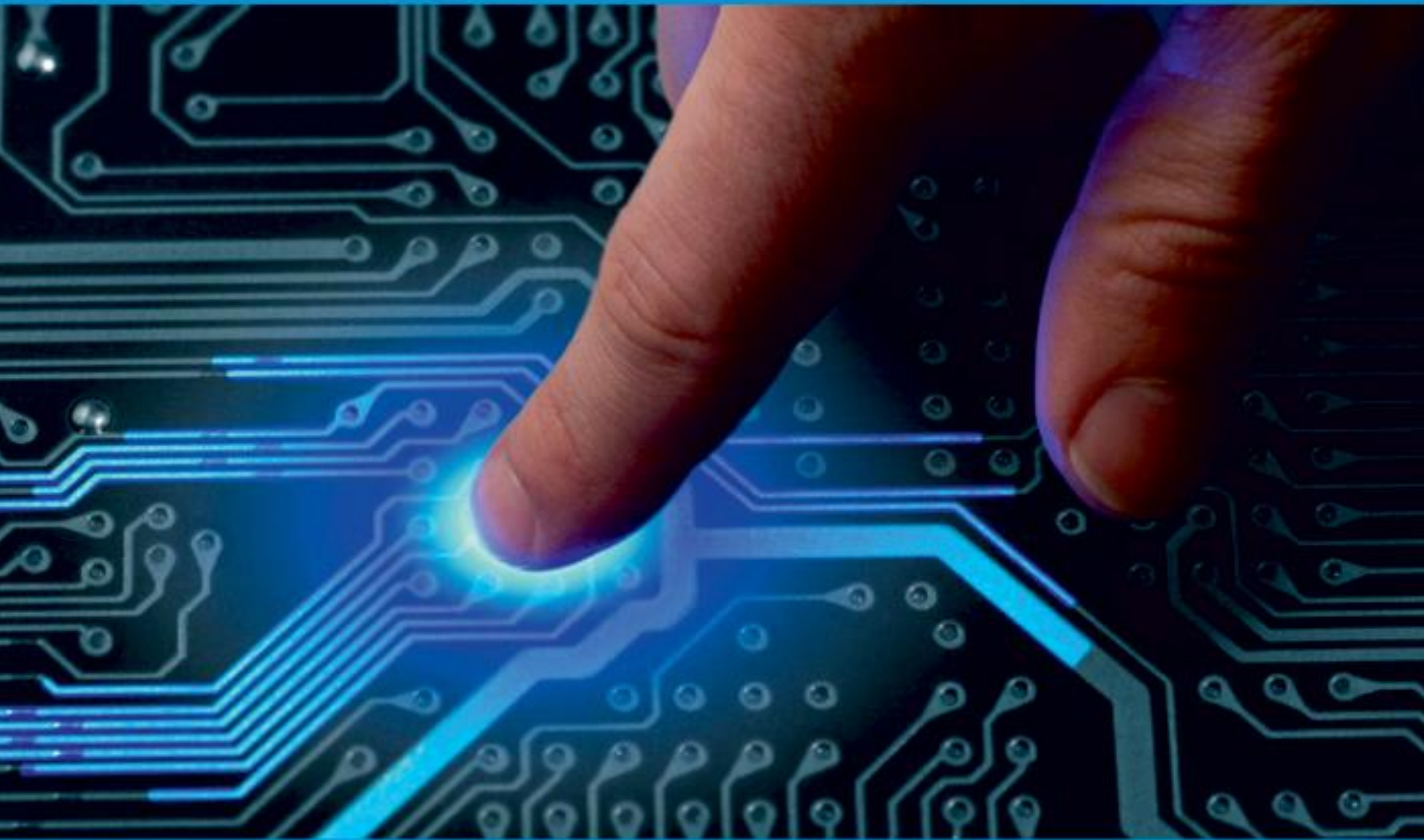
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205277 |

Air Canvas – Writing in air

Dr.(Mrs) S.K. Wagh, Arya Tamhane, Disha Dalvi, Sanvidha Haribhakta, Sakshi Malvadkar

Associate Professor, Department of Computer Engineering, Modern Education Society's College of Engineering, Pune, India

U.G. Student, Department of Computer Engineering, Modern Education Society's College of Engineering, Pune, India

U.G. Student, Department of Computer Engineering, Modern Education Society's College of Engineering, Pune, India

