Virtual Air Canvas Application using OpenCV and Numpy in Python

[1]Asst Prof.Jahnavi S [2]K Sai Sumanth Reddy [2]Abhishek R
[2]Abhinandan Heggde [2]Lakshmi Prashanth Reddy
[1] Asst. Professor,Department of Computer Science and Engineering
[2]BE students,Department of Computer Science and Engineering
[1][2]Dayananda Sagar Academy of Technology and Management,Bangalore,Karnataka,INDIA

Abstract- Writing is an integrated form of communication that can convey our thoughts. Typing and writing are standard ways to record information today. Letters or words are written in a relaxed space by marker or finger. These wearable devices can see and understand our actions.A computing process that attempts to recognize and interpret human gestures through the use of mathematical algorithms is known as gesture recognition.To track finger movement, the application will employ computer vision. Other uses for the created text include sending emails and texts, among others. It will be a helpful means of communication for the deaf.

Keywords-Gesture recognition , Air writing, wearable devices.

1.INTRODUCTION

Writing has evolved majorly over the years. Writing was first invented in 2000 BC by neolithic people. They first started writing on walls then it was replaced by stones. Stones were replaced by cloth and presently we use paper for communication. With the help of QWERTY keyboards we are moving towards a more digitalized form of writing. These electronic devices are slowly taking the place of traditional forms of writing with pen and paper.

The need to develop human machine interactions is rapidly growing with the surge in the usage of augmented and virtual reality. Applications using hand gestures have gained popularity over the coming years.Automotive interfaces (Ohn-Bar and Trivedi, 2014),Economical Air Writing system(Pavithra and Prabhu,2016) and Handwriting recognition in Free Space(Shashidhar, Kim and Chai, 2015) have developed systems for hand gestures recognition. However hand gestures recognition is not enough for writing in air. It also involves fingertip detection, tracking and tracing of it. Fingertip detection or LED light detection system is developed by [pavithra], where they first detect LED light and then capture the movement and by using Optical Character Recognition(OCR) and display the alphabet on the screen. However these methods which have the usage of devices have some limitations.

The proposed system eliminates the usage of cell phones for taking notes. Fingertip detection and finger movement techniques are used to develop the system.Using Python, OpenCV and CNN techniques fingertip is first detected and then the trajectory of fingertip is traced and displayed on the screen.

Python programming is used for developing the virtual mouse system, and also, the library for computer vision that is OpenCV is used in the AI virtual mouse system. In the proposed system, the model uses the MediaPipe package for the tracking of the hands and for tracking the tip of the fingers, also, Pynput, Autopy, and PyAutoGUI packages are used for moving around the window screen of the computer for performing various functions such as left click, right click, and scrolling functions. The results of the proposed AI virtual mouse model showed very high accuracy level, and the model can work very well in real-world application with the use of a CPU without the use of a GPU.

2. PROBLEM STATEMENT

The existing system only works with your fingers and no highlighters, paints, or relatives.Identify and distinguish something like a finger from RGB image without depth sensor great challenge. Another problem is lack of top and

movement under the pen. The system uses a one RGB camera that you can overwrite. From the depths discovery is impossible, jobs up and down of the pen cannot be traced. So, everything finger path is drawn, and the result the image will be abstract and unseen by model. Using real-time hand touch to change position the process from one region to another requires a lot of code care. In addition, the user should know many movement to control his plan adequately. The project focuses on solving some of the most important social issues Problems. First of all, there are many hearing-impaired people problems in everyday life. While listening again listening is taken for granted, people don't have this communicating with a disability using sign language. Most countries in the world cannot understand yours feelings and emotions outside the middle translator. Second, overuse Smartphones: causes accidents, stress, distractions, and other illnesses that people can no longer tolerate find out. Although its portability and ease of use exist.very popular, its obstacles include life terrifying events. Waste paper is not uncommon. Many papers are wasted on writing,writing, drawing, etc. A4 paper production requires about 5 liters of water. 93% sources come from trees, 50% of commercial waste is paper, 25% of landfills are paper, and the list goes to. Waste paper harms the environment through use of water and trees and produce tons waste. On-air writing can solve these problems quickly. It will serve as a communication tool for the deaf.Your online text can be displayed in AR.

