IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

Air Handwriting by Using CNN Model

Kalyani Sampat Gawande[1], Dr. D. M. Yadav[2], Prof. Gade S. A[3]
Student, Department of Computer Engineering[1]
Principal, ME (Electronics), Ph.D(Electronics)[2]
Professor & HOD, Department of Computer Engineering[3]
SND College of Engineering and Research Center, Yeola, Nashik, India

Abstract: *Gesture recognition has been a popular research field under the trend of IoT and intelligent devices. Air-writing is the most challenging and crucial topic in the gesture recognition field. In this paper, we propose a wearable airwriting system that makes users can write the English alphabet in the three-dimensional space without any write rules. The proposed system is based on the Inertial Measurement Unit (IMU), and it uses dynamic time warping (DTW) as the main recognition algorithm. In addition, to improve the recognition accuracy and take a better advantage of the DTW algorithm, we present an adjustment system that gives some new optimization methods to the application of IMU and DTW. In the experiment, the accuracy of recognition is 84.6% for the uppercase alphabet (from „A" to „Z") in user-dependent case. And we also confirmed that the recognition method only based on the DTW algorithm is one kind of user-dependent methods, which means this method is heavily dependent on personalization.*

Keywords: AI (Artificial Intelligence),Air-writing, Inertial Measurement Unit, Dynamic Time Warping, Gesture Recognition.

I. INTRODUCTION

Air handwriting, also referred to as virtual handwriting or gesture-based writing, represents a groundbreaking technological advancement that facilitates the act of writing or drawing in the air through the use of natural hand movements. This innovative technique integrates the capabilities of gesture recognition with augmented reality, resulting in a distinctive and intuitive means of engaging with digital devices.

At its core, air handwriting enables users to create letters, numbers, and a variety of symbols simply by maneuvering their hands through the air. The technology relies on sensors or cameras that diligently track these movements and accurately interpret them as digital writing. These representations can then be effortlessly displayed on screens or projected onto various surfaces in real time. Such a method paves the way for a writing experience that does not depend on conventional physical tools like pens or keyboards, thus providing users with a more immersive and seamless interaction.

The implications of air handwriting extend across multiple domains, leading many to believe it has the potential to revolutionize fields such as virtual reality, augmented reality, and human-computer interaction. In virtual reality environments, for instance, air handwriting equips users with the ability to write or annotate directly within these digital spaces, significantly enhancing their capacity for communication and collaboration. This feature is particularly valuable in creative and design industries, where teams can mark-up designs or brainstorm ideas in a shared virtual space, streamlining the creative process.

Moreover, air handwriting can also be utilized in educational settings, enabling teachers and students to engage in interactive lessons without the constraints of traditional writing tools. Imagine a classroom where students can easily illustrate concepts or collaborate on projects by writing in the air, making learning both engaging and dynamic.

In summary, air handwriting is an innovative technology that combines gesture recognition and augmented reality to transform the way we write and interact with digital environments. With its vast potential applications across various fields, it is poised to redefine communication and collaboration in our increasingly digital world. The promise of a more intuitive, efficient, and engaging method of writing is just the beginning of what air handwriting can achieve as it continues to evolve and integrate into our daily lives.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 3, November 2024

II. OVERVIEW

The concept of air handwriting leverages gesture recognition and augmented reality to allow users to write or draw in a three-dimensional space. In this study, we introduce a CNN-based model to interpret air handwriting gestures, offering a novel and intuitive method for user interaction with digital environments without relying on traditional physical interfaces. This paper describes a systematic approach to implementing air handwriting by leveraging a Convolutional Neural Network (CNN) model designed specifically for recognizing hand-drawn gestures in the air.

The proposed air handwriting system begins with the collection of gesture data, capturing a series of hand movements that represent characters, symbols, and numerals. This data is essential for training the CNN model, allowing it to recognize diverse handwriting styles and perform consistently across varying conditions. By employing data augmentation techniques, we expand the dataset to include numerous variations of each symbol, enhancing the model's ability to generalize and recognize gestures accurately despite differences in hand motion speed, angle, and orientation.

