UDC 004.932
DOI: 10.15587/2706-5448.2025.325423

Olesia Barkovska,
Igor Ruban,
Daria Tymoshenko,
Oleksandr Holovchenko,
Oleksandr Yankovskyi

RESEARCH ON MOBILE MACHINE LEARNING PLATFORMS FOR HUMAN GESTURE RECOGNITION IN HUMAN-MACHINE INTERACTION SYSTEMS

The subject of this research is mobile machine learning platforms for human gesture recognition within human-machine interaction systems, specifically for managing smart home components.

One of the key challenges in gesture recognition is ensuring high accuracy, efficiency, and robustness of algorithms under real-world operating conditions. The problem lies in selecting optimal machine learning platforms capable of balancing local and cloud computing, processing speed, and adaptability to changing environmental conditions.

The study presents a comparative analysis of the ML platforms Create ML (Apple) and Google Cloud AI Platform, which are used for gesture detection and recognition in smart home control systems. The obtained results demonstrate that Create ML achieves an accuracy of 95.81 %, while Google Cloud AI Platform reaches 89.43 %, justifying their selection for further research. Additionally, experimental testing of sensor placement topology revealed that diagonal camera positioning increases accuracy by 0.62 % compared to parallel placement.

The increased efficiency of Create ML is due to its ability to process data locally, reducing latency and dependence on an internet connection. In contrast, Google Cloud AI Platform relies on cloud resources, enabling the processing of large volumes of data but making it dependent on data transmission speed.

The proposed gesture control algorithms can be used to enhance the accessibility of technology for people with disabilities, particularly in rehabilitation centers. Additionally, the research findings can be applied to contactless interfaces in medical facilities and public spaces, reducing the need for physical interaction with surfaces and improving hygiene levels. The use of mobile ML platforms in such scenarios allows for the optimization of computational resources and ensures the effective integration of gesture control into modern human-machine systems.

Keywords: *human-machine interaction, HMI, Create ML, Google Cloud AI Platform, image processing, contactless control, ML platforms.*

Received: 18.12.2024
Received in revised form: 24.02.2025
Accepted: 14.03.2025
Published: 28.03.2025

How to cite
Barkovska, O., Ruban, I., Tymoshenko, D., Holovchenko, O., Yankovskyi, O. (2025). Research on mobile machine learning platforms for human gesture recognition in human-machine interaction systems. Technology Audit and Production Reserves, 2 (2 (82)), 6–14. https://doi.org/10.15587/2706-5448.2025.325423

1. Introduction

The study of methods for interpreting human movements in remote control systems for smart home elements is highly relevant as an example of human-machine interaction. In today's world, smart home technologies have become an integral part of everyday life, providing automation and remote control of household devices.

According to the data presented in [1], all segments of the smart home market are continuously growing, as evidenced by statistical indicators (Fig. 1). This confirms the relevance of developing the smart home industry and implementing smart technologies in various aspects of modern housing to ensure comfort, convenience, energy efficiency, and enhanced security.

The main reasons for such growth are the development of artificial intelligence, the Internet of Things, computer vision, and contactless control interfaces, among which one of the most promising is gesture control (Fig. 2).

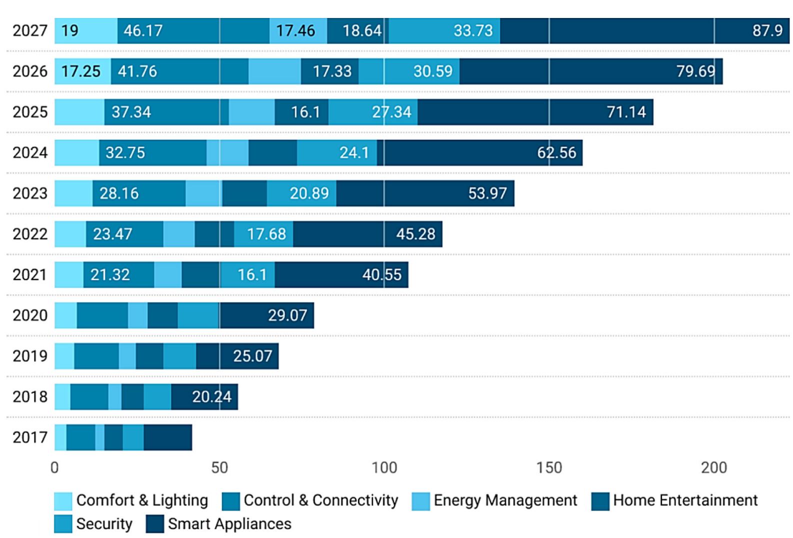


Fig. 1. Statistical indicators of revenue distribution in the global smart home market by individual segments

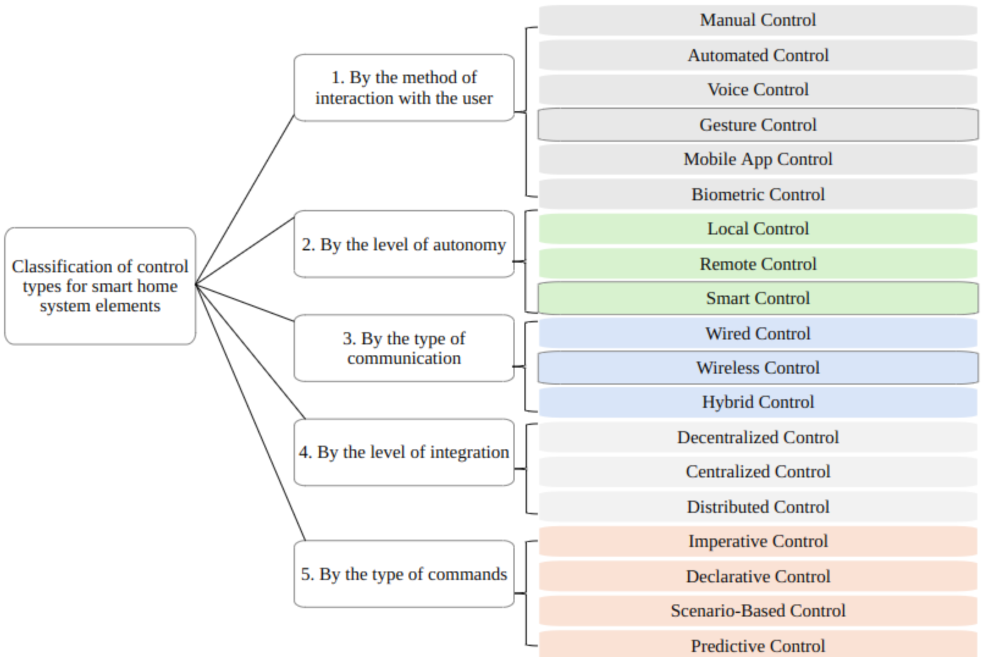


Fig. 2. Classification of control types for smart home system elements [2–6]

Gesture control is an example of the integration of intelligent management and wireless technologies, making smart homes even more adaptive and autonomous.

