GESTURITE: AI-POWERED AIR WRITING & DOCUMENT ENHANCEMENT SYSTEM

Project Reference Number: *49S_BE_1390*

College : K.L.E. College of Engineering and Technology, Chikodi
Branch : Department of Computer Science and Engineering
Guide(S) : Prof. Balachandra Rao.
Student(s) : Mr. Sushant Jabade
Mr. Sanket Dadde
Mr. Shubham Savant
Mr. Shubham Wadkar

Keywords

Air Writing, Computer Vision, MediaPipe, CNN, Gesture Recognition, AI, OpenCV

Introduction

AirWriting is an innovative computer vision-based application that enables users to write in the air using simple hand gestures captured through a webcam. Traditional input devices such as keyboards and touchscreens require physical interaction, which may not be suitable in all environments. In places like smart classrooms, medical facilities, and public systems, contactless interaction is becoming increasingly important.

This project transforms the space in front of a camera into a virtual canvas, allowing users to write naturally using finger movements. By leveraging MediaPipe for real-time hand tracking, the system detects hand landmarks and interprets gestures with high precision.

The application integrates machine learning models to recognize handwritten characters and digits drawn in the air. These predictions are further enhanced using AI-based text correction, ensuring accuracy and readability.

The system is designed to provide a seamless, intuitive, and futuristic user experience without requiring any additional hardware apart from a standard webcam. This approach makes it cost-effective and accessible.

AirWriting contributes to the advancement of Human-Computer Interaction (HCI) by enabling a natural and touch-free method of communication between humans and machines.

Objectives

1. To develop a contactless writing system using a webcam.
2. To implement real-time hand tracking using MediaPipe.
3. To recognize characters and digits using CNN models.

4. To improve accuracy using AI-based text correction.
5. To provide a smooth and user-friendly interface.

Methodology

1. The system begins by capturing real-time video input from a webcam. MediaPipe is used to detect and track 21 hand landmarks, which help in identifying finger positions and gestures. Based on these landmarks, the system distinguishes between different modes such as selection mode and drawing mode.
2. In drawing mode, the user writes in the air using a single finger. The coordinates of the finger movement are continuously tracked and plotted on a virtual canvas. To ensure smooth writing, filtering techniques are applied to reduce noise and jitter in the motion.
3. Once the writing is completed, the captured strokes are processed and converted into images. These images are passed to pre-trained Convolutional Neural Network (CNN) models, which classify them into alphabets or digits. Separate models are used for better accuracy.
4. After prediction, the recognized text is sent to the Google Gemini API for grammar correction and refinement. This improves the quality of the final output by correcting spelling and sentence structure.
5. The system also includes gesture-based controls such as clearing the screen, confirming input, and navigating options. A visually appealing interface is designed using overlays and real-time feedback mechanisms.
6. The backend is managed using Django, which handles session management and data storage. SQLite is used as the database for storing user data and results.

Results & Conclusions

1. The AirWriting system successfully demonstrates real-time air writing using computer vision techniques. The hand tracking module accurately detects finger movements, allowing users to write naturally without physical contact.
2. The CNN-based recognition system provides reliable predictions for both alphabets and digits. Integration with AI-based correction further enhances the readability and correctness of the output text.
3. The system performs efficiently in real-time and provides a smooth user experience. It eliminates the need for traditional input devices and offers a modern alternative for human-computer interaction.
4. Overall, the project proves that air writing can be a practical and effective method for data input in various applications such as education, healthcare, and smart systems.

Project Outcome & Industry Relevance

1. The project results in a fully functional air writing system that can be used in real-world applications. It demonstrates the integration of computer vision, machine learning, and AI technologies into a single system.
2. This technology is highly relevant in industries focusing on touchless interaction, such as healthcare, smart classrooms, and augmented reality systems. It can also be extended to assist differently-abled individuals in communication.
3. The project showcases skills in AI, real-time processing, and system integration, making it valuable for software development and research domains.

Working Model vs Simulation / Study

This project is a **working model**, as it uses a real-time webcam to capture hand gestures and produce actual output. It is not just a simulation or theoretical study.

Project Outcomes and Learnings

1. Developed a real-time air writing system using computer vision.
2. Built and trained CNN models for character recognition.
3. Learned hand tracking and gesture detection using MediaPipe.
4. Gained experience in integrating AI APIs like Google Gemini.
5. Improved understanding of real-time data processing and UI design.

Future Scope

1. The future scope of AirWriting includes several enhancements to improve its functionality and usability. The system can be extended to support sign language recognition, enabling communication for hearing-impaired individuals.
2. Integration with AR/VR technologies can provide an immersive writing experience in virtual environments. The system can also be enhanced to support multiple languages, making it more accessible globally.
3. Further improvements can include recognition of complex mathematical expressions and diagrams. Optimization of models using lightweight frameworks like TensorFlow Lite can enable deployment on mobile and embedded devices.
4. Multi-user collaboration features can allow multiple users to write on a shared virtual canvas. Advanced document formatting using gestures can transform the system into a complete virtual writing platform.
5. With continuous advancements, air writing technology has the potential to replace traditional input devices in many applications.