ISSN: 2583-7346
DOI: https://doi-ds.org/doilink/01.2024-86347548/IJAEAST.2024.01.0001, January 2024, Volume-2, Issue-1, pp.1-12
International Journal of Advances in Engineering Architecture Science and Technology

Aerosculpt: A Symphony of Creativity on the Air Canvas-Revolutionizing Digital Art and Design through Intuitive Hand Gestures

Hamzathul Favas E[1*], Abu Thahir M[2], Jaseela Thesni N[3], Mohammed Shahil P[4], Sabina N[5]

[1,2,3,4] Student, Department of Computer Science Engineering, MGM College of Engineering and Pharmaceutical Sciences, Kerala, India.

[5]Assistant Professor, Department of Computer Science Engineering, MGM College of Engineering and Pharmaceutical Sciences, Kerala, India.

Received: 02-01-2024 **Accepted:** 25-01-2024 **Published:** 26-01-2024

Abstract

Background: In the ever-evolving landscape of digital interaction, this research presents the development and implementation of the Virtual Air Canvas, a novel system that combines hand gesture recognition and machine learning for digital art and design applications. As technology progresses, the need for intuitive and inclusive interfaces becomes paramount, motivating our exploration into a hands-free approach for creative expression and communication.
Objectives: To design a robust hand gesture recognition system capable of translating dynamic gestures into digital commands with the integration of machine learning algorithms for enhanced gesture recognition accuracy and adaptability to explore the practical applications of the Virtual Air Canvas, with a focus on its utility as a digital whiteboard in educational settings, a communication tool for differently-abled individuals, and beyond.
Methods: The development of a custom hand gesture recognition model using a combination of depth sensing technology and machine learning algorithms. The system was pretrained with a MediaPipe pipeline for hand gestures to ensure adaptability to various user interactions. The machine learning component enhances the system's ability to learn and adapt to users' unique gestures over time.
Statistical Analysis: Assessing the accuracy and efficiency of the hand gesture recognition system through metrics such as precision, recall, and F1 score. Statistical methods are employed to evaluate the performance and robustness of the system across diverse gestures and user scenarios.
Findings: Results indicate a high level of accuracy in recognizing a wide range of hand gestures, demonstrating the effectiveness of the Virtual Air Canvas. The system's adaptability and learning capabilities contribute to its reliability in real-world applications.
Applications and Improvements: Particularly in educational environments as a digital whiteboard and as a communication tool for differently-abled individuals. Beyond these primary applications, the system holds promise for a variety of interactive and hands-free digital design scenarios.

*Hamzathul Favas E, Student, Department of Computer Science Engineering, MGM College of Engineering and Pharmaceutical Sciences, Kerala, India.
Email: ehamzathulfavas@gmail.com

Keywords: Machine Learning, Computer Vision, MediaPipe, Hand Recognition.

1. Introduction

In the rapidly advancing realm of human-computer interaction, the development of intuitive and inclusive interfaces has become paramount. This paper introduces the Virtual Air Canvas, a ground-breaking system that merges hand gesture recognition and machine learning to revolutionize digital art and design applications. As technology evolves, the demand for hands-free, yet precise, interfaces has grown, motivating our exploration into a system that allows users to interact with digital environments effortlessly through dynamic hand gestures. The primary motivation behind this research is to bridge the gap between traditional interfaces and innovative, gesture-based solutions. We aim to present a system that not only recognizes diverse hand gestures but also adapts and learns from user interactions, ensuring a seamless and personalized experience.

This research involved the intricate design of our hand gesture recognition model, the integration of machine learning algorithms for enhanced adaptability, and the statistical analyses validating the system's robust performance. Furthermore, the findings showcase the system's accuracy across a spectrum of gestures, demonstrating its potential for transformative applications. The Virtual Air Canvas is not merely a technological innovation; it signifies a shift towards more inclusive, creative, and accessible digital interactions. Through this research, we aspire to contribute to the growing field of gesture-based interfaces, shedding light on the immense possibilities such systems offer. Building upon the finger tracking capability, the next objective is to enable users to write and erase on a virtual screen without the need for any external devices writing and erasing, the project aims to enhance the functionality of the system by allowing users to draw shapes using their thumb as a reference point.

This feature will enable precise and controlled drawing, providing users with greater flexibility and artistic freedom. By exploring applications in educational settings, where the Virtual Air Canvas can serve as a dynamic digital whiteboard, and by enhancing communication for differently-abled individuals, we aim to underscore the practical relevance and impact of our work. The introduction sets the stage for an in-depth exploration of this innovative system, laying the foundation for a comprehensive understanding of its design, capabilities, and real-world applications.

