



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 Issue: IV Month of publication: April 2019

DOI: <https://doi.org/10.22214/ijraset.2019.4194>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

A Mechanism for Offline Character Recognition

Salman Khan

Department of Computer Science and Engineering , Jamia Hamdard , New Delhi-110062

Abstract: Character recognition is an exciting and interesting computer vision research field due to distinct human handwriting and can be adapted to recognize characters. Several upgrades to analyze handwritten data have been made, still data cannot be analyzed with 100% accuracy by the system. In order to improve the accuracy, we propose a method to recognize the characters of English language. This proposal is limited to English characters only. In our proposal, a Convolutional Neural Network (CNN) based on TensorFlow, an open source library for building machine intelligence applications, is designed for character recognition. Experimental outcomes demonstrate that the proposed model has finer accuracy and models are deployed quickly and easily.

Keywords: Character recognition, Machine learning, CNN, LSTM, CTC, TensorFlow

I. INTRODUCTION

Institutions, agencies, banks, medical management, and law courts keep a number of handwritten records such as answers booklets, cheques, datasheets etc. All these data are in a huge amount which is difficult to handle simultaneously and can take immense time. With passing time documents gets older and may get damaged. So in order to prevent from all these, there is a need to build an efficient handwriting recognition system which can transform all those documents and scripts into digital form and make our work easier, faster and stable. Developing an efficient handwritten character recognition system which can maintain a high accuracy rate, is a complex task. Accuracy is the key point in many of the character recognition system. Accuracy of input document is directly proportional to its quality. In any handwritten document, we can see there are many characters which seem to be touched to one another or may be superimposed [1]. Therefore, it becomes necessary to segment the characters individually. The need for segmentation is due to distinct human handwriting such as their rapidness of writing, gripping of a pen and tone. Different regional languages like Urdu, Arabic, Hindi makes recognition task more complicated to analyze and interpret the characters from the images. Thus, it stands as a challenging area for the researchers. Offline Handwriting Recognition is a process in which system take scanned inputs such as paper scripts, text images, handwritten character and perform sampling over them to convert into electronic or digital form. Different image processing techniques for eliminating noise, irregularities of an image are applied in addition. This improves the precision rate and facilitates interpretation. Computer can mimic the brain through neural network. Neural networks are especially useful for complex and large problem-solving. On the basis of character sets, we train the neural network algorithm and that training data is used for classification. In offline character recognition, different image processing techniques [2] are applied to the text to restructure its geometry. Few of them are noise removal, opening, image averaging, binarization, thickening, closing. After preprocessing, segmentation phase begins where an image of character sequence is broken down into sub-images of a single word module. Then, features are extracted from this image module based on the text structure. At last, neural network are used for the classification and we get the digitized image as an output. Accuracy depends on the type of algorithms used for classification, and on the quality and size of the image. Figure 1, shows various stages of the offline character recognition process.

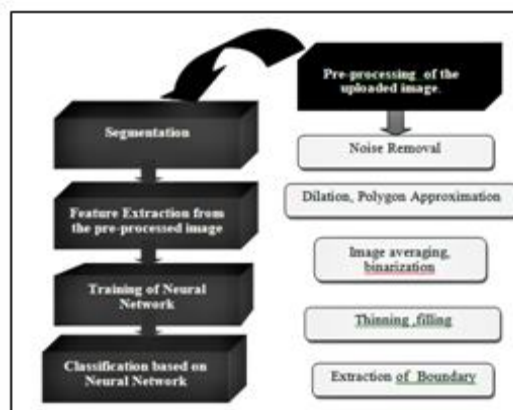


Fig 1: A Handwritten Character Recognition System

The handwritten character recognition system is somewhat difficult to design but it has a great impact on information technology in every aspect. It is used in biometric verification, and signature verification. In offline biometric verification, data are recorded through cameras or scanners, preprocessings techniques are performed then feature extraction and classification are done to extract the relevant features. Online shopping companies use OCR for customer's signature, in post-offices for postcards interpretation where data is recorded at run time on digital devices such as tablets, smartphones. Beside this, it is used in the mobile device in various application like ShareIt, CamScanner, Whatsapp to scan the QR code. Some of the best OCR software for PCs are OmniPage17, Abbyy Fine Reader, Readiris [3] and for Android device Text Fairy, Google Keep [4]. Character recognition also has a few deficiencies. The large image can create network congestion as there is an increase in the number of neurons resulting in a slow recognition process. Most of the documents become too old and their ink fades, preventing the system from properly recognizing and interpreting characters. In addition, many documents come into contact with water and become difficult for the system to recognize it. It may be difficult to distinguish when writing "i" as people sometimes put a circle instead of a dot. Also, it is difficult to distinguish between O and 0 [5]. This impacts recognition accuracy.

