



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

№ 1 (35), 2026

Затверджений до друку
Науково-технічною Радою
Харківського національного університету радіоелектроніки
(Протокол № 2 від 24 лютого 2026 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. 1 (35), 2026

Approved for publication
by the Scientific and Technical Council
of the Kharkiv National University of Radio Electronics
(Protocol No. 2 d.d. 24 February 2026)

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 (категорія А), в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора і кандидата наук наказом Міністерства освіти і науки України від 23.12.2025 № 1683

Харків – 2026

© 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;
Косєнко Віктор Васильович, д-р техн. наук, професор;
Костін Юрій Дмитрович, д-р екон. наук, професор;
Лєвашєнко Віталій, д-р техн. наук, професор (Словаччина);
Лємєшкє Олєксандр Віталійович, д-р техн. наук, професор;
Малєєва Олєга Володимирівна, д-р техн. наук, професор;
Момот Тєтяна, д-р екон. наук, професор (США);
Музикє Катєрина Миколєївнє, д-р техн. наук, професор;
Назарєва Гєлєна Валєнтинівнє, д-р екон. наук, професор;
Нєвлюдєв Ігор Шєкирєвич, д-р техн. наук, професор;
Опанєсєук Анєтєлєй Сєргїєвич,
д-р фіз.-мат. наук, професор;
Пєвлєв Сєргїєй Володимирєвич, д-р техн. наук, професор;
Пєржєнн Юрїєй, д-р техн. наук, професор (США);
Пєрєвє Ірїнє Гєннєдїєвнє, д-р техн. наук, професор;
Пєтлєнкєв Едурєд, канд. техн. наук (Єстєнїє);
Пєтрїшїн Лєбємїр, д-р техн. наук, професор (Польщє);
Пїтєрськє Варвєрє Мїхєйлїєвнє, д-р техн. наук, професор;
Рубєн Ігор Віктєрєвич, д-р техн. наук, професор;
Сємєнєць Валєрїєй Васїльєвич, д-р техн. наук, професор;
Сємєнєв Сєргїєй, д-р техн. наук, професор (Польщє);
Сєтлєк Гєлєнє, д-р. техн. наук, професор (Польщє);
Тєрзїєн Вєгєн, д-р техн. наук, професор (Фїнлєндїє);
Тєлєтєв Олєксєндр Сєргїєєвич, д-р екон. наук, професор;
Тїмєфєєв Володимир Олєксєндрєвич,
д-р техн. наук, професор;
Фїлєтєв Валєнтїн Олєксєндрєвич, д-р техн. наук, професор;
Чумєчєнкє Ігор Володимирєвич, д-р техн. наук, професор;
Чухрєй Нєтєлїєнє Івєнїєвнє, д-р екон. наук, професор;
Юн Джєн,
канд. фіз.-мат. наук, професор (КНР);
Ястрємськє Олєнє Миколєївнє, 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 Victors, 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 Yuriy, 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);
Piterska Varvara, Dr. Sc. (Engineering), Professor
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);
Iastremaska Olena, Dr. Sc. (Economics), Professor (Ukraine)

ЗМІСТ

Інформаційні технології

- 6 **Вакулік Є. В., Смеляков К. С., Чуприна А. С.**
Модель і метод автоматизованого високоточного вимірювання тривимірних об'єктів за даними комп'ютерної томографії (ua)
- 17 **Досталова К., Чижмарова, О., Клімо М., Ярошевич Р.О.**
Виявлення пухлин головного мозку на МРТ-зображеннях з можливістю пояснення використання YOLOv8 у LIME (en)
- 28 **Єна М.В., Погудіна О.К.**
Моделювання адаптивного управління маршрутами БПЛА на основі алгоритмів навчання з підкріпленням (en)
- 39 **Коломійцев О. В., Гулевич М. В., Дмитрієв О. М., Левченко А. О., Балабуха О. С.**
Оцінювання ефективності вдосконалених жадібного й дельта-дебагінг алгоритмів оптимізації тестових сценаріїв для C++ бібліотек (ua)
- 55 **Крайчі О., Сидоренко Л., Барковська О.Ю.**
Надійне прогнозування серцево-судинного ризику на невеликих наборах даних з використанням передової інженерії ознак (en)
- 65 **Кучук Г. А., Косенко Н. В., Кучук Н. Г., Гонєєнко В., Косенко В. В.**
Адаптивний метод розподілу ресурсів мобільного туманного шару високошільного промислового Інтернету речей у мережах Industry 5.0 (en)
- 79 **Саченко А. О., Радюк П. М., Лисий М. І., Мельниченко О. В., Заставний О. М.**
Архітектура кіберфізичних систем для виявлення дефектів у фотоелектричних модулях на основі пізнього злиття даних із БПЛА (en)
- 100 **Jian Yu, Сітнікова О.О.**
Метод прогнозування траєкторій та ухилення БПЛА для промислових автономних місій у динамічному середовищі (en)
- 115 **Соловей Б. А., Терентьєв О. О.**
Підхід до створення фізично достовірних даних для систем моніторингу безпеки кранів на основі скінченно-елементного моделювання (ua)
- 127 **Шапиро О.К., Сотник І.С.**
Метод семантичного пошуку навчального контенту на основі мультиагентних технологій (en)

Сучасні технології управління підприємством

- 139 **Матківська Х.С., Зачко О.Б., Тригуба А.М.**
Система цифрового HR-менеджменту безпекових структур на основі е-кабінету працівника (ua)

Інженерія та промислові технології

- 155 **Невлюдов І. Ш., Рубан О. А., Нікітін Д. О., Місан Б. С., Іохов О. Ю.**
Дослідження демпфувальних властивостей 3D-друкованих лайнерів з керованою перфорацією (en)
- 169 **Алфавітний покажчик**

За достовірність викладених фактів, цитат та інших відомостей відповідальність несе автор

CONTENTS

Information Technology

- 6 **Vakulik E., Smelyakov K., Chupryna A.**
Model and method of automated high-precision measurement of three-dimensional objects based on computed tomography data (ua)
- 17 **Dostalova K., Cizmarova A., Klimo M., Yaroshevych R.**
Explanation Detected Brain Tumours in MRI Images Using YOLOv8 with LIME-Based Interpretation (en)
- 28 **Yena M., Pohudina O.**
Modeling of adaptive UAV route control based on reinforcement learning algorithms (en)
- 39 **Kolomiitsev O., Hulevych M., Dmitriiev O., Levchenko A., Balabukha O.**
Evaluation of the effectiveness of the enhanced greedy and delta-debugging test case optimization algorithms for C++ libraries (ua)
- 55 **Krajči A., Sidorenko L., Barkovska O.**
Predicting Risks of Cardiovascular Disease on Small Datasets using Feature Engineering (en)
- 65 **Kuchuk H., Kosenko N., Kuchuk N., Gopejenko V., Kosenko V.**
Adaptive Resource Allocation Method for the Mobile Fog Layer of High-Density Industrial Internet of Things in Industry 5.0 Networks (en)
- 79 **Sachenko A., Radiuk P., Lysyi M., Melnychenko O., Zastavnyy O.**
Architecture of cyberphysical systems for UAV-based late-fusion defect detection in photovoltaic modules (en)
- 100 **Jian Yu, Sitnikova O.**
Method for predicting UAV trajectories and evasion for industrial autonomous missions in a dynamic environment (en)
- 115 **Solovei B., Terentyev O.**
An approach to generating physically realistic data for crane safety monitoring systems based on finite element modeling (ua)
- 127 **Shapyro O., Sotnyk I.**
A method of semantic search for educational content based on multi-agent technologies (en)

Modern Enterprise Management Technology

- 139 **Matkivska K., Zachko O., Tryhuba A.**
Digital HR management system for security structures based on the employee's e-cabinet (ua)

Engineering & Industry Technologies

- 155 **Nevlyudov I., Ruban O., Nikitin D., Misan B., Iokhov O.**
Investigation of the damping properties of 3D-printed liners with controlled perforation (en)
- 169 **Alphabetical index**

The author is responsible for the accuracy of the facts, quotations and other information



Головний редактор журналу, доктор технічних наук, професор
Євгеній БОДЯНСЬКИЙ
 Editor-in-Chief, Doctor of Technical Sciences, Professor
Yevgeniy BODYANSKIY

Шановні колеги!

З великою радістю та гордістю відкриваю цей випуск нашого журналу "Сучасний стан наукових досліджень та технологій в промисловості" в особливий для нас час. Нещодавно наш колектив досяг двох знакових результатів, що підтверджують високий науковий рівень, відкритість і міжнародну визнаність видання.

По-перше, журнал офіційно прийнято до міжнародної наукометричної бази Scopus. Це відкриває нашим авторам шлях до міжнародної наукової спільноти.

По-друге, наказом Міністерства освіти і науки України журнал включено до категорії А Переліку наукових фахових видань. Це важливе визнання відповідності національним стандартам якості.

Ці досягнення – заслуга й плід наполегливої праці всієї команди. Я щиро дякую:

- членам редакційної колегії за експертне керівництво й стратегічне бачення;
- науковим рецензентам за безкорисливу, ретельну працю й конструктивні відгуки;
- шановним авторам, без чийх інноваційних досліджень і довіри до нашого видання цей успіх був би неможливим;
- технічному персоналу за бездоганну організацію процесів і підтримку.

Цей випуск журналу, як і попередні, містить цінні наукові статті, що висвітлюють сучасні тенденції в промисловості й технологіях. Кожна публікація – це новий крок у розвитку знань.

Бажаю всім нам не втрачати наукового пориву! Разом ми можемо підтримувати високий стандарт і далі розвивати наше спільне наукове діло.

З повагою та вдячністю, головний редактор,
 професор Євгеній Володимирович Бодянський

Dear colleagues!

It is with great joy and pride that I open this issue of our journal "The Current State of Scientific Research and Technology in Industry" at a special time for us. Recently, our team has achieved two significant results that confirm the high scientific level, openness, and international recognition of the publication.

First, the journal has been officially accepted into the international scientometric database Scopus. This opens the way for our authors to the international scientific community.

Second, by order of the Ministry of Education and Science of Ukraine, the journal has been included in category A of the List of Scientific Professional Publications. This is an important recognition of compliance with national quality standards.

These achievements are the result of the hard work of the entire team. I sincerely thank:

- the members of the editorial board for their expert guidance and strategic vision;
- the scientific reviewers for their selfless, diligent work and constructive feedback;
- esteemed authors, without their innovative research and trust in our publication, this success would not have been possible;
- technical staff for their impeccable organization of processes and support.

This issue of the journal, like the previous ones, contains valuable scientific articles highlighting current trends in industry and technology. Each publication is a new step in the development of knowledge.

I wish all of us not to lose our scientific momentum! Together, we can maintain high standards and further develop our joint scientific endeavors.

With respect and gratitude, Editor-in-Chief,
 Professor Yevgeniy Bodyanskiy

Вакулік Є. В., Смеляков К. С., Чуприна А. С.

МОДЕЛЬ І МЕТОД АВТОМАТИЗОВАНОГО ВИСОКОТОЧНОГО ВИМІРЮВАННЯ ТРИВИМІРНИХ ОБ'ЄКТІВ ЗА ДАНИМИ КОМП'ЮТЕРНОЇ ТОМОГРАФІЇ

Предметом вивчення є лінійні розміри тривимірних об'єктів за даними комп'ютерної томографії, зокрема металевих чужорідних тіл у тканинах і органах людини. **Мета дослідження** – розробити математичну модель і метод автоматизованого високоточного вимірювання лінійних розмірів тривимірних об'єктів за результатами комп'ютерної томографії з реалізацією у вигляді програмного забезпечення. **Завдання:** сформувати математичну модель подання тривимірного об'єкта у воксельному просторі; розробити метод сегментації металевих фрагментів за результатами комп'ютерної томографії; запропонувати метод просторового вирівнювання сегментованих об'єктів на підставі аналізу основних компонент; розробити метод визначення максимального лінійного розміру об'єкта в новій координатній системі; реалізувати запропоновану модель і метод у вигляді програмного забезпечення й експериментально перевірити точність вимірювання. **Методи дослідження:** аналіз томографічних показників, порогова сегментація з адаптивним вибором порогових значень, хвильовий алгоритм пошуку зв'язаних компонент, аналіз основних компонент для визначення орієнтації об'єкта, воксельне моделювання й обчислення евклідових відстаней між граничними точками тривимірного об'єкта. **Досягнуті результати.** Запропоновано математичну модель подання тривимірного об'єкта й метод автоматизованого високоточного визначення його максимального лінійного розміру. Метод реалізовано у вигляді програмного модуля й апробовано на 72 зразках металевих фрагментів шести типів, розміщених у біологічних тканинах органів свині. Середнє відхилення не перевищує 3%, а в найскладніших випадках залишається в межах 5%, що свідчить про високу точність і стабільність запропонованого підходу. **Висновки:** розроблена модель і метод забезпечують автоматизоване й об'єктивне визначення лінійних розмірів чужорідних тіл за даними комп'ютерної томографії без участі оператора. Запропоноване програмне забезпечення може бути використане у військовій медицині, судово-медичній експертизі, медицині катастроф і закладах охорони здоров'я, де критично важливими є швидкість і достовірність діагностичних рішень.

Ключові слова: комп'ютерна томографія; сегментація; метод основних компонент; лінійні розміри; DICOM; воксель.

Вступ

У сучасних умовах збройних конфліктів, терористичних актів і технологічних катастроф значно зростає кількість випадків ураження людини чужорідними тілами, зокрема металевими фрагментами. Такі ураження мають складну морфологію, а процес їх діагностики потребує високоточної візуалізації та аналізу внутрішніх структур організму. Одним із найбільш ефективних засобів такої візуалізації є комп'ютерна томографія (КТ), яка дає змогу отримати тривимірне зображення анатомічних об'єктів з високою просторовою роздільною здатністю. Однак навіть за наявності якісних зображень виникає потреба в автоматизованих засобах оброблення даних, що допомагають швидко й точно оцінювати параметри виявлених фрагментів, зокрема їх лінійні розміри. Традиційні методи аналізу КТ-зображень здебільшого передбачають візуальне оцінювання лікарем-рентгенологом, що обмежує точність і відтворюваність результатів, особливо в умовах стресу, обмеженого часу або відсутності відповідного досвіду. Крім того, особливості

металевих фрагментів – різноманітність форм, відсутність чітких контурів, артефакти візуалізації – суттєво ускладнюють процес ідентифікації та вимірювання [1, 2]. У практиці трапляються ситуації, коли фрагмент розміщується під кутом до площини сканування або частково перекритий іншими анатомічними структурами. У цьому разі навіть незначна похибка у вимірюванні може мати критичне значення для вибору клінічної тактики.

Наразі існує низка програмних рішень, орієнтованих на оброблення КТ-знімків, однак більшість з них не забезпечує автоматичного вимірювання параметрів чужорідних тіл або не адаптована до особливостей тканин різних органів і завдань військової медицини [3]. Крім цього, доступні на ринку програми часто є комерційними, складними у використанні або не мають відкритого доступу до алгоритмів, що унеможливило їх модифікацію та інтеграцію у власні інформаційно-аналітичні системи. Отже, постає актуальне завдання – розробити спеціалізовану методiku, яка б давала змогу повністю автоматизувати процес ідентифікації, сегментації та вимірювання

лінійних розмірів металевих фрагментів на основі КТ-даних.

Запропоноване дослідження є частиною міждисциплінарного проекту, спрямованого на створення засобів підтримки медичних рішень у сфері військової та екстреної медицини [4, 5]. У межах цієї роботи було поставлено завдання – створення математичної моделі та відповідного програмного забезпечення для автоматичного аналізу КТ-знімків, що містять металеві чужорідні тіла (далі – металеві фрагменти). Ключовими вимогами стали: висока точність, повна автоматизація, відсутність необхідності у втручанні користувача й можливість масштабування для різних типів тканин і об'єктів. Практична цінність таких рішень поширюється на сфери судово-медичної експертизи, медицини катастроф, а також інформаційних технологій для оборонного сектору.

Водночас запропонований підхід не обмежується зазначеним класом об'єктів і може бути застосований для автоматизованого вимірювання лінійних розмірів інших тривимірних об'єктів, виявлених за показниками КТ.

Аналіз останніх досліджень і публікацій

Проблема точного визначення параметрів сторонніх предметів у тілі людини, зокрема в умовах військово-польової медицини й екстреної діагностики, привертає увагу фахівців не перше десятиліття. Основними напрямками досліджень є методи автоматизованої сегментації об'єктів на КТ-зображеннях, алгоритми відновлення форми й розмірів за томографічними даними, а також способи зменшення впливу артефактів, зумовлених присутністю металу.

Сучасні узагальнювальні огляди демонструють збалансоване поєднання класичних підходів і новітніх технологій, зокрема тих, що базуються на глибокому навчанні, для зменшення металевих артефактів у КТ-зображеннях. Розрізняють методи MAR, алгоритмічні та реконструкційні стратегії [6–8].

Паралельно з розвитком MAR-алгоритмів продовжуються роботи над універсальними рішеннями із сегментації та реконструкції тіл на КТ-зображеннях. Додаються принципи подвійного домену, що дає змогу підвищити точність визначення форми й геометричних властивостей, зберігаючи фізичну трактовку моделей [9, 10].

Не менш важливим є напрям, спрямований на усунення варіацій, спричинених різними виробниками обладнання, а також стандартизацію метаданих (формати DICOM). Такі підходи закладають основу для побудови міжсистемних платформ машинного аналізу, які забезпечують стабільність результатів незалежно від джерела даних [11].

Мета й завдання дослідження

Незважаючи на значні досягнення у сфері медичної візуалізації, автоматизоване визначення параметрів чужорідних тіл, зокрема їх лінійних розмірів, залишається технічно складним завданням. Особливо це актуально тоді, коли йдеться про металеві об'єкти з довільною просторовою орієнтацією, складною формою та наявністю візуальних артефактів на КТ-зображеннях.

Метою цього дослідження є розроблення математичної моделі й методу автоматизованого високоточного вимірювання лінійних розмірів тривимірних об'єктів за результатами комп'ютерної томографії з подальшою програмною реалізацією. У межах роботи основну увагу приділено вимірюванню металевих фрагментів у біологічних тканинах як найбільш складному з практичного погляду випадку.

Для досягнення окресленої мети необхідно виконати такі завдання: сформувати математичну модель подання тривимірного об'єкта у воксельному просторі; розробити метод сегментації металевих фрагментів на базі томографічних даних; створити метод просторового вирівнювання сегментованих об'єктів на основі аналізу основних компонент; розробити метод визначення максимального лінійного розміру об'єкта в новій координатній системі; реалізувати запроповану модель і метод у вигляді програмного забезпечення й експериментально перевірити точність вимірювання.

Матеріали й методи

Узагальнений алгоритм запропонованого методу передбачає послідовне оброблення томографічних показників. На першому етапі дані комп'ютерної томографії у форматі DICOM перетворюються у тривимірну воксельну матрицю щільності. Далі виконується первинна сегментація металевих фрагментів на основі порогового значення щільності,

що дає змогу виокремити потенційні ділянки об'єктів. На наступному етапі застосовується додатковий адаптивний поріг, який визначається залежно від максимального значення щільності в межах зв'язаної ділянки, що допомагає відокремити безпосередньо фрагмент від зони тінювих артефактів. Після цього формується уточнене геометричне подання об'єкта способом додавання граничних точок вокселів, що підвищує точність опису його форми в просторі. Для визначення орієнтації об'єкта впроваджено метод основних компонент, який дає змогу побудувати нову систему координат, узгоджену з напрямками найбільшого поширення фрагмента. У цій системі координат виконується обчислення максимального лінійного розміру об'єкта. Подальші підрозділи розкривають особливості реалізації зазначених етапів.

Підготовка даних

Дані комп'ютерної томографії (КТ), що використовуються в дослідженні, зберігаються у форматі DICOM – загальноприйнятому міжнародному стандарті медичної візуалізації [3, 12]. Кожний файл містить серію поперечних знімків і повний набір метаданих: розміри вокселів у площині, товщину шару, відстань між знімками, шкалу градацій, тип томографа та інші параметри. Ця інформація критично важлива для точного перетворення координат і метричних оцінок.

На початковому етапі зображення перетворюються у тривимірну воксельну матрицю щільності $P(i, j, k)$, задану в шкалі Хаунсфілда [13]. Така форма подання

$$x_1 = 0.5 \cdot d + i \cdot d, x_2 = 0.5 \cdot d + j \cdot d, x_3 = 0.5 \cdot h + k \cdot h, \quad (1)$$

де d – розмір вокселя; h – товщина шару; (i, j, k) – індекси. Щоб уникнути втрат точності на межах об'єкта, додаються координати вершин, центрів

$$x_1 - 0.5 \cdot d, x_1, x_1 + 0.5 \cdot d; x_2 - 0.5 \cdot d, x_2, x_2 + 0.5 \cdot d; x_3 - 0.5 \cdot h, x_3, x_3 + 0.5 \cdot h. \quad (2)$$

Сукупність додаткових точок визначається як декартовий добуток зазначених множин координат. Загальна кількість просторових комбінацій становить 27, одна з яких відповідає центру вокселя, що вже врахований у формулі (1). Отже, додатково додаються 26 нових точок, що уточнюють геометричне подання меж об'єкта. Це дає змогу точніше відтворювати форму, об'єм і граничні ділянки фрагмента в просторі.

визначає математичну модель тривимірного об'єкта як дискретну просторову структуру, елементи якої визначаються значеннями щільності та геометричними координатами у воксельному просторі. У разі потреби застосовуються методи попереднього оброблення – градаційна корекція, пригнічення шумів, покращення локального контрасту, – що покращує видимість меж металевих тіл і знижує похибки сегментації.

Сегментація відбувається у два етапи. Спочатку вокселі, що перевищують певне порогове значення щільності T_{seg} , вважаються частинами потенційних фрагментів. Цей поріг встановлюється експертно й залежить від типу тканин і особливостей томографії. Далі за допомогою хвильового алгоритму будується множина зв'язаних вокселів Q , кожна з яких відповідає окремому фрагменту.

Нехай v – значення щільності відповідного вокселя КТ-зображення. Оскільки фрагмент оточений зоною змішаної щільності (тінь), вводиться другий поріг $T_{ob} = c \cdot v_{max}$, де v_{max} – максимальна щільність у межах Q , а $c \in (0,1)$ – емпірично підібраний коефіцієнт. Фрагмент вважається множиною $O \subset Q$, де $v \geq T_{ob}$, решта $Q \setminus O$ – тінь. Це дає змогу мінімізувати похибки, зумовлені впливом навколишнього середовища.

Координати x_1, x_2, x_3 , отримані з КТ-зрізів, надалі розглядаються як декартові координати точок у тривимірному просторі й позначаються у векторній формі.

Координати вокселів перетворюються у фізичні значення за формулами:

граней і середини ребер кожного вокселя. Для цього по кожній осі розглядаються такі множини можливих значень координат:

Базова математична модель оцінювання лінійного розміру металевго фрагмента

У контексті медичної діагностики й клінічного аналізу надзвичайно важливим є визначення максимального лінійного розміру чужорідного тіла, зокрема металевго фрагмента. Саме довжина часто є критерієм прийняття медичних рішень, адже її значення може безпосередньо впливати на тактику

лікування, необхідність хірургічного втручання чи інші клінічні дії.

Одним із традиційних підходів до оцінювання довжини об'єкта в тривимірному просторі є використання поняття діаметра множини точок, що

$$d = \max_{p,q} \left\{ \sqrt{(x_{1p} - x_{1q})^2 + (x_{2p} - x_{2q})^2 + (x_{3p} - x_{3q})^2} \right\}. \quad (3)$$

Такий підхід є достатньо надійним в оцінюванні великих фрагментів із близькою до еліпсоїдної формою. У подібних випадках крайні точки об'єкта лежать уздовж його основної осі, а відстань між ними добре наближає дійсну довжину.

Однак зі зменшенням розміру фрагмента або за наявності артефактів, зокрема тінювих зон на КТ-зображеннях, ефективність цієї моделі знижується [14]. Якщо форма фрагмента близька до паралелепіпеда або має складну форму, визначення довжини лише за допомогою діаметра призводить до значних похибок. Також варто брати до уваги, що зона тіні, яка оточує металевий об'єкт на знімках, може істотно впливати на результат, створюючи хибні екстремальні точки.

У такому разі необхідне застосування більш гнучких і точних методів, які давали б змогу зважати на орієнтацію та реальну просторову структуру об'єкта. Одним з найбільш ефективних підходів у цьому контексті є метод основних компонент (*Principal Component Analysis*, PCA) [15, 16]. Він забезпечує об'єктивну побудову нової системи координат, узгодженої з формою об'єкта. PCA дає змогу орієнтувати осі системи координат відповідно до напрямків найбільшого поширення об'єкта, що є ключовим для точного визначення його габаритів.

Отже, метод основних компонент є універсальним інструментом для визначення просторових розмірів тіл довільної форми. Його застосування допомагає мінімізувати вплив асиметрії об'єкта, зменшити похибки, зумовлені артефактами візуалізації, та забезпечити високу відтворюваність результатів. Надалі розглянемо математичні особливості використання PCA в тривимірному просторі R^3 для оцінювання орієнтації та лінійних розмірів фрагментів.

Упровадження методу основних компонент

Метод основних компонент (PCA) є ключовим етапом в алгоритмі оцінювання розмірів металевих фрагментів, оскільки дає змогу визначити орієнтацію

відтворюють об'єкт. У цьому разі довжина d визначається як максимальна евклідова відстань між будь-якими двома точками об'єкта (x_{1p}, x_{2p}, x_{3p}) та (x_{1q}, x_{2q}, x_{3q}) :

об'єкта в просторі незалежно від його положення у вихідній системі координат. Це особливо важливо в разі складної або асиметричної форми фрагментів, де стандартні евклідові методи можуть давати значні похибки.

Сутність PCA полягає в побудові нової ортогональної системи координат, осі якої відповідають напрямкам найбільшої дисперсії набору точок. У межах цієї системи координат об'єкт повертається таким чином, що його основні розміри відповідають проєкціям на відповідні осі.

Знаходження компонент перетворення до нових основних осей зводиться до обчислення власних значень $\lambda \in R$ та відповідних їм власних векторів $v = (v_1, v_2, v_3)^T \in R^3$ коваріаційної матриці C координат (1) і (2). Для коректності подальших перетворень дані попередньо центровано. Це забезпечує правильне визначення початку координат у тривимірному просторі.

У разі тривимірного простору задано множину з n точок $x_i = (x_{1i}, x_{2i}, x_{3i})$ $i = 1, \dots, n$, що описують форму об'єкта.

Спочатку обчислюються середні значення координат:

$$M_1 = \frac{1}{n} \sum_{i=1}^n x_{1i}, M_2 = \frac{1}{n} \sum_{i=1}^n x_{2i}, M_3 = \frac{1}{n} \sum_{i=1}^n x_{3i}. \quad (4)$$

Після цього виконується центрування даних способом перерахунку координат точок:

$$x_i = (x_{1i} - M_1, x_{2i} - M_2, x_{3i} - M_3), i = 1, \dots, n. \quad (5)$$

Допустимо, що E – це одинична матриця вигляду

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (6)$$

Тоді власні значення λ і власні вектори v коваріаційної матриці C знаходяться способом розв'язання такого рівняння:

$$(C - \lambda \cdot E) \cdot v = 0. \quad (7)$$

Відомо, що це рівняння є нічим іншим, як системою лінійних рівнянь, нетривіальні рішення якої є шуканими власними векторами v матриці C . У процесі розв'язання цієї задачі цікаво знайти нетривіальні рішення. Множина таких рішень

для (7) – це власний підпростір матриці C . Він належить власному значенню λ . Для тривимірного випадку власні вектори v , які відповідають власним значенням λ , знаходяться за допомогою розв'язання системи рівнянь:

$$\left(\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} - \lambda_j \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} v_{j1} \\ v_{j2} \\ v_{j3} \end{bmatrix} = 0. \quad (8)$$

Тут c_{ij} – елементи коваріаційної матриці C . Вони обчислюються наперед і потім використовуються як коефіцієнти.

Отримана після розв'язання рівняння (8) система векторів визначає нову базу у вигляді $v = (v_1, v_2, v_3)$. Припускається, що вона нормована. Крім того, ймовірно, що всі ці вектори розташовані в порядку спадання відповідних їм власних значень.

У такій ситуації матриця перетворення координат до основних компонент має такий вигляд:

$$A = v^T = \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix}^T. \quad (9)$$

Отже, припустимо, що $x_1 = (x_{11}, x_{12}, x_{13})$ і $x'_1 = (x'_{11}, x'_{12}, x'_{13})$ – вектори, які визначають координати точок у старій і новій системах координат відповідно. У такому разі перехід до нових координат виконується так:

$$x' = A^{-1}x. \quad (10)$$

Після перетворення координат у новий базис форма множини й відстані між точками залишаються незмінними – зберігаються всі метричні властивості. Однак координати точок набувають оптимальної орієнтації, що спрощує та пришвидшує визначення лінійних розмірів складних тривимірних об'єктів, як-от фрагменти. Здебільшого об'єкти можна швидко вирівняти вздовж однієї осі, що дає змогу застосовувати метод основних компонент у двовимірному вигляді. Це особливо корисно для інтерактивних і напівавтоматичних систем просторового вирівнювання.

У такій ситуації (після центрування, відповідно до (4) і (5)), коваріаційна матриця в просторі R^2 матиме вигляд

$$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}. \quad (11)$$

Елементи цієї матриці обчислюються так:

$$c_{11} = \frac{1}{n} \sum_{i=1}^n (x_{1i} - M_1)^2, \quad c_{22} = \frac{1}{n} \sum_{i=1}^n (x_{2i} - M_2)^2, \quad (12)$$

$$c_{12} = c_{21} = \frac{1}{n} \sum_{i=1}^n (x_{1i} - M_1)(x_{2i} - M_2). \quad (13)$$

Тут вирази (12) визначають дисперсії ознак (координат точок), а вирази (13) – взаємні коваріації ознак (координат точок).

Завдяки тому, що дані були попередньо центровані, їх середні значення тепер дорівнюють нулю, і їх можна вилучити з виразів (12) і (13). У цьому разі ці вирази можна переписати таким чином:

$$c_{11} = \frac{1}{n} \sum_{i=1}^n (x_{1i})^2, \quad c_{22} = \frac{1}{n} \sum_{i=1}^n (x_{2i})^2, \quad (14)$$

$$c_{12} = c_{21} = \frac{1}{n} \sum_{i=1}^n (x_{1i})(x_{2i}). \quad (15)$$

Тепер ці значення можна швидко обчислити й використовувати далі у вигляді числових коефіцієнтів. Припустимо, що це вже зроблено.

Рівняння (8) для двовимірного випадку матиме вигляд

$$\left(\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0. \quad (16)$$

Після виконання тривіальних перетворень цей вираз можна записати явно у вигляді системи лінійних рівнянь

$$\begin{cases} (c_{11} - \lambda)v_1 + c_{12}v_2 = 0, \\ c_{21}v_1 + (c_{22} - \lambda)v_2 = 0. \end{cases} \quad (17)$$

Ця система лінійних рівнянь досить просто розв'язується, що визначає так зване характеристичне рівняння. Унаслідок його розв'язання можна знайти два власних значення. Після відповідних перетворень (підстановки й обчислення дискримінанта) знайдемо два корені характеристичного рівняння у такому вигляді:

$$\begin{cases} \lambda_1 = 0.5[(c_{11} + c_{22}) + Q], \\ \lambda_2 = 0.5[(c_{11} + c_{22}) - Q], \end{cases} \quad (18)$$

$$Q = (c_{11} - c_{22})^2 + 4c_{12}^2. \quad (19)$$

Для побудови нової ортогональної та нормованої системи використовуватимемо вектори одиничної довжини. У цьому разі зазвичай припускається, що перший вектор відповідає більшому власному значенню, а другий – меншому власному значенню.

З практичного погляду перетворення координат точок у нову систему координат відповідає простому обертанню у вигляді

$$\begin{cases} x'_1 = \cos(\theta)x_1 + \sin(\theta)x_2, \\ x'_2 = -\sin(\theta)x_1 + \cos(\theta)x_2. \end{cases} \quad (20)$$

Тут

$$\begin{cases} \cos(\theta) = v_{11}, \\ \sin(\theta) = v_{12}. \end{cases} \quad (21)$$

Значення v_{11} і v_{12} , які визначаються за формулою (18) та умовою нормування, обчислюються таким чином:

$$\begin{cases} v_{11} = \sqrt{\frac{(c_{12})^2}{(c_{12})^2 + (\lambda_1 - c_{11})^2}}, \\ v_{12} = \sqrt{\frac{(\lambda_1 - c_{11})^2}{(c_{12})^2 + (\lambda_1 - c_{11})^2}}. \end{cases} \quad (22)$$

Щодо реалізації обчислень доцільно знайти значення v_{11} і v_{12} один раз і потім використовувати їх у ітераційному перетворенні координат точок.

Обчислення лінійних розмірів металевого фрагмента

Після центрування координатного простору й застосування методу основних компонент (МГК) кожна точка набору даних подана в новій системі координат, осі якої орієнтовані відповідно до напрямків найбільшої дисперсії точок. Такий підхід допомагає значно спростити визначення габаритів об'єкта, зокрема металевого фрагмента.

У новій системі координат обчислення лінійних розмірів зводиться до визначення відстані між крайніми точками (мінімуми й максимуми) вздовж кожної з трьох основних осей. Для кожної осі $i \in \{1, 2, 3\}$ обчислюється різниця між найбільшим і найменшим значенням координати:

$$\{L^{(i)} = X_{max}^{(i)} - X_{min}^{(i)}\}, i = 1, 2, 3. \quad (23)$$

Максимальне значення – це довжина фрагмента, менше – його ширина, а найменше – висота.

Такий підхід дає змогу однозначно визначити габарити об'єкта за умов, коли дані є очищеними від шуму, а вокселі точно відповідають металевому фрагменту. У цьому разі модель оцінки є максимально простою, прозорою для реалізації та дає високу точність навіть за мінімальних обчислювальних витрат.

Проте в практичних умовах реконструкції зображень з КТ-даних об'єкти нерідко супроводжуються ефектом тіні, який утворюється внаслідок взаємодії металевих структур із рентгенівським випромінюванням. Така тінь призводить до розмиття контурів і долучення до множини точок об'єкта вокселів, що не належать фрагменту, але мають проміжну щільність. Це спотворює реальні розміри об'єкта в новій системі координат і знижує точність оцінки.

У зв'язку з цим модель лінійного вимірювання, подана в (23), потребує вдосконалення способом врахування зон тіні та побудови коригувальних прогнозних алгоритмів, які дають змогу відфільтрувати невірні точки на основі їх просторового розташування, щільності й топології, що і є в планах авторів.

Програмна реалізація

У межах цього дослідження створено спеціалізоване програмне забезпечення для автоматизованого оброблення КТ-даних і оцінювання лінійних розмірів металевих фрагментів.

Запропонований метод автоматизованого високоточного вимірювання лінійних розмірів реалізовано у вигляді програмного комплексу, що забезпечує оброблення медичних КТ-зображень і обчислення просторових властивостей об'єкта. Програмна реалізація виконується із застосуванням мови C# для основної логіки взаємодії з даними, інтерфейсу користувача та управління процесом, а також Python – для математичного оброблення даних, зокрема реалізації методу основних компонент (PCA).

Архітектура програмного забезпечення модульна. До основних компонентів належать:

- модуль імпорту й декодування DICOM-даних;
- модуль тривимірного відтворення воксельної матриці щільності;
- модуль автоматичної сегментації металевих тіл із зображень;
- модуль обчислень розмірів та орієнтації;
- блок виведення результатів і підготовки звітів.

Усі дані обробляються автоматично. Після завантаження DICOM-файлу система самостійно виконує сегментацію, центровано вирівнює дані, формує коваріаційну матрицю, обчислює основні компоненти, визначає реальні координати вокселів і обчислює лінійні розміри об'єкта. Результати надаються користувачу у вигляді числових значень, графічних візуалізацій і збережених файлів.

Програмне забезпечення гарантує високу надійність, масштабованість і придатність до інтеграції в більші інформаційно-аналітичні рішення, що застосовуються у військовій медицині та судово-медичній експертизі.

Візуалізація просторових структур металевих фрагментів є невід'ємним складником запропонованої системи. Вона дає змогу не лише підтвердити правильність виконаної сегментації та обчислень, а й забезпечити інтуїтивно зрозуміле сприйняття результатів медичними фахівцями. Для реалізації візуального подання об'єктів використано платформу Unity, яка забезпечує тривимірну інтерактивну візуалізацію воксельних структур і уможливує

масштабування, обертання й аналіз моделі в реальному часі (рис. 1).

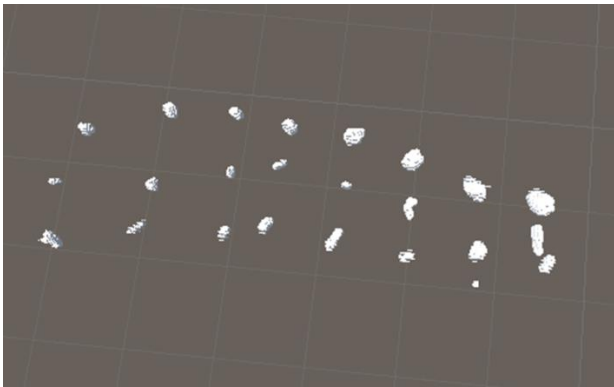


Рис. 1. Загальний план результату експерименту № 2 (печінка): усі 24 металевих фрагменти в органі успішно знайдені під час сегментації

Сцена в середовищі Unity будується на основі отриманої після сегментації воксельної моделі, до якої застосовано попереднє оброблення: вилучення тінювих зон, виділення ядра фрагмента й побудова його оболонки. Колірне кодування дає змогу розрізнити щільні частини фрагмента, його тінь і фон. Також реалізовано можливість увімкнення й вимкнення окремих шарів візуалізації (наприклад, межі, тінь, координатна сітка).

Ці можливості є важливими для верифікації алгоритмів, інтерактивного аналізу об'єктів і презентації результатів у наукових або клінічних звітах. У разі потреби Unity-сцену можна експортувати у форматі зображень, відео або

3D-моделей (наприклад, *.obj, *.fbx), що сприяє інтеграції з іншими системами візуалізації або медичного планування.

Хоча за основу взято Unity, розроблена система не обмежена конкретною платформою: всі елементи візуалізації можуть бути адаптовані до будь-якого іншого програмного середовища, яке підтримує тривимірну графіку й оброблення воксельних даних. Завдяки цьому система є гнучкою та легко вбудовується в наявні робочі процеси медичних установ або дослідницьких лабораторій.

Результати експериментального дослідження

Для перевірки точності й ефективності запропонованої моделі та методу було проведено серію експериментів із використанням реальних медичних зображень комп'ютерної томографії. Експерименти здійснювалися на зразках біологічної тканини (легені, печінка, кишківник свині), в які були імплантовані металеві фрагменти завдовжки від 5 до 15 мм. Як фрагменти застосовувалися чавунні й залізні уламки, а також нарізані цвяхи та дрiт

Загалом було досліджено 72 фрагменти, розташовані в різних просторових положеннях і під різними кутами до площини сканування (рис. 2). Томографічні зображення зберігалися у форматі DICOM і оброблялися автоматично за допомогою запропонованого програмного комплексу. Усі етапи – від імпорту до отримання результатів – виконувалися без участі оператора.

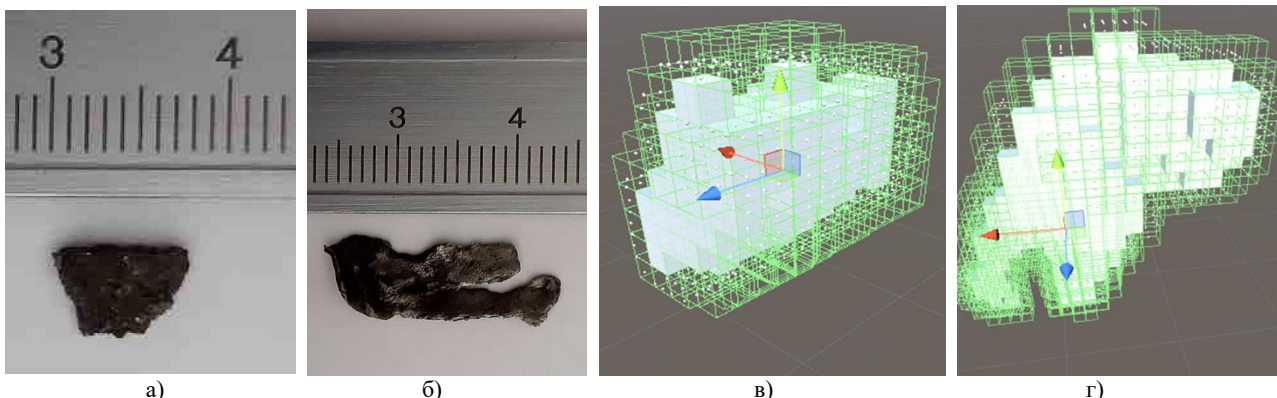


Рис. 2. Чавунний (а) і залізний (б) фрагменти, які використовувались під час експерименту; візуалізація у Unity (в, г)

Для перевірки точності результати автоматичного оцінювання порівнювалися з еталонними розмірами фрагментів, які були заздалегідь виміряні експертами вручну перед імплантацією в біологічні тканини.

Основним показником точності була відносна похибка оцінювання.

У табл. 1 наведено узагальнені результати експериментальних досліджень. Для кожного зразка

матеріалу надано середнє значення максимальної щільності об'єктів за шкалою Хаунсфілда в різних органах (легені, печінка, кишківник), а також обчислено відносну похибку вимірювань. Досягнуті результати дають змогу оцінити розбіжності

в обчисленні відносної похибки залежно від анатомічної ділянки й типу матеріалу, а також підтверджують стабільність розробленого методу оброблення зображень.

Таблиця 1. Узагальнені експериментальні результати похибки вимірювання в різних органах

	Легені		Печінка		Кишківник	
	щільність	похибка	щільність	похибка	щільність	похибка
Чавун	8282	0,026	9296	0,034	8337	0,023
Залізо	19579	0,018	19000	0,027	18655	0,026
Цвях	16582	0,055	18051	0,056	20120	0,056
Алюміній	2044	0,027	2063	0,054	2159	0,057
Мідь	23620	0,060	26597	0,059	25262	0,078

Аналіз результатів продемонстрував, що у понад 90% випадків похибка не перевищує 5%, а іноді досягається точність 98–99% відповідно до еталонних значень. Середнє відхилення за вибіркою становило менше ніж 3%, що свідчить про високу надійність і відтворюваність методу.

Також було перевірено вплив різних порогових значень T_{seg} та T_{ob} на результати вимірювання. Дослідження підтвердило, що автоматичне адаптивне визначення цих параметрів забезпечує стабільність методу внаслідок зміни умов сканування й типу тканини.

Досягнуті результати демонструють високу точність, надійність і практичну придатність розробленої системи для використання в медичній практиці, зокрема в умовах військової медицини, де потрібне оперативне й надійне оцінювання властивостей чужорідних тіл.

Обговорення результатів

Результати експериментального дослідження продемонстрували суттєве підвищення точності вимірювання лінійних розмірів тривимірних об'єктів за даними КТ. Відомо, що в разі застосування традиційних підходів без спеціальних алгоритмів просторового вирівнювання й адаптивної сегментації похибки, зумовлені металевими артефактами й довільною орієнтацією об'єкта, можуть досягати десятків відсотків. Застосування запропонованого методу дало змогу зменшити похибку вимірювання до рівня меншого ніж 5%, що підтверджує ефективність поєднання просторового вирівнювання на основі аналізу основних компонент і двоступеневої адаптивної порогової сегментації.

Крім цього, встановлено, що на точність вимірювання можуть впливати металеві артефакти й ефект тіні, що визначає доцільність подальшого вдосконалення алгоритмів фільтрації граничних вокселів. Незважаючи на це, досягнуті результати підтверджують практичну ефективність розробленого підходу для автоматизованого аналізу КТ-даних.

Висновки

Дослідження дало змогу розробити й програмно реалізувати метод автоматизованого оцінювання лінійних розмірів металевих фрагментів за показниками комп'ютерної томографії. Запропонований підхід поєднує автоматичну сегментацію, орієнтацію об'єкта за методом основних компонент і визначення його просторових властивостей у новій системі координат.

Розроблене програмне забезпечення пройшло апробацію на наборі з 72 зразків, що містять металеві фрагменти в біологічних тканинах, і продемонструвало високу точність вимірювання. Здебільшого похибка не перевищувала 5%, що свідчить про надійність розробленого методу в умовах практичного застосування. Важливо також, що дані обробляються автоматично. Це значно підвищує оперативність прийняття медичних рішень у складних або критичних ситуаціях.

Візуалізаційний складник рішення дає змогу не лише підтверджувати правильність геометричної оцінки, але й подати результати у формі, зручній для подальшого клінічного або експертного аналізу. Завдяки використанню сучасного програмного середовища візуалізації система є адаптованою до потреб як практичної медицини, так і наукових досліджень.

Подальший розвиток проекту передбачає поглиблення алгоритмічної частини, зокрема розширення підходів до аналізу об'єктів у складних середовищах, а також розроблення механізмів автоматичного налаштування параметрів оброблення. Одним із важливих напрямів також є інтеграція результатів у повноцінні медичні інформаційно-аналітичні системи для підвищення ефективності клінічної діагностики.

Отже, запропонована модель і алгоритм демонструють високу практичну цінність і мають значний потенціал для подальшого застосування в умовах медицини катастроф, військової та судової медицини.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів, зокрема фінансового, особистого, авторського чи будь-якого іншого характеру, який міг би вплинути на дослідження, а також на результати, опубліковані в цій статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

References

- Lee, C., Lee, A., Choi, Y.J. et al. (2021), "Accuracy of digital model generated from CT data with metal artifact reduction algorithm", *Scientific Reports*, Vol. 11, 10332. DOI: <https://doi.org/10.1038/s41598-021-89298-x>
- Nagayama, Y., Tanoue, S., Oda, S. et al. (2020), "Metal Artifact Reduction in Head CT Performed for Patients with Deep Brain Stimulation Devices: Effectiveness of a Single-Energy Metal Artifact Reduction Algorithm", *AJNR American Journal of Neuroradiology*, Vol. 41, No. 2, pp. 231–237. DOI: <https://doi.org/10.3174/ajnr.A6375>
- Pereira, H., Romero, L., Miguel Faria, P. (2025), "Web-Based DICOM Viewers: A Survey and a Performance Classification", *Journal of Imaging Informatics in Medicine*, Vol. 38, No. 3, pp. 1304–1322. DOI: <https://doi.org/10.1007/s10278-024-01216-5>
- Khoroshun, E., Smelyakov, K., Chupryna, A., Makarov, V., Nehoduiko, V., Vakulik, Y. (2024), "Improving of computed tomography images for effective diagnosis of gunshot wounds", *Emergency Medicine*, Vol. 20, No. 7, pp. 577–583. DOI: <https://doi.org/10.22141/2224-0586.20.7.2024.1776>
- Bunin, Y., Vakulik, Y., Mykhailusov, R., Nehoduiko, V., Smelyakov, K., Yasynskyi, O. (2020), "Assessment of foreign body size in the lung using computer vision algorithms", *Experimental and Clinical Medicine*, Vol. 89, No. 4, pp. 87–94. DOI: <https://doi.org/10.35339/ekm.2020.89.04.13>
- Arabi, H., Zaidi, H. (2021), "Deep learning-based metal artefact reduction in PET/CT imaging", *European Radiology*, Vol. 31, No. 8, pp. 6384–6396. DOI: <https://doi.org/10.1007/s00330-021-07709-z>
- Selles, M., van Osch, J.A.C., Maas, M., Boomsma, M.F., Wellenberg, R.H.H. (2024), "Advances in metal artifact reduction in CT images: A review of traditional and novel metal artifact reduction techniques", *European Journal of Radiology*, Vol. 170, 111276. DOI: <https://doi.org/10.1016/j.ejrad.2023.111276>
- Kleber, C.E.J., Karius, R., Naessens, L.E., van Toledo, C.O., van Osch, J.A.C., Boomsma, M.F., Heemskerk, J.W.T., van der Molen, A.J. (2024), "Advancements in supervised deep learning for metal artifact reduction in computed tomography: A systematic review", *European Journal of Radiology*, Vol. 181, 111732. DOI: <https://doi.org/10.1016/j.ejrad.2024.111732>
- Wang, H., Li, Y., Zhang, H., Chen, J., Ma, K., Meng, D., Zheng, Y. (2021), "InDuDoNet: An Interpretable Dual Domain Network for CT Metal Artifact Reduction", *arXiv*, available at: <https://doi.org/10.48550/arXiv.2109.05298>

Доступність даних

Дані, використані у цьому дослідженні, будуть надані авторами за обґрунтованим запитом.

Використання засобів штучного інтелекту

Під час підготовки рукопису автори застосовували інструменти штучного інтелекту тільки для допоміжної перевірки правопису й граматики тексту. Зокрема використано мовні інструменти Grammarly й мовну модель GPT-5.3 (OpenAI) для технічної перевірки окремих фрагментів тексту.

Інструменти штучного інтелекту не застосовувалися для наукового аналізу, оброблення експериментальних показників, побудови математичних моделей або формування результатів дослідження.

Усі наукові положення, методи, результати експериментів і висновки сформульовано самостійно. Автори перевірили коректність тексту після застосування інструментів штучного інтелекту й несуть повну відповідальність за зміст статті. Використання інструментів штучного інтелекту не вплинуло на результати дослідження та його висновки.

10. Yu, L., Zhang, Z., Li, X., Ren, H., Zhao, W., Xing, L. (2021), "Metal artifact reduction in 2D CT images with self-supervised cross-domain learning", *Physics in Medicine & Biology*, Vol. 66, No. 17, 175003. DOI: <https://doi.org/10.1088/1361-6560/ac195c>
11. Zou, H., Wang, Z., Guo, M., Peng, K., Zhou, J., Zhou, L., Fan, B. (2025), "Metal artifact reduction combined with deep learning image reconstruction algorithm for CT image quality optimization: a phantom study", *PeerJ*, Vol. 13, e19516. DOI: <https://doi.org/10.7717/peerj.19516>
12. Larobina, M. (2023), "Thirty Years of the DICOM Standard", *Tomography*, Vol. 9, No. 5, pp. 1829–1838. DOI: <https://doi.org/10.3390/tomography9050145>
13. DenOtter, T.D., Schubert, J. (2023), "Hounsfield Unit", in: *StatPearls*, StatPearls Publishing, Treasure Island, available at: <https://www.ncbi.nlm.nih.gov/books/NBK547721/>
14. Hashimov, E., Pashayev, A., Khaligov, G. (2025), "Camera control algorithm and image quality assessment method to obtain a quality image", *Advanced Information Systems*, Vol. 9, No. 3, pp. 50–56. DOI: <https://doi.org/10.20998/2522-9052.2025.3.06>
15. Greenacre, M., Groenen, P.J.F., Hastie, T., D'Enza, A.I., Markos, A., Tuzhilina, E. (2022), "Principal component analysis", *Nature Reviews Methods Primers*, Vol. 2, No. 1. DOI: <https://doi.org/10.1038/s43586-022-00184-w>
16. Jolliffe, I.T., Cadima, J. (2016), "Principal component analysis: a review and recent developments", *Philosophical Transactions of the Royal Society A*, Vol. 374, No. 2065, 20150202. DOI: <https://doi.org/10.1098/rsta.2015.0202>

Received (Надійшла) 08.09.2025

Accepted for publication (Прийнята до друку) 10.02.2026

Publication date (Дата публікації) 30.03.2026

Відомості про авторів / About the Authors

Вакулік Євген Вікторович – аспірант, Харківський національний університет радіоелектроніки, кафедра програмної інженерії, Харків, Україна;

Eugen Vakulik – PhD Student, Kharkiv National University of Radio Electronics, Department of Software Engineering, Kharkiv, Ukraine;

e-mail: yevhen.vakulik@nure.ua

ORCID ID: <http://orcid.org/0000-0002-4940-0529>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57816562200>

Смеляков Кирило Сергійович – доктор технічних наук, професор, Харківський національний університет радіоелектроніки, професор кафедри програмної інженерії, Харків, Україна;

Kyrylo Smelyakov – Doctor of Sciences (Engineering), Professor, Kharkiv National University of Radio Electronics, Professor of the Department of Software Engineering, Kharkiv, Ukraine;

e-mail: kyrylo.smelyakov@nure.ua

ORCID ID: <http://orcid.org/0000-0001-9938-5489>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57203149663>

Чуприна Анастасія Сергіївна – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри програмної інженерії, Харків, Україна;

Anastasiya Chupryna – PhD, Associate Professor, Kharkiv National University of Radio Electronics, Associate Professor of the Department of Software Engineering, Kharkiv, Ukraine;

e-mail: anastasiya.chupryna@nure.ua

ORCID ID: <http://orcid.org/0000-0003-0394-9900>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57202997528>

MODEL AND METHOD OF AUTOMATED HIGH-PRECISION MEASUREMENT OF THREE-DIMENSIONAL OBJECTS BASED ON COMPUTED TOMOGRAPHY DATA

The subject of study is the linear dimensions of three-dimensional objects based on computed tomography data, specifically metallic foreign bodies in human tissues and organs. **The purpose of the study** is to develop a mathematical model and a method for automated, high-precision measurement of the linear dimensions of three-dimensional objects based on computed tomography results, with implementation in the form of software. **Objectives:** to formulate a mathematical model for representing a three-dimensional

object in voxel space; to develop a method for segmenting metal fragments based on computed tomography results; to propose a method for spatial alignment of segmented objects based on principal component analysis; develop a method for determining the maximum linear dimension of an object in a new coordinate system; implement the proposed model and method as software and experimentally verify the measurement accuracy. **Research methods:** analysis of tomographic parameters, threshold segmentation with adaptive selection of threshold values, wave-based algorithm for finding connected components, principal component analysis to determine the object's orientation, voxel modeling, and calculation of Euclidean distances between boundary points of a three-dimensional object. **Results.** A mathematical model for representing a three-dimensional object and a method for automated high-precision determination of its maximum linear dimension are proposed. The method was implemented as a software module and tested on 72 samples of metal fragments of six types embedded in the biological tissues of pig organs. The average deviation does not exceed 3%, and in the most complex cases remains within 5%, which indicates the high accuracy and stability of the proposed approach. **Conclusions:** The developed model and method ensure automated and objective determination of the linear dimensions of foreign bodies based on computed tomography data without operator intervention. The proposed software can be used in military medicine, forensic medical examination, disaster medicine, and healthcare facilities where the speed and reliability of diagnostic decisions are critically important.

Keywords: computed tomography, segmentation, principal component analysis, linear dimensions, DICOM, voxel.

Бібліографічні описи / Bibliographic descriptions

Вакулік Є. В., Смеляков К. С., Чуприна А. С. Модель і метод автоматизованого високоточного вимірювання тривимірних об'єктів за даними комп'ютерної томографії. *Сучасний стан наукових досліджень та технологій в промисловості*. 2026. № 1 (35). С. 6–16. DOI: <https://doi.org/10.30837/2522-9818.2026.1.006>

Vakulik, E., Smelyakov, K., Chupryna, A. (2026), "Model and method of automated high-precision measurement of three-dimensional objects based on computed tomography data", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (35), P. 6–16. DOI: <https://doi.org/10.30837/2522-9818.2026.1.006>

Kristina Dostalova, Alexandra Cizmarova, Marek Klimo, Roman Yaroshevych

EXPLANATION DETECTED BRAIN TUMOURS IN MRI IMAGES USING YOLOV8 WITH LIME-BASED INTERPRETATION

Relevance. Precise identification of brain tumours in magnetic resonance imaging (MRI) is a critical task in medical image analysis. Although deep learning-based object detectors achieve high localisation accuracy, their limited transparency restricts trust and routine adoption in clinical practice, highlighting the need for explainable artificial intelligence (XAI) approaches.

Object of research. The object of this research is the automated detection of brain tumours in MRI scans using convolutional neural network – based object detection models. **Subject of research.** The subject of the research is the integration of YOLOv8 object detection models with the Local Interpretable Model-Agnostic Explanations (LIME) method to interpret individual detection outputs in medical imaging.

Purpose. The aim of this paper is to develop and evaluate an explainable framework for brain tumour detection in MRI images by integrating YOLOv8-based object detection with LIME-based interpretation and by quantitatively assessing the quality of the generated explanations. **Methods.** Two YOLOv8 variants (YOLOv8n and YOLOv8s) were trained and evaluated on a publicly available MRI dataset containing glioma, meningioma, and pituitary tumour classes. LIME was applied to generate superpixel-based, box-conditioned local explanations for individual detections. Detection performance was assessed using precision, recall, mAP@50, and mAP@50–95. Explanation quality was quantitatively evaluated using stability, sparsity, maximum superpixel weight, and entropy metrics. **Results.** Experimental results demonstrate that both YOLOv8 models achieve high detection performance, with YOLOv8s providing slightly improved accuracy. LIME successfully highlights image regions that most influence model decisions, and the proposed quantitative metrics confirm that the generated explanations are stable, informative, and aligned with clinically relevant tumour regions. **Conclusions.** The proposed framework provides a practical approach for combining accurate tumour localisation with interpretable and quantitatively validated explanations, supporting reliability-oriented evaluation of AI-based clinical decision-support systems.

Keywords: Explainable Artificial Intelligence (XAI); LIME; medical image analysis, brain tumour detection; MRI; YOLOv8.

Introduction

Relevance and an overview of scientific works

The use of artificial intelligence (AI) and machine learning has become increasingly widespread in modern healthcare [1]. It aligns with trends in precision and smart medicine developed under Industry 4.0 [2, 3]. This approach has the potential to reduce human error in diagnosis and treatment selection [4, 5]. This progress is particularly evident in medical image analysis, including ultrasound, X-ray, and magnetic resonance imaging (MRI) [6]. One of the key challenges in medical image processing is the accurate and timely identification of brain tumours in magnetic resonance imaging (MRI) [7]. While deep learning has significantly advanced localisation and classification performance in medical images, its limited transparency restricts trust and routine adoption in clinical practice [8, 9, 10]. Object detection models such as the YOLO family offer fast, end-to-end tumour localisation and class prediction, yet their decision paths are difficult to interrogate and validate by clinicians [11]. This motivates explainable artificial intelligence (XAI) approaches that expose which image regions and visual cues drive algorithmic decisions,

thereby enabling critical scrutiny, error analysis, and safer integration into clinical workflows [10, 12, 13].

In neuro-oncology, the tension between performance and transparency is particularly acute. MRI appearances of glioma, meningioma, and pituitary tumours can overlap, and downstream clinical actions (biopsy planning, therapy selection, and follow-up strategies) depend on how much confidence care teams place in algorithmic suggestions [9]. Beyond general “interpretability”, healthcare decision support systems must satisfy broader reliability requirements: stable behaviour under minor perturbations, robustness to distributional shifts, traceability of contributing factors, and explicit treatment of uncertainty. These requirements have been widely recognised in the reliability engineering community, including for healthcare applications where human factors and organisational contexts shape outcomes, and where aleatory and epistemic uncertainties must be disentangled and managed [4, 5, 14].

Existing XAI research in medical imaging has largely focused on classification or segmentation networks (e.g., U-Net variants for BRATS-style tasks), with explanation techniques such as Grad CAM/Grad CAM++ or relevance propagation providing saliency

maps over the full image [1,3]. However, object detection, where the model first proposes bounding boxes and then assigns class confidences, poses distinct explanation challenges. Post hoc methods must attribute predictions to a specific detection rather than the whole image; they should reflect both localisation and classification components and be compatible with non-differentiable post-processing steps used by modern detectors (e.g., NMS). As a result, despite encouraging reports of YOLO-based medical detectors, high-quality, box-conditioned, and quantitatively assessed explanations for detection outputs remain underexplored [10, 12, 15].

To address these limitations, the present study integrates a modern one-stage detector, YOLOv8, with the LIME framework to obtain local, model-agnostic, superpixel-based explanations that are explicitly tied to individual detections (bounding boxes) in brain MRI. We evaluate two YOLOv8 variants (YOLOv8n and YOLOv8s) to study the trade-off between computational efficiency and both accuracy and interpretability, following a consistent training and evaluation protocol on a public MRI dataset comprising glioma, meningioma, and pituitary tumours. We further complement qualitative overlays with a set of quantitative explanation metrics: stability, sparsity, maximum superpixel weight, and entropy. These metrics are designed to assess how concentrated, consistent, and interpretable the explanations are in practice. This combined perspective aims to bridge performance metrics (precision, recall, mAP@50, mAP@50–95) with reliability-oriented interpretability indicators meaningful for clinical decision-making.

From a reliability engineering perspective, quantifying explanation behaviour enhances the robustness of AI systems in healthcare. Stable, sparse, and clinically aligned explanations facilitate expert verification, support effective handover between algorithm and clinician, and mitigate risks associated with spurious correlations [5, 16].

Setting objectives

The aim of this paper is to develop and evaluate an explainable framework for brain tumour detection in MRI images by integrating YOLOv8-based object detection with LIME-based interpretation and by quantitatively assessing the quality of the generated explanations. The proposed solution integrates YOLOv8-based object detection models with the LIME explainable artificial intelligence technique to improve the transparency and reliability of model predictions.

This paper makes the following contributions:

- Conditioned, model-agnostic explanations for MRI tumour detection. LIME is adapted to generate box-conditioned superpixel explanations for YOLOv8 detections in brain MRI, enabling clinicians to verify why a specific bounding box and tumour class were proposed rather than only viewing global image saliency.
- Quantitative evaluation of explanation quality. A compact set of interpretability metrics (Stability, Sparsity, MaxWeight, and Entropy) is introduced and computed to assess robustness, concentration, and distribution of importance across superpixels. These metrics operationalise reliability-oriented interpretability by linking explanation behaviour to desirable properties such as robustness and focus.
- Empirical study across model capacities. YOLOv8n and YOLOv8s are compared under identical training conditions and dataset splits to analyse how model capacity affects detection performance (precision, recall, mAP@50, mAP@50–95) and the structure and stability of generated explanations.
- Reliability-aware perspective for clinical AI. Interpretability is situated within a broader reliability engineering framework for healthcare, providing a principled rationale for adopting quantitative XAI criteria in deployment-oriented evaluations.

The methodology

Dataset

The experimental evaluation was performed using a publicly available brain tumour MRI dataset designed for automated tumour detection tasks. The dataset contains magnetic resonance images annotated with the position of pathological areas and each image is labelled according to the tumour type, including glioma, meningioma, and pituitary tumours [17].

The dataset was selected due to its compatibility with object detection frameworks and its suitability for training YOLO-based models [18]. All images were resized to a uniform resolution to ensure consistency during training and inference. The dataset was divided into training, validation, and test subsets, enabling objective evaluation of detection performance and explainability results.

This dataset provides a representative collection of MRI scans with varying tumour sizes, locations, and visual characteristics, which enables robust evaluation of both detection accuracy and model interpretability.

YOLOv8-based brain tumour detection

Brain tumour detection was performed using the YOLOv8 [19] object detection framework, which is designed to localize objects within images by predicting bounding boxes in a single forward pass. YOLOv8 was selected for its balance between detection accuracy, computational efficiency, and flexibility for training on custom datasets.

In this study, two YOLOv8 model variants were utilised: YOLOv8n and YOLOv8s. The YOLOv8n variant is a lightweight model optimised for fast inference and reduced computational requirements, whereas YOLOv8s provides higher detection accuracy at the cost of increased model complexity. Comparing these two variants enables analysis of the trade-off between efficiency and performance in medical image applications.

Both models were trained under identical conditions to ensure a fair comparison. Training was conducted on resized MRI images with a fixed input resolution, and standard data augmentation techniques were applied to enhance model generalisation. The models were optimised using stochastic gradient-based learning and evaluated using established object detection metrics.

The detection outputs consist of bounding boxes with associated confidence scores and tumour class predictions, which serve as inputs for subsequent explainability analysis using the LIME method.

Explainable Artificial Intelligence (XAI) [12]

In recent years, a wide range of explainability methods has been proposed to address the limited transparency of deep learning models in medical imaging. Gradient-based approaches, such as Grad-CAM and Grad-CAM++, are among the most frequently used techniques. These methods visualize important regions by analyzing gradients in convolutional layers, which makes them relatively fast and easy to apply. However, their explanations are often coarse and heavily dependent on the model's internal architecture. Moreover, they are mainly designed for classification tasks and may not fully capture localized decision regions in object detection scenarios.

Another group of methods is based on relevance propagation, such as Layer-wise Relevance Propagation (LRP). These approaches aim to trace the model's prediction back to individual input pixels, providing detailed relevance maps. While this can provide fine-grained explanations, LRP requires direct access to the model's structure and careful adaptation to specific

architectures, limiting its flexibility when working with complex detection models.

Perturbation-based methods, such as SHAP and LIME, explain predictions by modifying parts of the input and observing changes in the model's output. SHAP is grounded in game theory and provides consistent feature importance estimates, but its computational cost becomes high for image-based data. This can limit its practical use in real-time or large-scale medical imaging applications.

LIME offers a more practical and intuitive alternative. It generates local explanations by approximating the complex model's behaviour for a single input using a simple, interpretable model. In the case of images, LIME operates on superpixels, which correspond to coherent image regions and are easier for humans to interpret. This allows LIME to clearly highlight areas that positively or negatively influence the model's decision.

The choice of LIME in this work was motivated by several factors. First, LIME is fully model-agnostic and can be directly applied to YOLOv8 without modifying the detection architecture. Second, its focus on local explanations aligns well with clinical requirements, where understanding individual predictions is more important than global model behaviour. Finally, LIME provides both visual explanations and numerical importance values, enabling not only qualitative assessment but also quantitative evaluation of explanation stability and relevance. For these reasons, LIME was considered a suitable explainability method for interpreting brain tumour detection results in MRI images.

LIME-based explanation of detection results [20]

To enhance the interpretability of the proposed brain tumour detection framework, the Local Interpretable Model-Agnostic Explanations (LIME) method was applied to the YOLOv8 model predictions. LIME is a perturbation-based explainable artificial intelligence technique that provides local explanations for individual predictions by approximating the behaviour of a complex model in the neighbourhood of a specific input instance.

In the context of MRI image analysis, LIME first segments the input image into superpixels, which represent homogeneous, spatially coherent regions. These superpixels serve as interpretable components that can be selectively modified. LIME then generates a large number of perturbed samples by randomly masking different combinations of superpixels, while keeping the remaining regions unchanged. Each

perturbed image is subsequently processed by the trained YOLOv8 detection model.

Based on the variations in the model's predictions across these perturbed samples, LIME estimates the contribution of each superpixel to the final detection outcome. A locally weighted linear model is fitted to approximate the complex decision function of YOLOv8 in the vicinity of the analysed prediction. The resulting coefficients of this surrogate model reflect the relative importance of individual superpixels, with positive weights indicating regions that support tumour detection and negative weights indicating regions that reduce the confidence of the prediction.

In this study, LIME was applied specifically to bounding box predictions for detected tumour regions. The generated explanations were visualised by overlaying the most influential superpixels onto the original MRI images, enabling intuitive qualitative assessment of the model's attention. This visualisation enables verification that the model relies on clinically meaningful tumour regions rather than irrelevant background structures.

Evaluation metrics

The performance of the proposed brain tumour detection framework was evaluated using standard object detection metrics and dedicated explainability metrics to assess the quality of LIME-based explanations.

Detection performance was measured using precision, recall, and mean average precision (mAP). Precision is the proportion of correctly detected tumour regions among all detections, while recall measures the model's ability to identify all relevant tumour regions. The mean average precision was computed at an intersection-over-union threshold of 0.50 (mAP@50) as well as across multiple thresholds ranging from 0.50 to 0.95 (mAP@50–95), providing a comprehensive assessment of detection accuracy [21].

To quantitatively evaluate the interpretability of LIME-generated explanations, several complementary metrics were used, each capturing a different aspect of explanation quality. Together, these metrics provide a more comprehensive understanding of how the model distributes attention across image regions and how reliable the generated explanations are.

Let x denote an input MRI image. Using LIME, the image is segmented into N superpixels

$$S = \{s_1, s_2, \dots, s_N\}.$$

Let d denote a single detection produced by YOLOv8, defined by a bounding box and its predicted tumour class.

For explanation purposes, LIME represents perturbed samples using an interpretable binary vector $z \in \{0, 1\}^N$, where $z_i = 1$ indicates that the superpixel s_i is present (not masked) and $z_i = 0$ indicates that it is masked.

LIME approximates the behaviour of the detection model in the neighbourhood of image x with respect to detection d using a local linear surrogate model:

$$g(z) = \beta_0 + \sum_{i=1}^N \beta_i z_i,$$

where $\beta_i \in \mathbb{R}$ represents the importance weight assigned to superpixel s_i .

Let $\beta = (\beta_1, \beta_2, \dots, \beta_N)$ denote the vector of superpixel importance weights for a single explanation.

Stability measures the robustness of explanations to small perturbations in the input image. For a given image x , LIME explanations are generated multiple times under slight input variations (e.g., noise injection, repeated perturbation sampling).

Let $\beta^{(k)} = (\beta_1^{(k)}, \beta_2^{(k)}, \dots, \beta_N^{(k)})$ denote the vector of superpixel importance weights obtained in the k -th run, $k = 1, \dots, K$.

Stability is defined as the mean standard deviation of superpixel weights across runs:

$$\text{Stability} = \frac{1}{N} \sum_{i=1}^N \text{Std}(\beta_i^{(1)}, \beta_i^{(2)}, \dots, \beta_i^{(K)}).$$

Sparsity quantifies the proportion of superpixels that have a non-negligible contribution to the explanation. Sparse explanations involve only a limited number of influential regions, making them easier to interpret and more clinically meaningful. In contrast, lower sparsity values indicate that the model relies on a broader spatial context.

Let $\tau > 0$ denote a small threshold defining “non-negligible” importance (e.g., $\tau = 0.01 \cdot \max_i |\beta_i|$).

Define the indicator:

$$\mathbb{I}_i = \begin{cases} 1, & |\beta_i| > \tau, \\ 0, & \text{otherwise.} \end{cases}$$

Sparsity is computed as:

$$\text{Sparsity} = \frac{1}{N} \sum_{i=1}^N \mathbb{I}_i.$$

MaxWeight represents the maximum absolute importance value assigned to a single superpixel.

This metric highlights the most influential image region contributing to the prediction and provides insight into whether the model relies strongly on a dominant area. A higher MaxWeight value suggests that a specific region plays a critical role in the detection decision, which is desirable when this region corresponds to the actual tumour location.

MaxWeight captures the strength of the most influential superpixel in the explanation:

$$\text{MaxWeight} = \max_{i=1,\dots,N} |\beta_i|.$$

Optionally, normalized by the sum of absolute weights:

$$\text{MaxWeight}_{\text{norm}} = \frac{\max_i |\beta_i|}{\sum_{j=1}^N |\beta_j|}.$$

Entropy evaluates how the importance weights are distributed across all superpixels. Lower entropy values indicate that importance is concentrated in fewer regions, leading to clearer, more interpretable explanations. Higher entropy, on the other hand, reflects a more uniform distribution of importance, suggesting that multiple regions jointly influence the model's decision.

First, normalize absolute weights to obtain a probability distribution:

$$p_i = \frac{|\beta_i|}{\sum_{j=1}^N |\beta_j|}, i = 1, \dots, N.$$

Entropy is then computed using Shannon entropy:

$$\text{Entropy} = - \sum_{i=1}^N p_i \log_2(p_i),$$

where terms with $p_i = 0$ are defined as contributing zero to the sum.

All explainability metrics were first computed at the level of individual detections and then averaged over correctly detected tumour instances in the test set. The practical computation of these metrics was performed under the following assumptions:

- LIME explanations were generated using the same number of superpixels and perturbation samples across all experiments.
- Stability was computed using repeated LIME runs on the same image.
- Sparsity threshold τ was defined relative to the maximum absolute weight to ensure scale invariance.
- Metrics were averaged over correctly detected tumour instances to obtain representative values.

From a reliability engineering perspective, stability relates to robustness under uncertainty (epistemic variability); sparsity reflects importance concentration

and cognitive interpretability; MaxWeight aligns with factor dominance in importance analysis; entropy captures uncertainty dispersion across decision factors.

The combination of detection and explainability metrics enables joint evaluation of model performance and transparency, which is essential for assessing the suitability of deep learning models for clinical decision-support applications [13, 15].

Experimental Results

The experimental evaluation was conducted to assess both the detection performance and the explainability of the proposed framework for brain tumour detection from MRI images. The experiments focused on comparing two YOLOv8 model variants and analysing the interpretability of their predictions using the LIME method. All experiments were conducted under identical settings to ensure reproducibility.

Software implementation and Python libraries

The proposed framework was implemented in Python due to its extensive ecosystem of libraries for deep learning, image processing, and explainable artificial intelligence, as well as its widespread adoption in medical image analysis research.

The YOLOv8 detection models were implemented using the Ultralytics library, which provides an efficient and user-friendly framework for training, evaluating, and deploying state-of-the-art object detection models. This library was used for model configuration, training, inference, and evaluation, including the computation of detection performance metrics.

The LIME library was employed to generate explainable visualizations of model predictions. Specifically, the `lime_image` module was used to produce local explanations based on superpixel perturbations, enabling interpretation of YOLOv8 detection results without requiring access to the internal architecture of the model.

For image preprocessing and superpixel segmentation, the `scikit-image` library was utilized, particularly the `skimage.segmentation` module. This library was used to divide MRI images into homogeneous regions required for LIME-based explanations. Additional numerical computations were performed using NumPy, while visualization of results was carried out with Matplotlib, which was used to generate figures, heatmaps, and annotated MRI images.

Together, these Python libraries enabled efficient implementation of the proposed detection and explainability pipeline, ensuring reproducibility of experiments and facilitating comprehensive analysis of both model performance and interpretability.

Training setup

The YOLOv8n and YOLOv8s models were trained using the Ultralytics framework on annotated MRI images under identical training conditions to ensure fair comparison. All input images were resized to a fixed resolution of 640×640 pixels prior to training.

Both models were trained for 40 epochs with a batch size of 16, using stochastic gradient-based optimization. Standard data augmentation techniques, including geometric transformations and intensity variations, were applied to improve model generalization. The dataset was divided into training, validation, and testing subsets.

Early stopping was applied with a patience of 5 epochs, meaning that training was terminated if no improvement in validation performance was observed over five consecutive epochs. Model weights were automatically saved during training, and the version achieving the best validation performance was selected for final evaluation on the test set.

Detection results

The detection performance was evaluated using two variants of the YOLOv8 object detection model, namely YOLOv8n (nano) and YOLOv8s (small), which differ in terms of model complexity, number of parameters, and computational requirements. The YOLOv8n model represents a lightweight architecture optimized for faster inference and lower hardware demands, whereas YOLOv8s offers increased representational capacity, enabling more accurate feature extraction at the cost of higher computational load.

Both models were trained under identical conditions using the same annotated MRI dataset and the same training and validation splits. To ensure a fair and objective comparison, performance evaluation was conducted on an independent held-out test set that was not used during training or validation. This approach minimises the risk of overfitting and allows assessment of the models' generalisation ability.

Table 1 summarizes the quantitative detection results obtained by the evaluated models. The YOLOv8s model achieved higher precision and recall values

compared to YOLOv8n, resulting in improved mean average precision across both evaluated IoU thresholds. This performance gain reflects the ability of the larger model to capture more complex visual patterns present in MRI images.

Table 1. Detection performance comparison of YOLOv8 models

Metric	YOLOv8n	YOLOv8s	Difference
Overall performance			
Precision (P)	0.889	0.906	+0.017
Recall (R)	0.863	0.871	+0.008
mAP@50	0.918	0.927	+0.009
mAP@50-95	0.700	0.713	+0.013
Glioma			
Precision (P)	0.802	0.823	+0.021
Recall (R)	0.713	0.768	+0.055
mAP@50	0.812	0.834	+0.022
mAP@50-95	0.519	0.545	+0.026
Meningioma			
Precision (P)	0.945	0.972	+0.027
Recall (R)	0.937	0.937	0.000
mAP@50	0.975	0.983	+0.008
mAP@50-95	0.823	0.836	+0.013
Pituitary tumour			
Precision (P)	0.921	0.922	+0.001
Recall (R)	0.941	0.908	-0.033
mAP@50	0.966	0.962	-0.004
mAP@50-95	0.757	0.758	+0.001

Despite its lower complexity, YOLOv8n demonstrated competitive detection performance, indicating that lightweight models can still provide reliable tumour localization when computational resources are limited. This trade-off between accuracy and efficiency is particularly relevant for practical deployment scenarios, such as real-time clinical decision-support systems.

The confusion matrix presented in Fig. 1 provides a detailed insight into the class-wise performance of the YOLOv8s model for brain tumour detection. Overall, the model demonstrates strong discriminatory capability across all three tumour classes, with many samples correctly classified.

For the glioma class, 282 samples were correctly classified, with only 3 misclassifications as meningioma and no confusion with pituitary tumours. The meningioma class achieved 136 correct predictions, with a small number of errors, including 4 samples misclassified as glioma and 2 as pituitary tumours, indicating partial visual similarity among these tumour types.

The pituitary tumour class showed the most reliable performance, with 184 correctly classified samples and only one misclassification as meningioma. Overall, misclassifications were rare and occurred mainly between

glioma and meningioma classes, reflecting overlapping MRI characteristics. These results confirm the robustness and stability of the proposed detection framework.

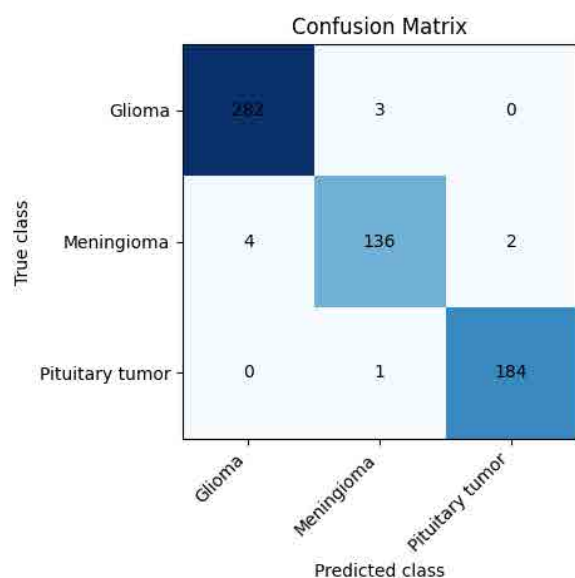


Fig. 1. Confusion matrix of the model

Explainability results

The explainability of the proposed brain tumour detection framework was analysed using the LIME method to identify which image regions contributed most to YOLOv8 model predictions. The objective of this analysis was to verify whether the detection decisions were based on clinically relevant tumour regions and to compare the interpretability of the model variants.

Figure 2 presents an MRI brain image with superpixels highlighted by their contribution to the YOLOv8 model's prediction. Although the tumour region is clearly included among the highlighted areas, LIME also marks several superpixels that do not visually correspond to pathological tissue.

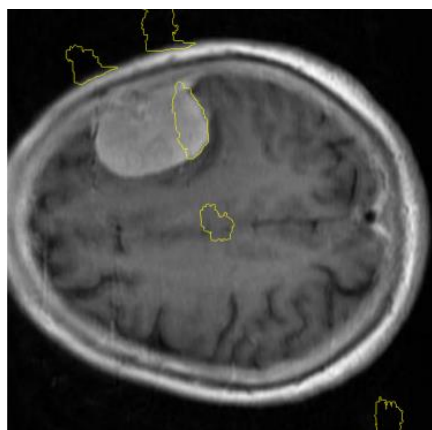


Fig. 2. Visualization of LIME explanation no. 1

This result indicates that the explanation encompasses multiple regions of the image, not all of which represent tumour-related structures, demonstrating the local nature of the LIME-based explanation.

Figure 3 shows the corresponding heatmap visualization of the same LIME explanation. The heatmap provides a continuous representation of feature importance, revealing that the influence of relevant regions is more dispersed compared to well-performing cases. This representation helps to better understand why the explanation is considered less reliable.

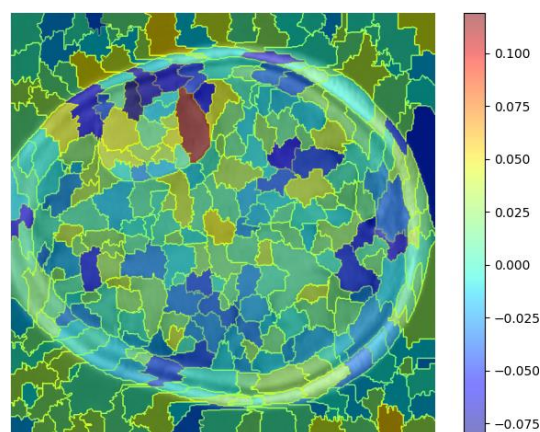


Fig. 3. Heatmap visualization of LIME explanation no. 1

Figure 4 presents an example of a LIME-based superpixel explanation generated for a correctly detected tumour. The highlighted superpixels are predominantly concentrated within the tumour area, indicating that the model focuses on pathological regions when making predictions. The clear alignment between the superpixel importance and the tumour boundaries suggests that the detection model relies on meaningful visual features rather than unrelated background structures.

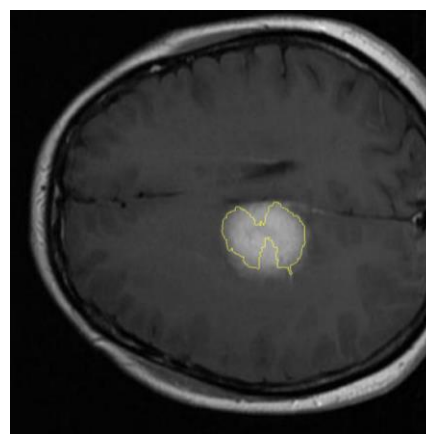


Fig. 4. Visualization of LIME explanation no. 2

To provide a more detailed view of the spatial distribution of feature importance, Fig. 5 illustrates a heatmap representation of the LIME explanation for the same detection case. Warmer colors indicate regions with higher influence on the model's prediction. As shown in the figure, the highest contribution values are localised within the tumour region, while the surrounding brain tissue exhibits significantly lower influence. Minor contributions outside the tumour area can be attributed to contextual information or intensity gradients present in MRI images.

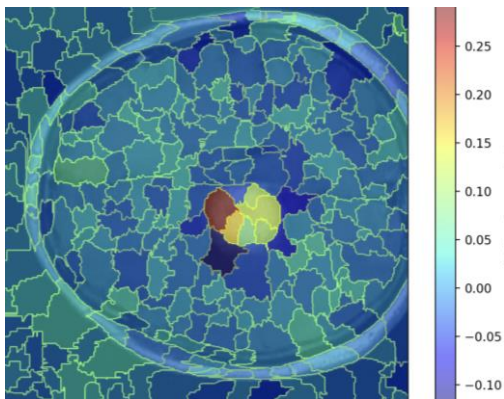


Fig. 5. Heatmap visualization of LIME explanation no. 2

Together, the superpixel-based visualisation (Fig. 4) and the heatmap representation (Figure 5) offer complementary insights into the YOLOv8 model's decision-making process. While the superpixel visualisation highlights discrete influential regions, the heatmap provides a continuous view of importance distribution. This combined analysis confirms that the model's predictions are primarily driven by tumour-related image features, supporting the transparency and reliability of the proposed explainable detection framework.

Qualitative evaluation of LIME explanations

To complement the qualitative analysis based on visual inspections of LIME explanations, a quantitative evaluation was conducted to objectively assess the quality and characteristics of the generated explanations. Quantitative interpretability metrics enable systematic comparison of explanation stability, spatial concentration, and importance distribution, providing deeper insight into the model's decision-making behaviour beyond visual interpretation alone.

Table 2 presents a quantitative evaluation of the explanations generated using the LIME method, providing insight into the stability, concentration, and

distribution of superpixel importance. These metrics enable objective assessment of explanation quality beyond qualitative visual inspection.

Table 2. Quantitative interpretability metrics for LIME

<i>Metric</i>	<i>Value</i>
<i>Stability</i>	0.0302
<i>Sparsity</i>	0.6442
<i>MaxWeight</i>	0.2840
<i>Entropy</i>	5.1498

The stability value of 0.0302 indicates low variability in superpixel weights, suggesting that the generated explanations are robust and consistent under small perturbations of the input image. This behavior is desirable in medical imaging applications, where reliable and repeatable explanations are essential for building trust in automated systems.

