Hand Gesture MediaPlayer

A PROJECT REPORT

Submitted by

Vishal Dolasia (23BHI10079)

BACHELORS OF TECHNOLOGY

COMPUTING SCIENCE & ENGINEERING
(HEALTH INFORMATICS)

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING AND

ARTIFICIAL INTELLIGENCE

VIT BHOPAL UNIVERSITY

KOTHRI KALAN, SEHORE
MADHYA PRADESH - 466114

NOVEMBER 2025

VIT BHOPAL UNIVERSITY, KOTHRI KALAN, SEHORE
MADHYA PRADESH – 466114

ABSTRACT

The Hand Gesture Media Player is a real-time application of computer vision that enables users to interact with their computer's Media Player using only their hand, which is captured via a standard webcam. Traditional physical input devices such as mouse and keyboard are not available everywhere and therefore cannot be used by everyone .This project aims to replace the traditional methods and create an intuitive, contactless interface, enhancing accessibility and usability—particularly in environments such as YT videos ,presentations, interactive installations, or for individuals with physical limitations.

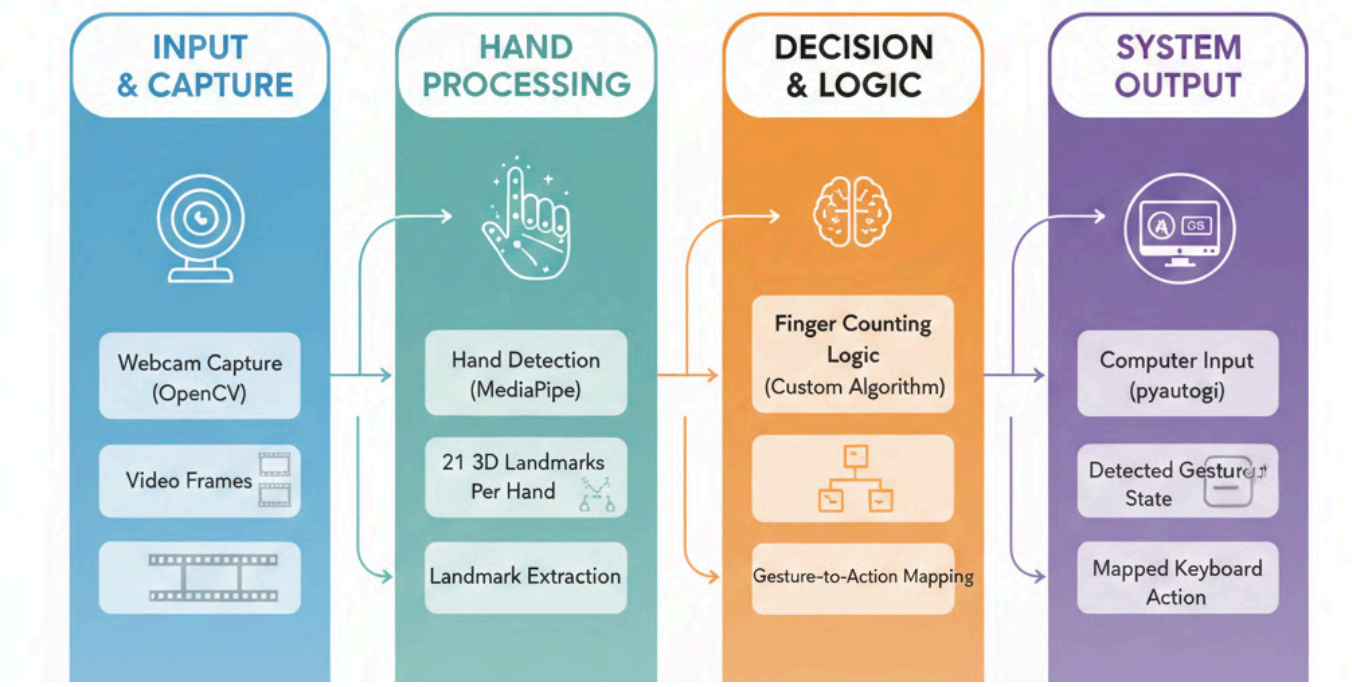
INTRODUCTION

Keyboards and mice are examples of physical devices that are crucial to traditional computer control systems. Despite this , there is a significant need for more accessible and practical hands-free options.This can be very much achieved by the help of computer vision, which can provide easily provide hand-free methods that are more accessible and are useful in vary situations like for watching YT videos , PPTs and for individuals with various disabilities.

System Architecture

The Hand Gesture Media Player follows a sequential data processing pipeline:

- **Webcam Capture**: A live video stream (frame by frame) using a camera is acquired using OpenCV.
- **Hand Detection**: Google's MediaPipe framework detects hands and extracts precise key points (21 landmarks per hand).
- **Landmark Extraction**: The model outputs 3D positional data for each landmark.
- **Finger Counting Logic**: A custom algorithm analyzes the distance between specific landmarks to determine which fingers are extended.
- **Gesture-to-Action Mapping**: With the help of the "**pyautogui**" library, each detected gesture is mapped to a keyboard action.
- **Computer Input**: "**pyautogui"** triggers corresponding keyboard actions on the computer.



