LSTM based Odia Handwritten Numeral Recognition

Abhishek Das, Gyana Ranjan Patra and Mihir Narayan Mohanty

***Abstract*—Recent works on character recognition involves in different applications like security, biometrics etc. In most of the works Convolutional neural networks (CNN) are used for image recognition with large dataset where convolutional filters with pooling layers are used for feature extraction. In this work we have utilized the concept of Long Short Term Memory (LSTM) for recognizing the Odia handwritten numerals. The loss in recognizing the digits is evaluated using categorical_crossentropy loss function and for optimization we have applied Adam optimization to minimize the error. The network is trained with supervised learning method. The accuracy found in this work is 97.93% which is a remarkable achievement in Odia numeral recognition.**

***Index Terms*—Character recognition, Deep learning, CNN, LSTM, Cross entropy, Adam Optimization, supervised learning**

I. Introduction

IN this digital world still some documents are there which are manually filled for safety purpose. One example of such document is bank check. Similarly different forms are being filled by hand. Handwritten reports are being collected for office work. There are many ancient handwritten documents which are to be scanned and restored in typed form so that it will be easy to read. Handwritten character and numeral recognition is helpful in such aspect. Deep learning is a part of machine learning which is developed on the basis of artificial neural networks. In artificial neural networks different nodes are used which are trained by different algorithms like a human brain but there are many differences among the artificial neural networks and biological brain. In deep learning large number of layers are used which are used for better accuracy value. Deep learning is used in different fields like automatic speech recognition, natural language processing, medical image analysis, natural language processing, visual art processing, image processing and in many other areas of research. In this paper we have used Long Short Term Memory network which is a type of deep learning in recognizing Odia numerals.

Odia (Previously known as Oriya) is the official language of the Odisha which is a state of India. This language is spoken in almost all districts of Odisha as well as in some places of nearby states. Different documents like old manuscripts, official files, are written in Odia. So, Odia character recognition is to be considered for different computer applications.

The main attributes of the paper are arranged in the following manner. The Section I provide the information about the deep learning in brief and its application in image processing for character recognition. The Section II provides the related works done in the field of character recognition as well as numeral recognition. The Section III provides the basic idea about the Odia scripts. The Section IV provides the basics of Long Short Term Memory network. The Section V is providing information regarding the proposed method in details. In Section VI the results obtained from the proposed method is given. In the Section VII the conclusion is drawn from this research work.

II. Literature Survey

Image classification is a part of optical character recognition if the image consists of different characters in it. Convolutional neural network (CNN) is used for image classification [1] and it shows better performance. In this work the authors have used MNIST dataset for numeral classification and CIFAR-10 dataset for classification of airplane, automobile, bird, cat, deer, dog, frog, horse, ship, truck. Medicine prescription recognition [2] is another form of hand written character recognition. In this work Peilun Wu *et. al.* have integrated three different classifiers. The three classifiers are CNN, PCA and KNN. A voting method is used to select the best features extractor from the different filters.

For Chinese clinical data recognition from the report Jiahui Qiu *et. al.* have proposed a method using convolutional neural network with conditional Random field [3]. Different researches have been done on Bangla handwritten digit recognition. In [4] the authors have used an improved version of CNN architecture which consists of 7 layers including convolutional layer, max pooling layer and fully connected layer. Handwritten Chinese character recognition and based on it different fonts are printed in [5]. For the recognition part the authors have used convolutional neural network (CNN). Now a day mobile phones are having online handwritten character input capability. To recognize the character different algorithms are being used. Authors have used CNN model for online handwritten recognition [6]. Birhanu Hailu Belay et. al. have utilized the beauty of CNN in recognizing Amharic characters[7]. The dataset used for this work is having 80000 numbers of Amharic characters. The authors have generated the Amharic text lines, characters using OCRopus.

Abhishek Das is with the Siksha 'O' Anusandhan University, Bhubaneswar, Odisha, (e-mail: abhishekdas225@gmail.com).

Gyana Ranajn Patra is with the Siksha 'O' Anusandhan University, Bhubaneswar, Odisha, (e-mail: gyana.patra@gmail.com).

*Mihir Narayan Mohanty is with the Siksha 'O' Anusandhan University, Bhubaneswar, Odisha, (e-mail: mihirmohanty@soa.ac.in).

