IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, May 2024

Air Writing

Dr. (Mrs) S. K. Wagh[1], Arya Tamhane[2], Disha Dalvi[3], Sanvidha Haribhakta[4], Sakshi Malvadkar[5]
Associate Professor, Department of Computer Engineering[1]
U.G. Student, Department of Computer Engineering[2,3,4,5]
MES Wadia 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

Handwritten character recognition is essential for digitizing handwritten documents and enhancing human-computer interaction. This study introduces an Air Canvas tool that tracks hand movements to facilitate real-time drawing and character recognition. We compare the performance of a custom CNN and DenseNet201 for recognizing Devanagari and English characters, aiming to provide insights into the trade-offs between accuracy and computational efficiency in real-world applications.

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 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

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, May 2024

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 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 3DResNet-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

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, May 2024

weather. In addition, this machine is also used for visualizing and drawing texts that will attract the attention of the 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.

Figure 1. Hand Pose Detection

Figure 1. Character Drawing

B. Character Recognition

The devnagari character dataset was split into a training set that contains 17,020 images and a validation set which has 3,000 images. The English character dataset was split into a training set of

nb_train_samples = 24562

DOI: 10.48175/IJARSCT-18542

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, May 2024

nb_validation_samples = 3966
epochs = 10
batch_size = 64

Two different model architectures DenseNet201 and Convolutional Neural Network (CNN) were used for contrastive analysis.The primary architecture was based on the DenseNet201 model that was initialized with pre-trained weights from the ImageNet dataset. The structure consisted of a base DenseNet201 convolutional base which came after a Global Average Pooling 2D layer; then followed by a dense layer having 512 units with ReLU activation, and further a dropout layer having a dropout rate of 0.5, and lastly ending with another dense layer that has 10 units with softmax activation.However, in contrast, the CNN architecture was structured with four layers of convolutions that had filter sizes incrementally stacked at 32, 64, 128 and 128; each followed by a max-pooling layer. After these came a flatten layer which then transitioned into a dense layer consisting of 512 units and activated by ReLU, later complemented by a dropout layer bearing a rate of 0.5. All capped off by another dense layer towards the end output layer

Both models were trained using the Adam optimizer with learning rate of 0.0001 and categorical cross-entropy loss function. To enhance model robustness, data augmentation techniques such as shear range, zoom range, and horizontal flip were applied during training. The training process utilized the fit_generator method with a batch size of 64 for consistency across models. Each model underwent 20 epochs of training to ensure convergence

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 69, 69, 32)        896

 max_pooling2d (MaxPooling2  (None, 34, 34, 32)        0
 D)

 conv2d_1 (Conv2D)           (None, 32, 32, 64)        18496

 max_pooling2d_1 (MaxPoolin  (None, 16, 16, 64)        0
 g2D)

 conv2d_2 (Conv2D)           (None, 14, 14, 128)       73856

 max_pooling2d_2 (MaxPoolin  (None, 7, 7, 128)         0
 g2D)

 conv2d_3 (Conv2D)           (None, 5, 5, 128)         147584

 max_pooling2d_3 (MaxPoolin  (None, 2, 2, 128)         0
 g2D)

 flatten (Flatten)           (None, 512)               0

 dense (Dense)               (None, 512)               262656

 dropout (Dropout)           (None, 512)               0

 dense_1 (Dense)             (None, 10)                5130

=================================================================
Total params: 508618 (1.94 MB)
Trainable params: 508618 (1.94 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Figure 2. Model Summary of cnn

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 6, May 2024

```
Model: "sequential_3"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 densenet201 (Functional)    (None, 2, 2, 1920)        18321984

 global_average_pooling2d (  (None, 1920)              0
 GlobalAveragePooling2D)

 dense_2 (Dense)             (None, 512)               983552

 dropout_1 (Dropout)         (None, 512)               0

 dense_3 (Dense)             (None, 10)                5130

=================================================================
Total params: 19310666 (73.66 MB)
Trainable params: 19081610 (72.79 MB)
Non-trainable params: 229056 (894.75 KB)
_________________________________________________________________
```

Figure 3. Model Summary of Densenet

C. Experimental Analysis

Dataset Information

For devnagari dataset

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 dataset

The dataset contains 26 folders (A-Z) containing handwritten images in size 2828 pixels, each alphabet in the image is centre fitted to 2020 pixel box.