The sparsity metric reached a value of 0.6442, indicating that approximately two-thirds of the superpixels contributed meaningfully to the model's decision. This result suggests that the model does not rely on a single isolated region but instead integrates information from a broader spatial context. While highly sparse explanations may indicate overly localized reasoning, the observed sparsity reflects a balanced attention pattern that captures both focal tumour regions and their immediate surroundings.

The maximum superpixel weight of 0.2840 confirms the presence of dominant regions with strong influence on the model output. These highly weighted superpixels correspond to visually salient tumour areas observed in the LIME visualizations, reinforcing the consistency between quantitative metrics and qualitative explanation maps.

Finally, the entropy value of 5.1498 indicates a relatively distributed allocation of importance across multiple superpixels. This suggests that the model's predictions are supported by several relevant image features rather than being dominated by a single region. Such behavior is particularly important in medical image analysis, where pathological patterns may span multiple spatial structures.

Overall, the quantitative results summarized in Table 2 demonstrate that the LIME explanations are stable, informative, and well-aligned with clinically relevant image regions. The combination of focused dominant areas and distributed contextual information supports the transparency and reliability of the proposed explainable brain tumour detection framework.

Discussion

The results confirm that YOLOv8-based object detection is a suitable approach for brain tumour localization in MRI images. Both evaluated models achieved high detection performance, with YOLOv8s consistently outperforming YOLOv8n across most metrics. The improved performance of YOLOv8s is mainly attributed to its higher model capacity, which enables more effective representation of complex tumour patterns, particularly for gliomas.

Analysis of the confusion matrix showed that misclassifications were infrequent and occurred primarily between glioma and meningioma classes, which exhibit similar visual characteristics in MRI scans. Pituitary tumours were detected with high reliability, likely due to their distinct anatomical location. These findings are consistent with observations reported in related medical imaging studies.

The explainability analysis demonstrated that LIME provides meaningful insight into the model's decision-making process. Both qualitative visualizations and quantitative metrics indicated that the models predominantly focus on clinically relevant tumour regions. The YOLOv8s model produced more stable and concentrated explanations, while YOLOv8n showed more distributed importance patterns, suggesting that increased model complexity contributes not only to higher accuracy but also to improved interpretability.

Several limitations should be considered. The analysis was performed on two-dimensional MRI scans, and tumour localization was limited to bounding boxes rather than precise segmentation. Additionally, LIME provides local explanations that may vary depending on perturbation settings. Despite these limitations, the proposed framework effectively combines detection performance with explainability and supports the development of transparent AI-based medical image analysis systems.

In terms of future directions, the proposed framework can be extended to include a wider range of brain tumour types, which would improve its applicability in more diverse clinical scenarios. Additionally, future work may focus on generating structured textual descriptions of detected or non-detected tumours, providing clinicians with interpretable summaries that complement visual explanations. Such extensions could further enhance the practical usability of explainable detection systems in clinical decision-support workflows.

Conclusions

This paper presented an explainable framework for brain tumour detection from MRI images based on YOLOv8 object detection models and the LIME explainable artificial intelligence method. Experimental results demonstrated that YOLOv8 models achieve reliable detection performance across multiple tumour types, with the YOLOv8s variant providing improved accuracy and more stable explanations.

The integration of LIME enabled transparent interpretation of model predictions by highlighting clinically relevant image regions. Quantitative interpretability metrics further confirmed the stability and informativeness of the generated explanations, supporting the trustworthiness of the proposed approach.

Despite certain limitations, including the use of two-dimensional MRI images and bounding box localization, the proposed framework effectively balances detection accuracy and explainability. Future work will focus on extending the approach to three-dimensional data, incorporating segmentation-based methods, and exploring additional explainability techniques to further enhance clinical applicability.

Acknowledgment

This work was supported in part by the Slovak Research and Development Agency under the grant "Development of a new approach for reliability analysis and risk assessment based on artificial intelligence" (reg.no. APVV-23-0033), and in part by the Ministry of Education, Science, Research, and Sport of the Slovak Republic "Creation of new methods and algorithms of eXplainable Artificial Intelligence based on Importance Analysis" under Grant VEGA 1/0090/25.

Conflict of interest

The authors declare that they have no conflicts of interest, including financial, personal, authorial, or any other nature, that could influence the research or the results published in this article.

Funding

The study was conducted without financial support.

Data availability

The manuscript has no associated data.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technology in the creation of this paper.

References

1. Di Fazio, N., Zanza, C., Longhitano, Y. *et al.* (2026), "Artificial intelligence for early diagnosis in emergency department", *J Anesth Analg Crit Care*, Vol. 6, No.7, 2026, DOI: <https://doi.org/10.1186/s44158-025-00334-y>
2. Zaitseva, E., Levashenko, V., Rabcan, J., Kvassay, M. (2023), "A New Fuzzy-Based Classification Method for Use in Smart/Precision Medicine", *Bioengineering*, Vol. 10(7), No. 838, DOI: <https://doi.org/10.3390/bioengineering10070838>
3. Francisco, K.K.Y., Apuhin, A.E.C., *et al.* (2026), "Personalized medicine and health equity: overcoming cost barriers and ethical challenges", *Int J Equity Health*, Vol.25, No. 4. DOI: <https://doi.org/10.1186/s12939-025-02710-0>
4. Zaitseva, E., Levashenko, V., Rabcan, J., Krsak, E. (2020), "Application of the structure function in the evaluation of the human factor in healthcare", *Symmetry*, Vol.12(1), No. 93. DOI: <https://doi.org/10.3390/SYM12010093>
5. Zaitseva, E., Levashenko, V. (2026), "Reliability engineering in healthcare: Opportunities and challenges", *Reliability Engineering and System Safety*, Vol. 267, No. 111933. DOI: <https://doi.org/10.1016/j.res.2025.111933>
6. Sharma, R.M. (2025), "Artificial intelligence in medical image analysis and molecular diagnostics: recent advances and applications", *J Med Artif Intell.*, Vol.8, 53 p. DOI: <https://doi.org/10.21037/jmai-24-412>
7. Menze, B.H., *et al.* (2015), "The Multimodal Brain Tumour Image Segmentation Benchmark (BRATS) ", *IEEE Trans. Med. Imaging*, Vol. 34, No. 10, Oct. 2015, pp. 1993–2024. DOI: <https://doi.org/10.1109/TMI.2014.2377694>
8. Aleid, A., Alhussaini, K., Alanazi, R., Altwaimi, M., Altwijri, O., Saad, A.S., (2023), "Artificial Intelligence Approach for Early Detection of Brain Tumours Using MRI Images", *Applied Science*, Vol. 13, No. 3808. DOI: <https://doi.org/10.3390/app13063808>
9. Gökmen, N. (2025), "AI techniques for brain tumour segmentation in MRI: a review (2019–2024)", *Netw Model Anal Health Inform Bioinforma*, Vol.14, 168 p. DOI: <https://doi.org/10.1007/s13721-025-00650-x>
10. Kolarik, M., Sarnovsky, M., Paralic, J., Babic, F. (2023), "Explainability of deep learning models in medical video analysis: a survey", *Peer J Comput. Sci.*, Vol. 9, 1253 p. DOI: <https://doi.org/10.7717/peerj-cs.1253>
11. Kondratenko, Y., Sidenko, I., Kondratenko, G., Petrovych, V., Taranov, M., Sova, I. (2021), "Artificial Neural Networks for Recognition of Brain Tumours on MRI Images". *Information and Communication Technologies in Education, Research, and Industrial Applications. ICTERI 2020*. Communications in Computer and Information Science, Vol 1308. Springer, Cham, DOI: https://doi.org/10.1007/978-3-030-77592-6_6
12. Ali, S., Abuhmed, T., El-Sappagh, Sh., *et al.* (2023), "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence", *Information Fusion*, Vol. 99, 101805 p. DOI: <https://doi.org/10.1016/j.inffus.2023.101805>
13. Hassan, S.U., Abdulkadir, S.J., Zahid, M.S.M., Al-Selwi, S.M. (2025), "Local interpretable model-agnostic explanation approach for medical imaging analysis: A systematic literature review", *Comput. Biol. Med.*, Vol. 185, 109569 p. DOI: <https://doi.org/10.1016/j.combiomed.2024.109569>
14. Zaitseva, E., Levashenko, V. (2025), "Reliability Analysis Based on Aleatory and Epistemic Uncertainty Using Binary Decision Diagrams", *International Journal of Intelligent Systems*, Vol.2025, No. 6471577. DOI: <https://doi.org/10.1155/int/6471577>
15. Zhou, Z., Hooker, G., Wang, F. (2021), "S-LIME: Stabilized-LIME for Model Explanation", in *Proc. of the 27th ACM SIGKDD Conf. on Knowledge Discovery & Data Mining*, pp. 2429-2438. DOI: <https://doi.org/10.1145/3447548.3467274>
16. Zaitseva, E., Rabcan, J., Levashenko, V., Kvassay, M. (2023), "Importance analysis of decision-making factors based on fuzzy decision trees", *Applied Soft Computing*, Vol. 134, No. 109988. DOI: <https://doi.org/10.1016/j.asoc.2023.109988>
17. "Medical Image DataSet: Brain Tumour Detection". available: <https://www.kaggle.com/datasets/pkdarabi/medical-image-dataset-brain-tumour-detection> (last accessed Feb. 04, 2026).
18. Alsufyani, A. (2025), "Performance comparison of deep learning models for MRI-based brain tumour detection", *AIMS Bioeng.*, Vol. 12, No. 1, pp. 1-21. DOI: <https://doi.org/10.3934/bioeng.2025001>
19. Ultralytics, 'Explore Ultralytics YOLOv8'. available: <https://docs.ultralytics.com/models/yolov8/> (last accessed Feb. 04, 2026).
20. Explainable AI (XAI) Using LIME', GeeksforGeeks. available: <https://www.geeksforgeeks.org/artificial-intelligence/introduction-to-explainable-ai-using-lime/> (last accessed Feb. 04, 2026).
21. Ultralytics, 'Performance Metrics Deep Dive'. available: <https://docs.ultralytics.com/guides/yolo-performance-metrics> (last accessed Feb. 04, 2026).

Received (Надійшла) 05.01.2026

Accepted for publication (Прийнята до друку) 10.02.2026

Publication date (Дата публікації) 30.03.2026

Відомості про авторів / About the Authors

Досталова Крістіна – аспірантка кафедри інформатики, факультет управлінських наук та інформатики, Жилінський університет; Жиліна, Словаччина;

Kristina Dostalova – postgraduate student at the Department of Informatics, Faculty of Management Science and Informatics, University of Žilina; Žilina, Slovakia;

e-mail: kristina.dostalova@st.fri.uniza.sk

ORCID ID: <https://orcid.org/0009-0002-8567-716X>

Чижмарова Олександра – аспірантка кафедри інформатики, факультет управлінських наук та інформатики, Жилінський університет; Жиліна, Словаччина;

Alexandra Cizmarova – postgraduate student at the Department of Informatics, Faculty of Management Science and Informatics, University of Žilina; Žilina, Slovakia;

e-mail: alexandra.cizmarova@st.fri.uniza.sk

ORCID ID: <https://orcid.org/0009-0007-1021-4722>

Клімо Марек – аспірант кафедри інформатики, факультет управлінських наук та інформатики, Жилінський університет; Жиліна, Словаччина;

Marek Klimo – postgraduate student at the Department of Informatics, Faculty of Management Science and Informatics, University of Žilina; Žilina, Slovakia;

e-mail: marek.klimo@fri.uniza.sk

ORCID ID: <https://orcid.org/0009-0004-2366-4325>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=57226717649>

Ярошевич Роман Олександрович – доктор філософії, старший викладач кафедри електронних обчислювальних машин, факультет комп'ютерної інженерії та інформаційних технологій, Харківський національний університет радіоелектроніки; Харків, Україна;

Roman Yaroshevych – PhD, senior lecturer at the Department of Electronic Computers, Faculty of Computer Engineering and Information Technology, Kharkiv National University of Radio Electronics; Kharkiv, Ukraine;

e-mail: roman.yaroshevych@nure.ua

ORCID ID: <https://orcid.org/0000-0002-7949-1513>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=58624172500>

ВИЯВЛЕННЯ ПУХЛИН ГОЛОВНОГО МОЗКУ НА МРТ-ЗОБРАЖЕННЯХ З МОЖЛИВІСТЮ ПОЯСНЕННЯ ВИКОРИСТАННЯ YOLOV8 У LIME

Актуальність. Точна ідентифікація пухлин головного мозку за допомогою магнітно-резонансної томографії (МРТ) є критично важливим завданням в аналізі медичних зображень. Хоча підходи глибокого навчання часто досягають відмінної продуктивності виявлення, їм часто бракує прозорості в процесах прийняття рішень. Ця відсутність інтерпретованості обмежує довіру в клінічній практиці та створює сильну потребу у впровадженні методів пояснювального штучного інтелекту (ХАІ). **Об'єкт дослідження.** Об'єктом цього дослідження є автоматизований процес виявлення пухлин головного мозку на МРТ-сканах з використанням підходів на основі згорткових нейронних мереж. **Предмет дослідження.** Предметом дослідження є застосування моделей виявлення об'єктів YOLOv8 разом з методом пояснювальності LIME для інтерпретації виходів моделі в контексті аналізу медичних зображень. **Мета.** Метою цієї статті є розробка та оцінка пояснювальної структури для виявлення пухлин головного мозку, яка інтегрує виявлення об'єктів на основі YOLOv8 з інтерпретованістю, керованою LIME, за підтримки кількісної оцінки якості пояснення. **Результати.** Експериментальна оцінка демонструє, що моделі YOLOv8 здатні точно виявляти вибрані типи пухлин головного мозку на зображеннях МРТ, тоді як LIME успішно визначає області зображення, які мають найбільший вплив на рішення моделі. Запропоновані кількісні метрики підтверджують стабільність та розрідженість згенерованих пояснень, тим самим покращуючи інтерпретованість та надійність запропонованої системи виявлення.

Ключові слова: виявлення пухлини головного мозку; МРТ; YOLOv8; зрозумілий штучний інтелект; LIME; аналіз медичних зображень.

Бібліографічні описи / Bibliographic descriptions

Досталова, К., Чижмарова, О., Клімо, М., Ярошевич Р.О. Виявлення пухлин головного мозку на МРТ-зображеннях з можливістю пояснення використання YOLOv8 у LIME. *Сучасний стан наукових досліджень та технологій в промисловості*. 2026. № 1 (35). С. 17–27. DOI: <https://doi.org/10.30837/2522-9818.2026.1.017>

Dostalova, K., Cizmarova, A., Klimo, M., Yaroshevych, R. (2026), "Explanation Detected Brain Tumours in MRI Images Using YOLOv8 with LIME-Based Interpretation", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (35), P. 17–27. DOI: <https://doi.org/10.30837/2522-9818.2026.1.017>

Maksym Yena, Olha Pohudina

MODELING OF ADAPTIVE UAV ROUTE CONTROL BASED ON REINFORCEMENT LEARNING ALGORITHMS

Subject matter is the reward function, action policy, and learning dynamics of the Proximal Policy Optimization (PPO) algorithm in the task of adaptive UAV navigation under dynamic airspace conditions and limited energy resources. **Goal** is to develop a simulation environment and a modified Proximal Policy Optimization (PPO) model for adaptive route management of a single UAV in 2D and 3D environments, considering the distance to the target, collision risk, and energy consumption. **Tasks:** to develop 2D and 3D simulation environments with different obstacle configurations and UAV motion parameters; to formulate a combined PPO reward function that incorporates distance to the target, collisions, and energy consumption; to implement and train PPO, DQN, and A2C algorithms in standardized navigation scenarios; to perform a comparative analysis of algorithm performance using key metrics: path length, number of collisions, reward, and energy consumption; to conduct statistical validation of the results using the t-test and confidence intervals; to analyze the influence of PPO hyperparameters on policy stability and learning convergence in 2D and 3D environments. **Methods:** deep reinforcement learning algorithms (PPO, DQN, A2C); two simulation models (2D and 3D) with randomly generated static obstacles were developed; a combined reward function was formulated, integrating distance-to-target progress, collision penalties, and an energy-related component; model performance was evaluated using average reward, path length, number of collisions, and total energy expenditure; statistical significance was assessed using the t-test and 95% confidence intervals. **Results:** The modified PPO model reduced the number of collisions in the 2D environment by 94,8% and shortened the route length by 94,3% compared to the baseline PPO, while exhibiting higher energy consumption due to more complex avoidance maneuvers. In the 3D environment, similar trends were confirmed, including improved navigation safety, more stable policy behavior, and statistically significant improvements across key metrics ($p < 0,05$). **Conclusions:** A unified 2D/3D simulation environment for adaptive UAV routing and a modified PPO model with a combined reward function were developed. In the 2D environment, the model achieved a $\approx 94,8\%$ reduction in collisions, a $\approx 94,3\%$ reduction in path length, and a $\approx 92,5\%$ increase in average reward compared to the baseline PPO. In the 3D environment, analogous improvements and statistically significant gains ($p < 0,05$) were obtained. A relationship between avoidance aggressiveness and energy consumption was identified, enabling selection of an optimal policy for BVLOS scenarios.

Keywords: adaptive control; Proximal Policy Optimization; reinforcement learning; simulation modeling; routing; 3D navigation.

1. Introduction

The increasing complexity of airspace and the rapid growth in the number of autonomous UAVs create a need for adaptive navigation methods capable of making real-time decisions in the face of dynamic obstacles and limited energy resources [1]. The problem is that existing adaptive routing methods work only in simplified 2D environments and do not account for energy constraints, which is critical for real-world UTM systems [2]. Classic rule-based navigation methods do not provide adaptability, especially in 3D space with a large number of degrees of freedom. Reinforcement learning (RL) methods, in particular Proximal Policy Optimization (PPO), have demonstrated effectiveness in dynamic evasion and trajectory prediction tasks. However, most studies are limited to two-dimensional environments or simplified obstacle models that do not reflect the actual structure of airspace [3].

2. Literature Review and Problem Definition

Among the most promising areas is the use of reinforcement learning algorithms, particularly Proximal Policy Optimization (PPO), which show good results in complex simulation environments [4]. Dai et al. (2023) demonstrate the effectiveness of PPO in safe agent control tasks; however, the main focus of their work is on two-dimensional (2D) spaces. Studies [5, 6] examine the classification of reinforcement learning algorithms and their ability to generalize in simulation models.

There is also a concept of a simulation environment for urban air traffic, but the implementation of adaptive navigation in three-dimensional (3D) space remains partially unresolved [7]. Most recent works [8, 9] highlight the importance of traffic forecasting that accounts for stochastic factors, but do not address their integration with learning algorithms.

For the most part, it is the combination of several algorithms that demonstrates the effectiveness of Monte Carlo and discrete-event modeling in traffic scenarios; however, these methods do not combine with dynamic collision-avoidance strategies [10]. Particular attention should be paid to studies analyzing the impact of hyperparameters on the effectiveness of agent training in prediction and dynamic collision avoidance tasks [11, 12].

Despite significant progress in UAV navigation, most modern approaches remain limited to two-dimensional models and simplified obstacles that do not account for vertical maneuvers and the multi-level structure of airspace. This reduces the algorithms' ability to generalize behavior in realistic 3D scenarios.

Thus, there is a need to create a simulation model of adaptive UAV route control based on PPO, focused on realistic scenarios in three-dimensional space, with a comparative analysis against classical 2D approaches.

3. Research Objectives and Tasks

The objective of this work is to create a simulation environment and develop a modified Proximal Policy Optimization (PPO) model for adaptive route control of a single UAV in 2D and 3D spaces, taking into account distance to the target, collision risk, and energy consumption. To achieve this goal, the following research tasks must be addressed:

- develop 2D and 3D simulation environments with varying obstacle structures and UAV motion parameters;
- formulate a combined PPO reward function that accounts for proximity to the target, collisions, and energy consumption;
- implement and train PPO, DQN, and A2C algorithms in standardized navigation scenarios;
- conduct a comparative analysis of algorithm performance based on key metrics: path length, number of collisions, average reward, and energy consumption;
- Perform a statistical test of the results using the t-test and confidence intervals;
- Analyze the impact of PPO hyperparameters on policy stability, convergence speed, and navigation characteristics in 2D and 3D environments.

Problem Formulation and Environment Model

The simulation model uses the following basic mathematical relationships: The distance between the

agent and the target in 2D and 3D environments is defined by the Euclidean metric:

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}, \quad (1)$$

where the component $(z_1 - z_2)$ is zero. Using a single metric allows for a fair comparison of the neural network's learning performance when transitioning from a 2D to a 3D navigation problem.

The route length L is defined as the sum of the distances between all consecutive points along the trajectory. It accounts for the sum of the distances between all consecutive points that form the motion trajectory [13]. This approach allows for an objective assessment of the efficiency of the constructed route in a three-dimensional environment.

$$L = \sum_{i=1}^{N-1} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 + (z_{i+1} - z_i)^2}, \quad (2)$$

where L – route length, N – number of points in the trajectory; x_i, y_i, z_i – the agent's coordinates at point i ; $x_{i+1}, y_{i+1}, z_{i+1}$ – coordinates at the next point $i+1$.

The agent's total reward at step t is defined as:

$$G_t = \sum_{k=0}^T \gamma^k r_{t+k}, \quad (3)$$

where G_t – the agent's total discounted reward, starting from step t , r_{t+k} – instant rewards at step $t+k$, $\gamma (0 \leq \gamma < 1)$ – discount rate, T – episode length.

Environment Model

A simplified energy consumption model of a multirotor UAV was used to evaluate route efficiency. Instantaneous energy consumption E_t (J) at time step t is calculated as the sum of the energy required for hovering and the linear component proportional to the speed of movement:

$$E_t = (P_t + P_{move} |v_t|) \Delta t, \quad (4)$$

where E_t – instantaneous energy consumption at step t , P_t – UAV hovering power, P_{move} – energy consumption coefficient for horizontal movement, v_t – agent velocity, Δt – simulation step time.

The agent's state is generally defined as:

$$s_t = (x_t, y_t, z_t, x_{goal}, y_{goal}, z_{goal}), \quad (5)$$

where x_t, y_t, z_t – current UAV coordinates, v_t – current UAV velocity, $x_{goal}, y_{goal}, z_{goal}$ – target point coordinates. For the two-dimensional case, $z_t = z_{goal} = 0$.

Action space:

- 2D: 5 actions ($\uparrow \downarrow \leftarrow \rightarrow$ stay);
- 3D: discrete displacements $(\Delta x, \Delta y, \Delta z) \in \{-1, 0, +1\}$.

In the 3D case, the action space is an extension of the 2D model and includes all possible discrete displacements along the three coordinate axes.

Formalization of the navigation problem

In a two-dimensional environment, the calculation of distance and reward is performed similarly to the three-dimensional case (see Formula 1), assuming $z_{agent} = z_{goal} = 0$.

To evaluate the agent's progress, a variable reward is used depending on the change in distance:

$$R_t = \alpha(d_{t-1} - d_t) - \beta I_{collision} + \gamma I_{goal} + \delta E_t, \quad (6)$$

where d_t is the Euclidean distance from the agent to the target at step t , defined by formula (1), $I_{collision} \in \{0, 1\}$ – obstacle collision indicator, $I_{goal} \in \{0, 1\}$ – goal achievement indicator, E_t – agent's instantaneous energy expenditure at step t determined by formula (4), $\alpha, \beta, \gamma, \delta$ – weight coefficients of the reward function components.

This approach allows the agent to receive feedback not only regarding the fact of goal achievement but also regarding the quality of individual actions. The positive component $\alpha(d_{t-1} - d_t)$ encourages movement toward the goal, while penalties for collisions and excessive energy expenditure steer the policy toward a safe and efficient trajectory.

A series of experiments established that optimal ratios between reward components are achieved at values of: $\gamma_{goal} = 50, \gamma_{collision} = -50$, which ensures a balance between convergence speed, learning stability, and the agent's ability to avoid obstacles. The reward configuration also includes a small negative penalty for each action, which incentivizes the agent to minimize path length and avoid unnecessary movements. Thus, the complete reward function at each episode step includes a variable reward for progress, penalties for undesirable events, and a fixed reward for reaching the goal, which ensures the learning of coordinated and goal-oriented agent behavior.

In the implemented model, the agent (UAV) makes movement decisions based on the current state, which includes: the agent's coordinates (x, y) , velocity, and the goal's position (x_{goal}, y_{goal}) . The action space is discrete and defines the possible directions of movement [14].

In a two-dimensional environment, the agent has access to a fixed set of five discrete actions: move up, down, left, right, or stay in place.

In a three-dimensional environment, the action space is described by a displacement vector $a_t = (\Delta x, \Delta y, \Delta z)$, where each coordinate can take the values $-1, 0, \text{ or } +1$, but a change is allowed along only one axis per step. Thus, the agent has six permissible directions of movement: forward, backward, left, right, up, and down.

After performing the selected action, the agent transitions to a new state and receives an instant reward, which is calculated using the reward function R_t , which accounts for changes in distance to the goal, the presence of collisions, and reaching the end point.

Policy updates are performed using the Proximal Policy Optimization (PPO) algorithm, which takes the preference function into account and applies a clipping mechanism to control policy changes. This prevents excessive updates and maintains the stability of the learning process.

4. Materials and Methods

To update the policy in both environments (2D and 3D), the Proximal Policy Optimization (PPO) algorithm [14] was used, which combines the efficiency of gradient methods with a mechanism for stabilizing updates. The main idea behind PPO is to limit the magnitude of policy updates during each iteration, preventing abrupt changes that could cause learning divergence.

Table 1. Main PPO training configuration parameters

Parameter	Value
Number of iterations	4
Frames per second (FPS)	1979 (frames per second)
Total number of steps	65,536
Algorithm	PPO (Proximal Policy Optimization)
Main hyperparameters	clip_range: 0.2; learning_rate: 0.0003; batch_size: 64

One of the components of the PPO algorithm is the critic's value loss function, defined as:

$$L_{value}(\theta) = E_t[(V_\theta(s_t) - V_t^{target})^2], \quad (7)$$

where $V_t^{target} = r_t + \gamma V(s_{t+1})$ is the target estimate for the value function.

The primary objective function of PPO for policy updating is the clipped policy objective:

$$L_{clip}(\theta) = E_t[\min(r_t(\theta)A_t, clip(r_t(\theta), 1-\varepsilon, 1+\varepsilon)A_t)], \quad (8)$$

where $r_t(\theta)$ is the ratio of action probabilities under the new and old policies, A_t is the preference score, and ε is the clip range parameter.

To maintain sufficient policy diversity during training, an entropy bonus is introduced:

$$L_{entropy}(\theta) = \beta_{ent} E_t[H(\pi_\theta(a_t | s_t))], \quad (9)$$

where $H(*)$ is the policy entropy, β_{ent} is the coefficient (0.01 is chosen in this work).

All components are combined into a final function:

$$L_{total}(\theta) = L_{clip}(\theta) + c_1 L_{value}(\theta) - c_2 L_{entropy}(\theta), \quad (10)$$

where $c_1 = 0.5$ and $c_2 = 0.01$ are the weight coefficients recommended by the authors of PPO and confirmed experimentally.

In all experiments, the clip range parameter ε is set to 0.2, which corresponds to standard recommendations for the PPO algorithm.

This formalization ensures robust policy updates and stable model convergence in both 2D and 3D environments.

5. Research Results

5.1 Simulation Scenarios (2D and 3D)

1. Scenarios: To verify the effectiveness of the PPO algorithm, two experimental simulation scenarios with different spatial complexities were implemented:

Scenario 1 – 2D environment:

- grid size: 20×20;
- number of static obstacles: 30;
- agent start coordinates: (0, 0);
- target coordinates: (19, 19);
- the environment contains randomly placed fixed obstacles;
- goal: teach the agent to minimize the number of steps and avoid collisions.

Scenario 2 – 3D environment:

- cubic space dimensions: 15×15×8;
- obstacles: 50 (including vertical objects);
- starting position: (0, 0, 0);
- target position: (14, 14, 7);
- Objective: to test the stability of PPO as the action space and environment complexity increase.

In each scenario, 5,000 simulation episodes were run. During each run, the following metrics were collected: average reward, episode length, number of collisions, and average time to reach the goal.

5.2 Metrics

1. The average path length is defined as the average value of L , calculated using formula (2) for all simulation episodes.

2. Number of collisions.

3. Average reward per episode.

Visualization allows for a better analysis of the agent's spatial behavior and its interaction with the environment. The figures show:

– The drone's location – red marker;

– Targets – green dots;

– Obstacles – white blocks with an "x" inside.

Figure 1 demonstrates the drone's trajectory in 2D space, emphasizing its ability to navigate around obstacles and reach targets.

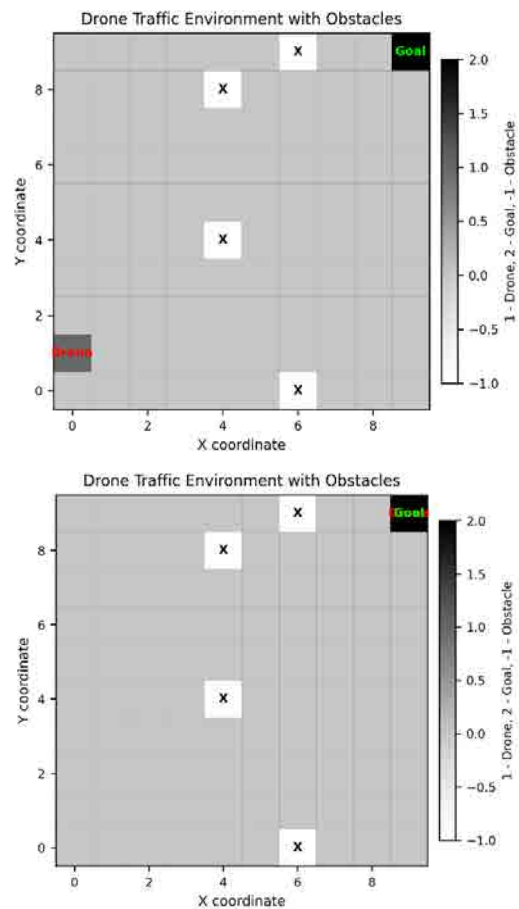


Fig. 1. Drone motion dynamics

The visualizations demonstrate the agent's evasion effectiveness and adaptability.

5.3 Results in a 2D environment

Training of the UAV adaptive route control model in a two-dimensional (2D) environment was performed

Table 2. Comparative average metrics for algorithms in a 2D environment (100 episodes)

Algorithm	Energy_Mean (J)	Mean_Reward	Collisions_Mean (count)	PathLength_Mean (m)
PPO_Custom	4.05	-18.89	4.21	4.63
PPO_base	1.71	-250.32	80.63	80.64
DQN	2.53	-253.64	81.57	81.57
A2C	5.28	-262.55	85.49	85.50

The Impact of the Learning Rate and Batch Size Hyperparameters on PPO Performance

During the training of models using the PPO algorithm, hyperparameters, particularly the learning rate and batch size, have a significant impact on the stability, speed, and quality of policy convergence [15].

A high learning rate (e.g., 0.01–0.1) implies rapid policy updates, which can be effective in simple environments.

A low value (0.0001–0.001) ensures stability but slows down adaptation. Similarly, a large batch size (256–1024) provides smoothness, while a small one (32–64) offers flexibility, though the risk of instability increases;

Thus, the coordinated tuning of PPO hyperparameters, particularly learning_rate and batch_size, plays a critical role in achieving stable and effective agent training. This is especially important when scaling from 2D to 3D environments with increased complexity in dynamics and spatial constraints.

Comparative Analysis Results (2D)

Calculation of percentage changes (PPO_custom vs PPO_base)

1) Energy Consumption

$$\Delta_{energy} = \frac{4.05 - 1.71}{1.71} \times 100\% \approx 136\%$$

PPO_custom uses 136% more energy than the base PPO.

2) Mean Reward Improvement

$$\Delta_{reward} = \frac{-18.89 - (-250.32)}{|-250.32|} \times 100\% \approx 92.5\%$$

The average reward increased by 92.5%.

3) Collision Reduction

$$\Delta_{collisions} = \frac{80.63 - 4.21}{80.63} \times 100\% \approx 94.8\%$$

The number of collisions decreased by 94.8%.

using the Proximal Policy Optimization (PPO) algorithm. The goal was to develop a stable policy capable of effectively avoiding obstacles and reaching targets in an environment with dynamic constraints. The main training configuration parameters are shown in Table 2.

4) Path Length Reduction

$$\Delta_{path} = \frac{80.64 - 4.63}{80.64} \times 100\% \approx 94.3\%$$

Path length decreased by 94.3%

Results of the 2D scenario

In a 2D environment, the modified PPO algorithm (PPO_custom) demonstrated the following improvements compared to the base version of PPO:

1. **+92.5%** increase in average reward (-18.9 vs. -250.3);

2. **-94.8%** reduction in the number of collisions (4.2 vs. 80.6);

3. **-94.3%** reduction in path length (4.6 vs. 80.6).

However, there is a **136% increase in energy consumption** (4.05 vs. 1.71), which requires further analysis and possible balancing of the reward environment.

5.4 Results in a 3D environment

The key differences between the reinforcement learning algorithm for 3D space and the base version are listed below.

The simulation environment models the drone's navigation process in three-dimensional space, taking into account obstacles and target coordinates. The main components of the model:

– state space – defined by formula (5);

– action space: $a = (\Delta x, \Delta y, \Delta z)$ – discrete displacements in three directions;

– environment: contains randomly generated obstacles that the drone must avoid.

Task Formalization and Key Problem

In three-dimensional space, the agent's progress is evaluated by the change in the Euclidean distance to the target, which is determined by formula (1).

This metric is used as the primary indicator of how close the agent is to the target in an environment with three degrees of freedom.

The reward function in a 3D environment retains the same structure as in 2D (Formula 8), where, d_t – current distance to the goal, $I_{collision} \in \{0, 1\}$ – collision indicator, $I_{goal} \in \{0, 1\}$ – goal attainment indicator, E_t – instantaneous energy expenditure of the agent at step t , $\alpha, \beta, \gamma, \delta$ – weight coefficients.

This formalization allows for a unified mechanism to evaluate progress in space, penalize dangerous maneuvers, and account for energy constraints. The 3D model architecture includes an extended space of states and actions:

- **State:** $(x, y, z, v, x_{goal}, y_{goal}, z_{goal})$
- **Action:** $(\Delta x, \Delta y, \Delta z)$
- **Objective:** Minimize the distance to the target while ensuring safety and energy efficiency.

Visualization

This section presents the results of a comparison of PPO performance in 2D and 3D environments.

5.5 Statistical Analysis of Performance

To confirm the model's training effectiveness under varying levels of complexity, a statistical analysis of key metrics was conducted: average reward, collision frequency, and episode duration. Comparisons between 2D and 3D environments were performed using mean values, standard deviation, confidence intervals, and the t-test. The results are visualized in Figures 2 and 3.

Figure 2 presents a comparison of reward distributions across environments. Figure 3 illustrates a training log file with a detailed representation of metric changes over time.

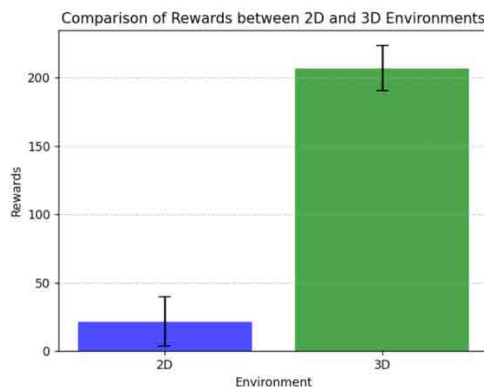


Fig. 2. Comparison of rewards between 2D and 3D spaces

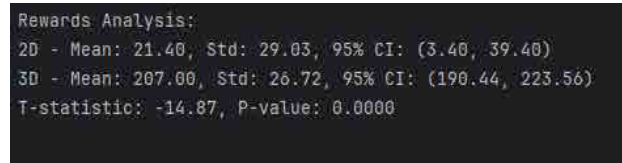


Fig. 3. Dynamics of reward changes during training depending on the environment type

Analyzing Figure 3, the following conclusions can be drawn:

1. **The average reward value** indicates a noticeable advantage of the 3D environment: in 2D, the model reached 21.40, while in 3D, it reached 207.0. This indicates significantly better policy adaptation in more complex conditions.

2. **The standard deviation** also demonstrates differences: in 2D – 29.03, in 3D – 26.72. The lower variance in 3D indicates more stable agent behavior.

3. **Confidence intervals (95%)** confirm the previous results: a wide interval in 2D (3.40–39.40) indicates high variability, while in 3D (190.44–223.56) it indicates greater statistical confidence.

4. The results of the t-test ($t = -14.87$, $p < 0.001$) indicate a statistically significant difference between the model's performance metrics in 2D and 3D environments, confirming the model's superiority in the more complex three-dimensional space with a significance level of $p < 0.05$.

The obtained data confirm the higher performance of PPO in a 3D environment, which is due to better adaptation to spatial complexity and flexibility of the navigation policy.

Thus, the PPO algorithm demonstrated not only a steady improvement in metrics during training but also a statistically confirmed advantage in the more complex three-dimensional environment. This demonstrates the potential of PPO for real-world air navigation scenarios under dynamic conditions.

Key comparison results

The model in the 2D environment demonstrated faster policy convergence, shorter episodes, and a higher average reward in the early stages of training. The smaller number of degrees of freedom ensured more stable collision dynamics and simpler trajectory behavior.

In the three-dimensional environment, during the initial training stages, the PPO algorithm was characterized by an increased number of collisions and longer episode durations, which is due to the increased

spatial complexity of the task. At the same time, a gradual policy adaptation and a reduction in the number of critical errors were observed during training.

Graphical Comparison

For a more visual analysis, the following metrics were presented in the graphs:

- **reward dynamics:** the graph shows that average rewards in 2D space increase faster compared to 3D;
- **collision frequency:** in 3D space, the collision frequency is higher in the early stages, but the model adapts over time;
- **episode length:** episodes in 3D space are typically longer, which is explained by the more complex trajectory.

Figure 4 illustrates the movement of drones in 2D space. The green dot represents the target's location, the red dots indicate where the drone reached the target, and the black blocks show the locations of obstacles. The graph demonstrates how effectively the algorithm avoids obstacles and reaches targets in a 2D environment.

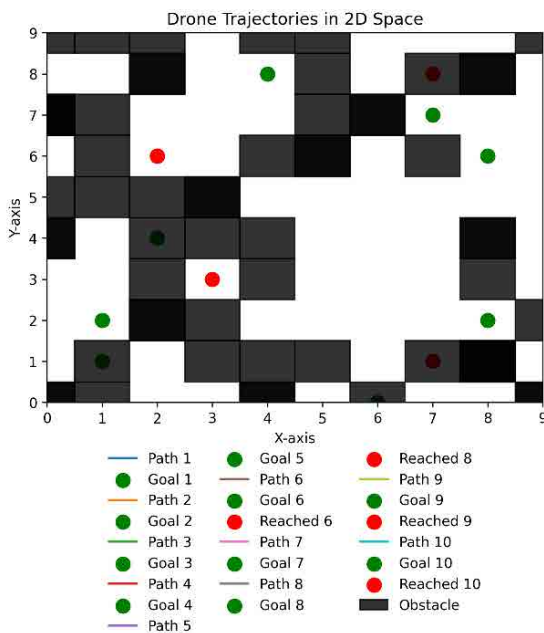


Fig. 4. Drone trajectories in 2D space

Figure 5 demonstrates movement in 3D space. Green dots mark the locations of targets, the red dot indicates where one of the drones reached the target, and the black dots represent obstacles in the space. The additional dimension complicates route planning, but the model demonstrates high adaptability in avoiding obstacles.

The generalized learning dynamics of the model in 2D and 3D environments are presented in Figures 6 and 7.

Drone Trajectories in 3D Space

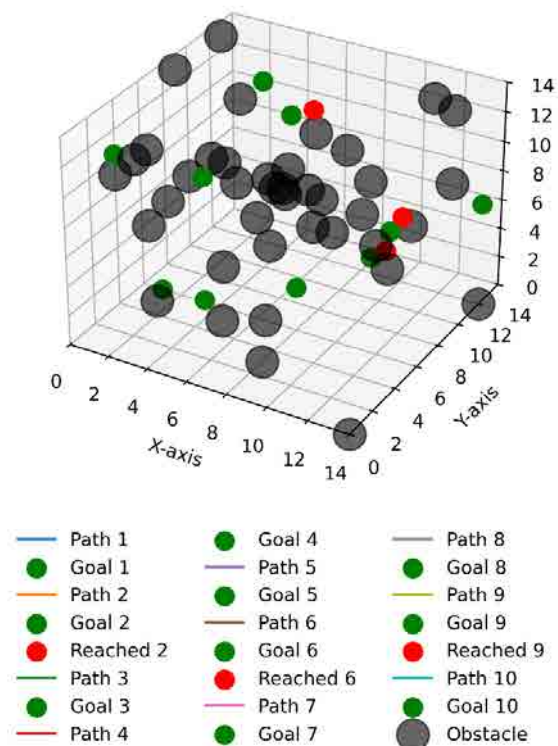


Fig. 5. Drone trajectories in 3D space

Figure 6 shows three main graphs for the 2D environment:

- the increase in the average reward per episode indicates the model's rapid adaptation to a simple topology;
- the decrease in the number of collisions confirms the effectiveness of the developed evasion policy;
- the stable duration of episodes reflects the stability of the agent's behavior.

Figure 7 shows similar graphs for the 3D environment. Here, a slower increase in rewards can be observed, which is due to a more complex trajectory structure and a greater number of obstacles. However, even under these conditions, the algorithm gradually reduces the number of collisions and stabilizes the episode length, indicating successful adaptation to the three-dimensional environment. Such visualization allows for a comprehensive assessment of learning effectiveness in both spaces and confirms the advantage of PPO in solving navigation problems of varying complexity.

Despite its complexity, the 3D approach offers broader practical applicability in real-world conditions, such as urban environments with tall buildings

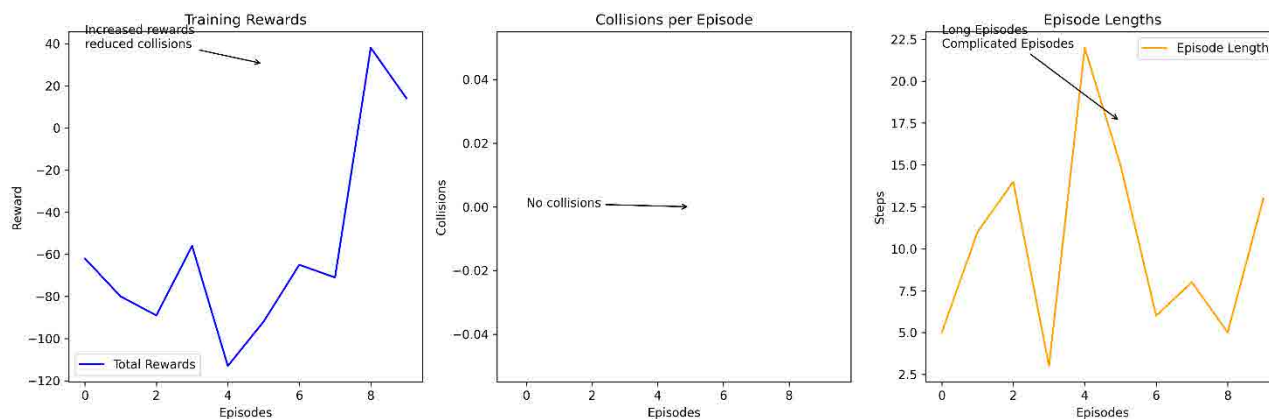


Fig. 6. Training dynamics for 2D space

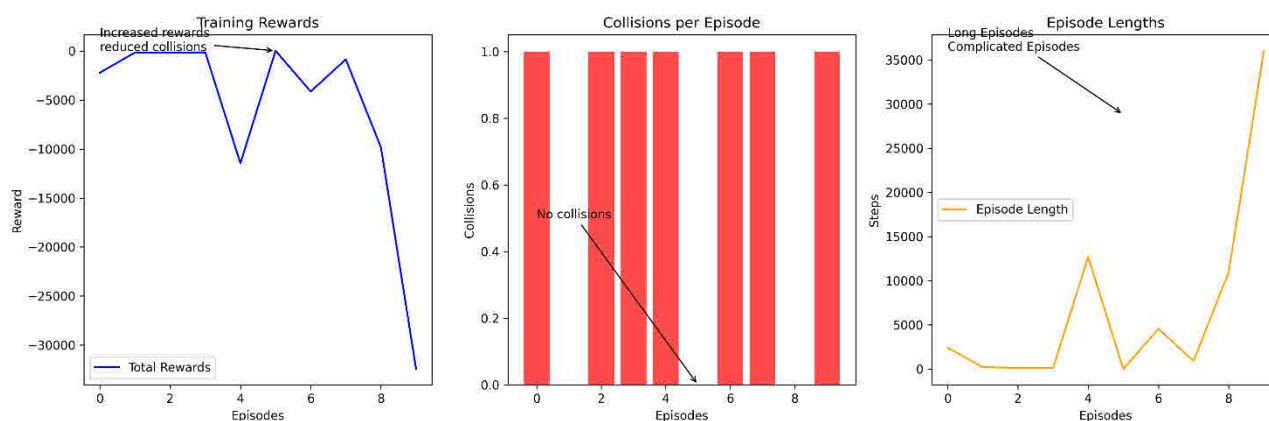


Fig. 7. Learning dynamics for 3D space

6. Discussion of Results

The results of the study confirm the effectiveness of PPO in adaptive routing tasks in both 2D and 3D environments. In 2D, the model achieved rapid convergence due to the lower complexity of the environment, whereas in 3D, greater variability was observed in the early stages. At the same time, as training progressed, the agent demonstrated stable behavior with a gradual decrease in the number of collisions. The modified reward function facilitated effective learning in both cases. A statistical t-test revealed a significant difference between the results, justifying the need to adapt strategies to each type of environment.

For the practical application of the proposed adaptive UAV route control algorithm, it is most appropriate to integrate it into existing Unmanned Traffic Management (UTM) concepts. Specifically:

– **NASA UTM (Unmanned Aircraft System Traffic Management)** aims to create an environment where multiple drones operate Beyond Visual Line of Sight (BVLOS) at altitudes up to 400 feet, coordinated through a suite of data exchange and route planning

services. The proposed PPO algorithm with an energy-based reward component can be implemented as a microservice within the Flight Information Management System (FIMS) module to dynamically update routes while accounting for obstacles and limited resources.

– **EU U-Space** regulates a set of automated services (Service Levels 1–4) for the safe access of large drone fleets to airspace. In particular, at the **U3 level (High-density UTM)**, the algorithm can provide adaptive real-time flight rerouting, reducing the risk of collisions and optimizing battery usage during peak load.

Thus, integrating the proposed modified PPO reward function into these UTM modules will allow for:

1. Ensure **proactive avoidance of obstacles** and airspace restrictions.
2. Dynamically **recalculate routes** based on the current battery charge level and remaining distance to the destination.
3. Optimize **cooperation with other participants** in the UTM ecosystem through standardized exchange protocols (REST, MQTT, AMQP).

This approach contributes to improving the reliability and efficiency of UAS operations in real-world UTM projects.

Limitations of the study. Agent training was conducted over 50,000 steps, which proved sufficient to demonstrate the convergence of the modified algorithm under the selected scenario conditions. However, we acknowledge that this may not be sufficient for more complex real-time navigation tasks (real-world applications). The results obtained should be considered a proof of concept, and further research will involve scaling up training to 10^6 epochs to improve the robustness of the policy.

7. Conclusions

Scientific novelty of the research. The scientific novelty of this work lies in the development of a combined reward function for the PPO algorithm, which simultaneously accounts for the dynamics of approach to the target, collision risk, and UAV energy consumption.

Unlike existing approaches, the proposed model provides a controlled trade-off between navigation safety and energy efficiency.

In addition, a unified 2D/3D simulation environment has been developed, allowing for a fair comparison of reinforcement learning algorithms in spaces of different dimensions. A comparative analysis and statistical verification of the results confirm the effectiveness of the approach and its advantages in three-dimensional navigation scenarios.

Main results of the study. In the course of this study, a simulation environment for UAV navigation in two- and three-dimensional spaces was developed, which takes into account the presence of static obstacles, target coordinates, and energy constraints on movement.

Based on this environment, a modified Proximal Policy Optimization (PPO) model with a combined reward function was implemented, combining incentives for approaching the target, penalties for collisions, and energy costs. A comparative analysis of the effectiveness of the PPO, DQN, and A2C algorithms in standardized 2D and 3D navigation scenarios demonstrated significant advantages of the proposed PPO model. Specifically, in the 2D environment, the number of collisions was reduced by 94.8%, and the route length by 94.3% compared to the baseline PPO version, while the average reward increased by 92.5%.

In three-dimensional space, similar trends toward improved safety and policy stability were confirmed, as evidenced by statistically significant results ($p < 0.05$).

Additionally, a relationship was established between the aggressiveness of obstacle-avoidance strategies and

energy consumption, allowing the agent's behavior to be interpreted as a trade-off between navigation safety and energy efficiency. The proposed visualization tools for trajectories and learning dynamics serve as a means of interpreting the agent's behavior and further confirm the convergence stability of the PPO algorithm in environments with varying spatial complexity.

Limitations of the study and prospects for further work. The main limitations of the study are the computational complexity of the reinforcement learning process, which requires significant resources as the environment's dimensionality increases, as well as the limited scalability of the proposed approach to multi-agent scenarios under real-world operating conditions. The results obtained should be considered within the framework of a simulation model with a single agent and static obstacles.

Further research should focus on expanding the model by integrating realistic airspace constraints, including meteorological factors, restricted zones, and variable traffic density, as well as on studying multi-agent interaction using multi-agent reinforcement learning approaches.

A promising direction is the automatic tuning of hyperparameters during training and the combination of the PPO algorithm with evolutionary methods or simulation-based learning methods to improve the stability and generalization ability of the policy.

Conflict of interest

The authors declare that they have no conflicts of interest, including financial, personal, authorial, or any other nature, that could influence the research or the results published in this article.

Funding

The study was conducted without financial support.

Data availability

Data will be provided upon reasonable request.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technology in the creation of this paper.

References

1. Debnath, D., Vanegas, F., Sandino, J., Hawary, A. F., Gonzalez, F. (2024), "A review of UAV path-planning algorithms and obstacle avoidance methods for remote sensing applications", *Remote Sensing*, Vol. 16 (21), 4019 p. DOI: <https://doi.org/10.3390/rs16214019>
2. Martins, F. G., Coelho, M. A. N. (2000), "Application of feedforward artificial neural networks to improve process control of PID-based control algorithms", *Computers & Chemical Engineering*, Vol. 24 (2-7). pp. 853-858. DOI: [https://doi.org/10.1016/S0098-1354\(00\)00339-2](https://doi.org/10.1016/S0098-1354(00)00339-2)
3. Liu, X., Peng, Z.R., Zhang, L.Y. (2019), "Real-time UAV rerouting for traffic monitoring with decomposition based multi-objective optimization", *Journal of Intelligent & Robotic Systems*, Vol. 94, pp. 491–501. DOI: <https://doi.org/10.1007/s10846-018-0806-8>
4. Almeida, E. N., Campos, R., Ricardo, M. (2022), "Traffic-aware UAV placement using a generalizable deep reinforcement learning methodology", *2022 IEEE Symposium on Computers and Communications (ISCC)*, pp. 1–6. DOI: <https://doi.org/10.48550/arXiv.2203.08924>
5. Madani, A., Engelbrecht, A. Ombuki-Berman, B., (2023), "Cooperative coevolutionary multi-guide particle swarm optimization algorithm for large-scale multi-objective optimization problems", *Swarm and Evolutionary Computation*. Vol. 82. 101262 p. DOI: <https://doi.org/10.1016/j.swevo.2023.101262>
6. Luo, J., Tian, Y., Wang, Z. (2024), "Research on unmanned aerial vehicle path planning". *Drones*, Vol. 8(2). 51 p. DOI: <https://doi.org/10.3390/drones8020051>
7. Li, C., Lian, J., (2007), "The Application of Immune Genetic Algorithm in PID Parameter Optimization for Level Control System", *Proceedings of the 2007 IEEE International Conference on Automation and Logistics (ICAL)*, Jinan, China, pp. 2670–2674. DOI: <https://doi.org/10.1109/ICAL.2007.4338670>
8. Yang, F., Lu, Q., Li, R., Xu, Y., Yuan, W., Wu, X. (2023), "Real-time optimal path planning and fast autonomous flight for UAV in unknown environments", *IEEE*. DOI: <https://doi.org/10.23919/CCC58697.2023.10240971>
9. Li, Q., Li, R., Ji, K., Dai, W. (2015), "Kalman filter and its application", *IEEE*. DOI: <https://doi.org/10.1109/ICINIS.2015.35>
10. Hooshyar, M., Huang, Y. (2023), "Meta-heuristic algorithms in UAV path planning optimization: A systematic review (2018–2022)", *Drones*, Vol. 7(12), 687 p. DOI: <https://doi.org/10.3390/drones7120687>
11. Li, H., Zhang, Z.-yu. (2012), "The application of immune genetic algorithm in main steam temperature of PID control of BP network", *Physics Procedia*, Vol. 25, pp. 80-86. DOI: <https://doi.org/10.1016/j.phpro.2012.02.013>
12. Zhang, M., Liu, Y., Wang, Y., Li, F., Chen, L. (2023), "Real-time path planning algorithms for autonomous UAV", *IEEE*. DOI: <https://doi.org/10.1109/CAC57257.2022.10054770>
13. Kim, D. H. (2003), "Comparison of PID controller tuning of power plant using immune and genetic algorithms", *The 3rd International Workshop on Scientific Use of Submarine Cables and Related Technologies*, Lugano, Switzerland, pp. 358-363. DOI: <https://doi.org/10.1109/CIMSA.2003.1227222>
14. Yena, M. (2024), "Optimizing air traffic control: Innovative approaches to collision avoidance in UAV operations", *Integrated Computer Technologies in Mechanical Engineering - 2023 (ICTM 2023)*. pp. 543–553. DOI: https://doi.org/10.1007/978-3-031-60549-9_41
15. Yena, M., & Pohudina, O. (2025), "Integrated simulation model of swarm control and adaptive routing of UAVS in a changing air environment", *Innovative technologies and scientific solutions for industries*, (4(34)), pp. 32-43. DOI: <https://doi.org/10.30837/2522-9818.2025.4.032>

Received (Надійшла) 23.11.2025

Accepted for publication (Прийнята до друку) 10.02.2026

Publication date (Дата публікації) 30.03.2026

Відомості про авторів / About the Authors

Єна Максим Вікторович – аспірант, Національний аерокосмічний університет «Харківський авіаційний інститут», кафедра «Інформаційних технологій проектування», Харків, Україна;

Maksym Yena – PhD Student, "Kharkiv Aviation Institute", Department of Information Technology Design, Kharkiv, Ukraine;

e-mail: yenamaksim98@gmail.com

ORCID ID: <https://orcid.org/0009-0006-0664-3244>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=59161031800>

Погудіна Ольга Костянтинівна – кандидат технічних наук, доцент, Національний аерокосмічний університет «Харківський авіаційний інститут», доцент кафедри інформаційних технологій проектування, Харків, Україна;

Olha Pohudina – Candidate of Technical Sciences, Associate Professor, National Aerospace University "Kharkiv Aviation Institute", Department of Information Technology Design, Kharkiv, Ukraine;

e-mail: olha.pohudina@poliba.it

ORCID ID: <https://orcid.org/0000-0001-5689-2552>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57204907264>

МОДЕЛЮВАННЯ АДАПТИВНОГО УПРАВЛІННЯ МАРШРУТАМИ БПЛА НА ОСНОВІ АЛГОРИТМІВ НАВЧАННЯ З ПІДКРІПЛЕННЯМ

Предмет дослідження – функція винагороди, політика дій та динаміка навчання алгоритму PPO у задачі адаптивної навігації БПЛА в умовах динамічного повітряного простору та обмежених енергетичних ресурсів. **Мета** – створення симуляційного середовища та розроблення модифікованої моделі Proximal Policy Optimization (PPO) для адаптивного управління маршрутом одиночного БПЛА у 2D та 3D просторах із урахуванням відстані до цілі, ризику зіткнень і енергоспоживання. **Завдання:** розробити 2D та 3D симуляційні середовища з різною структурою перешкод і параметрами руху БПЛА; сформувати комбіновану функцію винагороди PPO, що враховує відстань до цілі, зіткнення та енергоспоживання; Реалізувати та навчити алгоритми PPO, DQN і A2C у стандартизованих сценаріях навігації.; провести порівняльний аналіз ефективності алгоритмів за ключовими метриками (довжина маршруту, кількість зіткнень, винагорода, енергоспоживання; виконати статистичну перевірку результатів за допомогою t-тесту та довірчих інтервалів; проаналізувати вплив гіперпараметрів PPO на стабільність політики та збіжність навчання у 2D і 3D середовищах. **Методи:** використано алгоритми глибинного навчання з підкріпленням (PPO, DQN, A2C). Розроблено дві симуляційні моделі (2D та 3D) із випадковими статичними перешкодами. Сформовано комбіновану функцію винагороди, що включає динамічну компоненту зближення до цілі, штрафи за зіткнення та енергетичний термін. Ефективність моделей оцінювалася за середньою винагородою, довжиною маршруту, кількістю зіткнень та енергетичними витратами. Статистичну достовірність перевірено за допомогою t-тесту та 95% довірчих інтервалів. **Результати:** модифікована PPO-модель у 2D середовищі зменшила кількість зіткнень на 94,8% та довжину маршруту на 94,3% у порівнянні з базовою PPO, при цьому спостерігалось збільшення енергоспоживання через складніші маневри ухилення. У 3D середовищі підтверджено аналогічні тенденції: підвищення безпеки навігації, стабілізація політики та статистично значущі покращення ключових метрик ($p < 0,05$). **Висновки:** розроблено уніфіковане 2D/3D симуляційне середовище адаптивної маршрутизації БПЛА та модифіковану PPO-модель з комбінованою функцією винагороди, що враховує зближення до цілі, зіткнення та енергоспоживання. У 2D-середовищі досягнуто зменшення кількості зіткнень на $\approx 94,8\%$, скорочення довжини маршруту на $\approx 94,3\%$ та зростання середньої винагороди на $\approx 92,5\%$ порівняно з базовою PPO. У 3D-середовищі підтверджено аналогічні тенденції та статистично значущі покращення ($p < 0,05$). Встановлено залежність між агресивністю ухилення та енергоспоживанням, що дозволяє вибирати оптимальну політику для сценаріїв BVLOS.

Ключові слова: адаптивне управління; Proximal Policy Optimization; навчання з підкріпленням; імітаційне моделювання; маршрутизація; 3D-навігація.

Бібліографічні описи / Bibliographic descriptions

Єна М.В., Погудіна О.К. Моделювання адаптивного управління маршрутами БПЛА на основі алгоритмів навчання з підкріпленням. *Сучасний стан наукових досліджень та технологій в промисловості*. 2026. № 1 (35). С. 28–38. DOI: <https://doi.org/10.30837/2522-9818.2026.1.028>

Yena, M., Pohudina, O. (2026), "Modeling of adaptive UAV route control based on reinforcement learning algorithms", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (35), P. 28–38. DOI: <https://doi.org/10.30837/2522-9818.2026.1.028>

Коломійцев О. В., Гулевич М. В., Дмитрієв О. М., Левченко А. О., Балабуха О. С.

ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ ВДОСКОНАЛЕНИХ ЖАДІБНОГО Й ДЕЛЬТА-ДЕБАГІНГ АЛГОРИТМІВ ОПТИМІЗАЦІЇ ТЕСТОВИХ СЦЕНАРІЇВ ДЛЯ C++ БІБЛІОТЕК

Ефективна оптимізація тестових сценаріїв (ТС) є необхідною умовою для підвищення ефективності тестування C++ бібліотек. **Предметом дослідження** є алгоритми оптимізації ТС для C++ бібліотек. **Мета роботи** – підвищення ефективності тестування C++ бібліотек завдяки вдосконаленню класичного алгоритму жадібного формування (оптимізації) ТС і алгоритму дельта-дебагінг мінімізації ТС для C++ бібліотек. **Завдання.** Удосконалити жадібний алгоритм оптимізації ТС і усунути його детермінованість, що забезпечить ефективне стиснення ТС зі збереження гілкового покриття коду. Удосконалити алгоритм дельта-дебагінг мінімізації тестових сценаріїв для ТС із розрідженим розташуванням надлишкових дій. Оцінити ефективності вдосконалених алгоритмів оптимізації ТС для C++ бібліотек порівняно з класичними алгоритмами. У роботі застосовано такі **методи**: жадібний алгоритм оптимізації тестових сценаріїв, алгоритм дельта-дебагінг мінімізації ТС, математичне моделювання та методи статистичного аналізу. Ефективність удосконалених алгоритмів оцінено на основі статистичного аналізу 100 моделювань для кожного алгоритму на двох C++ бібліотеках із відкритим вихідним кодом різної структурної складності. **Результати** оцінювання ефективності вказують на те, що вдосконалені алгоритми забезпечують збереження (збільшення) гілкового покриття з коефіцієнтом до 1,058, підвищення коефіцієнта стиснення ТС до 0,86 та скорочення часу на виконання тестового набору (ТН) від 1,5 до 2,5 раза, якщо порівнювати з класичними алгоритмами. **Висновки.** Удосконалені алгоритми істотно зменшують час на тестування C++ бібліотек без втрати гілкового покриття коду. Запропоновано вдосконалений жадібний алгоритм оптимізації тестових сценаріїв, у якому вибір дій ТС здійснюється з огляду на їх майбутню користь, визначену за очікуваним приростом гілкового покриття коду. Такий підхід усуває детермінованість класичного жадібного алгоритму, підвищує інформативність тестових сценаріїв і забезпечує збалансованість між збереженням гілкового покриття та стисненням ТС. Запропоновано вдосконалений алгоритм дельта-дебагінгу, який виконує групове вилучення неунікальних дій ТС за внеском у гілкове покриття, що дає змогу суттєво знизити довжину тестових сценаріїв без втрати ефективності тестування.

Ключові слова: програмне забезпечення; тестовий сценарій; алгоритм; оптимізація; покриття; часові витрати; дефект.

1. Вступ

Якість програмного забезпечення (ПЗ) неможливо підтримувати без ефективного тестування. Проте в процесі розвитку й супроводження проєктів тестові набори (ТН) швидко зростають, накопичуючи надмірність. Така надмірність призводить до збільшення часу на їх виконання, підвищення обчислювальних витрат і ускладнення їх підтримки. Тому завдання оптимізації та мінімізації тестових сценаріїв (ТС) залишається ключовою з практичного й наукового погляду.

У роботі [1] запропоновано інструмент CIDER, призначений для автоматизації тестування C++ бібліотек. Аналіз методів автоматизованої генерації ТС [2] демонструє переваги еволюційних, стохастичних і навчальних методик оптимізації тестових сценаріїв.

Класичними підходами щодо оптимізації ТН є жадібні й дельта-дебагінг алгоритми. Жадібні алгоритми основані на покроковому відборі ТС, що забезпечують найбільший внесок у покриття й

використовуються як основна лінія для оцінювання сучасних методів і алгоритмів. Дельта-дебагінг алгоритм дає змогу скорочувати ТС завдяки поетапному вилученню частин ТС із перевіркою збереження необхідної поведінки. Крім стандартної версії, застосовуються варіації, які вилучають одразу групи тестових сценаріїв, що пришвидшує мінімізацію. Однак наявні варіанти залишаються схильними до детермінованості й не беруть до уваги динамічні властивості ТС.

Отже, удосконалення класичних жадібного й дельта-дебагінг алгоритмів оптимізації ТС для C++ бібліотек і оцінювання їх ефективності є актуальним науковим завданням.

2. Аналіз літературних джерел і визначення проблеми

2.1. Класичні методи оптимізації ТН

Задача мінімізації ТС (*Test Suite Minimization*, TSM) є однією з основних у галузі програмного тестування.

Вона полягає в пошуку піднабору тестових сценаріїв, який зберігає необхідне покриття або здатність виявляти дефекти ПЗ, проте водночас має менший розмір. Доведено, що задача TSM тісно пов'язана із задачею покриття множин і є NP-складною [3], що привело до появи різних евристичних і наближених алгоритмів.

Важливий систематичний огляд запропоновано в роботі [4], де узагальнено понад 113 досліджень, присвячених технікам скорочення ТН (*Test Suite Reduction*, TSR). Порівняно ефективність жадібних, кластеризаційних і еволюційних методів. Оцінено якість проведених досліджень. Запропоновано практичні рекомендації щодо відтворюваності результатів. Визначено стандарти оцінювання ефективності алгоритмів мінімізації, що застосовуються в більшості наукових робіт.

У систематичному огляді [5] проведено класифікацію методів регресійного тестування, зокрема з методами щодо селекції, пріоритизації та мінімізації ТН. Наголошено, що TSM має особливе значення у великих системах, де тестові набори можуть містити тисячі ТС.

У праці [6] емпірично оцінено техніки TSM, де продемонстровано, що прості евристичні алгоритми забезпечують скорочення розміру ТН на понад 70% у середньому. Спостерігається зниження здатності виявляти помилки приблизно на 12,5%, що обмежує практичну застосовність досліджених методів.

У роботі [7] окреслено використання еволюційних і локальних алгоритмів для скорочення ТН. Водночас продемонстровано, що, хоча еволюційні методи, зокрема генетичні алгоритми, демонструють високу ефективність, жадібні підходи також мають конкурентні результати в мультимодальному пошуковому просторі.

У дослідженні [8] запропоновано кластеризаційний метод щодо пріоритизації ТС, який формує групи сценаріїв за їх динамічною поведінкою під час виконання. Метод дає змогу скоротити кількість парних порівнянь і підвищує ефективність виявлення збоїв, якщо порівнювати з методами, орієнтованими на збереження покриття.

У статті [9] продемонстровано можливість інтеграції пошукових алгоритмів для промислових продуктових ліній, доведено, що поєднання відбору, пріоритизації та мінімізації забезпечує найкращий баланс між часом на виконання й рівнем покриття.

Розглянуті дослідження заклали концептуальне підґрунтя сучасних методів TSM, визначивши

оптимізацію покриття як багатокритеріальну задачу, у якій компроміс між точністю та швидкістю є ключовим чинником якості алгоритму.

2.2. Жадібні алгоритми оптимізації ТН

Жадібні алгоритми (*Greedy algorithms*) залишаються найпоширенішим інструментом для скорочення ТН. Основна ідея полягає в покроковому виборі ТС, які дають найбільший приріст критерію ефективності серед невідвіданих елементів.

У роботі [10] використано багаторівневу цільову функцію, яка дає змогу балансувати між різними цілями мінімізації, зокрема збереження (збільшення) покриття й усунення надмірності.

У статті [11] жадібні алгоритми оцінювалися за допомогою мутаційного тестування з огляду на кількість дефектів, що були виявлені сформованим тестовим набором. Ефективність жадібних алгоритмів підтверджено для регресійного тестування, де їх застосування забезпечує суттєве скорочення розміру ТН і часу на виявлення дефектів.

Зокрема результати в роботі [12] вказують на скорочення ТН на понад 75%, що демонструє практичну придатність жадібної пріоритизації для промислових програмних проєктів.

Крім того, у смарт-системах запропоновано шаблонно орієнтоване скорочення ТН для смарт-контрактів [13].

У дослідженні [14] проведено комплексне емпіричне оцінювання жадібних алгоритмів TSM за умови різних рівнів їх структурної складності. Доведено, що ефективність таких методів суттєво залежить від співвідношення кількості ТС і тестових вимог.

Систематичний огляд [15] підтверджує значущість жадібних алгоритмів порівняно з іншими евристичними й пошуковими методами.

Подальші дослідження зосереджені на вдосконаленні жадібних алгоритмів. Зокрема запропоновано прискорений варіант жадібного алгоритму (*Accelerated Greedy Additional Algorithm*, AGA), який підвищує ефективність пріоритизації ТН [16].

Більш нові роботи поєднують жадібні алгоритми з алгоритмами машинного навчання, наприклад, у скороченні ТН на основі подібності [17] або застосуванні механізмів часткової уваги для покращення відбору ТС [18].

Отже, жадібні алгоритми формують основу багатьох сучасних досліджень оптимізації ТН, і водночас є ключовими евристичними або складниками гібридних

методів, які інтегрують додаткові критерії покриття, подібності або ефективності. Роль у літературі засвідчує як стабільність результатів, так і гнучкість для адаптації в нових доменах тестування ПЗ.

Отже, жадібні алгоритми є надійною основою для методів оптимізації ТС і залишають простір для вдосконалень.

2.3. Дельта-дебагінг алгоритми мінімізації ТН

Алгоритм дельта-дебагінгу (*Delta Debugging*) вперше запропонований Zeller і Hildebrandt [19] для спрощення й локалізації вхідних даних, які викликають програмні збої. Згодом його почали застосовувати для мінімізації ТС. Класична версія алгоритму працює за принципом послідовного вилучення частин тестових сценаріїв із перевіркою щодо збереження необхідної поведінки.

У дослідженні [20] розширено цей алгоритм. Застосовано дельта-дебагінг для локалізації причин збоїв у складних програмних системах, що дає змогу визначати мінімальні підмножини інструкцій, відповідальних за помилку. Такий підхід є відправною точкою для наряду локалізації дефектів (*Fault Localization, FL*).

У статті [21] продемонстровано, що техніка програмного зрізання (*Program Slicing, PS*) має спільні з дельта-дебагінгом принципи редукції й може використовуватись для мінімізації ТС без втрати релевантних способів виконання. Такий зв'язок між програмним зрізанням і дельта-дебагінгом задав основу для гібридних методів редукції в сучасних системах тестування.

Класичний алгоритм дельта-дебагінгу набув подальшого розвитку в методі ієрархічного дельта-дебагінгу (*Hierarchical Delta Debugging, HDD*), який бере до уваги структурованість вхідних даних і застосовує послідовне спрощення на рівнях їх ієрархії. Такий підхід забезпечує значне прискорення мінімізації та покращує якість отриманих скорочених вхідних даних порівняно з основним алгоритмом [22].

У роботі [23] запропоновано алгоритм *Generalized Tree Reduction*, який поєднує принципи дельта-дебагінгу з ієрархічними перетвореннями деревоподібних структур. Запропонований алгоритм автоматично скорочує складні вхідні дані до кількох відсотків від їх початкового розміру, істотно перевершуючи класичний дельта-дебагінг за ефективністю мінімізації.

Автори праці [24] запропонували ймовірнісний підхід до дельта-дебагінгу (*Probabilistic Delta Debugging, ProbDD*), який поєднує класичний алгоритм *ddmin* із побудовою ймовірнісної моделі на основі абстрактного синтаксичного дерева. Метод бере до уваги синтаксичні зв'язки між елементами й результати попередніх випробувань, що скорочує середній час на оброблення на 27% і зменшує розмір результатів у 3,4 раза порівняно з наявними методами.

Серед сучасних модифікацій дельта-дебагінгу варто згадати алгоритм дельта-дебагінгу *Weighted Delta Debugging (WDD)* [25], у якому під час мінімізації береться до уваги вага елементів вхідних даних. Така модифікація дає змогу розрізняти фрагменти за їх внеском у структуру тесту, що підвищує ефективність алгоритмів *ddmin* і *ProbDD*.

Зокрема в середньому час на мінімізацію скорочується до 51%, а розмір результатів на 13% менший, якщо порівнювати з основними методами.

Ідею *hoisting*-операцій для HDD розглянуто в статті [26], а глибший теоретичний аналіз стохастичного дельта-дебагінгу виконано в роботі [27].

Дельта-дебагінг адаптовано для регресійного тестування на основі атрибутів [28].

Окремо варто згадати нові модифікації, спрямовані на підвищення ефективності дельта-дебагінгу. Зокрема в роботі [29] запропоновано MIP-DD, що є першим відкритим інструментом на основі дельта-дебагінгу для задач змішаного цілочисельного програмування (*Mixed Integer Programming, MIP*).

У роботі [30] сучасна модифікація дельта-дебагінгу основана на ідеї ітерації до фіксованої точки (DDMIN*). Автори продемонстрували алгоритм, який додатково скорочує вхідні дані на 48% порівняно з класичним DDMIN, а також підвищує стабільність процесу мінімізації та зменшує кількість зайвих ітерацій.

Крім того, сучасні дослідження вказують на те, що навчання з підкріпленням може стати логічним продовженням еволюції жадібних і дельта-дебагінг алгоритмів [31-33]. Зокрема в студіях [34, 35] застосовано RL для формування ефективних ТС на основі агента з Q-навчанням.

2.4. Визначення проблеми

Отже, попри значний розвиток методів TSM, низка ключових питань залишається нерозв'язаною. Класичні жадібні та дельта-дебагінг алгоритми довели ефективність у зменшенні розміру ТН і часу

на виконання ТС, однак залишаються детермінованими й не використовують адаптивних стохастичних механізмів. Більшість наявних модифікацій оптимізують лише структурне покриття, не зважаючи на семантичну значущість дій ТС або взаємозалежність базових блоків виконання C++ бібліотек у складі ПЗ, що знижує діагностичну цінність отриманих результатів для C++ бібліотек.

Новітні модифікації, як-от імовірнісний і зважений дельта-дебагінг, підвищують ефективність мінімізації. Однак вони ґрунтуються на фіксованих евристичних і не пристосовані до тестування C++ бібліотек із високою структурною складністю.

Крім того, більшість методів передбачають послідовне оброблення елементів, тоді як реальні ТС містять розсіяні або контекстно залежні дії, важливі для збереження повноти покриття.

3. Мета й завдання дослідження

Метою роботи є підвищення ефективності тестування C++ бібліотек завдяки вдосконаленню класичного алгоритму жадібного формування (оптимізації) ТС і алгоритму дельта-дебагінг мінімізації ТС для C++ бібліотек.

Для досягнення поставленої мети в роботі сформульовано задачу формування (оптимізації) ТС для C++ бібліотек та визначено такі часткові наукові завдання:

1) удосконалення жадібного алгоритму оптимізації ТС та усунення його детермінованості із забезпеченням ефективного стиснення ТС для збереження гілкового покриття коду;

2) оптимізація алгоритму дельта-дебагінг мінімізації тестових сценаріїв для ТС із розрідженим розташуванням надлишкових дій;

3) оцінювання ефективності вдосконалених алгоритмів оптимізації ТС для C++ бібліотек порівняно з класичними алгоритмами.

З метою оцінювання ефективності необхідно попередньо визначити критерії ефективності алгоритмів формування (оптимізації) ТС для C++ бібліотек.

4. Матеріали й методи дослідження

Задачу формування ТС для C++ бібліотек без специфікацій API можна формалізувати як марковський процес прийняття рішень (*Markov Decision Process*, MDP) [36], де ТС будується крок за кроком,

обираючи наступну функцію з допустимого простору дій (API) у такий спосіб:

$$\langle S, A, \mathcal{P}(s'|s, a), \mathcal{R}(s, a), \gamma \rangle, \quad (1)$$

де S – множина станів, що відповідає стану s ; A – множина допустимих дій у стані s ; $\mathcal{P}(s'|s, a)$ – функція ймовірності переходу до нового стану після виконання дії a у стані s ; $\mathcal{R}(s, a)$ – функція винагороди, яка надається за дію a в стані s ; $\gamma \in (0, 1)$ – коефіцієнт дисконтування.

Задача формування ТС описується таким чином. ТС довжини $n \in \mathbb{N}$, який формується, може бути впорядкованою послідовністю дій ТС:

$$S = (a_1, a_2, \dots, a_n), \quad (2)$$

де $a_i \in A$ – виклики API-бібліотеки з множини A .

Щоб керувати формуванням ТС, додатково використовується покриття базових блоків (*basic-block coverage*) як евристика. Для кожної дії a , яка виконується у префіксі сценарію $S_{1:t}$, де $t \in \mathbb{N}$ – це поточний крок формування ТС, визначається множина блоків $BB(S_{1:t})$. За умови, якщо остання дія відкрила блок, який ніколи не виконувався, то дія відкриває унікальний базовий блок виконання програми. Індикатор унікальності можна записати таким чином:

$$u_t(a) = \begin{cases} 1, & \text{якщо } BB(S_{1:t}) / BB(S_{1:t-1}) \neq \emptyset \\ 0, & \text{якщо } BB(S_{1:t}) / BB(S_{1:t-1}) = \emptyset \end{cases} \quad (3)$$

Миттєва корисність дії a на кроці t визначається за зростанням покриття гілок:

$$g_t(a) = Cov(S_{1:t} \oplus a) - Cov(S_{1:t}). \quad (4)$$

Тобто миттєва корисність дорівнює кількості нових гілок, які відкриються завдяки додаванню дії a на поточному кроці.

Задача мінімізації довжини ТС за умови збільшення (збереження) гілкового покриття. Необхідно знайти найкоротший тестовий сценарій S' на основі оригінального ТС S_{orig} , який задовольняє задану величину покриття коду:

$$\begin{cases} \min_{S' \in \mathcal{A}} Len(S') \\ Cov(S') \geq Cov(S_{orig}) \end{cases} \quad (5)$$

Розв'язати задачу (5) можна за допомогою формування нового тестового сценарію або мінімізації оригінального ТС за умови збереження оригінального покриття коду.

4.2. Удосконалення жадібного алгоритму формування ТС

Одним із найбільш поширених підходів щодо формування або мінімізації ТС є використання жадібних алгоритмів. Основна ідея полягає в тому, щоб на кожному кроці обирати дію, яка забезпечує найбільший безпосередній приріст покриття. У класичній постановці це покриття може обчислюватися за різними критеріями. Наприклад, за кількістю нових вимог, які задовольняються, або – нових елементів коду, що активуються.

На кожному кроці t жадібного алгоритму – обчислюється миттєва корисність кожної, ще не використаної дії за формулою (3). Далі обирається дія з максимальною миттєвою корисністю:

$$a' = \arg \max_{a \in A_t} g_t(a). \quad (6)$$

За умови, якщо $g_t(a') = 0$, то алгоритм завершує роботу, оскільки жодна з дій не дає змоги розширити гілкове покриття. В іншому разі обрана дія додається до сценарію та процес продовжується до досягнення оригінального покриття або до зупинки за іншим критерієм (наприклад, обмеження на довжину ТС). Псевдокод класичного жадібного алгоритму подано на рис. 1.

Попри свою простоту й практичну ефективність, класичний жадібний алгоритм має низку обмежень, а саме:

– *локальні плато* – якщо жодна з доступних дій не дає приросту покриття безпосередньо на поточному кроці, алгоритм зупиняється. Водночас він може втратити можливість досягти вищого покриття, яке відкривається лише після виконання певної "проміжної" дії;

– *детермінізм і відсутність диверсифікації* – алгоритм завжди обирає одну найкращу дію за миттєвим приростом, що робить його схильним до потрапляння в пастку локальних максимумів і обмежує різноманітність побудованих ТС;

– *ігнорування додаткової інформації* – класичний жадібний алгоритм не використовує жодної допоміжної інформації про потенційно корисні дії. Наприклад, він не бере до уваги, що деякі дії можуть відкривати нові базові блоки виконання програми, які згодом призведуть до зростання покриття гілок.

Класичний жадібний алгоритм обмежений тим, що на кожному кроці він відбирає лише дію з максимальним миттєвим приростом покриття гілок.

Як наслідок, якщо всі доступні дії мають нульовий приріст, то алгоритм зупиняється, навіть якщо деякі з них могли б відкрити нові можливості для майбутнього покриття. Таке обмеження призводить до стагнації на локальних плато.

Алгоритм 1.

Класичний жадібний алгоритм формування ТС.

Вхідні дані:

S_{orig} – оригінальний ТС.

Вихідні дані:

S_{out} – сформований ТС.

Тіло алгоритму:

1. $S_{out} \leftarrow \emptyset$
2. $A \leftarrow \text{Extract}(S_{orig})$
3. $Cov_{cur} \leftarrow \emptyset$
4. $Cov_{orig} \leftarrow Cov(S_{orig})$
5. **repeat**
6. $Gain_{best} \leftarrow 0$
7. $a' \leftarrow \text{None}$
8. **foreach** $a \in A \setminus S_{out}$ **do**
9. **local** $g(a) \leftarrow Cov(S_{out} \oplus a) - Cov_{cur}$
10. **if** $g(a) > Gain_{best}$ **then**
11. $Gain_{best} \leftarrow g(a)$
12. $a' \leftarrow a$
13. **end if**
14. **if** $a' = \text{None}$ **or** $Gain_{best} = 0$ **then**
15. **break**
16. **end if**
17. **end foreach**
18. $S_{out} \leftarrow S_{out} \oplus a'$
19. $Cov_{cur} \leftarrow Cov(S_{out})$
20. **until** $Cov_{cur} < Cov_{orig}$
21. **return** S_{out}

Рис 1. Псевдокод жадібного алгоритму формування ТС

Для подолання цього обмеження запропоновано вдосконалений жадібний алгоритм, що комбінує такі два параметри:

- 1) миттєвий приріст гілкового покриття $g_t(a)$;
- 2) майбутня корисність дії $f_t(a)$, яка оцінюється на основі появи нових унікальних базових блоків після виконання цієї дії.

Сутність удосконалення полягає в тому, що, якщо дія не збільшує покриття гілок у певний час, то вона може стати "ключем", після якого інші дії починають давати приріст. На кожному кроці вдосконалений алгоритм оцінює зважений бал дії:

$$F_t(a) = g_t(a) + \lambda \cdot f_t(a), \quad (7)$$

де $\lambda \in [0;1]$ – вага, що визначає пам'ять про майбутню корисність дій ТС.

За умови, якщо виконання $S_{1,t} \oplus a$ породжує нові базові блоки, що не з'являлись раніше, то отримується індикатор унікальності дії (3) – $u_t(a) \in \{0,1\}$. Припускається, що після дії a з'явилась деяка множина дій:

$$\mathcal{N}(a) = \{b \in \mathcal{A} | u_t(b) = 1\}, \quad (8)$$

де $u_t(b) = 1$ означає, що дія a є "ключем".

Для таких дій зберігається майбутня корисність дії $f_t(a)$, яка оновлюється за таким правилом:

$$f_t(a) \leftarrow \max(f_t(a), \max_{b \in \mathcal{N}(a)} F_t(b)). \quad (9)$$

Отже, обрано оновлення майбутньої корисності за принципом максимального значення, оскільки для задачі формування ТС важливим є факт, що дія хоча б один раз відкрила нові унікальні базові блоки виконання програми. Використання максимального значення гарантує збереження найсильнішого спостереженого впливу дії, робить алгоритм більш стійким і спрощує його збіжність без потреби в додаткових параметрах. На кожній ітерації алгоритму формується множина кандидатів $\mathcal{A}_{cand} \in \mathcal{A}$, яка містить усі дії a , для яких виконується $u_t(a) = 1$ або $F_t(a) > 0$. З метою диверсифікації алгоритму можна застосовувати формулу *softmax* для вибору такої дії:

$$P_t(a) = \frac{e^{\frac{F_t(a)}{\tau}}}{\sum_{a^* \in \mathcal{A}_{cand}} e^{\frac{F_t(a^*)}{\tau}}}, \quad (10)$$

де τ – коефіцієнт, який керує балансом між жадібним вибором і дослідженням майбутньої корисності дій ТС.

Псевдокод удосконаленого жадібного алгоритму подано на рис. 2.

Удосконалений жадібний алгоритм бере до уваги майбутню користь дій і здатний перевищити класичний жадібний алгоритм у задачі формування ТС. Завдяки збереженню інформації про дії, що відкривають доступ до нових унікальних станів, алгоритм має більший потенціал покриття гілок і очікувано формує коротші й ефективніші сценарії, ніж класичний алгоритм, що орієнтується лише на миттєвий приріст покриття.

Алгоритм 2.

Жадібний алгоритм формування ТС з огляду на майбутню користь дій ТС

Вхідні дані:

S_{orig} – оригінальний ТС;

λ – вага майбутньої корисності;

τ – коефіцієнт *softmax* вибору дії;

\cup – кількість кращих кандидатів для вибору.

Вихідні дані:

S_{out} – сформований ТС.

