JURNAL UPI

Jurnal Pendidikan Multimedia (EDSENCE)

Jurnal Pendidikan Multimedia EDSENCE Education, Science, and Creative Technology

Journal homepage: https://ejournal.upi.edu/index.php/Edsence

Real-Time Hand Gesture Recognition for Controlling Web Video of Traditional Dance Utilizing MediaPipe

Mahardika Pratama[1], Mochamad Fathur Rabbani[2], Reza Maulana Aziiz[3], Trisna Gelar[4], Muhammad Rizqi Sholahuddin[5], Iwan Awaludin[6], Aprianti Nanda Sari[7]

[1-7] Politeknik Negeri Bandung, Bandung, Indonesia

Correspondence: E-mail: trisna.gelar@polban.ac.id

ABSTRACT

The interaction between humans and digital devices has significantly evolved with a notable progression in Human-Computer Interaction (HCI) being the increasing adoption of contactless device control through hand gesture recognition. This research details the development and evaluation of a real-time hand gesture recognition system utilizing the MediaPipe framework for controlling web video players, with a specific focus on its application in managing the playback of traditional dance videos. The system enables users to manipulate video playback functions via hand gestures. The research methodology comprised key stages: a literature review, system workflow design, practical implementation, and performance evaluation. The MediaPipe framework facilitated real-time detection of hand landmark coordinates, processed, and mapped to specific video player control commands (e.g., play, pause, volume adjustment) using a rule-based approach. The findings indicate that the system performs very well in bright lighting conditions and at close user distances (0.5 meters) from the camera, accurately recognizing gestures with high responsiveness. However, recognition accuracy significantly decreases at further distances (>3 meters) and in dark lighting conditions, highlighting the importance of optimal camera placement and sufficient illumination for effective system operation. The evaluation results indicate effective control of web video players, demonstrating optimal performance in environments with bright illumination and close user proximity.

ARTICLE INFO

Article History:
Submitted/Received 18 Apr 2015
First Revised 03 Jun 2025
Accepted 12 Jun 2025
First Available online 27 Jun 2025
Publication Date 27 Jun 2025

Keyword:
hand gesture recognition, mediapipe framework, web video player control, contactless interaction, traditional dance.

1. INTRODUCTION

The interaction between humans and digital devices has undergone a significant evolution with the rapid advancements in computer technology. A notable progression in Human-Computer Interaction (HCI) is the increasing adoption of contactless device control facilitated by hand gesture recognition technologies. This technology enables users to manipulate diverse device operations solely through hand gestures, thereby substituting conventional input peripherals such as the mouse and keyboard. Hand gesture recognition offers a more natural, intuitive, and effective modality for interacting with computer systems (Roy & Akif, 2022).

Hand gesture recognition has emerged as a significant advancement with a diverse range of applications, including presentation control, interactive gaming, and smart environments. In the domain of presentation control, this technology empowers users to manipulate applications such as PowerPoint through hand gestures, facilitating actions like slide transitions and real-time activation of full-screen (Agustiani et al., 2024; Khairianto & Firdaus, 2024). Within interactive gaming, particularly for children with special needs, hand gesture recognition has been shown to enhance hand-eye coordination and motor skills, providing an intuitive and engaging experience (Dewangga & Subianto, Mochamad Swastika, 2024). Concurrently, in smart home automation systems, hand gestures are increasingly employed to regulate functionalities such as lighting and temperature, thereby improving efficiency and user convenience (Saputra & Hariyanto, 2024).

The MediaPipe framework stands as a widely recognized and robust solution for real-time hand gesture recognition. Leveraging deep learning algorithms, MediaPipe demonstrates the capability to accurately detect and track a wide range of hand gestures (Kriznar et al., 2021). This versatility has led to its application in diverse contexts, including the development of intuitive and adaptive camera-based user interfaces.

Despite the advancements in hand gesture recognition and its applications across various domains, its utilization in the specific context of managing web video players, particularly within dynamic presentation settings such as artistic performances or conference seminars, remains relatively underexplored. Given that video content, ranging from introductory performances of traditional dances to multimedia presentations, constitutes a fundamental component of such events, the development of more flexible and efficient control modalities is increasingly pertinent. As highlighted in related work on immersive video mapping for conference seminars, enhancing audience engagement through interactive technologies is a key objective (Awaludin et al., 2024). This research specifically aims to bridge this gap by developing a real-time MediaPipe-based hand gesture recognition system tailored for controlling web video players in such dynamic presentation environments, particularly for traditional dance videos.