The existing system only works with your fingers and no highlighters, paints, or relatives.Identify and distinguish something like a finger from RGB image without depth sensor great challenge. Another problem is lack of top and movement under the pen. The system uses a one RGB camera that you can overwrite. From the depths discovery is impossible, jobs up and down of the pen cannot be traced. So, everything finger path is drawn, and the result the image will be abstract and unseen by model. Using real-time hand touch to change position the process from one region to another requires a lot of code care. In addition, the user should know many movement to control his plan adequately. The project focuses on solving some of the most important social issues Problems. First of all, there are many hearing-impaired people problems in everyday life. While listening again listening is taken for granted, people don't have this communicating with a disability using sign language. Most countries in the world cannot understand yours feelings and emotions outside the middle translator. Second, overuse Smartphones: causes accidents, stress, distractions, and other illnesses that people can no longer tolerate find out. Although its portability and ease of use exist.very popular, its obstacles include life terrifying events. Waste paper is not uncommon. Many papers are wasted on writing,writing, drawing, etc. A4 paper production requires about 5 liters of water. 93% sources come from trees, 50% of commercial waste is paper, 25% of landfills are paper, and the list goes to. Waste paper harms the environment through use of water and trees and produce tons waste. On-air writing can solve these problems quickly. It will serve as a communication tool for the deaf.Your online text can be displayed in AR

3.LITERATURE SURVEY

Real-time hand gesture recognition employs a variety of techniques. A system created by Shomi Khan, M. Elieas Ali[3], and Sree Sourav Das employs a skin colour identification algorithm to translate American Sign Language (ASL) from real-time video into text. It could be difficult to identify the hand because skin tone and hand form vary from person to person. The technology uses two neural networks to overcome this. The SCD (Scalable color descriptor) neural network is the first algorithm. The picture pixels are fed into the SCD neural network, which determines whether or not they are skin pixels. The second is HGR (Hand gesture recognition) neural network in which the extracted features will be added. The features are to be extracted by two distinct algorithms namely Finding the fingertip and Pixel segmentation algorithm.

[5] To accomplish mouse actions such as moving the cursor, clicking left and clicking right with hand gestures, an impalpable interface is conceived and implemented utilizing computer-vision-based real-time dynamic hand gestures. MATLAB is used for the implementation of the system. S. Belgamwar and S. Agrawal [6] have collaboratively developed a new HCI technique that inculcates a camera, an accelerometer, a pair of Arduino microcontrollers and an Ultrasonic Distance Sensors. The main concept behind this interface is to capture motions using Ultrasonic Distance Sensors. To record the gestures, the distance between the hand and the distance sensor is determined.

By simply waving their finger over an LED light source, the user can create the alphabet or type anything they wish, according to a new technology put forth by Pavitra Ramasamy and Prabhu G [4].To extract the movement of the finger sketching the alphabet, only the colour of the LED is tracked. The background is black, and the tracked

object's colour is converted to white. The user wanted to draw an image of the alphabet in black and white, so they stitched together several black and white frames to make it.

For 3D hand gesture detection, Quentin De Smedt, Hazem Wannous, and Jean-Philippe Vandeborre [8] used a skeletonbased model. They used the geometric shape of hand to gsin an effective descriptor from the Intel Real-Sense depth camera's hand skeleton linked joints. The skeleton-based approach is far more better than the depth-based approach. In [9] Prajakta Vidhate, Revati Khadse and Saina Rasal they have developed a virtual paint application that uses ball-tracking technology to track the hand movement and write on the screen. They have used a glove with a ping-pong ball attached to it as a contour.[10] Ruimin Lyu, Yuefeng Ze, Wei Chen, and Fei Chen demonstrated an airbrush model that is easily adaptable and that employs the Leap Motion Controller, which tracks hands, to produce an immersive freehand painting experience.

