброблення текстової інформації, зокрема токенізації, кластеризації, семантико-релевантного пошуку й застосування нейромережових архітектур. Особливу увагу приділено підходам, що дають змогу оптимізувати обчислювальні витрати без суттєвого зниження якості результатів аналізу, що є критично важливим для роботи в умовах обмежених ресурсів. На основі аналізу розроблено багаторівневу методику організації машинного аналізу текстових даних. Методика передбачає попередню класифікацію текстових масивів за типами документів, групування текстів методом кластеризації для підвищення релевантності оброблення та застосування нейромережових моделей глибокого навчання. Для глибокого аналізу текстової інформації реалізовано архітектуру на основі двонаправленої рекурентної нейронної мережі (*Bidirectional LSTM*) із використанням регуляризації *Dropout* та механізмів раннього припинення навчання. З метою практичної перевірки запропонованої методики розроблено застосунок для автоматизованого аналізу стислих текстових повідомлень природною мовою. Подано результати навчання моделі, побудовано графіки динаміки зміни функції втрат на тренувальних і валідаційних вибірках, розроблено матриці помилок та візуалізацію результатів прогнозування. Продемонстровано стабільне зниження функції втрат без суттєвого збільшення обчислювальних витрат системи. Запропонована методика може бути застосована в інформаційних системах різного призначення для автоматизованого оброблення текстових повідомлень у режимах з обмеженими ресурсами, а також має перспективи подальшого розвитку в напрямі аналізу мультимодальних даних і впровадження в реальні інформаційно-аналітичні комплекси.

Ключові слова: машинний аналіз; текстові дані; природна мова; програмні алгоритми; нейромережева архітектура.

Вступ

Розвиток інформаційних технологій [1], упровадження нейромережових алгоритмів глибокого навчання [2, 3] та поширення концепції *Big Data* [4, 5] дали змогу автоматизувати оброблення великих масивів текстових даних, поданих природною мовою. Застосування програмних і нейромережових алгоритмів машинного аналізу текстів стає важливим завданням як для оброблення складних структурованих документів (наукових, технічних, юридичних текстів), так і для стислих текстових повідомлень і відгуків, що потребують швидкого оброблення в умовах обмежених обчислювальних ресурсів.

Оброблення невеликих текстів потребує уваги до таких особливостей: обмеженого обсягу інформації, високої варіативності лексичних конструкцій та необхідності збереження контексту за мінімальної довжини вхідних послідовностей. А для аналізу складних документів важливо забезпечити оброблення структурних і термінологічних особливостей тексту.

Сучасні системи машинного аналізу текстових даних мають бути адаптивними до різних форматів вхідної інформації – від лаконічних повідомлень до великих структурованих текстів – з огляду на обмеження обчислювальних ресурсів.

Зазначимо, що сучасні апаратно-програмні платформи центрів оброблення даних та мережеві сервіси на їх основі надають широкий інструментарій для автоматизованої роботи з великими текстовими масивами в галузях, що цілком охоплюють життєдіяльність сучасної людини.

1. Соціальні мережі та пошукові системи, що передбачають семантичний аналіз запитів користувачів, виявлення громадської думки й актуальних трендів, індексацію та класифікацію контенту.

2. Лінгвістичні та філологічні дослідження, що передбачають автоматичний переклад текстів, аналіз стилю та авторства, а також лексикографію та семантичний аналіз.

3. Наука, інженерія, медицина й навчання, що передбачають систематизацію знань і автоматизацію рецензування, формування адаптивних систем

персоналізованого навчання, виявлення симптомів і автоматизацію діагностики, а також управління проектами й технічну підтримку.

4. Державне управління та судова практика, що передбачає моніторинг і оцінювання впровадження політик, визначення громадських потреб та пріоритетів, виявлення прецедентів і правових аргументів.

5. Фінансовий аналіз, електронна комерція та маркетинг, що передбачає аналіз і прогнозування

ринкових трендів, вивчення поведінкових патернів споживачів, а також персоналізацію рекомендацій.

Багаторівнева класифікація, що може бути побудована на основі відповідного переліку (рис. 1), дає змогу надалі оцінити актуальні завдання щодо побудови цілісної методології організації та адаптації програмних і нейромережових алгоритмів машинного аналізу текстових даних, поданих природною мовою.



Рис. 1. Багаторівнева класифікація актуальних галузей застосування засобів машинного аналізу текстових даних

Як показав аналіз наукових досліджень, присвячених проблемам упровадження систем автоматизації процедури аналізу великих масивів

текстових даних, організація та адаптація відповідних програмних і нейромережових алгоритмів є комплексним завданням [6–10]. Дослідники зазначають,

що в розробленні математичної моделі та нейромережевої архітектури, що може лягти в основу алгоритмів машинного аналізу текстових даних, важливо визначити ресурс пам'яті, обчислювальний ресурс і перепускність інформаційних каналів апаратної платформи [6], а отже, в основі проєктування має лежати компроміс між продуктивністю інформаційної системи та рівнем витрат. У впровадженні нейромережевих алгоритмів також мають бути оцінені затрати на навчання великих мовних моделей (*Large Language Models*; LLM) з пріоритизацією важливості взаємодії між відповідними алгоритмами та наявним програмним забезпеченням апаратної платформи [6, 7] згідно з концепцією попередньо тренованої мовної моделі (*Pre-Trained Language Model*; PLM).

Зі свого боку ефективність процедури навчання штучної нейромережі (*Artificial Neural Network*; ANN) основана на підходах, що використовуються в дослідженні методів моделювання мови [7]. Завдання полягає в побудові алгоритмів, здатних із мінімальною затримкою виконувати поточні запити з необхідним рівнем точності, а також піддаватись масштабуванню, адаптації та налаштуванню [6, 7] за умов розширення поставленого набору завдань. Зазначимо, що оптимізація відповідних нейромережевих алгоритмів щодо збільшення продуктивності та зменшення ресурсоемності передбачає низку сучасних підходів, що лежать в основі концепції крайового машинного навчання (*Edge Machine Learning*; EML), зокрема низькорозрядна квантизація (*Low-Bit Quantization*), параметроефективне донавчання (*Parameter-Efficient Fine-Tuning*; PEFT), а також аналіз гіперпараметрів [8, 11]. Крім того, постійне оновлення PLM відповідно до актуальних даних і подолання таких обмежень, як погана інтерпретованість результатів машинного аналізу й потреба в підготовці великих масивів анотованих даних [9], пропонується виконати за допомогою формування нейромережевої архітектури мовної моделі з попереднім навчанням та підсиленням на основі зовнішньої бази знань (*Knowledge-Enhanced Pre-trained Language Models*; KEPLMs) з визначенням методів адаптації ANN [9, 12]. Нарешті, основна увага в галузі машинного аналізу текстової інформації, поданої природною мовою, приділяється моделям трансформерів [10], зокрема таким,

що лежать в основі архітектури генеративного трансформера попереднього навчання (*Generative Pre-Trained Transformer*; GPT) та використовується в онлайн-сервісах на базі *ChatGPT*. Підтверджено ефективність застосування моделей змішаної точності (*Mixed-Precision Model*; MPM) для зменшення застосування ресурсу пам'яті, а також збільшення точності та швидкості оброблення вхідних запитів [10, 12–14]. Необхідно зазначити, що з метою адаптації LLM для виконання практичних завдань, зокрема аналізу фондових ринків, необхідно розширити інструментарій нейромережевого аналізу за допомогою впровадження в межах інформаційної системи різних форматів даних, що є нерозв'язаною частиною загального дослідження.

Окремо варто наголосити, що в багатьох випадках використання повноцінних обчислювальних потужностей для машинного аналізу текстових повідомлень є недоступним через обмеження ресурсів кінцевих пристроїв або інфраструктури. Тому особлива увага в дослідженні приділяється підходам до розроблення систем, здатних працювати в середовищах з низькою обчислювальною потужністю без суттєвого зниження точності аналізу.

Мета роботи полягає в розробленні комплексної методики організації програмних і нейромережевих алгоритмів машинного аналізу текстових даних природною мовою з огляду на обмеження ресурсів та особливості оброблення різних форматів текстової інформації.

Для досягнення поставленої мети необхідно розв'язати такі **завдання**:

- проаналізувати сучасні методи машинного аналізу текстових даних, поданих природною мовою, зважаючи на їх переваги та обмеження;
- розробити концепцію семантико-релевантного пошуку для оброблення стислих і великих текстів;
- побудувати узагальнену схему організації інформаційної системи машинного аналізу текстів;
- визначити особливості адаптації програмних алгоритмів для стислих текстових повідомлень;
- дослідити можливості застосування нейромережевих архітектур для глибокого оброблення текстів;
- мінімізувати обчислювальні витрати інформаційної системи без суттєвого зниження якості аналізу текстових даних.

1. Постановка завдання машинного аналізу масивів текстових даних, поданих природною мовою

Зростання обсягів текстових повідомлень природною мовою, що потребують автоматизованого оброблення, зумовлює необхідність створення систем машинного аналізу тексту. Текстові масиви містять як стислі повідомлення (відгуки, запити, коментарі), так і структуровані документи (наукові публікації, технічну, юридичну та фінансову документацію, новинні статті), кожен з яких має особливості подання та вимоги до оброблення.

Постановка завдання машинного аналізу текстових даних полягає в побудові системи, що забезпечує:

- оброблення великих масивів текстів з огляду на їх контекст і семантичні зв'язки;
- підтримку оброблення як стислих, так і складних текстових форматів;
- оптимізацію обчислювальних ресурсів у процесі збереження точності результатів аналізу;

– мінімізацію часу оброблення та обсягу пам'яті, необхідної для роботи із текстовими масивами.

Основні складнощі розв'язання завдання пов'язані з експоненційним зростанням обсягів текстових повідомлень, необхідністю уваги до супутніх контекстних характеристик текстів та потребою підтримки різних форматів структурованої інформації.

Зменшення вимог щодо обчислювального ресурсу, перепускності інформаційних каналів та ресурсу пам'яті апаратної платформи одночасно зі збільшенням продуктивності інформаційної системи машинного аналізу текстових даних може бути виконано завдяки впровадженню концепції семантико-релевантного пошуку [15–18], що ґрунтується на визначенні контекстного значення окремих складників вхідного запиту. Упровадження зазначеної концепції в цьому разі може базуватись на алгоритмах, що визначаються мінімальним навантаженням на ресурси апаратної платформи та реалізації водночас розширеного функціоналу, як це показано на діаграмі (рис. 2).

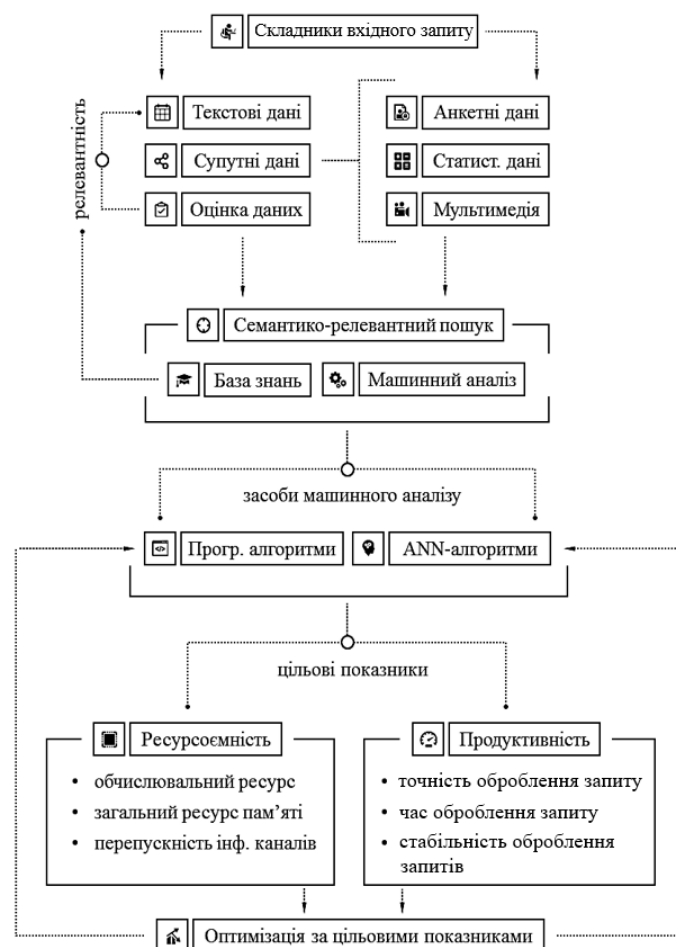


Рис. 2. Узагальнена схема організації системи машинного аналізу текстових даних на основі семантико-релевантного пошуку

Відповідно до схеми першим етапом проєктування системи машинного аналізу текстових повідомлень на основі семантико-релевантного пошуку є визначення складників вхідного запиту, а саме текстової інформації з оцінкою її релевантності та супутніх даних, що містять такі категорії, як анкетні, статистичні та мультимедійні дані. Семантико-релевантний пошук проводиться на основі бази знань та машинного аналізу із застосуванням нейромережових та програмних алгоритмів. Оптимізація та налаштування сервісу проводиться відповідно до набору цільових показників точності та швидкодії в процесі оброблення вхідних запитів, що визначаються на кількісному рівні (як на етапі проєктування способом розрахунку згідно з математичною моделлю, так і на етапі експлуатації програмної системи).

Організація системи машинного аналізу текстових повідомлень на основі семантико-релевантного пошуку, отже, оснований на отриманні максимально широкого набору даних відповідно до складників вхідного запиту, а також мети оброблення інформації для досягнення релевантного результату. У низці випадків збільшення продуктивності машинного аналізу за умов зменшення навантаження на ресурс апаратної платформи можливе завдяки долученню в аналіз фіксованої структури документа, властивої для відповідної категорії текстових даних.

У межах дослідження пропонується виокремити категорії організації масивів текстової інформації, що є найбільш актуальними для машинного аналізу.

1. Наукові публікації, структура яких зазвичай містить послідовний набір таких текстових блоків: назва публікації; автори; перелік наукових установ; тези; ключові слова; вступ; аналіз публікацій у профільних виданнях; основна частина; висновки; перелік джерел посилання; назва видання, у якому розміщена публікація.

2. Технічна документація, структура якої зазвичай містить послідовний набір таких текстових блоків: назва документа; назва організації; автори; дата випуску; вступ; зміст; основна частина (загальні відомості, опис продукту, інструкції з використання, технічні характеристики, процедури обслуговування та ремонту); безпека та попередження; додатки.

3. Юридична документація, структура якої зазвичай передбачає послідовний набір таких текстових блоків: назва документа; перелік авторів /

учасників; дата документа; преамбула; основна частина (визначення термінів, умови, зобов'язання сторін, санкції та відповідальність); завершальні положення; додаткові умови; підписи сторін.

4. Фінансова документація, структура якої зазвичай містить послідовний набір таких текстових блоків: назва документа; організація; перелік відповідальних осіб; дата звіту; вступ; основна частина (баланс, звіт про прибутки та збитки, звіт про рух грошових коштів, звіт про зміни у власному капіталі); примітки до фінансової звітності; аудиторський висновок; додатки.

5. Новинні статті, структура яких зазвичай передбачає послідовний набір таких текстових блоків: заголовок; автор; дата публікації; короткий опис; вступ; основна частина (факти, цитати, події); додаткова інформація; підсумки; джерела.

6. Стислі повідомлення із соціальних мереж, відгуків, коментарів, які часто мають завуальований або емоційно забарвлений контекст, містять скорочення, сленг або неструктуровану інформацію. Структура таких текстів, як правило, є фрагментарною та мінімально формалізованою: автор повідомлення, дата або час публікації, текст повідомлення, іноді – мультимедійні вкладення або реакції інших користувачів. Особливу складність для машинного аналізу становить необхідність розпізнавання прихованого змісту, емоційної тональності та інтерпретації контексту за мінімальної кількості лінгвістичних підказок.

Наявність фіксованої структури масиву текстових даних, що підлягає аналізу, надає додаткові можливості для пошуку компромісу між продуктивністю системи машинного аналізу, часом оброблення вхідного запису та навантаженням на ресурс апаратної платформи. Так, для категорії "Наукові публікації" такі структурні елементи, як "назва видання, у якому розміщена публікація", "перелік авторів" і "перелік наукових установ" дає змогу оцінити релевантність і достовірність поданої інформації без необхідності детального аналізу всього текстового масиву, а структурні елементи "тези" та "висновки" можуть бути використані для зменшення загального обсягу даних, що підлягає аналізу без суттєвого зменшення точності оброблення вхідного запиту. Це сприяє значному послабленню навантаження на обчислювальний ресурс, перепускність інформаційного каналу та ресурс пам'яті.

Розроблення системи машинного аналізу масивів текстових даних природною мовою передбачає розв'язання таких основних завдань:

- побудова концепції семантико-релевантного пошуку текстової інформації, яка базується на оцінці смислової близькості між текстовими блоками та контекстних властивостей стислих повідомлень;

- упровадження програмних і нейромережових алгоритмів попереднього оброблення, токенизації, кластеризації та глибокого аналізу текстів для підвищення якості інтерпретації змісту даних;

- розроблення підходів до оптимізації машинного аналізу текстової інформації з огляду на обмеження обчислювальних ресурсів системи (процесорний час, обсяг пам'яті, пропускну здатність каналів оброблення запитів);

- визначення критеріїв якості аналізу текстових повідомлень за кількісними показниками, такими як точність класифікації, середня похибка прогнозування та обсяг використаних ресурсів.

Запропонований підхід дає змогу адаптувати машинний аналіз до особливостей як стислих неструктурованих повідомлень, так і великих структурованих документів, що розширює можливості побудови гнучких і продуктивних інформаційних систем оброблення текстових даних природною мовою.

2. Методика організації

програмного оброблення текстових даних із використанням нейромережових моделей

З метою побудови системи машинного аналізу текстової інформації природною мовою, що поєднує оброблення як стислих повідомлень, так і складних структурованих документів, у роботі запропоновано багаторівневу методику організації програмних і нейромережових алгоритмів. Методика основана на послідовному застосуванні етапів класифікації текстів, попереднього оброблення, кластеризації за семантичними ознаками, релевантного пошуку та глибокого аналізу текстових даних із використанням нейронних мереж. У цьому разі особливу увагу приділено мінімізації обчислювальних витрат без суттєвого зниження точності аналізу. Належна адаптація програмних алгоритмів для організації системи машинного аналізу текстових повідомлень дає змогу зменшити навантаження на ресурс

апаратної платформи та збільшити швидкодію оброблення вхідної інформації.

2.1. Етапи організації

машинного аналізу текстових даних

У межах цього дослідження пропонуємо структурування процесу аналізу текстових даних у вигляді послідовності етапів, кожен з яких розв'язує конкретне завдання попереднього оброблення, класифікації або глибокого аналізу текстових масивів.

Методика передбачає п'ять основних етапів.

1. Аналіз структури вхідних текстових даних,

де аналізується характер вхідної інформації з метою її класифікації за типами документів. Береться до уваги різниця між стислими повідомленнями (соціальні мережі, відгуки) та великими структурованими документами (наукові статті, технічна й фінансова документація). Визначення типу тексту дає змогу попередньо обрати відповідні підходи до його оброблення, оптимізувати побудову моделі оброблення інформації та сформувати очікувану структуру вихідних результатів. Для кожної категорії текстових повідомлень формується типова модель структури, що зважає на формат подання, наявність метаданих і можливі супутні елементи (заголовки, авторські поля, дати). У межах запропонованої методики тексти поділяються на основі їх функціонального призначення, що наведено в табл. 1.

Таблиця 1. Класифікація типів текстових документів

№	Тип документа	Характеристика	Приклади
1	Наукові публікації	Використання термінології, структурованість	Статті, монографії
2	Технічна документація	Регламентовані формати опису процесів	Інструкції, технічні паспорти
3	Юридична документація	Формалізовані структури, правові терміни	Контракти, договори, рішення суду
4	Фінансова документація	Числова інформація, аналітичні висновки	Баланси, фінансові звіти
5	Новинні статті	Динамічні, оперативні тексти	Новини, репортажі
6	Стислі повідомлення з соціальних мереж, відгуків, коментарів	Неструктуровані тексти, емоційна забарвленість, лаконічна форма	Відгуки, дописи в соцмережах, коментарі

2. Попереднє оброблення текстових масивів, де здійснюється токенизація текстових даних (поділ тексту на окремі значущі одиниці: слова, підслова або символи). Для великих за обсягом документів особливу увагу приділяють збереженню логічних розділів і контексту між абзацами, тоді як для стислих повідомлень важливішою є нормалізація тексту та вилучення шуму. Паралельно застосовуються методи попереднього виділення ключових слів і тематичних груп завдяки частотного аналізу лексем. Для підвищення релевантності оброблення застосовується кластеризація текстів за змістовими ознаками, що дає змогу згрупувати тексти в тематично однорідні підмножини, мінімізуючи складність подальшого аналізу.

3. Семантико-релевантний аналіз текстових даних, де реалізується процедура визначення змістової близькості текстових елементів на основі семантичних ознак. Оброблення здійснюється із застосуванням концепції семантико-релевантного пошуку, що передбачає формування простору ознак для текстових блоків та оцінювання їх відносної релевантності до запиту чи бази знань. Релевантність розраховується як за класичними статистичними методами (наприклад, TF-IDF або BM25), так і за допомогою векторного подання текстів (Word2Vec, GloVe, BERT-ембединги). Семантична релевантність визначає пріоритетність текстів для подальшого глибокого аналізу.

4. Глибокий аналіз за допомогою нейромережевих моделей, де в запропонованому рішенні використовується архітектура двонаправленої рекурентної нейронної мережі (*Bidirectional LSTM*), що дає змогу моделювати залежності між словами як у прямому, так і у зворотному напрямках. Застосування двонаправленого оброблення особливо актуальне для невеликих за обсягом текстів, де кожен елемент має вагоме контекстне значення. Модель налаштовується на розв'язання завдання класифікації текстів за тональністю або іншими ознаками за мінімальної кількості епох і обмеженого розміру вхідного словника.

5. Оптимізація обчислювальних ресурсів на всіх етапах оброблення, де мінімізації витрат ресурсів інформаційної системи реалізуються за допомогою використання методів регуляризації моделі (*Dropout*, *Batch Normalization*), раннього припинення навчання (*EarlyStopping*), адаптивного регулювання розміру пакетів даних. Також передбачено

обмеження обсягу словника ознак і скорочення довжини послідовностей для текстів з низькою інформаційною насиченістю. Сукупність зазначених заходів сприяє зменшенню обсягів споживаної оперативної пам'яті та скороченню часу оброблення без суттєвого зниження точності результатів.

Запропоновані етапи дають змогу формалізувати оброблення текстової інформації природною мовою таким чином, щоб поєднати адаптивність системи до різних типів текстових масивів з оптимальним використанням обчислювальних ресурсів. Це забезпечує можливість застосування системи в реальних умовах, зокрема в середовищах з обмеженими ресурсами або для оброблення великих обсягів текстових даних.

2.2. Використання кластеризації та токенизації під час попереднього оброблення текстових даних

На етапі попереднього оброблення та структуризації великих масивів текстової інформації найбільшу ефективність відповідно до швидкодії та мінімізації використання ресурсів демонструють такі групи методів:

- методи кластеризації масивів текстових повідомлень, а саме кластеризація зв'язності [19], кластеризація методом *k*-середнього [20], зокрема кластеризація із середнім зсувом [21], а також кластеризація за схемою DBSCAN [22];

- методи токенизації [23, 24], зокрема токенизація на основі слів, токенизація на основі підслів і токенизація на основі символів;

- методи ймовірнісної класифікації складників масивів текстових повідомлень із застосуванням моделі гауссової суміші [25].

Кластеризація як група методів поділу масиву текстових даних на підмножини (кластери) визначається найбільш широким функціоналом, гнучкістю та масштабованістю. У межах дослідження пропонуємо поділити методи кластеризації на жорсткі, у межах яких ділянки кластерів не перетинаються, і м'які, де окремі елементи текстового блоку можуть одночасно належати до кількох кластерів [19–23]. Базовими підходами кластеризації, що використовується в межах концепції семантико-релевантного пошуку, є визначення ділянок кластерів на основі онтології семантичного словника, що адаптується для текстової інформації, поданої природною мовою [5, 19], та виділення ключових слів, що за типовим сценарієм визначаються програмними алгоритмами відповідно

до показника частоти вживання зазначених термінів і структури документа [30]. Водночас зауважимо, що нині під час роботи з текстовими даними найбільш широко використовуються алгоритми кластеризації, основані на методі k -середнього [20–22]. Згідно із зазначеним методом для кожної ділянки кластера з огляду на середнє арифметичне ключових ознак його елементів циклічно розраховується центроїд. Кількість циклів залежить від стійкості центроїда, а також вимог щодо мінімізації дисперсії між елементами в середині кластера та максимізації

дисперсії складників між окремими кластерами (рис. 3). Відповідні алгоритми кластеризації визначаються мінімальним навантаженням на ресурс апаратної платформи, але мають низку недоліків, що здебільшого пов'язано з випадковим характером початкового вибору центроїдів ділянок кластерів [20–22]. Розв'язання зазначеної проблеми полягає в упровадженні методів кластеризації із середнім зсувом як непараметричного аналізу простору ознак масиву текстових даних із визначенням місця розташування ділянок високої щільності ймовірності [21].