U.G. Student, Department of Computer Engineering, Modern Education Society's College of Engineering, Pune, India

ABSTRACT: The Air Writing Recognition System is a exploration of machine learning and computer vision technology, redefining the landscape of human-computer interaction. This system enables users to convey messages by shaping letters in the air, and its recognition. The primary objective is the accurate recognition of individual letters, both English and Devanagari scripts. The system is created with Python as the primary language, employing TensorFlow and Keras for machine learning models, and integrating MediaPipe for precise hand tracking and detection. The development environment includes widely-used IDEs Jupyter Notebook and Visual Studio Code. This project holds diverse applications, in education and language learning to creative expression and cross-cultural communication.

KEYWORDS: Air-writing, Tenserflow, Keras, Machine Learning, MediaPipe

I. INTRODUCTION

In an era characterized by the ever-changing landscape of interactive technologies, the "Air Canvas Writing in Air" initiative emerges as a pioneering endeavor at the intersection of gesture-based interaction and state-of-the-art projection systems. This visionary project aims to reimagine the traditional canvas by suspending it in the ethereal realm of mid-air, surpassing the limitations of physical touch and introducing a new paradigm for artistic expression. At its essence, the Air Canvas project exemplifies the seamless integration of depth-sensing cameras and advanced motion tracking algorithms, creating a harmonious connection between human gestures and digital strokes. This fusion results in an incredibly responsive and tactile drawing experience that not only blurs the boundaries between the physical and virtual worlds but also sets the stage for an exploratory journey into interactive possibilities.

Delving into the technical complexities, the system's ability to interpret hand movements in three-dimensional space opens up a realm of creative possibilities that extend beyond the conventional boundaries of artistry. As users navigate the virtual canvas through intuitive gestures, the technology translates these movements into a dynamic array of digital strokes, forming a visually captivating tapestry in real-time. The canvas, suspended in mid-air, transforms into a dynamic playground for imaginative minds, offering not only a fresh avenue for artistic expression but also a bridge between the tangible and the digital.

However, the Air Canvas project is not limited to the realm of solitary artistic creation; it serves as a catalyst for collaborative ingenuity. With customizable tools and brushes that enhance creative capabilities, the system encourages multiple users to contribute simultaneously, fostering collaborative creativity in various domains such as art installations, educational environments, and team-based projects.

II. RELATED WORK

The literature review explores gesture-based interaction, including research on algorithms and technologies that enable hand tracking for interactive systems. It also examines virtual canvases and augmented reality systems for artistic expression and educational contexts. The review covers studies on human-computer interaction, with a focus on user experience, usability, and design principles, as well as collaborative technologies and shared virtual spaces. The integration of technology into artistic practices, particularly in digital art and interactive installations, is analyzed, along with advancements in projection systems such as calibration, interactive mapping, and real-time rendering. Finally, the review investigates the impact of virtual canvases on learning experiences and entertainment content in educational and entertainment settings. In [1], We introduce a new method called AirWriting, which allows for the creation, recognition, and visualization of documents in the air. The approach utilizes a unique algorithm called 2-DifViz, which converts hand movements captured by a Myo-armband into x, y coordinates on a 2D Cartesian plane and displays them on a

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |
|| Volume 12, Issue 5, May 2024 ||
| DOI: 10.15680/IJIRCCE.2024.1205277 |

