Journal of Engineering Sciences

Vol 11, Issue 4 , April/2020

ISSN NO:0377-9254

HANDWRITTEN TEXT RECOGNITION USING TENSORFLOW

Anuradha[1] Ms. Nidhi Sengar[2]
Student[1] Assistant Professor[2]
Department Of Information Technology
Maharaja Agrasen Institute Of Technology, Rohini Delhi India

Abstract- Machine Learning comes under the field of computer science and technology that provides computers the ability to learn things without being programmed explicitly. Machine learning is used where it is not feasible to program and design algorithms with better performance; like a number of applications include filtering of emails, detection of malicious insiders working towards a data breach, optical character recognition (OCR) learning to rank, face recognition and computer vision. It became important to recognize handwriting as we start moving towards automated world. Deep learning is showing unbelievable results in the domain of visual and speech recognition. But still we lack in matching the accuracy of human vision and advancements will be continuing until we cross that limit. TensorFlow is one of the Google's open source machine learning as well as deep learning framework, which is convenient to build the standard deep learning model. Convolutional neural network is a unique model of deep learning; the advantage of using CNN is its powerful feature extraction capabilities of convolutional blocks. Based on the TensorFlow platform, a convolutional neural network model with five-convolution-layers has been created. The proposed system has been trained on samples of large collection of IAM database images and tested on sample images from user defines data set and in this experiment we found the highest recognition results.

Keywords: -Handwritten Character Recognition, Convolutional Neural Network, Feature extraction, TensorFlow

I. INTRODUCTION

A. OVERVIEW

Handwritten Text Recognition is a technique that is much needed in this world these days. Before implementing this technique properly, we have relied on writing with our hands which can lead to errors. Storing and accessing physical data efficiently is a difficult task [2]. Manual labor is required to maintain the correct organization of the data. Modern day technique is facilitating people to store the data over machines, where it is much easier to store, manage and access data. In addition, it also provides high security to the data. An example of Handwritten Text Recognition software is the Google Lens.

Handwritten documents are easy for humans to understand because we have the power to learn [3]. The Machines can also be induced with the same ability by the using Deep Learning and Artificial Intelligence. The research that deals with such kind of problem is known as the OCR or also known as Optical Character Recognition. It is a system that is used to convert electronic and image documents into digital text for machine readability.

In the last few decades, some of the feature extraction technology have been proposed such as histogram of oriented gradients and many techniques such as image recognition, character recognition, face detection, etc. used as prominent feature extraction methods [1].

OCR is further of two variants, HCR (Handwritten Character Recognition) which is intelligent identification of the handwritten documents and PCR (Printed Character Recognition). Due to the different handwriting styles of humans, we need greater recognition power. Many times the writing style of single person is different as many times he writes. Additionally, OCR is characterized into two types as

Offline and Online recognition systems on the basis of acquiring of the documents. Offline recognition system deals with the already written texts received through various input methods. Whereas in Online recognizing system, the writing is recognized at the same time it is written.

B. CONVOLUTIONAL NEURAL NETWORK

Convolutional Neural Network (CNN) is the most commonly used Deep Learning algorithm for image classification and also is the idea behind computer vision applications [4].

The most important advantage of using Convolutional Neural Network (CNN) as compared to its predecessors is that it automatically detects the useful features from the inputs without the supervision of human which is called feature map. Then for training the data, the spatial dimensions are reduced which gives the model faster computing power. Finally, the hidden layers give some values which are changed to a range of values to exactly classify the classes using activation function.

Convolutional Neural Network (CNN) has been successfully applied to recommender systems, natural language processing and many more. It is a very powerful and an efficient model which performs automatic feature extraction process to achieve very high accuracy like humans.

C. TENSORFLOW

TensorFlow officially became open source in November 2015. It is the deep learning system of second generation. Compared to the first generation system DistBelief, it has high speed of calculation, highly supported computer platforms and deep learning algorithms [5]. The stability of the system is also higher, so this is of great concern and is increasingly used by the users. It is a relatively high-level deep learning library, and C++ and Python are the supported languages. Users do not need to write complex code to build a neural network structure. With the rapid improvement of computing power, even simple computers can also build deep learning models on the TensorFlow platform, which reduces the cost of deep learning and the verification of algorithms becomes easy.