The core architecture of our air handwriting model is a deep CNN specifically configured to analyse and classify spatial features of each gesture sequence. The CNN model's design involves multiple layers—convolutional, pooling, and fully connected—which together provide an effective method for extracting features from the gesture data and mapping these features to corresponding characters. Convolutional layers detect and preserve essential spatial information from the hand movements, while pooling layers reduce computational complexity, making the model suitable for real-time applications. Softmax activation is used in the output layer for multi-class classification, enabling the model to differentiate between a wide range of characters.

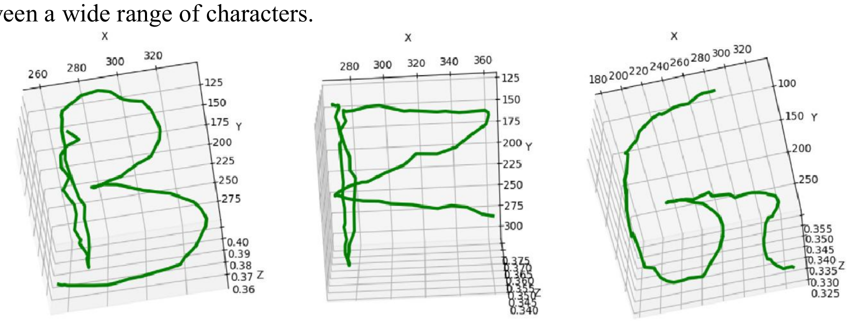

Fig. 1. Example of Air – Writing (3D Character).

To evaluate the CNN model's performance, the dataset is divided into training, validation, and test sets. During training, the model iteratively optimizes weights and biases through backpropagation, using cross-entropy as the loss function to minimize prediction errors. Model validation occurs at the end of each epoch to ensure robust learning and prevent overfitting. The final accuracy of the model is assessed by measuring its performance on the test dataset, with additional evaluation metrics such as precision and recall to analyse its classification effectiveness.

In the real-time application phase, the CNN model is integrated into a gesture recognition system where it can interpret hand movements on the fly. This system utilizes a camera to capture gestures, which the model then processes to predict and display corresponding characters. The display system projects recognized characters onto a screen, allowing users to view their handwriting in real-time, providing an immersive and immediate feedback loop. This real-time interaction is crucial for a seamless user experience, making the system practical for applications in virtual classrooms, collaborative digital workspaces, and creative design settings.

Additionally, the study addresses system optimization for real-time performance, exploring methods like model compression to reduce latency and improve responsiveness. Future enhancements include user personalization to adapt the system to individual writing styles, further improving recognition accuracy and user satisfaction.

Overall, this paper presents a comprehensive approach to developing an air handwriting recognition system using CNNs. With applications spanning virtual reality, augmented reality, and education, this technology has the potential to redefine digital writing and user interaction by offering a more natural and efficient means of engaging with digital platforms.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

III. ARCHITECTURE

This System presents air-writing problem generally relied upon depth sensors such as Kinect and LEAP Motion, or wearable gesture control and motion control device such as Myo, or multi-camera setup to estimate depth information. While these approaches account for easier tracking and better accuracy, they suffer from cost-effective general purpose usage due to essential dependency on the external hardware.

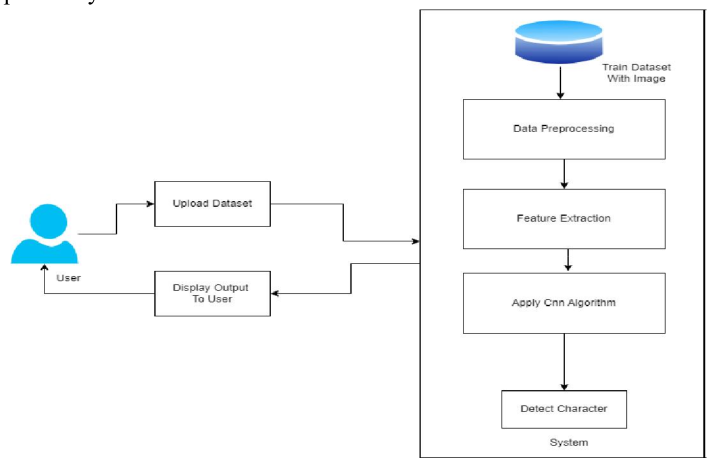

Fig. 2. System Architecture.

Image Processing:

Image processing is a technique for applying various procedures to an image in order to improve it or extract some relevant information from it. It is a kind of signal processing where the input is an image and the output can either be another image or features or characteristics related to that image.