2. Background

The origin of the art of writing can be traced back to 2000 BC, when Neolithic people first pioneered the practice. Initially, they etched their writings onto walls, eventually transitioning to stones as their medium. Over time, stones gave way to cloth, and currently, paper serves as the primary means of communication. The advent of QWERTY keyboards has propelled us into an era of digitalized writing. Traditional methods, such as pen and paper, are gradually being supplanted by electronic devices. As technology advances, there is a burgeoning demand for the development of human-machine interactions, particularly with the increasing utilization of augmented and virtual reality. Hand gesture applications have gained popularity, with notable examples being automotive interfaces (Ohn-Bar and Trivedi, 2014), the economic air writing system (Pavithra and Prabhu, 2016), and handwriting recognition in free space (Shashidhar, Kim, and Chai, 2015), all contributing to the field of hand gesture recognition. While hand gesture recognition is a pivotal aspect of air writing, it extends beyond mere gestures to encompass fingertip detection, tracking, and tracing. A LED light detection system for fingertip detection has been developed by Pavithra, where LED lights are initially identified, their movements captured, and the trajectory displayed on-screen using optical character recognition (OCR) to recognize and present the corresponding alphabet.

3. Literature Survey

In [1] An Economical Air Writing System: Converting Finger Movements to Text Using a Web Camera P. Ramasamy, G. Prabhu, and R. Srinivasan, 2016.

This presents an innovative solution for text input using finger movements captured by a web camera. The system offers an accessible and affordable alternative to traditional input methods, catering to a wide range of users, including those with limited resources and individuals with disabilities. The core concept of the proposed system revolves around utilizing a standard web camera to capture the finger movements of the user as they write in the air. By analyzing the video stream in real-time, the system translates these movements into corresponding text, allowing for seamless and efficient text input. The absence of physical keyboards or touchscreens eliminates the need for additional hardware, making the system cost-effective and accessible to a broader audience. One of the notable advantages of this system is its affordability. Traditional input devices such as keyboards or touchscreens can be costly, especially for individuals or organizations with limited financial resources. By leveraging the ubiquity of web cameras, the proposed system significantly reduces the financial barrier to entry.

This opens up opportunities for adoption in various contexts, including schools, libraries, and community centers, where budget constraints may limit the availability of conventional input devices. Moreover, the intuitive nature of air writing contributes to the system's practicality. Writing in the air using finger movements is a familiar action that most individuals can perform effortlessly. This reduces the learning curve associated with adopting new input methods, making the system accessible to users with varying levels of technological literacy. The ease of use also extends to individuals with disabilities, such as those with mobility impairments or conditions affecting their fine motor skills. The air writing system offers an alternative text input method that accommodates their needs and promotes inclusivity. In addition to affordability and intuitiveness, the real-time text conversion capability of the system enhances the user experience. As users write in the air, they can instantly see the corresponding text on a display or a connected device. This immediate visual feedback allows for quick error correction, ensuring the accuracy of the entered text. The real-time conversion also enables the system to keep up with the speed of the user's finger movements, ensuring a seamless and responsive writing experience. While the paper lays the groundwork for the proposed air writing system, there are potential areas for further exploration and development. For instance, research could focus on improving the accuracy and robustness of the text conversion algorithm, ensuring reliable performance in various lighting conditions and hand orientations. The system's potential integration with virtual reality environments or artistic expression platforms could expand its applications beyond traditional text input, unlocking new possibilities for interaction and creativity. The paper introduces an economical air writing system that utilizes a web camera to convert finger movements into text. The system's affordability, intuitive nature, and real-time text conversion make it a promising alternative to conventional input methods. By addressing the limitations of traditional hardware-based devices, this system has the potential to empower users with limited resources and individuals with disabilities, fostering inclusivity and accessibility in information technology.

ISSN: 2583-7346
DOI: https://doi-ds.org/doilink/01.2024-86347548/IJAEAST.2024.01.0001, January 2024, Volume-2, Issue-1, pp.1-12
International Journal of Advances in Engineering Architecture Science and Technology

In [2] A Pointing Gesture-Based Egocentric Interaction System: Dataset, Approach, and Application., Y. Huang, X. Liu, X. Zhang, and L. Jin, 2016.

