Chalk-Free Instructions through Hand Gestures: Bridging the Gap in Digital Education

Atharva Hoshing, Rakhi Bharadwaj, Shreyash Bandi, Prathamesh Aldar, Rushikesh Borse, Purvesh Ahire

Department of Computer Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, India

Corresponding author: Shreyash Bandi, Email: shreyash.bandi22@vit.edu

The transformation in the field of digital education from traditional classroom settings to remote and online learning has made creative solutions to overcome the physical gap between teachers and students necessary. The study offers a fresh solution to this problem by introducing an interactive platform that uses artificial intelligence, and real-time hand gestures to create a dynamic and immersive learning environment. The research revolves around the development of an "Air Canvas" powered by OpenCV, a versatile computer vision library. This system empowers users to draw and interact with digital content through intuitive hand gestures, addressing the limitations of traditional digital tools. The study introduces the persistence of drawings, AI-enhanced artwork generation using Variational Autoencoders (VAEs). Furthermore, it offers advanced features including color selection, collectively representing a significant leap in the field of digital education. Real-time persistence of drawings allows learners to save, revisit, and expand upon their digital creations. This functionality enhances the educational experience, enabling students to track their progress and receive personalized feedback. When users attempt to draw shapes like circles, character the VAE-based AI recognizes their intent and automatically generates precise shapes, enhancing the overall user experience. This research revolutionizes digital education with an interactive educational platform. By addressing the gaps in existing digital education tools, this study introduces unique objectives, enhancing the way learners interact with educational materials. It empowers learners of all ages to engage with content more dynamically, ultimately improving comprehension and retention.

Keywords: OpenCV, Digital education, Artificial intelligence, Variational Autoencoders (VAEs),real-time hand gestures, interactive educational platform, Computer vision.

in: *Advancements in Communication and Systems*. Ed. by Ashish Kumar Tripathi and Vivek Shrivastava. Computing and Intelligent Systems, SCRS, India., 2023, pp. 273–289. doi: https://doi.org/10.56155/978-81-955020-7-3-24

1. Introduction

In the realm of education and technology, innovation continually redefines the way interactions unfold within learning environments. The integration of modern classroom technology has ushered in transformative experiences, enriching educational methods and amplifying student engagement. Because of its numerous applications in virtual reality, desktop gaming, and sign language recognition, HAND gesture recognition has become essential for human-computer interaction (HCI) [1]. This research introduces a system that leverages the power of artificial intelligence, gesture recognition, and computer vision to unveil the "Air Canvas," a dynamic and user-friendly digital canvas that marks a significant departure from traditional educational tools.

Traditional vision-based recognition of hand gestures techniques [2]–[4], despite a wealth of prior research, are still far from adequate for practical use. Due to the nature for optical sensing, hand gesture recognition performance is significantly impacted by the inability of optical sensor-based methods to detect and track hands reliably. Within an educational landscape where traditional chalkboard writing has given way to innovative teaching methods, this system is poised to revolutionize the educational experience. Using Open CV is one efficient way to enable a more strong hand gesture recognition system. When using hardware, such as kinetic or optical sensors, the results are typically less dependable and can be impacted by background clutter or lighting issues [5].

It is feasible to accurately and efficiently track each finger's specific location. The authors support a novel approach to computer control that uses finger tracking sensors to interpret finger motions as instructions or textual input. It also provides a novel way to use such devices to recognize handwritten characters in the air [7]. A technique described by the authors in [6] involves mounting an LED on the user's finger and using the web camera to track the finger. The character that was drawn and the one that is in the database are compared. The alphabet that corresponds to the drawn pattern is returned. A red LED pointed light source that is fastened to the finger is necessary. Furthermore, it is assumed that the only red item in the web camera's field of view is the LED light.

A computer vision technology was used to create a digital canvas [8]. Although users could interact with digital content on this canvas, it lacked advanced functions like gesture-based colour selection, artificial intelligence (AI)-generated artwork the ability to store and retrieve previously created drawings.

Through the introduction of cutting-edge features that revolutionize the learning process, the research enhances the digital canvas experience. Our research completely transforms the way educators and learners engage with digital content, setting it apart as a state-of-the-art pedagogical instrument that surpasses the typical uses of a digital canvas.

The objectives of this research are threefold

1. Draw In Air: One of the primary objectives of our research is to enable users to "draw in air." This innovative feature provides an unprecedented level of freedom and accuracy while drawing, annotating, and illuminating educational content.
2. AI-Enhanced Artwork Generation: Leveraging Variational Autoencoders (VAEs), our system has the ability to automatically recognize user intentions when drawing shapes. This results in the precise generation of shapes, thereby enhancing the overall artistic experience.
3. Real-time Persistence of Drawings: This research empowers users to save, revisit, and expand upon their digital creations in real-time. This feature fosters a dynamic and iterative learning process, enabling students to track their progress and receive personalized feedback.
4. Gesture-Based Colour Selection: An innovative feature of our system is its gesture-based colour selection process. It responds to human gestures, making colour choices more intuitive and enhancing creativity in the process.