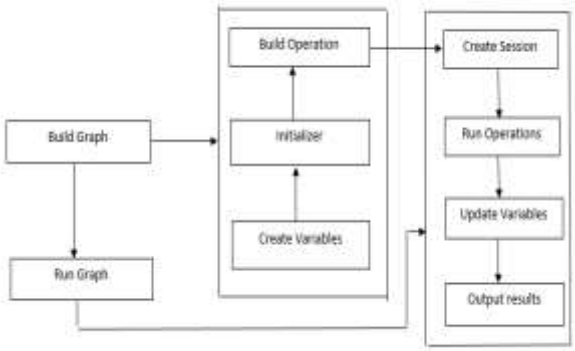

Figure 1. TensorFlow program development flow chart

TensorFlow is a product of Google Brain Team in Google's Machine Intelligence research organization for conducting research in the area of Machine Learning and Deep Neural Networks [4]. It has proven to be a scalable, clean, flexible and efficient library to work on. The operations of TensorFlow are arranged in a graph which is known as Computational Graph. TensorFlow gives some optimization nodes to search for parameters that can minimizes the cost function.

II. LITERATURE REVIEW

In 1959, Research from Grimsdale in the field of word recognition, is soonest endeavor to perceive the handwritten character. This research exhibited the utilization of examination by combination strategy being proposed by Eden. He demonstrates that the role of individual handwriting is limited to number of schematic highlight. This hypothesis was later used as a part of almost all strategies to support the methodologies in the field of text recognition.

Amit Choudhary [6] demonstrated an Off-Line Handwritten Character Recognition using Features Extracted by using Binarization Technique. It helps to extract features obtained by Binarization technique for recognition of English language handwritten characters. This algorithm delivers outstanding classification accuracy of 85.62 %.

Sonu Varghese K et al. [7] demonstrated Tri-Stage Recognition Scheme for Handwritten Malayalam Character Recognition. In the first step we will start setting up character groups in different classes based on the number of corners, loops, bifurcations and endings. In the second step we identify exact character

in the class based on the various feature extraction technique. In the final step we are checking the probability of occurrence of the character in the given position on the basis of defined rules for the making of words. Recognition conducted in different stages improves the efficiency, rate of recognition and accuracy of the system.

Parshuram M. Kamble [8] demonstrated handwritten Marathi character recognition using R-HOG Feature. The system has been tested with a large quantity of handwritten Marathi characters. From the results it can be concluded that the use of R-HOG based feature extraction method and FFANN based classification with high processing speed and accuracy is more accurate.

III. THE PROPOSED RECOGNITION SYSTEM

In Deep Learning the word 'deep' points to the number of hidden layers present in the Neural Network. Recognition of Handwritten Text is based on Deep Learning technique i.e., Convolutional Neural Network (CNN) which uses five hidden layers. It includes four main phases- Convolution, Pooling, Flattening and Fully-Connected Layer.

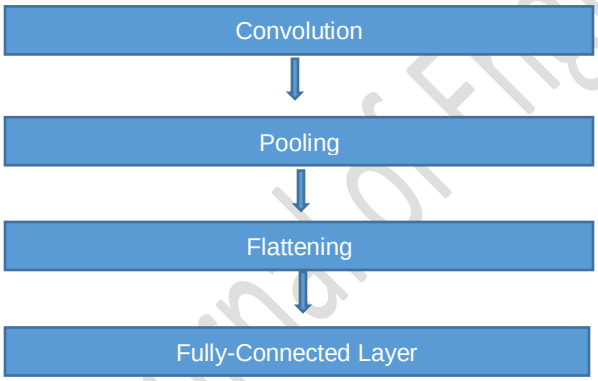

Figure 2. Steps involved in CNN

Convolution

Convolutional layer is the important building block of the whole Convolutional Neural Network (CNN) model. Convolutional layer creates a feature map by doing the Filtering process to the input images. The images are taken from the IAM dataset and then sent to the very first layer that is the Convolutional Layer. There is no separate preprocessing step for the input images from the dataset in Deep Learning model. The images are filtered using a convolutional filter and after that sent to the next layer for pooling.

Pooling

Pooling layer decreases the spatial dimensions of the image that makes it easy to lowers the number of parameters which both reduces the training time and combats overfitting [9]. There are two important pooling layers known as max-pooling and min pooling. In the max-pooling, it takes the maximum values from the selected region and in the min-pooling, it takes the minimum values from the selected region. The main advantages of reducing the spatial dimensions are:

- When the spatial information decreases, computational performance increases.
- Also, when the spatial information decreases simultaneously the parameters which are used to train the model gets reduced, thus it decreases the chances of overfitting.