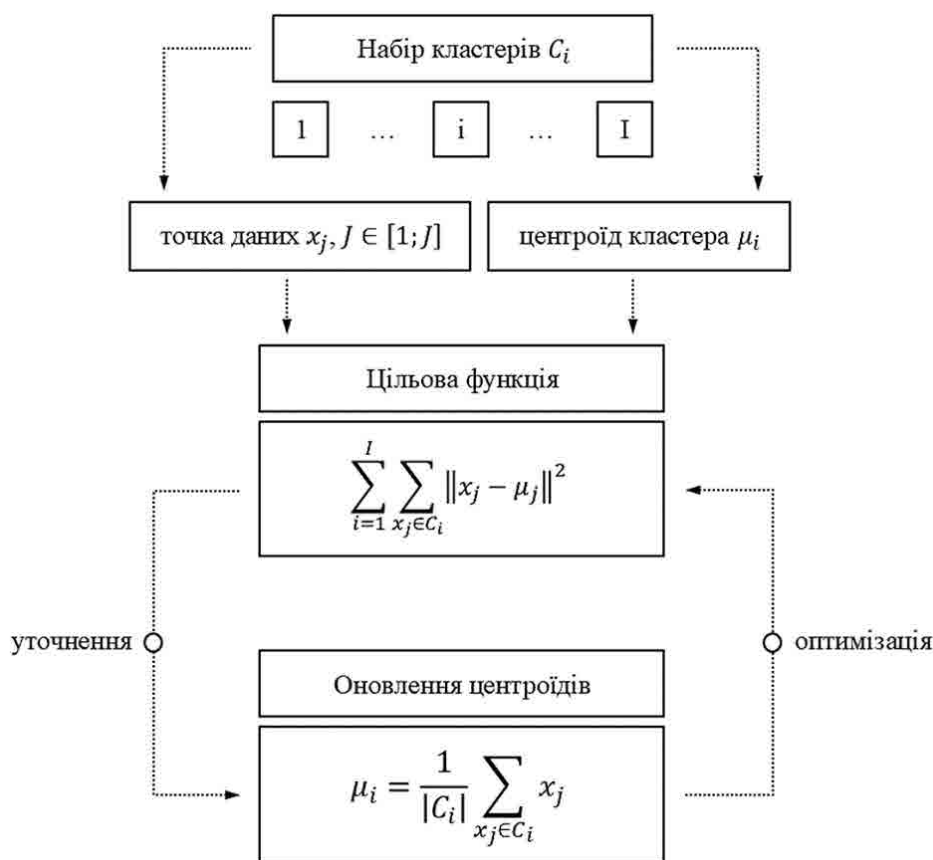


Рис. 3. Схема кластеризації на основі методу k -середнього

Важливо зауважити, що програмні алгоритми на основі методів кластеризації із середнім зсувом, зокрема такі, що базуються на схемі DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) можуть також ефективно застосовуватись у кластерному аналізі супутніх даних [22]. Відповідна схема кластеризації передбачає визначення таких категорій (рис. 4):

– ділянка щільності, що встановлюється радіусом;

– показник, який визначає мінімальну кількість точок у ділянці щільності, що є необхідною для формування точки ядра;

– точки ядра, що встановлюються за допомогою показника;

– приграничні точки, що знаходять у ділянці щільності принаймні однієї точки ядра;

– шумові точки, що лежать поза ділянки щільності всіх точок ядра.

Недоліки методів кластеризації із середнім зсувом полягають у невизначеності низки елементів,

що лежать на межі ділянок кластерів і зазвичай належать до викидів. Це виконується за допомогою вибору кластеризації на підставі зв'язності між елементами [19], алгоритми на основі якої формують ділянки кластерів у вигляді деревоподібної структури, де кожен кластер поданий у вигляді вузла загальної ієрархічної структури. У цьому разі для алгоритму машинного аналізу текстової інформації може бути обрана метрика для обчислення відстані між елементами масиву, що надалі відіграє ключову роль у кластеризації. Методи кластеризації зв'язності застосовуються для ідентифікації та класифікації параметрів текстових даних, водночас відповідна архітектура є достатньо гнучкою для модифікацій та масштабування.

Крім того, у використанні методів імовірнісної класифікації застосовується ранжування відповідно до оцінки за критерієм максимальної схожості складників (*Estimation Maximization*; EM) з використанням імовірнісної моделі гауссової суміші (*Gaussian Mixture Models*, GMM). Ці моделі передбачають, що дані можуть бути подані як комбінація кількох гауссових розподілів, кожен з яких відповідає окремій групі або кластеру. На відміну від детермінованих підходів, імовірнісні моделі беруть до уваги невизначеність та розподіл даних у просторі. Перелічимо основні етапи роботи.

1. Ініціалізація параметрів, де задаються початкові значення параметрів, таких як середнє значення та коваріаційна матриця для кожного гауссового компонента.

2. Розрахунок очікувань, де для кожного зразка даних обчислюється ймовірність належності до кожного з компонентів гауссової суміші. Ці ймовірності використовуються для розрахунку очікуваного значення функції правдоподібності. Замість жорсткої класифікації, де кожен зразок чітко належить до одного кластера, ймовірнісна класифікація дає змогу розподілити зразки між кількома кластерами на основі ймовірності.

3. Максимізація параметрів, де після розрахунку очікуваних значень на основі отриманих імовірностей оновлюються параметри моделі (середнього, коваріаційної матриці і вагових коефіцієнтів для кожного компонента). Мета полягає в максимізації функції правдоподібності на основі оновлених параметрів.

4. Циклічне повторення кроків 1–3 доти, доки функція правдоподібності не досягне збіжності, тобто

коли подальші зміни параметрів стануть незначними. Цей підхід дає змогу поступово покращувати модель, підлаштовуючи її під дані.

Алгоритми ймовірнісної класифікації використовуються з метою пошуку максимальної схожості відповідно до параметрів імовірнісної моделі. В основі цього підходу лежить припущення, що вихідні значення цілком залежать від набору прихованих змінних. Виконання процедури ймовірнісної класифікації передбачає циклічне визначення очікуваного значення функції правдоподібності та оцінювання показника максимальної схожості до отримання збіжності згідно з поставленими умовами. Імовірнісна класифікація найбільш ефективна в роботі з текстовими повідомленнями, що належать до декількох груп (м'яка класифікація). Порівняно із середнім зсувом відповідні алгоритми є більш гнучкими й масштабованими, а також не обмеженими метрикою класичних методів кластеризації.

Нарешті, токенизація масивів текстової інформації використовується для збільшення ефективності подальшого застосування нейромережових алгоритмів як на етапі навчання, так і на етапі машинного аналізу [23, 24]. Залежно від обраного підходу для слів, підслів або символів надаються показники, подані на кількісному рівні, що визначають положення складника в текстовому блоці, а також наявність відповідного елемента в словнику.

Важливо зауважити, що ефективність токенизації елементів залежить від однорідності тексту, а отже, машинний аналіз має проводитись не для всього масиву даних, а для структурних складників. Це дає змогу зменшити розмір набору токенів і ефективно адаптувати модель відповідно до вхідного запиту.

Зі свого боку програмні алгоритми, що можуть бути використані на основному етапі машинного аналізу текстових даних, містять такі категорії:

- кількісна оцінка важливості окремих термінів [26, 27] у масиві текстових даних (*Term Frequency-Inverse Document Frequency*; TF-IDF);

- ранжування складників текстового запиту відповідно до показників релевантності пошуковому запиту [28, 29].

В основі методів TF-IDF [26, 27] лежить розрахунок вагового коефіцієнта, що визначає ступінь важливості окремого терміна в масиві текстових повідомлень, який підлягає машинному аналізу. Відповідна величина розраховується через

добуток частоти появи терміна в окремому текстовому блоці (*Term Frequency*; TF) та оберненої частоти для всього масиву текстової інформації (*Inverse Document Frequency*; IDF). Застосування цього підходу сприяє зниженню ваги загальних частовживаних, але малоінформативних слів, наприклад сполучників і прийменників. Програмні алгоритми на основі методу TF-IDF визначаються простотою реалізації за умов постановки завдання початкового аналізу великих обсягів текстових даних, оскільки обчислювальна складність є лінійною щодо розміру загального масиву. Натомість основним недоліком зазначеного методу є обмеження функціональних можливостей. Зазначається, що алгоритми TF-IDF проводять лише лексичний аналіз, тоді як для позиціювання термінів необхідно застосувати інші підходи. Зі свого боку за умов ранжування складників текстового запиту за показниками релевантності пошуковому запиту [28, 29] програмний алгоритм оцінює близькість запиту до масиву текстової інформації без необхідності додаткових розрахунків з метою визначення відповідності між векторними репрезентаціями вхідного запиту й текстових повідомлень. Математичний апарат, що лежить в основі алгоритму, також містить обчислювальні метрики методу TF-IDF, але відповідні вагові коефіцієнти термінів обчислюються не для створення векторного подання документа, а для співвідношення документа й вхідного запиту.

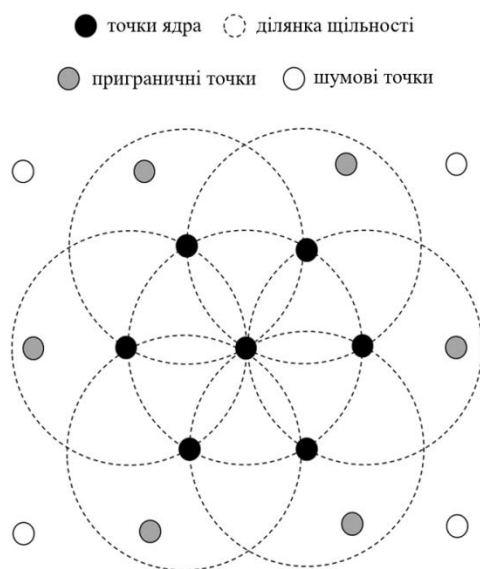


Рис. 4. Кластеризація на основі схеми DBSCAN

Перевага нейромережевих алгоритмів машинного аналізу масивів текстових даних полягає

в розширенні функціоналу програмної системи без необхідності розроблення точної математичної моделі мовного корпусу як структурованої збірки текстів, що використовується для семантичного аналізу з метою дослідження мовних закономірностей. Недоліком зазначеного підходу є зростання навантаження на ресурс апаратної платформи й складність упровадження схеми оптимізації програмної системи відповідно до цільових показників продуктивності машинного аналізу [31–39].

2.3. Практична реалізація методики машинного аналізу текстових даних

На основі розробленої багаторівневої методики створено експериментальну програмну систему для автоматизованого машинного аналізу текстової інформації природною мовою для визначення тональності відгуків користувачів. Основною метою практичної реалізації є перевірка працездатності запропонованого підходу на прикладі завдання аналізу стислих текстових повідомлень.

Реалізація передбачає модулі попереднього оброблення текстових даних (токенізація, нормалізація, кластеризація), механізми семантико-релевантного аналізу й нейромережеву модель на основі архітектури *Bidirectional LSTM* для глибокого аналізу змісту повідомлень. Особливу увагу приділено мінімізації ресурсних витрат унаслідок впровадження регуляризації, оптимізації структури мережі та адаптації довжини вхідних послідовностей до характеру оброблюваних текстових даних.

Практична реалізація дала змогу оцінити можливості запропонованої методики в роботі з текстовими масивами різної структури та довжини, а також сформулювати рекомендації щодо подальшого вдосконалення систем оброблення природної мови в ресурсно обмежених середовищах.

Створено експериментальну модель глибокого аналізу з використанням архітектури *Bidirectional LSTM* та налаштовано на розпізнавання тональності тексту способом оброблення токенізованих і нормалізованих послідовностей вхідної інформації (рис. 5). Конструкція моделі передбачає векторизацію ознак (*Embedding*), двонаправлену LSTM-структуру для оброблення контекстних залежностей, кілька повнозв'язних шарів для кінцевої регресії та механізми регуляризації для запобігання перенаванчання.



Рис. 5. Налаштування нейронної мережі для розпізнавання тональності тексту

Як показали дослідження, алгоритми на основі нейромережевої архітектури типу генеративного трансформера й моделей, побудованих на її основі, зокрема BERT, GPT та T5, мають найбільш широкий спектр застосувань в аналізі текстових даних. Завдяки механізму *self-attention* трансформери ефективно

обробляють великі текстові масиви, зважаючи на контекст кожного терміна в межах всього документа [31–33]. Це дає змогу досягати високої точності у виконанні завдання щодо роботи з природними мовами, таких як переклад, класифікація, витяг релевантної інформації та генерація тексту. Зі свого боку архітектура RNN та сучасні моделі на їх основі, наприклад LSTM та GRU, також широко застосовуються в аналізі текстових повідомлень завдяки своїй здатності обробляти послідовності інформації [34–38]. Алгоритми на їх основі використовуються для завдань, що потребують визначення контексту та порядку слів, зокрема мовне моделювання, машинний переклад, розпізнавання мовлення та класифікація тексту. У цьому разі RNN мають обмеження в обробленні дуже довгих послідовностей через проблеми з градієнтним загасанням.

Нарешті, архітектура CNN більшою мірою асоціюється з обробленням матриці зображення, але вони також ефективні для аналізу текстових даних, поданих природною мовою [37–39]. У текстових завданнях CNN використовують згорткові фільтри, що допомагають виявляти суттєві особливості в тексті, наприклад комбінації слів або фраз, що мають важливе значення для конкретного завдання.

Для підвищення точності розпізнавання тональності стислих текстових повідомлень розроблено процедуру навчання нейронної мережі на основі оптимізованої структури шарів і механізмів регуляризації. Навчання здійснювалося з використанням функції втрат середньоквадратичної помилки (MSE) та оптимізатора *Adam* із початковою швидкістю навчання 0.001.

Для контролю перенавчання моделі застосовано механізм раннього припинення тренування (*EarlyStopping*), а для адаптації швидкості навчання – механізм *ReduceLROnPlateau*. Попередню токенизацію даних проведено з обмеженням розміру словника до 10 000 ознак і фіксацією довжини вхідних послідовностей на рівні 600 tokenів.

Динаміка навчання нейронної мережі фіксувалася у вигляді протоколу, що відтворює зміну значень помилки на тренувальній та валідаційній вибірках після кожної епохи навчання.

На рис. 6 наведено фрагмент протоколу тренування, де продемонстровано кількість епох, значення функції помилки (*loss*) для тренувальної вибірки (*loss*), значення функції помилки для валідаційної вибірки (*val_loss*) та швидкість навчання (*lr*).


```

48/48 [=====] - 82s 2s/step - loss: 14.4485 - val_loss: 9.9785 - lr: 0.0010
Epoch 2/20
48/48 [=====] - 79s 2s/step - loss: 9.8172 - val_loss: 7.3289 - lr: 0.0010
Epoch 3/20
48/48 [=====] - 75s 2s/step - loss: 6.3964 - val_loss: 6.2269 - lr: 0.0010
Epoch 4/20
48/48 [=====] - 78s 2s/step - loss: 4.4186 - val_loss: 5.5421 - lr: 0.0010
Epoch 5/20
48/48 [=====] - 69s 1s/step - loss: 3.4891 - val_loss: 5.5777 - lr: 0.0010
Epoch 6/20
48/48 [=====] - 69s 1s/step - loss: 2.8274 - val_loss: 4.4453 - lr: 0.0010
Epoch 7/20
48/48 [=====] - 69s 1s/step - loss: 2.4941 - val_loss: 4.4178 - lr: 0.0010
Epoch 8/20
48/48 [=====] - 68s 1s/step - loss: 2.1417 - val_loss: 4.4678 - lr: 0.0010
Epoch 9/20
48/48 [=====] - 69s 1s/step - loss: 1.8395 - val_loss: 4.7220 - lr: 0.0010
Epoch 10/20
48/48 [=====] - 67s 1s/step - loss: 1.7584 - val_loss: 4.7917 - lr: 0.0010
Epoch 11/20
48/48 [=====] - 70s 1s/step - loss: 3.7593 - val_loss: 8.3930 - lr: 0.0010
Epoch 12/20
48/48 [=====] - 74s 2s/step - loss: 3.6373 - val_loss: 4.8200 - lr: 0.0010
Epoch 13/20
48/48 [=====] - 74s 2s/step - loss: 2.3809 - val_loss: 4.9578 - lr: 0.0010

```

Рис. 6. Фрагмент протоколу навчання нейронної мережі з показниками помилки на тренувальній та валідаційній вибірках

Для візуалізації результатів роботи нейронної мережі розроблено інтерфейс з метою виведення результатів прогнозування тональності стислих повідомлень (рис. 7), де кожному обробленому відгуку надається числова оцінка тональності в діапазоні від 1 до 5. Відгуки відтворюються

в структурованому вигляді із зазначенням оцінки й тексту повідомлення. Для забезпечення зручності імпорту текстових даних у застосунок передбачено кнопку "Імпорт відгуків", що дає змогу автоматично додавати нові повідомлення для аналізу без необхідності ручного введення.

Імпорт відгуків

Оцінка тональності (1–5): 4.8654556

Відгук:

Компанія відповідально ставиться до замовлень. Якість продукції чудова, планує співпрацювати надалі.

Оцінка тональності (1–5): 1.0154552

Відгук:

Сервіс дуже поганий. Замовлення не виконали вчасно, постійні затримки і відсутність зворотного зв'язку.

Оцінка тональності (1–5): 4.1304451

Відгук:

Дуже вдячний за швидке оформлення і доставку. Приємно працювати з професіоналами!

Оцінка тональності (1–5): 3.1012891

Відгук:

Все добре, але хотілося б більше варіантів оплати. Загалом сервіс якісний.

Рис. 7. Інтерфейс відтворення оцінювальних текстових відгуків у програмному застосунку

Особливістю практичної реалізації є орієнтація на мінімальні обчислювальні вимоги, що забезпечує можливість використання розробленої системи на

стандартних персональних комп'ютерах середнього класу або у віртуалізованих середовищах із обмеженим доступом до ресурсів процесора й пам'яті.

Для оцінювання якості запропонованого рішення були проведені експериментальні вимірювання ключових показників продуктивності. Модель тренувалася й тестувалася на апаратній платформі з обмеженими ресурсами: ноутбук з низькопотужним процесором Intel Core i3-8130U, 8 ГБ оперативної пам'яті та інтегрованою графікою. Основні результати навчання подано в табл. 2.

Таблиця 2. Параметри тренування моделі машинного аналізу текстових даних

Параметр	Значення
Кількість епох тренування	20
Кількість кроків на одну епоху	48
Середній час оброблення однієї епохи	≈73 с
Загальний час тренування	≈16 хв
Середнє завантаження процесора під час тренування	65–70 %
Максимальне використання оперативної пам'яті	≈3,2 ГБ
Середнє значення функції втрат (<i>Loss</i>) на тренувальній вибірці	4,82
Середнє значення функції втрат (<i>Loss</i>) на валідаційній вибірці	5,63

За результатами експериментів встановлено, що запропонована архітектура нейронної мережі забезпечує стабільне зниження функції втрат без суттєвого збільшення обчислювальних витрат. Отримані показники завантаження процесора та використання оперативної пам'яті свідчать про можливість впровадження розробленої системи навіть на апаратних платформах середнього класу без потреби в спеціалізованих графічних прискорювачах.

Список літератури

1. Petrov V.V., Zichun L., Kryuchyn A.A., Shanoylo S.M., Mingle F., Beliak I.V., Manko D.Y., Lapchuk A.S., Morozov E.M. Long-term storage of digital information. *Akademperiodyka*, Kyiv. 148 p. 2018. DOI: <https://doi.org/10.15407/akademperiodyka.360.148>
2. Giordano V., Spada I., Chiarello F., Fantoni G. The impact of ChatGPT on human skills: A quantitative study on Twitter data. *Technological Forecasting and Social Change*. 2024. No. 203. 124 p. DOI: <https://doi.org/10.1016/j.techfore.2024.123389>
3. Dahri N. A., Extended Tam based acceptance of ai-powered ChatGPT for supporting metacognitive self-regulated learning in education: A mixed-methods study / Dahri N. A., Yahaya N., Al-Rahmi W. M., Aldraiweesh A., Alturki U., Almutairy S., Shutaleva A., Soomro R. B. *Heliyon*. 2024. No. 10(8). DOI: <https://doi.org/10.1016/j.heliyon.2024.e29317>
4. Malhotra A., Bajaj K. A hybrid pattern based text mining approach for malware detection using DBScan. *CSI Transactions on ICT*, 4 (2-4), 2016. P. 141-149. DOI: <https://doi.org/10.1007/s40012-016-0095-y>
5. The Trustees of Princeton University. What is WordNet? Princeton University. Retrieved January 12, 2022, URL: <https://wordnet.princeton.edu>.
6. Marcos T. Efficient Methods for Natural Language Processing: A Survey. / Marcos Treviso, Ji-Ung Lee, Tianchu Ji. et al. *Transactions of the Association for Computational Linguistics*, 11. 2023. P. 826–860. DOI: 10.1162/tacl_a_00577
7. Zhao W. X., A Survey of Large Language Models. / Zhao W. X., Zhou, K., Li, J. et al. *Computation and Language*. 144 p. 2023. DOI: <https://doi.org/10.48550/arXiv.2303.18223>
8. Lialin V., Deshpande V., Rumshisky A. Scaling down to scale up: A guide to parameter-efficient fine-tuning. *Computation and Language*. 2023. DOI: <https://doi.org/10.48550/arXiv.2303.15647>

Висновки

Унаслідок проведеного дослідження розроблено та експериментально перевірено багаторівневу методику організації програмних і нейромережових алгоритмів для машинного аналізу текстових даних, поданих природною мовою. Основну увагу приділено мінімізації обчислювальних витрат, що дало змогу забезпечити стабільну роботу системи на апаратних платформах з обмеженими ресурсами.

Запропонована система поєднує методи кластеризації, токенизації, ймовірнісної класифікації та глибокого аналізу за допомогою нейронних мереж типу *Bidirectional LSTM*. Проведене тестування показало здатність моделі зберігати високу точність визначення тональності стислих текстових повідомлень за мінімального навантаження на обчислювальні потужності.

Серед основних переваг розробленого підходу необхідно наголосити на можливості адаптації до різних типів текстових даних, зниженні витрат ресурсів без погіршення якості аналізу, а також на придатності для розгортання в середовищах з низькою обчислювальною спроможністю.

Надалі дослідження планується спрямувати на розширення функціональності системи за допомогою інтеграції оброблення мультимодальних даних, поглибленого використання зовнішніх баз знань і вдосконалення архітектур для роботи в режимі реального часу.