The integration of MediaPipe-based hand gesture recognition offers a promising avenue for presenters and event organizers to manipulate essential video player functionalities—including playback initiation and cessation, navigation between different video segments (e.g., transitioning from a traditional dance performance to a presentation), and audio control—without the constraints of direct physical interaction with input devices, thereby potentially creating a more seamless and engaging experience for both presenters and the audience.

DOI: https://doi.org/10.17509/edsence.v7i1.82344
p-ISSN:2685-2489 | e-ISSN:2685-2535

This study endeavors to develop a real-time hand gesture recognition system, leveraging the MediaPipe framework, for the control of web video players, with a specific application focus on enhancing the presentation of artistic content, such as traditional dance performances during events like the ISSAT conference. This research seeks to contribute by: (1) offering a potential alternative to conventional input peripherals in web video player interaction within dynamic presentation scenarios, (2) providing a more intuitive and natural control mechanism for presenters through hand gestures, thereby facilitating a smoother flow of content, and (3) addressing the inherent limitations of physical input devices in such live event settings. Furthermore, this study investigates the performance and effectiveness of the developed system across various operational conditions relevant to presentation environments. Ultimately, the findings of this research are anticipated to lay a foundation for the advancement of more sophisticated and inclusive Human-Computer Interaction technologies, particularly in the realm of live event management and artistic presentations, potentially complementing and expanding upon interactive technologies like immersive video mapping.

2. METHODS

This study employs an experimental approach to design and assess a real-time hand gesture recognition system that leverages the MediaPipe framework for controlling web video players. The research methodology includes a systematic sequence of stages: a thorough literature review, system workflow design, practical implementation, and performance evaluation. MediaPipe is utilized for the real-time detection of hand landmark coordinates, detailing the precise positions of the fingers and palm. The acquired landmark coordinate data are processed to link specific gestures with application control functions, including play initiation, pause, and volume adjustment. The relationship between gestures and control commands is established using a rule-based approach, where specific configurations of hand landmark coordinates serve as representations of executable system commands.

2.1 Literature Review

MediaPipe is a framework utilized for processing sensory data, including audio and video, aimed at producing outputs like object detection and facial landmark recognition. MediaPipe enables developers to create pipelines that consist of various modular components, such as artificial intelligence models, media processing algorithms, and data transformations. This framework is tailored for researchers, educators, and developers aiming to quickly prototype or create production-ready applications driven by artificial intelligence. MediaPipe offers a cohesive interface and reusable components, enabling efficient application development across various platforms, including desktop and mobile devices, while maintaining consistent processing (Lugaresi et al., 2019).

Hand Gesture Recognition is a system that identifies and interprets hand movements, facilitating non-verbal communication between humans and computers. These systems provide a contemporary and innovative method for non-verbal communication. These systems are applicable in various domains, including human-computer interaction and sign language interpretation. Hand gesture recognition systems may utilize a shape-based

DOI: https://doi.org/10.17509/edsence.v7i1.82344
p-ISSN:2685-2489 | e-ISSN:2685-2535

approach to identify unique patterns in hand movements, encompassing finger positions, hand orientation (horizontal or vertical), and additional notable features. Hand gestures are classified according to distinct hand shapes, orientations, or finger configurations (Panwar & Singh Mehra, 2011).

Hand Gesture Recognition systems present an innovative input method for human-computer interaction, facilitating a more intuitive and convenient form of engagement. These systems employ images obtained from a camera and convert them into recognizable data through processes including segmentation, orientation detection, feature extraction, and classification. The proposed algorithms can operate independently of user characteristics and may not necessitate sample data training. These systems facilitate more intuitive human-computer interaction without requiring additional hardware, such as specialized gloves, thereby providing a more natural user experience in diverse applications, including sign language translation and software control (Haria et al., 2017).

Hand landmarks denote the identification of 21 essential keypoints on the hand, established through the 3D coordinates of the hand joints identified in an image. The process utilizes regression, with the MediaPipe hand landmark model producing direct coordinate predictions for each point within the detected hand region, thereby offering an accurate representation of the hand's structure in the identified image, as shown in Figure 1. Each joint in the hand landmark corresponds to x, y, and z coordinates. The x and y values are normalized within the range of [0.0, 1.0] according to the image's width and height, whereas z indicates the depth of the landmark. The depth of these landmarks is established with respect to the wrist as the origin, where decreasing z values signify proximity to the camera (Indriani et al., 2021).