4.ALGORITHM USED FOR HAND TRACKING

The MediaPipe framework handles hand gesture identification and tracking, and the OpenCV library handles computer vision. The application uses machine learning principles to track and identify hand movements and hand tips.

4.1 MediaPipe

The user experience can be significantly enhanced across a range of technological domains and platforms by being able to recognise the shape and motion of hands. For instance, it can provide as the foundation for hand gesture control and sign language comprehension. It can also make it possible for digital information and material to be superimposed on top of the real world in augmented reality. Although it comes effortlessly to individuals, robust real-time hand perception is an extremely difficult computer vision problem due to the fact that hands frequently occlude themselves or each other and lack high contrast patterns. A high-fidelity hand and finger tracking solution is MediaPipe Hands. It uses machine learning (ML) to deduce 21 3D hand landmarks from a single image.Our solution delivers real-time performance on a cell phone, and even scales to several hands, unlike existing state-of-the-art systems, which mostly rely on powerful desktop environments for inference. We anticipate that making this hand perception functionality available to a larger research and development audience will lead to the creation of innovative use cases, igniting new research directions.

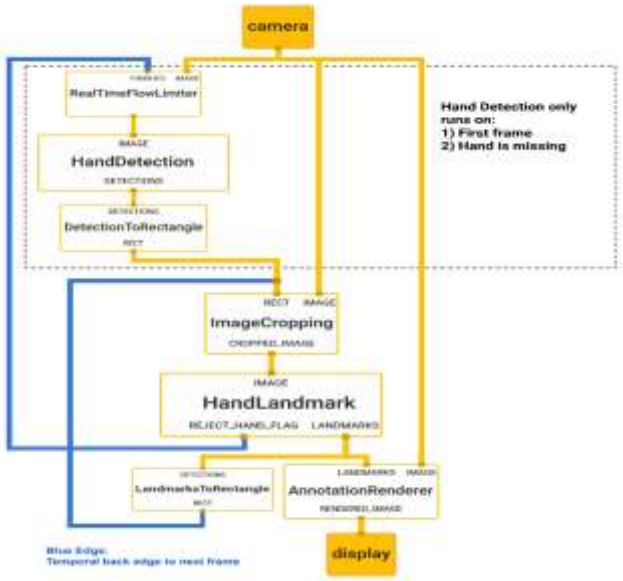
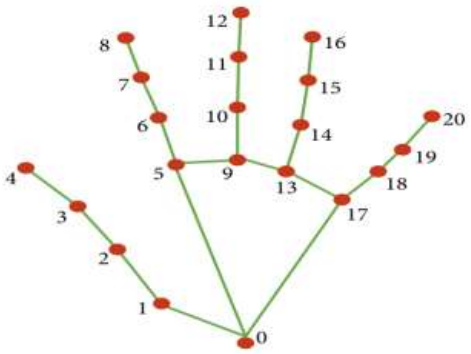


Fig -1:Hand detection model

4.2Hand Landmark Model

After detecting the palm over the whole image, our subsequent hand landmark model uses regression, or direct coordinate prediction, to accomplish precise keypoint localization of 21 3D hand-knuckle coordinates inside the detected hand regions. The model acquires a reliable internal hand posture representation and is unaffected by self-occlusions or partially visible hands. We manually added 21 3D coordinates to around 30K real-world photos to obtain ground truth data, as shown below (we take Z-value from image depth map, if it exists per corresponding coordinate). We additionally render a high-quality synthetic hand model over a variety of backgrounds and map it to the associated 3D coordinates in order to better cover the range of possible hand poses and provide additional supervision on the nature of hand geometry.



0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP

Fig-2:landmarks in the hand

4.3 OpenCV

Object detection image processing methods are included in the OpenCV computer vision library [14]. Real-time computer vision applications can be created by utilising the OpenCV library for the Python programming language. The processing of images and videos as well as analytical techniques like face and object detection use the OpenCV library.

