International Journal of Future Generation Communication and Networking
Vol. 13, No. 2s, (2020), pp. 564–569

Handwritten Cursive English Text Recognition Using Deep CNN-RNN based CT

Sagar Vyavahare[1],Manisha sagade[2],Karansingh Hajari[3],Shraddha Surwase[4]

[1, 3, 4] *UG student, Dept. of E & TC Engg, SKNCOE, SPPU, Pune.*
[2]*Assistant Professor, Dept. of E & TC Engg , SKNCOE, SPPU, Pune.*
[1]vyavaharesagar303@gmail.com
[2] Manishajs@gmail.com
[3]karanhajari7@gmail.com
[4]Surwaseshraddha1998@gmail.com

Abstract

The Cursive Handwriting Recognition is employed in applications like reading aid for the blind, recognition of historical documents, processing of bank cheques, pattern recognition then on which defines ability of a system to spot characters. These systems transcribes human written texts into digital texts. Back propagation algorithm, binary segmentation algorithm, Optical Character Recognition systems are commonly used for translating any handwriting into plain text format. the main drawback of those techniques is that their efficiency is a smaller amount. Hence neural networks are widely adopted for classification and performance approximation tasks. during this proposed technique, offline HWRs is completed using CNN-RNN and Tensorflow - Keras Functional API .The proposed technique used segmentation free approach to create a system ready to">which can be able to recognize the handwritten characters with highest accuracy. The proposed DCNN-CTC are able to do character error rate of 4.07%, yielding a relative CER reduction of 30.8%. therefore the main motive of the project is that we are getting to convert the unrecognizable cursive handwritten texts into an easier text.
Keywords*— DCNN, CER, WER, Tensorflow-Keras-Functional API.*

I INTRODUCTION

This project, ‘Handwritten Character Recognition’ may be a software algorithm project to acknowledge any handwritten character efficiently on a computer with input is either an image in which handwritten characters are written. Character recognition using CNN-RNN, is electronic translation of images of handwritten into a machine-editable text. It's a field of research in pattern recognition, AI , and machine Learning. CNN-RNN using handwritten character recognition may be a scheme which enables a computer to find out , understand, improvise, and interpret the written or printed character in their own language. RNN using handwritten Character Recognition uses the image processing technique to spot any character which is handwritten. tons of labor has been doing during this field.

II Literature Review

Image-based text sequence recognition has been a wide research topic in computer vision. Modeling Handwritten Character Recognition as a sequence to sequence problem using RNN and CTC has been a widely used approach. With the recognition and robustness of CNN, most of the works uses convolutional layers as their feature extractors.

[1] P. Voigtlaender, P. Doetsch, and H. Ney, Multidimensional long short-term memory recurrent neural networks achieve impressive result's for handwriting recognition, in ICFHR, 2016.

This paper proposes an efficient GPU based implementation which greatly reduces training times by processing the input in a diagonal-wise fashion. They used this presentation to explore stronger and very broad level architectures than previously used for handwriting recognition and shown that the depth plays an important role. They have outperformed the state of the art results on two databases with a deep multidimensional network.

[2] P. Doetsch, M. Kozielski, and H. Ney, “Fast and sturdy coaching of repeated neural networks for offline handwriting recognition,” in International Conference on Frontier's in Handwriting Recognition, Sep. 2014, pp. 279–284.

The project consists of a modified topology for LSTM recurrent neural networks that controls the shape of the squashing functions in gating units. They propose an efficient training framework which is depend upon the mini batch training on sequence level platform. The framework is estimated on publicly available databases containing English and French handwriting by using a GPU based execution.

[3] B. Shi, X. Bai, and C. Yao, An end-to-end trainable neural network for image-based sequence recognition and it's application's to scene text recognition, PAMI, 2016.

This paper informing about recognition of the problem of real time scene texts , which is among the most important and difficult tasks in image based character recognition. A neural network architecture, which is used for feature extraction, character modeling and transcription into a incorporated framework, is proposed. The experiments on standard level, including the IIIT-5K, Street View Text and ICDAR datasets, demonstrate the highness of the proposed algorithm over the prior steps.

