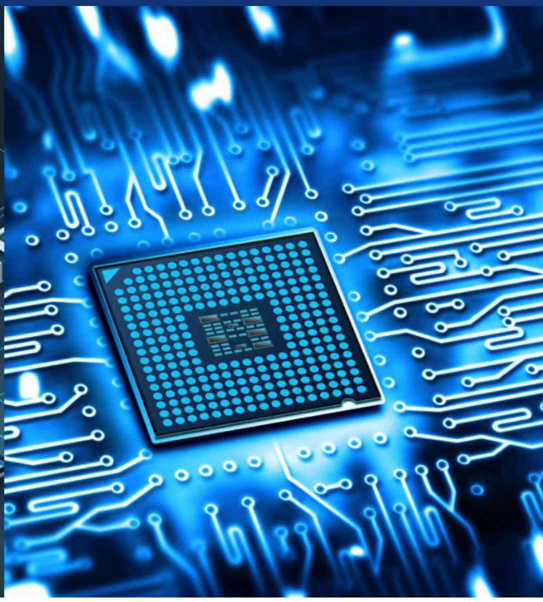
IJIRCCE

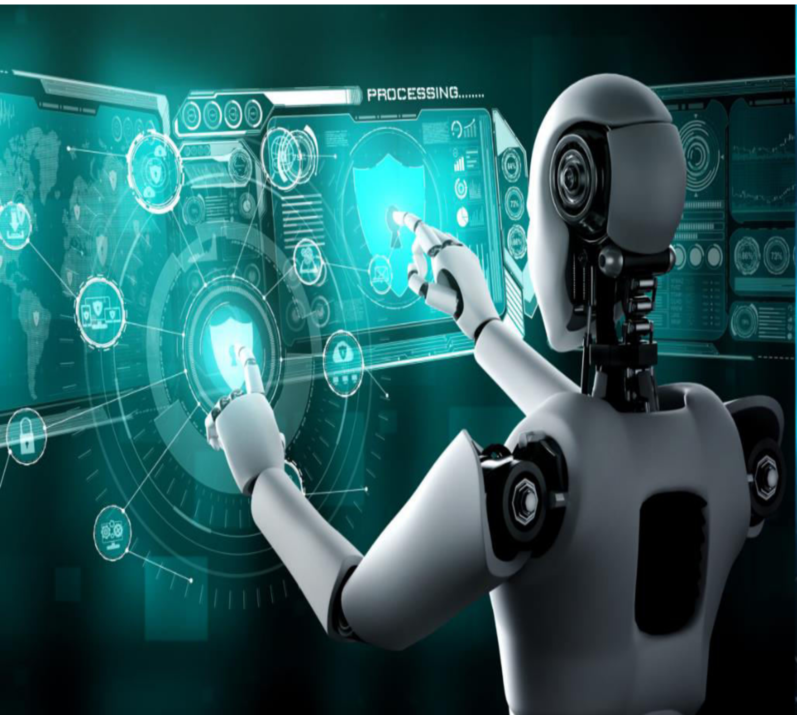
ISSN(O): 2320-9801

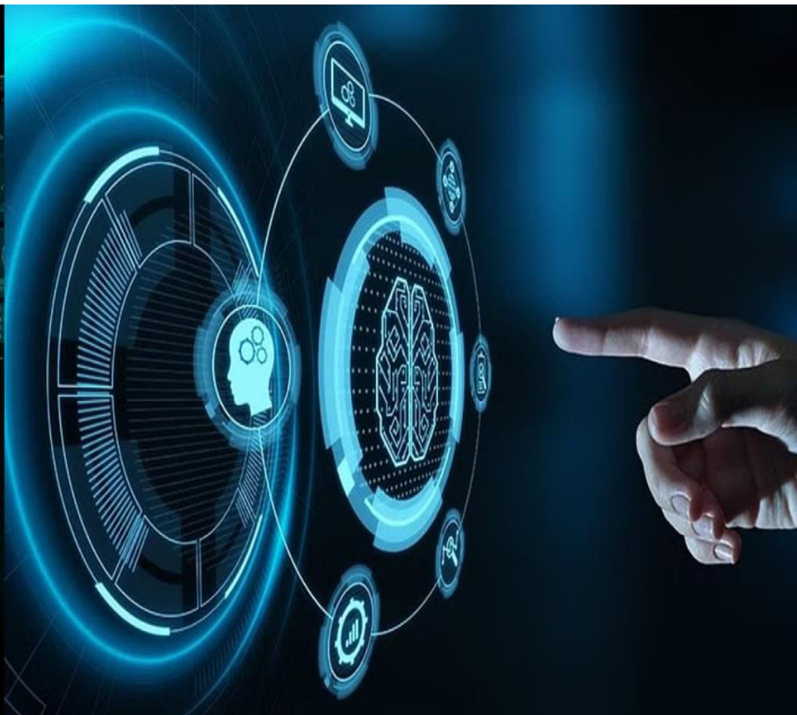
ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