II. INTRODUCTION TO TENSOR FLOW

Tensor Flow is a Google-published open source software library to make the development of deep- learning models or neural networks easier for developers. Tensor flow is probably one of the most popular deep learning frameworks, it is an open source library which is used for numerical computation and extensive machine learning. The tensor flow outpaces everything and it's an intrinsic component of almost all machine learning and deep learning things [6]. It's obviously the brainchild of Google's brilliant people and is a major part of all its applications. As it is open-source it has many opportunities for development and improvement so it is available on PCs and mobile platforms. This clearly is a huge positive and brings more people together and therefore it leads to one of the largest community of machine learning coming together under one platform. So, tensorflow has the biggest community of learners and associates in comparison to all other machine learning frameworks at present and most of the big tech companies from social media till an airline company makes use of tensor flow. In tensor flow, tensors are assigned with a rank i.e. unit of dimensionality. Ranks can be 0,1,2,3,n and so on. Rank zero tensor is scalar-tensor and rank one tensor is vector tensor. Scalar-tensor has the only magnitude whereas vector tensor has both magnitude and direction. If we increase the tensor rank, then this means we add a further dimension to the tensor and it becomes a row and column matrix with a two rank. Likewise, the tensor rank would increase with an increase in dimension. Tensors have multiple data types like integer, float, of 8 bits,16 bits, and 32 bits, string and boolean. In tensor- flow, however, the data type should not be specified for a tensor as the correct data type is allocated automatically.

III. IMPLEMENTATION

Our aim is to develop a neural network in a tensor flow which can identify English characters from an image.

A. Data Acquisition

This experiment is conducted on datasets taken from IAM handwriting database [7] and training is carried out on IAM datasets. We can also add images other than IAM sets as per our desire. Smaller size image should be used for better classification and accuracy. Here, umlaut alphabets ("Ä") and symbols are not taken into consideration. IAM database is a collection of enormous sets of handwritten English words used for different recognition tasks. A handwritten word from IAM datasets is shown below in Figure 2.

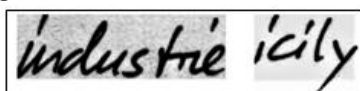


Fig 2: A handwritten word from IAM dataset [7]

B. Pre-processing

This is a very crucial process of recognition. It tends to improve the structure of an image. There are mainly four pre-processing techniques that are applied to an image those are cropping to resize the image, thresholding or contrast stretching to convert the image into binary form, thickening (dilation) to thicken the characters, opening to remove noise and closing to fill the narrow gaps. All these techniques are applied to an input image so that it looks similar to those images in IAM datasets. If a word is too lean vertically in either of the direction, then normalization algorithm can be performed as proposed by Vinciarelli and Luetttin [8], figure 3 shows an input sample image and morphological operations cropping, thresholding, thickening and performed over it.



Fig 3: Pre-processing is applied on a word 'Salman'

C. Word Segmentation

In word segmentation, sentence is decomposed into several words. Now, each word behaves as a sub image. Various word segmentation have been proposed in the past , of which scale-space algorithm [9] is highly efficient for detection of blobs. Gaussian low pass filter are used for formation of blobs.

A Gaussian Low Pass Mask [10] can be given by the equation,

$$G(x, y) = -e^{-\frac{x^2+y^2}{2\sigma^2}}$$

where σ is the cut-off value , it is inversely proportional to the rate of Gaussian low pass filtering . If cut-off value is decreased, there would be more blurring and vice-versa. Before applying the mask, image averaging should be done to remove noises. After applying Gaussian low pass filter, binarization is performed so that blobs can be graphed to sample input image to separate the connected pixels. Laplacian of Gaussian (LOG) [10] is used for the detection of blobs in an image. A Laplacian of Gaussian Filter can be given by the equation,

$$G(x, y) = -\left[\frac{x^2 + y^2 - 2\sigma^2}{\sigma^4}\right] e^{-\frac{x^2+y^2}{2\sigma^2}}$$