canvas. Unlike existing sensor-based methods that lack visual feedback or rely on fixed templates for character recognition, AirScript sets itself apart by offering users freedom of movement and real-time visual feedback, resulting in a more natural interaction. Additionally, AirScript incorporates a recognition module that utilizes deep learning techniques, leveraging both the sensor data and visualizations generated by 2-DifViz. This module consists of a Convolutional Neural Network (CNN) and two Gated Recurrent Unit (GRU) Networks, with the outputs from these networks combined to predict the characters written in the air.AirScript finds applications in advanced environments such as smart classrooms, smart factories, and smart laboratories, enabling individuals to annotate text anywhere without the need for a reference surface. To evaluate its performance, AirScript was compared against various established learning models (HMM, KNN, SVM, etc.) using data from 12 participants. The evaluation results demonstrate that AirScript's recognition module significantly outperforms these models, achieving an accuracy of 91.7 \% in person-independent evaluation and 96.7 % accuracy in person-dependent evaluation. In[2], We introduce that the computer vision air canvas enables users to interact with material and present it on the screen using an application. The use of varied color schemes enhances the user's ability to identify and understand the information presented. To utilize this technology, access to a built-in or independent web camera is necessary. This technology can be used for text visualization and drawing, providing a stepping stone for future innovative streams of material. By simply moving a finger through the air, creative ideas can be drawn using computer vision technology. Our paper outlines the process of constructing a screen that displays information or text drawn by waving a finger, similar to a touch screen. The objectives of this technology include color detection, marker tracking, and coordinate establishment. In[3], Air-writing refers to the act of writing characters or words in empty space using hand or finger movements. The issue of recognizing air-writing is discussed in two related papers. In Part 2, the focus is on detecting and recognizing air-writing activities that are seamlessly integrated into a continuous motion trajectory without clear boundaries. The challenge lies in distinguishing intended writing activities from extraneous finger movements that are not related to letters or words. To address this, we introduce a dataset that includes a combination of writing and nonwriting finger motions in each recording. The LEAP from Leap Motion is utilized for marker-free and glove-free finger tracking. We propose a window-based approach that automatically identifies and extracts the air-writing event from a continuous stream of motion data, which may contain stray finger movements unrelated to writing. Sequential writing events are then transformed into a writing segment. The recognition performance is evaluated based on the identified writing segment. Our primary contribution is the development of an air-writing system that encompasses both detection and recognition stages, and provides insights into how the identified writing segments impact the recognition outcome. Through leave-one-out cross validation, the proposed system achieves an overall segment error rate of 1.15 % for recognition based on words, and 9.84 % for recognition based on individual letters. Writing on a touch-based interface using a finger is considered intuitive because it mimics the act of writing with a pen. Recent advancements in tracking technology have eliminated the need for user-worn devices, allowing hand and finger motions to be tracked without any physical restrictions. This has paved the way for air-writing, which serves as a viable alternative for text input when traditional input devices like keyboards or mice are not available or suitable. Unlike other nontraditional input methods such as typing on a virtual keyboard, air-writing offers the advantage of "eye-free" execution, requiring minimal attention. When we write in the air using our fingertip and utilize a controller-free tracking system like LEAP, the motion data captures every aspect of the finger movement in a continuous stream. However, this poses a challenge in detecting and extracting the writing signal from the continuous motion data stream, as the intended writing activity is no longer explicitly located. The Leap device's precise finger tracking allows users to easily write in the air with their fingertip. However, to make Leap a practical writing interface, an intelligent system capable of detecting and recognizing air-writing mixed with other stray movements must be designed. While certain finger movements can be used as delimiter signals to indicate the endpoints of a writing activity, relying on these explicit delimiters hampers the user experience of air-writing. In this study, we propose a system that automatically detects, segments, and recognizes the writing part from the continuous motion tracking signal. In[4], The Challenge-response (CR) method is a reliable way to authenticate users, even if the communication channel is not secure. However, this method is vulnerable to insider attacks where a user can obtain the secret response from a legitimate user. To address this issue, a biometric-based CR authentication scheme called MoCRA has been designed. MoCRA uses the motions of a user operating depth-sensor-based input devices, such as a Leap Motion controller, to authenticate the user. To authenticate a user, MoCRA randomly selects a string and the user has to write the string in the air. MoCRA captures the user's writing movements using Leap Motion and extracts their handwriting style. After verifying that the user's writing matches what is asked for, MoCRA uses a Support Vector Machine (SVM) with co-occurrence matrices to model the handwriting styles and authenticate users reliably, even if what they write is different every time. MoCRA has been evaluated on data from 24 subjects over 7 months and managed to verify a user with an average of 1.18\% (Equal Error Rate) EER and reject impostors with 2.45% EER. User authentication is a crucial and complex task in computer security. The main challenge arises from the vulnerability of communication, which allows for eavesdropping, man-in-the-middle attacks, and replay attacks. To combat these threats, challenge-response (CR) authentication has proven to be effective. In a

