

AI - POWERED AIR CANVAS USING MEDIAPIPE HAND TRACKING FOR INTERACTION USING MACHINE LEARNING TECHNIQUES

Mrs. B. Koti Ratnam^{#1}, *Assistant Professor, Department of Computer Science and Engineering, Tirumala Engineering College, Jonnalagadda, Andra Pradesh, India - 522601.*

J. Sai Akshaya^{#2}, K. Gnaneswar^{#3}, K. Mahesh Reddy^{#4}, K. Surekha^{#5}

Abstract— The Air Canvas project is an innovative application that leverages hand gesture recognition and computer vision techniques to create an intuitive virtual drawing platform. This system enables users to draw, write, and interact with a digital canvas using hand movements detected by a webcam, eliminating the need for physical input devices. Built with Python, OpenCV, and MediaPipe, the application detects hand gestures in real-time and translates them into dynamic interactions such as freehand drawing, shape creation, and erasing. The application incorporates a comprehensive user interface, including a navigation bar for selecting tools and colors, and supports advanced functionalities such as fingertip detection, gesture-based tool switching, and shape resizing. The integration of these features provides an immersive and accessible experience, making the system suitable for creative expression, interactive teaching, and design applications. The Air Canvas demonstrates the potential of gesture-based human-computer interaction, offering an engaging, paperless, and versatile solution for digital content creation.

Keywords- Air Canvas, Object detection, Object tracking, mediapipe.

I. INTRODUCTION

With the rapid advancement of technology, the way humans interact with machines has undergone a paradigm shift.

The advent of gesture-based systems has opened up new avenues for intuitive and natural user interactions. Among these innovations, the concept of "Air Canvas" stands out as a revolutionary approach that allows users to draw, write, and interact with digital systems without the need for physical input devices such as pens, styluses, or touch screens. The fundamental idea revolves around using hand gestures to manipulate a virtual canvas, leveraging advancements in computer vision.

Traditional drawing methods often require physical surfaces and tools, such as paper, pens, or tablets. While these methods are effective, they pose limitations in terms of mobility, sustainability, and versatility. For instance, the reliance on physical tools contributes to environmental concerns such as paper wastage and resource consumption. In contrast, digital solutions often involve cumbersome hardware, complex setup processes, or steep learning curves, making them less accessible to a wider audience. The Air Canvas system addresses these challenges by introducing a simple yet powerful interface that requires only a webcam and a computer.

The primary objective of this project is to create an intuitive and versatile virtual canvas application that enables users to:

1. Draw and write freely in the air using natural hand gestures.

2. Select from a variety of tools and colors for enhanced creative expression.

3. Create predefined shapes such as rectangles, circles, and triangles with ease.

3. Erase or modify existing drawings without physical interaction with the screen.

II. LITERATURE SURVEY

1. S.V. Ashwin Kumar et al. (2021), *An Implementation of Virtual Whiteboard Using OpenCV for Virtual Classes*. International Journal of Research in Engineering and Technology, 10th March 2021 .

He proposed a virtual whiteboard system utilizing OpenCV for object detection and tracking. The study highlighted the feasibility of creating a digital writing platform where users can write in the air using a webcam. By leveraging Python for programming and OpenCV for real-time video processing, the authors demonstrated how hand movements could be tracked and displayed as written strokes on a screen. This foundational work underscored the potential of using computer vision techniques for virtual writing applications.

2. Akash Kumar Chaudhary et al. (2021), *Air Canvas Application Using OpenCV and NumPy in Python*. International Journal of Research in Engineering and Technology, Volume 8.

He introduced an air canvas application focusing on motion-to-art transformation. This study explored the use of fingertip tracking as an alternative to marker-based systems, enabling users to interact with the virtual canvas without any physical tools.

The authors utilized OpenCV for video processing and NumPy for computational tasks, showcasing a significant step towards markerless interactions.

3. Pranavi Srungavarapu et al. (2021), *Virtual Sketch Using OpenCV*. International Journal of Innovative Technology and Exploring Engineering, Volume 10.

She presented a system for creating virtual sketches using OpenCV. This work focused on morphological operations to enhance the detection of hand movements, allowing users to draw and erase directly on the screen. By introducing dynamic shape selection and gesture recognition, the authors expanded the scope of virtual canvas applications.

