

Valerii Yarovy, Olesia Barkovska, Dmytro Maksymov, Yana Ni, Daniil Raptanov

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

## METHODS FOR RECOGNISING HAND GESTURES USING COMPUTER VISION TO ASSIST PEOPLE WITH MOVEMENT OR SPEECH IMPAIRMENTS

**Abstract. Relevance.** Real-time hand gesture recognition systems play an important role in the development of modern technologies for human interaction with the “smart environment.” They create conditions for contactless device control and open up new communication opportunities for people with speech or motor impairments. **The object of research.** The object of research is a real-time hand gesture recognition system using computer vision and machine learning methods. The research problem is to create an accessible and effective means of communication for people with speech or motor impairments, as well as to expand the possibilities of controlling devices and elements of a “smart home” without physical contact or voice commands. **Purpose of the article.** The purpose of this article is to investigate and compare different methods for real-time hand gesture recognition using computer vision and machine learning. Specifically, the study evaluates the performance of MediaPipe Hands versus YOLOv8 for detecting and classifying hand gestures. The aim is to identify which approach provides higher accuracy and better adaptability for applications such as assistive communication for people with speech or motor impairments. **Research results.** In the course of the work, a prototype system was developed that combines video processing technologies (GStreamer or FFmpeg), the MediaPipe Hands library for determining key points of the hand, and CNN or LSTM deep learning models for classifying gestures and movement sequences. The results demonstrated the possibility of accurate gesture recognition in real time, as well as flexible adaptation to individual user characteristics. **Conclusions.** Interpretation of the results showed that the effectiveness of the system is due to the use of a comprehensive approach: the combination of 3D coordinate detection of key hand points with a neural network allowed high stability to be achieved even under changing lighting or background conditions. A distinctive feature of the proposed solution is the ability to recognise sequences of gestures that form phrases or commands (for example, “I want to drink”), as well as a function for teaching user gestures to expand the system’s individual vocabulary. The developed technology can be used in assistive communication systems for people with disabilities, in medical and rehabilitation facilities, as well as for controlling smart home elements. Its implementation contributes to increasing the accessibility of digital technologies and improving the quality of life of users.

**Keywords:** gesture, recognition, computer vision, MediaPipe Hands, neural networks, assistive technologies.

### Introduction

In today’s world of rapid technological development, systems that provide comfortable HCI (human-computer interaction) occupy a special place [1]. One of the most promising areas in this field is hand gesture recognition using computer vision. The scientific subject of the work combines methods of artificial intelligence, machine learning, biomechanics, and cognitive sciences. This area opens up new opportunities for creating interfaces that perceive a person not as

a user or operator of a device, but as an active participant in the communication process, since non-verbal communication plays a decisive role in our daily interactions, conveying approximately 65% of human messages, compared to verbal communication, which accounts for only 35% [2].

The scientific relevance of this topic lies in the fact that gesture recognition is one of the key elements in the development of intelligent assistive systems, human-centred technologies and the “smart environment” (Fig. 1).

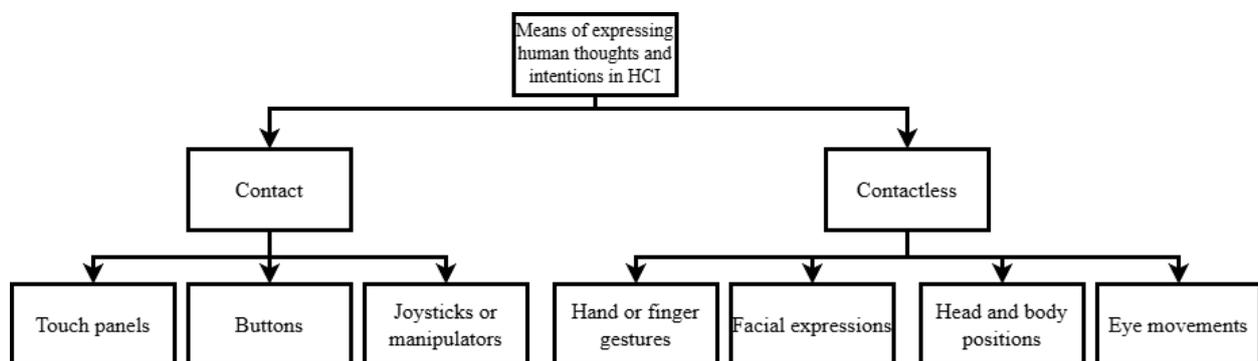


Fig. 1. Classification of means of expressing human thoughts and intentions in HCI

Among the new forms of interaction, gesture-based interfaces are becoming increasingly popular due to their intuitiveness, efficiency, and flexibility [3].

