Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

VisionDraw: Hand Gesture Interactive Canvas

Ms. Archana Nikose [1], Mr. Deepak Tete [2], Mr. Pravin Padole [3], Mr. Harsh Bawane [4], Mr. Arpan Pandey [5], Mr. Yassir Sheikh [6]

[1] Assistant Professor, Department of Computer Science & Engineeering, Priyadarshini Bhagwati College of Engineering
[2,3,4,5,6] UG Scholar, Department of Computer Science & Engineeering, Priyadarshini Bhagwati College of Engineering

Abstract
In recent years, touch-free interaction has gained attention as users seek more convenient and natural ways to engage with computers. While traditional input tools such as the mouse and keyboard remain effective, they restrict creativity and always require direct physical contact. To overcome these limitations, this study presents VisionDraw: Hand Gesture Interactive Canvas, a lightweight and cost- efficient system that enables users to draw or write in the air using finger gestures instead of physical devices.The system employs a standard webcam combined with MediaPipe and OpenCV to detect and follow hand movements, using the index fingertip as a virtual pen. Through intuitive gestures, users can switch colors, erase strokes, or clear the digital workspace. During experimentation, the system delivered consistent and smooth tracking under normal illumination, though darker or cluttered backgrounds slightly reduced accuracy.Overall, VisionDraw demonstrates that by integrating computer vision and real-time hand tracking, it is possible to develop an accessible, low-cost, and engaging method for creative digital interaction.

Keywords: Computer Vision, Hand Gesture Recognition, MediaPipe, OpenCV, Virtual Canvas

1. Introduction

Technology has changed the way people interact with computers [1]. In the early days, the main methods of communication were simple devices like keyboards and later the mouse. These tools worked well but have always depended on physical contact. Over time, touchscreens, stylus pens, and voice recognition became popular, making human–computer interaction more natural. Even then, most of these systems still need direct contact or expensive hardware. With the growing interest in touchless and hygienic solutions, hand gesture recognition has started to attract attention as a practical alternative [4].

The development of human–computer interaction has evolved alongside advances in computing, moving from physical input devices such as keyboards and mice toward more intuitive and natural interfaces [1].Gestures provide an inherent and expressive way for humans to convey information and emotions, making them a suitable medium for interaction between people and machines [1].With the progress of computer vision and machine learning, it has become feasible to detect and track hand and finger movements accurately using a regular camera, without the need for additional sensors [8].This has opened

new opportunities for creating virtual drawing systems, interactive games, and even controlling smart devices without touching them.

This research work introduces VisionDraw: Hand Gesture Interactive Canvas, a project that allows users to draw on a digital screen simply by moving their index finger in the air. The system does not require special sensors or gloves — a standard webcam is enough. The video feed from the camera is processed with MediaPipe, which provides hand landmarks, and OpenCV, which is used to track the fingertip and display the drawings on a canvas [9]. The fingertip acts like a pen, and extra gestures are added for switching colors, erasing, or clearing the canvas.

The motivation behind this project is to provide a low- cost and simple solution that anyone can try. It is especially useful in classrooms, online teaching, or for people who want to draw digitally without buying extra hardware. During testing, the system performed well under normal light, though a darker background sometimes caused small tracking errors [7]. Despite this, the idea shows great promise for creative fields and for interactive learning tools.

In short, VisionDraw explores how computer vision can turn natural hand movements into a creative digital tool. It combines accessibility, simplicity, and interactivity, giving a small but meaningful step toward the future of touchless interfaces [10].

Literature Review / Background

With time, gesture recognition has turned out to be one of the most interesting areas in computer vision and human-computer interaction. The work by Mitra and Acharya [1] was among the first ones to describe how it is possible to communicate with computers by means of gestures. They showed that the motion of hands is a natural way for humans to express thoughts, and this idea encouraged many researchers to explore it further.

After that came a number of studies focused on how computers could actually recognize hand gestures. Molchanov et al. [2] introduced a deep learning-based approach using 3D convolutional neural networks to improve the accuracy of gesture detection in dynamic environments. Their study emphasized the ability of neural models to capture spatial and temporal information from gesture sequences. In another approach, Ahmed et al. [3] applied computer vision techniques with a standard webcam, demonstrating that reliable hand gesture recognition can be achieved even with low-cost hardware setups. This finding highlighted that efficient gesture recognition does not always require expensive equipment.

