Handwritten Odia Digits Recognition Using Residual Neural Network

MrinmoySen[1], Shaon Bandyopadhyay[1], Palash Ray[1], Mahuya Sasmal[1], Rajesh Mukherjee[1]

1(Computer Science and Engineering, Haldia Institute of Technology, India)

ABSTRACT
Handwritten digit recognition is a highly evolved research domain of pattern recognition. Handwritten digits are segmented first and then they are classified using the handwritten digit recognition technique. The Odia script is one of the writing systems in Odisha. In this paper, an efficient Handwritten Odia numeral digit recognition using ResNet is proposed. Deep learning is a recent research trend in this field. Architectures like Residual neural Networks (ResNet) are being used for classification. ResNet is an architecture that is computationally expensive and normally used to provide high accuracy in classification problems. The structural design of the network consists of sacks of two convolutional (Conyv2D) layers with Batch Normalization and an activation function called Relu. We evaluated our scheme on 4970 handwritten samples of Odia numerals from the ISI database and from the experiment we have achieved 99.20% recognition rate.

Keywords - CNN, ResNet, Handwritten digits recognition, Odia, ISI database.

I. INTRODUCTION

In the area of image processing, Computer Vision and Pattern Recognition (CVPR) is a major growing field, where pattern recognition is one of the most important needs in Natural language processing (NLP). Handwritten digit recognition is a system that can recognize the characters from a digitalized or scanned handwritten document. This system has become an important part of various applications like office document automation, signature authentication, handwritten postcodes, auto-check, immigration data processing, health data record into digital format and many other applications [1]. This system becomes complicated because of challenges like characters written by the different writer are not identical in different aspects such as font, size, shape, and styles. Most of the previously proposed models are based on traditional pattern recognition where human expertise is required for feature engineering [2], [3].

The recent success of deep learning, especially Residual Neural Network (ResNet) for computer vision [4], [5], [6] [7] is used to recognize handwritten characters and digits as a computer vision problem. This ResNet is a deep Convolutional Neural Network (CNN) based model [8]. This is one of the current state-of-the-art deep learning models for image classification.

II. LITERATURE REVIEW

In the recognition of handwritten digits, various approaches have been proposed with very high accuracy rates [1][9][10] [11][12]. Various sets of classification techniques have been applied to this problem like Linear K-Nearest Neighbor, Random forest, Decision Tree, SVMs, Neural Network, and Convolutional Neural Network [9].

A deep learning technique for recognition of Arabic handwritten digits is proposed by Ahmed et.al [1]. The method uses CNN with LeNet-5 is trained and tested on the MADbase database that consists of 60000 training and 10000 testing images.

A Telugu printed numeral digits recognition system is proposed by Ravi et.al [13], where different feature extractions techniques like number of contours, skeleton feature, water reservoir features is used. This method trained and tested on database of 3150 printed multi-font printed Telugu numerals.

U. Pal et.al [14] proposed a method where off-line Bengali handwritten numerals were recognized which are unconstrained. This method is applied to their collected dataset of size 12000 and obtained an accuracy of 92.80%.

An approach for isolated Digit Recognition system is proposed by Vijay Kumar et.al [10]. In this approach features from digit image are extracted using Geometrical and Hosts pot features. This method used MNIST database which contains 60,000 training and 10000 testing samples.

Classifier Combination through Dempster-Shafer (DS) for Handwritten Bangla Digit Recognition is proposed by Basu et.al [22]. In this approach DS technique and MLP classifier for classification is used. Their method achieved 95.1% test accuracy.

A Digit Recognition system is proposed by Sarkar et.al [12]. In this method image features is extracted using Projection Histogram Features and Chain code Histogram Features. This method is trained and tested MNIST for Hindi numerals database that is more than 200,000 Hindi numerals.

III.PROPOSED WORK

The proposed method consists of five steps. In the first step, we have collected the data from the Handwritten Oriya numeral database. After collecting the data which is in grayscale we resize each image into 64×64 pixels. After that we used normalization techniques to convert the gray level values to the range of 0 to 1 values. In the second step, we reshape it from 2-dimensions to 3-dimensions. In the third step, we used a residual block which consists of 3×3 convolutional layers with the same number of output channels followed by a batch normalization layer and a ReLU activation function. In the residual block, we skip two convolution operations and add the input directly before the final ReLU activation function shown in figure 1. In this step, we extracted the features of image data atomically and used the algorithm for recognition of handwritten numerals. In fourth step, we use Adam Optimizer [19] as an optimization technique to maximize efficiency. In final step, diversity of training data set is increased by using real time augmentation. The proposed method block diagram is shown below in figure 2