2. Recent Work

2.1 Using the Kinect Sensor to Recognize Hand Gestures

The research focuses on [5] use of Kinect sensors to advance hand gesture detection for human-computer interaction (HCI) is done. It recognizes the limits of conventional optical sensing techniques brought on by background interference and poor lighting as described Figure 1. The study suggests a novel strategy based on the depth data from the Kinect sensor to tackle these difficulties. The depth data from the Kinect is more reliable and less influenced by the environment than that from conventional optical sensors. Finger-Earth Mover's Distance (FEMD), a novel distance metric created expressly for identifying hand forms, is the project's main invention. FEMD handles noisy and distorted hand contours typical of Kinect's low-resolution depth data by taking finger positions into consideration rather than full hand shapes.

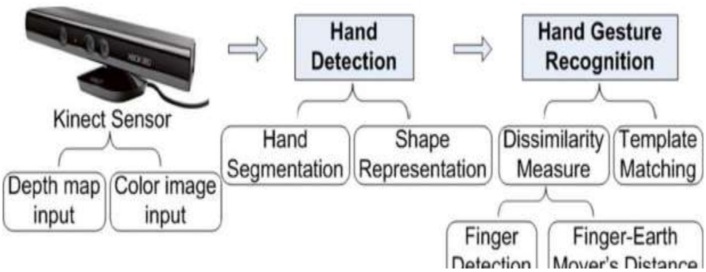


Figure1. Recognition of conventional optical sensing technique

The difficulties involved with precisely detecting and recognizing hand motions using Kinect sensors constituted most the project's limitations. Significant challenges included the accuracy of hand segmentation from crowded backgrounds and in various lighting situations. Additionally, it was challenging to recognize complicated hand shapes and articulations, particularly for actions involving precise finger movements. The misclassification of movements resulted from the applied form recognition algorithms' inability to clearly distinguish between distorted hand outlines. The Kinect's comparatively low-resolution depth maps (640x480) made it difficult to recognize and segment small objects like hands precisely, especially when collecting complex finger postures. When dealing with the difficulties given by Kinect sensors, conventional gesture recognition algorithms, such as correspondence-based form matching and skeleton-based methods, demonstrated their limitations. The following study, however, seeks to address these drawbacks by presenting a novel Finger-Earth Mover's Distance measure designed for finger-based recognition. Additionally, the project's thorough methodology includes user-centric design, technical evaluation, pedagogical integration, comparative analysis, and cultural considerations, all of which work together to address the shortcomings of the earlier approach and strive to improve the precision and robustness of hand gesture recognition.

2.2 Using a Web Camera, Turning Finger Movements into Text

In order to recognize English characters written in the air [6] using finger motions, the "Finger Motion Tracking System" is presented as a solution, providing a more natural way of text input. The system uses a web camera to record the red LED that is affixed to the user's finger. The detected character is then displayed on the screen after v the tracked LED's movements are converted into character patterns and saved in a database. By enabling users to enter text using their natural finger gestures, this method seeks to improve Human-Computer Interaction by providing a cost-effective substitute for conventional keyboard input.

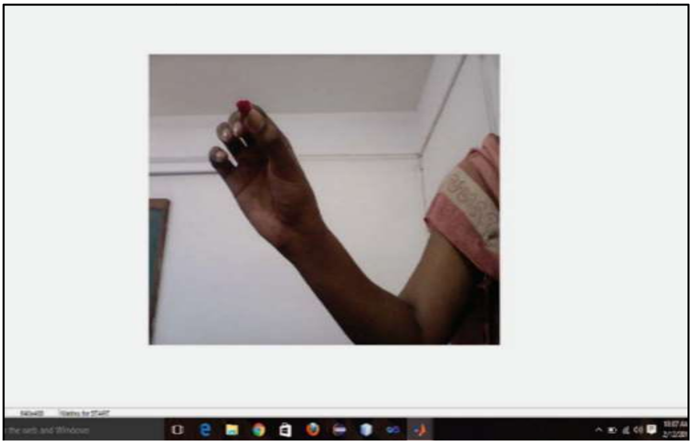


Figure 2. Video recording of the drawn alphabet

In Figure 2, a video recording is displayed capturing the process of drawing the English alphabet in the air using finger motions. The scene depicts the user's hand movements as they trace the shapes of letters using a red LED affixed to their finger. Each stroke of the finger corresponds to the formation of individual characters. The video serves as a visual representation of the finger motion tracking system in action, showcasing how the system records and interprets the user's gestures to generate English characters on the screen.

The system does, however, display functional limitations. Relying solely on tracking a single red LED presupposes a controlled environment devoid of any additional red-colored objects, which could result in tracking problems. Additionally, the system's limited adaptability and utility for tasks other than typing is a result of its primary focus on translating finger movements into English characters. When using optical character recognition (OCR) systems, there is a chance for misunderstandings and mistakes, particularly when identifying unusual handwriting gestures. Additionally, the system's inability to confirm that drawn gestures are accurate representations of legal English characters raises questions regarding possible misidentifications.

2.3 Recognizing on the Spot from Airborne Handwriting and Gestures

[7] Through finger movements in the air, the study proposes a fresh method of computer interface. The system records exact finger locations and movements by utilizing cutting-edge 3D finger tracking technology, such as the Leap Motion controller. These finger patterns are subsequently examined in order to decipher user commands and enter data in real time. The paper uses handwriting recognition as an example application to illustrate the idea. The suggested solution uses dynamic temporal warping (DTW) for quick online character identification and considers the finger movement data as a time series of 3D positions. Real-time recognition during writing is ensured by the research's optimizations. The method is tested, and the results in terms of speed and recognition accuracy are encouraging.