III Methodology

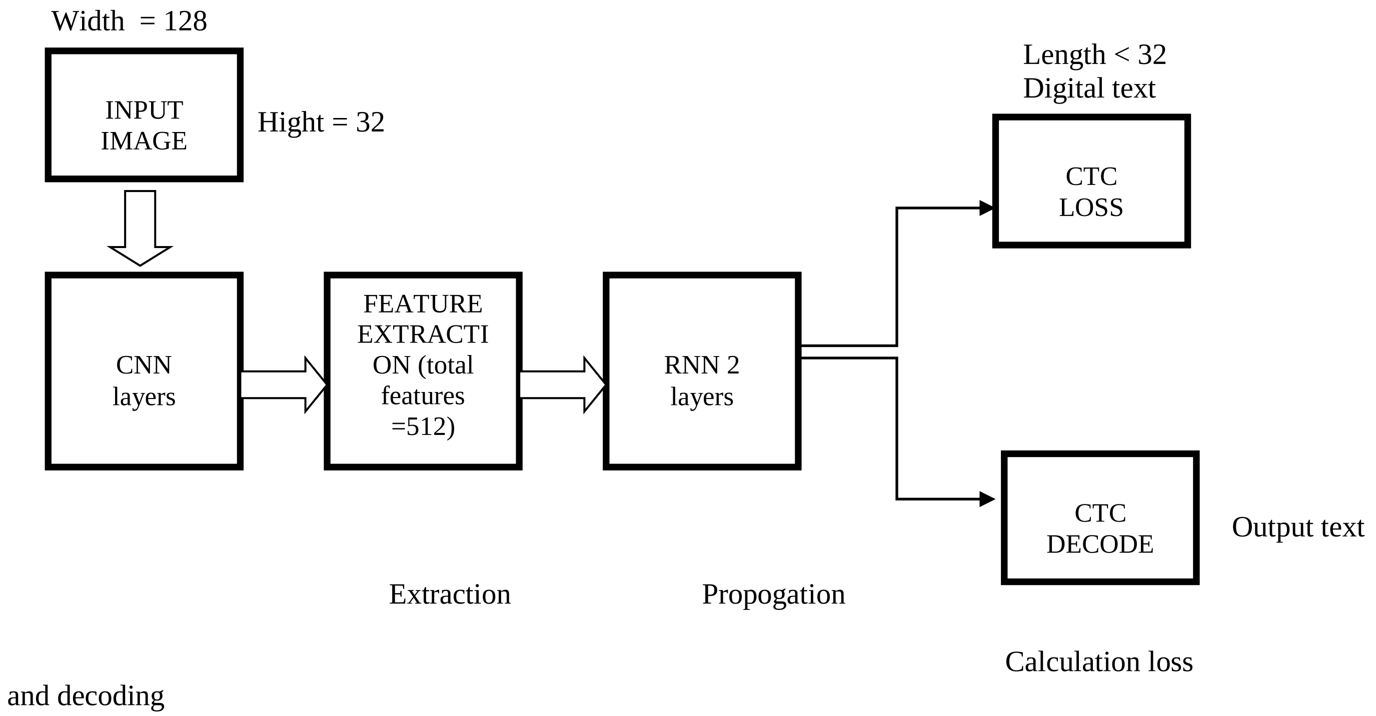

Fig. 1 Block diagram of proposed system

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	(None, 32, 128, 1)	0	
conv2d_1 (Conv2D)	(None, 32, 128, 64)	640	input_1[0][0]
max_pooling2d_1 (MaxPooling2D)	(None, 16, 64, 64)	0	conv2d_1[0][0]
conv2d_2 (Conv2D)	(None, 16, 64, 128)	73856	max_pooling2d_1[0][0]
max_pooling2d_2 (MaxPooling2D)	(None, 8, 32, 128)	0	conv2d_2[0][0]
conv2d_3 (Conv2D)	(None, 8, 32, 256)	295168	max_pooling2d_2[0][0]
conv2d_4 (Conv2D)	(None, 8, 32, 256)	590080	conv2d_3[0][0]
max_pooling2d_3 (MaxPooling2D)	(None, 4, 32, 256)	0	conv2d_4[0][0]
conv2d_5 (Conv2D)	(None, 4, 32, 512)	1180160	max_pooling2d_3[0][0]
batch_normalization_1 (BatchNor	(None, 4, 32, 512)	2048	conv2d_5[0][0]
conv2d_6 (Conv2D)	(None, 4, 32, 512)	2359808	batch_normalization_1[0][0]
batch_normalization_2 (BatchNor	(None, 4, 32, 512)	2048	conv2d_6[0][0]
max_pooling2d_4 (MaxPooling2D)	(None, 2, 32, 512)	0	batch_normalization_2[0][0]
conv2d_7 (Conv2D)	(None, 1, 31, 512)	1049088	max_pooling2d_4[0][0]
lambda_1 (Lambda)	(None, 31, 512)	0	conv2d_7[0][0]
bidirectional_1 (Bidirectional)	(None, 31, 512)	1574912	lambda_1[0][0]
bidirectional_2 (Bidirectional)	(None, 31, 512)	1574912	bidirectional_1[0][0]
dense_1 (Dense)	(None, 31, 79)	40527	bidirectional_2[0][0]
the_labels (InputLayer)	(None, 16)	0	
input_length (InputLayer)	(None, 1)	0	
label_length (InputLayer)	(None, 1)	0	
ctc (Lambda)	(None, 1)	0	dense_1[0][0] the_labels[0][0] input_length[0][0] label_length[0][0]

Total params: 8,743,247
Trainable params: 8,741,199

Fig.2 **–** Proposed CNN-RNN-CTC Achitecture

Dataset:

We used IAM handwritten word dataset. the standard partition for training, testing, and validation provided along with the dataset.

A. *Pre-Processing*:

The pre-processing is the series of operation performed on scanned input images. It extracts the required pattern from the background. The filtering and smoothing of noise is done here. Then we applied contrast enhancement techniques from open-cv image processing library. With the help of this library we applied some thresholding techniques. finally we resized the image and feed it to the model.

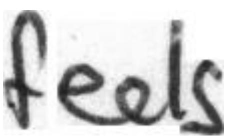

Fig.3 – Sample Image From IAM Database

