Grenze International Journal of Engineering and Technology, June Issue

Real-time Air-written English Numeral String Recognition using Deep Learning

Dr. Chandrashekhar H. Patil[1], Harshal V. Patil[2], Meenal Jabde[3], Koustubh Deodhar[4], Amaan Sayyed[5] and Hitank Shah[6]
[1]Department of Computer Science and Application, Dr. Vishwanath Karad MIT World Peace University Pune, Maharashtra, India
Email: chpatil.mca@gmail.com
[2]Department of Computer Science, Bhusawal Arts Science and P O Nahata Commerce College, Bhusawal Maharashtra, India
Email: harshal4patil@gmail.com
[3-6]Department of Computer Science and Application, Dr. Vishwanath Karad MIT World Peace University Pune, Maharashtra, India
Email: meenal82@gmail.com, koustubhdeodhar23@gmail.com, amaaansayyed1@gmail.com, hitankjain@gmail.com

***Abstract*—Air writing is gesturing to draw 3D letters without tools, which could allow devices to work in new ways. Single letters are easy to recognize; however, continuous writing is difficult because there are no clear breaks between letters. Stray hand moves also add noise that confuses re cognition systems. Eliminating noise and accurately sorting letters are crucial for better accuracy. To help solve these issues, we aim to use Recurrent Neural Networks (RNNs), which can handle time-base d data like gestures. Numbers often share parts too, making this another tricky case. Digits look alike, so telling them apart is hard. Our approach uses common ways people write digits to figure out which digit came first. It take s x-y coordinates as input that any camera can capture, so anyone can use it. We tested our system with English numerals from combined MNIST, Pen-digits, and our ISI-Air Dataset of air-written English numerals. Under standard conditions, it correctly identified 98.75% single -digit English numerals and 84 % multiple digit sequence.**

***Index Terms*— Handwritten Pattern, Optical Scanning, Computer Vision, Pattern Recognition.**

I. INTRODUCTION

These days, converting handwritten stuff into digital form is vital. Handwritten Character Recognition is tough because everyone write s differently. This issue impacts education, office work, and more. Accurately reading handwritten text helps information be come more accessible , and processes smoother. Still, today's HCR methods often can't quite grasp handwritten text well. We use advance d tech like Machine Learning and computer vision to improve this. Our models achieve big accuracy: 98.75% in Single Character Se t, and 84% in Multiple Character sets. Such innovations offer hope for better HCR systems, making paper-to-digital transitions easier. Written words have power in our technological age.

We look at new ways to use handwriting digitally. gesture tech adds some thing extra to handwriting software. Users can talk to compute rs with body motions, making interactions feel easier and smoother. These upgrades improve user fun but also ope n doors for new human-computer links across many fields.

Grenze ID: 01.GIJET.10.2.539_1

II. LITERATURE REVIEW

Air typing, defined as tracing characters in three-dimensional free space using hand gestures, is a path to peripheral- independent virtual communication through devices. While single-line character recognition is fairly simple, continuous writing recognition is difficult because there are no separators between characters. In addition, the diffuse movement of the hand while typing adds noise to the input, making accurate detection difficult.[1]

The capacity to recognize the motion and shape of hands can play a key role in enhancing user experience across a range of technology platforms and areas. For instance, it can provide the foundation for hand gesture control and sign language comprehension. It can also make it possible for digital information and material to be superimposed on top of the actual environment in augmented reality. Although it comes effortlessly to individuals, a computer vision problem, robust real-time hand perception is a distinctly difficult challenge because hands frequently occlude themselves or each other (for example, finger/palm occlusions and handshaking) and lack high contrast patterns.[2]

In pattern recognition applications, Handwritten Character Recognition is considered a central issue in a practical manner. Reading handwritten information like examination answer sheets is still difficult for many of us because each has a different interpretation style. As the world progresses towards digitization.[4] In the era of the digital world, the traditional art of writing is being replaced by digital art. Digital art refers to forms of expression and transmission of art form with digital form. Relying on modern science and technology is the distinctive characteristic of the digital manifestation.[3].

The classification of handwritten digit recognition using Machine Learning algorithms such as DT, SVM, ANN, and Deep Learning algorithm CNN with the help of the Vgg16 transfer learning model. This system uses the MNIST Eng-lish Handwritten digit dataset to evaluate the proposed system. The dataset is split into 60000 images for training and 10000 for testing purposes. The analysis shows that in the Deep Learning-based approach, the Vgg16 transfer learning algorithm outper-forms the Machine Learning algorithm. The training accuracy and loss obtained are 0.9984 and 0.06, respectively, while the validation accuracy and loss are 0.9602 and 0.2219, respectively[17].

