



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 5, May 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423



Recognition of Text writing in Air Using Deep Neural Network

**Dr.BHUVANESHWARI K V, AASTHIKN M, RAJA BABU SINGH, SACHIN,
HARSHITH K S**

Associate Professor, Department of ISE, BIET, Davangere, India

B.E Student, Department of ISE, BIET, Davangere, India

B.E Student, Department of ISE, BIET, Davangere, India

B.E Student, Department of ISE, BIET, Davangere, India

B.E Student, Department of ISE, BIET, Davangere, India

ABSTRACT:The innovative technique of creating linguistic symbols or phrases in open space with hand or finger motions is known as "air-writing." A System is created that can read digits and characters written in the air in three dimensions and classify the recognised character into one of several categories. With users object, system can scribble something down in an extremely 3D environment by writing in the air. Proposed system might combine machine learning for handwriting recognition with computer vision object chasing. The computer's digital camera is used by the air writing recognition system to track characters and digits that users have written in the air. A convolutional neural network is then used to classify the characters and digits into one of the possible classes. While many current systems use on complex, expensive pursuit configurations to realise gesture recognition, system have developed a method that could accomplish the same task with a much simpler setup.

KEYWORDS: Computer vision, machine learning, Convolutional neural networks, hand writing recognition, and hand writing classification.

I. INTRODUCTION

Desktop Interfaces using gestures are popular in many fields of application. But gestures and mobile devices is still an emerging trend. Computer vision is a technology that can be useful in this. It makes use of camera to recognize objects, people or events and perform analysis of the images. The present attempts to solve the problem of gesture recognition include use of mechanical devices and use of gloves for gestural interfaces which required the user to system a device, and generally make use of loads of cables that connect the device to a computer but this isn't feasible for a hand-held device like mobiles. Use of these devices can be replaced by using camera-based interaction techniques. Here, system propose a practical way to make new use of the video camera of the Android based Smart phones and track a single finger using the camera as sensor and recognize text written in air. They had only proposed a method to track fingertip on mobiles phones but the next task is to make it useful to give commands to system mobile device and enhance HCI. This task of text detection, tracking and recognition is a difficult problem. While the recognition of text in printed documents has been studied extensively in the form of Optical Character Recognition (OCR) Systems, and also for handwritten text on paper, but using these methods directly when a person writes in air is difficult and hence, they cannot be generalized also while detecting text or gesture in air factors like inconsistent lighting conditions, variable fonts, background noise and image distortions add to the problem complexity. The proposed system provides a method of interacting with a mobile device which has a touch-sensitive display without a touch. The convolutional neural network (CNN) is a category of deep learning neural networks. CNN represents a huge advance in image recognition. They are used to analyses the visual images and image classification. A convolutional neural network (CNN) is used to extract features from images. This eliminates the need of manual feature work extraction. The features from the set of images are not trained they are learned while the network trains on a set of images. It makes extreme accuracy for the deep learning models. Documents in the training set involvement of the learned features. A

particular amount dataset will be provided to detecting character recognition.

II. RELATED WORK

Application supports drawing of gestures and writing of English text over air in front of the mobile camera by using bare fingertip. Convolution neural networks has been used for character recognition. The method presented will be applicable to other fields requiring hand gesture recognition for distant interaction with mobiles and computers [1]. Drawing such gestures in air is a recently emerging field in recent times. In this system, a combination of computer vision and convolution neural networks is used for detecting drawn gesture and recognizing it. Generally, video analysis procedure has three major steps: firstly, detecting of the object, secondly tracking its movement from frame to frame and lastly analyzing the behavior of that object. For object tracking, fsystem different issues are taken into account; selection of suitable object representation, feature selection for tracking, object detection and object tracking. Few of the methods employed, use colored fingertip for tracking the motion of the finger as proposed in [2]. CNN and LSTM are used for feature extraction and recognize sequence data, respectively. The complete network diagram is shown in the network consisted of input, convolution, pooling, LSTM unit, dense layers. The input is one-dimensional sequential data. The input layer dimension is set to the maximum length of the trajectory sequence which is 300 and 800 for RTD and RTC, respectively. The first convolution layer contains a filter size of 256 filesystem by a pooling layer. Max pooling is used as a pooling layer that down samples the CNN network. Filter size 512 is used for the next convolution layer.