ISSN: 2233-7857 IJFGCN

B. *CNN*:

Convolution Layer:

Convolution is that the first layer to extract features from an input image. Convolution preserves the connection between pixels by learning image features using small squares of input file . It is a mathematical process that takes two inputs like image matrix and a filter or kernel

Max Pooling Layer:

It is a sample-based discretization process. It selects contributed value. We used it to downsize the image.

Normalization Layer :

Data normalization is done by subtracting the mean from each pixel and then dividing the result by the standard deviation.

Relu Layer:

Rectified Linear Unit (RELU). The rectified linear activation function may be a piecewise linear function which will output the input directly if is positive, otherwise, it'll output zero.

Output Layer :

Where the input from the opposite layers is flattened and sent so because the transform the output into the amount of classes as desired by the network.

C. RNN:

the feature sequence contains 512 choices per time-step, the RNN propagates relevant information through this sequence. the popular Long memory implementation of RNN's is used as a result's of it's able to propagate information through longer distances and provides many sturdy training-characteristics. the RNN's output sequence is mapped in to a matrix of size 32×80. The IAM dataset consists of seventy 9 utterly completely different characters, anyone additional character is needed for the authority operation (CTC blank label), so there area unit eighty entries for each time-steps.

D. CTC:

Conventional methods of handwriting recognition require to segment the input word image into sub-words and characters. In these methods, the segmentation step plays a vital role and every segment is considered as an observation which probability of need to be calculated. But, we aimed to solve the problem without segmentation. In the powerful connectionist temporal classification (CTC) method is proposed for sequence labeling without segmenting the input. In practice, CTC is a softmax layer following an RNN. Its outputs are the probabilities corresponding to all the label alignments to the input sequence with length T for all time steps.