978-1-7281-4988-2/20/$31.00 ©2020 IEEE

The application of CNN is found for number plate recognition [8-12] which is a part of intelligent transportation system. In the field of clinical named entity recognition [13] Guohua Wu *et. al.* have used bidirectional LSTM network along with conditional random field to recognize the words related to health issues i.e. body parts, different diseases, symptoms etc. For Nastaliq text recognition [14] authors have used multi-dimensional long short term memory. By using MD-LSTM the authors achieved an accuracy of 98%. A pixel based MDLSTM is used for this purpose. For in-air Chinese handwritten recognition [15]

Haiqing Ren et. al. have used Recurrent neural network as we know that sequence data can be processed easily using RNN. The X and Y coordinates of the dot locations are used as sequential data for RNN network. Two hidden layer based LSTM network [16] is used by Nogra *et. al.* for Baybayin handwritten recognition. Kalyan S Dash *et. al.* have proposed a hybrid method of feature extraction [17] in which they have used Kirsch gradient operator and properties based on curvature of handwritten numerals, followed by a feature dimension reduction using Principal Component Analysis (PCA). Om Prakash Jena *et. al.* have implemented linear discriminant Analysis [18] for Odia Numeral Recognition. Sanjibani Sudha Pattanayak have used Support vector machine to recognize Odia hand written recognition [19]. The characters are classified which gives accuracy up to 85%.

Most of the recent works for character recognition are developed using Convolutional Neural Network. No work is found using long short term memory for Odia numeral recognition. We have developed a network using LSTM for Odia numeral recognition.

III. About Odia Numerals

The basic Odia numeral contains 10 numbers as in other languages i.e. from 0 to 9 which is shown in Fig. 1.

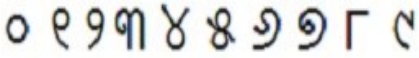

Fig. 1. Showing All Odia Numbers

But the structure for some numbers from 1 to 9 in Odia language are different from other language except 0.The numbers in Odia which are having similarity with other languages is shown in Fig. 2.

Language	Similar numbers
Odia	୪
Gurumukhi	੪
Hindi	४

Fig. 2. Showing similarity among Odia, Hindi and Gurumukhi digit four

IV. Long Short Term Memory

Recurrent neural networks (RNN) are a type of artificial network designed to recognize patterns in sequence of data, such as text, genomes, handwritten data, etc. RNNs use back propagation algorithm to train the network for each timestamp. It is commonly known as Back Propagation Through Time (BTT). RNN can model the sequence of data so that each sample can be assumed to be dependent on previous ones. But simple RNNs are having some disadvantages like vanishing Gradient and exploding gradient because of long term dependencies. Long Short Term Memory (LSTM) networks are capable of learning Long-term dependencies. The basic structure of LSTM cell is shown in Fig. 3.

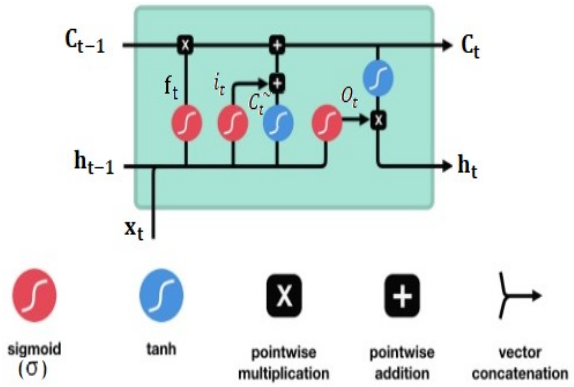

Fig. 3. Basic LSTM cell

The relation among input and output to the sigmoid activation function is given by,

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

Where, W_f = Weight

h_{t-1} = Output from the previous time stamp

x_t = New input

b_f = Bias

The relation among the input and output of the second sigmoid and tanh activation functions are given by,

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

$$C_t^{\sim} = \tanh(W_c[h_{t-1}, x_t] + b_c) \; C_t = f_t * C_{t-1} + i_t * C_t^{\sim}$$

Where, C_t = Cell state

W_i, W_c = Weights

b_i, b_c = Bias

The relation among the input and output of last sigmoid and tanh activation functions are given by,

$$O_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$
$$h_t = O_t * \tanh(C_t)$$

Where O_t = output of third sigmoid function

W_o = Weight b_o = bias of the third sigmoid function

V. Proposed Methodology

To recognize the Odia handwritten numerals we have taken one input layer, four LSTM layers and one fully connected layer at the end with 10 nodes. The LSTM layers are activated by using sigmoid activation function and the final fully connected layer is having Softmax function to classify the recognized digit to any one of the 10 nodes. Loss during the training is calculated by categorical_crossentropy function and the loss is minimized by using Adam optimizer.