Core Technologies

- **Python**: The main programming language for the project implementation.
- **OpenCV**: Handles webcam access and to capture video stream (frame by frame) which is displayed in real time.
- **MediaPipe Hands**: Provides highly accurate hand keypoint to detect which gesture is being played.
- **pyautogui**: Automates keyboard control based on recognized gestures

Finger Counting Algorithm

The finger counting procedure is based on the analysis of MediaPipe's keypoint data:

- **Threshold Calculation**: For non-thumb fingers, a dynamic “thresh” value is derived from the y-coordinates of the wrist and middle finger MCP (**Metacarpophalangeal Joint**).

$$\text{thresh} = \frac{(y_0 \times 100 - y_9 \times 100)}{2}$$

- **Non-Thumb Fingers**: Each finger is considered "up" if its MCP-to-tip vertical distance exceeds the threshold:
  - Index: $y5 - y8 > \text{thresh}$
  - Middle: $y9 - y12 > \text{thresh}$
  - Ring: $y13 - y16 > \text{thresh}$
  - Pinky: $y17 - y20 > \text{thresh}$

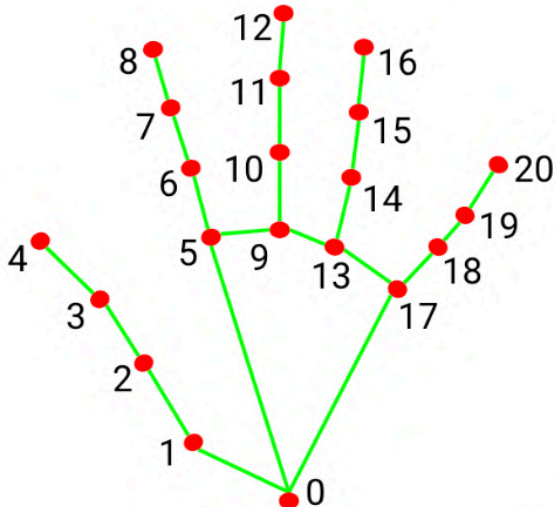


Fig:- Landmark

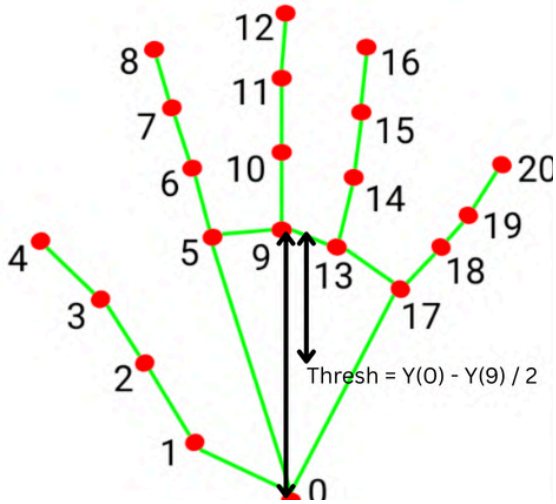


Fig:- Threshold Value

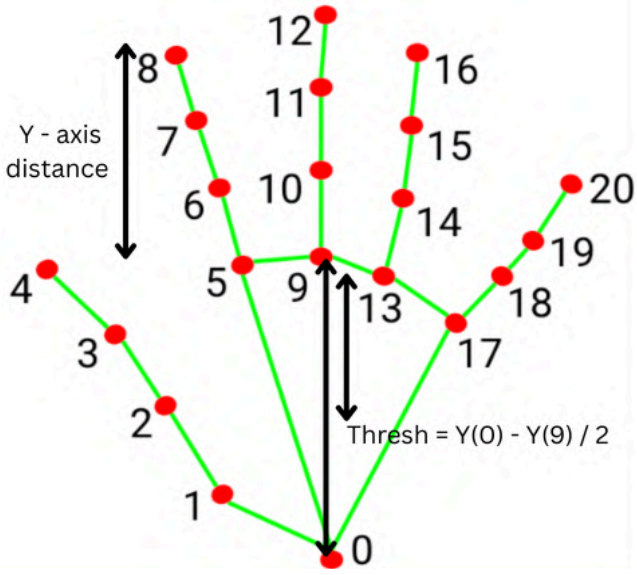


Fig:- Y-Axis Distance

- **Thumb**: The thumb's status uses the horizontal distance between the Index MCP and Thumb tip. Thumb is up if:
  $x5 - x4 > 5$