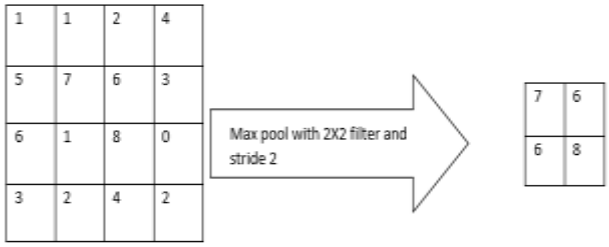

Figure 3. Max Pooling

Flattening

The last part of convolutional neural network (CNN) is a classifier that is the artificial neural network (ANN). For ANN we require input layer as a 1D feature vector. The process of conversion to 1D feature vector is known as flattening [9]. It takes the output of the convolutional layers, flattens all its structure to create a single long feature vector which is later used by the dense layer for the final classification process.

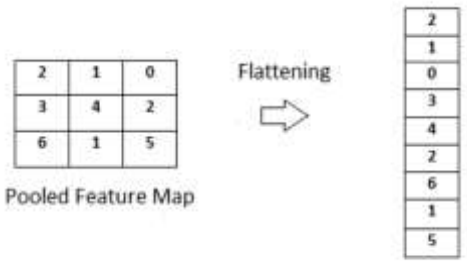

Figure 4. Flattening

Fully-Connected Layer

In a Convolutional Neural Network (CNN), the output layer is the Fully-Connected layer where the inputs from the various layers are flattened and are sent to transform the output into number of classes in the Neural Network. In this model, the activation function used is Rectified Linear Unit (ReLu). The activation function squashes the values into a range of value. Softmax is the last layer that predicts the accurate output among all the classes.

IV. RESULT AND DISCUSSION

This model is a stripped-down version of the Handwritten Text Recognition system. The implementation of this model only depends on numpy, openCV and TensorFlow imports. This model consists of 5 layers of CNN, 2 layers of RNN (LSTM) and the CTC loss and decoding layers. An overview of the Neural Networks is given below and following is the description in short:

- The input images are a gray-value images which are 128x32 in size.
- 5 layers of CNN map the input images to a feature sequence of size 32x256.
- 2 layers of RNN (LSTM) with 256 units propagate information through the sequence and then maps the sequence to a 32x80 sized matrix. Each matrix-element shows a score for one of the 80 characters in 32 time-steps.
- The CTC layer on the one hand calculates the loss value and the actual text (when training), and on the other hand it decodes the final text from the matrix with beam search decoding (when testing).
- Batch size is taken as 50.

The output of the CNN layers which is a sequence of length 32 is calculated. Every entry has 256 features. Then, these features are again processed by the RNN layers, however, some of the features already shows a high correlation with certain high-level properties of the input images. There are some features that are having a high correlation with characters, with duplicate characters and with character-properties such as looping.

The document to be scanned is

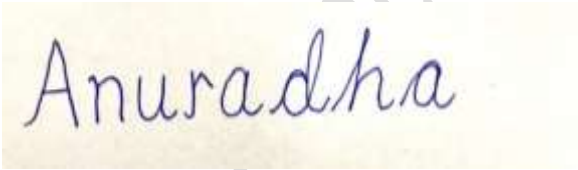

Figure 5. Test document 1

So the result is

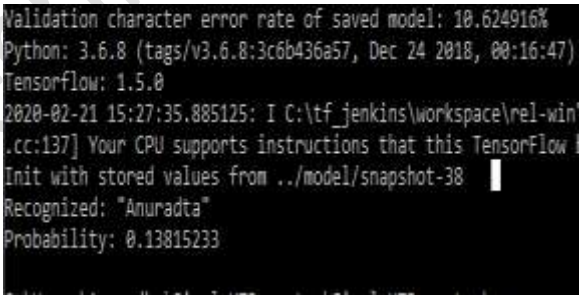

Figure 6. Result Image 1

It is evident from the results that text recognition is difficult in this type of condition.

From the resultant images, we can see that the text recognition failed to recognize “h” in the word.

Fig. 8 shows a result of the RNN output matrix for an image containing the text “project”. The matrix-entries, from top to bottom, corresponding to the characters: “ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefg hijklmnopqrstuvwxyz“!”()*+,&./#1234567890:-;?”. In most of the cases it is concluded that many times, the characters are predicted exactly at the same position as they are present in the input images.