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





Impact Factor: 8.771

Volume 13, Issue 4, April 2025

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

Virtual Drawing Board Using OpenCV and NumPy in Python

Mr. P .Sai Krishna[1], K. Kowshik[2], S. Krishna Raja[3], G. Jayachandran[4], R. Mohan Kumar[5]

Assistant Professor, Department of ECE, Mahendra Institute of Technology, Namakkal, Tamil Nadu, India[1]

UG Students, Department of ECE, Mahendra Institute of Technology, Namakkal, Tamil Nadu, India[2,3,4,5]

ABSTRACT: Recent years have witnessed significant developments in air-based writing technology, establishing it as a complex yet promising research domain within image processing and pattern recognition. This field plays a crucial role in enhancing automation and improving human-computer interaction across diverse applications. Researchers continually strive to optimize processing speed while maximizing recognition accuracy. Object tracking, fundamental to Computer Vision, has gained prominence due to advancements in computational power, affordable high-quality imaging devices, and growing demand for automated video analysis. The video analysis process typically encompasses three sequential phases: object detection, frame-to-frame tracking, and behavior interpretation. Successful object tracking requires addressing four essential components: selecting appropriate object representation, identifying key tracking features, implementing effective object detection, and maintaining accurate trajectory information. The resulting text can be utilized for various applications including message composition and email creation. Moreover, the system serves as a valuable communication aid for individuals with hearing impairments. By providing an intuitive, hands-free text input method, this approach significantly reduces dependence on conventional keyboards, offering a more accessible and natural form of digital communication.

KEYWORDS: Air Writing, Computer Vision, Object Tracking, Motion-to-Text Conversion, Gesture Recognition, Image Processing, Pattern Recognition, Human-Computer Interaction, Wearable Technology, Automated Text Input, Assistive Communication, Video Analysis.

I. INTRODUCTION

The traditional craft of writing is progressively being replaced by digital art in the current digital era. Various artistic expressions and messages delivered in a digital format are included in the category of digital art[3]. Digital art's reliance on modern science and technology is one of its main characteristics. On the other hand, traditional art refers to works of art produced before the advent of digital art. From the viewpoint of the audience, this encompasses literature, painting, sculpture, architecture, music, dance, theater, and other artistic creations that fall under the categories of visual art, audio art, audio-visual art, and audio-visual imaginary art. Traditional and digital art are interdependent and interconnected.

Human requirements have an impact on social growth rather than personal preferences. This idea applies to art as well. Nowadays, traditional and digital art coexist in a mutually beneficial relationship, thus it is essential to methodically investigate the underlying relationships between them. Conventional writing methods often require a pen and paper, as well as a board and chalk. Developing a hand gesture recognition system for digital writing is the primary objective of digital art.

Numerous writing methods are available in digital art, such as using a keyboard, touch-screen displays, digital pens, styluses, and electronic hand gloves[4]. We use a Python-based machine learning method to recognize hand gestures in this system. This method encourages organic human-machine connection. The need for creating natural Human-Computer Interaction (HCI) solutions to replace conventional approaches is increasing quickly due to technological improvements.

The remainder of this paper is structured as follows:

- Section 2 presents the literature we referred to before working on this project.
- Section 3 discusses the challenges faced during system development.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

- Section 4 defines the problem statement.
- Section 5 outlines the system methodology and workflow, including Fingertip Recognition Dataset Creation and Fingertip Recognition Model Training.
- Section 6 describes the workflow algorithm.