9. Yang J. A Survey of Knowledge Enhanced Pre-trained Models. / Yang J., Xiao G., Shen Y., Jiang W., Hu X., Zhang Y., Peng, J. et al. *Computation and Language*. 2021. 32 p. <https://doi.org/10.48550/arXiv.2110.00269>
10. Fournie Q., Caron G. M., Aloise D. A Practical Survey on Faster and Lighter Transformers. *ACM Computing Surveys*, 55(14s), P. 1-40. 2023. DOI: <https://doi.org/10.1145/3586074>
11. Zhang X. Edge intelligence optimization for large language model inference with batching and quantization. / Zhang X., Liu J., Xiong Z., Huang Y., Xie G., Zhang R. et al. *IEEE Wireless Communications and Networking Conference (WCNC)*. 2024. DOI: <https://doi.org/10.1109/wcnc57260.2024.10571127>
12. Wei X. Knowledge Enhanced Pretrained Language Models: A Comprehensive Survey. / Wei X., Wang S., Zhang D., Bhatia P., Arnold A.O. et al. *Computation and Language*. 2021. DOI: <https://doi.org/10.48550/arXiv.2110.08455>
13. Nagamatsu N., Hara-Azumi Y. Dynamic split computing-aware mixed-precision quantization for efficient deep edge intelligence. *IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*. 2023. DOI: <https://doi.org/10.1109/trustcom60117.2023.00355>
14. Bao Y., Xu Y., Xiong H. Feature map alignment: Towards efficient design of mixed-precision quantization scheme. *IEEE Visual Communications and Image Processing (VCIP)*. 2019. DOI: <https://doi.org/10.1109/vcip47243.2019.8965724>
15. Shamaeva I., Galley D. Simple and advanced Google Search. *Custom Search – Discover More*: P. 7–27. 2021. DOI: <https://doi.org/10.1201/9781003100133-2>
16. Vo N.P., Popescu O. A multi-layer system for semantic textual similarity. *Proceedings of the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*. P. 56-67. 2016. DOI: <https://doi.org/10.5220/0006045800560067>
17. Yıldız E., Findik Y. Question similarity detection in Turkish using semantic textual similarity methods. *2019 27th Signal Processing and Communications Applications Conference (SIU)*. 2019. DOI: <https://doi.org/10.1109/siu.2019.8806308>
18. Nel W., de Wet L., Schall R. Randomised controlled trial of the usability of major search engines (Google, Yahoo! and Bing) when using ambiguous search queries. *Proceedings of the 4th International Conference on Computer-Human Interaction Research and Applications*. P. 152-161. 2020. DOI: <https://doi.org/10.5220/0010133601520161>
19. Oladipo F. O., Ohiani A. B.A. Text summarization system: An extractive approach using hierarchical text clustering. *International Journal of Computer Applications*, 174 (23), P. 15–19. 2021. DOI: <https://doi.org/10.5120/ijca2021921015>
20. Bindal A. Pathak A. A survey on K-means clustering and web-text mining. *International Journal of Science and Research (IJSR)*, 5(4), P. 1049–1052. 2016. DOI: <https://doi.org/10.21275/v5i4.nov162776>
21. Shaposhnikov A. I. Feature-vector for the meanshift. *Proceedings of Tomsk State University of Control Systems and Radioelectronics*, 24(2), P. 34–38. 2021. DOI: <https://doi.org/10.21293/1818-0442-2021-24-2-34-38>
22. Tingting S. Application and research of DBSCAN optimization algorithm in big data analysis of experimental text. *Computer Science and Application*, 10 (05), P. 906-913. 2020. DOI: <https://doi.org/10.12677/csa.2020.105093>
23. Otani N. Pre-tokenization of multi-word expressions in cross-lingual word embeddings / Otani N., Ozaki S., Zhao X., Li Y., St Johns M., Levin L. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. P. 4451–4464. 2020. DOI: <https://doi.org/10.18653/v1/2020.emnlp-main.360>
24. Ribeiro E., Ribeiro R., de Matos D. A multilingual and Multidomain Study on Dialog Act recognition using character-level tokenization. *Information*, 10(3), 94 p. 2019. DOI: <https://doi.org/10.3390/info10030094>
25. Slimane F., Margner V. A new text-independent GMM writer identification system applied to Arabic handwriting. *International Conference on Frontiers in Handwriting Recognition*. 13 p. 2014. DOI: <https://doi.org/10.1109/icfhr.2014.124>
26. Belogorskaya D. V. Summarizing news texts using quantitative methods (TF-IDF). *Proceedings of the VII (XXI) International Scientific and Practical Conference of Young Scientists*. 2020. DOI: <https://doi.org/10.17223/978-5-94621-901-3-2020-30>
27. Choi E.A., Han Y.E., Lee S., Oh M. A comparison of TF and TF-IDF analysis for trends of Blockchain in health and Welfare. *Journal of the Korean Data And Information Science Society*, 30(5), P. 1025–1036. 2019. DOI: <https://doi.org/10.7465/jkdi.2019.30.5.1025>
28. Ghawi R., Pfeffer J. Efficient hyperparameter tuning with grid search for text categorization using KNN approach with BM25 similarity. *Open Computer Science*, 9(1), P. 160–180. 2019. DOI: <https://doi.org/10.1515/comp-2019-0011>
29. Tinega G. A., Mwangi P. W., Rimuru D. R. Text mining in digital libraries using okapi BM25 model. *International Journal of Computer Applications Technology and Research*, 7(10), P. 398–406. 2018. DOI: <https://doi.org/10.7753/ijcatr0710.1003>
30. Ma X., Hovy E. End-to-end sequence labeling via bi-directional LSTM-cnns-CRF. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. P. 1064–1074. 2016. DOI: <https://doi.org/10.18653/v1/p16-1101>
31. Khalil F., Pipa P. D. Transforming the generative pretrained transformer into augmented business text writer. *CC BY 4.0*. 2021. DOI: <https://doi.org/10.21203/rs.3.rs-1170589/v1>
32. Soyalp G., Alar A., Ozkanli K., Yildiz B. Improving text classification with Transformer. *2021 6th International Conference on Computer Science and Engineering (UBMK)*. 12 p. 2021. DOI: <https://doi.org/10.1109/ubmk52708.2021.9558906>

33. Shen Y., Liu J. Comparison of text sentiment analysis based on Bert and word2vec. *2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer (ICFTIC)*. 17 p. 2021. DOI: <https://doi.org/10.1109/icftic54370.2021.9647258>
34. Pyllkkönen J., Ukkonen A., Kilpikoski J., Tamminen S., Heikinheimo H. Fast text-only domain adaptation of RNN-Transducer Prediction Network. *Interspeech*. 2021. DOI: <https://doi.org/10.21437/interspeech.2021-1191>
35. Nismi Mol E. A., Santosh Kumar M. B. Study on impact of RNN, CNN and Han in text classification. *2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA)*. 2020. DOI: <https://doi.org/10.1109/accthpa49271.2020.9213231>
36. Zouzou A., Azami I. E. Text sentiment analysis with CNN & GRU model using glove. *2021 Fifth International Conference On Intelligent Computing in Data Sciences (ICDS)*. 2021. DOI: <https://doi.org/10.1109/icds53782.2021.9626715>
37. Park P.W. Text-CNN based intent classification method for automatic input of intent sentences in chatbot. *The Journal of Korean Institute of Information Technology*, 18 (1), P. 19–25. 2020. DOI: <https://doi.org/10.14801/jkiit.2020.18.1.19>
38. Sun S., Gao Z., Huang C., Yu H. Glove-FRCNN: Comprehensive network algorithm for Vespa Mandarin image-text extraction and classification. *2021 International Conference on Communications, Information System and Computer Engineering (CISCE)*. 2021. DOI: <https://doi.org/10.1109/cisce52179.2021.9445902>
39. Pan N., Yao W., Li X. Friends recommendation based on KBERT-CNN Text Classification Model. *International Joint Conference on Neural Networks (IJCNN)*. 2021. DOI: <https://doi.org/10.1109/ijcnn52387.2021.9533618>

References

1. Petrov, V. V., Zichun, L., Kryuchyn, A. A., Shanoylo, S. M., Mingle, F., Beliak, I. V., Manko, D. Y., Lapchuk, A. S., Morozov, E. M. (2018), *Long-term storage of digital information*. Akadempriodyka, Kyiv. 148 p. DOI: <https://doi.org/10.15407/akadempriodyka.360.148>
2. Giordano, V., Spada, I., Chiarello, F., Fantoni, G. (2024), "The impact of ChatGPT on human skills: A quantitative study on Twitter data". *Technological Forecasting and Social Change*. No. 203. 124 p. DOI: <https://doi.org/10.1016/j.techfore.2024.123389>
3. Dahri, N. A., (2024), "Extended Tam based acceptance of ai-powered ChatGPT for supporting metacognitive self-regulated learning in education: A mixed-methods study" / Dahri, N. A., Yahaya, N., Al-Rahmi, W. M., Aldraiweesh, A., Alturki, U., Almutairy, S., Shutaleva, A., Soomro, R. B. *Heliyon*. 2024. No. 10(8). DOI: <https://doi.org/10.1016/j.heliyon.2024.e29317>
4. Malhotra, A., Bajaj, K. (2016), "A hybrid pattern based text mining approach for malware detection using DBScan". *CSI Transactions on ICT*, 4 (2-4), P. 141-149. DOI: <https://doi.org/10.1007/s40012-016-0095-y>
5. "The Trustees of Princeton University. What is WordNet? Princeton University". Retrieved January 12, 2022, available at: <https://wordnet.princeton.edu>.
6. Marcos, T. (2023), "Efficient Methods for Natural Language Processing: A Survey". / Marcos Treviso, Ji-Ung Lee, Tianchu Ji. et al. *Transactions of the Association for Computational Linguistics*, 11. P. 826–860. DOI: 10.1162/tacl_a_00577
7. Zhao, W. X., (2023), "A Survey of Large Language Models". / Zhao W. X., Zhou, K., Li, J. et al. *Computation and Language*. 144 p. DOI: <https://doi.org/10.48550/arXiv.2303.18223>
8. Lialin, V., Deshpande, V., Rumshisky, A. (2023), "Scaling down to scale up: A guide to parameter-efficient fine-tuning. *Computation and Language*. DOI: <https://doi.org/10.48550/arXiv.2303.15647>
9. Yang, J. (2021), "A Survey of Knowledge Enhanced Pre-trained Models". / Yang J., Xiao G., Shen Y., Jiang W., Hu X., Zhang Y., Peng, J. et al. *Computation and Language*. 32 p. <https://doi.org/10.48550/arXiv.2110.00269>
10. Fournie, Q., Caron, G. M., Aloise, D. (2023), "A Practical Survey on Faster and Lighter Transformers. *ACM Computing Surveys*, 55(14s), P. 1-40. DOI: <https://doi.org/10.1145/3586074>
11. Zhang, X. (2024), "Edge intelligence optimization for large language model inference with batching and quantization". / Zhang X., Liu J., Xiong Z., Huang Y., Xie G., Zhang R. et al. *IEEE Wireless Communications and Networking Conference (WCNC)*. DOI: <https://doi.org/10.1109/wcnc57260.2024.10571127>
12. Wei, X. (2021), "Knowledge Enhanced Pretrained Language Models: A Comprehensive Survey". / Wei X., Wang S., Zhang D., Bhatia P., Arnold A.O. et al. *Computation and Language*. DOI: <https://doi.org/10.48550/arXiv.2110.08455>
13. Nagamatsu, N., Hara-Azumi, Y. (2023), "Dynamic split computing-aware mixed-precision quantization for efficient deep edge intelligence". *IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*. DOI: <https://doi.org/10.1109/trustcom60117.2023.00355>
14. Bao, Y., Xu, Y., Xiong, H. (2019), "Feature map alignment: Towards efficient design of mixed-precision quantization scheme". *IEEE Visual Communications and Image Processing (VCIP)*. DOI: <https://doi.org/10.1109/vcip47243.2019.8965724>
15. Shamaeva, I., Galley, D. (2021), "Simple and advanced Google Search". *Custom Search – Discover More*: P. 7–27. DOI: <https://doi.org/10.1201/9781003100133-2>
16. Vo, N.P., Popescu, O. (2016), "A multi-layer system for semantic textual similarity". *Proceedings of the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*. P. 56-67. DOI: <https://doi.org/10.5220/0006045800560067>

17. Yıldız, E., Findik, Y. (2019), "Question similarity detection in Turkish using semantic textual similarity methods". *2019 27th Signal Processing and Communications Applications Conference (SIU)*. DOI: <https://doi.org/10.1109/siu.2019.8806308>
18. Nel, W., de Wet, L., Schall, R. (2020), "Randomised controlled trial of the usability of major search engines (Google, Yahoo! and Bing) when using ambiguous search queries". *Proceedings of the 4th International Conference on Computer-Human Interaction Research and Applications*. P. 152-161. DOI: <https://doi.org/10.5220/0010133601520161>
19. Oladipo, F. O., Ohiani, A. B.A. (2021), "Text summarization system: An extractive approach using hierarchical text clustering". *International Journal of Computer Applications*, 174 (23), P. 15–19. DOI: <https://doi.org/10.5120/ijca2021921015>
20. Bindal, A. Pathak, A. (2016), "A survey on K-means clustering and web-text mining". *International Journal of Science and Research (IJSR)*, 5(4), P. 1049–1052. DOI: <https://doi.org/10.21275/v5i4.nov162776>
21. Shaposhnikov, A. I. (2021), "Feature-vector for the meanshift". *Proceedings of Tomsk State University of Control Systems and Radioelectronics*, 24(2), P. 34–38. DOI: <https://doi.org/10.21293/1818-0442-2021-24-2-34-38>
22. Tingting, S. (2020), "Application and research of DBSCAN optimization algorithm in big data analysis of experimental text". *Computer Science and Application*, 10 (05), P. 906-913. DOI: <https://doi.org/10.12677/csa.2020.105093>
23. Otani, N. (2020), "Pre-tokenization of multi-word expressions in cross-lingual word embeddings" / Otani N., Ozaki S., Zhao X., Li Y., St Johns M., Levin L. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. P. 4451–4464. DOI: <https://doi.org/10.18653/v1/2020.emnlp-main.360>
24. Ribeiro, E., Ribeiro, R., de Matos, D. (2019), "A multilingual and Multidomain Study on Dialog Act recognition using character-level tokenization". *Information*, 10(3), 94 p. DOI: <https://doi.org/10.3390/info10030094>
25. Slimane, F., Margner, V.(2014), "A new text-independent GMM writer identification system applied to Arabic handwriting". *International Conference on Frontiers in Handwriting Recognition*. 13 p. DOI: <https://doi.org/10.1109/icfhr.2014.124>
26. Belogorskaya, D. V. (2020), "Summarizing news texts using quantitative methods (TF-IDF)". *Proceedings of the VII (XXI) International Scientific and Practical Conference of Young Scientists*. DOI: <https://doi.org/10.17223/978-5-94621-901-3-2020-30>
27. Choi, E.A., Han, Y.E., Lee, S., Oh, M. (2019), "A comparison of TF and TF-IDF analysis for trends of Blockchain in health and Welfare". *Journal of the Korean Data And Information Science Society*, 30(5), P. 1025–1036. DOI: <https://doi.org/10.7465/jkdi.2019.30.5.1025>
28. Ghawi, R., Pfeffer, J. (2019), "Efficient hyperparameter tuning with grid search for text categorization using KNN approach with BM25 similarity". *Open Computer Science*, 9(1), P. 160–180. DOI: <https://doi.org/10.1515/comp-2019-0011>
29. Tinega, G. A., Mwangi, P. W., Rimiru, D. R. (2018), "Text mining in digital libraries using okapi BM25 model". *International Journal of Computer Applications Technology and Research*, 7(10), P. 398–406. DOI: <https://doi.org/10.7753/ijcatr0710.1003>
30. Ma, X., Hovy, E. (2016), "End-to-end sequence labeling via bi-directional LSTM-cnns-CRF". *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. P. 1064–1074. DOI: <https://doi.org/10.18653/v1/p16-1101>
31. Khalil, F., Pipa, P. D. (2021), "Transforming the generative pretrained transformer into augmented business text writer". *CC BY 4.0*. DOI: <https://doi.org/10.21203/rs.3.rs-1170589/v1>
32. Soyalt, G., Alar, A., Ozkanli, K., Yildiz, B. (2021), "Improving text classification with Transformer". *2021 6th International Conference on Computer Science and Engineering (UBMK)*. 12 p. DOI: <https://doi.org/10.1109/ubmk52708.2021.9558906>
33. Shen, Y., Liu J. (2021), "Comparison of text sentiment analysis based on Bert and word2vec". *2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer (ICFTIC)*. 17 p. DOI: <https://doi.org/10.1109/icftic54370.2021.9647258>
34. Pylkkönen, J., Ukkonen, A., Kilpikoski, J., Tamminen, S., Heikinheimo, H. (2021), "Fast text-only domain adaptation of RNN-Transducer Prediction Network". *Interspeech*. DOI: <https://doi.org/10.21437/interspeech.2021-1191>
35. Nismi, Mol, E. A., Santosh, Kumar M. B. (2020), "Study on impact of RNN, CNN and Han in text classification". *2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA)*. DOI: <https://doi.org/10.1109/accthpa49271.2020.9213231>
36. Zouzou, A., Azami, I. E. (2021), "Text sentiment analysis with CNN & GRU model using glove". *2021 Fifth International Conference On Intelligent Computing in Data Sciences (ICDS)*. DOI: <https://doi.org/10.1109/icds53782.2021.9626715>
37. Park, P.W. (2020), "Text-CNN based intent classification method for automatic input of intent sentences in chatbot". *The Journal of Korean Institute of Information Technology*, 18 (1), P. 19–25. DOI: <https://doi.org/10.14801/jkiit.2020.18.1.19>
38. Sun, S., Gao, Z., Huang, C., Yu, H. (2021), "Glove-FRCNN: Comprehensive network algorithm for Vespa Mandarin image-text extraction and classification". *2021 International Conference on Communications, Information System and Computer Engineering (CISCE)*. DOI: <https://doi.org/10.1109/cisce52179.2021.9445902>
39. Pan, N., Yao, W., Li, X. (2021), "Friends recommendation based on KBERT-CNN Text Classification Model". *International Joint Conference on Neural Networks (IJCNN)*. 2021. DOI: <https://doi.org/10.1109/ijcnn52387.2021.9533618>

Відомості про авторів / About the Authors

Шкурко Вячеслав Вікторович – Харківський національний університет радіоелектроніки, аспірант кафедри прикладної математики, Харків, Україна; e-mail: viacheslav.shkurko@nure.ua; ORCID ID: <https://orcid.org/0009-0003-3274-3941>

Поляков Андрій Олександрович – кандидат технічних наук, доцент, Харківський національний економічний університет імені Семена Кузнеця, доцент кафедри інформаційних систем, Харків, Україна; Харківський національний університет радіоелектроніки, доцент кафедри прикладної математики, Харків, Україна; e-mail: Andrii.Poliakov@nure.ua; ORCID ID: <https://orcid.org/0000-0003-1805-9011>

Shkurko Viacheslav – Kharkiv National University of Radio Electronics, Postgraduate Student Department of Applied Mathematics (AM), Kharkiv, Ukraine.

Poliakov Andrii – PhD (Engineering Sciences), Associate Professor, Simon Kuznets Kharkiv National University of Economics, Associated Professor at the Department of Information Systems, Kharkiv, Ukraine; Kharkiv National University of Radio Electronics, Associated Professor at the Department of Applied Mathematics, Kharkiv, Ukraine.

ORGANIZATION OF SOFTWARE AND NEURAL NETWORK ALGORITHMS FOR MACHINE ANALYSIS OF TEXTUAL DATA PRESENTED IN NATURAL LANGUAGE

The article addresses the problem of organizing software and neural network algorithms for natural language processing. Given the rapid growth of information flows and the limitations of computational resources, optimizing methods for processing short, unstructured messages and complex structured documents has become especially urgent. This study aims to develop a comprehensive method for organizing computer-aided analysis of text data, ensuring a balance between the accuracy of results and the efficiency of computational resource use. The research systematically examines modern approaches to tokenization, clustering, semantic-relevant search, and deep learning architectures, with particular attention to their adaptability under resource-constrained conditions. A multilevel methodology is suggested, combining the preliminary classification of text arrays, semantic clustering, and the use of a Bidirectional LSTM neural network model. Practical implementation of the method was tested through an automated text analysis application, demonstrating stable reduction in the loss function and acceptable resource consumption. The ability to adapt to different types of text data, reduced resource consumption while maintaining high-quality analysis, and suitability for deployment in environments with low computing capacity should be considered the main advantages of the developed approach. The scientific novelty of the article is substantiated by the integration of semantic-relevant clustering and lightweight deep learning techniques into a single optimized framework. The practical value of the study is the possibility of applying the proposed methodology in real-world information systems where limited hardware capabilities require efficient and adaptive text processing solutions.

Keywords: computer-aided analysis; text data; natural language; software algorithms; neural network architecture.

Бібліографічні описи / Bibliographic descriptions

Шкурко В. В., Поляков А. О. Організація програмних і нейромережевих алгоритмів машинного аналізу текстових повідомлень, поданих природною мовою. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 2 (32). С. 151–167. DOI: <https://doi.org/10.30837/2522-9818.2025.2.151>

Shkurko, V., Poliakov, A. (2025), "Organization of software and neural network algorithms for machine analysis of textual data presented in natural language", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (32), P. 151–167. DOI: <https://doi.org/10.30837/2522-9818.2025.2.151>

С. Бушуєв, Н. Бушуєва, Є. Лобок, Г. Мурованський

МЕНЕДЖМЕНТ ІННОВАЦІЙНИХ ПРОЄКТІВ НА ОСНОВІ ЗАСТОСУВАНЬ ШТУЧНОГО ІНТЕЛЕКТУ У ТУРБУЛЕНТНОМУ ОТОЧЕННІ

Предметом дослідження є менеджмент інноваційних проєктів на основі штучного інтелекту (ШІ) у турбулентному оточенні, що є надзвичайно актуальним для України станом через унікальне поєднання викликів і можливостей. **Метою дослідження** є створення концептуальної та математичної моделі ефективного управління інноваційними проєктами в умовах розвитку ШІ та турбулентного оточення. ШІ дозволяє прогнозувати ризики, оптимізувати ресурси і прискорювати адаптацію до змін, що є ключовим для підвищення конкурентоспроможності та стійкості в умовах війни та постконфліктного відновлення. **Задачі дослідження** охоплюють кілька етапів – вивчення глобальних трендів та особливостей стану України, визначення ключових принципів ефективного інноваційного менеджменту, таких як гнучкість, управління ризиками і технологічна інтеграція. **Методи дослідження** це аналіз останніх напрацювань, формування принципів управління, побудова та застосування концептуальної моделі створення цінностей інноваційних проєктів. **Результати дослідження:** розроблено концептуальну та математичну моделі менеджменту інноваційних проєктів та продуктів на основі ШІ в турбулентному оточенні. Концептуальна модель включає базові принципи такі як орієнтація на клієнта, цінності, гнучкість, управління ризиками, міждисциплінарність, експериментування, використання технологій, сталість і інноваційність. Запропонована математична модель спрямована на практичне застосування українськими компаніями для створення інноваційних продуктів і підвищення їхньої конкурентоспроможності. Розроблені моделі доводять, що ШІ підвищує ефективність інноваційних проєктів через автоматизацію, прогнозування і оптимізацію ресурсів на основі системи цінностей. Умови війни підкреслюють значущість ШІ для аналізу поля бою та логістиці. **Висновки.** Дослідження підтверджує, що менеджмент інноваційних проєктів на основі ШІ в турбулентному оточенні України є не лише актуальним, але й стратегічно важливим для відновлення та розвитку. Концептуальна та математична моделі, як продукт дослідження, забезпечують теоретичну та практичну основу для створення цінних інновацій, а результати підкреслюють потенціал ШІ для економії ресурсів, адаптації до змін і підвищення стійкості. Подальші дослідження можуть поглибити ці аспекти, сприяючи інноваційному прориву України.

Ключові слова: штучний інтелект; інноваційні проєкти; турбулентне середовище; управління ризиками; мультимодальні системи.

Вступ

Менеджмент інноваційних проєктів на основі штучного інтелекту у турбулентному оточенні в умовах України є актуальною темою досліджень. Україна, яка переживає складні часи через повномасштабне вторгнення та економічну нестабільність, водночас демонструє значний потенціал у сфері штучного інтелекту (ШІ). Управління інноваційними проєктами та продуктами на основі ШІ у таких умовах стає ключовим фактором для забезпечення конкурентоспроможності та стійкості [1, 2]. Розглянемо основні аспекти актуальності досліджень у цій галузі. В умовах війни та економічної нестабільності управління інноваційними проєктами вимагає нових підходів, зокрема використання ШІ для прогнозування ризиків, оптимізації ресурсів та прискорення адаптації до змін. Цифрова трансформація в Україні активно впроваджується ШІ у державне управління, оборону,

медицину та освіту, що підкреслює необхідність досліджень у сфері управління інноваціями. Сьогодні Україна посідає друге місце за кількістю ШІ-компаній у Центральній та Східній Європі за даними Міністерства цифрової трансформації України, що свідчить про значний потенціал для розвитку інноваційних проєктів.

ШІ дозволяє аналізувати великі обсяги даних для прогнозування ризиків та прийняття оперативних рішень у турбулентному середовищі. Дослідження спрямовані на розробку алгоритмів, які враховують специфіку військового часу, зокрема аналіз поля бою та розпізнавання ворожої техніки. Умови обмежених ресурсів вимагають ефективного управління фінансовими, людськими та технічними ресурсами. ШІ допомагає оптимізувати їх розподіл, що особливо актуально для оборонного та медичного секторів.

Дослідження зосереджуються на впровадженні ШІ у такі проєкти, як "Дія" та "Мрія", що дозволяє покращити якість державних послуг для громадян.

© С. Бушуєв, Н. Бушуєва, Є. Лобок, Г. Мурованський, 2025

Актуальним є дослідження механізмів підтримки стартапів, розвитку національної мовної моделі (LLM) та створення сприятливого правового середовища для інновацій.

Управління інноваційними проєктами вимагає врахування етичних норм та правових рамок, що особливо важливо в умовах використання ШІ у державному управлінні та обороні. Відтік кадрів через війну та недостатнє фінансування освіти у сфері ШІ. Складність залучення інвестицій у ШІ-проєкти через військову ситуацію.

Метою дослідження є створення концептуальної та математичної моделі ефективного управління інноваційними проєктами в умовах розвитку ШІ та турбулентного BANI оточення.

Управління інноваційними проєктами в умовах турбулентного середовища вимагає адаптації до швидких змін, високої невизначеності та складності. Застосування штучного інтелекту (ШІ) у таких проєктах може значно підвищити ефективність управління, оптимізувати ресурси та зменшити ризики.

Аналіз останніх досліджень

Аналіз ґрунтується на огляді наукових статей, матеріалів академічних репозиторіїв і вебресурсів. Використано методи узагальнення, порівняння та групування для виділення ключових тем.

Застосування штучного інтелекту в управлінні інноваційними проєктами, особливо в таких галузях, як будівництво, розробка програмного забезпечення та управління знаннями. В роботах [3, 4] визначено, що ШІ допомагає в управлінні ризиками, витратами та термінами, що є критично важливими для успіху проєктів у мінливих умовах.

Покращення прийняття рішень та автоматизація на основі ШІ дозволяє підприємцям приймати

рішення на основі даних, проактивно виявляти ризики, оптимізувати розподіл ресурсів та прогнозувати результати проєктів [4, 5].

Оптимізація ресурсів визначена у працях [6, 7] для ефективного розподілу ресурсів, прогнозування потреб у технічному обслуговуванні та підвищення загальної продуктивності проєктів на основі ШІ.

Підвищення конкурентоспроможності на основі інтеграції ШІ з гнучкими методологіями управління проєктами сприяє успішній реалізації проєктів та підвищенню конкурентоспроможності компаній [8, 9].

Важливо враховувати етичні та правові питання при використанні ШІ, щоб забезпечити відповідальне використання технологій [10, 11].

Успішний менеджмент інноваційних проєктів залежить від технічної інфраструктури, інтеграції з існуючими системами та готовності даних. В роботах [12, 13] визначено, що застосування штучного інтелекту в управлінні інноваційними проєктами в турбулентному середовищі відкриває нові можливості для підвищення ефективності, зменшення ризиків та оптимізації ресурсів. Проте, для досягнення успіху важливо враховувати етичні, правові та технічні аспекти, а також забезпечити розвиток відповідних навичок у командах [14, 15].

Базові принципи менеджменту інноваційних проєктів та продуктів

Для успішного управління інноваційними проєктами необхідно враховувати як теоретичні основи, так і практичні інструменти, які дозволяють ефективно реагувати на виклики сучасного бізнес-середовища (Таблиця 1).