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205277 |

typical CR authentication process, a server sends a random challenge to the user. The user must then respond with a valid response, usually a hash of the challenge and a pre-shared secret between the two parties. This method ensures security over insecure communication channels because the challenge is randomly generated and extracting the password from the response is difficult. However, CR authentication is solely based on knowledge rather than identity. Therefore, anyone who possesses the shared secret can pass the authentication, making it vulnerable to insider attacks. Insider attacks pose a significant threat to systems with strict security requirements, such as enterprises or government organizations that only allow security guards who have undergone extensive background checks to patrol their premises. Allowing unauthorized individuals, even if they are friends or colleagues of the authorized guards, can lead to an insider attack. To address this issue, we explore biometric-based challenge-response schemes that authenticate based on the user's identity. Incorporating biometrics into the authentication process adds an additional verification step, which may increase the overall authentication time. In [5], Air-writing refers to the act of writing characters or words in the empty space using hand or finger movements. In two companion papers, we address the challenges of air-writing recognition. Part 2 specifically focuses on detecting and recognizing air-writing activities that are seamlessly integrated into a continuous motion trajectory without clear boundaries. The detection of intended writing activities amidst extraneous finger movements unrelated to letters or words poses a unique challenge that requires a separate treatment from traditional pattern recognition problems.

To tackle this, we introduce a dataset that contains a combination of writing and nonwriting finger motions in each recording. The LEAP from Leap Motion is utilized for marker-free and glove-free finger tracking. We propose a window-based approach that automatically identifies and extracts the air-writing event from a continuous stream of motion data, which may include stray finger movements unrelated to writing. Sequential writing events are then transformed into a writing segment. The performance of the recognition system is evaluated based on the detected writing segments. Our primary contribution lies in the development of an air-writing system that encompasses both detection and recognition stages. Additionally, we provide insights into how the identified writing segments impact the overall recognition outcome. Through leave-one-out cross validation, our proposed system achieves an impressive segment error rate of 1.15% for word-based recognition and 9.84% for letter-based recognition. In [6] 1. Cognitive coding has become popular in human computer interface (HCI) applications. With the emergence of new mobile devices, the demand for user-friendly interfaces continues to increase. Previous methods of analyzing airborne text relied on cameras and sensors, but these methods have limitations in terms of cost and deployment. Recent research has shown that wireless signals can be used to recognize different directions. In this paper, we present a wireless recording device called Wri-Fi that uses Channel State Information (CSI) provided by wireless signals. . Knowing the characters of the alphabet becomes challenging due to their diversity and complexity. We use Principal Component Analysis (PCA) for effective noise removal and Fast Fourier Transform (FFT) for continuous detection. Characteristic CSI waveforms are created by writing patterns of 26 letters at specific positions. Finally, we use Hidden Markov Models (HMM) for modeling and classification. Our lab tests showed that the average Wi-Fi accuracy across the two typing areas was 86.75 % and 88.74 %, respectively.

In [7], we propose a new benchmark dataset for the challenging task of Write-in-Air (WiTA). vision and natural language (NLP). WiTA uses an intuitive, natural typing method that uses finger gestures for human-computer interaction (HCI). Our WiTA dataset will contribute to the development of data-driven WiTA systems, whose performance has so far been unsatisfactory due to lack of suitable data and reliance on statistical models. The database consists of five Korean and English sub-datasets containing a total of 209,926 video samples from 122 participants. To ensure wide and effective reach, we capture WiTA's finger movement using an RGB camera. Furthermore, we propose a 3D ResNet-inspired spatiotemporal residual network architecture for unconstrained fingerprint recognition. This model guarantees instant performance (>100 FPS) and meets the criteria. Impact Statement — Live Authoring (WiTA) is a technology that makes HCI new. As modern technology continues to permeate many aspects of people's daily lives, the demand for new articles suitable for this technology continues to increase. However, most existing scripts do not cover all users and have their own limitations; We will consider them in more detail in the article. The WiTA analysis method proposed in this study overcomes the previous limitations and completely frees HCI from limitations. Our network achieved an overall English error rate (CER) of 29.24 % and managed to maintain 697 FPS; this provided a good starting point for further research on WiTA. WiTA provides a contactless way for people to communicate with computers and has great potential in applications such as augmented reality (AR) and virtual reality (VR). In [8], computer vision air canvas was used to change the information that will appear on the screen and make it an important part of the interaction. The addition of a variety of colors further enhances this interaction, making it easier for users to identify products and providing greater clarity. To do this, you need to access your computer's built-in website or the website itself. This not only improves the overall experience but also provides a more detailed description of the weather. In addition, this machine is also used for visualizing and drawing texts that will attract the attention of the

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205277 |

audience. In addition, it forms the basis for new and interesting product streams in the future. You can harness the power of computer vision to bring your ideas to life by simply moving your fingers in the air. In this research paper, we created a display that can create graphical data or hand gestures by capturing finger movements using a digital webcam. This process is similar to how a touch screen works. The main purpose of this article is color control, character tracking, and collaboration design.

