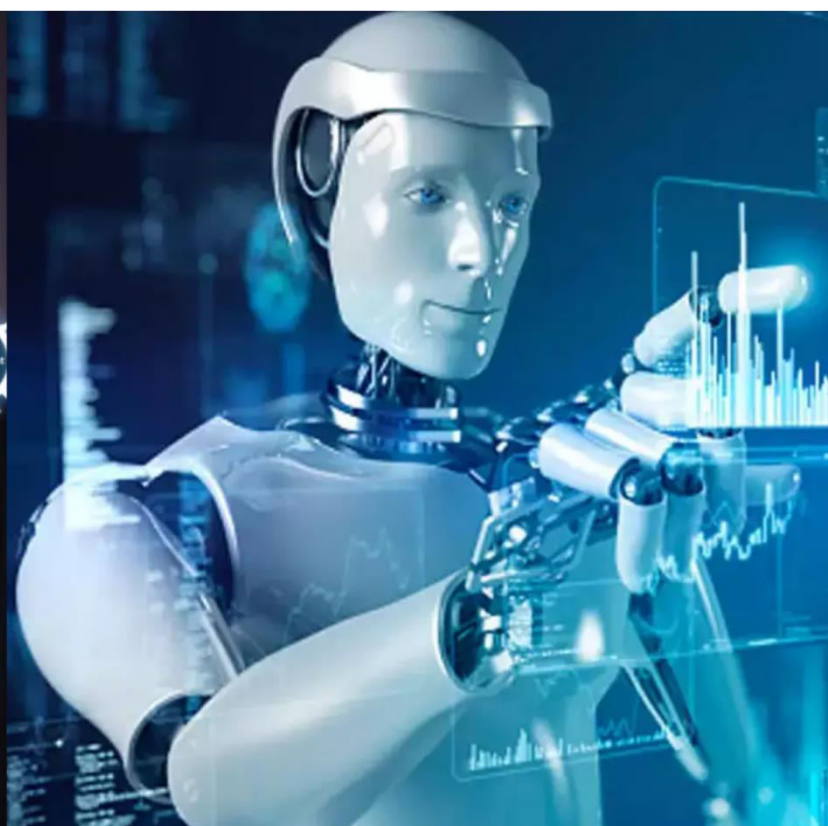




International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 4, April 2025

Air-Writing Recognition and Real-Time Text Conversion with Computer Vision

D. Suneetha¹, Y. Baby Laxmi², J. Akhil Kumar³, M. Akash Uday⁴, B. Rushikesh Reddy⁵

Professor, Department of Computer Science and Engineering (CSE), Malla Reddy University, Maisammaguda,
Hyderabad, India

Student, Department of Computer Science and Engineering (CSE), Malla Reddy University, Maisammaguda,
Hyderabad, India

ABSTRACT: In The increasing demand for touchless interaction systems in modern computing environments, the project "Air-Writing Recognition and Real-Time Text Conversion with Computer Vision" introduces a novel approach to contact-free text input. By employing computer vision techniques such as OpenCV for image processing, MediaPipe for precise hand tracking, and Tesseract OCR for robust text extraction, the system effectively translates air-written gestures into digital text in real-time. An LSTM (Long Short-Term Memory) model enhances recognition accuracy by capturing sequential hand movement patterns. The user-friendly web interface, developed using Flask, HTML, and CSS, ensures accessibility without the need for specialized hardware. This innovative system addresses the limitations of traditional text input methods, offering significant potential for applications in education, accessibility, and creative digital design.

KEYWORDS: Hand Gesture Recognition, Computer Vision, Optical Character Recognition (OCR), Real-time Text Recognition, Deep Learning, Handwriting Recognition.

I. INTRODUCTION

In The project "Air-Writing Recognition and Real-Time Text Conversion with Computer Vision" introduces a touchless text input system that allows users to write in the air using hand gestures. By leveraging computer vision and deep learning, the system tracks hand movements and converts them into digital text in real time. MediaPipe enables accurate hand tracking, while OpenCV processes and visualizes the strokes on a virtual canvas. Pytesseract OCR, powered by an LSTM-based model, extracts text from the gestures, ensuring precise recognition. A user-friendly web interface built with Flask, HTML, and CSS enhances accessibility. This project eliminates the need for traditional input devices, making it ideal for applications in education, accessibility, and creative digital interactions, offering a novel approach to modern human-computer interaction.

II. LITERATURE REVIEW

Air-writing and real-time gesture recognition have gained significant attention in human-computer interaction (HCI), leveraging computer vision, deep learning, and optical character recognition (OCR) technologies. Several studies have proposed innovative approaches to improve accuracy, efficiency, and real-time performance in air-writing systems.