In the CTC method, in addition to the alphabet labels (A), an additional label is defined as blank (denoted by "–"), that shows there is no label at a specific time in the output sequence. The CTC method has demonstrated its outperformance in the sequence labeling tasks such as speech recognition and online and offline handwriting recognition.

IV EXPERIMENTATION RESULT

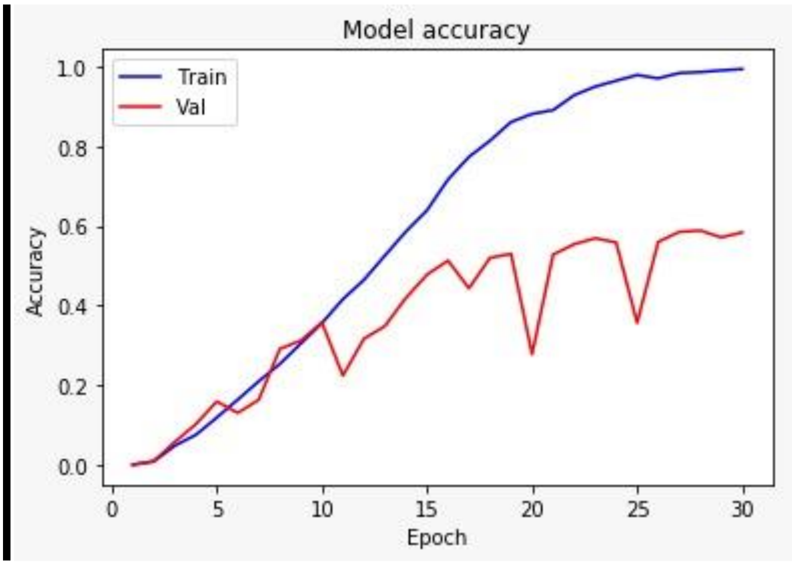

Fig.4- Traning accuracy and Validation accuracy

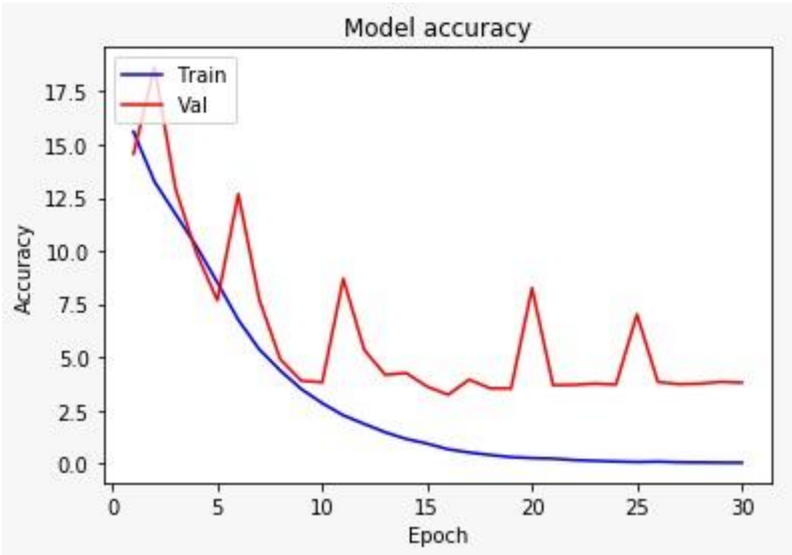

Fig.5 – Training loss and Validation losss

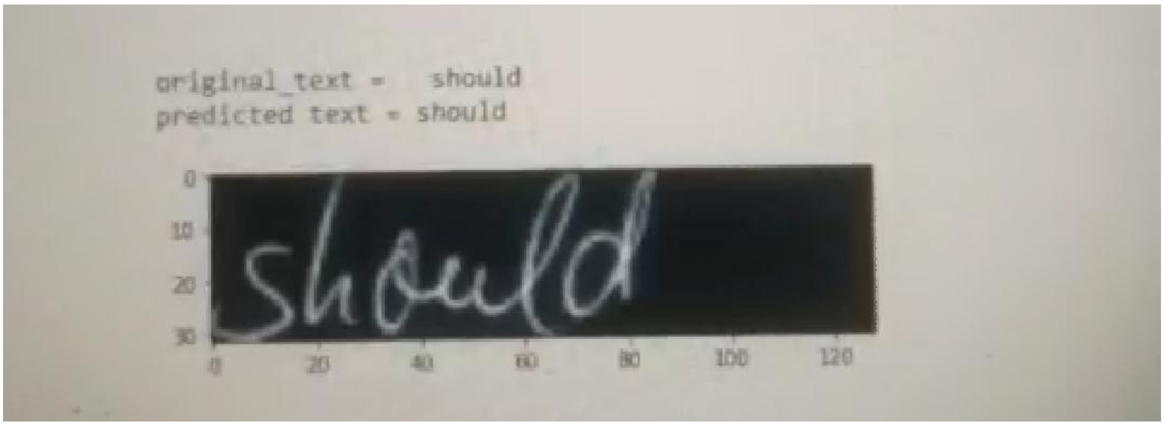

Fig.6 – predicted answer

ADVANTAGES

- It automatically detects the important features without any human supervision.
- It is computationally efficient.
- Better accuracy

V CONCLUSION

In this work, we presented effective ways to train a CNN-RNN architecture using synthetic data and domain specificimage normalization and augmentation. We also showed the individual contributions of each of these modules for improving the recognition rates at word levels. In future, we would also like to integrate line-level recognition model with language model based decoding so as to further enhance our recognition performance.

REFERNCES

[1] P. Voigtlaender, P. Doetsch, and H. Ney, Multidimensional long short-term memory recurrent neural networks achieve impressive result's for handwriting recognition, in ICFHR, 2016.
[2] P. Doetsch, M. Kozielski, and H. Ney, “Fast and sturdy coaching of repeated neural networks for offline handwriting recognition,” in International Conference on Frontier's in Handwriting Recognition, Sep. 2014, pp. 279–284.
[3] B. Shi, X. Bai, and C. Yao, An end-to-end trainable neural network for image-based sequence recognition and it's application's to scene text recognition, PAMI, 2016.