The diagram in Fig. 2 illustrates the classification of control types, dividing them based on various characteristics. It contains several hierarchical levels, with each level detailing the management approach. The proposed structure helps to understand which approaches can be applied for building control systems and the role of gesture control in human-machine interaction.

As a justification for the choice of mobile machine learning platforms for human gesture recognition in human-machine interaction systems, let's present the advantages of these approaches:

– mobile devices are widespread and portable, allowing the integration of machine learning algorithms for gesture control without the need for complex infrastructure;
– the use of mobile neural network frameworks (TensorFlow Lite, MediaPipe, Core ML) enables the optimization of recognition speed without the need to transmit data to a server;
– mobile platforms are optimized for energy-efficient computations, which is crucial for continuous gesture recognition without rapidly draining the device's battery;
– gesture control on mobile devices can be used for smart interfaces, assistive technologies for people with disabilities, virtual reality, robotics, and military contactless device control systems.

According to the World Health Organization, over 1 billion people worldwide have some form of disability, and contactless human-machine interaction can significantly improve their lives. For example, rehabilitation centers use systems that allow patients with limited mobility to control wheelchairs or computers through gestures [7]. Additionally, during the COVID-19 pandemic, the demand for contactless interfaces increased by 30 % as they help minimize physical contact with surfaces that may contain viruses. Such systems can be used in public places, for instance, for contactless door opening or elevator control.

In addition to human-machine interaction, such as systems for controlling smart home elements, intellectual assistants are also beneficial for people with disabilities. One of their tasks is to assist users in navigating the city, thereby ensuring safety and greater inclusivity, especially for users with visual impairments [8]. Monitoring the condition and progress of rehabilitation for people with musculoskeletal disabilities is also a demanding task. However, today, in the context of full-scale war, the requirements for such systems include the ability for remote monitoring and portability of the system [9].

In the context of the war in Ukraine, gesture interpretation technology is also gaining importance. According to the Ministry of Health of Ukraine, since the start of the full-scale invasion, more than 50,000 people have been injured, resulting in the loss of limbs or limited mobility. For such patients, contactless interaction and device control are critical elements of rehabilitation. Additionally, gesture control technologies can significantly enhance safety. For example, the system could allow automatic lighting or signaling to be activated with certain gestures, which is especially important for people with disabilities or in emergency situations.

The paper focuses on the study of mobile ML platforms for gesture recognition in human-machine interaction systems, as this type of control offers several advantages, namely:

– ease of use (the user can activate the required function with a single hand gesture, without the need to open an app or search for a remote control);
– high interaction speed (a hand movement takes less time than giving a voice command or opening a mobile app);
– accessibility for people with disabilities (the ability to control for elderly or disabled individuals who may not be able to quickly interact with mobile apps or buttons);
– enhanced security (unlike voice commands, which can be accidentally activated by conversation, gestures are harder to perform unintentionally);
– reliability in challenging conditions (local gesture recognition systems can operate without access to servers or cloud technologies);
– integration with other technologies (modern systems allow for training algorithms to recognize new gestures, making control more flexible).

Thus, the relevance of the chosen topic is determined not only by the active development of systems for controlling smart home elements, as an example of human-machine interaction, but also by the widespread use of gesture control in various areas of people's lives.

In [10], the advantages and limitations of using radar sensors for gesture recognition are discussed, particularly their effectiveness in low-light conditions and insensitivity to obstacles. However, the issue

ISSN-L 2664-9969; E-ISSN 2706-5448

of adapting radar recognition for mobile platforms has not been sufficiently explored, which has limited the potential for implementation in portable devices.

In the study [11], various machine learning algorithms for gesture recognition, such as CNN, Random Forest, and XGBoost, are compared, with CNN identified as the most accurate method. However, the issues of computational resource consumption and energy usage on mobile devices were not adequately addressed, which complicates real-world application for portable systems.

In the work [12], the MediaPipe library is used for gesture recognition in mobile applications, with its effectiveness and ease of implementation highlighted. However, the issue of recognition accuracy under low-light conditions or dynamic hand movement was not considered, which could significantly impact system performance in real-world scenarios.

In the study [13], YOLOv3 is used for real-time gesture recognition, providing high processing speed. However, the resource consumption of the algorithm and its adaptation for mobile platforms were not sufficiently explored, limiting the possibilities for integration into compact devices.

In [14], the use of convolutional neural networks (CNN) for real-time gesture recognition is explored. It is determined that pre-trained models improve accuracy. However, the issue of balancing performance and energy consumption on mobile devices remains unresolved, limiting the potential for integration into portable systems.

The work [15] examines the development of wearable devices for gesture recognition, including built-in sensors and deep learning algorithms. However, the issue of standardization between different wearable platforms and integration with mobile devices has not been adequately studied, which complicates practical implementation.

An analysis of current human-machine interaction technologies based on gesture recognition is presented in [16]. However, the issue of adapting systems to work in changing environmental conditions remains open, limiting the universality of recognition methods.

The analysis of the effectiveness of mobile platforms is presented in [17–20]. Specifically, the results highlighted in [17] determine that adapting face recognition algorithms on mobile devices, considering user movement, improves accuracy in changing conditions. However, the possibility of applying similar methods for gesture recognition has not been explored, which limits their practicality for other types of recognition.

The study [18] examines methods for training neural networks directly on mobile devices without using cloud services. It is determined that this approach improves autonomy. However, the issues of energy efficiency and real-time performance of such models remain unresolved, limiting their widespread application.

The paper [19] presents an analysis of the dependency of AI systems on cloud infrastructure, particularly its impact on performance and costs. It is found that cloud services provide scalability but create dependency on large providers. However, the compromise between cloud and mobile computing for gesture recognition has not been sufficiently explored, which limits the selection of the best approach for mobile solutions.

The study [20] analyzes cloud services for deploying generative AI, such as AWS, Google Cloud, Microsoft Azure, and others. It is determined that cloud services simplify the management and deployment of AI models. However, the issue of the effectiveness of these services in mobile environments and with low-power devices has not been addressed, which could impact the choice of appropriate architecture.

The problem discussed in the study is related to the efficiency and accuracy of human gesture recognition on mobile platforms, which can be integrated into human-machine interaction systems.

The effectiveness and accuracy of gesture recognition can be reduced due to a number of factors. Firstly, traditional methods have limitations: hands and fingers often overlap, and dynamic gestures are performed quickly or unclearly, which creates difficulties for computer vision algorithms. Secondly, many modern gesture recognition technologies (such as OpenPose) require significant computational resources, making them less suitable for mobile devices. The use of cloud platforms (e. g., Google Cloud AI Platform) leads to dependence on a stable internet connection, which is not always convenient. Thirdly, existing datasets for training models are limited, making it difficult to generalize results. Furthermore, hybrid solutions that combine local training (Create ML) and cloud computing (Google Cloud AI Platform) have been insufficiently studied. Solving these issues is possible through the evaluation of the advantages of different device placement topologies within a hybrid approach.