- 1. Vision-Based Gesture Recognition:** A significant portion of research in air-writing recognition focuses on vision-based techniques. Researchers have utilized convolutional neural networks (CNNs) and long short-term memory (LSTM) networks to classify air-written characters based on spatial and temporal patterns. Studies have explored the use of MediaPipe Holistic, CNN, and VGG16 for real-time gesture recognition, achieving high accuracy [6 - 7]. Additionally, various works have proposed frameworks that utilize fingertip tracking and 3D spatial modeling to enhance recognition rates [1 - 9].
- 2. Deep Learning Approaches :** Several studies have explored deep learning frameworks to enhance air-writing and gesture recognition. The Inception-v3 model with LSTM has been used for dynamic hand gesture recognition, achieving an accuracy of 89.7% [9]. A CNN-LSTM hybrid architecture has also been proposed to capture

spatiotemporal dependencies in hand movements [7]. Furthermore, some research has integrated YOLOv5 for coordinate correction and LGR-CNN for classification, achieving a 95.24% accuracy rate [12]. Other approaches have leveraged deep convolutional neural networks (CNNs) for improved text recognition [5].

3. **Air-Writing and Optical Character Recognition (OCR)** : Air-writing systems often rely on OCR for text conversion. Studies have highlighted the importance of OCR advancements, emphasizing deep learning and edge processing to enhance real-time text recognition [11 - 2]. Some works have integrated OCR with gesture recognition to create interactive writing systems that convert air-written text into a digital format [12]. Additional methods include real-time text extraction using OpenCV and Pytesseract for multilingual support [2].
4. **Gesture Recognition Using Sensors and Alternative Modalities** : While vision-based approaches dominate air-writing research, alternative methods incorporating sensor-based techniques have also been explored. A study employed FMCW radar to detect and classify hand gestures using deep learning models [10]. Another work developed a system leveraging electromagnetic sensing for continuous hand gesture recognition [10]. Additionally, air-writing recognition using IMU-based wearable sensors was explored to enable real-time text entry in dynamic environments [5]. Other works have introduced wearable character recognition systems with edge computing [12]. Air-writing techniques leveraging deep learning-based gesture modeling and multimodal sensor fusion have also been studied [8][3].
5. **Applications and Future Directions** : The application of air-writing recognition spans multiple domains, including education, assistive technology, and virtual reality. Some studies have explored the role of air-writing in improving accessibility for individuals with disabilities and enhancing interactive learning experiences [8]. Additionally, air-writing has been implemented in Air-Canvas applications, enabling real-time digital sketching and educational enhancements [9]. Air-writing has also been integrated with virtual keyboard applications, allowing efficient and intuitive non-touch-based text input [4]. Future research is expected to focus on improving model generalization across diverse user populations, integrating multimodal inputs, and enhancing real-time performance through optimized deep learning models.

Challenges: Gesture Variability and User Adaptability, Real-Time Processing Constraints, and OCR Accuracy in Air-Writing.

Future Research: Adaptive and Personalized Air-Writing Models, Lightweight Deep Learning for Edge Computing, Improved OCR for Air-Written Text, Enhanced Hand Tracking and Gesture Segmentation, and Real-Time Noise Reduction and Occlusion Handling.

III. PROBLEM STATEMENT

3.1 Existing System: Traditional text input methods, such as keyboards, touchscreens, and styluses, require physical contact, making them inconvenient and inaccessible for certain users. While some gesture-based systems exist, they often suffer from accuracy issues, hardware dependencies, and processing delays.

3.2 Limitations:

1. **Requires Physical Contact:** Keyboards and touchscreens demand direct interaction, which can be inconvenient for individuals with disabilities or in environments requiring touch-free interaction.
2. **Limited Accessibility:** Users with motor impairments may find it difficult to use conventional input devices effectively, limiting their ability to interact with digital systems.
3. **Hardware Dependency:** Some gesture-based systems require additional external sensors, styluses, or specialized hardware, increasing costs and reducing portability.
4. **Accuracy Issues in Gesture Recognition:** Existing solutions often fail to accurately track hand movements in real-time, especially in dynamic environments with varying lighting conditions.
5. **Processing Delays:** Some air-writing systems suffer from latency issues, resulting in slow or inaccurate text conversion, making real-time interaction difficult.

