Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 7.056

IJCSMC, Vol. 10, Issue. 8, August 2021, pg.41 – 45

HANDWRITTEN CHARACTER RECOGNITION USING CONVOLUTIONAL NEURAL NETWORKS

Dr. D.J. Samatha Naidu[1]; T. Mahammad Rafi[2]
[1]Principal, APGCCS, Rajampet, Kadapa, A.P. India
[2]MCA Department & APGCCS, Rajampet, Kadapa, A.P. India
[1] samramana44@gmail.com; [2] rafiqrafiq7613@gmail.com
DOI: 10.47760/ijcsmc.2021.v10i08.007

Abstract: Handwritten character Recognition is one of the active area of research where deep neural networks are been utilized. Handwritten character Recognition is a challenging task because of many reasons. The Primary reason is different people have different styles of handwriting. The secondary reason is there are lot of characters like capital letters, small letters & special symbols. In existing were immense research going on the field of handwritten character recognition system has been design using fuzzy logic and created on VLSI(very large scale integrated)structure. To Recognize the tamil characters they have use neural networks with the Kohonen self-organizing map(SOM) which is an unsupervised neural networks. In proposed system this project design a image segmentation based hand written character recognition system. The convolutional neural network is the current state of neural network which has wide application in fields like image, video recognition. The system easily identify or easily recognize text in English languages and letters, digits. By using Open cv for performing image processing and having tensor flow for training the neural network. To develop this concept proposing the innovative method for offline handwritten characters. detection using deep neural networks using python programming language.
Keywords— Handwritten, Character, segmentation, EMIST Dataset, Self organizing map

I. INTRODUCTION

Handwriting Recognition is one of the areas where mostly researcher uses deep neural networks. Recognizing handwriting is a simple task for human beings but a daunting task for computers. Handwriting recognition systems are categories into two types: Online and Offline. In online handwriting recognition system, the handwriting of the user is recognized based on the user is writing. But in offline handwriting recognition system, the handwriting of user is available as an image. Handwriting recognition is a complicated task because of many reasons. The primary reason is that different people have different styles of writing in different languages. The secondary reason is there are lot of characters like Capital letters, small letters, Digits and Special symbols. A large data samples are required to train a near-accurate neural

network model. To develop a good system an accuracy of at least 98% is required. However even the most modern and commercially 1available systems have not been able to achieve such a high accuracy.

We used the Convolutional Neural Network Model in our system. We used the publicly available EMNIST Dataset which contains samples of handwritten characters from thousands of writers in many other languages. The neural network model which we have used is Convolutional Neural Network. CNN's are neural networks which have huge applications is field of Computer Vision. The neural network uses some model for trained using TensorFlow which is an open-source library used for Machine learning applications. OpenCV (computer vision) was used to perform various image processing operations like segmentation, thresholding and Morphological Operations. OpenCV is an open-source library which is used for Image processing.

II. RELATED WORK

Immense research is going on the field of handwritten character recognition using neural networks. Many researchers have developed systems for handwritten character recognition. We have studied some of the systems: A character recognition system has been designed using fuzzy logic. The system developed by them can be created on a VLSI structure. They have made use of hamming neural network in their system. An innovative method for recognition of handwritten Tamil characters using Neural Networks has been developed in different styles. They have made use of Kohonen Self Organizing Map(SOM) which is an unsupervised neural network. The system developed by them can be used for recognition of Tamil characters as well as for the recognition of other Indic languages. Their system produces near accurate results but sometimes produces errors if the handwritten characters are not properly segmented. One of the Authors has presented a unique method for authenticating a person based on their handwriting. The author has used the Multi-layer feed forward neural network in their system. The author has proposed in this paper that the height and width of a handwritten alphabet is unique for each and every person. The author has presented a method for recognition and identification of a person from their handwriting. A novel method for handwritten character recognition has been designed which does not use feature extraction. They have implemented their system in Matlab. Their system uses a feed forward neural network with backpropagation. One of the authors have proposed a unique method for handwriting recognition. Their system uses Self Organizing Map for feature extraction. They have used a recurrent neural network for learning based on Deep learning. Researchers are conducted their experiment on recognition of Japanese characters.