The paper introduces a system that enables natural interaction between humans and devices in egocentric vision. This innovative system leverages pointing gestures, which are recognized as crucial interaction patterns. Comprising three main modules, namely gesture recognition, fingertip localization, and character recognition, it offers a comprehensive framework for seamless human-device interaction. The gesture recognition module utilizes a two-stage Faster R-CNN-based approach for hand detection and dual-target fingertip detection. This framework exhibits remarkable accuracy in detecting hands and fingertips in egocentric videos. The fingertip localization module employs the detected fingertips to track their trajectory, enabling precise tracking of hand movements. The character recognition module employs a deep learning model to accurately identify characters written in the air during interaction. To facilitate the development and evaluation of the system, the authors compiled a large-scale dataset of pointing gestures for egocentric vision, known as EgoFinger. This dataset encompasses over 100,000 images and videos capturing pointing gestures in diverse environments and featuring various hand shapes. The EgoFinger dataset serves as a valuable resource for training and assessing the performance of the gesture recognition and fingertip localization modules. The authors conducted thorough evaluations of the system's performance using the EgoFinger dataset. Notably, they achieved a fingertip detection error of approximately 12.22 pixels within a 640px × 480px video frame. Moreover, the system demonstrated a remarkable accuracy of over 90% in recognizing characters written in the air, showcasing its proficiency in gesture-based character input. The presented system holds great promise for facilitating natural human-device interaction in egocentric vision. Its ability to accurately detect and track fingertips empowers users to effortlessly write characters in the air, opening up exciting possibilities for interaction with wearable devices. Designed to be compatible with wearable devices like smart glasses, the system allows users to interact without interrupting their ongoing tasks. This feature is particularly valuable in applications where users need to engage with the system while on the move, such as augmented reality or virtual reality experiences. Additionally, the system's versatility enables interaction with a wide array of applications, including augmented reality, virtual reality, and gesture-based control. Its capabilities encompass pointing to real-world objects, selecting virtual items, and controlling devices through intuitive hand gestures. This flexibility positions the system as a valuable tool across diverse domains. While still in the development phase, the system exhibits significant potential for natural human-device interaction. The authors are actively working on enhancing the system's accuracy and introducing new features, indicating a commitment to refining its performance and expanding its capabilities. By addressing these aspects, the system can evolve into a robust and effective tool for seamless and intuitive human-device interaction.

In [3] Bare Finger 3D Air-Touch System Using an Embedded Optical Sensor Array for Mobile Displays., Guo-Zhen Wang, Yi-Pai Huang, Tian-Sheeran Chang, and Tsu-Han Chen (2014).

The paper introduces a novel touch input system that enables touchless interaction on mobile displays. The system aims to provide a more intuitive and natural user experience by eliminating the need for physical contact with the screen. The touchless system consists of three main components: an embedded optical sensor array, a gesture recognition algorithm, and a mobile display. The optical sensor array is integrated into the mobile device, typically positioned around the edges of the display. It comprises a grid of optical sensors that detect changes in light intensity caused by the movements of the user's finger in the vicinity of the

ISSN: 2583-7346
DOI: https://doi-ds.org/doilink/01.2024-86347548/IJAEAST.2024.01.0001, January 2024, Volume-2, Issue-1, pp.1-12
International Journal of Advances in Engineering Architecture Science and Technology

display. The system operates based on the principles of light intensity variation and triangulation. When the user's finger approaches the display, it interrupts the light paths between the optical sensors and the display surface. By analyzing the pattern of these interruptions, the system can determine the position and movement of the user's finger relative to the display, enabling touchless gesture recognition and tracking in three dimensions (3D). To accurately interpret the user's intended actions, a gesture recognition algorithm is developed to process the data obtained from the optical sensor array. This algorithm maps the detected finger movements to specific touch commands or gestures. It takes into account factors such as finger position, velocity, and direction to ensure a precise interpretation of the user's interactions. To evaluate the performance of the touchless system, experiments were conducted comparing it with traditional touch input methods. Various factors, including accuracy, response time, and user satisfaction, were assessed.

The results demonstrated that the system was capable of accurately detecting and interpreting a wide range of touch gestures, providing a viable alternative to conventional touch interfaces. The potential applications of the touchless system are diverse, particularly in areas such as mobile gaming, virtual reality, and augmented reality. In these contexts, touchless interaction can enhance user immersion and interactivity by eliminating the need for physical controllers or touchscreens, allowing for more intuitive and natural interactions. The paper also explores potential improvements and optimizations for the system in the future. For instance, reducing power consumption and enhancing recognition accuracy are identified as important areas for further development. The authors also suggest considering broader applications beyond mobile devices, such as larger display screens or interactive surfaces. In summary, the paper presents a novel touchless input system for mobile displays. By utilizing an embedded optical sensor array and a gesture recognition algorithm, the system enables users to interact with the display without physical contact. The paper provides experimental evidence of the system's effectiveness, discusses potential applications, and highlights avenues for future improvements and developments.

In [4] Hand Gesture Recognition in Real-Time for Automotive Interfaces., Eshed Ohn-Bar, Mohan Manubhai Trivedi, 2014.

