IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Hand Gesture Recognition Using Python

Prof Pragati Akre, Abinaya J, Viranch R, Tanmay M and Om W

T.E., Department of Artificial Intelligence and Data Science
SIES Graduate School of Technology, Nerul, Navi Mumbai.

Abstract: This project presents a hand gesture recognition system that uses Python, Mediapipe, Pytesseract, PyAutoGUI, and OpenCV to enable hands-free interaction with digital devices. It has features like mouse control, PowerPoint control, media player, and whiteboard functionalities, the system utilizes real-time hand tracking and gesture recognition for an intuitive user experience. It can also convert written handwriting to text. Additionally, the handwriting model is trained on a dataset of labeled images and uses transfer learning to improve its accuracy. This project enhances human-computer interaction by combining advanced computer vision techniques with versatile Python libraries, making it valuable for presentations, media control, and collaborative whiteboarding. This project could be used by handicapped people, could be used in school or office presentations, medical sectors where hygiene is important or anywhere where a touch free environment is required.

Keyword: Hand gesture, Computer vision, Mediapipe, OpenCV

I. INTRODUCTION

In the present day, as technology is evolving, there is a growing demand for a hands-free interface for hygiene, any disability or speedy work purposes. And gesture control technology has not only proven successful but also provided a possible solution by enabling users to connect themselves to devices via natural hand gestures. Hence, to address the issue this project aims to contribute to a touchless era. It describes the innovative technique of controlling media players, power points, whiteboards etc. with the help of hand signs. It also has numerous other features like whiteboard , handwriting to text and so on. The technology leverages computer vision in combination with depth-sensing cameras to create an immersive environment where users can effortlessly play, pause, adjust the volume, present their work, draw, explain, through simple gestures. This system uses mediapipe, python, opencv and machine learning algorithms for gesture recognition that provides immediate understanding of hand movements with the degree of accuracy and dependability. In addition to that, the interface does not limit itself to media control but allows users to interact with many different other apps and functions so now they benefit from a rich and immersive experience. This project has immense power to turn the paradigm of user interaction and to make the interface more accessible and convenient by opening new possibilities.

Literature Review

A literature survey on controlling media players and other functions through hand gestures reveals a growing interest in this innovative technology. Several research papers have delved into this area. Some of the earlier systems are studied as follows:

1.Controlling Media Player with Hand Gestures:
A paper written by Parshav Maloo (2022) looks at the assembling of a hand gesture recognition system for the media player direction and UI element gesture recognition and control. It relies on the sphere of the real communication, including both the limits and chances. This sphere can be implemented in different fields.[1]

2.Hand Gesture Recognition:
In the study published recently by Swetha Kotavenuka, Harshitha Kodakandla, Nimmakayala Sai Krishna, Subba Rao (IJRASET-2023), availability of hand gesture recognition devices is addressed considering machine learning approaches and computer vision for predicting gestures from array of frames, with edge detection having more to do with improved accuracy with well-tun.[2]

3.A systematic review on hand gesture recognition techniques, challenges and applications:
The research paper Lucy Fadiga,Mais Yasen, and Jusoh Shaydah (Peerj-2019) indicates a comprehensive review of hand gesture recognition issues including acquisition techniques, feature extraction, transfer models classification and environment challenges. It presents the use of the sEMG sensors, focusing on the contribution of the ANN as the classifier used mostly, the preference for hand gesture while designing the applications of sign language, the impacts of the background color on classification accuracy and the common existence of overfitting in the datasets.[3]

4. Media Controlling Using Hand Gestures:
The article published by R H Sapat, Shruti Tibhe, Ashwini Joshi, Aishwarya Warulkar, Aishwarya Sonawane, Miss T U Ahirrao (available on IJRCT.org – 2023) is about the development of a hand gestures recognizing system for controlling media player through computer vision technique, which offers beneficial advantages for people.[4]

