International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 03 | March - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Virtual Air Canvas: A Review

Akhil Mohan[1], Leshma Mohandas[2], Prabath P U[3], Saeedh Mohammed[4], Sreema E R[5]

[1,2,3,4]Student, Department of computer science and engineering, Universal Engineering College, Kerala, India.
[5]Assistant Professor, Department of computer science and engineering, Universal Engineering College, Kerala, India.

---***---

Abstract - This creative method of utilizing open-source frameworks to create gesture-controlled whiteboard devices is impressive. The device tracks and interprets finger movements collected by a camera to allow users to engage with a virtual whiteboard. The implementation includes effective data processing, hand landmark estimate, and real-time hand detection by utilizing computer vision algorithms. By providing a modern form of communication and going beyond regular typing and writing habits, the system seeks to expand beyond traditional writing methods and create new opportunities for connection. The system contributes to the changing field of human-computer interaction by offering an alternate mode of communication, which improves people's quality of life overall. It also demonstrates the versatility of gesture-controlled systems in terms of enhancing communication accessibility.

***Key Words*:** Data Processing, Handmark Estimation, Computer Vision algorithms, MediaPipe, OpenCV.

1. INTRODUCTION

In recent times, the science of image processing and pattern recognition has witnessed the exciting and captivating research of writing in the air. In many different operations, it can improve the interface between a machine and a human and makes a significant contribution to the automation process's advancement. Numerous research studies have focused on novel approaches and designs that offer enhanced recognition accuracy at a lower processing time.

The system's virtual whiteboard, which sets a new benchmark for the accuracy and clarity of lines in digital airwriting, is one of its standout features. It does this by combining hand movements and contemporary fonts. This virtual canvas changes the way users interact with digital art by giving them a fluid and organic medium for artistic expression and enabling them to use their hands as virtual paintbrush. Through natural hand movements, users can express themselves through drawing, writing, and other media,redefining the creative process.

The accuracy of Virtual Air Canvas is what sets it apart. It records the precise movement of your fingertips in addition to recognizing gestures, enabling intricate and subtle digital creations. With the use of state-of-the-art technologies like MediaPipe and OpenCV, it can track finger movements and interpret hand gestures with remarkable accuracy, resulting in an incredibly lifelike experience.

Users can precisely and subtly create their digital creations because of the system's comprehensive tracking of fingertip movements. This level of detail not only expands the platform's creative possibilities but also creates new opportunities for exact control and manipulation in the virtual environment. Virtual Air Canvas is driven by state-of-the-art technology and makes use of MediaPipe and OpenCV. With surprising accuracy, the system can comprehend and react to the user's hand motions thanks to OpenCV, a popular computer vision framework. In addition, MediaPipe technology makes it easier to track finger movements, which guarantees a smooth and organic extension of the user's intentions onto the digital canvas.

Character recognition in the context of digital art creation has undergone a revolutionary change with the system's incorporation of Optical Character Recognition (OCR) techniques. The accuracy and dependability of character recognition are increased to new heights with this creative enhancement. The technology can accurately separate individual letters from continuous air-written movements by using OCR, giving users a more responsive and sophisticated experience. The addition of OCR not only improves character recognition's technical elements but also gives users more creative options. The capacity to effectively combine textual components with graphic aspects in digital artwork creates a strong link between written and visual expression. Now, users can easily add text, messages, or complex designs to their artistic works, adding to the story.

By skillfully combining computer vision, gesture recognition, and OCR technologies, the system, which prides itself on being at the forefront of innovation, redefines the creative process in digital art. As a result, users may create aesthetically appealing compositions with just their hand motions on an immersive and user-friendly platform. With its contemporary typefaces, crisp lines, and markerless virtual whiteboard features, this system is a pathfinder that has raised the bar for interactive and expressive digital art experiences.

This system stands out as a trailblazing force, pushing the limits of what is possible at the nexus of technology and artistic creativity, as the terrain of virtual artistic expression continues to change.

