UDC 378.018.43:004.932

Olesia Barkovska
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
ORCID ID 0000-0001-7496-4353
olesia.barkovska@nure.ua

Artem Bakulin
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
ORCID ID 0009-0003-8885-7673
artem.bakulin@nure.ua

Volodymyr Mytiai
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
ORCID ID 0009-0003-1239-6727
volodymyr.mytiai@nure.ua

Oleksandr Yankovskyi
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
ORCID ID 0000-0002-1268-0029
oleksandr.yankovskyi@nure.ua

Ihor Novoseltsev
Independent Researcher, Carlsbad, CA, USA
ORCID ID 0009-0004-7353-7498
i.v.novoseltsev@gmail.com

Vitalii Serdechnyi
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
ORCID ID 0009-0007-8828-5803
vitalii.serdechnyi@nure.ua

COMPUTER VISION-BASED CONTACTLESS CONTROL INTERFACE FOR EDUCATIONAL ENVIRONMENTS

Abstract. The relevance of this study is determined by the growing role of distance learning systems and the need to expand interaction modalities with them, particularly in situations of limited access to traditional input devices or within the context of inclusive education. The aim of this study is to design and assess the practical feasibility of a computer vision-based contactless interface for LMS control through operating-system input events, considering robustness under varying lighting, distance, viewpoint, and resolution and comparison with conventional mouse-and-keyboard interaction. To achieve this aim, the following tasks were formulated: to develop a model of a gesture-based LMS control method; to implement both a single-hand gesture mode and a bimanual interaction mode for different control scenarios; to conduct a series of experiments evaluating the accuracy and robustness of gesture recognition under varying recording conditions; and to perform a comparative analysis of the effectiveness of gesture-based control versus the classical "mouse and keyboard" interface. The research methods include computer vision techniques based on hand landmark detection, algorithmic interpretation of gestures using geometric and dynamic features, as well as experimental measurements of recognition accuracy, task execution time, and the number of interaction errors. The evaluation involved five static hand poses, 30 repetitions per pose, and three representative interaction scenarios. The results demonstrate that the proposed system provides stable recognition of large static hand poses across a wide range of lighting conditions and camera distances. Changes in hand orientation were identified as the most critical factor affecting recognition accuracy. All three scenarios were successfully completed in the baseline trials at a resolution of 1280×720, with no critical interaction errors recorded. The practical feasibility of gesture-based control was demonstrated through its integration with a real learning management system, dl.nure.ua, via operating system–level input events. The practical significance lies in its applicability for controlling educational web platforms in distance learning environments, during presentations, and in inclusive education scenarios.

© Olesia Barkovska, Artem Bakulin, Volodymyr Mytiai,
Oleksandr Yankovskyi, Ihor Novoseltsev, Vitalii Serdechnyi, 2026.

Keywords: learning management system; gesture-based interaction; distance learning; inclusive interfaces.

1. INTRODUCTION

Statement of the problem. Modern Learning Management Systems (LMS) play a key role in the implementation of distance and blended learning, providing a centralized environment for organizing the educational process, interaction between participants, distribution of educational materials, and assessment of learning outcomes [1]. At the same time, the effectiveness of LMS usage depends significantly not only on the platform's functional capabilities but also on the convenience and flexibility of interaction interfaces, especially in the conditions of a real educational session. This highlights the implementation of artificial intelligence methods at various stages of the educational process [2], [3].

An important factor is the growing attention to touchless and hygienic interaction interfaces, as well as to the inclusivity of digital educational environments [4], [5]. For users with motor impairments, temporary physical limitations, or specific working conditions, traditional control tools can create additional barriers to accessing LMS functionality. In this context, the search for alternative or complementary interfaces capable of reducing reliance on fine motor skills and providing a more natural form of interaction is relevant.

Analysis of recent publications indicates a research gap between the development of alternative interaction interfaces and their systemic application for remote management of the educational process in LMS. This necessitates a systemic analysis of approaches to LMS management in the form of a generalized classification by interaction interface type and usage context, which will allow defining the place of gesture interfaces in this structure (Figure 1).

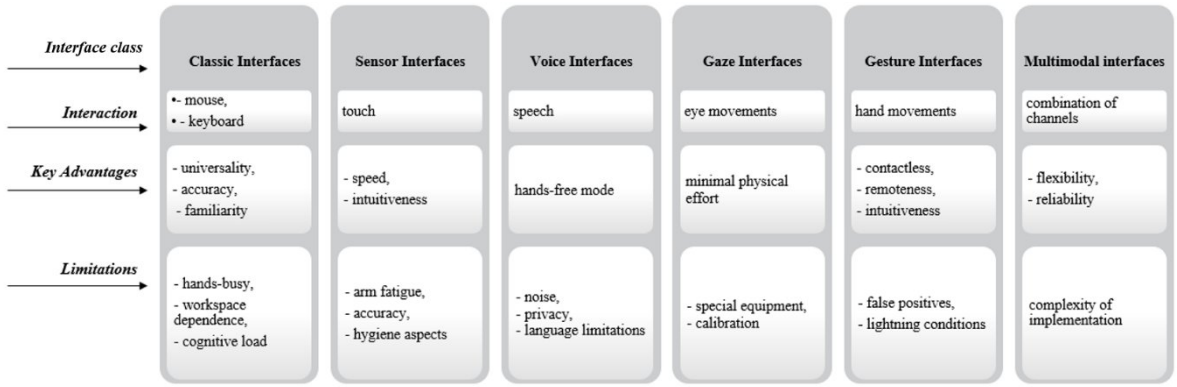


Figure 1. Generalized classification of approaches to LMS management by interaction interface type

Gesture interaction interfaces are considered a promising approach to expanding the capabilities of remote LMS management, as they allow for contactless, intuitive, and spatially independent control of key elements of the educational process [6], [7]. Unlike voice commands, gesture control does not depend on the acoustic conditions of the environment, noise levels, or the user's speech characteristics, which is critically important in group classes and online sessions with a large number of participants.

The use of gesture interfaces opens the possibility of differentiating control actions by their level of criticality: from quick navigation commands to operations requiring confirmation of the user's intent. Such an approach is particularly relevant for LMS, where erroneous execution of certain actions (e.g., publishing materials, changing grades, or access settings) can have significant consequences for the educational process. Accordingly, gesture control can be

integrated into LMS not as a replacement for conventional mouse-and-keyboard interaction, but as an additional control channel within a multimodal interaction model.

Within the broader field of educational ICT, this problem belongs to the development of intelligent, multimodal, and inclusive human–computer interaction mechanisms. While current LMS research largely focuses on platform functionality, learning analytics, content delivery, and personalization, the interaction layer through which users access these functions remains equally important. The development of alternative input modalities contributes to reducing dependence on a single interaction channel, improving accessibility, and adapting educational systems to different physical, environmental, and instructional contexts. In this scientific direction, gesture-based control represents not an isolated computer vision task, but a component of the broader transition from conventional interface models toward adaptive and multimodal educational environments.

Thus, the necessity of researching and developing gesture interfaces for remote management of learning management systems is driven by a combination of organizational, pedagogical, and technological factors:

- the growing role of LMS in the digital educational environment;
- the complication of interaction scenarios during distance learning;
- the need to reduce cognitive and operational load on the instructor;
- requirements for inclusivity and flexibility of interfaces.