Gestures are a special means of expressing thoughts and intentions that do not require words or physical contact. It is a type of body language in which the

position and shape of the palm and fingers convey certain information [4]. Systems capable of interpreting hand gestures can provide a new level of accessibility and autonomy for people with disabilities, as well as increase the efficiency of controlling technical devices in everyday life, medicine, or even industry.

Gesture recognition methods are usually divided into two main types (Fig. 2).

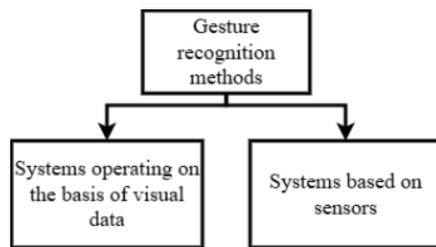


Fig. 2. Existing gesture recognition methods

Visual approaches use cameras to track hand movements or poses and, if necessary, depth sensors to provide more accurate three-dimensional analysis [5].

Gesture recognition has a number of specific features that must be taken into account when developing intelligent interaction systems.

Unlike standard users, the gestures of persons with disabilities may:

- differ in amplitude, trajectory, speed, and stability of movement;
- have a limited range of motion;
- be compensatory (e.g., in Tourette's syndrome), which complicates classification.

Therefore, systems must adapt to the individual motor skills of the user, analysing not only the absolute coordinates of the fingers, but also the relative proportions and dynamics of movements over time.

At the same time, analysis of existing solutions reveals a number of systemic limitations (Table 1).

Table 1 – Limitations of existing gesture recognition systems

No.	Limitation	Explanation
1	System flexibility	The number of gestures needs to be optimised, as experimental prototypes typically use a fairly limited vocabulary of gestures.
2	Accuracy and reliability of recognition	Ensuring high accuracy and reliability of algorithms in real operating conditions is one of the key problems in gesture recognition [6].
3	Response time	This manifests itself in delays between the execution of a gesture and the system's response, creating a feeling of "slow" interaction when the user is forced to wait for the system's reaction or repeat a certain gesture.
4	Lack of context for gesture execution	The system can recognise the shape of the hand, but does not understand in which environment or for what purpose the gesture is being used.
5	Not adapted for people with disabilities	The system's algorithms are designed for standard range and speed of movement, which makes them inaccessible to people with motor impairments.
6	Limitations on gesture sequence recognition	The systems are unable to correctly interpret complex or rapid combinations of movements – they "lose" part of the gestures, confuse their order, or respond only to individual movements rather than their logical sequence.

The practical significance of gesture recognition lies in its wide range of applications:

- in virtual environments, gesture recognition technologies create more natural and realistic interfaces;
- in robotics, it provides contactless control of manipulators;
- in medicine and rehabilitation, it forms the basis of assistive solutions for people with speech or motor impairments;
- in smart home systems, such technologies increase the comfort, safety, and level of personalisation of living space, allowing the user to interact with appliances without touch or voice commands.

In automated solutions, the hand is considered the main area for data collection [6]. Thus, research in the field of hand gesture recognition using computer vision and machine learning is an extremely relevant area of modern science. It is important both for the development of the theoretical foundations of computer vision and for the practical creation of technologies that improve the quality of life of people, especially those with physical limitations. That is why the study, improvement, and implementation of gesture recognition methods is a promising and socially significant task that requires further scientific research.

#### Review of recent studies and publications.

Among the external factors, technical characteristics of the hardware, and physical characteristics of the user that limit the accuracy of gesture recognition systems, the following can be highlighted:

- analysis of the shape of the object, which makes it possible to track the position of the hands, changes in the configuration of the fingers and the trajectory of movements in space;

- the parameters of the camera's optoelectronic path – resolution, frame rate, dynamic range, and sensor sensitivity determine image quality. For example, if the camera is located outside the room, in conditions of excessive or uneven lighting, the YOLOv8 model may experience failures in gesture recognition [7].