2. RELATED WORK

Taiki Watanabe and their team [1] investigate air-writing as a kind of human-computer interaction, an area of study that has attracted a lot of interest lately. The key component of air-writing that bridges the gap between natural language and digital interfaces is gesture recognition. The researchers propose robust character recognition systems as a solution to the problems caused by differences in writing styles and articulation speeds. Due to the shortcomings of earlier approaches utilizing gyroscopes, accelerometers, or computer vision, the authors suggest a unique strategy utilizing a web camera-based air-writing system. They provide a hybrid deep learning model that combines BiLSTM models for sequential data processing with CNNs for image-based recognition. The work intends to improve air-writing recognition, contribute to human-computer interaction paradigms, and solve the existing lack of comprehensive and accurate air-writing recognition systems. It emphasizes the value of publicly available datasets, such as the 6DMG dataset.

Saez-Mingorance, Mendez-Gomez, Mauro, Castillo Morales, Pegalajar Cuellar, and Morales-Santos [2] discuss the changing environment of human-computer interaction, focusing on the incorporation of natural communication interfaces. They concentrate on air-writing, a captivating application in which users express characters using hand movements in open space. The work takes a novel approach to air-writing recognition by using ultrasonic transceivers to capture hand movement trajectories, deviating from existing methods and providing a more efficient solution. The main contribution is a proposal for an air-writing device that uses ultrasonic technology to improve character identification accuracy. Convolutional LSTM is the most effective deep learning algorithm, outperforming CNN, LSTM, convolutional autoencoder, and convolutional LSTM, according to the authors' evaluation. Its remarkable accuracy of 99.51% makes it the clear winner. Validation in real environments demonstrates the usefulness of the suggested system and validates its effectiveness in gathering and classifying data in real-world situations. The development of reliable and effective air-writing recognition systems has improved significantly with the use of ultrasonic transceivers and sophisticated deep- learning algorithms.

Chen. H and their team [3] ground-breaking acoustic wave air-writing method transforms human-machine interaction. Their method provides a natural means of communication by interpreting hand movements traced in the air into instructions for machinery. Motion tracking, which is essential for these kinds of devices, is accomplished using a new technique that combines Direction of Arrival (DOA) data with an ultrasonic receiver array that tracks a wearable transmitter. With the introduction of the PDP method, precise tracking using a three-sensor receiver array is made possible by its efficient estimation of 2-D DOA. This strategy has the potential to improve air-writing systems' effectiveness.

Alam, Kwon, Md Imtiaz, Hossain, Kang, and Kim [4] present the TARNet architecture for air-writing recognition, in which CNN and LSTM perform different roles as feature generators and recognizers, respectively. The network extracts local contextual elements from low-level trajectory data using 1-dimensional separable convolution and then captures high- level dependencies using LSTM. TARNet performs well on air- writing datasets, with accuracies of 99.6% for RTD, 98.74% for RTC, and 95.62% for smart-band datasets. The architecture balances time efficiency and performance, demonstrating its success in a variety of air-writing circumstances. TARNet beats solo CNN and LSTM models by offering a complete solution for trajectory-based air-writing recognition.

Grigoris Bastas and Kosmas Kritsis [5] describe a significant development in human-computer interaction by allowing air-writing with off-the-shelf smart bands, breaking away from the previous dependency on specialist devices. The breakthrough resides in recording motion signals using a smart band worn by the user, removing the need for additional equipment. The strategy prioritizes user-friendly interfaces, in line with current trends. Extensive data collection involving 55 subjects substantiates the effectiveness of their methods, with the user-dependent approach achieving an average accuracy of 89.2%, and the user-independent method displaying robustness with an average accuracy of 83.2%. This study, when viewed within the larger framework of air-writing recognition, highlights the unique contribution of using smart bands to create a globally applicable interface. The study's results highlight the potential of motion data analysis from smart- bands, paving the way for further advancements in gesture- based human-computer interaction.