Тіло алгоритму:

1. $S_{out} \leftarrow \emptyset$, $\mathcal{A} \leftarrow \text{Extract}(S_{orig})$
2. $Cov_{orig} \leftarrow Cov(S_{orig})$
2. **foreach** $a \in \mathcal{A}$ **do** $f_t(a) \leftarrow 0$ **end**
3. **repeat**
4. $\mathcal{A}_{cand} \leftarrow \emptyset$
5. **foreach** $a \in \mathcal{A} \setminus S_{out}$ **do**
6. $g_t(a) \leftarrow Cov(S_{out} \oplus a) - Cov(S_{out})$
7. **if** $g_t(a) > 0$ **or** $f_t(a) > 0$ **then**
8. $F_t(a) = g_t(a) + \lambda \cdot f_t(a)$
9. $\mathcal{A}_{cand} \leftarrow \mathcal{A}_{cand} \cup \{a\}$
10. **end if**
11. **end foreach**
12. **if** $\mathcal{A}_{cand} = \emptyset$ **then break**
13. **foreach** $best \cup candidates a \in \mathcal{A}_{cand}$ **do**
14. $P_t(a) \leftarrow \text{Equation_10}(a, \tau)$
15. **end foreach**
16. $a' \sim P_t(a)$
17. $S_{out} \leftarrow S_{out} \oplus a'$
18. **foreach** $b \in \mathcal{A} \setminus S_{out}$ **do**
19. $F_t(b) \leftarrow g_t(b) + \lambda \cdot f_t(b)$
20. **if** $u_t(b) = 1$ **and** $F_t(b) > 0$ **then**
21. $f_t(a) \leftarrow \max(f_t(a), F_t(b))$
22. **end if**
23. **end foreach**
24. **until** $Cov(S_{out}) < Cov_{orig}$
25. **return** S_{out}

Рис. 2. Псевдокод жадібного алгоритму формування ТС з огляду на майбутню користь дій

4.3. Удосконалення алгоритму дельта-дебагінг мінімізації тестових сценаріїв

Дельта-дебагінг є одним з найпоширеніших алгоритмів мінімізації ТС. Його сутність полягає в поступовому вилученні частин ТС і перевірці щодо збереження бажаної властивості, відтворення

помилки або покриття коду. За умови, якщо скорочений ТС залишається валідним, то вилучені кроки вважаються надлишковими й відкидаються.

Псевдокод класичного алгоритму запропоновано на рис. 3, а додаткова процедура до алгоритму – на рис. 4.

Алгоритм 3.

Класичний алгоритм дельта-дебагінг мінімізації ТС

Вхідні дані:

- S_{orig} – оригінальний ТС.

Вихідні дані:

- S_{out} – мінімізований ТС.

Тіло алгоритму:

1. $Cov_{orig} \leftarrow Cov(S_{orig})$
2. $S_{out} \leftarrow S_{orig}$
3. $g \leftarrow 2$
4. **while** $Len(S_{out}) \geq 2$ **do**
5. $\{S_1, \dots, S_n\} \leftarrow PartitionContinous(S_{out}, g)$
6. $local\ success \leftarrow false$
7. **for** $i \leftarrow 1$ **to** g **do**
8. $S_{trial} \leftarrow S_{out} \setminus S_i$
9. **if** $Cov(S_{trial}) \geq Cov_{orig}$ **then**
10. $S_{out} \leftarrow S_{trial}$
11. $g \leftarrow \max(g-1, 2)$
12. $success \leftarrow true$
13. **break**
14. **end if**
15. **end for**
16. **if** $success$ **then continue**
17. **for** $i \leftarrow 1$ **to** g **do**
18. $S_{trial} \leftarrow S_i$
19. **if** $Cov(S_{trial}) \geq Cov_{orig}$ **then**
20. $S_{out} \leftarrow S_{trial}$
21. $g \leftarrow \max(g-1, 2)$
22. $success \leftarrow true$
23. **break**
24. **end if**
25. **end for**
26. **if** $success$ **then continue**
27. **if** $g \geq Len(S_{out})$ **then**
28. **break**
29. **else**
30. $g \leftarrow \min(2 \cdot g, Len(S_{out}))$
31. **end if**
32. **end while**
33. **return** S_{out}

Рис. 3. Псевдокод класичного алгоритму дельта-дебагінг мінімізації ТС

Процедура 1.

Процедура розбиття ТС на послідовні блоки дій
PartitionContinous

Вхідні дані:

- S – послідовність дій;
- n – бажана кількість блоків.

Вихідні дані:

- $\{S_1, \dots, S_n\}$ – суміжні блоки без перетинів.

Тіло процедури:

1. $Len_s \leftarrow Len(S)$
2. $n \leftarrow \min(n, Len_s)$
3. $Len_{base} \leftarrow floor(Len_s / n)$
4. $Len_{rem} \leftarrow mod(Len_s / n)$
5. $base \leftarrow 1$
6. **for** $i \leftarrow 1$ **to** n **do**
7. $local\ size \leftarrow base + (i \leq Len_{rem} ? 1 : 0)$
8. $S_i \leftarrow Cut(S, base, -1]$
9. $base \leftarrow base + size$
10. **end for**
11. **return** $\{S_1, \dots, S_n\}$

Рис. 4. Псевдокод процедури розбиття ТС на послідовні блоки дій ТС

Алгоритм дельта-дебагінг працює за принципом поділу на підмножини. Тоді нехай вхідний ТС містить n дій. Алгоритм починає з розбиття множини дій на два блоки й перевіряє, чи зберігається покриття після вилучення одного з них.

За умови, якщо так, то відкидає блок. Якщо ні, тоді ділить на дрібніші частини (наприклад, чотири блоки) та повторює процес.

Отже, алгоритм гарантує поступову мінімізацію ТС до підмножини, яка забезпечує необхідний результат.

Для мінімізації ТС запропоновано вдосконалений варіант алгоритму дельта-дебагінгу, який поєднує класичне вилучення елементів із стратегією множинного вилучення дій та врахуванням унікальності базових блоків.

Отже, для кожної дії a_i визначається множина унікальних базових блоків програми:

$$U(a_i) = \left\{ q \in B(a_i) \mid q \notin \bigcup_{j \neq i} B(a_j) \right\}, \quad (11)$$

де $B(a_i)$ – множина базових блоків програми, покритих дією a_i .

За умови, якщо $U(a_i) \neq \emptyset$, то дія позначається як "захищена" й не потрапляє до кандидатів на

вилучення поточної ітерації алгоритму. На кожній ітерації алгоритм формує підмножини дій без захищених елементів. Підмножини, вилучення яких не зменшує покриття нижче, ніж базовий рівень, вилучаються остаточно.

Псевдокод алгоритму запропоновано на рис. 5.

Процедуру формування множини захищених дій подано на рис. 6.

Алгоритм 4.

Алгоритм дельта-дебагінг мінімізації тестових сценаріїв з огляду на унікальність покриття базових блоків програми.

Вхідні дані:

- S_{orig} – оригінальний ТС;
- $g_0 \in (0,1)$ – коефіцієнт зернистості;
- g_{min} – мінімальна зернистість.

Вихідні дані:

- S_{out} – мінімізований ТС.

Тіло алгоритму:

1. $Cov_{orig} \leftarrow Cov(S_{orig})$
2. $S_{out} \leftarrow S_{orig}$
3. $g \leftarrow g_0$
4. $st \leftarrow \max(Len(S_{out}), g, g_{min})$
5. **while** $st \geq 1$ **do**
6. $P \leftarrow ComputeProtected(S_{out})$
7. **local** $i \leftarrow 1$
8. **while** $i \leq Len(S_{out})$ **do**
9. **local** $S_{batch} \leftarrow \emptyset$
10. $j \leftarrow i$
11. **while** $Len(B) < step$ **and** $j \leq Len(S_{out})$ **do**
12. **if** $j \notin P$ **then**
13. $S_{batch} \leftarrow B \cup \{a_j\}$
14. **end if**
15. $j \leftarrow j + 1$
16. **end while**
17. $S_{trial} \leftarrow S_{out} \setminus S_{batch}$
18. **if** $Cov(S_{trial}) \geq Cov_{orig}$ **then**
19. $S_{out} \leftarrow S_{trial}$
20. $P \leftarrow ComputeProtected(S_{out})$
21. **else**
22. $i \leftarrow j$
23. **end if**
24. **end while**
25. $st \leftarrow \max(st / 2, g_{min})$
26. **end while**
27. **return** S_{out}

Рис. 5. Псевдокод алгоритму дельта-дебагінг мінімізації ТС з огляду на унікальність покриття базових блоків програми

Алгоритм 4 керується $g \in \mathbb{R}$ – коефіцієнтом зернистості, що визначає кількість підмножин, на які розбивається вхідна послідовність дій.

На початкових ітераціях алгоритму це дає змогу виконувати більш грубі спроби вилучення, а в разі невдачі коефіцієнт зернистості поступово зростає, забезпечуючи більш дрібнозернисте розбиття. Такий підхід дає змогу балансувати між швидкістю мінімізації та її ефективністю.

Процедура 2.

Процедура формування множини дій за унікальним внеском у покриття базових блоків програми *ComputeProtected*

Вхідні дані:

- S – послідовність дій.

Вихідні дані:

- P – множина "захищених" дій.

Тіло процедури:

1. $P \leftarrow \emptyset$
2. **for** $j=1$ **to** $Len(S)$ **do**
3. **if** $U(a_j) \neq \emptyset$ **then**
4. $P \leftarrow P \cup \{a_j\}$
5. **end if**
6. **end for**
7. **return** P

Рис. 6. Псевдокод процедури формування множини "захищених" дій

Запропонований алгоритм має потенційні переваги над класичним дельта-дебагінгом у задачі формування ТС для C++ бібліотек завдяки адаптивному керуванню значенням коефіцієнта g врахуванню унікальності дій ТС за допомогою обчислення покриття базових блоків виконання програми. Використання змінної зернистості дає змогу спочатку швидко відсіювати великі непотрібні фрагменти ТС, а потім поступово переходити до більш точного локального аналізу, що забезпечує баланс між швидкістю мінімізації та якістю кінцевого результату.

Додаткове врахування дій із унікальним внеском у покриття базових блоків запобігає втраті критичних елементів ТС. Така комбінація дає змогу досягати високого ступеня стиснення ТС без зниження рівня покриття, особливо в разі довгих і надлишкових ТС.

4.4. Критерії ефективності алгоритмів оптимізації ТС

ТН C++ бібліотеки може бути поданий як впорядкована послідовність тестових сценаріїв:

$$TS = (S_1, S_2, \dots, S_m), \quad (12)$$

де m – кількість ТС у ТН.

Для оцінювання ефективності алгоритмів оптимізації ТС використано три основні метрики [4].

1. Коефіцієнт збереження покриття (*Retention of Coverage, RC*) – оцінює приріст гілкового покриття ТН, що відтворює втрату (збереження) ефективності тестування:

$$RC = \frac{Cov(TS_{out})}{Cov(TS_{orig})}, \quad (13)$$

де $Cov(TS_{out})$ – покриття ТН після оптимізації;
 $Cov(TS_{orig})$ – покриття оригінального ТН.

2. Коефіцієнт стиснення (*Compression Ratio, CR*) – оцінює зменшення кількості інструкцій ТН:

$$CR = \frac{Len(TS_{orig}) - Len(TS_{out})}{Len(TS_{orig})}, \quad (14)$$

де $Len(TS_{orig})$ – кількість інструкцій оригінального ТН; $Len(TS_{out})$ – після оптимізації.

3. Коефіцієнт збереження часу (*Execution Cost Reduction, ECR*) – оцінює зменшення часу на виконання ТН:

$$ECR = \frac{T(TS_{orig}) - T(TS_{out})}{T(TS_{orig})}, \quad (15)$$

де $T(TS_{orig})$ – час на виконання оригінального тестового набору; $T(TS_{out})$ – час на виконання ТН після оптимізації.

5. Результати дослідження та їх обговорення

5.1. Налаштування середовища оцінювання

Для оцінювання ефективності вдосконалених алгоритмів обрано дві C++ бібліотеки з відкритим кодом, які різняться за масштабом і структурною складністю.

Основні властивості подано в табл. 1.

Таблиця 1. Властивості C++ бібліотек, обраних для дослідження

Бібліотека	API	LoC	m	IC	BC %	Avg. CCN
<i>Bitmap PlusPlus</i>	33	760	7	71	27,0	2,5
<i>Hjson</i>	122	3936	57	1259	36,1	5,4

У табл. 1 використані такі позначки:

- API – кількість публічних функцій;
- LoC – кількість рядків у вихідному коді;

- m – кількість ТС у ТН;
- IC – загальна кількість інструкцій у ТН;
- BC – гілкове покриття ТН;
- Avg. CCN – середня цикломатична складність функцій бібліотеки.

Оцінку виконано на робочій станції з процесором Intel Core i7 (6 ядер, 2.6 ГГц) та 16 Гб оперативної пам'яті. Компіляцію бібліотек і ТС здійснено за допомогою AppleClang 15.0.0 (LLVM 15) з увімкненими параметрами профілювання покриття `-fprofile-instr-generate fcoverage-mapping`.

Для збору та аналізу покриття використано утиліту

`llvm-profdata` та `llvm-cov`.

Крім того, застосовано механізм збору інформації для ідентифікації базових блоків, що активовані під час виконання кожної дії ТС:

`__llvm_profile_instrument_target`.

Зібрані дані перетворено у двійкові вектори покриття, які асоціюються з відповідними діями ТС.

Для аналізу кількості рядків коду й цикломатичної складності використано утиліту `lizard`.

5.2. Конфігурації досліджуваних алгоритмів

Для порівняльного аналізу використано дві групи конфігурацій алгоритмів і два базові алгоритми:

- конфігурації вдосконаленого жадібного алгоритму формування ТС (GRR-TD1, GRR-TD2, GRR-TD3), що реалізують алгоритм 2;
- конфігурації вдосконаленого алгоритму дельта-дебагінгу оптимізації ТС (DSL-FM1, DSL-FM2, DSL-FM3), що реалізують алгоритм 4;
- конфігурація класичного жадібного алгоритму формування ТС (GR), що реалізує алгоритм 1;
- конфігурація класичного алгоритму дельта-дебагінг мінімізації ТС (DSL), що реалізує алгоритм 3.

Для кожної конфігурації алгоритму проведено 100 незалежних математичних моделювань послідовно для кожного тестового сценарію з ТН цільових C++ бібліотек.

Конфігурації вдосконаленого жадібного алгоритму формування ТС подано в табл. 2.

У табл. 2 значення параметра τ у механізмі *softmax* вибору дій зафіксовано на рівні 1,5, що обґрунтовано компромісом між жадібністю та можливістю дослідити вплив майбутньої корисності.

Конфігурації вдосконаленого алгоритму дельта-дебагінг мінімізації ТС подано в табл. 3.

Таблиця 2. Конфігурації вдосконаленого жадібного алгоритму оптимізації TC

Конфігурація	Вага майбут. корисності, λ	Розмір множ. канд. ν	Параметр <i>Softmax</i> , τ
GRR-TD1	0,35	3	1,5
GRR-TD2	0,65	4	1,5
GRR-TD3	0,9	5	1,5

Таблиця 3. Конфігурації вдосконаленого алгоритму дельта-дебаїнг мінімізації TC

Конфігурація	Коефіцієнт зернистості, g_0	Мінімальна зернистість, g_{min}
DSL-FM1	0,20	5
DSL-FM2	0,25	3
DSL-FM3	0,30	1

5.3. Аналіз збереження покриття TH

На рис. 7 зображено розподіл показників гілкового покриття коду залежно від конфігурації алгоритму оптимізації TC для C++ бібліотек *BitmapPlusPlus* та *Hjson*. Пунктирна лінія позначає рівень покриття оригінального TH.

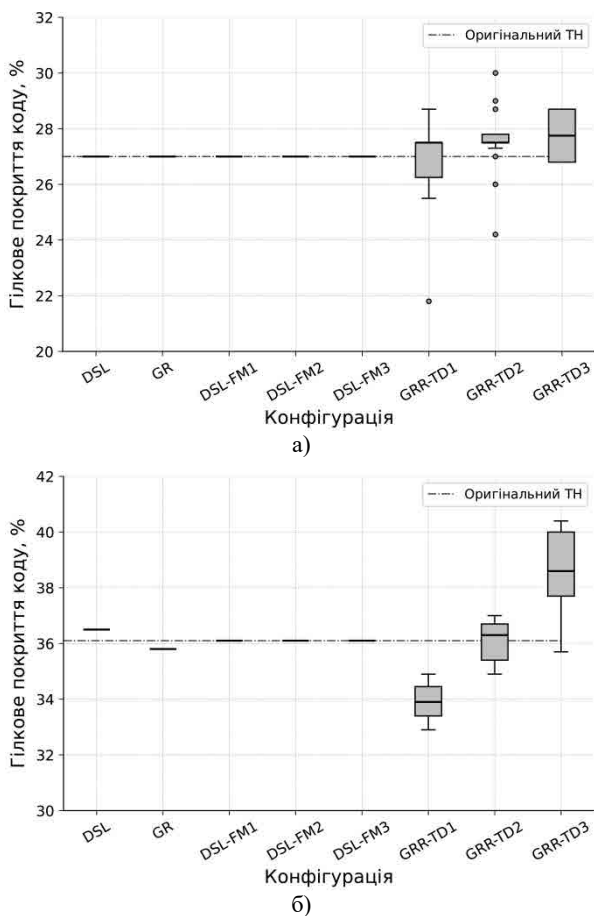


Рис. 7. Розподіл показників гілкового покриття коду залежно від конфігурації алгоритму оптимізації TC для бібліотек *BitmapPlusPlus* (а) та *Hjson* (б)

Розподіл покриття для бібліотеки *BitmapPlusPlus* (рис. 7, а) вказує на те, що класичні алгоритми DSL і GR, а також усі конфігурації DSL-FM стабільно зберігають початковий рівень гілкового покриття. Натомість удосконалені конфігурації GRR-TD демонструють більшу варіативність результатів і підвищене медіанне покриття. Отже, врахування майбутньої корисності дій дає змогу алгоритму компенсувати втрату покриття під час скорочення TC, утримуючи або навіть підвищуючи досяжність гілок.

На рівні розподілу покриття можна спостерігати стабільність DSL і адаптивну поведінку вдосконалених GRR-TD конфігурацій, що підвищують гілкове покриття коду зі зростанням складності середовища тестування.

5.4. Аналіз стиснення TH

На рис. 8 наведено залежність коефіцієнта стиснення TC від конфігурації алгоритмів для бібліотек *BitmapPlusPlus* і *Hjson*. Червоні відрізки на стовпчиках позначають стандартну похибку.

Для бібліотеки *BitmapPlusPlus* (рис. 8, а) спостерігається стабільне зростання коефіцієнта стиснення від базових конфігурацій (DSL, GR) до запропонованих (DSL-FM, GRR-TD). Значення підвищуються приблизно від 0,63 до 0,86, що свідчить про збільшення стиснення TC. Особливо ефективними виявилися конфігурації жадібного формування з огляду на майбутню корисність дій (GRR-TD), які забезпечують не лише найбільший коефіцієнт стиснення, а й високий ступінь збільшення гілкового покриття коду.

Для бібліотеки *Hjson* (рис. 8, б) результати виявляються більш варіативними. Базовий DSL демонструє високий коефіцієнт стиснення (~0,83), тоді як базовий GR значно нижчий (~0,6), що

пояснюється більш складною структурою викликів API в обраній бібліотеці та основними недоліками жадібного алгоритму. Удосконалені конфігурації DSL-FM поступово підвищують коефіцієнт до $\sim 0,86$, що вказує на ефективність фільтрації неінформативних дій.

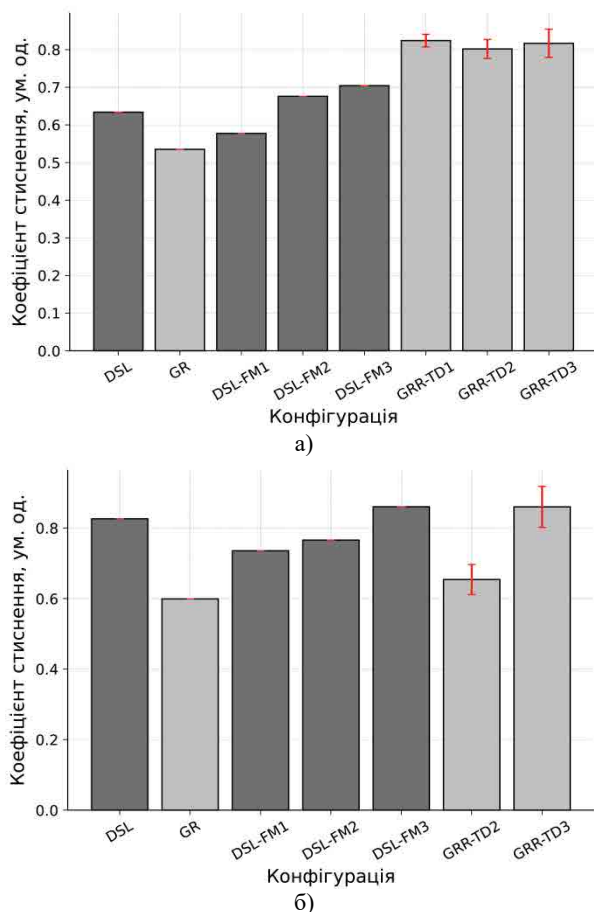


Рис. 8. Коефіцієнт стиснення залежно від конфігурації оптимізації TC для бібліотек *BitmapPlusPlus* (а) та *Hjson* (б)

Натомість для конфігурацій GRR-TD отримано такі результати:

- GRR-TD1 – не подана на графіку через нездатність забезпечити адекватний рівень гілкового покриття;
- GRR-TD2 – досягає кращого стиснення, ніж класичний GR, але визначається збільшенням похибки стиснення;
- GRR-TD3 – має найвище значення коефіцієнта стиснення (0,86), проте визначається великою похибкою.

5.5. Аналіз зменшення часу на виконання ТН

Результати аналізу часу на виконання ТН подано на рис. 9. Вони показують середній час на виконання ТН для різних конфігурацій під час тестування

цільових бібліотек. Червоні відрізки на стовпчиках позначають стандартну похибку. Пунктирна лінія відповідає часу на виконання оригінального ТН, що використаний як орієнтир.

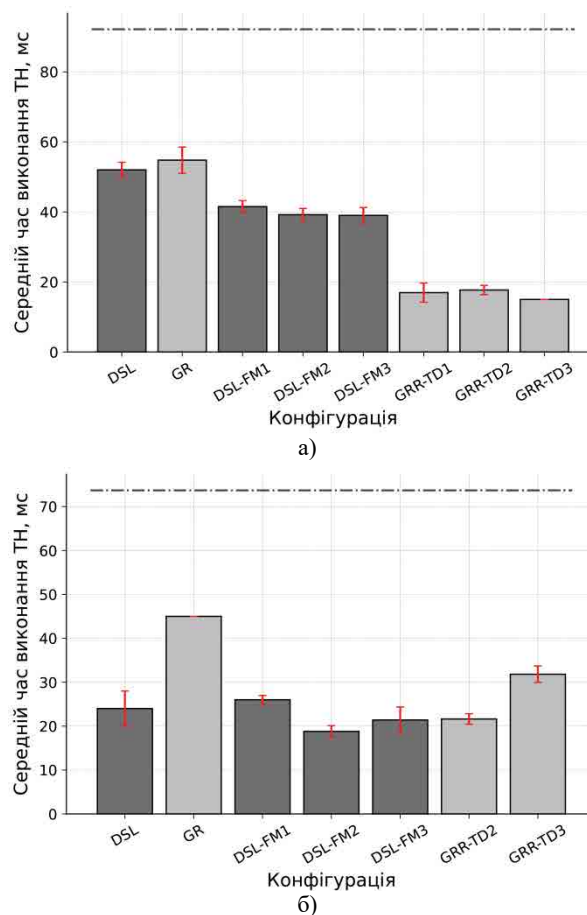


Рис. 9. Середній час на виконання ТН залежно від конфігурації оптимізації TC для бібліотек: *BitmapPlusPlus* (а) та *Hjson* (б)

На рис. 9, а для бібліотеки *BitmapPlusPlus* спостерігається зменшення часу на виконання від базових конфігурацій (DSL, GR) до вдосконалених конфігурацій (DSL-FM, GRR-TD). Середній час на виконання знижується приблизно з 92 мс до 20 мс, тобто більше ніж утричі.

Аналогічна тенденція спостерігається й для бібліотеки *Hjson* (рис. 9, б), де час на виконання скорочується понад удвічі.

Конфігурація GR має найбільший середній час (~ 45 мс), тоді як удосконалені конфігурації DSL-FM і GRR-TD2 забезпечують стабільно нижчі показники – у межах від 19 мс до 25 мс.

Конфігурація GRR-TD3 має дещо більший час (~ 32 мс), що може бути наслідком складнішої логіки

перевірок під час виконання. Однак вона залишається загалом ефективнішою за базові конфігурації.

5.6. Загальна оцінка ефективності алгоритмів

У табл. 4 й 5 наведено зведені результати оцінювання ефективності конфігурацій досліджуваних алгоритмів для C++ бібліотек *BitmapPlusPlus* і *Hjson* відповідно.

Таблиця 4. Оцінка ефективності конфігурацій досліджуваних алгоритмів для бібліотеки *BitmapPlusPlus*

Конфіг.	Коеф. збер. покриття, RC	Коеф. стиснення, CR	Коеф. збер. часу, ECR
DSL	1	0,63	0,34±0,013
GR	1	0,54	0,28±0,048
DSL-FM1	1	0,58	0,44±0,02
DSL-FM2	1	0,68	0,55±0,021
DSL-FM3	1	0,70	0,51±0,024
GRR-TD1	0,92±0,026	0,82±0,016	0,84±0,04
GRR-TD2	1,023±0,03	0,80±0,025	0,82±0,02
GRR-TD3	1,0285±0,038	0,82±0,037	0,86±0,001

Таблиця 5. Оцінка ефективності конфігурацій досліджуваних алгоритмів для бібліотеки *Hjson*

Конфіг.	Коеф. збер. покриття, RC	Коеф. стиснення, CR	Коеф. збер. часу, ECR
DSL	1,011	0,83	0,61±0,07
GR	0,991	0,60	0,228±0,001
DSL-FM1	0,99	0,74	0,61±0,011
DSL-FM2	0,99	0,77	0,67±0,012
DSL-FM3	1	0,85	0,63±0,042
GRR-TD1	0,93±0,02	0,00	0,66±0,038
GRR-TD2	0,988±0,02	0,65±0,04	0,7±0,024
GRR-TD3	1,058±0,046	0,86±0,06	0,55±0,04

Для бібліотеки *BitmapPlusPlus* (табл. 4) базові конфігурації DSL і GR мають однакове збереження покриття ($RC = 1$), але відносно невелике стиснення ($CR = 0,63$ та $0,54$) і низьке зменшення часу на виконання ($ECR \approx 0,34$ та $0,28$).

Удосконалені конфігурації DSL-FM забезпечують покращення CR до 0,7 та збільшення ECR до 0,55, що свідчить про ефективне вилучення надлишкових дій.

Найкращі результати демонструють жадібні конфігурації з огляду на майбутню корисність GRR-TD2 і GRR-TD3, які досягають максимальних значень:

$$RC \approx 1,0285, CR \approx 0,82 \text{ та } ECR \approx 0,86.$$

Отже, спостерігається динамічне скорочення TC після досягнення оригінального гілкового покриття

й забезпечення найбільшого зменшення часу на виконання ТН за умови повного збереження ефективності тестування.

Для бібліотеки *Hjson* (табл. 5) конфігурація DSL зберігає (збільшує) повне покриття ($RC = 1,011$) та забезпечує високе зменшення часу на виконання ($ECR \approx 0,61$), що робить його основним орієнтиром.

Жадібна конфігурація GR має трохи менший коефіцієнт збереження покриття ($RC = 0,991$) і помітно нижчий коефіцієнт $ECR \approx 0,228$, завдяки чому він поступається за ефективністю.

Конфігурації DSL-FM демонструють стабільні результати. У цьому разі значення коефіцієнтів ефективності лежать у таких інтервалах:

$$RC \in [0,99; 1], CR \in [0,74; 0,85] \text{ та } ECR \in [0,63; 0,67].$$

Отримані дані підтверджують ефективність запропонованих конфігурацій порівняно з базовими. Для конфігурацій GRR-TD2 і GRR-TD3 значення RC перевищують 1 ($\approx 1,058$), що означає підвищення гілкового покриття. Конфігурація GRR-TD2 досягає найвищого показника ECR у середньому 0,7, а GRR-TD3 – найвищого коефіцієнта стиснення у середньому 0,86, але з великим розкидом значень.

6. Висновки

У статті запропоновано вдосконалення класичного алгоритму жадібного формування (оптимізації) TC і алгоритму дельта-дебагінг мінімізації TC, а також оцінено їх ефективність.

Удосконалений жадібний алгоритм формування (оптимізації) TC дає змогу підвищити коефіцієнт стиснення TC з 0,63 до 0,86 і скоротити час на виконання ТН у $\sim 2,5$ раз, якщо порівнювати з класичним жадібним алгоритмом, за умови повного зростання (збереження) гілкового покриття з коефіцієнтом до 1,058.

Удосконалений алгоритм дельта-дебагінг, на відміну від класичного алгоритму, демонструє зменшення часу на виконання ТН у $\sim 1,5$ раз, з коефіцієнтом стиснення до 0,85, а також повним збереженням покриття.

Отже, до основних наукових досягнень дослідження належать такі:

– вдосконалено жадібний алгоритм, у якому дії TC обираються з огляду на їх майбутню корисність, визначену за очікуваним приростом гілкового покриття коду. Такий підхід усуває детермінованість класичного жадібного алгоритму, підвищує

інформативність тестового сценарію та забезпечує збалансованість між збереженням гілкового покриття та стисненням ТС;

– вдосконалено алгоритм дельта-дебагінгу, який виконує групове вилучення неунікальних дій ТС за внеском у гілкове покриття, що дає змогу суттєво знизити довжину тестового сценарію без втрати властивостей ефективності тестування.

З практичного погляду, ефективна оптимізація ТС є важливою для тестових сценаріїв, спрямованих на виявлення вразливостей у смарт-контрактах на основі моделей і методів глибокого машинного навчання [37], а також для тестування коректності та стійкості криптографічних алгоритмів шифрування в прикладному ПЗ [38].

Отже, отримані результати можна використати для підвищення ефективності тестування C++ бібліотек із застосуванням CI/CD-інструментів автоматизованого тестування в промислових середовищах розробки ПЗ, зокрема для оптимізації ТН у процесах безперервної інтеграції та регресійного тестування C++ бібліотек у складі ПЗ з високими вимогами до якості й безпеки.

Подальші дослідження необхідно спрямувати на поєднання запропонованих алгоритмів з агентними

методами на основі навчання з підкріпленням для створення адаптивних систем формування ТС для C++ бібліотек [31–33].

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів, зокрема фінансового, особистого, авторського або будь-якого іншого характеру, який міг би вплинути на дослідження, а також на результати, опубліковані в цій статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Дані будуть надані за обґрунтованим запитом.

Використання засобів штучного інтелекту

Автори підтверджують, що не застосовували технології штучного інтелекту для написання статті.

References

- Hulevych, M. (2024), "CIDER: Assisted automation tool for C++ libraries testing", *Control, Navigation and Communication Systems*, Vol. 2, No. 76, pp. 74–77. DOI: <https://doi.org/10.26906/sunz.2024.2.074>
- Hulevych, M., Kolomiitsev, O. (2025), "Automated test generation techniques for C++ software", *Control, Navigation and Communication Systems*, Vol. 2, No. 80, pp. 102–107. DOI: <https://doi.org/10.26906/SUNZ.2025.2.102>
- Tallam, S., Gupta, N. (2005), "A concept analysis inspired greedy algorithm for test suite minimization", in *Proceedings of the 6th ACM SIGPLAN-SIGSOFT Workshop on Program Analysis for Software Tools and Engineering (PASTE '05)*, Association for Computing Machinery, New York, NY, USA, pp. 35–42. DOI: <https://doi.org/10.1145/1108792.1108802>
- Rehman Khan, S.U., Lee, S.P., Javaid, N., Abdul, W. (2018), "A Systematic Review on Test Suite Reduction: Approaches, Experiment's Quality Evaluation, and Guidelines", *IEEE Access*, Vol. 6, pp. 11816–11841. DOI: <https://doi.org/10.1109/ACCESS.2018.2809600>
- Yoo, S., Harman, M. (2012), "Regression testing minimization, selection and prioritization: A survey", *Software Testing, Verification and Reliability*, Vol. 22, No. 2, pp. 67–120. DOI: <https://doi.org/10.1002/stvr.430>
- Noemmer, R., Haas, R. (2020), "An Evaluation of Test Suite Minimization Techniques", in Winkler, D., Biffl, S., Mendez, D. and Bergsmann, J. (Eds.), *Software Quality: Quality Intelligence in Software and Systems Engineering (SWQD 2020)*, Springer, Vol. 371, pp. 51–66. DOI: https://doi.org/10.1007/978-3-030-35510-4_4
- Li, Z., Harman, M., Hierons, R.M. (2007), "Search Algorithms for Regression Test Case Prioritization", *IEEE Transactions on Software Engineering*, Vol. 33, No. 4, pp. 225–237. DOI: <https://doi.org/10.1109/TSE.2007.38>
- Yoo, S., Harman, M., Tonella, P., Susi, A. (2009), "Clustering test cases to achieve effective and scalable prioritisation incorporating expert knowledge", in *Proceedings of the 18th International Symposium on Software Testing and Analysis (ISSTA 2009)*, Association for Computing Machinery, New York, NY, USA, pp. 201–212. DOI: <https://doi.org/10.1145/1572272.1572296>
- Wang, S., Ali, S., Gotlieb, A., Liaaen, M. (2016), "A systematic test case selection methodology for product lines: Results and insights from an industrial case study", *Empirical Software Engineering*, Vol. 21, pp. 1586–1622. DOI: <https://doi.org/10.1007/s10664-014-9345-5>

10. Parsa, S., Khalilian, A. (2009), "A Bi-objective Model Inspired Greedy Algorithm for Test Suite Minimization", in Lee, Yh., Kim, Th., Fang, Wc. and Ślęzak, D. (Eds.), *Future Generation Information Technology (FGIT 2009)*, Springer, Vol. 5899, pp. 208–215. DOI: https://doi.org/10.1007/978-3-642-10509-8_24
11. Jehan, S., Wotawa, F. (2023), "An Empirical Study of Greedy Test Suite Minimization Techniques Using Mutation Coverage", *IEEE Access*, Vol. 11, pp. 65427–65442. DOI: <https://doi.org/10.1109/ACCESS.2023.3289073>
12. Putra, A.W., Legowo, N. (2025), "Greedy Algorithm Implementation for Test Case Prioritization in the Regression Testing Phase", *Journal of Computer Science*, Vol. 21, No. 2, pp. 290–303. DOI: <https://doi.org/10.3844/jcssp.2025.290.303>
13. Górski, T. (2025), "Pattern-Based Test Suite Reduction Method for Smart Contracts", *Applied Sciences*, Vol. 15, No. 2, Article 620. DOI: <https://doi.org/10.3390/app15020620>
14. Lin, C.-T., Tang, K.-W., Wang, J.-S., Kapfhammer, G.M. (2017), "Empirically evaluating Greedy-based test suite reduction methods at different levels of test suite complexity", *Science of Computer Programming*, Vol. 150, pp. 1–25. DOI: <https://doi.org/10.1016/j.scico.2017.05.004>
15. Habib, A.S., Khan, S.U.R., Felix, E.A. (2023), "A systematic review on search-based test suite reduction: State-of-the-art, taxonomy, and future directions", *IET Software*, Vol. 17, No. 2, pp. 93–136. DOI: <https://doi.org/10.1049/sfw2.12104>
16. Li, F., Zhou, J., Li, Y., Hao, D., Zhang, L. (2022), "AGA: An Accelerated Greedy Additional Algorithm for Test Case Prioritization", *IEEE Transactions on Software Engineering*, Vol. 48, No. 12, pp. 5102–5119. DOI: <https://doi.org/10.1109/TSE.2021.3137929>
17. Pan, R., Ghaleb, T.A., Briand, L.C. (2024), "LTM: Scalable and Black-Box Similarity-Based Test Suite Minimization Based on Language Models", *IEEE Transactions on Software Engineering*, Vol. 50, No. 11, pp. 3053–3070. DOI: <https://doi.org/10.1109/TSE.2024.3469582>
18. Zhang, Q., Fang, C., Sun, W., Yu, S., Xu, Y., Liu, Y. (2022), "Test case prioritization using partial attention", *Journal of Systems and Software*, Vol. 192, Article 111419. DOI: <https://doi.org/10.1016/j.jss.2022.111419>
19. Zeller, A., Hildebrandt, R. (2002), "Simplifying and isolating failure-inducing input", *IEEE Transactions on Software Engineering*, Vol. 28, No. 2, pp. 183–200. DOI: <https://doi.org/10.1109/32.988498>
20. Cleve, H., Zeller, A. (2005), "Locating causes of program failures", in *Proceedings of the 27th International Conference on Software Engineering (ICSE 2005)*, Association for Computing Machinery, New York, NY, USA, pp. 342–351. DOI: <https://doi.org/10.1145/1062455.1062522>
21. Harman, M., Hierons, R. (2001), "An overview of program slicing", *Software Focus*, Vol. 2, pp. 85–92. DOI: <https://doi.org/10.1002/swf.41>
22. Mishserghi, G., Su, Z. (2006), "HDD: Hierarchical delta debugging", in *Proceedings of the 28th International Conference on Software Engineering (ICSE 2006)*, Association for Computing Machinery, New York, NY, USA, pp. 142–151. DOI: <https://doi.org/10.1145/1134285.1134307>
23. Herfert, S., Patra, J., Pradel, M. (2017), "Automatically reducing tree-structured test inputs", in *Proceedings of the 32nd IEEE/ACM International Conference on Automated Software Engineering (ASE 2017)*, IEEE, Urbana, IL, USA, pp. 861–871. DOI: <https://doi.org/10.1109/ASE.2017.8115697>
24. Wang, G., Wu, Y., Zhu, Q., Xiong, Y., Zhang, X., Zhang, L. (2023), "A Probabilistic Delta Debugging Approach for Abstract Syntax Trees", in *Proceedings of the 34th IEEE International Symposium on Software Reliability Engineering (ISSRE 2023)*, IEEE, Florence, Italy, pp. 763–773. DOI: <https://doi.org/10.1109/ISSRE59848.2023.00060>
25. Zhou, X., Xu, Z., Zhang, M., Tian, Y., Sun, C. (2025), "WDD: Weighted Delta Debugging", in *Proceedings of the 47th IEEE/ACM International Conference on Software Engineering (ICSE 2025)*, IEEE, Ottawa, ON, Canada, pp. 1592–1603. DOI: <https://doi.org/10.1109/ICSE55347.2025.00071>
26. Vince, D., Hodován, R., Bársony, D., Kiss, Á. (2022), "The effect of hoisting on variants of Hierarchical Delta Debugging", *Journal of Software: Evolution and Process*, Vol. 34, No. 11, Article e2483. DOI: <https://doi.org/10.1002/smr.2483>
27. Zhang, M., Xu, Z., Tian, Y., Cheng, X., Sun, C. (2025), "Toward a Better Understanding of Probabilistic Delta Debugging", in *Proceedings of the 47th IEEE/ACM International Conference on Software Engineering (ICSE 2025)*, IEEE, Ottawa, ON, Canada, pp. 2024–2035. DOI: <https://doi.org/10.1109/ICSE55347.2025.00117>
28. Pontolillo, G.J., Mousavi, M.R. (2024), "Delta Debugging for Property-Based Regression Testing of Quantum Programs", in *Proceedings of the 5th ACM/IEEE International Workshop on Quantum Software Engineering (Q-SE 2024)*, Association for Computing Machinery, New York, NY, USA, pp. 1–8. DOI: <https://doi.org/10.1145/3643667.3648219>
29. Hoen, A., Kamp, D., Gleixner, A. (2025), "MIP-DD: Delta Debugging for Mixed-Integer Programming Solvers", *INFORMS Journal on Computing*. Advance online publication. DOI: <https://doi.org/10.1287/ijoc.2024.0844>
30. Vince, D., Kiss, Á. (2024), "Evaluation of the fixed-point iteration of minimizing delta debugging", *Journal of Software: Evolution and Process*, Vol. 36, No. 10, Article e2702. DOI: <https://doi.org/10.1002/smr.2702>

31. Kuchuk, N., Kashkevich, S., Radchenko, V., Andrusenko, Y., Kuchuk, H. (2024), "Applying edge computing in the execution IoT operative transactions", *Advanced Information Systems*, Vol. 8, No. 4, pp. 49–59. DOI: <https://doi.org/10.20998/2522-9052.2024.4.07>
32. Durmaz, E., Tümer, M.B. (2022), "Intelligent software debugging: A reinforcement learning approach for detecting the shortest crashing scenarios", *Expert Systems with Applications*, Vol. 198, Article 116722. DOI: <https://doi.org/10.1016/j.eswa.2022.116722>
33. Kuchuk, H., Kalinin, Y., Dotsenko, N., Chumachenko, I., Pakhomov, Y. (2024), "Decomposition of integrated high-density IoT data flow", *Advanced Information Systems*, Vol. 8, No. 3, pp. 77–84. DOI: <https://doi.org/10.20998/2522-9052.2024.3.09>
34. Semenov, S., Kolomiitsev, O., Hulevych, M., Mazurek, P., Chernyk, O. (2025), "An Intelligent Method for C++ Test Case Synthesis Based on a Q-Learning Agent", *Applied Sciences*, Vol. 15, No. 15, Article 8596. DOI: <https://doi.org/10.3390/app15158596>
35. Hulevych M. (2025), "Evaluation of the effectiveness of the test scenarios forming method for C++ libraries based on a Q-learning agent", *Management Information Systems and Devises*, No. 4 (187). pp. 20–46. DOI: <https://doi.org/10.30837/0135-1710.2025.187.020>
36. Puterman, M.L. (1994), *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, New York. DOI: <https://doi.org/10.1002/9780470316887>
37. Shmatko O., Kolomiitsev O., Rekova N., Kuchuk N., Matvieiev O. (2023), "Designing and Evaluating DL-Model for Vulnerability Detection in Smart Contracts", *Advanced Information Systems*, Vol. 7, No. 4, pp. 41–51. DOI: <https://doi.org/10.20998/2522-9052.2023.4.05>
38. Fedorchenko V., Yeroshenko O., Shmatko O., Kolomiitsev O., Omarov M. (2024), "Password Hashing Methods and Algorithms on the .NET Platform", *Advanced Information Systems*, Vol. 8, No. 4, pp. 82–92. DOI: <https://doi.org/10.20998/2522-9052.2024.4.11>

Received (Надійшла) 20.12.2025

Accepted for publication (Прийнята до друку) 10.02.2026

Publication date (Дата публікації) 30.03.2026

Відомості про авторів / About the Authors

Коломійцев Олексій Володимирович – заслужений винахідник України, доктор технічних наук, професор, Національний технічний університет "Харківський політехнічний інститут", професор кафедри комп'ютерної інженерії та програмування, Харків, Україна;

Oleksii Kolomiitsev – Honored Inventor of Ukraine, Doctor of Technical Sciences, Professor, National Technical University "Kharkiv Polytechnic Institute", Professor at the Department of Computer Engineering and Programming, Kharkiv, Ukraine;

e-mail: alexus_k@ukr.net

ORCID ID: <http://orcid.org/0000-0001-8228-8404>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57211278112>

Гулевич Михайло Володимирович – аспірант, Національний технічний університет "Харківський політехнічний інститут", кафедра комп'ютерної інженерії та програмування, Харків, Україна;

Mykhailo Hulevych – PhD Student, National Technical University "Kharkiv Polytechnic Institute", Department of Computer Engineering and Programming, Kharkiv, Ukraine;

e-mail: gulevich30misha@gmail.com

ORCID ID: <http://orcid.org/0009-0003-8622-3271>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=60039563700>

Дмитрієв Олег Миколайович – доктор технічних наук, професор, Державний науково-дослідного інститут випробувань і сертифікації озброєння та військової техніки, провідний науковий співробітник, Черкаси, Україна;

Oleh Dmitriiev – Doctor of Technical Sciences, Professor, Leading Researcher, State Scientific Research Institute of Armament and Military Equipment Testing and Certification, Leading Researcher, Cherkasy, Ukraine;

e-mail: dmitriievoleh@gmail.com

ORCID ID: <http://orcid.org/0000-0003-1079-9744>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57208337326>

Левченко Андрій Олександрович – кандидат технічних наук, доцент, Військова академія (м. Одеса), професор кафедри застосування підрозділів військової розвідки та Сил спеціальних операцій, Одеса, Україна;

Andrii Levchenko – Candidate of Technical Sciences, Associate Professor, Odessa Military Academy, Professor at the Department of Training of Intelligence Units and Special Operations Forces, Odessa, Ukraine;

e-mail: katyaandreylev@gmail.com

ORCID ID: <https://orcid.org/0000-0001-5550-0027>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57220805603>

Балабуха Олексій Сергійович – кандидат технічних наук, Харківський національний університет Повітряних Сил ім. І. Кожедуба, докторант, Харків, Україна;

Oleksiy Balabukha – Candidate of Technical Sciences, Ivan Kozhedub Kharkiv National Air Force University, Doctoral Student, Kharkiv, Ukraine;

e-mail: alex15054444@gmail.com

ORCID ID: <http://orcid.org/0000-0002-5263-9485>

EVALUATION OF THE EFFECTIVENESS OF THE ENHANCED GREEDY AND DELTA-DEBUGGING TEST CASE OPTIMIZATION ALGORITHMS FOR C++ LIBRARIES

Efficient optimization of test cases (TCs) is a necessary condition for improving the efficiency of testing C++ libraries. **The subject of the research** is test case optimization algorithms for C++ libraries. **The purpose of** this work is to improve the efficiency of C++ libraries testing by enhancing classical algorithms for greedy TC optimization and delta-debugging TC minimization. **Goals.** To improve the greedy TC optimization algorithm and eliminate its determinism, ensuring effective TC compression while preserving branch code coverage. To improve the delta-debugging TC minimization algorithm under conditions where redundant actions are sparsely distributed within TCs. To evaluate the effectiveness of the proposed algorithms in comparison with the classical algorithms. **Methods.** The study applies a greedy TC optimization algorithm, a delta-debugging TC minimization algorithm, mathematical modeling, and statistical analysis methods. The effectiveness of the improved algorithms is evaluated based on statistical analysis of 100 simulation runs for each algorithm on two open-source C++ libraries of different structural complexity. **Results.** The evaluation results indicate that the improved algorithms provide preservation (and possible increase) of branch coverage with coverage retention coefficient of up to 1.058, increase the TC compression ratio up to 0.86, and reduce the TS execution time by 1.5–2.5× compared to the classical algorithms. **Conclusions.** The improved algorithms significantly reduce the testing time of C++ libraries without loss of branch coverage. An improved greedy TC optimization algorithm is proposed, in which the selection of TC actions accounts for their future utility estimated by the expected increase in branch coverage. This approach removes the determinism of the classical greedy algorithm, increases the informativeness of TCs, and provides a balance between preserving branch coverage and TC compression. An improved delta-debugging algorithm is proposed that performs group removal of non-unique TC actions according to their contribution to branch coverage, which makes it possible to substantially reduce the length of TCs without losing testing effectiveness.

Keywords: software; test case; algorithm; optimization; coverage; time consumption; fault.

Бібліографічні описи / Bibliographic descriptions

Коломійцев О. В., Гулевич М. В., Дмитрієв О. М., Левченко А. О., Балабуха О. С. Оцінювання ефективності вдосконалених жадібного й дельта-дебагінг алгоритмів оптимізації тестових сценаріїв для C++ бібліотек. *Сучасний стан наукових досліджень та технологій в промисловості*. 2026. № 1 (35). С. 39–54. DOI: <https://doi.org/10.30837/2522-9818.2026.1.039>

Kolomiitsev, O., Hulevych, M., Dmitriyev, O., Levchenko, A., Balabukha, O. (2026), "Evaluation of the effectiveness of the enhanced greedy and delta-debugging test case optimization algorithms for C++ libraries", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (35), P. 39–54. DOI: <https://doi.org/10.30837/2522-9818.2026.1.039>

Alexander Krajči, Ludmila Sidorenko, Olesia Barkovska

PREDICTING RISKS OF CARDIOVASCULAR DISEASE ON SMALL DATASETS USING FEATURE ENGINEERING

Relevance. Cardiovascular diseases remain a leading cause of mortality globally, creating a high demand for automated diagnostic systems. However, developing reliable machine learning models for electrocardiogram (ECG) analysis is often hindered by the availability of only small-scale and imbalanced datasets, which limits the effectiveness of deep learning approaches. **The object of research** is the process of automated processing and classification of electrocardiographic signals for diagnostic purposes. **The subject of the research** includes methods of beat-centric feature extraction, patient-level aggregation strategies, and machine learning algorithms for cardiovascular risk prediction. **The purpose of this paper** is to develop and evaluate a reliable classification framework, optimized for small datasets, that increases prediction accuracy by leveraging patient-level feature aggregation and explainable machine learning models. To achieve this goal, the following tasks were solved: 1) implementation of a robust preprocessing pipeline using a refined Pan-Tompkins algorithm for precise beat-centric segmentation; 2) development of a statistical feature aggregation strategy to mitigate local signal variability; and 3) optimization and validation of a Random Forest classifier. The methodology employed includes digital signal processing (Butterworth filtering), advanced feature engineering (HRV, Wavelets analysis), and rigorous 10-fold Stratified Cross-Validation to ensure generalization on limited data. **Research results.** The study proposes a pipeline initiating with standard signal preprocessing, followed by precise R-peak detection and beat-centric segmentation. Physiological features (HRV, wavelet, morphological) are then extracted from individual segments and statistically aggregated at the patient level. Experiments on a dataset of 164 subjects demonstrated that the proposed patient-level aggregation strategy significantly outperformed traditional segment-based analysis. The final Random Forest model achieved an ROC-AUC score of 0.84. Feature importance analysis confirmed the critical role of Heart Rate Variability (HRV) metrics, particularly SDNN and RMSSD, in differentiating between healthy and high-risk subjects.

Keywords: classification; risk prediction; feature importance; machine learning; ECG signal; Random Forest.

Introduction

Relevance

Cardiovascular diseases (CVDs) remain the leading cause of global mortality, driving the urgent demand for automated electrocardiogram (ECG) analysis systems [1]. In practice, however, model development is frequently constrained by small, imbalanced, and partially noisy datasets, a setting in which end-to-end deep learning often underperforms or becomes fragile without extensive data curation and augmentation. Consequently, feature-engineering-driven pipelines combined with interpretable machine learning remain highly relevant for clinical screening scenarios and low-data environments [2].

A persistent methodological gap arises from how ECG segments are handled. Many studies classify unaligned fixed-length windows, implicitly assuming that a random window is representative of a patient's global cardiac status. In small datasets, this assumption is brittle: local artifacts within a single window can dominate predictions, and segment-level labels may leak patient-specific idiosyncrasies across folds unless evaluation is performed strictly at the patient level. Addressing these

issues, this work adopts a beat-centric segmentation anchored at R peaks, aggregates features at the patient level, and trains an explainable ensemble model to improve generalization in small cohorts.

An overview of scientific works

Many studies focus on automated ECG signal processing and CVD detection. Classical surveys summarize the landscape of ECG analysis, covering preprocessing, morphological characterization, and feature extraction strategies for arrhythmia and risk stratification tasks (e.g., statistical and spectral descriptors, wavelets, entropy-based measures) [3, 4, 5]. Deep Learning approaches in cardiovascular diagnostics are reviewed in [6], which highlights both their potential and the difficulties posed by data needs. Foundational signal processing elements include QRS detection (e.g., Pan-Tompkins) and more recent advances in precise R peak localization, which underpin robust beat-centric workflows [7–10]. Studies [7, 11] focus on current trends in feature extraction methods and their applications. Higher-order wavelet packet decomposition (WPD) statistics are considered in [8]. In terms of signal preprocessing, [9]

establishes the basic real-time QRS detection algorithm, and [10] presents more recent developments in accurate R-peak localization. In addition to preprocessing, classifiers are important in processing ECG signals [12, 13]. In parallel with signal processing, the interpretability of decision systems has motivated the adoption of tree-based models and fuzzy inference [14, 15]. In particular, fuzzy decision trees and fuzzy classifiers have been shown to provide transparent rule-based reasoning and to support importance analysis of decision-making factors, which is vital for safety-critical healthcare applications [13, 16]. These efforts range from early formulations of fuzzy decision tree induction and information measures, to reliability assessment of healthcare systems via fuzzy trees, and to modern, large scale ECG classification with fuzzy approaches. Collectively, this body of work motivates our preference for interpretable, feature-based modeling in low data regimes.

Problem Statement and Approach

We target cardiovascular risk prediction on a small proprietary cohort of 164 subjects (92 high risk, 72 healthy), recorded at 512 Hz, reflecting a mild class imbalance representative of real-world settings. Rather than classifying unaligned windows, we first perform standard artifact suppression, precise R peak detection, and beat centric segmentation; then we extract a comprehensive set of physiologically grounded features (time/frequency domains, HRV metrics, wavelet energy, morphology, non-linear entropy). To stabilize subject profiles and mitigate local variability, we aggregate segment-level features per patient via summary statistics (mean, standard deviation, min, max), and train an interpretable Random Forest with systematic feature selection and cross-validated hyperparameter tuning. All validation is carried out with stratified 10-fold cross validation at the patient level to prevent leakage; no synthetic oversampling (e.g., SMOTE) is used in final evaluation, ensuring performance reflects the natural class distribution.

Our emphasis on interpretable features and transparent model behavior is consistent with the broader trend of leveraging fuzzy decision trees and fuzzy classifiers for explainability in medical AI. Fuzzy trees support rule extraction and factor importance analysis, which can complement or extend tree ensemble importance measures and SHAP style attributions. Recent research demonstrates that fuzzy frameworks can successfully handle uncertainty, vagueness, and limited data, especially in clinical contexts. We view our

pipeline as a strong baseline for small-cohort ECG studies, which can be further hybridized with fuzzy decision tree induction to yield explicit clinical rules while retaining the robustness of patient-level aggregation [13, 14].

Setting objectives

A key difficulty in implementing machine learning on small ECG groups is the approach for managing segmented data. Classifying unaligned segments independently often introduces variability, as local artifacts in a single segment may not reflect the patient's overall cardiac status. Therefore, it becomes necessary to optimize the processing pipeline by implementing a robust patient-level aggregation strategy. Based on this, the goal of this paper is to develop and evaluate a reliable classification framework, optimized for small datasets, that increases prediction accuracy by leveraging patient-level feature aggregation and explainable machine learning models. This paper makes the following contributions to ECG based risk prediction under small data constraints:

1. ***Beat-centric, patient-level pipeline for small cohorts.*** We formalize and evaluate an end-to-end workflow that combines zero-phase bandpass filtering, refined Pan–Tompkins–inspired R-peak detection, beat-anchored segmentation, and patient-level feature aggregation. The design explicitly targets low sample settings where segment-only classification is unstable.

2. ***Patient level aggregation as a principled regularizer.*** We introduce a simple yet effective statistical aggregation schema (mean, std, min, max) over segment features per subject and demonstrate its superiority to segment-based analysis on the same data, improving ROC AUC from 0.64 (baseline) to 0.84 and raising accuracy from 0.57 to 0.80. This quantifies the benefit of aggregation for reducing local noise and enhancing generalization in small populations.

3. ***Rigorous, leakage free evaluation.*** All experiments use stratified 10-fold cross validation at the patient level, with no oversampling in final reporting, and a feature selection step driven by Random Forest importance to curb dimensionality and improve stability – practices aligned with clinical ML standards for limited datasets.

4. ***Clinically meaningful interpretability.*** We provide feature importance analysis showing that HRV metrics (e.g., SDNN, RMSSD) dominate the model's decisions, in agreement with physiological understanding of autonomic regulation – thereby enhancing trust and facilitating clinical translation.

5. Pathway to fuzzy explainable hybrids.

By situating our results within the literature on fuzzy decision trees and fuzzy classifiers, we outline a natural extension toward rule based, linguistically interpretable models on top of patient level features, potentially combining the robustness of ensembles with the transparency of fuzzy rules.

Materials and Methods

ECG Signal Morphology

Before discussing the dataset, it is essential to define the structural characteristics of the analyzed signals. A typical ECG cycle (Fig. 1) consists of distinct waves reflecting the electrical activity of the heart: the P-wave (atrial depolarization), the QRS complex (ventricular depolarization), and the T-wave (ventricular repolarization). In the context of cardiovascular risk prediction, the QRS complex is particularly significant as it represents the main contraction of heart ventricles. Its accurate detection is a prerequisite for extracting heart rate variability (HRV) metrics, which are derived from the R-R intervals (the time distance between consecutive R-peaks). This morphological structure is susceptible to distortion and noise, which might result in inaccurate feature extraction and call for thorough preprocessing.

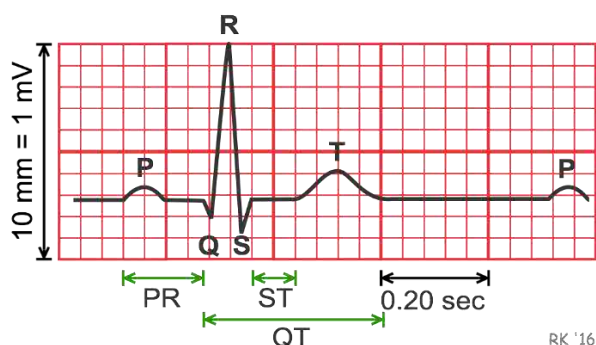


Fig. 1. Morphology of a standard ECG cycle composed of P,Q,R,S and T waves [17]

Dataset description

The experimental evaluation was conducted using a proprietary dataset of raw electrocardiographic signals obtained from 164 subjects. The cohort consists of two classes:

1. *Healthy Control Group (Label 0)*: 72 subjects with no history of cardiovascular pathology.
2. *High-Risk Group (Label 1)*: 92 subjects with cardiovascular disorders or exhibiting high-risk factors.

The class distribution is visualized on Fig. 2. The dataset exhibits a mild class imbalance (44% healthy and 56% high-risk), which reflects real-world clinical scenarios. The signals were recorded at a sampling frequency of 512 Hz, ensuring sufficient temporal resolution for precise morphological analysis.

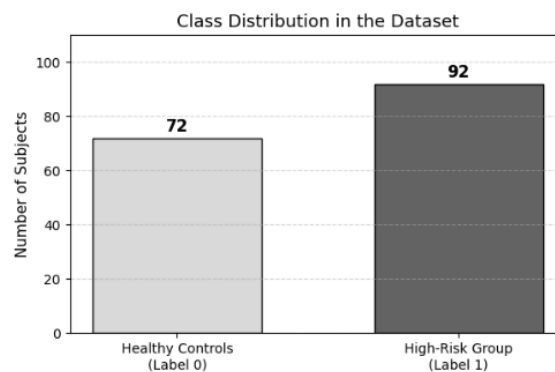


Fig. 2. Distribution of classes in the experimental dataset

Proposed Classification System

Building upon the defined methodological framework, the technical implementation of the proposed solution is divided into two distinct phases. **Phase 1** includes signal processing and feature engineering, initiating with raw ECG data and concluding with the creation of an aggregated patient-level dataset. **Phase 2** focuses on model design and classification, including data splitting, feature selection, hyperparameter tuning and the final model training.

Phase 1: Signal Preprocessing and Feature Engineering. To maximize information extraction while minimizing noise, we developed a multi-stage pipeline (Fig. 3). The workflow converts raw recordings into a structured dataset through the following physiologically grounded steps.

Signal Preprocessing. The input raw ECG data were first processed to remove artifacts such as baseline wander and muscle noise. We utilized a **4th-order Butterworth bandpass filter** designed via the *SciPy signal processing library*. The cutoff frequencies were set to 0.5 Hz and 40 Hz; the lower bound effectively eliminates baseline drift, while the upper bound suppresses high frequency electromyographic noise. To ensure numerical stability, the filter was implemented using Second-Order Sections (SOS). Crucially, we applied a bidirectional filter (*sosfiltfilt* from SciPy) to ensure zero-phase distortion. This preserves the precise shape and timing of the ECG waves, which is essential for accurate morphological analysis.

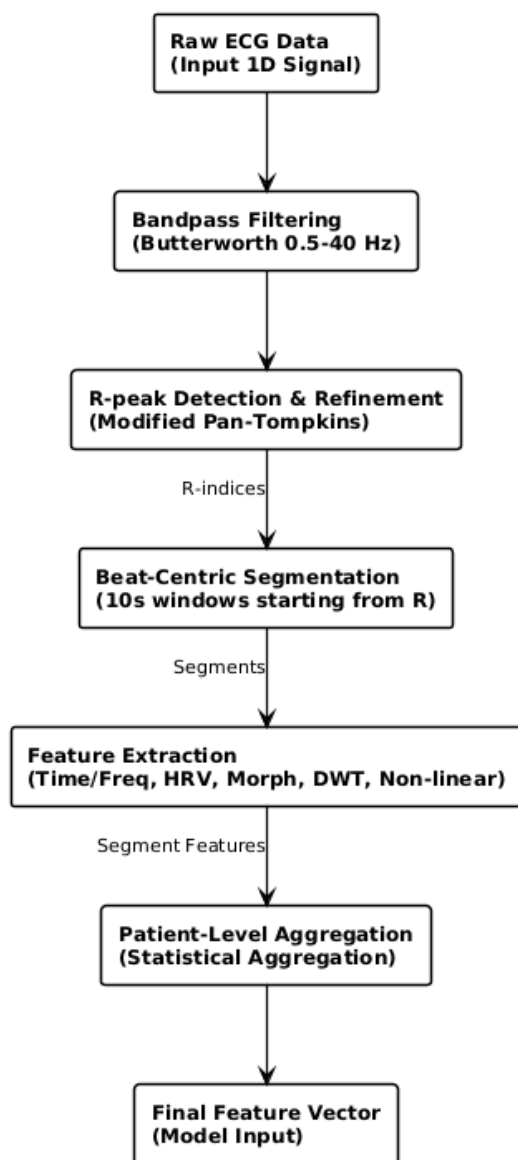


Fig. 3. Workflow of the signal processing and feature engineering pipeline (Phase 1)

R-Peak Detection. Following filtration, precise localization of R-peaks was performed using a multi-stage method inspired by the **Pan-Tompkins algorithm**. The process involved three distinct steps:

1. **Signal Enhancement:** The signal underwent a series of transformations to highlight the QRS complex. This included narrow-band filtration (5-15 Hz), differentiation to highlight slope changes, squaring (y^2) to emphasize high-energy segments, and moving window integration (150 ms window) to smooth the waveform.

2. **Candidate detection:** Peaks were identified on the transformed signal using an adaptive thresholding technique. The threshold was dynamically set as the signal median plus four times the **Median Absolute**

Deviation (MAD). This approach ensures robustness against outliers and varying signal amplitudes. A minimum distance of 250 ms was enforced to prevent multiple detections within a single heartbeat.

3. **Refinement and Relocalization:** Since the enhancement steps can slightly shift peak positions, each candidate was relocalized in the original filtered signal. The algorithm searched for the precise local maximum within a ± 100 ms window around the candidate. Finally, a refractory period check eliminated any duplicate detections that were physiologically impossible (closer than 250 ms).

Beat-centric Segmentation and Feature Extraction. Leveraging the validated R-peak indices, the continuous ECG signal was segmented into fixed 10-second windows. Crucially, we employed a beat-centric alignment strategy, ensuring that each segment initiates exactly at an R-peak. This synchronization guarantees temporal consistency across all samples. Subsequently, a comprehensive feature set was extracted from each segment, spanning multiple physiological domains:

- **Time-Domain Statistics:** Basic statistical descriptors of the signal amplitude, including mean, variance, skewness, kurtosis and signal energy.

- **Frequency-Domain Metrics:** Computed using Welch's method to estimate Power Spectral Density (PSD). Key features include Spectral Entropy, Dominant Frequency and relative power in Low-Frequency (LF) and High-Frequency (HF) bands.

- **Heart Rate Variability (HRV):** Derived from the inter-beat (R-R) intervals within the segment. We calculated standard metrics such as SDNN (total variability), RMSSD (short-term variability) and pNN50 to capture autonomic nervous system activity.

- **Morphological Features:** Direct measurements of the cardiac cycle shape, specifically the mean R-peak amplitude and the average QRS complex width.

- **Wavelet Features:** To analyze non-stationary characteristics, we utilized the Discrete Wavelet Transform (DWT). The signal was decomposed using the Daubechies 4 (db4) wavelet, and energy statistics were computed for different decomposition levels.

- **Non-linear Features:** We calculated Sample Entropy to quantify the complexity and regularity of the time series data.

Patient-Level Aggregation: The final step involved transitioning from segment-level data to a patient-level profile. The extracted segments were grouped by patient

ID. To create a stable feature vector that minimizes the impact of local artifacts, we applied statistical aggregation. For every feature column F_j (e.g., RMSSD), four descriptors were computed across all segments of the patient: Mean, Standard Deviation, Minimum and Maximum. This resulted in a high-dimensional, robust representation of each patient, ready for the classification phase.

Phase 2: Model Design and Classification.

The second phase focused on optimizing the predictive model using the aggregated dataset prepared in Phase 1. The systematic workflow of this phase, covering feature selection and hyperparameter tuning, and validation, is illustrated in Fig. 4.

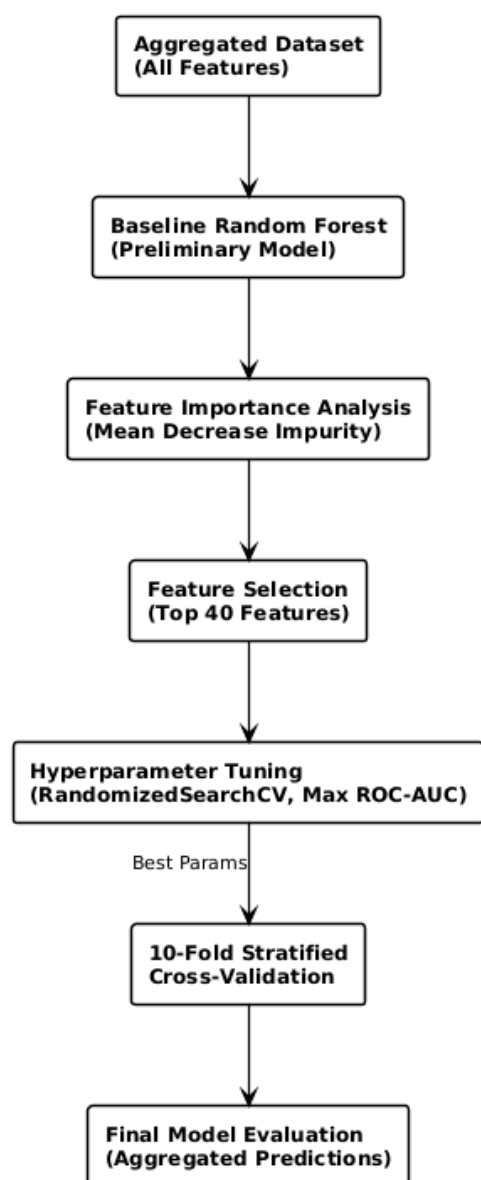


Fig. 4. Workflow of the model optimization and evaluation process (Phase 2)

Feature Selection Strategy: To reduce dimensionality and remove irrelevant predictors, we first trained a baseline Random Forest model on the complete feature set. We utilized the model's built-in feature importance metrics (Mean Decrease Impurity) to quantify the contribution of each feature to the decision-making process. Based on this analysis, we identified and selected the Top 40 most informative features. This step was crucial for focusing the model on the most predictive physiological markers while eliminating noise.

Hyperparameter Tuning: To find the optimal model configuration, we utilized RandomizedSearchCV integrated with 10-fold Stratified Cross-Validation. Unlike exhaustive search methods (e.g., Grid Search), this approach samples a fixed number of parameter settings from specified distributions. This offers a computationally efficient way to explore a high-dimensional hyperparameter space. The optimization process focused on maximizing the ROC-AUC metrics. The search space covered a wide range of hyperparameters, including:

- Number of estimators: 100 – 500
- Maximum tree depth: 2 – 20
- Minimum samples split: 2 - 10
- Minimum samples per leaf: 2 – 10
- Max features: Square root or Log2

Final Evaluation Strategy: To ensure an unbiased assessment of the model's generalization capability, we employed 10-fold Stratified Cross-Validation using optimized hyperparameters. This technique is essential for small and imbalanced datasets, as it preserves the class proportions across all training and validation folds. Final performance metrics were derived by aggregating predictions from the validation sets. Unlike a single train-test split, this approach ensures that every data point is evaluated exactly once as unseen data, providing a realistic estimate of performance while preventing data leakage. Synthetic oversampling (SMOTE) was not applied during this evaluation phase; therefore, the reported metrics (ROC-AUC, Sensitivity, Specificity) accurately reflect the model's performance on the original, real-world class distribution.

Research Results

Comparison of Classification Strategies

To validate the hypothesis that precise beat-centric alignment combined with patient-level aggregation is essential for robust classification, we compared the

proposed method against a baseline segment-based approach. The baseline model served as a control experiment to quantify the impact of our proposed feature engineering pipeline. In this baseline setup, the processing differed in three key aspects:

1. **Arbitrary Segmentation:** The ECG signals were divided into fixed 10-second windows without synchronization to R-peaks. Consequently, segments could start in the middle of a QRS complex, introducing noise and artifacts.

2. **No Aggregation:** Each segment was classified independently as a separate sample, ignoring the fact that multiple segments belong to the same patient.

3. **Raw Feature Set:** The classification was performed on the full set of extracted features without the specialized selection step used in the final model.

As shown in **Table 1**, this lack of synchronization and patient-level context resulted in suboptimal performance.

Table 1. Performance comparison between the baseline and the proposed method

Approach	Recall (0)	Recall (1)	Accuracy	ROC-AUC
Baseline	0.29	0.80	0.57	0.64
Proposed	0.72	0.86	0.80	0.84

The baseline segment-based method demonstrated poor generalization capability, achieving an overall Accuracy of only 57% and an ROC-AUC of 0.64. Crucially, the Recall for the healthy control group was only 0.29, indicating a high rate of false positives caused by local signal artifacts and misalignment. In contrast, the proposed patient-level aggregation strategy significantly improved the model's discriminative power. By statistically aggregating features across the patient's recording, the ROC-AUC increased to 0.84 and Accuracy reached 80%. This improvement in the AUC score confirms that the proposed aggregation method effectively filters out local noise and provides stable subject profile.

Performance of the Final Model

The final Random Forest model was evaluated using Stratified 10-Fold-Cross-Validation. The quantitative results demonstrate the model's robust discriminative capability, achieving an overall Accuracy of 80% and an aggregated ROC-AUC score of 0.84. Detailed performance metrics for each class are summarized in Table 2. The model exhibits a consistent precision

profile, achieving a Precision of 0.80 for both Healthy and High-Risk classes. This indicates that when the model predicts a specific label, it is correct in 80% of cases, suggesting a high degree of trustworthiness in positive detections. It should be noted that the support value of 164 in the final row of Table 2 represents the total number of subjects in the dataset.

Table 2. Detailed classification report of the final optimized mode

Class	Precision	Recall	F1-Score	Support
Healthy (0)	0.80	0.72	0.76	72
High-Risk (1)	0.80	0.86	0.83	92
Weighted Avg	0.80	0.80	0.80	164

Sensitivity and Specificity Analysis

A critical finding is observed in the Recall metrics. The model achieved a Recall of 0.86 for the High-Risk class, which corresponds to Sensitivity. In medical screening scenarios, maximizing Sensitivity is often prioritized over Specificity to minimize Type II errors (False Negatives). The obtained result confirms that the system successfully identified 86% of all high-risk patients, missing only a small fraction. Conversely the Recall for Healthy class, representing Specificity, reached 0.72. While lower than Sensitivity, this is an acceptable trade-off, as a moderate rate of false alarms (Type I errors) is preferable to missing a pathological condition.

Confusion Matrix Interpretation

The distribution of correct and incorrect predictions is further visualized in the Confusion Matrix (Fig. 5).

True Class	Healthy (0)	52	20
	High-Risk (1)	13	79
		Healthy (0)	High-Risk (1)
		Predicted Class	

Fig. 5. Confusion Matrix of the final classification model

- True Positives (TP): Out of 92 high-risk subjects, the model correctly identified 79 cases.
- False Negatives (FN): Only 13 high-risk patients were misclassified as healthy.
- True Negatives (TN): The model correctly recognized 52 healthy subjects.
- False Positives (FP): There were 20 cases where healthy individuals were flagged as high-risk.

Feature Importance Analysis

To gain insight into the model's decision-making process, we analyzed feature contributions using the Mean Decrease Impurity metric. The analysis was performed on the final optimized model. As illustrated in Fig. 6, the ranking of the top 10 features

reveals a distinct dominance of **Heart Rate Variability (HRV)** metrics. The top five positions were occupied exclusively by aggregated statistics derived from R-R intervals. This confirms that the variability of the heart rate, rather than just the raw morphology, is the strongest predictor of cardiovascular risk in this dataset. Beyond HRV, the model also leveraged **Wavelet-based** features, which capture signal energy in specific frequency bands, and time-domain amplitude descriptors. This hierarchy supports the clinical validity of the proposed approach. The fact that the model prioritized interpretable physiological markers over complex, abstract patterns suggests that the classifier relies on medically relevant phenomena rather than learning noise.

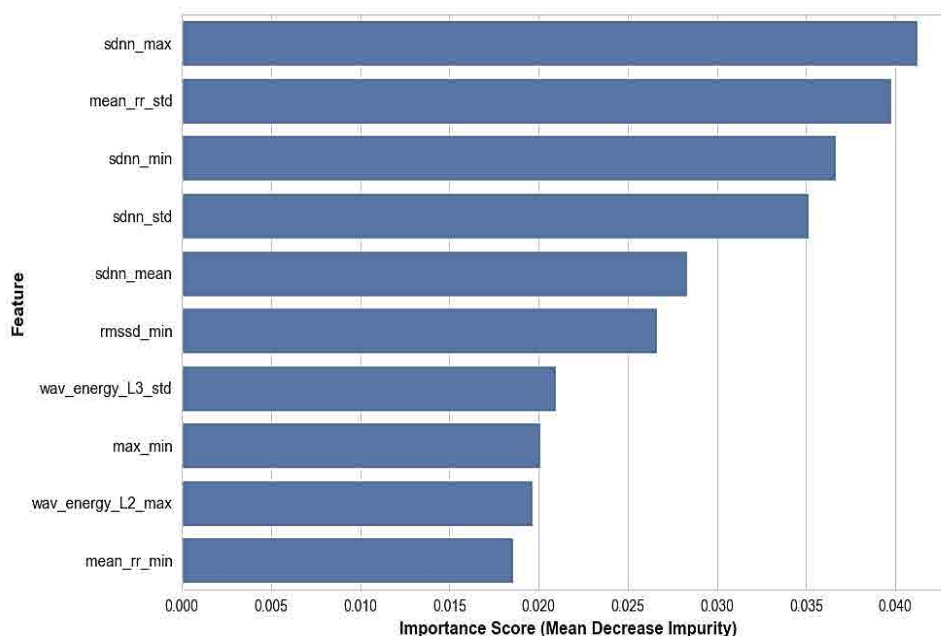


Fig. 6. Top 10 most significant features identified by the Random Forest model

Conclusions

The following results were obtained in this paper:

1. A multi-stage processing pipeline was implemented, utilizing a 4th-order Butterworth filter and a refined Pan-Tompkins algorithm to ensure precise R-peak detection and beat-centric segmentation.
2. A patient-level feature aggregation strategy was developed, which successfully mitigated local signal variability. This approach increased classification accuracy from 57% (baseline segment-based method) to 80%.
3. The Random Forest classifier was optimized using feature selection and hyperparameter tuning.

The final model achieved a ROC-AUC of 0.84 and a Sensitivity of 0.86 for high-risk subjects, demonstrating strong screening potential.

4. Feature importance analysis confirmed that Heart Rate Variability (HRV) metrics (SDNN, RMSSD) are the most critical predictors for differentiating between healthy and high-risk subjects, confirming the clinical interpretability of the proposed model.

Despite the encouraging results, this study has several limitations that should be acknowledged. First, the experimental evaluation was conducted on a relatively small, single-center proprietary dataset, which may limit the generalizability of the findings to broader

and more heterogeneous populations. Second, although strict patient-level cross-validation was applied to prevent data leakage, external validation on an independent cohort is required to fully assess the robustness of the proposed framework. Third, the current model relies on tree-ensemble feature importance, which, while informative, provides only a global view of decision drivers.

Future work will therefore focus on two complementary directions. From a modeling perspective, the proposed patient-level feature representation creates a natural foundation for fuzzy decision trees and fuzzy classifiers, which can translate aggregated physiological features into explicit, linguistically interpretable rules, enhancing clinical transparency and trust [18]. From an explainability perspective, integrating local explanation techniques such as SHAP would enable subject-specific analysis of risk factors, supporting individualized clinical decision-making. Together, these extensions aim to combine robust patient-level aggregation, formal uncertainty handling,

and fine-grained interpretability, further advancing ECG-based risk prediction in small-data clinical settings.

Conflict of interest

The authors declare that they have no conflicts of interest, including financial, personal, authorial, or any other nature, that could influence the research or the results published in this article.

Funding

The study was conducted without financial support.

Data availability

The manuscript has no associated data.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technology in the creation of this paper.

References

1. "World Heart Report 2023: Confronting the World's Number One Killer. Geneva, Switzerland. World Heart Federation". available at: <https://medbox.org/document/world-heart-report-2023-confronting-the-worlds-number-one-killer>
2. Global, Regional, and National Burden of Cardiovascular Diseases and Risk Factors in 204 Countries and Territories, 1990-2023. *JACC Journals*, Vol. 86, No. 22, 2025, pp. 2167-2243. DOI: <https://doi.org/10.1016/j.jacc.2025.08.015>
3. Berkaya, S. K. et al. (2018), "A Survey on ECG Analysis", *Biomedical Signal Processing and Control*, vol. 43, pp. 216-235, doi: <https://doi.org/10.1016/j.bspc.2018.03.003>
4. Yong O., Yang L., Kardos A., Zhao Y. (2026), "Non-invasive cardiovascular and vital signs monitoring techniques: review, challenges, and perspectives, Measurement", Vol. 258, part E, 119472 p. DOI: <https://doi.org/10.1016/j.measurement.2025.119472>
5. Ghasad, P.P., Vegivada, J.V.S., Kamble, V.M., et al. (2025), "A systematic review of automated prediction of sudden cardiac death using ECG signals", *Physiological Measurement*, Vol. 46. DOI: <https://doi.org/10.1088/1361-6579/ad9ce5>
6. Wu, Z., Guo, C. (2025), "Deep learning and electrocardiography: systematic review of current techniques in cardiovascular disease diagnosis and management", *BioMed Eng OnLine*, Vol. 24, No. 23, DOI: <https://doi.org/10.1186/s12938-025-01349-w>
7. Singh, A.K, Krishnan, S. (2023), "ECG signal feature extraction trends in methods and applications", *Biomed Eng Online*, Vol. 22(1):22. DOI: <https://doi.org/10.1186/s12938-023-01075-1>
8. Kutlu, Y., Kuntalp, D. (2012), "Feature Extraction for ECG Heartbeats Using Higher Order Statistics of WPD Coefficients", *Computer Methods and Programs in Biomedicine*, Vol. 105, No. 3, pp. 257-267, DOI: <https://doi.org/10.1016/j.cmpb.2011.10.002>
9. Pan, J., Tompkins, W. J. (1985), "A Real-Time QRS Detection Algorithm", *IEEE Trans. on Biomedical Engineering*, Vol. 32, No. 3, pp. 230-236, DOI: <https://doi.org/10.1109/TBME.1985.325532>
10. Zhai, D., Bao, X., Long, X., Ru, T. and Zhou, G. (2023), "Precise Detection and Localization of R-Peaks From ECG Signals", *Mathematical Biosciences and Engineering*, Vol. 20, No. 11, pp. 19191-19208, DOI: <https://doi.org/10.3934/mbe.2023848>
11. Safdar, M.F., Nowak, R.M., Palka, P. (2023), "Pre-Processing techniques and artificial intelligence algorithms for electrocardiogram (ECG) signals analysis: A comprehensive review", *Computers in Biology and Medicine*, Vol. 170, 107908 p. DOI: <https://doi.org/10.1016/j.compbiomed.2023.107908>
12. Breiman, L. (2001), "Random Forests", *Machine Learning*, Vol. 45, No. 1, pp. 5-32, DOI: <https://doi.org/10.1023/A:1010933404324>
13. Rabcan, J., Zaitseva, E., Levashenko, V., Kvassay, M. (2025), "Advancing ECG Signal Classification with a Fuzzy Classifier Approach", *IEEE Access*, Vol.13, 2025, pp. 83840-83856. DOI: <https://doi.org/10.1109/ACCESS.2025.3568086>

14. Zaitseva, E., Rabcan, J., Levashenko, V., Kvassay, M. (2023), "Importance analysis of decision-making factors based on fuzzy decision trees", *Applied Soft Computing*, Vol. 134, 109988 p. DOI: <https://doi.org/10.1016/j.asoc.2023.109988>
15. Levashenko, V., Zaitseva, E. (2002), "Usage of New Information Estimations for Induction of Fuzzy Decision Trees". In: Yin, H., Allinson, N., Freeman, R., Keane, J., Hubbard, S. (eds) *Intelligent Data Engineering and Automated Learning - IDEAL 2002*. IDEAL 2002. *Lecture Notes in Computer Science*, Vol 2412. Springer, Berlin, Heidelberg, DOI: https://doi.org/10.1007/3-540-45675-9_74
16. Zaitseva, E., Levashenko, V., Puuronen, S. (2007), "Fuzzy classifier based on fuzzy decision tree", *Proc. of the Int. Conf. on Computer as a Tool (EUROCON)*, 2007, pp. 823-827, DOI: <https://doi.org/10.1109/EURCON.2007.4400614>
17. Klabunde, R.E. (2021), "Normal Sinus Rhythm", *Cardiovascular Physiology Concepts*. available at: <https://cvphysiology.com/arrhythmias/a009>
18. Zaitseva, E., Levashenko, V., Kvassay, M., Deserno T. (2026), "Reliability estimation of healthcare systems using Fuzzy Decision Trees", *Proc. of the Fed. Conf. on Computer Science and Information Systems (FedCSIS)*, 2016, pp. 331-340, DOI: <https://doi.org/10.15439/2016F150>

Received (Надійшла) 01.01.2026

Accepted for publication (Прийнята до друку) 10.02.2026

Publication date (Дата публікації) 30.03.2026

Відомості про авторів / About the Authors

Крайчі Олександр – аспірант, Жилінський університет, кафедра інформатики, факультет управлінських наук та інформатики; Жиліна, Словаччина;

Alexander Krajčí – postgraduate student, University of Žilina, Department of Informatics, Faculty of Management Science and Informatics; Žilina, Slovakia;

e-mail: alexander.krajci@st.fri.uniza.sk

ORCID ID: <https://orcid.org/0009-0003-7951-3203>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=60174647200>

Сидоренко Людмила – Державний університет медицини та фармації імені Ніколае Тестемічану, доцент кафедри молекулярної біології та генетики людини; Кишинів, Молдова;

Ludmila Sidorenko – State University of Medicine and Pharmacy “Nicolae Testemițanu”, Associate Professor of the Department of Molecular Biology and Human Genetics; Chișinău, Moldova;

e-mail: ludmila.sidorenco@usmf.md

ORCID ID: <https://orcid.org/0000-0003-0382-4542>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=57190659134>

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

Olesia Barkovska – 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>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=24482907700>

НАДІЙНЕ ПРОГНОЗУВАННЯ СЕРЦЕВО-СУДИННОГО РИЗИКУ НА НЕВЕЛИКИХ НАБОРАХ ДАНИХ З ВИКОРИСТАННЯМ ПЕРЕДОВОЇ ІНЖЕНЕРІЇ ОЗНАК

Актуальність. Серцево-судинні захворювання залишаються провідною причиною смертності в усьому світі, що створює високий попит на автоматизовані діагностичні системи. Однак розробка надійних моделей машинного навчання для аналізу електрокардіограми (ЕКГ) часто ускладнюється наявністю лише невеликих за масштабом та незбалансованих наборів даних, що обмежує ефективність підходів глибокого навчання. **Об'єктом дослідження** є процес автоматизованої обробки та класифікації електрокардіографічних сигналів для діагностичних цілей. **Предметом дослідження** є методи

вилучення ознак на основі серцево-судинних захворювань, стратегії агрегації на рівні пацієнта та алгоритми машинного навчання для прогнозування серцево-судинного ризику. **Метою цієї статті** є розробка та оцінка надійної системи класифікації, оптимізованої для невеликих наборів даних, яка підвищує точність прогнозування шляхом використання агрегації ознак на рівні пацієнта та моделей машинного навчання, що пояснюються. **Результати дослідження.** У дослідженні пропонується конвеєр, що починається зі стандартної попередньої обробки сигналу, а потім виконується точне виявлення R-піку та сегментація на основі серцево-судинних захворювань. Фізіологічні ознаки (BCP, вейвлет, морфологічні) потім витягуються з окремих сегментів та статистично агрегуються на рівні пацієнта. Експерименти на наборі даних із 164 суб'єктів показали, що запропонована стратегія агрегації на рівні пацієнта значно перевершує традиційний аналіз на основі сегментів. Остаточна модель випадкового лісу досягла ROC-AUC балу 0,84. Аналіз важливості ознак підтвердив критичну роль показників варіабельності серцевого ритму (HRV), зокрема SDNN та RMSSD, у диференціації здорових суб'єктів та суб'єктів з високим ризиком.