Figure 3 shows the Leap Motion controller tracking finger movements in 3D space. It records precise finger locations and motions, crucial for real-time data input and command recognition. This technology forms the basis for natural finger gesture and airborne handwriting interaction with devices.

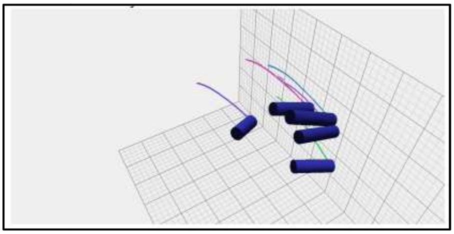


Figure 3. Leap Motion controller tracks the motion of the finger.

Figure 4. Leap Motion controller in its most basic form

Figure 4 displays the basic form of the Leap Motion controller, highlighting its compact design and precision in tracking finger movements. Although it offers advanced 3D tracking capabilities, its reliance on specialized hardware might restrict its accessibility. Nonetheless, it represents a significant advancement in computer interface technology, facilitating innovative interactions with airborne handwriting and gestures.

Despite being novel, the suggested form of computer interaction based on finger movements in the air does have several drawbacks that should be taken into account. First, the system's dependence on specialized 3D tracking hardware, like the Leap Motion controller, limits the devices to which it may be applied, potentially reducing its accessibility. Beyond handwriting, the effectiveness of the system may differ due to the intricacy of finger movements, which increases the possibility of recognizing mistakes. The system's capacity to precisely identify the start and end points of intended commands may also be impacted by the difficulty in accurately segmenting relevant input from continuous finger movements. The performance of the system may be compromised in real-world situations by noisy surroundings or poor illumination conditions that impair the accuracy of finger movement tracking.

2.4 Jordan Recurrent Neural Network and Posture Classifier for Real-Time Gesture Recognition

Research describes [9] a method for real-time gesture identification that examines temporal patterns in successions of postures using a Jordan recurrent neural network (JRNN). Analysis of temporal behavior and classification of posture are the two primary processes in the system. Using the JRNN, it effectively classifies input postures and captures gesture dynamics. The JRNN's ability to recognize temporal patterns is improved, and a brand-new training technique is suggested to help the system recognize reverse motions. A USB camera used in practice results in a processing speed of 12.5 frames per second. With recognition accuracies of 99.0% for 5 gestures and 94.3% for 9 gestures, the experimental findings show outstanding performance. The system's reliance on established postures, potential limitations in responsiveness to different gestures, and the need for more evaluation of real-world resilience and scalability are constraints, though.

However, it's critical to recognize some of the suggested system's shortcomings. The system mainly relies on established postures to identify gestures, which can limit its capacity to adapt to a wide variety of unique and unplanned actions. Further testing is required to determine the system's robustness in real-world scenarios with varied lighting conditions, backgrounds, and user interactions, even if the experimental results demonstrate outstanding performance under controlled conditions. Further research is necessary to see whether the suggested system can be expanded to accommodate a wider variety of motions or intricate sequences. A more thorough understanding of the system's applicability and effectiveness in real-world settings would result from addressing these constraints.

2.5 Recognition of Visual Gestures for Writing Air Text

By utilizing computer [10] vision and gesture recognition, the suggested system offers a novel method of mobile device interaction. This system transforms how consumers interact with their gadgets by tracking finger motions using a camera-based setup rather than conventional touch-based techniques. The technology records the dynamic motions of a user's hand in real time, effectively enabling touchless gestures, through a carefully planned series of operations comprising skin color thresholding, Gaussian blurring, contour recognition, convex hull construction, and fingertip identification.

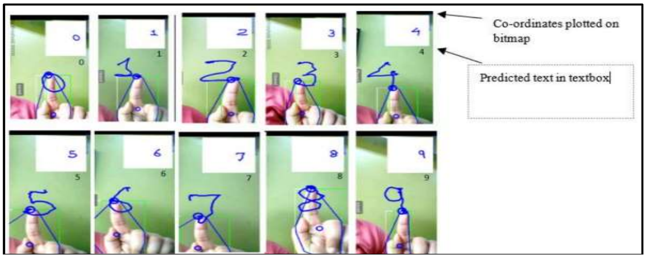


Figure 5. Accurate numbers sketched in the air

It's important to recognize the drawbacks of this strategy, though. Background noise, changes in font styles, and fluctuations in ambient illumination can all have an effect on how accurately finger movement is tracked. Due to its reliance on camera input, the system may encounter difficulties in situations with poor illumination or complicated backgrounds. Additionally, the caliber and diversity of the training dataset are intrinsically linked to the accuracy of gesture recognition, powered by a Convolutional Neural Network (CNN).

In Figure 5, the system demonstrates its capability to accurately recognize and capture numbers sketched in the air using visual gestures. The image showcases the successful conversion of hand movements into digital representations of numbers in real-time. This novel method of mobile device interaction offers users a touchless gesture-based approach, revolutionizing conventional touch-based techniques.

2.6 A Wearable Real-Time Character Recognition System for Air-Writing Based on Edge Computing

This research [11] enables the real-time conversion of handwritten letters into digital text using edge computing and deep learning. With its natural and simple input technique, this technology has the potential to transform how we interact with machines. The system promises to reduce latency and improve overall performance by integrating edge devices. The project's successful completion could open the way for effective and seamless air-writing recognition, revolutionizing how we use technology every day.