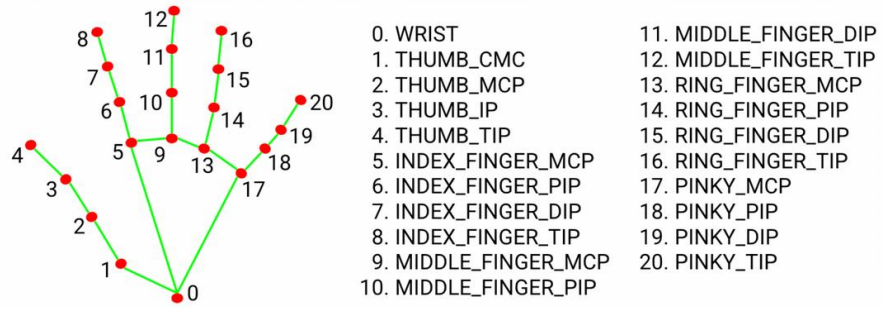

Figure 1. Hand Landmark on Mediapipe (Google Developers, 2025)

2.1 Workflow Design

The workflow design phase begins with the acquisition of hand gesture data through a camera, followed by processing via the MediaPipe framework to detect and track 21 hand landmarks. The data are processed to identify patterns of hand movements using a Neural Networks-based algorithm. Upon identifying movement patterns, the results are converted into relevant control commands, including play, pause, or volume adjustment on the web-based video player.

DOI: https://doi.org/10.17509/edsence.v7i1.82344
p-ISSN:2685-2489 | e-ISSN:2685-2535

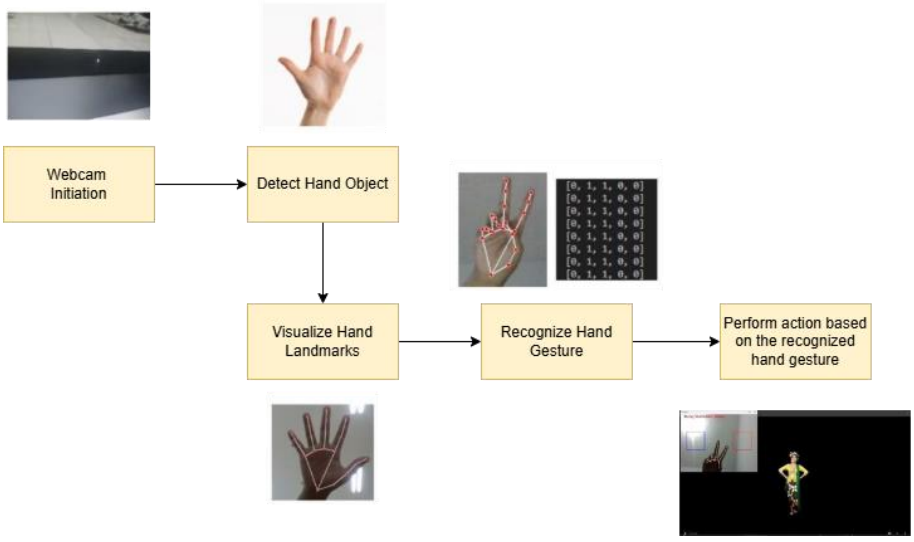

Figure 2. Workflow Design for System

2.3 Implementation

In the implementation phase, the MediaPipe-based hand gesture recognition system for controlling web video players was developed through the following steps:

1. Data Acquisition and Environment Setup
   The initial stage of the system encompasses data acquisition and the configuration of the environment. A camera captures visual input of the user's hand movements, which are processed by the MediaPipe framework for real-time detection and tracking of 21 hand landmark positions. The accurate positioning of the camera is essential for the clear detection of the user's hand within the image frame. The system's initiation necessitates the installation of the MediaPipe framework and its requisite dependencies. In a Python programming environment, this involves the installation of necessary libraries, including mediapipe, opencv-python, and other critical packages for image and video processing capabilities.
2. Hand Landmark Recognition
   The system utilizes the pre-trained MediaPipe Hand Landmark Model for hand landmark recognition, facilitating real-time detection of 21 specific landmarks that include the positions of the fingers, palm, and wrist. This study did not utilize a custom dataset for training the gesture recognition model from scratch; instead, it leveraged MediaPipe's inherent pre-trained model, which has been robustly trained on diverse hand datasets, thereby enabling robust landmark detection.The detection process occurs on each frame of the video stream, facilitating continuous tracking of the user's hand movements. The spatial information of these landmarks, indicated by their x, y, and z coordinates, is extracted from the image data, and used as features for the identification of hand movement patterns.
3. Mapping Hand Gestures to Control Commands
   Mapping hand gestures to control commands entails two fundamental aspects. Gesture recognition involves analysing the relative positions and movements of detected hand landmark points to identify specific hand configurations that correspond

DOI: https://doi.org/10.17509/edsence.v7i1.82344
p-ISSN:2685-2489 | e-ISSN:2685-2535

to predefined control commands. An open hand can signify the "play" command, a clenched fist indicates "pause," and vertical hand movements represent volume adjustments. The gesture mapping is executed via a rule-based approach, where each identified pattern of hand landmark positions and movements is converted into a specific control command according to established criteria, including the Euclidean distances between landmarks and the angular variations among the fingers and the hand.