A. Dataset

The IIT-Bhubaneswar dataset of handwritten Odia numerals is used for the purpose of training and testing. A Sample of the dataset is shown in Fig. 4.

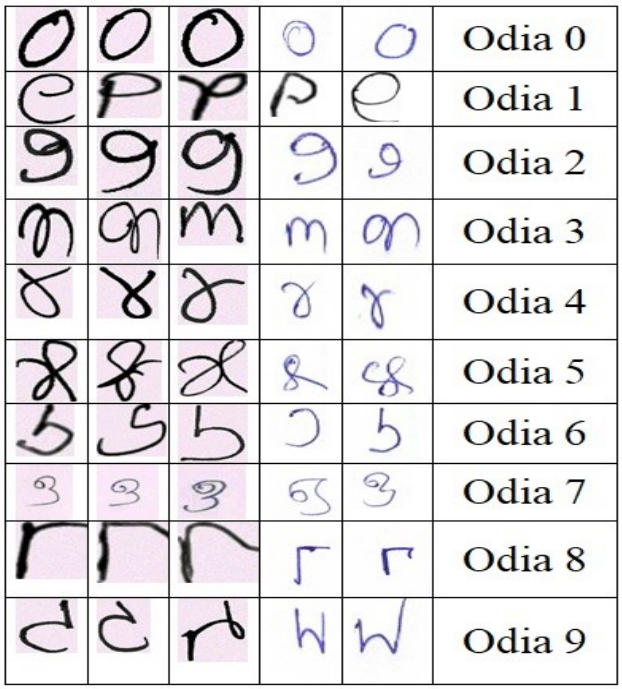

Fig. 4. Sample of IIT Bhubaneswar dataset for Odia handwritten Numerals

The dataset consists of 5166 numbers of samples which are of Jpg and Tiff formats. All the images are of different size in gray scale and RGB. The whole dataset is converted to binary form to reduce the number of channels from three to one.

B. Proposed LSTM network

The image data of handwritten numerals are converted to timestamp dependent input to the network so that the LSTM layers will process them and extract features depending on time. The time dependency in the proposed task makes the ability of image recognition in LSTM like recurrent neural network. In case of Convolutional neural networks it is not required to convert the input images to time dependent pixel values as the CNN directly extract the features from images using different filter windows. The proposed LSTM network for the Odia handwritten numeral recognition is shown in Fig. 5. In this figure all the layers, activation function, loss function and optimizer are clearly stacked which shows the direction of data flow to perform the recognition task.

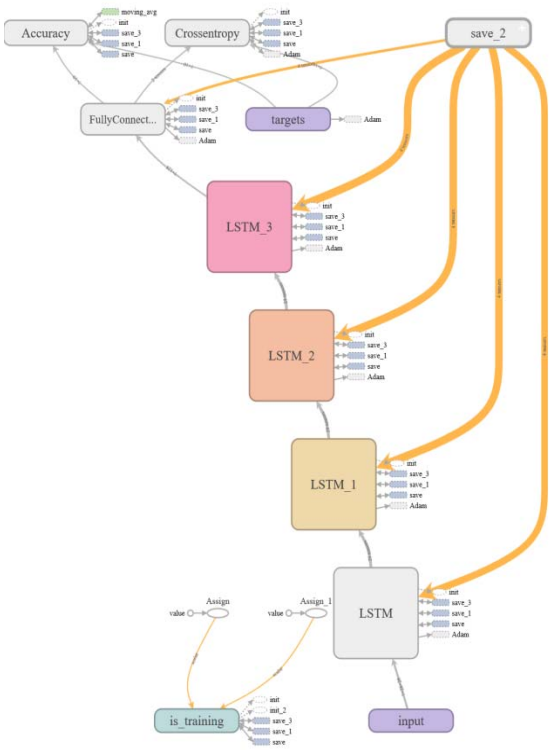

Fig. 5. Proposed LSTM network

VI. Simulation and Results

The proposed method is analyzed practically using Python 3.7 in Jupyter Notebook. The software was loaded in the operating system Microsoft Windows 10, i3 Generation with 8Gb RAM capacity enabled Laptop. After 20 epochs most of the Odia numerals are recognized accurately. A sample of the test result is shown in Fig. 6.