III. METHODOLOGY

A. Writing Detection and Tracking

In this paper, we propose a tracking module that uses mediapipe hand module for robust hand tracking and opencv for image processing tasks. The module captures frames from a webcam using opencv. These frames are then resized to a standardized resolution (640x480) to ensure consistency in processing .detected hands are also annotated with landmarks and connections. After localizing hand landmarks, the finger counting algorithm identifies the position of specific landmarks corresponding to fingertips and calculates the number of extended fingers. The algorithm checks the positions of landmarks associated with fingertips against those of the adjacent landmarks to determine whether a finger is extended or not. Finger's state is stored in a list, and the total count of extended fingers is calculated. Hand detection and finger counting process occur continuously in real-time as new frames are captured from the webcam. Before drawing begins, the system initializes a canvas or drawing area. This canvas can be a virtual space within the application. The system continuously monitors the position and movement of the fingertips, particularly the index finger. When the index finger is detected in an extended state and moves across the canvas, it indicates a drawing action. As the index finger moves across the canvas, the system translates its position into coordinates on the canvas. These coordinates represent the trajectory of the drawing action. The system uses these coordinates to draw lines or strokes on the canvas. Switching between drawing and other interactions is done by changing their finger configurations.

Fig.1 Hand pose Detection

Fig.2 Character Drawing

B. Character Recognition

Our CNN model for english character recognition consists of convolutional layers with rectified linear unit (ReLU) activation, max-pooling layers for spatial down-sampling, batch normalization layers to stabilize training, and global average pooling for dimensionality reduction. The model is topped with dense layers with ReLU activation followed by a softmax layer for multi-class classification. This architecture is choosed for its ability to capture spatial hierarchies in the input images. The model is trained using the Adam optimizer with a learning rate of 0.0005,10 epochs and sparse categorical cross-entropy loss function.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205277 |

For Devnagri character recognition ,(CNN) architecture consists of convolutional layers with ReLU activation, max-pooling layers for spatial down-sampling, and dropout layers for regularization. The final layers include dense layers with ReLU activation and a softmax layer for multi-class classification. The model is compiled with the Adam optimizer,50 epochs with a batch size of 256. and categorical cross-entropy loss function. The models are trained on the CPU due to GPU memory constraints.

IV. EXPERIMENTAL RESULTS

1. Dataset Information

The model uses an open source devnagari character dataset[16], a repository of handwritten images .It comprises of 46 characters with 2000 examples each, total of 92,000 images, this dataset provides a robust foundation for training our system. Each image, is in grayscale with a resolution of 32x32 pixels and stored in PNG format, offers a glimpse into the rich tapestry of Devanagari script. Additionally, a padding of 2 pixels ensures uniformity and clarity in our training data .For English character recognition EMNIST Dataset[17] is used, It features grayscale images centered within a compact 28x28 pixel frame. With 785 columns capturing the essence of each character, this dataset serves as a crucial resource for model training.

2. Hand Recognition and Tracking

The implemented hand tracking and trajectory recognition system operates optimally under medium hand speeds and in environments with optimal light exposure that isn't too bright. It is specifically designed to accurately detect and track the right hand while disregarding the left hand. The system's detection mechanism ensures proper identification of the hand centroid, particularly recognizing extended fingertips as they play a crucial role in initiating writing actions. Moreover, the system imposes a constraint where only one finger, typically the forefinger, should be extended to trigger the writing action. If more than one finger is extended, the system refrains from writing to prevent unintended input. While the system demonstrates robustness in tracking multiple hands, there are challenges when it comes to trajectory recognition, especially when dealing with overlapping or closely positioned hands. Despite this, the overall performance of the system in hand tracking and trajectory generation remains satisfactory. Through careful calibration and parameter tuning, the system effectively tracks hand movements and generates accurate trajectories, providing valuable input for applications requiring precise hand motion analysis. Continued refinement and optimization efforts could further enhance the system's capability to handle complex hand interactions and improve trajectory recognition, ultimately enhancing its usability across various interactive environments and applications.