4. Priyanshu Mishra et al. (2021), *Virtual Ink Using Python*. EasyChair Preprint No. 5707. Retrieved from EasyChair.

He developed a system called Virtual Ink, which utilized Python, OpenCV, and NumPy to create a virtual canvas for writing in the air. The authors emphasized the use of a webcam to detect the movement of a marker, translating it into real-time strokes on a virtual canvas. The integration of NumPy allowed for efficient handling of multidimensional arrays during the processing of visual data.

5. Harneet Kaur et al. (2021), *A Comprehensive Overview of AR/VR by Writing in Air*. International Journal of Scientific Research in Computer Science, Engineering, and Information Technology, Volume 10. He provided a detailed analysis of the applications of augmented reality (AR) and virtual reality (VR) in air writing systems. The authors explored the use of AR/VR to visualize and manipulate strokes

in a 3D space. This study demonstrated how air writing could serve as a foundation for AR/VR applications, such as interactive learning and immersive design.

6. M.S. Alam et al. (2019), *Trajectory-Based Air-Writing Character Recognition Using Convolutional Neural Network*.

Proceedings of the 4th International Conference on Control, Robotics, and Cybernetics

He explored the use of deep learning, specifically convolutional neural networks (CNNs), for air writing character recognition. The authors emphasized the importance of trajectory analysis in accurately identifying handwritten characters in the air. This study paved the way for incorporating machine learning techniques into air writing systems.

III. EXISTING SYSTEM

Traditional Art refers to the fine art which is made before the digital Art. The Traditional way incorporates pen and paper, chalk also, board for writing.

Once the digital artwork got picked up, the usage of computer vision techniques had become popular.

In the existing system, OpenCV plays a main role in object recognition i.e., a cap. The user can select any color and draw or write.

But it become complex when they try to draw any shape and even the object tracking is so inaccurate to detect by the webcam.

Complex to detect the object movement: When a user is interacting with an air canvas, it is often difficult to distinguish between the user's hands and other objects in the background or on the canvas. For example, if a user is holding an object in their hand while drawing, the system might have trouble recognizing where the hand ends and the object begins.

Complexity of Gesture Recognition: While MediaPipe supports a range of gestures, implementing advanced gesture recognition can be challenging. For example, differentiating between similar gestures (like a finger point vs. a fist) may require additional processing or fine-tuning of the system. Designing an intuitive and responsive gesture interface can be time-consuming and prone to errors.

Used for only Writing Purposes: If the system is designed only for writing, the interaction options are quite narrow. It restricts the user's ability to engage with the digital whiteboard in a more interactive way, such as moving objects, selecting tools, or navigating menus using gestures

User feel difficult to draw the shapes: Drawing precise geometric shapes (like circles, squares, or straight lines) is difficult because the hand moves freely in the air without any physical guidance. Small hand tremors or unintended movements can distort the shapes.

IV. PROPOSED SYSTEM

The project starts by initializing the video interface through webcam. The

system uses OpenCV which is open-source computer vision library to process the video. Once the video has been processed, system uses MediaPipe to detect and identify landmarks on the user's hand, the system can detect the hand landmarks.

We can draw or write on the canvas or screen when the tools, they are activated with two fingers and used to operate with one finger. The tools we have are brush, eraser, different shapes and colour customization.

We provided a variety of shapes, including rectangle, line, triangle and oval. Moreover, we can remove the artwork by simply choosing the erase tool which makes the process easy.

The user can exit the application by pressing 'Q' key in keyboard.

Cost-Effective: Requires minimal hardware, making it affordable for a wide range of users. Since MediaPipe's hand tracking only requires a camera, it's a low-cost solution compared to traditional touch-based devices or specialized input hardware (like styluses or smartboards). This can make hand tracking accessible to a wider audience or reduce the need for expensive equipment.

Sustainable and Accessible: Reduces reliance on paper and physical tools, promoting eco-friendly practices. For users with disabilities, hand tracking can offer a more accessible alternative to traditional input methods. It allows for easy navigation and drawing with gestures, without the need for a mouse, keyboard, or touch screen.