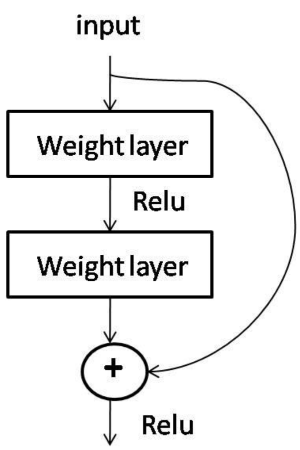

Figure 1: Residual Building Blocks

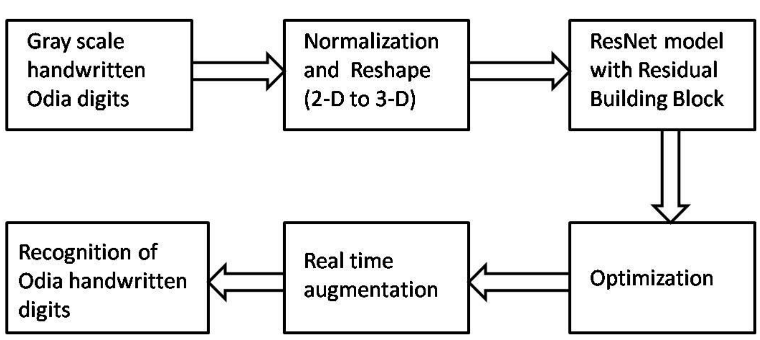

Figure 2: Block diagram of proposed method

A. Collection of Odia Dataset

The different handwritten digits of Odia numerical datasets are taken from ISI Kolkata Handwritten Oriya numeral database [29]. In this dataset we have collected 10 classes of numerical characters with total 4970 grayscale images. Then we converted each image into 64×64 pixels so that we can give fixed size input to our proposed model. Sample of digits from the database is shown in figure 3.

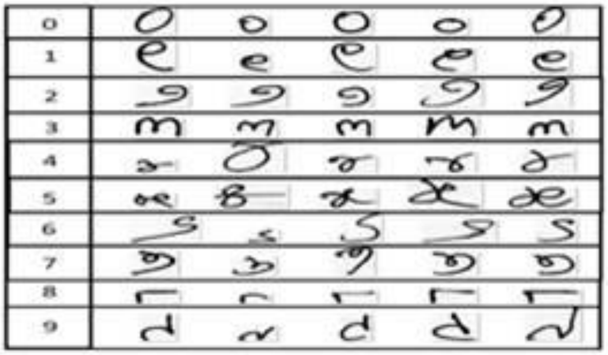

Figure 3: Different sample of dataset

B. Residual Neural Network (ResNet)

Computer Vision and pattern recognition is a major growing field in the area of image processing. ResNet plays a major role in computer vision. ResNet consists of Convolutional layers which are the core of most Computer Vision and pattern recognition systems today [16]. ResNet can be thought of various residual blocks containing convolutional layer followed by batch normalization and Relu activation function. There is also a skip layer in ResNet which helps us to overcome the vanishing gradient problem [17]. We have used Keras API with TensorFlow as a backend to implement this model. In this model we have used 50 layers where after the first layer we used maxpool (maxpool 2D), shown in figure 4. After that residual blocks are used and the output is added with the output of the first layer which is also called skip connection. Finally activation function Relu is shown in figure 5 and it is used on the output. Model architecture is shown in figure 6. We continued this setting until the last layer where we used Flatten layer and connected dense layer with 10 classes and finally softmax classifier for probability distribution.

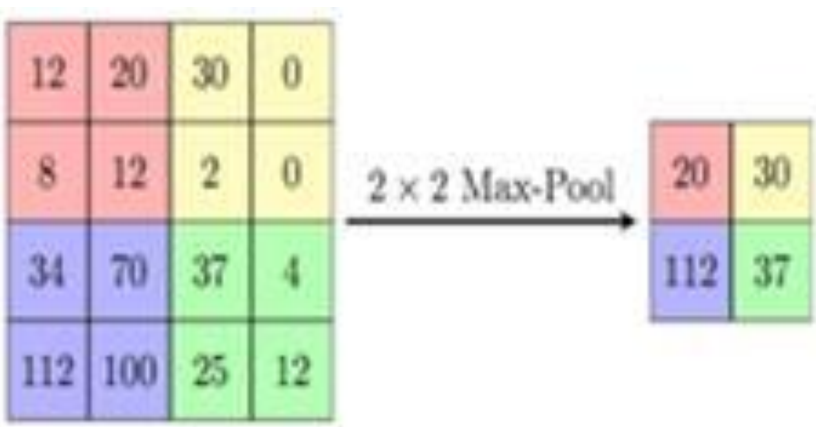

Figure 4: Realization of maxpool.