3.3 Proposed System:

The proposed system, "Air Scribbling and Real-Time Text Conversion using Computer Vision and OCR," provides a touchless, real-time method for text input. It integrates:

- Hand Tracking (MediaPipe) for precise gesture detection.
- Real-Time Video Processing (OpenCV) for capturing and interpreting air-writing movements.
- Text Recognition (Tesseract OCR) to extract readable text from air-drawn strokes.
- Deep Learning (LSTM Algorithm) to improve recognition accuracy.
- Web-Based Interface (Flask) to ensure broad accessibility.

Touchless and Intuitive Input: Users can write in the air using natural hand gestures, eliminating the need for physical contact with devices.

High Accuracy with OCR and Deep Learning: The system integrates Tesseract OCR with LSTM-based deep learning to improve recognition accuracy, even for irregular handwriting styles.

Real-Time Performance: The system processes gestures instantly, providing a smooth and interactive user experience.

No Additional Hardware Required: Works with a standard webcam, making it cost-effective and widely accessible without specialized hardware.

Improved Accessibility: Helps individuals with disabilities, especially those unable to use keyboards or touchscreens, by providing an alternative text input method.

Multi-Domain Applications: Can be used in education (learning tools), accessibility (assistive technology for disabled users), and creative digital design (air-drawing applications).

IV. METHODOLOGY

1. System Overview: This project aims to develop a touchless, real-time handwriting recognition system where users can "scribble" in the air using their hand gestures. The system tracks the hand movements, converts them into strokes on a virtual canvas, and then applies Optical Character Recognition (OCR) to extract readable text.

Key Technologies Used:

- Computer Vision: OpenCV for video capture and image processing.
- Hand Tracking: MediaPipe Hands for gesture recognition.
- OCR Processing: Tesseract OCR for text recognition.
- Web Interface: Flask for backend, HTML/CSS/JavaScript for frontend.

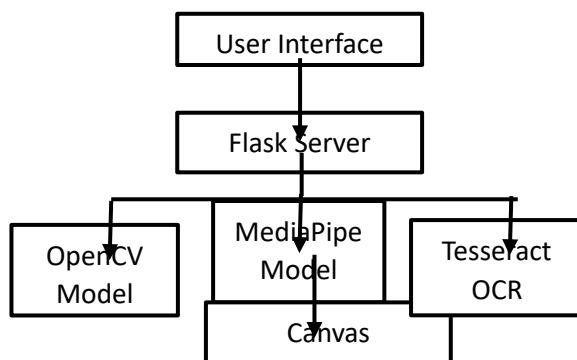
2. System Architecture:

Fig-1 Architecture Diagram

The system consists of the following main modules:

2.1 Input Module (Video Capture)

- Captures real-time video feed from a webcam using OpenCV.
- Ensures 30 FPS for real-time processing.

2.2 Hand Gesture Recognition

This module tracks the index finger movement to capture strokes.

- Library Used: MediaPipe Hands
- Landmarks Used:
 - Index finger tip (Landmark ID = 8)
 - Middle finger tip (Landmark ID = 12)
 - Middle finger knuckle (Landmark ID = 10)

Algorithm:

1. Detect the hand and extract landmarks.
2. Check the position of the middle finger:
 - If the middle finger is raised, disable drawing.
 - If the middle finger is lowered, enable drawing.
3. Track the index finger's movement to create strokes.

Mathematical Formula for Finger Tracking

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Where:

- (x_1, y_1) = Previous index finger position.
- (x_2, y_2) = Current index finger position.
- d = Distance moved by the index finger.

If $d > \text{Threshold}$, draw a line between the previous and current positions.

2.3 Virtual Canvas for Air Writing:

The Virtual Canvas is a key component of the system that captures air-written strokes and displays them digitally. It acts as an overlay where the user's gestures are translated into visible drawings. This module is implemented using OpenCV and NumPy.

Mathematical Formulas for Virtual Canvas

- Canvas Initialization:

$$\text{Canvas}(i, j) = 255$$

for all pixels (i, j) where 255 represents a white pixel in RGB format.

Drawing Mechanism:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Where:

- = Previous index finger position.
- = Current index finger position.
- Merging Video and Canvas:

Merging Video and Canvas:

Output = $0.7 \times \text{VideoFrame} + 0.3 \times \text{Canvas}$

where:

- 0.7 weight for VideoFrame to maintain visibility.
- 0.3 weight for Canvas to blend drawings smoothly.

2. .4 OCR Processing (Text Recognition):

Once the user completes air writing, the system extracts text using Tesseract OCR.