4.4 NumPy

The N-dimensional array type known as ndarray is the most significant object defined in NumPy. The collection of identically categorised things is described. A zero-based index can be used to access items in the collection. A ndarray's items all take up the same amount of space as a memory block. Every item in ndarray is a data-type object object (called dtype). A Python object of one of the array scalar types represents each item that is retrieved from an ndarray object (via slicing). The relationship between ndarray, data-type object (dtype), and array scalar type is depicted in the picture below.

5. METHODOLOGY

Based on the web camera frames that were captured, a virtual paint programme was offered. The web camera sends the system the frames that it has received. Until the application is finished, the approach uses a web camera to collect each frame.

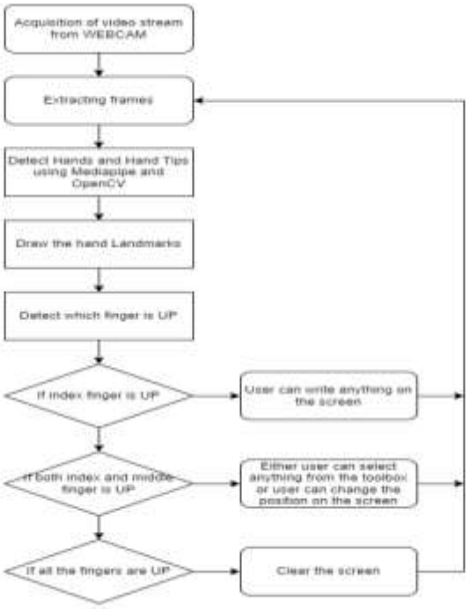


Fig-3:flow chart of air canvas application

5.1The Camera Used in the AIR CANVAS

The frames that have been recorded by a laptop or PC's webcam serve as the foundation for the proposed AIR CANVAS. As seen in Figure 4, the web camera will begin recording video after the video capture object is created using the Python computer vision package OpenCV. The virtual AI system receives frames from the web camera and processes them.

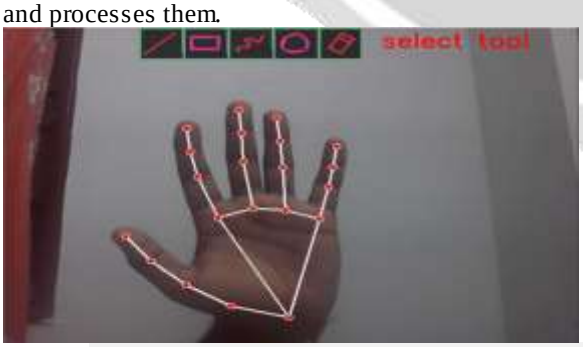


Fig-4: Capturing video using the webcam (computer vision).

5.2 Capturing the Video and Processing

The webcam is used by the AIR CANVAS system, and every frame is recorded up until the end of the application. As illustrated in the accompanying code, the video frames are converted from BGR to RGB in order to find the hands in the video frame by frame.

```
def findHands(self, img, draw = True):

imgRGB = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)

self.results = self.hands.process(imgRGB)
```

5.3 Detecting Which Finger Is Up and Performing the Particular Function

Using the tip Id of the specific finger that we located using the MediaPipe and the corresponding co-ordinates of the fingers that are up, as shown in Figure 6, we are able to determine which finger is up at this point. Then, in accordance with that determination, the specific mouse function is carried out.

Fig-5: selecting an option

5.4 For the Mouse to Perform ACTION

The pynput Python module is used to programme the computer to execute the right mouse button click if the middle finger with tip Id = 2 and the index finger with tip Id = 1 are both up and the distance between the two fingers is less than 40 px, as illustrated in Figure 10.