III. RELATED WORK

The suggested framework for multi-digit uni-stroke numeral recognition in an air-writing environment is a reliable and hardware-independent solution for virtual interaction with devices that is not dependent on peripherals. The system uses a sliding window-based technique to isolate a tiny portion of the spatiotemporal input from air-writing activity for noise reduction and digit segmentation. Recurrent Neural Networks (RNN), especially RNN-LSTM networks, are used to eliminate noise and recognize digits. The system was evaluated on English numbers using a combination of the MNIST and Pen digits datasets, as well as the ISI-Air Dataset, and achieved accuracy of 98.45% and 82.89% for single and -multiple digit English numerals, respectively. The buildup of transition noise has an influence on the system's performance, although bidirectional scanning mitigates its effects. The system's source code is accessible for further research and development.[1]

Air-writing is the method of writing a character or words in free space using finger or hand motions rather than a portable instrument. Because of its easy writing style, it has a significant advantage over traditional pen-and-paper systems. However, the absence of a single criterion, non-uniform characters, and diverse writing styles make it a challenging endeavor. In this paper, we present an air-written mathematical expression recognition system that accepts webcam footage as input. We deployed a novel hand identification model to recognize writing hands and monitor fingertip movement. Trajectories were collected and a convolutional neural network (CNN) was used for recognition. Through our concept, children may study fundamental ME assessment on their own without the support of an educator and learn additional information. With minimum effort. Experiments were done using a mix of MNIST, ISI Air-Written English numbers, RTD, and our own air-written Math operator's datasets (MAIR). To assess the robustness of our model, we tested it on a group of youngsters who provided the input by writing in the air, with the input data acquired using a system webcam. In both situations, we produced encouraging results in digit recognition and ME assessment.[6]

This study proposes a trajectory-based air writing character identification system called CNN-LSTM, which combines convolution neural networks (CNN) with long short-term memory (LSTM). Two publicly available datasets, RealSense trajectory digit (RTD) and RealSense trajectory digit (RTC), were used. RTD and RTC datasets include 20,000 digits and 30,000 characters, respectively. The proposed network consists of two CNNs with successive pooling, two LSTMs, and two thick layers. The input specifies the maximum number of character lengths (300 and 800 for RTD and RTC, respectively). To prevent overfitting, the dense layer has a dropout rate of 0.4. The output layer is changeable here, with output sizes of 10 and 26 for the RTD and RTC datasets, respectively. The accuracy of the RTD and RTC datasets is 99.63% and 98.74%, respectively. The identification time was 14 milliseconds per character, which is fast enough for real-time applications.[11]

A comprehensive review of the current work done in offline/online multilingual and air writing multilingual numeral recognition with and without sensors. The systematic architecture of datasets, segmentation, feature extraction, and classification is provided with current applications and challenges. The paper reports that deep learning-based models have recently superseded conventional machine learning methods.[18]

IV. METHODOLOGY

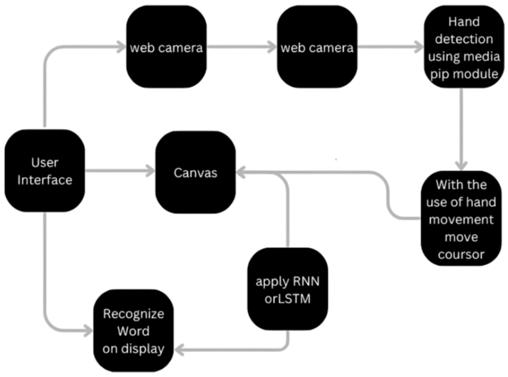


Fig 1. Proposed System Structure

We utilized the isi-air-dataset for our RNN model, analyzing and cleaning the dataset for testing. Our approach involves understanding hand movements, capturing writing on canvas, recognizing hand-drawn patterns in the air, and displaying the text. Hand movement recognition utilizes Media Pipe's hand model, which employs machine learning to track landmarks on the hand for each frame, facilitating gesture recognition. [7,8,9,10,11,12,16,19]

The primary step within the Discuss Composing extend includes downloading the fundamental datasets from GitHub These datasets contain hand-drawn digits captured in discussion utilizing sensors or other input gadgets. The datasets are curated to incorporate as if they were numerical digits from 0 to 9.