User-Friendly: Intuitive gestures and a straightforward interface ensure ease of use for individuals of all ages and technical expertise. By eliminating the need for complex menus or physical tools, the interface is simplified. Users can focus on drawing and interacting with the content in a more natural and comfortable manner, improving the overall experience.

Hands-Free Interaction: Users can control the whiteboard without needing a mouse, keyboard.

Real-Time Gesture Recognition: The system provides accurate, real-time tracking of hand movements. MediaPipe allows for fine-grained hand gesture recognition, enabling features like zooming in/out, selecting objects, or erasing parts of the canvas using specific gestures. This adds interactivity and convenience, enhancing the user experience by reducing the need for additional buttons or menu options.

V. METHODOLOGY

Development of a real-time tracker of hand movements can be a very complex task. Fortunately, there exists certain libraries of python that can enable us to accomplish the tasks with same accuracy. One such library is called OpenCV. The CV in OpenCV is short for computer vision and is a key library used for numerous computer vision projects. Components to generate the project are:

Gesture Tracking: uses algorithms to interpret human hand and body movements to execute specific features of a computer

program. Users can use this feature to perform simple actions and interact with devices or a computational device & never have to get in contact with them themselves.

Motion Tracking: is another facet that will be used extensively in this project. It is a bit like gesture tracking and holds a vital role in many animation and 3D software. It interprets real-time information in chunks with the compatible software and processes it which can be later edited.

Laplace Algorithm: It is a measure in two dimensions of the special derivation of any image. It can highlight any region which undergoes rapid intensity change and hence is an important step in edge detection.

Morphological Transforms: Morphological Transformations are operations which are simple in nature and are based on the image shape. Usually, it is performed on images which are binary in nature. It takes in two inputs, one is a element called kernel, also called as a structuring element. It is responsible for deciding what nature, the operation will take.

Color Space: Humans perceive color more in tune with the color space called HSV than CMYK & RGB color space. Consequently, HSV consists of three values i.e., H which stands for Hue, S which stands for Saturation, and V which stands for Value. Generally used to deploy colors based on the amount of gray level present in them and the brightness or sharpness of these colors, it is generally considered the best for color detection purposes.

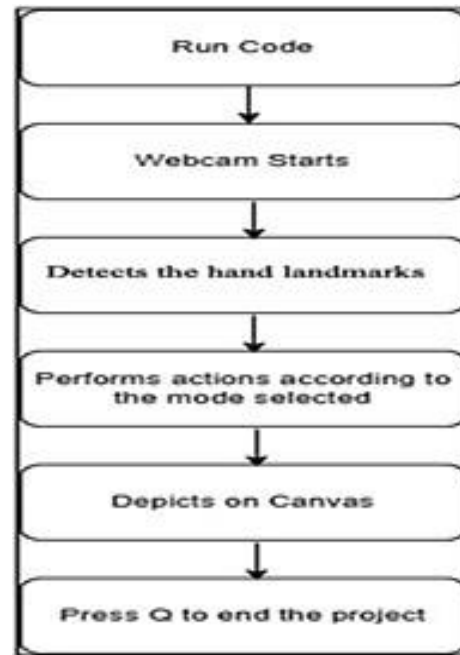


Fig. 1. Flowchart of the project.

I. ALGORITHM USED

A. Techniques of Fingertip Recognition Dataset Creation: 1) A: Video to Images: In various environments motion of a hand of a person is taken in a two-second video form. Then different 30 images were taken from these videos. Overall, 2000 pictures were collected. 99% of accuracy is obtained from the dataset we used on the trained model. For the ones where backgrounds are different in the dataset, the accuracy of model decreased because the video and environment where same for the taken 30 images [11].

B. Take Pictures in different Backgrounds: Now we created a new dataset to tackle the diversity issue in the previous method. We worked on controlling the system with use of gestures [1]. So, two different hand poses we used. The key is that the two fingers fingertip should be recognized by the model. By this with the number of fingers a user uses can be used for controlling the

system. Basically, two fingers for selecting the color of brush and one finger for drawing. 1800 images were there in the dataset for this. The obtained accuracy was 94%. In Copyright © IEEE–2023 ISBN: 979-8-3503-6996-0 various backgrounds this model worked better with respect to the previous one [12].