The aim of the paper is to study and conduct a comparative analysis of mobile machine learning platforms (Create ML and Google Cloud AI Platform) for human gesture recognition, which are integrated into human-machine interaction systems. The research aims to determine the balance between local and cloud computing, assess their performance in conditions with limited computational resources, and evaluate their effectiveness in a variable operating environment. The optimal platform is determined based on experimental testing of recognition accuracy, gesture processing performance, and adaptability to changing environmental conditions.

2. Materials and Methods

The object of research is mobile machine learning platforms for human gesture recognition in human-machine interaction systems, specifically in the control of smart home elements.

Despite significant progress in this field, there are a number of challenges that limit the effectiveness and accuracy of gesture control. The difficulty in gesture recognition for AI lies in the fact that hands and fingers may obscure each other, gestures can be too fast or unclear, making it difficult for algorithms to recognize them accurately and instantly.

With the advent of deep learning, particularly using deep convolutional neural networks (CNN), the paradigm of gesture recognition has significantly changed. Modern research demonstrates that the application of 3D-CNN allows for effective analysis of video sequences for recognizing dynamic hand gestures [21]. For example, in [22], a method for fast gesture recognition is proposed based on an improved lightweight network Lite-HRNet, transformed into 3D-CNN for processing video data.

An overview of modern methodologies for identifying hand movements and gestures is presented in the paper [23], where the author highlights methodologies based on:
– visual assessment;
– sensor-based approaches;
– hybrid methodologies.

In this work, methodologies based on visual assessment of hand fingers are considered and analyzed.

The results of modern studies have been implemented into a number of existing technologies and platforms for gesture detection, among which the following can be highlighted:
– Create ML (Apple);
– Google Cloud AI Platform;
– Microsoft Azure Cognitive Services;
– OpenPose (Carnegie Mellon University);
– MediaPipe (Google);
– Intel RealSense SDK;
– etc.

It is impossible to quantitatively assess the mentioned technologies in terms of accuracy, as no studies have been conducted under identical training and testing conditions on the same benchmarks. Therefore, generalized system characteristics that can address the problem described above were selected for comparison (Table 1). Create ML enables gesture recognition without relying on an internet connection, which eliminates the limitations of cloud platforms and optimizes real-time performance on mobile devices [24]. Google Cloud AI Platform compensates for the shortcomings of local solutions by using deep neural networks for more complex gestures and ensuring scalability, although it requires a stable internet connection [25]. Combining these platforms allows for evaluating the benefits of a hybrid solution, determining the balance between local and cloud processing, and optimizing accuracy and performance under varying operational conditions. OpenPose, although demonstrating high accuracy through body key-point analysis, has computational complexity and the need for a GPU, making it unsuitable for mobile platforms, as it requires significant resources for real-time operation [12]. Intel RealSense provides detailed 3D motion analysis, but this technology is focused on hardware solutions with deep cameras, limiting its integration into portable mobile devices [26]. Microsoft Azure Cognitive Services is not suitable for research within the hybrid technology framework, as, according to the analysis in the table, this platform does not support mobile integration, and its key advantage – streaming data processing – is less relevant to the specific task of the research.

Since the selected technologies, Create ML (Apple) and Google Cloud AI Platform, do not have built-in models for hand gesture recognition, they need to be trained on a suitable dataset. For this purpose, a review of datasets appropriate for this task was conducted (ASL Alphabet, Hand Gesture Recognition Database, Dynamic Hand Gesture Dataset, EgoHands Dataset, SHREC'17 Track Dataset, Diverse Hand Gesture Recognition Dataset, LaRED, OUHANDS, HANDS, SHAPE). Among the reviewed datasets, the OUHANDS dataset [27] was selected with the following characteristics:

– total dataset size (training and testing) – 4 GB;
– number of classes – 10 (gestures from everyday communication), with 300 images per class in .png format, resolution 640*480. Examples of training images are shown in Fig. 3.

This dataset contains images of hands in various poses, captured from different angles and under different lighting conditions, ensuring data diversity for training gesture recognition models. The presence of various perspectives and background details enhances the robustness of computer vision algorithms to the variability of gestures and the surrounding environment.

Table 1

Analysis of technologies and platforms for gesture detection

Technology	Cloud support	Real-time support	Mobile integration	Recognition method	Supported devices
Create ML (Apple)	No	Yes	Yes (iOS/macOS)	Classification of images	iPhone, iPad, Mac
Google Cloud AI Platform	Yes	No (cloud processing)	Yes (via TensorFlow Lite)	Video and image analysis	Android, iOS, Web
Microsoft Azure Cognitive Services	Yes	Yes	No	Video and streaming data analysis	Windows, Web
OpenPose	No	Yes (on GPU)	No	Identifying key points on the body	PC (Windows, Linux)
Intel RealSense	No	Yes	No	3D motion analysis	PC (Windows, Linux)