4. Integration with Web Video Player
   Upon recognizing a specific hand gesture, the system transmits the corresponding control commands to the HTML5-based web video player (Earl & Neal, 2016). This control is achieved via JavaScript, enabling the manipulation of various video player functions, such as initiating and terminating playback (play/pause), adjusting volume, and navigating the video. JavaScript can be used to interact with the video player's Application Programming Interface (API), allowing for direct programmatic manipulation of the video elements embedded in the webpage (Acosta et al., 2020). The "pause" command corresponds to the video.pause() method, whereas the "play" command triggers the video.play() method.

2.4 Evaluation

The system's performance was assessed during the evaluation phase across diverse environmental conditions, such as changes in lighting, user distance from the camera, and system latency. The tests aimed to evaluate gesture recognition accuracy, system responsiveness, and user comfort levels. The testing results were analyzed to assess the system's effectiveness in realistic usage scenarios and to identify areas requiring further refinement.

3. RESULTS AND DISCUSSION

This research yielded the development of a hand gesture recognition-based control system for web video players, implemented utilizing the MediaPipe framework. The outcomes of this study are presented in two primary sections: Results and Evaluation.

3.1 Results

3.1.1 Gesture Recognition Screenshot Samples

In this section, we present a selection of screenshots illustrating the system's capability to recognize various hand gestures performed by users. By employing the MediaPipe framework, the system is able to detect hand movements and translate them into corresponding control commands. The following are examples of the hand gesture recognition outcomes. The samples on Figure3-4 illustrate the system's ability to recognize various hand gestures, which are integral to motion-based device control.

DOI: https://doi.org/10.17509/edsence.v7i1.82344
p-ISSN:2685-2489 | e-ISSN:2685-2535

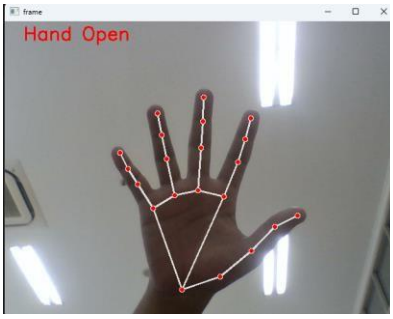

Figure 3a. Hand Open Sign

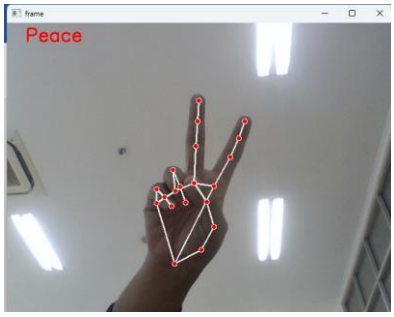

Figure 3b. Peace Sign

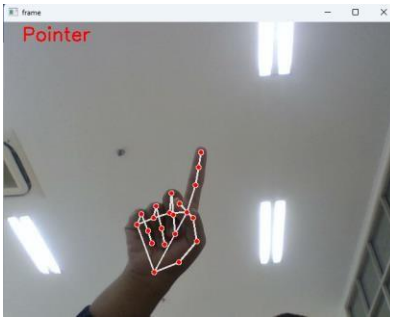

Figure 3c. Pointer Prev

Figure 4a. Thumbs up Sign

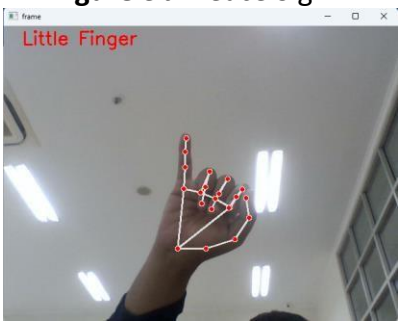

Figure 4b. Little Finger Sign

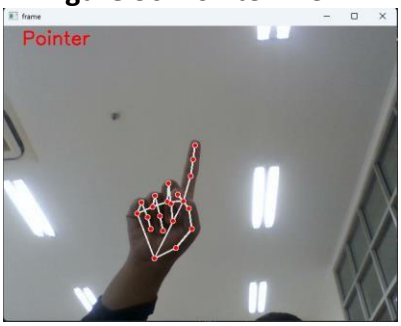

Figure 4c. Pointer Next

3.1.2 Integration of MediaPipe in the Web Video Player

Following gesture recognition, the next essential step is the integration of this functionality into the web video player application. This implementation relies on a defined mapping between specific hand gestures and control commands, as outlined in Table 1.