II. LITERATURE REVIEW

A. Robust Hand Recognition with Kinect Sensor

The technique suggested in [3] detects hand forms by using color and depth information from the Kinect sensor. Gesture detection remains a challenging issue despite the Kinect sensor. The sensor is helpful for tracking bigger objects, including the human body, because of its 640x480 resolution. Tracking tiny things, like fingers, is far more difficult, though.

B. LED-Fitted Finger Movements

A method that involves attaching an LED to the user's finger and using a webcam to track its movement was shown, as recommended in [5]. The machine then outputs the matching alphabet after comparing the drawn character with an existing database. A red LED that is affixed to the finger is needed for this technique. It also presumes that the webcam's focus region is free of any other red-colored items.

C. Augmented Desk Interface

In [5], an improved desk interface technique for user interaction was proposed. With the use of a video projector and a charge-coupled device (CCD) camera, this technology allows users to manipulate desktop applications with their fingertips. Each hand has a distinct function; the left hand selects from radial menus, while the right hand manipulates objects. The gadget uses an infrared camera to record fingertip movements. However, because determining fingertip positions requires a lot of calculation, the system incorporates preset search windows to enhance speed.

III. CHALLENGES IDENTIFIED

A. Fingertip Detection

The current system is designed specifically for finger detection and tracking without support for alternative input tools such as highlighters, paints, or related implements[6]. A primary technical limitation is the challenge of accurately identifying and isolating fingers from standard RGB images without the assistance of depth sensor technology. This creates significant difficulties in distinguishing finger contours from complex backgrounds and similar-colored objects, particularly in varying lighting conditions and environments. The absence of depth information requires more sophisticated image processing algorithms to achieve reliable finger detection and tracking performance.

B. Lack of Pen-Up and Pen-Down Motion

The system relies on a single RGB camera for writing from above. Due to the absence of depth sensing, the system cannot track up and down pen movements. Consequently, the entire trajectory of the fingertip is recorded, often producing an inaccurate output that the model fails to recognize. The distinction between handwritten and air-written ‘G’ is illustrated

C. Real-Time System Control

Implementing a system that transitions between states based on real-time hand gestures requires meticulous coding and thorough attention to detail. The complexity arises from the need to accurately detect and differentiate between various gestures in real-time while minimizing false positives and maintaining system responsiveness.

Figure 1: Original letter and trajectory formet

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

Furthermore, the system places a cognitive burden on users, who must learn and remember multiple distinct gestures to effectively operate the interface. This learning curve can impact adoption rates and user experience, particularly for novice users or those with limited dexterity. Users must develop muscle memory and spatial awareness to perform gestures consistently and accurately, which may require significant practice before achieving proficiency with the system.

IV. PROBLEM DEFINITION

Several important socioeconomic challenges are intended to be addressed by this project:

1. **Hearing Impairment:** Although listening and hearing are frequently taken for granted, people who have hearing impairments communicate through sign language. Unfortunately, most people need a translator to help them understand their ideas and feelings.

2. **Overuse of Smartphones:** Although smartphones are frequently praised for their portability and convenience, they can lead to a number of problems, including distractions, accidents, and mental health issues. Although some of their adverse effects are still being investigated, they may have major, even fatal, repercussions.

3. **Paper Wastage:** The overuse of paper is still a major environmental issue. After being used for writing, sketching, and scrawling, large quantities of paper are thrown away. Approximately five litres of water are needed to produce one A4-sized piece of paper; approximately 93% of paper comes from trees; paper makes up 50% of company trash; and wasted paper makes up 25% of landfill debris[10]. These are some concerning figures. This waste produces serious environmental damage, depletes natural resources, and aids in deforestation.

Air writing is one practical way to address these problems. It serves as an assistive communication tool for those with hearing impairments by converting printed text into speech or Augmented Reality (AR) displays. It also reduces distractions by allowing users to write in the air while continuing to focus on their tasks. Additionally, because air writing is entirely digital and eliminates the need for paper, it encourages a more ecologically friendly form of writing and recording.