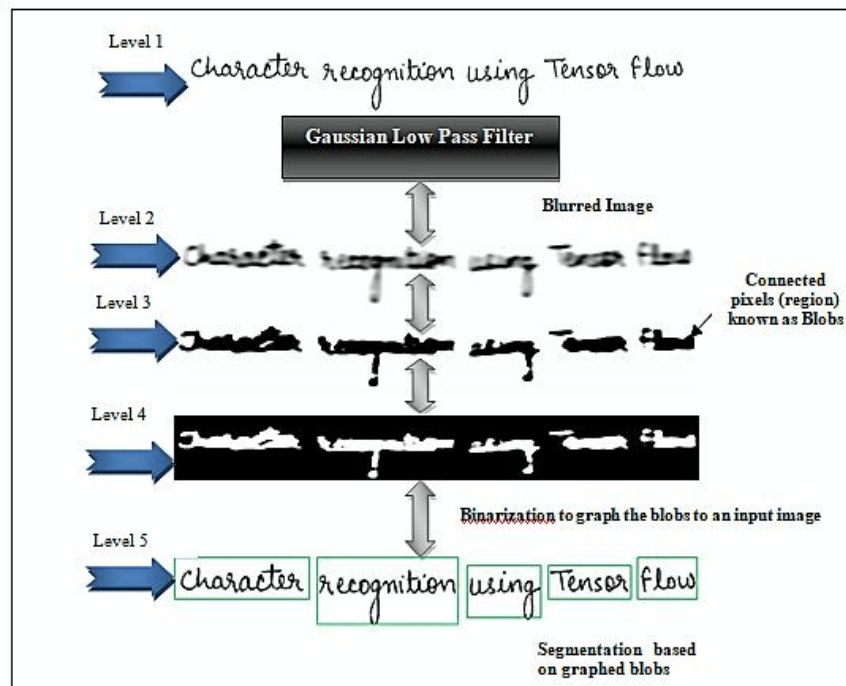


Fig 4: Segmentation

In figure 4, segmentation is divided into five levels. Level 1 shows an input image which has to be segmented and for this purpose, we make use of Gaussian low pass filter to create the blurring image. Level 2 shows a blurred image. Level 3 shows a thresholding image consisting of blobs. In order to extract those blobs, we can binarize the image as shown in level 4. After binarization, the image consisting of blobs is mapped to an input image to segment the words as shown in level 5. The segmented words are then feeded into the model for further operations.

IV.CHARACTER RECOGNITION

In Figure 5, the model consists of three main layers which are Convolutional Layers, Recurrent Layers and Connectionist Temporal Classification Layer.

A. Convolution Neural Network

CNN is a kind of neural network, in which the neuron of a layer is connected only to the small area of the layer just before it and does not connect completely to all neurons. A Simple Convolutional neural network model consists of a convolution layer, ReLu layer (activation function), Pooling layer and Fully connected layer [11].

B. Long Short Term Memory

LSTM is a specific kind of recurrent neural network, which is designed to acknowledge sequence patterns of information like characters, speech and knowledge from sensors and organization. LSTM (Long Short Term Memory) can be thought as of a kind of RNN which has a ability to learn long term dependencies [11].

C. Connectionist Temporal Classification

CTC is a kind of artificial neural network, which is designed to train recurrent neural network as well as to compute conditional probability and loss function of a text. It uses the concept of back-propagation. Also, it can be use as a decoder [12].

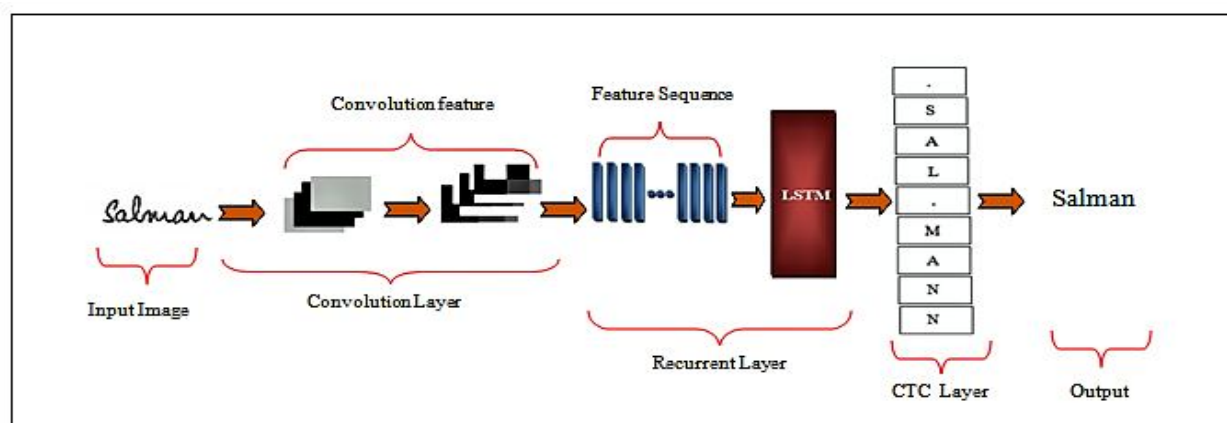


Fig 5: Basic Model of CNN, LSTM and CTC for Character Recognition System

D. Model Details

In figure 5, the process of feature extraction is shown, the input image is set to pixel size 118*30. Other layers used in convolution are activation function pooling layer and batch normalization to increase training speed. CNN consists of five layers. After that, an activation function is used. The use of activation function is to remove all negative values from the features and activate only that node which has some certain threshold value, also known as a non-RELU layer. Pooling layer of 2x2 with a stride of 2 is applied to shrink the image. RNN (LSTM) consists of two layers, that are used to yield better performance. Finally, CTC layer is used to train LSTM networks and to decode output matrix to a word.