III.PROPOSED WORK

In the proposed system, the EMNIST dataset is extended by adding some more characters from the English language. First, the input image is provided and is converted into a gray-scale image and normalized in such a way that it represents the same resolution (28 x 28) as that of the EMNIST dataset. CNN is trained using the EMNIST dataset and uses it as a classifier which can yield better results in comparison with other machine learning algorithms. The feature vectors are extracted from the input image and provided to the trained model of Convolution Neural Network which recognizes and provides the specified output.

1. PRE-PROCESSING

Pre-processing of the input image is carried out by converting the given image into a gray-scale image. Usually, a normal colored image consists of three channels- red channel, green channel, blue channel commonly known as RGB. Then the colored image is converted to the gray-scale image which consists of a single monochrome channel in order to avoid unwanted noise in the image. The given input image would be of a varied size which maybe leads to loss of accurate prediction when the image is compared with that of a trained convolutional neural network. So the image is resized and placed upon an empty 28 x 28 pixel blank image so that the image resolution matches the resolution of the EMNIST dataset.

2. FEATURE EXTRACTION

Feature extraction is that the process of remodelling the input file into a group of features which will alright represent the input file. Feature extraction is related to dimensionality reduction. When the input file is just too large to be processed, then it are often transformed into a reduced set of features (also named a feature vector). Determining a subset of the initial features is named feature selection. The selected features are expected to contain the relevant information from the input file in order that the specified task are often performed by using this reduced representation rather than the complete initial data. After resizing the

image, pixel values are obtained in the form of the 1D array which represents values between 255 and 0 based on pixel intensity.

3. MIN MAX SCALER

The min-max scalar sort of normalization uses the mean and variance to box all the info into a variety lying between a particular min and max value. It transforms features by scaling each feature to a given range. This estimator scales and translates each feature individually such it's within the given range on the training set, i.e. between zero and one. This transformation is usually used as an alternate to zero mean, unit variance scaling. It essentially shrinks the range such the range is now between 0 and 1 (or -1 to 1 if there are negative values). The Min Max Scaler is that the probably the foremost famous scaling algorithm, and follows the subsequent formula for every feature: (xi–min(x))/ (max(x)–min(x)).

4. IMAGE NORMALIZATION

Normalization may be a process that changes the range of pixel intensity values. Normalization is usually called contrast stretching or histogram stretching. In this input image, the normalization is administered by removing the background pixels and therefore the character alone are going to be provided because it is within the image. This can be done by using a random value so that the background pixels will have a value certainly less than the pixel values of shades of the character. In this way, the image is normalized such that the image is similar to the values in the EMNIST dataset. In this image, the pixel values are more than 0 for the region where the character 'A' is written and all other regions have pixel values 0 after image normalization.

5. CLASSIFICATION

Convolutional neural network is used as a classifier for classifying the handwritten character from the input image. A CNN consists of an input and an output layer, also as multiple hidden layers. The hidden layers of a CNN typically contains convolutional layers, pooling layers, fully connected layers, and normalization layers. A CNN consists of three major components which are the convolutional layer, pooling layer, and output layer. The activation function that is commonly used with CNN is ReLU which stands for Rectified Linear Unit. The convolution layer will compute the output of neurons that are connected to local regions in the input, each computing a dot product between their weights and a small region they are connected to in the input volume. A pooling layer may be a sort of non-linear down-sampling. Max pooling is that the commonest which partition the input image into a group of non-overlapping rectangles and, for every such sub-region, outputs the utmost. RELU applies the non-saturating activation function. It increases the nonlinear properties of the choice function and of the general network without affecting the receptive fields of the convolution layer. A rectified linear measure has output 0 if the input is a smaller amount than 0, and raw output otherwise. Its value is obtained based on the formula which is as follows f(x) = max(x).

IV. METHODOLOGY

The backend of our system performs two important things. The first thing is hosting the pre-trained neural network model to serve predictions. The second thing is performing image processing operations on the image of handwritten text which is to be recognized. At the backend, we have a Neural network model trained using Tensorflow and a python script that is equipped with the OpenCV library. We have used the Convolutional Neural network model.