This focuses on the development of a system that enables real-time hand gesture recognition for automotive interfaces. Propose a multimodal vision-based approach that utilizes cameras and computer vision algorithms to capture and interpret hand gestures. The primary objective of the research is to allow drivers to interact with in-vehicle systems without the need for physical contact or distraction from driving. To achieve this, the authors outline the different stages involved in the hand gesture recognition process. The first stage is hand detection, where the system identifies the presence of a hand within the captured images or video stream. The authors employ sophisticated computer vision techniques to accurately detect hands in various positions and orientations. Once the hand is detected, the next stage is hand tracking, where the system continuously tracks the movement of the hand over time. This allows the system to maintain a consistent representation of the hand despite changes in lighting conditions or occlusions. The following stage is hand gesture modeling, where the system creates a model of the hand and its pose based on the tracked data. The authors utilize a depth camera, which provides depth information along with the RGB image, to improve the accuracy of hand modeling.

Finally, the last stage is gesture recognition, where the system identifies and classifies the performed hand gestures. The authors employ machine learning algorithms to train the system to recognize a variety of gestures, such as pointing, swiping, and rotating motions. To evaluate the performance of their system, the authors conducted extensive experiments using a dataset of hand gestures performed by different individuals. They compared the accuracy and efficiency of their approach with existing methods and demonstrated its superiority in terms of real-time gesture recognition and tracking capabilities. The paper also discusses the integration of the proposed hand gesture recognition system with in-vehicle interfaces. Highlight the potential applications of the system, such as controlling infotainment systems, navigation systems, and climate control, without the need for physical buttons or touchscreens. In addition to presenting their approach and evaluation results, the authors address the challenges and limitations of the proposed system. These challenges include handling occlusions, different lighting conditions, and variations in hand shapes and sizes. Propose potential solutions and discuss future research directions to further improve the system's performance. The paper provides a comprehensive exploration of a multimodal vision-based approach for real-time hand gesture recognition in automotive interfaces. The proposed system offers a promising solution to enhance driver safety and the user experience by enabling intuitive and non-distracting interaction with in-vehicle systems.

4. Existing System

Traditional hand gesture recognition systems often rely on two predominant methods for capturing and interpreting hand movements: using LED lights for motion tracking and using color-tipped fingers as markers for gesture identification. In LED-based systems, small LED lights are attached to the user's fingertips or other parts of the hand. These lights emit infrared or visible light, and their movements are tracked by cameras to interpret hand gestures. LED-based systems require precise calibration to establish a correspondence between the position of the LEDs in the physical space and their representation in the captured images. Another approach involves having users wear colored tips on their fingers. Cameras then capture the distinct color patterns, allowing the system to identify and interpret specific hand gestures. Calibration is also essential in color-tipped finger systems to accurately associate each color with a particular finger or gesture.

Disadvantages

- Limited precision
- Dependence on lighting conditions
- Intrusiveness
- Complex calibration
- Limited adaptability

5. Proposed System

Problem Statement

The virtual air canvas project aims to develop an interactive system that enables users to draw in the air using their hand movements. The objective is to create a seamless and intuitive drawing experience without relying on physical tools such as pens or brushes. By leveraging computer vision algorithms, the system will track and interpret the user's hand movements, translating them into drawings on a virtual canvas. The first key component of the project involves hand tracking. The system needs to accurately detect and track the user's hand in real-time. This functionality can be achieved using the Mediapipe library, which provides robust hand-tracking capabilities.

By identifying hand landmarks, such as the positions of the fingertips and palm, the system can precisely track the user's hand movements throughout the drawing process. Gesture recognition is another essential aspect of the virtual air canvas project. The system should be able to recognize specific gestures made by the user's hand to perform various actions. For instance, closing the hand into a fist can signify the start of drawing, while opening the hand can indicate the cessation of drawing. Additional gestures can be employed to change colors or brush sizes. By analyzing the relative positions and movements of the hand landmarks, the system can identify and interpret these gestures accurately. Creating a virtual canvas is another crucial step in the project. The system needs to generate a digital canvas on which the user can draw. This canvas can be represented as an image or a video frame. OpenCV, a popular computer vision library, can be used to create a blank canvas or load an existing image or video frame that serves as the drawing surface.

To facilitate the drawing process, the system requires a mechanism to interpret the user's hand movements and translate them into drawings on the canvas. By mapping the positions of the hand landmarks to their corresponding locations on the canvas, the system can accurately reproduce the user's gestures as lines, circles, or other shapes. OpenCV provides various drawing functions that can be used to render these shapes on the canvas. Real-time visualization is a crucial aspect of the project, ensuring that users can witness their drawings as they create them. The system should continuously update and render the canvas, allowing users to visualize their drawings in real-time. OpenCV can be employed to achieve this functionality, ensuring a smooth and interactive drawing experience. The virtual air canvas addresses the challenge of creating an immersive and interactive drawing experience without the need for physical tools. By employing hand tracking, gesture recognition, canvas creation, drawing mechanisms, and real-time visualization, the system enables users to express their creativity by drawing in the air.