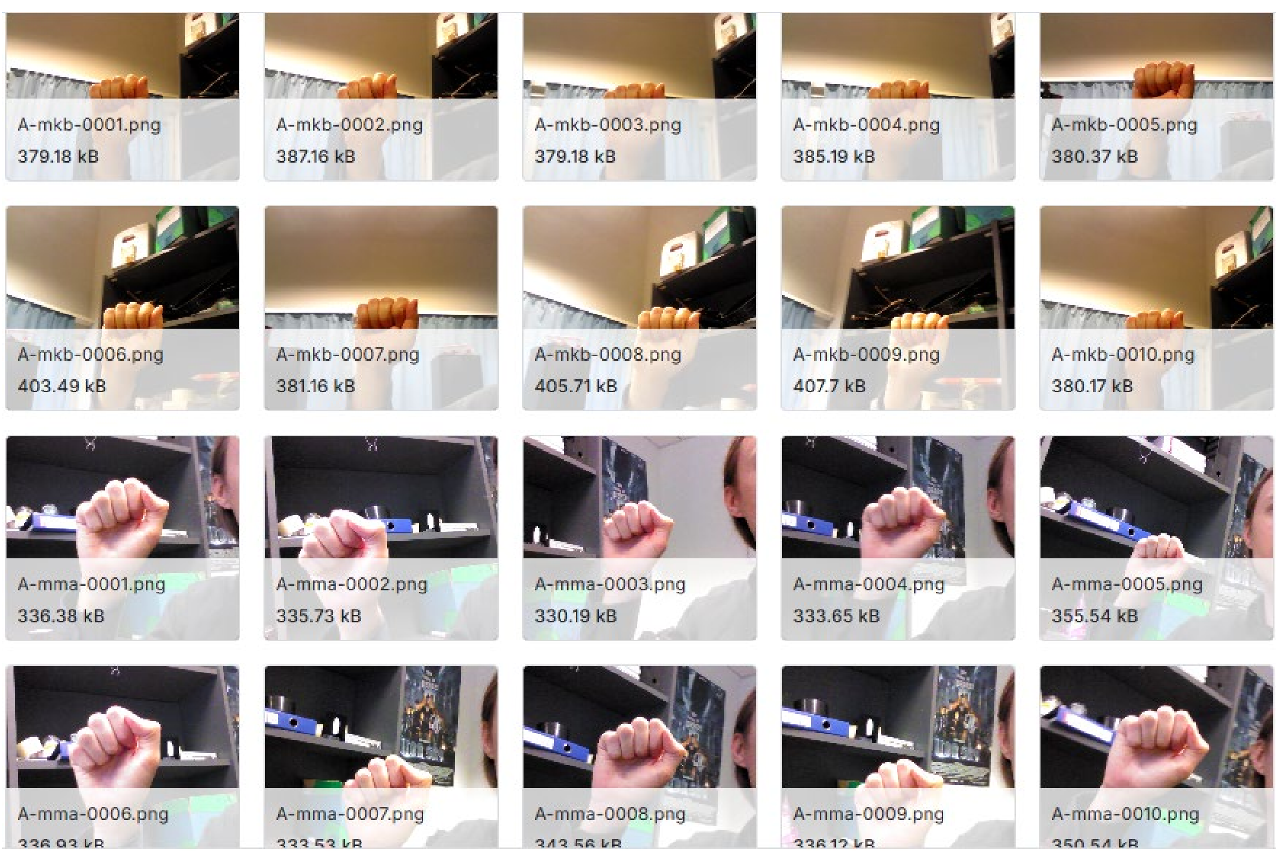


Fig. 3. Examples of training images from the OUHANDS static hand gesture dataset

The workflows for the above-mentioned technologies are as follows:

– for Create ML (Apple): Video stream capture → Hand detection using Vision API model → Gesture classification using Create ML model → Execute command in HomeKit;

– for Google Cloud AI Platform: Video stream capture → Hand contour detection using MediaPipe Hands → Gesture classification using Google AI → Execute command in Smart Home.

In the first experiment, the Create ML tool from Apple is used, which is applied on four computing devices placed at different corners of the studied area. All the devices that make up the gesture recognition system are trained on a single neural network based on Create ML. This tool is available through Xcode and allows local training of a model on devices like MacBook, iMac, etc., using a specified image collection (OUHANDS).

In the second experiment, the Google Cloud AI Platform service is used, which is applied on all four devices placed at the corners of the studied area. Thus, all the devices that form the human gesture detection and recognition system are trained on the same neural network model, specifically using Google Cloud AI Platform. The trained model is then converted into a .mlmodel format through the Google service, which works with CoreML Framework.

The third experiment involves testing the proposed hybrid system based on Create ML and Google Cloud AI Platform. The final experiment focuses on using hybrid technologies for human gesture recognition by combining the trained models. Initially, two diagonal devices use the same technologies, and then two parallel devices are employed.

3. Results and Discussion

The paper proposes a functional model of the user gesture recognition system (Fig. 4).

The distinction of the proposed model from existing ones lies in the use of a hybrid approach to gesture processing, where some operations are performed locally on the mobile device, while more resource-intensive tasks, including preprocessing and classification, can be offloaded to the cloud for processing using Core ML. The image preprocessing module is responsible for preparing images for further analysis, including resizing and normalization. This enhances the image quality by adjusting brightness, contrast, and colors. Unlike traditional models that rely on fixed data collection points, the proposed system allows for adaptive sensor placement, as demonstrated in subsequent experiments.

The feature extraction module is responsible for extracting various features from the preprocessed images using convolutional neural networks (CNN) or other algorithms.

The feature analysis module performs gesture classification using deep learning models, determining the type of gesture and its meaning.

The module responsible for training the neural network operates based on a dataset that contains images of hands in different lighting conditions and environments.

After training, the network performs gesture classification, and the result is passed to the interpretation module. This module assesses the reliability of the classification and provides recommendations to the user. The system includes communication between the sensors and the server for data processing and decision-making. The system structure is depicted in the organizational diagram in Fig. 5.

The sensor devices (Processing Slave 1–4) record and process video to detect gestures, transmitting the data to the decision-making server. Upon receiving the processed information, the server uses it to control smart home devices via a wireless connection. The system is flexible and portable, allowing it to be used in various locations and on different devices. It ensures reliability and fault tolerance when operating with both 2 and 4 gadgets.

To test the functionality of the proposed system based on the mentioned technologies, a test bench was set up for conducting experimental research (Fig. 6).

The test bench allowed a series of experiments to be conducted aimed at determining the optimal placement of sensors to achieve the highest accuracy in recognizing the user's hand gestures based on Create ML (Apple) technologies and/or Google Cloud AI Platform. The testing was carried out considering the following variable conditions, as shown in Fig. 6, *a*, *b*:

– different lighting conditions;

– selected hands for training based on classes: left hand, right hand;

– photos of the user's own hands (approximately 200 images for each of the two classes – fist and open palm);

– presence of accessories in the photos (ring, watch);

– environment – room for conducting experiments;

– hand rotation of about 50' is taken into account.