Lee and Hong [4] proposed a vision-based hand gesture recognition framework that relied on image processing algorithms to achieve smooth, real-time interaction. Their contributions laid a foundation for later systems that combined efficiency with usability. Building on this, Khan and Kumar [5] developed a real-time “drawing in air” interface using finger motion as input, confirming that such systems can perform fluid drawing operations through simple webcam tracking.

Thakur and Sharma [6] created a gesture-controlled virtual drawing model using OpenCV, focusing on fingertip tracking to draw on-screen shapes. However, their system’s reliability decreased under poor

lighting or unclear backgrounds. Singh and Verma [7] further enhanced the stability of such models by integrating artificial intelligence algorithms for improved hand detection and responsiveness.

To make hand tracking easier, Google introduced MediaPipe [8], a ready-to-use framework that detects multiple hand landmarks in real time using a webcam. OpenCV [9] has also been crucial for image processing and shape detection in live video input. Finally, Zhang [10] presented a flexible camera calibration method that significantly improved precision in computer vision applications dependent on camera input.

Research Gap Even after extensive exploration in this domain, existing gesture-based and air-drawing systems continue to face persistent challenges. Many models perform poorly in low-light or complex background conditions, while others depend on external sensors or wearable devices, increasing both cost and complexity [5], [6], [7]. Furthermore, most prior implementations struggle to maintain real-time responsiveness and user convenience on standard consumer hardware.To address these shortcomings, VisionDraw has been developed as a lightweight, real- time hand gesture drawing system that operates using only a webcam. The objective is to make digital drawing more intuitive, smooth, and accessible without additional sensors or specialized tools.

Problem Statement

Most people still use a mouse, keyboard, or touchscreen to control a computer [1]. These methods are fine, but they always need physical contact. In many places, like classrooms or offices, it can be uncomfortable or even unhygienic to touch devices again and again. Because of this, many researchers have tried to make systems that work without any touch, using hand gestures instead [2].

Some of these systems do a good job, but many have big drawbacks. A lot of them need special hardware such as gloves or sensors that cost a lot of money. Others depend too much on lighting or background, so they stop working if the conditions change [7]. Even computer vision–based systems that use a normal camera often have trouble tracking the hand properly or showing results in real time. These issues make them hard for normal users to rely on.

There’s still a clear gap in having a simple and low- cost system that works smoothly with just a webcam [9]. It should be easy to use and accurate enough to draw or write on a screen without delay.
This project, called VisionDraw, tries to solve that issue. It allows people to draw in the air using their index finger as a pen. The system tracks the movement of the hand using MediaPipe and OpenCV. With this, users can create drawings, erase, or change colors without touching any surface. The main goal is to make digital drawing more natural, hygienic, and accessible to everyone.

Methodology

This project mainly works on a simple idea — drawing on a screen without touching it [1]. Everything is done using the hand’s motion in front of a webcam. To make this possible, the whole system was built in a few simple steps, where each step handles one part of the process.

1. Capturing the Hand in Real Time: The first part of the work was to use the webcam to capture the hand movement. The camera keeps taking live frames one after another. Each frame represents a momentary visual input that allows the system to continuously monitor the hand's position and motion in real time [9]. The video stream is processed frame by frame, ensuring that every small movement of the hand can be analyzed for gesture tracking.
2. Detecting and Tracking the Hand: After getting the video, the next step is to find the hand inside each frame. The system looks for the hand shape and focuses on key points like the fingers. By using the MediaPipe framework, the algorithm identifies 21 landmark coordinates on the hand and isolates the index fingertip as the active drawing point [8]. This enables consistent fingertip localization even when the hand moves, rotates, or changes distance from the camera.The main point used for drawing is the tip of the index finger, which acts like a pen. When the finger moves, its position keeps changing, and that is used for drawing [8].
3. Drawing Process: Once the fingertip is detected, the system starts joining one point to another as the finger moves. This creates a continuous line, which looks like drawing with an invisible pen [5]. If the finger is lifted or moved away, the line stops automatically.
4. Gesture-Based Actions: A few hand gestures are used to control simple things. For example, one gesture can be for changing color, another for erasing, and one more for clearing the canvas. This makes the system feel more interactive and doesn't require any buttons or touch [6].
5. Testing and Observation: After building the system, it was tested in different lighting and background conditions. It worked quite well when the room had normal light. However, when the light was low or the background was too bright, the system sometimes struggled to detect the hand properly. Still, for normal conditions, the drawing came out smooth and accurate [7].