Концептуальна модель менеджменту інноваційних проєктів та продуктів у сучасному оточенні наведена на рис. 1.

Таблиця 1. Аналіз тенденцій та особливостей України в управлінні інноваціями

Аспект	Характеристика	Значення для України
1	2	3
Глобальні тенденції	Швидкий розвиток ШІ Цифрова трансформація Посилення конкуренції	Необхідність адаптації до глобальних технологічних змін Можливість використання ШІ для підвищення конкурентоспроможності Загроза відставання у разі ігнорування тенденцій
Особливості України	Турбулентність та невизначеність Обмежені ресурси Потреба у відновленні та розвитку Євроінтеграційний курс	Ускладнення управління інноваційними проєктами Необхідність ефективного використання ресурсів Можливість за допомогою інновацій прискорити відновлення та розвиток Вимога адаптації до європейських стандартів

Продовження таблиці 1

1	2	3
Переваги використання ШІ в управлінні інноваціями	Автоматизація та оптимізація Аналіз даних та прогнозування Генерація ідей та креативність Прийняття рішень Управління ризиками	Підвищення ефективності та зменшення витрат Краще розуміння ринку та потреб споживачів Створення більш інноваційних продуктів Прийняття обґрунтованих рішень Зменшення ризиків та підвищення ймовірності успіху
Практичне значення дослідження	Допомога українським компаніям в управлінні інноваціями Підвищення конкурентоспроможності Розробка стратегій розвитку інноваційного підприємництва	Сприяння розвитку інноваційної економіки України Залучення інвестицій та створення нових робочих місць Інтеграція України у світовий інноваційний простір



Рис. 1. Концептуальна модель менеджменту інноваційних проєктів

Управління інноваційними проєктами вимагає унікального підходу, враховуючи високу невизначеність, ризики та необхідність швидкої адаптації. Нижче, у таблиці 2, наведено ключові принципи, які формують основу ефективного інноваційного менеджменту.

Базові принципи управління інноваційними проєктами та продуктами ґрунтуються на гнучкості, орієнтації на клієнта, управлінні ризиками та інтеграції сучасних технологій. Успішні інновації

вимагають не лише креативності, але й системного підходу, який балансує експерименти з дисципліною. Ключовим завжди залишається створення цінності, яка перевищує очікування ринку.

Концептуальна модель (рис.1) є основою для розробки практичних рекомендацій та методологій, які допоможуть компаніям успішно створювати інноваційні проєкти та продукти на основі ШІ в умовах BANI-середовища.

Таблиця 2. Ключові принципи ефективного інноваційного менеджменту

Принцип	Суть	Реалізація
Орієнтація на клієнта та цінність	Інновації мають вирішувати конкретні проблеми клієнтів або створювати нову цінність.	Використання методів дизайн-мислення для виявлення потреб користувачів. Постійний зворотний зв'язок через MVP (мінімально життєздатний продукт) та ітеративне тестування.
Гнучкість та адаптивність	Інноваційні проєкти потребують гнучких методологій управління для швидкого реагування на зміни.	Використання Agile, Scrum або Lean Startup. Короткі ітерації (спрінти) для швидкої корекції помилок.
Управління ризиками та невизначеністю	Інновації завжди супроводжуються ризиками, які потрібно систематично аналізувати.	Ідентифікація ризиків на ранніх етапах (наприклад, за допомогою SWOT-аналізу). Стратегії зменшення ризиків: диверсифікація, резервування ресурсів, сценарне планування.
Міждисциплінарні команди	Інновації народжуються на стику різних галузей знань.	Формування крос-функціональних команд (розробники, маркетинги, аналітики). Заохочення відкритої комунікації та спільного прийняття рішень.
Експериментування та швидке прототипування	Тестування гіпотез на ранніх етапах зменшує витрати на невдалі рішення.	Використання MVP для перевірки ринкового попиту. A/B-тестування різних версій продукту.
Використання сучасних технологій	Технології (AI, Big Data, IoT) прискорюють інноваційні процеси.	AI для прогнозування ринкових трендів або оптимізації логістики. Інструменти колаборації (наприклад, Jira, Trello) для координації команд.
Фокус на стійкості (sustainability)	Інновації мають враховувати екологічні, соціальні та економічні аспекти.	Розробка продуктів з низьким вуглецевим слідом. Використання циклічної економіки (переробка матеріалів).
Постійне навчання та інноваційна культура	Організації мають стимулювати творчість та експерименти.	Внутрішні хакатони та ідеї-марафони. Система винагород за інноваційні пропозиції.

Ціннісний підхід є ключовим для управління інноваційними проєктами та продуктами. Ланцюг цінностей це формат створення цінностей інновацій.

Формування ланцюгів створення цінностей для інноваційних проєктів є комплексним процесом, який вимагає інтеграції різних етапів, учасників та технологій. Використання штучного інтелекту дозволяє оптимізувати цей процес, забезпечуючи

швидкість, ефективність та адаптивність. Для успішної реалізації інноваційних проєктів необхідно враховувати специфіку ринку, залучати кваліфіковані кадри та забезпечувати фінансову підтримку. Це дозволить створити стійкі ланцюги цінностей, які забезпечать довгостроковий успіх інновацій.

Основні етапи формування ланцюга створення цінностей наведені у табл. 3.

Таблиця 3. Основні етапи формування ланцюга створення цінностей

Етап	Мета	Дії	Результат	Використання ІІІ
1	2	3	4	5
Дослідження та розробка (R&D)	Генерація нових ідей, технологій та продуктів.	Проведення фундаментальних та прикладних досліджень. Розробка прототипів та тестування.	Створення інноваційного продукту або технології.	Аналіз даних для виявлення трендів та потреб ринку. Оптимізація процесів розробки за допомогою машинного навчання. Автоматичне генерування коду та документації.
Дизайн та інженерія	Перетворення ідей у готові до впровадження рішення.	Розробка технічної документації. Створення MVP (мінімально життєздатного продукту).	Готовий до виробництва продукт.	Автоматизація проектування та моделювання за допомогою CAD/CAM систем з підтримкою ІІІ. Оптимізація дизайну для кращої продуктивності та ергономіки. Прогнозування та попередження помилок проектування.

Продовження таблиці 3

1	2	3	4	5
Виробництво	Масштабування інноваційного продукту.	Організація виробничих процесів. Впровадження індустрії 4.0 (IoT, роботизація, Big Data).	Вироблений продукт, готовий до комерціалізації.	Оптимізація логістики та управління якістю за допомогою ШІ. Прогнозування поломок обладнання та їх попередження. Автоматизація виробничих процесів за допомогою роботів та систем з ШІ.
Комерціалізація	Виведення продукту на ринок.	Маркетингові дослідження та стратегії. Побудова каналів збуту.	Продукт, який досягає цільової аудиторії.	Персоналізація маркетингу та реклами за допомогою аналізу даних про клієнтів. Прогнозування попиту та оптимізація запасів. Аналіз конкурентів та ринкової ситуації.
Післяпродажна підтримка	Забезпечення довгострокової цінності для клієнтів.	Технічна підтримка та оновлення продукту. Збір зворотного зв'язку для вдосконалення продукту.	Задоволені клієнти та довгострокові стосунки.	Аналіз зворотного зв'язку клієнтів для виявлення проблем та потреб. Персоналізована технічна підтримка за допомогою чат-ботів та віртуальних асистентів. Прогнозування потреб клієнтів та пропозиція відповідних послуг та продуктів.

Модель цінностей інноваційних проєктів та продуктів

Для структурованого аналізу запропоновано використовувати багатокритеріальний підхід, де кожна цінність оцінюється за шкалою (наприклад, від 0 до 10) з урахуванням вагових коефіцієнтів, що відображають пріоритети зацікавлених сторін.

Кроки аналізу.

1. Ідентифікація проєкту/продукту. Визначити конкретний інноваційний проєкт чи продукт (наприклад, Smart Logistics).

2. Визначення зацікавлених сторін. Хто отримує цінність (користувачі, бізнес, держава, суспільство)?

3. Оцінка за категоріями. Присвоїти бали за кожною цінністю на основі даних або експертних оцінок.

4. Формування ланцюга цінностей – дослідження, дизайн, виробництво, комерціалізація, підтримка та їх зважування. Призначити ваги категоріям залежно від контексту.

5. Інтегральна оцінка. Розрахувати загальний індекс цінності.

Запропонуємо модель створення цінностей інноваційних проєктів та продуктів.

$$V = w_s S + w_e E + w_n N + w_t T + w_m M, \quad (1)$$

де V – загальна цінність; S, E, N, T, M – бали за соціальну, економічну, екологічну, технологічну та етичну цінності; w_s, w_e, w_n, w_t, w_m – ваги (сума = 1).

Розглянемо приклад аналізу інноваційного проєкту.

Зробимо оцінку за категоріями:

1. Соціальна цінність ($S = 6/10$):

– Покращення доступності товарів, але прямий вплив на населення обмежений.

– Вага – $w_s = 0.2$.

2. Економічна цінність ($E = 9/10$):

– Зниження витрат на логістику, підвищення прибутковості бізнесу.

– Вага – $w_e = 0.3$.

3. Екологічна цінність ($N = 7/10$):

– Зменшення викидів через оптимізацію маршрутів, але виробництво IoT-датчиків має екологічний слід.

– Вага – $w_n = 0.25$.

4. Технологічна цінність ($T = 8/10$):

– Інтеграція ШІ та IoT, висока новизна.

– Вага – $w_t = 0.15$.

5. Етична цінність ($M = 7/10$):

– Прозорість і справедливість у розподілі ресурсів, але можливі питання щодо даних водіїв.

– Вага – $w_m = 0.1$.

Розрахунок.

$$V = 0.2 \cdot 6 + 0.3 \cdot 9 + 0.25 \cdot 7 + 0.15 \cdot 8 + 0.1 \cdot 7 = 7.55$$

Результат.

Індекс цінності = $7.55/10$ – високий рівень, з акцентом на економічну та технологічну цінність.

Висновки

Менеджмент інноваційних проєктів на основі штучного інтелекту (ШІ) у турбулентному оточенні є критично важливим для України в умовах війни, економічної нестабільності та цифрової трансформації. Дослідження підкреслює, що ШІ відіграє ключову роль у прогнозуванні ризиків, оптимізації ресурсів і прискоренні адаптації до змін, що особливо актуально для таких секторів, як оборона, медицина, державне управління та освіта. Друге місце України за кількістю ШІ-компаній у Центральній та Східній Європі свідчить про значний потенціал для розвитку інновацій.

Визначено базові принципи менеджменту інноваційних проєктів, які включають орієнтацію на клієнта, гнучкість, управління ризиками, міждисциплінарність, експериментування, використання технологій, фокус на стійкості та інноваційну культуру. Ці принципи, підкріплені такими методами, як Agile, Scrum, Lean Startup і ШІ-аналітика, формують основу для ефективного управління в умовах BANI-середовища.

Використання ШІ в управлінні інноваціями забезпечує автоматизацію процесів, точне прогнозування (80–85%), генерацію ідей, оптимізацію ресурсів і зниження ризиків. Практичні приклади, такі як "Дія" та "Мрія", демонструють покращення якості державних послуг, а аналіз поля бою та логістики підкреслюють адаптивність ШІ до воєнних умов.

Запропонована модель ланцюга цінностей показує, як ШІ оптимізує кожен етап – від аналізу даних і прогнозування до автоматизації виробництва та персоналізації маркетингу. Приклад "Smart Logistics" (індекс цінності 7.55/10) ілюструє високу економічну (9/10) і технологічну (8/10) цінність, що підтверджує потенціал ШІ для інноваційних проєктів.

Турбулентність (війна, обмежені ресурси) ускладнює управління інноваціями, але водночас створює можливості для прискорення відновлення та євроінтеграції. ШІ дозволяє ефективно використовувати обмежені ресурси, підвищувати конкурентоспроможність і залучати інвестиції, що є критично важливим для інноваційної економіки.

Основними перешкодами є відтік кадрів через війну, недостатнє фінансування освіти в сфері ШІ, складність залучення інвестицій і потреба в етичних та правових рамках. Ці виклики вимагають системного підходу для підтримки стартапів і створення сприятливого середовища.

Розроблена концептуальна модель менеджменту інноваційних проєктів на основі ШІ є основою для практичних рекомендацій, спрямованих на підвищення конкурентоспроможності українських компаній, інтеграцію в глобальний інноваційний простір і створення нових робочих місць.

Напрямки подальших досліджень

Подальші дослідження мають зосередитися на створенні алгоритмів ШІ, адаптованих до умов військового часу (наприклад, аналіз поля бою, розпізнавання техніки) та обмежених ресурсів.

Необхідно провести аналіз повернення інвестицій (ROI) для ШІ-проєктів в Україні, враховуючи витрати на технології і потенційну економію, щоб обґрунтувати їхню доцільність для бізнесу та держави.

Аналіз можливостей застосування концептуальної моделі в інших галузях (сільське господарство, енергетика, логістика) для прискорення відновлення України та підвищення її стійкості.

Дослідження механізмів державної та міжнародної підтримки ШІ-стартапів (гранти, інкубатори), включаючи оцінку впливу на залучення інвестицій і створення робочих місць.

Список літератури

1. Godoy M. F. Ribas Filho D. Facing the BANI World. *International Journal of Nutrology*, 14(2), 33 p. 2021. DOI: <https://doi.org/10.1055/s-0041-1735848>
2. Бушуєв С. Д., Тихонович Ю., Черниш О., Сухонос Н., Халілов А. Креативні принципи управління інноваційними проєктами в середовищі BANI. *Управління розвитком складних систем*, (57), Р. 6–11. DOI: 2024. <https://doi.org/10.32347/2412-9933.2024.57.6-11>

3. Fernández J., Moreno J., Vergara-González E., Iglesias G. Bibliometric Analysis of the Application of Artificial Intelligence Techniques to the Management of Innovation Projects. *Applied Sciences*. 12(22). 2022. DOI: <https://doi.org/10.3390/app122211743>
4. Aliyev A. Project Management Revolution: How Startups Integrate Artificial Intelligence to Succeed. *Ancient land*. 2024. DOI: <https://doi.org/10.36719/2706-6185/42/146-149>
5. Bushuyev S., Ivko A. Construction of models and application of syncretic innovation project management in the era of artificial intelligence. *Eastern-European Journal of Enterprise Technologies*. 2024. 33 (129), P. 44–54. DOI: <https://doi.org/10.15587/1729-4061.2024.306436>
6. Bushuev S., Ivko A., Tikhonovych Y. Syncretic project management in the era of artificial intelligence explosion. *Environmental safety and natural resources*. 2024. 49(1), P. 85–98. DOI: <https://doi.org/10.32347/2411-4049.2024.1.85-98>
7. Prasetyo M., Peranginangin R., Martinovic N., Ichsan M., Wicaksono H. Artificial intelligence in open innovation project management: A systematic literature review on technologies, applications, and integration requirements. *Journal of Open Innovation: Technology, Market, and Complexity*. 2025. Volume 11, Issue 1, 100445 p. DOI: <https://doi.org/10.1016/j.joitmc.2024.100445>
8. Reznikov R. Enhancing Project Management Success through Artificial Intelligence. *International Journal of Multidisciplinary Research and Growth Evaluation*. 2025. DOI: <https://doi.org/10.54660/ijmrge.2025.6.1.1036-1046>
9. Taboada I., Daneshpajouh A., Toledo N., De Vass T. Artificial Intelligence Enabled Project Management: A Systematic Literature Review. *Applied Sciences*. 2023. 13(8), 5014 p. DOI: <https://doi.org/10.3390/app13085014>
10. J. M. Mesa Fernández Bibliometric analysis of the application of artificial intelligence techniques to the management of innovation projects / J. M. Mesa Fernández et al. *Applied sciences*. 2022. Vol. 12, No. 22. 11743 p. DOI: <https://doi.org/10.3390/app122211743>
11. El Khatib M., Al Falasi A. Effects of artificial intelligence on decision making in project management. *American journal of industrial and business management*. 2021. Vol. 11, No. 03. P. 251–260. DOI: <https://doi.org/10.4236/ajibm.2021.113016>
12. Elmousalami H. H. Comparison of artificial intelligence techniques for project conceptual cost prediction: a case study and comparative analysis. *IEEE transactions on engineering management*. 2021. Vol. 68, No. 1. P. 183–196. DOI: <https://doi.org/10.1109/tem.2020.2972078>
13. L. Kahn Modelling hybrid human-artificial intelligence cooperation: a call center customer service case study / L. Kahn et al. *IEEE Xplore logo*. 2020. 3075 p. DOI: <https://doi.org/10.1109/BigData50022.2020.9377747>
14. Prifti V. Optimizing project management using artificial intelligence. *European journal of formal sciences and engineering*. 2022. Vol. 5, No. 1. 29 p. DOI: <https://doi.org/10.26417/667hri67>
15. Wachnik B. Analysis of the use of artificial intelligence in the management of Industry 4.0 projects. *The perspective of Polish industry. Production engineering archives*. 2022. Vol. 28, No. 1. P. 56–63. DOI: <https://doi.org/10.30657/pea.2022.28.07>

References

1. Godoy, M. F. Ribas, Filho D. (2021), "Facing the BANI World". *International Journal of Nutrology*, 14(2), 33 p. DOI: <https://doi.org/10.1055/s-0041-1735848>
2. Bushuyev, S., Tykchonovych, Ju., Chernysh, O., Sukhonos, N., Khalilov, A., (2024). "Creative principles for managing innovation projects in BANI environment". *Management of Development of Complex Systems*, 57, P. 6–11. DOI: [10.32347/2412-9933.2024.57.6-11](https://doi.org/10.32347/2412-9933.2024.57.6-11)
3. Fernández, J., Moreno, J., Vergara-González, E., Iglesias, G. (2022), "Bibliometric Analysis of the Application of Artificial Intelligence Techniques to the Management of Innovation Projects". *Applied Sciences*. 12(22). DOI: <https://doi.org/10.3390/app122211743>
4. Aliyev, A. (2024), "Project Management Revolution: How Startups Integrate Artificial Intelligence to Succeed". *Ancient land*. DOI: <https://doi.org/10.36719/2706-6185/42/146-149>
5. Bushuyev, S., Ivko, A. (2024), "Construction of models and application of syncretic innovation project management in the era of artificial intelligence". *Eastern-European Journal of Enterprise Technologies*. 33 (129), P. 44–54. DOI: <https://doi.org/10.15587/1729-4061.2024.306436>
6. Bushuev, S., Ivko, A., Tikhonovych, Y. (2024), "Syncretic project management in the era of artificial intelligence explosion". *Environmental safety and natural resources*. 49(1), P. 85–98. DOI: <https://doi.org/10.32347/2411-4049.2024.1.85-98>

7. Prasetyo, M., Peranginangin, R., Martinovic, N., Ichsan, M., Wicaksono, H. (2025), "Artificial intelligence in open innovation project management: A systematic literature review on technologies, applications, and integration requirements". *Journal of Open Innovation: Technology, Market, and Complexity*. Volume 11, Issue 1, 100445 p. DOI: <https://doi.org/10.1016/j.joitmc.2024.100445>
8. Reznikov, R. (2025), "Enhancing Project Management Success through Artificial Intelligence". *International Journal of Multidisciplinary Research and Growth Evaluation*. DOI: <https://doi.org/10.54660/ijmrge.2025.6.1.1036-1046>
9. Taboada, I., Daneshpajouh, A., Toledo, N., De Vass, T. (2023), "Artificial Intelligence Enabled Project Management: A Systematic Literature Review". *Applied Sciences*. 2023. 13(8), 5014 p. DOI: <https://doi.org/10.3390/app13085014>
10. J. M. Mesa, Fernández (2022), "Bibliometric analysis of the application of artificial intelligence techniques to the management of innovation projects" / J. M. Mesa Fernández et al. *Applied sciences*. Vol. 12, No. 22. 11743 p. DOI: <https://doi.org/10.3390/app122211743>
11. El Khatib, M., Al, Falasi A. (2021), "Effects of artificial intelligence on decision making in project management". *American journal of industrial and business management*. Vol. 11, No. 03. P. 251–260. DOI: <https://doi.org/10.4236/ajibm.2021.113016>
12. Elmousalami, H. H. (2021), "Comparison of artificial intelligence techniques for project conceptual cost prediction: a case study and comparative analysis". *IEEE transactions on engineering management*. Vol. 68, No. 1. P. 183–196. DOI: <https://doi.org/10.1109/tem.2020.2972078>
13. Kahn, L. (2020), "Modelling hybrid human-artificial intelligence cooperation: a call center customer service case study" / L. Kahn et al. *IEEE Xplore logo*. 3075 p. DOI: <https://doi.org/10.1109/BigData50022.2020.9377747>
14. Prifti, V. (2022), "Optimizing project management using artificial intelligence". *European journal of formal sciences and engineering*. Vol. 5, No. 1. 29 p. DOI: <https://doi.org/10.26417/667hri67>
15. Wachnik, B. (2022), "Analysis of the use of artificial intelligence in the management of Industry 4.0 projects". *The perspective of Polish industry. Production engineering archives*. Vol. 28, No. 1. P. 56–63. DOI: <https://doi.org/10.30657/pea.2022.28.07>

Надійшла (Received) 08.05.2025

Відомості про авторів / About the Authors

Бушувєв Сергій Дмитрович – доктор технічних наук, професор, Київський національний університет будівництва і архітектури, професор кафедри управління проектами, Київ, Україна; e-mail: sbushuyev@ukr.net; ORCID ID: <https://orcid.org/0000-0002-7815-8129>

Бушувєва Наталія Сергіївна – доктор технічних наук, Київський національний університет будівництва і архітектури, професор кафедри управління проектами, Київ, Україна; e-mail: Natbush@ukr.net; ORCID ID: <https://orcid.org/0000-0002-4969-7879>

Лобок Євген Анатолійович – Київський національний університет будівництва і архітектури, аспірант кафедри управління проектами, Київ, Україна; e-mail: ownereugene@gmail.com, ORCID ID: <https://orcid.org/0009-0006-9841-1132>

Мурованський Глеб Анатолійович – Київський національний університет будівництва і архітектури, аспірант кафедри управління проектами, Київ, Україна; e-mail: 4648800@gmail.com, ORCID ID: <https://orcid.org/0009-0003-4610-9087>

Bushuyev Sergiy – Doctor of Sciences (Engineering), Professor, Kyiv National University of Construction and Architecture, Professor management department, Kyiv, Ukraine.

Bushuyeva Natalia – Doctor of Technical Sciences (Engineering), Professor, Kyiv National University of Construction and Architecture, Professor of the Department of Project Management, Kyiv, Ukraine.

Lobok Yevgen – Kyiv National University of Construction and Architecture, PhD student, Department of Project Management, Kyiv, Ukraine.

Murovansky Hleb – Kyiv National University of Construction and Architecture, PhD student, Department of Project Management, Kyiv, Ukraine.

MANAGEMENT OF INNOVATIVE PROJECTS BASED ON ARTIFICIAL INTELLIGENCE APPLICATIONS IN A TURBULENT ENVIRONMENT

The subject of this study is the management of innovative projects based on artificial intelligence (AI) in a turbulent environment, which is a particularly relevant situation for Ukraine due to its unique combination of challenges and opportunities. The full-scale invasion of Russia, economic instability in Ukraine, the outflow of personnel, and limited resources create a complex BANI environment (Brittle, Anxious, Nonlinear, Incomprehensible), which complicates traditional management approaches. The purpose of the study is to create a conceptual and mathematical model of effective management of innovative projects in the conditions of AI development and a turbulent BANI environment. At the same time, Ukraine demonstrates significant potential in the field of AI application, ranking second in the number of AI companies in Central and Eastern Europe according to the Ministry of Digital Transformation. The objectives of the study are to develop models for managing innovative projects in the context of digital transformation, covering the state administration "Diya", defence, medicine and education. AI allows you to predict risks, optimise resources and accelerate adaptation to changes, which is key to increasing competitiveness and resilience in war and post-conflict recovery. The relevance is also enhanced by the need to develop an ethical and legal framework for the use of AI, support start-up's and create a national language model (LLM) that meets Ukraine's European integration aspirations. The product of the study is a conceptual and mathematical model for the management of innovative projects and products based on AI in a turbulent environment. A mathematical model is proposed that is aimed at practical application by Ukrainian companies to create innovative products and increase their competitiveness. This model includes basic principles such as customer orientation, values, flexibility, risk management, interdisciplinarity, experimentation, use of technology, sustainability and innovation culture. The research process covers several stages – studying global trends and specifics of the state of Ukraine, identifying key principles of effective innovation management, such as flexibility, risk management and technological integration. The results of the study include models that prove that AI increases the efficiency of innovation projects through automation, forecasting and optimisation of resources based on a value system. War conditions emphasise the importance of AI for battlefield analysis and logistics. Conclusions – the study confirms that AI-based innovation project management in the turbulent environment of Ukraine is not only relevant, but also strategically important for recovery and development. The conceptual and mathematical model, as a product of the study, provides a theoretical and practical basis for creating valuable innovations, and the results emphasise the potential of AI to save resources, adapt to change and increase resilience. Further research can deepen these aspects, contributing to Ukraine's innovative breakthrough.

Keywords: artificial intelligence, innovative projects, turbulent environment, risk management, multimodal systems.