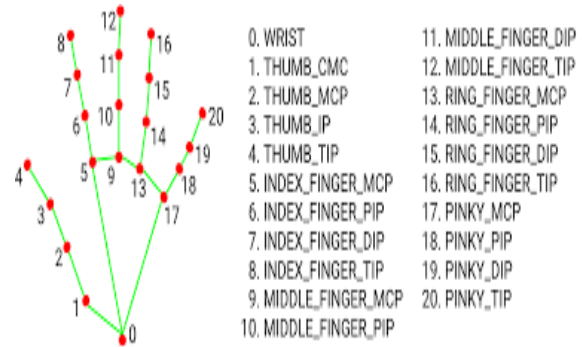


Fig. 2. Hand landmarks.

VI. RESULT ANALYSIS

A. Hand Tracking: First, our project of Air Canvas requires an efficient hand tracking in real time. So first the hand tracking has been implemented in python by using OpenCV and mediapipe packages. In various platforms and devices, inference of model, processing of media, transformations of data like problems are solved with the available calculators set in Mediapipe. Acceleration of GPU is used to enhance the rendering, cropping and computations of neural network which all are individual calculators [13].

A lot of computation is saved with Mediapipe because it optimizes the model to only run the palm detector when it is required and not unnecessarily [14]. To achieve that in every frame palm detector is not applied, the previous frame hand landmarks are taken and from the current frame we derive the location of hand. To make sure that in the input crop the hand is aligned and present an extra scaler is given by the model of hand tracker for robustness. In the next frame the model is applied again only when the confidence rate is below the given required threshold.

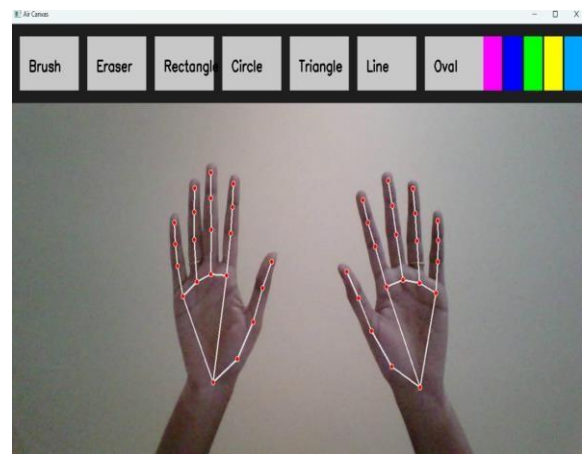


Fig. 3. Result of hand tracking in real time.

B. Air Canvas: Now using the hand tracking module, the air canvas module is implemented. The main idea is we use two fingers as a selection mode and one finger as a drawing mode (in command prompt it will be displayed also in which mode we are). When the two fingers are used, we can select the available three brush colors and the eraser while with one finger we can use the selected brush or eraser. In total we are using three canvas: first one is the main canvas where we use the webcam to display our video and have all the selection items in the header, second one is a black canvas where whatever we draw with a specific color chosen will be displayed in this canvas and third one is a white canvas where whatever we draw will be displayed but only in black color.



Fig.4. Output od Air Canvas.

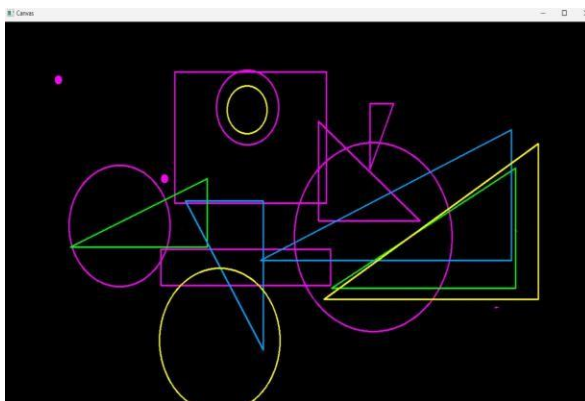


Fig.5. Output od Air Canvas.

VII. CONCLUSION

The Air Canvas project exemplifies a groundbreaking approach to human computer interaction by leveraging advanced gesture recognition and computer vision technologies. By enabling users to draw, write, and interact with a digital canvas through simple hand gestures, the system eliminates the need for traditional input devices. This not only enhances accessibility and user convenience but also promotes sustainability by reducing reliance on paper and other physical resources.

