

EXPERT MATHEMATICIAN

Ansari Ariba Tabrez¹, Khan Areeba Ashfaq², Khot Zaiba Arshad³, Zaibunnisa Malik⁴

^{1,2,3,4}Dept of Computer Engg, India.

ansariariba29@gmail.com khanareeba3007@gmail.com zaibakhot1@gmail.com zebamalik@yahoo.com

DOI: <https://www.doi.org/10.58257/IJPREMS36955>

ABSTRACT

Recent advancements in hardware and internet connectivity have enabled the widespread application of AI and big data in fields like computer vision. This thesis presents a stereo vision-based handwriting recognition system that tracks finger movements in 3D space using depth data, improving accuracy with techniques such as Particle Swarm Optimization. Additionally, it explores the growing role of gesture-controlled systems in human-computer interaction (HCI), showcasing a Virtual Calculator that leverages hand gestures for calculations and a personalized interface for custom gesture input. The thesis also introduces a deep learning approach for air-writing recognition, which aids communication for the deaf, and highlights gesture-driven creative applications like the Air Canvas. Finally, it reviews optimization strategies for large language models (LLMs) and the role of code-based self-verification in enhancing computational tasks

1. INTRODUCTION

Air writing is a dynamic gesture recognition method where users write alphanumeric characters using hand or finger movements in 3D space, ideal for scenarios where typing or touch interfaces are impractical. Gesture recognition systems are broadly categorized as device-based and device-free, with the latter offering greater convenience [1][10]. While 3D camera systems like Kinect or Leap Motion provide superior finger tracking, they come with higher costs. To address this, a cost-effective solution using CNNs and web cameras has been proposed, achieving over 99% real-time accuracy [11][28]. Hand gesture recognition (HGR) interprets palm and finger movements for human-computer interaction (HCI), with applications in robotics, virtual reality, and sign language. While user-defined gestures increase flexibility, challenges persist in 3D tracking with 2D cameras. A Mediapipe-based detection system with a lightweight neural network showed effectiveness in a user study [2]. Air writing, where users trace letters or words in the air, is gaining traction in user interfaces, particularly where traditional typing isn't feasible [3]. Tracking continuous finger motion for air-writing recognition is challenging due to stray movements. A proposed system enhances accuracy by using motion streams, segmentation, and stray movement filtering [5][6]. The role of gesture recognition in HCI has shifted control from traditional devices to speech, emotions, and body gestures [21] [30]. A CNN-based classifier using transfer learning from VGG16 was used to recognize static and dynamic gestures for computer control [7]. Neo-cybernetics bridges logic-based AI and behavior-based robotics, integrating non-linear dynamics and feedback for generative AI (GAI) [8]. Large language models (LLMs) like ChatGPT and GPT-4 now assist in math problem-solving by enhancing reasoning and step-by-step calculations [14]. A novel calculator application based on handwriting recognition was developed to improve user experience by allowing real-time adaptation to individual writing styles without preliminary samples [13]. Early gesture recognition relied on gloves or sensors, but modern systems now use skin color for hand segmentation. SVM methods with morphological operations outperform neural networks in both static and dynamic gesture recognition [15].

The rapid advancement of technology, particularly Virtual Reality (VR), is transforming education, making tools like the Virtual Mouse and Virtual Calculator accessible via gesture recognition for hands-free, hygienic interaction [16][18]. Innovations like asymmetric VR promote collaborative experiences, even without a headset for each student [17]. Gesture-based tools, including a virtual calculator designed for the metaverse, highlight immersive, intuitive interaction beyond traditional methods [19]. Human-Computer Interaction (HCI) is undergoing a significant transformation with advancements in Natural Hand Gesture Recognition (HGR), enabling users to interact with machines without the need for additional devices [20] [22]. Handwritten Mathematical Expression (HME) Recognition, which is essential for fields such as education and engineering, continues to challenge researchers due to the complexity of recognizing symbols in two-dimensional space [27] [30].