3. Character Recognition

- **Model Evaluation:** Upon evaluating the EMNIST dataset trained model on the test dataset, we achieve a test accuracy of 85.71 percent. Evaluation of Devnagari test set also shows a high accuracy of 99.25 percent .This performance demonstrates the effectiveness of our approach in accurately recognizing handwritten characters drawn in the air. The test accuracy closely aligns with the validation accuracy, indicating good generalization capability.
- **Training Performance:** From fig.3 ,we can observe that both training and validation accuracy improve as the number of epochs increases. The training accuracy starts at around 0.8 and increases to just above 0.85, while the validation accuracy starts slightly lower but converges towards the training accuracy, ending up just below it. This convergence suggests that the model is generalizing well and not overfitting significantly to the training data.

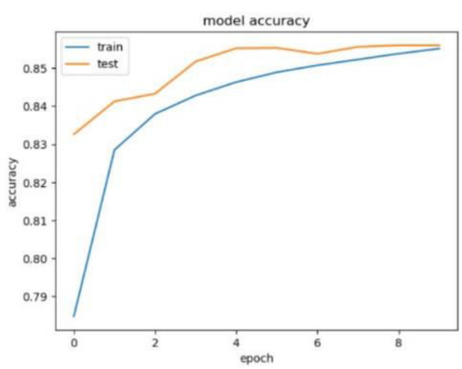


Fig.3 Accuracy graph for English dataset

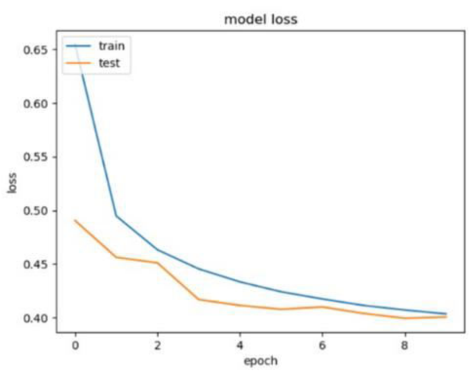


Fig.4 Loss graph for English dataset

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205277 |

The fig. 4 shows that both training and validation loss decrease as the number of epochs increases. The training loss starts at around 0.65 and drops sharply to below 0.4 within the first few epochs, then continues to decrease more gradually. The validation loss follows a similar trend but with a slightly higher value, starting just below 0.65 and decreasing to around 0.35. The close proximity of the training and validation loss lines towards the end of the training indicates that the model is not overfitting .

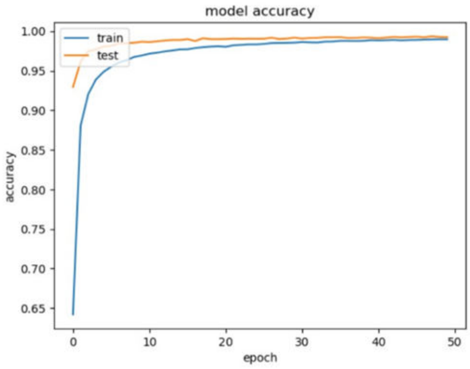


Fig.5 Accuracy graph for Devanagri dataset

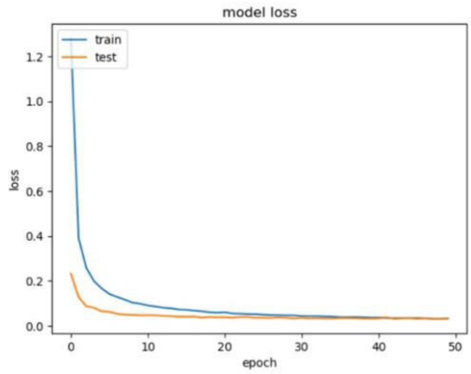


Fig.6 Loss graph for Devanagri dataset

For Devnagri character recognition :From (fig.5) ,we can observe that The training accuracy begins close to 0.65 and rapidly increases to nearly 1.00, indicating that the model is performing very well on the training data. The test accuracy follows a similar trajectory, starting slightly lower than the training accuracy but quickly catching up and plateauing close to the training accuracy, just below 1.00.The high level of accuracy and the convergence of training and test lines suggest that the model is generalizing well and not overfitting.

