Air Canvas: Survey On Hand Tracking Using MediaPipe

[1]Mr.Shreyash Gawande, [2]Mr.Ganesh Hiwarkar, [3]Mr.Ninad Dange,
[4]Dr.P.S.Gawande [1,2,3] Student, Department of Computer Science and Engineering,
Anuradha Engineering College, Chikhli
[4] Head of Department of Computer Science and Engineering,
Anuradha Engineering College, Chikhli
[1]shreyashmgawande@gmail.com , [2]ganeshhiwarkar9460@gmail.com, [3]dangeninad82@gmail.com,
[4]prashant.gawande@aecc.ac.in

ABSTRACT: Air Canvas is a system that lets people draw in the air using only their hand. You don't need a pen, paper, or any device to draw. It works by using a camera, which watches your hand movements. Then, with the help of special computer tools called MediaPipe and OpenCV, it understands your hand's position and draws on the screen. It's like having an invisible pen in your hand. This system is helpful for digital artists, teachers in online classes, people playing games, and especially for people who cannot use normal tools due to disabilities. It is easy, clean (no touching), and fun to use.

Keywords: Virtual Drawing, Contactless Interaction, Air Canvas, Gesture Recognition, Computer Vision.

Corresponding Author: *Mr.Shreyash Gawande*
Student, Department of CSE, Anuradha Engineering College Chikhli, Maharashtra, India
Mail: shreyashmgawande@gmail.com

INTRODUCTION:

Technology is changing. Earlier, people used a mouse or keyboard to control computers. Now, people want faster and easier ways, like using their hands to give commands. Air Canvas is one of those new ways. It lets you draw in the air with your hand while the camera captures your hand's motion. This system is very useful during times like the COVID-19 pandemic, where people avoid touching things to stay safe. It's also good for students and teachers in online learning, for artists who want to try new ways of creating art, and for disabled people who find it hard to use tools like a mouse or touchscreen.

It uses computer vision and machine learning This makes drawing and using the computer fun and easy.

The development of such systems combines multiple domains of computer science, including image processing, machine learning, and HCI. Challenges such as accurate hand detection, minimizing latency, ensuring smooth stroke rendering, and interpreting diverse gestures are addressed through advanced algorithms and techniques for smoother hand tracking. These innovations ensure a seamless and responsive user experience.

As interest in immersive technologies grows, the Air Canvas system stands as a practical and scalable solution for virtual interaction. This paper explores the technical architecture, gesture recognition techniques, use cases, and future improvements of Air Canvas systems, aiming to provide a comprehensive understanding of their capabilities and impact on the evolving landscape of human-computer interaction.

LITERATURE REVIEW

Hand gesture recognition and virtual drawing interfaces have gained significant traction due to the advancement of computer vision techniques and the accessibility of deep learning models. Zhang et al. introduced MediaPipe Hands, a highly efficient and real-time hand tracking system that operates across platforms and devices, enabling the development of hand gesture-based applications without the need for external sensors [1]. Similarly, Chaudhary et al.developed a virtual drawing system using hand gestures and color detection, focusing on creating an interactive interface for real-time sketching [2].Jain and Sahu presented an air canvas system leveraging OpenCV and MediaPipe, which allows users to draw virtually using finger gestures. This system emphasizes robustness and real-time performance [3]. Patel and Sharma explored a similar virtual painting interface using real-time hand detection, enhancing accuracy through optimized image processing techniques [4].Google AI Blog highlighted the capabilities of MediaPipe for real-time hand tracking and its integration potential with various AI-driven applications [5]. Kaur and Gill proposed a gesture-based drawing interface by combining machine learning with computer vision to create a seamless user experience [6].Singh and Gupta enhanced the virtual drawing experience by integrating Kalman filters with AI-based gesture control, improving the stability and smoothness of drawings [7]. Al Farid et al. provided a comprehensive review of vision-based hand gesture recognition systems, discussing methodologies, datasets, and challenges faced by researchers [8].Sahoo et al. employed a fine tuned Convolutional Neural Network (CNN) for real-time hand gesture recognition, achieving high accuracy on benchmark datasets [9]. Gadekallu et al.optimized CNN models using the Harris Hawks algorithm to further improve recognition performance [10].

Lim et al.developed 'Handmate,' a browser-based controller utilizing AI hand-tracking for interactive web applications such as audio and MIDI control [11]. Ahmad et al. proposed a landmark classification-based approach for hand gesture recognition using MediaPipe's hand landmarks [12].Liu et al. combined multisensing and information processing capabilities in a stretchable neuromorphic transistor, pushing the boundaries of wearable gesture recognition systems [13]. Cai et al. introduced the LET procedure for improving gesture recognition across varying forearm angles using sensor data [14].Wang et al. integrated deformable convolutional networks with sEMG signals for robust hand gesture recognition [15]. Bhushan et al.analyzed multiple machine learning algorithms to evaluate their effectiveness for gesture recognition tasks.[16].Bhaumik et al. presented HyFiNet, a hybrid feature attention network that combines different features to enhance hand gesture recognition accuracy [17]. Subramanian et al. utilized a GRU model optimized with MediaPipe for Indian sign language recognition, achieving promising results [18].Garg et al. used CNN and MediaPipe for yoga pose classification, demonstrating how hand tracking can extend to posture and pose recognition [19]. Amprimo et al. merged MediaPipe with RGB-depth cameras to assess hand motor skills remotely, contributing to healthcare applications [20].