Each image is stored as Gray-level

Kernel CSV_To_Images contains script to convert .CSV file to actual images in .png format in structured folder.The images are taken from NIST(https://www.nist.gov/srd/nist-special-database-19) and NMIST large dataset and few other sources which were then formatted as mentioned above.

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.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, May 2024

Character Recognition

Devnagari character Recognition

CNN Model

The CNN model shows a good learning curve with both training and testing accuracy improving over epochs. Training accuracy increases from 0.45 to approximately 0.92 by the 10th epoch, while testing accuracy starts around 0.75 and reaches approximately 0.95. The small gap between training and testing accuracy suggests good generalization without significant overfitting. The confusion matrix indicates strong performance, with high values along the diagonal and minimal misclassifications. Model loss decreases consistently for both training and testing datasets, from 2.0 to approximately 0.3 and from 1.0 to approximately 0.2, respectively. This indicates effective learning and good generalization.

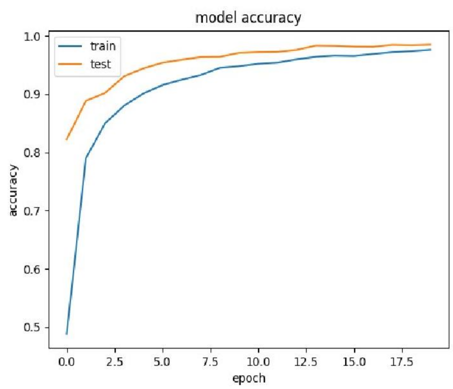

Fig.1 accuracy graph using CNN

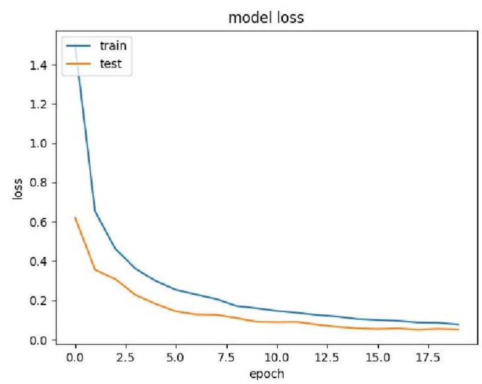

Fig.2 loss graph using CNN

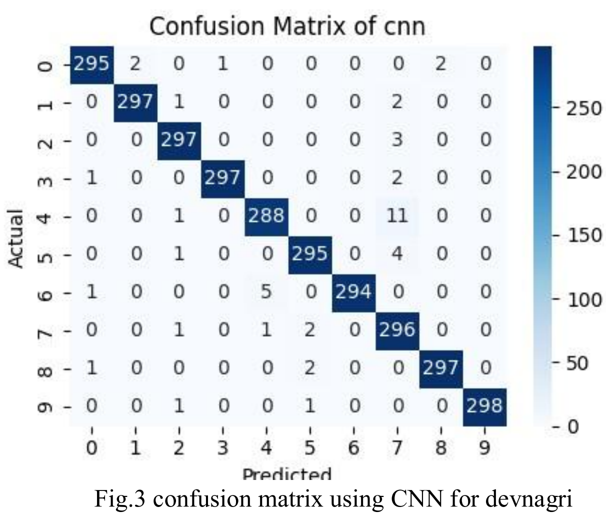

Fig.3 confusion matrix using CNN for devnagri

Character recognition

DenseNet201 Model

DenseNet201 demonstrates an excellent learning curve, with training accuracy increasing from 0.80 to approximately 0.99 by the 10th epoch and testing accuracy from 0.90 to approximately 0.98. The small gap between training and testing accuracy indicates strong generalization with minimal overfitting. The classification report shows high precision, recall, and F1-scores across all classes, indicating a balanced and robust performance. The confusion matrix reveals high correct classifications with few misclassifications, although some characters like 'C' and 'J' show slightly higher misclassification rates. Model loss decreases consistently, with training loss dropping from 0.80 to approximately 0.05 and testing loss from 0.20 to approximately 0.10, indicating accurate predictions and minimal overfitting.