This determines the relevance of further analysis of methods and algorithms for gesture control, as well as their adaptation to the specifics of educational scenarios in LMS.

Implementing gesture interfaces in LMS is a strategic step towards creating an intuitive, inclusive, and effective educational environment. Arguments for the benefits of implementing gesture control in LMS are presented in Table 1. This necessity stems from a number of key challenges and opportunities in modern distance education.

Table 1

The relationship between e-learning challenges and the potential of gesture control

№	Area of educational impact	Underlying educational need	Potential contribution of gesture-based interaction
1.	Increasing accessibility and inclusivity	Gesture control is a powerful tool for bridging the digital divide and ensuring equal access to education	- simplification of interaction for users with disabilities; - new opportunities for people with sensory impairments; - naturalness for a generation accustomed to touch technologies.
2.	Improving ergonomics and reducing load	Prolonged work at a computer leads to physical fatigue	- possibility of changing body position during work; - acceleration of routine operations; - focus on content instead of control mechanics.
3.	Increasing engagement and interactivity	Passive information consumption is one of the main problems of distance learning	- transformation of the user from a passive observer into an active participant; - basis for gamification of the learning process; - dynamization of online lectures and webinars.
4.	Adaptability to different contexts and devices	Modern learning takes place everywhere: at home, in the classroom, in the laboratory	- universality of experience on different devices; - key interface for AR/VR environments; - possibility of remote practical demonstrations.
5.	Preparation for the future of educational technologies	Integration of gesture interfaces is an investment in the future of LMS	- foundation for natural " human-machine " communication in education.

In modern scientific literature, the issue of integrating gesture control directly into the web environment of Learning Management Systems (LMS), such as Moodle, Canvas, or Google Classroom, remains insufficiently covered.

To implement the task of LMS gesture management, it is proposed to use knowledge about basic human-machine interactions during LMS usage, existing gesture datasets as a basis for research, as well as an understanding of external factors that can affect the accuracy of gesture command recognition by neural network methods.

Analysis of recent studies and publications. The problem of improving Human-Computer Interaction (HCI) in educational environments is increasingly considered one of the key research directions in the development of intelligent educational technologies, since interaction quality directly influences the accessibility, efficiency, and practical usability of digital learning platforms [8]–[10]. Gesture-based interaction is considered one of the modalities of natural user interfaces. Its usability depends on recognition performance, on the consistency of the gesture vocabulary, on correspondence between gestures and interface actions, on empirical evaluation, and the context in which gesture control is applied [10]. Despite progress in detection algorithms (YOLO, SSD, MediaPipe Hands), most existing developments are either standalone desktop applications or highly specialized solutions for presentation management (PowerPoint), which justifies the relevance of this research - ensuring contactless navigation in LMS using a standard webcam and optimized computer vision algorithms.

To date, there is a wide range of technical solutions for hand tracking and gesture recognition, each having different operating principles, hardware requirements, and suitability for use in distance learning systems based on regular PCs [11]. The choice of the optimal method is determined by the «accuracy–availability–convenience» triad and directly affects the user experience. This section proposes a classification of the main approaches, an analysis of their operating principles, and their suitability for creating a gesture interface in LMS [12]. Existing methods of hand tracking and interpretation of gesture commands can be divided based on the principle of obtaining input data into:

- contact (marker-based) [13];
- contactless (visual) [14].

A more detailed classification of methods is shown in Figure 2.

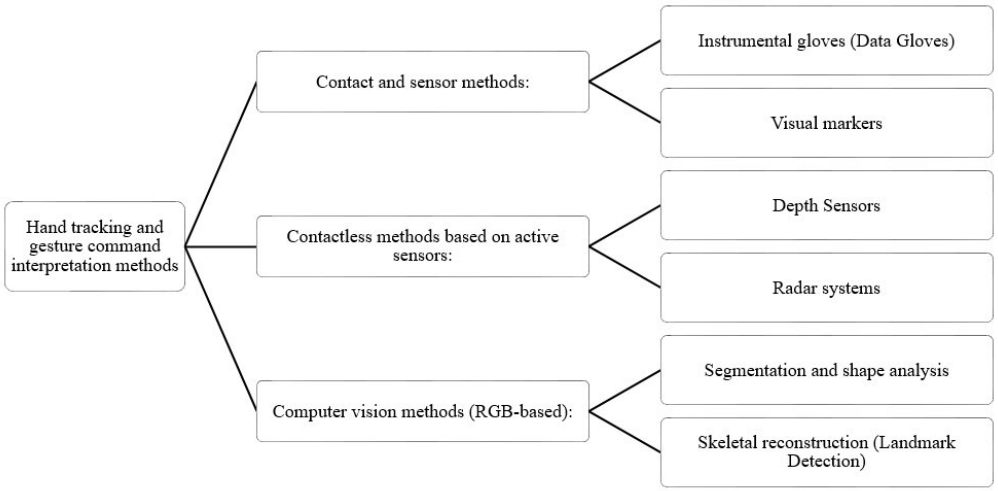


Figure 2. Methods of hand tracking and interpretation of gesture commands by the principle of obtaining input data

For integrating gesture control into a classic PC-oriented LMS, the most viable and practical solution is the landmark-based approach, implemented through libraries such as MediaPipe Hands. This ensures a low entry barrier (webcam), cross-platform compatibility, and sufficient accuracy for navigation and command gestures. Specialized systems like UltraLeap provide exemplary accuracy and can be used in stationary educational laboratories. The experience of XR platforms (Meta, Apple) should be studied to borrow ergonomic patterns of gesture interaction, rather than as a direct technical basis for PC. Other technologies (radar, hybrid) remain promising but are not currently the basis for a mass solution.

The most promising direction for web solutions is the analysis of the video stream from a standard monocular camera. The Landmark Detection method provides a structured representation of the hand suitable for real-time interpretation (Figure 3) [15]. Recent on-device gesture-recognition frameworks demonstrate that hand-landmark representations can be used to construct and deploy customizable gesture vocabularies from RGB camera input through MediaPipe-based processing pipelines [16]. The operation algorithm of such systems (e.g., MediaPipe Hands) includes the following stages:

- detection (Palm Detection) - localization of the region of interest (ROI) in the image and filtering out background noise;
- landmark regression - restoration of the coordinates of reference nodes (usually 21 points) forming the kinematic "skeleton" of the hand;
- feature extraction - calculation of geometric descriptors (angles between phalanges, Euclidean distances between points) and dynamic parameters (speed, acceleration);
- classification - determination of the gesture semantics using heuristic rules or neural network classifiers.

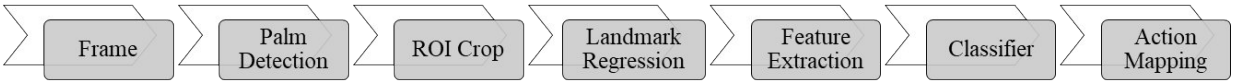


Figure 3. Generalized scheme of the landmark-based gesture recognition algorithm

For the development of the LMS control system prototype, the skeletal reconstruction method (Landmarks) was chosen. Its key advantage is the ability to work in real-time (Real-time inference) on the central processor (CPU) without the need for graphic accelerators or peripheral sensors. This ensures cross-platform compatibility and accessibility of the technology for the end user.