V. DATA-PREPROCESSING

Upon downloading the datasets, the other stage includes preprocessing the information to get it ready for preparing and testing. This incorporates steps such as:

Data Cleaning: Evacuating any clamor or artifacts from the captured pictures to guarantee clarity and consistency.

Labeling: Relegating suitable names or classes to each picture, showing the digit it speaks to (0 to 9).

Data Expansion: Creating extra preparing tests through procedures like revolution, scaling, and flipping to upgrade the strength of the model.

Training and Testing: Once the information is preprocessed, it is separated into preparing and testing sets. The preparing set is utilized to prepare the machine learning model, whereas the testing set assesses its execution.

The taking after steps diagrams the preparing and testing process

Feature Extraction: Extricating important highlights from the preprocessed pictures to speak to each digit effectively.

Model Choice: Choosing a fitting machine learning show engineering for the acknowledgment errand. This may involve convolutional neural networks (CNNs)[6,12,13,14,19], repetitive neural systems (RNNs), or other profound learning architectures.[8,10,11,12,15,18,19,20]

Model Preparing: Preparing the chosen demonstration utilizing the prepared dataset. While preparing, the demonstrator learns to relate input pictures with their comparing names (digits), slowly progressing its exactness over numerous cycles (epochs).

Model Assessment: Assessing the prepared model's execution utilizing the testing dataset. This includes measuring measurements such as exactness, exactness, review, and F1-score to survey how well the show generalizes to concealed data.
Fine-Tuning: Iteratively refining the demonstrated design, and hyperparameters, and preparing handle based on the assessment comes about to optimize execution further.
Deployment: Sending the prepared demonstration for real-time discussion composing acknowledgment applications. This includes joining the demonstration into computer programs or equipment frameworks competent of capturing hand-drawn digits and preparing them utilizing the prepared model.
Integration of OpenCV, TensorFlow, NumPy, and Matplotlib
Throughout the complete prepare, different libraries and systems are utilized to encourage distinctive tasks:
OpenCV: Utilized for picture prepossessing, including extraction, and visualization of results.
TensorFlow: Employed for building, preparing, and assessing machine learning models, particularly profound learning architectures.
NumPy: Utilized for numerical computing and taking care of multidimensional clusters, basic for information control and processing.
Matplotlib: Utilized for visualizing preparation and testing, counting exactness bends, perplexity networks, and test predictions.
By coordination of these techniques, we can successfully recognize hand-drawn numerical digits in real-time and discuss composing scenarios, advertising a commonsense arrangement for different applications.

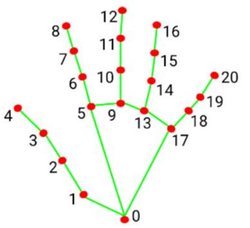


Fig 2: Finger Landmark's [5]

After all the recognition of patterns the actual word will print on the screen.

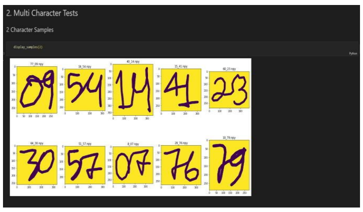


Fig 3: Recognize output

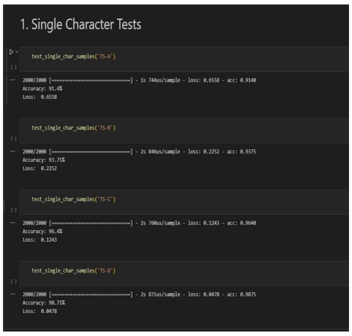
VI. Result and Discussion

We present outcomes from tests evaluating the RNN system's ability to recognize continuous motion numerals using the ISI Air-Data set.
The RNN achieves high accuracy rates for both single-digit English numerals and multiple-digit sequences, demonstrating its effectiveness in capturing air-written numeral intricacies.