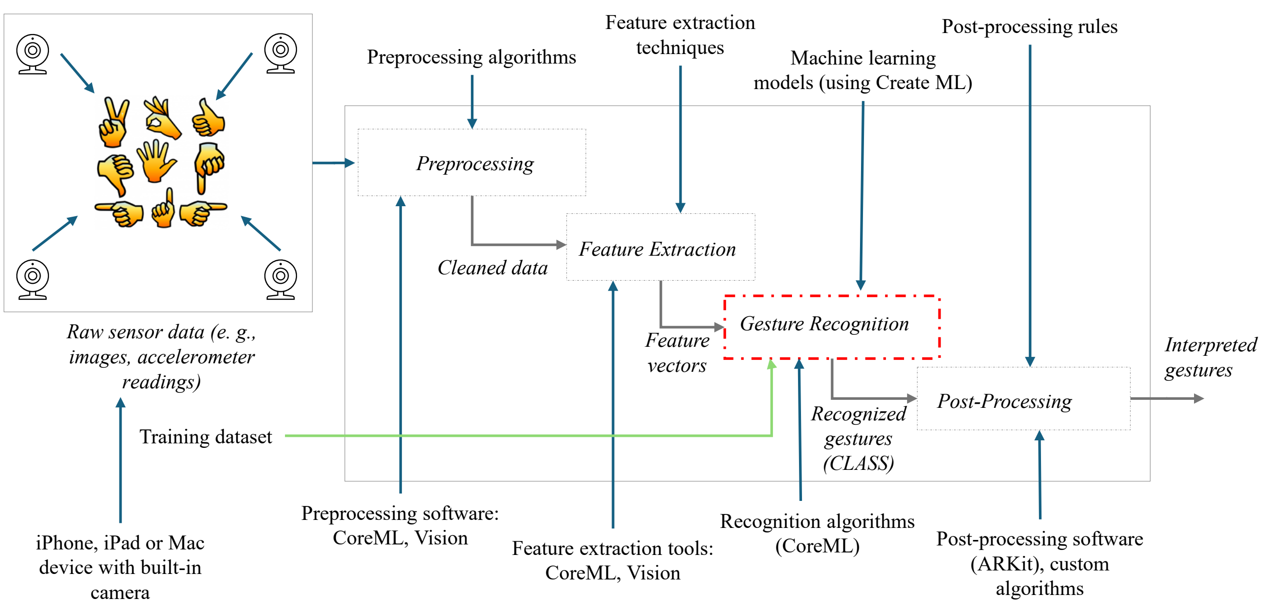


Fig. 4. The proposed functional model of the user gesture recognition system (using Apple's Create ML technology as an example)

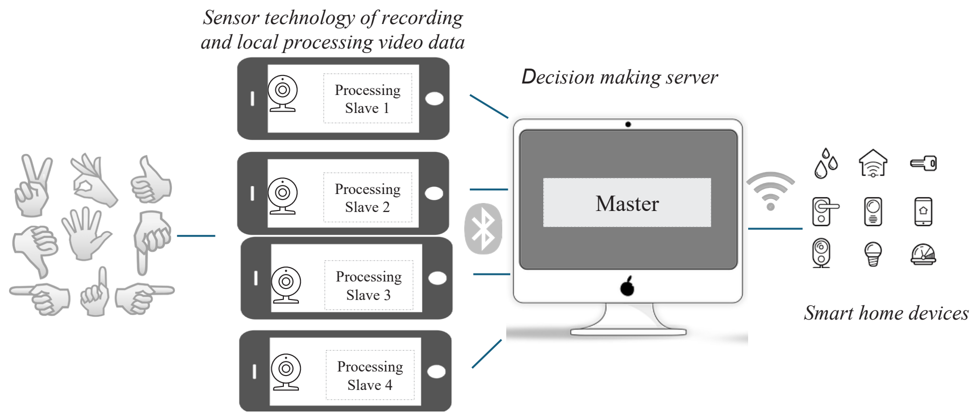
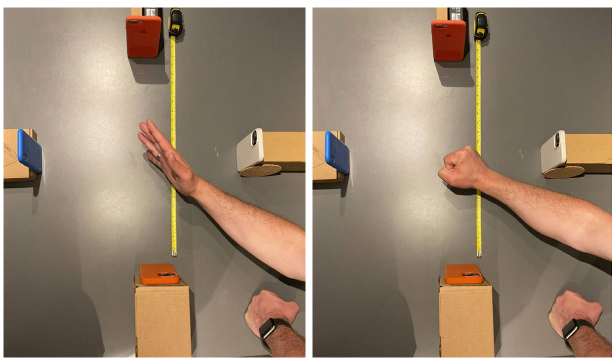


Fig. 5. The organizational diagram of the proposed system



a *b*

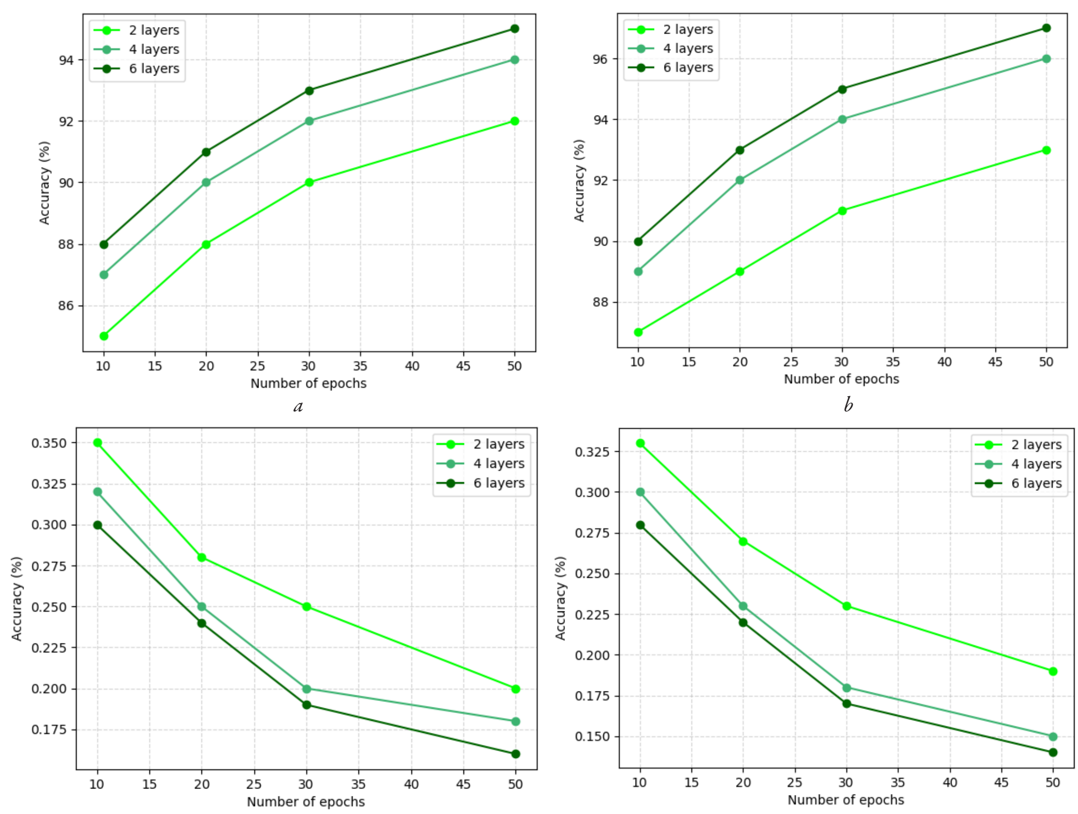
Fig. 6. The test bench for conducting experimental research:
a – for recognizing an open palm; *b* – for recognizing a clenched fist

The results of the first experimental research showed that the dataset distribution in a 70:15:15 ratio for training: testing: validation samples, respectively, yields a higher test accuracy, reaching 97 % with the lowest accuracy loss.

At the same time, the average recognition accuracy for the two gestures under study is 96.66 % for the open palm and 95.94 % for the clenched fist.

The technology selected for the first experiment showed better results with the use of a larger number of layers and epochs, especially with the 70:15:15 ratio, where the test accuracy reached its highest values – 97 % accuracy with 6 layers and 50 training epochs (Fig. 7).