CNN Algorithm:

A CNN is a kind of network architecture for deep learning algorithms and is specifically used for image recognition and tasks that involve the processing of pixel data. There are other types of neural networks in deep learning, but for identifying and recognizing objects, CNNs are the network architecture of choice. The term "CNN" commonly refers to Convolutional Neural Networks, which are a type of deep learning algorithm specifically designed for processing and analysing visual data. CNNs are widely used in various computer vision tasks, such as image classification, object detection, and image segmentation. Convolutional Neural Networks are inspired by the structure and functionality of the visual cortex in animals. They consist of multiple layers, including convolutional layers, pooling layers, and fully connected layers. fig.2 shows the architecture of CNN.

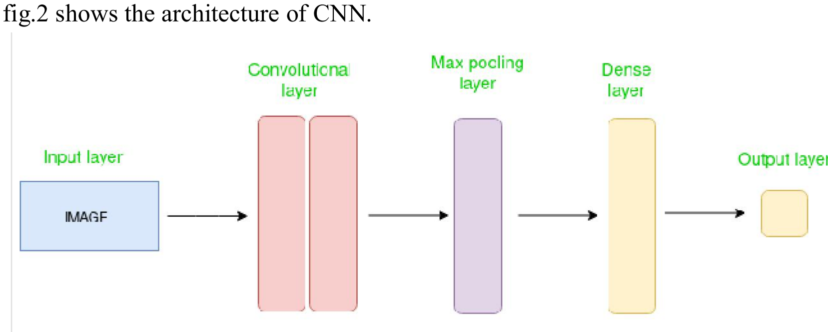

Fig. 3. CNN Architecture.

IV. METHODOLOGY OF THE PROJECT

The development of an air handwriting system based on a Convolutional Neural Network (CNN) involves a series of systematic steps to ensure accurate gesture recognition, classification, and real-time processing. This methodology outlines the key stages of the project, from data collection to model training, evaluation, and implementation within an interactive environment.

A. Data Collection and Preprocessing:

Gesture Data Acquisition:

The initial phase involves capturing gesture data through sensors or cameras capable of tracking hand movements in real time. For this study, a high-resolution camera setup was used, capturing hand motion sequences representing different letters, numbers, and symbols.

Data Labelling and Augmentation:

The captured data is labelled to correspond with the intended symbols or characters. Data augmentation techniques, such as rotation, scaling, and translation, are applied to enhance the dataset and improve model robustness against varying user styles.

Preprocessing:

Each gesture is converted into a format compatible with the CNN model, which may include resizing, grayscale conversion, and normalization. This step ensures uniformity in the input dimensions and improves model accuracy.

B. Model Architecture:

CNN Model Selection:

A Convolutional Neural Network is chosen for its efficacy in image-based pattern recognition. The architecture is designed to include multiple convolutional layers to capture spatial hierarchies in the gestures, followed by pooling layers for dimensionality reduction.

Layers and Configuration: The model includes three primary components:

- Convolutional Layers to extract features.
- Pooling Layers to reduce computational load and preserve important features.
- Fully Connected Layers to perform classification.

Activation Functions and Loss Calculation:

ReLU is applied as the activation function in convolutional layers, while Softmax is used in the output layer to classify gestures into respective categories. Cross-entropy is applied as the loss function to evaluate model performance.

C. Real-Time Implementation:

Gesture Recognition in Real Time:

After successful training, the CNN model is integrated into a system capable of recognizing gestures in real time. The camera captures hand movements, and the model processes these sequences to predict characters or symbols accurately.

Projection and Display:

The recognized characters are displayed on a screen or projected onto a surface, providing an interactive platform for users to see the results of their handwriting instantly.