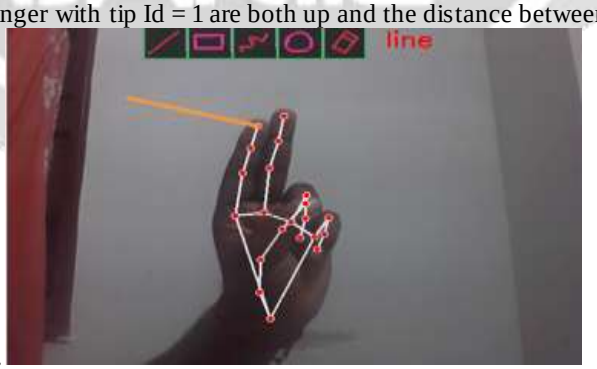


Fig-6:performing an action

6.PROJECT SCOPE

The scope of this systemis mainly used as a

powerful means of communication for the deaf,
which means implementing this project can help in:
1. An effective communication method that
reduces mobile and laptop usage by eliminating the need to write.
2. It helps people with hearing impairments to
communicate well.
3. Various purposes, such as sending messages, e-mails, etc., as the generated text can also be used for that.

7. CONCLUSION

This program has the potential to challenge traditional writing methods. Eliminates the need to carry a cell phone in hand to take notes, to give an easy way on the go to do the same. It will again work towards a greater purpose in helping especially those who know them to communicate easily. Even adults who find it difficult to use the keyboard can easily use the program. Expanding functionality, this program can also be used to control IoT devices soon. Air painting can also be made happen. This program will be very good smart clothing software using which people can work better with the digital world. The unpopular reality of taxpayers we see can make the text come alive. Wind-writing programs should listen only to their master's control touch and should not be misled by people all around. Such discovery algorithms are as follows as YOLO v3 can improve fingerprint recognition accuracy and speed. In the future, progress on Artificial Intelligence will improve the efficiency of writing in the air.

8.REFERENCES

[1] Y. Huang, X. Liu, X. Zhang, and L. Jin, "A Pointing Gesture Based Egocentric Interaction
System: Dataset, Approach, and Application," 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Las Vegas, NV, pp. 370-377, 2016.
[2] P. Ramasamy, G. Prabhu, and R. Srinivasan, "An economical air writing system is converting finger movements to text using a web camera," 2016 International Conference on Recent Trends in Information Technology (ICRTIT), Chennai, pp. 1-6, 2016.
[3] S. Khan, M. E. Ali, S. Das and M. M. Rahman, "Real Time Hand Gesture Recognition by Skin Color Detection for American Sign Language," 2019 4th International Conference on Electrical Information and Communication Technology (EICT), 2019, pp. 1-6, doi: 10.1109/EICT48899.2019.9068809.

[4] P. Ramasamy, G. Prabhu and R. Srinivasan, "An economical air writing system converting finger movements to text using web camera," 2016 International Conference on Recent Trends in Information Technology (ICRTIT), 2016, pp. 1-6, doi: 10.1109/ICRTIT.2016.7569563.

[5] P. Ramasamy, G. Prabhu and R. Srinivasan, "An economical air writing system converting finger movements to text using web camera," 2016 International Conference on Recent Trends in Information Technology (ICRTIT), 2016, pp. 1-6, doi: 10.1109/ICRTIT.2016.7569563

[6] Yuan-Hsiang Chang, Chen-Ming Chang, "Automatic Hand-Pose Trajectory Tracking System Using Video Sequences", INTECH, pp. 132- 152, Croatia, 2010
[7]EshedOhn-Bar, Mohan ManubhaiTrivedi, "Hand Gesture Recognition In Real-Time For Automotive Interfaces," IEEE Transactions on Intelligent Transportation Systems, VOL. 15, NO. 6, December 2014, pp 2368-2377

[8] Q. De Smedt, H. Wannous and J. Vandeborre, "SkeletonBased Dynamic Hand Gesture Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2016, pp. 1206-1214, doi: 10.1109/CVPRW.2016.153.

[9] Prajakta Vidhate, Revati Khadse, Saina Rasal, “Virtual paint application by hand gesture recognition system”, 2019 International Journal of Technical Research and Applications, 2019, pp. 36-39

[10] R. Lyu et al., "A flexible finger-mounted airbrush model for immersive freehand painting," 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS), 2017, pp. 395-400, doi:10.1109/ICIS.2017.7960025.