Preprocessing Steps:

1. Convert to grayscale:

$$I_{gray}(x, y) = 0.2989R + 0.5870G + 0.1140B$$

2. Denoising using Gaussian Blur:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

3. Adaptive Thresholding for binarization:

$$T(x, y) = \frac{1}{S} \sum_{(i,j) \in S} I(i, j)$$

4. OCR Processing:

$$Text = OCR(PreprocessedImage)$$

3. Performance Evaluation

To measure system efficiency, we use three key performance metrics: OCR Accuracy to Measures text recognition efficiency, Frame Rate to Determines real-time performance, Gesture Precision Evaluates hand tracking accuracy.

$$OCR \text{ Accuracy} = \frac{\text{Correct Characters}}{\text{Total Characters}} * 100$$

$$\text{Frame Rate} = \frac{\text{Total Frame Processed}}{\text{Time Taken}}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

Where

TP = True Positives

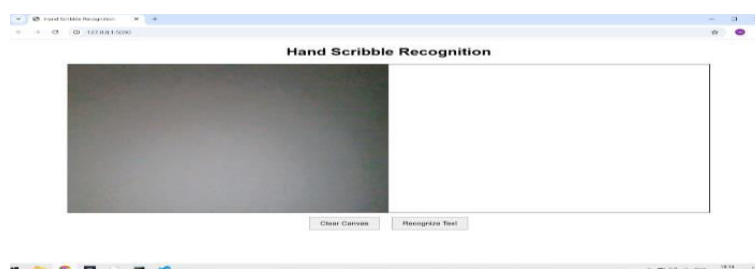
FP = False Positives

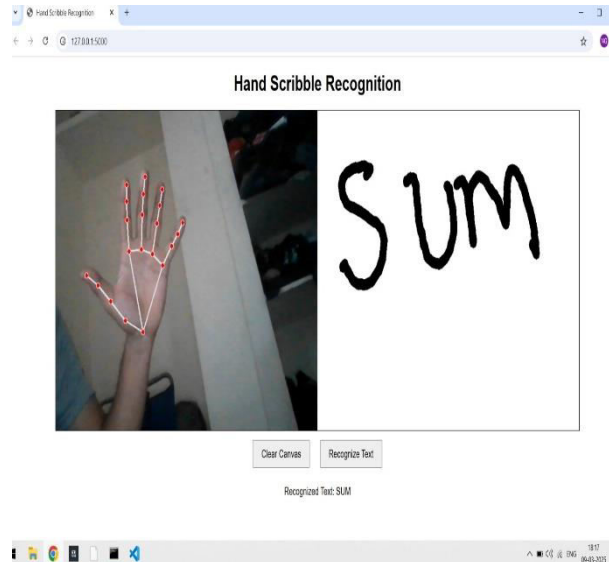
4. Deployment:

The **Air Scribbling and Real-Time Text Conversion** application is deployed **locally** within a Python environment, ensuring smooth execution without the need for an internet connection. This approach offers greater flexibility, faster processing, and enhanced data privacy since all computations are performed on the user's device rather than relying on cloud-based services.

V. EXPERIMENT RESULT

OUTPUT:





VI. CONCLUSION

The project "Air Scribbling and Real-Time Text Conversion using Computer Vision and OCR" successfully demonstrates an innovative and touchless approach to human computer interaction. By combining MediaPipe for hand tracking, OpenCV for real time video processing, and Pytesseract OCR with LSTM for accurate text recognition, the system converts air-drawn gestures into digital text with efficiency and ease. This project addresses the limitations of traditional input methods, offering a flexible, accessible, and intuitive solution for real-time text conversion.

VII. FUTURE WORK

To further improve the usability, accuracy, and versatility of the Air Scribbling and Real-Time Text Conversion system, several key enhancements can be implemented in future iterations. These enhancements include multi-hand support, gesture-based UI controls, cloud integration, and multi-language OCR support.

1. Multi-hand Support (Dual Writing)

Currently, the system processes handwriting from a single hand. Future enhancements can include support for dual-hand input, allowing users to write with both hands simultaneously.

- Improved Writing Speed: Users can write using both hands for faster text input.
- Bimanual Interaction: One hand can be used for writing while the other performs gestures or edits.
- Left-hand Compatibility: Ensuring the system works efficiently for left-handed users.

2. Gesture-based UI Controls

A gesture-based user interface can be integrated to improve interactivity and efficiency. Users can perform specific gestures to execute commands like erasing text, changing stroke thickness, or navigating the interface.

Example Gestures:

- Swipe Left/Right – Undo/Redo actions.
- Swipe Down – Erase the last stroke or clear the screen.
- Pinch Gesture – Adjust stroke thickness.
- Hold & Tap – Convert handwritten strokes into text instantly.