Test losses decrease as the number of layers and epochs increases, indicating an improvement in the overall performance of the model and its ability to generalize the acquired knowledge to new data.



a *b* *c* *d*

Fig. 7. Evaluation of test accuracy and test losses for the Create ML model: *a* – change in classification accuracy depending on the number of epochs for models with 2, 4, and 6 layers at a training sample ratio of 50:25:25; *b* – change in classification accuracy depending on the number of epochs for models with 2, 4, and 6 layers at a training sample ratio of 70:15:15; *c* – change in classification loss depending on the number of epochs at a training sample ratio of 70:15:15; *d* – change in classification loss depending on the number of epochs at a training sample ratio of 50:25:25

ISSN-L 2664-9969; E-ISSN 2706-5448

The analysis of the results from the second experiment showed that the dataset distribution in a 70:15:15 ratio for the training, testing, and validation samples, respectively, yields higher test accuracy and lower test losses, reaching 97 % with network training over 50 epochs and test losses of 0.17 (Fig. 8).

At the same time, the average recognition accuracy for the two studied gestures is 90.56 % and 88.3 %, respectively, for the open and clenched palm (Table 2).

Regardless of the chosen technology for gesture detection and recognition, all experiments show higher recognition accuracy for the open palm compared to the clenched fist. This may be because the open palm has a larger surface area and more distinct features, such as fingers, palm lines, and hand shape, providing more key points for computer vision algorithms. In the clenched fist, these features are less pronounced, making recognition more difficult.

The results of the third experiment showed that the diagonal placement has an advantage of 0.62 % compared to the parallel placement (Table 3).

The overall conclusions from Experiment 4 are that, due to the combination of two different gesture recognition technologies on local processors, the overall accuracy of gesture detection for diagonal placement was 96.43 %. This is 0.62 % higher than for parallel placement (95.81 %). This indicates the advantage of the combined approach and the optimal placement of sensor devices to enhance the accuracy of the gesture recognition system.

Apple's Create ML technology demonstrated better gesture recognition accuracy compared to Google Cloud AI Platform in all scenarios. It achieved an average accuracy of 95.81 %, while the accuracy of Google Cloud AI Platform was 89.43 %. Using a single technology for all cameras yielded good results, but the hybrid approach proved to be more effective.

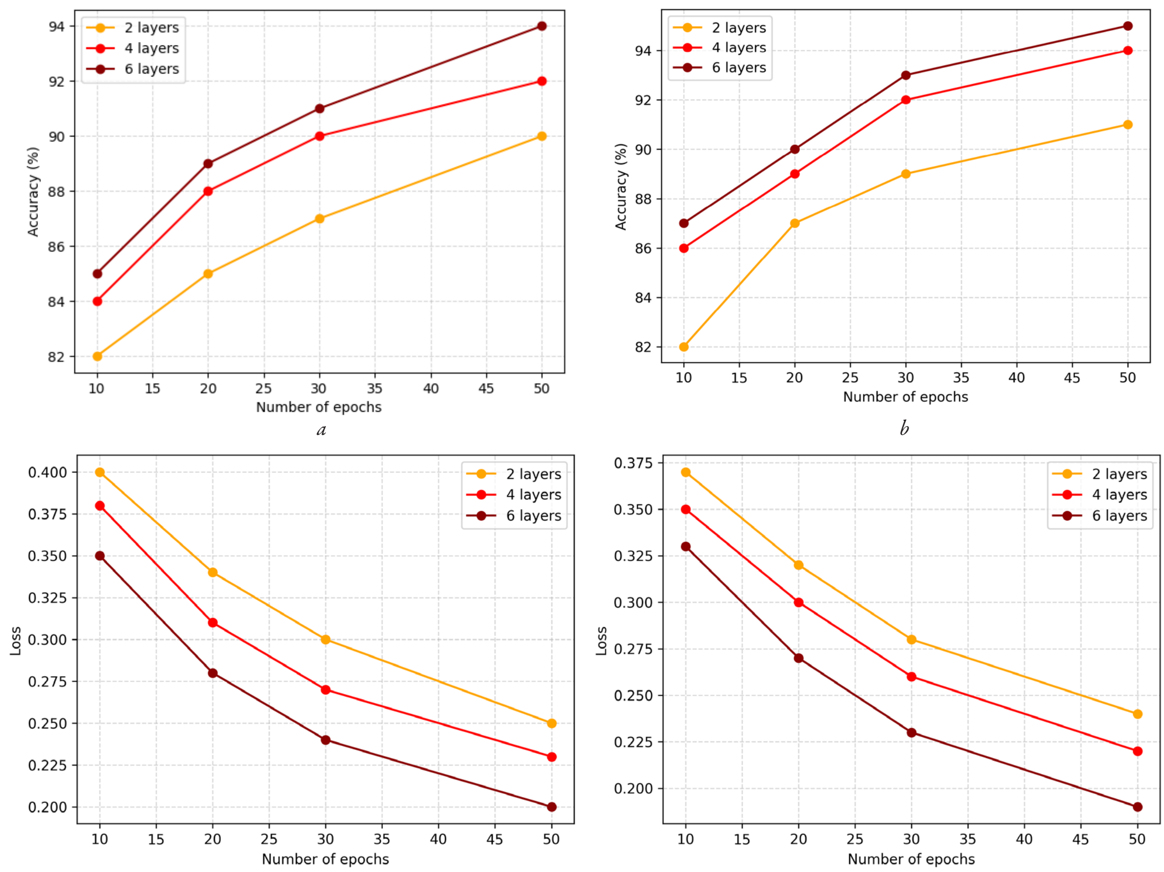


Fig. 8. Evaluation of test accuracy and test losses for the Google Cloud AI Platform service: *a* – change in classification accuracy depending on the number of epochs for models with 2, 4, and 6 layers at a training sample ratio of 50:25:25; *b* – change in classification accuracy depending on the number of epochs for models with 2, 4, and 6 layers at a training sample ratio of 70:15:15; *c* – change in classification loss depending on the number of epochs at a training sample ratio of 70:15:15; *d* – change in classification loss depending on the number of epochs at a training sample ratio of 50:25:25

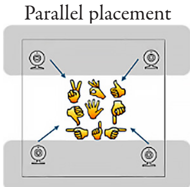
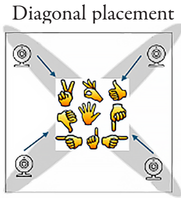
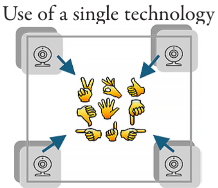
Table 2

The study of system accuracy on different gestures for the Apple Create ML tool and Google Cloud AI Platform (for the defined architecture)

