Comparing Deep Neural Network Models for Handwriting Recognition

Ritesh Jeeban
School of Innovative Technologies and Engineering
University of Technology
La Tour Koenig, Pointe-aux-Sables, Republic of Mauritius
ritesh.jeeban@gmail.com

Geerish Suddul
School of Innovative Technologies and Engineering
University of Technology
La Tour Koenig, Pointe-aux-Sables, Republic of Mauritius
g.suddul@umail.utm.ac.mu

***Abstract* — Handwriting recognition is a technique used to interpret intelligible handwritten input and convert them into digital text using Machine Learning tools. This research paper provides a comparison of the application of CRNN and RNN for handwriting recognition, using a dataset containing about 400,000 handwritten names. Our experiments demonstrate that the CRNN model produces the highest accuracy compared to RNN model.**

Keywords — Handwriting recognition, CRNN, RNN

I. INTRODUCTION

Handwriting Recognition is an interesting and demanding research based in Artificial Intelligence, computer vision and pattern recognition [6]. A computer performing handwriting recognition is defined as a system capable of acquiring and detecting characters or words in a paper documents, images and others converting them into machine encoded form. In order to perform these tasks, machine learning algorithm has to be implemented for more advanced intelligent Handwriting recognition. It has contributed immensely to the advancement of automation process in many fields and made improvement to the interface between man and machine in numerous applications. During the past years, main focus was on the implementation of new techniques and methods to reduce the processing time while ensuring higher recognition accuracy.

II. OBJECTIVES

In an attempt to find an accurate machine learning model for Handwriting Recognition, the main objective of this research work is to compare the accuracy of a hybrid Convolutional Recurrent Neural Network (CRNN) model against the Recurrent Neural Network (RNN) model.

III. METHODOLOGY

Model A: CRNN

The CRNN model which is a hybrid model is created using the tensor flow and Keras library of python. The model consists of different layers. Layer 1 where the input is fed and is reshaped into (256,64,1) by the reshape layer. Layer 2,3 and 4 are Convolutional Layer. Layer 2 creates 32 feature maps as output using 32 different filters. Layer 3 is another Convolutional layer having 64 feature maps while Layer 4 uses 128 features map. For the 3 layers of CNN, MaxPooling is used to reducing the spatial size of the image i.e. (2,2), (2,2) and (2,1) respectively, uses a filter of size (3,3) and uses ReLU as the activation function. Layer 5 is RESHAPE Layer of size (64, 1024) and Layer 6 is the DENSE layer using ReLU for the activation. Layer 7 and 8 is the RNN where LSTM will output 256 parameters. Layer 9 is the DENSE layout. Layer 10 is where the neurons are activated using SOFTMAX as the activation function. Layer 11 is the last layer that outputs the results.

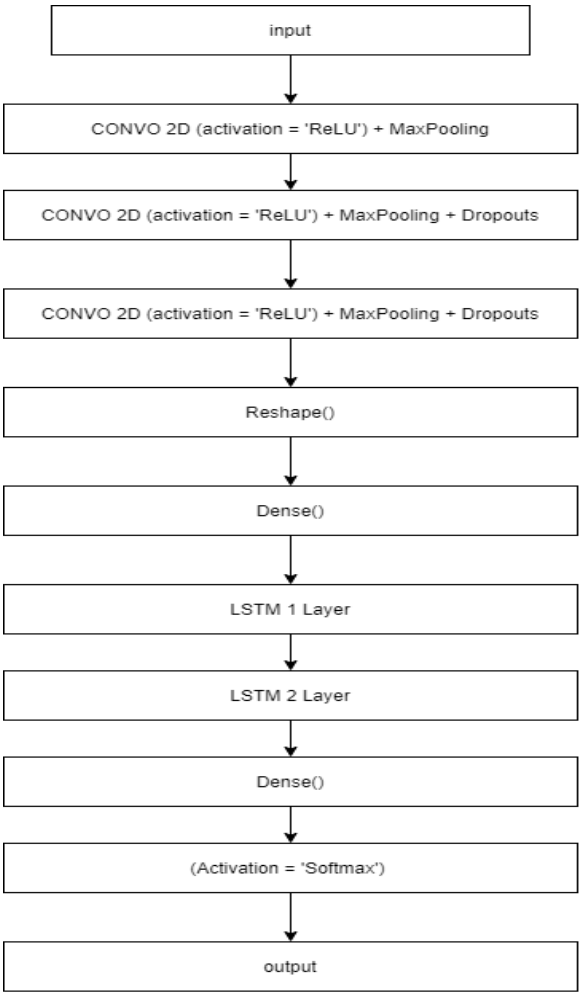

Fig. 1: shows the structure of the first model.

Model B: RNN

The second model is an RNN that is recurrent neural network, and it consists of 6 layers only. Layer 1 is where the input is processed and are shaped as (256, 64). Layer 2 and 3 are LSTM Layer where both will output 256 parameters. Layer 4 is DENSE layer that will feeds outputs from the previous layer to all its neurons. Layer 5 uses SOFTMAX for the activation. Layer 6 is the final layer that will produce the outputs.

