IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 5, May 2025

DOI: 10.17148/IJIREEICE.2025.13515

Gesture-Controlled Music Player Using Hand Gestures

Akshat Singh[1], Ankit Rai[2], Sher Mohammad[3], MR. KUMAR BIBHUTI BHUSHAN SINGH[4]

Department of Computer Science & Engineering, Goel Institute of Technology and Management, Lucknow[1-3]

Assistant Professor, Department of Computer Science & Engineering, Goel Institute of Technology and Management, Lucknow[4]

Abstract: This paper presents a gesture-controlled music player system that interprets real-time hand gestures to control media functions such as play, pause, next/previous track, and volume adjustment. The project aims to deliver a touchless user experience, offering intuitive control of digital media via simple hand movements detected through a webcam. The system utilises computer vision technologies, including Opencv, MediaPipe, and Python's gesture recognition algorithms. Designed with accessibility and hygiene in mind, this music player is particularly relevant in environments where hands-free interaction is preferred or necessary. Testing showed gesture accuracy of over 90% under good lighting conditions, confirming the system's practical usability. Future enhancements will focus on integrating machine learning for personalised gesture mapping and expanding crossplatform support.

I. INTRODUCTION

In an era where touchless interactions are increasingly significant, particularly in healthcare, public utilities, and consumer electronics, gesture recognition presents a compelling alternative to conventional input methods. Music players, being among the most frequently used digital tools, stand to benefit from gesture-based interfaces that offer both convenience and accessibility. This research introduces a gesture-controlled music player using real-time hand tracking to execute media control operations. The proposed system addresses user needs for intuitive, hygienic, and efficient multimedia control, especially in scenarios like kitchens, laboratories, or during workouts where physical contact with devices may be impractical or undesired.

II. LITERATURE REVIEW

Hand gesture recognition has seen considerable advancement with the development of computer vision and deep learning technologies. Early systems relied on hardware like Leap Motion or Kinect, but recent solutions favour camera-based tracking using libraries such as Opencv, TensorFlow, and Google's MediaPipe. Research by Zhang et al. (2021) on dynamic gesture classification achieved high accuracy using CNNS with depth data. MediaPipe's hand landmark tracking has been widely adopted due to its lightweight real-time performance and platform independence. In music control applications, gesture recognition has been mostly limited to research prototypes, with few open-source or commercial implementations achieving seamless integration with existing media players.

III. METHODOLOGY

3.1 System Architecture

The music player system consists of four core modules:

Video Capture: Utilises a webcam to stream live input.

Hand Landmark Detection: MediaPipe's pre-trained hand detection model extracts 21 key landmarks per hand.

Gesture Classification: Custom Python scripts interpret landmark positions to detect gestures like fist (pause/play), thumb up/down (volume), and swipe left/right (previous/next track).

Media Control: Uses Python's pyautogui or keyboard library to simulate keyboard shortcuts for music control.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526
International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 5, May 2025
DOI: 10.17148/IJIREEICE.2025.13515