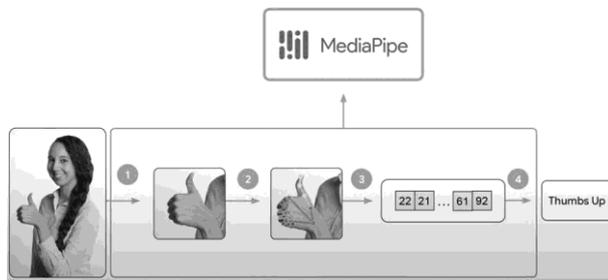
- lighting conditions and optical distortions – changes in light intensity, the appearance of shadows, blurring or distortion of perspective can significantly affect the accuracy of hand contour detection;

- Skin texture and colour – recognising colour and tonal characteristics makes it possible to separate hands from the background or other objects in the frame, which increases the stability of the algorithm in real conditions.

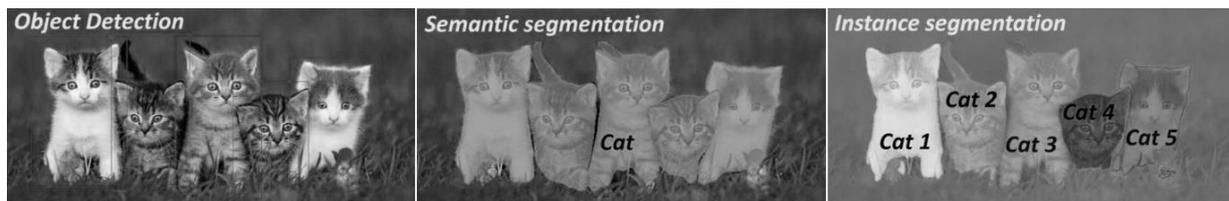
Given that the camera parameters are known before the start of the study, the results need to be interpretable and a decision needs to be made as quickly as possible for objects such as human gestures.

The MediaPipe Hands method is based on a combination of geometric principles and elements of neural image analysis. The algorithm operates in two consecutive stages: first, the BlazePalm Detector module determines the area of the frame where the palm is located, after which the Hand Landmark Model calculates the coordinates of 21 control points of the hand

in three-dimensional space. Unlike classic deep neural networks, MediaPipe supports both 2D and 3D key points and requires few computational resources, making it ideal for lightweight, real-time applications [8].



**Fig. 3.** Sequence of the MediaPipe Hands library gesture recognition method



**Fig. 4.** Sequence of the YOLOv8 model object recognition method

The structure of YOLOv8 is divided into two key components: the base network and the detection head. The base network extracts various rich features from the input image at different scales. On the other hand, the detection head takes on the task of combining these features and generating various high-quality predictions for bounding boxes [10].

YOLOv8 is distinguished by its high recognition accuracy, thanks to the improved CSPDarknet and PAN-FPN architectures, which ensure effective hand detection even in cases of overlap, lighting fluctuations, or small object sizes. Another important feature of YOLOv8 is its ability to work in real time. Optimisation of the inference process for graphics processing units (GPUs) provides speeds of up to 60-120 frames per second, allowing gestures to be tracked with high accuracy even when the user is actively moving.

At the same time, YOLOv8 requires significant computing resources and does not provide direct determination of three-dimensional joint coordinates, as implemented in MediaPipe.

**The purpose of this work** is to compare two methods of hand gesture recognition and analyse the influence of external factors on the accuracy of gesture detection indoors. To do this, we plan to use a heuristic method based on the MediaPipe library and the YOLOv8 neural network, which implements a modern approach to object recognition.

The main idea is to determine which of the approaches - traditional (MediaPipe, using rules and coordinates of key points of the hand) or neural network (YOLOv8, with training on images/videos) provides higher accuracy, stability, and speed under different shooting conditions (lighting, background, resolution, and the peculiarities of gesture display by a person with a disability). To achieve the set goal, the following tasks must be solved:

The MediaPipe Hands algorithm is characterised by highly interpretable results, as it returns the coordinates of key points on the hand, allowing the position of joints and the angular relationships between fingers to be visualised and possible recognition errors to be controlled. The system remains stable to changes in lighting and noise thanks to the spatial structure of the hand analysis, which ensures correct operation even when the background changes or is partially shaded – from 1200 to 100 lumens.

The YOLOv8 (You Only Look Once, version 8) method implements a deep learning approach based on convolutional neural networks (CNN) for fast and accurate object recognition [9]. CNNs, in turn, allow you to simultaneously detect objects and determine their key points (pose estimation) in real time.

1. Justify the choice of gesture recognition methods, conduct an analytical review of modern approaches to gesture detection and classification based on computer vision. Consider the advantages and limitations of MediaPipe (heuristic approach using key points) and YOLOv8 (deep learning based on images/videos).