Бібліографічні описи / Bibliographic descriptions

Бушуєв С. Д., Бушуєва Н. С., Лобок Є. А., Мурованський Г. А. Менеджмент інноваційних проєктів на основі застосувань штучного інтелекту у турбулентному оточенні. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 2 (32). С. 168–176. DOI: <https://doi.org/10.30837/2522-9818.2025.2.168>

Bushuyev, S., Bushuyeva, N., Lobok, Y., Murovansky, H. (2025), "Management of innovative projects based on artificial intelligence applications in a turbulent environment", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (32), P. 168–176. DOI: <https://doi.org/10.30837/2522-9818.2025.2.168>

УДК 621.891.

DOI: <https://doi.org/10.30837/2522-9818.2025.2.177>

О. БИЗКРОВНИЙ, К. СМЕЛЯКОВ

МОДЕЛІ КЛАСИФІКАТОРА ПЕРЕДУМОВ ВИНИКНЕННЯ ДТП ДЛЯ ПЕРЕДБАЧЕННЯ НЕБЕЗПЕЧНИХ СИТУАЦІЙ НА ПЕРЕХРЕСТЯХ

Предметом статті є передумови виникнення дорожньо-транспортних пригод на перехрестях та ділянках з обмеженою видимістю; використання моделей комп'ютерного зору для завдань класифікації передумов виникнення ДТП і визначення ефективності їх застосування для роботи в режимі реального часу. **Мета дослідження** – порівняння моделей комп'ютерного зору для завдання класифікації передумов виникнення ДТП для роботи в режимі реального часу; порівняння моделей з використанням платформи *Jetson TX2*; визначення ефективності роботи зазначеного підходу з метою створення сигналів для водія про небезпеку в режимі реального часу. **Завдання:** дослідження моделей комп'ютерного зору *Detectron2*, *YOLOv7* для завдання класифікації передумов ДТП на зображенні в розрізі швидкодії моделей, зручності створення набору даних, тренування моделей та їх використання; порівняння *YOLOv8* і *DetectNet_v2* на одноплатному комп'ютері *Jetson TX2* у розрізі швидкодії. **Методи дослідження:** тренування та використання моделей машинного навчання та методи моделювання небезпечних ситуацій за допомогою програмного забезпечення *BeamNG.tech*, *CARLA*; порівняльний аналіз результатів застосування моделей із застосуванням метрик для оцінювання їх ефективності. Основними **результатами дослідження** є виявлення найбільш ефективної моделі для завдань класифікації на одноплатному комп'ютері *Jetson TX2-DetectNet_v2*; позитивна оцінка ефективності використання запропонованого підходу для попередження водія про небезпеку в режимі реального часу, проте з огляду на розмір навчального набору даних та складність його підготовки. **Висновки.** Розглянуто алгоритми комп'ютерного зору, зокрема *Detectron2*, *YOLOv7* та *DetectNet_v2*. Виявлено, що модель *YOLOv7* є кращою порівняно з *Detectron2* для завдання виявлення передумов ДТП на зображенні, проте зі свого боку *DetectNet_v2* є більш ефективним для застосування на одноплатному комп'ютері *Jetson TX2*, якщо порівнювати з *YOLOv7*. Додатково на підставі експериментів з'ясовано, що застосування окресленого підходу для передбачення передумов ДТП є проблематичним через складнощі у створенні навчального набору даних – варіативність сценаріїв передумов.

Ключові слова: класифікація об'єктів на зображенні; комп'ютерний зір; машинне навчання; штучний інтелект; інформаційні технології; *Nvidia Jetson*; передбачення ДТП; *YOLO*; *DetectNet_v2*; *Detectron2*.

Вступ

Зростання кількості транспортних засобів та їх швидкості є одним із чинників підвищення частоти дорожньо-транспортних пригод, що зі свого боку також є фактором збільшення кількості жертв і постраждалих унаслідок ДТП. Сучасні методи запобігання небезпечних ситуацій на дорогах спрямовані переважно на безпеку окремих автомобілів і не беруть до уваги учасників дорожнього руху, які перебувають поза зоною дії вбудованих механізмів безпеки транспортних засобів.

Однією з проблем індивідуальних засобів безпеки автомобіля є неможливість попередження водія про перешкоди, розташовані за межами досяжності окремих засобів безпеки. Отже, виникає необхідність у розробленні технологій, що дають змогу виявляти та попереджати водіїв про потенційні небезпеки на дорогах у режимі реального часу, які можуть працювати поза межами зони

застосування індивідуальних засобів забезпечення безпеки автомобіля.

З огляду на сучасні тенденції у сфері технологій зазначена система має бути розроблена з використанням машинного навчання. Ще одним важливим фактором безпеки дорожнього руху є здатність водія швидко й адекватно реагувати на зміну ситуації та отриману інформацію. Адже швидкість виявлення та спосіб повідомлення водія також має вплив на ефективність системи.

Отже, розроблення та дослідження нових методів прогнозування небезпечних ситуацій і оперативного оповіщення про них є актуальним і важливим завданням.

Аналіз останніх досліджень і публікацій

Методи виявлення передумов і чинників, що викликають автомобільні аварії, використовують досягнення в сфері комп'ютерного зору та глибокого

навчання. Зростання кількості досліджень у цьому напрямі свідчить про важливість раннього виявлення потенційно аварійно небезпечних умов. Визначені передумови застосовують для управління безпекою дорожнього руху. Інноваційне дослідження [1] зосереджене на виявленні різних видів дорожніх перешкод, таких як лежачі поліцейські та вибоїни, зображення, отримані за допомогою камер ZED і глибокої нейронної мережі.

Окреслений підхід може використовуватись для виявлення перешкод на певних ділянках руху й за допомогою додаткового механізму, сповіщати

інших водіїв про виявлені небезпеки на дорожньому полотні, що дасть змогу учасникам дорожнього руху знати більше про дорожні умови, що сприяє кращому прийняттю рішень і зменшенню аварійності.

Крім цього, у роботі [2] було застосовано акселерометричні дані в поєднанні з обробленням зображень для покращення аналізу дорожнього покриття та виявлення перешкод, що можуть бути причинами аварій. У зазначеному дослідженні наголошено на важливості як візуальних, так і фізичних чинників для аналізу, які можуть попередити водія про потенційну небезпеку (рис. 1).

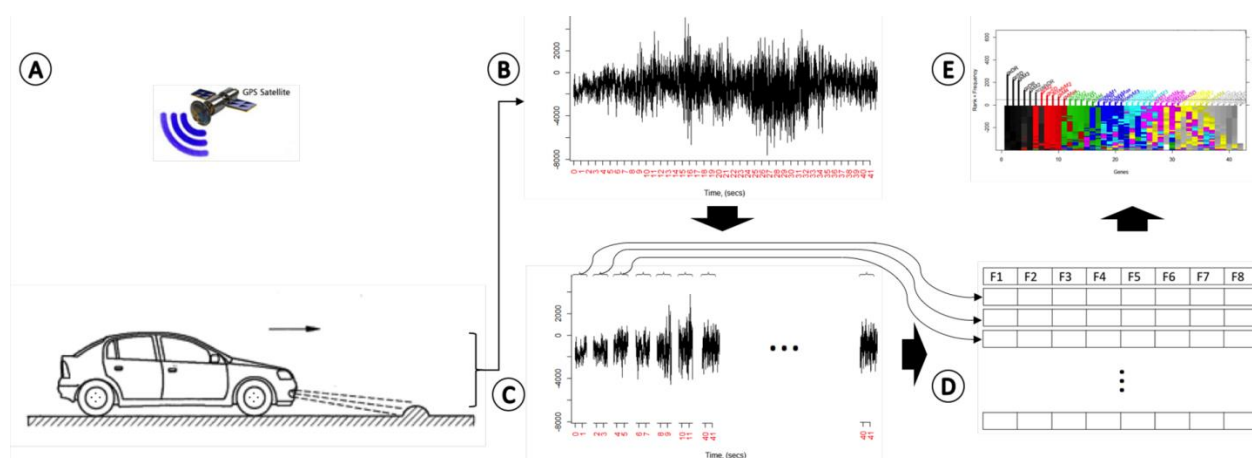


Рис. 1. Поєднання даних із сенсорів та візуального складника для завдань прогнозування небезпек [2]

Такі інтегровані методи дають змогу здійснювати моніторинг і допомагають водієві реагувати на перешкоди в режимі реального часу. Наявні дослідження сприяють розвитку та створенню автономних транспортних засобів.

У праці [3] використовується непараметричний метод оцінювання випадкової величини [4] для аналізу просторового розподілу ДТП на основі інформації, отриманої за допомогою камер спостереження. Цю методологію використовують для визначення кластерів аварій, виявлення ділянок, що потребують більш суворого контролю, і потенційних поліпшень інфраструктури. Зазначене наголошує на необхідності поєднання оцінювання історичних даних із сучасними методами оброблення зображень для більш повного розуміння аварійно небезпечних середовищ.

У роботі [5] запропоновано новий метод автоматичного виявлення дорожньо-транспортних пригод за допомогою машинного зору, зокрема аналізу записів із придорожніх камер спостереження для встановлення пошкоджених транспортних засобів, що свідчать про аварію. Цей тип аналізу дає

змогу швидко повідомляти аварійні служби про ДТП порівняно з традиційними методами інформування. Аналогічно підхід, розроблений у роботі [6], упроваджує глибоке навчання для вилучення візуальних ознак із подальшою ідентифікацією часових патернів, досягаючи вражаючого рівня точності у виявленні дорожньо-транспортних пригод на основі відеоданих. Ці дослідження надають важливого значення обробленню зображень у реальному часі для підвищення ефективності реагування на надзвичайні ситуації на дорогах. Описана в дослідженні [7] інтеграція підходів сегментації зображень з *fuzzy*-теорією [8] для оцінювання зон пошкоджень автомобілів у аваріях також показує, як оброблення зображень може бути використане для детального аналізу пошкоджень після аварії, забезпечуючи більш чітку картину наслідків зіткнення. Поєднання ефективних методів сегментації з *fuzzy*-логікою підвищує точність ідентифікації пошкоджених ділянок транспортними засобами й цим надає цінну інформацію для реконструкції та аналізу аварії.

Застосування систем кооперативної транспортної інфраструктури, як зазначено в дослідженні [9], відтворює, як дані зображень і датчиків можуть допомагати в автоматичному виявленні ДТП і внаслідок цього підвищувати ефективність реагування й безпеку дорожнього руху загалом.

У праці [10] запропоновано метод, що інтегрує аналіз зображень за допомогою згорткових нейронних мереж (CNN) для оптимізації механізмів реагування на ДТП, покращуючи як управління дорожнім рухом, так і заходи безпеки.

Інтеграція різних типів даних за допомогою передових технологій має важливе значення для сучасних систем управління дорожнім рухом, спрямованих на зниження аварійності. Оскільки ця галузь продовжує розвиватися, долучення оброблення зображень у методологію запобігання аваріям обіцяє не лише покращити прогнозування потенційних інцидентів, але й оптимізувати екстрене реагування на аварії. У розглянутих дослідженнях наголошено на вирішальній ролі технологічного прогресу в зменшенні ризиків, пов'язаних із дорожньо-транспортними пригодами.

Нерозв'язані проблеми та мета дослідження

У проаналізованих вище роботах розглянуто ДТП та реагування на неї як на постфактум-подію, тобто коли дорожньо-транспортна пригода вже відбулася. Необхідна технологія, яка зможе попередити водія про небезпеку в режимі реального часу, щоб запобігти аварії. Такий підхід дасть змогу суттєво зменшити кількість ДТП, оскільки інформація надходитиме безпосередньо з місця розташування автомобіля й транслюватиме ситуацію, що виникає безпосередньо в певний проміжок часу, а не базуватиметься на історичних даних певних ділянок руху.

Мета дослідження – порівняння моделей комп'ютерного зору для завдання класифікації передумов виникнення ДТП з метою роботи в режимі реального часу. Порівняння моделей з використанням платформи *Jetson TX2* [11]; визначення ефективності роботи зазначеного підходу для створення сигналів для водія про небезпеку в режимі реального часу. Для досягнення мети необхідно **виконати такі завдання**: дослідити моделі комп'ютерного зору *Detectron2*, *YOLOv7* для завдання класифікації передумов ДТП на зображенні в розрізі швидкодії

моделей, зручності створення набору даних, тренування моделей та їх використання; порівняння *YOLOv8* та *DetectNet_v2* на одноплатному комп'ютері *Jetson TX2* у розрізі швидкодії.

Матеріали й методи

З метою порівняння моделей комп'ютерного зору для завдань класифікації в режимі реального часу розглядаються моделі *Detectron 2* [12] та *YOLOv7* [13]. Після визначення кращого алгоритму в цій статті обрану модель або сімейство буде розгорнуто на *Jetson TX2* для оцінювання ефективності роботи на одноплатному комп'ютері.

Створення набору даних

Для створення типових сценаріїв дорожніх ситуацій було використано програмне забезпечення *BeamNG.tech* [14], де розроблялися скрипти, за якими автомобілі мали рухатись і влаштовувати небезпечні ситуації. Для спрощення класифікації небезпечних ситуацій було визначено, що для експерименту застосовуватиметься небезпечна ситуація повороту ліворуч на перехресті через декілька зустрічних смуг, де одна з них пропускає водія, що повертає, проте крайня права смуга пуста, або нею хтось рухається і не знає, чому сусідня смуга зупинена (рис. 2).

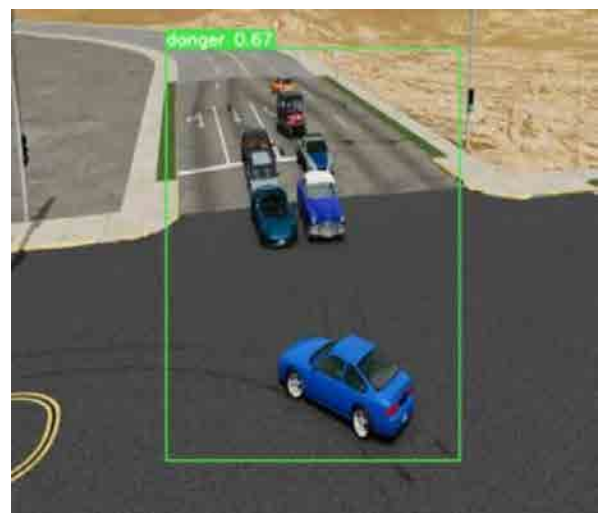


Рис. 2. Типова ситуація, під час якої водії, що мають пріоритет у русі, призупиняються, щоб дати іншому транспортному засобу зробити поворот

Описана ситуація поширена та керується більш етичними нормами, ніж правилами дорожнього руху.

У деяких ситуаціях такий вияв поваги до водія, який хоче повернути ліворуч, завершується трагічно.

За допомогою створених сценаріїв ДТП з використанням програмного забезпечення *BeamNG.tech* було записано відео, яке потім застосовувалося для формування тренувального набору даних. Стратегія створення набору даних для дослідження передбачала такі кроки:

- симуляція щонайменше п'яти різних перехресть із різними конфігураціями автомобільного руху та різною кількістю зустрічних смуг;
- класифікація зображень лише за двома категоріями: "небезпека" й "відсутність небезпеки".

Навчальний набір даних містить 300 зображень, з яких 150 належать до класу "небезпека", а 150 – до класу "відсутність небезпеки". На рис. 3 наведено приклад зображення з набору даних.

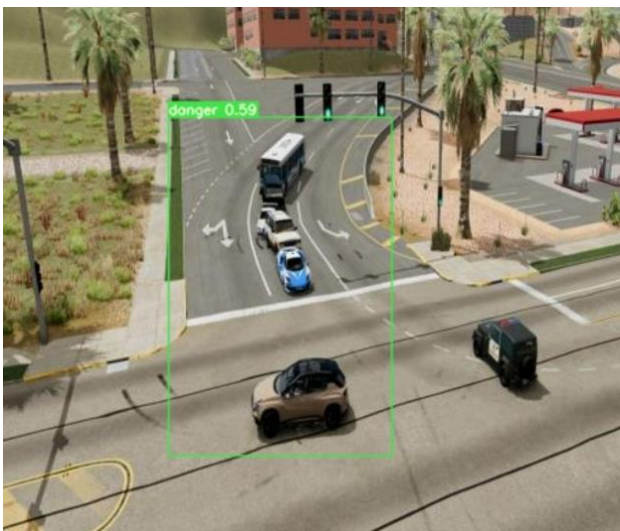


Рис. 3. Приклад зображення з тренувального набору даних, який описує клас "небезпека"

Тобто створений набір даних слугує для визначення патернів того, що є передумовами виникнення зазначеного типу ДТП.

Метрики валідації моделей

Результати проведеного експерименту були визначені за допомогою наведених далі метрик: *mAP* (*Mean Average Precision*) *with IoU*, час тренування моделей, використана пам'ять відеокарти, розмір тренуваної моделі, *Loss Box*, *Loss Cls*, час для розпізнавання класів на зображенні, *Recall*.

Mean Average Precision (*mAP*) *with IoU* – це метрика, що узагальнює якість класифікації об'єкта за кількома критеріями: наскільки правильно модель ідентифікує об'єкти (класифікація) та наскільки просторово точно вона їх локалізує (локалізація через *IoU*).

IoU між передбаченим об'єктом B_p та істинною розміткою B_g визначається як відношення:

$$IoU(B_p, B_g) = \frac{area(B_p \cap B_g)}{area(B_p \cup B_g)}, \quad (1)$$

де $area(B_p \cap B_g)$ – площа перетину двох прямокутників; $area(B_p \cup B_g)$ – площа об'єднання двох прямокутників.

IoU (*Intersection over Union*) є показником оцінювання ступеня перекриття двох обмежувальних рамок або сегментаційних масок. Якщо передбачення ідеальне, тоді *IoU* дорівнює 1, а за повної невідповідності – 0. Значення *IoU* визначається рівнем перекриття між цими межами. Тому для оцінювання точності модель аналізують у разі різних значень *IoU*, щоб з'ясувати, за яких умов досягається найкраща продуктивність.

Метрика точності (*Precision*) має такий вигляд:

$$P = \frac{TruePositive}{TruePositive + FalsePositive}, \quad (2)$$

де *TruePositive* – кількість правильно передбачених об'єктів; *FalsePositive* – кількість хибнопозитивних передбачень, коли модель передбачила об'єкт, якого насправді немає.

Формула *Recall* (повноти) використовується для оцінювання того, наскільки добре модель виявляє всі релевантні об'єкти. Вона визначається як відношення кількості правильно передбачених позитивних випадків (*True Positives*) до загальної кількості фактичних позитивних випадків (сума *True Positives* і *False Negatives*):

$$Recall = \frac{TP}{TP + FN}, \quad (3)$$

де *TP* – кількість правильно передбачених об'єктів; *FN* – кількість втрачених об'єктів. Реальний об'єкт існує, проте модель не змогла його класифікувати.

Loss Box – це функція втрат, що оцінює, наскільки точно передбачені обмежувальні рамки прилягають до реального об'єкта. Зазвичай використовується регресійна втрата, така як *L1* або

smooth L1. Формула для обчислення цієї втрати має такий вигляд:

$$L_{box}(t^u, v) = \sum_{i \in \{x, y, w, h\}} L_1^{smooth}(t_i^u - v_i), \quad (4)$$

де $t^u = (t_x^u, t_y^u, t_w^u, t_h^u)$ – позначає вектор передбачених параметрів рамки для об'єкта класу u ;

$v = (v_x, v_y, v_w, v_h)$ – вектор відповідних істинних координат;

$i \in \{x, y, w, h\}$ – описує ітерування над компонентами координат;

L_1^{Smooth} – функція втрат.

Loss Cls (classification loss) оцінюють правильність класифікації кожного передбаченого класу. Зазвичай для цього використовується функція втрат на основі крос-ентропії. Формула для обчислення цієї втрати має такий вигляд:

$$L_{cls}(p, y) = -[y \log(p) + (1 - y) \log(1 - p)], \quad (5)$$

де $p \in \{0, 1\}$ – передбачена класифікація; $y \in \{0, 1\}$ – істинна класифікація, де 1 – позитивний клас, 0 – негативний.

Для порівняння моделей після розгортання їх на одноплатному комп'ютері *Jetson TX2* було використано

інструментарій *trtexes*, що дає змогу досягти швидкодії та інші характеристики після запуску моделей комп'ютерного зору на *Jetson TX2* для порівняння.

Результати досліджень

У табл. 1 подано результати тренування та порівняння моделей *YOLOv7* і *Detectron2*. Для тренування було використано *Google Colab* середовище з GPU: *Tesla T4*.

Графіки метрик втрат моделі *YOLOv7* та *Average Precision with IoU* зображені на рис. 4 і 5.

Таблиця 1. Порівняння отриманих метрик під час тренування моделей *YOLOv7* та *Detectron2*

Характеристика	YOLOv7	Detectron2
System RAM, GB	5.7	3.7
GPU RAM, GB	11.4	8.5
SSD, GB	24.9	25.9
Training time, mins	15.9	51
AP@.5:95	0.583	0.651
AP@.5	0.926	0.909
Inference time, s	0.1916668	0.198093
Model size, MB	74.8	815

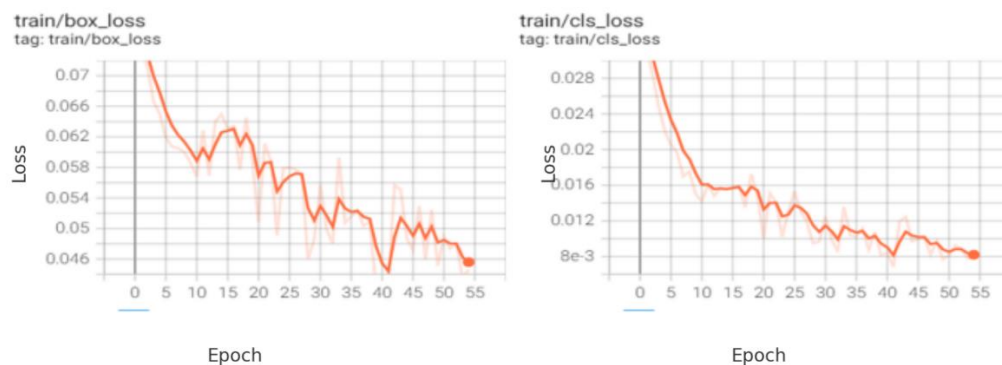


Рис. 4. Графіки функцій втрат моделі *YOLOv7*

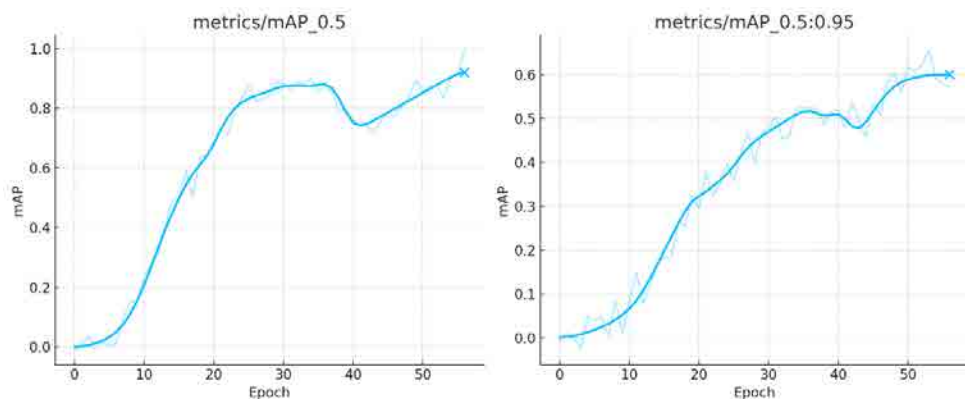


Рис. 5. Графіки mAP з $IoU = 0.5$ та $IoU = 0.95$ моделі *YOLOv7*

Протягом сорокової ітерації тренування більшість метрик мали спад у показниках. Такі значення можуть свідчити про надмірне пристосовування моделі до тренувальних даних, що може негативно

впливати на здатність розпізнавання класів, що загалом знижує точність моделі, але й фіктивно зменшує кількість помилок. Результати верифікації моделі подані на рис. 6.



Рис. 6. Результат розпізнавання моделлю *YOLOv7* повороту ліворуч як небезпечної ситуації

Графіки *Loss Box* моделі *Detectron2* зображені на рис. 7.

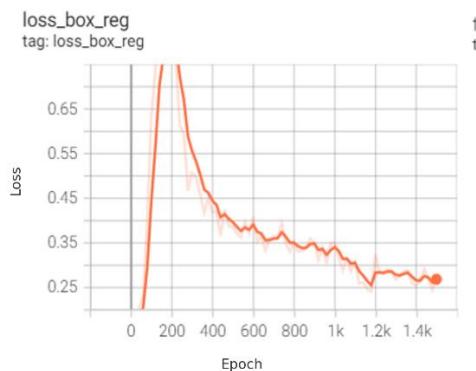


Рис. 7. Графік функції втрат отриманої під час тренування моделі *Detectron*

Результат валідації тренуваної моделі *Detectron2* поданий на рис. 8.



Рис. 8. Результат валідації моделі *Detectron2* [15]

Висновки щодо тренування та порівняння моделей *YOLOv7* і *Detectron2*

З огляду на результати тренувань, валідації моделей і метрик можна сказати, що на цьому наборі даних важко виявити кращу модель за точністю. Швидкість роботи / класифікації є дуже близькою одна до одної. Проте можна зауважити, що модель *YOLOv7* значно швидше навчається, та їй необхідно менше ітерацій для досягнення того результату, що був поданий в експерименті. Також розмір моделі *YOLOv7* суттєво менший, що може вплинути на її розгортання на апаратному забезпеченні, що має обмежені можливості з боку оперативної пам'яті та пам'яті зберігання.