System Outline

The system consists of several interconnected components that enable users to draw in the air using hand movements. The project utilizes MediaPipe for hand tracking, NumPy for efficient numerical computations, and OpenCV for image processing and rendering. Hand tracking is a crucial component of the system, as it accurately detects and tracks the user's hand in real-time. By leveraging MediaPipe's hand tracking capabilities, the system identifies the landmarks of the hand, including the fingertips and palm positions. This allows for precise tracking of the user's hand movements throughout the drawing process. Gesture recognition plays a significant role in the virtual air canvas system, enabling users to perform specific actions using hand gestures. For instance, closing the hand into a fist can signify the start of drawing, while opening the hand can indicate the cessation of drawing. Additional gestures can be employed to change colors or brush sizes. By analyzing the relative positions and movements of the hand landmarks, the system accurately interprets these gestures. The virtual canvas is created using OpenCV, which provides functionalities for image manipulation and rendering. The canvas can be represented as an image or a video frame, acting as the drawing surface for the user. OpenCV's drawing functions allow the system to render the user's hand movements as lines, circles, or other shapes on the canvas. Real-time visualization is a vital aspect of the system, ensuring that users can see their drawings as they create them. By continuously updating and rendering the canvas using OpenCV, the system provides users with immediate feedback, allowing them to visualize their drawings in real-time. The combined functionality of hand tracking, gesture recognition, canvas creation, drawing mechanisms, and real-time visualization creates an immersive and interactive drawing experience. The system enables users to express their creativity by drawing in the air, eliminating the need for physical tools such as pens or brushes.

ISSN: 2583-7346
DOI: https://doi-ds.org/doilink/01.2024-86347548/IJAEAST.2024.01.0001, January 2024, Volume-2, Issue-1, pp.1-12
International Journal of Advances in Engineering Architecture Science and Technology

Use-Case Model

The system's use case diagram consists of several key components: initializing variables, reading frames, detecting contours, starting drawing, tracking coordinates, and executing the output. Let's delve into the implementation details of each component to provide a comprehensive understanding of the virtual air canvas system. To begin with, the system initialization involves setting up variables and parameters required for the subsequent steps. This includes initializing the canvas, defining color and brush size, and establishing flags for drawing and gesture recognition. These variables will be used throughout the system to store and manipulate the necessary information. The next step in the implementation is reading frames from the input source, such as a webcam or video file. This is typically achieved using OpenCV, which provides convenient methods for capturing and processing frames. By continuously capturing frames, the system can track the user's hand movements and update the canvas in real-time. Detecting contours plays a vital role in identifying the user's hand in the captured frames. OpenCV offers various image processing techniques, such as thresholding and contour detection algorithms, to extract the hand region from the frame. These techniques analyze the pixel intensities and boundaries to isolate the hand, providing the necessary input for further processing. Once the hand contours are detected, the system can initiate the drawing process.

This is typically triggered by recognizing a specific hand gesture, such as closing the hand into a fist. When the gesture is detected, the system sets the drawing flag to true, indicating that the user is ready to draw. Conversely, opening the hand or performing another gesture can stop the drawing process by setting the drawing flag to false. Tracking the coordinates of the hand is a crucial step in the virtual air canvas system. By continuously monitoring the position of the hand contours in subsequent frames, the system can track the movement of the user's hand. This allows for translating the hand movements into corresponding drawing actions on the canvas. The tracked coordinates are typically stored and updated in variables, enabling real-time interaction with the canvas. The final step in the use case diagram is executing the output. This involves rendering the tracked hand movements on the canvas, creating the illusion of drawing in the air. OpenCV provides various drawing functions, such as line or circle drawing, that can be utilized to represent the hand movements as strokes or shapes on the canvas. The output is continuously updated and rendered in real-time, providing immediate visual feedback to the user.