Table 1. Hand Gesture and Control Features

Hand Gesture	Control Features
Hand Open	Play / Pause
Peace	Mute / UnMute
Thumbs Up	Volume Up
Little Finger	Volume Down
Pointer Blue	Next Video
Poiter Red	Prev Video

Figures 5-10 provide a visual representation of these hand gestures and highlight their key characteristics, which are then recognized by the MediaPipe framework. The recognized hand movements are subsequently translated into the corresponding control commands for the web-based video player application, enabling users to control different functions (Play/Pause, Mute/UnMute, Volume Up/Down, Next/Previous Video) exclusively through these intuitive hand gestures, eliminating the need for physical input devices like a mouse or keyboard.

DOI: https://doi.org/10.17509/edsence.v7i1.82344
p-ISSN:2685-2489 | e-ISSN:2685-2535

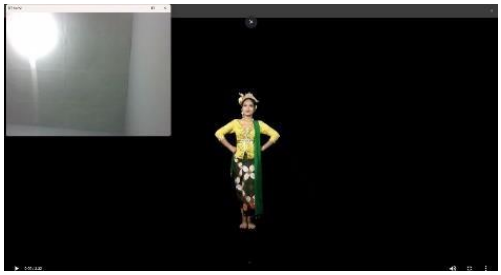

Figure 5a. Before Hand Open

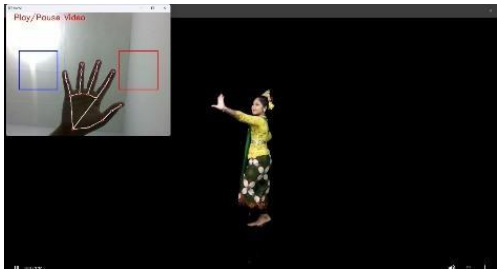

Figure 5b. After Hand Open

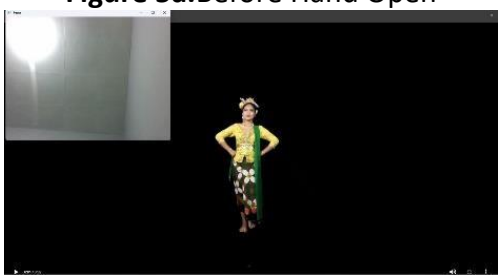

Figure 6a. Before Peace

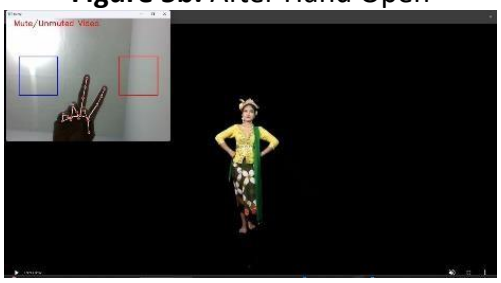

Figure 6b. After Peace

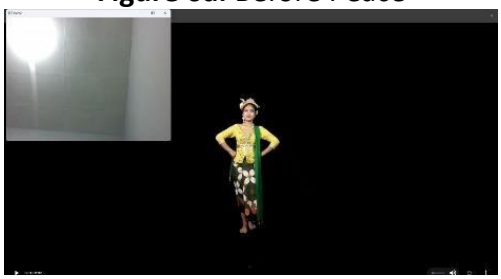

Figure 7a. Before Thumbs up

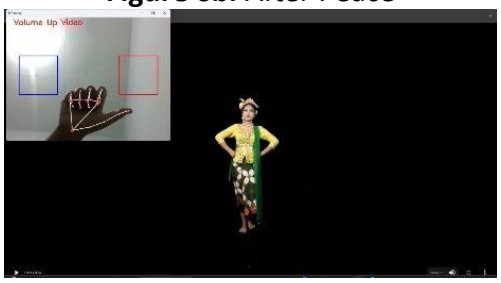

Figure 7b. After Thumbs up

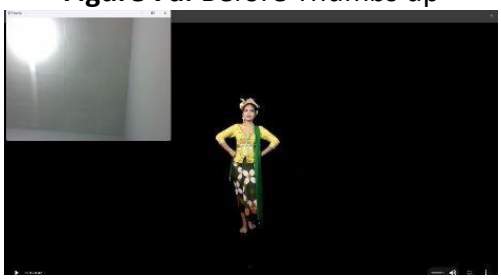

Figure 8a. Before Little Finger

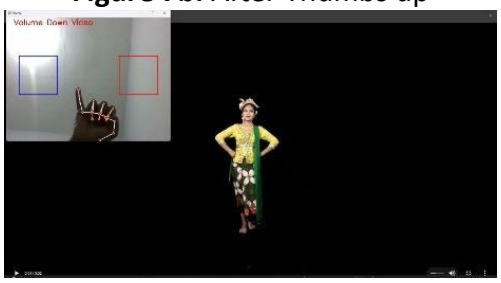

Figure 8b. After Little Fingger

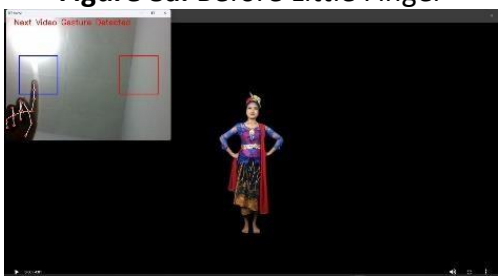

Figure 9a. Before Pointer Blue

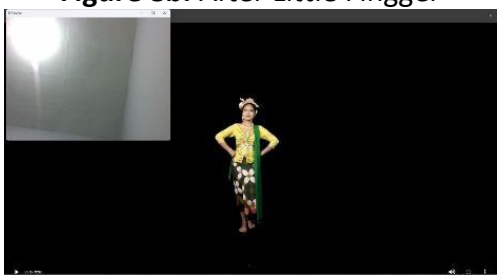

Figure 9b. After Pointer Blue

DOI: https://doi.org/10.17509/edsence.v7i1.82344
p-ISSN:2685-2489 | e-ISSN:2685-2535

Figure 10a. Before Pointer Red

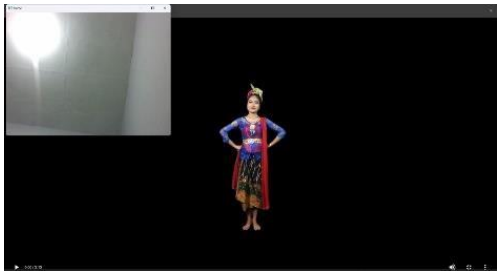

Figure 10b. After Pointer Red

3.2 Evaluation