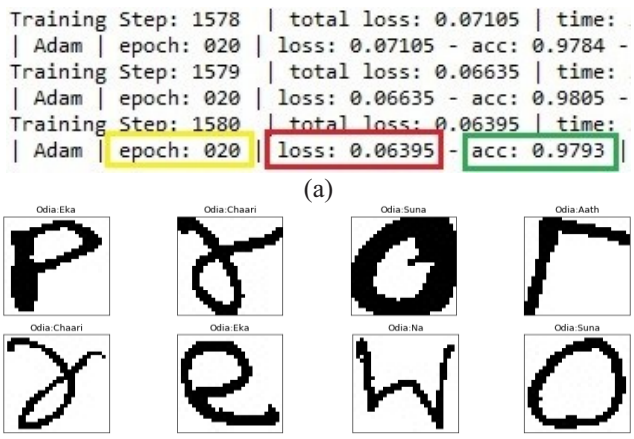

Fig. 6. (a) Screenshot of final result after 20 epochs
(b) Recognition of Handwritten Odia numbers using proposed network

The training accuracy after 20 epochs is 97.93% which is shown in Fig. 7. The Graph shown in the figure is between the accuracy in Y-axis and the number of steps in X-axis. As the number of steps increases it trains the network with more number of images and the accuracy also increases.

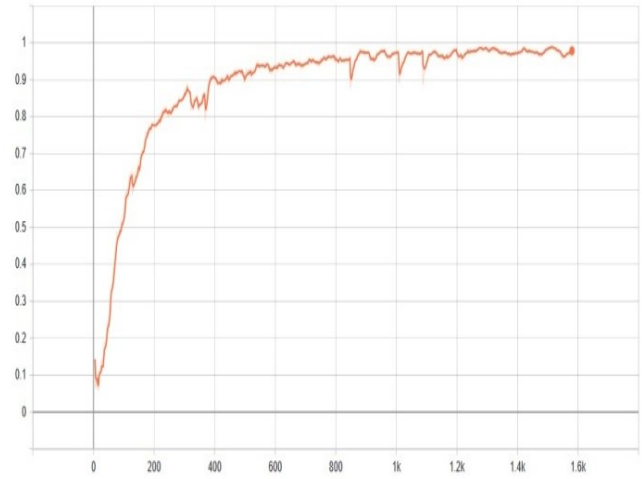

Fig. 7 Accuracy Plot

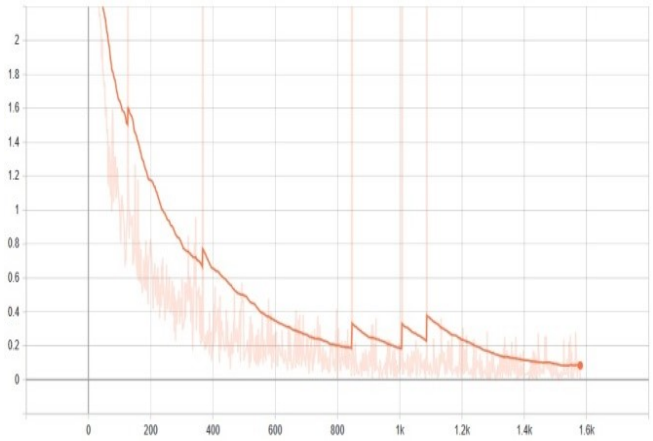

Fig. 8. Loss Plot

The relation between loss in recognition and number of steps is shown in Fig. 8. From this figure it is found that as the training increases the error in recognition decreases

VII. Conclusion

Recognizing Odia numerals is a challenging work because of similarity among some of the texts. In this paper we have develop a method to recognize Odia hand written numerals by training the network using Long Short Term Memory network to correctly recognize the digits. The database with more number of handwritten numbers can be considered for more accuracy value. While developing the method the accuracy and time factor was also considered for better performance. An accuracy of 97.93% is achieved in this method to recognize Odia handwritten numerals which is a remarkable result. The steps followed in this method are described in details which will be helpful for researchers for implementing their research work.

References

[1] Xin, M., & Wang, Y. Research on image classification model based on deep convolution neural network. *EURASIP Journal on Image and Video Processing*, *2019*(1), 40.