Ключові слова: класифікація; прогнозування ризику; важливість ознак; машинне навчання; ЕКГ; випадковий ліс.

Бібліографічні описи / Bibliographic descriptions

Крайчі О., Сидоренко Л., Барковська О.Ю. Надійне прогнозування серцево-судинного ризику на невеликих наборах даних з використанням передової інженерії ознак. *Сучасний стан наукових досліджень та технологій в промисловості*. 2026. № 1 (35). С. 55–64. DOI: <https://doi.org/10.30837/2522-9818.2026.1.055>

Krajčí, A., Sidorenko, L., Barkovska, O. (2026), "Predicting Risks of Cardiovascular Disease on Small Datasets using Feature Engineering", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (35), P. 55–64. DOI: <https://doi.org/10.30837/2522-9818.2026.1.055>

Heorhii Kuchuk, Nataliia Kosenko, Nina Kuchuk, Viktors Gopejenko, Victor Kosenko

ADAPTIVE RESOURCE ALLOCATION METHOD FOR THE MOBILE FOG LAYER OF HIGH-DENSITY INDUSTRIAL INTERNET OF THINGS IN INDUSTRY 5.0 NETWORKS

Relevance of the article. The modern concept of Industry 4.0 laid the foundation for complete digitalization through the industrial Internet of Things. However, the transition to Industry 5.0 requires greater flexibility and resilience of systems. High-density mobile industrial IoT with a fog layer is a critical element of this transformation, as it provides not only automation but also the adaptability of production to human needs and environmental standards. **The object of study** is the process of pre-processing transactions of the HDIoT edge layer. **The main hypothesis of the study:** the implementation of a new adaptive method of resource allocation for mobile devices of fog clusters will reduce the average pre-processing time of transactions of the HDIoT edge layer. **The goal of the work** is to reduce the average time a transaction of the HDIoT peripheral layer spends in the fog layer by developing an adaptive method for distributing the resources of mobile devices in fog clusters. **Research objectives:** to identify the architectural features of fog computing in HDIoT networks; to create a mathematical model of the process of optimal resource allocation for mobile cluster devices in the fog layer; to formalize a multi-agent approach to cluster resource allocation; to develop and investigate a theoretical game model for managing the resources of a mobile fog cluster of a multi-layer IoT. **Methods used:** multi-agent approach, game theory, in particular, optimization of a cooperative stochastic game, computer modeling. **Results.** An adaptive method for distributing resources of mobile devices in fog clusters has been developed. Within the framework of the method, the architecture of a mobile fog cluster has been proposed and a mathematical model of the process of optimal distribution of its resources has been created. In addition, a multi-agent approach is used to find an approximate solution to the formulated two-parameter nonlinear optimization problem, and a game-theoretical approach is implemented to reduce computational complexity and accelerate the search for an approximate solution. **Conclusion.** As a result of applying the developed method, the average time a transaction of the peripheral layer of a high-density IoT spends in the fog layer has been reduced, which, given the high density of mobile devices, has made it possible to meet QoS requirements.

Keywords: Industry 5.0; Internet of Things; fog cluster; mobile node; computer system; agent-based approach; intelligent agent; stochastic game.

1. Introduction

The Internet of Things (IoT) is one of the key paradigms for the development of modern information and communication technologies, enabling the integration of physical objects into the digital space [1]. Thanks to the IoT, billions of sensors, actuators, and embedded systems are capable of collecting, transmitting, and processing data in near real time [2]. This opens up vast opportunities for building intelligent environments in various areas, such as industry [3], transportation [4], healthcare [5], energy [6], and others. IoT systems generate enormous amounts of data, which in some cases requires instant collection, transmission, and analysis for operational decision-making [7]. Traditional cloud computing, despite its power, faces critical delays when transporting traffic from the periphery to remote data centers [8]. In addition, network bandwidth limitations and data transfer delays reduce the efficiency of centralized processing [9]. That is why there is an urgent

need to decentralize computing power and bring it closer to the sources of information.

To overcome these problems, a fog computing layer is actively used, which is located between peripheral devices and the cloud [10]. Fog computing allows part of the data processing to be moved closer to its source. This reduces delays, increases reliability, and reduces the load on backbone networks [11]. In addition, the fog layer provides better support for mobility and geographic distribution, which is critical for modern dynamic systems [12]. The use of fog node resources enables the implementation of real-time computing, which is vital for industrial automation and critical infrastructure [13]. Thus, the fog layer becomes a key element of scalable IoT systems.

Fog computing is particularly relevant in the context of High-Density IoT (HDIoT) [14]. High-density networks are characterized by an extremely large number of active nodes in a limited area. For example, the concept of high-density Industrial IoT (IIoT) involves the deployment of thousands of sensors in a limited area,

which creates a huge load on the network. The addition of a Mobile Fog Layer (MFL) allows data to be processed directly near the source, even if that source or computing node is constantly moving. For example, this could be a large automated warehouse covering 50,000 m², where thousands of pallets with RFID tags, vibration and temperature sensors, and hundreds of autonomous mobile robots are operating simultaneously. Each unit of goods and each rack is equipped with sensors. Due to their large number (high density), transferring all data directly to the cloud would cause delays (latency) and overload communication channels. The fog node (robot) makes most decisions without connecting to the cloud, coming to the sensors itself, which allows the sensors to operate at very low power (saving battery life).

Mobile fog environments HDIoT are characterized by high transaction intensity, dynamic topology, and traffic heterogeneity [15]. The huge number of simultaneous requests from sensors and actuators requires the infrastructure to be highly adaptable and capable of rapid scaling [16]. High-density conditions make centralized management virtually impossible due to the exponential growth in coordination complexity. In addition, high-density IoT significantly increases the requirements for response speed and data processing stability [17]. At the same time, the resources of individual fog nodes remain limited. This makes it difficult to ensure guaranteed quality of service.

Under such conditions, there is a need for territorial clustering of fog layer devices. Clustering allows fog nodes and peripheral devices to be grouped according to spatial proximity and load characteristics [18]. Division into territorial segments helps to localize traffic within a certain area, preventing overload of the entire network. This contributes to more efficient use of available computing and network resources [19]. Territorial clusters also simplify data flow management and load balancing [20].

A complicating factor is that a mobile fog layer cluster functions as a decentralized system with limited resources [21]. The mobility of nodes leads to constant changes in available resources and network connections. The lack of a single control center forces nodes to self-organize and make decisions based on local information [22]. In such conditions, static approaches to resource management prove ineffective. The system must adapt to the current state of the cluster and the intensity of transaction flow.

Therefore, preprocessing transactions at the edge layer requires a change in traditional approaches to

resource allocation. Adaptive allocation methods focused on the dynamics of mobile devices in a fog cluster are becoming a prerequisite for the effective functioning of a high-density Internet of Things. This determines the relevance of research on the adaptive allocation of fog cluster resources in modern HDIoT systems.

2. Analysis of literature and problem statement

Recently, there have been many works addressing the issue of optimal distribution of communication and computing resources in mobile systems. Article [23] addresses the problem of optimizing computations at the network edge. The main optimization task is to minimize the total energy consumption of mobile devices, taking into account the current requirements of Internet of Things transactions. However, the issue of optimal resource allocation is not addressed, as in the classical approach, the nearest available resource is always allocated. In [24], the problem of minimizing the total delay of HDIoT transactions that arises during data processing is considered. An algorithm for making decisions on resource allocation is presented, taking into account the presence of several base stations and computing servers in the network. However, the HDIoT mobile cluster is a decentralized system. In [25], the task of minimizing delay is considered for decentralized systems, but the proposed approach does not allow for the requirements of HDIoT transactions to be taken into account. In [26], procedures are given for organizing resource allocation in decentralized computing networks that use a weighted utility function for energy and transaction delays. However, these procedures are also not focused on optimal resource allocation. To solve the problem of dynamic changes in the mobile IoT environment, there are also other computing models described in [27, 28], but they also do not take into account all the requirements of HDIoT transactions.

In [29], the process of data transfer between the edge and fog layers of the IoT ecosystem is investigated. The proposed method is effective for managing the data transfer process. However, issues related to the mobility of fog devices, energy consumption for data transfer, and the high density of IoT devices remain unresolved. This is due to restrictions imposed on the number of elements in a virtual cluster. In [30], there are no such restrictions on the number of cluster elements. However, as in [29], this algorithm does not take into account the characteristics of mobile devices.

This approach is proposed in [31, 32], which present the results of research on the transmission of data using mobile devices. Work [31] proposes a resource planning method for the mobile Internet of Things, focused on energy efficiency. The method of forming a fog layer cluster developed in work [32] is also focused on mobile IoT devices. However, these works leave unresolved the issue related to the high density of IoT devices. The main reason is the significant increase in the computational complexity of the proposed algorithms as the density of IoT devices increases.

In [33], data transmission is controlled using a multi-criteria optimization problem. The authors propose a fast resource allocation algorithm for many nodes based on a deterministic gradient. However, this algorithm is aimed at minimizing the age of information in the mobile Internet of Things. A similar problem is found in [34], which analyzes in detail the problems of data transmission when performing mobile computing for the Internet of Things. The proposed algorithm improves the balance between system latency and energy consumption, which reduces transmission delays. However, as in [33], the issue of optimal resource allocation for mobile devices takes a back seat.

Thus, all the analyzed works leave unresolved issues related to the optimal distribution of limited resources of the fog layer of the high-density Internet of Things. resources of the fog cluster of the high-density Internet of Things.

This gives reason to argue that it is expedient to conduct research aimed at reducing the time indicators associated with the transmission and pre-processing of information sent from the IoT periphery to the fog layer.

3. Purpose and objectives of the study

The purpose of the study is to reduce the average time a HDIoT peripheral layer transaction spends in the fog layer by developing an adaptive method for allocating resources of mobile devices in fog clusters. This will make it possible to meet quality of service (QoS) requirements even with a high density of mobile devices.

The object of the study is the process of preprocessing transactions of the HDIoT peripheral layer. The main hypothesis of the study is that the implementation of a new adaptive method for allocating resources of mobile devices in fog clusters will reduce the average preprocessing time of transactions in the HDIoT peripheral layer.

The following conditions were used as a basis for developing the method.

Condition 1. Mobile fog devices belonging to the same HDIoT territorial cluster are considered.

Condition 2. Mobile fog devices of the territorial cluster are considered as a completely decentralized system.

Condition 3. The peripheral layer transaction is transmitted to the nearest available mobile fog device.

Condition 4. Additional computing resources of another device located within the range of the base device may be used to pre-process the transaction.

To achieve the goal, the following tasks were set:

- to identify the architectural features of fog computing in HDIoT networks;
- to develop a mathematical model of the process of optimal resource allocation for mobile fog cluster (MFC) devices;
- to formalize a multi-agent approach for MFC resource allocation;
- to develop a theoretical game model for managing MFC HDIoT resources;
- to describe in the form of a sequence of steps and investigate the developed adaptive method for distributing resources of mobile devices in fog clusters.

4. Development and investigation of an adaptive method for distributing resources of mobile devices in fog clusters

4.1. Architectural features of fog computing in HDIoT networks

The development of the Internet of Things (IoT) concept has led to a transition to a three-tier architecture through the introduction of a fog layer [35].

This layer acts as an intermediary between end devices and cloud data processing centers (DPC), where the IoT level is responsible for data collection and the cloud is responsible for long-term storage. Fog layer nodes provide pre-processing and instant transmission of critical information to operational control points. Vertical integration of such computing significantly improves quality of service (QoS) [36], ensuring minimal response time, reduced load on communication channels, and increased system fault tolerance through decentralization.

Architecturally, the fog layer is based on the principles of modularity and virtualization [37], which allows for flexible system scaling. The hardware platform includes computing, memory, security, and management modules, as well as containerization tools for service

deployment [38]. Communication with data centers is maintained through a distributed decentralized core that combines the resources of IoT devices into a single aggregation network. In such conditions, the clustering of spatially dispersed nodes becomes critical, optimizing the use of network resources and increasing the energy efficiency of the system.

Promising IoT networks are characterized by high node mobility, which involves placing fog capabilities on mobile platforms, including unmanned aerial vehicles. In such systems, fog devices can act as both last-mile network components and user terminals, which facilitates network slicing. The formation of mobile clusters is based on critical parameters: the speed vector of devices or the time of task execution. For ultra-high-density architectures with mobile components, a four-level hierarchy is mandatory: cloud level (DPC), fog layer core servers, clustered fog devices, and the lower level of direct information collection.

A fragment of the ultra-dense IoT architecture with a dedicated mobile fog layer cluster is shown in Fig. 1.

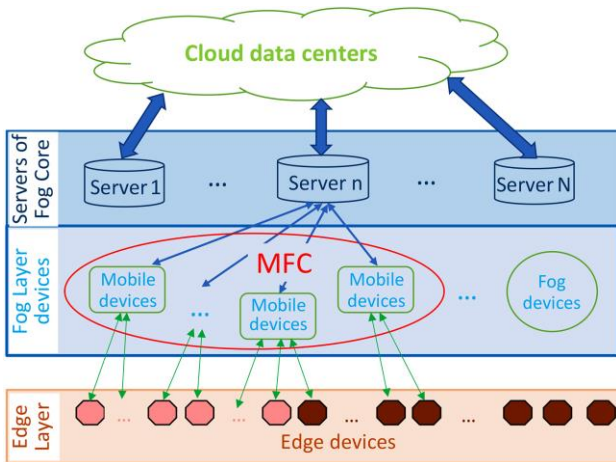


Fig. 1. MFC's place in the vertical structure of HDIoT

One of the main factors affecting the efficiency of HDIoT is the quality of clustering of mobile fog devices. For mobile clusters, stability indicators, which are analyzed in detail in article [39], are essential. However, given the limited computing resources of mobile devices, time criteria are no less important, especially for operational transactions of the peripheral layer. Among the many time criteria for evaluating the quality of multi-node decentralized mobile systems, the criterion of minimizing the maximum data processing delay has been selected [40, 41]. This criterion helps to reduce the time transactions spend in the fog layer.

4.2. Mathematical model

of the optimal MFC resource allocation process

A set of mobile computing nodes (Mobile Fog Computing Node, MFCN) MFC $N = \{N_1, N_2, \dots, N_k, \dots, N_K\}$, represented by K mobile devices, card $N = K$, is considered. Each node $N_k \in N$ has the ability to accept transactions from the edge layer of HDIoT, which, depending on the source, constitute a set of types $M = \{M_1, M_2, \dots, M_s, \dots, M_S\}$, card $M = S$. During the operation of HDIoT, the flow of transactions of type M_s arriving at MFC is distributed according to the Poisson distribution with an average intensity of λ_s , and the average intensity of task processing at the k -th node is μ_k .

Let node N_k accept transaction HDIoT M_s , which is characterized by the following parameters:

G_s – computing resource required to execute transaction M_s ;

B_s – size of transaction M_s .

When processing this transaction, node N_k provides the required computing resource at a rate of f_k . Then, the delay in computing during the execution of transaction M_s on node k is determined as follows:

$$t_{k,s} = G_s / f_k. \quad (1)$$

Let the following conditions be satisfied:

$$f_k \cdot t_{k,s} \leq F_{k,\max}; \quad (2)$$

$$\lambda_s < \mu_k, \quad (3)$$

where $F_{k,\max}$ – the maximum possible amount of computing resources that node N_k can provide per unit of time.

When these conditions are met, the transaction can be fully resolved on the node that accepted it, and the time the transaction spends on this node is determined as

$$T_k(M_s) = t_{k,s} \cdot \left(\frac{1}{\mu_k - \lambda_s} \right). \quad (4)$$

However, if at least one of conditions (2) and (3) is not met, the computing resources of node k are insufficient to execute the transaction, let node j be selected as the auxiliary node. Let the share of transaction M_s processed on node k be denoted as $z_{k,s}$, $0 \leq z_{k,s,j} \leq 1$. When conditions (2) and (3) are met, $z_{k,s,j} = 1$, and if $z_{k,s,j} = 0$, the transaction is completely transferred to another mobile node j for execution.

In the case of $0 < z_{k,s,j} < 1$, the transaction is executed jointly by mobile nodes k and j . The delay in transmitting the necessary information is calculated using the formula

$$\tau_{k,s,j} = (1 - z_{k,s,j}) \cdot \frac{B_s}{C_{k,j}}, \quad (5)$$

$$T_k(M_s | z_{k,s,j}) = z_{k,s,j} \cdot t_{k,s} \cdot \left(\frac{1}{\mu_k - z_{k,s,j} \cdot \lambda_s} \right) + (1 - z_{k,s,j}) \cdot \left(\tau_{k,s,j} + t_{j,s} \cdot \left(\frac{1}{\mu_j - (1 - z_{k,s}) \cdot \lambda_s} \right) \right). \quad (6)$$

Let us introduce a set of possible states of the system that satisfy conditions (1)–(5), depending

$$\Gamma = \{\gamma_\eta\}, \quad \gamma_\eta = \{z_{k,s,j_s}\}, \quad \text{card}(\gamma_\eta) = S, \quad s \in \overline{1, S}; k_s, j_s \in \overline{1, K},$$

where γ_η – η -th option of fixed distribution of a set of transactions between computing nodes of a mobile cluster.

Using expression (6) and the criterion of minimizing the maximum transaction processing delay, we can formulate the objective function of the problem of optimal distribution of MFC computing resources:

$$\gamma_\eta^* = \arg(T(\gamma_\eta) | T(\gamma_\eta) < T(\gamma_\eta^*) \quad \forall \gamma_\eta \in \Gamma; \gamma_\eta \neq \gamma_\eta^*), \quad (7)$$

where $T(\gamma_\eta) = \max_s(T_k(M_s | z_{k,s,j_s}))$, $z_{k,s,j_s} \in \gamma_\eta$.

At the same time, a number of restrictions must be observed:

$$0 \leq z_{k,s} \leq 1 \quad \forall s \in \overline{1, S}; \quad (8)$$

$$\lambda_s < z_{k,s} \cdot \mu_k, \quad \forall s \in \overline{1, S}, \forall k_s \in \overline{1, K}; \quad (9)$$

$$(1 - z_{k,s}) \cdot \lambda_s < \mu_j, \quad \forall s \in \overline{1, S}, \forall j_s \in \overline{1, K}; \quad (10)$$

$$f_{j_s} \cdot \tau_{j_s,s} \cdot (1 - z_{k,s}) \leq F_{j_s, \max}, \quad \forall s \in \overline{1, S}, \forall j_s, k_s \in \overline{1, K}. \quad (11)$$

Thus, based on the resource allocation model (expressions (8)–(11)), it is possible to optimize the

where $C_{k,j}$ – the throughput capacity of the channel between the k -th and j -th mobile nodes.

Then, in the general case, the time spent by transaction M_s in the MFC environment is calculated using the following formula:

on the distribution of transactions and the options for their processing:

process of allocating MFC computing resources relative to the criterion of minimizing the maximum transaction processing delay by minimizing the objective function (7).

4.3. Formalization of a multi-agent approach for MFC resource allocation

MFC HDIoT is a stochastic decentralized system. Therefore, to find an approximate solution to a two-parameter nonlinear optimization problem, it is advisable to use an agent-based approach.

In the work of M. Wuldrige and N. Jennings, Intelligent Agents: Theory and Practice [42], the concept of an intelligent agent (IA) was systematized for the first time in the context of multi-agent systems. The authors define a separate IA as an autonomous system capable of perceiving the environment, making decisions, and behaving purposefully, which is ensured by reactivity, proactivity, and the ability to interact socially, i.e., an IA must meet a number of specific properties. These properties in HDIoT systems correspond to mobile fog computing nodes (MFCN, Table 1).

Table 1. Compliance of mobile fog HDIoT nodes with the canonical definition of IA

IA characteristics	Canonical interpretation	Compliance of MFCN with HDIoT
Autonomy	Independent choice of actions	The node itself controls transmission, sleep modes, and offloading
Reactivity	Function of reaction to the environment	Reaction to events, channel changes, load
Pro-activeness	Purposeful behavior	Local traffic prediction, task migration, fog node change
Social ability	Communication with other agents	MQTT/CoAP/DDS, coordination with fog nodes
Situatedness	Placement in the environment	Physical location + network topology + decentralized system element
Rationality	Maximization of utility	Energy/latency/QoS optimization

Therefore, any MFCN that has accepted a peripheral layer transaction for processing can be considered as an IA, and the set of nodes with which it interacts at this time can be considered as the operating environment of this IA [43].4 For MFC devices, there is no reward

criterion for IA actions (utility function). This requires the IA to implement random actions and learn to select those that maximize the utility function. It should be noted that none of the MFC mobile devices can control the entire process and has no complete picture of the

global state of the system [44]. Therefore, MFC HDIoT cannot be classified as a fully observable environment, so a reinforcement learning approach should be used.

Single-agent reinforcement learning by trial and error shapes the behavior of the IA that will be necessary to achieve the set goal [45]. A generalized diagram of this process is shown in Fig. 2.

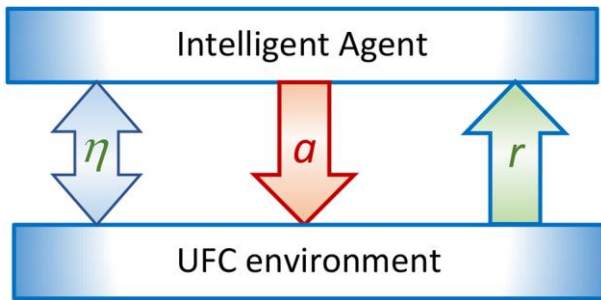


Fig. 2. Interaction IA and UFC

According to the concept of reinforcement learning, the AI interacts with the environment, having some observation η , on the basis of which it performs action a and receives reward r , which depends on the results of the impact on the environment. The reward received in accordance with the reinforcement learning algorithm increases or decreases the probability of performing the action under the same conditions in the future.

The formal model of such interaction is given by the tuple:

$$\theta = \langle S, A, H, R \rangle, \quad (12)$$

where set $S = \{s\}$ is the set of all possible states of the environment; set $A = \{a\}$ is the set of all possible actions of the IA; $H = \{h\}$ is the set of all possible observations of the environment available to the computing node; $R = \{r\}$ is the set of all possible values of the IA's gain, where each element is a reward function.

$$r = r(s_t, a_t, s_{t+1}): S \times A \times S \rightarrow R, \quad (13)$$

which is determined at a discrete moment in time t and which is obtained by the IA when performing action a_t in state s_t of the environment and transitioning the environment to state s_{t+1} .

Tuple θ elements at a discrete moment in time t determine the transition function

$$v = v(s_t, a_t, s_{t+1}): S \times A \rightarrow [0;1], \quad (14)$$

which returns the probability of the environment transitioning from state s_t to state s_{t+1} when the agent performs action a_t , and the observation function

$$v = v(s_t, a_t, s_{t+1}): S \times A \rightarrow [0;1], \quad (15)$$

which returns the probability of an agent receiving observation o_t in state s_t .

When considering several IA models (12)–(15), it is transformed into multiagent reinforcement learning (MRL) [46]. When transitioning from single-agent reinforcement learning to multiagent reinforcement learning, the corresponding mathematical model transitions from scalar values to vectors, which are elements of the corresponding sets.

The number of states of the environment S grows exponentially with the increase in the number of IA, because the states of the environment, except for state σ (without mobile nodes that have become intelligent agents), are supplemented by the states of the set of IA. Thus, in the tuple (48)

$$S = \{s\} = \left\{ \left\langle \sigma, s^{(1)}, \dots, s^{(i)}, \dots, s^{(n_k)} \mid s^{(i)} \in S_k, i = \overline{1, n_k} \right\rangle \right\}, \quad (16)$$

where S_k – the set of possible states of mobile node k , which belongs to the set of intelligent agents, $\text{card } S_k = n_k$.

The set of all possible actions of an IA becomes a set of action tuples, i.e.

$$A = \{a\} = \left\{ \left\langle a_i \mid a_i \in A_k \right\rangle \right\}, \quad (17)$$

where each tuple is a set of actions of a separate IA of the corresponding mobile computing node N_k , which form the corresponding set of actions A_k .

Similarly, when transitioning from single-agent to multi-agent reinforcement learning, sets H and R are formed with the corresponding tuples:

$$H = \{h\} = \left\{ \left\langle h_i \mid h_i \in H_k \right\rangle \right\}, \quad (18)$$

$$R = \{r\} = \left\{ \left\langle r_i \mid r_i \in R_k \right\rangle \right\}. \quad (19)$$

The power of the sets $A_k = \{a_i\}$ of possible actions of IA, which is the k -th mobile computing node, increases due to the additional capabilities of interaction between agent nodes. The number of components of a single element and $h_i \in H_k$ possible observations of the environment of the k -th IA increases because the nodes observe not only the state of the environment without intelligent agents, but also the states of other agents.

Thus, multi-agent reinforcement learning (MARL), described by formal models (12), (16)–(19) and

functions (13)–(15), allows us to find an approach to solving the optimization problem (7)–(11). Given the characteristics of MFC HDIoT, it is necessary to take into account such features of the process as complete decentralization, heterogeneity, the possibility of simultaneity, and individualization of gains [47]. Due to the absence of information in the observations about the actions that other computing nodes plan to take, tasks were selected without IA knowledge of the planned actions of other agents. Based on the simultaneity of actions a_i , a multi-agent reinforcement learning task was selected, where the actions of all computing nodes are performed simultaneously. Based on the criterion of individualization of gains, tasks with different individual gains were selected. In order to have information generated by other agents in the observation, tasks with communication were selected. In order to have homogeneity of the sets of available actions and observations q , tasks with physically heterogeneous computing nodes were selected.

4.4. Theoretical game model of MFC HDIoT resource management

Mobile computing nodes (MCN) MFC can pursue different interests. Therefore, the above-considered MFC functioning scheme, taking into account MARL, does not fully reflect the dynamics of its functioning. In order to refine this scheme, it is proposed to move to a modification of the MARL scheme, within which a cooperative stochastic game (CSG) is implemented.

The priority is calculated as a win for the AI, which allows MCN tasks to be redistributed according to their priority. Thus, the higher the priority of the MCN, the faster its task will be assigned for distribution and solution in the system.

Let $z^{(k)} = (z_1, \dots, z_{k-1}, z_{k+1}, \dots, z_K)$ is the decision to provide its resources by all nodes except node k . Taking into account the decisions $z^{(k)}$ for other nodes, computing node k would seek to choose the correct decision z_k regarding whether to provide its resources to other nodes in order to achieve a common goal, so as to minimize resource costs and the time the task spends in the network, i.e.

$$\min_{z_k} \psi_k(z_k, z^{(k)}) \forall k \in \overline{1, K}, \quad (20)$$

where $z_k = \emptyset$, if node k does not provide computing resources to other nodes, in this case $\psi_k(z_k, z^{(k)}) = E^{(k)}$ –

computing resource costs when performing a transaction independently; $z_k = \emptyset$ if node k is ready to provide its own resources, in this case $\psi_k(z_k, z^{(k)}) = Y^{(k)}$ – computing resource costs when using the assistance of other mobile nodes.

Let's formulate problem (20) as a strategic game

$$\mathfrak{G} = (K, \{F'_k\}, \{\psi_k\}) \forall k \in \overline{1, K}, \quad (21)$$

where the set of K computing nodes is considered as a set of players; F'_k is a set of strategies for user k ; $\psi_k(z_k, z^{(k)})$ is considered as a service function, which is a cost function that player k must minimize. The game \mathfrak{G} will perform a decentralized distribution of resources. The solution to the game lies in the region of rational solutions to the optimization problem (7)–(11) that are close to the optimal. The solution to this game is an equilibrium point that can be found using Nash equilibrium [48, 49]. Let there be a strategy profile

$$z^* = (z_1^*, z_2^*, \dots, z_K^*), \quad (22)$$

which is a Nash equilibrium in a game with decentralized resource allocation, if, in a state of equilibrium, no computing node can further reduce its own computing resource costs by unilaterally changing its own strategy, i.e.,

$$\psi_k(z_k^*, z^{*(k)}) \leq \psi_k(z_k, z^{*(k)}) \forall z_k \in F'_k, k \in \overline{1, K}, \quad (23)$$

Nash equilibrium has the property of self-stability, so computing nodes that are in equilibrium can reach a mutually acceptable solution, and no node has an incentive to deviate. Considering the strategies $z^{(k)}$ of other players, the strategy $z^{*(k)} \in F'_k$ of the k -th player is the best if

$$\psi_k(z_k^*, z^{*(k)}) \leq \psi_k(z_k, z^{(k)}) \forall z_k \in F'_k, k \in \overline{1, K}. \quad (24)$$

According to (23) and (24), in Nash equilibrium, all players use the best response strategies in relation to each other. Due to the property of Nash equilibrium, no player has an incentive to deviate from the decisions reached. Thus, the iterative algorithm for finding the Nash equilibrium of the game allows us to quickly find a resource allocation that is close to optimal in terms of minimizing the maximum data processing delay.

4.5. Research on an adaptive method for distributing resources among mobile devices in fog clusters

The results obtained in sections 4.1–4.4 allow us to propose an adaptive method for distributing resources

among mobile devices in fog clusters, which can be described as a sequence of the following steps:

Stage 0. Determination of initial time characteristics $t = t_0$ and discretization step $\Delta t = \Delta t_0$, fixation of a stable core of a fuzzy cluster.

Stage 1. Determination of average values of key characteristics of mobile devices and HDIoT transaction flows using a mathematical model of the MFC optimal resource allocation process.

Stage 2. Determination of sets of possible states of a fog cluster and possible actions of mobile nodes that accept HDIoT transactions as intelligent agents using the proposed multi-agent approach for MFC resource allocation.

Stage 3. If $\text{card}(\Gamma) \neq \emptyset$, then the transition to the next stage of the method is carried out. Otherwise, the discretization step must be reduced by 2 times. If the obtained discretization step value is less than the minimum permissible value, then control is transferred to the classical method of resource allocation of the fuzzy cluster of the high-density Internet of Things.

Stage 4. Use of the MFC HDIoT resource management theoretical game model, which implements a cooperative stochastic game based on the data generated in the previous stages of the method.

Stage 5. The distribution of fog cluster resources is carried out based on game equilibrium strategies.

Stage 6. We form the beginning of the next interval $t = t + \Delta t$.

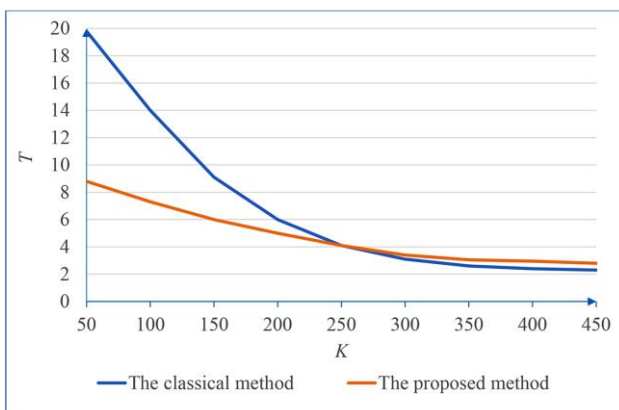
Stage 7. Correction of multiple possible states of a fog cluster and possible actions of mobile nodes that accept HDIoT transactions. Transition to stage 3.

Unlike the methods discussed above, the proposed method does not require the use of cloud technologies and remote server-type nodes. The method involves solving computational tasks directly on mobile devices in the fog layer. To do this, the computing resources of devices located in the same cluster are used. If a mobile device lacks its own resources, such as memory or processor time, it redirects its task or part of it to another device or devices that have the necessary resources to solve it. To evaluate the performance of the method, the criterion of minimizing the maximum data processing delay is used. Therefore, the worst delay caused by data transfer and processing should be as small as possible.

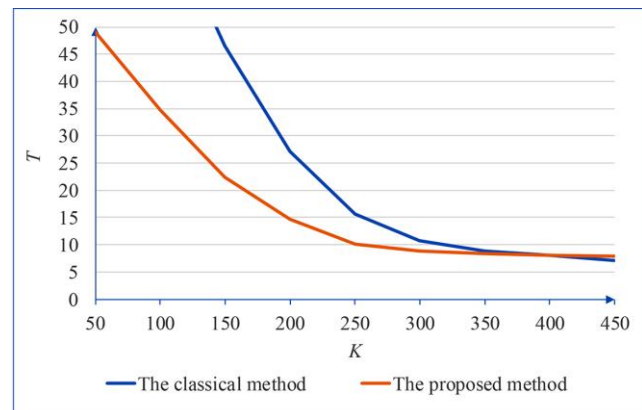
To evaluate the effectiveness of the proposed resource allocation method, a simulation model of the MFC HDIoT mobile cluster operation was used. The average time a transaction from the edge layer spends in the HDIoT fog layer during a fixed time interval was chosen as the performance indicator. To compare transaction processing methods, two methods were modeled:

- the classic method, which distributes the transaction for execution to a mobile device that has sufficient computing resources to execute it and sent the first response to the request; the transaction is fully processed by this device;
- the proposed method.

The generalized results of modeling the dependence of the transaction residence time in the fog layer on the number of mobile devices in the cluster and the intensity of transaction arrival are shown in Fig. 3.



a) $\lambda = 100$ trans./s



b) $\lambda = 250$ trans./s

Fig. 3. Dependence of the transaction residence time in the fog layer (T, s) on the number of mobile devices in the cluster (K) and the intensity of transaction arrival

In high-density IoT, even stable mobile fog clusters have a variable structure with some stable core. Therefore, important characteristics when analyzing the dynamic stability of a cluster are both the intensity of adding new nodes to the system and the intensity of nodes leaving the cluster. Fig. 4 shows the dependence of the average transaction residence time in the fog layer on the ratio of the intensity of adding nodes to the intensity of nodes leaving the cluster.

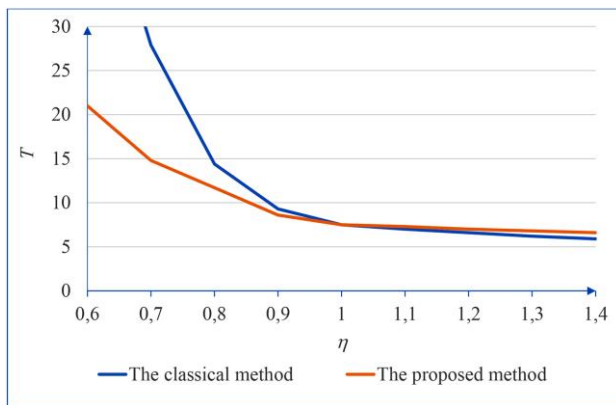


Fig. 4. Dependence of the transaction residence time in the fog layer (T, s) on the ratio of node addition intensity to node exit intensity (η)

When modeling the dependence shown in Fig. 4, system parameters were selected for which, for a constant fog cluster structure, the time value T coincides for both methods under analysis: $\lambda = 250$ trans./s; $K = 350$ nodes.

5. Discussion of the results of research on an adaptive method for distributing resources among mobile devices in fog clusters

The architecture of the high-density IoT ecosystem subsystem has been developed, which performs preprocessing of peripheral layer transactions on fog layer mobile devices (Fig. 1). A distinctive feature of this architecture is its focus on the cluster structure of the fog layer. After receiving a transaction, the fog mobile device, unlike the classical method, can select another available node to participate in the preprocessing of the transaction. This allows for more flexible use of the limited computing resources of fog layer devices.

The proposed mathematical model of the MFC optimal resource allocation process allows minimizing the maximum delay in preprocessing transactions of the HDIoT peripheral layer. A distinctive feature of this

model is that it takes into account the possibility of distributing preprocessing between fog mobile devices. The use of an agent-based approach made it possible to take into account the decentralized structure of the fog mobile cluster. The use of the mathematical apparatus of cooperative stochastic games made it possible to significantly speed up the search for an approximate solution to the formulated optimization problem (7)–(11) in the context of implementing an agent-based approach for a decentralized system.

The evaluation of the effectiveness of the proposed method (Figs. 2–4) showed the following results:

- when the mobile cluster is loaded by no more than 30%, the proposed method is not significantly inferior to the classical one in terms of time indicators;
- when the mobile cluster is loaded by more than 50%, the proposed method provides better results in terms of time indicators than the classical one; at the same time, the higher the load of the mobile cluster devices, the better the time indicators become when using the proposed method;
- in the case of an unstable mobile cluster structure, the proposed method provides better results in most cases, especially with a relative decrease in the number of mobile devices.

The results of the study can be explained by a more rational use of the limited resources of mobile fog devices, which reduces the volume of transaction queues to the fog layer.

Unlike [23, 24], which proposes optimization algorithms for centralized systems, the proposed method is focused on decentralized systems. In addition, the proposed method, unlike [25–28], takes into account the requirements of HDIoT transactions.

Unlike [29, 30], which uses methods for controlling the IoT data transmission process, the proposed method takes into account the characteristics of mobile devices. Also, unlike [31, 32], where methods for resource planning for the mobile Internet of Things are developed, the proposed method solves issues related to the high density of IoT devices by reducing computational complexity. Unlike [33, 34], where the energy consumption of mobile devices is optimized, the proposed method is also focused on the optimal distribution of limited resources.

Thus, the developed adaptive method for allocating resources of mobile devices of fog clusters of high-density Internet of Things made it possible to reduce the average transaction residence time of the HDIoT

peripheral layer in the fog layer. This made it possible to meet QoS requirements even with a high density of mobile devices.

However, the proposed results should be applied with the following restrictions:

- high density of IoT mobile devices;
- presence of territorial clustering of fog devices.

It is also advisable to use the proposed method when the average load of fog layer devices is at least 50%.

A drawback of this study is that it does not consider the analysis of the possible interaction of mobile virtual clusters of the HDIoT fog layer. To eliminate this, the simulation model of the fog layer of the HDIoT support ecosystem should be expanded with the appropriate procedure.

6. Conclusions

An adaptive method for distributing the resources of mobile devices in fog clusters has been developed, which has made it possible to reduce the average time a HDIoT peripheral layer transaction spends in the fog layer. The following results were obtained in the process of developing the method:

- the architecture of the high-density IoT ecosystem subsystem has been formed, which pre-processes peripheral layer transactions on fog layer mobile devices and allows for more flexible use of the limited computing resources of fog layer devices;
- a mathematical model of the MFC optimal resource allocation process was proposed, which allows minimizing the maximum delay in preprocessing HDIoT peripheral layer transactions. A distinctive feature of this model is that it takes into account the possibility of distributing preprocessing between fog mobile devices;
- a multi-agent approach was used to find an approximate solution to the formulated two-parameter

nonlinear optimization problem, which made it possible to take into account the decentralized structure of the fog mobile cluster;

– a theoretical game model for managing the resources of a mobile fog cluster of a multi-layer IoT has been developed and researched. The use of its mathematical apparatus has made it possible to significantly speed up the finding of an approximate solution to the formulated optimization problem in the context of implementing an agent-based approach for a decentralized system.

The proposed method can be used to create flexible and scalable production facilities in accordance with the Industry 5.0 concept, such as a smart logistics hub with autonomous robots. As a development of this research, the development of a method for priority processing of HDIoT operational transactions can be considered.

Conflict of interest

The authors declare that they have no conflict of interest with respect to this research, including financial, personal, authorship, or other conflicts that could influence the research and its results presented in this article.

Funding

The study was conducted without financial support.

Data availability

The manuscript has no associated data.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technology in the creation of this work.

References

1. Fatlawi, A., Al Dujaili, M.J. (2023), "Integrating the Internet of Things (IoT) and Cloud Computing Challenges and Solutions: A Review", *AIP Conference Proceedings*, Vol. 2977(1), 020067. DOI: <http://dx.doi.org/10.1063/5.0181842>
 2. Kuchuk, H., Mozhaiev, O., Tiulieniev, S., Mozhaiev, M., Kuchuk, N., Lubentsov, A., Onishchenko, Yu., Gnusov, Yu., Brendel, O., Roh, V. (2025), "Devising a method for energy-efficient control over a data transmission process across the mobile high-density Internet of Things", *Eastern European Journal of Enterprise Technologies*, Vol. 4(4(136)), pp. 46–57. DOI: <https://doi.org/10.15587/1729-4061.2025.336111>
 3. Pardo, C., Wei, R., Ivens, B.S. (2022), "Integrating the business networks and Internet of Things perspectives: A system of systems (SoS) approach for industrial markets", *Industrial Marketing Management*, Vol. 104, pp. 258–275. DOI: <https://doi.org/10.1016/j.indmarman.2022.04.012>
-

4. Ajay, P., Nagaraj, B., Pillai, B.M., Suthakorn, J., Bradha, M. (2025), "Intelligent ecofriendly transport management system based on IoT in urban areas", *Environment Development and Sustainability*, Vol. 27(10), pp. 24127–24134. DOI: <https://doi.org/10.1007/s10668-021-02010-x>
5. Lin, Y., Lin, C.-C., Chang, C.-C., Chang, C.-C. (2025), "An IoT-Based Electronic Health Protection Mechanism With AMBTC Compressed Images", *IEEE Internet of Things Journal*, Vol. 12(3), pp. 2430–2444. DOI: <https://doi.org/10.1109/JIOT.2024.3467152>
6. Gupta, S., Varshney, R., Tiwari, D.K., Varshney, T., Shukla, P.K. (2025), "Role of IoT and IIoT in Energy System Automation", *Optimizing Automation in Engineering with Energy Systems and Communication Networks*, pp. 1–32, IGI Global Scientific Publishing. DOI: <https://doi.org/10.4018/979-8-3373-2737-2.ch001>
7. Kuchuk, N., Kashkevich, S., Radchenko, V., Andrusenko, Y., Kuchuk, H. (2024), "Applying edge computing in the execution IoT operative transactions", *Advanced Information Systems*, Vol. 8, No. 4, pp. 49–59. DOI: <https://doi.org/10.20998/2522-9052.2024.4.07>
8. Hu, N. (2024), "Internet of things edge data mining technology based on cloud computing model". *International Journal of Innovative Computing, Information and Control*, Vol. 20(6), pp. 1749–1763. DOI: <http://doi.org/10.24507/ijicic.20.06.1749>
9. Qayyum, T., Trabelsi, Z., Waqar Malik, A., Hayawi, K. (2022), "Mobility-aware hierarchical fog computing framework for Industrial Internet of Things", *Journal of Cloud Computing*, Vol. 11, article number 72. DOI: <https://doi.org/10.1186/s13677-022-00345-y>
10. Thomas, P., Jose, D.V. (2023), "Towards Computation Offloading Approaches in IoT-Fog-Cloud Environment: Survey on Concepts, Architectures, Tools and Methodologies", *Lecture Notes in Networks and Systems*, 613 LNNS, pp. 37–52. DOI: https://doi.org/10.1007/978-981-19-9379-4_4
11. Kuchuk, H., Husieva, Y., Novoselov, S., Lysytsia, D., Krykhovetskyi, H. (2025), "Load Balancing of the layers Iot Fog-Cloud support network", *Advanced Information Systems*, Vol. 9, No. 1, pp. 91–98. DOI: <https://doi.org/10.20998/2522-9052.2025.1.11>
12. Bhajantri, L.B., Gangadharaiah, S. (2023), "Heuristic-Based Resource Allocation for Internet of Things in Gateway Centric Multi-layer Fog Computing", *Lecture Notes in Networks and Systems*, Vol. 516, pp. 567–579. DOI: https://doi.org/10.1007/978-981-19-5221-0_54
13. Rezanov, B., Kuchuk, H. (2023), "Model of elemental data flow distribution in the Internet of Things supporting Fog platform", *Innovative Technologies and Scientific Solutions for Industries*, Vol. 2023(3), pp. 88–97. DOI: <https://doi.org/10.30837/ITSSI.2023.25.088>
14. Lee, B.M. (2025), "Efficient Resource Management for Massive MIMO in High-Density Massive IoT Networks", *IEEE Transactions on Mobile Computing*, Vol.24(3), pp. 1963–1980. DOI: <https://doi.org/10.1109/TMC.2024.3486712>
15. Kuchuk, H., Kalinin, Y., Dotsenko, N., Chumachenko, I., Pakhomov, Y. (2024), "Decomposition of integrated high-density IoT data flow", *Advanced Information Systems*, Vol. 8, No. 3, pp. 77–84. DOI: <https://doi.org/10.20998/2522-9052.2024.3.09>
16. Lee, B.M. (2024), "Leveraging Massive MIMO for Enhanced Energy Efficiency in High-Density IoT Networks", *Mathematics*, Vol. 12(22), 3539. DOI: <https://doi.org/10.3390/math12223539>
17. Jang, H.-C., Li, T.-C. (2024), "Enhancing Edge Computing in High-Density IoT for Improved Service Quality and Privacy Protection", *Iet Conference Proceedings*, Vol. 22, pp. 142–143. DOI: <https://doi.org/10.1049/icp.2024.4321>
18. Kuchuk, H., Mozhaiev, O., Tiulieniev, S., Mozhaiev, M., Kuchuk, N., Tymoshchuk, L., Lubentsov, A., Gnusov, Y., Klivets, S., Kuleshov, A. (2025), "Devising a method for stabilizing control over a load on a cluster gateway in the internet of things edge layer", *Eastern-European Journal of Enterprise Technologies*, Vol. 2(9(134)), pp. 24–32. DOI: <https://doi.org/10.15587/1729-4061.2025.326040>
19. Abkenar, F.S., Khan, K.S., Jamalipour, A. (2021), "Smart-Cluster-Based Distributed Caching for Fog-IoT Networks", *IEEE Internet of Things Journal*, Vol. 8(5), pp. 3875–3884, 9205197. DOI: <https://doi.org/10.1109/JIOT.2020.3026322>
20. Kuchuk, H., Mozhaiev, O., Tiulieniev, S., Mozhaiev, M., Kuchuk, N., Tymoshchuk, L., Lubentsov, A., Onishchenko, Y., Gnusov, Y., Tsuranov, M. (2025), "Devising a method for increasing data transmission speed in monitoring systems based on the mobile high-density Internet of Things", *Eastern-European Journal of Enterprise Technologies*, Vol. 3(4(135)), pp. 52–61. DOI: <https://doi.org/10.15587/1729-4061.2025.330644>
21. Zheng, Z., Nazif, H. (2024), "An Energy-aware Technique for Resource Allocation in Mobile Internet of Thing (MIoT) Using Selfish Node Ranking and an Optimization Algorithm", *IETE Journal of Research*, Vol. 70(4), pp. 3546–3571. DOI: <https://doi.org/10.1080/03772063.2023.2202163>
22. Tran-Dang, H., Kim, D.-S. (2023), "Online Learning based Matching for Decentralized Task Offloading in Fog-enabled IoT Systems", *Proceedings 2023 28th Asia Pacific Conference on Communications Apcc 2023*, pp. 231–236. DOI: <https://doi.org/10.1109/APCC60132.2023.10460738>
23. Wang, F., Xu, J., Wang, X., Cui, S. (2018), "Joint offloading and computing optimization in wireless powered mobile-edge computing systems", *IEEE Transactions on Wireless Communications*, Vol. 17, No. 3, pp. 1784–1797. DOI: <https://doi.org/10.1109/TWC.2017.2785305>
24. Liu, Q., Han, T., Ansari, N. (2018), "Joint radio and computation resource management for low latency mobile edge computing", *Proceedings IEEE Global Communications Conference Globecom*, 8647792, pp. 1–7. DOI: <https://doi.org/10.1109/GLOCOM.2018.8647792>

25. Du, J., Zhao, L., Chu, X., Yu, F., Feng, J., Chih-Lin, I. (2019), "Enabling low-latency applications in LTE-A based mixed fog/cloud computing systems", *IEEE Transactions on Vehicular Technology*, Vol. 68, No. 2, pp. 1757–1771. DOI: <https://doi.org/10.1109/TVT.2018.2882991>
26. Tran, T. X., Pompili, D. (2019), "Joint task offloading and resource allocation for multi-server mobile-edge computing networks", *IEEE Transactions on Vehicular Technology*, Vol. 68, No. 1, pp. 856–868. DOI: <https://doi.org/10.1109/TVT.2018.2881191>
27. Zhang, Y., Lan, X. Y., Li, L., Cai, J., Pan, (2019), "Efficient computation resource management in mobile edge-cloud computing", *IEEE Internet Things Journal*, Vol. 6, No. 2, pp. 3455–3466. DOI: <https://doi.org/10.1109/JIOT.2018.2885453>
28. Chen, T., Ling, Q., Shen, Y., Giannakis, G. (2018), "Heterogeneous online learning for “thing-adaptive” fog computing in IoT", *IEEE Internet Things Journal*, Vol. 5, No. 6, pp. 4328–4341. DOI: <https://doi.org/10.1109/JIOT.2018.2860281>
29. Kuchuk, H., Mozhaiev, O., Kuchuk, N., Tiulieniev, S., Mozhaiev, M., Gnusov, Y., Tsuranov, M., Bykova, T., Klivets, S., Kuleshov, A. (2024), "Devising a method for the virtual clustering of the Internet of Things edge environment", *Eastern-European Journal of Enterprise Technologies*, Vol. 1(9(127)), pp. 60–71. DOI: <https://doi.org/10.15587/1729-4061.2024.298431>
30. Lee, B.M. (2025), "Efficient Resource Management for Massive MIMO in High-Density Massive IoT Networks", *IEEE Transactions on Mobile Computing*, Vol. 24(3), pp. 1963–1980. DOI: <https://doi.org/10.1109/TMC.2024.3486712>
31. Yu, J., Hou, K., Zhang, H., Kostic B., Yang, M., Nazif, H. (2025), "A new energy-aware resources scheduling method for mobile internet of things using a hybrid optimisation algorithm", *International Journal of Mobile Communications*, Vol. 25(2), pp. 176–207. DOI: <https://doi.org/10.1504/IJMC.2025.144192>
32. Kuchuk, H., Mozhaiev, O., Tiulieniev, S., Mozhaiev, M., Kuchuk, N., Tymoshchuk, L., Onishchenko, Yu., Tulupov, V., Bykova, T., Roh, V. (2025), "Devising a method for forming a stable mobile cluster of the Internet of things fog layer. Eastern-European Journal of Enterprise Technologies", Vol. 1(4(133)), pp. 6–14. DOI: <https://doi.org/10.15587/1729-4061.2025.322263>
33. Zheng, K., Luo, R., Liu, X., Qiu, J., Liu, J. (2024), "Distributed DDPG-Based Resource Allocation for Age of Information Minimization in Mobile Wireless-Powered Internet of Things", *IEEE Internet of Things Journal*, Vol. 11(17), pp. 29102–29115. DOI: <https://doi.org/10.1109/JIOT.2024.3406044>
34. Liu, J., Wei, X., Fan, J. (2019), "Tolerable Data Transmission of Mobile Edge Computing under Internet of Things", *IEEE Access*, Vol. 7, pp. 71859–71871, 8728032. DOI: <https://doi.org/10.1109/ACCESS.2019.2920442>
35. Joshi, N., Srivastava, S. (2023), "QoS-Aware Task Allocation and Scheduling in Three-Tier Cloud-Fog-IoT Architecture Using Double Auction", *International Conference on Cloud Computing and Services Science Closer Proceedings*, pp. 253–260. DOI: <https://doi.org/105220/0011967400003488>
36. Toghyani, M., Khorsand, R., Khaksar, H. (2025), "QoS-SLA-aware Optimization Framework for IoT-Service Placement in Integrated Fog-Cloud Computing", *Journal of Grid Computing*, Vol. 23(1), article number 1. DOI: <https://doi.org/10.1007/s10723-024-09787-x>
37. Muñoz, L.A., Berná Martínez, J.V., Asensi, C.C., Pastor, D.S. (2025), "Research Notes: Design of a Distributed and Highly Scalable Fog Architecture for Heterogeneous IoT Infrastructures", *Int. Journal of Software Engineering and Knowledge Eng.*, Vol. 35(2), pp. 195–215. DOI: <https://doi.org/10.1142/S0218194025430016>
38. Wang X., Sui Y., Wang J., Yuen C., Wu W. (2021), "A Distributed Truthful Auction Mechanism for Task Allocation in Mobile Cloud Computing", *IEEE Transactions on Services Computing*, Vol. 14(3), pp. 628–638. DOI: <https://doi.org/10.1109/TSC.2018.2818147>
39. Kuchuk, H., Mozhaiev, O., Tiulieniev, S., Mozhaiev, M., Kuchuk, N., Khorobrykh, P., Gnusov, Yu., Horelov, Yu., Svitlychnyi, V., Bilyk, O. (2025), "Devising a method for managing computing resources in a fog layer of the mobile high-density Internet of Things", *Eastern European Journal of Enterprise Technologies*, Vol. 6, No. 4(138), pp. 15–25. DOI: <https://doi.org/10.15587/1729-4061.2025.344553>
40. Semenov, S., Mozhaiev, O., Kuchuk, N., Mozhaiev, M., Tiulieniev, S., Gnusov, Yu., Yevstrat, D., Chyrva, Y., Kuchuk, H. (2022), "Devising a procedure for defining the general criteria of abnormal behavior of a computer system based on the improved criterion of uniformity of input data samples", *Eastern-European Journal of Enterprise Technologies*, Vol. 6(4-120), pp. 40–49. DOI: <https://doi.org/10.15587/1729-4061.2022.269128>
41. İbrahimov B.G., Hasanov A.H., Hashimov E.G. (2024), "Research and analysis of efficiency indicators of critical infrastructures in the communication system", *Advanced Information Systems*, Vol. 8, No. 2, pp. 58–64. DOI: <https://doi.org/10.20998/2522-9052.2024.2.07>
42. Wooldridge, M., Jennings, N. R. (1995), "Intelligent agents: Theory and practice", *The Knowledge Engineering Review*, Vol. 10(2), pp. 115–152. DOI: <https://doi.org/10.1017/S0269888900008122>
43. Nimmala, S., Sena, P.V., Inturi, S., Janbhasha, S., Narsimhulu, P., Manoranjini, J. (2025), "Multi-Agent Deep Reinforcement Learning for Intelligent Industrial Iot Networks", *Proceedings of the 9th International Conference on Inventive Systems and Control Icisc 2025*, pp. 455–459. DOI: <https://doi.org/10.1109/ICISC65841.2025.11187915>
44. Kuchuk, H., Malokhvii, E. (2024), "Integration of IoT with Cloud, Fog, and Edge Computing: A Review", *Advanced Information Systems*, Vol. 8, No. 2, pp. 65–78. DOI: <https://doi.org/10.20998/2522-9052.2024.2.08>

45. Heik, D., Bahrpeyma, F., Reichelt, D. (2024), "Study on the application of single-agent and multi-agent reinforcement learning to dynamic scheduling in manufacturing environments with growing complexity: Case study on the synthesis of an industrial IoT Test Bed", *Journal of Manufacturing Systems*, Vol. 77, pp. 525–557. DOI: <https://doi.org/10.1016/j.jmsy.2024.09.019>
46. Oliveira, H.D., Kaneko, M., Boukhatem, L. (2024), "Federated Multiagent Deep Reinforcement Learning for Intelligent IoT Wireless Communications: Overview and Challenges", *IEEE Vehicular Technology Magazine*, Vol. 19(4), pp. 73–82. DOI: <https://doi.org/10.1109/MVT.2024.3451191>
47. Gharbi, A., Ayari, M., Albalawi, N. (2025), "Intelligent Caching in IoT Sensing Networks: A Decentralized Multi-Agent Reinforcement Learning Approach with Entropy-Driven Exploration", *Procedia Computer Science*, Vol. 270, pp. 4696–4703. DOI: <https://doi.org/10.1016/j.procs.2025.09.595>
48. Wang, Z., Fu, S., Pan, J., Zhao, J., Wang, Z. (2024), "Nash equilibrium, dynamics and control of congestion games with resource failures", *Nonlinear Dynamics*, Vol. 112(18), pp. 16587–16599. DOI: <https://doi.org/10.1007/s11071-024-09885-1>
49. Hou, J., Zeng, X. (2024), "Distributed Convergence to Nash Equilibria in a Zero-Sum Resource Allocation Game", *Lecture Notes in Electrical Engineering*, 1203 LNEE, pp. 77–88. DOI: https://doi.org/10.1007/978-981-97-3324-8_7

Received (Надійшла) 23.01.2026

Accepted for publication (Прийнята до друку) 10.02.2026

Publication date (Дата публікації) 30.03.2026

Відомості про авторів / About the Authors

Кучук Георгій Анатолійович – доктор технічних наук, професор, Національний технічний університет "Харківський політехнічний інститут", професор кафедри комп'ютерної інженерії та програмування; Харків, Україна;

Heorhii Kuchuk – Doctor of Technical Sciences, Professor, National Technical University "Kharkiv Polytechnic Institute", Professor at the Computer Engineering and Programming Department; Kharkiv, Ukraine;

e-mail: kuchuk56@ukr.net

ORCID ID: <http://orcid.org/0000-0002-2862-438X>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57057781300>

Косенко Наталія Вікторівна – кандидат технічних наук, доцент, Харківський національний університет міського господарства ім. О. М. Бекетова, доцент кафедри управління проектами у міському господарстві і будівництві; Харків, Україна;

Nataliia Kosenko – Candidate of Technical Sciences, Associate Professor, Beketov National University of Urban Economy in Kharkiv, Associate Professor at the Department of Project Management in Urban Economy and Construction; Kharkiv, Ukraine;

e-mail: kosnatalja@gmail.com

ORCID ID: <https://orcid.org/0000-0002-5942-3150>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57196219605>

Кучук Ніна Георгіївна – доктор технічних наук, професор, Національний технічний університет "Харківський політехнічний інститут", професор кафедри комп'ютерної інженерії та програмування; Харків, Україна;

Nina Kuchuk – Doctor of Technical Sciences, Professor, National Technical University "Kharkiv Polytechnic Institute", Professor at the Computer Engineering and Programming Department; Kharkiv, Ukraine;

e-mail: nina_kuchuk@ukr.net

ORCID ID: <http://orcid.org/0000-0002-0784-1465>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57196006131>

Гопєєнко Віктор – доктор технічних наук, професор, проректор з наукової роботи, директор навчальної програми "Комп'ютерні системи" Ризького північного університету прикладних наук, кафедра природничих наук та комп'ютерних технологій; Рига, Латвія; провідний науковий співробітник Вентспілського університету прикладних наук, Вентспільський міжнародний радіоастрономічний центр; Вентспілс, Латвія;

Viktors Gopejenko – Doctor of Technical Sciences, Professor, Vice-Rector for Research, Director at the Study Programme Computer Systems (MSc) Riga Nordic University of Applied Sciences, Department of Natural Science and Computer Technologies; Riga, Latvia; Leading Researcher at the Ventspils University of Applied Sciences, Ventspils International Radio Astronomy Centre; Ventspils, Latvia;

e-mail: viktors.gopejenko@isma.lv

ORCID ID: <https://orcid.org/0000-0002-7783-4519>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=55038229400>

Косенко Віктор Васильович – доктор технічних наук, професор, Національний університет "Полтавська політехніка імені Юрія Кондратюка", професор кафедри автоматизації, електроніки та телекомунікацій; Полтава, Україна; Харківський національний університет радіоелектроніки, професор кафедри комп'ютерно-інтегрованих технологій, автоматизації, робототехніки та безпекової інженерії; Харків, Україна;

Viktor Kosenko – Doctor of Technical Sciences, Professor, National University "Yuri Kondratyuk Poltava Polytechnic", Professor at the Department of Automation, Electronic and Telecommunication; Poltava, Ukraine; Kharkiv National University of Radio Electronics, Professor at the Department of Computer-Integrated Technologies, Automation, Robotics and Safety Engineering; Kharkiv, Ukraine;

e-mail: kosvict@gmail.com

ORCID ID: <https://orcid.org/0000-0002-4905-8508>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57190443921>

АДАПТИВНИЙ МЕТОД РОЗПОДІЛУ РЕСУРСІВ МОБІЛЬНОГО ТУМАННОГО ШАРУ ВИСОКОЩІЛЬНОГО ПРОМИСЛОВОГО ІНТЕРНЕТУ РЕЧЕЙ У МЕРЕЖАХ INDUSTRY 5.0

Актуальність дослідження. Сучасна концепція Індустрії 4.0 заклала фундамент для повної цифровізації через індустріальний Інтернет речей. Але перехід до Індустрії 5.0 вимагає більшої гнучкості та стійкості систем. Високощільний мобільний індустріальний IoT з туманним шаром є критичним елементом цієї трансформації, оскільки забезпечує не лише автоматизацію, а й адаптивність виробництва до людських потреб і екологічних стандартів. **Об'єктом вивчення** є процес передоброблення транзакцій периферійного шару HDIoT. **Основна гіпотеза дослідження:** впровадження нового адаптивного методу розподілу ресурсів мобільних пристроїв туманних кластерів дасть змогу зменшити середній час передоброблення транзакції периферійного шару HDIoT. **Метою роботи** є зменшення середнього часу перебування транзакції периферійного шару HDIoT в туманному шарі завдяки розробленню адаптивного методу розподілу ресурсів мобільних пристроїв туманних кластерів. **Завдання дослідження:** визначити архітектурні особливості туманних обчислень у мережах HDIoT; створити математичну модель процесу оптимального розподілу ресурсів пристроїв мобільного кластера туманного шару; формалізувати багатоагентний підхід для розподілу ресурсів кластера; розробити й дослідити теоретико-ігрову модель управління ресурсами мобільного туманного кластера багатощільної IoT. **Застосовані методи:** багатоагентний підхід, теорія ігор, зокрема оптимізація кооперативної стохастичної гри, комп'ютерне моделювання. **Досягнуті результати.** Розроблено адаптивний метод розподілу ресурсів мобільних пристроїв туманних кластерів. У межах методу запропоновано архітектуру мобільного туманного кластера, створено математичну модель процесу оптимального розподілу його ресурсів. Крім цього, застосовано багатоагентний підхід для знаходження наближеного рішення сформульованої двопараметричної нелінійної оптимізаційної задачі, а для зменшення обчислювальної складності та прискорення знаходження наближеного рішення впроваджено теоретико-ігровий підхід. **Висновок.** Унаслідок застосування розробленого методу зменшено середній час перебування транзакції периферійного шару високощільного IoT в туманному шарі, що за умови високої щільності мобільних пристроїв уможливило виконання вимог QoS.

Ключові слова: Індустрія 5.0; Інтернет речей; туманний кластер; мобільний вузол; комп'ютерна система; агентний підхід; інтелектуальний агент; стохастична гра.

Бібліографічні описи / Bibliographic descriptions

Кучук Г. А., Косенко Н. В., Кучук Н. Г., Гопесенко В., Косенко В. В. Адаптивний метод розподілу ресурсів мобільного туманного шару високощільного промислового Інтернету речей у мережах Industry 5.0. *Сучасний стан наукових досліджень та технологій в промисловості*. 2026. № 1 (35). С. 65–78. DOI: <https://doi.org/10.30837/2522-9818.2026.1.065>

Kuchuk, H., Kosenko, N., Kuchuk, N., Gopejenko, V., Kosenko, V. (2026), "Adaptive Resource Allocation Method for the Mobile Fog Layer of High-Density Industrial Internet of Things in Industry 5.0 Networks", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (35), P. 65–78. DOI: <https://doi.org/10.30837/2522-9818.2026.1.065>

Anatoliy Sachenko, Pavlo Radiuk, Mykola Lysyi, Oleksandr Melnychenko, Oleg Zastavnyy

ARCHITECTURE OF CYBERPHYSICAL SYSTEMS FOR UAV-BASED LATE-FUSION DEFECT DETECTION IN PHOTOVOLTAIC MODULES

The **subject** of research is the architectural advancement of inspection systems for large-scale solar power plants. As global solar infrastructure expands, reliance on manual or offline analytical methods creates significant operational bottlenecks. The **goal** of research is to improve the operational utility of unmanned aerial vehicle (UAV)-based photovoltaic module inspection by developing a cyber-physical system (CPS) architecture. It integrates onboard deep learning, edge nodes, cloud analytics, and Supervisory Control and Data Acquisition (SCADA)-aware decision-making into a single coordinated workflow. The **tasks** of research: 1) formalise a multi-tiered CPS architecture (UAV-edge-cloud) and define interfaces for data, geo-tags, and alarms; 2) develop and validate an onboard thermographic detection pipeline with palette-aware fusion; and 3) integrate detection results with a SCADA-aware logic layer for hazard inference and fire risk mitigation. The **methods** of research: Computer vision and deep learning (YOLOv11) are used for onboard defect segmentation. Model ensembling via a late-fusion strategy for M2 and M3 thermal palettes mitigates domain shift. RTK-supported spatial clustering algorithms ensure precise geo-indexing and deduplication, and deterministic Boolean logic assesses fire risks based on bypass diode states. The **results** obtained with five-fold cross-validation shows the proposed architecture significantly outperforms single-modality baselines. The onboard YOLOv11 model achieved a macro mAP@0.5 of 0.91 and 0.90 for M2 and M3 palettes, respectively. The late-fusion ensemble elevated mAP@0.5 for cracks to 0.96 and delamination to 0.95. It reduced end-to-end per-frame processing latency from 4.235 s to 2.858 s. Field validation demonstrated an error of 0.71 defects per inspected string compared to manual counts. Sensitivity analysis highlighted that a 10 m flight altitude provides an optimal balance, yielding 93% precision and 90% recall. **Conclusions:** Treating UAV inspection as an integrated cyber-physical service improves defect detection. This offers a scalable, real-time solution for preventive maintenance and automated fire-risk mitigation in renewable energy.

Keywords: cyber-physical systems; photovoltaic modules; unmanned aerial vehicles; infrared thermography; edge computing; deep learning.

Introduction

Photovoltaic (PV) energy has evolved from a niche technology into a vital component of modern power systems. At the same time, photovoltaic power plants remain vulnerable to harsh environmental conditions and a long list of faults that are commonplace but collectively costly [1]. Microcracks can form during manufacturing, transportation, installation, or hail; they can lead to inactive cell areas, hot spots, and energy losses [2]. Aging and delamination of the encapsulant can lead to moisture accumulation and accelerate corrosion. Contamination reduces the intensity of solar radiation on the cells and alters the thermal regime, while damaged connections or bypass diodes can cause abnormal heating, posing a safety risk. Therefore, reliability-focused maintenance is mandatory [3]: predictive and preventive inspections directly impact energy output and the plant's service life.

Infrared (IR) thermography is one of the most informative methods for non-contact diagnostics of photovoltaic systems. It provides spatial temperature maps that reveal electrical mismatches, increased series

resistance, shunting, and localized degradation. Reviews show that thermography can detect a wide range of photovoltaic system anomalies, but also emphasize that its diagnostic value depends on data acquisition conditions, assumptions regarding emissivity, inspection geometry, and the ability to distinguish defect signatures from reflections and environmental gradients [4, 5]. Comparative studies also show that defects can manifest differently under lighting conditions and in darkness, meaning that the inspection protocol is just as important as the camera itself [6].

Unmanned aerial vehicles (UAVs) have recently become the standard platform for large-scale photothermal imaging. UAVs can cover industrial-scale fields in a matter of minutes and inspect hard-to-reach rooftop installations [7, 8]. Many studies report the operational advantages of UAV thermography, including improved coverage and reduced inspection time [9], but also note practical limitations: flight altitude and speed control spatial resolution and motion blur; time of day and weather affect thermal contrast; and specular reflections from glass surfaces can create misleading hot spots [10, 11]. These limitations are not mere nuisances;

they are the reason why “laboratory” image analysis pipelines often fail when used in the field.

Automation is the second major trend. Deep learning has achieved high performance in classifying and localizing defects in photovoltaic systems on infrared and RGB images, and surveys describe a rapidly growing set of architectures ranging from classical convolution-based classifiers to modern single-stage detectors and segmentation networks [12, 13]. Recent UAV-oriented approaches use variants of the YOLO family to detect modules and anomalies in thermal and visible data [14, 15], while methods based on thermal video demonstrate that automation can be extended from individual frames to temporal streams [16]. However, a persistent limitation is that many published solutions implicitly assume an offline workflow: data is collected, downloaded, and processed in batches [12, 17]. In operational photovoltaic power plants, the inspection process is a closed-loop system that must meet requirements for latency, throughput, and reliability, as well as provide results that can be traced back to physical modules and utilized by operators [18].

This observation calls for the creation of a system-level architecture. The photovoltaic plant inspection pipeline is not merely a computer vision model; it is a cyber-physical system (CPS) [19] that integrates sensing, computing, communication, and control functions. The UAV captures physical measurements, but the station requires cyber actions: alerts, maintenance tickets, and safety procedures. Edge nodes offer a natural architectural solution: they can reduce bandwidth by processing data close to the source and provide fast responses without relying on a cloud connection [20]. However, despite the maturity of edge nodes in industrial monitoring, UAV solutions for inspecting photovoltaic stations rarely provide an explicit KFS architecture that integrates on-board inference, edge aggregation, long-term cloud analysis, and SCADA-aware decision-making logic [21]. Similarly, the literature on photovoltaic thermography typically emphasizes detection accuracy while underestimating the georeferencing and deduplication steps necessary to avoid false positives and alarm fatigue.

Therefore, the problem addressed in this article lies in the need to develop and validate an accurate and operationally implementable UAV-based PV plant defect monitoring architecture. The technical gap lies not simply in “detecting more defects,” but in creating a reliable, comprehensive service with sequential georeferencing, real-time calculations on onboard

hardware, and direct integration with security surveillance and monitoring systems. This is particularly relevant for fire prevention, as abnormal heating and faults in shunt diodes can quickly turn local defects into critical incidents. The integration of UAV-based detection with SCADA logic aligns with recent trends that view unmanned technologies and artificial intelligence as integral components of power plant safety monitoring systems [22, 23].

The aim of this study is to improve the operational utility of UAV-based photovoltaic module inspections by developing a cyber-physical architecture that integrates embedded deep learning, edge nodes, cloud analytics, and SCADA-aware decision-making logic into a single coordinated workflow.

To achieve this goal, the paper proposes three main scientific contributions, namely:

- 1) a multi-level CFA (UAV-periphery-cloud) architecture for inspecting photovoltaic modules, which clearly defines data flow, responsibilities, and decision-making points, including RTK-based geotagging and deduplication to create traceable defect events at the module level;
- 2) a thermography-oriented on-board inference pipeline based on the YOLOv11 segmentation model, with post-processing that accounts for the color palette and an original late-merging ensemble between two representations of thermograms, ensuring high detection quality and significantly reducing end-to-end processing time;
- 3) integration of detection results with an interpretable logic layer that accounts for SCADA states (in particular, shunt diodes) for automated hazard inference and rapid fire risk mitigation.

Related Works

Research on UAV-based monitoring of photovoltaic module defects spans sensor physics, computer vision, and systems engineering. This section briefly discusses the place of this work in this context and explains how the proposed approach to cyber-physical operation differs from approaches based solely on detectors.

Thermography is widely used to detect hot spots, anomalies, and signs of degradation in photovoltaic modules. Comprehensive reviews emphasize that thermal signals are influenced by solar irradiance, ambient temperature, wind, and viewing angle, and discuss recommended inspection procedures and limitations [4, 5].

Gallardo-Saavedra et al. [6] compared lighting and darkness conditions and demonstrated that defect detection depends on operating conditions, highlighting the need for consistent flight protocols and contextual information, such as solar irradiance and environmental conditions. The practical application of UAVs has been evaluated in field studies that highlight how flight altitude and motion control affect spatial resolution and, consequently, the ability to detect anomalies [7, 11]. These findings prompt a focus on the sensitivity of flight parameters and reflection-aware planning.

Deep learning for the diagnosis of photovoltaic systems has progressed from the classification of cropped module images to detection and segmentation at the power plant level. Reviews summarize the expanding use of neural networks, object detectors, and hybrid models in the monitoring of photovoltaic systems [17, 24]. Early work by Dotenko et al. [25] demonstrated the automatic detection of photovoltaic modules and defect analysis in aerial infrared images using statistical tests and classical vision, reporting high defect classification accuracy while emphasizing the importance of robust preprocessing. Later works increasingly rely on single-stage detectors. Di Tommaso et al. [14] proposed a multi-stage pipeline using YOLOv3 for detecting modules and anomalies in both infrared and visible images, which improved automation but still required careful separation of tasks and datasets. Xi et al. [18] introduced ST-YOLO for detecting faults in photovoltaic systems on infrared images and reported a very high mAP@0.5 under their experimental conditions. Other studies focus on investigating thermographic features and reliability, including deep learning strategies for fault diagnosis from thermograms [26] and segmentation-based methods for identifying modules and defects [27]. These works prompt the use of a detector with segmentation capabilities (YOLOv11-seg) and attention to palette-induced domain shift.

Although thermography is central, many implementations use RGB images to provide context and reduce false positives. RGB-based deep learning is used to classify signs of faults and operational issues in photovoltaic stations [28]. Multimodal pipelines can also support human interpretation, as operators can check the RGB context when thermal features are ambiguous [29]. Therefore, the proposed architecture involves synchronized RGB and thermal imaging, although the defect detector in this paper focuses on thermal masks.

From a systems perspective, UAV inspection is an example of distributed sensing with severe constraints [30]. Edge nodes are often recommended for the Industrial Internet of Things because they reduce dependence on the cloud and provide low-latency solutions [20]. Tang et al. [22] demonstrated a deep learning architecture at the cloud edge for detecting linear defects in large photovoltaic power plants, showing that distributed computing can support plant-level monitoring and reduce data transmission volume. However, as noted by Ferlito et al. [23], many edge-of-the-cloud studies focus solely on “edge detection” and do not explicitly model feedback to plant operations. This study extends this line of research by adding RTK-based georeferencing, duplicate suppression, and SCADA-aware fire risk logic, motivated by the observation that photovoltaic system defects have implications for both productivity and safety [21]. Thus, the technical gap lies not simply in detecting a greater number of defects, but in creating a comprehensive cyber-physical architecture capable of functioning in real time and integrating with the production process.

Thus, the literature review confirms that, to overcome existing limitations, it is necessary to develop and validate a UAV-based defect monitoring architecture for photovoltaic systems that is accurate, georeferenced, and fully functional under real-world operating conditions. Based on the analysis, the overall research objective was broken down into 5 interrelated tasks.

- 1) Formalize a multi-level UAS architecture (UAS-periphery-cloud) and define protocols and interfaces for the efficient exchange of data, geotags, and alarms.

- 2) Develop and adapt a lightweight deep learning model (YOLOv11) to perform accurate on-board segmentation of photovoltaic module defects directly in real time.

- 3) Propose and validate an embedded thermal detection pipeline that uses an ensemble strategy (late merging of M2 and M3 palettes) to mitigate domain shift and improve the accuracy of recognizing different types of damage.

- 4) Develop spatial clustering algorithms with RTK navigation support to ensure reliable georeferencing and eliminate duplication (deduplication) of detected defects at the edge computing level.

- 5) Integrate the obtained detection results with a deterministic logic layer based on SCADA data for automated fire risk assessment and instant generation of critical alerts to operators.

Methods and Materials

This section describes the proposed approach in detail. It covers the KFS architecture, sensor and communication design, flight planning with mapping, dataset construction, on-board deep learning, edge/cloud analytics, and SCADA-informed decision-making. The core principle is to treat defect detection as a cyber-physical workflow: sensors, computations, communications, and actions must be designed in concert.

System Architecture and Operational Cycle

In this study, the inspection of photovoltaic systems is implemented as a closed-loop cycle, starting with measurement and ending with action at the power plant level. The CPS is organized into three computational levels (Fig. 1): the UAV level, the peripheral level deployed at the photovoltaic power plant, and the cloud level for long-term analysis. The operator level interacts with the CPS via a control panel and SCADA/HMI interfaces.

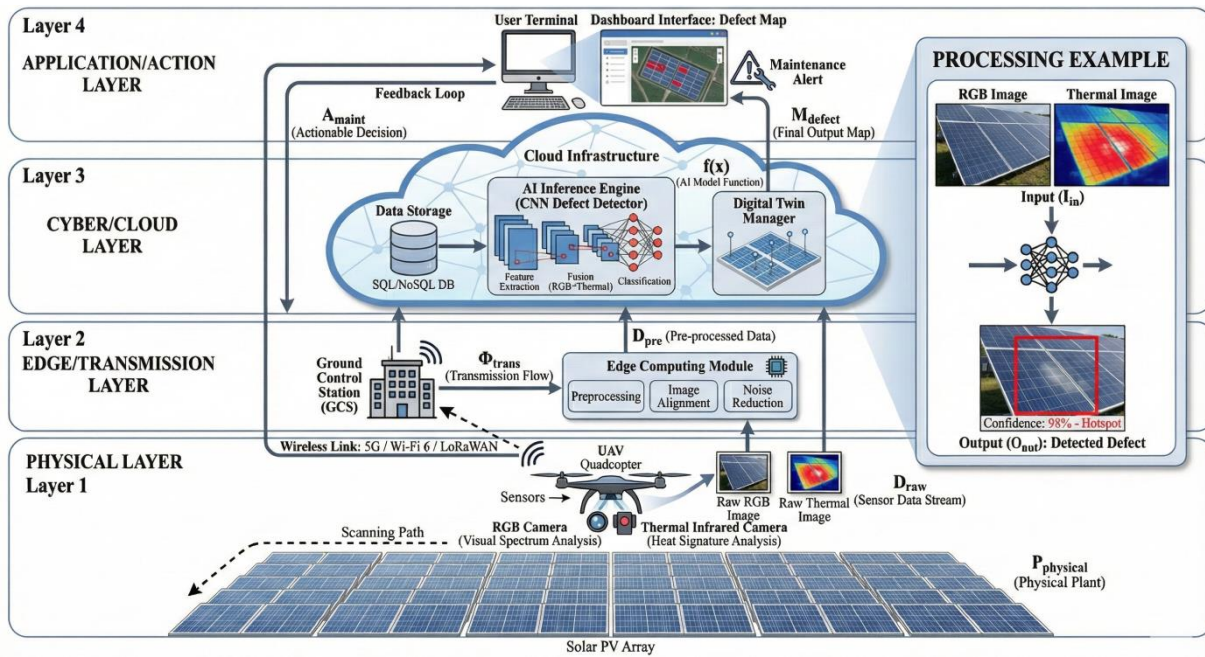


Fig. 1. Architecture of a cyber-physical system for monitoring defects in photovoltaic systems based on UAVs

The UAV level is responsible for tasks requiring rapid response. It captures synchronized RGB and thermal images and performs on-board computations to detect potential defects. The edge layer is responsible for tasks related to the plant context: it performs geotagging using RTK, removes duplicates of recurring detections, aggregates them into events at the module level, and interacts with SCADA to trigger alarms and initiate maintenance. The cloud stores mission artifacts and provides dashboards for trend analysis, module history, and inspection reports.

Let the task generate a continuous stream of frames $\{I_t\}_{t=1}^T$, where each frame contains an RGB image, I_t^{rgb} a thermal image I_t^{th} , and a navigation record n_t (RTK position, altitude, heading, and timestamp). The built-in inference system performs a set of detections as follows:

$$\mathcal{D}_t = \left\{ (c_k, M_k, s_k, g_k) \right\}_{k=1}^{K_t}, \quad (1)$$

where c_k is the defect class (crack, contamination, delamination), M_k is the segmentation mask, s_k is the confidence score, and g_k is the geotag (calculated location of the module in the object's coordinates).

The edge server receives \mathcal{D}_t and performs (i) spatial clustering and temporal deduplication, (ii) association of detections with module IDs and strings, and (iii) generation of a status report that maps detections and SCADA signals to operational actions. The cloud stores mission artifacts and provides dashboards for trend analysis, module history, and audit reports.

UAV Platform, Sensors, and Communication

KFS provides a UAV platform capable of stable flight along predefined patterns (grid or corridor scanning) and equipped with a combined RGB/thermal payload. The thermal channel provides the primary signal

for detecting abnormal heating patterns; the RGB channel provides contextual information and supports operator verification. Although the architecture is platform-independent, the stated implementation is comparable to a Matrice-class UAV with a dual-sensor payload gimbal (RGB + thermal imager) and RTK positioning.

From an ORM perspective, georeferencing is just as important as the imagery. RTK positioning reduces drift and ensures consistent mapping of detected objects onto physical modules throughout missions. In addition, precise timestamps allow for the synchronization of the thermal camera, RGB camera, and flight controller. The UAV records navigation data for each frame, including RTK position and altitude. This data is subsequently used for deduplication and associating detections with the plant layout.

Communication is treated as a limited resource. Streaming thermal video at full resolution is often impossible via standard UAV channels, especially in large factories. Instead, the UAV transmits compact

detection messages according to formula (1) to a peripheral server in near real-time. Full-resolution images can be cached on board and downloaded after the flight or selectively downloaded when the edge server requests evidence. This design aligns with the general logic of edge nodes in industrial monitoring: local processing reduces bandwidth, and decisions at the plant level can be made independently of cloud connectivity [20].

Choosing a viewpoint to account for reflections

Thermography of photovoltaic modules using UAVs is susceptible to specular reflections from glass surfaces. If the camera is pointed toward a specular reflection of the sun, the thermogram may be obscured by glare, leading to the appearance of false hot spots and distortion of the thermal contrast required for defect detection. Therefore, to mitigate this effect, an observation point selection algorithm that accounts for reflection has been added to the RTK geometry and the sun's position (Fig. 2).

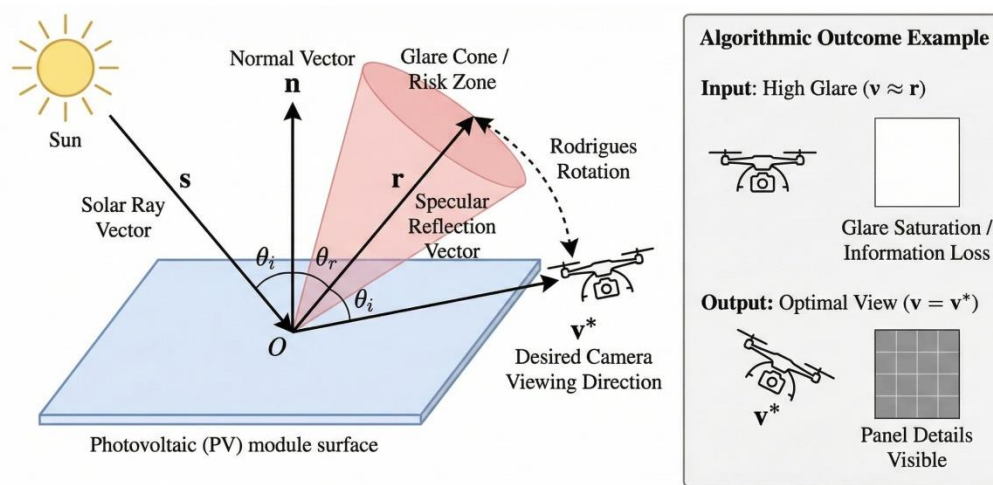


Fig. 2. Geometric model for selecting a viewing point that accounts for reflection. Given the plane of the module n and the direction of the sunbeam s , the direction of specular reflection r is calculated. The desired camera viewing direction v^* is obtained by deviating from r (Rodrigues rotation) to reduce glare while maintaining a sufficient angle of incidence for thermal contrast

Let the three points on the module plane that have an RTK relationship be $p_1, p_2, p_3 \in \mathbb{R}^3$ (e.g., the corners of a representative module). The normal vector of the module is calculated as follows:

$$n = \frac{(p_2 - p_1) \times (p_3 - p_1)}{\|(p_2 - p_1) \times (p_3 - p_1)\|}. \quad (2)$$

Let s be the single vector pointing from the module to the Sun (obtained from the solar ephemeris model at time t). The direction of specular reflection is as follows:

$$r = s - 2(n^\top s)n. \quad (3)$$

If the camera's viewing direction coincides with r , specular reflections dominate. The target's viewing direction v^* is determined by rotating r around the normal n by an angle θ , chosen such that the reflection direction lies outside the camera's field of view. Using the Rodrigues rotation formula, the inverse vector takes the following form:

$$v^* = r \cos \theta + (n \times r) \sin \theta + n(n^\top r)(1 - \cos \theta). \quad (4)$$

Algorithm 1 summarizes the procedure for selecting the viewpoint. In practice, Θ may be a small set of possible

angles (e.g., $\pm 5^\circ$, $\pm 10^\circ$), and the glare assessment may be approximated by the angular distance between v_θ and r relative to the camera's field of view. The key

point is that the algorithm is lightweight and compatible with real-time decision-making at the edge.