Recent studies in educational HCI additionally emphasize that the design of interaction mechanisms should be considered together with the adaptability of LMS interfaces and cognitive load [17], the use of combined gesture and voice modalities in virtual teaching environments [18], accessibility requirements for educational software supporting spatial interaction [19], and the broader development of multimodal interfaces integrating gesture, speech, gaze, touch, and other communication channels [20].

The research goal and the research objectives. The aim of this study is to design and evaluate a computer vision-based contactless interface for LMS control through operating-system input events, intended for instructors and learners who require or benefit from hands-free or alternative input, and to assess its robustness under varying lighting, distance, viewpoint, and resolution in comparison with conventional mouse-and-keyboard interaction..

The following research tasks were formulated:

- to design a landmark-based gesture control method for LMS interaction supporting single-hand and bimanual control modes;
- to implement a software prototype integrating gesture-based control with the operating system without modifying the LMS web interface;

- to evaluate the robustness of static hand pose recognition under variations in lighting conditions, camera distance, hand orientation, and input resolution;
- to conduct a comparative performance analysis of gesture-based control versus the classical “mouse and keyboard” interface using real user scenarios;
- to validate the operation of the proposed interface through its integration with a real LMS used in an educational institution.

2. RESEARCH METHODS

The solution proposed in the work is based on the idea of context-oriented gesture management of the educational process in Learning Management Systems (LMS), where the method of gesture interpretation depends on the interaction mode and the type of actions performed. Unlike most existing approaches, in which gestures are considered as isolated commands or as direct analogues of traditional controls, in this work, gesture control is considered as a multi-level system supporting various usage scenarios in an educational environment.

The scientific novelty of the approach lies in the introduction and formalization of several gesture interaction modes, which differ in semantics, complexity, and the level of responsibility of the commands performed. In particular, two modes are proposed:

- single-hand control mode, oriented towards executing discrete commands without simulating a manipulator (e.g., "tap", "swipe");
- synchronous bimanual interaction mode, in which mouse simulation is implemented with a clear division of functions between hands (one hand is responsible for aiming/positioning, and the other for confirmation (click/hold)).

From an interaction-design perspective, the gesture vocabulary was organized according to the functional significance and potential consequences of LMS commands. Single-hand gestures were assigned to routine navigation actions, whereas operations capable of causing substantial changes in educational content or settings required bimanual confirmation. This separation is intended to improve the correspondence between gestures and interface actions and to reduce unintended activation of critical commands [10].

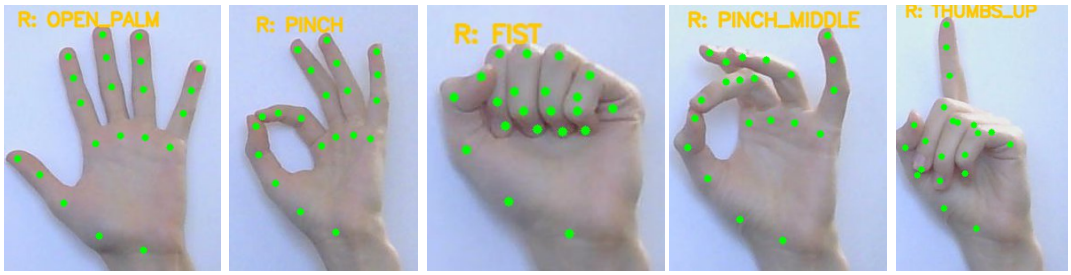


Figure 4. Investigated palm positions: a) open palm (OPEN_PALM); b) pinch with thumb and index finger (PINCH); c) fist (FIST); d) pinch with thumb and middle finger (PINCH_MIDDLE); e) thumb up (THUMBS_UP)

The practical significance of the proposed solution lies in its orientation towards real usage scenarios of LMS during distance and blended learning. The single-hand control mode ensures quick access to frequently used commands (navigation, confirmation, function calls) without being tied to cursor simulation, which is appropriate in hands-free or hands-light interaction situations. Conversely, the bimanual mode allows reproducing the familiar "cursor - click" control model, maintaining compatibility with existing interfaces of the operating

system and LMS, but supplementing it with gesture mechanisms for confirming the user's intent.

The proposed method works as a linear pipeline «frame – key points – interpretation – command». The video stream from the camera enters the input, the frame is converted to the required format, after which palms and hand landmarks are determined. The obtained coordinates are normalized, and movement features are calculated based on them. Next, gesture interpretation is performed. Under the condition of a valid sequence, a command for the operating system is formed, and hints and states are updated on the screen. The context structure of the solution is reflected in the IDEF0 diagram shown in Figure 5.

The internal processing is presented as a linear chain from A1 to A5.

Block A1 «Capture frame» accepts the video frame stream of the scene from the webcam and outputs a frame in RGB format.

Block A2 «Determine key points» receives this frame and identifies the palm and 21 hand landmarks for each found hand along with a confidence score.

Block A3 «Detect tokens and sequences» converts points into movement features and identifies basic actions.

Block A4 «Decision making» applies the sequence vocabulary and intent confirmation rules. At the output, a high-level command is formed taking into account the active OS window.

Block A5 «Submit action» converts this command into specific operating system actions (key combinations, clicks, cursor movements), and updates on-screen hints.