In simple words, the project follows an easy flow — the webcam watches the hand, the system finds the fingertip, and the finger's motion becomes a line on the screen. It's like using your finger as a digital brush. The goal was to make it work smoothly, without extra devices, so anyone can try it.

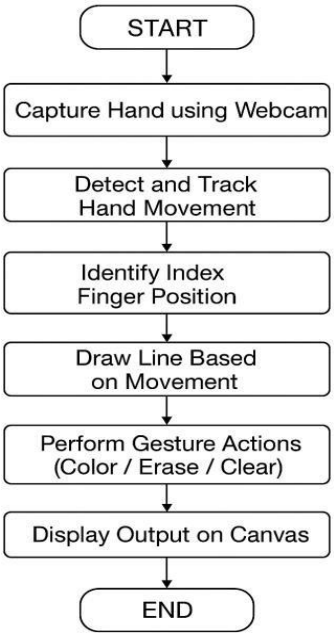

Fig 1 Flowchart of VisionDraw System

Implementation steps:
Step 1: Capturing the Hand I started by using the webcam to take live video of my hand [1]. The camera keeps capturing frames, and each frame helps the system find where the hand is. I made sure the lighting and angle were fine so the hand was easy to detect.
Step 2: Detecting and Tracking the Hand After that, I used MediaPipe to detect and track the hand [8]. It marks different points on the hand, and I focused on the index finger because it works like a digital pen. The fingertip position changes as I move my hand, and the system keeps tracking it.
Step 3: Drawing on the Screen Once the fingertip was detected, OpenCV was used to draw lines following the finger's path [9]. When the finger moved, a line appeared, and when I lifted my hand or moved out of view, it stopped automatically. It felt like drawing in the air.
Step 4: AddingGestures To make it interactive, I added simple gestures [5]. Moving the hand to certain spots on the screen changed the color, and another gesture cleared the canvas. It made the project easier and more fun to use.

Step 5: Testing In the end, I tested it under different lighting. It worked best in normal light but was a bit less accurate in dark or cluttered backgrounds [7]. Still, the overall result was smooth and responsive.

Result & discussion

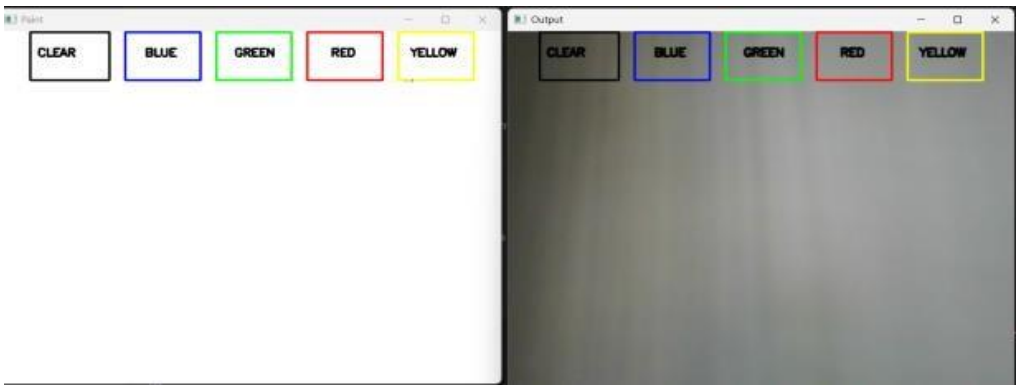

Fig 2 Result window-1

1. What you see at a glance: Two windows side-by- side. The left window is a plain white Paint canvas with the control buttons along the top. The right window is the live Output from the webcam — also showing the same control buttons overlaid on the video [1].
2. Purpose of each window: The left window keeps a clean copy of whatever is drawn.The right window is the live feedback so the user knows where their hand is and can interact with the buttons [8].
3. Top controls: The five boxes (CLEAR, BLUE, GREEN, RED, YELLOW) act as simple on-screen tools. CLEAR wipes the left canvas; the color boxes switch drawing color.