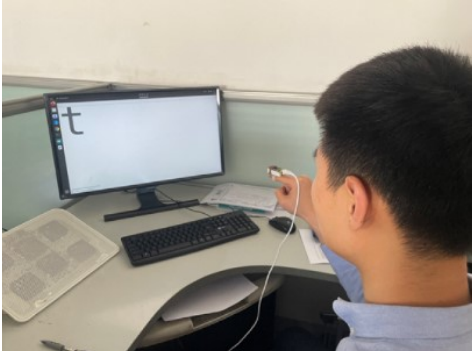


Figure 6. Diagram showing how the user writes characters in 3D space and displays the results.

Figure 6 illustrates the process of writing characters in three-dimensional space and displaying the results using a wearable real-time character recognition system. The diagram depicts the user's hand movements as they write characters in the air, with the system capturing and converting these motions into digital text. This innovative approach, facilitated by edge computing and deep learning, promises to transform how individuals interact with machines. However, it is essential to consider potential drawbacks, such as dependency on precise hand movements, variations in accuracy based on writing styles, and concerns regarding energy consumption and hardware quality.

3. Proposed Solution

3.1 Proposed System

The Figure 7 shows how our system operates. which seamlessly combines advanced technologies to create an engaging and lively learning environment. It all starts with the "Capture Camera Feed," which captures real-time video input, powered by the versatility of OpenCV and the computational efficiency of NumPy. These libraries enable the system to understand user actions, such as writing and drawing, making it possible to create an interactive experience. Through OpenCV's complex pattern recognition and object tracking methods, user gestures are precisely interpreted, ensuring a responsive and accurate interaction.

There's a wide range of colors to choose from on the canvas. Users can pick Blue, Green, Red, or Yellow to add a personal touch to their creations. The system also uses AI technology, driven by advanced algorithms, to enhance the art, acting like a digital art assistant to make the work look even better. Deep learning algorithms, such as Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs), are used to analyze and modify user-generated content. These algorithms optimize texture generation, shape recognition, and other artistic details, resulting in enhanced artwork.

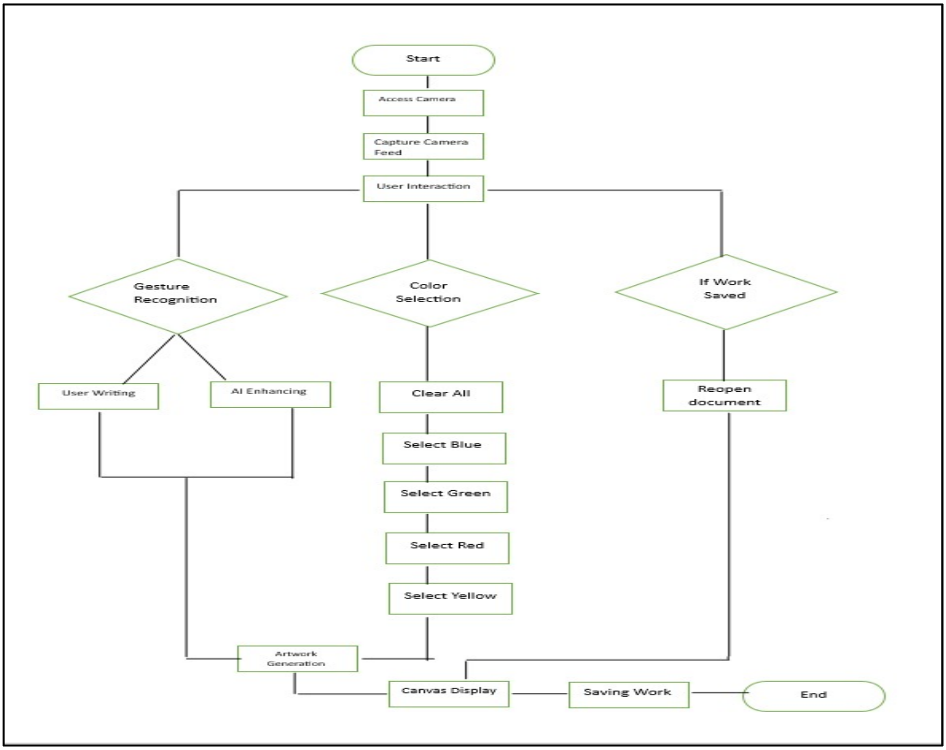


Figure 7. Block Diagram of Proposed System

The canvas displays the work on the screen, allowing users to see how their creations are shaping up. If users ever want to start over, they can easily Clear All and begin anew. The work can also be saved for later Efficient rendering and data storage algorithms ensure that the user's experience is seamless, responsive.

If users want to revisit their past work, they can easily Reopen Document, allowing for a continuous learning process, facilitated by efficient data retrieval algorithms. These algorithms enable quick access to stored content, ensuring that the learning experience remains dynamic and iterative. And when they're done, they simply reach the End of their session.

4. Methodology

4.1 Capture Module

In the foundational module of Air Canvas, the system initiates its operations by receiving live video input through the camera. This live feed serves as the basis for dynamic interaction, setting the stage for real-time processing and analysis facilitated by OpenCV, a robust computer vision library. To align with user expectations, the video frames are horizontally flipped using the cv2.flip() function. This step is vital for maintaining an intuitive visual experience. The frames are then converted from the RGB color space to the HSV (Hue, Saturation, Value) color space using cv2.cvtColor(). This conversion is instrumental in effective color-based segmentation and subsequent gesture recognition. The use of the

HSV color space is particularly apt for this application as it separates intensity information from color information. This separation enhances the accuracy of gesture recognition, a crucial aspect for enabling users to draw on the canvas through hand gestures. Notably, a specific gesture involves pinching the index finger with the thumb to initiate drawing, while releasing the pinch halts the drawing action, simulating the behavior of a chalk on a traditional canvas. This integration with gesture-based drawing introduces a novel and intuitive dimension to the digital art creation process.