The "Training and Validation Loss" graph(fig.6) shows that both training and validation loss decrease as the number of epochs increases The close proximity of the training and validation loss lines towards the end of the training indicates that the model is not overfitting .Overall, the fig indicate that the model has learned the patterns in the data effectively, achieving high accuracy and low loss on both the training and test datasets.

IV. ADVANTAGES AND LIMITATIONS

1. **Advantages**

- **Education and Training:** The Air Canvas technology can be employed in educational settings to facilitate interactive learning. It provides a hands-on platform for teaching art, design, and spatial concepts, enhancing engagement and comprehension for students.
- **Accessible Art for Individuals with Disabilities:** The intuitive and gesture-based interface of Air Canvas makes it accessible for individuals with disabilities, offering a means of artistic expression that goes beyond traditional physical limitations.
- **Gesture-Controlled Presentations:** The technology behind Air Canvas can be adapted for gesture-controlled presentations, offering dynamic way for presenters to interact with content in real-time. This can enhance engagement during lectures, workshops, or business presentations.
- **Entertainment and Gaming:** The Air Canvas technology can be integrated into entertainment and gaming applications, providing users with innovative ways to interact with virtual environments. This can lead to more immersive and engaging gaming experiences.

2. **Limitations**

- **Environmental Constraints:** External factors such as ambient lighting, background clutter, or the presence of reflective surfaces can interfere with the accuracy of the computer vision algorithms, leading to suboptimal performance in certain environments.
- **Gesture Recognition Accuracy:** Achieving precise gesture recognition proves challenging, especially under varied lighting conditions or when users execute complex gestures. The system may struggle to accurately interpret intricate hand movements, impacting the fidelity of the virtual artwork**.**
- **Ambiguity in Gestures:** Deciphering air writing gestures may be challenging due to potential ambiguity in the shapes and movements, leading to inaccuracies in letter recognition. Security Concerns: Air writing recognition systems may raise security concerns, as capturing and interpreting gestures in public spaces could lead to unintentional privacy breaches or unauthorized access.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205277 |

- **Dependency on Technology:** The system is dependent on the availability and reliability of technology, and disruptions or failures in sensors or computing components may hinder its functionality.
- **Limited Vocabulary Recognition:** Recognizing a wide vocabulary of words and phrases through air writing may pose challenges, especially when dealing with complex sentences or technical terms.

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 28, 28, 32)	320
max_pooling2d (MaxPooling2D)	(None, 14, 14, 32)	0
batch_normalization (BatchNormalization)	(None, 14, 14, 32)	128
conv2d_1 (Conv2D)	(None, 14, 14, 64)	18496
batch_normalization_1 (BatchNormalization)	(None, 14, 14, 64)	256
max_pooling2d_1 (MaxPooling2D)	(None, 7, 7, 64)	0
batch_normalization_2 (BatchNormalization)	(None, 7, 7, 64)	256
conv2d_2 (Conv2D)	(None, 7, 7, 256)	147712
batch_normalization_3 (BatchNormalization)	(None, 7, 7, 256)	1024
conv2d_3 (Conv2D)	(None, 7, 7, 256)	590080
global_average_pooling2d (GlobalAveragePooling2D)	(None, 256)	0
dense (Dense)	(None, 256)	65792
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 62)	15934

Total params: 839998 (3.20 MB)
Trainable params: 839166 (3.20 MB)
Non-trainable params: 832 (3.25 KB)

Fig. 7 Model Summary for EMNIST dataset

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 30, 30, 32)	896
conv2d_1 (Conv2D)	(None, 28, 28, 64)	18496
max_pooling2d (MaxPooling2D)	(None, 14, 14, 64)	0
conv2d_2 (Conv2D)	(None, 12, 12, 64)	36928
conv2d_3 (Conv2D)	(None, 10, 10, 64)	36928
max_pooling2d_1 (MaxPooling2D)	(None, 5, 5, 64)	0
dropout (Dropout)	(None, 5, 5, 64)	0
flatten (Flatten)	(None, 1600)	0
dense (Dense)	(None, 128)	204928
dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 46)	5934

Total params: 304110 (1.16 MB)
Trainable params: 304110 (1.16 MB)
Non-trainable params: 0 (0.00 Byte)

Fig.8 Model Summary for Devanagri dataset

V. CONCLUSION

In conclusion, our study presents a robust methodology for air-writing recognition using hand tracking and gesture recognition techniques. By leveraging computer vision tools such as the MediaPipe Hand module and OpenCV, we have developed a system capable of accurately tracking hand movements in real-time and interpreting them as characters of respective languages on a virtual canvas. This approach offers a natural and intuitive way for users to interact with digital content, opening up new possibilities for creative expression and interactive applications. As technology advances, air-writing recognition systems hold significant potential for enhancing user experience and driving future innovation in human-computer interaction.