Algorithm 1. Selection of a viewing point accounting for reflection for UAV thermography

Requirements: Points on the RTK plane p_1, p_2, p_3 , sun direction s , possible rotation angles Θ

Implementation: Select the viewing angle v^* that minimizes glare

- 1: Calculate the surface normal n using formula (2)
- 2: Calculate the reflection direction r using formula (3)
- 3: **For all** θ in Θ , **perform**
- 4: Calculate the possible v_θ using formula (4)
- 5: Estimate the glare level (expected reflection in the camera's field of view)
- 6: end for
- 7: Select $v^* = \arg \min \{ \text{GlareScore}(v_\theta) \}$
- 8: **return** v^*

Building a dataset and presenting thermograms

Thermal cameras often support multiple color palettes that represent temperature ranges as colors for visualization. Although these palettes are designed for human interpretation, they also alter the statistical distribution of pixel values fed into the training model. The experiments consider two representations of thermograms: a two-color palette (labeled M2) and a three-color palette (labeled M3). The key difference lies in the effective temperature range highlighted by the visualization. In the described configuration, M3 uses a narrower dynamic range of approximately 15°C , which increases sensitivity to small temperature contrasts (on the order of $0.1\text{--}0.5^\circ\text{C}$) but may saturate extreme anomalies. M2 uses a wider range of approximately 35°C , which reduces sensitivity to small hot spots but covers the entire spectrum from minor heating inconsistencies to serious damage. M2 and M3 are considered as two reference representations of thermograms obtained from the same physical scenes to assess robustness to palette-induced domain differences and to motivate selective merging.

To develop and validate the proposed approach, two publicly available UAV thermography datasets were additionally used to provide additional defect coverage:

1. STHS-277 [31]: This dataset contains 277 full-frame thermal images capturing defects in the form of “snail trails” and hotspots, and includes environmental metadata. We extend the provided annotations by adding bounding boxes for all instances of photovoltaic panels to ensure comprehensive detector training.

2. PVF-10 [32]: This is a large-scale, high-resolution UAV thermal imaging dataset containing 5,579 annotated patches of individual solar panels from eight power plants and a detailed taxonomy of ten defect classes.

For the on-board segmentation experiments discussed in this article, we prepared a compact dataset of masks with three classes (cracks, contamination, delamination) by combining field thermography with selected samples from the aforementioned publicly available datasets. The images were annotated using the Computer Vision Annotation Tool (CVAT) [33] and exported as polygonal masks suitable for segmentation training.

Since the minority classes are small, the evaluation is based on stratified five-fold cross-validation to avoid inaccurate point estimates from a small test set. Table 1 shows the composition of the dataset and the approximate partitioning used in the evaluation.

Table 1. Dataset structure and approximate partitioning used for stratified 5-fold cross-validation (number per partition)

Class	Defect Type	Total images	Training/Validation/Testing per component (count)
1	Crack	20	14/2/4
2	Contamination	10	7/1/2
3	Delamination	170	122/14/34
Total		200	143/17/40

On-board deep learning: YOLOv11 segmentation

The UAV layer uses a lightweight deep learning model for defect detection and segmentation. A segmentation model from the YOLO family was chosen because it provides a good balance between accuracy and speed and has well-developed deployment tools. Specifically, YOLOv11-seg [34] is used as the base network. The model outputs bounding boxes and pixel-level masks for each detected defect. Masks are valuable in the inspection of photovoltaic systems because they allow for the estimation of the defect's area and shape, which improves prioritization and supports temporal tracking.

The detector is trained on the annotated dataset described in Table 1. To mitigate overfitting under conditions of severe class imbalance, training employs transfer learning and class-aware sampling: the network is initialized from a pre-trained YOLO checkpoint and fine-tuned using strong geometric and photometric augmentations (random affine transformations, flipping, blurring, and contrast/brightness fluctuations), while instances of minority classes (cracks and contaminants) are downsampled during mini-batch construction. Public thermal imaging datasets (STHS-277 and PVF-10) [31, 32] are used to expand the thermal distribution during fine-tuning.

During inference, the model generates a set of potential objects $\{(b_k, M_k, s_k, c_k)\}$, where b_k is the bounding box and M_k is the mask. Non-maximum suppression (NMS) removes duplicate bounding boxes, and a similar suppression at the mask level is applied to reduce overlap between instance masks. The post-processing stage is particularly important because UAV video streams contain many nearly identical frames, and thermal palettes can create color artifacts that cause false detections.

Mask similarity, size-based routing, and post-processing

The built-in output provides instance masks that must be filtered and merged to obtain stable defect events. In this study, mask overlap is based on the intersection over union (IoU), which is calculated as follows:

$$\text{IoU}(A, B) = |A \cap B| / |A \cup B|. \quad (5)$$

Algorithm 2. Embedded processing of a three-color thermogram (M3)

Input: Thermogram I in a three-color palette, YOLOv11-seg model, thresholds $(\tau_c, \tau_{\text{IoU}})$

Execution: Filtered set of defects $\mathcal{D}_{\text{small}}$

Mask-level noise suppression (Mask-NMS) removes duplicates when the IoU exceeds a threshold, retaining predictions with a higher confidence level. Mask-NMS is important in UAV inspection because adjacent frames often contain the same defect with slightly offset masks.

In addition to removing duplicates, the FCS performs defect routing by size between two thermograms. Let S_{def} denote the area (in pixels) of the detected defect mask, and S_{cell} denote the area of a single photodiode at the same image scale. The relative defect area is determined as follows:

$$R_d = S_{\text{def}} / S_{\text{cell}}. \quad (6)$$

Intuitively, one can assume that defects smaller than a single cell behave as localized hot spots and benefit from the higher contrast of M3, whereas larger anomalies are more reliably captured in the M2 thermogram with a wider range.

Using R_d , two mutually exclusive sets of defects are formed as follows:

$$\mathcal{D}_{\text{small}} = \{d \in \mathcal{D}^{\text{M3}} \mid R_d(d) < 1.0\}, \quad (7)$$

$$\mathcal{D}_{\text{large}} = \{d \in \mathcal{D}^{\text{M2}} \mid R_d(d) \geq 1.0\}, \quad (8)$$

and the final set of detections is determined by the formula:

$$\mathcal{D}_{\text{final}} = \mathcal{D}_{\text{small}} \cup \mathcal{D}_{\text{large}}. \quad (9)$$

Since the routing criteria are complementary and Mask-NMS is applied within each branch, the resulting sets of $\mathcal{D}^{\text{M3}} \cap \mathcal{D}^{\text{M2}} = \emptyset$ meet practical requirements, and the number of duplicate alarms is reduced. The application of Mask-NMS and selective merging allowed for a reduction in the number of false alarms by approximately 8% in sequences resembling real-world conditions.

This study identifies two palette-specific post-processing pipelines.

M3 processing (three-color palette). The three-color palette increases sensitivity to small anomalies but can cause banding. Therefore, the M3 branch targets small defects according to formula (7) and uses a conservative IoU threshold (0.4) in Mask-NMS. Algorithm 2 summarizes the steps.

- 1: Convert the thermogram into a multi-channel image (palette separation)
- 2: Run YOLOv11-seg to obtain candidate masks and confidence scores.
- 3: Apply confidence filtering: keep masks with $s_k \geq \tau_c$ (e.g., $\tau_c = 0.25$)
- 4: Compute R_d for each mask using formula (6); keep only instances with $R_d < 1$
- 5: Apply Mask-NMS using an IoU threshold of τ_{IoU} (e.g., 0.4)
- 6: Detect geotags using the RTK navigation log n_i
- 7: **Return** \mathcal{D}_{small}

M2 processing (two-color palette). The two-color palette covers a wider temperature range and supports the detection of larger anomalies. The M2 branch targets

large defects according to formula (8) and uses a slightly higher Mask-NMS IoU threshold (0.5). Algorithm 3 outlines the steps.

Algorithm 3. Built-in processing of a two-color thermogram (M2)

Input: Thermogram I in a two-color palette, YOLOv11-seg model

Execution: Filtered set of defects \mathcal{D}_{large}

- 1: Normalize the thermal image and run YOLOv11-seg
- 2: Apply confidence filtering and Mask-NMS (IoU threshold 0.5)
- 3: Calculate the R_d for each mask using formula (6); keep only instances with $R_d \geq 1$
- 4: Detect geotags using RTK navigation logs n_i
- 5: Return \mathcal{D}_{large}

Implementation of late merging in thermograms

Domain differences caused by the palette prompt the merging of models. Instead of training a large multi-domain model, a lightweight late-merging ensemble is implemented based on size-routed branches defined in formulas (7)–(9). The detector is applied to both representations, generating \mathcal{D}_i^{M2} and \mathcal{D}_i^{M3} , and the edge level combines them according to the set union rule as follows:

$$\mathcal{D}_i^{ens} \equiv \mathcal{D}_{final} = \mathcal{D}_{small} \cup \mathcal{D}_{large}. \quad (10)$$

This selective integration preserves the branch that best suits the expected thermal contrast regime: M3 provides a set of small defects, while M2 provides a set of large defects. The edge level further suppresses duplicates between frames using RTK georeferencing (Section 3.8). The design is compatible with real-time operation, as it does not require joint training or a more powerful foundation; additional costs are limited to running the detector on two representations.

Georeferencing and deduplication at the edge

The UAV mission enables repeated observations of the same modules in consecutive frames and along overlapping flight paths. Without deduplication, a single defect could be counted multiple times, leading to an excessive number of alerts and reduced operator confidence. Therefore, georeferencing and deduplication using RTK are performed at the edge.

Each detection in \mathcal{D}_i is mapped to the approximate location of the module using the RTK UAV's position, altitude, and camera geometry. The exact projection depends on camera calibration and gimbal orientation; the current implementation uses a simplified mapping sufficient for clustering, with module-level granularity. Additionally, the association is refined using a site layout map.

To calculate distances between geotagged detections, the Haversine distance between latitude/longitude coordinates is used as follows:

$$d = 2R \arcsin \left(\sqrt{\sin^2 \left(\frac{\Delta \varphi}{2} \right) + \cos \varphi_1 \cos \varphi_2 \sin^2 \left(\frac{\Delta \lambda}{2} \right)} \right), \quad (11)$$

where r is the Earth's radius, φ is the latitude, and λ is the longitude.

Detections are grouped using a distance threshold ϵ , derived from the ground sampling distance and the expected RTK error.

Within each cluster, the edge server aggregates detections over a specific time interval, saves representative frame evidence, and associates the cluster with a module identifier or a string based on the plant layout. The result is a set of defect events at the module level, rather than detections at the frame level.

Condition Assessment Incorporating SCADA and Fire Risk Logic

Defect detection becomes operationally useful when it translates into actions at the plant level. KFS includes a decision-making layer that combines UAV detections with electrical and controlled signals to determine operational states and hazardous conditions. The core idea is to express safety-related conditions as

understandable logic that engineers can verify and integrate into SCADA rule engines.

Binary variables are defined as X_1 , X_2 , and X_3 . These variables indicate the presence of a fault in circuits 1–3 of the module group (or the corresponding section of the electrical diagram). The variables X_{11} , X_{12} , and X_{13} indicate whether the bypass diodes in the respective circuits are functioning correctly (1 indicates correct operation). The hazard output Y is defined as:

$$Y = X_1 \overline{X_{11}} + X_2 \overline{X_{12}} + X_3 \overline{X_{13}}, \quad (12)$$

which reflects the engineering intuition that defects combined with faulty bypass diodes increase the risk of overheating. A complete report on the research results is presented in Table 2.

Table 2. $X_1 - X_3$ indicate defects in three circuits; $X_{11} - X_{13}$ indicate that the bypass diode is functioning correctly; $Y = 1$ denotes a hazardous state

#	X_1	X_2	X_3	X_{11}	X_{12}	X_{13}	Y	Status description
1	0	0	0	0	0	0	0	normal operation
2	0	0	0	0	0	1	0	normal operation
3	0	0	0	0	1	0	0	normal operation
4	0	0	0	0	1	1	0	normal operation
5	0	0	0	1	0	0	0	normal operation
6	0	0	0	1	0	1	0	normal operation
7	0	0	0	1	1	0	0	normal operation
8	0	0	0	1	1	1	0	normal operation
9	1	0	0	0	0	0	1	Danger; bypass diode failure
10	0	1	0	0	0	0	1	Danger; bypass diode failure
11	0	0	1	0	0	0	1	Danger; bypass diode failure
12	1	1	0	0	0	0	1	Danger; bypass diode failure
13	1	0	1	0	0	0	1	Danger; failure of the bypass diode
14	0	1	1	0	0	0	1	Danger; bypass diode failure
15	1	1	1	0	0	0	1	Danger; bypass diode failure
16	1	0	0	1	0	0	0	The bypass diode blocks the circuit
17	0	1	0	0	1	0	0	The bypass diode blocks the circuit
18	0	0	1	0	0	1	0	The bypass diode blocks the circuit
19	1	1	1	1	1	1	0	Bypass diodes block the circuits

To enable additional safety sensors (such as smoke or flame detectors), the binary variables X_{21} and X_{22} are added, representing the sensor triggers. The extended hazard logic is formalized as follows:

$$Y = X_1 \overline{X_{11}} + X_2 \overline{X_{12}} + X_3 \overline{X_{13}} + X_1 X_{21} + X_2 X_{22}. \quad (13)$$

The edge server evaluates these expressions in real time and forwards alarm signals to the SCADA/HMI level, where standard procedures (e.g., dispatch, shutdown recommendation) can be executed.

Temperature Context Using the NOCT Approximation

To interpret thermal anomalies, the KFS can optionally estimate the expected module temperature

under nominal conditions using the nominal operating cell temperature (NOCT) approximation. Given an ambient temperature of T_{amb} and solar irradiance G (W/m^2), the expected module temperature is approximated by the formula:

$$T_m = T_{amb} + \frac{NOCT - 20}{800} G. \quad (14)$$

This estimate does not replace thermography; it provides context for whether the observed temperatures are plausible under current insolation and weather conditions, which can help identify instances where reflections or temporary shading are prevalent.

Research and Results

This section briefly describes a practical example of the presented study and provides quantitative results for the proposed defect detection process using KFS. This study focuses on experiments with real data and excludes synthetic experiments. The results are organized into the following categories: (i) detection quality on two thermograms (M2 and M3), (ii) the ensemble averaging effect, (iii) hardware characteristics and a comparison of YOLO models, (iv) practical validation compared to manual inspection, and (v) sensitivity to flight and weather parameters.

Experiment setup

The experiments were conducted in a closed-loop KFS cycle: (i) UAV surveying, (ii) on-board inference modeling, (iii) aggregation at the edge, and (iv) SCADA-related events, as described in Section 3.1. An industrial-grade corporate-class UAV (DJI Matrice 300 RTK) was used for surveying, providing a flight range of up to 15 km. The aircraft is equipped with a gimbal-mounted dual sensor system comprising a thermal infrared camera and a 4K RGB camera. The sensor captures synchronized thermal images and high-resolution contextual RGB frames, each of which is geotagged using the UAV's RTK GNSS with centimeter-level accuracy. Similar UAV-based thermal imaging configurations have been documented for the inspection and quality control of photovoltaic power plants [35]. The gimbal imaging system is controlled by a specialized low-power video processor (Ambarella H2 SoC) [36], which supports 4K 60 FPS video encoding. All image frames were acquired and preprocessed using OpenCV v4.9.0 [37] prior to analysis.

The flight mission was configured using specialized software (e.g., DJI Pilot 2) as a grid mission, ensuring 80% frontal and 70% lateral frame overlap, taking into account the tilt angle of the panels. The UAV control mode was semi-automatic: the operator performed the launch and general visual monitoring of flight safety, while the UAV carried out object search, positioning, and task execution (imaging along a specified route) completely autonomously. Return to base (RTH) was initiated automatically upon completion of the scanning route or when the battery charge reached a critical level (below 20%).

A real-time object detection model ran continuously on board the UAV to identify anomalies in photovoltaic panels from the thermal video stream. Ultralytics v8.3.193 YOLO models [34] were deployed and

implemented using the PyTorch v2.7.0 deep learning library [38]. Prior to deployment, the segmentation model was finely tuned on the selected thermal dataset in Table 1 using stratified five-fold cross-validation; all detection metrics listed correspond to the mean \pm standard deviation across all folds, unless otherwise specified. Training was performed on a workstation equipped with a single NVIDIA RTX 3060 (12 GB) graphics processing unit, using 300 epochs, an input resolution of 640, and a batch size of 16. Inference was performed on an embedded artificial intelligence computing platform integrated into the UAV's payload. In our configuration, the onboard inference engine uses the Qualcomm QCS605 system-on-chip [39], which provides hardware acceleration for neural processing of computer vision tasks.

During flight, the UAV transmitted compact detection messages, as defined in Equation (1), via a secure wireless communication channel (e.g., OcuSync Enterprise at 2.4/5.8 GHz) to a ground server. The mission's effectiveness was demonstrated under real-world conditions involving adverse weather factors and radio interference. In particular, even in strong winds (up to 15 m/s) or in the presence of significant electromagnetic interference from high-voltage power lines – which typically cause delays or temporary loss of video stream transmission to the operator's console – on-board processing of YOLOv11 on the Qualcomm QCS605 processor ensured stability. The system reliably cached segmentation results and full frames locally for subsequent auditing. This ensured uninterrupted task execution and the preservation of critical defect data regardless of the quality and stability of the radio link.

A ground server running on the NVIDIA Jetson Orin platform [40] collected incoming detection data and performed spatial clustering to eliminate duplicates (Section 3.8). The clustering radius was set to account for the positioning uncertainty of the UAV's RTK navigation; specifically, it was chosen to correspond to an accuracy of ~ 2 cm, which can be achieved using inexpensive dual-frequency RTK receivers such as the u-blox ZED-F9P [41]. To maximize throughput and minimize latency, the YOLO inference model was optimized on a Jetson edge server using NVIDIA TensorRT v10.9.0 [42]. Finally, the consolidated and deduplicated detection events were transmitted to the SCADA system in real time, enabling automatic monitoring and the generation of alerts for the operator.

Defect detection performance is evaluated using standard detection metrics [43]. For a given class, precision

and recall are used. Detection quality is summarized using average precision (AP) and mean average precision (mAP). This study reports the average precision (mAP@0.5), i.e., AP, calculated at an IoU threshold of 0.5 and averaged across classes. For field validation against manual monitoring, the root mean square error (RMSE) between automatic and manual defect counts is reported as defects per verified photovoltaic circuit.

Performance characteristics are determined by inference throughput (frames per second), end-to-end processing time, and sensitivity to flight and

environmental parameters (altitude, speed, time of day, weather). In Section 4.5, the accuracy metric is used to represent the proportion of correct defect detections among all detected defects based on selected IoU/confidence thresholds.

Quality as detected on two reference thermal images

Table 3 summarizes the results by class for YOLOv11-seg on M2 and M3, presented as the mean \pm standard deviation based on stratified five-fold cross-validation.

Table 3. Comprehensive comparison of YOLOv11-seg productivity on the M2 (two-color palette) and M3 (three-color palette) thermograms, presented as mean \pm standard deviation based on stratified 5-fold cross-validation

Class	M2 thermogram (2 colors)			M3 thermal image (3 colors)		
	mAP@0.5	Accuracy	Reproduction	mAP@0.5	Accuracy	Reproduction
Crack	0.93 \pm 0.04	0.91 \pm 0.05	0.89 \pm 0.06	0.87 \pm 0.07	0.86 \pm 0.07	0.84 \pm 0.08
Pollution	0.90 \pm 0.06	0.88 \pm 0.07	0.86 \pm 0.08	0.90 \pm 0.06	0.88 \pm 0.07	0.86 \pm 0.08
Stratification	0.89 \pm 0.02	0.87 \pm 0.02	0.85 \pm 0.03	0.93 \pm 0.02	0.90 \pm 0.02	0.88 \pm 0.03
Macro-average value	0.91 \pm 0.03	0.89 \pm 0.03	0.87 \pm 0.04	0.90 \pm 0.04	0.88 \pm 0.04	0.86 \pm 0.05

On the two-color palette (M2), the model achieves a macro-mAP@0.5 of 0.91 \pm 0.03, a precision of 0.89 \pm 0.03, and a reproducibility of 0.87 \pm 0.04. On the three-color palette (M3), the macro-mAP@0.5 is 0.90 \pm 0.04 with accuracy of 0.88 \pm 0.04 and reproducibility of 0.86 \pm 0.05. These average values are close, but the behavior across classes differs. Crack

detection is stronger on M2 (mAP@0.5 0.93 \pm 0.04) than on M3 (0.87 \pm 0.07), whereas delamination detection is strongest on M3 (0.93 \pm 0.02). This pattern confirms the practical observation that thermogram palettes act as separate domains and that one representation may emphasize certain defect features at the expense of others (Fig. 3).

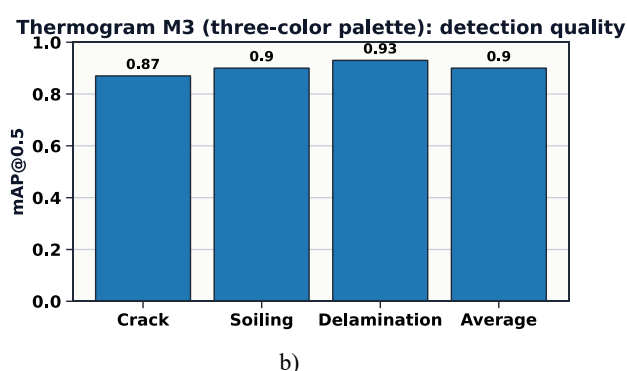
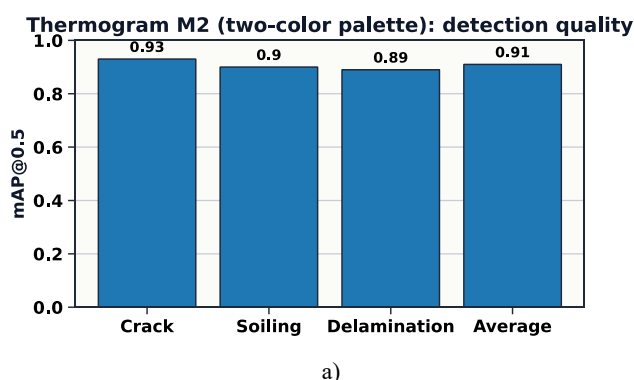


Fig. 3. Detection quality assessment (average mAP@0.5 across five runs) for YOLOv11-seg on different thermograms: a) the two-color palette (M2) is excellent for crack detection; b) the three-color palette (M3) demonstrates excellent results in delamination detection

Impact of late merging set and runtime

The late merging mode improves reliability without requiring a more powerful network. Table 4 shows the mAP@0.5 metrics before and after merging (mean \pm standard deviation across runs).

mAP@0.5 benefits the most, improving from 0.93 \pm 0.04 to 0.96 \pm 0.03. This is of practical importance because crack-like patterns can be barely visible and can be confused with thermal noise or reflections. Delamination improves from 0.93 \pm 0.02 to 0.95 \pm 0.02, while contamination remains stable. Figure 4 shows these changes.

Table 4. Effect of merging on mAP@0.5 (before and after), presented as mean ± standard deviation across stratified 5-fold cross-validation

Class	Before	After Merger
Crack	0.93 ± 0.04	0.96 ± 0.03
Soiling	0.90 ± 0.06	0.90 ± 0.06
Delamination	0.93 ± 0.02	0.95 ± 0.02

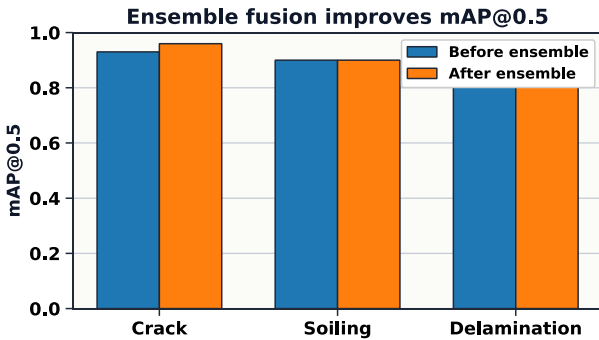


Fig. 4. Improvement in the mAP@0.5 after late merging of M2 and M3 detections

The pipeline approach also improves execution time. Thanks to early detection merging and the avoidance of repeated post-processing, the processing

Table 5. Comparison of typical edge hardware for embedded inference. The mAP@0.5 listed correspond to the deployment configuration (processor and numerical precision), not just the hardware

Platform	Output Accuracy	Throughput (FPS)	mAP@0.5	Accuracy
NVIDIA Jetson Orin Nano	FP16 (TensorRT)	100	0.95	0.93
Ambarella H2	INT8 (compressed)	60	0.70	0.75
Qualcomm QCS605	INT8 (SNPE)	80	0.85	0.82

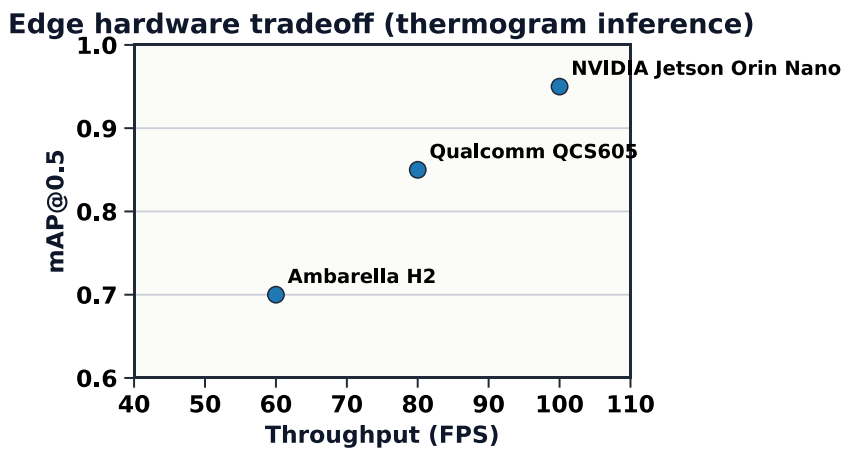


Fig. 5. Trade-off between throughput and detection quality for typical embedded platforms

To further justify the choice of YOLOv11, several generations of YOLO were trained and evaluated

time for each frame was reduced from 4.235 s to 2.858 s. Although this latency depends on the hardware and implementation, its reduction demonstrates that careful pipeline design can improve both accuracy and performance, which is critical for the operation of the KFS.

Comparison of Edge Hardware and the YOLO Model

Implementing deep learning on UAVs imposes strict constraints on computational power, weight, and power consumption. Table 5 provides an overview of typical embedded platforms and highlights a practical trade-off: throughput determines the maximum possible flight speed and frame rate (see Fig. 5).

It is important to note that detection quality (mAP/accuracy) is a property of the deployed model and numerical precision, not just the hardware. Therefore, the values of mAP@0.5 in Table 5 correspond to a deployment configuration specific to the platform (e.g., FP16 versus INT8 quantization and, if necessary, additional compression to meet memory/latency constraints).

under a unified protocol. Table 6 shows accuracy, reproducibility, and mAP@0.5, as well as training time.

YOLOv11 demonstrates the highest overall productivity (accuracy 0.96, reproducibility 0.95, mAP@0.5 0.93), although training time is longer than for some smaller

models. The results indicate that for embedded inference, YOLOv11 provides a favorable balance between accuracy and manageable deployment complexity.

Table 6. Quantitative comparison of YOLO model generations (segmentation). Training time is measured over 300 epochs on a single NVIDIA RTX 3060 (12 GB) GPU with an input resolution of 640 and a batch size of 16

Model	Accuracy	Recall	mAP@0.5	Training time (min)
YOLOv5	0.92	0.90	0.86	55
YOLOv8	0.91	0.88	0.83	45
YOLOv9	0.94	0.80	0.90	59
YOLOv10	0.94	0.86	0.90	132
YOLOv11	0.96	0.95	0.93	106

On-site data verification versus manual monitoring

The KFS, designed for plant operation, must match the results of a human-conducted field inspection. The performance of the AIC is verified through manual monitoring at two photovoltaic installations: a rooftop and a ground-mounted site. The evaluation compares the automatic defect count after edge-level deduplication with the manual count obtained from expert inspection. The test set consisted of approximately 2,500 photovoltaic modules across two sites and included 145 expert-verified defect cases; the agreement was quantified using an RMSE of 0.71 defects per inspected photovoltaic string.

Beyond the overall error, field tests highlight the role of KFS components that are not visible in offline metrics. The built-in inference alone yields repeated detections of the same defect in adjacent frames; RTK-based clustering and edge-level deduplication reduce this redundancy and generate module-level events suitable for reporting. Similarly, when abnormal heating is detected, the logic supporting SCADA can elevate the event's priority, especially if the bypass diode status indicates an elevated risk. These elements are critical for operational implementation, as operators require stable counts, clear module identifiers, and actionable alerts, rather than a stream of raw detections.

Sensitivity to Flight and Environmental Parameters

Thermal imaging is limited by physical factors. Therefore, flight altitude, speed, time of day, and weather conditions are analyzed in terms of their impact

on detection quality. The results quantitatively evaluate the operating parameters of the proposed thermal imaging system.

Altitude provides an optimal balance between resolution and coverage (Table 7). At an altitude of 5 m, detection accuracy reaches 98% with a reproducibility of 96% (Fig. 6, a), but coverage is limited, and mission duration increases. At a height of 10 m, accuracy remains high (93%, reproducibility – 90%), and coverage increases approximately 4–5 times, making it acceptable for operation. At a height of 15 m, accuracy decreases to 84% (reproducibility 79%), indicating insufficient spatial resolution for detecting minor defects.

Flight speed affects motion blur and temporal redundancy (Table 7, Fig. 6, b). At speeds of 3–7 m/s, accuracy remains above 91% and reproducibility above 88%, indicating that the system can handle moderate speeds when the camera shutter and stabilization are sufficient. At a speed of 10 m/s, accuracy drops to 85%, and reproducibility to 82%, which may be unacceptable for safety-critical monitoring.

Environmental conditions are also important (Table 8). Surveys conducted during midday (12:00–14:00) yield the best results (96% accuracy, 94% reproducibility), corresponding to stronger radiation and thermal contrast (Fig. 6, c). Cloudy conditions improve performance (96% accuracy, 94% reproducibility) compared to clear skies (92% accuracy, 89% reproducibility), likely because diffuse lighting reduces specular reflections (Fig. 6, d). These results are consistent with previous observations that reflections may dominate thermograms under clear conditions and that careful flight planning is necessary [7, 11].

Table 7. Sensitivity to height and speed (accuracy/reproducibility)

Height			Speed		
m	Accuracy (%)	Reproducibility (%)	m/s	Accuracy (%)	Repeatability (%)
5	98	96	3	94	92
10	93	90	5	93	90
15	84	79	7	91	88
			10	85	82

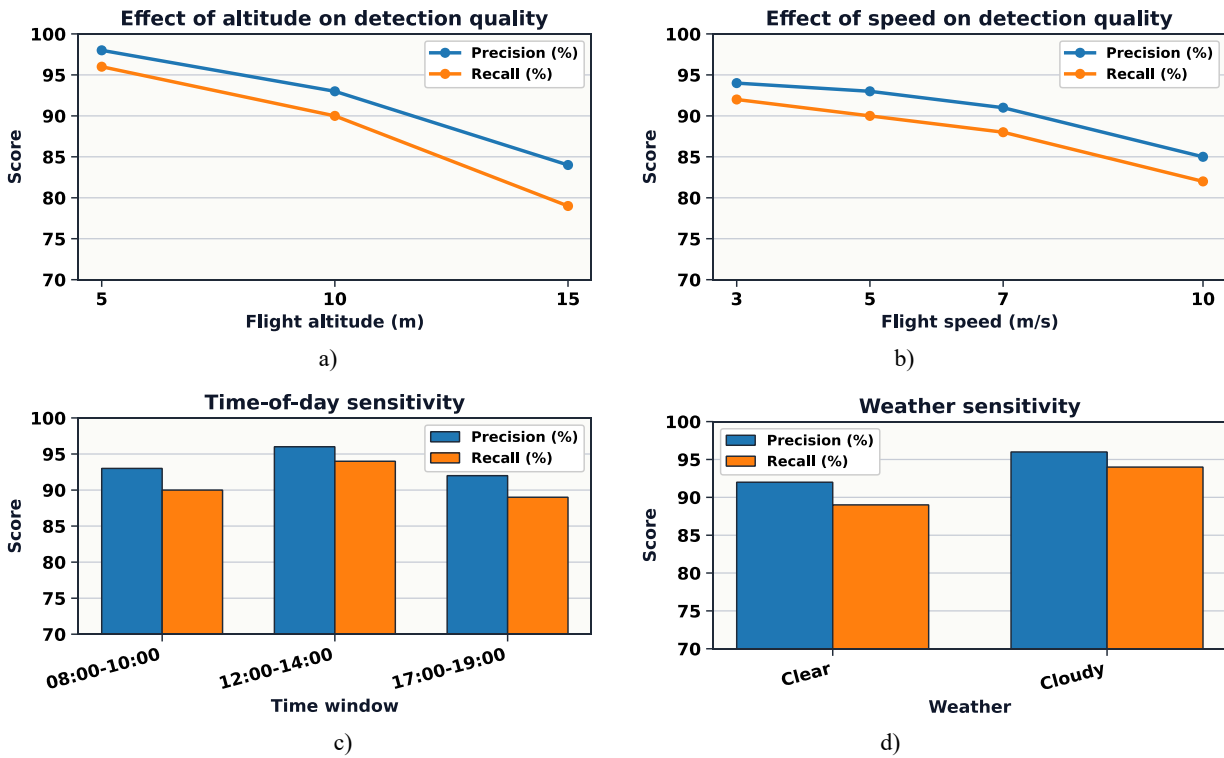


Fig. 6. Sensitivity of detection to changes in operating conditions and the environment. The results show that an increase in a) flight altitude and b) speed degrades detection quality, requiring strict adherence to flight profiles. Environmental factors also affect performance: c) the time from 12:00 to 14:00 and d) cloudy weather provide optimal accuracy and reproducibility

Table 8. Sensitivity to time of day and weather (accuracy/reproducibility)

Time of day			Weather		
Window	Accuracy (%)	Reproducibility (%)	Conditions	Accuracy (%)	Reproducibility (%)
8:00–10:00	93	90	Clear	92	89
12:00–14:00	96	94			
5:00 PM–7:00 PM	92	89	Cloudy	96	94

Quantitative comparison with the current state of the art

The obtained results are compared with typical approaches described in the literature. A direct comparison

is imperfect, as the datasets, defect definitions, and evaluation protocols differ. However, Table 9 provides a quantitative basis and explains how the proposed KFS differs in scope, which is further illustrated in Fig. 7.

Table 9. Quantitative comparison with published approaches (according to the respective authors)

Work	Modality	Model / approach	Task	Reported metric
Dotenko et al. [25]	IR	Statistical tests + vision	Defect classification	F1 = 93.88%
Di Tommaso et al. [14]	IR	Multi-stage YOLOv3	Hotspot detection	AP@0.5: 66.9%
Xi et al. [18]	IR	ST-YOLO	Detection of faults in photovoltaic systems	mAP@0.5 = 96.6%
This work	IR + RGB	YOLOv11-seg ensemble + KFS	Defect segmentation + decision making	mAP@0.5s up to 96% (cracks)

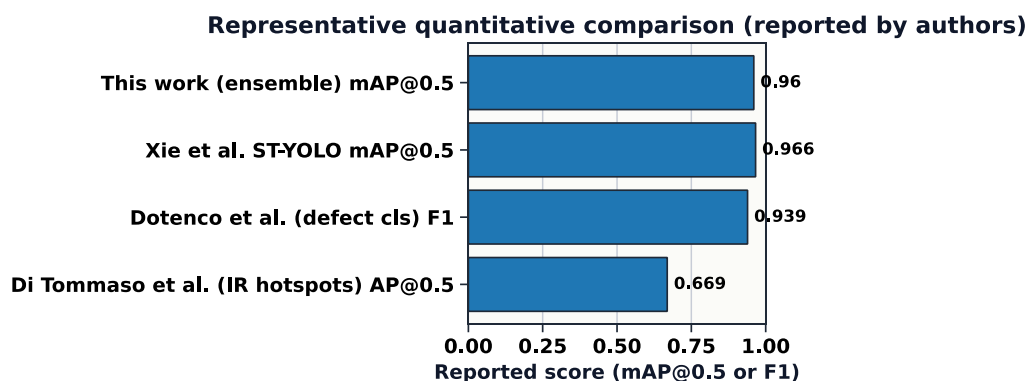


Fig. 7. Illustrative quantitative comparison (indicators shown). Values are not strictly comparable across datasets; the figure contextualizes the magnitude of the results

The literature notes that a high mAP can be achieved under favorable conditions, for example, using ST-YOLO on infrared images [18]. The proposed ensemble achieves a comparable mAP@0.5 for the crack class (0.96), incorporating system-level components necessary for the facility's operation: georeferencing, deduplication, and SCADA-aware decision-making logic. Multi-stage approaches, such as those by Di Tommaso et al. [14], report a significantly lower AP@0.5 anomaly detection rate on hotspots (66.9%), illustrating that anomaly detection can be more challenging than module detection and that performance is strongly influenced by the dataset and defect type. Dotenko et al. [25] reported high efficiency in defect classification using classical vision and statistical tests, which remains relevant as a baseline metric and demonstrates that robust preprocessing is crucial in infrared aerial photography.

Discussion

The presented results confirm two interrelated theses. First, a modern detector with a segmentation function can achieve high accuracy in defect localization on UAV thermal images even with a small training dataset, provided that post-processing requirements and operational constraints are met. Second, and more importantly, the operational utility of UAV inspection depends on a cyber-physical architecture that links detection results with georeferenced, deduplicated events and plant monitoring.

From a computer vision perspective, palette-specific results illustrate a subtle but important phenomenon: the representation of the thermogram is part of the model's input domain. The two-color palette (M2) highlights

certain gradients and enables more effective crack detection, while the three-color palette (M3) enhances delamination detection. This is consistent with broader observations in photothermal thermography that the choice of acquisition and visualization methods affects defect contrast and, consequently, their detection [4, 6]. The late-merging ensemble improves reliability by leveraging this complementarity. The improvement for cracks (from 0.93 to 0.96 mAP@0.5) is particularly important, as cracks are often early-stage defects whose thermal signals can be weak. The reduction in runtime (from 4.235 s to 2.858 s) further underscores that pipeline design can yield tangible benefits.

From a systems perspective, the peripheral layer is not a luxury. Without RTK-based georeferencing and deduplication, raw UAV detections would yield inflated metrics and unstable alarms. Therefore, the architecture treats georeferencing as a first-class signal, consistent with field studies that highlight the practical difficulty of tracing anomalies back to modules and chains [7, 11]. The edge layer also provides a natural integration point for SCADA signals, such as bypass diode status. The hazard Boolean logic, formalized by equations (12)–(13), is intentionally interpretable: plant engineers can verify, adjust, and integrate it into existing rule mechanisms. This contrasts with “black box” alarm models, which can be difficult to certify for safety monitoring purposes.

Sensitivity analysis offers another argument in favor of KFS framing. The sharp drop in productivity at a height of 15 m and a speed of 10 m/s indicates that the detector's productivity is limited by physics and optics, not just the model's power. Similarly, better productivity in cloudy conditions suggests that reflections may be the primary source of false alarms.

These observations prompt the inclusion of reflection-aware planning (Section 3.3) and the use of contextual temperature models (Section 3.10) as part of the overall system. In other words, the “best model” alone is insufficient; the system must actively generate its own data through flight planning and quality control.

A key limitation of this evaluation is the small and unbalanced labeled dataset for minority classes. Even with stratified five-fold cross-validation, the crack and contamination classes contain only 20 and 10 labeled images, respectively, leading to greater variance (Tables 3–4). Therefore, the reported mean values should be interpreted as feasibility indicators rather than definitive benchmarks. Expanding labeled datasets and adopting standardized public benchmarks for UAV thermography with consistent metadata would significantly improve statistical reliability and reproducibility.

The proposed approach also has drawbacks and trade-offs. The architecture complicates deployment: a peripheral server must be maintained on-site, and reliable communication between the UAV and the periphery must be ensured. The approach relies on RTK positioning and continuous camera calibration; failure of these subsystems can degrade georeferencing and reduce confidence in module-level mapping. Although transfer learning and augmentation mitigate the need for retraining, generalization to new defect types, new thermal cameras, or different palettes remains a challenge. Future work should explore palette-independent representations or temperature-calibrated models that operate directly on radiometric data rather than on reconstructed palettes.

Several limitations point to open research directions. First, thermography remains sensitive to environmental conditions; a key challenge is to incorporate radiation intensity, wind, and ambient temperature into training models in a principled manner, potentially using physically grounded characteristics or uncertainty-aware inference. Second, current SCADA logic is based on explicit Boolean rules; hybrid approaches combining learned risk assessment with interpretable rules can increase flexibility while maintaining auditability. Third, a simplified projection is currently used during the georeferencing stage; improvements in geometric modeling (e.g., through camera calibration and 3D facility models) could reduce association errors in dense installations. Finally, there is a broader issue

of comparative analysis: the photovoltaic installation community would benefit from standardized, publicly available UAV thermography datasets with consistent labels and metadata (flight altitude, solar irradiance, camera parameters), which would enable reproducible comparisons between systems.

Conclusions

This paper proposes an architecture for cyber-physical systems designed to monitor photovoltaic power plants, which successfully bridges the gap between raw computer vision data and practical operational maintenance. By integrating real-time UAV-based YOLOv11 segmentation with an intelligent edge computing layer for RTK-supported geotagging and duplicate suppression, the system effectively transforms fragmented output results into trackable defect events at the module level. Experimental results confirm the superior performance of the proposed late fusion ensemble, which significantly improved crack detection to 0,96mAP@0.5 and delamination detection to 0,95mAP@0.5, while reducing the processing latency per frame to 2.858 s. Field validation further demonstrated consistency with manual inspection records, achieving an RMSE of 0.71 defects per inspected PV row and identifying a flight altitude of 10 m as the most balanced operating point with 93% accuracy and 90% reproducibility. The unique integration of SCADA-aware interpretive logic enables real-time fire risk assessment based on the status of the bypass diode. Despite these successes, notable limitations of this study include sensitivity to specular reflections under clear sky conditions, reduced productivity at flight speeds exceeding 7 m/s, and statistically unstable estimates for a minority of defect classes due to the limited amount of labeled data. The dependence on precise RTK calibration for module association also poses a challenge for deployment in dense array configurations. Overall, the proposed architecture offers a scalable, safety-oriented solution for automating the lifecycle management of large-scale solar assets.

Future research will focus on palette-independent radiometric models to reduce the impact of ambient noise, larger-scale benchmark-based assessments to reduce uncertainty, and predictive maintenance algorithms utilizing historical SCADA trends.

Conflict of interest

The authors declare that they have no conflicts of interest, including financial, personal, authorship, or any other conflicts that could influence the research or the results published in this article.

Funding

This research was supported by the Ministry of Education and Science of Ukraine and funded by the European Union's external aid instrument as part of Ukraine's commitments under the EU Framework Program for Research and Innovation "Horizon 2020". The work was carried out as part of the research project "Intelligent Defect Detection System for Green Energy Facilities Using UAVs", state registration number 0124U004665 (2024–2026).

Data availability

Data will be provided upon reasonable request. The base datasets (STHS-277 and PVF-10) are publicly

available. Analysis scripts may be provided by the authors via email upon reasonable request. The proprietary dataset obtained during field validation is not publicly available due to commercial confidentiality.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technology in the creation of this paper.

Acknowledgments

The authors express their sincere gratitude to the Ministry of Education and Science of Ukraine for its support. This research was funded by the European Union's external aid instrument under the "Horizon 2020" research and innovation program. The work was carried out as part of the project "Intelligent Defect Detection System for Green Energy Facilities Using UAVs" (state registration number 0124U004665).

All authors have read and approved the published version of this manuscript.

References

1. Sun, L., Sun, Y. (2022), "Photovoltaic power forecasting based on artificial neural network and ultraviolet index", *International Journal of Computing*, Vol. 21, No. 2, 1531158 p. DOI: <https://doi.org/10.47839/ijc.21.2.2583>
2. Sinha, A., Sulas-Kern, D. B., Owen-Bellini, M., Spinella, L., Uličná, Š., Ayala, R., Murphy, R., Walker, C., Hacke, P. (2024), "A review of degradation and reliability analysis of a solar PV module", *IEEE Access*, Vol. 12, pp. 185036-185056. DOI: <https://doi.org/10.1109/ACCESS.2024.3432394>
3. Lynnyk, R., Vysotska, V., Matseliukh, Y., Burov, Y., Demkiv, L., Sachenko, A., Shylinska, I., Yevseyeva, I., Bihun, O. (2020), "DDoS attacks analysis based on machine learning in challenges of global changes", *Proceedings of the 2nd International Workshop on Modern Machine Learning Technologies and Data Science (MoMLet+DS 2020), Volume 2631*, CEUR.org, Aachen, pp. 159-171, available at: <https://ceur-ws.org/Vol-2631/paper12.pdf> (last accessed 16.01.2025).
4. Aghaei, M., Kolahi, M., Nedaei, A., Venkatesh, N. S., Esmailifar, S. M. (2025), "Autonomous intelligent monitoring of photovoltaic systems: An in-depth multidisciplinary review", *Progress in Photovoltaics: Research and Applications*, Vol. 33, No. 3, pp. 381-409. DOI: <https://doi.org/10.1002/pip.3859>
5. Khatri, A., Khadka, S., Lamichhane, N., Shrestha, R. (2025), "A comprehensive review of infrared thermography and deep learning applications for solar photovoltaic systems", *Infrared Physics & Technology*, Vol. 148, 105878 p. DOI: <https://doi.org/10.1016/j.infrared.2025.105878>
6. Gallardo-Saavedra, S., Hernández-Callejo, L., Alonso-García, M. d. C., Muñoz-Cruzado-Alba, J., Ballestín-Fuertes, J. (2020), "Infrared thermography for the detection and characterization of photovoltaic defects: Comparison between illumination and dark conditions", *Sensors*, Vol. 20, No. 16, 4395 p. DOI: <https://doi.org/10.3390/s20164395>
7. Prasshanth, C. V., Deepak, B. B. V. L., Vijayakumar, N., Prabadevi, B., Alhussen, F. A. (2025), "Fault detection in photovoltaic systems using unmanned aerial vehicle-captured images and rough set theory", *Solar Energy*, Vol. 290, 113348 p. DOI: <https://doi.org/10.1016/j.solener.2025.113348>
8. Moctezuma, D., Valentín-Coronado, L. M. (2025), "Deep-learning-based electrical fault detection in photovoltaic modules through aerial infrared imaging: Addressing data complexity", *Computación y Sistemas*, Vol. 29, No. 1, pp. 65-75. DOI: <https://doi.org/10.13053/cys-29-1-5531>

9. Svystun, S., Scislo, L., Pawlik, M., Melnychenko, O., Radiuk, P., Savenko, O., Sachenko, A. (2025), "DyTAM: Accelerating wind turbine inspections with dynamic UAV trajectory adaptation", *Energies*, Vol. 18, No. 7, 1823 p. DOI: <https://doi.org/10.3390/en18071823>
10. Rouibah, N., El Hammoumi, A., Bouttout, A., Haddad, S., Oukaci, S., Limam, A., Benghanem, M. (2025), "Smart monitoring of photovoltaic energy systems: An IoT-based prototype approach", *Scientific African*, Vol. 30, e02973 p. DOI: <https://doi.org/10.1016/j.sciaf.2025.e02973>
11. Rodriguez-Vazquez, J., Prieto-Centeno, I., Fernandez-Cortizas, M., Pérez-Saura, D., Molina, M., Campoy, P. (2024), "Real-time object detection for autonomous solar farm inspection via UAVs", *Sensors*, Vol. 24, No. 3, 777 p. DOI: <https://doi.org/10.3390/s24030777>
12. Noura, H., Sun, Y., Taghezouit, B., Dairi, A. (2024), "Explainable artificial intelligence of tree-based algorithms for fault detection and diagnosis in grid-connected photovoltaic systems". *Engineering Applications of Artificial Intelligence*, Vol. 139, 109503 p. DOI: <https://doi.org/10.1016/j.engappai.2024.109503>
13. Bodyanskiy, Y., Deineko, A., Skorik, V., Brodetskiy, F. (2022), "Deep neural network with adaptive parametric rectified linear units and its fast learning", *International Journal of Computing*, Vol. 21, No. 1, pp. 11-18. DOI: <https://doi.org/10.47839/ijc.21.1.2512>
14. Di Tommaso, A., Betti, A., Fontanelli, G., Michelozzi, B. (2022), "A multi-stage model based on YOLOv3 for defect detection in PV panels based on IR and visible imaging by unmanned aerial vehicle", *Renewable Energy*, Vol. 193, pp. 941-962. DOI: <https://doi.org/10.1016/j.renene.2022.04.046>
15. Melnychenko, O., Scislo, L., Savenko, O., Sachenko, A., Radiuk, P. (2024), "Intelligent integrated system for fruit detection using multi-UAV imaging and deep learning", *Sensors*, Vol. 24, No. 6, 1913 p. DOI: <https://doi.org/10.3390/s24061913>
16. Aljafari, B., Satpathy, P. R., Thanikanti, S. B., Nwulu, N. (2024), "Supervised classification and fault detection in grid-connected PV systems using 1D-CNN: Simulation and real-time validation", *Energy Reports*, Vol. 12, pp. 2156-2178. DOI: <https://doi.org/10.1016/j.egy.2024.08.008>
17. Setiawan, E. A., Fathurrahman, M. (2025), "Exploring insights on deep learning-based photovoltaic fault detection for monofacial and bifacial modules using thermography", *International Journal of Cognitive Computing in Engineering*, Vol. 6, pp. 495-507. DOI: <https://doi.org/10.1016/j.ijcce.2025.04.001>
18. Xie, H., Yuan, B., Hu, C., Gao, Y., Wang, F., Wang, C., Wang, Y., Chu, P. (2024), "ST-YOLO: A defect detection method for photovoltaic modules based on infrared thermal imaging and machine vision technology", *PLOS ONE*, Vol. 19, No. 12, pp. 1-19. DOI: <https://doi.org/10.1371/journal.pone.0310742>
19. Carni, D. L., Grimaldi, D., Lamonaca, F., Nigro, L., Sciammarella, P. F. (2017), "From distributed measurement systems to cyber-physical systems: A design approach", *International Journal of Computing*, Vol. 16, No. 2, pp. 66-73. DOI: <https://doi.org/10.47839/ijc.16.2.882>
20. Boucif, O. H., Toral-Cruz, H. (2025), "Artificial intelligence of things for solar energy monitoring and control", *Applied Sciences*, Vol. 15, No. 11, 6019 p. DOI: <https://doi.org/10.3390/app15116019>
21. Lysyi, A., Sachenko, A., Radiuk, P., Lysyi, M., Melnychenko, O., Ishchuk, O., Savenko, O. (2025), "Enhanced fire hazard detection in solar power plants: An integrated UAV, AI, and SCADA-based approach", *Radioelectronic and Computer Systems*, Vol. 2025, No. 2, pp. 99-117. DOI: <https://doi.org/10.32620/reks.2025.2.06>
22. Tang, W., Xu, J., Wang, X., Li, H., Chen, J., Zhang, Y. (2022), "Deep learning-based detection of linear defects in large-scale photovoltaic plants using an edge-cloud architecture", *Solar Energy*, Vol. 237, pp. 144-161. DOI: <https://doi.org/10.1016/j.solener.2021.11.016>
23. Ferlito, S., Ippolito, S., Santagata, C., Schiattarella, P., Di Francia, G. (2024), "A study on an IoT-based SCADA system for photovoltaic utility plants", *Electronics*, Vol. 13, No. 11, 2065 p. DOI: <https://doi.org/10.3390/electronics13112065>
24. Masita, K., Hasan, A., Shongwe, T., Abu Hilal, H. (2025), "Deep learning in defects detection of PV modules: A review", *Solar Energy Advances*, Vol. 5, 100090 p. DOI: <https://doi.org/10.1016/j.seja.2025.100090>
25. Dotenco, S., Dalsass, M., Winkler, L., Würzner, T., Brabec, C., Maier, A., Gallwitz, F. (2016), "Automatic detection and analysis of photovoltaic modules in aerial infrared imagery", *Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, IEEE, New York, NY, USA, pp. 1-9. DOI: <https://doi.org/10.1109/WACV.2016.7477658>
26. Ebied, M. A., Munshi, A., Alhuzali, S. A., El-sotouhy, M. M., Shehta, A. I., Elborlsy, M. S. (2025), "Advanced deep learning modeling to enhance detection of defective photovoltaic cells in electroluminescence images", *Scientific Reports*, Vol. 15, 31640 p. DOI: <https://doi.org/10.1038/s41598-025-14478-y>
27. Barraz, Z., Sebari, I., Lamrini, N., Ait El Kadi, K., Ait Abdelmoula, I. (2025), "Fast and automatic solar module geo-labeling for optimized large-scale photovoltaic systems inspection from UAV thermal imagery using deep learning segmentation", *Cleaner Engineering and Technology*, Vol. 28, 101048 p. DOI: <https://doi.org/10.1016/j.clet.2025.101048>
28. Abdelsattar, M., AbdelMoety, A., Emad-Eldeen, A. (2025), "ResNet-based image processing approach for precise detection of cracks in photovoltaic panels", *Scientific Reports*, Vol. 15, 24356 p. DOI: <https://doi.org/10.1038/s41598-025-09101-z>

29. Svystun, S., Melnychenko, O., Radiuk, P., Savenko, O., Sachenko, A., Lysyi, A. (2024), "Thermal and RGB images work better together in wind turbine damage detection", *International Journal of Computing*, Vol. 23, No. 4, pp. 526-535. DOI: <https://doi.org/10.47839/ijc.23.4.3752>
30. 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, No. 3, 927 p. DOI: <https://doi.org/10.3390/su17030927>
31. Alfaro-Mejía, E., Loaiza-Correa, H., Franco-Mejía, E., Restrepo-Girón, A. D., Nope-Rodríguez, S. E. (2019), "Dataset for recognition of snail trails and hot spot failures in monocrystalline Si solar panels", *Data in Brief*, Vol. 26, 104441 p. DOI: <https://doi.org/10.1016/j.dib.2019.104441>
32. Wang, B., Chen, Q., Wang, M., Chen, Y., Zhang, Z., Liu, X., Gao, W., Zhang, Y., Zhang, H. (2024), "PVF-10: A high-resolution unmanned aerial vehicle thermal infrared image dataset for fine-grained photovoltaic fault classification", *Applied Energy*, Vol. 376, 124187 p. DOI: <https://doi.org/10.1016/j.apenergy.2024.124187>
33. CVAT.ai Corporation, (2024), "CVAT: Computer vision annotation tool (Version 2.16.1)", DOI: <https://doi.org/10.5281/zenodo.12771595>
34. Jocher, G., Qiu, J., Chaurasia, A. (2025), "Ultralytics YOLO", available at: <https://github.com/ultralytics/ultralytics> (last accessed 16.01.2026).
35. European Commission, (2018), "Standards for the assessment of the environmental performance of photovoltaic modules, power conversion equipment and photovoltaic systems", Publications Office of the European Union, Luxembourg, available at: 10.2760/89830
36. Ambarella, Inc. (2016), "Ambarella introduces low power 4K ultra HD SoCs for sports and flying cameras", available at: <https://www.ambarella.com/news/ambarella-introduces-low-power-4k-ultra-hd-socs-for-sports-and-flying-cameras/> (last accessed 16.01.2026).
37. Bradski, G. (2000), "The OpenCV library", *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, Vol. 25, No. 11, pp. 120-123.
38. Paszke, A. et al. (2019), "PyTorch: An imperative style, high-performance deep learning library", *Advances in Neural Information Processing Systems, Volume 32*, Curran Associates, Inc., pp. 8024-8035, available at: https://proceedings.neurips.cc/paper_files/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf (last accessed 16.01.2026).
39. Qualcomm Technologies, Inc. (2020), "Qualcomm QCS603/QCS605 SoCs for IoT: Product brief", available at: https://www.qualcomm.com/content/dam/qcomm-martech/dm-assets/documents/qcs603.605-socs-product-brief_87-pg764-1-c.pdf (last accessed 16.01.2026).
40. NVIDIA Corporation, (2024), "Jetson AGX Orin: Technical brief", available at: <https://www.nvidia.com/en-us/lp/embedded-computing/robotics-edge-ai-tech-brief/> (last accessed 16.01.2026).
41. Robustelli, U., Cutugno, M., Pugliano, G. (2023), "Low-cost GNSS and PPP-RTK: Investigating the capabilities of the u-blox ZED-F9P module", *Sensors*, Vol. 23, No. 13, 6074 p. DOI: <https://doi.org/10.3390/s23136074>
42. NVIDIA Corporation, (2024), "NVIDIA TensorRT documentation", available at: <https://docs.nvidia.com/deeplearning/tensorrt/10.9.0/getting-started/release-notes.html> (last accessed 16.01.2026).
43. Rainio, O., Teuvo, J., Klén, R. (2024), "Evaluation metrics and statistical tests for machine learning", *Scientific Reports*, Vol. 14, No. 1, 6086 p. DOI: <https://doi.org/10.1038/s41598-024-56706-x>

Received (Надійшла) 18.01.2026

Accepted for publication (Прийнята до друку) 18.02.2026

Publication date (Дата публікації) 30.03.2026

Відомості про авторів / About the Authors

Саченко Анатолій Олексійович – доктор технічних наук, професор, Західноукраїнський національний університет, директор Науково-дослідного інституту інтелектуальних комп'ютерних систем; Тернопіль, Україна; Радомський університет імені Казімежа Пуласького; Радом, Польща;

Anatoliy Sachenko – Doctor of Technical Sciences, Professor, West Ukrainian National University, Director of the Research Institute for Intelligent Computer Systems; Ternopil, Ukraine; Kazimierz Pulaski University of Radom; Radom, Poland;

e-mail: as@wunu.ua

ORCID ID: <https://orcid.org/0000-0002-0907-3682>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=35518445600>

Радюк Павло Михайлович – доктор філософії, Хмельницький національний університет, доцент кафедри комп'ютерних наук; Хмельницький, Україна;

Pavlo Radiuk – PhD in Computer Science, Khmelnytskyi National University, Associate Professor at the Department of Computer Science; Khmelnytskyi, Ukraine;

e-mail: radiukp@khnmu.edu.ua

ORCID ID: <https://orcid.org/0000-0003-3609-112X>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=57216894492>

Лисий Микола Іванович – доктор технічних наук, професор, Національна академія Державної прикордонної служби України імені Богдана Хмельницького; Хмельницький, Україна;

Mykola Lysyi – Doctor of Technical Sciences, Professor, Bohdan Khmelnytskyi National Academy of the State Border Guard Service of Ukraine; Khmelnytskyi, Ukraine;

e-mail: lisiy3152@ukr.net

ORCID ID: <https://orcid.org/0000-0002-9858-706X>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=59482244100>

Мельниченко Олександр Вікторович – доктор філософії, Хмельницький національний університет, старший викладач кафедри комп'ютерної інженерії та інформаційних систем; Хмельницький, Україна;

Oleksandr Melnychenko – PhD in Computer Science, Khmelnytskyi National University, Senior Lecturer at the Department of Computer Engineering and Information Systems; Khmelnytskyi, Ukraine;

e-mail: melnynchenko@khnmu.edu.ua

ORCID ID: <https://orcid.org/0000-0001-8565-7092>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=45961382900>

Заставний Олег Михайлович – кандидат технічних наук, Західноукраїнський національний університет, старший викладач кафедри спеціалізованих комп'ютерних систем; Тернопіль, Україна;

Oleg Zastavnyi – Candidate of Technical Sciences, West Ukrainian National University, Senior Lecturer at the Department of Specialized Computer Systems; Ternopil, Ukraine;

e-mail: o.zastavnyi@wunu.edu.ua

ORCID ID: <https://orcid.org/0000-0001-8630-8791>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=8366871500>

АРХІТЕКТУРА КІБЕРФІЗИЧНИХ СИСТЕМ ДЛЯ ВИЯВЛЕННЯ ДЕФЕКТІВ У ФОТОЕЛЕКТРИЧНИХ МОДУЛЯХ НА ОСНОВІ ПІЗЬНОГО ЗЛИТТЯ ДАНИХ ІЗ БПЛА

Предмет дослідження – архітектурне вдосконалення систем інспектування великих сонячних електростанцій на основі БПЛА. **Метою дослідження** є поліпшення операційної корисності інспектування фотоелектричних модулів на основі БПЛА через розроблення кіберфізичної архітектури, яка інтегрує вбудоване глибоке навчання, периферійні вузли, хмарну аналітику та логіку прийняття рішень з урахуванням SCADA в єдиний скоординований робочий процес.

Завдання дослідження: 1) формалізувати багаторівневу архітектуру КФС (БПЛА-периферія-хмара) та визначити інтерфейси для даних, гео-тегів та тривоги; 2) розробити та валідувати вбудований конвеєр термографічного виявлення з обробкою та злиттям з урахуванням палітри; 3) інтегрувати результати виявлення з логічним шаром, що враховує SCADA, для інференції небезпеки та зниження ризику пожеж. **Методи дослідження:** методи комп'ютерного зору та глибокого навчання (YOLOv11) для бортової сегментації дефектів; ансамблювання моделей (стратегія пізнього злиття M2 та M3 палітр) для нівелювання зсуву домену; алгоритми просторової кластеризації з підтримкою RTK для геоіндексації та дедуплікації; детермінована булева логіка для визначення пожежних ризиків на основі станів шунтуючих діодів.

Результати п'ятикратної перехресної валідації свідчать, що запропонована архітектура значно перевершує одномодальні підходи. Бортова модель YOLOv11 досягла показників mAP@0.5 на рівні 0,91 та 0,90 для палітр M2 та M3 відповідно. Ансамбль пізнього злиття підвищив mAP@0.5 для тріщин до 0,96, а розшарування – до 0,95. Запропонована архітектура зменшила наскрізну затримку оброблення кадру із 4,235 с до 2,858 с. Польова перевірка продемонструвала RMSE 0,71 дефекту на інспектований рядок порівняно з ручним підрахуванням. Аналіз чутливості

виявив, що висота польоту 10 м забезпечує оптимальне узгодження між роздільною здатністю та охопленням (точність 93 %, відтворюваність 90 %). **Висновки:** дослідження демонструє, що розгляд інспектування БПЛА як інтегрованого кіберфізичного сервісу суттєво покращує точність виявлення дефектів. Запропонована архітектура кіберфізичних систем є масштабованим рішенням для профілактичного обслуговування та автоматизованого зниження ризику пожеж у відновлюваній енергетиці.

Ключові слова: кіберфізичні системи; фотоелектричні модулі; безпілотні літальні апарати; інфрачервона термографія; периферійні обчислення; глибоке навчання.

Бібліографічні описи / Bibliographic descriptions

Саченко А. О., Радюк П. М., Лисий М. І., Мельниченко О. В., Заставний О. М. Архітектура кіберфізичних систем для виявлення дефектів у фотоелектричних модулях на основі пізнього злиття даних із БПЛА. *Сучасний стан наукових досліджень та технологій в промисловості*. 2026. № 1 (35). С. 79–99. DOI: <https://doi.org/10.30837/2522-9818.2026.1.079>

Sachenko, A., Radiuk, P., Lysyi, M., Melnychenko, O., Zastavnyy, O. (2026), "Architecture of cyberphysical systems for UAV-based late-fusion defect detection in photovoltaic modules", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (35), P. 79–99. DOI: <https://doi.org/10.30837/2522-9818.2026.1.079>

Yu Jian, Oksana Sitnikova

METHOD FOR PREDICTING UAV TRAJECTORIES AND YAW FOR INDUSTRIAL AUTONOMOUS MISSIONS IN A DYNAMIC ENVIRONMENT

This paper addresses the problem of predicting the trajectory and anticipatory evasion of an unmanned aerial vehicle (UAV) in industrial autonomous missions in a dynamic environment, subject to noise in navigation measurements and energy constraints, which pose risks of delayed or excessive maneuvering and increased deviation from the route. **Objective.** To develop and verify, using a simulation model, a method that ensures prediction-based obstacle avoidance with control of deviation from the reference trajectory and maneuvering energy consumption. **Tasks.** Formulate the architecture of the avoidance system; develop a predictor of future coordinates based on a recurrent neural network with long-term short-term memory; determine a method for assessing collision risk using a safety zone; implement a trajectory correction algorithm taking into account the “safety–deviation–energy consumption” trade-off; perform a comparative evaluation with baseline methods. **Methods.** Coordinate predictions are constructed based on time sequences of coordinates and motion parameters; collision risk is assessed by analyzing the intersection of the predicted trajectory with obstacle safety zones; trajectory correction is formalized as an optimization problem to select a maneuver that minimizes the total tracking error and the proximity penalty. The effectiveness was verified in a Python environment on standard trajectories (straight, circular, and polygonal) by comparison with pure tracking, line-of-sight, vector field, and nonlinear stabilization methods. **Results.** The proposed approach achieved the smallest mean deviation from the trajectory (14.95 m), the lowest maneuver energy consumption (72 conventional units), the highest tracking success rate (86.08%), and the highest overall productivity coefficient (0.494) among the algorithms considered; a trade-off was observed regarding the minimum distance to obstacles. **Conclusions.** The prediction-oriented evasion method improves the overall navigation efficiency in industrial mission models; further research involves field validation on real platforms and optimization of the predictor’s computational costs.

Keywords: unmanned aerial vehicle; industrial autonomous missions; trajectory prediction; obstacle avoidance; recurrent neural network; trajectory correction; energy efficiency.

Introduction

Unmanned aerial vehicles (UAVs) have become a standard tool in industrial autonomous missions, such as inspections of power lines and substations, monitoring of pipelines and tank farms, surveys of wind farms, monitoring of construction sites, quarries and industrial zones, as well as the rapid assessment of facility conditions following incidents. In these scenarios, the value of UAVs is determined not only by the quality of the data collected, but above all by their ability to fly predictably and safely.

This task is complicated by existing industrial environments, which are characterized by restricted flight corridors, complex obstacle geometries, local radio interference zones, and multipath effects, as well as a high level of dynamic activity from surrounding objects (equipment, vehicles, personnel). In practice, this translates into the need to simultaneously ensure high trajectory tracking accuracy, collision avoidance, controlled energy consumption, and system stability in the face of sensor errors.

In industrial missions, UAVs operate in environments where navigation measurements are often

noisy, and the environment’s configuration can change during flight. The most common practical consequence is instability in position and trajectory estimates (e.g., coordinate oscillations), leading to unwanted corrective maneuvers, increased route deviation, and energy waste. Under such conditions, reactive approaches that correct motion only “after the fact” of a dangerous proximity can trigger delayed or excessive evasion, compromising both safety and mission efficiency.

An additional practical requirement for industrial scenarios is reproducible and energy-efficient behavior. The control system must ensure the minimization of unnecessary maneuvers and the maintenance of an acceptable distance from obstacles while keeping the UAV on the specified trajectory. This is particularly important for missions where flight time is a limited resource, and any increase in energy consumption directly reduces operational duration and diminishes route coverage.

Literature Review

In recent years, a number of reviews have been published that systematize algorithms for route planning, obstacle detection, and avoidance for UAVs. Work [1]

presents a comprehensive classification of forecasting approaches and discusses the suitability of methods for remote sensing application scenarios. Review [2] focuses on obstacle detection and avoidance (ODA) and compares approaches based on obstacle type, sensor configuration, and mission environment. Additionally, in [3], recent comprehensive reviews highlight the trend toward integrating artificial intelligence with classical planners and the growing demands for computational efficiency and real-time robustness.

Summarizing these scientific works, it can be concluded that the issues of ensuring obstacle prediction and avoidance processes, as separate components of UAV mission execution algorithms, remain relevant. At the same time, for industrial autonomous missions, it is necessary to coordinate actions such as “state estimation”, “forecasting”, “decision-making”, and “maneuvering” in the presence of noise and sensor data degradation.

Classical trajectory tracking methods are widely represented by approaches based on vector fields. Reference [4] is one of the fundamental sources where a vector field is used to asymptotically reduce the tracking error to zero for small UAVs. Further development in this direction [5] aims to combine trajectory tracking and obstacle avoidance with minimal deviation from the route. The strength of vector and geometric methods lies in the simplicity and predictability of control. However, in industrial environments with noisy coordinates and changing obstacle configurations, reactive corrections increase the amplitude of maneuvers and, consequently, energy consumption. This is why there is a growing need for anticipatory control and motion prediction.

Model Predictive Control (MPC) is a common paradigm for generating safe trajectories while accounting for system constraints. Classical MPC formulations for obstacle avoidance [6] demonstrate the ability to formalize the problem as an optimization over the prediction horizon. In applied inspection scenarios [7], MPC is combined with Simultaneous Localization and Mapping (SLAM) and employs path smoothing to reduce flight time and memory requirements. This indirectly impacts the mission’s energy efficiency. Modern variants [8] also integrate safety barrier functions to handle dynamic obstacles. At the same time, MPC architectures depend on model quality and can be sensitive to sensor errors. Under unstable measurements, state errors translate into inadequate control decisions, and computational complexity may limit their use on low-power onboard systems.

To improve resilience to navigation signal degradation, approaches [9–11] based on Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) are being actively developed.

In particular, for the task of compensating for errors during GPS signal outages, hybrid models [10] have been proposed that combine prior noise suppression with CNN–LSTM forecasting. Another direction, discussed in [12, 13], involves the optimization and selection of LSTM component parameters in combination with adaptive variants of the Extended Kalman Filter to improve 3D positioning during GPS outages.

Additionally, the paper [14] investigates LSTM-based approaches for sensor fusion in orientation estimation tasks based on inertial measurements, which is relevant for UAVs under unstable positioning conditions. Overall, neural network filters demonstrate potential, but the key challenge remains reconciling data and physical constraints.

Predicting UAV trajectories over a time horizon is used as a tool for early conflict detection and proactive risk assessment. Paper [15] examines recurrent LSTM-based trajectory prediction in the context of detecting potential conflicts in low-altitude airspace. Concurrently, review and methodological works [16–18] have emerged, systematizing UAV trajectory prediction algorithms, including variants combining neural networks and filtering schemes. A practically significant limitation of this class of methods is the accumulation of errors over long horizons and dependence on the quality of input measurements, which brings the problem back to the need for robust filtering and adaptive selection of the prediction horizon.

Reinforcement learning is applied to obstacle avoidance tasks in both discrete and continuous action spaces. In [19], a DRL approach for obstacle avoidance is demonstrated, comparing different categories of algorithms in simulation environments. For large-scale 3D scenarios, approaches [20] with a “human-in-the-loop” mechanism are being developed, which improve training controllability and agent behavior consistency.

A common practical choice is Proximal Policy Optimization (PPO). Studies [21, 22] have been conducted where PPO is applied to autonomous UAV navigation in unknown environments.

Of particular interest are comparisons [23] of algorithms (e.g., PPO vs. Deep Q-Network) in obstacle avoidance tasks in 3D simulations.

For industrial inspection missions, the articles [24, 25] are noteworthy, as they are directly focused on infrastructure scenarios, in particular the LSTM-DDPG (Long Short-Term Memory – Deep Deterministic Policy Gradient) approach, validated in simulations and field tests.

At the same time, it should be noted that while the DRL approach provides high adaptability, it often requires a significant number of training episodes and demonstrates sensitivity to input data. This is precisely why the role of hybrid approaches is growing in engineering systems, where learning complements, rather than replaces, stabilizing control logic and safety control [26, 27].

A separate body of work addresses energy constraints and resource-aware mission planning. In particular, the paper [28] proposes an energy-efficient coverage route planning method, where energy limits are an explicit constraint during route generation. Concurrently, data-driven models [29] for predicting performance degradation and energy consumption are being developed, which are crucial for estimating flight time reserves. In the context of industrial missions, this reinforces the need to minimize “unnecessary” maneuvers that may be caused by the instability of navigation measurements and reactive obstacle avoidance.

As can be seen, most of the approaches considered focus on isolated components: route planning, obstacle avoidance, data filtering, or adaptive control. Against this backdrop, systemic gaps emerge, such as insufficient coordination within a unified loop for state estimation, trajectory prediction, and trajectory correction for noisy navigation measurements. Furthermore, there is insufficient attention to energy consumption as a metric for evasion quality. This is particularly true in scenarios where smooth behavior and maneuver reproducibility are required for industrial missions.

Objectives and Contribution

The objective of this study is to develop a method for predicting UAV trajectories and deviations for industrial autonomous missions in dynamic environments under conditions of noisy navigation measurements. The proposed approach is geared toward scenarios in which the navigation decision must be robust to instability in coordinate estimates, and the visual channel (camera) is not considered a critical prerequisite for decision-making during maneuvering.

Main research objectives

1. Develop a method for predicting UAV trajectories based on RNN/LSTM for proactive obstacle avoidance in a dynamic environment in the presence of sensor noise.
2. Develop the architecture of a UAV obstacle avoidance system.
3. Investigate methods for predicting future UAV coordinates and checking for collision risks.
4. Implement a trajectory correction algorithm that ensures safe avoidance while minimizing deviation from the route and controlling the energy cost of the maneuver.
5. Conduct a comparative study with common basic UAV motion control methods on standard trajectories using applied metrics and a generalized indicator of productivity.

Main Section

1. Construction of evasion trajectories based on path coordinates using RNN/LSTM for prediction

The proposed method implements a prediction-oriented evasion cycle of “prediction, risk assessment, trajectory correction,” which is repeated at each discrete time step t . To generate the forecast, a window of the length of the previous measurements is used $(x_{t-n}, y_{t-n}) \dots (x_t, y_t)$, as well as current driving parameters v_t and θ_t and information about the nearest obstacles. The prediction horizon is determined by the number of steps N_p . Here t is the discrete time index (simulation step number), n This is the length of the input window (the number of previous trajectory points fed into the model).

Step 1 (input data preparation). The sequence of the last n trajectory points and the motion parameters are collected $(v_t$ and $\theta_t)$ along with data on nearby obstacles.

Step 2 (coordinate prediction). The LSTM model generates the predicted coordinates of the UAV on the horizon N_p : $(\hat{x}_{t+1}, \hat{y}_{t+1}) \dots (\hat{x}_{t+N_p}, \hat{y}_{t+N_p})$.

Step 3 (collision risk check). The predicted trajectory is checked for intersections with the safety zones of obstacles within a radius of R . If an intersection is detected, potential points of conflict are identified.

Stage 4 (Trajectory Correction). If the risk of a collision is confirmed, a correction algorithm is activated, which generates an evasive maneuver that

minimizes deviation from the route and controls energy consumption.

The UAV performs a corrected maneuver; in the next step t , the cycle repeats with updated sensor data.

1.1. Development of an architecture for a UAV obstacle avoidance system

In complex navigation scenarios where rapid and adaptive evasion is required, traditional methods may prove insufficient. The use of RNNs opens up new possibilities for predicting UAV trajectories by analyzing past coordinates and motion dynamics.

Unlike conventional methods, which calculate the trajectory based on current data, LSTM (Long Short-Term Memory) can take historical information into account and use it to predict the future position of the UAV. Thanks to its ability to remember significant events over long time intervals, LSTM is an ideal tool for motion prediction tasks.

The LSTM model consists of a cascade of memory blocks that filter the input data, removing insignificant information and retaining only important patterns. This allows the model to identify patterns in the UAV's motion and predict its coordinates with high accuracy.

When avoiding obstacles, the neural network receives the following input:

- a sequence of previous coordinates $(x_{t-n}, y_{t-n}) \dots (x_t, y_t)$;
- velocity and direction of motion;
- information about nearby obstacles.

The network generates predicted coordinates as output (x_{t+1}, y_{t+1}) , which allows for determining the probability of a collision and adjusting the course before the UAV approaches an obstacle.

A block diagram of the proposed model is shown in Fig. 1.

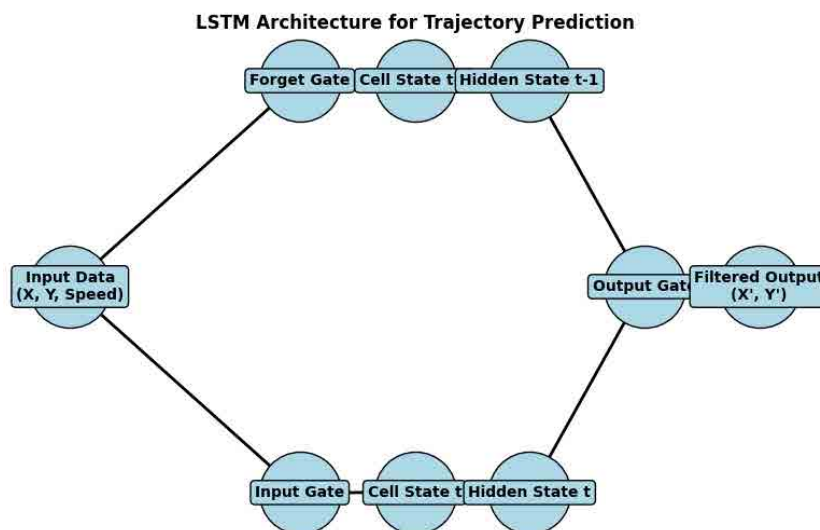


Fig. 1. Block diagram of an LSTM model for trajectory prediction

The main advantage of LSTM in forecasting tasks is its ability to account for nonlinear interactions in motion. This means that such a model can learn to adjust its trajectory even when obstacles appear unexpectedly or change their position.

In addition to directly predicting future coordinates, RNN/LSTM can be used to optimize evasion maneuvers by determining the safest and most energy-efficient route. For example, if a standard evasion algorithm suggests a sharp maneuver that could lead to energy loss, the neural network can adapt the maneuver to ensure an optimal balance between safety and flight efficiency.

Thanks to these advantages, the use of RNN/LSTM in UAV navigation systems allows for the simultaneous resolution of several important tasks:

- reducing delays in the evasion process – the network predicts hazards in advance, minimizing the risk of delayed reactions;
- adaptation to environmental changes – the system can respond to new threats without having to recalculate the route from scratch;
- motion optimization – smooth evasion is ensured, which minimizes deviation from the mission and saves energy.

Thus, the use of LSTM in obstacle avoidance not only enhances flight safety but also ensures efficient, flexible, and intelligent UAV trajectory control.

Specifying the model shown in Fig. 1, it can be noted that the architecture of the UAV evasion system is based on a deep neural network that analyzes the environment, predicts potential threats, and generates an optimal maneuvering trajectory. The foundation of the architecture is the sensory perception module, which receives information about the space in front of the UAV, taking into account data from GPS, inertial measurement units (IMUs), lidars, and optical cameras. The sensor module transmits input data to the preprocessing subsystem, which performs noise filtering and measurement delay compensation. For this task, the study employs an LSTM, which allows for the recognition of temporal dependencies and the elimination of measurement drift.

The next component of the architecture is the UAV trajectory prediction module, which performs multi-step prediction of future coordinates based on time sequences of navigation data. The module receives the last n coordinate $(x_{t-n}, y_{t-n}) \dots (x_t, y_t)$ measurements, as well as motion parameters $(v_t$ i $\theta_t)$ and information about nearby obstacles. The prediction module is implemented using a recurrent neural network with an LSTM architecture, which allows for the consideration of temporal dependencies and reduces the impact of sensor noise when estimating the future position of the UAV.

Next, the collision risk assessment module is executed: the predicted trajectory is compared with

a map of obstacles (static or those that change configuration over time), taking into account the safety zone radius R . If the predicted trajectory crosses safety zones, potential conflict points are identified and a motion correction is initiated.

The trajectory correction module generates a corrective maneuver to ensure safe avoidance of the obstacle with minimal deviation from the initial route and with control over the maneuver's energy expenditure. The final element is the feedback module, which updates the forecast and decisions at every step based on new sensor measurements, ensuring continuous real-time trajectory adaptation.

The final component of the architecture is the feedback module, which provides deviation correction based on continuous monitoring of flight parameters. The use of recurrent neural networks in this module reduces prediction errors through adaptive real-time updating of the deviation model. This enables the UAV to adjust its behavior in response to dynamic environmental changes, avoiding sudden obstacles or changes in the trajectories of other objects.

The overall architecture of the evasion system combines modern methods of deep learning, recurrent forecasting, and real-time motion optimization. The interaction between its modules ensures high evasion accuracy, rapid response to threats, and flight stability even under challenging conditions.

Fig. 2 shows a block diagram of the UAV obstacle avoidance system architecture.

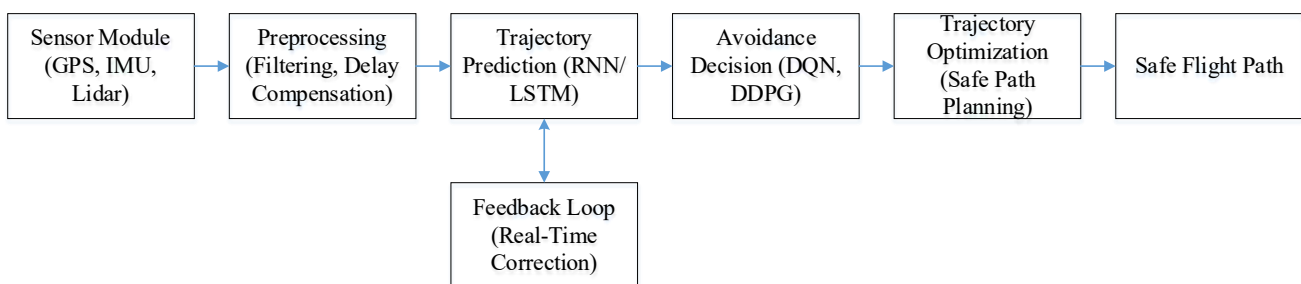


Fig. 2. Block diagram of the UAV obstacle avoidance system architecture

1.2. A method for predicting the future coordinates of a UAV

To ensure autonomous obstacle avoidance, a UAV must be able to predict its future position, taking into account its previous coordinates, speed, surrounding objects, and possible changes in flight conditions. The use of recurrent neural networks, specifically

LSTMs, allows for the analysis of temporal dependencies in the flight path and the determination of predicted coordinates.

The basis of the prediction is the formation of an input data sequence used to train the neural network. The following are passed as input parameters to the *LSTM*: historical UAV coordinates $(x_{t-n}, y_{t-n}), \dots, (x_t, y_t)$,

that reflect the device's previous position; speed and direction of movement v_t, θ_t , which enable the assessment of flight inertial characteristics; sensor data on surrounding objects that may affect course changes.

The task of the neural network is to construct a prediction function that estimates future coordinates (x_{t+1}, y_{t+1}) based on previous data:

$$(x_{t+1}, y_{t+1}) = f_{\theta}((x_{t-n}, y_{t-n}), \dots, (x_t, y_t), v_t, \theta_t)$$

where f_{θ} – This is a function parameterized by the neural network's weights, which are optimized during training.

An important consideration is selecting the optimal prediction horizon. If the prediction horizon is too short, evasive maneuvers may be too late, increasing the risk of collision. On the other hand, an overly long horizon can lead to increased prediction error. The optimal value depends on the UAV's speed, the flight environment, and dynamic changes in obstacles.

The LSTM network is trained to minimize the loss function:

$$L = \frac{1}{N} \sum_{i=1}^N ((x_{t+1}^{(i)} - \hat{x}_{t+1}^{(i)})^2 + (y_{t+1}^{(i)} - \hat{y}_{t+1}^{(i)})^2), \quad (1)$$

where L – the value of the loss function (loss), which is minimized during the training of an LSTM/RNN model; N – the number of training examples in the sample (or in the batch, if the computation is performed using a mini-batch); $x_{t+1}^{(i)}, y_{t+1}^{(i)}$ – the true (target) coordinates of the UAV in the next step, for example i ; $\hat{x}_{t+1}^{(i)}, \hat{y}_{t+1}^{(i)}$ – the coordinates specified by the model for the next step, for example i .

$$e(N_p) = \frac{1}{N} \sum_{i=1}^N ((x_{t+N_p}^{(i)} - \hat{x}_{t+N_p}^{(i)})^2 + (y_{t+N_p}^{(i)} - \hat{y}_{t+N_p}^{(i)})^2), \quad (2)$$

the larger N_p is, the greater the error $e(N_p)$ becomes, which can render the forecast unreliable.