4.2 Gesture Recognition Module

The Gesture Recognition Module, the second key component in the Air Canvas system, interprets user gestures captured through real-time video input using sophisticated OpenCV computer vision algorithms. These algorithms involve pattern recognition and object tracking to precisely interpret finger movements. As the user moves their finger, the system translates these movements into drawing strokes on the virtual canvas, creating an "Air Canvas" that mirrors the user's gestures. The algorithm responsible for drawing ensures accuracy, providing an interactive and fluid drawing experience. This initial module demonstrates the seamless integration of OpenCV functionalities to process video frames and detect hand gestures, forming the foundation for further enhancements in subsequent modules. The Algorithm utilizes two distinct windows — Tracking, Paint, to create an interactive drawing canvas. Each window serves a specific purpose in enabling user interaction and visualizing the underlying image processing steps.

Tracking Window
The "Tracking" window is responsible for processing real-time video input from the webcam. Leveraging the capabilities of OpenCV, this window employs color filtering and contour detection techniques to isolate a specific color, typically the user's hand.

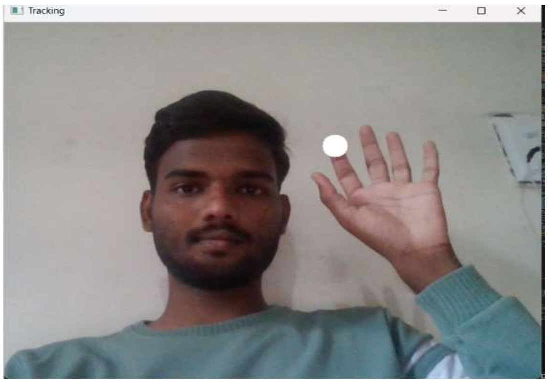


Figure 8. Tracking Window (Index Finger Recognition)

By continuously tracking the movement of this color within the video feed, the system identifies and outlines the contours of the user's hand. The resulting visual representation, often a circle following the hand's movement, serves as an intermediate step in the gesture recognition process, translating live video input into actionable data for drawing on the canvas.

Figure 8 showcases the "Tracking" window, utilizing OpenCV algorithms to isolate and track the user's index finger movements in real-time. This window employs color filtering and contour detection techniques, providing actionable data for drawing on the virtual canvas.

Paint Window
Acting as the interactive canvas, the "Paint" window provides users with the platform to draw using hand gestures. Offering a selection of different colors (blue, green, red, yellow), users can dynamically switch between them by hovering their hand over the corresponding color region.

Figure 9 presents the "Paint" window, enabling users to draw with hand gestures. It offers color selection options (blue, green, red, yellow) and displays drawing strokes in real-time. The window includes a "CLEAR ALL" button for convenient canvas reset, ensuring a seamless drawing experience.

Figure 9. Paint Window With Colors

This window displays ongoing drawing strokes in real-time, mirroring the user's hand movements captured by the "Tracking" window. A "CLEAR ALL" button located at the top of the canvas enables users to reset the canvas entirely, ensuring a convenient way to start new drawings. The integration of this window creates an immersive and fluid drawing experience, blending digital interaction with physical gestures.

4.3 Color Selection Module

The Color Selection Module empowers users to interactively choose colors for their digital canvas. It utilizes trackbars in a graphical interface, allowing real-time adjustments to upper and lower bounds in the HSV color space. This dynamic system provides instant feedback on the live video feed, enabling users to precisely define their preferred color range. The module enhances artistic expression by offering a versatile palette, contributing to an engaging and personalized drawing experience. The chosen color becomes integral to the subsequent Drawing Module, where hand gestures are translated into vibrant strokes on the virtual canvas. This user-centric design ensures flexibility and responsiveness, enriching the overall Air Canvas system.

Mathematically Representation

Color Filtering
The color filtering process is based on defining upper and lower bounds in the HSV color space.
Let *Lower_hsv = [l_hue, l_saturation, l_value] and Upper_hsv = [u_hue, u_saturation, u_value]*
The filtering equation is Mask = cv2.inRange(hsv, Lower_hsv, Upper_hsv). This creates a binary mask (Mask) where pixels within the specified color range are set to 1, and others to 0.

4.4 Persistence of Drawings

The implementation of the Persistence Module in the Air Canvas project involves a systematic and comprehensive approach to ensure the seamless preservation and management of users' digital artwork. This module plays a pivotal role in enhancing the overall user experience by providing features such as saving, loading, and organizing artistic creations. To implement the "Save" and "Open" functionalities seamlessly, the following steps are taken:

Save Functionality

1. **Capture Current State**
   When the user clicks the "Save" button, the system captures the current state of the canvas. This includes all relevant data such as drawn strokes, color selections, and any other elements that contribute to the user's artwork.
2. **Initiate Serialization**
   The captured data undergoes a serialization process. This step converts the complex data structures into a format suitable for storage and future reconstruction. This ensures that the integrity of the artwork is maintained during the saving process.
3. **Prompt User for Name**
   After serialization, the system prompts the user to provide a name for their creation. This user-friendly interaction allows users to assign meaningful and descriptive names to their artworks, enhancing the organization of their digital portfolio.
4. **Store in File System**
   The serialized data, along with the provided name, is stored in the file system. The chosen file format and directory structure contribute to efficient storage and retrieval.