2. PROPOSED METHODS

The proposed method develops an air-handwriting recognition system using stereoscopic vision, grayscale imaging, and deep learning to enhance real-time hand gesture detection in indoor environments. The goal is to create a reliable, touchless system for accurately tracking finger movements and interpreting handwriting gestures without physical input devices. Key steps involved include:

A. Trajectory Acquisition

The first stage of the method involves trajectory acquisition by capturing 2D images of the user's writing motion with a standard webcam to track hand movement. The system uses skin color detection and moving feature analysis, applying the Camshift algorithm for real-time tracking under varying lighting and skin tones [1]. Activation occurs upon face detection, enabling frame capture as the user writes. It accurately identifies stroke start and end points, addressing the "push-to-write" problem and eliminating the need for explicit delimiters [2][3].

B. Data Processing

After trajectory acquisition, the captured data undergoes processing for machine learning. The first transformation converts the original 640x480 image to 360x360 pixels, then further reduces it to 36x36 pixels for training efficiency [4]. The second transformation normalizes coordinates to a range of (-1, 1), preserving character aspect ratios. The trajectory data is organized in two formats: 1D padding, creating a (1, 256) array, and 2D no-pad, forming a (2, 100) array without padding [5][6][7]. These steps prepare the data for the neural network.

C. Network Design

The final stage involves designing a convolutional neural network (CNN) to recognize user-defined hand gestures from trajectory data. Both 1D and 2D CNN architectures include multiple convolutional blocks with layers for convolution, max pooling, batch normalization, and ReLU activation to enhance feature extraction [8]. Training utilizes the Adam optimizer, which adapts learning rates based on historical gradients [9]. This structured approach creates a comprehensive framework for robust hand gesture recognition [10][22][29].

D. Dynamic Hand State & Gesture Recognition

Recognizing dynamic hand gestures is inherently challenging; however, the Kinect sensor's ability to capture depth images significantly aids this process. For instance, when a user waves their hand, the Kinect effectively tracks the joint positions, marking any detected joint as "tracked," thus enabling real-time gesture recognition [11][12].

E. Hand Area and Movement Detection

This section outlines motion functions like "Go Left," "Go Right," and others that users can perform via hand gestures, with the system using Euclidean distance to validate gesture commands effectively [13]. A complementary study developed a system for air handwriting that distinguishes the palm from background noise through continuous image subtraction and morphological analysis [14]. While established printed character recognition systems like Omnipage and Tesseract exist, an Android-based system utilizes Artificial Neural Networks (ANN) and Dynamic Time Warping (DTW) for efficient character classification [15]. Although large language models (LLMs) have improved complex reasoning, challenges remain in mathematical reasoning, prompting techniques like code-based self-verification (CSV) to enhance LLM performance [16]. Gesture recognition systems, especially those using Support Vector Machines (SVM), involve feature extraction, image preprocessing, and classification, employing Fourier descriptors for accurate gesture recognition [17] [18].

F. Gesture Recognition Research Field Analysis Using Bibliometrics

A gesture-controlled virtual mouse system showcases the versatility of hand tracking technology in gaming and accessibility, utilizing OpenCV and MediaPipe for intuitive finger movement control [19][20]. Scientific mapping methods like co-authorship and keyword co-occurrence analysis identify trends and gaps in gesture recognition research, offering insights for future innovations [21].

G. Workflow and Fingertip Detection Model for Air Canvas

The use of hand gestures for artistic creation is gaining popularity, exemplified by an air-canvas system through marker-based tracking. This system utilizes OpenCV for live video capture and NumPy arrays to create a digital canvas, translating hand movements into drawings [24]. Additionally, fingertip detection models have been developed to improve accuracy in gesture recognition [25].

H. Large Language Model (LLM) Optimization Systematic Review

Integrating large language models (LLMs) with gesture recognition systems is an emerging research area. The review emphasizes the potential for this integration through efficient algorithms and hardware optimizations, especially for voice-gesture hybrid interfaces in smart devices and robotics [26].

I. Gesture Recognition in Air-Writing Systems and Motion-to-Text Conversion for Wearable Devices

Air-writing systems convert hand gestures into written text, merging gesture recognition with natural language processing. Methods like hybrid motion tracking, RFID, and radar-based systems have achieved over 90% accuracy in recognizing air-written characters [28]. Motion-to-text systems for wearable devices allow users to input text or control smart devices by tracking finger movements with computer vision techniques [30].