5. Hand Gesture Recognition Using Automatic Feature Extraction and Deep Learning Algorithms with Memory:
The research paper published by Robail Yasrab, Md Mostafa, Kamal Sarker, Rubén E Nogales, Marco E Benalcázar (MDPI-2023) explores hand gesture recognition using manual and automatic feature extraction techniques with a dataset collected from 56 volunteers using the Leap Motion Controller, evaluating different classifiers and proposing a hybrid model for improved recognition accuracy.These research works collectively demonstrate the advancements and potential of using hand gestures to control media players, highlighting the benefits of intuitive and hands-free interactions with technology.[5]

II. PROPOSED SYSTEM

We have developed a hand recognition system that can control media, power points, whiteboard, and normal mouse movements as well. The system uses Python, MediaPipe, Libraries like Pytesseract, PyAutoGUI and OpenCV etc. Additionally, the system includes a feature that allows to convert drawn text into typed text. The system also includes a frontend by which user can select the desired action.

Dataset:
To train our handwriting model, we used a dataset of images labeled with the correct class of handwriting. The dataset contains images of several handwriting. We used approximately 3000 images for training and 750 images for validation.

Model Architecture:
We developed a Hand Recognition System using Python, integrating libraries such as Mediapipe, Pytesseract, PyAutoGUI, and OpenCV. The architecture includes:

1. Hand Tracking Module: Utilizes the Mediapipe library for real-time hand tracking and landmark detection.Employs OpenCV for image processing to enhance hand gesture recognition accuracy.
2. Gesture Recognition Module:Incorporates machine learning techniques for hand gesture classification.Custom-trained model using a labeled dataset of hand gesture images.
3. Text Extraction Module:Integrates Pytesseract for optical character recognition (OCR) to convert handwritten text to machine-readable text.
4. Control Module:Uses PyAutoGUI for mouse control and specific algorithms for PowerPoint, media player, and whiteboard functionalities.
5. Transfer Learning for Handwriting Recognition:Trains a handwriting recognition model using transfer learning on a labeled dataset, improving accuracy.Adapts pre-trained models and fine-tunes them on the specific handwriting dataset.
6. Application Scenarios:Designed for scenarios such as presentations, medical sectors, and environments requiring touch-free interaction.Enhances human-computer interaction through a combination of computer vision and versatile Python libraries.

Software used:- HTML,CSS for frontend .VS studio is used to make webpage. Flask is used for the integration of frontend and python scripts. Machine Learning for handwriting recognition

III. SYSTEM IMPLEMENTATION

This integrated system incorporates multiple features for diverse interactive applications, each utilizing a combination of Python libraries to enhance user interaction. Firstly, the system employs the cvzone library, along with OpenCV functionalities, to facilitate hands-free control over presentations. Hand gestures, such as swiping left or right and lifting the index finger to annotate, are detected and translated into commands, providing an intuitive method for managing slides. Secondly, the system leverages the cv2, numpy, mediapipe, and deque libraries to create an interactive drawing application. Hand gestures are interpreted to draw lines in various colors on a virtual canvas, offering users a dynamic and engaging drawing experience. It also has a model trained for detecting handwriting and converting into text. The third feature utilizes OpenCV, Mediapipe, and PyAutoGUI for gesture-based mouse control. Hand landmarks are detected, and specific gestures trigger predefined keyboard commands, providing hands-free computer control. The count_fingers function calculates the raised fingers, ensuring accurate gesture recognition. Finally, the fourth feature integrates OpenCV, Mediapipe, and PyAutoGUI to create a gesture-controlled media interface, enabling users to navigate videos and applications intuitively.It can also be used for controlling volume and brightness as well. Currently we have 4 features:
Feature 1(PPT Control): It utilizes the cvzone library, OpenCV, and Mediapipe for hands-free control of presentations. It employs hand tracking and gesture recognition to interpret specific movements, enabling functionalities such as swiping to navigate slides and lifting the index finger for annotations.
Feature 2(White board): It leverages cv2, numpy, and deque to create an interactive drawing application. The script captures webcam frames, detects hand landmarks using Mediapipe, and interprets gestures to draw lines in various colors on a virtual canvas. The dynamic color-changing and clearing functionalities enhance the drawing experience.
Feature 3(Mouse Control): It utilizes OpenCV, Mediapipe, and PyAutoGUI to enable gesture-based mouse control. It captures webcam frames, detects hand landmarks, and interprets gestures to trigger keyboard events, offering hands-free computer control. The introduction of a time-based mechanism ensures accurate and deliberate gesture recognition.
Feature 4(Media Control): It utilizes OpenCV and PyAutoGUI to create a gesture-controlled media interface. It captures video frames from the camera, detects hand landmarks using the Mediapipe library, and interprets specific hand gestures to trigger keyboard events. The system then processes the hand gestures, counting the number of raised fingers, and simulates corresponding keyboard inputs for play/pause, volume up/down, and forward commands.