The integration of OpenCV and MediaPipe ensures real-time tracking

and responsiveness, while features such as a dynamic navigation bar, color selection, and shape creation provide users with a comprehensive creative toolkit. Applications of the Air Canvas span across education, creative design, entertainment, and accessibility, showcasing its versatility and potential impact.

While the project has addressed many challenges inherent in gesture-based systems, such as real-time processing and intuitive controls, further enhancements could focus on improving gesture recognition accuracy, expanding the toolset, and integrating multimodal input options. Overall, the Air Canvas stands as a testament to the possibilities of modern technology in creating innovative and user-friendly solutions for diverse applications.

VIII. REFERENCES

- [1] **S.V. Ashwin Kumar et al.** (2021). *An Implementation of Virtual Whiteboard Using OpenCV for Virtual Classes*. International Journal of Research in Engineering and Technology, 10th March 2021.
- [2] **Priyanshu Mishra et al.** (2021). *Virtual Ink Using Python*. EasyChair Preprint No. 5707. Retrieved from [EasyChair](#).
- [3] **Akash Kumar Chaudhary et al.** (2021). *Air Canvas Application Using OpenCV and NumPy in Python*. International Journal of Research in Engineering and Technology, Volume 8.

- [4] **Pranavi Srungavarapu et al.** (2021). *Virtual Sketch Using OpenCV*. International Journal of Innovative Technology and Exploring Engineering, Volume 10, Pages 107-108.
- [5] **Harneet Kaur et al.** (2021). *A Comprehensive Overview of AR/VR by Writing in Air*. International Journal of Scientific Research in Computer Science, Engineering, and Information Technology, Volume 10, Pages 477-482.
- [6] **M.S. Alam et al.** (2019). *Trajectory-Based Air-Writing Character Recognition Using Convolutional Neural Network*. Proceedings of the 4th International Conference on Control, Robotics, and Cybernetics, pp. 86-90. DOI: [10.1109/CRC.2019.00026](https://doi.org/10.1109/CRC.2019.00026).
- [7] **Yuan-Hsiang Chang et al.** (2019). *Automatic Hand Pose Trajectory Tracking System Using Video Sequences*. INTECH, Pages 132-152.
- [8] **J. Patel et al.** (2021). *Text Recognition by Air Drawing*. Proceedings of the Fourth International Conference on Computational Intelligence and Communication Technologies (CCICT), DOI: [10.1109/CCICT53244.2021.00061](https://doi.org/10.1109/CCICT53244.2021.00061).
- [9] **Z.-T. Liu et al.** (2020). *An IMU-Based Wearable Ring for On-Surface Handwriting Recognition*. Proceedings of the International Symposium on VLSI Design, Automation, and Test, DOI: [10.1109/VLSIDAT49148.2020.9196479](https://doi.org/10.1109/VLSIDAT49148.2020.9196479).
- [10] **Tanaka, S. et al.** (2014). *Recognition of Finger Alphabet Using Remote Image Sensor*. Proceedings of the 76th National Convention (IPJS), Volume 76, Pages 2275-2276.
- [11] **Ohkura, M. et al.** (2011). *A Recognition Algorithm of Numerals Written in the Air by a Finger Tip*. Journal of IPSJ, Volume 52, Pages 910-916.
- [12] **Murata, T. et al.** (2014). *Hand Gesture and Character Recognition Based on Kinect Sensor*. International Journal of Distributed Sensor Network, Volume 10, DOI: [10.1155/2014/278460](https://doi.org/10.1155/2014/278460).
- [13] **Ramachandra, V. et al.** (2022). *Virtual Air Canvas Using OpenCV and MediaPipe*. Proceedings of the International Conference on Futuristic Technologies (INCOFT), DOI: [10.1109/INCOFT55651.2022.10094385](https://doi.org/10.1109/INCOFT55651.2022.10094385).
- [14] **Rai, P. et al.** (2022). *Virtual Canvas for Interactive Learning Using OpenCV*. Proceedings of the IEEE 3rd Global Conference for Advancement in Technology (GCAT), DOI: [10.1109/GCAT55367.2022.9971903](https://doi.org/10.1109/GCAT55367.2022.9971903).