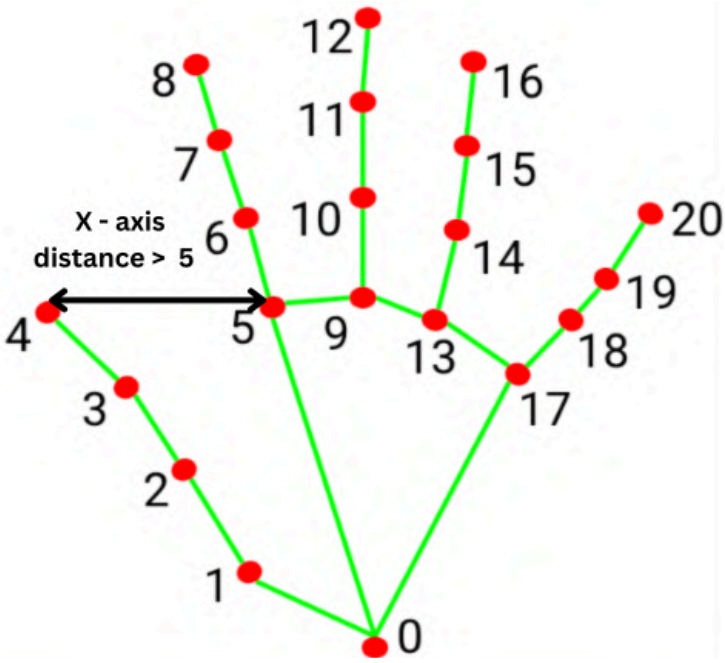


Fig:- X-Axis Distance

Gesture Mapping

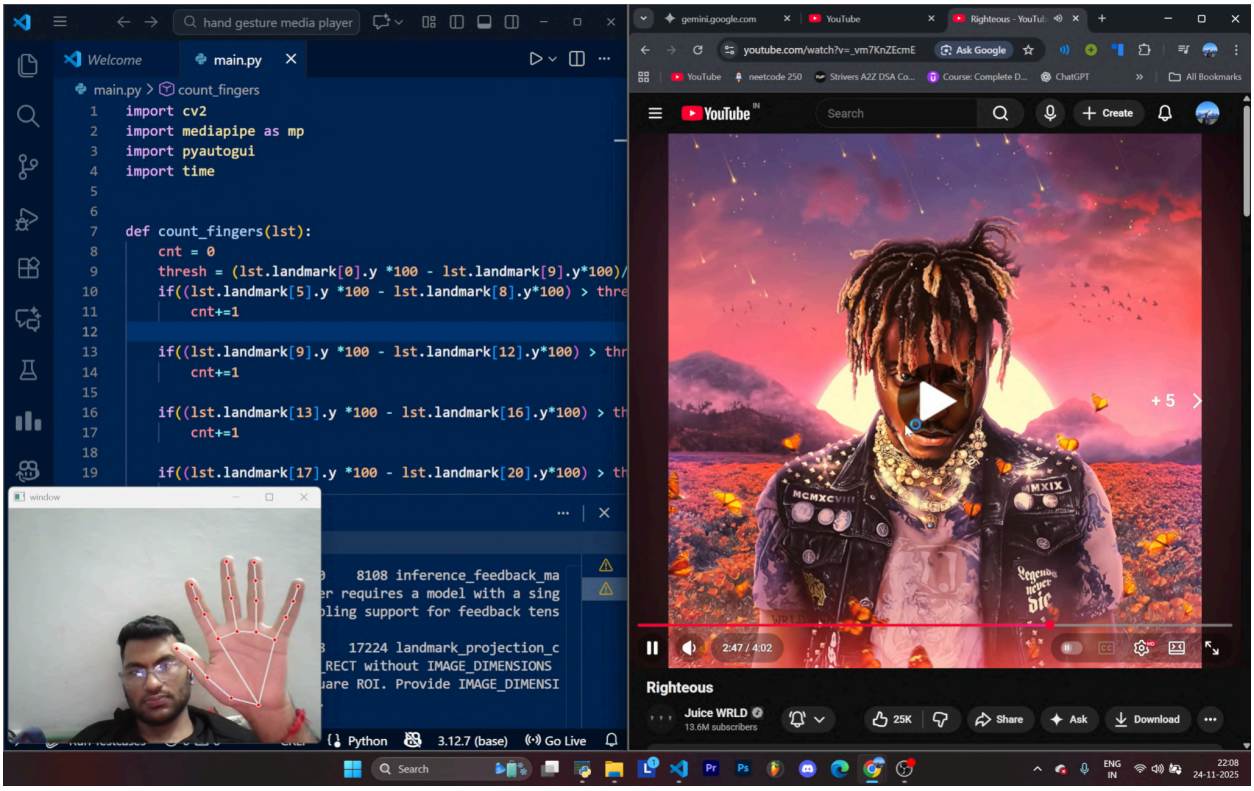
Finger Count	Key Mapped	Application
1	right	Navigation
2	left	Navigation
3	up	Navigation
4	down	Navigation
5	space	Select/Action

Challenges and Solutions

ID	Challenge	Cause/Description	Solution
1	Gesture Flicker/Inaccuracy	Quick transitions caused unstable gesture recognition	Added a debouncing mechanism using time library (action triggers only if held > 0.2s)
2	Repetitive Key Presses	Loop spams input if finger count unchanged	Implemented state management with previous count to trigger only on change

Results and Impact

The Hand Gesture Media Player provides a functional, hands-free hassle-less input method for basic computer control. By integrating the logic by which the accuracy in the project was increasing, the system addresses common challenges in gesture recognition, ensuring stable and predictable performance



Demo Video :- https://youtu.be/LLp23ujYa40