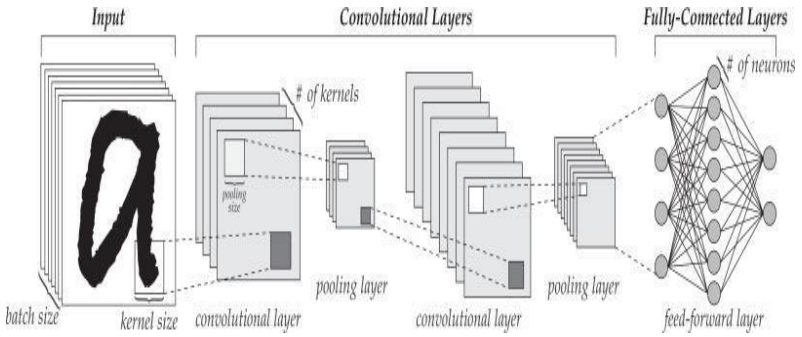


Figure: 4.1 A Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) is the current state-of-art neural network that has wide applications in fields like Image and Video Recognition, Natural Language Processing, Recommender systems. CNN's are biologically inspired neural networks. CNN's are very good at image recognition. In the case of CNN, the input is a multi-channelled image(Often an image having Red, Green, and Blue channels). A CNN comprises a stack of Convolutional layers and a Max-pooling layer followed by a fully connected layer. The convolutional layer is that the most vital layer of the network. It performs the convolution operation. The pooling layer comes after the convolutional layer. This layer is needed because, in the case of larger images, the number of trainable parameters can be very large. This increases the time taken to coach a neural network and isn't practical. The pooling layer is used to reduce the size of the image. We used the NIST database which contains thousands of images of handwritten characters. Some of them are shown below. However, these images were original of size 128x128 pixels. The images within the training set were cropped to a size of 28x28. Reducing the dimensions of images decreases the general time is taken to coach the neural network model. After training the Neural network model, and accuracy

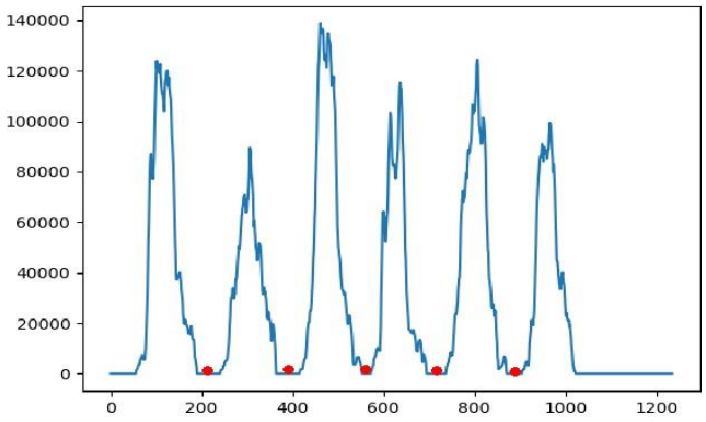
1) Pre-processing: this is often the primary step performed in image processing. In this step, the noise from the image is removed by using median filtering. Median filtering is one of the most widely used noise reduction techniques. This is because in median filtering the edges in the image are preserved while the noise is still removed.

2) Conversion to Gray-Scale: After the pre-processing step, the image is converted into grayscale. Conversion into grayscale is necessary because different writers use pens of different colors with varying intensities. Also performing on grayscale images reduces the general complexity of the system.

3) Thresholding: When an image is converted into grayscale, the handwritten text is darker as compared to its background. With the assistance of thresholding, we will separate the darker regions of the image from the lighter regions. Thus due to thresholding, we will separate the handwritten text from its background.

4) Image Segmentation: A user can write text within the sort of lines. Thus the thresholded image is first segmented into individual lines. Then each individual line is segmented into individual words. Finally, each word is segmented into individual characters. Segmentation of image into lines is carried out using the Horizontal projection method. First, the thresholded image is inverted in order that background becomes foreground and vice-versa.