2. Determine the specifics of the influence of external factors. Investigate how changes in shooting conditions – lighting, camera resolution, noise level, viewing angle, as well as individual characteristics of users (in particular, people with motor impairments) affect the accuracy of gesture recognition.

3. Develop a software prototype of a gesture recognition module based on the results of a comparative analysis of recognition methods.

4. Conduct a series of comparative tests of both methods under controlled and variable conditions (with different levels of lighting, noise, and different positions of a given gesture). Determine the accuracy, speed, and stability of the systems.

5. Perform a comparative analysis of the effectiveness of the two approaches, determine the optimal conditions for their use, and formulate recommendations for the practical application of each method in assistive systems for people with disabilities.

Further development of the system – analyse potential additional functions (e.g., voice messages that communicate the user's intentions based on the semantic sequence of gesture commands). Consider integrating the system with assistive technologies or a "smart home" to ensure user autonomy.

### Main part

A detailed model of the hand gesture recognition process is shown in figure 3. The model includes five interrelated functional modules:

- image capture – a video stream from a camera is used, which provides continuous data updates for further processing. At this stage, individual frames suitable for analysis are formed;

- hand detection – each frame is analysed to determine the coordinates of the hand within the image. For this purpose, detection algorithms based on neural network models are used, which allow the area of the hand to be accurately localised even in difficult lighting conditions;

- context detection – the scene is analysed to take into account the environment, the position of the hand relative to other objects or the user. This allows for increased recognition accuracy by taking additional factors into account;

- gesture detection – performed based on hand coordinates and scene context. The feature description algorithm generates descriptors that reflect finger configuration, palm orientation, and movement characteristics.

- gesture classification – completes the recognition process. At this stage, the system uses a classifier and computational models to analyse the obtained descriptors and determine the corresponding gesture class. The result of the module’s work is the identification of a sequence of hand gestures over time.

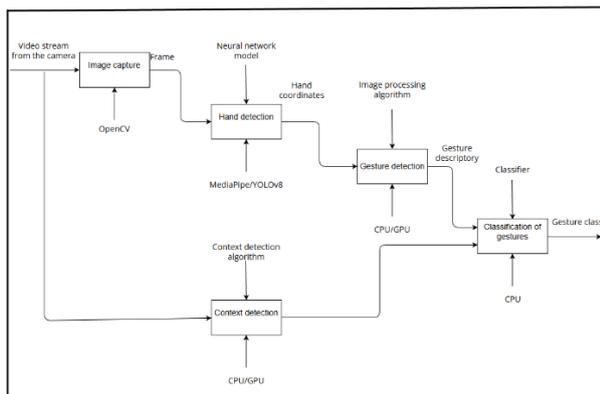


Fig. 5. IDEF0 diagram of the system

The order of the experiments is as follows: we record the actual distance from the camera to the hand

(using a ruler or laser rangefinder), place the camera in a stable position (tripod or fixed surface), launch the gesture recognition algorithm (MediaPipe Hands or YOLO), perform a set of standard gestures at different distances and under different lighting conditions, record the recognition results and compare them with real gestures, determining the accuracy and error (Fig. 6).

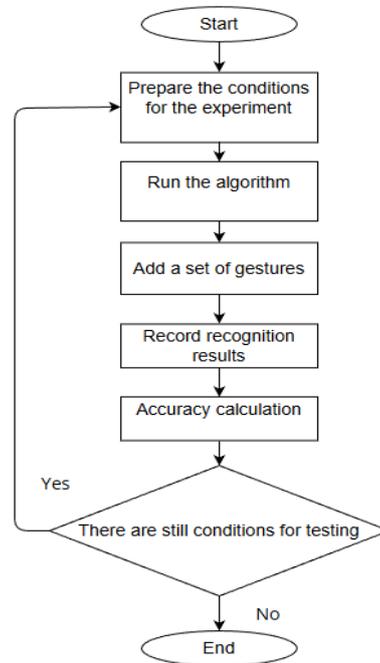


Fig. 6. Flowchart of the experiment algorithm

The experiment investigates the effect of camera resolution (320x240, 640x480) and lighting conditions (bright – 1200 lumens, medium – 500 lumens, low – 100 lumens) on the accuracy of gesture recognition (Fig. 7). The hypothesis of the experiment is that the further the object is from the camera, the more difficult it is for the system to correctly identify the gesture, as the number of pixels that “describe” the hand decreases. In addition, in poor lighting conditions, the algorithm may incorrectly recognise contours and key points, which also reduces the accuracy of gesture recognition.