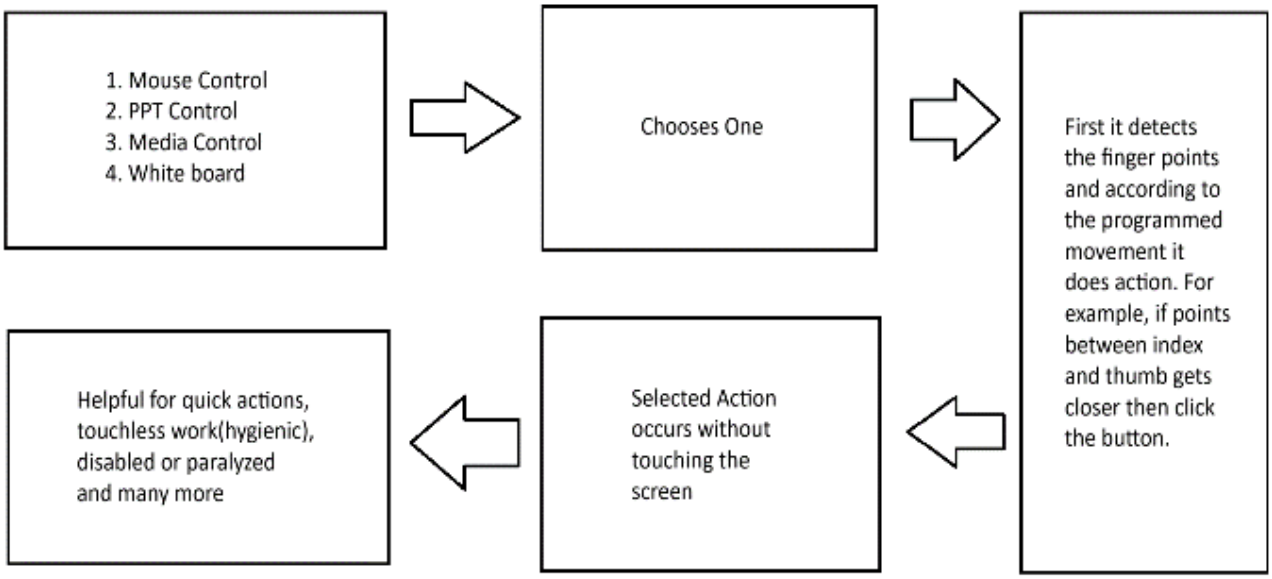

Fig 1: Block diagram

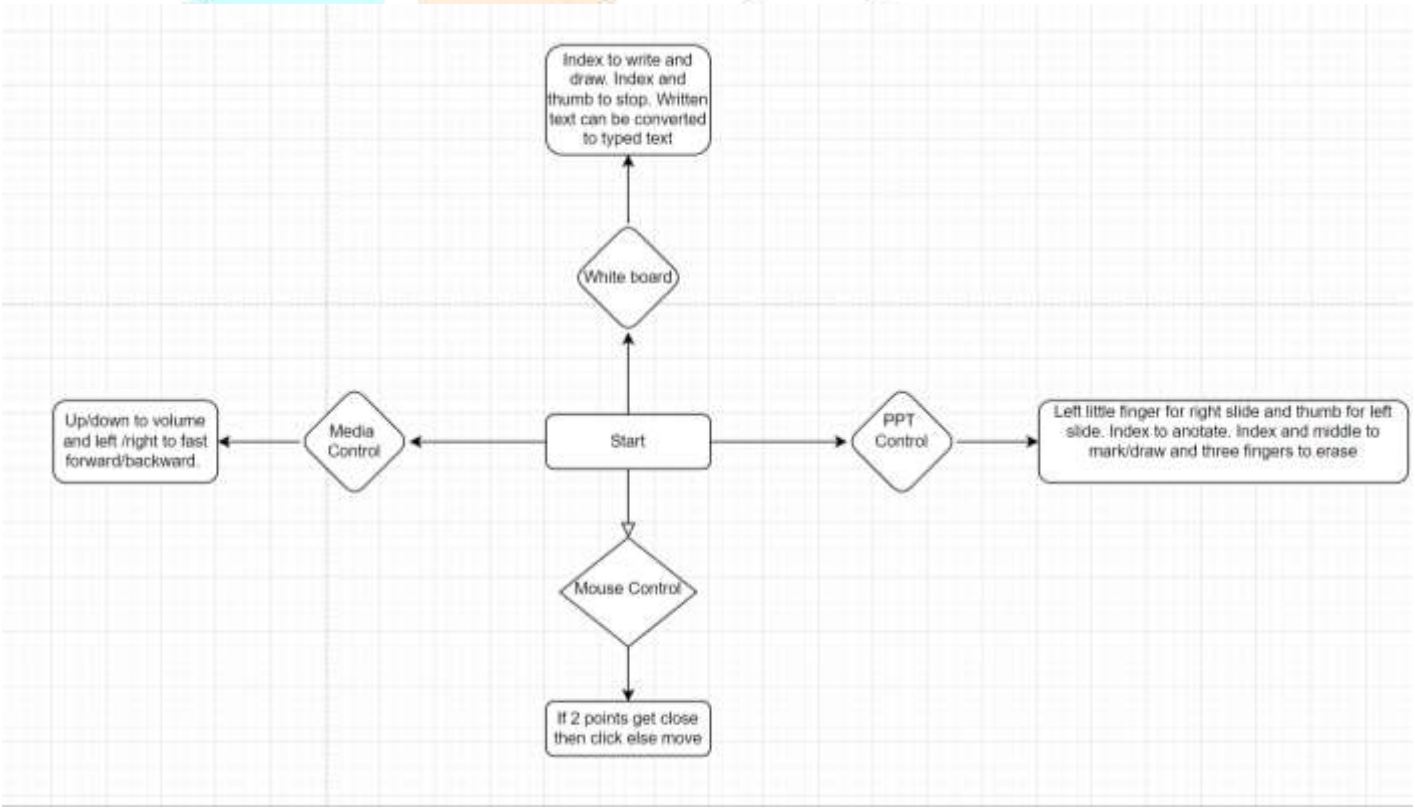

Fig 2: Architecture

IV. ANALYSIS AND RESULTS

Proposed system is implemented with the various technologies, the screen shots exhibit the layouts of its components.