User Feedback Mechanism:

The system includes a feedback loop, allowing users to verify the accuracy of each symbol in real-time. Incorrect predictions prompt the model to recalibrate through user input, refining recognition accuracy over time.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, November 2024

V. SOFTWARE INTERFACES

Python:

Python is an interpreted, high-level and general-purpose programminglanguage. Python's design philosophy emphasizes code readability withits notable use of significant indentation. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for smalland large-scale projects. Python is dynamically-typed and garbage-collected.It supports multiple programming paradigms, including structured (particularly, procedural), object-oriented and functional programming. Python isoften described as a" batteries included" language due to its comprehensive standard library. Python was created in the late 1980s, and first releasedin 1991, by Guido van Rossum as a successor to the ABC programming language. Python 2.0, released in 2000, introduced new features, such aslist comprehensions, and a garbage collection system with reference counting,and was discontinued with version 2.7 in 2020.Python 3.0, released in2008, was a major revision of the language that is not completely backwardcompatible and much Python 2 code does not run unmodified on Python 3.With Python 2's end of-life (and pip having dropped support in 2021), onlyPython 3.6.x and later are supported, with older versions still supporting e.g.Windows 7 (and old installers not restricted to 64-bit Windows). Python interpretersare supported for mainstream operating systems and available fora few more (and in the past supported many more). A global community ofprogrammers develops and maintains Python, a free and open-source reference implementation. A non-profit organization, the Python Software Foundation,manages and directs resources for Python and CPython development.As of January 2021, Python ranks third in TIOBE's index of most popular programming languages, behind C and Java, having previously gained secondplace and their award for the most popularity gain for 2020.

Spyder:

Spyder is an open-source cross-platform integrated developmentenvironment (IDE) for scientific programming in the Python language.Spyder integrates with a number of prominent packages in the scientificPython stack, including NumPy, SciPy, Matplotlib, pandas, IPython,SymPy and Cython, as well as other open-source software.It is released underthe MIT license. Initially created and developed by Pierre Raybaut in 2009,since 2012 Spyder has been maintained and continuously improved by a teamof scientific Python developers and the community. Spyder is extensible withfirst-party and third-party plugins, includes support for interactive tools fordata inspection and embeds Python-specific code quality assurance and introspectioninstruments, such as Pyflakes, Pylint and Rope. It is availablecross-platform through Anaconda, on Windows, on macOS through Mac-Ports, and on major Linux distributions such as Arch Linux, Debian, Fedora,Gentoo Linux, openSUSE and Ubuntu. Spyder uses Qt for its GUI and isdesigned to use either of the PyQt or PySide Python bindings.QtPy, thisabstraction layer developed by the Spyder project and later adopted by multiple
other packages, provides the flexibility to use either backend.

DB SQLite:

DB Browser for SQLite (DB4S) is a high quality, visual,open-source tool to create, design, and edit database files compatible withSQLite. DB4S is for users and developers who want to create, search, and editdatabases. DB4S uses a familiar spreadsheet-like interface, and complicatedSQL commands do not have to be learned. Controls and wizards are availablefor users to: Create and compact database files Create, define, modify anddelete tables Create, define, and delete indexes Browse, edit, add, and deleterecords Search records Import and export records as text Import and exporttables from/to CSV files Import and export databases from/to SQL dumpfiles Issue SQL queries and inspect the results Examine a log of all SQLcommands issued by the application Plot simple graphs based on table orquery data.

Algorithm & flowchart detailed

To fully realize air handwriting recognition, we have developed an algorithm using a Convolutional Neural Network (CNN) model optimized for recognizing and classifying hand-drawn gestures in real time. This algorithm covers all essential steps, from capturing hand movements to interpreting them as digital handwriting displayed on a screen or projected surface. Using a systematic data flow and structured CNN model architecture, the algorithm captures spatial information in hand movements.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 3, November 2024

A. Algorithm for Air Handwriting Recognition Using CNN:

Data Acquisition:

- Set up a high-resolution camera to capture hand gestures representing letters, numbers, and symbols.
- Collect gesture data in real-time or from a predefined dataset with labelled gestures for each character or symbol.

Data Preprocessing:

- Convert the captured gesture images into grayscale to reduce computational complexity.
- Resize each image to a consistent input size for the CNN (e.g., 64x64 pixels).
- Normalize the images to scale pixel values between 0 and 1, improving model performance.

Data Augmentation:

- Apply augmentation techniques such as rotation, scaling, and shifting to simulate various handwriting styles.
- Label each augmented gesture appropriately to maintain the dataset's integrity.

CNN Model Initialization:

- Define a CNN architecture with several convolutional layers, pooling layers, and fully connected layers.
- Configure the input layer to match the gesture image dimensions (e.g., 64x64x1).
- Set up the output layer with Softmax activation to classify gestures into different categories (e.g., A-Z, 0-9).