REFERENCES

[1] A. Dash et al., "Air Script - Creating Documents in Air," 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 2017, pp. 908-913, doi: 10.1109/ICDAR.2017.15 \textit{2018 IEEE 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)}

[2] B. A. Kumar, T. Vinod and M. S. Rao, "Interaction through Computer Vision Air Canvas," 2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC), Bhubaneswar, India \textit{ 2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC)}

[3] M. Chen, G. AlRegib and B. -H. Juang, "Air-Writing Recognition—Part II: Detection and Recognition of Writing Activity in Continuous Stream of Motion Data"\textit{M. Chen, G. AlRegib and B. -H. Juang, "Air-Writing Recognition—Part II: Detection and Recognition of Writing Activity in Continuous Stream of Motion Data," in IEEE Transactions on Human-Machine Systems, vol. 46, no. 3, pp. 436-444, June 2016, doi: 10.1109/THMS.2015.2492599.}

[4] Jing Tian, Yu Cao, Wenyuan Xu, Song Wang's "Challenge-Response Authentication using In-Air Handwriting Style Verification"\textit{IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, VOL. XXX, NO. XXX, XXXXX 2017}

[5] Mingyu Chen, Ghassan AlRegib, and Biing-Hwang Juang's "Air-Writing Recognition—Part II: Detection and Recognition of Writing Activity in Continuous Stream of Motion Data" \textit{IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS}

[6] Zhangjie Fu, Jiashuang Xu, Zhuangdi Zhu, Alex X. Liu and Xingming Sun's “Writing in the Air with WiFi Signals for Virtual Reality Devices” \textit{ IEEE Transactions on Mobile Computing}

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205277 |

[7] Ue-Hwan Kim*, Yewon Hwang*, Sun-Kyung Lee and Jong-Hwan Kim's "Writing in The Air: Unconstrained Text Recognition from Finger Movement Using Spatio-Temporal Convolution"
\textit{IEEE TRANSACTIONS ON ARTIFICIAL INTELLIGENCE}
[8] B Anand Kumar, T.Vinod, M Srinivas Rao's "Interaction through Computer Vision Air Canvas" \textit{}{2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC) }
[9] M.Y. Chen, G. Alregib, B.-H. Juang, Air-writing recognition—\textit{Part I: modeling and recognition of characters, words, and connecting motions, IEEE Trans. Hum.-Mach. Syst. 46 (3) (2016) 403–413.}
[10] S. Mitra, T. Acharya, Gesture recognition:\textit{ a survey, IEEE Trans. Syst., Man, Cybern. C Appl. Rev. 37 (3) (2007). }
[11]L. Gupta, S. Ma, Gesture-based interaction and communication: \textit{automated classification of hand gesture contours, IEEE Trans. Syst. Man Cybern. C Appl. Rev. 31 (1) (2001) 114–120. }
[12]I. Infantino, R. Rizzo, S. Gaglio, A framework for sign language sentence recognition by commonsense context, IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., 37(5) (Sep. 2007) 1034–1039.
[13]X. Zhang, et al.,\textit{ A framework for hand gesture recognition based on accelerometer and EMG sensors, IEEE Trans. Syst., Man, Cybern., A, Syst., Humans, 41(6) (Nov. 2011) 1064-1076. }
[14] K.M. Lim, A.W.C. Tan, S.C. Tan, \textit{Block-based histogram of optical flow for isolated sign language recognition, J. Vis. Commun. Image R. 40 (2016) 538–545. }
[15] T.-H. S. Li, M.-C. Kao, P.-H. Kuo, Recognition System for Home-Service-Related \textit{Sign Language Using Entropy-Based K -Means Algorithm and ABC-Based HMM, IEEE Trans. Systems, Man, and Cybernetics, 46(1) (Jan. 2016).}
[16]https://bit.ly/Devnagrihandwrittencharacterdataset
[17]https://www.kaggle.com/datasets/crawford/emnist

.