III. METHODOLOGY

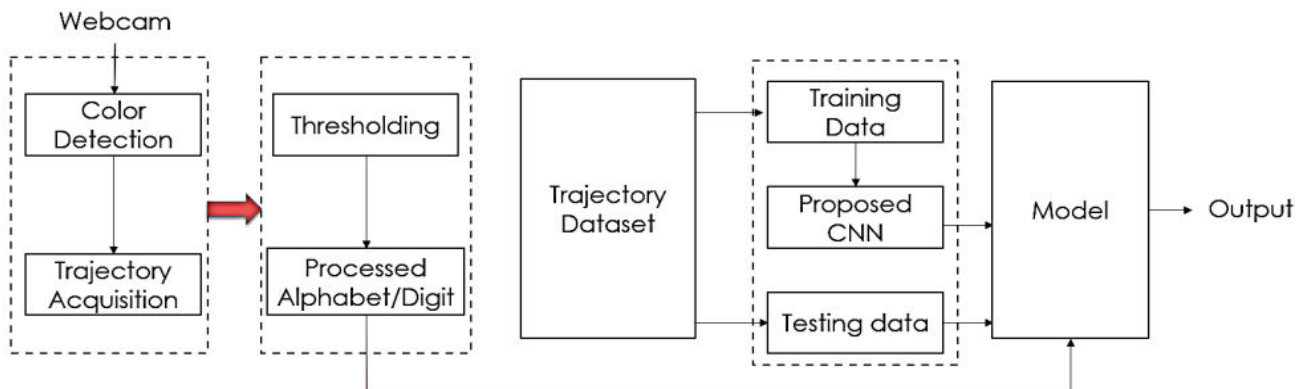


Fig.1. Block Diagram

From the above fig.1. system developed a model using an object as a reference to track motion of colored object. System proposed a model by tracking a colored object through system webcam from system system. The images captured by the camera module are processed to detect the motion of the object. This may involve detecting the edges of the object and tracking their movement over time. The motion of the object is analyzed to detect the gestures made by the user. This may involve segmenting the motion into discrete time intervals and detecting the presence or absence of written characters/digits within each interval. Once a characters/digits is detected, it is analyzed to recognize the character or digit being written by the user. This involved the use of deep learning techniques such as convolutional neural networks to classify the written character/digits into different categories based upon on the trajectory dataset. The recognized text is displayed on a screen or output device, allowing the user to see the characters/digits they have written in the air.

III.A.CNN MODEL

The first layer of the model is a convolutional layer with 32 filters of size 3x3 and a ReLU activation function. This layer takes an input image of size 28x28 pixels with a single channel (grayscale). The second layer is another convolutional layer with 64 filters of size 3x3 and a ReLU activation function. After each convolutional layer, a max pooling layer with a 2x2 pool size is added to reduce the spatial dimensions of the feature maps. A dropout layer with a rate of 0.25 is also added to prevent overfitting during training by randomly dropping out a fraction of the connections. The output of the second max pooling layer is then flattened and passed through a fully connected layer with 128 neurons and a ReLU activation function. Another dropout layer with a rate of 0.5 is added before the final output layer, which consists of 26 neurons (one for each letter of the English alphabet) and a SoftMax activation function. The SoftMax function produces a probability distribution over the 26 classes, and the predicted class is the one with the highest probability. The model is trained using a categorical cross-entropy loss function and the Adam optimization algorithm.

III.B.COLOR DETECTION

Color detection can be used in recognition of text writing in air using deep learning to isolate the foreground (object doing the writing) from the background, and then extract the information necessary for recognition. One common approach is to use computer vision techniques to detect the hand or object doing the writing by tracking a specific color range in the video frames captured by a camera.