Розгортання моделей *YOLOv8* і *DetectNet_v2* на *Jetson TX2*

Під час розгортання моделі *YOLOv8* на *Jetson TX2* виникли складнощі із застосуванням натренованої моделі. Розгортання моделей *YOLOv8* безпосередньо на платформі *Jetson TX2 NX* є неможливим. Розробник *YOLOv8* – компанія *Ultralytics* – використовує *Python 3.10*, тоді як остання доступна версія *Python* на *Jetson TX2* – 3.6.9. Хоча існує можливість зібрати *Python 3.10* з вихідного коду для архітектури *aarch64*, відсутні готові *pip*-пакети для встановлення *torchvision*, сумісного з цією

версією *JetPack* і оновленим *Python*. Крім того, для роботи *Ultralytics* потрібен *CUDA Toolkit 11*, тоді як *Jetson TX2* підтримує лише версію 10.2. Оновлення *CUDA* на цій платформі є складним завданням через сильну залежність компонентів на низькому рівні. Однак існує спосіб розгортання моделі *YOLOv8* на *Jetson TX2*. Загалом він полягає у використанні плагіна *DeepStream* для запуску моделі. Процес передбачає два основні етапи:

1) генерація ваг моделі з *.pt*-файлу за допомогою скрипту *get_wts_yoloV8.py* на іншому комп'ютері, що підтримує *Python 3.10*;

2) використання згенерованих файлів для виконання моделі в тестових застосунках безпосередньо на *Jetson TX2 NX*.

Порівняння швидкодії *YOLOv8* та *DetectNet_v2*

Під час проведення експерименту швидкодії моделі *YOLOv8* [16] на відео було виявлено суттєву деградацію її швидкодії. Використання ресурсів відеокарти становило 100 %, проте швидкодія – п'ять кадрів за секунду.

Це свідчить про неможливість використання запропонованої моделі на одноплатному комп'ютері *Jetson TX2* для завдань класифікації в реальному часі, оскільки швидкодія не дає змогу обробляти зображення вчасно.

Проте компанія *Nvidia* надає розроблені нею моделі, які є оптимізованими для застосування

на одноплатних комп'ютерах *Nvidia Jetson*. До такої моделі належить *DetectNet_v2*, яку можна натренувати за власним набором даних.

Швидкодія тренованої моделі, яка працює на основі *ResNet-10*, що виявляє автомобілі, є досить високою: 74 кадри за секунду, що може вважатися більш ніж достатнім для завдань розпізнавання в реальному часі (рис. 9).



Рис. 9. Швидкодія тренованої моделі *DetectNet_v2*

Із використанням інструментарію *trtexec* було обчислено, наскільки модель *DetectNet_v2* краще працює на *Jetson TX2*, ніж *YOLOv8* (табл. 2).

Таблиця 2. Порівняння *YOLOv8* і *DetectNet_v2*

Metric Name	Yolov8m	DetectNet_v2	DetectNet_v2 Nano
Average on 10 runs – GPU latency, ms	155.703	6.927	4.180
Average on 10 runs – Host latency, ms	155.991	7.088	4.270
Throughput, qps	6.407	140.517	233.145
Latency: min, ms	153.808	7.052	4.244
Latency: max, ms	157.84	8.735	5.289
Latency: mean, ms	156.049	7.109	4.282
End-to-End Host Latency: min, ms	153.817	7.057	4.250
End-to-End Host Latency: max, ms	157.854	8.741	5.295
End-to-End Host Latency: mean, ms	156.058	7.115	4.288
GPU Compute Time: min, ms	153.52	6.892	4.154
GPU Compute Time: max, ms	157.556	8.575	5.196
GPU Compute Time: mean, ms	155.761	6.948	4.192

Метрики, що надає інструментарій *trtexec*, мають таке пояснення:

– *GPU latency (average on 10 runs, ms)*: середній час, необхідний для обчислень класифікації винятково

на графічному процесорі, незважаючи на перед- і післяоброблення або копіювання даних між CPU та GPU;

– *Host latency (average on 10 runs, ms)*: середній час, заміряний на стороні CPU, що містить виклики

до GPU, передавання вхідної інформації, очікування результату та його зчитування;

- *Throughput (qps)*: продуктивність системи, що визначається як кількість запитів (інференсів), які можуть бути оброблені за секунду (*queries per second*);

- *Latency (min / max / mean, ms)*: мінімальна, максимальна та середня затримка під час класифікації, що дає змогу оцінити варіативність часу виконання;

- *End-to-End Host Latency (min / max / mean, ms)*: загальна затримка, виміряна на стороні хоста, яка передбачає повний цикл оброблення – від моменту надсилання вхідної інформації до досягнення кінцевого результату;

- *GPU Compute Time (min / max / mean, ms)*: час фактичного виконання обчислень на GPU (ядро виконання), що демонструє завантаження та ефективність використання графічного процесора.

Порівняння моделей *YOLOv8m*, *DetectNet_v2* та *DetectNet_v2 Nano* показало суттєву різницю у швидкодії. *YOLOv8m* має найвищу затримку й найнижчу пропускну здатність, тому потребує потужних GPU та не є оптимальною для вбудованих систем. *DetectNet_v2* забезпечує кращий баланс між точністю та продуктивністю й може застосовуватись

на одноплатних комп'ютерах *Jetson*. Найефективнішою для *Jetson*-платформ є *DetectNet_v2 Nano* завдяки найменшій затримці (~4.2 мс) та найвищій продуктивності (~233 qps).

Обговорення досягнутих результатів

Проведені експерименти вказують на те, що є можливість створення моделі комп'ютерного зору, яка зможе класифікувати передумови ДТП в режимі реального часу. Для покращення цього підходу можна застосувати такі спрощення:

- вимкнення фонові інформації на зображенні, оскільки вона не впливає на процес виконання маневру;

- ігнорування типу транспортного засобу, адже характер маневру залишається незмінним незалежно від того, чи його здійснює легковий автомобіль, вантажівка або велосипед;

- додавання характеристик, які сприятимуть кращій класифікації зображень, зокрема інформації про дорожню розмітку, яка може бути відсутньою або малопомітною.

Приклад описаних покращень зображень на рис. 10.

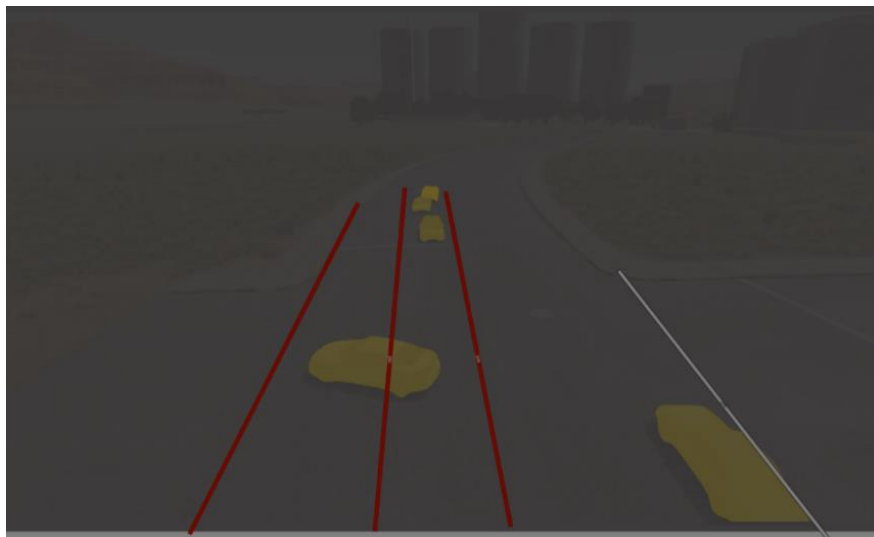


Рис. 10. Приклад змінених елементів зображення для покращення результатів класифікації

Однак цей підхід потребує значних ресурсів для підготовки, оскільки анування зображень виконується вручну, і тільки людина може коректно визначити, які лінії та об'єкти необхідно позначати, а які не мають значення. Також значна кількість різних варіацій небезпечних ситуацій зумовлює розширення навчального набору даних.

Висновки

У межах цього дослідження розглянуто сімейства алгоритмів комп'ютерного зору *YOLO*, *Detectron2* та *DetectNet_v2*, а також їх можливість у застосуванні на одноплатному комп'ютері *Jetson TX2*.

Було створено навчальний набір даних для класифікації небезпечної ситуації повороту ліворуч, а також проведено навчання моделей і порівняно їх. За результатами експериментів визначено, що *DetectNet_v2* є найкращою моделлю для роботи на одноплатному комп'ютері *Jetson TX2* з-поміж

розглянутих. Подальші кроки в розробленні цього підходу можна здійснювати в розрізі зміни навчального набору даних, щоб оптимізувати й допомогти моделі у визначенні того, які ситуації вважаються небезпечними, а які ні, а також зміни гіперпараметрів моделей під час тренування.

Список літератури

1. Speed Bump and Pothole Detection Using Deep Neural Network with Images Captured through ZED Camera / J.-E. Peralta-López та ін. *Applied Sciences*. 2023. Vol. 13, № 14. 8349 p. DOI: <https://doi.org/10.3390/app13148349>
2. Speed Bump Detection Using Accelerometric Features: A Genetic Algorithm Approach / J. Celaya-Padill. *Sensors*. 2018. Vol. 18, № 2. 443 p. DOI: <https://doi.org/10.3390/s18020443>
3. Eslami Nezhad S. A., Delavar M. R. An integrated network-constrained spatial analysis for car accidents: a case study of tehran city, iran. *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2019. XLII-4/W18. P. 335–342. DOI: <https://doi.org/10.5194/isprs-archives-xxii-4-w18-335-2019>
4. Contributors to Wikimedia projects. Kernel density estimation – Wikipedia. Wikipedia, the free encyclopedia. URL: https://en.wikipedia.org/wiki/Kernel_density_estimation (дата звернення: 19.04.2025).
5. Ravindran V., Viswanathan L., Rangaswamy S. A Novel Approach to Automatic Road-Accident Detection using Machine Vision Techniques. *International Journal of Advanced Computer Science and Applications*. 2016. Vol. 7, № 11. DOI: <https://doi.org/10.14569/ijacsa.2016.071130>
6. Robles-Serrano S., Sanchez-Torres G., Branch-Bedoya J. Automatic Detection of Traffic Accidents from Video Using Deep Learning Techniques. *Computers*. 2021. Vol. 10. № 11. 148 p. DOI: <https://doi.org/10.3390/computers10110148>
7. Amirfakhrian M., Parhizkar M. Integration of image segmentation and fuzzy theory to improve the accuracy of damage detection areas in traffic accidents. *Journal of Big Data*. 2021. Vol. 8, № 1. DOI: <https://doi.org/10.1186/s40537-021-00539-2>
8. Zimmermann H. J. Fuzzy set theory. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2010. Vol. 2, № 3. P. 317–332. DOI: <https://doi.org/10.1002/wics.82>
9. An Automatic Car Accident Detection Method Based on Cooperative Vehicle Infrastructure Systems / D. Tian та ін. *IEEE Access*. 2019. Vol. 7. P. 127453–127463. DOI: <https://doi.org/10.1109/access.2019.2939532>
10. Wang S. Traffic Accident Prediction based on CNN and Time of Commuting after Accident. *Highlights in Science, Engineering and Technology*. 2024. Vol. 118. P. 65–71. DOI: <https://doi.org/10.54097/5t9mba88>
11. Suzen A. A., Duman B., Sen B. Benchmark Analysis of Jetson TX2, Jetson Nano and Raspberry PI using Deep-CNN. *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Ankara, Turkey, 2020 p. 2020. DOI: <https://doi.org/10.1109/hora49412.2020.9152915>
12. Detection, instance segmentation, and classification for astronomical surveys with deep learning (DeepDISC): Detectron2 implementation and demonstration with hyper supprime-cam data / G. Merz та ін. *Monthly Notices of the Royal Astronomical Society*. 2023. DOI: <https://doi.org/10.1093/mnras/stad2785>
13. Wang C.-Y., Bochkovskiy A., Mark Liao H.-Y. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *Computer Vision and Pattern Recognition*. 2022. DOI: <https://doi.org/10.48550/arxiv.2207.02696>
14. BeamNG.tech Contributors. BeamNG.tech. *BeamNG.tech*. URL: <https://beamng.tech/> (дата звернення: 25.02.2025).
15. Comparison of potential road accident detection algorithms for modern machine vision system / O. Byzkrovnyi та ін. *Environment. technologies. resources. Proceedings of the International Scientific and Practical Conference*. 2023. Vol. 3. P. 50–55. DOI: <https://doi.org/10.17770/etr2023vol3.7299>
16. Comparison of Object Detection Algorithms for the Task of Person Detection on Jetson TX2 NX Platform / O. Byzkrovnyi та ін. *2024 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream)*, Vilnius, Lithuania, 2024 p. 2024. DOI: <https://doi.org/10.1109/estream61684.2024.10542592>

References

1. Peralta-López, J., Morales-Viscaya, J., Lázaro-Mata, D., Villaseñor-Aguilar, M., Prado-Olivarez, J., Pérez-Pinal, F., Gutiérrez, A. (2023), "Speed bump and pothole detection using deep neural network with images captured through zed camera". *Applied Sciences*, № 13(14), 8349 p. DOI: <https://doi.org/10.3390/app13148349>
2. Celaya-Padilla, J. M., Galván-Tejada, C. E., López-Monteagudo, F. E., Alonso-González, O., Moreno-Báez, A., Martínez-Torteya, A., Gamboa-Rosales, H. (2018), "Speed bump detection using accelerometric features: a genetic algorithm approach". *Sensors*, № 18(2), 443 p. DOI: <https://doi.org/10.3390/s18020443>
3. Eslaminezhad, S. A. and Delavar, M. R. (2019), "An integrated network-constrained spatial analysis for car accidents: a case study of tehran city, iran". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-4/W18, P. 335–342. DOI: <https://doi.org/10.5194/isprs-archives-xlii-4-w18-335-2019>
4. Contributors to Wikimedia projects, (2005), "Kernel density estimation – Wikipedia. *Wikipedia, the free encyclopedia*", available at: https://en.wikipedia.org/wiki/Kernel_density_estimation (last accessed 25.02.2025)
5. Ravindran, V., Viswanathan, L., & Rangaswamy, S. (2016), "A novel approach to automatic road-accident detection using machine vision techniques". *International Journal of Advanced Computer Science and Applications*, №7(11). DOI: <https://doi.org/10.14569/ijacsa.2016.071130>
6. Robles-Serrano, S., Torres, G. S., & Branch, J. W. (2021), "Automatic detection of traffic accidents from video using deep learning techniques". *Computers*, №10(11), 148 p. DOI: <https://doi.org/10.3390/computers10110148>
7. Amirfakhrian, M. and Parhizkar, M. (2021), "Integration of image segmentation and fuzzy theory to improve the accuracy of damage detection areas in traffic accidents". *Journal of Big Data*, №8(1). DOI: <https://doi.org/10.1186/s40537-021-00539-2>
8. Zimmermann, H. (2010), "Fuzzy set theory". *WIREs Computational Statistics*, №2(3), P. 317–332. DOI: <https://doi.org/10.1002/wics.82>
9. Tian, D., Zhang, C., Duan, X., & Wang, X. (2019), "An automatic car accident detection method based on cooperative vehicle infrastructure systems". *IEEE Access*, № 7, P. 127453-127463. DOI: <https://doi.org/10.1109/access.2019.2939532>
10. Wang, S. (2024), "Traffic accident prediction based on cnn and time of commuting after accident". *Highlights in Science, Engineering and Technology*, №118, P. 65–71. DOI: <https://doi.org/10.54097/5t9mba88>
11. Süzen, A. A., Duman, B., & Şen, B. (2020). "Benchmark analysis of jetson tx2, jetson nano and raspberry pi using deep-cnn". *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, № 1-5. DOI: <https://doi.org/10.1109/hora49412.2020.9152915>
12. Merz, G., Liu, Y., Burke, C. J., Aleo, P., Liu, X., Kind, M. C., Liu, Y. (2023), "Detection, instance segmentation, and classification for astronomical surveys with deep learning (deepdisc): detectron2 implementation and demonstration with hyper supprime-cam data". *Monthly Notices of the Royal Astronomical Society*, № 526(1), P. 1122–1137. DOI: <https://doi.org/10.1093/mnras/stad2785>
13. Wang, C., Bochkovskiy, A., & Liao, H. M. (2022), "Yolov7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors". *Computer Vision and Pattern Recognition*. DOI: <https://doi.org/10.48550/arxiv.2207.02696>
14. "BeamNG.tech Contributors, BeamNG.tech". 2023. available at: <https://beamng.tech/> (last accessed: 25.02.2025)
15. Byzkrovnyi, O., Smelyakov, K., Chupryna, A., Savulionienė, L., & Sakalys, P. (2023), "Comparison of potential road accident detection algorithms for modern machine vision system". *Environment. technologies. resources*. Proceedings of the International Scientific and Practical Conference, №3, P. 50–55. DOI: <https://doi.org/10.17770/etr2023vol3.7299>
16. Byzkrovnyi, O., Smelyakov, K., Chupryna, A., & Lanovyy, O. (2024), "Comparison of object detection algorithms for the task of person detection on jetson tx2 nx platform". *IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream)*. DOI: <https://doi.org/10.1109/estream61684.2024.10542592>

Надійшла (Received) 15.04.2025

Відомості про авторів / About the Authors

Бизкровний Олександр Миколайович – Харківський національний університет радіоелектроніки, аспірант кафедри програмної інженерії, Харків, Україна; e-mail: oleksandr.byzkrovnyi@nure.ua; ORCID ID: <https://orcid.org/0000-0001-9335-442X>

Смеляков Кирило Сергійович – доктор технічних наук, професор, Харківський національний університет радіоелектроніки, завідувач кафедри програмної інженерії, член НТР, Харків, Україна; e-mail: kyrylo.smelyakov@nure.ua; ORCID ID: <https://orcid.org/0000-0001-9938-5489>

Byzkrovnyi Oleksandr – Kharkiv National University of Radio Electronics, PhD Student at the Department of Software Engineering, Kharkiv, Ukraine.

Smelyakov Kyrylo – Doctor of Sciences (Engineering), Professor, Kharkiv National University of Radio Electronics, Head at the Department of Software Engineering, Kharkiv, Ukraine.

MODELS OF THE CLASSIFIER OF PREREQUISITES FOR THE OCCURRENCE OF ROAD ACCIDENTS TO PREDICT DANGEROUS SITUATIONS AT INTERSECTIONS

The **subject matter**: this study focuses on the preconditions for the occurrence of road traffic accidents at intersections and areas with limited visibility; the use of computer vision models for classifying the preconditions of road traffic accidents and evaluating the effectiveness of their use in real-time operation. The **goal** of the study is to compare computer vision models for the task of classifying the preconditions of traffic accidents for real-time operation. The study involves comparing models using the Jetson TX2 platform and determining the effectiveness of this approach for generating real-time warning signals for drivers. **Tasks**: explore computer vision models such as Detectron2 and YOLOv7 for the task of classifying traffic accident preconditions in terms of model performance, ease of dataset creation, model training, and deployment. Compare YOLOv8 and DetectNet_v2 on a single-board computer Jetson TX2 in terms of processing speed. The **methods** used include training and using machine learning models, as well as simulating hazardous situations using software such as BeamNG.tech and CARLA. A comparative analysis of the application results of the models was conducted using performance evaluation metrics. The **main results** of the study include identifying the most effective model for classification tasks on the Jetson TX2 single-board computer – DetectNet_v2; a positive evaluation of the effectiveness of this approach for real-time driver warning, although certain limitations were noted regarding the size of the training dataset and the complexity of its preparation. **Conclusions**. The following computer vision algorithms were examined: Detectron2, YOLOv7, and DetectNet_v2. It was found that YOLOv7 outperforms Detectron2 in detecting the preconditions of traffic accidents in images. However, DetectNet_v2 was found to be better suited for deployment on the Jetson TX2 single-board computer compared to YOLOv7. Additionally, based on experimental findings, it was concluded that the application of this approach to predicting accident preconditions is problematic due to difficulties in creating a training dataset – specifically, the variability of precondition scenarios.

Keywords: Object classification on an image; Computer vision; Machine Learning; Artificial Intelligence; Information Technologies; Nvidia Jetson; road traffic accidents prediction; YOLO; DetectNet_v2; Detectron2.

Бібліографічні описи / Bibliographic descriptions

Бизкровний О. М., Смельяков К. С. Моделі класифікатора передумов виникнення ДТП для передбачення небезпечних ситуацій на перехрестях. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 2 (32). С. 177–187. DOI: <https://doi.org/10.30837/2522-9818.2025.2.177>

Byzkrovnyi, O., Smelyakov, K. (2025), "Models of the classifier of prerequisites for the occurrence of road accidents to predict dangerous situations at intersections", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (32), P. 177–187. DOI: <https://doi.org/10.30837/2522-9818.2025.2.177>

М. ЗЕНКІН, А. ІВАНКО, М. ЧЕРНИШ

ВПЛИВ ВІБРАЦІЙ РІЗАЛЬНОГО ІНСТРУМЕНТА НА ЯКІСТЬ ПОВЕРХНІ ОБРІЗАНИХ ЛИСТОВИХ МАТЕРІАЛІВ ТА МЕТОДИ ЇХ МІНІМІЗАЦІЇ

Предметом дослідження є вплив вібрацій різального інструмента на якість поверхні обрізаних листових матеріалів у процесі механічного оброблення, а також аналіз механізмів виникнення коливань і розроблення методів їх мінімізації для підвищення ефективності технологічних операцій у машинобудуванні. **Мета роботи** – обґрунтувати наукові підходи до прогнозування ресурсу різального інструмента й оптимізації режимів різання способом зменшення вібраційних навантажень, що дасть змогу підвищити якість обробки поверхні листових матеріалів, знизити виробничі витрати та запобігти дефектам обробки. **Завдання:** оцінити вібраційні властивості інструмента залежно від режимів різання та ступеня зносу; проаналізувати вплив коливань на шорсткість і дефекти поверхні (хвилястість, мікротріщини); розробити та протестувати методи мінімізації вібрацій, зокрема демпфувальні системи й ультразвукові коливання; визначити оптимальні умови оброблення для різних матеріалів. **Методи:** експериментальний аналіз із застосуванням акселерометрів для тривісного моніторингу вібрацій; профілометрія та мікроскопія з метою оцінювання якості поверхні; стохастичне моделювання зносу інструмента на основі статистичних показників; тестування демпфувальних систем (пружні муфти, демпфери) та ультразвукових коливань (20–30 кГц); порівняльний аналіз ефективності змащувально-охолоджувальних рідин. **Результати дослідження:** встановлено, що вібрації по осі X мають найбільший вплив на формування дефектів поверхні, особливо внаслідок збільшення глибини різання та зносу інструмента; ультразвукові коливання (22 кГц, амплітуда 10 мкм) знижують тертя та стабілізують процес різання, покращуючи якість оброблених поверхонь; демпфувальні системи ефективно зменшують амплітуду коливань у критичних частотах (до 1,5 кГц); використання води як ЗОР виявилось ефективнішим за мастило для міді, оскільки знижує нерівності поверхні; розроблена вдосконалена модель торцевої фрези з демпфувальними елементами зменшує вібраційні навантаження та дефекти під час оброблення тонколистових матеріалів. **Висновки:** з досліджень випливає, що моніторинг вібрацій є ефективним інструментом для діагностики стану різального інструмента й оптимізації процесу різання; інтеграція демпфувальних систем і ультразвукових коливань дає змогу значно знизити вібраційні впливи, підвищити якість поверхні листових матеріалів і подовжити ресурс інструмента, що сприяє економічній ефективності виробництва та мінімізації дефектів обробки.

Ключові слова: вібрації; різальний інструмент; якість поверхні; листові матеріали; знос; мінімізація коливань.

Вступ

Активний розвиток машинобудівної галузі сприяє тому, що ключовим напрямом забезпечення високої конкурентоспроможності вітчизняної продукції є зниження її собівартості внаслідок підвищення ефективності всіх технологічних процесів, що супроводжують виготовлення деталей і вузлів. Серед таких процесів оброблення матеріалів різанням посідає провідне місце, оскільки, за експертними оцінками, близько 15% вартості виробів машинобудування у світовій практиці припадає саме на операції механічного оброблення з використанням лезового інструмента. Проте, незважаючи на тривале впровадження усталених методик призначення режимів різання та регламентів заміни інструмента, значним бар'єром на шляху до підвищення якості оброблення та зменшення витрат залишається проблема вібрацій, що неминуче виникають під час різання, незалежно від типу обладнання або матеріалу заготовки.

Особливої уваги набуває дослідження впливу вібрацій різального інструмента на якість поверхні листових матеріалів, що часто застосовують у відповідальних конструкціях, адже навіть незначні коливання можуть спричинити погіршення геометричної точності, збільшення шорсткості поверхні, передчасний знос інструмента та зниження стабільності технологічного процесу. Тоді як незначні вібрації зазвичай не беруть до уваги під час аналізу, значні автоколивання можуть призводити до втрати стійкості різання, що потребує впровадження систематичних підходів до їх виявлення, контролю та мінімізації. Відомі режими різання, описані в довідниках, часто є результатом експериментів, що проводилися без уваги до впливу фонових коливань, а отже, не відтворюють повною мірою поведінку системи внаслідок зміни пружних властивостей верстата або інструмента.