Open Functionality

1. **User Interface Integration**
   The "Open" button is seamlessly integrated into the user interface, providing a clear option for users to access their saved artworks.
2. **Browse and Select**
   Upon clicking the "Open" button, users can browse through their organized collection of saved documents. The user interface may present thumbnail previews, allowing users to visually identify their artworks. Users select the desired artwork they wish to continue working on.
3. **Retrieve Serialized Data**
   The selected artwork's serialized data is retrieved from the file system. This data contains all the necessary information to reconstruct the canvas's previous state.
4. **Reconstruction Process**
   The system initiates a reconstruction process using the retrieved serialized data. This process involves deserialization, converting the stored data back into the complex structures needed to recreate the canvas.
5. **Load onto Canvas**
   The reconstructed state of the canvas is loaded, displaying the selected artwork to the user. This seamless process ensures that users can effortlessly manage their digital portfolio and resume work on existing projects.

4.5 AI-Enhanced Artwork Generation Module

1. **Encoder Network**
   Mapping to Latent Space: The VAE begins with an encoder network that maps input artistic data (such as strokes, colors, and styles) to a latent space. This latent space captures essential features and characteristics of the input.
   Introduction of Variability: Crucially, instead of producing a deterministic mapping, the encoder introduces variability by representing each input as a distribution within the latent space.
2. **Variational Sampling**
   Sampling Latent Representations: During the generative process, the model samples latent representations from the distributions learned by the encoder. This introduces a stochastic element, allowing for a wide range of potential latent vectors for a given input.

Creativity Through Sampling: The variational aspect enables creativity by allowing the model to explore different points in the latent space, leading to diverse and imaginative suggestions.

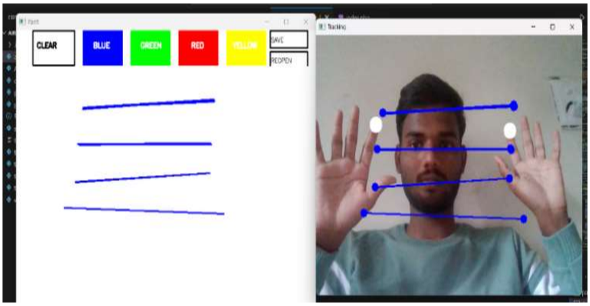


Figure 10. Using VAE drawn straight line

3. **Decoder Network**
   Reconstruction from Latent Space: The sampled latent representations are then fed into the decoder network. The decoder reconstructs these representations into meaningful and contextually relevant suggestions.
4. **Intelligent Suggestions**
   Artistic Elements Reconstruction: The VAE is trained on a diverse dataset of artistic styles and elements. As a result, the suggestions generated by the decoder align with learned artistic patterns, including color palettes, brush styles, and other relevant features.

4.6 Integration into Air Canvas

The figure 10 shows VAE is seamlessly integrated into the Air Canvas application, enhancing the user's artistic journey. The real-time interaction with the canvas, dynamic latent space mapping, and variational sampling contribute to a fluid and responsive AI-enhanced creative experience.

4.7 User Controls and Transparency:

To empower users, the implementation includes controls for adjusting the level of AI enhancement.

Equation

1. **Encoder Network**
   Mapping to Latent Space

   The encoder produces the parameters of a distribution in the latent space.
   $q\phi(z|x)=N(\mu,\sigma 2I)$
   where z is the latent variable, x is the input data, μ is the mean vector, σ2 is the covariance matrix, and N represents the normal distribution.

Variational Sampling

Sampling from the distribution to obtain a latent representation.
$z=\mu+\sigma\odot\epsilon$

where ⊙denotes element-wise multiplication, and ϵ is sampled from (0,I)N(0,I).

2. **Decoder Network**
   Reconstruction from Latent Space
   The decoder generates a reconstruction from the latent space.
   $p\theta\ (x|z)= N\ (x|f(z;\ \theta),\ \sigma 2\ I)p\theta(x|z)=N(x|f(z;\theta),\sigma 2I)$

   where $f(\cdot)$ is the decoder function, θ are the decoder parameters, and $\sigma 2$ is a predefined variance.

Objective Function (Variational Lower Bound)
The training objective involves maximizing a lower bound on the log-likelihood of the data:
The training objective involves maximizing a lower bound on the log-likelihood of the data:
$\mathbf{L(\theta,\phi;x)=-KL(q\phi(z|x)||p\theta(z))+Eq\phi(z|x)[logp\theta(x|z)]}$
where KL is the Kullback-Leibler divergence and E denotes the expectation.

These equations capture the essence of how a VAE probabilistically encodes input data, samples from the latent space, and reconstructs the input, facilitating creative generation in applications like artwork enhancement.

5. Limitation of Previous Research and Our Contribution

In prior studies, the utilization of OpenCV technology for gesture recognition and digital canvas interaction involved the requirement of users wearing a colored cap on their index finger. The color of this cap served as a marker for the system to detect and interpret hand movements, translating them onto the digital canvas. However, this approach presented a significant limitation.