It should be noted that a UAV maneuver requires time to change its trajectory. This, in turn, depends on the aircraft's dynamic characteristics. The minimum required forecast horizon is determined by the formula:

$$T_{maneuver} = \frac{\Delta\theta}{\omega}, \quad (3)$$

$$P_{collision}(T_p) = \int_S p(x, y) \cdot \delta(x - x_{t+T_p}, y - y_{t+T_p}) dx dy, \quad (4)$$

where $p(x, y)$ – the probability of obstacles appearing in a given area; $\delta(x - x_{t+T_p}, y - y_{t+T_p})$ – a UAV position marker that mathematically selects a point (x, y) in space.

The final optimization problem in the process of predicting the future coordinates of the UAV is defined as:

$$T_p^* = \arg(\alpha e(N_p) + \beta P_{collision}(T_p) - \gamma T_{maneuver}), \quad (5)$$

This helps reduce the root mean square error of the forecast.

By using a neural network, the evasion system is able to anticipate a potential collision in advance, rather than merely reacting to the current situation.

Selecting the prediction horizon is a critical task, since a horizon that is too short may not allow enough time for maneuvering, while a horizon that is too long may increase the prediction error and lead to instability in the evasion system.

Optimal forecasting horizon T_p is defined as a trade-off between forecast accuracy, system response time, and maneuverability. Formally, the choice of the forecasting horizon can be described as an optimization problem that minimizes the sum of the forecast error and the collision risk.

Let us formalize the task of determining the forecasting horizon and consider it as an optimization problem.

The forecasting horizon determines how far ahead the LSTM network must predict the future coordinates of the UAV. Let us denote it as T_p , and the sampling interval as Δt . In that case, the prediction horizon in seconds is defined as: $T_p = N_p \cdot \Delta t$ of predicted points in the time series.

When selecting the optimal N_p , the following must be taken into account:

Forecast error. The forecast error increases with N_p , since the LSTM model operates in a recurrent mode and accumulates errors

where $\Delta\theta$ – a necessary change in the orientation angle; ω – UAV turn speed.

The probability of a collision $P_{collision}$ depends on the predicted trajectory of the UAV and surrounding objects. The safe prediction horizon must be sufficiently large to avoid obstacles:

where α, β, γ – weighting factors that determine the balance between prediction accuracy, collision risk, and maneuverability.

Depending on flight conditions, the prediction horizon can adapt dynamically. If the trajectory is stable (no sudden changes in direction, low probability of collision), the prediction horizon is extended to allow for more forward-looking avoidance.

$$T_p = \min(T_p + \Delta T, T_{max}).$$

If the predicted trajectory indicates a high risk of collision, the prediction horizon is shortened so that the evasion model can adapt more quickly to changes.

$$T_p = \max(T_p - \Delta T, T_{min})$$

where T_{max} , T_{min} – boundary values of the forecast horizon.

1.3. Method for checking the possibility of collision

As shown in Fig. 2, after predicting the UAV's coordinates, it is necessary to assess the possibility of collisions with obstacles in the surrounding space. To do this, a map of the surrounding environment is

$$P_{collision} = \int_S p(x, y) \cdot \delta(x - x_{t+1}, y - y_{t+1}) dx dy, \quad (6)$$

where $p(x, y)$ – probability distribution of the presence of an obstacle; $\delta(x, y)$ – a function that determines the UAV's position.

To ensure safety, a safety zone is defined around the device. This zone defines the area into which no obstacles should enter and can be represented as an ellipsoidal or spherical zone:

constructed based on data from LiDAR, cameras, radars, and other sensors. The map can be represented discretely as a three-dimensional grid or as a continuous function describing the location of objects.

Based on this data, the next step involves assessing the collision risk. This assessment is based on comparing the predicted coordinates of the UAV with the spatial locations of static and dynamic objects. If the predicted trajectory (x_{t+1}, y_{t+1}) intersects an area occupied by an obstacle, the system determines the collision probability $P_{collision}$:

$$(x - x_i)^2 + (y - y_i)^2 \leq R^2, \quad (7)$$

where R – safety zone radius.

Fig. 3 (a, b) shows illustrations of the results of implementing the presented method for checking for potential collisions.

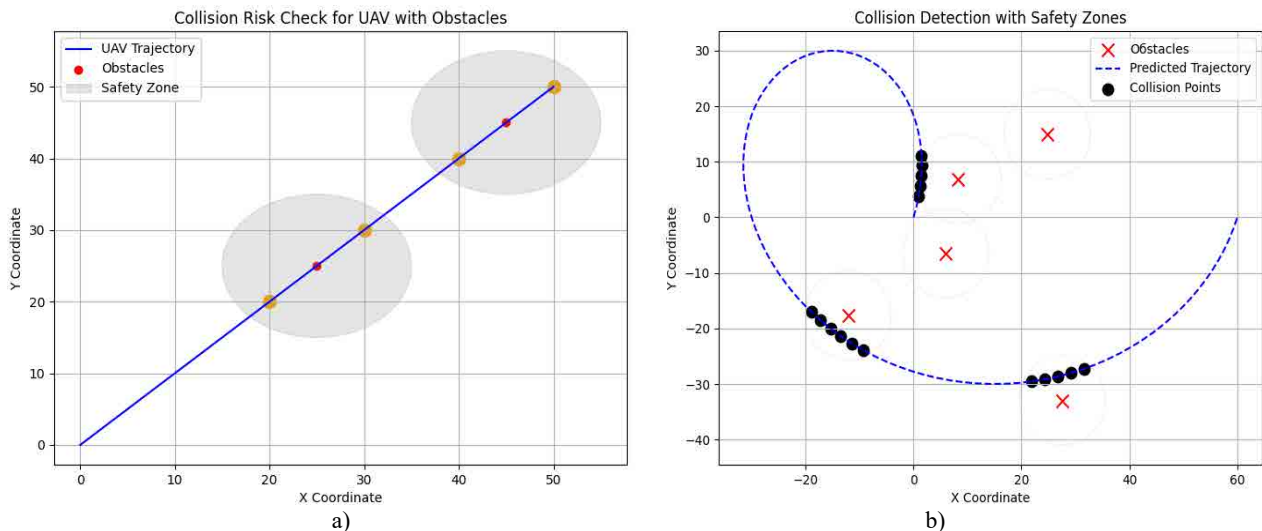


Fig. 3. Results of the collision detection method simulation

As can be seen in Fig. 3, the model generates random obstacles and displays them on the map. Obstacles can be static or dynamic objects. A safety zone is depicted around each obstacle, which likely defines the area that the UAV must not enter.

Next, the UAV's planned trajectory is generated, and the intersections of the future route with obstacles are checked using the safety zone. If an obstacle is detected, potential collision points with the obstacle are identified. If there is no risk of collision, the UAV's flight path remains unchanged. If the path passes

too close to an obstacle, the model monitors the possibility of a collision and marks potential collision points on the diagram.

In Fig. 3, you can see the points where the UAV's path intersects the safety zones of the obstacles. These points are marked as "collision risk." This indicates that the program has successfully detected potential collisions and marked them on the diagram.

If the predicted coordinates go beyond the safety zone, the evasion mechanism is activated.

1.4. Trajectory Correction Algorithm

When the system detects a collision threat, a new motion vector must be generated to ensure a safe evasion. To do this, the direction of motion is adjusted to

$$Q^* = \underset{Q'(\cdot)}{\operatorname{argmin}} \int_t^{t+\tau} (\|r(Q'(s)) - r_{goal}\|_2 + \lambda \|r(Q'(s)) - r_0(s)\|_2) ds, \quad (8)$$

where Q' – new direction of movement; $r(Q'(s))$ – current trajectory; r_{goal} – initial route; $r_0(s)$ – reference (initial) trajectory at the moment s ; λ – a weighting factor that balances the approach to the target with the minimum deviation from the reference route.

Course correction without losing the main heading is achieved using a parameter λ that determines the balance between collision avoidance and maintaining the optimal path.

The final choice of evasion maneuver depends on current conditions:

If the obstacle is stationary, the system can perform a smooth maneuver to fly around it. If the obstacle is moving, dynamic trajectory prediction is used to select the optimal moment for evasion. If there is no safe path, the system activates emergency braking and hovers in place.

Thanks to this approach, the evasion algorithm ensures continuous real-time trajectory adaptation and minimizes deviations from the mission.

The conducted modeling and research of the developed model allowed for the formation of a set of resulting data. Among them, Fig. 4 stands out, which presents the results of modeling the UAV's obstacle avoidance process using trajectory prediction and energy consumption reduction.

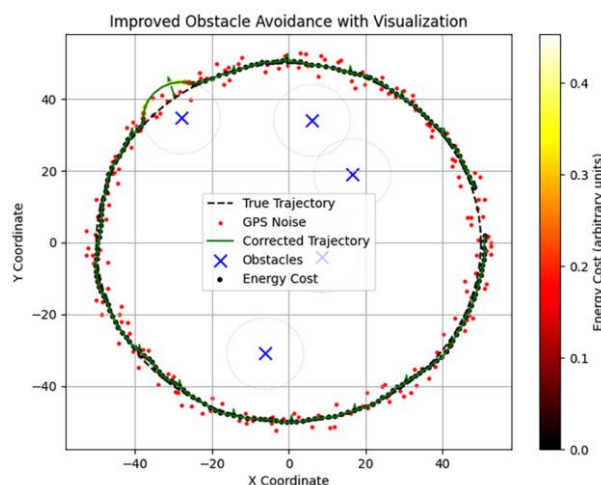


Fig. 4. Results of modeling the UAV obstacle avoidance process using trajectory prediction

avoid the obstacle without deviating significantly from the main route.

The trajectory correction task is formalized as an optimization problem:

The figure illustrates key navigation elements, including the actual trajectory, noisy GPS coordinates, the corrected trajectory after applying a neural network, as well as the location of obstacles and the energy costs of avoidance.

In Fig. 4, the actual trajectory line (indicated by a black dashed line) represents the ideal UAV flight path without accounting for noise, obstacles, or other external influences. It takes the form of a uniform circle, indicating stable flight along predefined coordinates.

GPS measurements (marked by red dots) exhibit random deviations from the true trajectory within ± 3 meters, which corresponds to typical GPS errors under real-world conditions. They demonstrate significant fluctuations and deviations, which can complicate safe flight.

The corrected trajectory line (green line) shows the results of the proposed correction algorithm. It is obtained by predicting future coordinates using an LSTM recurrent neural network. Thanks to this, the model smooths out GPS errors and corrects the UAV's movement in response to detected threats.

The blue "x" markers indicate static obstacles located within the operational space. They are used to verify the effectiveness of evasion and to simulate collision scenarios. Each obstacle has a safety zone with a radius of 10 meters ($\text{safe_distance} = 10$) around it. If the UAV enters this zone, the evasion algorithm is activated.

Additionally, Fig. 4 illustrates the evasion directions (green arrows). They show the change in the UAV's trajectory in response to obstacle detection. The correction occurs in the direction that minimizes the risk of collision, while simultaneously accounting for energy expenditure.

Additionally, the colored dots on the corrected trajectory in the figure represent the energy expenditure required for evasion. The brighter the color (from yellow to dark red), the higher the energy expenditure. As shown in the figure, maximum energy expenditure occurs in areas of sharp maneuvering near obstacles.

As shown in Fig. 4, the use of a recurrent neural network allowed for significant smoothing of GPS noise distortions and the attainment of a more stable trajectory.

The evasion algorithm adjusts the route in risk zones, ensuring safe navigation around obstacles.

Data on energy consumption in the proposed example are presented in Table 1.

Within the framework of simulation modeling, energy consumption is evaluated not as battery consumption, but as a conditional indicator of “maneuvering effort,” which directly depends on the intensity of changes in direction of motion. At each step t , the course is calculated based on the sequence of coordinates:

$$\theta_t = \text{atan2}(y_t - y_{t-1}, x_t - x_{t-1}).$$

Table 1. Summary data on energy consumption in the proposed example

Total energy consumption (units)	Mean energy consumption (units)	Максимальні витрати енергії (у.о.)
2.3569	0.2143	0.3486

These results demonstrate the effectiveness of the proposed trajectory correction method using LSTM and adaptive obstacle avoidance.

1.5. Practical Application Example and Integration Option into an Autopilot

Let's consider a typical industrial scenario. Inspection of linear infrastructure (power lines/pipelines) or monitoring of a construction site, where the UAV flies along a specified trajectory and must avoid local obstacles (pylons, structures, equipment, temporary objects). Under such conditions, the evasion system receives:

- navigation estimates of position and motion parameters (coordinates, speed, heading) from GPS/IMU;
- information about nearby obstacles in the form of coordinates or distances to the nearest objects, obtained from onboard avoidance sensors or an external surveillance module.

The proposed method is used as a guidance-level add-on to the standard flight stabilization loop. At each step, a window of the last n navigation points and motion parameters is formed, an LSTM forecast of coordinates at horizon N_p is performed, after which the intersection of the predicted trajectory with safety zones of radius R is checked. If a risk is detected, a trajectory correction is generated, while the autopilot ensures stabilization and command execution.

The practical implementation within the scope of this article is presented as a software prototype in the Python environment (simulation modeling), which demonstrates the full loop of prediction, risk checking, and correction. Field testing is not within the scope of this work and is defined as the next validation stage,

In that case, the instantaneous maneuvering “energy loss” is defined as the absolute change in heading:

$$e_t = |\text{wrap}(\theta_t - \theta_{t-1})|,$$

where $\text{wrap}(\cdot)$ reduces the angle difference to an interval $(-\pi; \pi]$. Total costs along a trajectory segment $E_{total} = \sum e_t$, are mean: $E_{mean} = (1/N) \sum e_t$, maximum: $E_{max} = \max e_t$. Values e_t are used for color-coding points on the corrected trajectory (Fig. 4), while Table 1 contains aggregated metrics derived directly from the simulation data.

taking into account hardware limitations, sensor update rates, and real-time requirements.

We will conduct comparative studies of the developed UAV trajectory and deviation prediction for industrial autonomous missions in a dynamic environment and evaluate its effectiveness.

2. Comparative studies of the proposed method for predicting UAV trajectory and deviation

To evaluate the effectiveness of the developed method and determine its practical value, this article conducts comparative studies of the proposed method for predicting UAV trajectories and deviations against existing methods [30–32]. The following research strategy is proposed for this purpose.

Comparisons were made between methods similar in purpose. These methods are:

- Pure Pursuit;
- Line of Sight;
- Vector Field;
- Nonlinear Stabilization;
- Intelligent control with prediction based on RNN/LSTM.

Performance evaluation was carried out using criteria that highlight the practical advantages (and disadvantages) of our development. The main parameters used for comparison are:

- average deviation from the trajectory (meters);
- adaptation time after deviation (seconds);
- number of obstacle avoidance maneuvers;
- minimum distance to an obstacle (meters);
- average energy consumption (conditional units);

– trajectory tracking success rate (%) (ratio of the distance traveled without exceeding the permissible deviation to the total distance);

– overall productivity coefficient.

To obtain the necessary data, simulation modeling was performed using Python and the required libraries *numpy*, *scipy*, and *pandas*.

The developed simulation model allowed for testing algorithms on standard trajectories (straight line, circle, polyline) and collecting the corresponding statistical data. The results of the simulation modeling are presented in Fig. 5–11 and in Table.

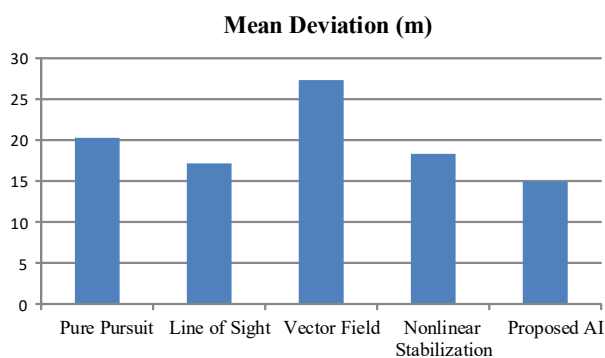


Fig. 5. Results of a comparative study of the “Proposed AI” method based on the “average deviation from the trajectory” metric

First, trajectory tracking accuracy was examined (see Fig. 5). The average deviation indicates the extent to which each method enables the UAV to maintain route stability. In this regard, the Proposed AI method proved to be the most effective, demonstrating the smallest deviation (14.95 m). This indicates the algorithm’s ability to quickly correct the trajectory and maintain flight stability. In contrast, the Vector Field method had the largest deviation (27.32 m), indicating the instability of this approach.

Another important factor is the adaptation speed after deviation, which assesses how quickly the UAV returns to the optimal trajectory (see Fig. 6). The Line of Sight and Nonlinear Stabilization methods demonstrated relatively short adaptation times of 26–30 seconds. The Proposed AI method, although it took 39 seconds to stabilize, compensated for this with flight accuracy and overall efficiency. The Vector Field method had the longest delay (47 seconds), which negatively affects its application in complex conditions.

Significant attention was paid to safety issues, namely obstacle avoidance and minimum distance to objects (see Fig. 7).

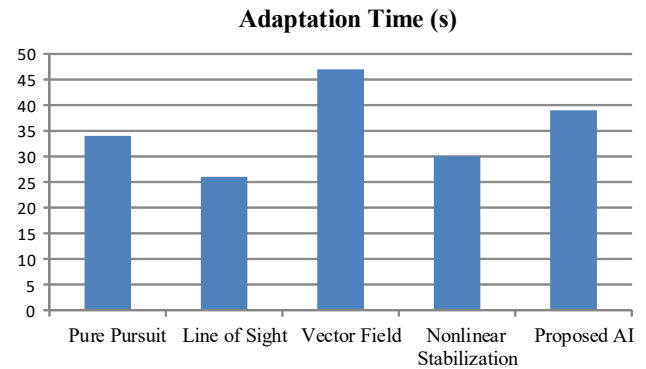


Fig. 6. Results of a comparative study of the “Proposed AI” method based on the “adaptation time after deviation (seconds)” metric

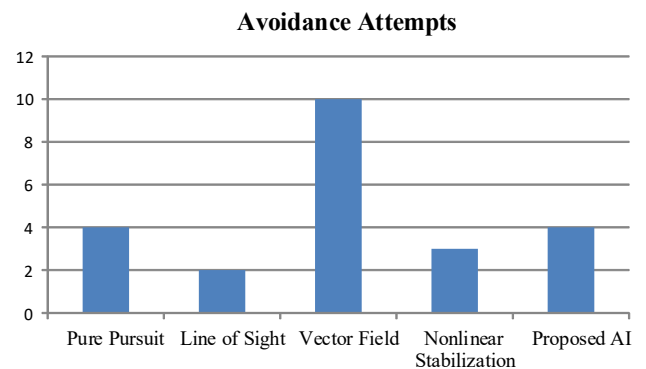


Fig. 7. Results of a comparative study of the “Proposed AI” method based on the “number of obstacle avoidance maneuvers” metric

The number of obstacle avoidance attempts indicates how frequently the algorithm detects a hazard and adjusts its course. In this regard, Vector Field had the highest number of attempts (10), indicating trajectory instability and potential erroneous maneuvers. In contrast, Proposed AI had only 4 attempts, demonstrating a balance between evasion and the minimization of unnecessary corrections.

The minimum distance to obstacles is another important criterion (see Fig. 8). In this regard, the Nonlinear Stabilization method (7.66 m) provided the highest level of safety, while Vector Field (5.81 m) demonstrated a risk of collisions. The Proposed AI method (6.12 m) fell within the optimal balance between evasion efficiency and stability maintenance.

Energy consumption also plays a significant role in selecting a UAV control method (see Fig. 9).

The maneuver energy consumption index E_{total} was calculated identically for all methods using the formulas given above, on the same standard trajectories (straight line, circle, polygon), and under identical

simulation parameters. Thus, the comparison in Fig. 9 reflects the relative “cost of maneuvering” of different algorithms under common experimental conditions.

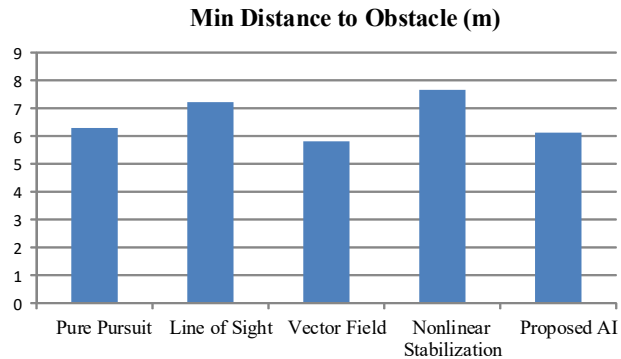


Fig. 8. Results of a comparative study of the “Proposed AI” method based on the “minimum distance to obstacle (meters)” metric

The Proposed AI (72 conventional units) proved to be the most economical, allowing for an increase in autonomous flight duration. By comparison, Vector Field (135 units) consumed the most energy, indicating its inefficiency for use in long-duration missions.

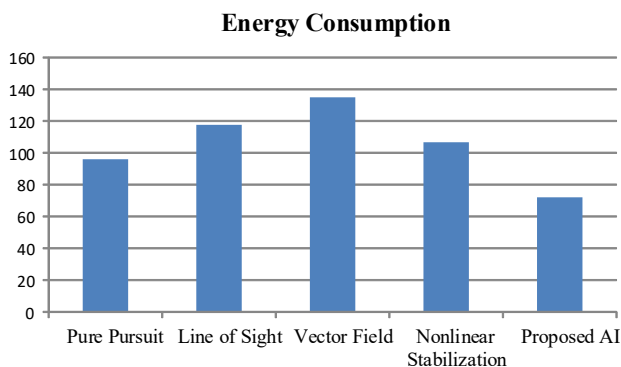


Fig. 9. Results of a comparative study of the “Proposed AI” method based on the “total maneuvering energy consumption” metric

An important factor is trajectory tracking performance, which reflects the percentage of the distance traveled without exceeding the permissible deviation (Fig. 10). In this regard, Proposed AI (86.08%) again demonstrated the best result among all methods. This indicates that the algorithm most effectively keeps the UAV within permissible motion parameters, which is a key factor for application in autonomous missions. The Nonlinear Stabilization (82.84%) and Line of Sight (84.31%) methods also performed well, while Vector Field (73.66%) lagged significantly, indicating trajectory instability and a higher probability of errors.

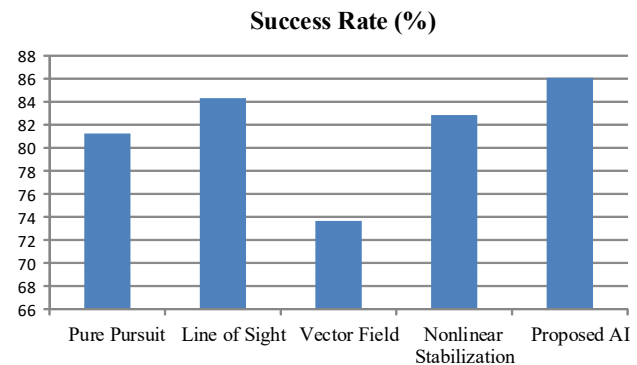


Fig. 10. Results of a comparative study of the “Proposed AI” method based on the “trajectory tracking success” metric

The final criterion of effectiveness is the overall productivity index (OPE), which combines all the aforementioned metrics (see Fig. 11). In this regard, Proposed AI achieved the best value (0.494), which is the highest among all the methods considered. This is 10% better than “Pure Pursuit,” 5% better than “Line of Sight,” 18% better than “Vector Field,” and 8% better than “Nonlinear Stabilization.” This demonstrates the comprehensive effectiveness of the method, which combines accuracy, adaptation speed, energy efficiency, and safety.

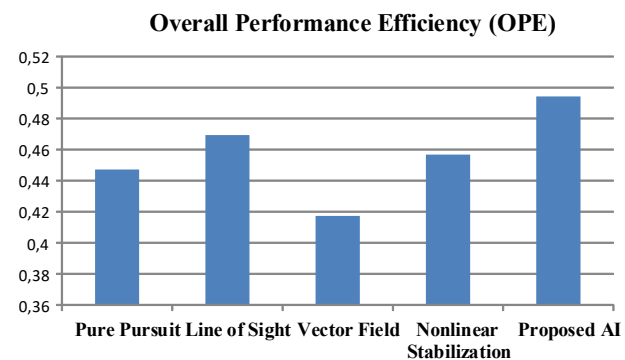


Fig. 11. Results of a comparative study of the “Proposed AI” method based on the “trajectory tracking success” metric

Thus, within the scope of the conducted simulation-based comparative study on standard trajectories and using selected evaluation metrics, the proposed method (Proposed AI) demonstrates the best overall performance (in particular, based on the generalized OPE metric) among the approaches considered, combining tracking accuracy, energy efficiency, and an acceptable level of evasion safety. The obtained results confirm the feasibility of using a prediction-oriented evasion scheme; field validation on real UAV platforms is considered a direction for further research.

Conclusions

In this article, within the framework of simulation modeling, the current scientific problem of developing a method for predicting UAV trajectories and evasion for industrial autonomous missions in a dynamic environment under conditions of noisy navigation measurements is solved. The method uses time series of coordinates and motion parameters for multi-step prediction of the UAV's position, which makes it possible to detect potentially dangerous approaches **before** entering the risk zone and avoid delayed reactions characteristic of purely reactive approaches.

An architecture for a UAV obstacle avoidance system has been developed, comprising modules for sensor data acquisition and preprocessing, prediction of future coordinates based on LSTM, collision risk assessment through formalization of the safety zone, and trajectory correction. This structure provides a coordinated “prediction, risk assessment, motion correction” loop.

A method for predicting future coordinates and checking for collision risks has been investigated. Conflict checking is performed by analyzing the intersection of the predicted trajectory with obstacle safety zones; if a threat is detected, potential risk points are identified, which initiates the evasion procedure. Additionally, the selection of the prediction horizon is considered as a key parameter that determines the balance between lead time and the accumulation of prediction error, with the possibility of adapting it depending on traffic conditions.

A trajectory correction algorithm has been implemented that ensures safe evasion while minimizing deviation from the route and controlling the energy costs of the maneuver. A simulation example demonstrates how the corrected trajectory is formed, and evasion maneuvers are performed in risk areas with the energy costs displayed on the trajectory.

A comparative study of the proposed method was conducted against common basic UAV motion control methods on standard trajectories using performance metrics and a generalized coefficient of productivity. According to the comparison results, the proposed method demonstrated the smallest average deviation from

the trajectory **14.95 m**, the lowest average energy consumption **72 u.o.**, the highest tracking success rate **86.08%**, as well as the highest value of the integral indicator **OPE = 0.494** among the approaches considered. At the same time, it was shown that trade-offs are possible for individual metrics: in particular, the minimum distance to the obstacle for the proposed method is **6.12 m**, whereas the maximum minimum distance in the experiment was achieved by the Nonlinear Stabilization method (**7.66 m**), which highlights the difference in priorities between methods and the importance of a comprehensive evaluation based on multiple criteria.

Promising directions for further research include extending the scenarios to cases with more pronounced obstacle dynamics, transitioning to a full three-dimensional framework accounting for altitude and altitude gain/loss constraints, and validation using experimental data or field tests. In particular, we plan to conduct field validation on a DJI Mavic-class UAV in a typical industrial scenario, measuring decision-making delays in real time and safety metrics. Separately, we will evaluate the computational budget of the LSTM module in inference mode and apply model optimization for on-board or companion computing.

Conflict of interest

The authors declare that they have no conflicts of interest, including financial, personal, authorial, or any other nature, that could influence the research or the results published in this article.

Funding

The study was conducted without financial support.

Data availability

Data will be provided upon reasonable request.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies to write this paper.

References

1. Debnath, D., Vanegas, F., Sandino, J., Hawary, A. F., Gonzalez, F. (2024), "A review of UAV path-planning algorithms and obstacle avoidance methods for remote sensing applications", *Remote Sensing*, Vol. 16 (21), 4019 p. DOI: <https://doi.org/10.3390/rs16214019>
2. Merei, A., Mcheick, H., Ghaddar, A., Rebaine, D. (2025), "A Survey on Obstacle Detection and Avoidance Methods for UAVs". *Drones*, 9(3), 203 p. <https://doi.org/10.3390/drones9030203>
3. Sangeeth V., Muralimohan G. (2025), "A comprehensive review of path planning algorithms for autonomous navigation", *Results in Engineering*, Vol. 28, 2025, 107750 p. DOI: <https://doi.org/10.1016/j.rineng.2025.107750>
4. Braquet, M., Bakolas, E. (2022), "Vector Field-based Collision Avoidance for Moving Obstacles with Time-Varying Elliptical Shape", *IFAC-PapersOnLine*, Vol. 55, Issue 37, 2022, pp. 587-592, DOI: <https://doi.org/10.1016/j.ifacol.2022.11.246>
5. Wilhelm, J., Clem, G (2019), "Vector Field UAV Guidance for Path Following and Obstacle Avoidance with Minimal Deviation". *Journal of Guidance, Control, and Dynamics*. pp. 1-9. DOI: <https://doi.org/10.2514/1.G004053>
6. Olcay, E., Meeß, H., Elger, G. (2024), "Dynamic Obstacle Avoidance for UAVs using MPC and GP-Based Motion Forecast", *European Control Conference (ECC)*, Stockholm, Sweden, pp. 1024–1031. DOI: <https://doi.org/10.23919/ECC64448.2024.10591083>
7. Tripicchio, P., Unetti, M., D'Avella, S., Avizzano, C.A. (2023), "Smooth Coverage Path Planning for UAVs with Model Predictive Control Trajectory Tracking". *Electronics*, 12(10), 2310 p. DOI: <https://doi.org/10.3390/electronics12102310>
8. Yoo, H.; Choi, S. (2025), "Improved Model Predictive Control for Dynamical Obstacle Avoidance". *Mathematics*, 13(22), 3624 p. DOI: <https://doi.org/10.3390/math13223624>
9. Zhuo, Z., Datong, L., Liansheng, L. (2022), "A performance compensation method for GPS/INS integrated navigation system based on CNN–LSTM during GPS outages", *Measurement*, Vol. 188, 110516 p. DOI: <https://doi.org/10.1016/j.measurement.2021.110516>
10. Semenov, S., Jian, Y., Jiang, H., Chernykh, O., Binkovska, A. (2025), "Mathematical model of intelligent uav flight path planning", *Advanced Information Systems*, 9(1), pp. 49-61. DOI: <https://doi.org/10.20998/2522-9052.2025.1.06>
11. Yena, M., Pohudina, O. (2025), "Integrated simulation model of swarm control and adaptive routing of UAVS in a changing air environment", *Innovative technologies and scientific solutions for industries*, 4(34), pp. 32-43. DOI: <https://doi.org/10.30837/2522-9818.2025.4.032>
12. Li, X., Wang, X., Pei, C. (2025), "Handling method for GPS outages based on PSO-LSTM and fading adaptive Kalman filtering", *Scientific Reports*, Vol.15, 11817 p. DOI: <https://doi.org/10.1038/s41598-025-95716-1>
13. Semenov, S., Wasiuta, O., Jammine, A., Golec, J., Krupska-Klimczak, M., Tarasenko, Y., Voronets, V., Breslavets, V., Lvov, S., Moskalenko, A. (2025), "Development of an Intelligent Method for Target Tracking in Radar Systems at the Initial Stage of Operation Under Intentional Jamming Conditions". *Applied Sciences*, 15(3), 7072 p. DOI: <https://doi.org/10.3390/app15137072>
14. Narkhede, P., Walambe, R., Poddar, S., Kotecha, K. (2021), "Incremental learning of LSTM framework for sensor fusion in attitude estimation". *PeerJ Computer Science*, DOI: <https://doi.org/10.7717/peerj-cs.662>
15. Zhang, Y., Jia, Z., Dong, C., Liu, Y., Zhang, L., Wu, Q. (2022), "Recurrent LSTM-based UAV Trajectory Prediction with ADS-B Information", *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, Rio de Janeiro, Brazil, pp. 1–6, DOI: 10.1109/GLOBECOM48099.2022.10000919
16. Cao, Y., Zhang, J., Shi, G., Yang, Q., Zhang, C. (2025), "Research on Trajectory Prediction Algorithm Based on Unmanned Aerial Vehicles Behavioral Intentions". *Drones* 2025, 9, 640 p. DOI: <https://doi.org/10.3390/drones9090640>
17. Semenov, S., Krupska-Klimczak, M., Mazurek, P., Zhang, M., Chernikh, O. (2025), "Improving Unmanned Aerial Vehicle Security as a Factor in Sustainable Development of Smart City Infrastructure: Automatic Dependent Surveillance–Broadcast (ADS-B) Data Protection". *Sustainability*, 17(4), 1553 p. DOI: <https://doi.org/10.3390/su17041553>
18. Kosarevskiy, B., Tetskiy, A. (2025), "Modern approaches to deploying the infrastructure of mobile intelligent systems". *Innovative technologies and scientific solutions for industries*, 2(32), pp. 33-48. DOI: <https://doi.org/10.30837/2522-9818.2025.2.033>
19. Shin, S., Kang, Y., Kim, Y. (2019), "Obstacle Avoidance Drone by Deep Reinforcement Learning and Its Racing with Human Pilot". *Applied Science*, 9(24), 5571 p. DOI: <https://doi.org/10.3390/app9245571>
20. Xuyang, L., Jianwu, F., Kai, D., Kuizhi, M., Jianru, X. (2023), "UAV Obstacle Avoidance by Human-in-the-Loop Reinforcement in Arbitrary 3D Environment", *42nd Chinese Control Conference (CCC)*, Tianjin, China, pp. 3589–3595. DOI: 10.23919/CCC58697.2023.10240962
21. Chikhaoui, K., Ghazzai, H., Massoud, Y. (2022), "PPObased Reinforcement Learning for UAV Navigation in Urban Environments". *2022 IEEE 65th International Midwest Symposium on Circuits and Systems (MWSCAS)*. DOI: <https://doi.org/10.1109/mwscas54063.2022.9859287>

22. Semenov, S., Zhang, M., Yenhalychev, S., Smidovych, L. (2023), "Generalized model of the ads-b unmanned aerial vehicle data transmission process in a steganographic system". *Innovative technologies and scientific solutions for industries*, 4 (22), pp. 14–19. DOI: <https://doi.org/10.30837/ITSSI.2022.22.014>
23. Júlio, C., Bruno, E., Fábio, L., Julio, C., Paulo, F. (2025), "Comparative Analysis of PPO and DQN for UAV Obstacle Avoidance in Simulated Environments", *IFAC-PapersOnLine*, Vol. 59, Issue 10, 2025, pp. 1540–1545. DOI: <https://doi.org/10.1016/j.ifacol.2025.09.259>
24. Dang, X., Eom, J., Vu, B., Shin, O. (2025), "A DDPG-LSTM Framework for Optimizing UAV-Enabled Integrated Sensing and Communication". *Drones*, 9(8), 548 p. DOI: <https://doi.org/10.3390/drones9080548>
25. Sun, S. (2025), "LSTM-DDPG-Based Dynamic Obstacle Avoidance for UAVs in Power Distribution Networks Using Velocity Obstacle Modeling", *Informatica*, 49(35). DOI: <https://doi.org/10.31449/inf.v49i35.12192>
26. Kuchuk, H., Kalinin, Y., Dotsenko, N., Chumachenko, I., Pakhomov, Y. (2024), "Decomposition of integrated high-density IoT data flow", *Advanced Information Systems*, Vol. 8, No. 3, pp. 77-84, DOI: <https://doi.org/10.20998/2522-9052.2024.3.09>
27. Semenov, S., Mozhaiev, O., Kuchuk, N., Mozhaiev, M., Tiulieniev, S., Gnusov, Yu., Yevstrat, D., Chyrva, Y., Kuchuk, H. (2022), "Devising a procedure for defining the general criteria of abnormal behavior of a computer system based on the improved criterion of uniformity of input data samples", *Eastern-European Journal of Enterprise Technologies*, Vol. 6(4-120), pp. 40-49, DOI: <https://doi.org/10.15587/1729-4061.2022.269128>
28. Wang, L., Zhuang, X., Zhang, W., Cheng, J., Zhang, T. (2024), "Coverage Path Planning for UAVs: An Energy-Efficient Method in Convex and Non-Convex Mixed Regions". *Drones*, 8(12), 776 p. DOI: <https://doi.org/10.3390/drones8120776>
29. Al-Haddad, LA, Łukaszewicz, A, Majdi, HS. (2025), "Energy consumption and efficiency degradation predictive analysis in unmanned aerial vehicle batteries using deep neural networks". *Advances in Science and Technology Research Journal*, 19(5), pp. 21-30. DOI: 10.12913/22998624/201346
30. Semenov, S., Krupska-Klimczak, M., Wasuta, O., Krzaczek, B., Mieczkowski, P., Głowacki, L., Yu, J., He, J., Chernykh, O. (2025), Intelligent Assurance of Resilient UAV Navigation Under Visual Data Deficiency for Sustainable Development of Smart Regions. *Sustainability*, 17(13), 6030 p. DOI: <https://doi.org/10.3390/su17136030>
31. Semenov, S., Voloshyn, D., Lymarenko, V., Semenova, A., Davydov, V. (2019), "Method of UAVs Quasi-Autonomous Positioning in the External Cyber Attacks Conditions," *2019 10th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Leeds, UK, pp. 149-153, DOI: 10.1109/DESSERT.2019.8770024
32. Yena, M. (2024), "Uav urban mobility control: swarm intelligence and collision avoidance". *Innovative technologies and scientific solutions for industries*, 4(30), pp. 59-66. DOI: <https://doi.org/10.30837/2522-9818.2024.4.059>

Received (Надійшла) 08.01.2026

Accepted for publication (Прийнята до друку) 03.03.2026

Publication date (Дата публікації) 30.03.2026

Відомості про авторів / About the Authors

Jian Yu – магістр, Zhongke Shuguang, старший інженер з розробки тестів; Тяньцзінь, Китай;

Yu Jian – Master, Zhongke Shuguang, Senior Test Development Engineer; Tianjin, China;

e-mail: ianyu220272@gmail.com

ORCID ID: <https://orcid.org/0009-0007-8990-8347>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=59681719600&origin=recordpage>

Сітнікова Оксана Олександрівна – кандидат технічних наук, Приватна установа "Університет науки, підприємництва та технологій"; Київ, Україна;

Oksana Sitnikova – PhD, Private Institution "University of Science, Entrepreneurship and Technology"; Kyiv, Ukraine;

e-mail: oasitnikova11@gmail.com

ORCID ID: <https://orcid.org/0000-0002-2417-8220>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=57208803509>

МЕТОД ПРОГНОЗУВАННЯ ТРАЄКТОРІЙ ТА УХИЛЕННЯ БПЛА ДЛЯ ПРОМИСЛОВИХ АВТОНОМНИХ МІСІЙ У ДИНАМІЧНОМУ СЕРЕДОВИЩІ

Розглянуто задачу прогнозування траєкторії та випереджувального ухилення безпілотного літального апарата у промислових автономних місіях у динамічному середовищі за наявності шуму навігаційних вимірювань і обмежень енергоресурсу, що зумовлює ризики запізнілого або надмірного маневрування та зростання відхилення від маршруту. **Мета.** Розробити та верифікувати на імітаційній моделі метод, який забезпечує прогнозно-орієнтоване ухилення від перешкод із контролем відхилення від опорної траєкторії та маневрових енерговитрат. **Завдання.** Сформувати архітектуру системи ухилення; розробити прогнозатор майбутніх координат на основі рекурентної нейронної мережі з довготривалою короткочасною пам'яттю; визначити спосіб перевірки ризику зіткнення із використанням зони безпеки; реалізувати алгоритм корекції траєкторії з урахуванням компромісу «безпека–відхилення–енерговитрати»; виконати порівняльне оцінювання з базовими методами. **Методи.** Прогноз координат будується за часовими послідовностями координат і параметрів руху; ризик зіткнення оцінюється шляхом аналізу перетину прогнозованої траєкторії із зонами безпеки перешкод; корекція траєкторії формалізується як оптимізаційний вибір маневру, що мінімізує сумарну похибку слідування та штраф за зближення. Ефективність перевірено у середовищі Python на стандартних траєкторіях (пряма, коло, ламана) з порівнянням із методами чистого переслідування, лінії візування, векторних полів і нелінійної стабілізації. **Результати.** Запропонований підхід забезпечив найменше середнє відхилення від траєкторії 14,95 м, найнижчі маневрові енерговитрати 72 умовні одиниці, найвищу успішність слідування 86,08 % та найбільший узагальнений коефіцієнт продуктивності 0,494 серед розглянутих алгоритмів; зафіксовано компроміс за мінімальною дистанцією до перешкод. **Висновки.** Метод прогнозно-орієнтованого ухилення підвищує інтегральну ефективність навігації у моделях промислових місій; подальші дослідження передбачають натурну валідацію на реальних платформах та оптимізацію обчислювальних витрат прогнозатора.

Ключові слова: безпілотний літальний апарат; промислові автономні місії; прогнозування траєкторій; уникнення перешкод; рекурентна нейронна мережа; корекція траєкторії; енергоефективність.

Бібліографічні описи / Bibliographic descriptions

Jian Yu, Sitnikova O.O. Метод прогнозування траєкторій та ухилення бпла для промислових автономних місій у динамічному середовищі. *Сучасний стан наукових досліджень та технологій в промисловості*. 2026. № 1 (35). С. 100–114. DOI: <https://doi.org/10.30837/2522-9818.2026.1.100>

Jian, Yu, Sitnikova, O. (2026), "Method for predicting UAV trajectories and evasion for industrial autonomous missions in a dynamic environment", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (35), P. 100–114. DOI: <https://doi.org/10.30837/2522-9818.2026.1.100>

Соловей Б. А., Терентьев О. О.

ПІДХІД ДО СТВОРЕННЯ ФІЗИЧНО ДОСТОВІРНИХ ДАНИХ ДЛЯ СИСТЕМ МОНІТОРИНГУ БЕЗПЕКИ КРАНІВ НА ОСНОВІ СКІНЧЕННО-ЕЛЕМЕНТНОГО МОДЕЛЮВАННЯ

Предметом дослідження є стійкість монорейкового пересувного крана, зокрема коефіцієнти на перекидання й на зсув. Проблема полягає в неадекватності статичних моделей, що базуються на припущенні про жорсткість конструкції, для визначення безпеки. Такі моделі ігнорують вплив деформацій, що призводить до неконсервативних оцінок стійкості та робить їх ненадійними для генерації даних для систем моніторингу на основі моделей штучного інтелекту. **Метою дослідження** є розроблення та обґрунтування підходу до створення даних для оцінювання стійкості монорейкових пересувних кранів, який ґрунтується на скінченно-елементному моделюванні та бере до уваги деформаційні ефекти конструкції. Для досягнення мети в статті визначено такі **завдання**: розроблення підходу до створення наборів даних для оцінювання стійкості монорейкового крана на основі скінченно-елементного моделювання; перевірка запропонованого підходу способом кількісного порівняльного аналізу згенерованих даних із даними, отриманими за статичним підходом, і доведення статистичної значущості розбіжностей. Для реалізації окреслених завдань використано такі **методи**: теорія експериментів, теорія балок Ейлера – Бернуллі, статистична теорія розподілів, теорія чисельного моделювання. **Досягнуті результати**. Отримано кількісні докази розбіжності між статичною та скінченно-елементною моделями. Для стійкості на перекидання середнє значення коефіцієнта β за скінченно-елементною моделлю становить 1.365 (на 2.4% нижче, ніж у статичній – 1.398). Для стійкості на зсув середній коефіцієнт γ зменшився майже удвічі – з 1.87 до 0.85. Статистичний аналіз підтвердив високу значущість цих розбіжностей ($p < 0.001$). **Висновки**. Скінченно-елементна модель бере до уваги геометричну нелінійність. Пружна деформація стріли змінює плече дії сили ваги вантажу, збільшуючи перекидальний момент і горизонтальну реакцію, що системно знижує запас стійкості. Відмінною рисою запропонованої моделі є її здатність виявляти "небезпечні зони" – режими роботи, які статична модель хибно ідентифікує як безпечні. Сферою практичного використання є розроблення систем моніторингу безпеки монорейкового крана на базі штучного інтелекту. Запропонована модель може бути застосована як "генератор" синтетичних даних для навчання та перевірки моделей машинного навчання, що дасть змогу створювати предиктивні системи для оцінювання стійкості монорейкового крана в реальному часі.

Ключові слова: чисельне моделювання; скінченно-елементна модель; генерація даних; моніторинг безпеки; монорейковий кран.

1. Вступ

Будівельні крани є невід'ємним складником сучасного будівництва, однак їх експлуатація пов'язана зі значними ризиками. Стійкість крана – одна з ключових властивостей, що визначає його надійність та безпеку. Допустимі значення коефіцієнтів стійкості чітко регламентовані нормативними документами, зокрема правилами охорони праці [1]. Тому точне прогнозування й моніторинг коефіцієнта стійкості в режимі реального часу є основою будь-якої сучасної системи забезпечення безпеки кранів [2].

Останні досягнення в галузі штучного інтелекту відкривають нові можливості для моніторингу безпечної експлуатації кранів. Рекурентні нейронні мережі (RNN), зокрема моделі з довгою короткочасною пам'яттю (LSTM) та вентильним рекурентним блоком (GRU), ефективні для задач із часовими рядами й навчаються на показниках

датчиків інтернету речей [3–5]. Проте якість навчання таких моделей залежить від точності вхідних даних, які мають відображати реальні динамічні процеси.

Отримати такі дані з натурних експериментів – процес дорий, тривалий і здебільшого неможливий. Традиційно для оцінювання стійкості кранів використовуються аналітичні моделі, що базуються на статистиці абсолютно жорсткого тіла. Такі моделі є спрощеними, оскільки вони не зважають на ключові фізичні явища, зокрема деформацію згину стріли під навантаженням. Це призводить до того, що розраховані коефіцієнти стійкості можуть не відповідати реальній поведінці крана в динамічних режимах, що знижує точність даних для навчання рекурентних нейронних мереж.

Тому виникає потреба в пошуку підходів, що поєднують сучасні методи машинного навчання з фізично обґрунтованими моделями стійкості крана. Це дасть змогу створювати великі обсяги достовірних

даних для навчання й перевірки моделей штучного інтелекту, призначених для систем моніторингу безпеки крана.

З огляду на важливість безпечної експлуатації кранів обмеження традиційних методів прогнозування й потенціал штучного інтелекту це наукове питання є актуальним і потребує подальших досліджень.

2. Аналіз літературних джерел і визначення проблеми

Метод скінченних елементів (МСЕ) є ефективним інструментом для розв'язання широкого кола інженерних задач, пов'язаних із крановими конструкціями.

У роботі [6] автори використовують МСЕ для перевірки методики створення масштабованих фізичних моделей кранових стріл. Порівнюючи МСЕ-моделі повнорозмірної та зменшеної стріл, вони доводять високу кореляцію результатів, що підтверджує адекватність їхнього підходу до масштабування.

Ці результати є важливим обґрунтуванням достовірності МСЕ як інструменту для моделювання поведінки кранових конструкцій. Однак, оскільки робота зосереджується на аналізі міцності окремого елемента (стріли), а не на глобальній стійкості всієї системи крана, її результати й модель не можуть бути безпосередньо застосовані для розв'язання задач моделювання стійкості крана.

Автори роботи [7] аналізують стійкість баштових кранів за умов великої висоти, вітрових і сейсмічних навантажень. Як рішення вони пропонують додаткові розтяжки, що кріплять щоглу крана до будівлі або фундаменту. МСЕ використано для оцінювання впливу цих розкріплень: спочатку визначено зони з підвищеними напруженнями й деформаціями, а потім оновлену модель перевірено з огляду на розкріплення. Порівняння результатів "до" й "після" доводить ефективність підходу.

У дослідженні [8] МСЕ застосовано для аналізу динамічної реакції морського крана-маніпулятора залежно від швидкості прикладання навантаження. Було встановлено експоненційну залежність між часом прикладання навантаження й динамічним фактором. Висновки роботи засвідчують, що навіть незначне скорочення часу підйому вантажу призводить до суттєвого нелінійного зростання динамічних ефектів, як-от: перенапруження, ризик перекидання й деформації конструкції.

У праці [9] на основі МСЕ розраховано максимальні напруження та деформації, що виникають під дією навантаження в різних формах поперечного перерізу вантажного гака. Отримані значення використовуються як критерій оптимізації для вибору оптимальної форми перерізу, яка забезпечує міцність гака. Ці значення використовуються як критерій оптимізації з метою знайти найкращу форму поперечного перерізу для вантажного гака для забезпечення його міцності.

У дослідженні [10] за допомогою МСЕ проаналізовано, як ексцентричне навантаження впливає на напружено-деформований стан кранової балки. У процесі роботи створюється скінченно-елементна модель, на якій імітуються експлуатаційні навантаження. Методологія моделювання цього ефекту, запропонована в [10], буде використана в подальшій роботі для підвищення точності власних розрахунків.

Автори дослідження [11] використали МСЕ для аналізу напружено-деформованого стану елементів баштового крана з горизонтальною стрілою під дією статичних і динамічних навантажень. Метою було глибоке вивчення розподілу напружень для подальшого вдосконалення наявних конструкцій.

У статті [12] МСЕ використано для аналізу втомної довговічності крана без додаткових опор. У дослідженні створено 3D-модель у середовищі САПР, побудовано скінченно-елементну сітку й накладено граничні умови, що імітують реальну експлуатацію. На етапі статичного аналізу розраховано НДС конструкції під дією максимального робочого навантаження. Отримані карти розподілу напружень (за критерієм фон Мізеса) дають змогу визначити потенційні зони зародження втомних тріщин.

Проаналізовані студії [6–12] підтверджують ефективність методу скінченних елементів для дослідження різних аспектів роботи кранових конструкцій, тоді як питання стійкості монорейкових пересувних кранів залишається менш вивченим. Зокрема відсутні моделі, що комплексно беруть до уваги вплив деформації згину балки стріли у вертикальній площині на загальну стійкість конструкції.

Складність отримання навчальних даних для інтелектуальних систем моніторингу також є обмеженням. У роботі [13] описано використання фізично-інформованих нейронних мереж (PINN) для прогнозування коливань навантаження в автомобільних маніпуляторних кранах. Також обґрунтовано, що

запропонована PINN зменшує потребу в кількості навчальних даних і підвищує їх надійність. Це досягається завдяки технології трансферного навчання, яка адаптує попередньо навчену модель до нових ситуацій на основі знань, здобутих на попередніх етапах навчання.

Як зазначено в статті [13], до етапу трансферного навчання дані для навчання було отримано способом моделювання руху автомобільного маніпуляторного крана за допомогою числових симуляцій. Запропонований підхід вимагає наявності початкового набору даних для навчання й перевірки моделі, а їх відсутність до етапу трансферного навчання може стати перешкодою для застосування методу.

Отже, нерозв'язаним залишається питання створення початкових наборів даних для моніторингу стійкості монорейкових кранів із використанням штучного інтелекту.

3. Мета й завдання дослідження

Мета дослідження – розробити й обґрунтувати підхід до створення даних для оцінювання стійкості монорейкових пересувних кранів, який ґрунтується на скінченно-елементному моделюванні й бере до уваги деформаційні ефекти конструкції.

Практичне значення роботи полягає у використанні розробленого підходу як інструменту для створення великих обсягів достовірних даних для навчання й перевірки моделей штучного інтелекту, призначених для систем моніторингу безпеки монорейкового крана.

Щоб досягти окреслену мету, було визначено такі завдання:

1) розроблення підходу до створення наборів даних для оцінювання стійкості монорейкового крана на основі скінченно-елементного моделювання;

2) перевірка запропонованого підходу способом кількісного порівняльного аналізу згенерованих даних, з даними, отриманими за статичним підходом, і доведення статистичної значущості розбіжностей.

4. Матеріали й методи дослідження

Стійкість монорейкового стрілоподібного пересувного крана визначається двома ключовими параметрами: стійкість на перекидання й стійкість на зсув. Стійкість на перекидання вказує на здатність

крана протистояти моментам, коли стріла направлена уздовж рейкового шляху. Стійкість на зсув оцінює стійкість крана до бокових сил, які можуть виникати під час переміщення стріли поперек рейкового шляху.

Умова стійкості на перекидання монорейкового крана має такий вигляд:

$$\beta = \frac{M_{yt}}{M_{пер}} \geq 1.4, \quad (1)$$

де β – коефіцієнт вантажної стійкості; M_{yt} – момент утримування, Н·м, що розраховується як

$$M_{yt} = G_{cr}(l - b) + G_{cw}(c + l) + G_t l; \quad (2)$$

$M_{пер}$ – момент перекидання, Н·м, що розраховується як

$$M_{пер} = G_B(a - l), \quad (3)$$

де a – відстань від осі обертання крана до лінії дії сили G_B , (м); $(a-l)$ – визначає плече дії сили вантажу щодо осі обертання, (м); G_B – вага вантажу, кН; G_{cr} – вага крана, кН; G_{cw} – вага противаги, кН; G_t – вага транспортного візка, Н; b – відстань від осі обертання крана до лінії дії сили G_{cr} , м; c – відстань від осі обертання крана до лінії дії сили G_{cw} , м; l – відстань від осі обертання крана до точки дотику колеса вантажного візка до рейок, м.

У положенні, коли стріла крана розташована поперек рейкового шляху, рейкова колія є основною опорою, а моменти діють щодо осі рейки. У верхній опорі виникає горизонтальна реакція H , що прагне зсунути кран

$$H = \frac{G_B a + G_{cr} b - G_{cw} c}{h}, \quad (4)$$

де h – відстань між нижньою та верхньою опорами, м.

У цьому разі перекидання не можливе, але виникає ризик зсуву крана. Тому стійкість перевіряють за умови, що горизонтальна реакція H не перевищує силу тертя, тобто $H \leq F_{fr}$:

$$F_{fr} = (G_B + G_{cr} + G_{cw} + G_t)\mu \quad (5)$$

$$\gamma = \frac{F}{H} \quad (6)$$

де γ – вантажний коефіцієнт стійкості крана на зсув; μ – коефіцієнт тертя коліс по рейках.

Умова стійкості на зсув має вигляд

$$\gamma \geq 1.5 \dots 1.6. \quad (7)$$

Наведені формули (2)–(6) описують розрахунок стійкості на основі моделі абсолютно жорсткого тіла [14].

4.1. Інструментальна та процедурна основа підходу до створення наборів даних для оцінювання стійкості монорейкового крана

4.1.1. Скінченно-елементна модель оцінювання стійкості монорейкового крана

Запропонуємо МСЕ моделі оцінювання стійкості монорейкового крана, в основі якої розглядається стріла крана довжини l_c із жорстким кріпленням біля опори (вузол 1) та вільним кінцем (вузол 2), на якому зосереджена сила вантажу.

Модель базується на теорії балок Ейлера – Бернуллі, яка припускає, що поперечні перерізи балки залишаються плоскими й перпендикулярними до її осі після деформації. Теорія Ейлера – Бернуллі ігнорує деформації зсуву, таким чином модель, розроблена в цьому дослідженні, буде коректна для тонких балок і малих прогинів; зсувні деформації та геометрична нелінійність не беруться до уваги. Вага вантажу прикладається у вільному вузлі; розподілені навантаження, маса балки й демпфування ігноруються. Динамічні ефекти (розгін / гальмування візка, коливання) не моделюються [15].

Стріла як балковий елемент з двома вузлами (і чотири ступенями свободи) дискретизується на скінченні елементи, що дає змогу описати її деформований стан за допомогою матричного підходу

$$u = \begin{bmatrix} w_1 \\ \theta_1 \\ w_2 \\ \theta_2 \end{bmatrix}, \quad (8)$$

де w_1, w_2 визначають прогини балки у вузлах; θ_1, θ_2 – кути повороту перерізів у вузлах.

Основне рівняння методу скінченних елементів

$$Ku = F, \quad (9)$$

де F – вектор вузлових сил; K – глобальна матриця жорсткості.

Для одного балкового елемента маємо

$$K = k^e; \quad (10)$$

$$k^e = \frac{EI}{l_e^3} \begin{bmatrix} 12 & 6l_e & -12 & 6l_e \\ 6l_e & 4l_e^2 & -6l_e & 2l_e^2 \\ -12 & -6l_e & 12 & -6l_e \\ 6l_e & 2l_e^2 & -6l_e & 4l_e^2 \end{bmatrix}, \quad (11)$$

де E – модуль Юнга, Па; I – момент інерції поперечного перерізу, (м⁴).

Коли балка зафіксована біля опори (вузол 1), тоді прогин (w_1) і кутовий поворот θ_1 дорівнюють нулю, і матриця жорсткості (4) описується тільки для вільного вузла 2 як

$$K_{22}u_2 = F_2, \quad u_2 = \begin{bmatrix} w_2 \\ \theta_2 \end{bmatrix}, \quad (12)$$

де $K_{22} \in R^{2 \times 2}$ – підматриця жорсткості (4); F_2 – вертикальна сила у вільному вузлі 2 описується як

$$F_2 = \begin{bmatrix} -G_B \\ 0 \end{bmatrix}, \quad (13)$$

де G_B – сила ваги вантажу, прикладена у вільному кінці.

Прогін вільного вузла 2 збільшує плече вантажу з огляду на цей скоригований перекидний момент і набуває вигляду

$$M_{\text{пер}}^* = G_B(a - l + |w_2|), \quad (14)$$

де w_2 – вертикальний прогин у вузлі 2 (додатний за умови опускання кінця стріли).

Коефіцієнт вантажної стійкості β крана (1) відповідно до (14) визначається відношенням утримувального моменту $M_{\text{ут}}$ до скоригованого перекидного моменту $M_{\text{пер}}^*$:

$$\beta_{fem} = \frac{M_{\text{ут}}}{M_{\text{пер}}^*} = \frac{M_{\text{ут}}}{G_B(a - l + |w_2|)}. \quad (15)$$

Момент утримання $M_{\text{ут}}$ у розрахунках залишимо в некоригованому вигляді (2), оскільки утримувальні моменти визначаються силами вага крана й противаги, що діють крізь жорсткі конструктивні елементи; їх плечі практично не змінюються внаслідок характерних прогинів стріли, і тому їх можна вважати незмінними в межах цієї моделі.

Оскільки $|w_2| \geq 0$, маємо $M_{\text{пер}}^* \geq M_{\text{пер}}$, звідки впливає

$$\beta_{fem} < \beta. \quad (16)$$

Отже, оцінка стійкості, отримана з огляду на деформації у межах запропонованої моделі, є більш консервативною порівняно зі статичною оцінкою.

Розраховані деформації впливають на плечі сил $a^* = a + u^2$. Отже, горизонтальна сила H (4), зважаючи на деформації, визначається як

$$H_{fem} = \frac{G_{\text{гр}}(a + u_2) + G_{\text{кр}} \cdot b - G_{\text{пр}} \cdot c}{h}. \quad (17)$$

Вираз (17) необхідно брати до уваги для формулювання скінченно-елементної моделі з метою оцінювання стійкості на зсув.

Вантажний коефіцієнт стійкості крана на зсув γ (7) з огляду на (17) обчислюється за виразом

$$\gamma_{fem} = \frac{F_{fr}}{H_{fem}} \geq 1,5 \dots 1,6. \quad (18)$$

Як видно з формули (18), прогин стріли вниз під дією вантажу ($u_2 > 0$) збільшує перекидний момент H_{fem} (17), якщо зіставити з моментом H (4), і, відповідно, зменшує вантажний коефіцієнт стійкості крана γ_{fem} (18) порівняно з коефіцієнтом γ (7).

У цій роботі не моделюються динамічні ефекти та вплив зовнішніх факторів, зокрема: динамічні

навантаження як сили інерції внаслідок розгону / гальмування візка, ривки під час підйому вантажу або коливання стріли від вітру, що можуть впливати на запас стійкості; температурні впливи, які змінюють механічні властивості матеріалу й спричиняють температурні деформації та напруження в конструкції.

Отже, запропонована модель є квазістатичною та бере до уваги лише деформації згину від статичного навантаження. Динамічні й температурні фактори потребують окремого дослідження й не є частиною цієї роботи.

4.1.2. Підхід до створення наборів даних для оцінювання стійкості монорейкового крана

Запропонуємо підхід до створення двох збалансованих наборів даних, кожен з яких містить 1000 унікальних конфігурацій.

Для створення цього набору даних застосовано спрощену модель, у якій утримувальний момент формується вагою крана, а перекидний – тільки вантажем на його плечі. Цільовий коефіцієнт стійкості β встановлено на рівні 1.4. Було згенеровано 500 стійких ($\beta > 1.4$) та 500 нестійких ($\beta < 1.4$) конфігурацій поблизу критичного порогу, що забезпечує достатнє подання як безпечних, так і небезпечних випадків.

Параметри для генерації даних обрано таким чином, щоб охопити широкий спектр реалістичних сценаріїв:

- властивості матеріалу: модуль пружності E було зафіксовано на рівні 2.1×10^8 Па для усунення варіацій, пов'язаних з матеріалом;

- геометрія: момент інерції I , визначено з високою розбіжністю в значеннях (мінімум = 2.0×10^{-6} м⁴, максимум = 5.8×10^{-4} м⁴) для моделювання як гнучких, так і жорстких перерізів балки. Виліт стріли a змінювався в межах від 5 до 20 м, а база опори k – в діапазоні 1.5–3.5 м. Їх поєднання визначає плече $l_e = a - k$, що змінюється від 1.6 до 18.3 м ($\sigma \approx 4.36$ м);

- навантаження: власна вага крана G_{cr} та вага вантажу G_B також змінювалися в значних межах (38–3745 кН та 50–150 кН відповідно), щоб охопити легкі й важкі промислові конфігурації.

Формування набору даних для оцінювання стійкості на зсув базувалося на визначенні значення противаги G_{cw} навколо цільового рівня стійкості $\gamma \approx 1.5$, що розраховується за формулою

$$G_{cw} \approx \frac{G_B a + G_{cr} b - \frac{(G_B + G_{cr} + \frac{G_{cr} a}{c} + G_t) \mu h}{1.5}}{c}, \quad (19)$$

де a , b , c , h , G_{cr} визначають за рівномірним розподілом з параметрами $U(10, 25)$, $U(2.0, 3.0)$, $U(4.0, 6.0)$, $U(7.0, 10.0)$, $U(100, 200)$ відповідно.

Для врахування природної невизначеності матеріальних властивостей, геометрії та умов навантаження після початкового точного визначення параметрів було введено випадкову варіацію на рівні $\pm 20\%$

$$x = x_0 \cdot (1 + \delta), \quad \delta \sim U(-0.2, +0.2). \quad (20)$$

Коефіцієнт тертя μ було зафіксовано.

Розроблений підхід дає змогу створювати контрольовані та репрезентативні набори даних, що є необхідною умовою для коректного порівняльного аналізу й перевірки дослідницьких гіпотез. Описові статистики згенерованих наборів даних подано в розділі "Результати досліджень та їх обговорення" (табл. 1, 2).

4.2. Методологія перевірки й порівняльного аналізу підходу до створення даних для оцінювання стійкості монорейкового крана

Для кількісного порівняння статичної та розробленої МСЕ-моделей сформулюємо дослідницькі гіпотези.

Гіпотеза H_1 : врахування деформацій у МСЕ-моделі систематично зменшує розрахунковий запас стійкості монорейкового крана порівняно зі статичною моделлю.

Гіпотеза H_2 : різниця між коефіцієнтами стійкості, розрахованими за МСЕ та статичною моделями, є статистично значущою.

Порівняння проводитимемо між двома наборами даних: створеного за допомогою базової статичної методики та створеного за допомогою запропонованого підходу на основі МСЕ.

Для перевірки статистичної значущості різниці між двома наборами парних даних застосовано парний t -тест Стьюдента. Він дає змогу визначити, чи є середня різниця між двома вибірками статистично відмінною від нуля. T -статистика розраховується за формулою

$$t = \frac{\bar{a}}{s_d / \sqrt{n}}, \quad (21)$$

де \bar{a} – середнє значення різниць між парними спостереженнями; s_d – стандартне відхилення цих різниць; n – обсяг вибірки.

Рівень значущості $p < 0.05$ буде використано як критерій для відхилення нульової гіпотези про відсутність різниці.

5. Результати дослідження

5.1. Результати розробленого підходу до створення даних для оцінювання стійкості монорейкового крана

Результатом є два збалансованих набори даних по 1000 конфігурацій для оцінювання стійкості на перекидання й на зсув. У табл. 1 подано описові

статистики для набору даних, згенерованого для оцінювання стійкості на перекидання.

Створений набір визначається широким діапазоном значень вхідних параметрів. У табл. 2 наведено описові статистики для набору даних, згенерованого для оцінювання стійкості на зсув.

Створені набори вихідних даних з описовими статистиками (табл. 1) призначені для обчислення утримувального моменту M_{yt} (1), перекидного моменту $M_{пер}$ (2), коефіцієнта вантажної стійкості β (3), скоригованого перекидного моменту $M_{пер}^*$ (14) і коефіцієнта вантажної стійкості β_{fem} крана з огляду на деформації (15).

Таблиця 1. Описові статистики створено набору даних для оцінювання стійкості крана на перекидання

Описові статистики	E, Па	I, м ⁴	G _{сг} , кН	G _в , кН	a, м	к, м	I _с , м
Кількість рядків	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00
середнє	2.1×10^8	1.48×10^{-4}	659.14	99.54	12.42	2.52	9.90
Стандартне відхилення	0.00	1.32×10^{-4}	528.52	28.90	4.29	0.59	4.36
мінімум	2.1×10^8	2.0×10^{-6}	38.39	50.00	5.01	1.50	1.64
0.25	2.1×10^8	4.3×10^{-5}	293.49	74.61	8.71	1.98	6.28
0.50	2.1×10^8	1.07×10^{-4}	511.82	100.38	12.38	2.53	9.92
0.75	2.1×10^8	2.18×10^{-4}	847.97	124.62	16.26	3.03	13.71
максимум	2.1×10^8	5.8×10^{-4}	3745.36	149.89	19.98	3.50	18.31

Таблиця 2. Описові статистики для формування набору даних для оцінювання стійкості крана на зсув

Описові статистики	I _с , м	I, м ⁴	G _в , кН	G _{св} , кН	G _{сг} , кН	G _т , кН	a, м	b, м	c, м	h, м
Кількість рядків	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
середнє	17.4	4×10^{-4}	148.9	444.9	104.2	14.9	17.4	2.5	5.03	8.49
Стандартне відхилення	4.4	2.4×10^{-4}	29.9	162.8	22.8	2.99	4.4	0.3	0.57	0.87
мінімум	10	5.6×10^{-5}	100.1	125.7	62.5	10	10	2	4	7
0.25	13.5	2×10^{-4}	122.1	320.3	84.9	12.2	13.6	2.2	4.5	7.7
0.50	17.5	3.4×10^{-4}	147.3	421.9	103.1	14.7	17.5	2.5	5.1	8.4
0.75	21.2	5.6×10^{-4}	176.0	539.7	121.8	17.6	21.2	2.8	5.5	9.2
максимум	25	1.2×10^{-3}	199.9	973.8	157.5	20	25	3	6	10

Створені набори вихідних даних з описовими статистиками (табл. 2) призначені для обчислення горизонтальної сили H (4), сили тертя F_{fr} (5), коефіцієнта стійкості крана на зсув γ (6), горизонтальної сили H_{fem} (17) та коефіцієнта стійкості крана на зсув γ_{fem} (18) з огляду на деформації.

5.2. Результати перевірки запропонованого підходу до створення наборів даних для оцінювання стійкості монорейкового крана

На рис. 1 подано карти розподілу коефіцієнта стійкості β , отримані за результатами статичного (а) та МСЕ (б) аналізу. У процесі порівняння графіків

необхідно звернути увагу на положення лінії критичної стійкості $\beta = 1.4$. На графіку для МСЕ-моделі (б) ця лінія зміщена вище й більше вправо. Між пунктирними лініями позначено зону, де для статичної моделі $\beta \geq 1.4$, а для МСЕ-моделі $\beta < 1.4$.

Гістограми розподілу коефіцієнта β для обох моделей запропоновано на рис. 2. Гістограма для МСЕ-моделі зміщена вліво щодо гістограми для статичної моделі.

У табл. 3 наведено статистичні показники для розподілів, поданих на рис. 2. Середнє значення β для статичної моделі становить 1.398, тоді як для МСЕ-моделі – 1.365.

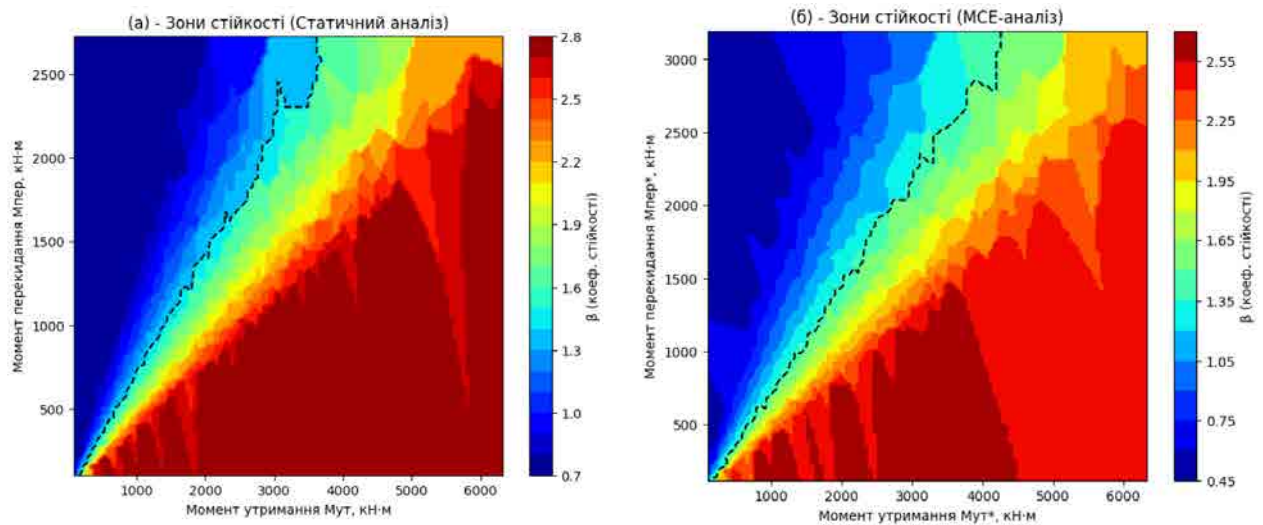


Рис. 1. Зона стійкості на перекидання: а – статичний аналіз; б – МСЕ-аналіз

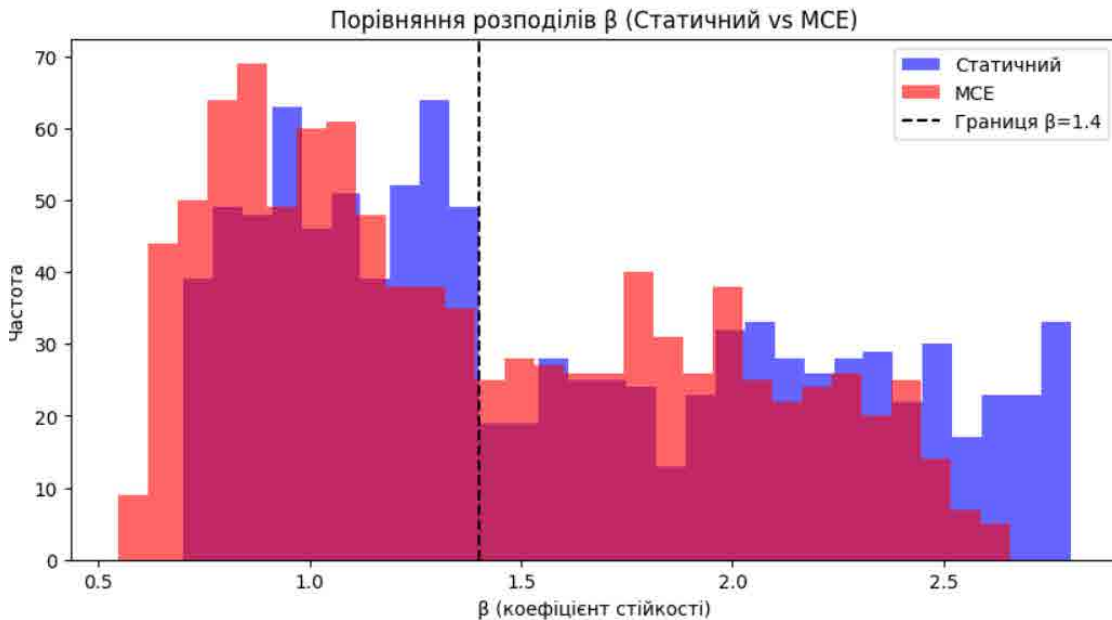


Рис. 2. Порівняння розподілів коефіцієнта стійкості на перекидання β

Таблиця 3. Статистика коефіцієнта стійкості β на перекидання

Модель	Середнє (mean)	Медіана (median)	Стандартне відхилення (std)
Статична	1.398	1.395	0.451
МСЕ	1.365	1.358	0.448

Отже, результати порівняння коефіцієнтів на рис. 1 та в табл. 3 підтверджують першу гіпотези дослідження для оцінювання стійкості монорейкового крана на перекидання. Результати парного тесту Стьюдента для перевірки статистичної значущості результатів моделювання для коефіцієнта стійкості β подано в табл. 4.

Таблиця 4. Парний тест Стьюдента для перевірки статистичної значущості результатів моделювання для β

Тест Стьюдента	t-статистика	p-value
$t(\beta, \beta_{fem})$	50.9	0

Розрахована t-статистика для коефіцієнта β становить 50.9, якщо значення p-value < 0.001.

На рис. 3, 4 запропоновано карти розподілу коефіцієнта стійкості на зсув γ . На графіку статичного аналізу (рис. 3) лінія критичної стійкості $\gamma = 1.5$ розташована інакше, ніж на графіку МСЕ-аналізу (рис. 4).

На графіку для МСЕ-моделі (рис. 4) більша частина простору параметрів розташована під лінією $\gamma = 1.5$.

На рис. 5 зображено гістограми розподілу коефіцієнта γ . Розподіл для МСЕ-моделі (червоний колір) зміщений вліво порівняно з розподілом для статичної моделі (синій колір).

Статистичні показники для розподілів коефіцієнта γ подано в табл. 5.

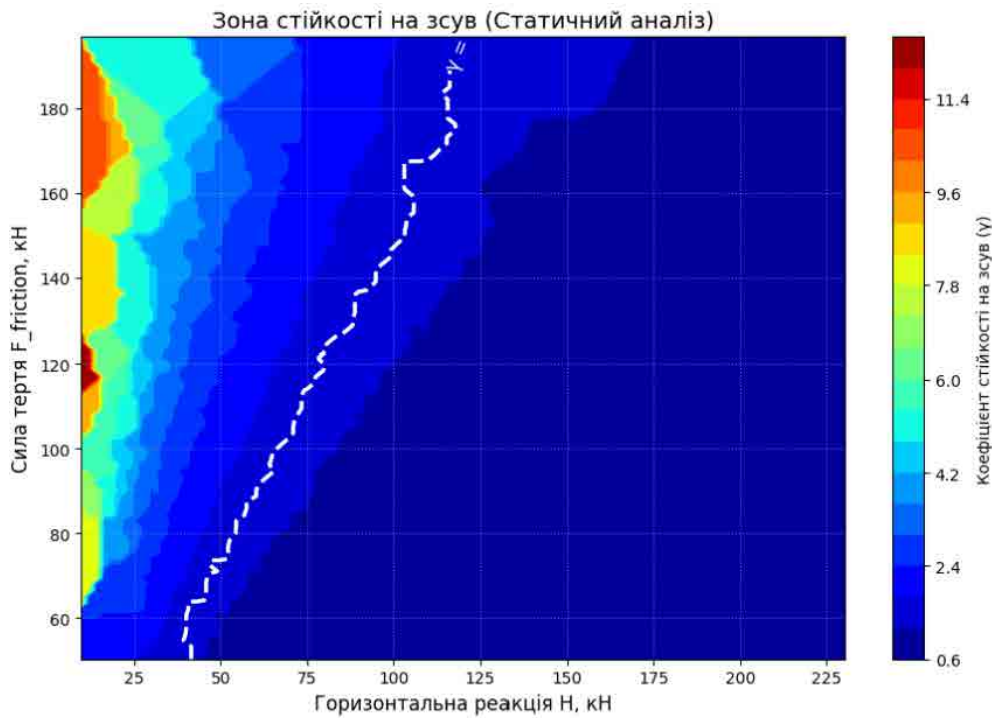


Рис. 3. Зона стійкості на зсув на основі статичного аналізу

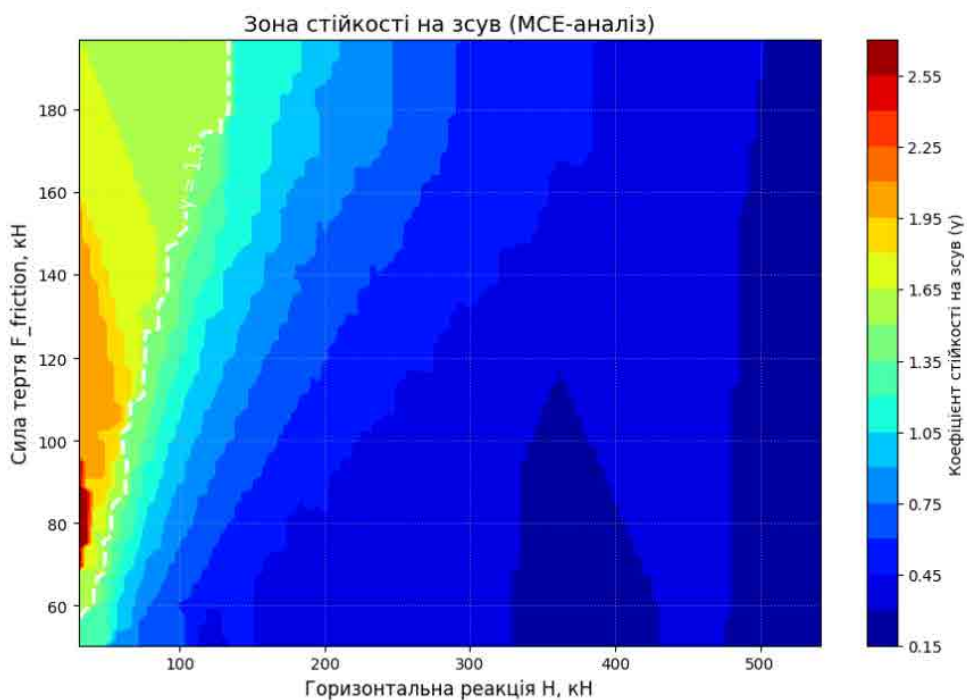
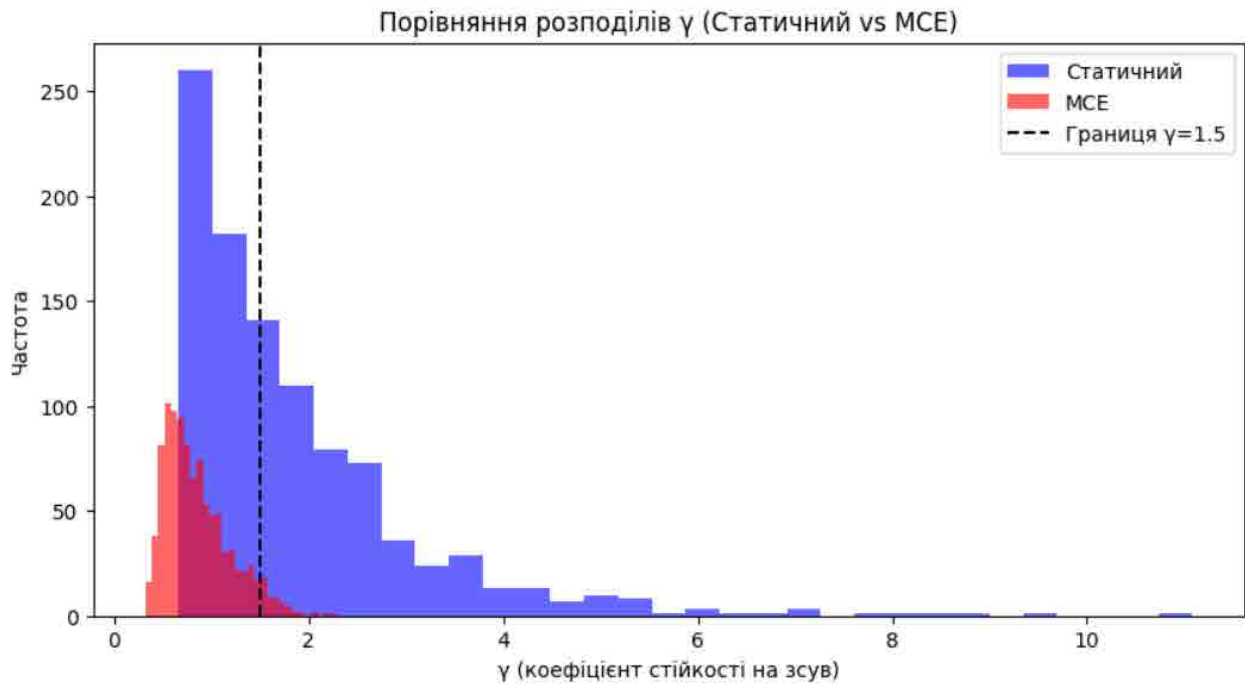


Рис. 4. Зона стійкості на зсув на основі МСЕ-аналізу

Рис. 5. Порівняння розподілів коефіцієнта стійкості на зсув γ Таблиця 5. Статистика коефіцієнта стійкості γ на зсув

Модель	Середнє (mean)	Медіана (median)	Стандартне відхилення (std)
Статична	1.87	1.505	1.516
MCE	0.85	0.774	0.326

Середнє значення γ для MCE-моделі (0.85) є нижчим, ніж для статичної моделі (1.87).

Результати проведеного парного тесту Стьюдента представлені в табл. 6.

Таблиця 6. Парний тест Стьюдента для перевірки статистичної значущості результатів моделювання для γ

Тест Стьюдента	t-статистика	p-value
$t(\gamma, \gamma_{fem})$	23.8	0

Розрахована t-статистика для коефіцієнта γ становить 23.8, а значення p-value – менше ніж 0.001.

6. Обговорення результатів

Досягнуті внаслідок обчислювальних експериментів результати демонструють систематичну розбіжність між статичною моделлю та розробленою скінченно-елементною моделлю (MCE). Ця різниця пояснюється ключовою особливістю MCE-моделі – врахуванням геометричної нелінійності, а саме пружної деформації (прогину) стріли під дією вантажу.

Для стійкості на перекидання (рис. 1, 2; табл. 3, 4) зниження коефіцієнта β в MCE-моделі є наслідком того, що прогин стріли призводить до збільшення реального плеча дії сили вантажу. Це зі свого боку збільшує перекидний момент $M_{per\ i}$, зрештою, зменшує коефіцієнт стійкості. Статична модель, розглядаючи стрілу як абсолютно жорстке тіло з фіксованим плечем, не здатна взяти до уваги цей дестабілізаційний ефект.

Для стійкості на зсув (рис. 3–5; табл. 5, 6) ефект є більш вираженим. Деформація стріли та всієї конструкції під навантаженням призводить до перерозподілу сил і значного збільшення горизонтальної реакції H , що прагне зсунути кран. Це зменшує коефіцієнт γ , що й пояснює практичне зникнення "безпечної зони" в MCE-моделі порівняно зі статичною.

Існування виявленої "небезпечної зони", де статична модель демонструє відповідність нормативам, а MCE – ні, є прямим доказом неадекватності спрощеного підходу. Отже, дослідження не лише підтверджує висунуті гіпотези, а й надає кількісне обґрунтування необхідності переходу до більш

досконалих, нелінійних моделей для реалістичного оцінювання безпеки.

Для обчислювальних експериментів використовувалася спрощена геометрія крана. Реальна конструкція має більш складні вузли, з'єднання та розподіл мас, що може вплинути на результати.

Висновки

У дослідженні розроблено підхід до створення наборів даних для оцінювання стійкості монорейкового крана. Інструментальна частина містить:

1) скінченно-елементну модель для оцінювання стійкості на перекидання, яка, на відміну від статичної моделі, бере до уваги деформацію згину стріли. Результатом її застосування є більш консервативна й реалістична оцінка: середнє значення коефіцієнта стійкості β зменшилося з 1.398 (статична модель) до 1.365 (МСЕ-модель). Це пояснюється тим, що прогин стріли збільшує реальне плече дії сили вантажу, посилюючи перекидний момент;

2) скінченно-елементну модель для оцінювання стійкості на зсув. Встановлено, що вплив деформацій на цей тип стійкості є ще більш критичним, ніж на перекидання. Кількісно це виражено в зменшенні середнього значення коефіцієнта γ майже удвічі – з 1.87 до 0.85. Такий значний ефект пояснюється перерозподілом сил внаслідок деформації конструкції, що суттєво збільшує горизонтальну реакцію, яка прагне зсунути кран. Процедурна основа описує послідовність дій для створення збалансованих наборів даних для оцінювання стійкості монорейкового крана.