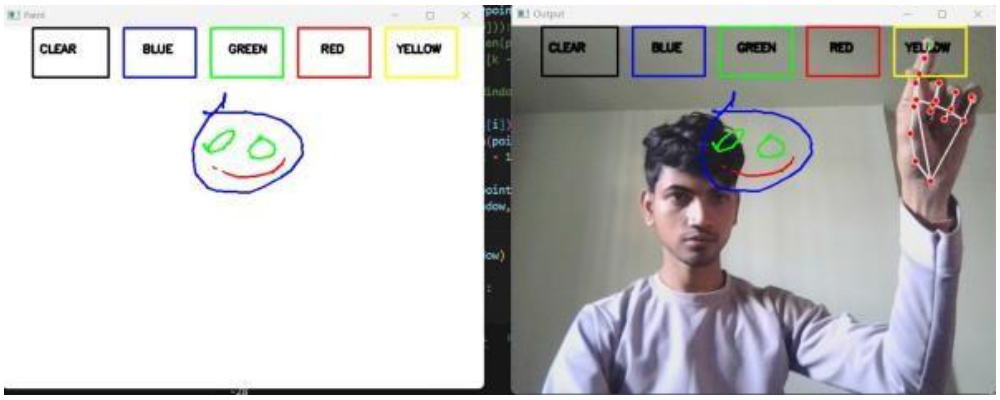

Fig 3 Result window-2

1. Immediate impression: The left Paint window now has a colored smiley (blue outline, green eyes, red mouth) — so multiple colors were used [5]. The right Output shows you the user with hand landmarks.
2. Step-by-step of what happened to produce this:

a. The webcam captured your hand and MediaPipe detected hand landmarks [9].
b. The system tracked the index fingertip and translated its screen coordinates to drawing coordinates.

3. What the landmarks tell us: The red dots and connecting lines on the right confirm that the hand tracker is locating joints reliably.
4. User experience notes: Drawing felt natural — the strokes are continuous and fairly smooth, which means the deque-based point buffering and line-joining are working. The overlay gives immediate feedback so you know when the system recognized your finger.

Conclusion

Working on this project, VisionDraw: Hand Gesture Interactive Canvas, was honestly a great learning experience. The main idea was to draw on the screen without touching anything — just by moving the hand in front of a camera. In the end, it actually worked the way I imagined.

While testing, I noticed it worked best when the light in the room was normal. If the lighting was too low or the background was a bit messy, the hand detection wasn't always perfect. Still, most of the time, the response was smooth and quick. It felt like using an invisible pen. Through this project, I got to understand how tools like MediaPipe and OpenCV can be used to make interactive systems.

References

1. S. Mitra and T. Acharya, "Gesture Recognition: A Survey," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 37, no. 3, pp. 311–324, 2007.
2. P. Molchanov, S. Gupta, K. Kim, and J. Kautz, "Hand Gesture Recognition with 3D Convolutional Neural Networks," IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1–7, 2015.
3. F. Ahmed, A. Haque, and M. Hasan, "Hand Gesture Recognition Using Computer Vision," International Journal of Computer Applications, vol. 975, pp. 8887–8959, 2018.
4. J. Lee and S. Hong, "Vision-Based Hand Gesture Recognition System for Human– Computer Interaction," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 10, no. 5, pp. 230–236, 2019.

5. R. U. Khan and P. Kumar, "Real-Time Drawing Using Hand Gesture Detection," International Research Journal of Engineering and Technology (IRJET), vol. 7, no. 6, pp. 2459–2464, 2020.
6. A. Thakur and N. Sharma, "Gesture Controlled Virtual Drawing Application Using OpenCV," International Journal of Innovative Research in Computer and Communication Engineering, vol. 8, no. 9, pp. 3421–3428, 2020.
7. M. K. Singh and R. Verma, "Hand Gesture- Based Human–Computer Interaction Using AI and Computer Vision," IEEE Access, vol. 9, pp. 122345–122358, 2021.
8. Google Research, "MediaPipe: Cross- Platform Framework for Building Perception Pipelines,"[Online].Available: https://developers.google.com/mediapipe.
9. OpenCV Team, "Open Source Computer Vision Library," [Online]. Available: https://opencv.org/.
10. Z. Zhang, "A Flexible New Technique for Camera Calibration," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pp. 1330–1334, 2000.