V. SYSTEM METHODOLOGY

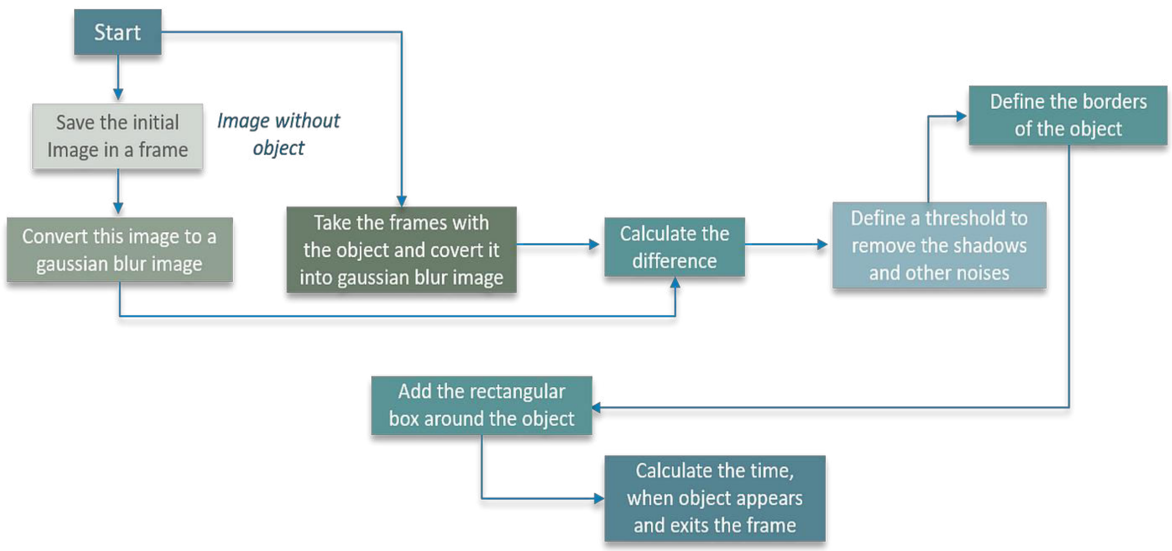


Figure 2: Flow diagram of the mythology

The system requires a dataset to train the Fingertip Detection Model, which is primarily designed to capture motion, enabling air character recognition.

A. Fingertip Detection Model

Air pens of a particular color or a stylus can be used for air writing [2]. However, by using fingertips, this approach does

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

away with the requirement for such instruments. Enabling users to write while in the air without having to lug along extra equipment is the aim. Each frame's fingertip is detected using deep learning algorithms, which produce a set of coordinates.

B. Techniques for Fingertip Recognition Dataset Creation

a. Video to Image Conversion

Using this method, short two-second films of hand movements in various contexts were captured. As seen in Figure 3, these movies were then split up into 30 distinct frames. Two thousand photos in all were gathered and given hand labels [13]. 99% accuracy was attained by the top-performing model that was trained on this dataset. Nevertheless, the dataset lacked variance because all 30 photos were captured from the same film and location, which had a detrimental effect on the model's performance in different situations.

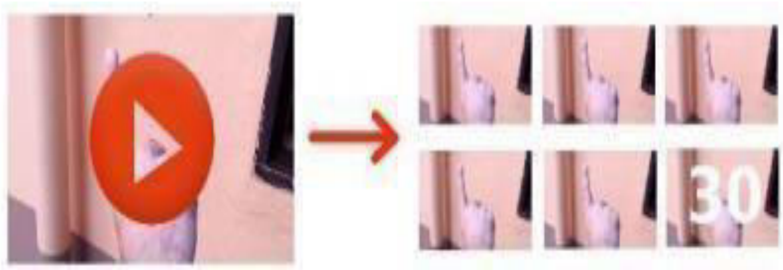


Figure 3: Recognition of continues image

b. Capturing Images in Varied Backgrounds

To address the limitations of the first approach, a more diverse dataset was created. This time, we ensured that specific gestures were included to facilitate system control. We captured four distinct hand poses, as shown in Figure 4. The objective was to train the model to accurately recognize fingertips from multiple fingers, allowing users to interact with the system through different hand gestures. The system was designed so that:

- Writing could be performed by showing one index finger.
- The text could be converted into e-text by displaying two fingers.
- A three-finger gesture was used to add spaces.
- Five fingers were used for backspacing.
- Entering prediction mode required showing four fingers, and then users could select the 1st, 2nd, or 3rd prediction by displaying one, two, or three fingers, respectively.
- To exit prediction mode, users needed to show five fingers.