III.C.TRAJECTORY ACQUISITION

This is achieved by capturing live video feed from a camera pointed towards the user's hand. The captured frames are then processed to extract the hand region using image processing techniques such as background subtraction, skin color detection, or motion detection. Once the hand region is extracted, the trajectory of the hand movement is obtained by tracking the movement of the hand across successive frames. The trajectory of the hand movement is then used as an input to a deep learning model to recognize the letters. The model is trained on a large dataset of hand-drawn letters and can recognize the letters with high accuracy. The trajectory acquisition process plays a crucial role in the accuracy of the system, as any errors or noise in the trajectory can affect the recognition performance of the system.

III.D.THRESHOLDING

In proposed system, the thresholding technique is applied on the input image to obtain a binary image. This is done using OpenCV's `cv2.threshold()` function, which applies a fixed level thresholding to the input image. A threshold value is selected such that pixel intensities below the threshold are set to 0 (black), and pixel intensities above the threshold are set to 255 (white). This binary image obtained from thresholding is then used for contours detection, which identifies the contours of the handwriting. These contours are then used to compute the trajectory of the handwriting, which is used for recognition.

IV. RESULTS

The recognition of text writing in the air can be evaluated based on several metrics, including accuracy, precision, recall. These metrics are commonly used in evaluating machine learning models for object detection, image classification, and text recognition tasks. In the case of recognizing text writing in the air, accuracy refers to the percentage of correctly recognized letters, numbers, and symbols. Precision measures the proportion of correctly detected gestures out of all the gestures detected, while recall measures the proportion of correctly detected gestures out of all the gestures that should have been detected. The result of the recognition system can also be evaluated by comparing it with the ground truth data. The ground truth data is the correct text that was actually written in the air during the test phase. The performance of the recognition system can be evaluated by comparing the recognized text with the ground truth data.

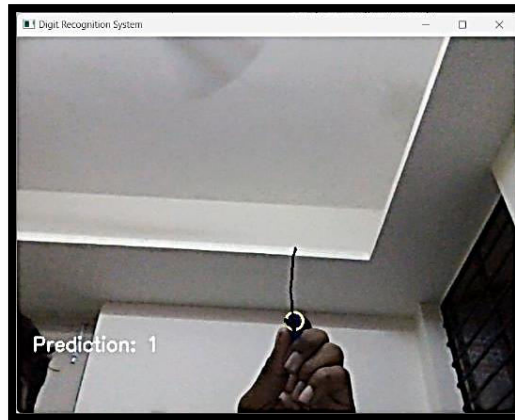


Fig.2. Digit Prediction

Result show prediction of digit '1' in the above frame

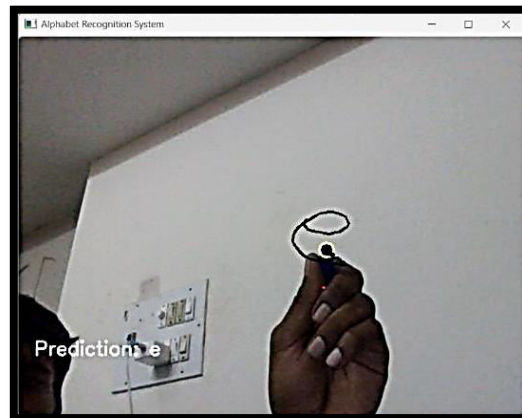


Fig.3. Alphabet Prediction

Result show prediction of alphabet 'e' in the above frame

IV.A. ACCURACY

Epoch	Loss	Accuracy
1	0.8378	0.7391
2	0.2912	0.9122
3	0.1987	0.9403
4	0.1553	0.9545
5	0.1285	0.9625
6	0.1104	0.9684
7	0.0980	0.9715
8	0.0876	0.9752
9	0.0798	0.9776
10	0.0733	0.9793

Fig.4. Digit Performance

The presented results demonstrate the performance of the proposed deep learning model for digit recognition, which has been trained and evaluated on a dataset of 70,000 samples of digits obtained from EMNIST. The model's accuracy and loss system tracked at every epoch during training, showing a clear trend of increasing accuracy and decreasing loss with each epoch.