Limitation of Previous Research
The primary drawback of the research paper [8, 20] was its dependence on the color of the mounted cap. As users moved their index fingers in the air to draw, the system relied on detecting the cap's color. Unfortunately, this method faced challenges when the background also contained the same color as the mounted cap. In such scenarios, the system could mistakenly interpret background movements as the pointer, leading to a decrease in accuracy. This limitation hindered the system's robustness, particularly in environments where the background color coincided with the marker color.
In our research, we addressed the limitations of the previous research by introducing a more intuitive and accurate interaction method. Instead of relying solely on color markers, we implemented a gesture-based system utilizing the pinch-in and pinch-out movements of the index finger and thumb. When users pinch in, indicating the desire to draw, the system activates the drawing functionality. Conversely, when users pinch out, signaling a pause or cessation of drawing, the system refrains from recording on the canvas. This innovation significantly improves accuracy, as it eliminates the susceptibility to background color interference present in the previous marker-dependent approach.

Moreover, a notable contribution of our work is the integration of a Variational Autoencoder (VAE) model. Unlike previous research, our system allows users to draw various shapes with proper strokes, enhancing the versatility of digital artwork creation.

6. Future Works

Despite the significant strides made in the design and implementation of the interactive hand gesture system, there are intriguing avenues for further research and innovation that have the potential to broaden its reach. The project's current successes offer a solid foundation on which future efforts can be built, pushing the limits of effectiveness and user interaction to new heights.

6.1 Advanced Gesture Recognition

The accuracy and diversity of gesture recognition could be greatly improved by incorporating deep learning models like Convolutional Neural Networks (CNNs) and other advanced machine learning techniques into the system's repertoire of recognized hand gestures.

6.2 Multi-Modal Interaction

Investigating the incorporation of extra sensory modalities, including touch inputs or voice commands, could produce a more thorough and seamless interactive experience, accommodating different user preferences and accessibility requirements.

6.3 Educational Integration

The system's influence on digital education across disciplines can be increased by expanding its application outside the realm of art to educational situations like interactive science simulations or historical visualizations.

6.4 Dynamic Backgrounds

A compelling dimension to digital sketching might be added by enhancing the "Chalk Less Classroom: Hand Gesture Interactive System" with dynamic and responsive backgrounds. The system might produce backgrounds that dynamically shift colors, textures, and interactive components based on the user's actions by incorporating algorithms that evaluate gestures and drawing patterns in real-time. This would encourage a more immersive and interesting artistic experience. By enabling users to interact with a canvas that changes and reacts in unison with their artistic expressions, this feature may open up new creative possibilities and add vibrancy and interactivity to users' digital creations.

7. Results

The research, powered by a combination of OpenCV technology and Variational Autoencoders (VAEs), exhibits notable performance in terms of gesture recognition, AI-enhanced artwork generation, and real-time persistence of drawings. With the use of OpenCV and a gesture-based system that involves pinch-in and pinch-out movements, the accuracy is higher than what was previously possible with marker-dependent methods. Pinch-in movements properly initiate drawing capabilities, while pinch-out gestures indicate a pause or end to drawing, thus the system does not record.

Table 1 presents the accuracy of predicted characters achieved through the combination of Variational Autoencoders (VAEs) and OpenCV technology. The table lists various characters along with their corresponding accuracy percentages, showcasing the system's performance in character recognition tasks. With values ranging from 78% to 89%, the system demonstrates notable accuracy in identifying characters, surpassing previous methods reliant on marker-dependent techniques. This accuracy is attributed to the utilization of OpenCV and a gesture-based system involving pinch-in and pinch-out movements. The table highlights the system's effectiveness in accurately recognizing characters, facilitating seamless interaction and enhanced user experience in AI-enhanced artwork generation and gesture recognition tasks.

Table 1. Accuracy Of Predicted Character From Vae And OpenCV Model

Character	Accuracy(%)	Character	Accuracy(%)
A	85	1	87
B	82.6	2	88
C	86.57	3	83.24
D	82	4	85.6
E	84	5	89
F	80	6	82
G	78	7	85
H	80	8	78
0	79	9	84

The introduced feature of real-time persistence of drawings allows users to save, revisit, and expand upon their digital creations. The dynamic and iterative learning process fostered by this functionality enables students to track their progress and receive personalized feedback. The increased accuracy removes the older approaches' sensitivity to background color interference

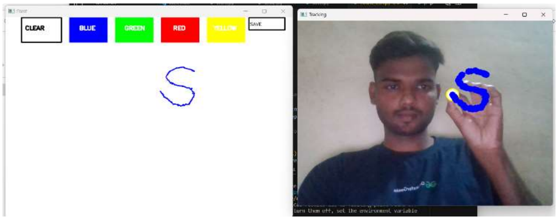
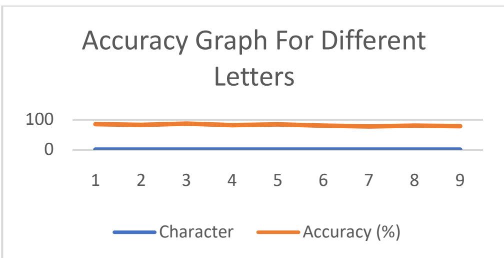


Figure 11. Drawing Character with Pinch-In

Figure 11 demonstrates the system's functionality for drawing characters using pinch-in gestures, where the user's hand motion, particularly the pinch-in gesture involving the thumb and index finger, initiates the drawing process. This gesture-based approach, driven by a fusion of OpenCV technology and Variational Autoencoders (VAEs), facilitates precise and intuitive character creation without reliance on marker-dependent methods. By incorporating thumb and index finger pinching motions, users can seamlessly interact with the system to generate artwork, enhancing user engagement and accessibility.