Results:

1. **Single Character:**

- 2000/2000 [==============================]
  - 1s 871us/sample - loss: 0.0478
  - acc: 0.9875
- **Accuracy: 98.75%**
- Loss: 0.6558



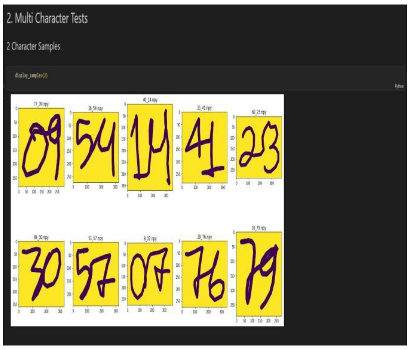


Fig 4. Single Character

Fig 5. Multiple Character

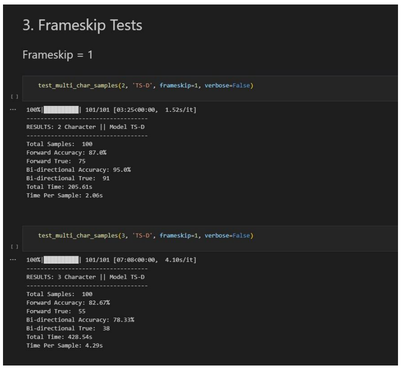
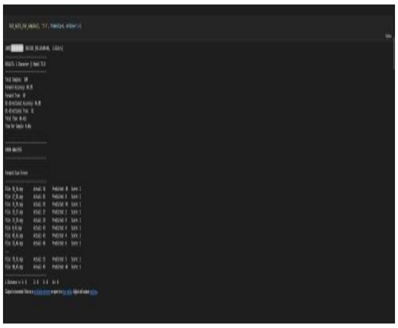


Fig 6. Multiple Character

Fig 7.Frame Skip



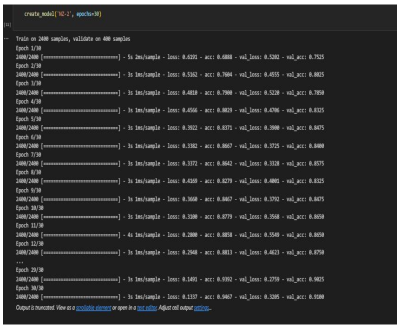


Fig 8. Frame Skip

Test Fig 9. Model

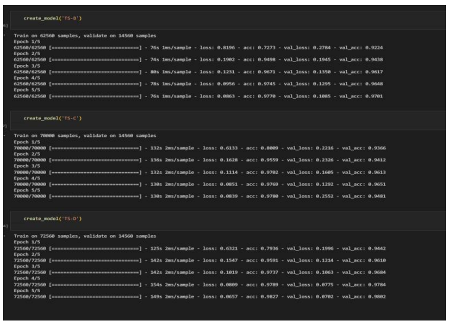


Fig 10. Model

2. Multiple Character :

- 100%|█ █ █ █ █ █ █ █| 101/101 [01:26<00:00, 1.61it/s]
- ----------------------------

- RESULTS: 2 Character || Model TS- D
- ---------------------------
- Total Samples: 100
- **Forward Accuracy: 84.5%**
- Forward True: 69
- **Bi-directional Accuracy: 96.0%**
- Bi-directional True: 92
- Total Time: 86.42s
- Time Per Sample: 0.86s Forward Scan Errors
- ---------------------------
- File: 94_34.npy Actual: 34 Predicted: 30 Score: 1
- File: 27_81.npy Actual: 81 Predicted: 8 Score: 1
- File: 31_95.npy Actual: 95 Predicted: 94 Score: 1
- File: 35_17.npy Actual: 17 Predicted: 1 Score: 1
- File: 33_29.npy Actual: 29 Predicted: 9 Score: 1
- File: 0_43.npy Actual: 43 Predicted: 4 Score: 1
- File: 85_42.npy Actual: 42 Predicted: 4 Score: 1
- File: 25_46.npy Actual: 46 Predicted: 6 Score: 1
- File: 78_51.npy Actual: 51 Predicted: 5 Score: 1
- File: 90_45.npy Actual: 45 Predicted: 46 Score: 1

L-Distance => 1: 8 2: 0 3: 0

4+: 0

VII. Conclusion

The use of RNN in continuous motion numeral recognition, as demonstrated with the ISI Air-Dataset, shows promising results. High accuracy rates and real- time recognition potential make this approach suitable for practical applications, particularly where conventional input methods are impractical. Further refinement of the model's design and integration of attention mechanisms could enhance recognition precision over extended sequences and swift digit transitions.

References

[1] Rahman, A., Roy, P., Pal, U. (2020). Continuous Motion Numeral Recognition Using RNN Architecture in Air-Writing Environment. In: Palaiahnakote, S., Sanniti di Baja, G., Wang, L., Yan, W. (eds) Pattern Recognition. ACPR 2019. Springer.