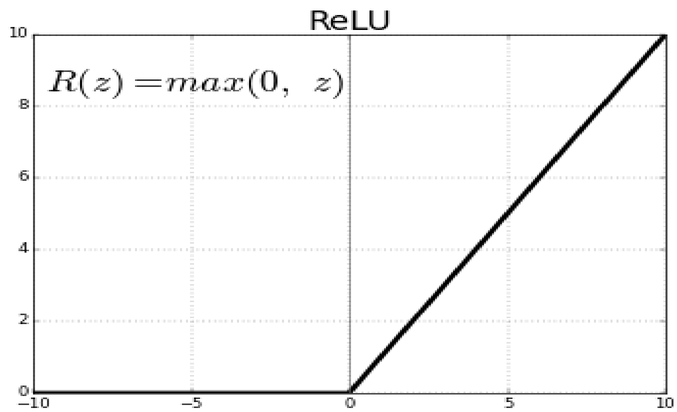

Figure 5: Relu Activation function

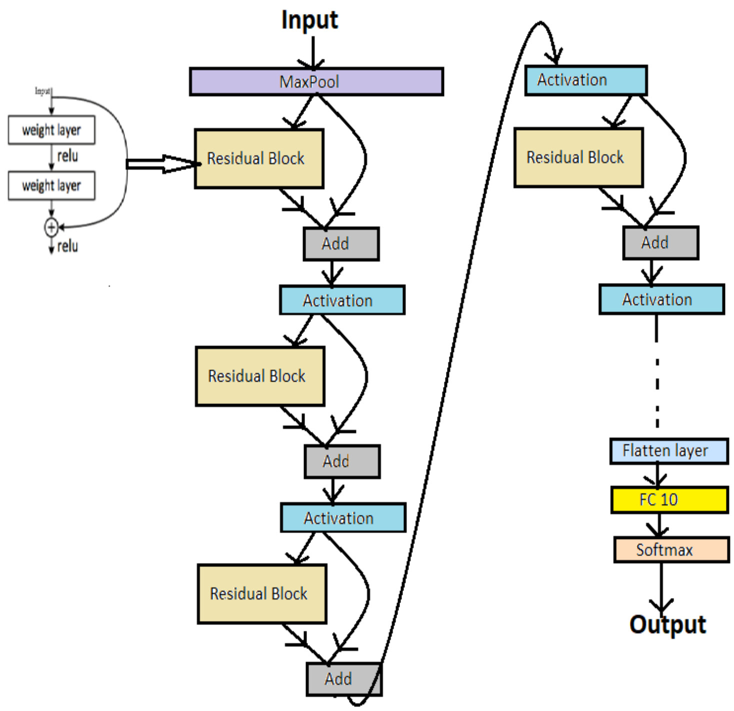

Figure 6: Model architecture

B. Softmax Classifier

It is a Logistic Regression classifier which is a simplification of the binary form of like hinge loss or squared hinge loss. It is used at the last dense layer. Mapping function is derived using

$$f(x_i;W)=Wx_i \tag{1}$$

Where x is input data itemsand w is weight [18].
Cross Entropy loss has the form

$$L_i = -\log\left(\frac{e^{f_{y_i}}}{\sum_j e^{f_j}}\right) \rightarrow L_i = -f_{y_i} + \log\sum_j e^{f_j} \tag{2}$$

Probabilistic interpretation can be defined as

$$P(y_i \mid x_i; W) = \frac{e^{f_{yi}}}{\sum_j e^{f_j}} \tag{3}$$

A single data point results in final loss function.

$$L_i = -\log(e^{S_{yi}} / \sum_j e^{S_j}) \tag{4}$$

Cross-entropy loss is computed by taking average of total dataset.

$$L = \frac{1}{N} \sum_{i=1}^{N} L_i \tag{5}$$

C. Adam Optimizer

In this model, Adam optimizer is used [19]. It uses to follow adaptive learning rate method, with which for different parameters, it computes individual learning rates. First and second moments of gradient to adapt the learning rate are estimated for each weight of the neural network using this technique. Random variable of N^{th} moment is defined as

$$m_n = E[X^n] \tag{6}$$

Where m is the moment and random variable is represented as X. The first moment is mean, and uncensored variance is the second moment. To estimates the moments, Adam utilizes exponentially moving averages which can be defined as

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \tag{7}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \tag{8}$$

Where m and v are moving averages, g is gradient on current mini-batch, and β is new introduced hyper-parameters of the algorithm.