3. Cloud Integration for Data Storage and Retrieval

To enhance accessibility, cloud integration can enable users to store, retrieve, and share their recognized text.

- Real-time Synchronization: Access handwritten notes across multiple devices.
- Backup & Recovery: Prevent data loss with secure cloud storage.
- Collaboration: Share notes and collaborate in real time.

Implementation Approach:

- Integrate with cloud services such as Google Drive, OneDrive, or Firebase.

- Allow exporting recognized text as PDF, Word, or TXT files.
- Implement REST API connections to enable seamless cloud interactions.

4. Multi-language OCR Support

Expanding beyond English handwriting recognition to support multiple languages will make the system more inclusive.

- Support for Complex Scripts: Recognizing non-Latin scripts like Chinese, Arabic, Hindi, etc.
- Adaptive OCR Models: Detect and process handwriting in different languages dynamically.
- Custom Training Data: Enhance accuracy by training OCR models with diverse handwriting datasets.

REFERENCES

Here are some references that you can explore for further information

- [1] Prasun Roy, Subhankar Ghosh, and Umapada Pal, "A CNN Based Framework for Unistroke Numeral Recognition in Air-Writing", February 2025, arXiv. Also available at github.com/prasunroy/air-writing.
- [2] Ayush Tripathi, Arnab Kumar Mondal, Lalan Kumar, Prathosh A.P., "ImAiR: Airwriting Recognition Framework using Image Representation of IMU Signals", September 2022, arXiv. Also available at github.com/ayushayt/ImAiR.
- [3] Koye Navya, Mallela Sowmyah, Shaik Ayesha Amreen, Vemireddy Sravani, Yarrakula Gayathri Devi, Ms. Perli Nava Bhanu, "A Machine Learning Approach for Air Writing Recognition", April 2023, IJRASET. Also available at ijraset.com.
- [4] Aseel Qedear, Aldanh AlMatrafy, Athary Al-Sowat, Abrar Saigh, Asmaa Alayed, "Real-Time Air-Writing Recognition for Arabic Letters Using Deep Learning", September 2024, Sensors. Also available at mdpi.com.
- [5] Bijon Mallik, Md Abdur Rahim, Abu Saleh Musa Miah, Keun Soo Yun, Jungpil Shin, "Virtual Keyboard: A Real-Time Hand Gesture Recognition-Based Character Input System Using LSTM and Mediapipe Holistic", March 2024, CSSE. Also available at techscience.com.
- [6] Yaseen, Oh-Jin Kwon, Jaeho Kim, Sonain Jamil, Jinhee Lee, Faiz Ullah, "Next-Gen Dynamic Hand Gesture Recognition: MediaPipe, Inception-v3 and LSTM-Based Enhanced Deep Learning Model", August 2024, Electronics. Also available at mdpi.com.
- [7] Prof. A.S. Khandagale, Mr. Soham Kale, Mr. Prasanna Khore, Mr. Sahil Bhosale, Mr. Siddhesh Kondhalkar, "Air-Canvas: Real-Time Gesture Recognition", November 2024, IJRPR. Also available at ijrpr.com.
- [8] Daniel Licea, "Advancements in Optical Character Recognition Systems for Real-Time Data Processing", December 2024, ResearchGate. Also available at researchgate.net.
- [9] Prof. Snehal Javheri, Yash Rajiv Ghute, Nandan Anerao, Akshay Sudhakar Shinde, "Air Gesture and Handwriting Recognition: Advancements and Applications in Human-Computer Interaction", November 2024, International Journal of Enhanced Research in Management & Computer Applications. Also available at researchgate.net.
- [10] Yajun Zhang, Zijian Li, Zhixiong Yang, Bo Yuan, Xu Liu, "Air-GR: An Over-the-Air Handwritten Character Recognition System Based on Coordinate Correction YOLOv5 Algorithm and LGR-CNN", January 2023, Sensors. Also available at mdpi.com.
- [11] B. Avinash, Shaik Ishrath Anjum, Syed Roohi, Vibudi Divya Priya, Venicharla Bhargavi, "Real-Time Text Detection and Recognition Based on Optical Character Recognition (OCR)", April 2024, Journal of Science and Technology. Also available at jst.org.in.
- [12] Chaur-Heh Hsieh, You-Shen Lo, Jen-Yang Chen, Sheng-Kai Tang, "Air-Writing Recognition Based on Deep Convolutional Neural Networks", October 2021, IEEE Access. Also available at iee.org.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details