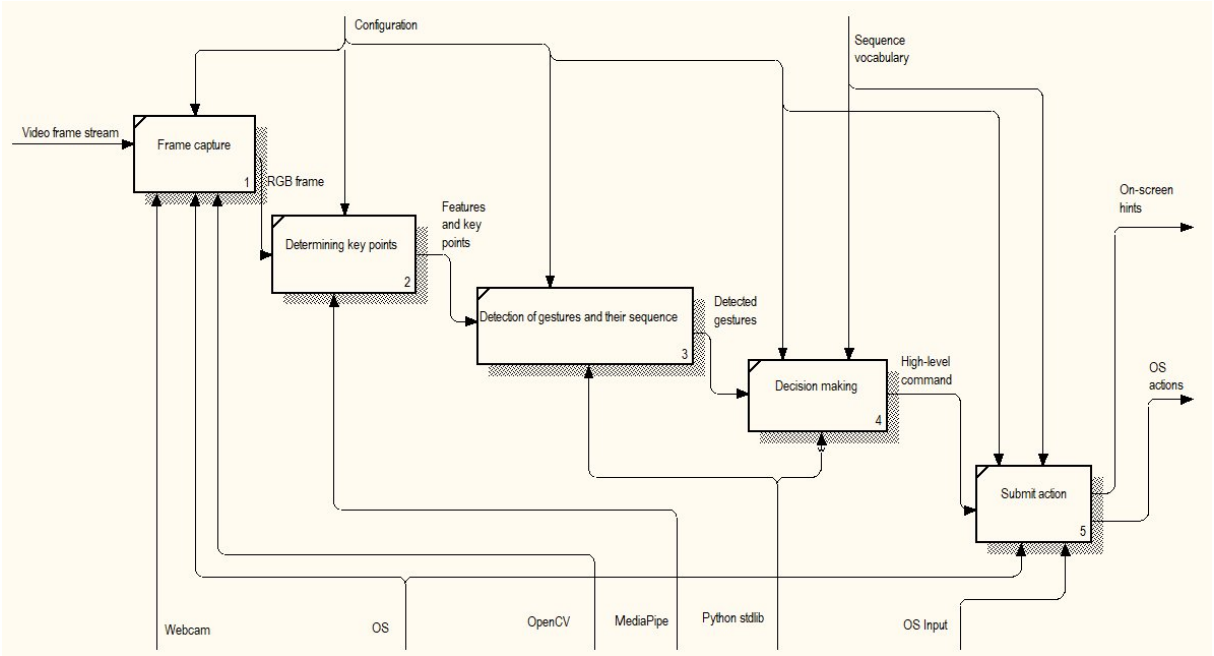


Figure 5. IDEF0 diagram of the LMS gesture control method

Methodologically, the proposed approach combines landmark-based hand tracking as a basic sensory level with its own gesture interpretation algorithm that works on top of the coordinates of reference points and does not rely on ready-made gesture classes. This allows forming flexible control scenarios, introducing new gestures, distinguishing static poses and dynamic movements, as well as adapting threshold values for a specific user. Such architecture is fundamentally important for LMS management, where controllability, predictability of system behavior, and minimization of false actions are critical.

In the project, a gesture is not issued by MediaPipe as a ready-made class but is defined by our algorithm based on 21 hand landmarks. MediaPipe is used as a tracker: it returns the

coordinates of the hand's key points on each frame. Further, it calculates features from landmarks and sets active pose states and forms events at the moment of triggering according to threshold rules.

For static poses, geometric conditions are used. For example, a pinch is defined as the distance between the tip of the thumb and index finger, normalized to the palm scale, with an additional check for finger bending.

The list of realized gestures in the system: PINCH, PINCH_MIDDLE, FIST, THUMBS_UP, OPEN_PALM, SWIPE_LEFT, SWIPE_RIGHT ; their thresholds are additionally adapted to the user in calibration mode (Figure 4). The proposed method is necessary because it allows introducing new gestures and transparent rules on features, whereas using a "ready-made set of gestures" as a complete interface is limited and insufficiently flexible for scenario commands, sequences, and individual accuracy adjustment.

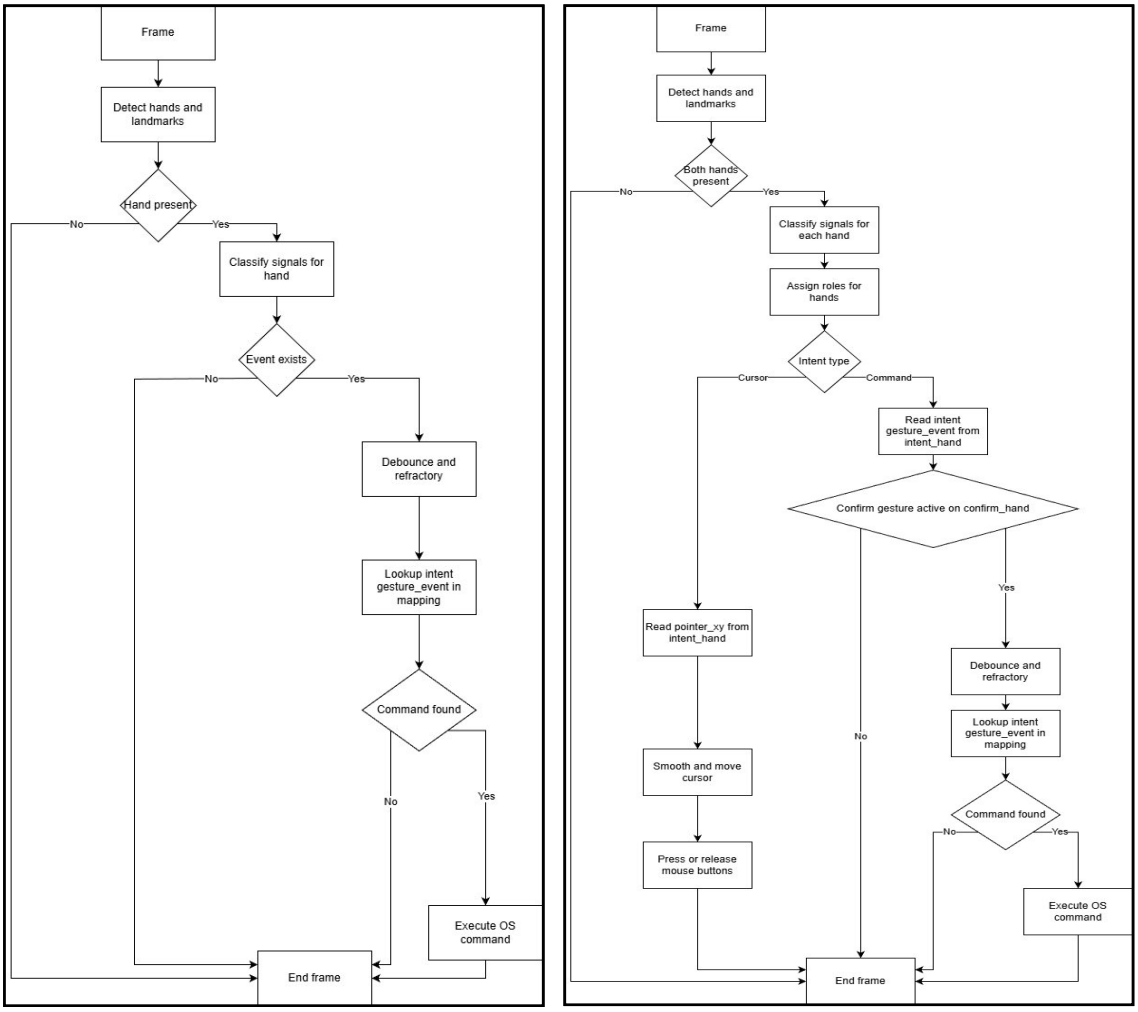


Figure 6. Algorithm of system operation in mode: a) single hand mode; b) bimanual gesture

As mentioned above, the system that includes the implemented method operates in two modes:

- mode 1 – single hand mode (1 hand) mainly for low-risk commands (open course, go to next page, show material, scroll, "play/pause" video, open/close panel);
- mode 2 – bimanual mode (2 hands) mainly for high-risk commands (delete/publish, change grades, finish test, "publish", bulk actions).

In this case, one hand controls/aims or forms an intent, and the second confirms (confirm gesture).

The operation algorithms of the system in the two modes differ, which can be seen in Figure 6.

3. FINDINGS

Within the framework of this study, a complex of experiments was conducted aimed at assessing the performance and efficiency of the contactless computer control system using hand gestures. The goal was to determine the optimal parameters ensuring stable interaction with the operating system, minimizing delays, error frequency, and maintaining an acceptable speed of user task execution for LMS management.

