CHARACTER RECOGNITION FROM IMAGE USING TENSORFLOW AND CONVOLUTIONAL NEURAL NETWORKS

Keshav Gupta
Department of CSE
IMSEC, Ghaziabad, UP, India

Mohit Kumar
Department of CSE
IMSEC, Ghaziabad, UP, India

Narayani Sachdeva
Department of CSE
IMSEC, Ghaziabad, UP, India

Abstract— **Machine Learning (ML) can be defined as the pc algorithms through which a particular system can learn on it's own with help of experiences. It can be consider as the subset of Artificial Intelligence (AI). Machines acquiring knowledge of various algorithms to build a mathematical model which is based upon given (sample) data, considered as "education statistics", in order to perform decisions or predictions. Machine learning algorithms are widely utilized in various programs encompass e-mail spam and malware filtering, virtual assistant, fraud detection, online customer support [9].**

Many practices have been made to classify information from handwritten statistics, however the laptop alone hasn't been capable of classify statistics efficiently. Hence, the requirement of handwriting recognition system has come up [13,14]. In this undertaking we purpose at constructing a handwritten character popularity system using convolutional Neural Network. Which will be capable of understand the handwritten characters and provide the output using the training dataset i.e(IAM) dataset.

Keywords— **Recurrent Neural Network (RNN) layers, Final Connectionist Temporal Classification (CTC) layer, Convolutional NN (CNN) layers, Tensorfow.**

I. INTRODUCTION

Handwritten text, characters or typed files are easy to understand for human being as we have the ability to learn. The similar property may be provided to the machines additionally with the help of **Machine Learning**. The area which can overcome this problem is known as the **OCR** or it can be called as "**Optical Character Recognition**". It is the area of take a look at among diverse fields inclusive of recognizing of pattern. With the help of the OCR in banking area it become easier to recognize characters from the various banking documents. Many vital and sensitive files may be processed quicker without human intervention. Hence, the want of handwriting popularity has come up. By schooling the laptop to apprehend the set of handwritten characters we will classify the handwritten characters and retrieve facts from them [1,8,10].

II. REQUIREMENT OF TENSORFLOW

The Tensorflow required for the implementation of following modules:-

A. **SamplePreprocessor.py** –

This file is used to manage the images retrieved from the IAM dataset for the neural network.

B. **DataLoader.py** –

This file is able to read sample data from the dataset, and puts them into particular batches.

C. **Model.py** –

This file is helpful to create the models, load the models and save them, and also manages the 'Tensorflow' session which can provides an interface to the NN for inference and training purpose.

D. **Main.py** –

This file can combine all the previous modules information together in order to execute the program and provide output.

III. INTRODUCTION TO TENSORFLOW

Tensor Flow is an open-source software library for dataflow programming which can perform various tasks. It is a kind of Symbolic math library, which is useful for 'Machine Learning' applications like as 'Neural Networks (NN)'. It is useful in both purpose production and research at Google. Tensor Flow was developed by the Google Brain team for internal Google use.
It was released under the Apache 2.0 open source license on November 9, 2015. Computations in TensorFlow are represented in the form of stateful dataflow graphs [7]. The name Tensor Flow derives from the operations that such neural networks perform on Multi-dimensional data arrays. These arrays are called as "tensors" [15].

IV. MODEL OVERVIEW

In this model we are using a neural network for our task. It consists of different layers such as:- Convolutional NN (CNN) layers, Recurrent NN(RNN) layers and a final Connectionist Temporal Classfication (CTC) layer. Below figure show an overview of our HTR system [3].
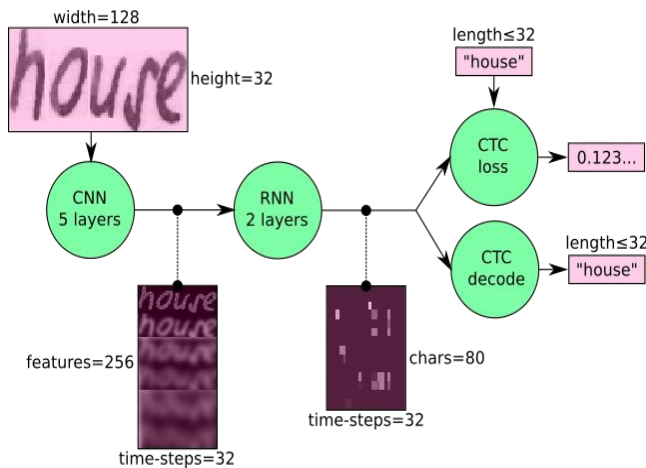

Fig. 1. **NN operations are represented with (green)** and thus the **information flow through the NN is represented with (pink***)*.