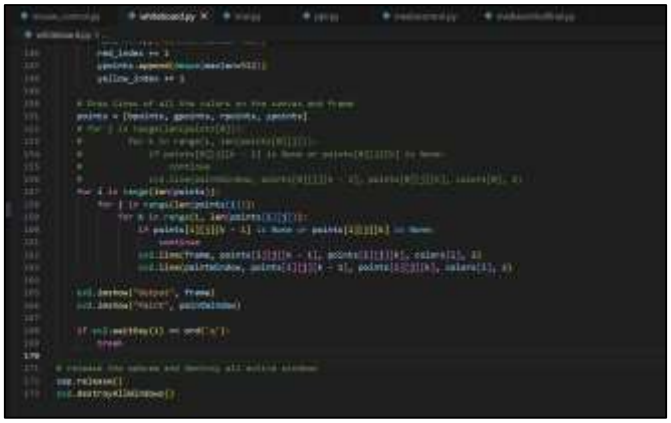

Fig 3: Snapshots of code in Python

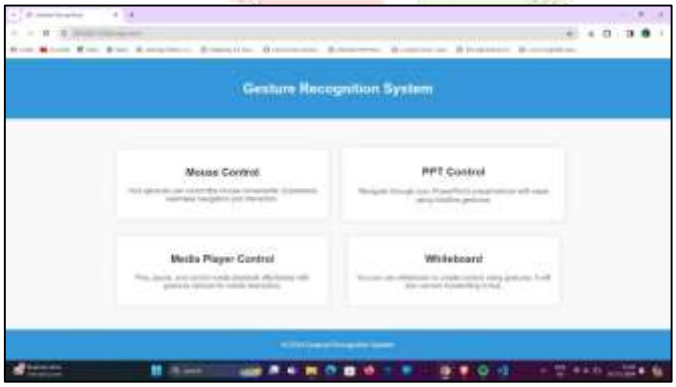

Fig 4: Software

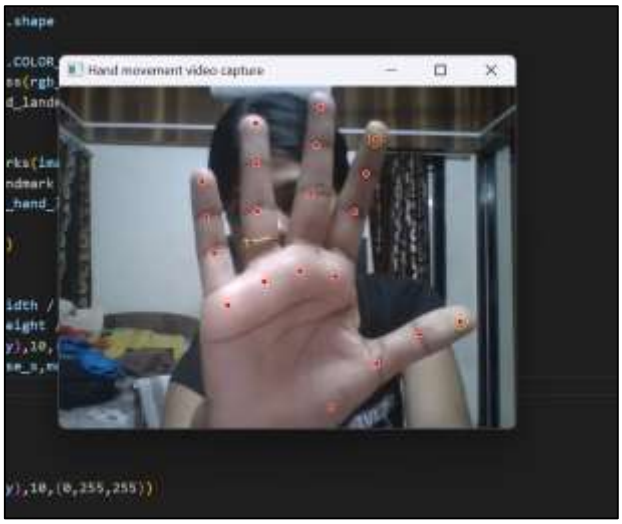

Fig 5: Mouse Control

Fig 6: PPT Control- Going to left Slide

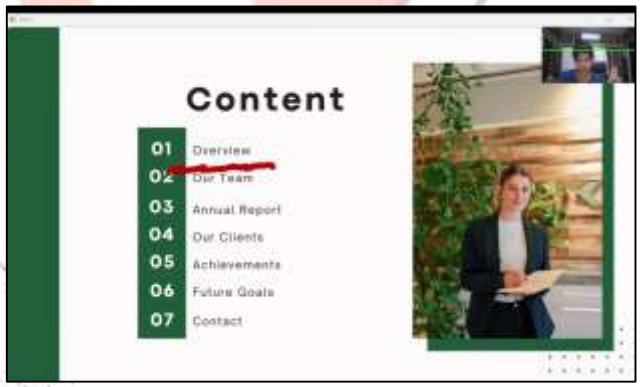

Fig 7:PPT Control- Annotating and Marking

Fig 8: White board

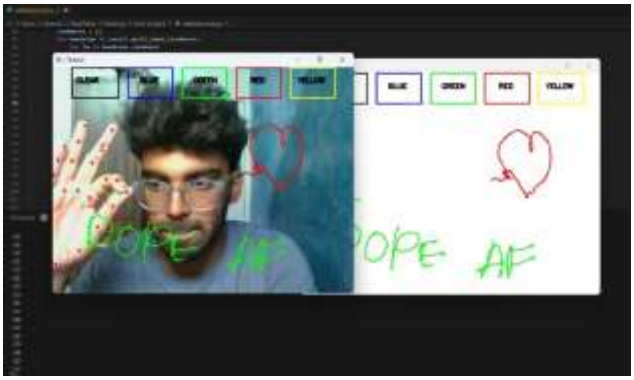

Fig 9: White board (Detects Text)