A prerequisite for the main experiment on performance evaluation in different scenarios was a series of basic calibration measurements. They aimed to:

- determine the baseline accuracy of recognizing five selected gestures (PINCH, PINCH_MIDDLE, FIST, THUMBS_UP, OPEN_PALM) under controlled conditions;
- assess the system's robustness to degrading factors, such as increased distance to the camera, worsened lighting conditions, presence of a complex background, and change of hand angle;
- investigate the dependence of recognition quality on the technical characteristics of the video stream - resolution (320x180, 1280x720, 1920x1080) and frame rate (FPS).

The efficiency of the system is evaluated by a set of quantitative and qualitative indicators. The key criteria for the system's suitability for everyday use are defined as speed (average frame rate per second, should not fall below 15 fps), reliability (to minimize actions leading to data loss and false positives) ; and task execution efficiency, which determines the time to execute a typical scenario using gestures. Task execution efficiency was assessed by comparing scenario completion times for gesture-based and mouse-and-keyboard interaction.

Within the scenario-based evaluation, an interaction error was defined as a false activation, a missed command, or an incorrect interpretation of the intended action. A non-critical error was a recoverable failure, such as unintended scrolling, inaccurate cursor positioning, a missed click, or opening an incorrect interface element without changing educational data. A critical error was an unintended action that caused or could directly cause a persistent change in the LMS state, including deletion or publication of educational content, modification of grades or access settings, premature test submission, or loss of an object during a compound operation.

Three scenarios were selected to represent the main classes of interaction required for LMS-related operation through an operating-system-level interface. Scenario 1 evaluates discrete pointing, selection, and context-command execution; Scenario 2 covers sequential navigation, browser interaction, and text entry; and Scenario 3 evaluates continuous compound manipulation requiring a sustained gesture state. Thus, the scenarios reflect increasing interaction complexity and jointly cover positioning, command activation, navigation, text input, and object-holding operations rather than three isolated application tasks. They are not intended to reproduce the full range of LMS functionality, but to evaluate representative interaction primitives on which LMS operations depend.

The experiment needs to be conducted for several main scenarios.

Scenario №1 «File operations»: Creating, renaming, and opening a folder (test for cursor positioning accuracy and context command execution).

Scenario №2 «Web surfing»: Opening a browser, entering text into the address bar, navigating to a web page (test for interaction with interface elements and virtual keyboard).

Scenario №3 «Clipboard»: Drag-and-Drop operations or copying files between directories (test for complex compound gestures, such as holding an object).

In this study, robustness refers to the ability of the interface to maintain recognition accuracy when recording conditions or technical parameters change. Stability refers to the continuity and predictability of command execution during an interaction scenario.

3.1. Evaluation of robustness of static hand pose recognition under different shooting conditions

This section presents the results of a series of controlled experiments aimed at assessing the accuracy and stability of the gesture-based LMS control interface to key variable environmental factors and technical parameters (lighting, distance to camera, changes in hand orientation, and input image resolution). The research was conducted on a set of five static poses: PINCH, PINCH_MIDDLE, FIST, THUMBS_UP, and OPEN_PALM.

Table 2

Comparative accuracy of static hand pose recognition under different operating conditions

Conditions Gesture	Baseline shooting conditions 1280x720, 0.5 м	Distance ~2 м	Poor lighting (130 Lux)	Changes in hand orientation (angle change)	Low resolution 320x180 (22.2 FPS)	High resolution 1920x1080 (14.4 FPS)
PINCH	0.733	0.433	0.467	0.300	0.667	0.867
PINCH_MIDDLE	0.867	0.133	0.633	0.433	0.767	0.867
FIST	0.967	0.800	0.967	0.167	0.900	1.000
THUMBS_UP	0.867	0.833	0.967	0.733	0.967	0.933
OPEN_PALM	0.933	0.967	0.967	0.933	0.967	0.967
Average accuracy	0.873	0.633	0.800	0.513	0.854	0.927

Each static hand pose was repeated 30 times under each recording condition. Every trial was classified as a correct recognition, misclassification, or missed detection. Recognition accuracy was calculated as the proportion of correct recognitions. Because each trial had a binary outcome, 95% confidence intervals were calculated using the Wilson score method.

For the aggregated results obtained from 150 trials per condition, the mean accuracy and corresponding 95% confidence intervals were: baseline conditions – 0.873 [0.81–0.92], distance of approximately 2 m – 0.633 [0.55–0.71], poor lighting – 0.800 [0.73–0.86], changed hand orientation – 0.513 [0.43–0.59], low resolution – 0.854 [0.79–0.90], and high resolution – 0.927 [0.87–0.96].

To comprehensively assess the stability (robustness) of the system, the results of all experiment series were summarized in a comparative table. Indicators obtained under ideal conditions were chosen as the baseline: distance 0.5 m, normal lighting (approx. 130 Lux), frontal angle, and resolution 1280x720. Other columns reflect accuracy when changing one of the parameters while keeping others unchanged.

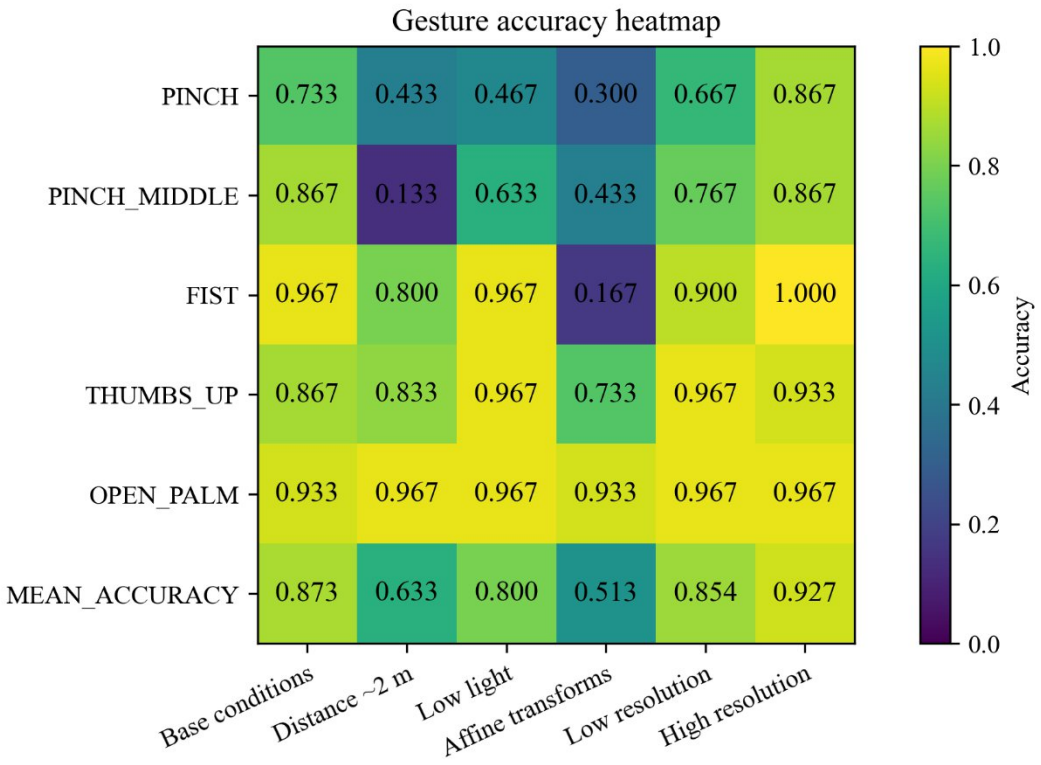


Figure 7. Heatmap of gesture recognition accuracy under different operating conditions