There were 1,800 photos in this collection. The dataset was auto-labeled by the previously trained model using a script. Before a new model was trained, all incorrectly labeled photos were manually fixed. Unlike the prior model, this method worked well over a range of backdrops and attained 94% accuracy.

Figure 4: Capturing Images in Different Backgrounds

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

D. Training of a Fingertip Recognition Model

Once the dataset was labelled, it was divided into training and validation sets (85%-15%). The dataset was trained using Single Shot Detector (SSD) and Faster R-CNN pre-trained models. While SSD offers faster detection by combining region proposal and classification into a single step, Faster R-CNN proved to be more accurate.

Refer to the Results Section for further information SSD is well-suited for real-time applications as it performs region proposal and classification simultaneously, accelerating object identification. Conversely, Faster R-CNN employs a Region Proposal Network (RPN) to extract potential regions of interest from feature maps. These regions undergo processing through a Region of Interest (RoI) pooling layer before being analyzed by fully connected layers for classification and bounding box regression [15]. For this specific application, Faster R-CNN's final fully connected layer was optimized to achieve high accuracy in finger detection within images. A Region of Interest (Roi) pooling layer processes these regions before sending them to fully connected layers for bounding box regression and classification [15]. Faster R-CNN's last fully connected layer was optimized to recognize fingers in pictures with accuracy.

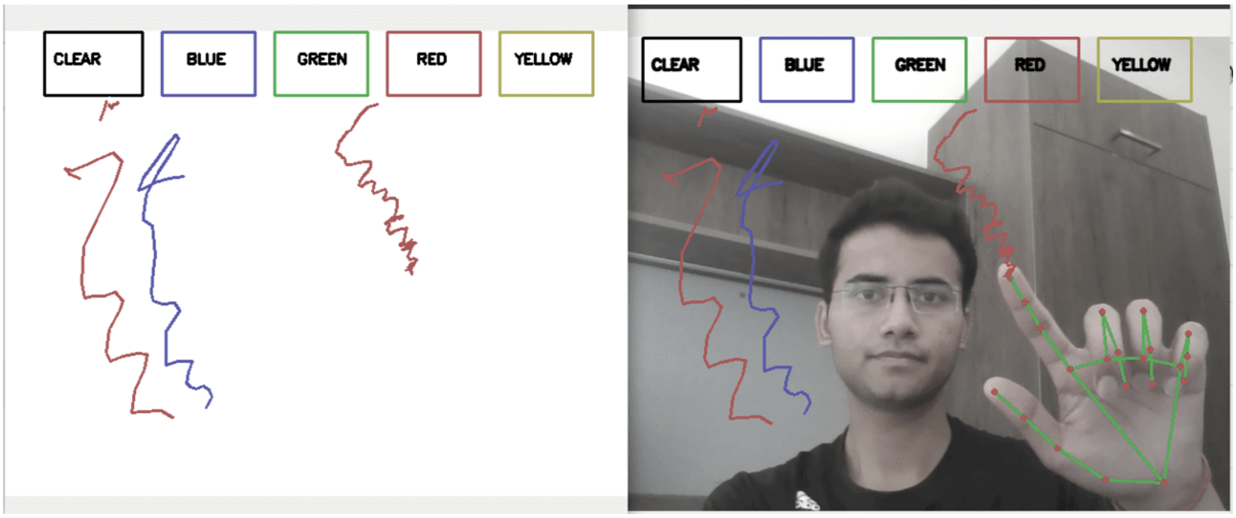


Figure 5: A traced word in the air on a black background

VI. WORKFLOW ALGORITHM

This section covers the most engaging aspect of our system. Writing requires multiple functionalities, so the number of gestures used to operate the system corresponds to these different actions. The essential features integrated into our system include:

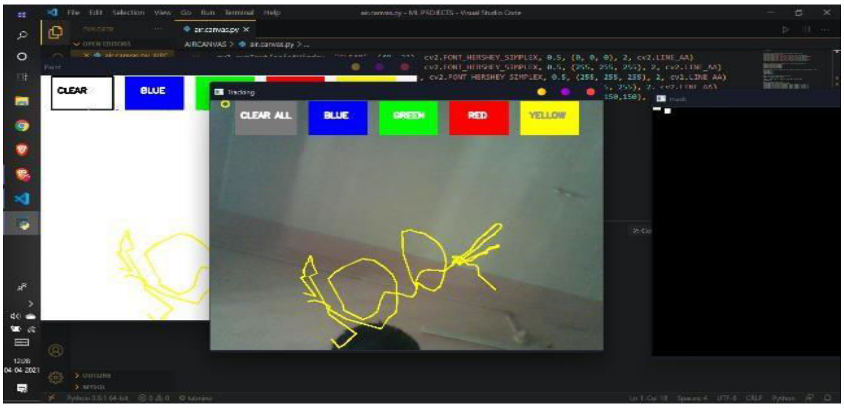


Figure 6: Real time image acquisition to system

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

1. **Writing Mode** – In this mode, the system tracks the fingertip's movement and records the corresponding coordinates.
2. **Colour Mode** – Users can switch between different text colours from a predefined selection.
3. **Backspace** – A designated gesture allows users to quickly delete text in case of mistakes.

VII. CONCLUSION

The system introduces a contemporary alternative to conventional writing methods, removing the necessity of physically holding a smartphone for note-taking purposes. It offers a convenient, mobile solution for writing that can be used anywhere, anytime. Beyond convenience, it serves as a significant communication aid for individuals with disabilities who may face challenges with traditional input methods.The technology also provides substantial benefits for senior citizens and users who experience difficulties with standard keyboards[19]. These demographics often struggle with the precision required for small touchscreen keyboards or the dexterity needed for physical keyboards. By allowing text input through natural hand movements in the air, the system creates a more accessible and intuitive interface that reduces the technical barriers often encountered by these user groups, potentially increasing their digital inclusion and independence. In the future, the system's functionality could be extended to manage Internet of Things devices. Additionally, features that pull air can be included. If the software is integrated into smart wearables, users might be able to interact with the digital world more effectively. Augmented Reality (AR), which makes text appear dynamically, can further enhance the experience.

Many of these limitations can be addressed in future studies. First off, word-by-word input would be possible with a handwriting recognizer rather than a character recognizer, which would increase efficiency. Second, real-time system controls could be controlled by hand gestures with pauses, as demonstrated in [21], rather than counting the number of fingers displayed. Thirdly, the system can occasionally change its state when it detects unwanted fingertips in the background. Air-writing systems should solely react to the gestures of the designated user and not be influenced by other people in the vicinity in order to guarantee accuracy.

The dataset is another area that needs work. The EMNIST dataset, which was not created especially for air-character recognition, is now used by the system. Advanced object detection models such as YOLO v3 could improve speed and accuracy in future deployments. The effectiveness and dependability of air-writing technology will greatly increase with ongoing developments in artificial intelligence.

REFERENCES

[1] Y. Huang, X. Liu, X. Zhang, and L. Jin, "A Pointing Gesture Based Egocentric Interaction System: Dataset, Approach, and Application," 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Las Vegas, NV, pp. 370-377, 2016.
[2] P. Ramasamy, G. Prabhu, and R. Srinivasan, "An economical air writing system converting finger movements to text using a web camera," 2016 International Conference on Recent Trends in Information Technology (ICRTIT), Chennai, pp. 1-6, 2016.
[3] Saira Beg, M. Fahad Khan, and Faisal Baig, "Text Writing in Air," Journal of Information Display, Vol. 14, Issue 4, 2013.
[4] Alper Yilmaz, Omar Javed, and Mubarak Shah, "Object Tracking: A Survey," ACM Computer Survey, Vol. 38, Issue 4, Article 13, pp. 1-45, 2006.
[5] Yuan-Hsiang Chang, Chen-Ming Chang, "Automatic Hand-Pose Trajectory Tracking System Using Video Sequences," INTECH, pp. 132-152, Croatia, 2010.
[6] Erik B. Sudderth, Michael I. Mandel, William T. Freeman, and Alan S. Willsey, "Visual Hand Tracking Using Nonparametric Belief Propagation," MIT Laboratory for Information & Decision Systems Technical Report P2603, IEEE CVPR Workshop on Generative Model-Based Vision, pp. 1-9, 2004.
[7] T. Grossman, R. Balakrishnan, G. Kurtenbach, G. Fitzmaurice, A. Khan, and B. Buxton, "Creating Principal 3D Curves with Digital Tape Drawing," Proc. Conf. Human Factors Computing Systems (CHI' 02), pp. 121- 128, 2002.
[8] T. A. C. Bragatto, G. I. S. Ruas, M. V. Lamar, "Real-time Video-Based Finger Spelling Recognition System Using Low Computational Complexity Artificial Neural Networks," IEEE ITS, pp. 393-397, 2006.
[9] Yusuke Araga, Makoto Shirabayashi, Keishi Kaida, and Hiroomi Hikawa, "Real-Time Gesture Recognition System

