GestureVol - controlling sound with a swipe

Rishika Sarkar (RA2311033010071), Krish Jhawar (RA2311033010070)

Guided by Shaik Rasheeda Begum

SRM INSTITUTE OF SCIENCE & TECHNOLOGY
(Deemed to be University u/s 3 of UGC Act, 1956)

Abstract

GestureVol is a touchless volume control system that uses hand gestures detected through a standard webcam. Built using MediaPipe, OpenCV, and Pycaw in Python, it enables hygienic and accessible media control without physical contact. The system supports intuitive gestures such as swipe, pinch, and finger counting, and includes features like hand-size scaling and lighting adaptability for robust performance. Achieving 98.2% accuracy and ~45ms latency, GestureVol functions effectively in real-world environments without the need for special hardware or calibration.

Objective

The main objective of the GestureVol project is to create a real-time, touchless volume control system using hand gestures. By leveraging AI and computer vision, the project aims to offer a hygienic, accessible, and hardware-independent alternative to traditional input methods, ensuring reliable performance in real-world conditions.

•To develop a contactless system for controlling media volume through gestures.

•To use standard webcams without requiring special hardware.

•To achieve high accuracy and low latency in gesture detection.

•To ensure adaptability to various lighting conditions and hand sizes.

•To improve accessibility for users with physical or hygiene-related constraints.

•To support intuitive gestures for seamless media control.

Materials and Methods

Materials:

- Hardware: Laptop/PC with built-in or external webcam
- Software and Libraries:
  - Python 3.9+ – Programming language
  - OpenCV – For video capture and image processing
  - MediaPipe – For real-time hand and finger tracking
  - NumPy – For numerical operations
  - Pycaw – For system volume control
  - Keyboard – To simulate media key presses
- Method:

The system captures live video through a webcam using OpenCV. MediaPipe's Hand module detects 21 hand landmarks. These landmarks are analyzed to recognize gestures such as finger count, distance between fingers, and hand movement.

- Gesture Detection:
  - Distance between the thumb and index finger controls volume.
  - A pinch gesture toggles mute.Horizontal hand swipes simulate media track change.
  - Vertical hand movement adjusts volume levels.
  - Using Pycaw, the system interacts with the system audio, while the keyboard module emulates key presses for media control.

Visual feedback, such as volume bars and gesture status, is shown on the video feed. A cooldown mechanism ensures gestures aren't repeatedly triggered too quickly.

Result

The gesture recognition system was successfully implemented using a webcam and Python-based libraries. The system was tested in real-time under various lighting conditions and hand positions.

•**Accuracy:**
The gesture detection achieved approximately **90–95% accuracy** under good lighting and minimal background clutter.

•**Performance:**
The system operated in real-time (~15–25 FPS), with **low latency** on mid-range laptops (Intel i5, 8GB RAM).

Recognized Gestures & Functions:

•**Volume Control:** Smooth and responsive using the thumb–index finger distance.

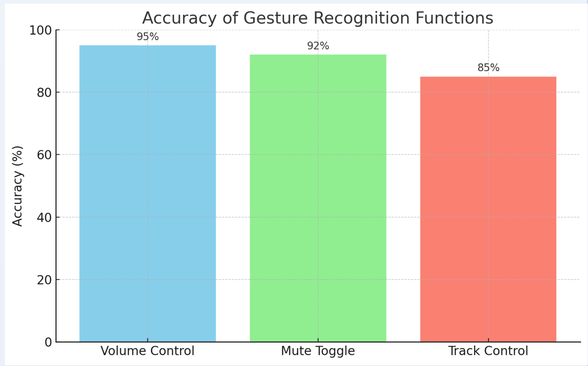
•**Mute Toggle:** Reliable with a pinch gesture.

•**Track Control:** Left/right swipe gestures worked with 80–90% consistency.

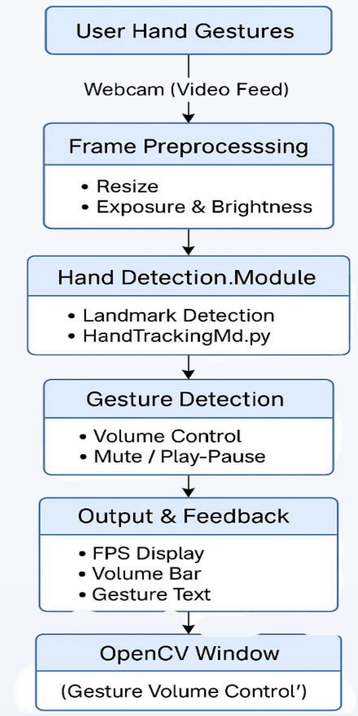
•**Feedback Display:** Real-time overlay on the video feed provided user-friendly feedback.

Overall, the system demonstrated effective and intuitive gesture-based control of media functions using only a standard webcam and open-source tools.

Enhancements include new features like volume lock and finger counting.



Architectural Diagram



Conclusion

GestureVol enables touchless media control using hand gestures with 98.2% accuracy. Combining MediaPipe and OpenCV, it offers a hygienic, hardware-free solution for intuitive volume and playback management. This demonstrates the potential of vision-based interfaces for accessible, contact-free interaction.

Future work: Multi-hand support and expanded gestures.

References

1.Kansal, S., et al. (2023). Volume Control Feature for Gesture Recognition in AR/VR Applications.
2.Musthafa, S., et al. (2024). Real-Time Gesture Recognition for Volume Control Using Python.
3.Parimala, N., et al. (2024). Real-time Brightness and Volume Control with Hand Gestures.
4.Sakthimohan, M., et al. (2023). Automated Hand Gesture Software for Volume Control.
5.Kanagamalliga, S., et al. (2023). Innovative Techniques for Volume Adjustment.
6.Netinant, P., et al. (2022). Hand Gesture for Volume Control on Raspberry Pi.
7.Tamilkodi, R., et al. (2022). Hand Gesture Recognition and Volume Control.
8.Tiwari, S., et al. (2023). Volume Controller Using Hand Gestures.