3. DESIGN AND ANALYSIS

3.1. 2D CAMERA-BASED AIR-WRITING IDENTIFICATION

The air-writing system collects character data and uses a hybrid deep learning strategy to train models by integrating CNN and BiLSTM models for character recognition, as well as hand joint estimation, data pre-processing, feature extraction, and model training.UsingMediaPipe, the system guesses hand joint coordinates from images captured by a web camera. To create sequence and picture data, character data must be normalized such that it is independent of character size. To recognize characters, a hybrid deep learning model is employed, which trains CNN on image data and BiLSTM on sequence data.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 03 | March - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

When the trained model's accuracy in recognizing both alphabetic and numeric characters is assessed, it scores highly.

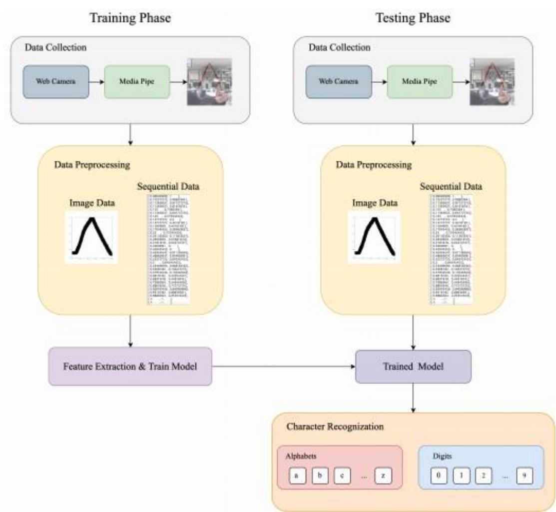

Fig-1: Proposed methodology of air-writing system.

3.2. UlTRASONIC TRANCEIVERS FOR AIR-WRITING CHARACTER RECOGNITION

The creation of an air-writing system represents a significant advancement in human-computer interaction, offering a novel means of input for users with diverse needs. This innovative system integrates cutting-edge technologies and sophisticated methodologies to accurately identify and estimate characters written in the air, thereby facilitating seamless communication and interaction.Central to the functionality of this system is its robust signal acquisition and processing framework. Leveraging the Time of Flight (ToF) approach, the system precisely determines the location of the hand marker, essential for tracking air-written characters in three-dimensional space. The utilization of advanced algorithms such as Multidimensional Scaling (MDS) and Limited-Memory Broyden Fletcher Goldfarb Shanno (LM-BFGS) further enhances the accuracy and efficiency of hand marker localization, ensuring precise input capture.The system's methodology encompasses two primary tasks: character estimation and recognition. For character estimation, ToF technology serves as a cornerstone, providing real-time spatial information that enables the system to infer the trajectory of air- written characters with remarkable precision. This approach not only supports a wide range of inputs but also ensures rapid and reliable estimation of characters, even in dynamic environments.Character recognition, a critical component of the system, is facilitated by Deep Neural Networks (DNNs), particularly Convolutional Neural Networks (CNNs).

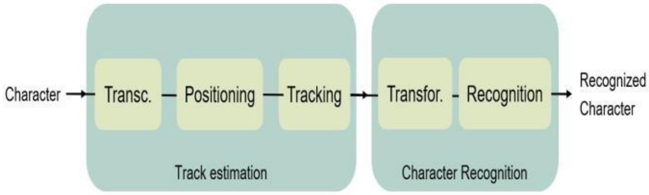

Fig -2: General block diagram for air-writing systems.

3.3. AIR WRITING USING ULTRASONIC SOURCE LOCALIZATION

The innovative air-writing system that combines text recognition via ultrasonic waves and motion tracking is presented in this study. The suggested PDP algorithm, which facilitates motion data translation, orientation feature extraction, and Direction of Arrival (DOA) estimate, is at the heart of its methodology. In addition, the ORM algorithm for text recognition is explained with simulation and experimental results. The study assesses the performance of several classifiers in letter recognition tasks, such as RD, DTW, and ORM. Every system component is outlined in separate parts that provide comprehensive information about its functions and results. This systematic method leads to promising directions for future research and development, while also improving knowledge and highlighting the potential for future breakthroughs in air-writing systems.