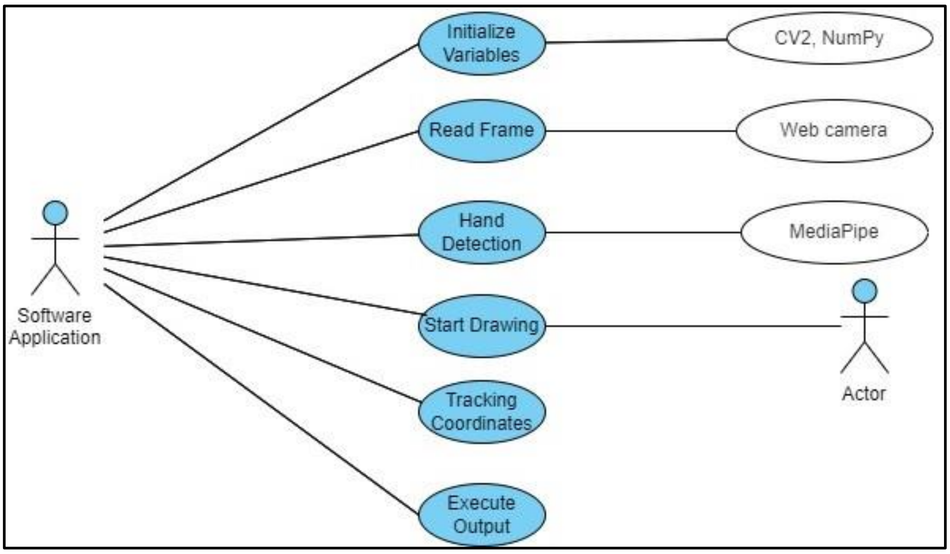

Figure 1. Use Case Diagram

6. Methodology

The methodology requires a large amount of data to be stored and sometimes leads to incorrect predictions due to background differences or skin color differences. Some of the existing papers followed the image processing process using thresholding and database data for different images. To overcome all these limitations and drawbacks, we designed the system using MediaPipe. In the proposed system, hand tracking is performed using MediaPipe, which first detects hand landmarks and then obtains positions according to them. The figure below describes the system flow or project process in which it works. This is a complete workflow that occurs when starting the system and drawing images with just an efficient hand wave.

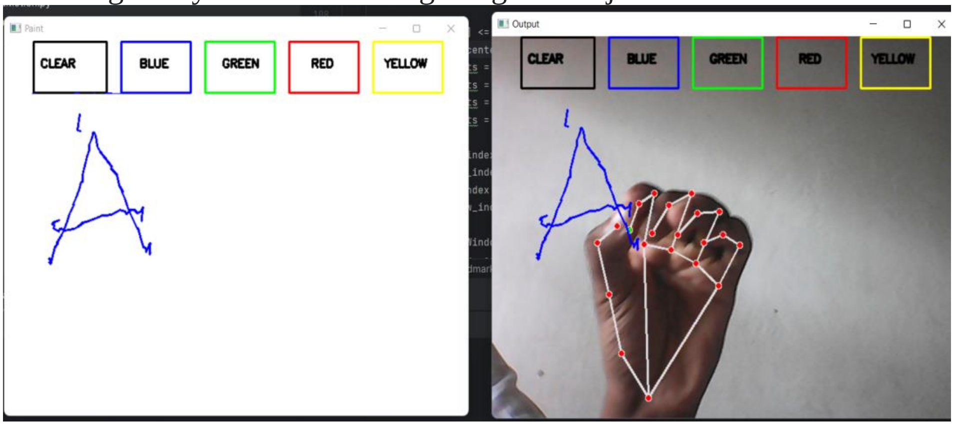

Figure 2. Draw Using Blue Color

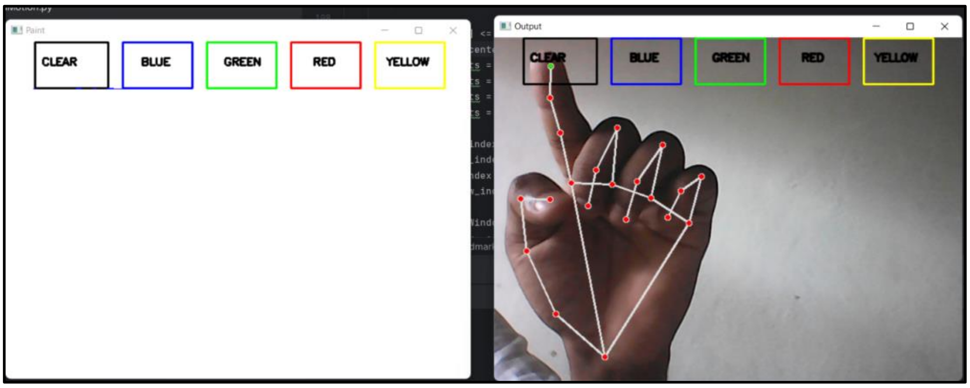

Figure 3. Clear the Entire Drawings from the Window

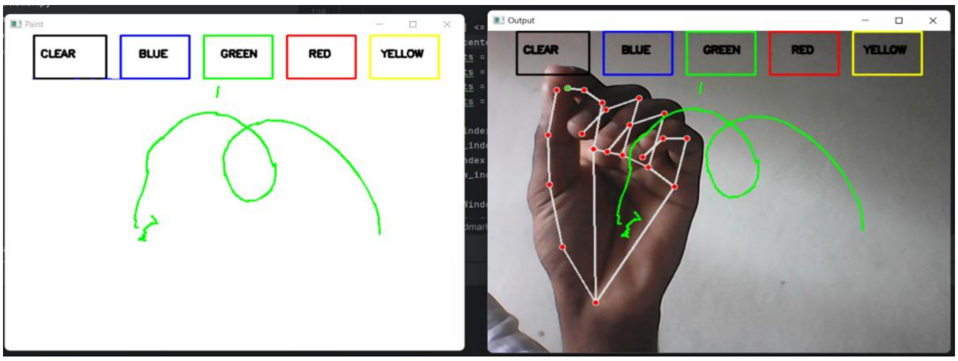

Figure 4. Draw Using Green Color

ISSN: 2583-7346
DOI: https://doi-ds.org/doilink/01.2024-86347548/IJAEAST.2024.01.0001, January 2024, Volume-2, Issue-1, pp.1-12
International Journal of Advances in Engineering Architecture Science and Technology