Training the CNN Model:

- Split the preprocessed dataset into training, validation, and test sets (e.g., 70:15:15).
- Train the model on the training set using a batch size and learning rate optimized through experimentation.
- For each epoch, calculate the loss using cross-entropy and update model parameters using the Adam optimizer.
- Validate the model after each epoch, tuning hyperparameters as necessary to prevent overfitting.

Model Evaluation:

- Use the test set to evaluate the model's accuracy, precision, and recall, ensuring the model generalizes well to unseen data.
- Check latency to confirm the model's performance for real-time applications.

Real-Time Gesture Recognition:

- Integrate the trained CNN model into a live application system that captures hand movements in real time.
- For each captured gesture, preprocess it (resize, grayscale, normalize) and feed it into the CNN for prediction.
- Display the recognized character or symbol on a screen or projection, providing immediate feedback.

User Feedback and Adaptation:

- Enable a feedback mechanism to capture corrections for misrecognized gestures.
- Adjust model parameters periodically to improve performance for personalized handwriting styles.

B. Flowchart of the Algorithm:

1. Start Air Writing
2. Capture Hand Gesture Data.
3. Preprocess Gesture Images.
4. Data Augmentation Techniques.
5. Initialize CNN Model.
6. Train CNN Model on Data.
7. Evaluate CNN Model.
8. Preprocess and Feed to CNN
9. Display Recognized Character.
10. User Feedback & Model Adaptation.
11. End.

VI. CONCLUSION

In conclusion, this paper introduces an innovative air-writing system driven by an inertial measurement unit (IMU) that enables users to write in the air using gestures. The system employs Dynamic Time Warping (DTW) as its primary algorithm for gesture recognition, facilitating the comparison of time series data and improving the accuracy of recognizing written characters.

To enhance the performance of the air-writing system, we have developed an adjustment mechanism specifically designed to increase the recognition accuracy while also minimizing the time required for processing each gesture input. This is crucial as users often engage with such systems in varied and dynamic environments.

A key feature of the adjustment system is the incorporation of an idle-cutter technique. This feature addresses a significant challenge in gesture recognition—namely, the impact that unintentional hand tremors, caused by the user's lack of control or unconscious movements, can have on the system's ability to accurately interpret gestures. Such involuntary movements have been a common hurdle for many researchers in the field of gesture recognition, and our solution provides a practical means to mitigate this issue.

Additionally, the paper presents a multi-template approach that introduces a novel concept to optimize the application of DTW in gesture recognition. This approach is versatile and extends beyond just the hand angle parameter, suggesting potential applications in various fields that require gesture recognition and similar technologies.

Results from our experiments demonstrate that in user-dependent scenarios, where the system is tailored to individual users, our air-writing system performs exceptionally well, showcasing both high accuracy and reliability. This indicates that the proposed system not only meets its intended objectives but also paves the way for future advancements in gesture recognition technology. Overall, the findings of this paper contribute significantly to the ongoing discourse in the field, presenting viable solutions to existing challenges while offering new avenues for research and application.

VII. FUTURE SCOPE

The future scope of air handwriting recognition is vast and holds the potential to transform multiple fields, especially as augmented reality (AR) and virtual reality (VR) technologies continue to advance. In education, air handwriting could revolutionize classroom interactions, enabling teachers and students to collaborate in a shared digital space without physical writing tools. This technology can enhance remote learning, as students and educators could write, draw, and annotate virtually, allowing for interactive and immersive educational experiences. Additionally, the integration of air handwriting into virtual collaborative platforms could make brainstorming sessions, project planning, and creative tasks more dynamic, promoting seamless interaction regardless of physical location.

Moreover, the application of air handwriting in accessibility and assistive technology is promising. For individuals with physical disabilities, the ability to communicate through hand gestures could provide a powerful and inclusive tool for expression and interaction with digital devices. With continuous advancements in gesture recognition accuracy and the ability to personalize models for different users, air handwriting could become a standard interface in smart homes, healthcare, and IoT devices, offering users a natural and intuitive way to interact with technology. As the model evolves with more complex algorithms, such as those leveraging deep learning and reinforcement learning, air handwriting could eventually lead to highly sophisticated, context-aware systems that understand not only individual characters but complex commands and expressions in real-time.