3. RESULTS

A. Air-Writing Gesture Recognition for Smart TV Control

A custom dataset was created due to the lack of a public one, containing digits and directional symbols, with data collected from volunteers aged 20-30. Using CNN models with trajectory data, a 5-fold cross-validation approach yielded the best recognition rate, optimizing for both recognition accuracy and reduced network complexity [1].

B. Touchless Fingerwriting System

Utilizes 3D finger tracking to segment writing activity from continuous motion data, achieving a 96.4% true positive rate and a 5.1% false alarm rate. The system effectively recognized words and letters, with error rates of 1.15% and 9.84% for word-based and letter-based recognition, respectively [5].

C. CNN for Gesture Recognition

With a 54x54 binary image dataset and preprocessing steps like background subtraction and skin filtering, the CNN model achieved 93.25% accuracy on the training set and 98.44% on validation, enabling mouse and keyboard operations through gesture recognition [6].

D. Dynamic Gesture Recognition Using Kinect

Implemented for recognizing Arabic numbers and letters using SVM, achieving a 95.42% average recognition rate, with improvements focused on feature extraction and advanced algorithms for robustness [10] [11].

E. Palm and Fingertip Detection System

This system uses particle swarm algorithms for stable palm detection and tracks palms accurately at distances of 0.8 and 1 meter, though fingertip recognition faced challenges due to interference [12].

F. Hand gesture-based virtual mouse and calculator system

A hand gesture-based virtual mouse and calculator system uses fingertip detection via MediaPipe to enable touchless interaction for functions like clicking, scrolling, volume, and brightness control. This system allows users to perform calculations and interact with digital interfaces intuitively, making it accessible for users with mobility challenges [17] [19].

G. The GREFIT system

The GREFIT system excels in fingertip detection with an accuracy of 90-95% for open fingers, though performance decreases with closed fingers due to visual overlap. Various techniques like HMM, RNN, and Gabor filters help enhance detection in diverse gesture applications [20].

H. Air-Writing System Developed With Opencv

Another air-writing system, developed with OpenCV in Anaconda, enables precise pen tracking and real-time gesture-based drawing, featuring interactive elements like clearing and color changes for enhanced responsiveness [28]. A recognition model trained on the EMNIST dataset achieved high accuracy—99% for digits and 98.77% for alphabets—effectively distinguishing similar characters (e.g., "O" and "0") and demonstrating robust performance across a large dataset [29].

4. DISCUSSIONS

A. Advancements in Air-Writing Systems and Real-Time Gesture Recognition Technologies

Recent innovations in air-writing systems for smart device control highlight the effectiveness of deep learning models in gesture recognition. One study created specialized datasets for digits and directional symbols, employing 1D and 2D Convolutional Neural Networks (CNNs) that outperformed traditional methods [1]. Another approach utilized a Mediapipe-based framework and a Multi-Layer Perceptron (MLP) architecture for customizable hand gestures, achieving high accuracy but encountering challenges with static gesture recognition and complex user customization [2][3].

B. Pre-processing Techniques

Pre-processing techniques are essential for improving gesture recognition accuracy. Systems using skin detection and noise reduction have achieved validation accuracies above 98%, highlighting the importance of effective pre-processing [6][7]. Gesture recognition engines, combined with pre-processing and feature extraction, allow users to perform algebraic operations using hand gestures [13].

C. Specialized Applications

Gesture recognition is expanding into specialized applications. The Monk system exemplifies this by utilizing image processing for large-scale word searches in historical handwritten texts, benefiting from user collaboration to refine its

capabilities [8]. In education, generative AI chatbots like ChatGPT and Wolfram Alpha are transforming math instruction by providing personalized support, thus enhancing student engagement [9][14].

D. Evolving Gesture Recognition Methodologies

The advancements in Kinect-based gesture recognition systems reflect improvements in hardware and software. Various methodologies, such as Support Vector Machines (SVMs) and Hidden Markov Models (HMMs), have shown notable progress in recognizing dynamic gestures [10]. Real-time gesture recognition systems have become increasingly effective in facilitating intuitive human-computer interaction [11]. Leading research efforts have focused on AI techniques, including Deep Learning (DL), Machine Learning (ML), and models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Support Vector Machines (SVMs) [21][25].