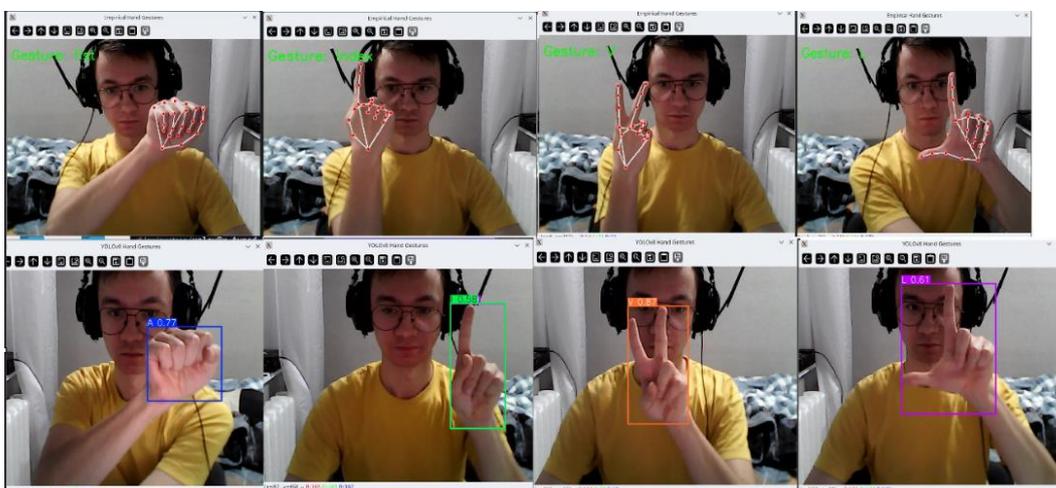
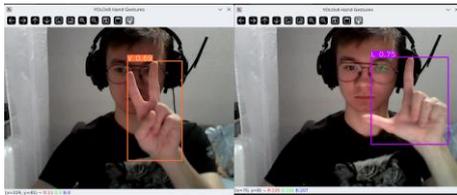
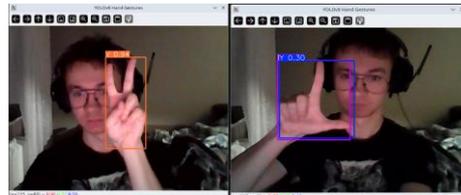


Fig. 7. Set of gestures recognised by the implemented systems based on MediaPipe Hands and YOLOv8 models



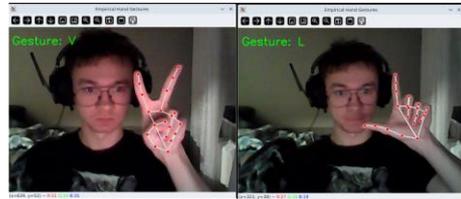
**Fig. 8.** Demonstration of the experiment conducted on recognising the V and L gestures using the YOLOv8 model in good lighting conditions (approximately 1200 lumens)



**Fig. 11.** Demonstration of the experiment conducted on the recognition of V and L gestures using the YOLOv8 model in average lighting conditions (approximately 500 lumens)

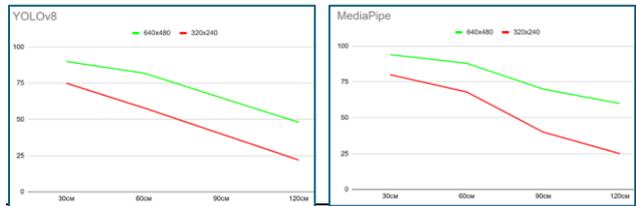


**Fig. 9.** Demonstration of the experiment conducted on recognising the V and L gestures based on MediaPipe methods in good lighting conditions (approximately 1200 lumens)



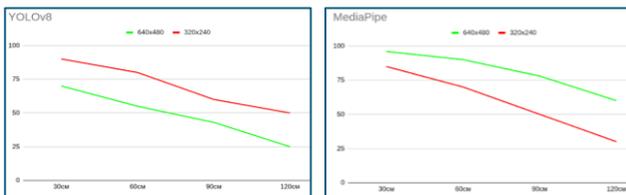
**Fig. 12.** Demonstration of the experiment conducted on the recognition of V and L gestures based on MediaPipe methods in average lighting conditions (approximately 500 lumens)

From the results of the experiments shown in figure 10, we can see that when the resolution is reduced, the YOLOv8 model demonstrates a higher recall rate – that is, it better detects the presence of hands or gestures even in less clear images. This is because YOLOv8 focuses on detecting objects as a whole, using contour and spatial features that remain visible even at low pixel counts. At the same time, MediaPipe Hands did not lose its ability to identify key points of the hand – even at lower resolutions, the algorithm continued to correctly construct the skeleton of the hand and accurately calculate the position of the fingers. Thanks to this, MediaPipe maintained high accuracy during further calculations and gesture classification, as it relied not only on the shape of the object, but also on the relative coordinates of the anatomical points of the hand.