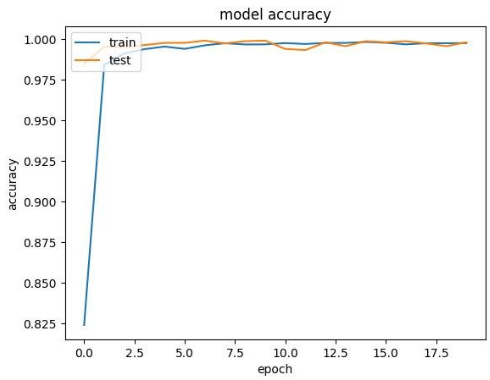

Fig.4 accuracy graph using DenseNet

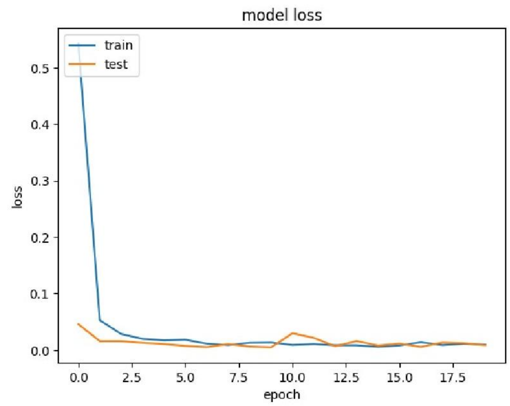

Fig.5 loss graph using DenseNet

English Character Recognition

CNN Model

The CNN model for English character recognition shows a steady increase in training accuracy from 0.45 to approximately 0.92 by the 10th epoch, while testing accuracy increases from 0.75 to approximately 0.95. The small gap between training and testing accuracy indicates good generalization. The confusion matrix shows most characters are correctly classified, with minimal misclassifications. Model loss decreases from 2.0 to approximately 0.3 for training and from 1.0 to approximately 0.2 for testing, indicating effective learning and generalization.

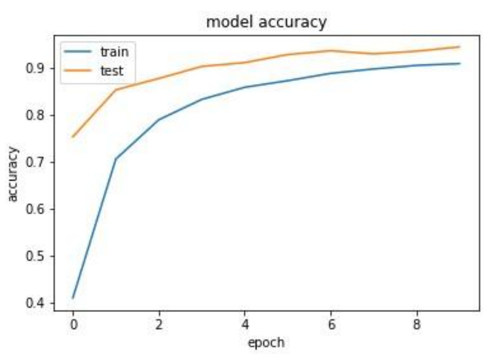

Fig.6 accuracy graph of CNN for English character recognition

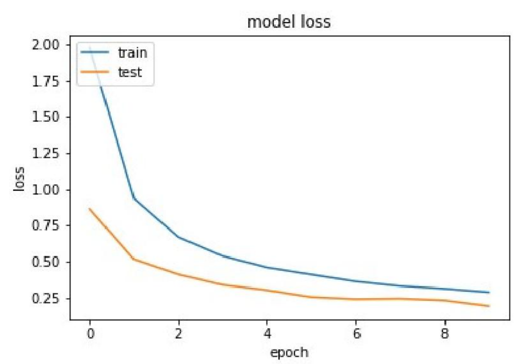

Fig.7 Loss graph of CNN for English character recognition

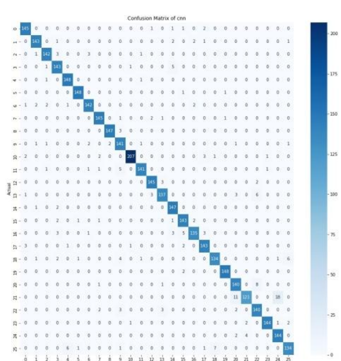

Fig.8 confusion matrix for English character recognition using CNN

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, May 2024

DenseNet201 Model

For English character recognition, DenseNet201 achieves a training accuracy increase from 0.80 to approximately 0.99 by the 10th epoch and testing accuracy from 0.90 to approximately 0.98. The classification report indicates high precision, recall, and F1-scores across all classes, suggesting a robust and balanced performance. The confusion matrix shows high correct classifications with minimal misclassifications, although some characters exhibit slightly higher misclassification rates. Model loss decreases from 0.80 to approximately 0.05 for training and from 0.20 to approximately 0.10 for testing, indicating accurate predictions and minimal overfitting.