Nitin Kumar et al. implemented an air writing recognition system utilizing MediaPipe and OpenCV for intuitive character recognition in air [21]. Haldera and Tayade developed a vernacular sign language recognition system by combining MediaPipe with machine learning classifiers [22].Deng et al.proposed a 3D hand pose estimation model based on depth cameras, enhancing accuracy in three-dimensional gesture recognition tasks [23]. Harish discussed a computer vision-based system for real-time hand gesture recognition applicable to various user interaction scenarios [24].Roy and Akif presented a gesture-based human-computer interaction system that prioritizes user-friendliness and real-time responsiveness [25]. Vasavi et al. explored painting applications using MediaPipe to facilitate virtual art creation [26].Zhou and Lv examined gesture recognition applications in virtual environments, using MediaPipe to improve interaction accuracy [27]. Farooq et al. reviewed advancements in machine translation for sign language, providing insights into the limitations and potential of these technologies [28].Schlüsener and Bücker introduced few-shot learning models for dynamic gesture recognition, addressing data scarcity challenges in training deep learning models [29]. Gupta et al. applied digital twin techniques and CNN fusion methods to human action recognition, expanding gesture recognition to broader activity recognition domains [30].Lastly, Khajuria et al. applied CNN and VGG16 models to facial emotion recognition, illustrating the versatility of computer vision techniques beyond hand gestures [31].

COMPARISON & DISCUSSION

1. **Air Canvas vs Mouse**

To use a mouse or stylus, you need to touch them and know how to use them properly. Styluses can be costly and are used on tablets. Air Canvas doesn't need any extra device—just a camera. This makes it cheap and easy for more people to use. But for very detailed drawing, a stylus might be more precise than Air Canvas.

2. Air Canvas vs Touchscreen

Touchscreens are easy to use and respond quickly when you touch them. But in public places, they can become dirty or unhygienic. Air Canvas doesn't require any touch, so it is clean and safe. However, sometimes it might have a slight delay or not be as smooth as touchscreens, especially in poor lighting.

3. Air Canvas vs Other Gesture Devices

Devices like Kinect or Leap Motion are very good at tracking movements, but they are expensive and need special equipment. Air Canvas only uses a regular camera and is much cheaper, but it may not be as accurate as those devices.

4. User Experience

People who tried Air Canvas said it was fun and easy to use. They liked drawing in the air. However, it works best when the room is well-lit. You can also change things like brush size and colors. Some users might need a little practice to use it well.

FUTURE WORK AND POTENTIAL IMPROVEMENTS

In the future, Air Canvas can be made even better by using AI to guess movements, storing drawings in the cloud, and working on phones or other devices. It can also be used in Virtual Reality (VR) or Augmented Reality (AR) to make 3D drawings, not just 2D. It can be helpful in schools, work-from-home, art, therapy, and even games.

CONCLUSION

Air Canvas is a new and exciting way to draw and interact with a computer without touching anything. It uses your hand and a webcam, with smart tools like MediaPipe and OpenCV, to follow your movements. It is useful for clean, safe, and easy digital drawing. It is perfect for teaching, online meetings, drawing, and helping people who can't use normal devices. Even though it might have small issues, like needing good lighting or not being perfect for detailed drawings, it can be improved using AI, better tracking, and maybe depth sensors. In the future, Air Canvas can work in VR/AR and allow users to draw in 3D space. It is a great step toward making computer use easier, more fun, and open to everyone, including people with special needs.

In conclusion, Air Canvas offers an innovative, user-friendly, and scalable platform for gesture based drawing and interaction, paving the way for more natural and immersive digital experiences. With continued development and optimization, it has the potential to redefine digital art and interface control, making technology more inclusive and engaging.

REFERENCE:

[1] Zhang, F., Bazarevsky, V., et al. (2020). MediaPipe Hands: On-Device Real-Time Hand Tracking. *arXiv preprint* arXiv:2006.10214.

[2] Chaudhary, A., Raheja, J. L., Das, K., & Raheja, S. (2021). Intelligent virtual drawing using hand gestures and color detection. *Multimedia Tools and Applications*, 80(1), 159-180. [3] Jain, P., & Sahu, P. (2023). Air Canvas Using Hand Gestures for Drawing in Real-Time Using OpenCV and MediaPipe. *International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)*, 11(2), 320–326.

[4] Patel, R., & Sharma, K. (2022). Virtual Painting Interface using Real-time Hand Detection. *International Journal of Computer Applications Technology and Research*, 11(4), 145–149. [5] Google AI Blog (2021). Real-Time Hand Tracking with MediaPipe.