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025| DOI: 10.15680/IJIRCCE.2025.1304022

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

Using Posture Classifier and Jordan Recurrent Neural Network," IEEE World Congress on Computational Intelligence, Brisbane, Australia, 2012.
[10] Ruiduo Yang, Sudeep Sarkar, "Coupled Grouping and Matching for Sign and Gesture Recognition," Computer Vision and Image Understanding, Elsevier, 2008.
[11] R. Wang, S. Paris, and J. Popovic, "6D Hands: Marker less Hand-Tracking for Computer-Aided Design," in Proc. 24th Ann.ACM Symp. User Interface Soft. Technol., 2011, pp. 549–558.
[12] Maryam Khosravi Nahouji, "2D Finger Motion Tracking, Implementation for Android- Based Smartphones," Master's Thesis, CHALMERS Applied Information Technology, 2012, pp. 1-48.
[13] Eshed Ohn-Bar, Mohan Manubhai Trivedi, "Hand Gesture Recognition In Real-Time For Automotive Interfaces," IEEE Transactions on Intelligent Transportation Systems, Vol. 15, No. 6, December 2014, pp. 2368-2377.
[14] P. Ramasamy, G. Prabhu, and R. Srinivasan, "An economical air writing system converting finger movements to text using a web camera," 2016 International Conference on Recent Trends in Information Technology (ICRTIT), Chennai, 2016, pp. 1-6.
[15] Kenji Oka, Yoichi Sato, and Hideki Koike, "Real-Time Fingertip Tracking and Gesture Recognition," IEEE Computer Graphics and Applications, 2002, pp. 64-71.
[16] H.M. Cooper, "Sign Language Recognition: Generalizing to More Complex Corpora," Ph.D. Thesis, Centre for Vision, Speech and Signal Processing, Faculty of Engineering and Physical Sciences, University of Surrey, UK, 2012.
[17] Y. Huang, X. Liu, X. Zhang, and L. Jin, "A Pointing Gesture-Based Egocentric Interaction System: Dataset, Approach, and Application," 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Las Vegas, NV, pp. 370-377, 2016.
[18] Vladimir I. Pavlovic, Rajeev Sharma, and Thomas S. Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, July 1997, pp. 677-695.
[19] Guo-Zhen Wang, Yi-Pai Huang, Tian- Sheeran Chang, and Tsu-Han Chen, "Bare Finger 3D Air-Touch System Using an Embedded Optical Sensor Array for Mobile
isplays," Journal of Display Technology, Vol. 10, No. 1, January 2014, pp. 13-18.
[20] Napa Sae-Bae, Kowsar Ahmed, Katherine Isbister, and Nasir Memon, "Biometric-Rich Gestures: A Novel Approach to Authentication on Multi- Touch Devices," Proc. SIGCHI Conference on Human Factors in Computing Systems, 2005, pp. 977-986.
[21] W. Makela, "Working 3D Meshes and Particles with Finger Tips, Towards an Immersive Artists' Interface," Proc. IEEE Virtual Reality Workshop, pp. 77-80, 2005.
[22] A.D. Gregory, S.A. Ehmann, and M.C. Lin, "InTouch: Interactive Multiresolution Modelling and 3D Painting with a Haptic Interface," Proc. IEEE Virtual Reality (VR' 02), pp. 45-52, 2000.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details