VIII. ACKNOWLEDGMENT

We are deeply grateful to our mentors and educators, whose support and guidance have been invaluable throughout the development of this project. Their expertise in AI and computer vision has been instrumental in shaping our approach and ensuring the success of this air handwriting recognition system. We would also like to thank our academic institution and department for providing the resources and environment necessary to explore and implement cutting-edge technologies.

Additionally, we extend our heartfelt appreciation to various online communities and open-source contributors who generously shared their knowledge and tools, which significantly aided us in overcoming technical challenges and refining our CNN model. The collaborative spirit of these communities enabled us to gain a deeper understanding of

gesture recognition, data processing, and model optimization techniques. This project stands as a testament to the collective effort of all who contributed to our learning and success.

REFERENCES

[1] Yuqi Luo, Jiang Liu," Wearable Air-writing Recognition System employing Dynamic Time Warping" Shigeru Shimamoto School of Fundamental Science and Engineering Waseda University Tokyo, Japan luoyuqi1997@toki.waseda.jp, jiang@waseda.jp, shima@waseda.jp.

[2] Aradhana Kumari Singh a , Deepika Koundal " A Temporal Convolutional Network for modeling raw 3D sequences and air-writing recognition" Received 1 October 2023; Received inrevised form 20 November 2023; Accepted 25 November 2023 Available online 28 November 2023.

[3] Prasun Roy, Subhankar Ghosh, and Umapada Pal "A CNN Based Framework for Unistroke Numeral Recognition in Air-Writing" 2018 16th International Conference on Frontiers in Handwriting Recognition Computer Vision and Pattern Recognition Unit, Indian Statistical Institute 203, B. T. Road, Kolkata, WB - 700108, India.

[4] Md. Shahinur Alam1, Ki-Chul Kwon, and Nam Kim2 "Trajectory-based Air-writing Character Recognition Using ConvolutionalNeuralNetwork" 2019 4th International Conference on Control, Robotics and Cybernetics (CRC Dept. of Computer and Communication Engineering Chungbuk National University.

[5] Z. -T. Liu, D. P. Y. Wong and P. H. Chou,"An Imu-Base WearableRing for On-Surface Handwriting Recognition," 2020 InternationalSymposium on VLSI Design, Automation and Test (VLSI-DAT), 2020,pp. 1-4, doi: 10.1109/VLSIDAT49148.2020.9196479.

[6] Savita Ahlawat 1, Amit Choudhary 2, Anand Nayyar 3, Saurabh Singh 4 and Byungun Yoon 4, "Improved Handwritten Digit Recognition Using Convolutional Neural Networks (CNN)" Received: 25 May 2020; Accepted: 9 June 2020; Published: 12 June 2020.

[7] J. Patel, U. Mehta, K. Panchal, D. Tailor and D. Zanzmera," TextRecognition by Air Drawing," 2021 Fourth International Conference onComputational Intelligence and Communication Technologies (CCICT),2021, pp. 292-295, doi: 10.1109/CCICT53244.2021.00061.

[8] S.V. Aswin Kumer, P. Kanakaraja, Sheik Areez, Yamini Patnaik, PamarthiTarun Kumar," An implementation of virtual white boardusing open CV for virtual classes.

[9] A. Schick, D. Morlock, C. Amma, T. Schultz, R. Stiefel Hagen, Vision-basedhandwriting recognition for unrestricted text input in mid-air, in: Proceedingsof the14th ACM International Conference on Multimodal Interaction, 2012, pp.217–220.

[10] C. Amma, M. Georgi, T. Schultz, Airwriting: a wearable handwriting recognitionsystem, Pers. Ubiquitous Comput. 18 (2014) 191–203.

[11] S.K. Behera, P. Kumar, D.P. Dogra, P.P. Roy, Fast signature spotting in continuousair writing, in: 2017 Fifteenth IAPR International Conference on Machine VisionApplications, MVA, IEEE, 2017, pp. 314–317.

[12] M.S. Alam, K.-C. Kwon, N. Kim, Trajectory-based air-writing character recognitionusing convolutional neural network, in: 2019 4th International Conferenceon Control, Robotics and Cybernetics, CRC, IEEE, 2019, pp. 86–90.