Перевірено запропонований підхід до створення наборів даних. Підтверджено дослідницьку гіпотезу (H_1) про те, що врахування деформацій систематично знижує розрахунковий запас стійкості. Особливістю цього результату є виявлення "небезпечної зони" – режимів роботи, які статична модель хибно ідентифікує як безпечні, тоді як МСЕ-модель вже фіксує втрату стійкості. Це пояснюється нездатністю статичної моделі зважати на нелінійні ефекти, що призводить до неконсервативних оцінок.

Підтверджено дослідницьку гіпотезу (H_2) про статистичну значущість виявлених розбіжностей. Застосування парного t-тесту Стьюдента дало високі значення t-статистики ($t = 50.9$ для β і $t = 23.8$ для γ) за умови $p < 0.001$. Це свідчить про те, що зниження запасу стійкості в МСЕ-моделі є системним,

а не випадковим ефектом, що пояснюється фундаментальними розбіжностями у фізичних припущеннях, закладених в основу кожної з моделей.

Недоліком можна вважати відсутність моделі на основі натурних експериментів. Порівняння результатів МСЕ-моделювання з даними, отриманими з реального фізичного крана, значно підвищило б достовірність висновків.

Наукова новизна дослідження полягає в запропонованому підході до створення даних для систем моніторингу монорейкового крана на основі скінченно-елементного моделювання (МСЕ). Новий підхід бере до уваги деформації згину й перерозподіл внутрішніх зусиль, що дає змогу точно визначити "небезпечні зони" режимів роботи, які статичні моделі вважають безпечними.

Напрямом подальших досліджень є розроблення архітектури рекурентної нейронної мережі типу GRU (*Gated Recurrent Unit*), призначеної для прогнозування ризику втрати стійкості монорейкового пересувного крана в реальному часі за двома ключовими критеріями: на перекидання й на зсув. Розроблений і перевірений у цій роботі підхід до створення даних слугуватиме інструментом для створення великих синтетичних наборів даних, необхідних для навчання, тестування й перевірки рекурентної нейронної мережі.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів, зокрема фінансового, особистого, авторського чи будь-якого іншого характеру, який міг би вплинути на дослідження, а також на результати, опубліковані в цій статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Дані будуть надані за обґрунтованим запитом.

Використання засобів штучного інтелекту

Автори підтверджують, що не застосовували технології штучного інтелекту для написання цієї роботи.

References

1. Вапнічна, В. В., Сергієнко, М. І., Фролов, О. О. (2025), *Технологія, механізація та організація геотехнічного будівництва: гірничі і будівельні машини*: навч. посіб. для здобувачів ступеня бакалавра за освіт. програмою «Геоінженерія» спец. G16 Гірництво та нафтогазові технології, КПІ ім. Ігоря Сікорського, Київ, 248 с., available at: <https://ela.kpi.ua/handle/123456789/75143> (last accessed 18.03.2026).
2. Awad, A. H., Alsabaan, M., Ibrahim, M. I., Saraya, M. S., Elksasy, M. S. M., Ali-Eldin, A. M., Abdelsalam, M. M. (2024), "Low-cost IoT-based sensors dashboard for monitoring the state of health of mobile harbor cranes: Hardware and software description", *Heliyon*, 10(22), e40239. DOI: <https://doi.org/10.1016/j.heliyon.2024.e40239>
3. Han, Z., Zhao, J., Leung, H., Ma, K.F. and Wang, W. (2019), "A review of deep learning models for time series prediction", *IEEE Sensors Journal*, 21(6), pp.7833-7848. available at: <https://ieeexplore.ieee.org/document/8742529>
4. Li, W., Law, K. E. (2024), "Deep learning models for time series forecasting: A review", *IEEE Access*, 12, pp. 92306–92327. DOI: <https://doi.org/10.1109/ACCESS.2024.3422528>
5. Терентьев, О., Соловей, Б. (2025), "Машинне навчання Байєсової нейронної мережі з гамма-розподілом для оцінювання стійкості монорейкового крана", *Управління розвитком складних систем*, (62), pp. 134–140. DOI: <https://doi.org/10.2412-9933.2025.62.134-140>
6. Liu, Z., Liu, K., Chen, C., Lu, S., Wang, H., Yin, L. (2025, June), "Similitude-based construction of scaled crane boom models via finite element analysis", *Journal of Physics: Conference Series*, 3033(1), 012034. DOI: <https://doi.org/10.1088/1742-6596/3033/1/012034>
7. Georgiev, C. (2024), "Enhancing tower crane stability through mast bracing and finite element analysis", *Proceedings of the Technical University of Sofia*, 74, pp. 8-14. DOI: <https://doi.org/10.47978/TUS.2024.74.01.002>
8. Tomasi, I., Solazzi, L. (2025), "Dynamic analysis of an offshore knuckle-boom crane under different load applications laws", *Applied Sciences*, 15(14), 8100. DOI: <https://doi.org/10.3390/app15148100>
9. Dipu, M. N. H., Apu, M. H., Chowdhury, P. P. (2024), "Identification of the effective crane hook's cross-section by incorporating finite element method and programming language", *Heliyon*, 10(9). DOI: <https://doi.org/10.1016/j.heliyon.2024.e29918>
10. Lu, C., Yang, Z., Li, P., Zhang, X., Huang, J., Wang, L. (2022), "Finite element analysis and reinforcement of steel crane beam under eccentric track loading", *Machines*, 10(9), 783. DOI: <https://doi.org/10.3390/machines10090783>
11. Zhang, Q., Mei, B., Yang, H., Hu, X., An, W., Yue, Y., Wang, Z. (2025), "Stress measurement and analysis of structural parameters of flat arm tower crane under different working conditions", *Buildings*, 15(7), 1137. DOI: <https://doi.org/10.3390/buildings15071137>
12. Buczkowski, R., Żyliński, B. (2021), "Finite element fatigue analysis of unsupported crane", *Polish Maritime Research*, 28(1), pp.127-135, DOI: <https://doi.org/10.2478/pomr-2021-0012>
13. Zhou, Z., Johns, B., Fang, Y., Bai, Y., Abdi, E. (2025), "Physics-informed neural network for load sway prediction in travelling autonomous mobile cranes", *Advanced Engineering Informatics*, 65, 103269. DOI: <https://doi.org/10.1016/j.aei.2025.103269>
14. Воляннюк, В. О., Горбатюк, Є. В. (2021). *Розрахунок механізмів вантажопідіймальних машин*: навч. посіб., КНУБА, Київ, 136 с.
15. Склярів, І. О., Тонкачєв, В. Г., Склярєва, Т. С. (2024), *Конструкції будівель і споруд. Металеві конструкції*: навч. посіб., Каравела, Київ, 174 с.

Received (Надійшла) 13.10.2025

Accepted for publication (Прийнята до друку) 10.02.2026

Publication date (Дата публікації) 30.03.2026

Відомості про авторів / About the Authors

Соловей Богдан Анатолійович – Київський національний університет будівництва і архітектури, аспірант кафедри інформаційних технологій проектування та прикладної математики; Київ, Україна;

Bohdan Solovei – Kyiv National University of Construction and Architecture, PhD Student at the Department of Information Technology for Design and Applied Mathematics; Kyiv, Ukraine;

e-mail: solovei_ba-2023@knuba.edu.ua

ORCID ID: <http://orcid.org/0009-0008-0328-1123>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=59531932700>

Терентьев Александр Александрович – доктор технічних наук, Київський національний університет будівництва і архітектури, професор кафедри інформаційних технологій проектування та прикладної математики; Київ, Україна;
Oleksandr Terentyev – Doctor of Technical Sciences, Kyiv National University of Construction and Architecture, Professor at the Department of Information Technology for Design and Applied Mathematics; Kyiv, Ukraine;
e-mail: terentiev.oo@knuba.edu.ua
ORCID ID: <https://orcid.org/0000-0001-9499-6635>
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57211341629>

AN APPROACH TO GENERATING PHYSICALLY REALISTIC DATA FOR CRANE SAFETY MONITORING SYSTEMS BASED ON FINITE ELEMENT MODELING

This article focuses on the stability of monorail mobile cranes, particularly tipping and sliding stability coefficients. Traditional static models, which assume rigid structures, fail to account for deformation effects, leading to non-conservative and unreliable stability assessments. This makes them unsuitable for generating data used in AI-based safety monitoring systems. **The aim** of the study is to develop and substantiate an approach to generating data for assessing the stability of monorail traveling cranes based on finite element modeling, taking into account structural deformation effects. To achieve this aim, the following **tasks** are addressed in the paper: developing an approach to generating datasets for assessing the stability of a monorail crane based on finite element modeling; validating the proposed approach through a quantitative comparative analysis of the generated data with data obtained using the static approach; demonstrating the statistical significance of the identified differences. **Methods** used: experimental design, Euler-Bernoulli beam theory, statistical distribution analysis, and numerical simulation. **Results** show that FEM produces significantly different outcomes. The average tipping coefficient β decreased by 2.4% (from 1.398 to 1.365), and the sliding coefficient γ dropped nearly 50% (from 1.87 to 0.85), with high statistical significance ($p < 0.001$). **Conclusions:** FEM captures geometric nonlinearity, showing that elastic deformation alters the load path, increasing tipping moments and reducing stability margins. Unlike static models, the proposed method identifies unsafe operating zones previously misclassified as safe. Practical application: the model can generate synthetic data to train machine learning systems for real-time crane stability prediction, enhancing the safety of AI-driven monitoring technologies.

Keywords: numerical modeling; finite element model; data generation; safety monitoring; monorail crane.

Бібліографічні описи / Bibliographic descriptions

Соловей Б. А., Терентьев О. О. Підхід до створення фізично достовірних даних для систем моніторингу безпеки кранів на основі скінченно-елементного моделювання. *Сучасний стан наукових досліджень та технологій в промисловості*. 2026. № 1 (35). С. 115–126. DOI: <https://doi.org/10.30837/2522-9818.2026.1.115>

Solovei, B., Terentyev, O. (2026), "An approach to generating physically realistic data for crane safety monitoring systems based on finite element modeling", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (35), P. 115–126. DOI: <https://doi.org/10.30837/2522-9818.2026.1.115>

Oleksii Shapryro, Ihor Sotnyk

A METHOD OF SEMANTIC SEARCH FOR EDUCATIONAL CONTENT BASED ON MULTI-AGENT TECHNOLOGIES

The digital transformation of industry is accompanied by the active adoption of new technologies and the rapid evolution of production processes. A significant portion of educational materials is distributed across various information sources, including internal corporate systems, open educational platforms, and specialized web resources. Such resources often contain duplicates, redundant information, and heterogeneous metadata, which complicates the timely retrieval of relevant learning materials. The **subject** of the study is a method of semantic search for educational content in a distributed information environment using ontology-based knowledge models. The **goal** of the work is to investigate a method of semantic search for educational content in a distributed information environment based on a multi-agent organization of information resource processing and the use of ontology-based knowledge models. The **objectives** of the study are: to investigate the architectural model of a multi-agent search system; to develop a semantic selection algorithm based on the comparator identification method and an ontology-based model; to formalize a relevance evaluation predicate considering weighted metadata coefficients; to develop a multi-agent software platform; and to experimentally evaluate performance and resource consumption under different agent operating modes. Research **methods** include: the method of multi-agent organization of information resource processing with non-blocking message exchange; three-level URL deduplication; ontology-based term matching and a formalized relevance evaluation predicate; and experimental measurement of processing time, the number of processed links, and system resource consumption. **Results**: a model of a multi-agent system with four types of agents and a semantic search algorithm eliminating loops and duplicate links has been proposed; a software platform based on Kotlin using coroutines and asynchronous interaction between agents has been implemented; experimental results demonstrate that the proposed organization of processing provides higher performance compared to the sequential mode. **Conclusions**: the integration of semantic search and a multi-agent architecture enables efficient organization of the process of discovering and processing educational content in a distributed environment. The proposed method ensures coordinated operation of agents, eliminates link duplication, and provides a rational balance between search completeness and the use of computational resources.

Keywords: semantic search; educational content; software agents; multi-agent systems; ontologies; scalability; performance; asynchrony; Kotlin.

Introduction

The current stage of information technology development in industry and related manufacturing sectors is characterized by a growing volume of educational content and increasing demands on the effectiveness of e-learning systems. The digital transformation of enterprises, rapid updates to production equipment, hardware and software systems, and technological regulations create a constant need for the continuous improvement of staff qualifications. Under these conditions, the availability of relevant educational content directly impacts the quality of technological operations, the level of workplace safety, compliance with standards, and the reduction of the risk of errors and downtime. The rapid growth in the volume of open educational content is accompanied by a sharp diversity of sources, formats, and the quality of descriptions. A significant portion of the materials is published outside of specialized platforms, with minimal or inconsistent

metadata, which complicates the automatic discovery, classification, and subsequent reuse of resources in learning environments. For industrial scenarios, this is further complicated by the fact that training materials are often scattered across internal enterprise knowledge bases, corporate LMSs, and external resources (suppliers, technical documentation, courses, standards), creating knowledge fragmentation and additional barriers to rapid learning.

Traditional keyword-based search approaches yield inconsistent results precisely in educational scenarios, where semantic alignment with the subject domain, learning objectives, and the expected level of proficiency is crucial. For example, in industry, the same terminology can have different meanings depending on the type of equipment or standard, and content requirements are determined not only by the general topic but also by the employee's role, access level, operational context, and other factors that must also be taken into account. Additional factors contributing to quality degradation

include duplicates and semi-duplicates, different language versions of the same material, time-based revisions without proper versioning, as well as noise in the form of advertising and navigation elements. In the absence of structured descriptions, even routine tasks require consideration of the context of the subject ontology, not just term matches.

At the same time, the requirements for the scalability and resilience of such systems are growing: data collection and preprocessing must handle arrays of tens of thousands of links, withstanding peak loads on memory, the processor, and the network, while maintaining controllable latency and fault tolerance. In practical terms, this is compounded by external constraints, which complicate the maintenance of consistent throughput and reproducibility of results within corporate infrastructure.

All these discrepancies between the needs of real-world educational applications for industrial workforce upskilling and the capabilities of existing search approaches highlight a problem requiring further research: search quality directly impacts the speed of learning, safety, and the effectiveness of production processes.

Literature Review and Problem Definition

The IEEE LTSC (Learning Technology Standards Committee) [1] was established to develop specifications and recommendations for the integration of modern information technologies in education. This committee develops architectures and technologies for educational systems in accordance with international standards. Separately, the IMS Global Learning Consortium is engaged in developing standards to ensure the interoperability of educational platforms [2]. According to the IEEE LTSA (Learning Technology System Architecture), the universality of this architecture ensures its use as a reference model for the development of educational technology systems of any type. It has been shown that these standards form the basis for the interoperability of various educational platforms and ensure a unified approach to organizing the educational process. However, issues remain unresolved regarding the adaptation of these general architectural recommendations to the specific needs of intelligent semantic search for educational content and the lack of concrete algorithmic solutions for the automated analysis of educational resources. This may be due to objective

difficulties related to the excessive abstractness of international standards, which focus on general principles rather than specific technical implementations, making their application challenging for practical semantic search tasks.

In [3], key trends in adaptive e-learning systems were investigated based on an analysis of 57 studies from recent years. It was shown that building effective adaptive systems requires learner, domain, and adaptation models that account for individual learning characteristics and dynamically adjust instructional materials. However, issues remain unresolved regarding the fact that most existing systems do not provide sufficient accuracy in predicting learner needs in real time and do not adapt to the context of learning activities. This is due to objective difficulties related to the complexity of collecting and processing large volumes of data on student behavior, as well as the inherent uncertainty of individual cognitive processes, which makes universal solutions economically impractical.

In [4], an ontology-driven framework for semantic search of educational content is proposed, utilizing natural language processing and machine learning methods for text classification. It is shown that the ontological approach allows for the identification of semantic relationships between concepts and the discovery of the most relevant materials compared to traditional keyword search. However, issues related to the automation of ontology creation and updating for various subject areas remain unresolved. The reason for this is the cost involved in engaging experts to develop high-quality ontologies and the fundamental complexity of automating the processes of identifying semantic relationships in dynamically evolving knowledge.

In [5], the application of the latest approaches to semantic search using ontological models and document relevance analysis methods was investigated. The potential of these methods for improving the accuracy of identifying relevant educational materials was demonstrated. However, issues remain unresolved regarding maintaining a balance between search accuracy and system scalability when processing large repositories of educational objects, as well as the integration of such systems with existing educational content metadata standards. The reason for this is the difficulty associated with the fact that complex machine learning models often require significant computational resources, making their application impractical in distributed environments with limited resources.

In [6], decentralized multi-agent systems for solving planning and optimization problems in manufacturing environments were investigated. It was shown that the multi-agent approach ensures scalability and resilience to failures of individual components. However, issues remain unresolved regarding ensuring effective communication between agents when there are a large number of participants and adapting agent-based systems to the specifics of educational applications with semantically rich data. The reason for this is the exponential growth in the complexity of communication protocols and the fundamental difficulty of predicting the behavior of a large number of interacting agents.

Additionally, many studies leave unresolved issues related to practical implementation on modern programming platforms and their adaptation to the requirements of high-performance real-time data processing. This is due to difficulties stemming from the fact that many existing implementations are based on outdated technologies and do not support modern asynchronous programming mechanisms. One way to overcome these difficulties is to use modern programming languages with built-in support for asynchronous programming and lightweight parallelism mechanisms. For example, some case studies describe the use of coroutines and asynchronous operations to organize effective communication in distributed systems. It has been shown that this approach allows for increasing system throughput and reducing resource management costs.

Research Objectives and Tasks

Systematizing the local problems identified in the analyzed sources allows us to formulate a general unresolved problem: there is currently no approach to searching for educational content that is simultaneously effective in terms of result relevance, performance, scalability, and resource efficiency. Its relevance stems from the need to quickly find relevant educational materials from a vast number of sources, taking into account subject-area ontologies, content duplication and noise, as well as computational resource constraints. Several factors are the main reasons this problem remains unsolved. **First**, the lack of solutions that combine multi-agent architecture with semantic technologies. **Second**, the inefficiency of traditional algorithms for processing heterogeneous metadata of educational resources. **Third**, the difficulty of scaling distributed

systems without compromising service quality. **Fourth**, many existing implementations are based on outdated technologies that do not support modern asynchronous programming mechanisms. All of this supports the need for relevant research to address this issue.

Accordingly, the objective of this work was defined as the need to develop an approach to educational content search that is effective in terms of both the relevance of results and performance, scalability, and resource efficiency.

To achieve this objective, the following tasks were set:

- to develop an architectural model of a multi-agent search system;
- create a semantic search algorithm using the comparator identification method and ontological models of subject domains;
- implement a software platform using modern technologies with built-in support for asynchronous programming and lightweight parallelism mechanisms;
- experimentally evaluate the system's performance and resource consumption in various agent operating modes.

Materials and Methods

In the context of semantic search for educational content, it is necessary to process large volumes of information resources located in a distributed web environment and characterized by high dynamism and structural heterogeneity. Traditional centralized information retrieval systems may face limitations in scalability, performance, and flexibility of resource processing control under such conditions. One effective approach to organizing search systems in a distributed environment is the use of multi-agent technologies, which allow the system's functionality to be decomposed into autonomous software components that perform specialized tasks and interact with one another to achieve a common goal.

A software agent is an autonomous software component capable of functioning in an information environment, perceiving its state, making decisions according to specified rules, and interacting with other agents to achieve a common goal.

This paper proposes an architectural framework for an ontology-driven semantic search of educational resources using a multi-agent system. The operation of agents is ensured and regulated by a naming service and an access control module. The agent registry is

designed to provide agents with information about the available interaction environment, specifically regarding the composition of the system's agents, their roles, and functional purposes. An ontological model is used to ensure a unified formalized description of the domain, common to all agents. The information resources supporting the system's operation are stored in a database and a knowledge base.

The multi-agent search system is implemented using four functional types of agents:

- the coordination agent manages the search process, generates initial query descriptors, distributes tasks among other system agents, controls the sequence of resource processing, and coordinates message exchange between agents;
- the resource collection agent follows the URLs specified by the coordination agent, retrieves metadata, and transmits it to the semantic evaluation agent for further processing;
- the semantic evaluation agent uses a relevance assessment mechanism based on the search descriptor model, which allows determining the feasibility of further expanding the search from the current web resource, taking into account its thematic relevance;
- the indexing agent ensures the automated generation of a structured description of the found

document in the form of metadata in accordance with the selected representation model.

The architecture of the multi-agent system for semantic search of educational content is shown in Fig. 1. The system operates in a distributed information environment that includes a set of software agents, a metadata database, a knowledge base containing an ontological model of the subject domain, as well as an external web environment that serves as a source of information resources.

The interaction of the system components is organized such that the coordination agent initiates the search process, forms initial query descriptors, and distributes tasks among the resource collection agents. Resource collection agents crawl web resources and transmit the obtained metadata to semantic evaluation agents. After semantic analysis is performed, the results are transmitted to the indexing agent, which generates a structured description of the documents and stores it in the database. The ontological model of the subject domain is used by semantic evaluation agents to determine the thematic relevance of resources to the query. This architecture ensures distributed processing of information resources, coordination of agent interactions, and the ability to scale the system through the parallel execution of search procedures.

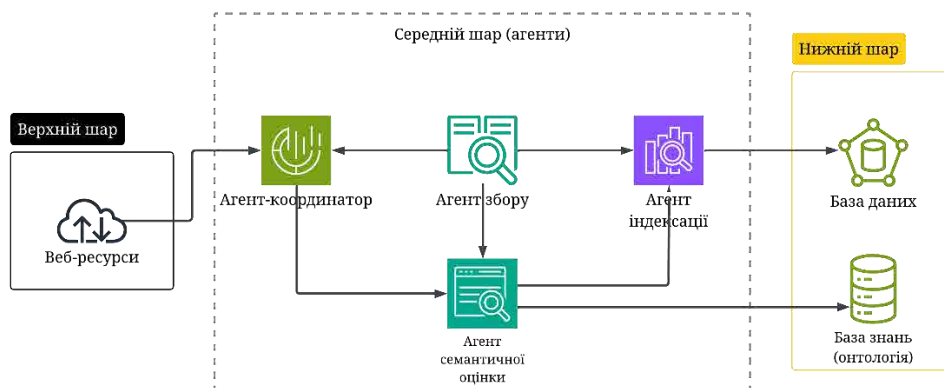


Fig. 1. Architecture of a multi-agent system for semantic search of educational content

The initial stage of the search process involves identifying web resources containing educational materials, which is carried out by a resource collection agent using automated hypertext space crawling technologies. Automated crawling refers to the process of retrieving web resources from the Internet for the purpose of their subsequent indexing and use in a search engine [7]. The main goal of this stage is the rapid formation of a set of relevant web resources along

with hyperlinks reflecting their interrelationships. When a URL is obtained that defines an entry point into the search space, the agent downloads the corresponding web resource, extracts the outgoing hyperlinks from it, and adds them to the queue for further processing [8]. The process continues until the queue of resources has been fully processed.

The key factor in the efficiency of a resource-gathering agent is the strategy for processing the

hyperlinks accumulated in the queue. In practice, two basic approaches are used: sequential traversal without prioritization and guided selection based on an evaluation function. In the case of sequential traversal, the next resource is selected without considering its content characteristics. In contrast, the guided approach involves the use of evaluation criteria that allow determining the appropriateness of moving to a specific resource. Such algorithms can be based on both additional domain-specific knowledge and formal characteristics of the link structure.

The algorithm is based on a multi-agent organization of the process and ensures distributed processing of the hypertext space, taking into account thematic focus. The search for data sources is carried

out by a set of agents distributed throughout the environment, which function as a coordinated system thanks to mechanisms of inter-agent interaction and coordination [9]. To ensure controlled scalability, three complementary mechanisms have been implemented. First, a multi-level duplicate removal system has been introduced, which performs filtering at the resource level, at the level of URLs stored in memory, and between different agents. Second, the procedure for verifying and adding URLs to the shared storage structure has been synchronized, which prevents the duplicate inclusion of identical links. Third, a limit has been imposed on the number of simultaneously active agents, which helps stabilize RAM consumption. The algorithm structure is shown in Fig. 2.



Fig. 2. Algorithm for creating agents

As a result of the resource collection agent's work, a set of URLs is generated in memory; some of these URLs are thematically relevant to the query, but not all resources directly contain educational materials. Therefore, the next stage involves semantic filtering of the collected resources to identify those that actually contain educational objects in the relevant subject area.

The document representation model defines the metadata structure, the query formation mechanism, and the relevance evaluation procedure.

The selected resources are transferred to the indexing agent, where their structured description is generated. Indexing involves constructing the corresponding metadata and storing the obtained information about educational objects for further use in the search process.

A metadata description unit is considered to be a pair (n_i, v_i) , where n_i – the name of the i -th element of the learning object's content description, and v_i – the

corresponding value of this element. The document's metadata description is then presented as a formula (1)

$$m(d) = \{((n_i, v_i), k_i) | i = 1 \dots l\}, \quad (1)$$

where k_i – weight coefficient of the i -th element of the metadata description, l – number of elements.

The process of determining semantic similarity between a document and a query is interpreted as establishing a mapping between the structured description of the document's metadata, as expressed by formula (1), and the result of evaluating its thematic relevance. In this case, the input data is a set of elements $m(d)$, and the output is a relevance score determined by comparing it with the search descriptor. This approach allows the use of the comparator identification method [10] to formalize the rule for transforming a set of document features into a semantic evaluation result.

Before initializing the search process, a domain ontology is formed that reflects the key concepts of the

relevant domain and their semantic relationships. Based on this, a set of terms relevant to the query topic is determined, which may be represented in the structural elements of a web resource, specifically in titles, keywords, and hyperlinks. The selection of these specific elements is due to the fact that in most electronic resources they have a formalized representation via corresponding markup tags, which simplifies their automated extraction. To determine the thematic affiliation of a web resource, the agent performs a comparison of terms from the title h , set of keywords k and the hyperlink text r with the concepts represented in the domain ontology [11]. The resulting thematic description of the resource is then compared with the query topic in accordance with the mappings (2–4).

$$P_k(k, q) = P_k(k, q_1) \vee P_k(k, q_2) \vee \dots \vee P_k(k, q_l), \quad (2)$$

$$P_h(h, q) = P_h(h, q_1) \vee P_h(h, q_2) \vee \dots \vee P_h(h, q_l), \quad (3)$$

$$P_r(r, q) = P_r(r, q_1) \vee P_r(r, q_2) \vee \dots \vee P_r(r, q_l), \quad (4)$$

where q – subject of the inquiry; h – heading terms; k – set of keywords; r – hyperlink text.

The relevance of a page to a search query can be expressed by a predicate (5)

$$S = P(k, h, r, q) = P_k(k, q_1) \vee P_h(h, q_1) \vee P_r(r, q_1), \quad (5)$$

where S – A page score that takes the values 0 or 1.

In the simplest case, a web resource is considered thematically relevant if at least one term from the title, set of keywords, or hyperlink text belongs to the ontologically defined topic of the query. This model corresponds to a binary membership predicate. At the same time, the structure of the predicate can be extended by introducing weighting coefficients for various structural elements of the document, allowing a transition from a strict logical condition to a weighted relevance assessment.

The application of the comparator identification method formalizes the thematic matching procedure within a multi-agent system. Combined with algebraic-logical models [12], this creates a foundation for accounting for the semantic structure of educational objects during their automated selection and indexing.

The Kotlin programming language was chosen for implementing the multi-agent platform due to its support for asynchronous programming and full compatibility with the JVM ecosystem [13]. This choice ensures the solution's scalability and the ability to utilize existing library infrastructure for building distributed systems.

A key factor is the use of coroutines [14], which implement a model of structured concurrency and allow for non-blocking interaction between agents. Unlike the classical thread model, coroutines have lower overhead for execution management and ensure efficient coordination of a large number of simultaneously active agents. This is particularly important for a multi-agent system, where parallel processing and asynchronous message exchange directly impact performance and scalability.

The inter-agent interaction mechanism is implemented using the Kotlin SharedFlow tool [15], which provides a non-blocking model for broadcast event distribution. This approach ensures asynchronous event delivery to dynamically created agents and supports their independent lifecycle. This allows the system to be scaled by adding or removing agents without modifying the communication mechanism. The chosen messaging model minimizes the overhead of synchronization and contributes to the system's fault tolerance in a distributed environment.

The “Factory Method” pattern is used for unified agent creation. Its use allows encapsulating the agent initialization process and ensuring system extensibility without modifying the platform's base code. Adding a new agent type boils down to implementing its behavioral logic and defining the conditions under which it can accept tasks for execution. This approach ensures the scalability of the architecture and simplifies system configuration when expanding functionality.

A mechanism for configurational parameterization of agents has been implemented, allowing their characteristics to be dynamically defined, including the identifier, number of instances, behavioral logic, and additional execution parameters. Agent configuration is performed by loading the corresponding behavioral classes during program execution using the JVM reflection mechanism, which enables the extension of system functionality without recompiling the code and facilitates the flexible configuration of the multi-agent platform according to the requirements of a specific application scenario.

To implement the mechanism for verifying an agent's ability to execute a task, the “Chain of Responsibility” pattern [16] is used. The verification consists of a sequence of handlers, each of which determines whether the agent meets the specified conditions and, if necessary, passes the request to the next element in the chain. This organization allows for

the modular formation of a set of task acceptance criteria without complicating the agent's basic logic.

To organize data access, the "Repository" pattern [17] is used, which abstracts the logic of data storage and retrieval from the agents' business logic.

The system's information environment includes:

- a metadata database;
- a knowledge base containing the domain ontology;
- a set of software agents;
- a coordination module;
- an external web environment containing sources

of educational resources.

The implemented software platform functions as a multi-agent runtime environment with a console-based control interface and ensures the initialization and interaction of agents within the experimental implementation of the proposed architecture.

Research Results

The following conditions were adopted within the scope of the study:

- educational resources have formalized metadata structured in accordance with IEEE LOM [18] or SCORM [19] standards;
- semantic analysis is performed within a single linguistic interpretation of terms without accounting for cross-cultural differences;

– agent interaction occurs in an environment with hardware-homogeneous infrastructure and a stable network with sufficient bandwidth;

– aspects of information security, authentication, and access control are outside the scope of this study and are treated as external to the proposed architecture.

The research was conducted on a computing platform with an 8-core x86-64 architecture processor operating at a clock speed of 2.5–4.8 GHz and 32 GB of RAM, running the Windows 11 operating system in a JVM environment.

Three agent operating modes were tested. The first mode involved the sequential operation of agents A0, A1, A2, and A3, where each agent completed its task before handing control over to the next. The second mode involved the coordinated sequential operation of 50 agents of type A1 under the control of a coordinator. The third mode implemented the coordinated parallel operation of a set of agents: 50 instances of agent A0, and one instance each of agents A1, A2, and A3.

Performance evaluation was conducted by recording the duration of request processing, the number of processed links, and system resource usage metrics (CPU load and RAM) within specified time intervals (15, 30, 45, and 60 minutes). The results are presented in Table 1, which lists the main system performance metrics.

Table 1. Testing results

Characteristics	Consistent work of agents A0, A1, A2, A3				Coordinated consistent work of 50 agents A1				Coordinated parallel work of a set of agents A0(50), A1, A2, A3			
	Database entry								Cache and database entry			
Time (min)	15	30	45	60	15	30	45	60	15	30	45	60
Number of links loaded	15	30	45	60	15	30	45	60	15	30	45	60
Number of links checked	2377	3576	4914	6342	9423	13276	25487	33278	14238	23487	32398	59213
Load	1132	1819	2453	2927	2691	3601	4657	6243	9231	17512	26831	39321
CPU	27	29	32	35	40	43	46	49	49	51	55	57
RAM load (MB)	25	27	28	30	101	103	110	110	439	601	963	965

The results show significant differences in system performance depending on the mode of agent operation. Figure 3 illustrates how the number of links found and verified varies with execution time in sequential mode. Over the course of 60 minutes, the system detected 6,342 links, of which 2,927 were verified, confirming the effectiveness of filtering duplicates and irrelevant resources even during sequential processing. In the second mode, with the coordinated operation of 50 A1-type agents, search efficiency increased significantly. As shown in Fig. 4, in 60 minutes the

system found 33,278 links and verified 6,243. Thus, the number of resources found increased approximately 5.24-fold compared to sequential mode, indicating an increase in search coverage and the completeness of educational resource discovery.

The third mode, featuring coordinated parallel operation of a set of agents A0(50), A1, A2, and A3, demonstrated the highest search performance over the same time interval. Figure 5 shows that, with the same execution time (60 minutes), the system found 59,213 links and verified 39,321 links.

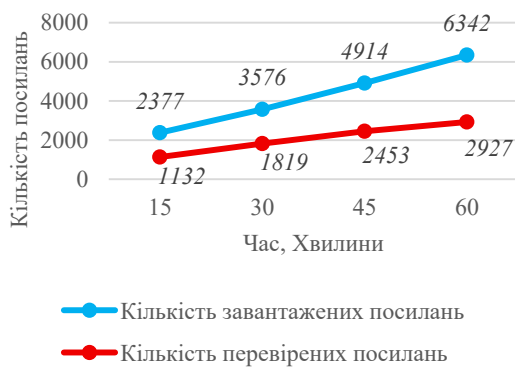


Fig. 3. The relationship between the number of links found (blue curve) and the number of links verified (red curve) as a function of execution time (minutes) in sequential agent mode A0, A1, A2, A3

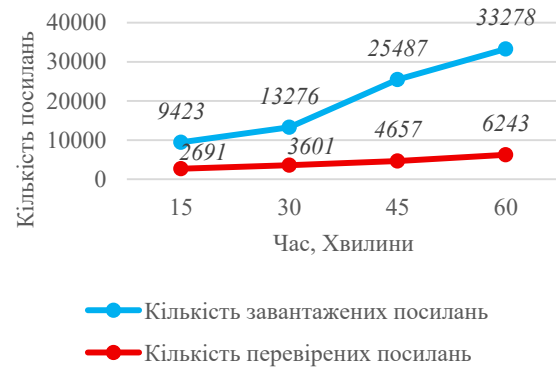


Fig. 4. The relationship between the number of links found (blue curve) and the number of links verified (red curve) as a function of execution time (minutes) in the coordinated sequential operation mode of 50 agents of type A1

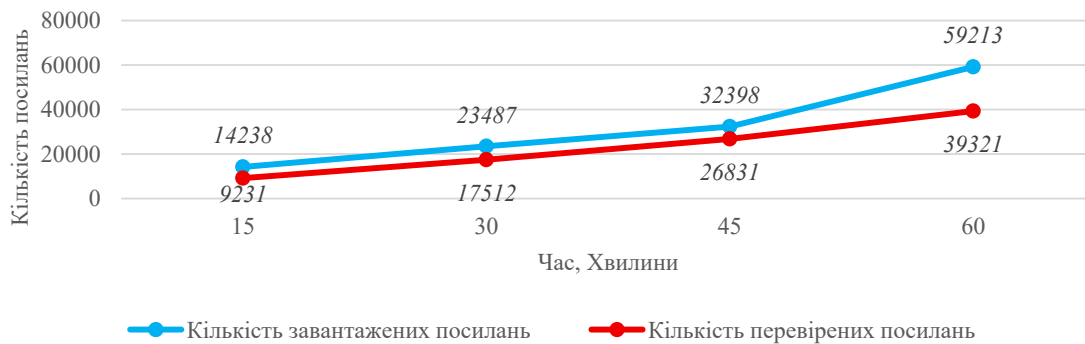


Fig. 5. The relationship between the number of links found (blue curve) and the number of links verified (red curve) as a function of execution time (minutes) in the coordinated parallel operation mode of the set of agents A0(50), A1, A2, A3

A comparison with the first mode reveals a 9.3-fold increase in the number of links found and a 13.4-fold increase in the number of links verified. This indicates an exponential impact of the parallel organization of agent operations on the efficiency of semantic search within the same time constraint.

An analysis of system resource usage metrics revealed significant RAM requirements depending on the operating mode. In serial mode, RAM usage was 25–30 MB with a CPU load of 27–35%. In the second mode, memory usage increased to 101–110 MB with 40–49% CPU utilization due to the simultaneous operation of 50 agents. In the third mode of maximum parallelism, memory usage reached 439–965 MB with 49–57% CPU utilization. It was experimentally established that the limiting factor for performance is the amount of RAM, not the CPU power. The research results demonstrate a significant improvement in the accuracy and completeness of identifying relevant educational resources when using the parallel coordinated operation of a set of agents, with a trade-off in the form of a linear increase in RAM consumption.

Discussion of Results

The obtained results can be explained by the effective use of Kotlin coroutines and optimized algorithms for inter-agent interaction. The experimental data in Table 1 demonstrate a significant difference in performance depending on the mode of agent operation. In the sequential mode, the system found 6,342 links in 60 minutes; in the second mode with 50 agents of type A1, 33,278 links were found; and in the third mode of coordinated parallel operation of a set of agents A0(50), A1, A2, A3, 59,213 links were found. This represents a 9.3-fold increase in efficiency compared to the sequential approach. This increase is attributed to the ability of multiple agents to operate simultaneously without blocking the main execution thread, thanks to the asynchronous nature of Kotlin coroutines and lightweight context-switching mechanisms.

The effectiveness of the semantic search described by equations (2–5) can be explained by the use of the comparator identification method and ontological models of subject domains. The mathematical model for

describing metadata according to formula (1) allows for the consideration of weight coefficients k_i for each element of the metadata description, which improves the accuracy of identifying relevant educational resources. The relevance evaluation predicate $P(k, h, r, q)$, defined by formula (5), provides a comprehensive evaluation of a document based on three criteria: the relevance of keywords, the title, and the links to the search topic. Such a multi-criteria evaluation significantly exceeds the accuracy of traditional one-dimensional methods based on simple word matching.

The solutions developed address the critical issue of the lack of an effective multi-agent platform that combines semantic search for educational content with modern asynchronous programming technologies. The problem of inefficient document ranking algorithms has been solved by introducing a relevance assessment predicate (5) that accounts for semantic relationships between the three components of a document: titles, keywords, and links. The problem of the complexity of creating universal algorithms to account for the diversity of educational standards (LOM, SCORM) is solved through the use of ontological models that provide a flexible representation of knowledge.

Limitations of the study include dependence on a stable network environment with sufficient bandwidth for message exchange between agents. Additionally, educational resource metadata must comply with LOM or SCORM standards, and assumptions about the uniformity of client device technical specifications are required. Experimental conditions involved testing on sets of up to 60 resources; scaling to larger volumes requires further research.

Conclusions and Prospects for Further Development

In the course of this study, its objective was achieved by completing all set tasks, namely:

1. Develop an architectural model of a multi-agent search system. An architectural model of a multi-agent system for semantic search of educational content, oriented toward four types of agents (scanner, evaluator, indexer, coordinator), was developed. The proposed structure ensures the platform's modularity, scalability, and extensibility.

2. Create a semantic search algorithm using the comparator identification method and ontological models of subject domains. A semantic search algorithm has been implemented, based on the use of a formal relevance

evaluation predicate and ontological models of the subject domain. The difference from standard metadata lies in the introduction of weighting coefficients and multi-level comparison (keywords, title, links). Analytical justification is provided by the use of a Boolean predicate function with three independent evaluation components. Experimental studies have confirmed a ninefold increase in semantic search performance compared to standard approaches. The relevance predicate has the structure of a logical conjunction of three independent conditions: the correspondence of the document's keywords to the user's query, the semantic proximity of the title to the query, and the relevance of the outgoing hyperlinks to the query's subject matter. This is explained by the flexibility of the ontological description and the comparator's ability to identify the subject matter of educational resources even with incomplete or diverse descriptors.

3. Implement a software platform using modern technologies with built-in support for asynchronous programming and lightweight parallelism mechanisms. The developed software platform is built using Kotlin coroutines, architectural patterns, and inter-agent communication via SharedFlow. This ensured efficient operation in both sequential and parallel scenarios with hundreds of active agents. The load on system resources remained moderate. RAM consumption ranged from thirty-four megabytes for sequential mode to nine hundred twenty megabytes for parallel execution with fifty agents.

4. Experimentally evaluate the system's performance and resource consumption in different agent operating modes. A series of experimental tests was conducted, confirming the advantages of parallel and coordinated agent operation. In 60 minutes of operation in parallel mode, 9.3 times more links were processed. In the first mode, the system processed 6,196 links, and in the coordinated parallel mode, 57,535 links. This increase is attributed to the resolution of the looping and URL duplication problem through multi-level filtering and flexible synchronization of agent queues. The achieved balance between search completeness and resource consumption is considered optimal for modern medium-scale educational information systems.

Further development of this research could proceed in several directions. **First**, ontological models could be expanded to account for cultural and linguistic differences, which would improve search accuracy in global educational systems. **Second**, it is necessary to

integrate encryption and access control mechanisms into the system architecture, which will ensure its use in critical educational content management systems. **Third**, it is advisable to develop a deterministic mode of system operation with guarantees of sequence and consistency, using distributed systems theory and formal verification models. **Fourth**, the system's adaptation to heterogeneous network conditions should be investigated, taking into account communication delays and failures, which is particularly relevant for educational systems in regions with limited infrastructure.

Conflict of interest

The authors declare that they have no conflicts of interest, including financial, personal, authorial, or any

other nature, that could influence the research or the results published in this article.

Funding

The study was conducted without financial support.

Data availability

The manuscript has no associated data.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technology in the creation of this paper.

References

1. IEEE LTSC (2003), IEEE Standard for Learning Technology – Learning Technology Systems Architecture (LTSA), IEEE Std 1484.1-2003, available at: <https://standards.ieee.org/ieee/1484.1/3397/> (last accessed 28.02.2026)
2. IMS Global Learning Consortium (2003), IMS Learning Resource Meta-data Specification v1.3, Final Specification, available at: <https://www.ledtech.org/standards/learning-resource-metadata> (last accessed 28.02.2026)
3. Wang, X., Li, R., Liu, Y. (2024), "Development and techniques in learner model in adaptive e-learning system: A systematic review", *Computers & Education*, 221, 105114 p. DOI: <https://doi.org/10.1016/j.compedu.2024.105184>
4. Ghembaza, M. I. E.-K., Moussaoui, A., Mansour, H. (2024), "Ontology-driven and semantic web-enabled framework for intelligent e-learning system", *International Journal of Advanced Trends in Computer Science and Engineering*, 13 (6), pp. 259-265, DOI: <https://doi.org/10.30534/ijatcse/2024/071362024>
5. Huettemann, S., Harpprecht, T., Scherer, R., et al. (2025), "Designing ontology-based search systems for research information", *International Journal of Information Management*, Vol. 83, 102901 p. DOI: <https://doi.org/10.1016/j.ijinfomgt.2025.102901>
6. Didden, J. B. H. C., Dang, Q.-V., Adan, I. J. B. F. (2023), "Decentralized learning multi-agent system for online machine shop scheduling problem", *Journal of Manufacturing Systems*, Vol. 67, pp. 338-360. DOI: <https://doi.org/10.1016/j.jmsy.2023.02.004>
7. Cho, J., Garcia-Molina, H. (2002), "Parallel crawlers", *Proceedings of the 11th International Conference on World Wide Web (WWW'02)*, New York, NY, USA, pp. 124-135. DOI: <https://doi.org/10.1145/511446.511464>
8. Boldi, P., Codenotti, B., Santini, M., Vigna, S. (2004), "UbiCrawler: A scalable fully distributed web crawler", *Software: Practice and Experience*, 34 (8), pp. 711–726. DOI: <https://doi.org/10.1002/spe.587>
9. Li, Z. et al. (2025) "Embodied Multi-Agent Systems: A Review", *IEEE/CAA Journal of Automatica Sinica*, 12(6), pp. 1095-1116. DOI: <https://doi.org/10.1109/jas.2025.125552>
10. Karataiev, O., Shubin, I. (2023), "Formal model of multi-agent architecture of a software system based on knowledge interpretation", *Radioelectronic and Computer Systems*, Vol. 4, pp. 53-64, DOI: <https://doi.org/10.32620/reks.2023.4.05>
11. Carbonaro, A. (2008), "Ontology-based video retrieval in a semantic-based learning environment", *Journal of e-Learning and Knowledge Society*, Vol. 4 (3), pp. 203–212, DOI: <https://doi.org/10.20368/1971-8829/290>
12. Shubin, I. (2021), "Development of conjunctive decomposition tools", *CEUR Workshop Proceedings*, 2870, pp. 890-900, available at: <https://ceur-ws.org/Vol-2870/paper67.pdf> (last accessed 28.02.2026)
13. Talib, M. S., Al-Najar, A. A.-M., Hassan, A. H., Talib, Z. S. (2024), "Kotlin Programming Language: A Comprehensive Overview of Definition, Applications, Advantages, and Limitations", *International Journal of Computer Applications Technology and Research*, Vol. 13 (7). DOI: <https://doi.org/10.7753/IJCATR1303.1002>

14. Gupta, N. et al. (2025), "Comparative Analysis of Kotlin, Kotlin Multiplatform, Swift, Flutter, and React Native in Cross-Platform App Development", *International Journal of Innovations in Engineering and Science*, Vol. 10(6), pp. 23-27. DOI: <https://doi.org/10.46335/ijies.2025.10.6.5>
15. Kim, J. (2023), "Derivation of Kotlin Flow API from Observer Pattern", *Journal of KIISE*, 50(9), pp. 821-826. DOI: <https://doi.org/10.5626/jok.2023.50.9.821>
16. Goel, R. (2025), "Design Patterns For Enterprise Application", *International Journal For Multidisciplinary Research*, 7(2). DOI: <https://doi.org/10.36948/ijfmr.2025.v07i02.41268>
17. Arun, N. (2025), "Creational Design Patterns – A Comprehensive Review And Their Role In Object-Oriented Design", *International Journal of Core Engineering & Management*, 8(2), pp. 15–34. DOI: <https://doi.org/10.5281/zenodo.15486190>
18. IEEE LTSC (2020), *IEEE Standard for Learning Object Metadata*, IEEE 1484.12.1-2020, DOI: <https://doi.org/10.1109/IEEESTD.2020.9262118>
19. Martin, H. and Dick, L. (2025) "Distributed Learning for Lasting Impact", *The Clinical Teacher*, 22(6). DOI: <https://doi.org/10.1111/tct.70225>

Received (Надійшла) 02.12.2025

Accepted for publication (Прийнята до друку) 14.03.2026

Publication date (Дата публікації) 30.03.2026

Відомості про авторів / About the Authors

Шапіро Олексій Костянтинович – аспірант, Харківський національний університет радіоелектроніки, кафедра програмної інженерії, Харків, Україна;

Oleksii Shapuro – PhD Student, Kharkiv National University of Radio Electronics, Department of Software Engineering, Kharkiv, Ukraine;

e-mail: oleksii.shapuro@nure.ua

ORCID ID: <https://orcid.org/0009-0005-3539-266X>

Сотник Ігор Сергійович – аспірант, Харківський національний університет радіоелектроніки, кафедра програмної інженерії, Харків, Україна;

Ihor Sotnyk – PhD Student, Kharkiv National University of Radio Electronics, Department of Software Engineering, Kharkiv, Ukraine;

e-mail: igor.sotnyk@nure.ua

ORCID ID: <https://orcid.org/0009-0000-5638-6975>

МЕТОД СЕМАНТИЧНОГО ПОШУКУ НАВЧАЛЬНОГО КОНТЕНТУ НА ОСНОВІ МУЛЬТИАГЕНТНИХ ТЕХНОЛОГІЙ

Цифрова трансформація промисловості супроводжується активним впровадженням нових технологій та швидким оновленням виробничих процесів. Значна частина навчальних матеріалів розміщується у різних інформаційних джерелах, включаючи внутрішні корпоративні системи, відкриті освітні платформи та спеціалізовані веб-ресурси. При цьому такі ресурси часто містять дублікати, надлишкову інформацію та різноманітні метадані, що ускладнює оперативний пошук релевантних навчальних матеріалів. **Предметом** дослідження є метод семантичного пошуку навчального контенту у розподіленому інформаційному середовищі з використанням онтологічних моделей знань. **Мета** роботи – дослідження методу семантичного пошуку навчального контенту у розподіленому інформаційному середовищі на основі мультиагентної організації обробки інформаційних ресурсів та використання онтологічних моделей знань. **Завдання** дослідження: дослідити архітектурну модель мультиагентної системи пошуку; розробити алгоритм семантичного відбору на основі методу компараторної ідентифікації та онтологічної моделі; формалізувати предикат оцінки релевантності з урахуванням вагових коефіцієнтів метаданих; розробити програмну мультиагентну платформу; експериментально оцінити продуктивність і ресурсоемність у різних режимах роботи агентів. **Методи** дослідження: метод мультиагентної організації обробки інформаційних ресурсів з неблокуючим обміном повідомленнями; тривірневу дедуплікацію URL-адрес; онтологічне зіставлення термінів та формалізований предикат оцінки релевантності; експериментальне вимірювання тривалості обробки,

кількості опрацьованих посилань та споживання системних ресурсів. **Результати:** запропоновано модель мультиагентної системи з чотирма типами агентів та алгоритм семантичного пошуку з усуненням зациклення і дублювання посилань; реалізовано програмну платформу на базі Kotlin із використанням корутин та асинхронної взаємодії агентів; експериментально встановлено, що запропонована організація обробки забезпечує більшу продуктивності порівняно з послідовним режимом. **Висновки:** поєднання семантичного пошуку та мультиагентної архітектури дозволяє ефективно організувати процес виявлення та обробки навчального контенту у розподіленому середовищі. Запропонований метод забезпечує узгоджену роботу агентів, усунення дублювання посилань та раціональний баланс між повнотою результатів пошуку і використанням обчислювальних ресурсів.

Ключові слова: семантичний пошук; навчальний контент; програмні агенти; мультиагентні системи; онтології; масштабованість; продуктивність; асинхронність; Kotlin.

Бібліографічні опису / Bibliographic descriptions

Шапиро О.К., Сотник І.С. Метод семантичного пошуку навчального контенту на основі мультиагентних технологій. *Сучасний стан наукових досліджень та технологій в промисловості*. 2026. № 1 (35). С. 127–138. DOI: <https://doi.org/10.30837/2522-9818.2026.1.127>

Shapuro, O., Sotnyk, I. (2026), "A method of semantic search for educational content based on multi-agent technologies", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (35), P. 127–138. DOI: <https://doi.org/10.30837/2522-9818.2026.1.127>

UDC 005.95

DOI: <https://doi.org/10.30837/2522-9818.2026.1.139>

Матківська Х.С., Зачко О.Б., Тригуба А.М.

СИСТЕМА ЦИФРОВОГО HR-МЕНЕДЖМЕНТУ БЕЗПЕКОВИХ СТРУКТУР НА ОСНОВІ Е-КАБІНЕТУ ПРАЦІВНИКА

Предметом є методи та моделі, операційні процеси цифровізації системи HR-менеджменту в безпеко-орієнтованих системах. **Метою** роботи є розробка методів і моделей управління проектами цифровізації операційних процесів системи HR-менеджменту в безпекових структурах для підвищення їх ефективності та адаптивності. У статті вирішуються такі **завдання**: розроблено концептуальну модель цифрової трансформації HR-процесів у безпеко-орієнтованих системах, розроблено модель цифровізації кадрових процесів на основі створення Е-кабінету працівника, розроблено та описано методи автоматизованої генерації електронних контрактів та їх підписання з використанням КЕП. Використовуються такі **методи**: методи моделювання зокрема графічне та структурно-логічне моделювання, що дозволило побудувати комплексну модель цифрової трансформації та централізовану модель управління. Здобуто такі **результати**: розроблено концептуальну модель цифрової трансформації HR-процесів у безпеко-орієнтованих системах, яка формалізує взаємодію кадрових, організаційних та інформаційно-безпекових компонентів. На основі запропонованої моделі обґрунтовано та розроблено метод цифровізації HR-процесів шляхом створення Е-кабінету працівника, що забезпечує централізацію та автоматизацію ключових HR-операцій. У межах реалізації методу розроблено і описано метод автоматизованої генерації, збереження та підписання електронних контрактів із використанням кваліфікованого електронного підпису, що гарантує юридичну значущість та цілісність електронного документообігу. Покращено процес прийняття управлінських рішень завдяки консолідованому доступу до даних у реальному часі та можливостям HR-аналітики, можливість використання розроблених моделей та методів як основу для планування та управління проектами цифрової трансформації в інших силових структурах України, адаптуючи їх до специфіки конкретних підрозділів. **Висновки**: впровадження узгодженої стратегічної політики з розробки та використання інформаційних даних сприятиме модернізації та прискоренню процесів ухвалення рішень у різних аспектах безпеко-орієнтованих систем. Таким чином, зазначене вище підкреслює важливість цифрових технологій у реалізації пріоритетів державних служб, зокрема їх доступність для особового складу, а також підвищення ефективності та впровадження інноваційних рішень.

Ключові слова: цифровізація, автоматизована комунікація, державні структури, HR-менеджмент, безпеко-орієнтовані системи.

Вступ

У сучасних умовах розвитку суспільства та технологічного прогресу особливої актуальності набуває цифрова трансформація систем, діяльність яких пов'язана із забезпеченням безпеки населення, реагування на надзвичайні ситуації та збереженням життєво важливих ресурсів держави. Ефективність роботи безпеко-орієнтованих систем значною мірою залежить від людського фактора – підготовленості, професіоналізму та психологічної стійкості персоналу. Цифрова трансформація дозволить створити інтегровану систему кадрового менеджменту, у якій дані, технології та аналітика підтримують прийняття управлінських рішень, оптимізують підбір, навчання, оцінку й мотивацію кадрів, підвищуючи при цьому рівень оперативності та надійності роботи всієї служби.

Аналіз проблеми й наявних методів

У межах даного дослідження під безпековими структурами (безпеко-орієнтованими системами)

розуміються органи державної влади та підрозділи, основною функцією яких є забезпечення національної безпеки, цивільного захисту та реагування на надзвичайні ситуації, зокрема правоохоронні органи та інші силові структури. Враховуючи стратегічну важливість цих підрозділів, в умовах сучасної цифрової трансформації в безпеко-орієнтованих системах залишається актуальною потреба вирішення наукової проблеми створення інтегрованої системи кадрового менеджменту. Основні труднощі полягають у відсутності уніфікованих моделей інформаційної взаємодії, ефективних алгоритмів обробки даних і надійних каналів цифрової комунікації, які враховували б особливості оперативної-службової діяльності, різний рівень цифрової підготовки персоналу та високі вимоги до забезпечення кібербезпеки. Для вирішення цього завдання необхідно розробити науково обґрунтовані методи та програмно-технічні рішення. Вони повинні інтегрувати існуючі цифрові платформи, мобільні сервіси та засоби автоматизації в єдиний комунікаційний простір для потреб безпекових

структур. Цифровізація є стійким трендом у безпеко-орієнтованих структурах України, що підтверджує впровадження державних цифрових сервісів, таких як «Я-доброволець», ЄАСКО, «Резерв+», Дія.

Табл. 1. Аналіз державних цифрових сервісів і програмних застосунків у сфері кадрового та безпекового управління

Назва програмного застосунку	Цільова аудиторія	Функціонал програмного застосунку
Я-доброволець	Волонтери та соціально активні громадяни, готові допомагати у ліквідації надзвичайних ситуацій, потенційно рятувальники-добровольці. 100+ завантажень Play Market станом на 2025 рік, фактична кількість може бути вищою при офіційному впровадженні	Платформа для організації та координації дій добровольців у надзвичайних ситуаціях; платформа надсилає інформацію про події, маршрути до місця події; включає календар доступності учасника та реєстрацію з перевіркою даних
ЄАСКО (єдина автоматизована система кадрового обліку)	Охоплює загальну чисельність особового. Функціонування системи здійснюється в умовах обмеженого доступу уповноваженим користувачам з дотриманням вимог інформаційної безпеки та захисту персональних даних, що є критично важливим для безпеко-орієнтовної структури	Функціональне призначення полягає у централізованому веденні, зберіганні та актуалізації кадрової інформації про особовий склад. Підтримує виконання внутрішніх кадрових процедур, підготовку службової звітності та аналітичних матеріалів для управлінських рішень, а також забезпечує контроль достовірності й цілісності персональних даних.
Дія	Громадяни України та іноземці з посвідкою на проживання, які користуються державними послугами в цифровому форматі. Понад 23 млн. користувачів станом на 2025 рік	Державний мобільний сервіс «держава в смартфоні»: електронні документи з юридичною силою, доступ до державних послуг онлайн; оформлення прав, техпаспорту, статусів (ветеранів, сімей загиблих)
Резерв+	Військовозобов'язані громадяни України віком 18-60 років, резервісти, призовники, особи, що бажають отримати відтермінування та брати участь у рекрутингових процесах, а також жінки з медичною/військовою спеціальністю. Понад 5 млн. користувачів станом на 2025 рік	Управління військовим обліком: електронний військовий документ, відтермінування за певними категоріями (інваліди, піклування) подача заяв на відстрочку мобілізації, пошук вакансій у ЗСУ, доступ до цифрової військової ідентифікації, оновлення реєстраційних даних.

Проведений аналіз державних цифрових сервісів і програмних застосунків у сфері безпеки та кадрового управління свідчить про активний розвиток цифровізації в безпеко-орієнтованих структурах України. Найвні рішення орієнтовані на вирішення окремих завдань, зокрема облік персоналу, координацію добровольців, надання державних послуг або ведення військового обліку. Водночас аналіз функціональних можливостей цих застосунків показав відсутність єдиного комплексного інструменту, спрямованого на інтегроване управління HR-процесами з урахування специфіки безпеко-орієнтованих систем та необхідності безперервного доступу особового складу до кадрових сервісів та інформаційних даних.

Наукова діяльність С. Д. Бушуєва, Д. А. Бушуєва та Р.Ф. Ярошенка сконцентрована на вдосконаленні методологій управління проектами, програмами та портфелями в умовах складних соціотехнічних систем [2, 3]. Дослідження О.І. Ковальчука та Д.С. Кобилкіна охоплюють вивчення принципів формування та розвитку компетенцій менеджерів проектів, інтеграцію інформаційних технологій

у процесі управління персоналом, а також аналіз впливу поведінкових, емоційних і когнітивних факторів на ефективність виконання інноваційних проєктів [8, 4]. Особливу увагу автори приділяють розробці концепцій управління, орієнтованих на цінності, формуванню життєвих циклів портфелів розвитку організацій і створенню адаптивних моделей управління в умовах невизначеності й цифрових змін. Роботи цих учених закладають основу наукової школи, яка інтегрує системний, когнітивний та поведінковий підходи до управління проектами, спрямовані на посилення стійкості та інноваційного потенціалу організацій. Дослідження Д.В. Боднаря присвячено тому, як глобальні інформаційні потоки, цифрові технології та інтернет-комунікації впливають на стратегічні напрями розвитку безпекових структур [5, 6]. У матеріалах авторів розглядаються основні тренди впровадження сучасних інформаційно-комунікаційних технологій у сферу HR-менеджменту безпекових структур. У дослідженні підкреслюється, що для ефективного управління кадрами в умовах глобального

інформаційного простору необхідно розвивати нові компетенції, посилювати цифрову культуру та впроваджувати гнучкі управлінські моделі, орієнтовані на інновації, оперативну комунікацію й обмін знаннями [1, 7, 11].

Водночас аналіз наукових джерел показує, що питання створення інтегрованих цифрових систем управління персоналом для безпеко-орієнтованих структур досліджено недостатньо. Більшість існуючих рішень орієнтовані на автоматизацію окремих кадрових процедур і не забезпечують комплексної цифрової трансформації HR-процесів.

У зв'язку з цим виникає наукова проблема розроблення моделей та методів цифровізації систем управління персоналом у безпеко-орієнтованих організаціях, які дозволять підвищити ефективність кадрових процесів та забезпечити їх адаптивність до змін зовнішнього середовища. Незважаючи на значну кількість досліджень у сфері цифровізації управління персоналом, питання створення інтегрованих цифрових сервісів для HR-менеджменту безпеко-орієнтованих систем потребує подальшого наукового опрацювання. **Метою** статті є розроблення моделей та методів цифровізації системи HR-менеджменту безпеко-орієнтованих організацій на основі створення електронного кабінету працівника, що дозволяє підвищити ефективність та адаптивність управління кадровими процесами.

Для досягнення поставленої мети необхідно вирішити такі завдання:

- проаналізувати сучасний стан цифровізації кадрових процесів у безпекових структурах;
- розробити концептуальну модель цифрової трансформації HR-процесів;
- формалізувати взаємозв'язок між параметрами управління цифровими HR-проектами;
- розробити метод цифровізації HR-процесів на основі створення електронного кабінету працівника.

Вирішення завдання

Для досягнення поставленої мети у роботі розглянуто теоретичні та прикладні аспекти цифровізації HR-процесів у безпеко-орієнтованих системах. У сучасному світі, де технології стрімко розвиваються, цифрова трансформація набуває особливого значення для систем, що займаються забезпеченням безпеки населення [9]. Результативність роботи безпеко-орієнтованих систем значною мірою

визначається людським фактором, а саме підготовленістю, професіоналізмом та психологічною стійкістю персоналу. У зв'язку з цим управління кадрами в системах, орієнтованих на безпеку, вимагає сучасного підходу, заснованого на принципах цифровізації, автоматизації та аналітичних методів управління. Теоретична частина статті присвячена формуванню наукових засад цифрової трансформації HR-процесів у безпеко-орієнтованих системах. У межах дослідження обґрунтовано доцільність розгляду HR-менеджменту як складної організаційно-управлінської системи, що функціонує в умовах підвищених вимог до надійності, безпеки та оперативності управлінських рішень. На основі аналізу сучасних підходів до цифровізації управління персоналом і проєктного управління в складних організаційних структурах визначено ключові принципи цифрової трансформації HR-процесів у безпекових структурах.

У теоретичній частині розроблено концептуальну модель цифрової трансформації HR-процесів у безпеко-орієнтованих системах, яка базується на інтеграції кадрових, організаційних та інформаційно-безпекових компонентів в єдине цифрове середовище. Запропонована модель відображає логіку переходу від фрагментованого паперового управління персоналом до централізованого цифрового HR-управління з використанням електронним сервісів самообслуговування. Концептуальна модель включає такі взаємопов'язані блоки: кадрові процеси (прийняття на службу, переведення, укладення контрактів, облік персоналу); інформаційні потоки (введення, перевірка, зберігання та актуалізація даних); цифрові інструменти управління (е-кабінет працівника, база даних, модулі електронного документообігу); механізми інформаційної безпеки та контролю доступу; користувачів системи (особовий склад, HR-служба, керівник). Запропонована модель забезпечує формалізацію HR-процесів та створює методологічну основу для розроблення практичного методу цифровізації кадрової діяльності у безпеко-орієнтованих структурах.

У центрі представленої діаграми знаходиться «Цифрова трансформація HR-процесів у безпеко-орієнтованих системах», що виступає ядром та стратегічним напрямом розвитку кадрової роботи у сфері безпеки. Цей елемент відображає процес глибоких змін у системі управління персоналом – від традиційних, паперових чи ручних процедур

до керованих цифрових платформ, які забезпечують рутинні HR-процеси.

Послідовний перехід від рутинних повторювальних кадрових дій до надзвичайно адаптивних інформаційних мереж та автоматизованої комунікації всередині державної служби, що базуються на стандартизації та спрощенні внутрішніх операційних процесів HR-менеджменту. Такий перехід можливий саме завдяки цифровій трансформації. Вона надає переваги, що виходять далеко за межі простого підвищення стійкості до кризових умов, як країні загалом, так і її державним службовцям, компаніям та сектору державного управління. Цифровізація збільшить адаптивність і ступінь впливу держави в мінливому середовищі країни. Цифрова трансформація операційних процесів HR-менеджменту покращить цілісне управління державою з акцентом на формування доступності особистих інформаційних даних особового складу та можливість онлайн-послуг для них. Цифрові інструменти дають державним службам змогу залишатися на зв'язку з особовим складом та отримувати зворотній зв'язок від них. Водночас процес цифрової трансформації супроводжується рядом організаційних і технологічних викликів на рис.1 наведено бар'єри, що перешкоджають розвитку HR-цифровізації в безпекових структурах. Високий рівень невизначеності зовнішнього середовища,

відсутність інноваційної культури, яка сприяє прийняттю ризиків, недостатній бюджет і розмитість корпоративного бачення, брак можливостей для навчання та розвитку персоналу, а також слабкі технологічні і аналітичні навички. До цього додається недостатня підтримка з боку керівництва на незадовільна якість і доступність персональних даних. Ці аспекти відображають поширені виклики, з якими стикаються державні установи в процесі цифровізації. Серед основних перешкод можна виділити обмежені ресурси, інерційність управлінських систем і дефіцит кваліфікованих кадрів.

У межах дослідження ефективність HR-системи розглядається як здатність забезпечувати досягнення організаційних цілей із мінімальними витратами часу та ресурсів. Адаптивність HR-системи визначається як здатність системи кадрового управління швидко реагувати на зміни зовнішнього та внутрішнього середовища, зокрема зміну структури підрозділів, кадрових потреб та оперативних завдань. Таким чином, ефективність характеризує результативність функціонування HR-процесів, тоді як адаптивність відображає їх гнучкість та здатність до швидкої перебудови.

Поряд із зазначеними бар'єрами існують також чинники, що сприяють успішній реалізації цифрової трансформації HR-процесів. Вони також відображені на рис.1.

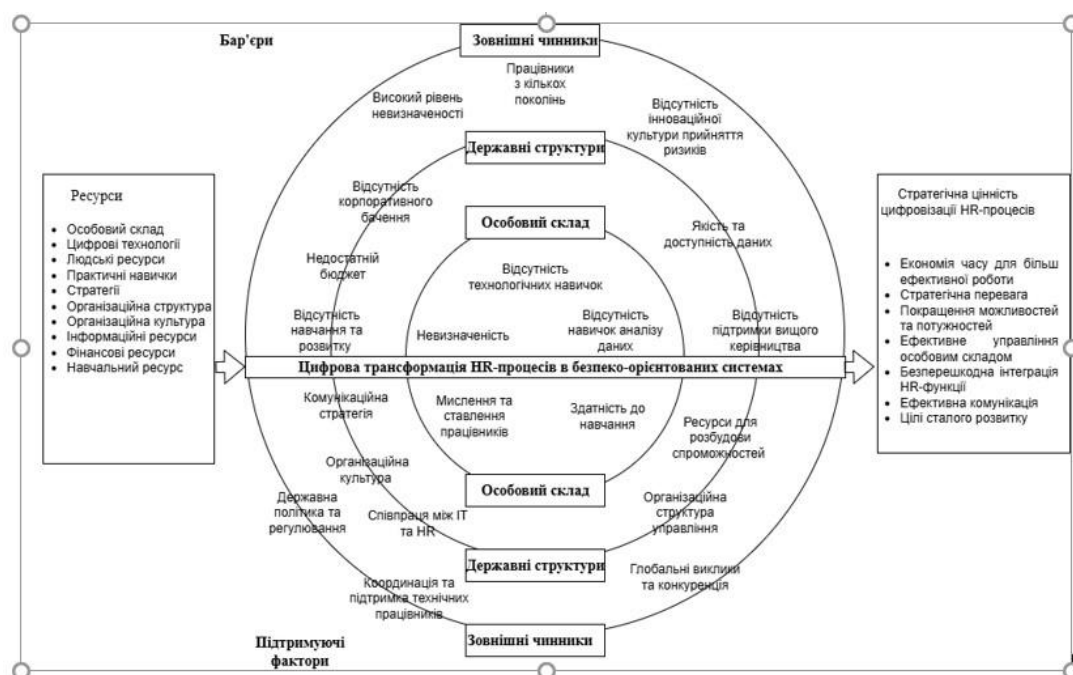


Рис. 1. Модель цифрової трансформації HR-процесів у безпеко-орієнтованих системах

Комунікаційна стратегія, що забезпечує ефективну внутрішню взаємодію між структурами, сприяє узгодженості стратегічних рішень та формуванню єдиного інформаційного простору для безпеко-орієнтованих систем. Організаційна структура, орієнтована на відкритість до змін та на підтримку інновацій, створення хорошого середовища для впровадження цифрових технологій. Допомагає працівникам вдосконалюватись, діяти та пристосовуватись до нових способів роботи. Державна політика та регуляторна база створюють правові умови для впровадження цифрових технологій у роботу з управління персоналом безпеко-орієнтованих систем. Законодавче забезпечення цього процесу, дозволить впровадити цифрові ініціативи з стратегіями розвитку та забезпечення виконання стандартів інформаційної безпеки. Співпраця між IT-підрозділами та безпеко-орієнтованих систем є необхідною передумовою для розроблення інтегрованих технологічних рішень, що дозволяє автоматизувати процеси кадрового менеджменту та підвищити ефективність аналітики кадрових даних [10]. Координація та технічна підтримка персоналу забезпечують стабільне функціонування цифрових систем, оперативне вирішення технічних питань та адаптацію працівників до нових технологічних змін. Розвиток навичок навчання та цифрової грамотності виступає важливою умовою успішної трансформації HR-процесів. Формування компетентностей у сфері використання

цифрових інструментів підвищує професійний рівень працівників і сприяє підвищенню ефективності кадрового управління в цілому. [16] Ці чинники виступають каталізаторами інновацій в безпеко-орієнтованих системах, сприяючи ефективнішому використанню ресурсів та покращенню управлінських процесів.

На рис.2 зображено очікувані результати цифрової трансформації, які мають стратегічне значення для безпеко-орієнтованих систем. Економія часу та підвищення ефективності роботи, досягається завдяки автоматизації рутинних процесів і скороченню адміністративного навантаження на персонал. Стратегічна перевага забезпечується швидким доступом до актуальної інформації, аналітики та можливістю оперативного прийняття рішень. Покращення кадрового потенціалу реалізується через використання цифрових інструментів оцінювання, прогнозування потреб у персоналі та планування кар'єрного розвитку [12–14]. Ефективне управління особовим складом досягається завдяки централізованим системам обліку, контролю та планування ресурсів. Безперешкодна інтеграція HR-функцій сприяє об'єднанню усіх аспектів кадрової роботи. Ефективна комунікація всередині системи дозволяє швидко передавати інформацію між структурами, підвищуючи рівень координації та взаємодії. Досягнення цілей сталого розвитку передбачає підвищення соціальної, економічної та екологічної стійкості безпеко-орієнтованих систем.



Рис. 2. Цифрова трансформація HR-процесів у безпеко-орієнтованих системах

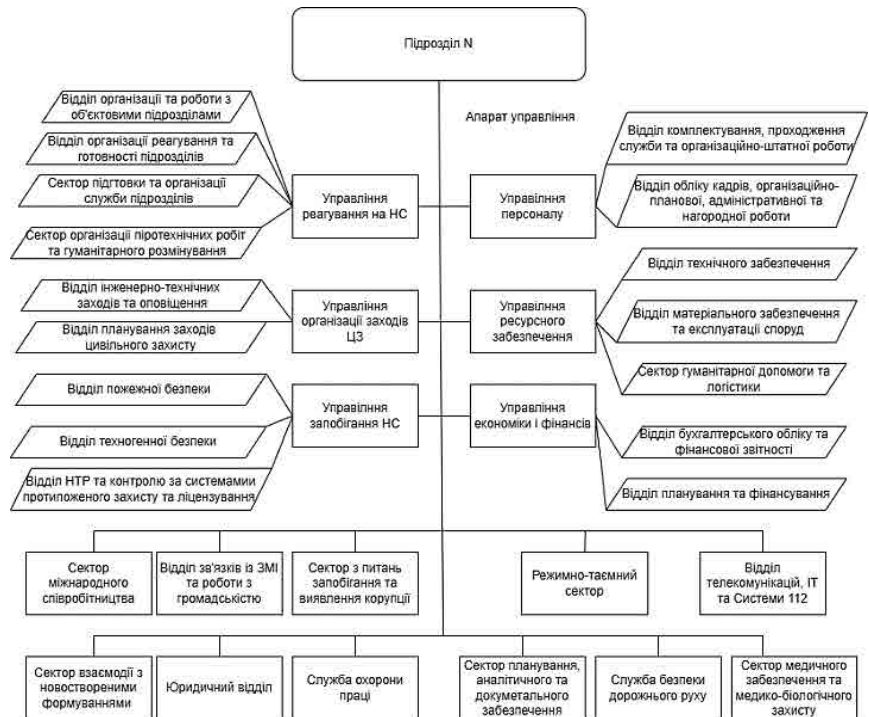
Таким чином, цифрова трансформація HR-процесів безпекових структур-це не лише технічне оновлення, а й стратегічний крок до створення сучасної, адаптивної та стійкої організації, де кожен

співробітник стає частиною єдиної цифрової екосистеми управління.

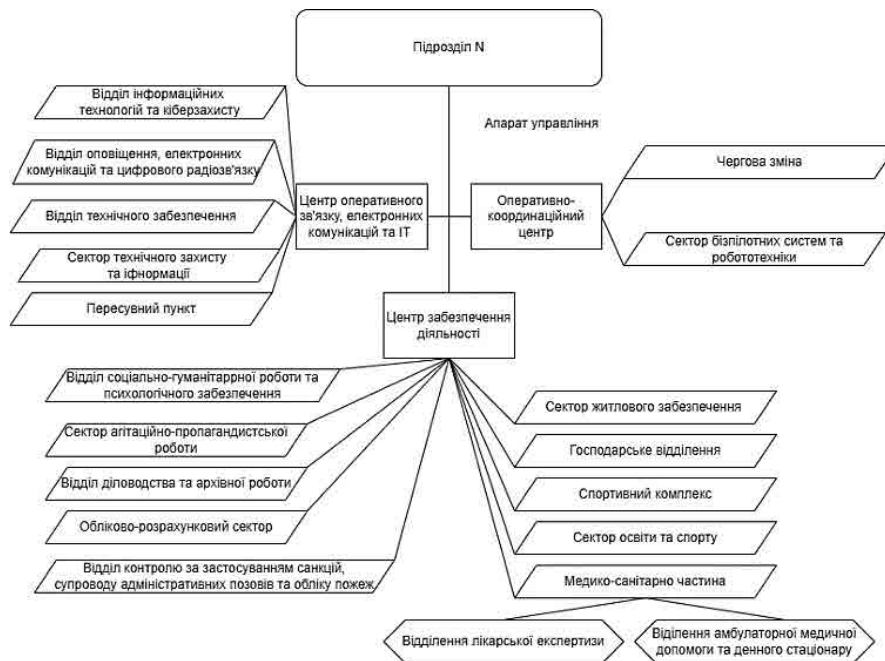
Для практичної реалізації зазначених підходів необхідно проаналізувати ресурси, що забезпечують

цифрову трансформацію HR-процесів у безпеко-орієнтованих системах. З аналізу даних, наведених на рис. 3, проаналізовано організаційну структуру охоплюючи всі категорії безпеко-орієнтованих систем. Така масштабність людського ресурсу

зумовлює високу складність організації та супроводу HR-процесів, зокрема обліку особового складу, укладання та продовження контрактів, управління кадровими переміщеннями, контролю соціальних гарантій і дотримання вимог законодавства.



а)



б)

Рис. 3. Організаційна структура безпекового підрозділу N: а) структура апарату управління підрозділу N; б) структура центру забезпечення діяльності підрозділу N

Особливої актуальності проблема ефективного управління персоналом набуває в умовах обмеженої чисельності кадрових підрозділів N , що підтверджується співвідношенням загальної кількості персоналу HR-менеджерів, як на центральному, так і на територіальному рівнях. У зв'язку з цим цифровізація кадрових процесів є не лише інструментом оптимізації, а необхідною умовою забезпечення ефективного функціонування безпеко-орієнтованих систем.

За відсутності цифрових інструментів виникають труднощі з веденням обліку особового складу, своєчасним формуванням документів, атестацією, управлінням відпустками, нарахуванням пільгової та календарної вислуги років та іншими обов'язковими процедурами. У системі, де домінує оперативний склад і де HR-інформація постійно оновлюється, затримки або помилки можуть мати серйозні наслідки для організаційної готовності та управлінських рішень. На рис. 3 відображено масштабну та багаторівневу організаційну структуру безпеко-орієнтованої системи, що характеризується значною кількістю рівнів управління і складною системою підпорядкованості. Структура охоплює апарат управління, численні управління, відділи, сектору та підпорядковані підрозділи, між якими існують як лінійні, так і функціональні зв'язки. Кожен рівень включає десятки структурних одиниць і сотні посад, а в сукупності – тисячі посад із різним статусом. Така розгалужена ієрархія ускладнює управління персоналом, облік посад, контроль підпорядкування та кадрові переміщення, що підкреслює необхідність системного та автоматизованого підходу до кадрових процесів у безпеко-орієнтованих системах.

У контексті забезпечення інформаційної безпеки (D) безпеко-орієнтованих систем, особливо в умовах інтенсивної цифровізації внутрішніх операційних процесів, ключове значення має структурна модель управління. [15] Ця модель визначає причинно-наслідкові зв'язки та залежності, необхідні для досягнення цільового стану безпеки. Фундаментальне відношення залежності, що регулює імплементацію безпекових інформаційних заходів, описується формулою 1

$$\left. \begin{aligned} E_i &\in \langle t_1, t_2, \dots, t_i; i = 1 \dots n \rangle \\ P_{s_i} &\in \langle c_1, c_2, \dots, c_j; j = 1 \dots n \rangle \\ IT_i &\in \langle q_1, q_2, \dots, q_k; k = 1 \dots n \rangle \end{aligned} \right\} \Rightarrow P_i \Rightarrow D_i, \quad (1)$$

де E_i – середовище i -го проекту цифровізації;
 P_i – проект цифровізації HR-процесу;

IT_i – IT-інфраструктура підрозділу; t_i – час реалізації проекту; c_j – вартість проекту; q_k – якість результату; D_i – ступінь цифровізації HR-системи. Індeksi i, j, k використовуються для відображення різних вимірів проектної тріади управління (час – вартість – якість) у програмі цифровізації HR-процесів.

Запропонована модель цифровізації HR-процесів базується на принципах проектно-орієнтованого управління, що відповідає сучасним стандартам РМВОК та Р2М. У межах моделі HR-процеси розглядаються як сукупність керованих операційних проектів, що мають визначені параметри часу, вартості та якості. Такий підхід дозволяє формалізувати процес цифрової трансформації та забезпечити можливість кількісної оцінки ступеня цифровізації HR-системи.

Цифрова трансформація державного управління є одним із ключових напрямів модернізації публічного сектору та підвищення ефективності діяльності безпекових структур. Особливої актуальності це питання набуває у сфері управління персоналом, оскільки ефективність функціонування безпеко-орієнтованих систем значною мірою залежить від якості кадрового потенціалу, швидкості прийняття управлінських рішень та оперативності обробки інформації.

У сучасних умовах кадрові процеси у безпекових структурах характеризуються значною складністю організаційних зв'язків, великою чисельністю особового складу та багаторівневою системою управління. Традиційні підходи до ведення кадрового обліку, що ґрунтуються на паперовому документообігу або фрагментованих інформаційних системах, не забезпечують достатньої оперативності, прозорості та інтегрованості кадрових процесів.

У межах дослідження ефективність HR-системи розглядається як здатність системи управління персоналом забезпечувати досягнення організаційних цілей із мінімальними витратами ресурсів, часу та управлінських зусиль. Натомість адаптивність HR-системи визначається як здатність системи швидко реагувати на зміни зовнішнього та внутрішнього середовища, зокрема зміну структури підрозділів, кадрових потреб, функціональних завдань та нормативних вимог. Таким чином, ефективність характеризує результативність функціонування системи управління персоналом, тоді як адаптивність відображає її гнучкість та здатність до трансформації в умовах невизначеності.

Одним із перспективних напрямів вирішення зазначених проблем є впровадження цифрових систем HR-менеджменту, які забезпечують інтеграцію кадрових даних, автоматизацію кадрових процедур та підтримку управлінських рішень на основі актуальної інформації.

На основі наданих джерел, висновок щодо причинно-наслідкового зв'язку між безпековим управлінням, проектами експлуатації та досягненням безпеки (описаний формулою 1) візуально репрезентується схематичним зображенням та описом ключових блоків [18, 19].

Практична частина статті присвячена реалізації запропонованих теоретичних положень і концептуальної моделі цифрової трансформації HR-процесів у безпеко-орієнтованих системах у вигляді прикладного методу цифровізації HR-діяльності. Практичну апробацію здійснено шляхом розроблення та впровадження методу цифровізації HR-процесів на основі створення е-кабінету працівника, який функціонує як інтегроване цифрове середовище управління персоналом. У межах практичної частини реалізовано поетапну автоматизацію ключових HR-процедур, зокрема формування цифрової організаційної структури, централізований облік особового складу, управління кадровими переміщеннями та електронний документообіг. Особливу увагу приділено реалізації методу автоматизованої генерації електронних

контрактів, їх збереження, верифікації та підписання з використанням кваліфікованого електронного підпису. Практична частина також включає реалізацію механізмів автентифікації користувачів, розмежування прав доступу відповідно до організаційно-організаційної структури, інтеграцію реляційної бази даних та вебінтерфейсу, а також застосування інструментів адміністрування і контролю цілісності кадрових даних.

У межах дослідження розроблено метод цифровізації кадрових процесів у безпеко-орієнтованих системах, реалізований на основі створення Е-кабінету працівника. Запропонований метод спрямований на формалізацію, автоматизацію та централізацію HR-процесів шляхом перенесення ключових кадрових операцій у захищене вебсередовище з використанням сучасних вебтехнологій, серверної логіки та реляційної даних. Метод цифровізації шляхом створення е-кабінету працівника ґрунтується на централізованому зберіганні кадрових даних у єдиному цифровому середовищі, поетапній автоматизації ключових кадрових процедур, розмежуванні прав доступу відповідно до організаційної та організаційної структури, інтеграції електронного документообігу з механізмами електронного підпису, а також забезпеченості цілісності та достовірності кадрової інформації впродовж усього життєвого циклу HR-процесів.



Рис.4. Схематичне зображення та опис ключових блоків

Запропонований метод цифровізації HR-процесів на основі створення Е-кабінету працівника зображено на рис. 5 послідовними етапами, що охоплюють формування цифрової організаційної структури, автентифікацію та розмежування прав доступу користувачів, введення й адміністрування HR-даних

через вебінтерфейс, автоматизовану генерацію та централізоване збереження електронних контрактів, а також їх верифікацію і підписання з використанням кваліфікованого електронного підпису, що забезпечує юридичну значущість і контрольованість HR-процесів у безпеко-орієнтованому цифровому середовищі.



Рис. 5. Етапи реалізації методу е-кабінету працівника

У рамках дослідження розроблено структуру бази даних інформаційної системи електронного кадрового документообігу, реалізовану на основі PostgreSQL. Для адміністрування та візуалізації структури бази даних використовується графічний інструмент DBaver рис. 6. Результати включають детально викладену архітектуру взаємодії користувача із вебінтерфейсом, механізми перевірки введених даних, а також технологічні процедури автоматичної генерації контрактів. У рамках дослідження був запропонований підхід до збереження створених електронних документів у базі даних, їх подальшого адміністрування, а також підтвердження з боку адміністратора та користувача. Проведений аналіз довів ефективність розробленого методу, який гарантує цілісність та достовірність, зручність обробки електронних контрактів, тим самим сприяючи автоматизації та оптимізації HR-процесів в електронному документообігу в цифровому середовищі.

Після цього реалізовано процес авторизації за допомогою мови програмування PHP з різними способами входу. У базі даних створено таблиці користувачів, які відповідають встановленій організаційній структурі. Крім того, було належним чином налаштовано зв'язки між цими таблицями, що забезпечує коректну інтеграцію даних та взаємодію між об'єктами системи. Розроблено організаційну структуру підрозділу N, яка дозволяє призначати посадових осіб та здійснювати їх переміщення як всередині підрозділу N, так і за його межами. Процес введення даних про переміщення проводиться через веб-інтерфейс з подальшою синхронізацією з базою даних.

Виведено таблицю із переліком особового складу із можливістю у Web вносити дані працівника.