The evaluation of the system aimed to assess the performance and reliability of the hand gesture recognition-based web video player control system, focusing on two main factors: the user's distance from the camera and the ambient lighting conditions (dark versus bright environments). The objective of this testing was to assess the system's effectiveness in various realistic usage scenarios.

3.2.1 Evaluation of User Distance from the Camera

The initial evaluations focused on assessing the accuracy of gesture recognition as the distance between the user's hand and the camera varied. Figure 11 illustrates the visual input at these different distances, ranging from a clear close-up at 0.5 meters to a more distant view at 3 meters, during which participants performed a range of hand gestures.

Figure 11a. Close range evaluation (0.5 meters)

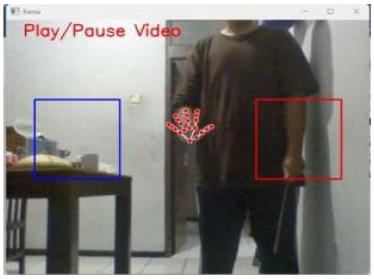

Figure 11b. Intermediate range evaluation (1.5 meters)

Figure 11c. Far range evaluation (3.0 meters)

Evaluation results:

1) Close Range (0.5 meters): The system demonstrated the ability to detect nearly all gestures with very high accuracy. The system's responsiveness was rapid, and hand movements were detected with good precision.
2) Intermediate Range (1.5 meters): The accuracy of gesture recognition experienced a slight decline; however, most gestures were still detectable with reasonable accuracy. Transitions between gestures were observed to be detected with a somewhat slower response time.
3) Far Range (3 meters): At this distance, gesture recognition exhibited a significant decrease in performance. Nearly all gestures were either difficult or impossible to detect. For example, the "Thumbs Up," "Little Finger," and "Pointer" gestures were challenging or even undetectable."

DOI: https://doi.org/10.17509/edsence.v7i1.82344
p-ISSN:2685-2489 | e-ISSN:2685-2535

3.2.2 Evaluation of Ambient Lighting Condition

The second set of evaluations focused on the system's ability to recognize hand gestures under varying illumination conditions. These tests were conducted under two primary conditions: bright lighting and a dark environment.