Indicators	Recognition accuracy
Create ML from Apple	
Accuracy of gesture 1 recognition (open palm)	96.15 %
Accuracy of gesture 2 recognition (clenched fist)	95.47 %
Average recognition accuracy	95.81 %
Google Cloud AI Platform	
Accuracy of gesture 1 recognition (open palm)	90.56 %
Accuracy of gesture 2 recognition (clenched fist)	88.3 %
Average recognition accuracy	89.43 %

Table 3

Study of the topology of gadget placement with installed trained models

Scheme of recorder placement	Selected configuration	Accuracy of gesture 1 recognition (open palm)	Accuracy of gesture 2 recognition (clenched fist)	Average recognition accuracy
Parallel placement 	Camera 1, 2 – Create ML technology from Apple	97.24 %	95.9 %	95.86 %
	Camera 3, 4 – Google Cloud AI Platform technology	95.78 %	94.52 %	
Diagonal placement 	Camera 1, 3 – Create ML technology from Apple	98.05 %	96.78 %	96.43 %
	Camera 2, 4 –Google Cloud AI Platform technology	96.5 %	94.39 %	
Use of a single technology 	Camera 1–4 – Create ML technology from Apple	96.15 %	95.47 %	95.81 %
	Camera 1–4 – Google Cloud AI Platform technology	90.56 %	88.3 %	89.43 %

The research results can be *applied in the field of* contactless device control, smart home automation, rehabilitation systems for individuals with motor impairments, as well as in the development of human-machine interaction interfaces. The use of a hybrid approach to gesture recognition and movement analysis improves the accuracy and speed of signal processing, which is critical for medical and rehabilitation applications. Additionally, the proposed methodology can be used in security and monitoring systems for recognizing specific gesture commands in critical situations.

The main limitations are the dependency on the quality of the dataset for training the models, the impact of external factors such as lighting, shadows, and camera positioning. Hardware limitations are also important: the performance of the computing devices affects the speed of recognition. Additionally, the proposed system may require further calibration to function correctly under different operating conditions.

The conditions of martial law in Ukraine partially impacted the research, as it became more difficult to organize joint studies and exchange information among the members of the research group. However, this did not affect the results, as testing and algorithm optimization were carried out in conditions close to real-life situations, namely at home.

The prospects for further research will focus on improving the system's adaptation to individual user characteristics, as well as expanding the set of gestures for interaction. Special attention will be given to the integration of movement recognition methods with augmented and virtual reality technologies for the rehabilitation of patients who have lost limbs. Additionally, a promising direction is the improvement of energy efficiency in machine learning models for use in mobile devices and embedded systems.

4. Conclusions

The methods of interpreting human movements in smart home control systems play a significant role in enhancing convenience, energy efficiency, and home security. This work addressed the following tasks: an overview of sensor devices for detecting human movements, analysis of technologies for hand gesture detection and recognition (specifically Create ML (Apple), Google Cloud AI Platform, and Microsoft Azure). A functional model of the gesture recognition system was created, with the main research focusing on the neural network analysis module of the enhanced input data.

The proposed system is based on a new hybrid approach with adaptive computation distribution, combining local processing (Create ML) and cloud computing (Google Cloud AI Platform). This allows for optimal balancing of performance and latency in gesture processing. The use of an optimized camera topology and adaptive gesture processing enhances accuracy and performance, as confirmed by experimental data. These results are significant for the further development of human-machine interaction in smart home systems and contactless control, as well as contributing to the scientific novelty of the conducted experimental studies.

The results of the conducted research showed that the highest accuracy in hand gesture recognition, with an open palm and a clenched fist, can be achieved with 6 layers and 50 training epochs, reaching an accuracy of 97 % with test losses of 0.17. The study on the placement of gadgets relative to the object under investigation demonstrated that the diagonal arrangement of gadgets with predefined technologies has an advantage over the use of a single technology and over the parallel arrangement of gadgets with identical technologies by 0.62 %.

The combined approach, which integrates Apple's Create ML technologies and Google Cloud AI Platform, demonstrates higher efficiency compared to using only one of them. Specifically, the diagonal camera arrangement ensures better gesture recognition accuracy. This confirms the advantages of the hybrid method and the importance of optimal placement of sensor devices to enhance the accuracy of the gesture recognition system.

Thus, the development and improvement of gesture recognition systems open new possibilities for contactless smart home control. The use of a hybrid approach based on Create ML and Microsoft Azure significantly improves gesture recognition accuracy, especially with optimal sensor placement. The obtained results confirm the potential of combining different technologies and can be used for further research and practical implementations in the field of gesture control.

ISSN-L 2664-9969; E-ISSN 2706-5448

Conflict of interest

The authors declare that they have no conflict of interest regarding this research, including financial, personal, authorship, or any other kind of conflict that could influence the research and its results presented in this article.

Financing

The study was conducted without financial support.

Data availability

Data will be provided upon reasonable request.

Use of artificial intelligence

The authors used artificial intelligence technologies within acceptable limits to provide their own verified data, as described in the methodology section of the research.