Based on the data in Table 2, the following conclusions can be drawn regarding influencing factors:

- changes in hand orientation (gesture presentation angle) turned out to be the most critical factor. The average accuracy across all gestures drops to 0.513. Gestures whose shape depends on the mutual arrangement of fingers in 3D space (FIST, PINCH) suffer the most. When rotating the hand, the projection loses informativeness due to self-occlusion, making this factor the most critical for UX;
- the distance factor shows that increasing the distance to 2 meters critically affects only fine motor skills (PINCH_MIDDLE dropped to 0.133). Global gestures (OPEN_PALM, THUMBS_UP) demonstrate high stability to scale changes;
- the system demonstrates high robustness in poor lighting conditions (accuracy 0.800), which allows it to be used in the evening or in darkened auditoriums during presentations;
- when assessing the impact of resolution, a direct correlation between resolution and accuracy is observed: the 320x180 (Low Res) mode ensures the highest FPS (22.2 fps), but accuracy for small gestures decreases (average accuracy 0.854). The 1920x1080 (High Res) mode showed the best results (average accuracy 0.927). For the FIST gesture, 100% recognition was achieved, and the PINCH gestures, which are difficult for the system, increased to the level of 0.867. This confirms that a larger number of pixels in the hand image allows the algorithm to more accurately build a skeletal model, compensating for the FPS drop to 14.4 fps.

The OPEN_PALM gesture showed almost absolute stability to any changes in shooting conditions (accuracy >0.93 in all tests). This makes it an ideal candidate for critically important commands (e.g., "Stop", "Reset", or "Activate menu").

3.2. Comparative analysis of gesture control efficiency with the classical «mouse + keyboard»

The goal of the final stage of the research is to assess the integral efficiency of the developed system under conditions of performing real user tasks. To ensure the validity of the results, three test scenarios were developed covering basic interaction patterns with the Windows operating system: file system navigation, working with a web browser, and clipboard operations.

The traditional «keyboard + mouse» interface was chosen as the baseline for comparison, as it is the «gold standard» of ergonomics for desktop OS.

For each run, three key metrics were recorded: total execution time (T_{exec}), number of control errors (N_{err}), and the presence of critical failures. The results of the comparative analysis are shown in Table 3.

Table 3

Comparative characteristic of interaction scenario execution

№ п/п	Mode / Resolutio	Scenario	Interface Type	Time (T_{exec}), s	Errors (N_{err})	Critical Error	Notes
1	1280x720 (15.6 FPS)	1	Gestures	48.692	1	No	
2	-	1	Mouse+Keyb.	15.127	0	No	Baseline
3	1280x720 (15.6 FPS)	2	Gestures	65.129	0	No	
4	-	2	Mouse+Keyb.	16.206	0	No	Baseline
5	1280x720 (15.6 FPS)	3	Gestures	80.317	3	No	Compound gesture
6	-	3	Mouse+Keyb.	13.401	0	No	Baseline
7	320x180 (22.2 FPS)	2	Gestures	80.433	5	No	Low accuracy
8	1920x1080 (14.3 FPS)	2	Gestures	59.379	2	No	
9	1920x1080 (14.3 FPS)	3	Gestures	58.862	6	Yes	Hypersensitivity

Analysis of the data in Table 3 demonstrates a significant discrepancy in the speed of the two interfaces. Calculation of the execution-time ratio ($K_{time} = \frac{T_{gesture}}{T_{mouse}}$) shows the following dynamics:

- Scenario 1: $K_{time} \approx 3{,}2$;
- Scenario 2: $K_{time} \approx 4{,}0$;
- Scenario 3: $K_{time} \approx 6{,}0$.

The increase in the coefficient in scenario №3 is explained by the complexity of performing Drag-and-Drop operations using contactless gestures. Holding a virtual object requires constant muscle tension from the user (PINCH or FIST gesture) and simultaneous hand movement, which is a kinematically more complex task compared to physically moving a mouse.

Analysis of the influence of resolution on scenario execution demonstrates that using low resolution (320x180) in Scenario 2 led to an increase in the number of errors to 5 and an increase in execution time to 80 s, making such a mode unsuitable for precise work.

At the same time, switching to Full HD (1920x1080) in Scenario 3, although it reduced the execution time to 58 s (almost 27% faster than HD), led to the occurrence of a critical error and 6 failures. This phenomenon is explained by the increased sensitivity of the detector at high resolution: hand micro-tremors or accidental lighting changes can be interpreted by the system as a break in the holding gesture ("dropping the file"), which is unacceptable in the context of file operations.

Thus, the 1280x720 mode is the optimal balance between object holding stability and positioning accuracy.

3.3 Demonstration of integration with OS for LMS management

Despite the fact that gesture control is on average slower than classical control, the system confirmed its viability. It successfully performs the functions assigned to it, and the detected errors (except for the hypersensitivity mode of 1080p) are not critical and are subject to correction by the user. The proposed solution should be considered not as a replacement for the mouse for intensive work, but as an auxiliary interaction layer (for example, for managing presentations, media content, or basic navigation when access to the keyboard is limited).

Table 4

Demonstration of activating the *dl.nure.ua* system and opening the browser using the gesture interface via operating system events

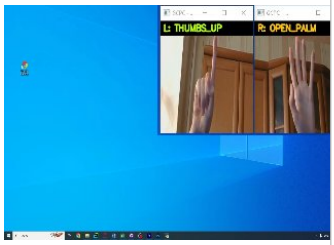
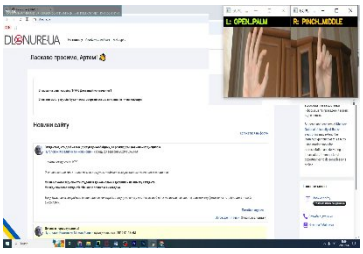
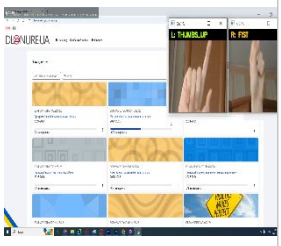
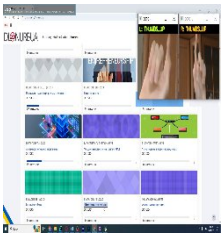
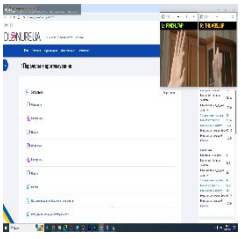
Gesture Command	Gesture THUMBS_UP (thumb up) with the left hand	Gesture THUMBS_UP (thumb up) with the right hand and moving this gesture in space	Double PINCH (double pinch)
Command Interpretation	activation of control mode	moving the cursor to the browser shortcut	opening the browser
Demonstration			

Table 5

Demonstration of working with the LMS dl.nure.ua interface

Gesture Command	Pointing via THUMBS_UP with the right hand and clicking PINCH with the left hand to open	FIST with the right hand and movement to scroll the page	Pointing via THUMBS_UP with the right hand	Double PINCH
Command Interpretation	hovering over the "My Courses" menu item and clicking	vertical scrolling of the course list	selecting a specific course	entering the course
Demonstration				

To confirm the practical suitability of the proposed approach, a demonstration of the integration of gesture control with the real distance learning platform of NURE - dl.nure.ua,

used in the educational process of the Kharkiv National University of Radio Electronics, was performed.