[6] Kaur, H., & Gill, M. (2023). Gesture-Based Drawing Interface Using Machine Learning and Computer Vision. *Procedia Computer Science*, 212, 180–187.

[7] Singh, V., & Gupta, R. (2022). Enhancing Virtual Drawing Experience Using Kalman Filters and AI-based Gesture Control. *International Journal of Computer Vision and Signal Processing*, 12(1), 35–42.

[8] Al Farid, F., et al. (2022). A Structured and Methodological Review on Vision-Based Hand Gesture Recognition System. *J. Imaging*, 8(6), 153.

[9] Sahoo, J.P., Prakash, A.J., Pławiak, P., & Samantray, S. (2022). Real-Time Hand Gesture Recognition Using FineTuned Convolutional Neural Network. *Sensors*, 22(3), 706. [10] Gadekallu, T. R., et al. (2022). Hand Gesture Recognition Based on a Harris Hawks Optimized Convolution Neural Network. *Computers & Electrical Engineering*, 100, 107836.

[11] Lim, M., Kotsani, N., & Hartono, P. (2022). Handmate: An Accessible Browser-based Controller for Web Audio and Midi using AI Hand-Tracking. *International Conference on New Interfaces for Musical Expression*.

[12] Ahmad, K. A., Silpani, D. C., & Yoshida, K. (2022). Hand Gesture Recognition by Hand Landmark Classification. *J-STAGE Conference Proceedings*.

[13] Liu, L., et al. (2022). Stretchable Neuromorphic Transistor That Combines Multisensing and Information Processing for Epidermal Gesture Recognition. *ACS Nano*, 16(2), 2282- 2291.

[14] Cai, S., et al. (2022). The LET Procedure for Gesture Recognition With Multiple Forearm Angles. *IEEE Sensors Journal*, 22(13), 13226-13233.

[15] Wang, H., Zhang, Y., Liu, C., et al. (2022). sEMG based hand gesture recognition with deformable convolutional network. *International Journal of Machine Learning and Cybernetics*, 13, 1729–1738.

[16] Bhushan, S., et al. (2022). An Experimental Analysis of Various Machine Learning Algorithms for Hand Gesture Recognition. *Electronics*, 11(6), 968.

[17] Bhaumik, G., Verma, M., Govil, M.C. et al. (2022). HyFiNet: Hybrid feature attention network for hand gesture recognition. *Multimedia Tools and Applications*. [18] Subramanian, B., et al. (2022). An integrated mediapipe-optimized GRU model for Indian sign language recognition. *Scientific Reports*, 12, 11964.

[19] Garg, S., Saxena, A., & Gupta, R. (2022). Yoga pose classification: a CNN and MediaPipe inspired deep learning approach for real-world application. *Journal of Ambient Intelligence and Humanized Computing*.

[20] Amprimo, G., et al. (2022). GMH-D: Combining Google MediaPipe and RGB-Depth Cameras for Hand Motor Skills Remote Assessment. *IEEE International Conference on Digital Health (ICDH)*, 132-141.

[21] Nitin Kumar, R., et al. (2022). Air Writing Recognition Using Mediapipe and OpenCV. In: Karuppusamy, P., García Márquez, F.P., Nguyen, T.N. (eds) *Ubiquitous Intelligent Systems*, Smart Innovation, Systems and Technologies, vol 302. Springer.

[22] Haldera, A., & Tayade, A. (2021). Real-time Vernacular Sign Language Recognition using MediaPipe and Machine Learning. *International Journal of Research Publication and Reviews*, 2(5), 9-17.

[23] Deng, Z., et al. (2023). A 3D hand pose estimation architecture based on depth camera. *SPIE Proceedings*, 12593.

[24] Harish, S. (2023). Computer vision-based Hand gesture recognition system. *NeuroQuantology*, 20(7), NQ33365.

[25] Roy, K., & Akif, M. A. H. (2022). Real Time Hand Gesture Based User Friendly Human Computer Interaction System. *International Conference on Innovations in Science, Engineering and Technology (ICISET)*, 260-265.

[26] Vasavi, R., et al. (2022). Painting with Hand Gestures using MediaPipe.

[27] Zhou, Z., & Lv, Y. (2023). The application of gesture recognition based on MediaPipe in virtual scene. *SPIE Proceedings*, 12636.

[28] Farooq, U., et al. (2021). Advances in machine translation for sign language: approaches, limitations, and challenges. *Neural Computing and Applications*, 33(21), 14357-14399. [29] Schlüsener, N., & Bücker, M. (2022). Fast Learning of Dynamic Hand Gesture Recognition with

Few-Shot Learning Models. *arXiv preprint* arXiv:2212.08363.

[30] Gupta, M., Kumar, R., & Dewari, S. (2021). Digital twin techniques in recognition of human action using the fusion of convolutional neural network. In *Digital Twin Technology* (pp. 165-186). CRC Press.

[31] Khajuria, O., Kumar, R., & Gupta, M. (2023). Facial Emotion Recognition using CNN and VGG16. In *2023 International Conference on Inventive Computation Technologies (ICICT), 472-477. IEEE.*