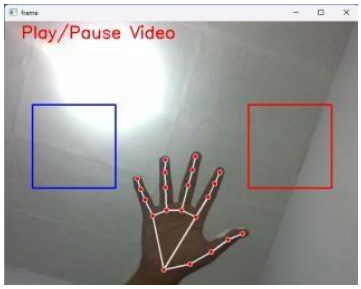

Figure 12a. Bright Illumination Evaluation

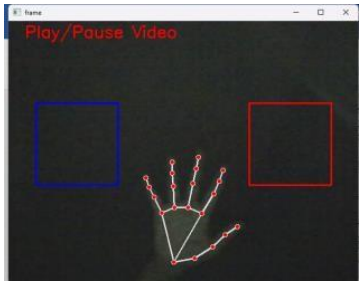

Figure 12b. Dark illumination evaluation

Figure 12c. Dark illumination evaluation

Evaluation results:

1) Bright Illumination: As illustrated in Figure 12a, which demonstrates the accurate recognition of all tested hand gestures, the system performed exceptionally well under bright lighting conditions. Hand gestures were recognized with high accuracy, and the system's response was very rapid.
2) Dark Illumination: In contrast, the system's accuracy decreased significantly in dark conditions, as depicted in Figure 12b, which shows examples of failed gesture recognition. Figure 12c further illustrates the challenging visual environment of the darker setting where several gestures were not detected. Consequently, a sufficiently bright minimum level of illumination is essential for the system to function optimally.

3.2 Discussions.

The incorporation of MediaPipe into the hand gesture-based control system for web video players significantly enhances various aspects of user interaction. Firstly, the responsiveness of the interface is significantly improved. Contactless control through hand gestures enables users to manipulate the web video player directly, eliminating the necessity for traditional physical input devices like mice and keyboards. This leads to quicker and more responsive interactions, enabling smoother and more intuitive control.

Furthermore, the system provides intuitive interactivity. The use of intuitive hand gestures, such as "Thumbs Up" for increasing volume and "Hand Open" for starting and stopping video playback, enhances the user experience. This non-contact interaction may enhance system usability and user engagement.

Thirdly, hand gesture-based control enhances the user experience. Users are able to control multiple video playback parameters, such as play, pause, volume adjustment, mute/unmute, and video navigation, exclusively via hand movements. This provides significant convenience, especially in situations where access to or the use of physical input devices is limited or ergonomically difficult.

The results of the system performance evaluation demonstrate that the user's proximity to the camera and the ambient lighting conditions significantly affect gesture recognition accuracy. At a distance of 0.5 meters, the system exhibits superior performance

DOI: https://doi.org/10.17509/edsence.v7i1.82344
p-ISSN:2685-2489 | e-ISSN:2685-2535

characterized by high recognition accuracy and swift response times. At distances greater than 1.5 meters, a notable decrease in accuracy is evident, especially for certain gestures including "Thumbs Up," "Little Finger," and "Pointer." Thus, the identification of optimal camera placement and management of user distance are critical factors in the implementation of the system.

Moreover, testing under different illumination conditions demonstrates that sufficient lighting is crucial for maintaining reliable system performance. In poorly illuminated settings, the accuracy of gesture recognition significantly declines, leading to inconsistent detection of specific gestures. This highlights the significance of evaluating the ambient lighting conditions within the system's operational environment.

In its current implementation, the system is designed to independently recognize single hand gestures, whether from the left or right hand. The MediaPipe model intrinsically supports landmark detection for both hands. However, the current gesture-to-control mapping mechanism is optimized for detecting an active single hand. If both hands are simultaneously detected, the system prioritizes the more clearly identified hand or the first hand to perform a defined gesture. Simultaneous use of two hands for distinct gestures has not yet been implemented and represents an area for future research.

4. CONCLUSION

This study developed a control system for web video players based on hand gesture recognition, employing the MediaPipe framework. This system allows users to control multiple video playback functions, including play, pause, volume adjustment, and video navigation, exclusively through real-time hand gestures. This method facilitates a more intuitive, natural, and contactless user interaction with the device, thereby improving convenience and efficiency, especially in situations where physical input devices are unavailable.

The test results demonstrated that the system operated at optimal performance levels under bright illumination conditions and when the user was positioned at a distance of approximately 0.5 meters from the camera. Gesture recognition accuracy decreased with increased distance and in low-light conditions. This highlights the significance of optimal camera positioning and adequate lighting to guarantee effective system operation. This study demonstrates that hand gesture-based video player control utilizing MediaPipe enhances user interactivity and responsiveness. This technology has potential for further development in various applications, including education, entertainment, and virtual conferencing. The system can be effectively implemented in large-scale events, such as conferences or artistic performances, where various types of videos with differentiated control are necessary, enhancing the audience's experience.