The present study was limited to the technical feasibility, robustness, and scenario-based performance of the proposed interface. A formal usability assessment involving a representative participant sample and standardized subjective measures was outside the scope of this stage. Therefore, the reported execution times and error counts characterize interaction performance but should not be interpreted as a complete assessment of usability, user satisfaction, cognitive workload, or long-term comfort.

The presented demonstration (Table 4 and 5) confirms the possibility of using gesture control as an alternative interface for interaction with educational web systems without the need for their modification. Such an approach is promising for scenarios of distance learning, inclusive education, and situations where the use of traditional input devices is limited or undesirable.

Further research will include a multi-user usability assessment using standardized questionnaires and measures of learnability, workload, and user comfort. Subsequent development may focus on multimodal interaction, adaptive personalization, and optimization for mobile devices.

3.4 Authors' contributions

The research was conducted by a multidisciplinary team of experts in artificial intelligence, computer vision, and educational technologies.

Olesia Barkovska led the conceptual design of the study, defined the research objectives, and coordinated the overall scientific structure of the work. She developed the methodology of gesture-based LMS control, supervised the experimental research, and was responsible for drafting and final revision of the manuscript.

Artem Bakulin contributed to the implementation of the gesture-based LMS control interface and participated in the development of interaction scenarios for LMS control. He was involved in software testing and assisted in conducting experimental evaluations.

Volodymyr Mytiai participated in the experimental setup and data collection, including the execution of gesture recognition tests under varying conditions. He also contributed to the analysis and validation of experimental results.

Oleksandr Yankovskyi provided scientific consultation on computer vision methods and intelligent interfaces. He contributed to the critical review of the methodology and supported the interpretation of experimental findings.

Ihor Novoseltsev designed and executed the experimental evaluation, including robustness testing and quantitative analysis. He prepared result visualizations and contributed to manuscript writing and technical revisions.

Anton Havrashenko contributed to the integration of the gesture-based control system with the operating system environment. He also assisted in software implementation and technical validation of the proposed solution.

4. CONCLUSIONS AND PROSPECTS FOR FURTHER RESEARCH

The conducted study demonstrated the technical feasibility of the proposed gesture-based control system for LMS interaction and identified the conditions under which its recognition and interaction performance remain acceptable. Quantitative evaluation showed that the average recognition accuracy of static hand poses reached 0.873 under baseline conditions, remaining above 0.80 in low-light environments, while decreasing to 0.513 under changes in hand orientation, which were identified as the most critical destabilizing factor.

Experimental results confirmed a direct correlation between input image resolution and recognition accuracy. The Full HD mode (1920×1080) achieved the highest average accuracy (0.927), including 100% recognition of the FIST gesture, whereas the low-resolution mode (320×180) resulted in a notable degradation of precision for fine-motor gestures such as PINCH and PINCH_MIDDLE. At the same time, the HD resolution (1280×720) was shown to provide the best balance between stability and performance for real interaction scenarios.

Gesture-based interaction required 3.22, 4.02, and 5.99 times more execution time than mouse-and-keyboard interaction in Scenarios 1–3, respectively. The largest performance gap was observed in drag-and-drop operations, due to the increased motor and cognitive load associated with sustained gestures. Nevertheless, no destructive errors were observed in the HD configuration, confirming the overall viability of the approach.

The results indicate that gesture-based control should be considered not as a replacement for traditional input devices, but as an auxiliary interaction layer suitable for presentation control, navigation, and hands-free or accessibility-oriented scenarios. The practical applicability of the proposed method was confirmed through successful execution of representative LMS interaction tasks on the real educational platform dl.nure.ua through operating-system-level input events.

Based on the limitations identified during the study and with the aim of further expanding the functionality and practical value of the system, promising directions are: implementation of a multimodal interface combining gestures with voice commands and gaze tracking to create a more natural and powerful way of interaction ; development of adaptive machine learning algorithms capable of personalizing recognition parameters to individual anatomical features and gesture execution style of a specific user, which will increase accuracy and comfort ; as well as optimization of the system architecture for efficient operation on mobile devices (tablets, smartphones), which will open up possibilities for its use in educational and domestic scenarios outside the stationary workplace.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

All scientific content was produced, verified, and approved by the authors, who take full responsibility for the originality, accuracy, integrity, and final content of the manuscript. AI-assisted tools were used solely for translation and linguistic editing and were not used to generate research data, experimental results, figures, or scientific conclusions.

CONFLICT OF INTEREST

The authors declare that there is no financial or non-financial conflict of interest that could influence the results of the study, their interpretation, or the publication of this article.

REFERENCES (TRANSLATED AND TRANSLITERATED)

[1] N.S. Alotaibi,"The impact of AI and LMS integration on the future of higher education: Opportunities, challenges, and strategies for transformation," *Sustainability*, 16, no. 23: 10357. https://doi.org/10.3390/su162310357. (in English)

[2] O. Barkovska, H. Ivashchenko, D. Rosinskiy, D. Zakharov, “Educational Training Simulator for Monitoring Reading Technique and Speed Based on Speech-To-Text (STT) Methods,” *Inf. Technologies and Learn. Tools*, vol. 103, no. 5, pp. 21–38, Oct. 2024, doi: 10.33407/itlt.v103i5.5647. (in English)