[4] Zi-Rui Wang, Jin Sui hu, “Deep Convolutional Neural Network primarily based Hidden Markov Model for Offline written Chinese Text Recognition”, 2017 fourth IAPR Asian Conference on Pattern Recognition.
[5] Durjoy fractional monetary unit Maitra, Ujjwal Bhattacharya, and Swapan K. Parui, “CNN primarily based Common Approach to written Character Recognition of Multiple Scripts." 2015 thirteenth International Conference on Document Analysis and Recognition (ICDAR).
[6] Chirag Patel, Ripal Patel, Palak Patel, “Handwritten Character Recognition exploitation Neural Networks”, International Journal of Scientific & Engineering analysis Volume two, Issue 5, May-2011.
[7] Shruti Shrikant Kharkar, Hoshank Jitendra African nation, Sakshi Satish Gadhari, Sarang Sunil Shrikhande, Prof. D.K. Chitre, “Handwriting Recognition exploitation Neural Network”, © 2017 IJEDR | Volume five, Issue four | ISSN: 2321-9939.
[8] Jun Du, Zi-Rui Wang, and Jian-Fang Zhai,” Deep Neural Network-based Hidden Markov Model for Offline written Chinese Text Recognition”, 2016 twenty third International Conference on Pattern Recognition (ICPR) urban center Atma
[9] Prakash Singh, Ravindra Nath, Santosh Kumar, “Handwritten English Character Recognition exploitation HMM, Baum-Welch and Genetic Algorithm”, Ravindra Nath et al, / (IJCSIT) International Journal of technology and knowledge Technologies, Vol.
[10] Yves, "Segmentation and committal to writing of Arabic written words," in Proceedings of thirteenth International Conference on Pattern Recognition, 1996.
[11] Ghassemlooy, Automatic segmentation for Arabic character's recognition in handwriting.

ISSN: 2233-7857 IJFGCN