This research extends beyond general applications to investigate the implementation of the hand gesture recognition system in specific contexts, including conferences and performances. Gesture-controlled videos extend beyond traditional video content, enabling the management of diverse media forms, including introductory videos, conclusions, narratives, and other multimedia presentations. The application of hand gesture recognition in events provides a dynamic and interactive approach to video control, improving audience engagement and facilitating a more seamless interaction between presenters and the content displayed(Maharana et al., 2024). For future work, this research can be further

DOI: https://doi.org/10.17509/edsence.v7i1.82344
p-ISSN:2685-2489 | e-ISSN:2685-2535

enhanced in several ways. Firstly, exploring the use of deep learning models trained on custom gesture datasets could improve recognition accuracy, particularly at greater distances and in suboptimal lighting conditions. Secondly, the system could be developed to recognize two-handed gestures or gesture combinations for more complex control functionalities. Thirdly, a procedural system latency test is needed to measure the average time required from hand gesture execution to the corresponding command being executed on the web video player. Fourthly, integrating with Augmented Reality (AR) or Virtual Reality (VR) technologies could be explored to create a more immersive multimedia interaction experience. Finally, further testing in real-world event scenarios with a more diverse range of participants would provide deeper insights into the system's scalability and usability in dynamic environments.

ACKNOWLEDGMENT

This research was supported by DIPA funds from the Bandung State Polytechnic, under the Applied Research funding agreement No. 235.4/PL1.R7/PG.00.03/2024. The authors extend their sincere gratitude to the dedicated students who assisted in the data collection and system implementation phases of this research.

AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. Authors confirmed that the paper was free of plagiarism.

REFERENCES

Acosta, T., Zambrano-Miranda, J., & Lujan-Mora, S. (2020). Techniques for the Publication of Accessible Multimedia Content on the Web. *IEEE Access*, *8*, 55300–55322.

Agustiani, A. D., Putri, S. M., Hidayatullah, P., & Sholahuddin, M. R. (2024). Penggunaan MediaPipe untuk Pengenalan Gesture Tangan Real-Time dalam Pengendalian Presentasi. *Media Jurnal Informatika*, *16*(2).

Awaludin, I., Sholahuddin, M. R., Gelar, T., Engineering, I., & Bandung, P. N. (2024). 3D Large-Screen Immersive Video Mapping Installations & Interactive Light Game for Conference Seminars. *Journal of Computer Engineering , Electronics and Information Technology (COELITE)*, *3*(2), 89–100.

Dewangga, S., & Subianto, Mochamad Swastika, W. (2024). Implementation of Hand Gesture Recognition as Smart Home Devices Controller. *INSYST*, *6*(2), 63–68.

Earl, C., & Neal, E. (2016). HTML5 accessible video player. *Proceedings of the 13th International Web for All Conference*, 1–4.

Google Developers. (2025). *Hand landmarks detection guide*. https://ai.google.dev/edge/mediapipe/solutions/vision/hand_landmarker

Haria, A., Subramanian, A., Asokkumar, N., Poddar, S., & Nayak, J. S. (2017). Hand Gesture Recognition for Human Computer Interaction. *Procedia Computer Science*, *115*, 367–374.

Indriani, Harris, M., & 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)*, *207*(Issat), 101–108.

Khairianto, D., & Firdaus, R. (2024). Penerapan Hand Gesture Recognition Sebagai Media Kontrol Presentasi Aplikasi Powerpoint. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(2), 1852–1860.

Kriznar, V., Leskovsek, M., & Batagelj, B. (2021). Use of Computer Vision Based Hand Tracking in Educational Environments. *2021 44th International Convention on Information, Communication and Electronic Technology, MIPRO 2021 - Proceedings*, 804–809.

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M. G., Lee, J., Chang, W.-T., Hua, W., Georg, M., & Grundmann, M. (2019). MediaPipe: A Framework for Building Perception Pipelines. *Arxiv*.

Maharana, S., Singh, R. K. B., Priya, M. P., & Aswini, M. (2024). SMART Virtual Painter and Mouse System to Improve Online Education with Computer Vision and Deep Learning Features. *ESIC 2024 - 4th International Conference on Emerging Systems and Intelligent Computing, Proceedings*, 768–773.

Panwar, M., & Singh Mehra, P. (2011). Hand gesture recognition for human computer interaction. *2011 International Conference on Image Information Processing*, 1–7.

Roy, K., & Akif, M. A. H. (2022). Real Time Hand Gesture Based User Friendly Human Computer Interaction System. *2022 International Conference on Innovations in Science, Engineering and Technology (ICISET)*, 260–265.

Saputra, W., & Hariyanto, S. (2024). Implementasi Hand Gesture Mediapipe Pada Game Interaktif Untuk Stimulasi Motorik Anak Down Syndrome Pada Sekolah Luar Biasa Autis Kalimantan Barat. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(5), 10776–10784.

DOI: https://doi.org/10.17509/edsence.v7i1.82344
p-ISSN:2685-2489 | e-ISSN:2685-2535

DOI: https://doi.org/10.17509/edsence.v7i1.82344
p-ISSN:2685-2489 | e-ISSN:2685-2535