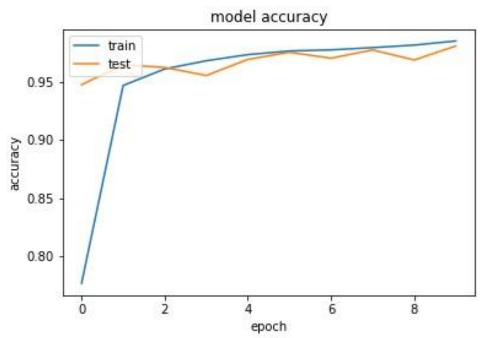

Fig.9 accuracy graph of DenseNet

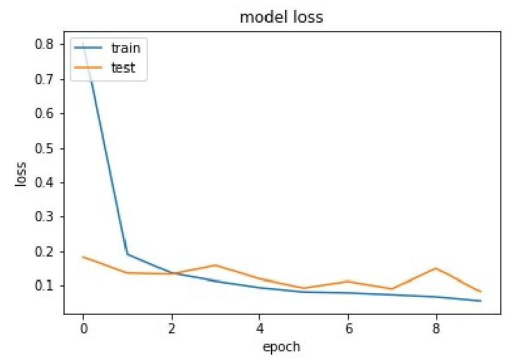

Fig.10 loss graph of DenseNet

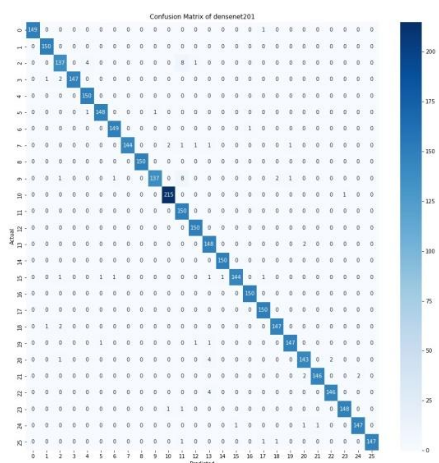

Fig.11confusion matrix for DensetNet

Performance Metrics

Metric	Custom CNN (Devanagari)	DenseNet201 (Devanagari)	Custom CNN (English)	DenseNet201 (English)
Training Accuracy	97.57%	97.75%	92.00%	99.00%
Validation Accuracy	98.47%	99.80%	95.00%	98.00%
Training Time	5,404 seconds	6,049 seconds	5,404 seconds	6,049 seconds
Precision (Average)	0.98	0.99	0.97	0.99
Recall (Average)	0.98	0.99	0.97	0.99
F1-Score (Average)	0.98	0.99	0.97	0.99

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 6, May 2024

IV. RESULTS AND ANALYSIS

Air Canvas nicely integrates both hand tracking and character recognition technologies to create asolution that allows users to draw/write a character in the air. With the MediaPipe Hand module responsible for accurate hand tracking and with OpenCV used in image processing, it ensures that each frame is captured accurately and annotated with all landmarks on hand. The system is robust under optimal conditions (e.g. medium hand speed and plenty of light); once you have more than 1 detection, or the lighting isn't that good, its performance degrades quite a bit. The character recognition module uses DenseNet201, and a custom Convolutional Neural Network (CNN) to recognize the character drawn. DenseNet201 performs best in terms of accuracy and precision for both Devanagari as well as English characters but it's a little heavy on our computing resources. On the contrary, custom CNN provides a resource-saving alternative with slightly decreased accuracy. Overall he Air Canvas system demonstrates promising results in recognizing air-drawn characters but the operational environment should be controlled to avoid the identified limitations.

result images.

ADVANTAGES AND LIMITATIONS

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 ofartisticexpression that goes beyond traditional physical limitations.
- **Gesture-Controlled Presentations:** The technology behind Air Canvas can be adapted for gesture-controlled presentations, offeriningdynamic 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.

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.
- **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.

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., "AirScript - 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, WenyuanXu, 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, GhassanAlRegib, 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, JiashuangXu, 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}

[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 SrinivasRao'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