References

1. Pangarkar, T. (2025). *Smart Home Statistics 2025 By Technology, Features, Security*. Scoop Market. Available at: https://scoop.market.us/smart-home-statistics/ Last accessed: 07.02.2025
2. Taiwo, O., Ezugwu, A. E., Oyelade, O. N., Almutairi, M. S. (2022). Enhanced Intelligent Smart Home Control and Security System Based on Deep Learning Model. *Wireless Communications and Mobile Computing, 2022,* 1–22. https://doi.org/10.1155/2022/9307961
3. Netinant, P., Utsanok, T., Rukhiran, M., Klongdee, S. (2024). Development and Assessment of Internet of Things-Driven Smart Home Security and Automation with Voice Commands. *IoT, 5 (1),* 79–99. https://doi.org/10.3390/iot5010005
4. Prieto González, L., Fensel, A., Gómez Berbís, J. M., Popa, A., de Amescua Seco, A. (2021). A Survey on Energy Efficiency in Smart Homes and Smart Grids. *Energies, 14 (21),* 7273. https://doi.org/10.3390/en14217273
5. Touqeer, H., Zaman, S., Amin, R., Hussain, M., Al-Turjman, F., Bilal, M. (2021). Smart home security: challenges, issues and solutions at different IoT layers. *The Journal of Supercomputing, 77 (12),* 14053–14089. https://doi.org/10.1007/s11227-021-03825-1
6. Malik, S., Shafqat, W., Lee, K.-T., Kim, D.-H. (2021). A Feature Selection-Based Predictive-Learning Framework for Optimal Actuator Control in Smart Homes. *Actuators, 10 (4),* 84. https://doi.org/10.3390/act10040084
7. Aksak, N. H., Barkovska, O. Yu., Ivashchenko, H. S. (2017). Development of the hand gesture recognition system on the basis of clonal selection model. *Systemy Obrobky Informatsii, 3 (149),* 76–80. https://doi.org/10.30748/soi.2017.149.15
8. Barkovska, O., Serdechnyi, V. (2024). Intelligent assistance system for people with visual impairments. *Innovative technologies and scientific solutions for industries, 2 (28)*, 6–16. https://doi.org/10.30837/2522-9818.2024.28.006
9. Barkovska, O., Oliinyk, D., Sorokin, A., Zabroda, I., Sedlaček, P. (2024). A System for monitoring the progress of rehabilitation of patients with musculoskeletal disorder. *Advanced Information Systems, 8 (3),* 13–24. https://doi.org/10.20998/2522-9052.2024.3.02
10. Ahmed, S., Kallu, K. D., Ahmed, S., Cho, S. H. (2021). Hand Gestures Recognition Using Radar Sensors for Human-Computer-Interaction: A Review. *Remote Sensing, 13 (3),* 527. https://doi.org/10.3390/rs13030527
11. Bhushan, S., Alshehri, M., Keshta, I., Chakraverti, A. K., Rajpurohit, J., Abugabah, A. (2022). An Experimental Analysis of Various Machine Learning Algorithms for Hand Gesture Recognition. *Electronics, 11 (6),* 968. https://doi.org/10.3390/electronics11060968
12. Indriani, Harris, Moh., Agoes, A. S. (2021). Applying Hand Gesture Recognition for User Guide Application Using MediaPipe. *Proceedings of the 2nd International Seminar of Science and Applied Technology (ISSAT 2021).* https://doi.org/10.2991/aer.k.211106.017
13. Mujahid, A., Awan, M. J., Yasin, A., Mohammed, M. A., Damaševičius, R., Maskeliūnas, R., Abdulkareem, K. H. (2021). Real-Time Hand Gesture Recognition Based on Deep Learning YOLOv3 Model. *Applied Sciences, 11 (9),* 4164. https://doi.org/10.3390/app11094164
14. Sahoo, J. P., Prakash, A. J., Pławiak, P., Samantray, S. (2022). Real-Time Hand Gesture Recognition Using Fine-Tuned Convolutional Neural Network. *Sensors, 22 (3),* 706. https://doi.org/10.3390/s22030706
15. Jiang, S., Kang, P., Song, X., Lo, B., Shull, P. (2022). Emerging Wearable Interfaces and Algorithms for Hand Gesture Recognition: A Survey. *IEEE Reviews in Biomedical Engineering, 15,* 85–102. https://doi.org/10.1109/rbme.2021.3078190
16. Guo, L., Lu, Z., Yao, L. (2021). Human-Machine Interaction Sensing Technology Based on Hand Gesture Recognition: A Review. *IEEE Transactions on Human-Machine Systems, 51 (4),* 300–309. https://doi.org/10.1109/thms.2021.3086003
17. Liu, H., Zhu, X., Lei, Z., Cao, D., Li, S. Z. (2021). Fast Adapting Without Forgetting for Face Recognition. *IEEE Transactions on Circuits and Systems for Video Technology, 31 (8),* 3093–3104. https://doi.org/10.1109/tcsvt.2020.3035890
18. Uddin, Md. S., Slany, W., Hasan, K. J. (2024). On-Device Neural Network for Object Train and Recognition using Mobile. *International Journal of Interactive Mobile Technologies (IJIM), 18 (12),* 99–111. https://doi.org/10.3991/ijim.v18i12.47895
19. Van der Vlist, F., Helmond, A., Ferrari, F. (2024). Big AI: Cloud infrastructure dependence and the industrialisation of artificial intelligence. *Big Data & Society, 11 (1).* https://doi.org/10.1177/20539517241232630
20. Patel, D., Raut, G., Cheetirala, S. N., Nadkarni, G. N., Freeman, R., Glicksberg, B. S. et al. (2024). *Cloud Platforms for Developing Generative AI Solutions: A Scoping Review of Tools and Services.* arXiv preprint arXiv:2412.06044. https://doi.org/10.48550/arXiv.2412.06044
21. Chornenkyi, V. Ya., Kazymyra, I. Ya. (2023). Research of the models for sign gesture recognition using 3D convolutional neural networks and visual transformers. *Ukrainian Journal of Information Technology, 5 (2),* 33–40. https://doi.org/10.23939/ujit2023.02.033
22. Zhang, G., Ma, N., Li, J., Jiang, B., Wu, Z. (2021). Gesture Recognition Algorithm Based on Lightweight 3DCNN Network. *2021 17th International Conference on Computational Intelligence and Security (CIS),* 217–221. https://doi.org/10.1109/cis54983.2021.00053
23. Osman Hashi, A., Zaiton Mohd Hashim, S., Bte Asamah, A. (2024). A Systematic Review of Hand Gesture Recognition: An Update From 2018 to 2024. *IEEE Access, 12,* 143599–143626. https://doi.org/10.1109/access.2024.3421992
24. Openja, M., Nikanjam, A., Yahmed, A. H., Khomh, F., Jiang, Z. M. J. (2022). An Empirical Study of Challenges in Converting Deep Learning Models. *2022 IEEE International Conference on Software Maintenance and Evolution (ICSME),* 13–23. https://doi.org/10.1109/icsme55016.2022.00010
25. Wakchaure, A., Kanawade, P., Jawale, M. A., William, P., Pawar, A. B. (2022). Face Mask Detection in Realtime Environment using Machine Learning based Google Cloud. *2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC),* 557–561. https://doi.org/10.1109/icaaic53929.2022.9793201
26. Abdul-Al, M., Kyeremeh, G., Qahwaji, R., Abd-Alhameed, R. (2022). The Impacts of Depth Camera and Multi-Sensors for Face and Fingerprint Recognition in 2D and 3D – A Case Study. *Proceedings of 2nd International Multi-Disciplinary Conference Theme: Integrated Sciences and Technologies, IMDC-IST 2021.* Sakarya. https://doi.org/10.4108/eai.7-9-2021.2314977
27. Matilainen, M., Sangi, P., Holappa, J., Silven, O. (2016). OUHANDS database for hand detection and pose recognition. *2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA).* Oulu, 1–5. https://doi.org/10.1109/ipta.2016.7821025

✉***Olesia Barkovska,*** *PhD, Associate Professor, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine, e-mail: olesia.barkovska@nure.ua, ORCID: https://orcid.org/0000-0001-7496-4353*

Igor Ruban, *Doctor of Technical Sciences, First Vice-Rector, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine, ORCID: https://orcid.org/0000-0002-4738-3286*

Daria Tymoshenko, *Assistant, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine, ORCID: https://orcid.org/0000-0001-5514-7075*

Oleksandr Holovchenko, *Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine, ORCID: https://orcid.org/0009-0002-7582-1746*

Oleksandr Yankovskyi, *PhD, Associate Professor, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine, ORCID: https://orcid.org/0000-0002-1268-0029*

✉*Corresponding author*