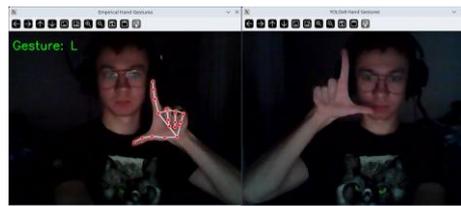
**Fig. 13.** Graphs showing the effect of input image resolution (480p and 240p) on gesture recognition accuracy under medium lighting conditions (approximately 500 lumens), left – YOLOv8 model, right – MediaPipe model

Fig. 13 shows that the YOLOv8 model performs significantly worse in recognising gestures under average lighting conditions at low resolution, while MediaPipe continues to demonstrate fairly stable results and maintains acceptable accuracy

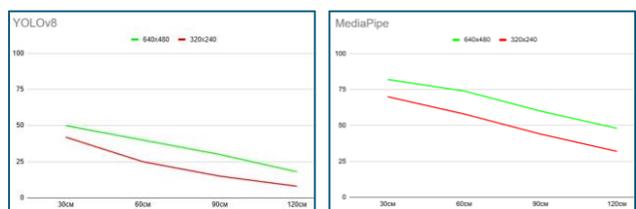


**Fig. 10.** Graphs showing the impact of input image resolution (480p and 240p) on gesture recognition accuracy under good lighting conditions (approximately 1200 lumens), left – YOLOv8 model, right – MediaPipe model

Based on the experimental data obtained, it can be concluded that both models – YOLOv8 and MediaPipe Hands – demonstrate a decrease in gesture recognition accuracy with an increase in distance to the camera and a decrease in resolution. At the same time, YOLOv8 shows higher stability at low resolution (320x240), maintaining relatively high accuracy values even at medium distances, which indicates the effectiveness of its object detector. However, at high resolution (640x480), MediaPipe significantly outperforms YOLOv8, providing recognition accuracy of over 90% at close and medium distances (up to 60 cm). This is because MediaPipe uses an anatomical model of the hand.



**Fig. 14.** Demonstration of the experiment conducted on gesture L recognition based on MediaPipe methods (left) and the YOLOv8 model (right) in low light conditions (approximately 100 lumens)



**Fig. 15.** Graphs showing the effect of lighting changes on gesture recognition accuracy, left – YOLOv8 model, right – MediaPipe model

At a lighting level of about 100 lumens, the accuracy of both methods decreases significantly. YOLOv8 loses

stability even at short distances and almost does not recognise gestures at a distance of 90-120 cm, especially at a resolution of 240p. MediaPipe, on the other hand, although it shows a drop in accuracy, still maintains an acceptable level of recognition at short and medium distances, which indicates its better stability in low light conditions. Thus, we can conclude that the MediaPipe model is better suited to changing shooting conditions and is a more versatile tool for recognition in complex visual environments.

The next stage of the study was to compare the accuracy of MediaPipe and YOLOv8 models in affine transformations. In this experiment, the hands were positioned at different angles relative to the camera – both frontally and at an angle, partially covering the fingers or extending beyond the frame.



**Fig. 16.** MediaPipe model (left) and YOLOv8 model (right) at different hand tilt angles

Fig. 16 shows that the MediaPipe model demonstrates stable results thanks to the use of a method based not only on visual detection of the hand, but also on calculations of distances and angles between key points. This approach allows for the analysis of spatial relationships between the fingers and the palm, ensuring high accuracy when lighting changes or the hand is partially obscured. However, it should be noted that MediaPipe performs worse with strong hand tilts, as its algorithm is empirical in nature – the angles and distances between points depend on the projection of the hand onto the image plane. Because of this, when the tilt angle changes significantly, the calculated parameters also change, which can lead to incorrect gesture recognition. At the same time, even at angles of 45 and 90°, MediaPipe still retains the ability to consistently track key points and correctly recognise gestures. In contrast, YOLOv8 demonstrates lower overall accuracy, likely due to insufficient model training or a limited number of variations in the training sample. If the model observed only a narrow set of hand positions during training, it generalises new situations less well – different angles,