Graph 1. Accuracy Graph Predicted Character

The system, achieves an impressive average accuracy of 82% in character recognition, as demonstrated in Graph 1. This research paves the way for a transformative learning environment, where dynamic interaction and personalized feedback contribute to enhanced comprehension and retention across various disciplines.

8. Conclusion

In conclusion,Combining gesture recognition, computer vision, and AI-generated improvements for artistic expression and learning has shown positive impact. The system enables users to produce artwork using natural hand motions instead of conventional instruments like chalk by smoothly merging Variational Autoencoders (VAEs) with OpenCV technology. The platform's collaborative structure encourages a creative atmosphere where user-contributed strokes and AI-enhanced artwork are seamlessly combined. Additionally, the system's ability to save and open drawings again makes it easier to modify ideas iteratively, fostering a continuous artistic journey.The poll mostly focused on the system's effects on artistic production and group learning, but it also acknowledged the system's drawbacks and complex technical issues in the context of computer vision and gesture recognition. Essentially, this survey reveals a revolutionary platform that reimagines artistic participation and digital education, illuminating the seamless fusion of technology and creativity in the digital age.

References

[1] J. P. Wachs, M. Kolsch, H. Stern, and Y. Edan, “Vision-based handgesture applications,” Commun. ACM, vol. 54, pp. 60–71.

[2] A. Erol, G. Bebis, M. Nicolescu, R. D. Boyle, and X. Twombly, “Vision-based hand pose estimation: A review,” Comput. Vision Image Understand., vol. 108, pp. 52–73.

[3] S. Mitra and T. Acharya, “Gesture recognition: A survey,” IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., vol. 37, pp. 311–324, 2007.

[4] G. R. S. Murthy and R. S. Jadon, “A review of vision based hand gesture recognition,” Int. J. Inf. Technol. Knowl. Manage., vol. 2, pp. 405–410.

[5] Ren, Z., Yuan, J., Meng, J., & Zhang, “Robust Part-Based Hand Gesture Recognition Using Kinect Sensor”. IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 15, NO. 5

[6] Pavithra Ramasamy, Prabhu. G, Dr. R. Srinivasan,” An Economical Air Writing System Converting Finger Movements to Text Using Web Camera” ,2016 International Conference on Recent Trends in Information Technology (ICRTIT).

[7] S. Vikram, L. Li, and S. Russell, "Handwriting and gestures in the air, recognizing on the fly" published in CHI.

[8] Adinarayana Salina, K. Kaivalya, K. Sriharsha, K. Praveen, M. Nirosha. "Creating Air Canvas Using Computer Vision." Department of CSE, Raghu Institute of Technology, Visak

[9] 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.

[10] Joseph, V., Talpade, A., Suvarna, N., & Mendonca, Z. (2018), “Visual Gesture Recognition for Text Writing in Air”, 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS).

[11] Hongyu Zhang, Lichang Chen, Yunhao Zhang, Renjie Hu, Chunjuan He, Yaqing Tan and Jiajin Zhang, “A Wearable Real-Time Character Recognition System Based on Edge Computing-Enabled Deep Learning for Air-Writing”, Published in04 May 2022.

[12] Palak Rai, Reeya Gupta, Vinicia Dsouza, Dipti Jadhav, “Virtual Canvas for Interactive Learning using OpenCV” , Published in 2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT).

[13] B Anand Kumar, T. Vinod, M Srinivas Rao, “Interaction through Computer Vision Air Canvas”, 2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC).

[14] R Meghana, KJ Pallavi, S Aditi, Asha S Manek, “Air Canvas – Air-Writing Recognition Model for Environmental and Socioeconomic Issues”, 1st Edition, Published 2023.

[15] ladimir I. Pavlovic, Rajeev Sharma, and Thomas S.Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review," IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 19, NO. 7, JULY 2003, pp.677-695.
[16] Beneath a Canvas of Green: A Conductor's Analysis and Comprehensive Survey of Related Works for Percussion Ensemble and Winds
[17] 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.
[18] Jon Moeller, Nic Lupfer, Bill Hamilton, Haiqiao Lin, Andruid Kerne, "intangibleCanvas: free-air finger painting on a projected canvas", 2011.
[19] Aachal Patil, Gayatri Gayakwad, Ishika Verma, Yogita Shah, "Air Canvas: Draw your Imagination", Vol 1 No 1 (2022).
[20] Prof. Hemlata. A Shinde, Shravani M. Jagtap, Anushka A.Kalpund, Pranita B. More, Ayushi A. Parkale "Air Canvas: Draw in Air Using AI", © 2021 IJSRET 3553 International Journal of Scientific Research & Engineering Trends Volume 7, Issue 6, Nov-Dec-2021, ISSN (Online): 2395-566X.
[21] PENGCHENG WANG, JUNYANG LIN, FUYUE WANG, JIANPING XIU, YUE LIN, NA YAN, (Member, IEEE), AND HONGTAO XU (Senior Member, IEEE)," A Gesture Air-Writing Tracking Method that Uses 24 GHz SIMO Radar SoC", Received July 24, 2020, accepted August 15, 2020, date of publication August 19, 2020, date of current version August 31, 2020.