We can represent the Neural Network in a form of function (see Eq. 1) which provide more formal way and also able to maps a matrix(or an image) M of size "W×H" to a sequence of character(c1, c2, …) with a length between '0' and 'L'.
In this model you are able to see that the text is recognized on the character-level, and also with the help of this model you are able to recognize words or texts which are not contained in the training dataset.

$$NN: \underset{W \times H}{M} \to \underset{0 \le n \le L}{(C_1, C_2, …, C_n)}$$

(1)

The NN is represent as a "mathematical function" which can maps an image M to a character sequence (c1, c2, …).

V. DATA ACQUISITION

The "IAM Handwriting Database" includes styles of handwritten English textual content which can be used to teach and take a look at handwritten text recognizers and to perform writer identification and verification experiments. The database changed into first published in the ICDAR 1999.The database consist of styles of unconstrained handwritten text,

which had been scanned at a resolution of 300dpi [2] and stored in the form of PNG photographs with 256 gray levels.

A. *Characteristics* –
The IAM Handwriting Database 3.0 is structured as follows:-

1. 657 writers contributed samples of their handwriting

2. 1'539 pages of scanned text

3. 5'685 isolated and labeled sentences

4. 13'353 isolated and labeled text lines

5. 115'320 isolated and labeled words [6]

VI. OPERATIONS

A. **CNN –**
The input image is provided to the CNN layers. These layers are trained in order to extract important (relevant) features from the given input image. Each CNN layer comprises of 3 operation. Initial operation is known as, the 'Convolution operation', which applies a filter out kernel of length '5×5' within the first layers and the last 3 layers to the input are consist of '3×3' filter out kernel. Then, the 'RELU' which is a non-linear function is applied (RELU is kind of an activation function). Finally, CNN layer provides (a downsized version of the input and the picture regions) as the output to the RNN layers. [11].

B. **RNN –**

These layers use the relevant features extracted by the CNN layers. In this layer feature sequence carried out 256 features per time-step, and therefore RNN layers able to propagates further the important information through this collection to the CTC layer. Generated output collection is mapped to a matrix of size 32×80 [12].

C. **CTC –**

In this layer RNN output matrix and the ground truth textual content is provided to the CTC layer during the training process of neural network and it also computes the cost occur during the loss. Both the textual content "ground truth" and "diagnosed" can be supposed to be of max '32' characters long [5].

VII. DATA(INPUT)

Grey-level photo is provided as the data input whose dimension is '128×32'. IAM dataset does not consist pictures of such size, therefore we have to resize the picture until the image has dimension of '128×32'. Then we generate the photo into (white)target picture has width of '128' and height of '32'.
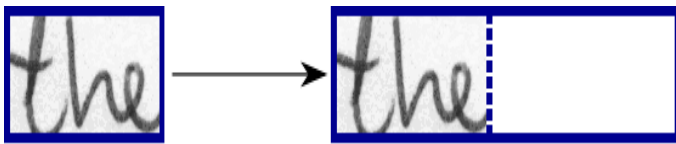
Fig. 2. On Left side: an arbitrary size input image is taken from the dataset. On Right side: input image is scaled in order to fit the target image of size '128×32' and the rest area of target image is consist of white color.

VIII. TRAINING PROCESS

For the training of neural network we require the mean of loss values of the batch elements.

A. After training –

"Character error rate":12.065659% and "Word accuracy rate":71.252174% we can easily see these rate values in the below given figure.

Fig. 3. Completion of training process
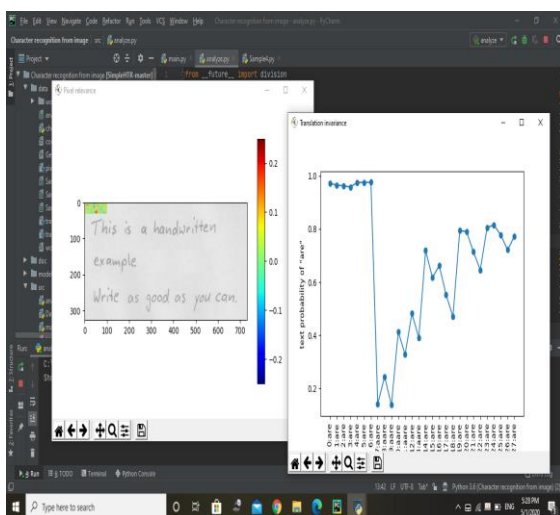
IX. OUTPUT

Fig. 4. Output can be provided in two forms, one is "**pixel relevance**" form and other one is "**translational invariance**" form.