lighting or partial overlap. This is especially noticeable at angles of 45° and even more so at 90°, where YOLOv8 quite often fails to recognise the gesture at all, highlighting its sensitivity to affine transformations and the lack of variability in the training data. The experiment showed that MediaPipe demonstrates better results in almost all conditions – it consistently determines the position of the hand even in average or poor lighting. However, its accuracy may decrease with significant hand tilts or partial finger overlap, when it is difficult for the algorithm to correctly construct a spatial model. In turn, YOLOv8 is better suited for rapid classification of visual images, so it is advisable to combine the advantages of both methods. The optimal solution may be a hybrid system in which MediaPipe is used to highlight the hand and determine key points, and YOLOv8 or another classifier is used to analyse these points and finally recognise the gesture. This approach will improve the accuracy and stability of the recognition system in real-world conditions.

## Conclusions

The study resulted in a comparative analysis of two models for hand gesture recognition – MediaPipe Hands and YOLOv8 – taking into account the influence of external factors such as lighting, resolution, and spatial position of the hand. The resulting graphs of the experiments show that the MediaPipe model provides higher stability and accuracy under various conditions due to the use of an anatomical hand model, while YOLOv8 demonstrates better resistance to low resolution but significantly loses accuracy in difficult lighting conditions and affine transformations. Analysis of the impact of external conditions confirmed that the quality of gesture recognition is determined by the availability of visual information: poor or overly bright lighting, low video signal quality, and a strong tilt of the hand reduce the accuracy of both models, but MediaPipe responds to such changes more confidently and predictably. The YOLOv8 model, on the other hand, showed a high dependence on the diversity of training data, which limits its performance in cases of non-standard gestures or when the system is used by people with motor impairments.

The software prototype confirmed that no single approach is universal, and the best results can be achieved by combining the strengths of both models. The optimal solution is a hybrid system in which MediaPipe Hands are used to accurately determine key points of the hand, and a neural network classifier is used to interpret gestures. This opens opportunities for the creation of adaptive assistive systems that can work in real-world conditions, supporting gesture sequence recognition, and increasing the level of autonomy and accessibility for people with disabilities.

## REFERENCES

1. Barkovska, O., Shapiro, A., Mavrynskyi, O., & Zhebin, P. (2025). Research on the sensitivity of the distance measurement method based on the facemesh algorithm. *Системи управління, навігації та зв'язку. Збірник наукових праць*, 2(80), 76-82. <https://doi.org/10.26906/SUNZ.2025.2.076>
2. J. Shin, A. S. M. Miah, M. H. Kabir, M. A. Rahim and A. Al Shiam, "A Methodological and Structural Review of Hand Gesture Recognition Across Diverse Data Modalities," in *IEEE Access*, vol. 12, pp. 142606-142639, 2024, <https://doi.org/10.1109/ACCESS.2024.3456436>.
3. Zhang, T.; Wang, Y.; Zhou, X.; Liu, D.; Ji, J.; Feng, J. Intelligent Human-Computer Interaction for Building Information Models Using Gesture Recognition. *Inventions* 2025, 10, 5. <https://doi.org/10.3390/inventions10010005>.

4. Al Farid, F.; Hashim, N.; Abdullah, J.; Bhuiyan, M.R.; Shahida Mohd Isa, W.N.; Uddin, J.; Haque, M.A.; Husen, M.N. A Structured and Methodological Review on Vision-Based Hand Gesture Recognition System. *Journal of Imaging* 2022, 8, 153. <https://doi.org/10.3390/jimaging8060153>.
5. Marques, P.; Váz, P.; Silva, J.; Martins, P.; Abbasi, M. Real-Time Gesture-Based Hand Landmark Detection for Optimised Mobile Photo Capture and Synchronisation. *Electronics* 2025, 14, 704. <https://doi.org/10.3390/electronics14040704>.
6. 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>
7. M. Al-Hammadi et al., "Deep Learning-Based Approach for Sign Language Gesture Recognition With Efficient Hand Gesture Representation," in *IEEE Access*, vol. 8, pp. 192527-192542, 2020, <https://doi.org/10.1109/ACCESS.2020.3032140>.
8. Li, H.-H., & Hsieh, C.-C. (2025). Dynamic Hand Gesture Recognition Using MediaPipe and Transformer. *Engineering Proceedings*, 108(1), 22. <https://doi.org/10.3390/engproc2025108022>
9. Sun, Y., Zhang, Y., Wang, H. et al. SES-YOLOv8n: automatic driving object detection algorithm based on improved YOLOv8. *SIVIP* 18, 3983–3992 (2024). <https://doi.org/10.1007/s11760-024-03003-9>
10. R. Varghese and S. M., "YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness," *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*, Chennai, India, 2024, pp. 1-6, <https://doi.org/10.1109/ADICS58448.2024.10533619>