[2] P. A. Bamanikar, “Air Writing and Recognition System.” International Journal for Research in Applied Science and Engineering Technology, vol. 11, no. 5, pp. 6291-6296, 2023.

[3] Ashutosh Kr. Pandey, Dheeraj, Manas Tripathi, and Vidyotma, “AIR WRITING USING PYTHON (2021-2022)”, May 2022.

[4] Prof. S.U. Saoji, Nishtha Dua, Akash Kumar Choudhary, and Bharat Phogat, “AIR CANVAS APPLICATION USING OPENCV AND NUMPY IN PYTHON”, Aug 2021.

[5] Fig 1.FingerLandMarks web: https://mediapipe.dev/images/mobile/hand_landmarks.png

[6] Kumar, Shobhan & Trivedi, (Dr.) Munesh & Chauhan, Arun. (2022). Real-Time Air-Written Mathematical Expression Recognition for Children’s Enhanced Learning. Neural Processing Letters. 55. 10.1007/s11063-022-11012-3

[7] Mohammadi, Shahram & Maleki, Reza. (2020). Air-writing recognition system for Persian numbers with a novel classifier. The Visual Computer. 36. 10.1007/s00371-019-01717-3.

[8] Yana, Buntueng & ONOYE, Takao. (2018). Air-Writing Recognition Based on Fusion Network for Learning Spatial and Temporal Features. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences. E101.A. 1737-1744. 10.1587/transfun.E101.A.1737.

[9] Mahmoudi, Maryam & Mojtahedi, Sara & Shams, Sepideh. (2017). AR-based value-added visualization of infographic for enhancing learning performance. Computer Applications in Engineering Education. 25. 10.1002/cae.21853.

[10] Alam, Md. Shahinur & Kwon, Ki-Chul & Alam, Md. Ashraful & Abbass, Mohammed Y. & Imtiaz, Shariar Md & Kim, Nam. (2020). Trajectory-Based Air-Writing Recognition Using Deep Neural Network and Depth Sensor. Sensors. 20. 376. 10.3390/s20020376.

[11] Alam, Md. Shahinur & Kwon, Ki-Chul & Imtiaz, Shariar Md & Hossain, Md & Shinde, Rupali & Kim, Joon & Kim, Nam. (2021). Air-writing recognition using a fusion CNN-LSTM neural network.
[12] Rahim, Md & Shin, Jungpil & Islam, Dr. MD Rashedul. (2020). Hand gesture recognition-based non-touch character writing system on a virtual keyboard. Multimedia Tools and Applications. 79. 10.1007/s11042-019-08448-6.
[13] De, Oyndrila & Deb, Puskar & Mukherjee, Sagnik & Nandy, Sayantan & Chakraborty, Tamal & Saha, Sourav. (2016). Computer vision-based framework for digit recognition by hand gesture analysis. 1-5. 10.1109/IEMCON.2016.7746361.
[14] Deng, Xiaoming & Yuan, Ye & Zhang, Yinda & Tan, Ping & Chang, Liang & Yang, Shuo & Wang, Hongan. (2016). Joint Hand Detection and Rotation Estimation by Using CNN. IEEE Transactions on Image Processing. PP. 10.1109/TIP.2017.2779600.
[15] ZHANG, Lei. (2021). Hand-Drawn Sketch Recognition With Double-Channel CNN. 10.21203/rs.3.rs-516601/v1.
[16] Dongdong, Li & Limin, Zhang & Deng, Xiangyang. (2019). Gesture Recognition Based on Multi-Channel Convolutional Neural Network. ICMAI 2019: Proceedings of the 2019 4th International Conference on Mathematics and Artificial Intelligence. 5-9. 10.1145/3325730.3325737.
[17] Meenal Jabde, Chandrashekhar Patil, Shankar Mali and Amol Vibhute, "Comparative Study of Machine Learning and Deep Learning Classifiers on Handwritten Numeral Recognition", International Conference on Innovations in Data Analytics ICIDA 2022.
[18] Meenal K. Jabde, Chandrashekhar H. Patil, Amol D. Vibhute and Shankar Mali1. (2024). A Comprehensive Literature Review on Air-written Online Handwritten Recognition. International Journal of Computing and Digital Systems ISSN (2210-142X) Int. J. Com. Dig. Sys. 15, No. 1 (Jan-24).