X. FUTURE SCOPE

Accuracy of recognition is improved by adding some new features which are more effective than the previous one. These algorithms are further implemented upon the large datasets of handwritten characters or text. The proposed system is further improved in order to work upon the distorted text or broken text. The proposed system can also be improved in future so that it can recognize the vast number of languages.

XI. CONCLUSION

The CNN is used for the process of features extraction and provide an appropriate output to the user [4]. The neural community consists of several layers which assist in the education procedure. The error turned into reduced upto '**12.065659%**' and accuracy advanced upto '**71.252174%**'. Hence, our **Handwritten Text Recognition** (HTR) system is generated with the help of NN model.

XII. REFERENCES

[1] Singh Sameer and Mark Hewitt (2000)."Cursive Digit And Character Recognition on Cedar Database", Pattern Recognition. 15th international conference on. Vol.2. IEEE 2000.

[2] Brunessaux Sylvie, Giroux Patrick, Grilheres Bruno, Manta Mathieu (2014). "The Maurdor Project: Improving Automatic Processing of Digital Documents", 11th IAPR International Workshop on Document Analysis Systems, 2014.

[3] Salvador España-Boquera, Maria J. C. B., Jorge G. M. and Francisco Z. M. (2011). "Improving Offline Handwritten Text Recognition with Hybrid HMM/ANN Models", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 33, No. 4, April 2011.

[4] Yadav Usha, Verma Satya, Xaxa Deepak Kumar, Mahobiya Chandrakant (2017). "A deep learning based character recognition system from multimedia document", Innovations in Power and Advanced Computing Technologies (i-PACT), 2017.

[5] Bajaj Renna, Dey Lipika, and Chaudhury S. (2002)."Devnagari numeral recognition by combining decision of multiple connectionist classifiers",Sadhana, Vol.27, part. 1, (pp.-59-72) ,2002.

[6] Impedovo S., Facchini G., Mangini F.M. (2012). "A New Cursive Basic Word Database for Bank-Check Processing Systems",10th IAPR International Workshop on Document Analysis Systems, 2012.

[7] Varalakshmi P, Rachel J Julanta Leela. (2019)."Recognition Of Fish Categories Using Deep Learning Technique", 3rd International Conference on Computing and Communications Technologies (ICCCT), 2019.

[8] Aparna K. H., Subramanian Vidhya, Kasirajan M., Vijay Prakash G.,Chakravarthy V. S., Madhvanath Sriganesh (2004). "Online Handwriting Recognition for Tamil", IWFHR, 2004, Proceedings. Ninth International Workshop on Frontiers in Handwriting Recognition, Proceedings. Ninth International Workshop on Frontiers in Handwriting Recognition 2004, (pp. 438-443), doi:10.1109/IWFHR.2004

[9] Surya Nath R S M, Afseena (2015). "Handwritten Recognition- A review", International Journal of Scientific and Research Publications, Volume 5, Issue 3, March 2015 1 ISSN 2250-3153

[10] Pal Anita and Singh Davashankar (2011)."Handwritten English Character Recognition Using Neural Network", International Journal of Computer Science and Communication, (pp: 141-144), 2011.

[11] J.Pradeep1 , E.Srinivasan2 and S.Himavathi3 (2011). "DIAGONAL BASED FEATURE EXTRACTION FOR HANDWRITTEN ALPHABETS RECOGNITION SYSTEM USING NEURAL NETWORK" International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No 1, Feb 2011

[12] Theingi Htike and Yadana Thein (2013). "Handwritten Character Recognition Using Competitive Neural Trees " IACSIT International Journal of Engineering and Technology, Vol. 5, No. 3, June 2013

[13] Purohit Ayush #1,Singh Chauhan Shardul #2 (2016). "A Literature Survey on Handwritten Character Recognition" Ayush Purohit et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol.(1),(pp:15),2016.

[14] Sonkusare Manoj and Sahu Narendra(2016). "A SURVEY ON HANDWRITTEN CHARACTER RECOGNITION (HCR) TECHNIQUES FOR ENGLISH ALPHABETS" Advances in Vision Computing: An International Journal (AVC) Vol.3, No.1, March 2016

[15] Abadi Martín, Barham Paul, Chen Jianmin, Chen Zhifeng, Davis Andy, Dean Jeffrey, Moore Sherry, G. Murray Derek, Steiner Benoit, Tucker Paul, Vasudevan Vijay, Warden Pete, Wicke Martin, Yu Yuan, and Zheng Xiaoqiang, (2016), "TensorFlow: A system for large-scale machine learning", 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16). ISBN 978-1-931971-33-1, (pp. 265-266).