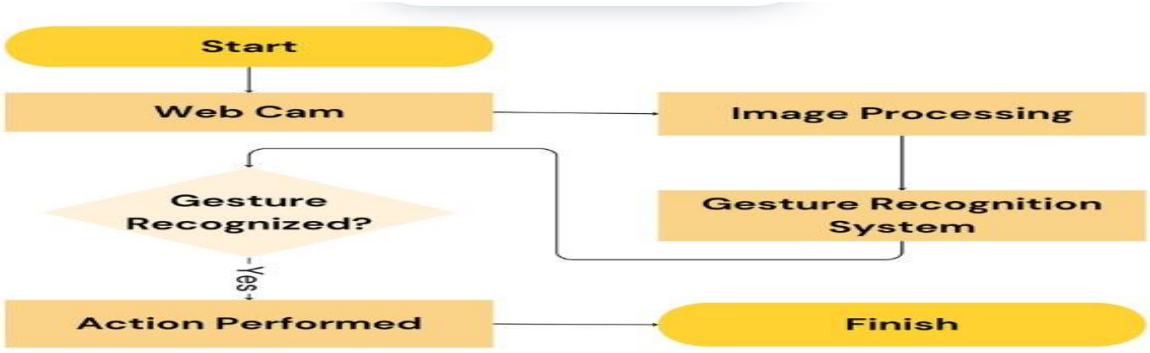


Figure No 1:- System Flow Diagram

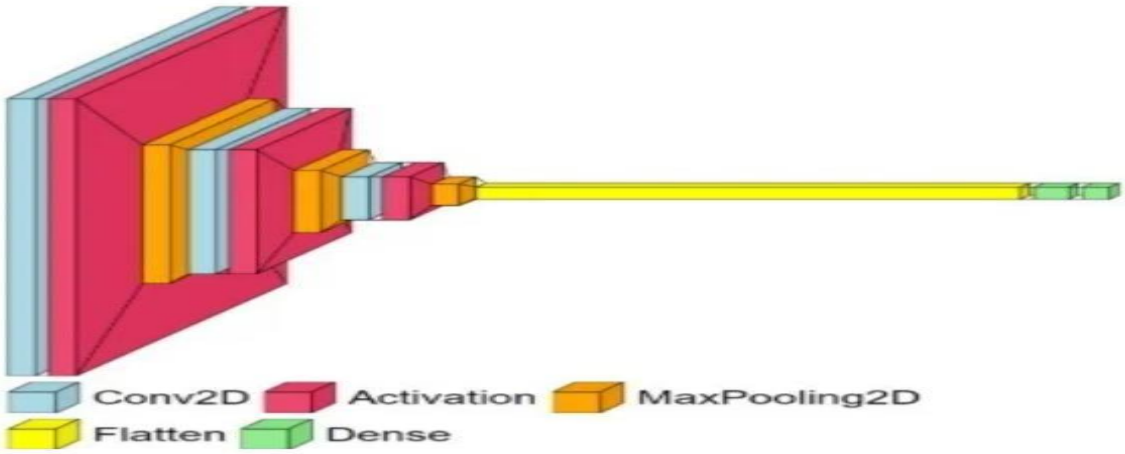


Figure No 2 :- Architecture Of Trained Mode.

3.2 Software Stack
Platform: Windows
Programming Language: Python
Libraries: Opencv, MediaPipe, PyAutoGUI
Hardware: Standard webcam (720p or higher)

3.3 Gesture Definitions

Gesture	Action
Open palm	Play
Closed fist	Pause
Thumb up	Volume Up
Thumb down	Volume Down
Swipe right	Next song
Swipe left	Previous song

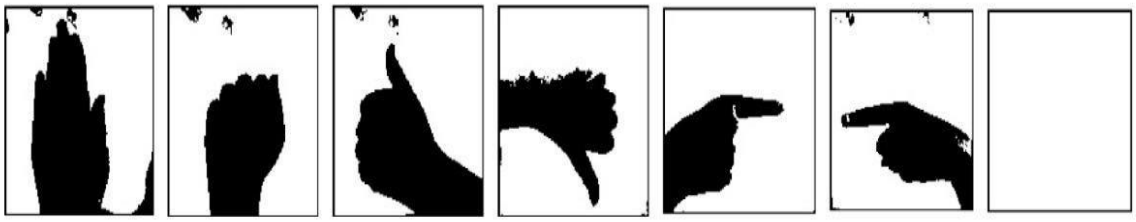
IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 5, May 2025

DOI: 10.17148/IJIREEICE.2025.13515

Gestures obtained after Data collection and preprocessing: Palm, fist, thumbs up, thumbs down, index pointing right, index pointing left and no gesture (Left to right)



Gestures are determined based on distances and angles between detected hand landmarks, calibrated for robustness to hand size and orientation.

IV. RESULTS AND DISCUSSION

The system was tested across various environments and lighting conditions with a user group of 20 individuals. Performance metrics are summarised below:

Play/Pause:	94%
Volume Control:	89%
Song Navigation:	91%
Response Time:	Average of 0.3 seconds from gesture to action.

User Feedback:	85% of participants rated the system as intuitive and effective. Users appreciated the handsfree control, especially in handsoccupied contexts.

Challenges included poor recognition in low-light environments and slight lag during rapid gesture transitions. Integration with native operating system media players using background processes worked effectively without disrupting ongoing tasks.

V. SYSTEM OVERVIEW

The system flow begins with webcam initialisation, followed by continuous frame analysis using MediaPipe. Once a gesture is recognised, it is mapped to a predefined media control command and executed using OS-level automation libraries. The interface includes a minimal GUI to provide visual feedback (gesture detected, action performed). A calibration module allows users to customise gesture sensitivity.

VI. FUTURE SCOPE

This prototype opens avenues for significant improvements:
Machine Learning Integration: Use of custom-trained deep learning models to recognise a wider range of gestures.
Gesture Customisation: User-defined gesture mappings for accessibility.
Mobile App Development: Extend support to Android/ios using crossplatform frameworks.
Multi-Hand Support: Enable simultaneous recognition of two hands for complex gesture commands.
Smart Home Integration: Control other devices like lights or appliances via the same gesture interface.

VII. CONCLUSION

The gesture-controlled music player introduces an innovative and userfriendly way to manage digital music playback. It demonstrates the practicality of computer vision-based gesture recognition in real-world applications and showcases how AI can enhance user experience by reducing dependence on physical touch. With continued refinement, such systems can become standard in smart environments, contributing to both convenience and public health.

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 5, May 2025

DOI: 10.17148/IJIREEICE.2025.13515

REFERENCES

[1] F. Zhang, L. Wang, "Real-time Hand Gesture Recognition using Deep Learning," IJCV, 2021.
[2] Google AI, "MediaPipe: Cross-Platform Framework for Building Perceptual AI," 2023.
[3] S. Sharma, "Gesture-Based Media Control using Opencv and Python," TechXplore, 2022.
[4] D. Kaur, "Touchless Interaction Using AI Vision," Springer AI Journal, 2024.
[5] GitHub – MediaPipe Hand Tracking Example.
[6] A. Jain, "PyAutoGUI for Automation," Python Weekly, 2023.