System Architecture

The system architecture consists of several interconnected components that work together to provide an immersive drawing experience. The key components include gesture recognition, camera input, image processing, gesture tracking and recognition, decision-making, command memory, a PC for decoding appropriate commands, and output interpretation based on gestures. The gesture recognition component serves as the interface between the user and the system. It captures the user's hand movements and recognizes specific gestures to trigger various actions. This component employs computer vision algorithms and machine learning techniques to analyze the captured hand movements and determine the corresponding gestures. The input module captures real-time video input through cameras, typically integrated into devices such as webcams or depth-sensing cameras. This raw video feed serves as the basis for hand gesture recognition.

The accuracy and robustness of the gesture recognition algorithms are crucial for ensuring a seamless and intuitive interaction with the system. The camera input component is responsible for capturing the user's hand movements in real-time. It can be a webcam or any other camera device that provides a video stream. The camera continuously captures frames that contain the user's hand, which are then processed by the subsequent components. The image processing component takes the frames from the camera input and applies various techniques to extract relevant information for gesture recognition. This involves tasks such as background subtraction, noise reduction, and hand region segmentation. These operations help isolate the user's hand from the background and other elements in the frame, enabling more accurate tracking and recognition.

The gesture tracking and recognition component analyzes the processed frames to track and recognize the user's hand movements. This step involves extracting features from the hand, such as hand shape, finger positions, and motion patterns. Machine learning algorithms, such as convolutional neural networks or support vector machines, can be employed to classify the gestures based on these features. The recognition results are then used for subsequent decision-making. The decision-making component processes the recognized gestures and determines the appropriate actions to be performed. For example, when a fist gesture is detected, the system may start drawing, and when an open-hand gesture is recognized, the system may stop drawing. This component acts as the control center, mapping the recognized gestures to the corresponding commands for the virtual air canvas system. The core of the system relies on the MediaPipe Hand Tracking module.

This module utilizes computer vision algorithms to identify and track the user's hand in the video feed. It robustly recognizes the hand's key landmarks, enabling accurate tracking of dynamic gestures. The command memory component stores and manages the recognized commands for further processing. It keeps track of the sequence of gestures and ensures that the system maintains the correct state during the drawing process. The command memory allows for smooth transitions between different gestures and ensures the system responds accurately to the user's input. The PC for decoding appropriate commands plays a crucial role in interpreting the recognized gestures and translating them into appropriate actions. This component receives the recognized gestures and maps them to the corresponding drawing commands, such as changing the color or brush size.

It communicates with the virtual air canvas system to execute the desired actions based on the decoded commands. The output interpretation component is responsible for rendering the user's drawings based on the recognized gestures. It takes the processed gestures and translates them into actual drawing strokes on the virtual canvas. This component works in real-time, continuously updating the canvas as the user moves their hand and creates drawings in the air.

ISSN: 2583-7346
DOI: https://doi-ds.org/doilink/01.2024-86347548/IJAEAST.2024.01.0001, January 2024, Volume-2, Issue-1, pp.1-12
International Journal of Advances in Engineering Architecture Science and Technology