Received (Надійшла) 18.11.2025

Accepted for publication (Прийнята до друку) 28.01.2026

Publication date (Дата публікації) 27.02.2026

#### ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

**Яровий Валерій Вікторович** – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

**Valerii Yarovy** – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine; e-mail: [valerii.iarovy@nure.ua](mailto:valerii.iarovy@nure.ua); ORCID Author ID: <https://orcid.org/0009-0008-3631-8586>.

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

**Olesia Barkovska** – Associate Professor, Department of Electronic Computers, Candidate of Technical Sciences, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine; e-mail: [olesia.barkovska@nure.ua](mailto:olesia.barkovska@nure.ua); ORCID Author ID: <https://orcid.org/0000-0001-7496-4353>.

**Максимов Дмитро Максимович** – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

**Dmytro Maksymov** – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Ukraine; e-mail: [dmytro.maksymov2@nure.ua](mailto:dmytro.maksymov2@nure.ua); ORCID Author ID: <https://orcid.org/0009-0002-5206-3233>.

**Ні Яна Самвелівна** – доцент кафедри електронних обчислювальних машин, кандидат технічних наук, Харківський національний університет радіоелектроніки, Харків, Україна;

**Yana Ni** – Associate Professor, Department of Electronic Computers, Candidate of Technical Sciences, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine; e-mail: [yana.movsesian@nure.ua](mailto:yana.movsesian@nure.ua); ORCID Author ID: <https://orcid.org/0000-0002-1352-700X>.

**Раптанов Данііл Андрійович** – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

**Daniil Raptanov** – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine; e-mail: [daniil.raptanov@nure.ua](mailto:daniil.raptanov@nure.ua); ORCID Author ID: <https://orcid.org/0009-0001-9564-0080>.

#### Методи розпізнавання жестів рук із використанням комп'ютерного зору для допомоги людям з порушеннями руху або мовлення

В. В. Яровий, О. Ю. Барковська, Д. М. Максимов, Я. С. Ні, Д. А. Раптанов

**Анотація. Актуальність.** Системи розпізнавання жестів рук у реальному часі відіграють важливу роль у розвитку сучасних технологій взаємодії людини з «розумним середовищем». Вони створюють умови для безконтактного керування пристроями та відкривають нові можливості комунікації для людей з порушеннями мовлення або руховими порушеннями. **Об'єктом дослідження** є система розпізнавання жестів рук у реальному часі з використанням методів комп'ютерного зору та машинного навчання. **Метою статті** є дослідження та порівняння різних методів розпізнавання жестів рук у реальному часі з використанням комп'ютерного зору та машинного навчання. Зокрема, у дослідженні оцінюється ефективність MediaPipe Hands та YOLOv8 для виявлення та класифікації жестів рук. Мета полягає у визначенні, який підхід забезпечує вищу точність та кращу адаптивність для таких застосувань, як допоміжна комунікація для людей з порушеннями мовлення або моторики. **Результати досліджень.** У процесі роботи розроблено прототип системи, що поєднує технології відеообробки (GStreamer або FFmpeg), бібліотеку MediaPipe Hands для визначення ключових точок руки, та моделі глибинного навчання типу CNN або LSTM для класифікації жестів і послідовностей рухів. Отримані результати засвідчили можливість точного розпізнавання жестів у реальному часі, а також гнучку адаптацію до індивідуальних особливостей користувача. **Висновки.** Інтерпретація результатів показала, що ефективність системи пояснюється використанням комплексного підходу: поєднання детекції 3D-координат ключових точок руки з нейронною мережею дозволило досягти високої стабільності навіть за змінних умов освітлення чи фону. Розроблена технологія може бути застосована в асистивних системах комунікації для людей з інвалідністю, у медичних та реабілітаційних закладах, а також для управління елементами «розумного будинку». Її впровадження сприяє підвищенню рівня доступності цифрових технологій і покращенню якості життя користувачів.

**Ключові слова:** жест, розпізнавання, комп'ютерний зір, MediaPipe Hands, нейронні мережі, асистивні технології.