E. Decoding

The word beam search method [13] is used to mapped the input sequence to the output sequence. The resulting output sequence is the required output image after decoding. CTC plays an enourmous role in decoding the output matrix to a word.

F. Training

Adam is a popular deep - learning algorithm because it gives good results immediately. One can use Root Mean Square Propagation (RMSProp) [14] Optimizer for training. This optimizer works fine on noisy images. The loss function of a text contained in batches is calculated and the loss value is used by Adam Optimizer to train the network.

V. EXPERIMENTAL RESULTS

I implemented the model on the Integrated Development Environment PyCharm Community Edition 2018.3.2. Experiments were conducted on Windows 10 with the hardware configuration Intel(R) Core™ i3 5005U CPU @ 2GHz, 4GB of RAM and on 64-bit operating system x 64-bit processor with updated version of Microsoft Visual C++ 2015 Redistributable (14.0.24212.0). This is necessary for specially Windows 10 users ,otherwise it can create run-time error on PyCharm. The version of tensor flow used is 1.12.0 and of OpenCV 4.0.0.12. Image dimension is kept small to 118*30 so as to lower training time. Apart from it ,batches are used. The model is trained with 75 IAM words with batch size of 15. Out of which 75 words are taken as training data and 25 testing data. Out of 25 testing data 18 words are recognized correctly by the model. Besides IAM dataset , model is tested with other 10 handwritten words by applying pre-processing to make them similar as IAM datasets ,out of them 4 words are recognized by the model. Overall, the model has achieved character accuracy of 62.85 %.

VI. CONCLUSION

This paper presents a character recognition model implemented using Tensor flow. In the proposed system, offline character recognition model is trained on datasets from IAM handwriting database. On empirical evaluations, it was found that it correctly recognize most of the words whose size are under given range and words are not too long. For the large sized words, some image processing techniques are applied and large space words are segmented using “scale space techniques”. Finally, the segmented words are feeded into the model. Word Beam Search method is used [13]. From the result, we have achieve 62.85 % accuracy in recognizing the characters.

REFERENCES

- [1] Anisha Priya, Surbhi Mishra, Saloni Raj, Sudarshan Mandal and Sujoy Datta, “Online and Offline Character Recognition: A Survey,” *IEEE Xplore* :pp 967-970.
- [2] Yasser Alginahi, "Preprocessing Techniques in Character Recognition. " *Intech Open*. <https://www.intechopen.com>.
- [3] https://www.simpleocr.com/OCR_Software_Guide.
- [4] <https://techwiser.com/5-best-ocr-app-for-android>.
- [5] Charles C. Tappert, Ching Y. Suen and Toru Wakahara. "The State of the Art in Online Handwriting Recognition." *IEEE* vol. 12 (1990).
- [6] <https://www.tensorflow.org/>.
- [7] U. Marti and H. Bunke "The IAM Database: An English Sentence Database for Offline Handwriting Recognition." *International Journal on Document Analysis & Recognition* vol. 5 (2002): pp 39-46.
- [8] Alessandro Vinciarelli, Juergen Luetin, "A New Normalization Technique For Cursive Handwritten Words," *Elsevier*, 2001:pp 1041-1050.
- [9] R. Manmatha, N. Srimal. "Scale Space Technique for Word Segmentation in Handwritten Documents." *Springer* vol.1682 (1999).
- [10] Rafael C. Gonzalez, Richard E. Woods. *Digital Image Processing*. Pearson Education International, 3rd Edition.
- [11] Radhika K, Bindu K. R. ,Latha Parmeshwaran, "A Text Classification Model Using Convolution Neural Network," *International Journal of Pure and Applied Mathematics*, 2018 Vol 119 :pp 1549-1554.
- [12] <https://towardsdatascience.com/intuitively-understanding-connectionist-temporal-classification-3797e43a86c>.
- [13] Scheidl, H. and Fiel, S. and Sablatnig, R, "Word Beam Search: A Connectionist Temporal Classification Decoding Algorithm," *IEEE*, 2018:pp 253-258.
- [14] A Gentle Introduction to Adams Optimizer, <https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/>.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)