ISSN 2756-9160 / November 2020.
International Conference on Advances in Computing and Technology (ICACT–2020) Proceedings

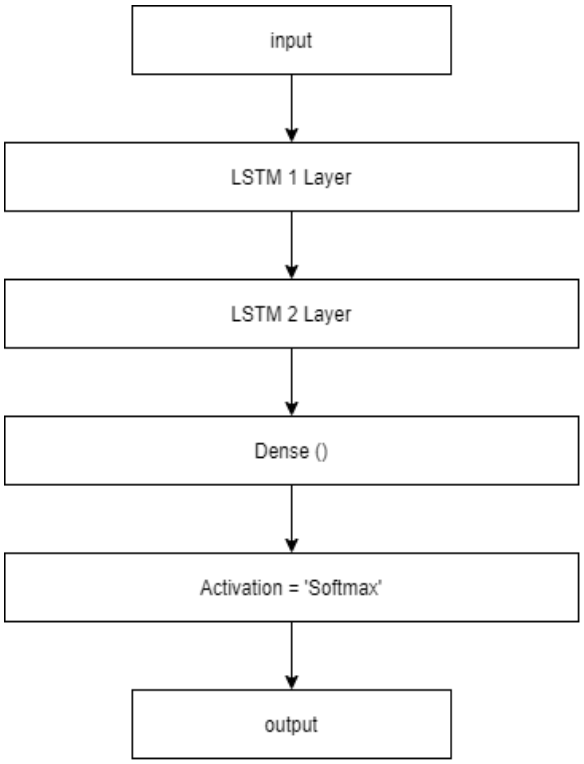

Fig. 2: shows the structure of the second model.

IV. RESULTS AND DISCUSSION

All the accuracy of the models (Deep CNN, CNN and CMAM, TPP-PHOCNet and AFDM) found in previous research papers was taken into considerations. The two model that I have created, tested and validated are included in the table with their accuracy.

For the model A and B, the accuracy was calculated by having an array of alphabets. CTC loss function is applied. The Machine learning model identifies the handwritten text in the image and predicts each character using the array. The percentage of number of correct character predicted is also calculated. The characters are concatenated and it is compared with the identity of the image. An algorithm is applied that is number of correct word * 100 and it is divisible by the total number of the validation size. Hence the total accuracy is achieved.

The accuracy for all model ranges from 32.89 % to 94.31 %. Model A: CRNN has an accuracy of 75.27 % and Model B: RNN has an accuracy of 32.89 %. Note that for the other models, they were trained using different datasets. Two of them are IAM and RIMES. We have train the 2 model using the same dataset.

Table 1: The accuracy of the 7 different model

	Accuracy Obtained
Deep CNN [1]	80.0
CNN [2]	87.1
CMAM [3]	74.45
TPP-PHOCNet[10]	94.31
AFDM[10]	92.94
Model A: CRNN	75.27
Model B: RNN	32.89

V. CONCLUSION & LIMITATIONS

This research paper provided a comparison of the application of CRNN and RNN for handwriting recognition. Based on the results obtained, it can be concluded that TPP-PHOCNet has the highest accuracy of 94.31 % compared to other models. The difference of accuracy for all the models varied because it depends on the number of dataset trained, tested and on different PC with different specifications. Hence it can be concluded that TPP-PHOCNet model is the best model compared to the two models (Model A: CRNN and Model B: RNN) that we have created.

Due to limited resources, only 30000 values were trained, and 3000 values were tested for both the model of CRNN and RNN. With better resources such more RAM and GPU, more values will be trained and tested for better accuracy and the training time also would have been faster.

REFERENCES

[1] M Fabian Stark, Caner Hazırba¸s, Rudolph Triebel, and Daniel Cremers, CAPTCHA Recognition with Active Deep Learning, Conference: German Conference on Pattern Recognition Workshop, 2015

[2] Shubham Sanjay Mor, Shivam Solanki, Saransh Gupta and al, HANDWRITTEN TEXT RECOGNITION: with Deep Learning and Android, International Journal of Engineering and Advanced Technology (IJEAT), 2019

[3] Duc Nguyen, Nhan Tran and Hung Le, Improving Long Handwritten Text Line Recognition with Convolutional Multi-Way Associative Memory, arXiv:1911.01577, 2019

[4] Tensorflow, Recurrent Neural Networks (RNN) with Keras, https://www.tensorflow.org/guide/keras/rnn

[5] Built-In, Artificial Intelligence. What is Artificial Intelligence? How Does AI Work?, https://builtin.com/artificial-intelligence

[6] Harald Scheidl, Build a Handwritten Text Recognition System using TensorFlow, https://towardsdatascience.com/build-a-handwritten-text-recognition-system-using-tensorflow-2326a3487cd5.