E. Advances in 3D Gesture Tracking for Handwriting Recognition

Recent advancements in stereo vision technology and depth-based tracking have enhanced the ability to accurately recognize gestures, including handwriting in 3D space. By overcoming the limitations of infrared-based systems, these solutions leverage advanced algorithms like Particle Swarm Optimization and Probability Density Functions for automatic gesture detection [12][5].

F. Gesture Recognition for Enhancing Accessibility

Hand gesture recognition technologies offer significant improvements in accessibility, especially for individuals with disabilities [17]. By replacing traditional input methods with gesture-based control systems, users can interact with devices more naturally [18]. The accuracy of gesture recognition systems is highly dependent on factors such as image quality, lighting, and system design [19]. While many systems achieve impressive accuracy rates, challenges remain in recognizing more complex, dynamic gestures, particularly under less-than-ideal conditions [20].

G. Optimization and Acceleration Techniques for Large Language Models (LLMs)

Training large language models (LLMs) is intensive, but optimization techniques have improved efficiency. Frameworks like Megatron-LM, ByteTransformer, and LightSeq2 enhance training by distributing tasks across multiple GPUs and using mixed-precision techniques [26]. Handwriting mathematical expression (HME) recognition has also advanced significantly over the past four decades, evolving from initial prototypes to sophisticated integrated solutions [27].

H. Evolution of Air Writing and Character Recognition

Air writing enables users to write in mid-air using hand gestures, becoming a key focus in gesture recognition. This technology has practical applications in areas where traditional input devices are impractical, such as medical rehabilitation and smart home controls [28]. Neural networks have shown exceptional character recognition performance, with a model trained on the EMNIST dataset achieving 99% accuracy for digits and 98.77% for alphabets, effectively generalizing to unseen data [29].

5. CONCLUSION

The review paper explores advancements in gesture recognition, particularly air-writing systems, and their applications in Human-Computer Interaction (HCI) and education. It highlights the shift from traditional device-based methods to more intuitive, device-free systems powered by deep learning and computer vision. Air-writing systems allow users to write alphanumeric characters in the air, with CNN-based frameworks enhancing real-time accuracy. The paper reviews various gesture recognition technologies, noting their effectiveness in environments where traditional input devices are impractical. Additionally, it discusses the impact of gesture recognition on education and HCI, particularly in enhancing virtual learning and math education using VR and AI-based systems. Despite these advancements, challenges such as high costs, infrastructure needs, and equitable access remain. The authors recommend future research to improve accuracy, develop adaptive systems, and optimize gesture recognition for edge devices. The paper concludes by emphasizing the potential of these technologies to revolutionize education, smart devices, and virtual reality while acknowledging barriers to widespread adoption.

6. REFERENCES

- [1] Chaur-Heh Hsieh, You-Shen Lo, Jen-Yang Chen, Sheng-Kai Tang, Air-Writing Recognition Based on Deep Convolutional Neural Networks, IEEE, 2021
- [2] Jindi Wang, Ioannis Lvrissimtzis, Zhaoxing Li, Lei Shi, Hand Gesture Recognition for User-Defined Textual Inputs and Gestures, SPRINGER, 2024
- [3] Ekta R S, Kajal C K, Manali V P, Vaishnavi M K, Shraddha S K, Pranali M S, Karad, A Survey on Air Writing Character Recognition and Translation, IJEAST, 2023