Now the image is scanned from top to bottom. While scanning, the sum of pixels in each row of image is calculated. The sum of pixels will be zero if all the pixels in one particular row are black. The sum will be non-zero if some white pixels are present in a row. After this a horizontal histogram is plotted in which the X-axis represents the Y-coordinate of image(Starting from Top to Bottom) and the Y-axis represents the sum of pixels in the row corresponding to the Y-coordinate. The horizontal histogram is plotted using MatPlotLib and is as shown in Fig



(a)

Figure: 4.2 Horizontal Histogram of Image

The points marked in red are the points corresponding to the rows where sum of pixels are zero. After identifying all such rows we can easily segment handwritten text into lines at these points. Now once the image is segmented into lines, each line must be further segmented into individual words. Segmentation of a line into words can be performed using the Vertical projection method. For segmenting line into words, we can make use of the fact that the spacing between two words is larger than the spacing between two characters. To segment a single line into individual words, the image is scanned from left to right and sum of pixels in each column is calculated. A vertical histogram is plotted in which the X-axis represents the X-coordinates of image and Y-axis represents the sum of pixels in each column. The vertical histogram is as shown below: As we can see the points which are marked as red in Fig.5(a) are the points corresponding to the columns where sum of pixels is zero. The region where the sum of pixels is zero is wider when it is a region separating two words as compared to the region which is separating two characters. After segmenting a line into words, each word can be separated into individual character using similar technique as explained earlier. Now these individual characters are given to the pre-trained neural network model and predictions are obtained. Using this the final predicted text is sent back as a response to the user.

V. IMPLEMENTATION

To design this offline handwritten character recognition system, we have used various tools like Python, OpenCV and Tensorflow. VI. CONCLUSION In this project I have proposed a deep learning architecture with training character images and testing made on different characters and that correctly classifies our test images. The number of epochs used was stopped at particular number because it receives a cut point after which the accuracy was not improving and the loss was not decreasing on both training and validation data.

ACKNOWLEDGEMENT

I would like to extend my gratitude to Dr. D.J.Samatha Naidu for helping me in developing this system and providing the infrastructure.

REFERENCES

[1]. Wei Lu, Zhijian Li,Bingxue Shi . ”Handwritten Digits Recognition with Neural Networks and Fuzzy Logic” in IEEE International Conference on Neural Networks, 1995. Proceedings.,

[2]. P. Banumathi , Dr. G. M. Nasira .”Handwritten Tamil Character Recognition using Artificial Neural Networks” in 2011 International Conference on Process Automation, Control and Computing.

[3]. B. V. S. Murthy. ”Handwriting Recognition Using Supervised Neural Networks” in International Joint Conference on Neural Networks, 1999. IJCNN ’99.

[4]. J.Pradeep, E.Srinivasan, S.Himavathi. ”Neural Network based Handwritten Character Recognition system without feature extraction” in International Conference on Computer, Communication and Electrical Technology ICCCET 2011.

[5]. Shun Nishide, Hiroshi G. Okuno, Tetsuya Ogata, Jun Tani. ”Handwriting Prediction Based Character Recognition using Recurrent Neural Network” in 2011 IEEE International Conference on Systems, Man, and Cybernetic.

[6]. T. Kohonen, Self-Organization and Associative Memory, 2nd Ed., New York, Springer, 1988.

[7]. Y. Yamashita and J. Tani, Emergence of Functional Hierarchy in a Multiple Timescales Recurrent Neural Network Model: A Humanoid Robot Experiment, PLoS Computational Biology, Vol. 4, e1000220, 2008.

[8]. Lei Wang, Lei Zhang , Yanqing Ma . ”Gstreamer Accomplish Video capture and coding with PyGI in Python language” in 2017 24th Asia Pacific Software Engineering Conference (APSEC) .

[9]. Rahul R. Palekar , Sushant U. Parab , Dhrumil P. Parikh , Prof. Vijaya N. Kamble. ”Real Time License Plate Detection Using OpenCV and Tesseract” in International Conference on Communication and Signal Processing.

[10]. “OpenCV” https://en.wikipedia.org,[Online] Available: https://en.wikipedia.org/wiki/OpenCV/. [Accessed 05 March 2018]. [11] Fatih Ertam, Galip AydOn. ”Data Classification with Deep Learning using Tensorflow” in 2017 International Conference on Computer Science and Engineering (UBMK).