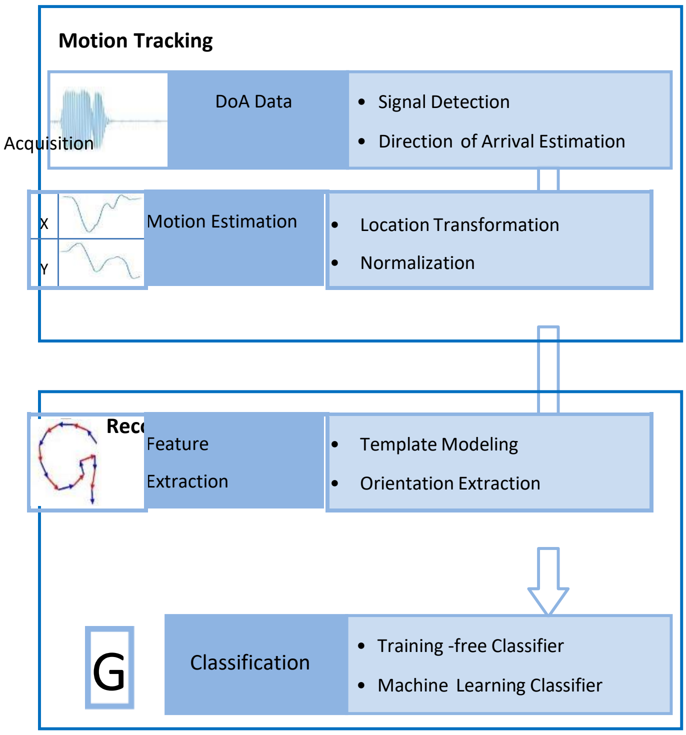

Fig-3: Block diagram of the proposed air-writing system with two main modules: motion tracking and text recognition.

3.4. EFFICIENT TARNet: CNN-LSTM FOR AIR-WRITING RECOGNITION

Each of the ten participants in the dataset wrote each digit at least ten times, yielding a total of about 1200 samples for performance evaluation. Air-written digits may be captured and classified efficiently using this methodology, which methodically investigates multiple deep learning architectures taking into account both temporal and spatial information. By comparing the effectiveness of various architectures on real-world data gathered from a variety of participants, the study seeks to develop air-writing recognition systems.

Choosing suitable three-dimensional (3D) tracking devices such as Leap Motion, Microsoft Kinect, or Intel RealSense is the first step in the methodology for implementing the air-writing technique. Of them, Intel RealSense is the most preferred due to its extensive APIs and effective trajectory tracking. After that, convolutional neural networks (CNNs) and long short-term memory (LSTM) networks are used to handle sequential input and generate features in a deep learning-based recognition system. RealSense trajectory digit (RTD), RealSense trajectory character (RTC), smart-band, and Abas datasets are among the many datasets that verify a novel approach that combines characteristics from both CNNs and LSTMs to improve recognition accuracy. The suggested model outperforms solo CNN or LSTM networks, as confirmed by benchmarking. This methodology covers device selection, air- writing technique efficacy, and adaptation in augmented reality (AR) and virtual reality (VR) applications.

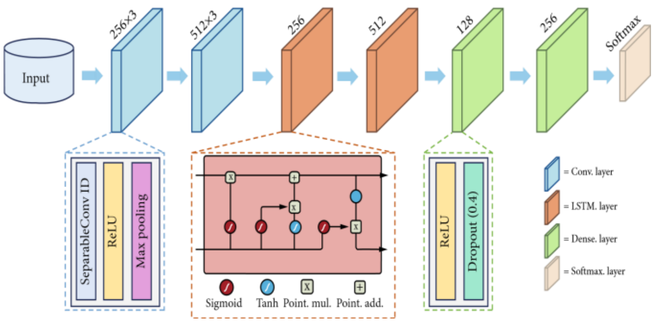

Fig- 4: The proposed TARNet architecture.