- [4] Christina M. Krause, The Mathematics in Our Hands: How Gestures Contribute to Constructing Mathematical Knowledge, SPRINGER, 2016
- [5] Mingyu Chen, Ghassan AlRegib, and Biing-Hwang Juang, Air-Writing Recognition—Part II: Detection and Recognition of Writing Activity in Continuous Stream of Motion Data, IEEE , 2015
- [6] Md. Milon Islam, Md. Repon Islam, Md. Saiful Islam, An Efficient Human-Computer Interaction through Hand Gesture Using Deep Convolutional Neural Network, SPRINGER, 2020
- [7] Soeb Hussain, Rupal Saxena, Xie Han, Jameel Ahmed Khan, Hyunchul Shin, Hand Gesture Recognition Using Deep Learning, IEEE, 2017
- [8] Tijin van der Zant, Matthijs Kouw, Lambert Schomaker, Generative Artificial Intelligence, SPRINGER, 2012
- [9] Niroj Dahal, Binod Prasad Pant, Bal Chandra Luitel, Exploring Capabilities and Limitations of Generative AI Chatbots in Solving Math, 2024
- [10] Niroj Dahal, Bal Chandra Luitel, Binod Prasad Pant, A Real-Time Dynamic Hand Gesture Recognition System Using Kinect Sensor, IEEE, 2016
- [11] Yanmei Chen, Bing Luo, Yen-Lun Chen, Guoyuan Liang, and Xinyu Wu, A Novel Dynamic Hand Gesture and Movement Trajectory Recognition model for Non-Touch HRI Interface, IEEE Access, 2019
- [12] Ying-Nong Chen, Chi-Hung Chuan, Kuo-Chin Fan, Stereo-Based 3D Space Handwriting Recognition, IEEE Access, 2018
- [13] Barna szocs, Szilard vajda, Judit robu, D.A.C Draw and Calc - the intuitive calculator, IEEE Access, 2012
- [14] Vamsi Krishna, Kumar Gompa, Solving Math Reasoning Using Various LLMs, International Journal of Research Publication and Reviews, 2023
- [15] Preeti Niranjana, Brijesh Pandey, F. Masooma Nigar, Virtual Calculator using Hand Gesture Recognition via Support Vector Machine, IJRSET, 2016
- [16] Lui Albæk Thomsen, Virtual Reality In Mathematics education (VRiME), Aalborg University Press, 2023
- [17] Dr. B. Esther Sunanda, Vaddi Bhargavi, Varikuti Yasasri, Batta Lahitha, Chappidi Suma Sri, Chellapalli Madhuri, Hand Gesture Based Virtual Mouse and Calculator, SSRN, 2024
- [18] Poornima R, Elakkiya M P, Kannan S, Akash B, Dhilip R M, HAND GESTURE BASED VIRTUAL CALCULATOR USING PYTHON FOR METAVERSE, IJSREM, 2023
- [19] M. Hemalatha, V. Sreeja, S. Aswathi, Creating Air Canvas Using Computer Vision, IJARCCCE, 2024
- [20] Ankit Chaudhary, J. L. Raheja, Karen Das, Sonia Raheja, Intelligent Approaches to interact with Machines using Hand Gesture Recognition in Natural way: A Survey, IJCSES, 2011
- [21] Juan Jesus Ojeda-Castelo, Maria De Las Mercedes Capobianco-Uriarte, Jose Antonio Piedra-Fernandez, Rosa Ayala, A Survey on Intelligent Gesture Recognition Techniques, IEEE, 2022
- [22] Sandeep Kumar, Mohit Tanwar, Anand Kumar, Prashant Chamoli, Hand Gesture Recognition Based Calculator, IJCSMC, 2019
- [23] Chaur-Heh Hsieh, You-Shen Lo, Jen-Yang Chen, Sheng-Kai Tang, Air-Writing Recognition Based on Deep Convolutional Neural Networks, IEEE, 2021
- [24] Prof. S.U. Saoji, Akash Kumar Choudhary, Nishtha Dua, Bharat Phogat, AIR CANVAS APPLICATION USING OPENCV AND NUMPY IN PYTHON, IRJET, 2021
- [25] Adinarayana Salina, K. Kaivalya, K. Sriharsha, K. Praveen, M. Nirosha, Creating Air Canvas Using Computer Vision, IJAEM, 2022
- [26] Zhyar Rzgar K. Rostam, Sándor Szénási, Gábor Kertész, Achieving Peak Performance for Large Language Models: A Systematic Review, IEEE, 2024
- [27] Dmytro Zhelezniakov, Viktor Zaytsev, Olga Radyvonenko, Online Handwritten Mathematical Expression Recognition and Applications: A Survey, IEEE, 2021
- [28] Amith K R, Nikhil Holla R, Prashanth J, Air Writing Detection and Recognition, IJARSC, 2024
- [29] Ms. M.Saranya, Kariketi Tharun Reddy, Madhumitha Raju, Manoj Kutala, Air-Writing Alphanumeric Character Recognition, IJCRT, 2022
- [30] Prof. Hemlata A. Shinde, Shravan M. Jagtap, Anushka A. Kalpund, Pranita B. More, Air Canvas: Drawing in Air using AI, IJRSET, 2022