[2] Wu, P., Wang, F., & Liu, J. (2018, November). An Integrated Multi-Classifier Method for Handwritten Chinese Medicine Prescription Recognition. In *2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS)* (pp. 1-4). IEEE.

[3] Qiu, J., Zhou, Y., Wang, Q., Ruan, T., & Gao, J. (2019). Chinese Clinical Named Entity Recognition Using Residual Dilated Convolutional Neural Network with Conditional Random Field. *IEEE Transactions on NanoBioscience*.

[4] Saha, C., Faisal, R. H., & Rahman, M. M. (2019, February). Bangla Handwritten Digit Recognition Using an Improved Deep Convolutional Neural Network Architecture. In *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)* (pp. 1-6). IEEE.

[5] Zhuang, H., Li, C., & Zhou, X. (2018, July). CCRS: Web service for Chinese character recognition. In *2018 IEEE International Conference on Web Services (ICWS)* (pp. 17-25). IEEE.

[6] Xiao, X., Yang, Y., Ahmad, T., Jin, L., & Chang, T. (2017, November). Design of a very compact cnn classifier for online handwritten chinese character recognition using dropweight and global pooling. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)* (Vol. 1, pp. 891-895). IEEE

[7] Belay, B. H., Habtegebrial, T. A., & Stricker, D. (2018, October). Amharic Character Image Recognition. In *2018 IEEE 18th International Conference on Communication Technology (ICCT)* (pp. 1179-1182). IEEE.

[8] Yang, Y., Li, D., & Duan, Z. (2017). Chinese vehicle license plate recognition using kernel-based extreme learning machine with deep convolutional features. *IET Intelligent Transport Systems*, *12*(3), 213-219.

[9] Bulan, O., Kozitsky, V., Ramesh, P., & Shreve, M. (2017). Segmentation-and annotation-free license plate recognition with deep localization and failure identification. *IEEE Transactions on Intelligent Transportation Systems*, *18*(9), 2351-2363.

[10] Nguyen, T. N., & Nguyen, D. D. (2018, November). A new convolutional architecture for Vietnamese car plate recognition. In *2018 10th International Conference on Knowledge and Systems Engineering (KSE)* (pp. 7-12). IEEE.

[11] Dhar, P., Guha, S., Biswas, T., & Abedin, M. Z. (2018, February). A System Design for License Plate Recognition by Using Edge Detection and Convolution Neural Network. In *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)* (pp. 1-4). IEEE.

[12] Mondal, M., Mondal, P., Saha, N., & Chattopadhyay, P. (2017, December). Automatic number plate recognition using CNN based self synthesized feature learning. In *2017 IEEE Calcutta Conference (CALCON)* (pp. 378-381). IEEE.

[13] Wu, Guohua, et al. "An Attention-Based BiLSTM-CRF Model for Chinese Clinic Named Entity Recognition." *IEEE Access* 7 (2019): 113942-113949.

[14] Naz, S., Umar, A. I., Ahmed, R., Razzak, M. I., Rashid, S. F., & Shafait, F. (2016). Urdu Nasta'liq text recognition using implicit segmentation based on multi-dimensional long short term memory neural networks. *SpringerPlus*, *5*(1), 2010.

[15] Ren, H., Wang, W., Lu, K., Zhou, J., & Yuan, Q. (2017, July). An end-to-end recognizer for in-air handwritten Chinese characters based on a new recurrent neural networks. In *2017 IEEE International Conference on Multimedia and Expo (ICME)* (pp. 841-846). IEEE.

[16] Nogra, James Arnold, Cherry Lyn Sta Romana, and Elmer Maravillas. "LSTM Neural Networks for Baybáyin Handwriting Recognition." *2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)*. IEEE, 2019.

[17] Dash, Kalyan S., N. B. Puhan, and Ganapati Panda. "A hybrid feature and discriminant classifier for high accuracy handwritten Odia numeral recognition." *2014 IEEE REGION 10 SYMPOSIUM*. IEEE, 2014.

[18] Jena, Om Prakash, et al. "Implementation of Linear Discriminant Analysis for Odia Numeral Recognition." *2018 International Conference on Information Technology (ICIT)*. IEEE, 2018.

[19] S. S. Pattanayak, "Recognizing ODIA handwritten scripts," *2017 Second International Conference on Electrical, Computer and Communication Technologies(ICECCT)*, Coimbatore, 2017, pp.1-4.doi: 10.1109/ICECCT.2017.8117879