Epoch	Loss	Accuracy
1	0.7472	0.8899
2	0.3185	0.9232
3	0.2289	0.9398
4	0.1811	0.9456
5	0.1486	0.9498
6	0.1259	0.9531
7	0.1068	0.9550
8	0.0918	0.9562
9	0.0788	0.9567
10	0.0681	0.9572

Fig.5. Alphabet Performance

The presented results demonstrate the performance of the proposed deep learning model for digit recognition, which has been trained and evaluated on a dataset of 124,800 samples of alphabets obtained from EMNIST. The model's accuracy and loss system tracked at every epoch during training, showing a clear trend of increasing accuracy and decreasing loss With each other

V. CONCLUSION

In conclusion, the recognition of text writing in air is a promising and challenging research area with a wide range of potential applications, such as hands-free text input for people with disabilities, virtual reality interfaces, and gesture-based communication. The methodology for proposed system involves collecting a large amount of training data, pre-processing the data, selecting a suitable deep neural network architecture, training and testing the model, and deploying it in a real-time system. Object detection can be used to detect the hand, gesture, and text in the air can be used to recognize the text. The success of proposed system will depend on several factors, such as the quality and quantity of training data, the choice of deep neural network architecture, the optimization of hyperparameters, and the accuracy of object detection. Handwritten characters have been recognized with more than 97% test accuracy. This can be also further extended to identifying the handwritten characters of other languages too.

REFERENCES

- [1] M. Chen, G. AlRegib, and B.-H. Juang, "Air-Writing Recognition—Part I: Modeling and Recognition of Characters, Words, and Connecting Motions," *IEEE Trans. Human-Machine Syst.*, vol. 46, no. 3, pp. 403–413, Jun. 2016.
- [2] M. Chen, G. AlRegib, and B. H. Juang, "Air-Writing Recognition - Part II: Detection and Recognition of Writing Activity in Continuous Stream of Motion Data," *IEEE Trans. Human-Machine Syst.*, vol. 46, no. 3, pp. 436–444, 2016.
- [3] Y. Luo et al., "An Attitude Compensation Technique for a MEMS Motion Sensor Based Digital Writing Instrument," in *2006 1st IEEE International Conference on Nano/Micro Engineered and Molecular Systems*, 2006, pp. 909–914.
- [4] J. S. Wang and F. C. Chuang, "An accelerometer-based digital pen with a trajectory recognition algorithm for handwritten digit and gesture recognition," *IEEE Trans. Ind. Electron.*, vol. 59, no. 7, pp. 2998–3007, 2012.
- [5] C. Amma, M. Georgi, and T. Schultz, "Airwriting: A systemarable handwriting recognition system," in *Personal and Ubiquitous Computing*, 2014.
- [6] C. Qu, "Online Kinect Handwritten Digit Recognition Based on Dynamic Time Warping and Support Vector Machine," *J. Inf. Comput. Sci.*, vol. 12, no. 1, pp. 413–422, 2015.
- [7] S. Poularakis and I. Katsavounidis, "Low-Complexity Hand Gesture Recognition System for Continuous Streams of Digits and Letters," *IEEE Trans. Cybern.*, vol. 46, no. 9, pp. 2094–2108, 2016.
- [8] X. Zhang, Y. Sugano, M. Fritz, and A. Bulling, "MPIIGaze: Real-World Dataset and Deep Appearance-Based Gaze



Estimation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 41, no. 1, pp. 162–175, Jan. 2019.

[9] C. Palmero, J. Selva, M. A. Bagheri, and S. Escalera, “Recurrent CNN for 3D Gaze Estimation using Appearance and Shape Cues,” May 2018.

[10] H. Ranganathan, S. Chakraborty, and S. Panchanathan, “Multimodal emotion recognition using deep learning architectures,” in 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), 2016, pp. 1–9.

[11] J. K. Sharma, R. Gupta, and V. K. Pathak, “Numeral Gesture Recognition Using Leap Motion Sensor,” in Proceedings - 2015 International Conference on Computational Intelligence and Communication Networks, CICN 2015, 2016.



SJIF Scientific Journal Impact Factor

Impact Factor: 8.423



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details