Наявність численних фізичних явищ, які можуть одночасно спричинити автоколивання в процесі різання, зокрема залежність сили різання від швидкості,

фазові зсуви або резонансні ефекти, вимагає комплексного підходу до аналізу та моделювання динаміки оброблення. Зважаючи на вищевикладене, актуальність дослідження зумовлена необхідністю глибокого вивчення природи вібрацій у процесі різання листових матеріалів, визначення їх впливу на кінцеву якість обробленої поверхні, а також розроблення ефективних методів мінімізації шкідливих коливань з метою забезпечення стабільності, точності та надійності оброблення в умовах сучасного машинобудівного виробництва.

Аналіз останніх досліджень і публікацій

У науковій літературі значна увага приділяється дослідженню впливу вібрацій різального інструмента на якість обробки листових матеріалів, а також методам їх мінімізації. Питанню стану різального інструмента присвячено численні роботи як вітчизняних, так і зарубіжних авторів [1, 2]. Зокрема, у праці [3] подано огляди, які детально висвітлюють сучасний стан окресленої проблеми, що залишається актуальною навіть у контексті автоматизованого виробництва. За результатами досліджень, наведеними в статті [4], до 20% простоїв технологічного обладнання спричинені поломками різального інструмента, що доводить важливість цього аспекту.

У роботі [4] описано універсальний механізм виникнення автоколивань, пов'язаний із координатною залежністю в технологічній системі та дискретністю стружкоутворення. Цей механізм вважається ключовим для появи інтенсивних високочастотних вібрацій, які супроводжуються властивим свистом і високим рівнем звукового тиску, а також менш помітних коливань, що завжди присутні під час різання полікристалічних матеріалів. Такі автоколивання суттєво впливають на якість поверхні оброблених матеріалів, продуктивність процесу та довговічність інструмента [5]. Виникнення вібрацій залежить від динамічних властивостей системи, зокрема від змінної сили, що діє під час різання. У процесі оброблення листових матеріалів можливі коливання різних елементів обладнання. Однак вібрації самого різального інструмента є найбільш критичними, оскільки впливають на знос інструмента, точність обробки, стан верстата й, основне, на якість поверхні обрізаних матеріалів [6].

Вібрації в технологічній системі можуть відбуватися у вертикальній або горизонтальній

площинах, паралельно чи перпендикулярно до напрямку подачі, що далі призводить до зміщення інструмента щодо заготовки, унаслідок чого змінюються режими різання (швидкість, глибина та подача) і, відповідно, виникає змінна сила [1]. Отже, вібрації різального інструмента є основним фактором, що погіршує якість поверхні листових матеріалів, а їх мінімізація вимагає комплексного підходу до контролю та оптимізації процесу різання.

Аналіз проблеми й наявних методів

Дослідження передбачає експериментальний аналіз вібраційних властивостей різального інструмента під час оброблення різних листових матеріалів з подальшим оцінюванням якості поверхні. Методика передбачає кілька етапів.

1. Підготовка експерименту:

- вибір матеріалів заготовок та інструментів (твердосплавні пластини, фрези з різною геометрією);
- налаштування обладнання (токарний / фрезерний верстат із системою ЧПК);
- установлення датчиків вібрації (акселерометри) на інструмент і заготовку для реєстрації коливань у трьох осях (X , Y , Z).

2. Проведення експериментів:

- оброблення заготовок у різних режимах (глибина різання, швидкість, подача) з фіксацією вібраційних сигналів;
- дослідження впливу зносу інструмента на амплітуду коливань;
- аналіз якості поверхні (профілометрія, мікроскопія) для виявлення дефектів (хвилястість, мікротріщини).

3. Методи мінімізації вібрацій:

- тестування демпфувальних систем (пружні муфти, демпфери);
- використання ультразвукових коливань (частота 20–30 кГц) для зменшення тертя;
- оцінювання ефективності різних ЗОР (мастила, емульсії, вода);
- оброблення отриманих показників.

Результати дають змогу визначити оптимальні режими різання та ефективні методи зменшення вібрацій для підвищення якості обробки.

Мета статті – дослідження з виявлення закономірностей впливу вібрацій різального інструмента на якість поверхні обрізаних листових матеріалів

та обґрунтування ефективних методів мінімізації вібраційних коливань для підвищення точності, стабільності та якості процесу механічного оброблення.

Розв'язання завдання

Точність передбачення ресурсу експлуатації різального інструмента, зокрема періоду його зносостійкості, відіграє визначальну роль у забезпеченні ефективності технологічних операцій оброблення різанням у процесі виготовлення високоточних та відповідальних елементів продукції сучасного машинобудування, оскільки раптовий вихід інструмента з ладу або його критичне зношення без належного контролю можуть спричинити суттєве підвищення виробничих витрат, затримку виконання технологічного циклу та зниження якості кінцевого виробу. З огляду на природну стохастичність процесів різання листів металу, пов'язану зі змінністю властивостей матеріалу заготовки, властивостями інструмента та умовами оброблення, все більшого поширення набуває застосування ймовірнісних моделей, що дають змогу оцінювати ризики відмов інструмента на основі законів розподілу тривалості його ефективного функціонування або інтенсивності зношення, а також зважають на залежність цих властивостей від параметрів режимів різання [7].

Практична реалізація подібних стохастичних моделей вимагає збору великого обсягу статистичних показників, що визначають процес поступового зношення інструмента, який відбувається впродовж тривалого часу роботи. Одним з ефективних та інформативних підходів до отримання такої статистики є моніторинг вібраційних параметрів технологічної системи, що містить елементи "станок – пристрій – інструмент – заготовка" (СПІЗ), оскільки вібрації є чутливим індикатором стану різальної частини інструмента й дають змогу виявити навіть незначні зміни, які передують його критичному зносу. Джерелом коливань у цій зоні необхідно вважати складні процеси тертя й деформування, що виникають унаслідок взаємодії інструмента з оброблюваним матеріалом і призводять до виникнення як сил пружних, так і пластичних деформацій [1]. Ці сили, зокрема пружні (P_{np}) та пластичні (P_{nl}) реакції, орієнтовані перпендикулярно до передньої поверхні різального інструмента,

а також сили аналогічного походження (P'_{np} і P'_{nl}), що діють з боку задньої поверхні інструмента, спричиняють виникнення сил тертя, які генерують коливання, що можуть набувати резонансного характеру й негативно позначатися на якості обробленої поверхні (рис. 1).

$$\begin{aligned} T &= \mu \cdot (P_{np} + P_{nl}), \\ T' &= \mu \cdot (P'_{np} + P'_{nl}), \end{aligned} \quad (1)$$

де T – сила тертя вздовж передньої поверхні різального інструмента;

T' – сила тертя вздовж задньої поверхні різального інструмента;

μ – коефіцієнт тертя.

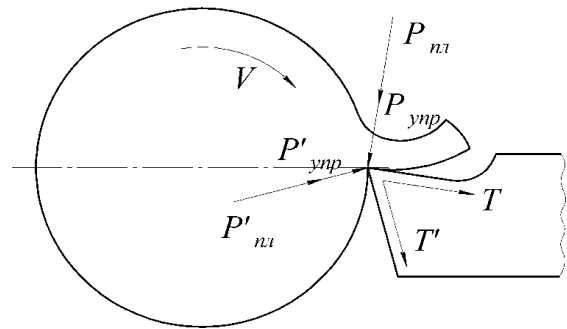


Рис. 1. Деформування оброблюваного матеріалу в процесі різання

Для спрощення розрахунків зазначена система сил наводиться до рівнодійної сили різання R . Рівнодійна сила різання R для зовнішнього точення зображена на рис. 2. Сила різання R у разі токарного оброблення розкладається на три взаємно перпендикулярні складники – P_x , P_y і P_z . Сила різання R та її проекції є нелінійними величинами, що залежать від режимів і фактичних умов різання [8]. Під час різання металу тангенціальний складник сили різання можна оцінювати за емпіричною формулою

$$P_z = C_{pz} \cdot a_p^{xz} \cdot f_n^{yz} \cdot v_c^{nz} \cdot K_{pz}, \quad (2)$$

де C_{pz} , xz , yz , nz – емпіричні коефіцієнти і показники ступеня;

a_p – глибина різання;

f_n – подання на оберт;

v_c – швидкість різання;

K_{pz} – корегувальний коефіцієнт, що бере до уваги фактичні умови різання, зокрема й геометрію різального інструмента.

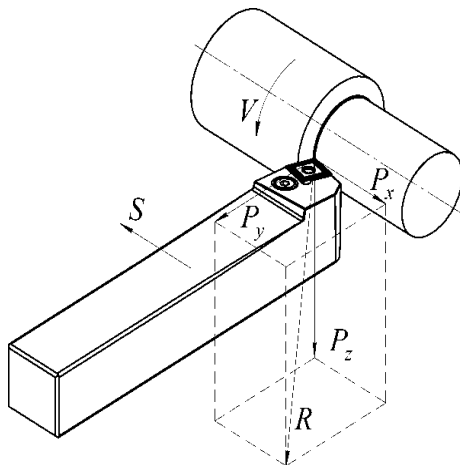


Рис. 2. Сила різання за умови зовнішнього впливу

Під дією сили різання машинна частина технологічної системи здійснює вимушені коливання. Для аналізу системи СПЗ, що здійснює інтенсивні поперечні коливання, можна скористатися принциповою схемою, зображеною на рис. 3.

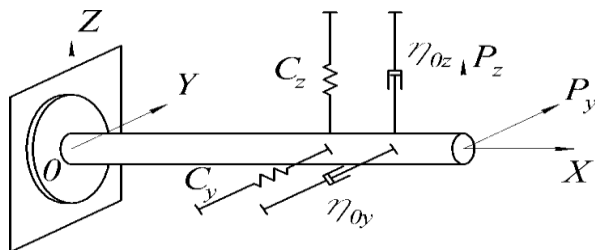


Рис. 3. Принципова схема коливальної системи СПЗ

Відповідно до принципу Даламбера має місце рівність

$$\vec{F}_i + \vec{F}_c + \vec{F}_y + \vec{R} = 0, \quad (3)$$

де $\vec{F}_i$ – сила інерції;

$\vec{F}_c$ – сила опору (демпфування);

$\vec{F}_y$ – сила пружності системи.

У наукових дослідженнях, присвячених питанням зменшення амплітуди коливань, що виникають у процесі оброблення різанням, увагу зосереджено переважно на коливаннях, які відбуваються в площині, перпендикулярній осі заготовки, зокрема в умовах токарного оброблення, де традиційно вважається, що коливання в напрямку осі Y мають найбільш суттєвий вплив на формування якості обробленої поверхні. Проте результати досліджень, виконаних із застосуванням трикомпонентного акселерометра для аналізу вібрацій вершинної частини різального інструмента,

свідчать про те, що зі зростанням глибини різання або зі збільшенням ступеня зношення різальної кромки інструмента амплітуди коливань по координатних осях X та Z можуть зростати випереджальними темпами порівняно з коливаннями по осі Y [5].

Особливо небезпечними для забезпечення високої якості оброблення листових матеріалів є коливання по осі X , тобто в напрямку, що збігається з подовжньою віссю заготовки, оскільки саме такі коливання спричиняють мікроскопічні вбивання інструмента в бокову поверхню оброблюваної кромки, що спричиняє періодичні дефекти на поверхні, які мають характер повторюваних слідів. Ці сліди, що виникають внаслідок кожного оберту інструмента з певним фазовим зсувом, формують хвилястий рельєф, що значно погіршує шорсткість та однорідність поверхні листового матеріалу, особливо в зонах з більшою глибиною різання (рис. 4).

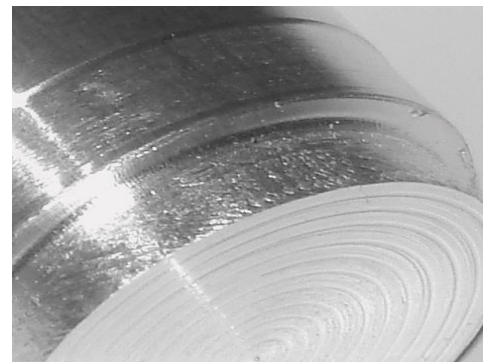


Рис. 4. Зовнішній вигляд поверхні обрізаного валика залежно від глибини різання: вплив параметрів оброблення на якість поверхні [8]

На основі фотоспостережень оброблення валика діаметром 30 мм із варіацією глибини різання було встановлено, що саме збільшення глибини різання приводить до інтенсивнішого зростання амплітуди коливань по осі X , що зі свого боку набуває вигляду видимих періодичних слідів на поверхні заготовки, отриманих у процесі різання з найбільшими глибинами. Водночас із поступовим зменшенням глибини різання спостерігалось зниження інтенсивності цих коливань, що корелювалося зі зменшенням шорсткості обробленої поверхні. Це є свідченням прямої залежності між вібраційними властивостями інструмента й кінцевими параметрами якості поверхні листових матеріалів [7].

Під час усіх етапів оброблення листового матеріалу різання відбувалося без помітних ознак

інтенсивних автоколивань, що є важливим для розуміння стабільності процесу, хоча накопичена потенційна енергія, яка виникає внаслідок взаємодії інструмента з матеріалом, найвиразніше виявляється в момент швидкого розвантаження різального інструмента, зокрема коли він виходить у зону різкої зміни геометрії заготовки, такої як шпонковий паз. На прикладі записів вібраційної швидкості вздовж осі X , отриманих під час оброблення алюмінієвих листів, видно, що вихід різця в паз спричиняє різкий перехід потенційної енергії в кінетичну, що нагадує ефект відриву стружки, до того ж для алюмінієвого сплаву це супроводжується значним стрибком вібраційної швидкості [9].

Експерименти з обрізання листових матеріалів продемонстрували, що навіть за умов роботи інструмента з надмірним зносом автоколивання не виникають у широкому діапазоні глибин різання, що свідчить про меншу схильність таких матеріалів до вібраційних збурень порівняно з металевими сплавами. На графіку вібраційної швидкості вершини різця під час виходу в паз на алюмінієвій заготовці видно, що вектор швидкості різко зростає вздовж осі X (перпендикулярно до поверхні різання за умови кута атаки $\pi/2$). І хоча в процесі формування зливної стружки зазвичай опір матеріалу та поверхні різання обмежують таке зростання, воно все одно присутнє, досягаючи максимуму, після чого швидкість V_x поступово знижується до мінімального значення, до того ж частота цих коливань, близька до власної частоти технологічної системи (приблизно 1,5 кГц), відіграє ключову роль у динаміці процесу.

Ці зміни вібраційної швидкості не відповідають стабільним умовам різання з утворенням зливної стружки, оскільки максимальне значення V_x , досягнуте за дуже нетривалий час (приблизно 0,25 мс), не призводить до значного відхилення вершини різця від тангенціального напрямку (відхилення становить лише 3–4 мкм), однак відрив задньої грані інструмента від поверхні різання руйнує адгезійні зв'язки, даючи змогу інструменту знову контактувати із зоною пластичної деформації з ударом, що за незначних відхилень сприяє розвитку тангенціальних тріщин у попередньо деформованому матеріалі, полегшуючи відокремлення стружки та впливаючи на якість поверхні. Оскільки радіальна вібраційна швидкість у разі зливної стружки значно менша за швидкість різання. Саме остання визначає силу

й напрямок удару, тоді як радіальні коливання відіграють допоміжну роль у руйнуванні адгезії між інструментом і заготовкою, а кут атаки залишається близьким до тангенціального, що підвищує ефективність удару завдяки концентрації енергії руйнування в точці контакту через обмежену швидкість поширення деформацій [10].

Коли в процесі оброблення виникають інтенсивні автоколивання, амплітуда радіальної швидкості, а також відповідні переміщення різального інструмента можуть зростати вдсятеро, що істотно ускладнює стабільність контакту між різальною кромкою та заготовкою. Такий різкий стрибок у динаміці інструмента призводить до того, що після чергового зсуву шару стружки та імпульсного збільшення радіальної швидкості точка нового контакту вершини інструмента може значно зміститися щодо попереднього положення, де вже були частково деформовані зерна матеріалу та сформована випереджальна мікротріщина. Такого плану зміщення ускладнює подальше формування основної тріщини зсуву, порушує сталу геометрію зрізу та сприяє зростанню навантаження на інструмент, що погіршує якість обробленої поверхні обрізаних листових матеріалів.

Крім того, внаслідок порушення стабільності траєкторії інструмента під дією автоколивань кут атаки різальної кромки може суттєво відхилитися від оптимального тангенціального напрямку, що ще більше ускладнює процес утворення стружки й підвищує ризик виникнення дефектів на поверхні. Ця складна динаміка взаємодії між інструментом і заготовкою створює передумови для вивчення альтернативних підходів до керованого введення коливань, зокрема застосування примусових мікроколивань високої частоти, зокрема ультразвукові коливання, які можуть бути ефективним інструментом для покращення процесу різання внаслідок стабілізації режимів оброблення та зниження рівня контактного тертя [8].

Зокрема використання ультразвукових тангенціальних коливань із частотою близько 22 кГц і амплітудою порядку 10 мікрометрів виявилось перспективним методом мінімізації негативного впливу вібрацій на якість обробленої поверхні за швидкості різання до 80 м/хв. У таких умовах між передньою поверхнею інструмента та стружкою періодично виникає мікрозазор, який руйнує адгезійні зв'язки, що спричиняють нарост на різальній кромці.

Внаслідок цього різальний інструмент не накопичує надлишкову потенціальну енергію, яка в інших випадках могла б викликати різкі стрибки й нерівномірні переміщення інструмента, а, навпаки, енергія розсіюється в кожному циклі мікроколивань, що позитивно впливає на стабільність оброблення [11].

Проте необхідно зауважити, що навіть за умови дії ультразвукових коливань у тангенціальному напрямку існує ризик погіршення якості поверхні через фрикційний контакт задньої грані інструмента з оброблюваною поверхнею, який може спричинити утворення мікроподрапин або втиснення частинок знімного матеріалу в поверхневий шар. Це було підтверджено в експерименті з обробленням сталеві заготовки марки 18ХГТ за допомогою різальної пластини з нульовим заднім кутом: результати продемонстрували, що без використання змащувально-охолоджувальної технологічної рідини (ЗОТР) фрикційний контакт зберігався, а якість поверхні залишалася незадовільною [12]. Водночас унаслідок додавання в зону різання індустріального мастила спостерігалось помітне покращення візуального та структурного стану обробленої поверхні, що підтверджує ефективність комплексного підходу до зменшення фрикційних впливів (рис. 5).

Однак застосування змащувальних речовин, хоча й покращує ситуацію, не є універсальним розв'язком усіх проблем, пов'язаних із вібраціями. Недостатня радіальна жорсткість технологічної системи може призводити до так званого "спливання" різального інструмента під час його ковзання у в'язкому середовищі, що спричиняє зміщення його позиції щодо заготовки, суттєво впливаючи на точність

оброблення та стан поверхні. На профілограмі поверхні мідного листового матеріалу, отриманої внаслідок оброблення з використанням ультразвукових вібрацій різального інструмента, чітко видно, що додавання мастила в зону різання викликає зміщення вершини інструмента в радіальному напрямку на відстань до 60 мкм, що свідчить про нестабільність процесу та може погіршувати якість поверхні через нерівномірний контакт із матеріалом. Зміна ситуації можлива завдяки збільшенню заднього кута різального інструмента, що посилює тиск на шар змащувально-охолоджувальної технологічної суміші (ЗОТС) на обробленій поверхні, запобігаючи "спливанню" інструмента, хоча в цьому разі виникає ризик пошкодження поверхні безпосередньо інструментом під час вібрацій, особливо якщо зважати, що поява фаски зносу на задній поверхні інструмента знову підвищує ймовірність його неконтрольованого зміщення (рис. 6).

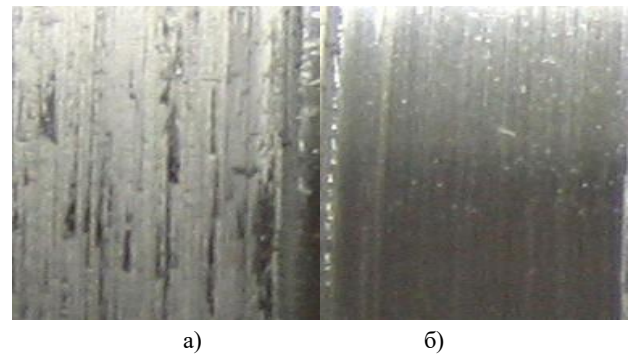


Рис. 5. Зовнішній вигляд поверхні листа сталі 18ХГТ у процесі оброблення з тангенціальними вібраціями: а – без застосування системи охолодження та змащення; б – з подачею мастила в зону різання

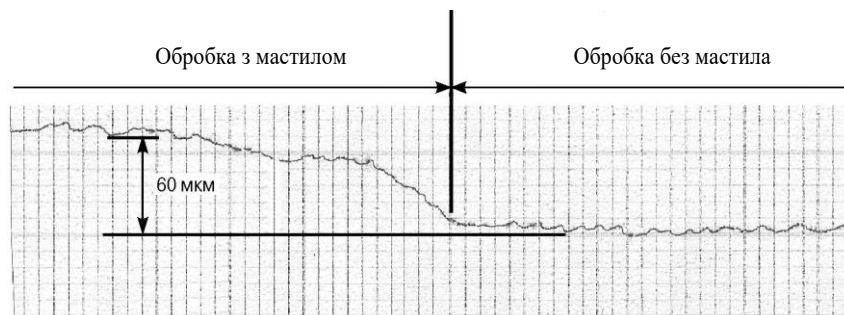


Рис. 6. Профілограма поверхні металевго листа в процесі різання з використанням ультразвукових вібрацій різального інструмента (ВУ = 200, ГУ = 50)

Оптимізація процесу ультразвукового оброблення, що розглядається як один із методів мінімізації вібрацій, ускладнюється можливістю варіювання в'язкості та інших властивостей ЗОТС, що потребує

ретельного підбору умов для забезпечення стабільності різання та покращення якості поверхні листових матеріалів. На прикладі профілограми видно, що нанесення води на поверхню заготовки замість

мастила значно покращує зовнішній вигляд обробленої поверхні. Це може бути пов'язано з меншою в'язкістю води, яка сприяє кращому контролю вібрацій і стабільності позиціонування різального інструмента, зменшуючи цим нерівності та дефекти на обрізаній поверхні (рис. 7) [13, 14].

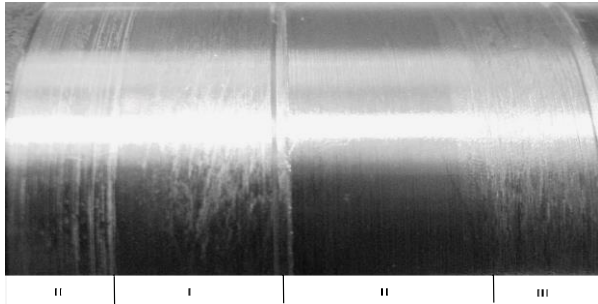


Рис. 7. Вплив різних типів систем охолодження та змащення (ЗОТС) на різання листів:
I – без ЗОТС; II – із застосуванням води;
III – з використанням мастила [10]

Однією з конструктивних переваг аналізованого технічного рішення є наявність можливості гнучкого регулювання демпфувальної здатності різального інструмента внаслідок зміни сили затягування фіксувального гвинта, що зі свого боку дає змогу впливати на величину сили тертя між верхнім і нижнім кільцями демпфувального вузла, яка разом із властивостями пружного елемента конструкції визначає загальні демпфувальні властивості інструмента. Як було встановлено в попередніх дослідженнях, підвищення рівня демпфування сприяє

суттєвому зниженню амплітуди коливань незалежно від маси інструментального вузла або жорсткості всієї вібраційної системи, що є надзвичайно важливим для стабілізації процесу оброблення листових матеріалів.

Жорсткість самого інструмента в цьому контексті залежить від властивостей пружного елемента, виготовленого зі спеціально підібраного матеріалу з визначеним модулем пружності, що забезпечує адаптацію системи до конкретних умов різання та типу оброблюваної заготовки [14]. Водночас варто наголосити, що запропонована конструкція не монолітна жорстка система, а є складальним вузлом, у якому верхнє та нижнє кільця демпфувального блоку з'єднані між собою за допомогою пружної муфти. Така архітектура дає змогу істотно зменшити масу коливної частини інструмента, що автоматично знижує власну частоту його вібрацій та сприяє уникненню резонансних явищ у критичних режимах оброблення [15].

Крім того, завдяки змінному співвідношенню між власною частотою інструментального вузла й частотою збурювального впливу, яка залежить від швидкості обертання шпинделя або режиму подачі, запропонована конструкція забезпечує вихід системи з резонансної ділянки, що є особливо цінним у процесі оброблення тонколистових металевих матеріалів, де навіть незначні резонансні коливання можуть призвести до появи хвилястих дефектів, мікротріщин або нерівностей на поверхні зрізу (рис. 8).