3.5. DEEP CNN AND RNN FOR AIR-WRITING RECOGNITION

The handwritten digits (0–9) acquired by a Leap Motion Controller (LMC) sensor are the focus of this paper's investigation into deep learning architectures for air-writing recognition in three dimensions. The work investigates motion trajectory modeling using both dynamic and static methodologies. We train and compare several cutting-edge convolutional and recurrent architectures. Specifically, input sequences are mapped to fixed-dimensional vectors for classification using Long Short-Term Memory (LSTM) networks, including their bidirectional variation (BLSTM). Furthermore, for feature encoding, a mix of 1D Convolutional Neural Networks (CNNs) and LSTM networks (CNN-LSTM) is used. Additionally, a Temporal Convolutional Network (TCN) with dilated causal convolutions is examined. Finally, automated feature learning and classification from unprocessed input data is demonstrated using a deep CNN architecture.

4. DISCUSSION

The Virtual Air Canvas system, which provides users with an immersive and user-friendly platform for creating digital art, marks a revolutionary step forward in the fusion of technology and artistic expression. The virtual whiteboard, which redefines the precision and legibility of lines in digital airwriting, is one of its most notable characteristics. This technology gives users a fluid and organic medium for artistic creation by fusing hand gestures with modern fonts, allowing them to use their hands as virtual paintbrushes. This invention promotes a more organic and intuitive creative process by radically altering the way users engage with digital art. The remarkable accuracy of the Virtual Air Canvas is one of its main advantages; it is attained by accurate fingertip movement tracking and gesture detection. With the help of cutting-edge technologies like MediaPipe and OpenCV, the system can understand hand gestures with astounding accuracy, producing a remarkably lifelike experience. The enhanced creative experience is further enhanced by the high degree of detail, which not only broadens the creative options for users but also permits precise control and manipulation within the virtual environment. In addition, the use of optical character recognition (OCR) methods transforms the field of digital art creation through character recognition. OCR improves the system's accuracy and reliability in character recognition by precisely isolating individual letters from continuous air-written movements. Its artistic improvement allows for the smooth blending of text and graphic elements in digital artwork while simultaneously enhancing the technical aspects of character identification and providing users with additional creative possibilities.

With its immersive and user-friendly platform for creating digital art, the Virtual Air Canvas system is a revolutionary step forward in the fusion of technology and artistic expression. The virtual whiteboard, one of its most notable characteristics, redefines the precision and legibility of lines in digital airwriting. This system allows users to use their hands as virtual paintbrushes by fusing hand movements with modern fonts to create a fluid and organic medium for artistic expression. The way people engage with digital art is radically altered by this breakthrough, which encourages a more organic and intuitive creative process.

In the future, there is a great deal of room for innovation and growth in the field of digital art creation with the Virtual Air Canvas. The incorporation of augmented reality (AR) technology, which enables users to superimpose their digital creations in real time in the real world, is one exciting future trend. This would create new opportunities for interactive and immersive artistic experiences by enabling artists to merge virtual and physical elements flawlessly. Furthermore, the system's ability to recognize characters and gestures may be improved by developments in artificial intelligence and machine learning. The Virtual Air Canvas could improve its ability to comprehend intricate hand movements and recognize handwritten characters by utilizing deep learning algorithms. This would improve the user experience and open up new creative possibilities. In addition, the addition of collaborative functionalities may convert the Virtual Air Canvas into a forum for group artistic expression, permitting numerous individuals to concurrently create and modify digital works of art in real time. Making the Virtual Air Canvas a center for creative experimentation and cooperation, this collaborative feature may encourage a sense of community and shared artistic expression. Virtual reality (VR) technology integration may also be explored as a way to improve Virtual Air Canvas's immersive qualities and enable users to engage with their digital works in a three-dimensional environment. With the potential for ongoing innovation and development at the cutting edge of technology and artistic expression, the Virtual Air Canvas has a huge amount of potential in the future.