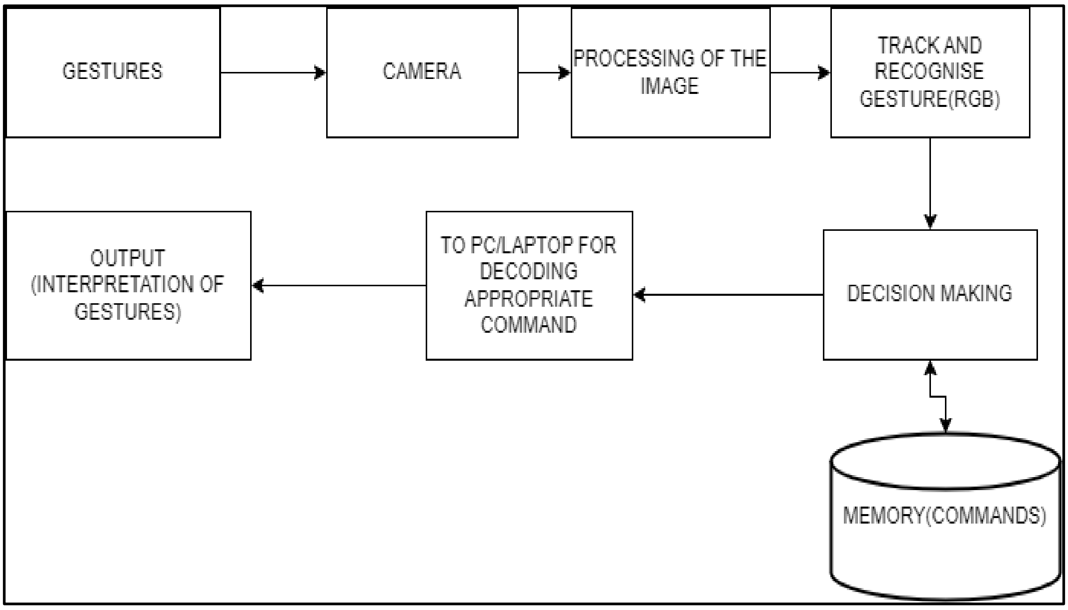

Figure 5. System Architecture

7. Conclusion

The virtual air canvas system has successfully developed an interactive system that allows users to draw in the air using hand movements. By leveraging technologies such as MediaPipe, NumPy, and OpenCV, the project has achieved a seamless and intuitive drawing experience, eliminating the need for physical tools and opening up new possibilities for creative expression. The implementation of hand tracking using MediaPipe ensures accurate detection and real-time tracking of the user's hand. This enables precise interpretation of hand movements, facilitating the creation of drawings on the virtual canvas. Gesture recognition further enhances the user experience by recognizing specific hand gestures, such as closing the hand into a fist to initiate drawing and opening the hand to stop.

Additionally, gestures can be used to change colors and brush sizes, providing versatility in the drawing process. Real-time visualization plays a crucial role in the virtual air canvas system, providing immediate visual feedback to users. The continuous update and rendering of the virtual canvas using OpenCV allows users to see their drawings as they create them, enhancing the interactive nature of the system. The project demonstrates the potential of computer vision techniques and gesture recognition to create an immersive and interactive drawing experience. It opens up new possibilities for artistic expression, interactive digital art, and educational tools. Further advancements in the system could include refining hand tracking accuracy, expanding the range of recognized gestures, and implementing advanced features like 3D drawing or collaborative drawing experiences.

The potential applications extend beyond the presented use cases. Its adaptability, precision, and user-friendly nature pave the way for innovative solutions in interactive digital environments. It successfully showcased the power of combining computer vision, gesture recognition, and real-time visualization to create an engaging and intuitive drawing experience. The research provides a foundation for future developments and applications in the fields of interactive digital art and human-computer interaction, offering exciting opportunities for creative expression and interaction.

References

1. 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, https://doi.org/10.1109/ICRTIT.2016.7569563.
2. 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, https://doi.org/10.1109/CVPRW.2016.53.
3. Guo-Zhen Wang, Yi-Pai Huang, Tian-Sheeran Chang, and Tsu-Han Chen, "Bare Finger 3D Air-Touch System Using an Embedded Optical Sensor Array for Mobile Displays", Journal of Display Technology, Vol. 10, No. 1, January 2014, pp.1, https://doi.org/10.1109/JDT.2013.2277567.
4. EshedOhn-Bar, Mohan Manubhai Trivedi, "Hand Gesture Recognition In Real-Time For Automotive Interfaces", IEEE Transactions on Intelligent Transportation Systems, Vol. 15, No. 6, December 2014, pp 2368-2377, https://doi.org/10.1109/TITS.2014.2337331.
5. Saira Beg, M. Fahad Khan and Faisal Baig, "Text Writing in Air", Journal of Information Display Volume 14, Issue 4, 2013, https://doi.org/10.1080/15980316.2013.860928.
6. Yusuke Araga, Makoto Shirabayashi, Keishi Kaida, Hiroomi Hikawa, "Real Time Gesture Recognition System Using Posture Classifier and Jordan Recurrent Neural Network", IEEE World Congress on Computational Intelligence, Brisbane, Australia, 2012, https://doi.org/10.1109/IJCNN.2012.6252595.
7. TS Angel, GS Kumar, VM Sehgal, G Nayak, "Effective bug processing and tracking system", Journal of Computational and Theoretical Nanoscience 15 (8), 2604-2606, 2018
8. John T Mesia Dhas, "The Functional and Storage Risks Associated to the Size Estimation of Parallel Computing Applications", Advances in Parallel Computing, 40, 373-379, 2022, doi:10.3233/APC220052
9. Badugu Srinivasu, Bellamkonda Meghana, "Design & Development of Emotion Recognition for Speech Data", IJAEAST, 1(3), June 2023, doi: https://doi-ds.org/doilink/06.2023-34882542/0004062023