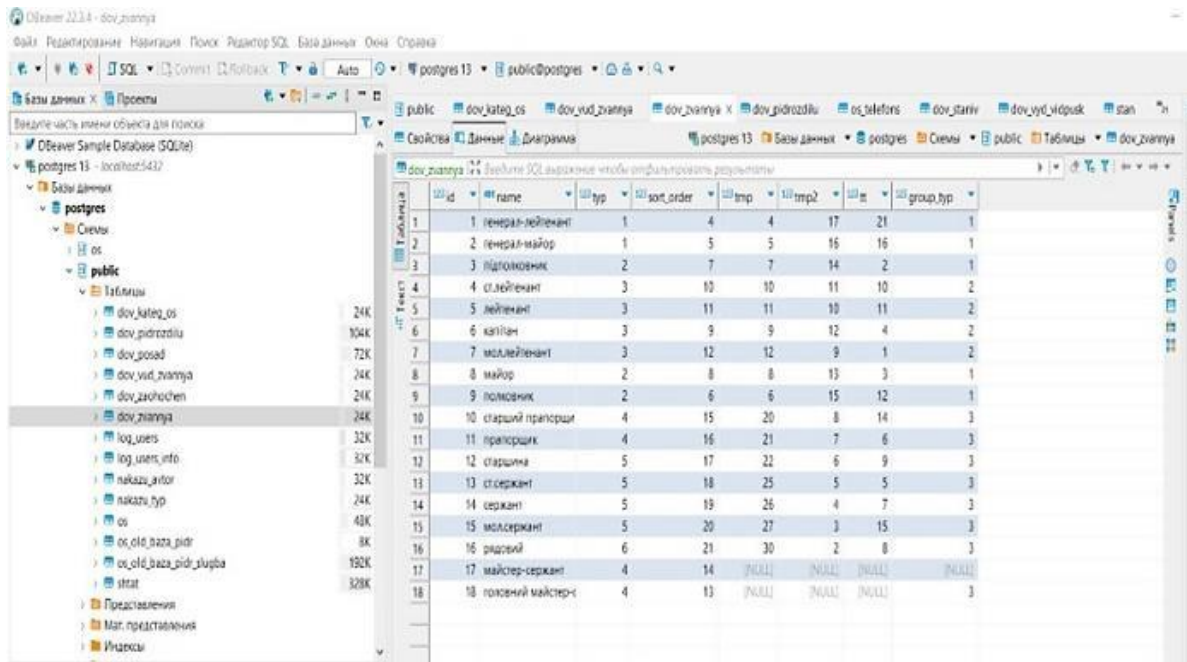
Налаштовано додавання фотографії. Фотографії зберігаються на сервері, у разі редагування фотографії зберігання здійснюється у файльовій системі сервера з унікальним ідентифікатором файлу. Також розроблено перелік та підпорядкованість структурних підрозділів у структурі «дерева» із можливістю редагувати та додавати у Web сторінці із подальшою синхронізацією з базою даних [17].

Схема ілюструє загальну логіку функціонування системи: від подання заявки через вебсайт до завершального етапу управління та верифікації електронних документів. Запропонований підхід забезпечує автоматизацію документообігу, підвищує ефективність обробки інформації та сприяє створенню зручного й інтуїтивного середовища для користувачів інформаційної системи.

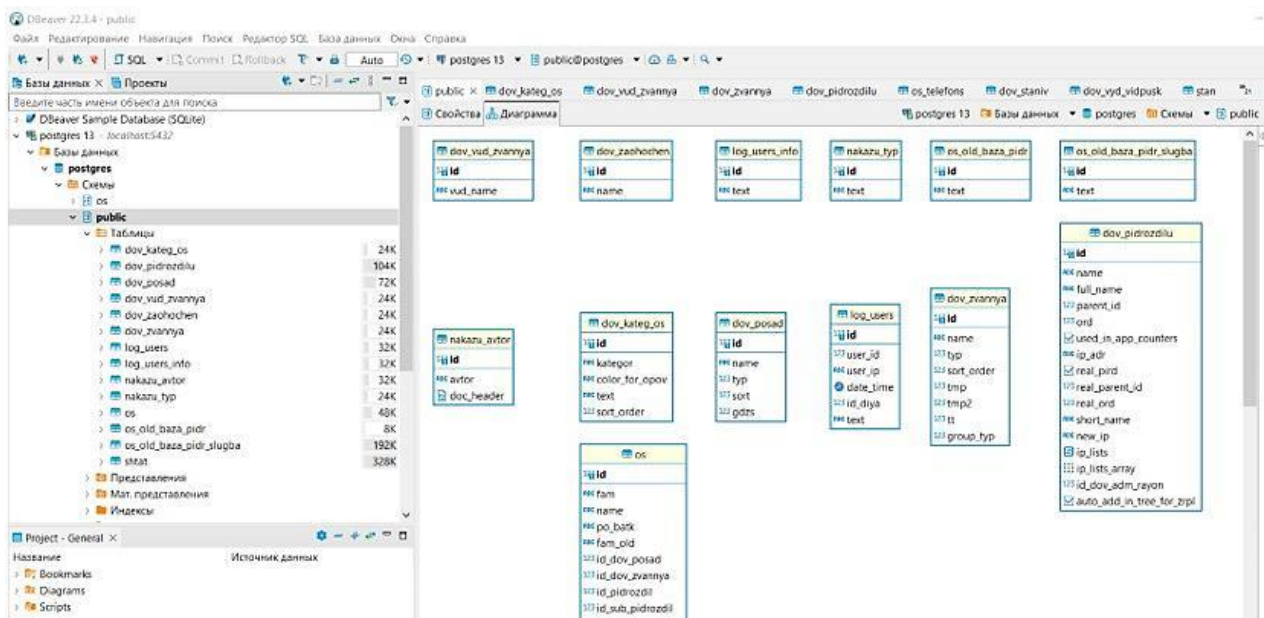
Процес розпочинається з того, що користувач, який бажає укласти контракт або є працівником кадрової служби, відкриває вебзастосунок у браузері. Це може бути захищений вебпортал, розміщений на локальному сервері або в інфраструктурі мережі органів цивільного захисту. Запит обробляється сервером Apache HTTP Server, який перенаправляє його до PHP-інтерпретатора. Система автентифікації для забезпечення безпеки даних може вимагати входу до особистого кабінету. Доступ до форми контракту дозволено лише авторизованим користувачам. Процес переходить у відображення форми подання заявки. На цьому етапі користувач отримує доступ до вебформи, призначеної для введення персональних і професійних даних, які потрібні для створення контракту. Форма розроблена за допомогою HTML, оформлена за допомогою CSS, а PHP використовується для динамічної генерації елементів, як-от завантаження переліку посад із бази даних.

Типові поля: вибір типу контракту (про проходження служби цивільного захисту; про навчання (проходження служби цивільного захисту); про перебування у резерві служби цивільного захисту); прізвище, ім'я,

по батькові; дата народження; спеціальне звання; найменування посади; адреса місця проживання; паспортні дані; вибір терміну контракту (від 1–5 років; на термін навчання; на 5 років).



а)



б)

Рис. 6. Графічний інструмент DBeaver для роботи з базою даних PostgreSQL: а) взаємодія бази даних; б) база даних таблиці користувачів відповідно до організаційної структури

Матківська Христина Степанівна

Підприємство: [Вибрати]

Сімейний стан: Одружений/Заміжня

Прізвище: Гончарук

Ім'я: Сергій

По батькові: Петрович

Телефон: 0988502954

Дата народження: 11.12.1985

Сімейний стан: Неодружений/Незаміжня

Звання: Капітан

Записок відпустки: 23

Адреса: Проспект Червоної Калини 70

Освіта: повна вища, 2007

Паспортні дані: КС 959345 від 12.12.2000

Новий пароль: [Залиште пустим, щоб не змінити]

Зберегти дані

Зберегти

а)

Електронний кабінет працівника

Редагувати ціну Редагувати людей Редагувати посади Звіт

Матківський Володимир

Підрозділ N

1 Державний пожежно-рятувальний загін

1 Державна пожежно-рятувальна частина

3 Державна пожежно-рятувальна частина

61 Державна пожежно-рятувальна частина

Управління персоналу

Відділ обліку кадрів, організаційно-планової, адміністративної та господарської роботи

Відділ комплектування, проходження служби та організаційно-штатної роботи

Організаційна структура підрозділу N

Матківська Христина
Функція

Матківський Володимир
Функція

Додати до списку

б)

Рис. 7. Веб-інтерфейс Електронного кабінету працівника безпекового підрозділу N: а) е-кабінет з особистою інформацією та можливістю редагувати та додавати нових працівників; б) список організаційної структури в базі даних

Форма також може включати елементи для підтвердження згоди з умовами, додавання електронного підпису або завантаження додаткових документів, таких як скани паспорта тощо. Валідація введених даних (РНР). У РНР-скрипті, який обробляє дані, отримані з форми, реалізується попередня перевірка (валідація) введеної інформації перед виконанням будь-яких операцій із базою даних або створенням документів. У разі виявлення помилок користувачу буде виведено повідомлення з детальним описом проблеми та рекомендації щодо її усунення. Наступним етапом є підключення до бази даних. Після перевірки даних скрипт розпочинає процес підключення до бази даних

Додатково може здійснюватись логування, яке включає історію змін та запис усіх дій користувача для подальшого аналізу або аудиту. Центральним етапом у процесі є створення електронного контракту. Інформація, отримана з форм та бази даних, інтегрується у шаблон договору. Цей процес може бути реалізований двома основними методами або використання шаблону HTML + CSS для перегляду документа безпосередньо в браузері або генерація PDF-файлу за допомогою бібліотек РНР. Після цього перехід до наступного етапу збереження контракту в базі даних. Після формування документ може бути оброблений збереженням у файлової системі сервера у форматах

.pdf або .html, при цьому в базу даних вноситься шлях до цього файлу, а також збереженням безпосередньо в базі даних у вигляді BLOB (бінарного об'єкта) або текстового поля.

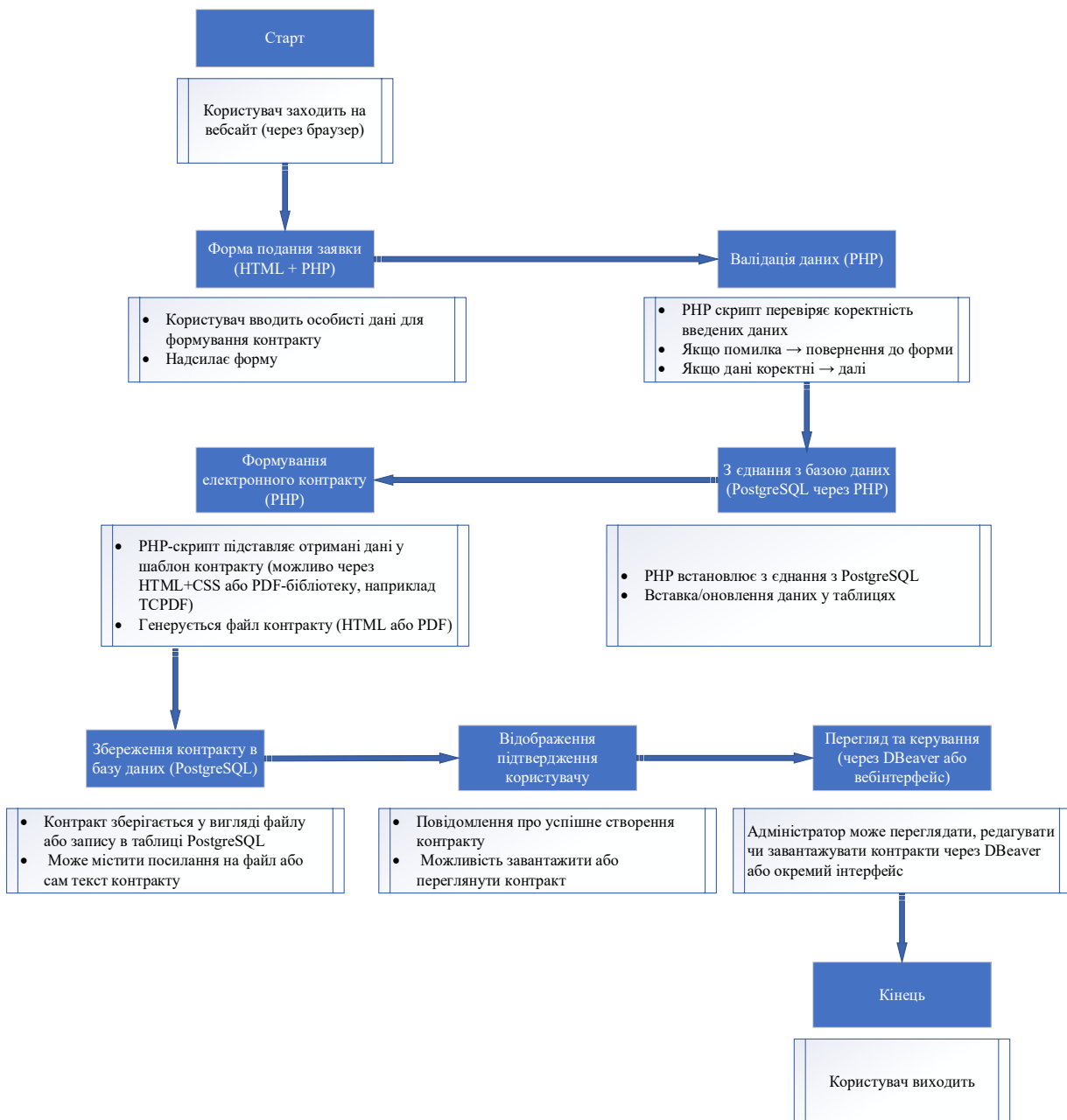


Рис. 8. Схема процесу генерування електронної версії контракту для безпекових структур

У таблиці contracts або documents створюється запис, який містить такі дані:

- Ідентифікатор користувача.
- Тип контракту.
- Статус (наприклад, створено, підписано, відхилено).
- Дата та час створення.
- Посилання на файл або вміст документа.
- Примітки чи коментарі від кадрової служби.

Це забезпечує централізоване збереження всіх документів і дозволяє ефективно відслідковувати їхній стан. Після успішного формування контракту система інформує користувача про завершення процесу та надає можливість перегляду контракту у форматі HTML безпосередньо в браузері; завантаження копії документа у PDF-форматі; збереження файлу локально на пристрої; відправлення контракту на підпис у разі інтеграції

з електронним цифровим підписом (ЕЦП) або кваліфікованим електронним підписом (КЕП).

Окрім того, користувач може обрати перехід до історії створених контрактів або на сторінку, що дозволяє підтверджувати чи редагувати документи. Для адміністраторів чи фахівців управління персоналу безпеко-орієнтованих систем доступна можливість ефективно переглядати, аналізувати та редагувати дані за допомогою інтерфейсу DBeaver – багатофункціонального SQL-клієнта. За необхідності можна також налаштувати окремий вебінтерфейс для адміністративних завдань, що включатиме авторизацію

користувачів і розмежування прав доступу. Користувач отримує доступ до сформованого документа, при цьому база даних оновлюється відповідно. Контракт стає доступним для перегляду, підписання, друку або проведення подальших операцій. За необхідності, систему можна інтегрувати з державними інформаційними сервісами, такими як системи електронного документообігу, державні реєстри чи платформи електронного цифрового підпису. Таке інтегрування значно підвищує юридичну чинність документа та полегшує процес його використання у формальних і правових відносинах.

Табл. 2. Технології у схемі генерування контракту

Етап процесу	Технологія	Опис
1–2	HTML + PHP	Створення форми введення
3	PHP	Перевірка даних
4	PHP + PostgreSQL	З'єднання з БД
5	PHP (TCPDF або шаблони)	Генерація контракту
6	PostgreSQL	Збереження
7	HTML + PHP	Вивід результату
8	DBeaver	Перегляд та управління

Висновки

У результаті проведеного дослідження обґрунтовано необхідність цифрової трансформації кадрових процесів у безпекових структурах, що зумовлено складною організаційною структурою, значною чисельністю персоналу та багаторівневою системою управління. Традиційні підходи до управління персоналом не забезпечують достатньої оперативності, прозорості та ефективності кадрових процедур, що актуалізує потребу у впровадженні цифрових інструментів управління персоналом. Розроблено та представлено такі наукові результати:

1. Запропоновано модель цифрової трансформації управління персоналом у безпекоорієнтованих організаціях, що інтегрує організаційні, інформаційні та аналітичні компоненти цифрового HR-менеджменту.

2. Проаналізовано дисбаланс між масштабом організації та кількістю HR-фахівців на основі спроектованих організаційних структур безпекового підрозділу N.

3. Сформульовано взаємодію між безпековим управлінням, проектами експлуатації та досягненням безпеки, також візуально репрезентується схематичним зображенням та описом ключових блоків.

4. Розроблено архітектуру програмного забезпечення е-кабінету працівника, яка забезпечує автоматизацію основних кадрових процесів.

Напрямки подальших досліджень:

1. Практичними результатами стала концепція стратегії в рамках проекту Е-кабінету працівника. Цей інструмент розроблений для забезпечення специфічних завдань державних структур. Його концепція передбачає створення цифрового середовища самообслуговування працівників завдяки автоматизації рутинних завдань, спрощенню доступу до особистих даних і документів та підвищенню загальної прозорості процесів. Важливим висновком дослідження стало те, що впровадження таких розробок значно підвищує ефективність роботи HR-систем та рівень задоволеності працівників своїм робочим середовищем.

2. Запропоновано метод цифровізації шляхом створення Е-кабінету працівника та деталізовано механізм його реалізації на прикладі автоматизації процесу укладання електронних контрактів з використанням сучасних веб-технологій та засобів електронної ідентифікації. Це створює міцну основу для подальшого впровадження HR-технологій у діяльність безпеко-орієнтованих систем.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів, зокрема фінансового, особистого,

авторського чи будь-якого іншого характеру, який міг би вплинути на дослідження, а також на результати, опубліковані в цій статті.

Доступність даних

Рукопис не має пов'язаних даних.

Фінансування

Дослідження проводилося без фінансової підтримки.

Використання засобів штучного інтелекту

Автори підтверджують, що не використовували технології штучного інтелекту для написання цієї роботи.

References

- Agres, O., Sodoma, R., Binert, O., Samilo, A., Romaniv, V. (2025), "Construction of a management model for financial security of business projects in the context of digitalization", *Eastern-European Journal of Enterprise Technologies*, 5(13 (137)), pp. 29-38. DOI: <https://doi.org/10.15587/1729-4061.2025.340920>
- Bushuev, S., Bushuev, D., Yaroshenko, R. (2018), "Project management in the context of the behavioural economy", *Management of complex systems development*. Issue 33, pp. 22-30. available at: http://nbuv.gov.ua/UJRN/Urss_2018_33_5
- Bushuev, S., Bushueva, N., Lobok, E., Murovansky, G. (2025), "Management of Innovation Projects Based on Artificial Intelligence Applications in a Turbulent Environment", *Current state of scientific research and technologies in industry*, 2(32), pp. 168-176. DOI: 10.30837/2522-9818.2025.2.168
- Balash, L., Sodoma, R., Kovalchuk, O. (2023), "Intelligent systems for forming project teams in security-oriented systems", *Current state of scientific research and technology in industry*, 2(24), pp. 5-15. DOI: 10.30837/ITSSI.2023.24.005
- Bondar, D. (2024), "The system of entities using information technologies in the field of life safety: administrative, legal and economic principles of effective functioning", *Baltic Journal of Economic Studies*. Vol. 10. № 5, pp. 131-136. DOI: <https://doi.org/10.30525/2256-0742/2024-10-5-131-136>
- Bondar, D. (2024), "Legal means of ensuring life safety in the context of the digitalization of public administration and the economy", *Baltic Journal of Economic Studies*. Vol. 10. № 4, pp. 101-105. DOI: 10.30525/2256-0742/2024-10-4-101-105
- Filatov, V., Zolotukhin, O., Kudryavtseva, M. (2025), "Intellectual data analysis in relational information and analytical systems", *Innovative Technologies and Scientific Solutions for Industries*, № 4 (34), pp. 101-111. DOI: <https://doi.org/10.30837/2522-9818.2025.4.10>
- Kovalchuk, O., Kobylkin, D. (2025), "Digitization of IT team management processes in project management: the role of information systems and artificial intelligence", *Innovative Technologies for Project and Program Management*, Collective monograph edited by I. Linde. European University Press ISMA University of Applied Sciences Riga, pp. 134-146. available at: <https://mmp-conf.org/uk/archive>
- Matkivska, H., Zachko, O. (2024), "Digitalization of HR-management in government agencies", ITPM 2024: Proceedings of the 5th International Workshop IT Project Management, May 22, 2024, Bratislava, Slovakia. available at: <https://sci.ldubgd.edu.ua/jspui/handle/123456789/14675>
- Zachko, O., Chalyy O., Kobylkin D. (2020), "Models of technical systems management for the forest fire prevention", *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*. № 5, pp. 129-135. DOI: <https://doi.org/10.33271/nvngu/2020-5/129>
- Bushuyev, S., Ivko, A. (2024), "Values spiral development method in the implementation of digitalization projects in syncretic methodology", *International Journal of Computing*, № 23(2), pp. 177-186. DOI: <https://doi.org/10.47839/ijc.23.4.3535>
- Hayat, K., Hafeez, M., Bilal, K., Shabbir, M. (2022), "Interactive Effects of Organizational Structure and Team Work Quality on Project Success in Project Based Non Profit Organizations", *IRASD Journal of Management*. Volume 4(1), pp. 84-103. DOI: 10.52131/jom.2022.0401.0064
- Piterska, V., Samoilovska, V., Shakhov, V., Tanaka, H. (2023), "Risk-oriented port management in the process of implementing concession projects", *Innovative Technologies and Scientific Solutions for Industries*. V. 2 (24), pp. 200-211. DOI: <https://doi.org/10.30837/ITSSI.2023.24.200>
- Prifti, V. (2022), "Optimizing project management using artificial intelligence", *European journal of formal sciences and engineering*. Vol. 5, No. 1. 29 pp.29-37. DOI: <https://doi.org/10.26417/667thri67>
- Strohmeier, S. (2020), "Digital human resource management: A conceptual clarification", *German Journal of Human Resource Management*. Vol. 34, No. 3. pp. 345-365. DOI: 10.1177/2397002220921131

16. Kobylkin, D., Havrys, A., Rogulia, A., Sodoma, R., Pavuk, I., Avdiyeva, K., Filippova, V. (2025), "Safety-oriented management of protection projects of critical infrastructure objects", *Management Theory and Studies for Rural Business and Infrastructure Development*, 47(4), pp. 537-548. DOI: <https://doi.org/10.15544/mts.2025.42>
17. PMI (2021), *The Standard for Project Management and a Guide to the Project Management Body of Knowledge (PMBOK® Guide)* Seventh Edition. USA. Project Management Institute. 274 p. ISBN: 978-1-62825-664-2
18. PMI (2017), *The Standard for Portfolio Management*. Fourth Edition. USA. Project Management Institute (PMI), 127 p.
19. PMAJ (2017), *P2M Bibelot* (Overview of P2M Third Edition). Japan. Project Management Association of Japan (PMAJ), 2017. 20 p. available at: [https://www.pmaj.or.jp/ENG/p2m/p2m_guide/P2M_Bibelot\(All\)_R3.pdf](https://www.pmaj.or.jp/ENG/p2m/p2m_guide/P2M_Bibelot(All)_R3.pdf)

Received (Надійшла) 08.12.2025

Accepted for publication (Прийнята до друку) 10.02.2026

Publication date (Дата публікації) 30.03.2026

Відомості про авторів / About the Authors

Матківська Христина Степанівна – ад'юнкт денної форми навчання ад'юнктури, докторнатури Львівського державного університету безпеки життєдіяльності; Львів, Україна;

Khrystyna Matkivska – adjunct full-time education Lviv State University of Life Safety; Lviv, Ukraine;

e-mail: matkivskahrystyna@gmail.com

ORCID ID: <https://orcid.org/0009-0007-6044-2387>

Scopus Author ID: 59197093100

Зачко Олег Богданович – заслужений діяч науки і техніки України, доктор технічних наук, професор, Львівський державний університет безпеки життєдіяльності, професор кафедри права та менеджменту у сфері цивільного захисту; Львів, Україна;

Oleh Zachko – honored Worker of Science and Technology of Ukraine, Doctor of Technical Sciences, Professor, Lviv State University of Life Safety, Professor of the Department of Law and Management in the Field of Civil Protection; Lviv, Ukraine;

e-mail: zachko@ukr.net

ORCID ID: <https://orcid.org/0000-0002-3208-9826>

Scopus Author ID: 57194169062

Тригуба Анатолій Миколайович – доктор технічних наук, професор, Львівський національний університет ветеринарної медицини та біотехнології ім. С.З. Гжицького, завідувач кафедри інформаційних технологій; Львів, Україна;

Anatoliy Tryhuba – Doctor of Technical Sciences, Professor, Lviv National University of Veterinary Medicine and Biotechnology named after S.Z. Gzhytsky, Head of the Department of Information Technologies; Lviv, Ukraine;

e-mail: trianamik@gmail.com

ORCID ID: <https://orcid.org/0000-0001-8014-5661>

Scopus Author ID: 57205225539

DIGITAL HR MANAGEMENT SYSTEM FOR SECURITY STRUCTURES BASED ON THE EMPLOYEE'S E-CABINET

The **subject** is methods and models, operational processes of digitalization of HR management systems in security-oriented systems. The **goal** of the work is to develop methods and models for managing projects of digitalization of operational processes of HR management systems in security structures to increase their efficiency and adaptability. The article addresses the following **tasks**: a conceptual model of digital transformation of HR processes in security-oriented systems has been developed, a model of digitization of personnel processes based on the creation of the Rescuer+ e-cabinet has been developed, and algorithms for the automated generation of electronic contracts and their signing using electronic digital signatures have been developed and described. The following **methods** are used: modeling methods, in particular, graphical and structural-logical modeling, which made it possible to build a comprehensive model of digital transformation and a centralized

management model. The following **results** were obtained: a conceptual model of digital transformation of HR processes in security-oriented systems was developed, which formalizes the interaction of personnel, organizational, and information security components. Based on the proposed model, a method for digitizing HR processes was substantiated and developed by creating the Rescuer+ employee e-cabinet, which ensures the centralization and automation of key HR operations. As part of the implementation of the method, an algorithm for the automated generation, storage, and signing of electronic contracts using a qualified electronic signature has been developed and described, which guarantees the legal significance and integrity of electronic document management. The process of managerial decision-making has been improved thanks to consolidated access to real-time data and HR analytics capabilities, the ability to use the developed models and methods as a basis for planning and managing digital transformation projects in other law enforcement agencies in Ukraine, adapting them to the specifics of particular departments. **Conclusions:** The implementation of a coordinated strategic policy for the development and use of information data will contribute to the modernization and acceleration of decision-making processes in various aspects of security-oriented systems. Thus, the above emphasizes the importance of digital technologies in the implementation of public service priorities, in particular their accessibility to personnel, as well as increasing efficiency and implementing innovative solutions.

Keywords: digitalisation, automated communication, government structures, HR management, security-oriented systems.

Бібліографічні описи / Bibliographic descriptions

Матківська Х.С., Зачко О.Б., Тригуба А.М. Система цифрового HR-менеджменту безпекових структур на основі е-кабінету працівника. *Сучасний стан наукових досліджень та технологій в промисловості*. 2026. № 1 (35). С. 139–154. DOI: <https://doi.org/10.30837/2522-9818.2026.1.139>

Matkivska, K., Zachko, O., Tryhuba, A. (2026), "Digital HR management system for security structures based on the employee's e-cabinet", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (35), P. 139–154. DOI: <https://doi.org/10.30837/2522-9818.2026.1.139>

UDC 004.932.2

DOI: <https://doi.org/10.30837/2522-9818.2026.1.155>

Igor Nevlyudov, Olena Ruban, Dmytro Nikitin, Bogdan Misan, Oleksandr Iokhov

INVESTIGATION OF THE DAMPING PROPERTIES OF 3D-PRINTED LINERS WITH CONTROLLED PERFORATION

This study focuses on thermoplastic polyurethane (TPU) liners with controlled internal perforation intended for use in the “residual limb–liner–socket” system of lower-limb prostheses. The influence of material stiffness, hole geometry, and degree of perforation on vibration damping efficiency under impact–dynamic loading is investigated. **The aim of the work** is the experimental determination and optimization of the damping characteristics of TPU liners by varying material stiffness, hole shape, and perforation percentage using the free-decay vibration method. **Tasks:** to analyze the functional role of the liner as a damping element in the prosthetic system; to fabricate a series of 3D-printed TPU specimens with different stiffness levels, hole geometries, and perforation degrees; to implement an impact-based method for measuring free damped vibrations for each specimen; to determine the damping ratio and vibration attenuation percentage; to process experimental data in order to identify optimal combinations of material and geometric parameters; and to establish the relationships between material stiffness, perforation level, hole pattern, and the damping properties of liners. **Results:** a method for evaluating the damping characteristics of 3D-printed TPU liners with controlled internal structures was implemented and experimentally validated. A nonlinear dependence of vibration damping efficiency on the degree of perforation was identified, along with a systematic decrease in damping as TPU stiffness increased. It was shown that hexagonal hole geometry provides a more uniform deformation distribution and slightly higher damping efficiency compared to rhombic and triangular patterns. The obtained relationships enable targeted design of damping liners with an optimal balance between stiffness and vibration attenuation capacity. These findings support the solution of the following practical challenges: reduction of impact loads by decreasing peak dynamic forces transmitted from the prosthetic socket to the soft tissues of the residual limb; pressure redistribution through the formation of more uniform contact stresses at the skin–liner interface; improved user comfort by reducing vibration, pain, and skin irritation during walking; individual optimization through personalized selection of liner geometry and material based on body parameters, activity level, and tissue condition; and engineering design of structurally optimized components for biomedical applications. **Conclusions:** the study experimentally confirms the feasibility of controlling the damping properties of TPU liners by adjusting material stiffness, degree of perforation, and hole geometry. An optimal perforation range was identified for each TPU stiffness level, as well as the advantages of hexagonal perforation in terms of deformation uniformity and vibration damping efficiency.

Key words: TPU liners; 3D printing; material stiffness; prosthetics; impact testing.

Introduction

In lower limb prostheses, the mechanical interaction between the patient’s residual limb and the rigid socket determines the level of comfort, gait stability, and long-term load tolerance. The central element of this interaction is the prosthetic liner (an elastic insert), which serves to cushion, redistribute pressure, and reduce shear stress in soft tissues. During walking, running, or descending stairs, the stump is subjected to impact and cyclic loads which, in the absence of effective damping, are transmitted directly to the tissues, causing pain, irritation, and microtrauma.

Modern liners are primarily made of silicone or polyurethane and have an almost solid structure. This approach provides basic elasticity but does not allow for targeted control of damping properties or adaptation to the patient’s body weight, activity level, and anatomical characteristics. Additionally, solid polymer

liners have limited breathability and do not ensure optimal stress distribution in the contact zone.

The use of 3D printing with thermoplastic polyurethane (TPU) allows for the creation of perforated, cellular structures in which properties are determined not only by the material’s stiffness but also by the product’s geometry. By changing the perforation pattern, stiffness, deformability, and energy dissipation capacity can be adjusted without replacing the material. This opens up the possibility of creating liners with controllable damping properties. The damping capability of TPU liners is determined by two mechanisms:

- viscoelastic losses in the TPU material;
- local deformations of the hole walls and the bridges between them.

As TPU stiffness decreases, internal friction and hysteresis increase, which enhances energy dissipation. At the same time, increasing the degree of perforation reduces the effective stiffness of the structure and

amplifies bending and shear deformations, which also contributes to damping. However, excessive perforation leads to a loss of load-bearing capacity and a reduction in the efficiency of load transfer to the material, which can result in reduced vibration damping efficiency.

The shape of the holes determines the stress distribution and the nature of local deformations. Structures with more isotropic geometry (e.g., hexagonal) provide a more uniform distribution of deformations, whereas structures with sharp angles (triangular) create local stress concentrations, which may limit damping efficiency.

Analysis of Recent Studies and Publications

An analysis of existing approaches to cushioning liners in foot prosthesis designs demonstrates that the interface between the distal residual femur-tibia and the prosthetic socket is a critical area in terms of contact pressure distribution, shock load damping, and the prevention of local soft tissue trauma [1]. Liners traditionally serve as shock absorbers, i.e., they cushion impacts and distribute pressure generated during the foot's contact phase with the support surface, helping to reduce peak reaction forces and minimize the risk of pressure sores, pain, or skin damage during prolonged prosthesis use.

A recent systematic review demonstrates that liners made of elastomeric materials such as silicone or thermoplastic elastomers (TPU) are widely used to provide cushioning properties and user comfort, while their design and material properties influence the ability to evenly distribute peak pressure across the contact surface and reduce dynamic loads on the residual limb. An important indicator of the effectiveness of such interfaces is pressure distribution along the residual thigh-shin: clinical studies using pressure sensors indicate that the use of different liner materials can significantly alter the pressure profile, reducing peak values and increasing distribution uniformity, which, in turn, is associated with higher comfort scores and better tolerance for wearing the prosthesis [2, 3]. Despite the prevalence of commercial gel or polymer liners, traditional materials have a number of limitations: they often have limited capacity for controlled stiffness modulation, their mechanical properties may degrade under multi-cycle loading, and they do not always provide an optimal combination of cushioning with

sufficient stability of the denture suspension, which can lead to increased shear stresses and discomfort.

Current research identifies two directions for improving the shock-absorbing function of liners: enhancing the material base through composite and hybrid materials, and structurally oriented approaches that allow for the control of stiffness and energy absorption distribution via geometric solutions, rather than solely through changes in material composition. Review articles emphasize that the development of new composite materials, including those based on thermoplastic, silicone, or polyurethane matrices, as well as combined systems with phase transitions or reinforcement, can provide better biomechanical adaptation to the contours of the residual limb and distribution of mechanical loads, in particular by reducing local pressure peaks, which is critical in areas of bony prominences [4]. Commercial solutions demonstrate that such materials provide increased flexibility and energy storage capacity, but questions regarding their durability, behavior under multi-cycle compression, and adaptation to various residual limb shapes remain the subject of active research.

Recently, the scientific community has focused on the implementation of bio-inspired metamaterials and cellular structures in the design of liners and sleeves, particularly using additive 3D printing technologies. Such structurally complex designs allow for the modification of local stiffness and energy absorption parameters, tailoring them to the anatomical features of the residual limb.

For example, a study demonstrates the use of auxetic metamaterials integrated into the liner and sleeve, resulting in a significant reduction in peak contact stresses by 60% to 65% compared to traditional silicone-elastomer solutions, as well as an increase in energy absorption capacity due to the specific geometry of the structure's cells [5]. This approach potentially allows not only for comfort and damping but also for the controlled distribution of stresses along the anatomical profile, which is critically important in personalized prosthetics.

However, although such structures with tunable parameters demonstrate significant advantages in laboratory testing and modeling, their clinical validation and standardized pressure assessments (e.g., contact pressure during walking, dynamic shear/friction profiles, tissue adaptation) require further research. Furthermore, reviews indicate a lack of data on the long-term behavior of such materials and structures in the context of cyclic

loading, material aging, and changes in body mass and residual limb shape during prosthesis use [6]. Among the most important areas for future research are the development of standardized methods for assessing comfort and the mechanical properties of the “limb-liner-sleeve” interface, rigorous protocols for multifactorial experimental validation, and a shift toward individualized solutions based on the specific user’s anatomical data.

This transition is key to optimizing the shock absorption of lower limb prosthesis liners and reducing the risk of secondary complications arising from excessive contact pressure or uneven mechanical loads.

Identification of Previously Unresolved Aspects of the Overall Problem.

Purpose of the Work, Objectives

Designing damping TPU liners – in the simplest case, solid elastomeric liners with fixed stiffness – does not pose significant engineering complexity. However, the task becomes significantly more complex when using 3D-printed perforated structures, in which the damping properties are determined by the combined effect of several factors: the stiffness of the TPU material, the shape of the holes, and the degree of perforation.

Unlike solid liners, which do not allow for targeted control of vibration damping, perforated structures make it possible to alter damping properties by modifying the internal geometry. At the same time, this geometric freedom leads to the emergence of nonlinear and ambiguous relationships, under which perforation can either increase or decrease damping efficiency depending on the combination of parameters.

An analysis of scientific publications indicates that, to date, there is no unified method for selecting the degree of perforation and the shape of the holes for a given TPU stiffness that would ensure maximum damping of impact and vibration loads. Most studies focus either on the material properties of elastomers or on the geometry of structures, without a systematic combination of these factors and their experimental verification under impact-dynamic loading conditions. This necessitates further research aimed at developing engineering-based recommendations for the design of TPU liners with controllable damping properties.

The main objective of the study is the experimental determination and optimization of the damping properties of 3D-printed TPU liners by analyzing the

influence of material stiffness, hole geometry, and degree of perforation under impact-dynamic loading conditions. To achieve this goal, the task was set to quantitatively establish the relationships between the liner’s design parameters and its vibration damping efficiency. To solve this problem, the following is required:

- analyze the functional role of the liner in the prosthetic system;
- manufacture a series of 3D-printed samples made of TPU with varying stiffness and perforation parameters;
- implement an impact testing method to measure the free damped oscillations of the samples;
- determine the logarithmic decrement, damping coefficient, and percentage of vibration damping for each combination of parameters;
- process the obtained results to establish optimal values for stiffness, perforation density, and hole geometry.

The final result is the development of engineering-based recommendations for the design of TPU liners with controllable damping properties for use in lower limb prostheses.

Materials and Methods

The shape of the holes in the perforated structure of the shock-absorbing liner determines the nature of the redistribution of mechanical stresses, local deformations, and the propagation path of elastic waves in the polymer material, which directly affects the effectiveness of vibration damping. During an impact load, elastic waves arise in the material, which are repeatedly reflected, refracted, and partially scattered at the “material-void” boundaries, forming a complex pattern of wave interaction within the perforated structure (Fig. 1).

To ensure the controllability of the geometric parameters of perforation and the subsequent analysis of their impact on damping characteristics, it is advisable to use automated software tools with a graphical user interface (HMI) that allow for real-time modification of model parameters and the integration of experimental measurement results. Approaches to building such interfaces and implementing object-oriented automation algorithms are described in [7], which provides the basis for creating a software suite for controlling 3D printing parameters and analyzing vibration processes in liners with controlled perforation.

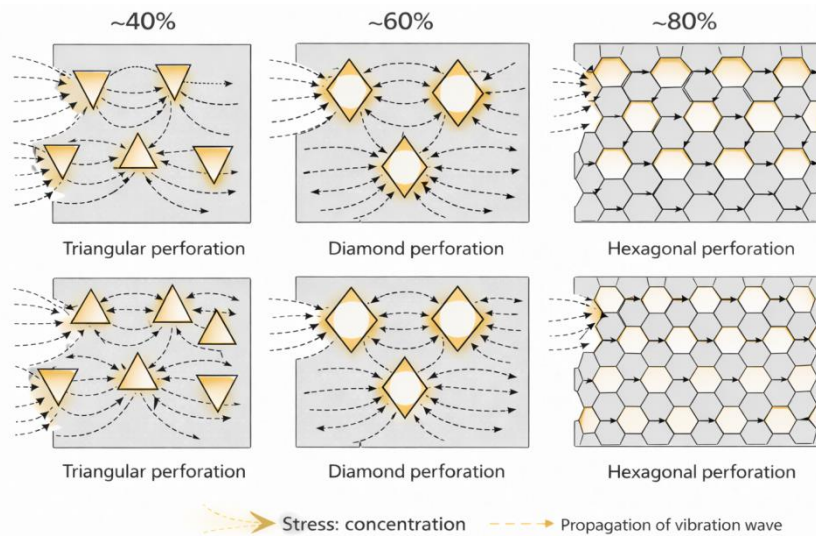


Fig. 1. Scheme of mechanical load distribution within the liner structure

Triangular apertures have sharp corners where stress concentrations occur. In these areas, waves are reflected unevenly, leading to a local increase in stiffness and a restriction of deformations. As a result, part of the mechanical energy is not converted into viscoelastic losses but is returned to the system in the form of reflected vibrations. This reduces the overall damping efficiency, especially for TPUs with increased stiffness [8].

Diamond holes have a more balanced geometry, which reduces local stress concentrations compared to triangular shapes. Wave propagation in such a structure is more uniform, though the presence of corners still leads to partial wave reflection. This provides an intermediate level of vibration damping, representing a compromise between triangular and hexagonal perforation.

Hexagonal holes form a quasi-isotropic cellular structure similar to honeycomb materials. With this geometry, mechanical waves propagate more uniformly in various directions, refract multiple times on the cell

walls, and dissipate more effectively. This promotes intense shear deformations and increased internal viscoelastic energy losses, ensuring the highest vibration damping efficiency among the considered shapes [9].

To evaluate the damping properties, the free damped oscillation method (impact method) was applied. This method is the standard for rapid evaluation of viscoelastic materials and allows the logarithmic decay rate to be determined without the use of complex vibration test bench equipment. The essence of the method lies in recording the damped oscillations of a mechanical system excited by a short-term impulse load, followed by an analysis of the rate of decrease in oscillation amplitude over time [10]. The method is particularly effective for studying elements with pronounced viscoelastic properties, which include polymer materials such as TPU. The operating principle of the test rig for impact testing of samples and the test rig developed according to this scheme are shown in Fig. 2.

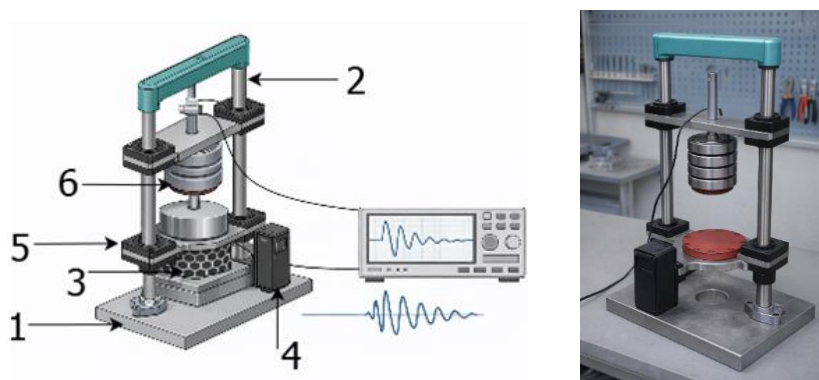


Fig. 2. Operating principle of the model for impact testing of samples: 1 – model base; 2 – guide for the impact hammer; 3 – test sample; 4 – accelerometer sensor block; 5 – impact platform; 6 – impact hammer

The setup consists of a massive, rigid base (weighing 20 kg) on which the test specimen is mounted. An inertial mass (weighing 2.5 kg) is placed on top of the specimen. The system constitutes a classical oscillator with a single degree of freedom.

The oscillations were excited by a calibrated vertical impact on the center of the inertial load, which ensured the reproducibility of the initial experimental conditions. The system's response was recorded by a piezoelectric accelerometer with a sensitivity of 100 mV/g; its signal was fed to an analog-to-digital converter for subsequent digital processing. A sampling rate of 10 kHz provided sufficient temporal resolution for analyzing transient processes and accurately determining damping parameters.

The approach to experimental verification and processing of measurement signals is based on the principles of automated control of dynamic parameters, similar to those presented in [11], where the feasibility of using digital recording and mathematical modeling for evaluating the characteristics of technical systems was confirmed. After impact excitation, the system's motion was approximated by a model of a single-mass mechanical system with viscous damping, for which the amplitude of oscillations decreases exponentially over time. This model allows determining the damping coefficient and the logarithmic decay rate, which are key parameters for evaluating the effectiveness of 3D-printed liners with controlled perforation.

This approach allows for a direct assessment of how effectively the liner reduces oscillations and is convenient for engineering interpretation and comparative analysis of different design variants. To ensure comparability of results, tests of all samples were performed using a standardized algorithm that included the following steps:

- Sample positioning: The TPU liner under study was placed in the test bench's working area, ensuring centering between the rigid support plate and the inertial mass.

- Installation of the measurement system: The accelerometer was rigidly fixed to the load surface, after which the reliability of the mounting was checked to avoid signal distortion.

- Excitation of vibrations: The system was brought out of equilibrium by applying a single vertical impact with controlled energy (impulse load).

- Data acquisition: The time-domain signal of the vibration acceleration $a(t)$ was recorded using an ADC.

- Signal analysis: The amplitude values of successive peaks of the decaying oscillations were identified.

- Calculation of characteristics: Based on the obtained data, the logarithmic decrement, damping coefficient, and vibration damping efficiency index were calculated.

For several successive oscillation peaks $x_1, x_2, x_3, \dots, x_n$, the decay rate is described by the logarithmic decrement [12,13].

The logarithmic decrement over several periods is determined by the formula:

$$\delta = \frac{1}{N} \ln \left(\frac{x_1}{x_{N+1}} \right) \quad (1)$$

where N is the number of periods between the measured peaks.

The ratio of adjacent peaks:

$$r = \frac{x_{n+1}}{x_n} = 1 - \frac{D}{100\%} \quad (2)$$

where D – vibration damping ratio, %.

For small and moderate damping values, the damping coefficient ζ can be determined using an approximate formula:

$$\zeta \approx \frac{\delta}{2\pi} \quad (5)$$

To provide a clear comparison of the effectiveness of various TPU gaskets in operation, a vibration damping coefficient expressed as a percentage has been introduced, which indicates the relative reduction in vibration amplitude over a single cycle:

$$D = \left(1 - \frac{x_{n+1}}{x_n} \right) \cdot 100\% \quad (4)$$

The automated control system (ACS) for filtering noise from the accelerometer when measuring the vibrations of a test specimen will consist of the following blocks, Fig. 3.

The developed automated control system suppresses external noise and extracts the informative component of the accelerometer signal, which corresponds to free damped oscillations following an impact excitation.

At the system input, an acceleration signal $a(t)$ is formed, to which external noise (measurement, mechanical, and electronic) is added, modeled as an additive random process. The resulting signal $a(t)+n(t)$ is fed to the input of the measurement path [14–16].

The overall transfer function of the accelerometer signal processing channel can be expressed as:

$$W_{(s)} = W_s(s) \cdot W_{DAQ}(s) \cdot W_f(s) \quad (5)$$

where $W_s(s)$ – transfer function of an accelerometer, $W_{DAQ}(s)$ – transfer function of the measurement path, $W_f(s)$ – transfer function of the filter.

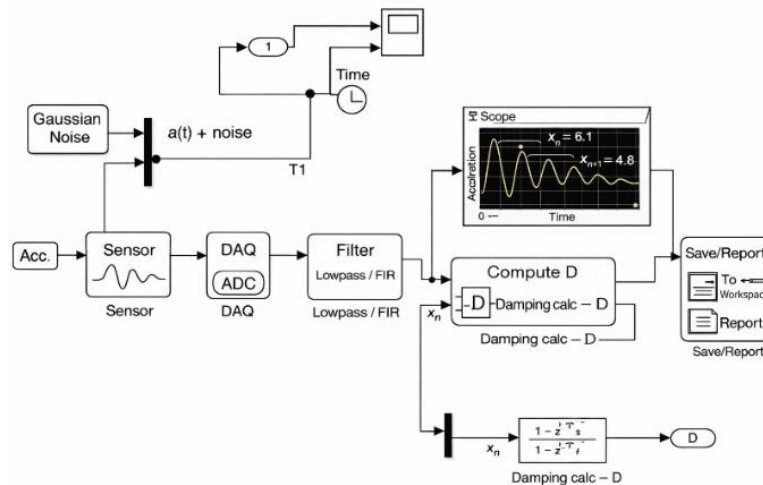


Fig. 3. Accelerometer signal processing

The transfer function of the accelerometer can be described, in a first approximation, as a first-order aperiodic system:

$$W_s(s) = \frac{K_s}{T_s s + 1}, \quad (6)$$

where K_s – accelerometer sensitivity coefficient, T_s – time constant.

The transfer function of a low-pass filter designed to suppress noise in the system will take the following form:

$$W_f(s) = \frac{1}{T_f s + 1} \quad (7)$$

where T_f – the filter time constant, which determines the cutoff frequency.

This transfer function of the digital filter smooths out the high-frequency noise components of the accelerometer's acceleration signal while preserving the system's primary low-frequency dynamics, which are necessary for the accurate determination of the test specimen's vibration damping parameters. A similar approach to numerical modeling and algorithmic processing of measured electrical sensor signals is described in [17], where the effectiveness of using digital filtering methods to improve the accuracy of experimental studies of dynamic systems is demonstrated.

To minimize experimental error, each measurement was repeated at least three times for each sample configuration, followed by statistical averaging of the results. This approach reduces the influence of random signal fluctuations and ensures the reproducibility of the obtained data.

The use of the impulse method (free oscillation method) in this study is due to its high information content, ease of implementation, and the ability to quickly determine damping characteristics without the

need for complex excitation equipment. In combination with digital signal processing, this method provides a reliable estimate of the damping coefficient and the logarithmic decay factor.

First and foremost, the method is easily implemented in laboratory conditions without the use of complex electrodynamic vibration test benches; it also allows for the rapid comparative analysis of a large number of samples within a short period of time. Furthermore, it does not require significant energy consumption or expensive equipment. At the same time, the limitations of the method were taken into account when interpreting the results.

These include dependence on initial conditions: response parameters may vary depending on the impact force and the magnitude of the static preload; the specifics of the mode: the obtained characteristics describe energy dissipation in the mode of free damped oscillations, which may differ from the material's behavior under steady-state harmonic excitation; processing requirements: the accuracy of the results significantly depends on the quality of the signal filtering and peak detection algorithms.

Research Results and Discussion

To investigate the influence of structural and material parameters of printed TPU spacers on vibration damping efficiency, ranges of variation for key factors were determined, taking into account operating conditions and the capabilities of FFF/FDM technology.

Polymer stiffness was selected as the primary material factor.

The study utilized TPU with different Shore A hardness values, specifically:

- 50A, a high-stiffness elastomer characterized by lower deformability and reduced internal energy losses;
- 80A, a medium-hardness elastomer with a balanced ratio of stiffness and damping properties;

– 90A, a low-hardness elastomer characterized by increased deformability and significant viscoelastic energy losses.

For the convenience of conducting the experiment, samples of different hardness will be printed in different colors (Fig. 4).

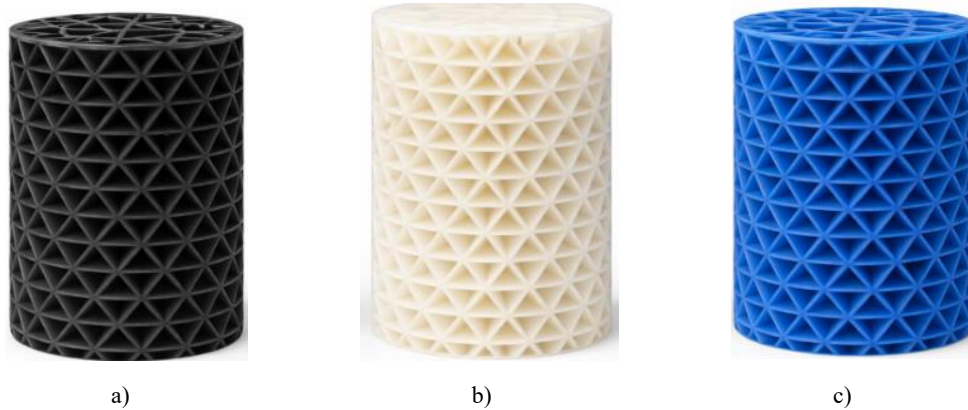


Fig. 4. Images of three products with different hardnesses on the Shore A. scale: a) sample with a hardness of 50A; b) sample with a hardness of 80A; c) sample with a hardness of 90A

The second parameter is the perforation ratio P , which defines the ratio of the area of the holes to the total cross-sectional area. The range under investigation: 40% – preservation of high effective stiffness; 60% – expected optimum damping; 80% – significant reduction in structural stiffness.

The selected range of values allows us to trace the nonlinear relationship between the liner's damping properties and its deformation capacity, and additionally to identify the optimal perforation parameters for materials of varying stiffness. Examples of sample perforation are shown in Fig. 5.

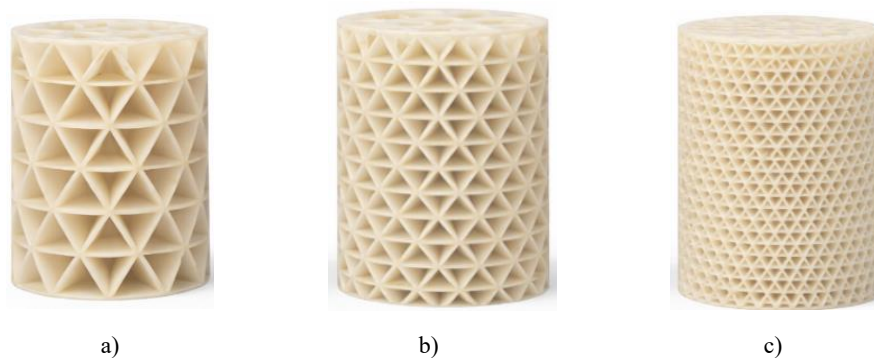


Fig. 6. Examples of different perforation percentages: a) sample with a 40% perforation rate; b) sample with a 60% perforation rate; c) sample with an 80% perforation rate

The third factor is the topology of the cells (perforation shape): Triangular: characterized by local stress concentration; Diamond: an intermediate option with moderate anisotropy; Hexagonal: provides the most uniform distribution of deformations. A comparative analysis of the proposed shapes allows us to assess the sensitivity of the damping properties to changes in the hole contour when other determining parameters

(stiffness and perforation percentage) remain constant. Examples of perforation patterns are shown in Fig. 7.

To eliminate the dimensional factor, all samples were cylindrical in shape (diameter 40 mm, height 60 mm). The experimental design included a sample of 45 specimens to characterize nonlinear effects [18]. The results of the study are presented in Table 1.

The results of the full sample of 45 samples are presented graphically in Fig. 8.

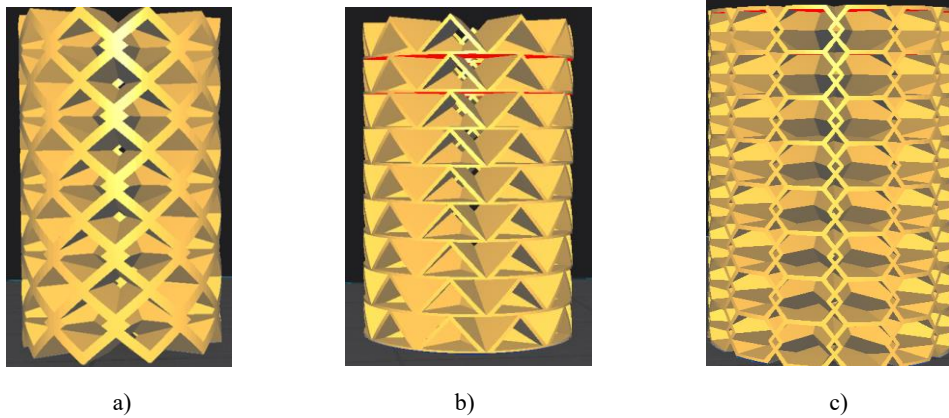
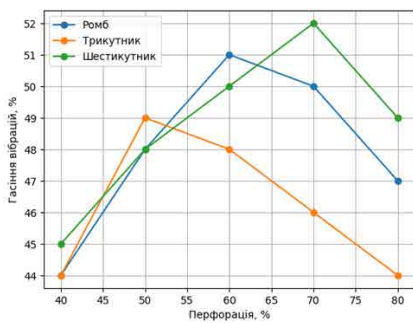


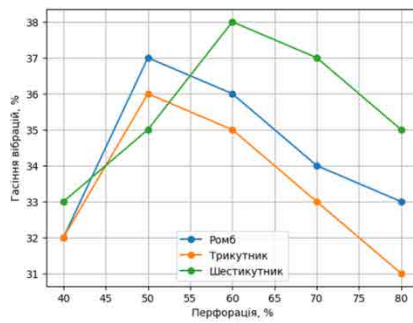
Fig. 7. Examples of different perforation percentages: a) diamond-shaped perforation; b) triangular perforation; c) hexagonal perforation

Table 1. Research results

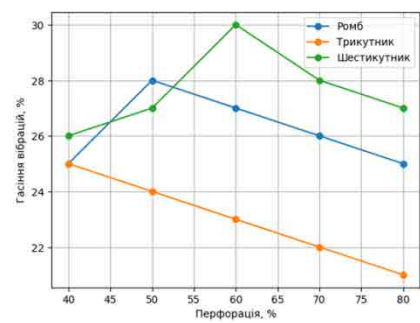
№	TPU	Shape	Perforation rate, %	D, %	№	TPU	Shape	Perforation rate, %	D, %
1	50A	Diamond	40	44	24	80A	Triangular	70	33
2	50A	Diamond	50	48	25	80A	Triangular	80	31
3	50A	Diamond	60	51	26	80A	Hexagon	40	33
4	50A	Diamond	70	50	27	80A	Hexagon	50	35
5	50A	Diamond	80	47	28	80A	Hexagon	60	38
6	50A	Triangular	40	44	29	80A	Hexagon	70	37
7	50A	Triangular	50	49	30	80A	Hexagon	80	35
8	50A	Triangular	60	48	31	90A	Diamond	40	25
9	50A	Triangular	70	46	32	90A	Diamond	50	28
10	50A	Triangular	80	44	33	90A	Diamond	60	27
11	50A	Hexagonal	40	45	34	90A	Diamond	70	26
12	50A	Hexagonal	50	48	35	90A	Diamond	80	25
13	50A	Hexagonal	60	50	36	90A	Triangular	40	25
14	50A	Hexagonal	70	52	37	90A	Triangular	50	24
15	50A	Hexagonal	80	49	38	90A	Triangular	60	23
16	80A	Hexagonal	40	32	39	90A	Triangular	70	22
17	80A	Diamond	50	37	40	90A	Triangular	80	21
18	80A	Diamond	60	36	41	90A	Hexagonal	40	26
19	80A	Diamond	70	34	42	90A	Hexagonal	50	27
20	80A	Diamond	80	33	43	90A	Hexagonal	60	30
21	80A	Diamond	40	32	44	90A	Hexagonal	70	28
22	80A	Triangular	50	36	45	90A	Hexagonal	80	27
23	80A	Triangular	60	35					



a)



b)



c)

Fig. 8. Research results: a) vibration damping of perforated TPU gaskets with a stiffness of 50A; b) vibration damping of perforated TPU gaskets with a stiffness of 80A; c) vibration damping of perforated TPU gaskets with a stiffness of 90A

Analysis of the experimental results showed that the key factor in the effectiveness of vibration damping by TPU gaskets is the material's stiffness. An increase in stiffness from 50A to 90A is accompanied by a steady decrease in the damping level.

Specifically:

- for TPU 50A, the maximum damping efficiency reaches approximately 52%;
- for TPU 80A — about 39%;
- for TPU 90A — does not exceed 31%.

This trend is due to a decrease in internal energy dissipation in stiffer elastomers, as well as an increase in the effective stiffness of the “support–pad–load” system, which limits the deformation of the damping element and, accordingly, reduces the level of energy dissipation.

For all investigated TPU stiffness values, a nonlinear dependence of damping on the degree of perforation was observed. Increasing perforation from 40% to a certain optimum leads to an increase in damping properties, while a further increase in perforation causes them to decrease. At the same time, the position of the damping maximum depends on the material stiffness:

- for TPU 50A, the optimal perforation range is 60–70%;
- for TPU 80A, 50–60%;
- for TPU 90A, 40–60%.

The obtained data indicate that as material stiffness increases, the optimal perforation level shifts toward lower values, since excessive weakening of the structure of stiff samples leads to a loss of contact stability and a decrease in damping capacity.

The perforation geometry has a secondary but reproducible effect on vibration damping. The analysis reveals the following trends:

- triangular holes provide maximum damping at lower perforation levels (40–50%), which is associated with stress concentration at the vertices and local deformation confinement;
- Diamond geometry is optimal at medium perforation levels (approximately 60%);
- the hexagonal structure demonstrates the highest damping values at higher perforation levels (60–70%) due to a more uniform distribution of stresses and deformations in the material.

Differences between the shapes typically do not exceed 2–3%, confirming their corrective rather than determining role compared to material parameters.

An analysis of 45 parameter combinations showed that the effectiveness of vibration damping by TPU gaskets is determined by the combined action of three

factors: material stiffness, degree of perforation, and hole geometry. Changing any one of these factors in isolation does not ensure maximum damping, indicating the need for comprehensive design optimization that takes into account both material and geometric characteristics.

As a result, it was established that the highest level of vibration damping is achieved for soft TPU gaskets with an optimally selected degree of perforation. As material stiffness increases, the optimal perforation values shift toward lower values, while the shape of the holes serves as a tool for fine-tuning the damping properties. The obtained patterns can be used in the design of damping elements for additively manufactured machine and mechanism components.

To identify the patterns of influence of physical-mechanical and geometric parameters on the vibration-damping effectiveness of TPU gaskets manufactured using additive technologies, correlation-regression analysis methods were applied. The key objective of this stage was to establish a quantitative relationship between the design features of the samples and their ability to reduce vibration amplitude. To optimize the research process and reduce the number of required measurements without losing statistical reliability, a multivariate experimental design was used instead of a full search of all possible combinations of factor levels. This approach involved selecting representative points within the studied range, which allowed for the construction of an adequate model based on a limited data sample.

The factorial space of the experiment was formed by the following independent variables:

- TPU material hardness on the Shore A scale (50, 80, 90);
- degree of gasket perforation, % (from 40% to 80% in 10% increments);
- perforation shape (Diamond, triangular, hexagonal);
- amplitude of free-mode vibrations after impact excitation;
- vibration damping percentage D.

The obtained values were used to calculate Pearson correlation coefficients and construct multivariate regression models.

To perform calculations in the IBM SPSS Statistics software environment, all variables were converted to numerical form and coded as follows:

- material hardness (Hardness);
- degree of perforation (Perforation), %;
- perforation shape (Shape): 2 – triangular; 1 – diamond; 3 – hexagonal;

– vibration damping (Damping), %.

To mathematically describe the influence of input parameters on the damping capacity of the gaskets, a multiple linear regression model was selected:

$$D = b_0 + b_1H + b_2P + b_3S + \varepsilon \quad (5)$$

where D – vibration damping percentage; H – material hardness TPU; P – perforation degree, %; S – perforation shape; b_0, b_1, b_2, b_3 – regression coefficients; ε – random model error.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	,957 ^a	,916	,910	2,845	,916	148,467	3	41	,000

a. Predictors: (Constant), Perforation, Shape, Hardness

Fig. 9. Model Summary

An ANOVA table was used to test the hypothesis regarding the statistical significance of the constructed regression model. In all cases, the significance level Sig. < 0.05, which indicates the reliability of the results

A quality check of the resulting model using the Model Summary module demonstrated high adequacy. The coefficient of determination R^2 is 0.916, which means that the constructed model explains over 91.6% of the variation in the experimental data [19, 20]. The adjusted coefficient of determination R^2 also exceeds the threshold of 0.91, confirming the presence of a strong and stable correlation between the selected design parameters of the liners and their damping properties. The results of the quality check are shown in Fig. 9.

obtained and allows us to reject the null hypothesis regarding the absence of an effect of the studied parameters. The results of the ANOVA significance calculations are presented in Fig. 10.

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3605,730	3	1201,910	148,467	,000 ^b
	Residual	331,915	41	8,095		
	Total	3937,644	44			

a. Dependent Variable: Damping

b. Predictors: (Constant), Perforation, Shape, Hardness

Fig. 10. Results of the ANOVA significance test

To rank the input parameters according to their degree of influence on the liner's integral energy dissipation capacity, an analysis of standardized regression coefficients ("β-coefficients") was conducted [21]. A comparison of the absolute values of these indicators allowed us to establish a hierarchy of factor influence:

– material hardness: acts as the dominant factor in the model. The highest absolute value of the β-coefficient indicates that it is the physical and mechanical properties of the polymer matrix that are decisive for vibration damping efficiency.

– degree of perforation: has a significant effect on damping characteristics; however, the nature of this effect is complex and likely nonlinear, reflecting changes in the structural stiffness of the sample.

– cell topology: is characterized by the lowest coefficient value, indicating its auxiliary role. The shape of the holes serves to "fine-tune" (adjust)

the product's properties, ranking significantly lower in importance than material hardness and perforation density. The results of the influence of the liner's perforation parameters are presented as Pearson correlations and shown in Fig. 11.

Correlations

		Damping	Hardness	Shape	Perforation
Pearson Correlation:	Damping	1,000	-,956	,049	,013
	Hardness	-,956	1,000	,000	,000
	Shape	,049	,000	1,000	,000
	Perforation	,013	,000	,000	1,000
Sig. (1-tailed):	Damping	.	,000	,373	,465
	Hardness	,000	.	,500	,500
	Shape	,373	,500	.	,500
	Perforation	,465	,500	,500	.
N	Damping	45	45	45	45
	Hardness	45	45	45	45
	Shape	45	45	45	45
	Perforation	45	45	45	45

Fig. 11. Pearson correlation

Conclusions

The stiffness of the TPU material, the degree of perforation, and the shape of the holes significantly affect the vibration damping efficiency of additively manufactured gaskets. The experimental dependencies obtained show that the damping properties result from the interaction between the viscoelastic characteristics of the material and the geometry of the gasket's internal structure. An excessive increase in material stiffness or the degree of perforation leads to a decrease in damping due to a reduction in deformation capacity and a loss of contact stability in the system.

It has been established that for each stiffness value, there is an optimal degree of perforation that ensures maximum vibration damping. For soft TPU 50A, the optimum lies in the range of 60–70%; for medium-stiffness 80A materials, from 50% to 60%; and for stiff TPU 90A, from 40% to 60%. Thus, as material stiffness increases, the optimal perforation level systematically shifts toward lower values.

The shape of the perforation affects the distribution of stresses and strains in the material and, accordingly, the level of damping; however, its influence is secondary compared to the material parameters. Hexagonal perforation provides a more uniform stress distribution and demonstrates slightly higher damping values; Diamond perforation occupies an intermediate position; and triangular perforation is characterized by lower efficiency due to stress concentration at the vertices of the holes. The difference between the shapes does not exceed 2–3%, which allows us to consider hole geometry as a fine-tuning tool rather than a determining parameter.

Based on the results of the study, the following tasks were accomplished:

- the influence of TPU stiffness on the damping properties of the gaskets was analyzed;
- the dependence of vibration damping on the degree of perforation for various materials was investigated;
- the influence of hole geometry on stress distribution and damping efficiency was evaluated;
- a regression-correlation analysis was performed and a multifactorial model was constructed describing the influence of structural and material parameters on vibration damping.

The constructed regression model showed that over 91% of the variation in damping level is explained by

the combined effect of material stiffness, the degree of perforation, and the shape of the holes. TPU stiffness makes the largest contribution; the degree of perforation has a significant but nonlinear effect, while the shape of the holes plays a moderating role.

Further research should focus on analyzing the damping properties of 3D-printed TPU insoles under various frequency loading conditions and cyclic fatigue typical of real-world walking. It is promising to study the influence of the anisotropy of additively manufactured structures, printing parameters, and external factors (temperature, humidity) on the viscoelastic behavior of the material. Combining experimental results with numerical modeling and optimization methods will allow for the automated selection of perforation parameters tailored to the individual characteristics of the prosthesis user and enhance the effectiveness of personalized shock-absorbing solutions.

Thus, to achieve maximum vibration damping, comprehensive optimization of the TPU insert design is required, taking into account both the material properties and the geometry of the internal structure. The results obtained can be used in the design of shock-absorbing and vibration-damping elements for additively manufactured machine and mechanism components, particularly in systems where effective absorption of dynamic loads is required within limited dimensions and mass.

Conflict of interest

The authors declare that they have no conflicts of interest, including financial, personal, authorial, or any other nature, that could influence the research or the results published in this article.

Funding

The study was conducted without financial support.

Data availability

Data will be provided upon reasonable request.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies to write this paper.

References

1. Ahn, S.J., Lee, H. and Cho, K.-J. (2024), "3D printing with a 3D printed digital material filament for programming functional gradients", *Nature Communications*, Vol. 15, 3605 p. DOI: <https://doi.org/10.1038/s41467-024-47480-5>
2. Wang, Y., Tan, Q., Pu, F., Boone, D.A. and Zhang, M. (2020) "A review of the application of additive manufacturing in prosthetic and orthotic clinics from a biomechanical perspective", *Engineering*, Vol. 6 (11), pp. 1258-1266. DOI: <https://doi.org/10.1016/j.eng.2020.07.019>
3. Baldock, M., Pickard, N., Prince, M., Kirkwood, S., Chadwell, A., Howard, D., Dickinson, A., Kenney, L., Gill, N., Curtin, S. (2023), "Adjustable prosthetic sockets: A systematic review of industrial and research design characteristics and their justifications", *Journal of NeuroEngineering and Rehabilitation*, Vol. 20, 147 p. DOI: <https://doi.org/10.1186/s12984-023-01270-0>
4. Nevliudov, I., Yevsieiev, V., Maksymova, S., Chala, O., (2023), "A Small-Sized Robot Prototype Development Using 3D Printing", *CAD In Machinery Design Implementation and Educational Issues (CADMD'2023): proceedings of the XXXI International Conference*. (Conference in memory of Professor Jerry Wrobel), Suprasl, 26-28 October, 2023. Suprasl, 2023. 12 p.
5. Xie J., Liu X., Tang J., Li X., Li W., (2021), "Study on friction behavior at the interface between prosthetic socket and liner", *Acta of Bioengineering and Biomechanics*, Vol. 23, No. 1. P. 83–93. DOI: <https://doi.org/10.37190/ABB-01751-2020-04>
6. Rossi, S.; Puglisi, A.; Benaglia, M., (2018) "Additive Manufacturing Technologies: 3D Printing in Organic Synthesis". *ChemCatChem*, Vol. 10, (7), pp. 1512– 1525. DOI: <https://doi.org/10.1002/cctc.201701619>.
7. Syed, K. M., Yevsieiev, V., Nevliudov, I., Lyashenko, V., Adel, R. A., Wahid, R. (2022), "HMI Development Automation with GUI Elements for Object-Oriented Programming Languages Implementation," *International Journal of Engineering Trends and Technology (IJETT)*, Vol. 70, No. 1, pp. 139-145. DOI: <https://doi.org/10.14445/22315381/IJETT-V70I1P215>
8. Hassan Beygi, B., Wong, M.S. (2023), "Contemporary and future development of 3D printing technology in the field of assistive technology, orthotics and prosthetics", *Canadian Prosthetics & Orthotics Journal*, 6(2), 42225. DOI: <https://doi.org/10.33137/cpoj.v6i2.42225>
9. Devin, K.M., Tang, J., Hamilton, A.R., Moser, D. and Jiang, L. (2024), "Assessment of 3D-printed mechanical metamaterials for prosthetic liners", *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 238(3), 348-357. DOI: <https://doi.org/10.1177/09544119231225529>
10. Nikitin, D. O., Nevlyudov, I. Sh., Zharikova, I. V., Bronnikov, A. I., Strilets, R. E. (2025), "Devising of a method for controlling products during photopolymer 3D printing". *Eastern-European Journal of Enterprise Technologies*, 4(1 (136), pp. 42–54. DOI: <https://doi.org/10.15587/1729-4061.2025.335706>
11. Sh. Nevlyudov, S.P. Novoselov, A.G. Resnichenko. (2012), "Experimental verification of theoretical foundations making the basis of the substrate roughness surface automated control technology". *Telecommunications and Radio Engineering*, 71(19), pp. 1791-1799. DOI: <https://doi.org/10.1615/TelecomRadEng.v71.i19.80>
12. Plesec, V., Hanželič, B. and Harih, G. (2024), "Development of a metamaterial numerical model for improving 3D-printed lower-limb prosthetic liners", *Human Systems Engineering and Design (IHSED 2024): Future Trends and Applications*, 158, pp. 228–238. DOI: <https://doi.org/10.54941/ahfe1005546>
13. Gritsyuk, V., Nevliudov, I., Zablodskiy, M. and Subramanian, P. (2022), "Estimation of eddy currents and power losses in the rotor of a screw electrothermomechanical converter for additive manufacturing", *Machinery and Energetics*, 13(2), pp. 41–49. DOI: [https://doi.org/10.31548/machenergy.13\(2\).2022.41-49](https://doi.org/10.31548/machenergy.13(2).2022.41-49)
14. Andriën, A.R.P., Guerreiro Tomé Antunes, D.J., van de Molengraft, M.J.G. and Heemels, W.P.M.H. (2018) "Similarity-based adaptive complementary filter for IMU fusion", *2018 European Control Conference (ECC)*, pp. 3044-3049. DOI: <https://doi.org/10.23919/ECC.2018.8550358>
15. Justa, J., Šmídl, V. and Hamáček, A. (2020) "Fast AHRS Filter for Accelerometer, Magnetometer, and Gyroscope Combination with Separated Sensor Corrections", *Sensors*, 20(14), 3824 p. DOI: <https://doi.org/10.3390/s20143824>
16. Osadchyi S. I., Zozulya, V. A., Kalich, V. M., Timoshenko, A. S. (2024), "The frequency method for optimal identification of close-loop system elements". *Radio Electronics, Computer Science, Control*, (4), 195 p. DOI: <https://doi.org/10.15588/1607-3274-2023-4-18>
17. Mamalis, A.G., Nevliudov, I., Romashov, Yu. (2021), "An approach for numerical simulating and processing of measured electrical signals from board sensors installed on wheeled electro-mechanical platforms", *Journal of Instrumentation*, 16(10), 10006 p. DOI: <https://doi.org/10.1088/1748-0221/16/10/P10006>
18. Rudenko, O., Bezsonov, O., Ilyunin, O., Demirskiy, O., Serdiuk, N., Arsenyeva, O., Semenenko, O. (2023), "Using a Neural Network Approach to Predict Deposits on the Surfaces of Heat Exchange Equipment", *Chemical Engineering Transactions*, 103, pp. 697-702. DOI: <https://doi.org/10.3303/CET23103117>
19. Brian, H., Daniel, T. (2017), *Essential MATLAB for Engineers and Scientists*, Academic Press, 411 p. DOI: <https://doi.org/10.1016/C2015-0-02182-7>

20. Han, S., Bhattacharyya, S. P. (2018), "PID controller synthesis using a v-Hurwitz stability criterion", *IEEE Control Syst. Lett.*, Vol. 2, No. 3, pp. 525-530. DOI: <https://doi.org/10.1109/LCSYS.2018.2842784>
21. *Statistical Methods in Psychiatry Research and SPSS*. 2nd Edition, By M. Venkataswamy Reddy. Apple Academic Press. 2019. 442 p. DOI: <https://doi.org/10.1201/9780429023309>

Received (Надійшла) 26.01.2026

Accepted for publication (Прийнята до друку) 14.03.2026

Publication date (Дата публікації) 30.03.2026

Відомості про авторів / About the Authors

Невлюдов Ігор Шакірович – доктор технічних наук, професор, заслужений діяч науки і техніки України, Лауреат Державної премії в галузі науки і техніки України; Лауреат Державної премії України в галузі освіти, Харківський національний університет радіоелектроніки, завідувач кафедри комп'ютерно-інтегрованих технологій, автоматизації, робототехніки та безпекової інженерії; м. Харків, Україна;

Igor Nevlyudov – Doctor of Technical Sciences, Professor, Honored Worker of Science and Technology of Ukraine, Laureate of the State Prize in Science and Technology of Ukraine; Laureate of the State Prize of Ukraine in the field of education, Kharkiv National University of Radio Electronics, Head of the Department of Computer Integrated Technologies, Automation, Robotics and Safety Engineering; Kharkiv, Ukraine;

e-mail: igor.nevlyudov@nure.ua

ORCID ID: <https://orcid.org/0000-0002-9837-2309>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57216434058>

Рубан Олена Анатоліївна – доктор фармацевтичних наук, професор, Національний фармацевтичний університет, завідувачка кафедри заводської технології ліків та косметичних засобів; м. Харків, Україна;

Olena Ruban – Doctor of Pharmaceutical Sciences, Professor, National University of Pharmacy, Head of the Department of Industrial Technology of Medicines and Cosmetics; Kharkiv, Ukraine;

e-mail: Ruban_elen@ukr.net

ORCID ID: <https://orcid.org/0000-0002-2456-8210>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=55976080800>

Нікітін Дмитро Олександрович – кандидат технічних наук, Харківський національний університет радіоелектроніки, доцент кафедри комп'ютерно-інтегрованих технологій, автоматизації, робототехніки та безпекової інженерії; м. Харків, Україна;

Dmytro Nikitin – Candidate of Technical Sciences, Kharkiv National University of Radio Electronics, Associate Professor of the Department of Computer Integrated Technologies, Automation, Robotics and Safety Engineering; Kharkiv, Ukraine;

e-mail: dmytro.nikitin@nure.ua

ORCID ID: <https://orcid.org/0000-0003-4388-4996>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=59412787100&origin=recordpage#>

Місан Богдан Сергійович – аспірант, Харківський національний університет радіоелектроніки, кафедра комп'ютерно-інтегрованих технологій, автоматизації, робототехніки та безпекової інженерії; м. Харків, Україна;

Bogdan Misan PhD Student, Kharkiv National University of Radio Electronics, Department of Computer Integrated Technologies, Automation, Robotics and Safety Engineering; Kharkiv, Ukraine;

e-mail: bohdan.misan@nure.ua

ORCID ID: <https://orcid.org/0009-0007-4905-9280>

Іохов Олександр Юрійович – доктор технічних наук, професор, Національна академія Національної гвардії України, начальник центру імітаційного моделювання; Харків, Україна;

Oleksandr Iokhov – Doctor of Technical Sciences, Professor, National Academy of the National Guard of Ukraine, Head of the Simulation Modeling Center; Kharkiv, Ukraine;

e-mail: iokhov@ukr.net

ORCID ID: <https://orcid.org/0000-0002-1718-0138>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=57220574439>

ДОСЛІДЖЕННЯ ДЕМПФУВАЛЬНИХ ВЛАСТИВОСТЕЙ 3D-ДРУКОВАНИХ ЛАЙНЕРІВ З КЕРОВАНОЮ ПЕРФОРАЦІЄЮ

Предметом дослідження цієї статті є демпфувальні властивості 3D-друкованих прокладок (лайнерів) з термопластичного поліуретану (TPU) з керованою внутрішньою перфорацією, призначених для використання в системі «кукса – лайнер – гільза» у протезах нижніх кінцівок. Досліджується вплив жорсткості матеріалу, геометрії отворів та ступеня перфорації на ефективність гасіння вібрацій при ударно-динамічному навантаженні. **Мета** роботи – експериментальне визначення та оптимізація демпфувальних характеристик TPU-лайнерів шляхом варіювання жорсткості матеріалу, форми отворів і відсотка перфорації з використанням методу вільних затухаючих коливань. **Завдання:** проаналізувати функціональну роль лайнера як демпфувального елемента у протезній системі; виготовити серію 3D-друкованих TPU-зразків з різною жорсткістю, геометрією отворів та ступенем перфорації; реалізувати ударний метод вимірювання затухаючих коливань для кожного зразка; визначити коефіцієнт демпфування та відсоток гасіння вібрації; провести обробку експериментальних даних з метою виявлення оптимальних комбінацій матеріалу і геометричних параметрів; виявлено закономірність впливу жорсткості матеріалу, перфорації та патерну отворів на демпфувальні властивості лайнерів. **Результати:** реалізовано та експериментально перевірено метод оцінки демпфувальних характеристик 3D-друкованих TPU-лайнерів із керованою внутрішньою структурою. Встановлено нелінійну залежність ефективності гасіння вібрацій від ступеня перфорації та систематичне зменшення демпфування зі зростанням жорсткості TPU. Показано, що шестикутна геометрія отворів забезпечує більш рівномірний розподіл деформацій і дещо вищу ефективність гасіння порівняно з ромбічною та трикутною. Отримані залежності дозволяють цілеспрямовано проектувати демпфувальні лайнери з оптимальним співвідношенням жорсткості та здатністю до гасіння вібрації. Це дозволить вирішити наступні питання, такі як: зниження ударних навантажень – зменшення пікових динамічних сил, що передаються від гільзи протеза до м'яких тканин кукси; перерозподіл тиску – формування більш рівномірних контактних напружень у зоні контакту виробу зі шкірою; підвищення комфорту користувача – зменшення вібрацій, болю та подразнення шкіри під час ходьби; індивідуальна оптимізація – персоналізований підбір геометрії та матеріалу лайнеру під параметри тіла, рівня активності та стану тканин пацієнта; інженерне проектування – створення структурно оптимізованих елементів для біомединого застосування. **Висновки:** за результатами роботи експериментально підтверджено можливість керування демпфувальними властивостями TPU-лайнерів шляхом зміни жорсткості матеріалу, ступеня перфорації та геометрії отворів. Встановлено наявність оптимального діапазону перфорації для кожного рівня жорсткості TPU, а також переваги шестикутної перфорації з точки зору рівномірності деформацій та ефективності гасіння вібрацій.

Ключові слова: TPU-лайнери; 3D-друк; жорсткість матеріалу; протези; ударні випробування.

Бібліографічні описи / Bibliographic descriptions

Невлюдов І. Ш., Рубан О. А., Нікітін Д. О., Місан Б. С., Іохов О. Ю. Дослідження демпфувальних властивостей 3D-друкованих лайнерів з керованою перфорацією. *Сучасний стан наукових досліджень та технологій в промисловості*. 2026. № 1 (35). С. 155–168. DOI: <https://doi.org/10.30837/2522-9818.2026.1.155>

Nevlyudov, I., Ruban, O., Nikitin, D., Misan, B., Iokhov, O. (2026), "Investigation of the damping properties of 3D-printed liners with controlled perforation", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (35), P. 155–168. DOI: <https://doi.org/10.30837/2522-9818.2026.1.155>

АЛФАВІТНИЙ ПОКАЖЧИК

Јіан Үі	100
Балабуха О. С.	39
Барковська О. Ю.	55
Вакулік Є. В.	6
Гопесенко В.	65
Гулевич М. В.	39
Дмітрієв О. М.	39
Досталова К.	17
Єна М. В.	28
Заставний О. М.	79
Зачко О. Б.	139
Іохов О. Ю.	155
Клімо М.	17
Коломійцев О. В.	39
Косенко В. В.	65
Косенко Н. В.	65
Крайчі О.	55
Кучук Г. А.	65
Кучук Н. Г.	65
Левченко А. О.	39
Лисий М. І.	79
Матківська Х. С.	139
Мельниченко О. В.	79
Місан Б. С.	155
Невлюдов І. Ш.	155
Нікітін Д. О.	155
Погудіна О. К.	28
Радюк П. М.	79
Рубан О. А.	155
Саченко А. О.	79
Сидоренко Л.	55
Сітнікова О. О.	100
Смеляков К. С.	6
Соловей Б. А.	115
Сотник І. С.	127
Терентьєв О. О.	115
Тригуба А. М.	139
Чижмарова О.	17
Чуприна А. С.	6
Шапіро О. К.	127
Ярошевич Р. О.	17

ALPHABETICAL INDEX

Jian Yu	100
Balabukha Oleksiy	39
Barkovska Olesia	55
Vakulik Eugen	6
Gopejenko Viktors	65
Hulevyeh Mykhailo	39
Dmitriiev Oleh	39
Dostalova Kristina	17
Yena Maksym	28
Zastavnyy Oleg	79
Zachko Oleh	139
Iokhov Oleksandr	155
Klimo Marek	17
Kolomitsev Oleksii	39
Kosenko Viktor	65
Kosenko Nataliia	65
Krajči Alexander	55
Kuchuk Heorhii	65
Kuchuk Nina	65
Levchenko Andrii	39
Lysyi Mykola	79
Matkivska Khrystyna	139
Melnychenko Oleksandr	79
Misan Bogdan	155
Nevlyudov Igor	155
Nikitin Dmytro	155
Pohudina Olha	28
Radiuk Pavlo	79
Ruban Olena	155
SachenkoAnatoliy	79
Sidorenko Ludmila	55
Sitnikova Oksana	100
Smelyakov Kyrylo	6
Solovei Bohdan	115
Sotnyk Ihor	127
Terentyev Oleksandr	115
Tryhuba Anatoliy	139
Cizmarova Alexandra	17
Chupryna Anastasiya	6
Oleksii Shapyro	127
Yaroshevyh Roman	17

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

№ 1 (35), 2026

No. 1 (35), 2026

Відповідальний секретар журналу *I.Г. Перова*
Відповідальний за випуск *A.A. Коваленко*
Відповідальний за ліцензування *V.V. Косенко*
Редактор *L.V. Кузьміна*
Комп'ютерна верстка *Л.Ю. Светайло*

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. 19,6.
Тираж 150 прим.

Format 60×84/8. Conventional printed sheets 19,6.
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