[3] O. Barkovska, Y. Liapin, T. Muzyka, I. Ryndyk, P. Botnar, “Gaze Direction Monitoring Model in Computer System for Academic Performance Assessment,” *Inf. Technologies and Learn. Tools*, vol. 99, no. 1, pp. 63–75, Feb. 2024, doi: 10.33407/itlt.v99i1.5503. (in English)
[4] N.D. Handayani, S. Robby Satria. "Inclusive Education System: A Path to Equal Digital Access and Quality Learning for University Students," *E3S Web of Conferences*, vol. 657. EDP Sciences, 2025, https://doi.org/10.1051/e3sconf/202565705007. (in English)
[5] G.A. Toto, *et al.* “What is the Role of Technologies for Inclusive Education? A Systematic Review”, *International Conference on Higher Education Learning Methodologies and Technologies Online*. Springer, Cham, 2024, vol 2076, Springer, Cham, https://doi.org/10.1007/978-3-031-67351-1_36. (in English)
[6] P. Markad, R. Chile, V. Thakare and S. Ghodke, "A Review on Gesture-Sensitive Interactive Learning for Preschoolers," *2024 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS)*, Gobichettipalayam, India, 2024, pp. 502-505, doi: 10.1109/ICUIS64676.2024.10866628. (in English)
[7] G. Cosentino and M. Giannakos, "Multisensory Interaction and Analytics to Enhance Smart Learning Environments: A Systematic Literature Review," in *IEEE Transactions on Learning Technologies*, vol. 16, no. 3, pp. 414-430, 1 June 2023, doi: 10.1109/TLT.2023.3243210. (in English)
[8] Y.Xie, L.Yang, M. Zhang, S.Chen & J. Li "A Review of Multimodal Interaction in Remote Education: Technologies, Applications, and Challenges", *Applied Sciences*, *15*(7), 3937, https://doi.org/10.3390/app15073937. (in English)
[9] H. Zhou,, D.Wang, Y.Yu, Z.Zhang, "Research progress of human–computer interaction technology based on gesture recognition". *Electronics*, *12*(13), 2805, https://doi.org/10.3390/electronics12132805. (in English)
[10] K. Stec and L. B. Larsen, “Gestural Interfaces in Human–Computer Interaction,” in *The Cambridge Handbook of Gesture Studies*, A. Cienki, Ed. Cambridge, UK: Cambridge University Press, 2024, pp. 661–682, doi: 10.1017/9781108638869.027 (in English)
[11] F. Sufyan, S. Sagar, Z. Ashraf, S. Nayel, M. S. Chishti, A. Banerjee, "A Novel and Lightweight Real-Time Continuous Motion Gesture Recognition Algorithm for Smartphones," in *IEEE Access*, vol. 11, pp. 42725-42737, 2023, https://doi.org/10.1109/ACCESS.2023.3255402. (in English)
[12] D. Noh, H. Yoon and D. Lee, "A Decade of Progress in Human Motion Recognition: A Comprehensive Survey From 2010 to 2020," in *IEEE Access*, vol. 12, pp. 5684-5707, 2024, https://doi.org/10.1109/ACCESS.2024.3350338 . (in English)
[13] R.Tchantchane, H.Zhou, S.Zhang, G.Alici, "A review of hand gesture recognition systems based on noninvasive wearable sensors," *Advanced intelligent systems* 5.10 (2023): 2300207, https://doi.org/10.1002/aisy.202300207. (in English)
[14] F.Al Farid, N. Hashim, J. Abdullah, M. R. Bhuiyan, W. N. Shahida Mohd Isa, J. Uddin, & M. N. Husen, "A structured and methodological review on vision-based hand gesture recognition system", *Journal of Imaging*, *8*(6), 153, https://doi.org/10.3390/jimaging8060153. (in English)
[15] J. S. Han, C. I. Lee, Y. H. Youn, & S. J. Kim, "A study on real-time hand gesture recognition technology by machine learning-based mediapipe", *Journal of System and Management Sciences*, *12*(2), 462-476, https://doi.org/10.33168/JSMS.2022.0225. (in English)
[16] E. Uboweja, D. Tian, Q. Wang, Y.-C. Kuo, J. Zou, L. Wang, G. Sung, and M. Grundmann, “On-device Real-time Custom Hand Gesture Recognition,” in *Proc. 2023 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, Paris, France, 2023, pp. 4275–4279, doi: 10.1109/ICCVW60793.2023.00461. (in English)
[17] M. Suryani, D. I. Sensuse, H. B. Santoso, R. F. Aji, S. Hadi, R. R. Suryono, and Kautsarina, “An initial user model design for adaptive interface development in learning management system based on cognitive load,” *Cognition, Technology & Work*, vol. 26, pp. 653–672, 2024. https://doi.org/10.1007/s10111-024-00772-8. (in English)
[18] K. Fang and J. Wang, “Interactive design with gesture and voice recognition in virtual teaching environments,” *IEEE Access*, vol. 12, pp. 4213–4224, 2024. https://doi.org/10.1109/ACCESS.2023.3348846. (in English)
[19] A. Kešelj Dilberović, K. Žubrinić, M. Miličević, and M. Kristić, “Evaluation of accessibility parameters in education software that supports three-dimensional interactions,” *Applied Sciences*, vol. 15, no. 10, Art. no. 5258, 2025. https://doi.org/10.3390/app15105258. (in English)
[20] E. Dritsas, M. Trigka, C. Troussas, and P. Mylonas, “Multimodal interaction, interfaces, and communication: A survey,” *Multimodal Technologies and Interaction*, vol. 9, no. 1, Art. no. 6, 2025. https://doi.org/10.3390/mti9010006. (in English)

Text of the article was accepted by Editorial Team 04.02.2026

БЕЗКОНТАКТНИЙ ІНТЕРФЕЙС КЕРУВАННЯ НА ОСНОВІ КОМП'ЮТЕРНОГО ЗОРУ ДЛЯ ОСВІТНІХ СЕРЕДОВИЩ

Олеся Барковська
Харківський національний університет радіоелектроніки, м.Харків, Україна
ORCID ID 0000-0001-7496-4353
olesia.barkovska@nure.ua

Артем Бакулін
Харківський національний університет радіоелектроніки, м.Харків, Україна
ORCID ID 0009-0003-8885-7673
artem.bakulin@nure.ua

Володимир Митяй
Харківський національний університет радіоелектроніки, м.Харків, Україна
ORCID ID 0009-0003-1239-6727
volodymyr.mytiai@nure.ua

Олександр Янковський
Харківський національний університет радіоелектроніки, м.Харків, Україна
ORCID ID 0000-0002-1268-0029
oleksandr.yankovskyi@nure.ua

Ігор Новосельцев
незалежний дослідник, Карлсбад, США
ORCID ID 0009-0004-7353-7498
i.v.novoseltsev@gmail.com

Віталій Сердечний
Харківський національний університет радіоелектроніки, м.Харків, Україна
ORCID ID 0009-0007-8828-5803
vitalii.serdechnyi@nure.ua

Анотація. Актуальність дослідження зумовлена зростаючою роллю систем дистанційного навчання в освітньому процесі та необхідністю розширення способів взаємодії з ними, зокрема в умовах обмеженого доступу до традиційних пристроїв введення або в контексті інклюзивної освіти. Метою роботи є розробка та експериментальне дослідження жестового інтерфейсу керування LMS, який інтегрується в операційну систему без необхідності модифікації самої платформи дистанційного навчання. Для досягнення поставленої мети були сформульовані такі задачі: розробити модель методу жестового управління LMS; реалізувати режим моножестикулювання та бімануальний режим взаємодії для різних сценаріїв керування; провести серію експериментів з оцінювання точності та стійкості розпізнавання жестів за різних умов зйомки; виконати порівняльний аналіз ефективності жестового керування та класичного інтерфейсу «миша + клавіатура». Методи дослідження охоплюють комп'ютерний зір на основі лендмарк-детекції кисті, алгоритмічну інтерпретацію жестів за геометричними та динамічними ознаками, експериментальні вимірювання точності, швидкості виконання сценаріїв та кількості помилок. Результати дослідження показали, що система забезпечує стабільне розпізнавання великих статичних поз кисті за широкого діапазону умов освітлення та відстані до камери. Найбільш критичним фактором впливу на точність виявились афінні перетворення (зміна ракурсу). Продемонстровано працездатність жестового керування як альтернативного шару взаємодії з реальною LMS dl.nure.ua через події операційної системи. Практична значущість полягає в можливості використання запропонованого підходу для керування освітніми вебплатформами в умовах дистанційного навчання, під час презентацій, а також у сценаріях інклюзивної освіти, для віддаленого керування різними типами контенту.

Ключові слова: система управління навчанням; взаємодія на основі жестів; дистанційне навчання; інклюзивні інтерфейси.

(cc) BY-NC-SA

This work is licensed under Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.