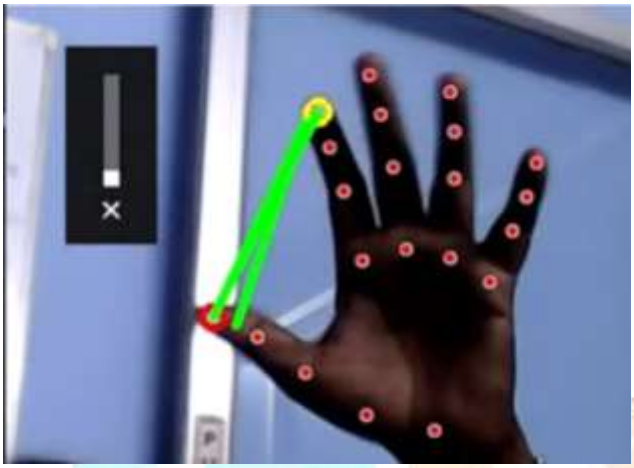

Fig 10: Media Player (Volume Control)

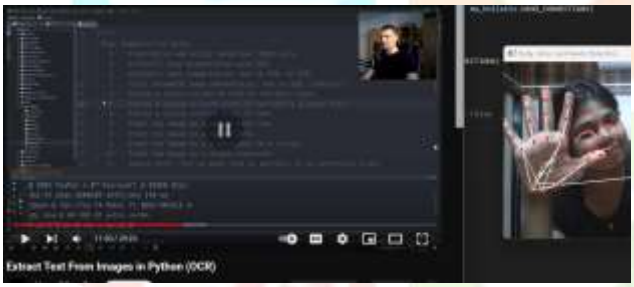

Fig 11: Media Player (Pause the video)

V.CONCLUSION AND FUTURE WORK

In conclusion, our project on Hand Gesture Recognition has successfully demonstrated its practical applications in enhancing user interaction with various systems. The implementation of this technology as a Media Controller, ppt Controller, and Mouse Controller using the Hand Gesture System, along with real-time detection through Python and OpenCV, has proven its effectiveness in making system interaction more seamless.

Looking ahead, the potential applications of this technology extend beyond the realms of conventional user interfaces. The integration of hand gesture recognition could play a pivotal role in diverse fields such as medical rehabilitation and virtual reality, promising advancements in how individuals interact with technology for improved healthcare and immersive experiences.

The versatility of this technology opens up possibilities for integration into Home Control systems, healthcare systems, gaming zones, automobiles, and robotics. The prospect of using hand gesture recognition to facilitate communication between users and devices holds great promise for creating intuitive and efficient interfaces across various domains.

As we move forward, there is ample room for improvement in the system. Enhancements in recognition accuracy and the integration of advanced controllers can further refine the user experience and broaden the scope of applications. Our project serves as a foundation for future developments in the field, paving the way for more sophisticated and widespread use of hand gesture recognition technology.

REFERENCES

[1] Parshav Maloo, 2022, Controlling Media Player with Hand
Gestures, INTERNATIONAL JOURNAL OF ENGINEERING
RESEARCH & TECHNOLOGY (IJERT) Volume 11, Issue 11
(November 2022),https://www.ijert.org/controlling-media-player-with-hand-gestures

[2] Kotavenuka, S., Kodakandla, H., Krishna, N. S., & Rao, D. S. P. V. S. (2023). Hand Gesture Recognition. International Journal for Research in Applied Science and Engineering Technology, 11(1), 331–335. https://doi.org/10.22214/ijraset.2023.48557

[3] Yasen, M., & Jusoh, S. (2019). A systematic review on hand gesture recognition techniques, challenges and applications. PeerJ Computer Science, 5, Article e218. https://doi.org/10.7717/peerj-cs.218

[4] Shruti Tivhe ,Media Controlling Using Hand Gestures (2023), https://ijcrt.org/viewfull.php?&p_id=IJCRT2312800

[5] Nogales, R. E., & Benalcázar, M. E. (2023). Hand Gesture Recognition Using Automatic Feature Extraction and Deep Learning Algorithms with Memory. Big Data and Cognitive Computing, 7(2), 102. https://doi.org/10.3390/bdcc7020102