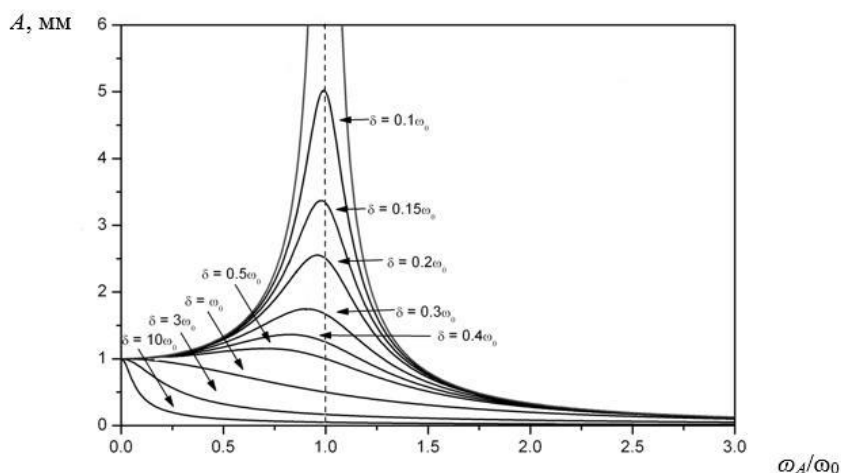


Рис. 8. Прояв резонансних коливань у системі різання за різних коефіцієнтів демпфування [5]

На основі проведених досліджень, спрямованих на виявлення механізмів впливу вібрацій різального інструмента на якість обробки поверхні листових

матеріалів, а також на пошук ефективних шляхів зменшення амплітуди коливань у процесі різання було запропоновано та сконструйовано вдосконалену

модель торцевої фрези, яка містить вбудовані демпфувальні елементи, функціональне призначення яких полягає в зниженні рівня вібраційних навантажень способом часткового поглинання енергії коливань у критичних діапазонах частот (рис. 9), що покращує чистоту обробленої поверхні та зменшує ризик формування дефектів під час обрізання листових металевих заготовок [16].

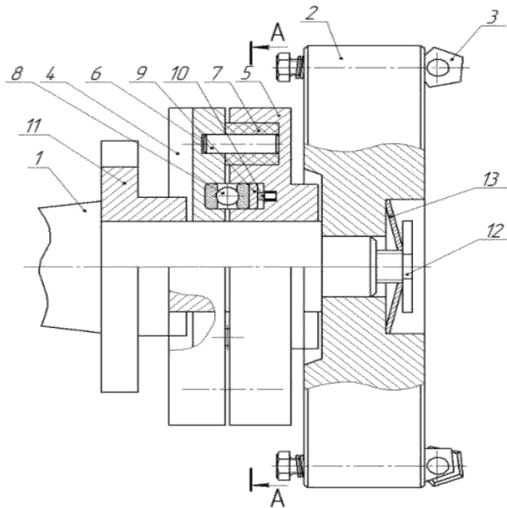


Рис. 9. Збірна торцева фреза з демпфувальними елементами для зменшення вібрацій у процесі оброблення: 1 – оправка; 2 – торцева фреза; 3 – різальні пластини; 4 – верхнє кільце демпфувального вузла; 5 – нижнє кільце демпфувального вузла; 6 – сталеві кульки; 7 – демпфувальний елемент; 8 – повідець; 9 – гвинт; 10 – пружна шайба [16]

Отже, моніторинг рівня вібрацій різального інструмента може слугувати ефективним методом оцінювання його стану під час оброблення листових матеріалів за особливих умов, що дає змогу брати до уваги мінливість процесу різання, оптимізувати тривалість використання інструмента, підвищити якість поверхні обрізаних заготовок і, як наслідок, знизити витрати на виробництво завдяки мінімізації вібраційних впливів.

Висновок

Унаслідок проведеного дослідження встановлено взаємозв'язок між вібраційними властивостями різального інструмента під час оброблення листових матеріалів та якістю обробленої поверхні. Зокрема виявлено, що амплітудно-частотні характеристики коливань, особливо в напрямку осі X , безпосередньо впливають на шорсткість, хвилястість та структурну

однорідність кромки заготовки. Досліджено механізми виникнення та поширення коливань, спричинених силами тертя та деформації в зоні різання. Визначено ключові фактори, що призводять до переходу системи в резонансний режим, особливо за умов недостатнього демпфування або критичних параметрів оброблення. Проаналізовано динамічні зміщення вершини різального інструмента під впливом раптових змін геометрії заготовки. Доведено, що навіть незначні мікроколивання можуть спричинити формування повторюваних дефектів на поверхні, що суттєво знижує її експлуатаційні якості.

Обґрунтовано ефективність застосування ультразвукових тангенціальних мікроколивань для зменшення контактної енергії в зоні різання. Цей метод дає змогу стабілізувати процес формування стружки, знизити ризик появи мікротріщин та покращити якість поверхні, навіть у складних умовах оброблення тонколистових матеріалів. Розроблено аналітичну модель взаємодії інструмента із заготовкою, яка бере до уваги деформаційні, інерційні та демпфувальні сили в динамічній системі. Модель забезпечує точне прогнозування поведінки інструмента за різних режимів різання та рівнів зношення. Запропоновано вдосконалену конструкцію торцевої фрези з вбудованими демпфувальними елементами. Нова конструкція дає змогу знизити рівень вібрацій у критичних частотних діапазонах, покращити стабільність оброблення та зменшити ймовірність появи хвилястих або структурно неоднорідних дефектів на поверхні листових матеріалів.

Подальші дослідження доцільно присвятити вдосконаленню математичних моделей, що описують динаміку системи "інструмент – заготовка – верстат" з огляду на зміни пружних і демпфувальних властивостей інструментального вузла в реальному часі. Перспективним є розроблення адаптивних систем керування різанням, які інтегрують показники з тривісних акселерометрів і миттєво регулюють режими оброблення для уникнення резонансних коливань. Важливим напрямом подальших робіт є дослідження комбінованого впливу ультразвукових мікроколивань і властивостей змашувально-охолоджувальних рідин на зниження шорсткості поверхні. Крім того, актуальним є вивчення впливу наноструктурованих покриттів на різальну кромку інструмента з метою зменшення тертя й стабілізації траєкторії різання. Доцільно також створити

цифрову платформу прогнозування зносу на основі вібраційних режимів та оптимізації технологічних машинного навчання для визначення критичних параметрів різання.

Список літератури

1. Бойко Ю. І., Литвиненко О. А., Яновський В. А. Оцінка впливу режимів різання на вібрацію розточувального різця і точність оброблених отворів на верстатах з ЧПК. *Технічна інженерія*. 2020. № 2(86). С. 34–40. DOI: [https://doi.org/10.26642/ten-2020-2\(86\)-34-40](https://doi.org/10.26642/ten-2020-2(86)-34-40)
2. Tsiakoumis V. I. An investigation into vibration assisted machining: Application to surface grinding processes. Doctoral thesis. Liverpool John Moores University. 2011. URL: <https://researchonline.ljmu.ac.uk/id/eprint/6113/1/555817.pdf>
3. Rauf A., Khan M. A., Jaffery S. H. I., Butt S. I. Effects of machining parameters, ultrasonic vibrations and cooling conditions on cutting forces and tool wear in meso scale ultrasonic vibrations assisted end-milling (UVAEM) of Ti-6Al-4V under dry, flooded, MQL and cryogenic environments. A statistical analysis. *Journal of Materials Research and Technology*. 2024. P. 8287–8303. DOI: <https://doi.org/10.1016/j.jmrt.2024.05.202>
4. Томашевський О. О., Балицька Н. О. Процес мікрофрезерування металів і сплавів: аналітичний огляд. *Технічна інженерія*. 2023. № 2(92). С. 74–88. DOI: [https://doi.org/10.26642/ten-2023-2\(92\)-74-88](https://doi.org/10.26642/ten-2023-2(92)-74-88)
5. Chamble P., Bharath M. R., Lokesh K., Ezhil Singh S. C. Machine tool vibration on dimensional accuracy and surface roughness during milling operation of Al6082 with indexable carbide inserts. *Journal of Applied Research and Technology*. 2020. 18(2). P. 69–76. DOI: <https://doi.org/10.22201/icat.24486736e.2020.18.2.999>
6. Zhang C., Cheung C., Bulla B., Zhao C. An investigation of the high-frequency ultrasonic vibration-assisted cutting of steel optical moulds. *Micromachines*. 2021. 12(4). 460 p. DOI: <https://doi.org/10.3390/mi12040460>
7. Химич Г. М., Куцевський М. О. Вібрація як джерело динамічних навантажень. Вісник Хмельницького національного університету. 2014. № 1 (209). С. 36–40.
8. Степчин Я. А., Отаманський В. В. Аналіз та вдосконалення алгоритмів модульного принципу створення металорізальних верстатів з ЧПК. *Технічна інженерія*. 2020. № 1(85). С. 23–29. DOI: [https://doi.org/10.26642/ten-2020-1\(85\)-23-29](https://doi.org/10.26642/ten-2020-1(85)-23-29)
9. Сліпчук А. М. Симуляція процесу power skiving для нарізання внутрішнього зубчастого колеса з моделюванням недеформованої стружки. *Інформаційні системи, механіка та керування*. 2023. № 57. С. 46–58. DOI: <https://doi.org/10.23939/istcipa2023.57.046>
10. Makedon V., Myachin V., Plakhotnik O., Fisunen N., Mykhailenko O. Construction of a model for evaluating the efficiency of technology transfer process based on a fuzzy logic approach. *Eastern-European Journal of Enterprise Technologies*. 2024. No 2(13(128)). P. 47–57. DOI: <https://doi.org/10.15587/1729-4061.2024.300796>
11. Пуховський Є. С. Вплив вібрацій на стійкість багатолезового інструменту. *Технічна інженерія*. 2022. № 2 (90). С. 44–51. DOI: [https://doi.org/10.26642/ten-2022-2\(90\)-44-51](https://doi.org/10.26642/ten-2022-2(90)-44-51)
12. Kolomy S., Maly M., Sedlak J., Zouhar J., Slany M., Hrabec P., Kouril K. Machinability of extruded H13 tool steel: Effect of cutting parameters on cutting forces, surface roughness, microstructure, and residual stresses. *Ain Shams Engineering Journal*. 2024. P. 394–407. DOI: <https://doi.org/10.1016/j.aej.2024.05.018>
13. Han G., Ye Z., Xu J., Ma Y. Investigation on cutting forces and tool wear in high-speed milling of Ti-6Al-4V assisted by longitudinal torsional ultrasonic vibrations. *The International Journal of Advanced Manufacturing Technology*. 2023. 129(1–2). P. 1–17. DOI: <https://doi.org/10.1007/s00170-023-12306-z>
14. Rigas K., Willers B., Eckert S., Glaser B. Investigations on vibrational interpretations of bubbles in metal-making processes. *Metallurgical and Materials Transactions*. 2023. № 54. P. 2105–2120. DOI: <https://doi.org/10.1007/s11663-023-02820-6>
15. Македон В. В., Холод О. Г., Ярмоленко Л. І. Модель оцінки конкурентоспроможності високотехнологічних підприємств на засадах формування ключових компетенцій. *Академічний огляд*. 2023. № 2 (59). С. 75–89. DOI: <https://doi.org/10.32342/2074-5354-2023-2-59-5>
16. Wang S.M., Tsou W.S., Huang J.W., Chen S.E., Wu C.C. Development of a method and a smart system for tool critical life real-time monitoring. *Journal of Manufacturing and Materials Processing*. 2024. 8(5). 194 p. <https://doi.org/10.3390/jmmp8050194>

References

1. Boiko, Yu. I., Lytvynenko, O. A., Yanovskyi, V. A. (2020), "Evaluation of the influence of cutting regimes on the vibration of the boring tool and the accuracy of machined holes on CNC machines" ["Otsinka vplyvu rezhymiv rizannia na vibratsiiu roztockhuvalnoho rizzia i tochnist obrobliuvanykh otvoriv na verstatakh z ChPK"], *Technical Engineering*, No. 2(86), P. 34–40. DOI: [https://doi.org/10.26642/ten-2020-2\(86\)-34-40](https://doi.org/10.26642/ten-2020-2(86)-34-40)

2. Tsiakoumis, V. I. (2011), "An investigation into vibration assisted machining: Application to surface grinding processes", Doctoral thesis, Liverpool John Moores University. available at: <https://researchonline.ljmu.ac.uk/id/eprint/6113/1/555817.pdf>.
3. Tomashevskiy, O. O., Balytska, N. O. (2023), "Micro-milling process of metals and alloys: analytical review" ["Protses mikrofrezeruvannia metaliv i splaviv: analitychnyi ohliad"], *Technical Engineering* [Tekhnichna Inzheneriia], No. 2(92), P. 74–88. DOI: [https://doi.org/10.26642/ten-2023-2\(92\)-74-88](https://doi.org/10.26642/ten-2023-2(92)-74-88)
4. Rauf, A., Khan, M. A., Jaffery, S. H. I., Butt, S. I. (2024), "Effects of machining parameters, ultrasonic vibrations and cooling conditions on cutting forces and tool wear in meso scale ultrasonic vibrations assisted end-milling (UVAEM) of Ti-6Al-4V under dry, flooded, MQL and cryogenic environments – A statistical analysis", *Journal of Materials Research and Technology*. DOI: <https://doi.org/10.1016/j.jmrt.2024.05.202>
5. Chamble, P., Bharath, M. R., Loksha, K., Ezhil Singh, S. C. (2020), "Machine tool vibration on dimensional accuracy and surface roughness during milling operation of Al6082 with indexable carbide inserts", *Journal of Applied Research and Technology*, Vol. 18, No. 2, P. 69–76. DOI: <https://doi.org/10.22201/icat.24486736e.2020.18.2.999>
6. Zhang, C., Cheung, C., Bulla, B., Zhao, C. (2021), "An investigation of the high-frequency ultrasonic vibration-assisted cutting of steel optical moulds", *Micromachines*, Vol. 12, No. 4, 460 p. DOI: <https://doi.org/10.3390/mi12040460>
7. Khimych, H. M., Kushchevskiy, M. O. (2014), "Vibration as a source of dynamic loads" ["Vibratsiia yak dzherelo dynamichnykh navantazhen"], *Visnyk Khmelnytskoho natsionalnoho universytetu*. [Herald of Khmelnytskyi National University], No. 1(209), P. 36–40.
8. Stepchyn, Ya. A., Otamanskyi, V. V. (2020), "Analysis and improvement of algorithms of the modular principle for creating CNC machine tools" ["Analiz ta vdoskonalennia alghorytmiv modulnoho pryntsyphu stvorennia metalozrizalnykh verstativ z ChPK"], *Technical Engineering* [Tekhnichna Inzheneriia], No. 1(85), P. 23–29. DOI: [https://doi.org/10.26642/ten-2020-1\(85\)-23-29](https://doi.org/10.26642/ten-2020-1(85)-23-29)
9. Slipchuk, A. M. (2023), "Simulation of the power skiving process for cutting an internal gear with undeformed chip modeling" ["Symuliatsiia protsesu power skiving dlia narizannia vnutrishnoho zubchastoho kola z modeliuvanniam nedeformovanoi struzhky"], *Information Systems, Mechanics and Control* [Informatsiini systemy, mekhanika ta keruvannia], No. 57, P. 46–58. DOI: <https://doi.org/10.23939/istcipa2023.57.046>
10. Makedon, V. V., Kholod, O. H., Yarmolenko, L. I. (2023), "Model for assessing the competitiveness of high-tech enterprises based on the formation of key competencies" ["Model otsinky konkurentospromozhnosti vysokotekhnolohichnykh pidpriemstv na zasadakh formuvannia kluchovykh kompetentsii"], *Academic Review* [Akademichnyi ohliad], No. 2(59), P. 75–89. DOI: [10.32342/2074-5354-2023-2-59-5](https://doi.org/10.32342/2074-5354-2023-2-59-5)
11. Pukhovskiy, Ye. S. (2022), "Influence of vibrations on the durability of multi-blade tool" ["Vplyv vibratsii na stiiikist bahatozozovoho instrumentu"], *Technical Engineering* [Tekhnichna Inzheneriia], No. 2(90), P. 44–51. DOI: [10.26642/ten-2022-2\(90\)-44-51](https://doi.org/10.26642/ten-2022-2(90)-44-51)
12. Kolomy, S., Maly, M., Sedlak, J., Zouhar, J., Slany, M., Hrabec, P., Kouril, K. (2024), "Machinability of extruded H13 tool steel: Effect of cutting parameters on cutting forces, surface roughness, microstructure, and residual stresses", *Ain Shams Engineering Journal*. P. 394–407. DOI: <https://doi.org/10.1016/j.aej.2024.05.018>
13. Han, G., Ye, Z., Xu, J., Ma, Y. (2023), "Investigation on cutting forces and tool wear in high-speed milling of Ti-6Al-4V assisted by longitudinal torsional ultrasonic vibrations", *The International Journal of Advanced Manufacturing Technology*, Vol. 129, No. 1–2, P. 1–17. DOI: <https://doi.org/10.1007/s00170-023-12306-z>
14. Rigas, K., Willers, B., Eckert, S., Glaser, B. (2023), "Investigations on vibrational interpretations of bubbles in metal-making processes", *Metallurgical and Materials Transactions*. Vol. 54, P. 2105–2120. DOI: <https://doi.org/10.1007/s11663-023-02820-6>
15. Makedon, V., Myachin, V., Plakhotnik, O., Fisunen, N., Mykhailenko, O. (2024), "Construction of a model for evaluating the efficiency of technology transfer process based on a fuzzy logic approach", *Eastern-European Journal of Enterprise Technologies*, No 2(13(128)), P. 47–57. DOI: <https://doi.org/10.15587/1729-4061.2024.300796>
16. Wang, S.-M., Tsou, W.S., Huang, J.W., Chen, S.-E., & Wu, C.-C. (2024), "Development of a method and a smart system for tool critical life real-time monitoring", *Journal of Manufacturing and Materials Processing*, Vol. 8, No. 5, 194 p. <https://doi.org/10.3390/jmmp8050194>

Відомості про авторів / About the Authors

Зенкін Микола Анатолійович – доктор технічних наук, професор, Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", професор кафедри машин та агрегатів поліграфічного виробництва, Київ, Україна; e-mail: nikolay_zenkin@ukr.net; ORCID ID: <https://orcid.org/0000-0002-8840-0572>

Іванко Андрій Іванович – кандидат технічних наук, доцент, Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", доцент кафедри машин та агрегатів поліграфічного виробництва, Київ, Україна, e-mail: ivanko-a@ukr.net; ORCID ID: <https://orcid.org/0000-0002-4735-9665>

Черниш Максим Леонідович – Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", аспірант кафедри машин та агрегатів поліграфічного виробництва, Київ, Україна; e-mail: maxim073074@gmail.com; ORCID ID: <https://orcid.org/0009-0003-9376-6055>

Zenkin Mykola – Doctor of Sciences (Engineering), Professor, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Professor at the Department of Printing Machines and Automated Complexes, Kyiv, Ukraine.

Ivanko Andrii – PhD (Engineering Sciences), Associate Professor, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Associate Professor at the Department of Printing Machines and Automated Complexes, Kyiv, Ukraine.

Chernysh Maksym – National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Postgraduate Student at the Department of Printing Machines and Automated Complexes, Kyiv, Ukraine.

INFLUENCE OF CUTTING TOOL VIBRATIONS ON THE SURFACE QUALITY OF CUT SHEET MATERIALS AND METHODS FOR THEIR MINIMIZATION

The subject of the study is the influence of cutting tool vibrations on the surface quality of cut sheet materials during mechanical processing, as well as the analysis of the mechanisms of vibration occurrence and the development of methods for their minimization to increase the efficiency of technological operations in mechanical engineering. **The purpose** of the work is to develop scientifically based approaches to predicting the resource of cutting tools and optimizing cutting modes by reducing vibration loads, which will allow to improve the quality of surface processing of sheet materials, reduce production costs and prevent processing defects.

Tasks: assessment of the vibration characteristics of the tool depending on the cutting modes and the degree of wear; analysis of the influence of vibrations on surface roughness and defects (waviness, microcracks); development and testing of vibration minimization methods, including damping systems and ultrasonic vibrations; determination of optimal processing conditions for various materials.

Methods: experimental analysis using accelerometers for three-axis vibration monitoring; profilometry and microscopy for surface quality assessment; stochastic modeling of tool wear based on statistical data; testing of damping systems (elastic couplings, dampers) and ultrasonic vibrations (20–30 kHz); comparative analysis of the effectiveness of lubricating and cooling fluids.

Research results: it was found that vibrations along the X axis have the greatest impact on the formation of surface defects, especially with increasing cutting depth and tool wear; ultrasonic vibrations (22 kHz, amplitude 10 μm) reduce friction and stabilize the cutting process, improving the quality of machined surfaces; damping systems effectively reduce the amplitude of vibrations at critical frequencies (up to 1.5 kHz); the use of water as a coolant turned out to be more effective than copper oil, reducing surface irregularities; an improved model of an end mill with damping elements was developed, which reduces vibration loads and defects when processing thin-sheet materials. **Conclusions:** the conducted studies show that vibration monitoring is an effective tool for diagnosing the condition of the cutting tool and optimizing the cutting process; the integration of damping systems and ultrasonic vibrations allows you to significantly reduce vibration effects, improve the surface quality of sheet materials and extend the tool life, which contributes to the economic efficiency of production and minimizing processing defects.

Keywords: vibrations, cutting tool, surface quality, sheet materials, wear, vibration minimization.

Бібліографічні описи / Bibliographic descriptions

Зенкін М. А., Іванко А. І., Черниш М. Л. Вплив вібрацій різального інструмента на якість поверхні обрізаних листових матеріалів та методи їх мінімізації. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 2 (32). С. 188–198. DOI: <https://doi.org/10.30837/2522-9818.2025.2.188>

Zenkin, M., Ivanko, A., Chernysh, M. (2025), "Influence of cutting tool vibrations on the surface quality of cut sheet materials and methods for their minimization", *Innovative Technologies and Scientific Solutions for Industries*, No. 2 (32), P. 188–198. DOI: <https://doi.org/10.30837/2522-9818.2025.2.188>

АЛФАВІТНИЙ ПОКАЖЧИК

Андрєєв С. М.	16
Артѣомов І. В.	111
Артѣомова А. В.	111
Артюх Р. В.	58
Барковська О. Ю.	5
Бизкровний О. М.	177
Бушуєв С. Д.	168
Бушуєва Н. С.	168
Головченко О. С.	5
Даншина С. Ю.	16
Євсєєв В. В.	58
Зенкін М. А.	188
Іванко А. І.	188
Косарєвський Б. В.	33
Костін А. О.	5
Крицький Д. М.	111
Легезін Н. В.	5
Лобок Є. А.	168
Максимова С. С.	58
Мурованський Г. А.	168
Мухін В. Є.	49
Невлюдов І. Ш.	58
Поддубний В. О.	69
Поляков А. О.	151
Прокопенко О. С.	79
Сєверінов О. В.	69
Сироткін К. В.	102
Смеляков К. С.	177
Смеляков С. В.	79
Сопов Є. О.	111
Сторчай Д. Г.	5
Тецький А. Г.	33
Удовенко С. Г.	129
Хабло Я. О.	49
Чала Л. Е.	129
Черниш М. Л.	188
Чухран І. Д.	129
Шергін В. Л.	129
Шкурко В. В.	151

ALPHABETICAL INDEX

Andrieiev Sergey	16
Artomov Ihor	111
Artomova Alina	111
Artiukh Roman	58
Barkovska Olesia	5
Byzkrovnyi Oleksandr	177
Bushuyev Sergiy	168
Bushuyeva Natalia	168
Holovchenko Oleksandr	5
Danshyna Svitlana	16
Yevsieiev Vladyslav	58
Zenkin Mykola	188
Ivanko Andrii	188
Kosarevskyi Bohdan	33
Kostin Anton	5
Krytskyi Dmytro	111
Lehezin Nikita	5
Lobok Yevgen	168
Maksymova Svitlana	58
Murovansky Hleb	168
Mukhin Vadym	49
Nevliudov Igor	58
Poddubnyi Vadym	69
Poliakov Andrii	151
Prokopenko Oleksandr	79
Sievierinov Oleksandr	69
Syrotkin Kostiantyn	102
Smelyakov Kyrlyo	177
Smelyakov Serhii	79
Sopov Iegor	111
Storchai Denis	5
Tetskyi Artem	33
Udoenko Serhii	129
Khablo Yaroslav	49
Chala Larysa	129
Chernysh Maksym	188
Chukhran Ivan	129
Shergin Vadim	129
Shkurko Viacheslav	151

НАУКОВЕ ВИДАННЯ

SCIENTIFIC PUBLICATION

Сучасний стан наукових досліджень та технологій в промисловості

Innovative technologies and scientific solutions for industries

Щоквартальний науковий журнал

Quarterly scientific journal

№ 2 (32), 2025

No. 2 (32), 2025

Відповідальний секретар журналу *І.Г. Перова*
Відповідальний за випуск *А.А. Коваленко*
Відповідальний за ліцензування *В.В. Косенко*
Редактор *Л.В. Кузьміна*
Комп'ютерна верстка *Л.Ю. Свєтайло*

Responsible secretary of journal *I. Perova*
Responsible for release *A. Kovalenko*
Responsible for licensing *V. Kosenko*
Editor *L. Kuzmina*
Computer layout *L. Svietailo*

Формат 60×84/8. Умов. друк. арк. 23,25.
Тираж 150 прим.

Format 60×84/8. Conventional printed sheets 23,25.
Edition of 150 copies.

Віддруковано з готових оригінал-макетів
в типографії ФОП Андреев К.В.
Єдиний державний реєстр юридичних осіб
та фізичних осіб-підприємців.
Запис №24800170000045020 від 30.05.2003.

Printed from ready-made original layouts
in the printing house
of Individual Entrepreneur Andreev K.V.
Unified State Register of Legal Entities
and Individual Entrepreneurs.
Entry No. 24800170000045020 of 30.05.2003.

61157, Харків, вул. Акад. Богомольця, 9, кв. 50,
тел. +38 (063) 993-62-73
e-mail: ep.zakaz@gmail.com

fl. 50, 9, Acad. Bogomolets Str., Kharkiv, 61157,
тел. +38 (063) 993-62-73
e-mail: ep.zakaz@gmail.com