The Virtual Air Canvas technology, while promising, comes with several limitations that need to be addressed for broader accessibility and usability. Firstly, its reliance on advanced technologies like MediaPipe and OpenCV may render it inaccessible to individuals lacking access to relevant hardware or software. Additionally, environmental factors such as background interference and lighting could affect the system's accuracy in character and gesture detection, leading to potential errors. Moreover, the system's current focus on handwritten numbers limits its applicability for users seeking to create more diverse digital artworks. This restrictiveness may hinder its adoption among those desiring to produce intricate designs. Lastly, there may be a learning curve for users unfamiliar with gesture-based interfaces or digital art tools, potentially excluding certain demographics from utilizing the system effectively. Addressing these limitations is imperative to enhance the Virtual Air Canvas's accessibility, reliability, and usability, ensuring it caters to a broader range of users.

5. CONCLUSIONS

To sum up, the Virtual Air Canvas project is an innovative endeavor that has the potential to completely transform how the digital era approaches creativity, communication, and efficiency. It redefines online collaboration by providing an immersive, hands-free experience that goes beyond conventional approaches and opens up countless opportunities for creative expression and teamwork. The Mark Air project's strength is its capacity to influence online cooperation in the future and act as a positive change agent in a society where digital communication and remote work are now essential. Its lively and smooth virtual whiteboarding experience encourages an atmosphere where ideas are free to flow and allows both individuals and groups to express themselves creatively and cooperatively. Additionally, the project's innovative method of taking notes—which allows users to do so comfortably and hands-free—simplifies the creative process and shows a dedication to efficient and user-centric design. Through its advocacy of accessibility as a core value, the Virtual Air Canvas project makes sure that cooperation is not only creative but also inclusive and efficient, taking into account different communication preferences and styles. The Virtual Air Canvas project, a monument to the revolutionary power of technology, establishes a new benchmark for accessibility in virtual collaboration and leaves a lasting impression on the digital world. Its user-centric design approach guarantees that it is more than simply a tool rather, it is a platform that boosts creativity and productivity, facilitating inclusive and easy collaboration for all.

SOME OF THE ADVANTAGES

a) Exceptional precision in hand gesture interpretation.

b) Text and graphic elements are seamlessly integrated.

c) Overlay of digital works in real-time.

d) The possibility of instantaneous collaborative invention.

e) Augmented Reality (AR) immersion.

f) Enhanced powers using algorithms for machine learning.

ACKNOWLEDGEMENT

The authors extend sincere gratitude to Mrs.Sreema E R for her Guidance and to project coordinators, Mrs. Najla Nazar, and Mrs. Gishma K M for mentorship. Their expertise played a vital role in shaping the direction and focus of this research.

REFERENCES

1. Taiki.Watanabe, Md. Maniruzzaman , Si-Woong .Jang, and Jungpil.Shin, "2D camera-based air-writing recognition using hand pose estimation and hybrid deep learning model," 2023.

2. Saez-Mingorance.B, Mendez-Gomez.J, Mauro.G, Castillo-Morales.E,Pegalajar-Cuellar.M, Morales-Santos.D.P, "Air-Writing Character Recognition with Ultrasonic Transceivers," Sensors 2021.

3. Chen. H, Ballal. T, Muqaibel. A.H, Zhang. X, and Al-Naffouri. T.Y, "Air writing via receiver array-based ultrasonic source localization," IEEE Trans. Instrum. Meas. 2020.

4. Alam.M, Kwon. K.C, Md Imtiaz.S, Hossain. M.B, Kang. B.G, and Kim. N, "TARNet: An Efficient and Lightweight Trajectory-Based Air-Writing Recognition Model Using a CNN and LSTM Network," Hum. Behav. Emerg.Technol.2022.

5. Grigoris Bastas, and Kosmas Kritsis, "Air writing recognition using deep convolutional and recurrent neural network architectures, "2020.