Figure.7 Test document 2

The last character "t" is not aligned. But this doesn't make large difference, because the CTC operation is segmentation-free and does not take into consideration the absolute positions. From the image showing the scores for the characters "p", "r", "o", "j","e", "c","t" and the CTC blank label, the characters can easily be decoded: we only have to take the most probable character from each time-step, this forms the best path, after that we throw away repeated characters and at last all blanks: "p---r--o-jj--e--c…-t" → "p---r--o-j--e—c-…-t" → "project".

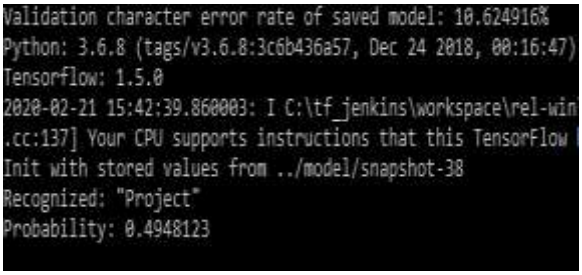

Figure 8. Result Image 2

V. CONCLUSION

This paper gives a detailed review of text recognition in the English language using TensorFlow and studies so many algorithms for recognition. The accuracy of text recognition fully depends on the quality and the nature of the image to be read. Current research does not deal with the cursive handwriting because it needs a high supervised system. In this paper, we have studied numerous papers with different algorithms to increase the accuracy of the result.

Apart from these, we have concluded that Recognition of Handwritten Characters based on Deep Learning with TensorFlow gives the most accurate classification and prediction values which can be taken for further researches. The training time taken by the Convolutional Neural Network (CNN) model is very less as compared with any other model. The error rate is also less as compared with the previous models. Convolutional Neural Network (CNN) is not a lot different from other machine learning models but it tries to find patterns in the dataset.

VI. FUTURE WORK

In the future, we need to improve the current performance of our project. For a better accuracy we will work on the following ideas:

- By increasing dataset-size and by applying various (random) transformations to the input images.
- By Removing all the cursive handwriting style images from the input images
- By Increasing the input size
- By adding more layers of CNN
- Replacing LSTM by two- dimensional LSTM.
- By using Adam optimizer because it gives more accurate results.
- If the recognized word is not present in a dictionary, then search for the most similar word.

REFERENCES

1 R. Vijaya Kumar Reddy, U. Ravi Babu, "Handwritten Hindi Character Recognition using Deep Learning Techniques", Dept. of CSE, Acharya Nagarjuna University, Guntur, India, International Journal of Computer Sciences and Engineering, 2019, pp 1

2. Shubham Sanjay Mor, Shivam Solanki, Saransh Gupta, Sayam Dhingra, Monika Jain, Rahul Saxena, "Handwritten text Recognition: with Deep Learning and Android", Dept. of Computer Science and Engineering, Manipal University Jaipur, International Journal of Engineering and Advanced Technology (IJEAT), 2019, pp 819

3. Megha Agarwal, Shalika, Vinam Tomar, Priyanka Gupta, "Handwritten Character Recognition using Neural Network and Tensor Flow", Computer Science and Engineering, SRM IST Ghaziabad, India, International Journal of Innovative Technology and Exploring Engineering (IJITEE), 2019, pp 1445

4. Jemimah K, "Recognition of Handwritten Characters based on Deep Learning with TensorFlow", Research Scholar, School of Computer

Science and Engineering, Bharathidasan University, Trichy, India, International Research Journal of Engineering and Technology (IRJET), 2019, pp 1164-1165

5. Liang Yu, Binbin Li and Bin Jiao, "Research and Implementation of CNN Based on TensorFlow", School of Electric engineering, Shanghai Dianji University, Shanghai, China, IOP Conf. Series: Materials Science and Engineering 490, 2019, pp 1-3

6 Choudhary, A., Rishi, R., and Ahlawat, S., "Off-Line Handwritten Character Recognition using Features Extracted from Binarization Technique", AASRI Conference on Intelligent Systems and Control, 2013, pp. 306-312.

7. Sonu Varghese K, Ajay James, Dr.Saravanan Chandran , "A Novel Tri-Stage Recognition Scheme for Handwritten Malayalam Character Recognition", International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST - 2015),2016 pp-1333-1340.

8. Parshuram M. Kamble, Ravinda S. Hegadi, "Handwritten Marathi character recognition using R-HOG Feature", International Conference on Advanced Computing Technologies and Applications (ICACTA-2015),2015 pp- 266 – 274.

9 . Utkarsh gupta (2k16/swe/19), "cnn based approach for handwriting recognition using TensorFlow", Department of computer science and engineering delhi technological university (Formerly Delhi College of Engineering) Bawana Road , India, 2018, pp 12-14.