D. Proposed algorithm

Input: Odia Numerical image

Output: Odia Numerical image recognition

Method: Handwritten Odia Digits Recognition Using Residual Neural Network

Step1: Gray scale images of Odia numerical are taken.

Step2: Pre-processing the gray scale images into 64×64 pixels.

Step3: Normalize the gray scale images i.e. from 0 to 255 into 0 to 1.

Step4: Reshape the images from 2D to 3D.

Step5: One hot encoding is done. i.e. 1 is represented as [0100000000].

Step6: Proposed model is used with Flatten layer and fully connected layers.

Step7: Input image is classified into suitable class using softmax classifier.

Step8: Accuracy is improved using Adam optimizer.

Step9: Real time augmentation technique is used to increase diversity of training data.

End

IV. RESULTS AND DISCUSSION

We evaluated the performance of our model on ISI Kolkata Handwritten Oriya numeral database using Google colaboratory [29]. We have selected 10 classes of numerical with a total 4970 images. Further the data set was split into a training set and a test set where 3479 images selected randomly for the training set and 1491 images selected randomly for testing. Our model reaches 99.20% accuracy on the validation dataset after 64 epochs. In our model total no of trainable parameter is 23,608,202 and non-trainable parameter is 53,120. Loss and accuracy curves for training and validation are presented in figure 7 and 8.

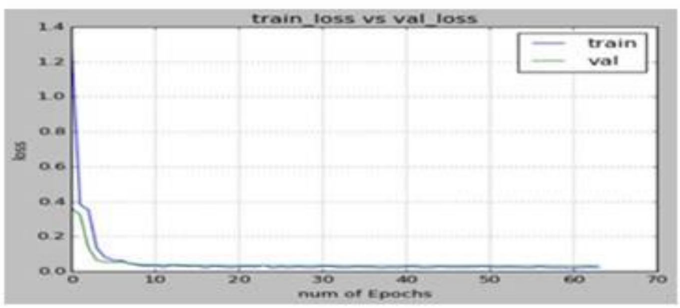

Figure 7: Training loss vs. validation loss curve

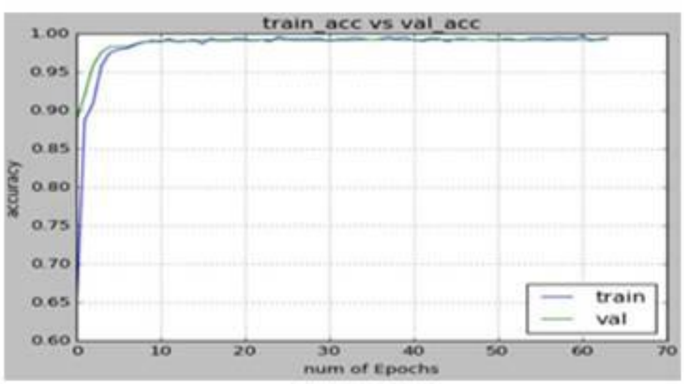

Figure 8: Training accuracy vs. validation accuracy

Authors	Method	Rec.Rate (%)
Bhowmik et.al [26]	Stroke-based features	90.50
K.Roy et.al [27]	NN and Quadratic	94.81
Sarangi et. al [28]	Hopfied Network	95.40
U. Pal et.al [3]	Quadratic Classifier	98.36
Proposed Method	**ResNet**	**99.20**

Table 1: Performance comparison with existing work

V. CONCLUSION

A deep learning approach for Odia Numeral Digit Recognition is been proposed in this paper. We evaluated the performance using ResNet on a standard ISI Kolkata Handwritten Oriya numeral database. From the results, it is observed that ResNet yields the best accuracy for Odia Numeral Digit Recognition compared to the alternative techniques. Our method achieved 99.20% recognition rate.

REFERENCES

[1] Ahmed El-Sawy, Hazem EL-Bakry , “CNN for Handwritten Arabic Digits Recognition Based on LeNet-5” Advances in Intelligent Systems and Computing, Springer International Publishing AG 2017.
[2] S. Bag, G. Harit, A survey on optical character recognition for bangla and devanagari scripts, Sadhana 38 (1) (2013) 133-168.
[3] U. Pal, R. Jayadevan, N. Sharma, Handwriting recognition in indian regional scripts: A survey of offline techniques 11 (1) (2012) 1:1-1:35.
[4] Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: F. Pereira, C. J. C. Burges, L. Bottou, K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 25, Curran Associates, Inc., 2012, pp. 1097-1105.
[5] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436-444.
[6] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, T. Chen, Recent advances in convolutional neural networks, Pattern Recognition 77 (2018) 354-377.
[7] F. Sultana, A. Sufian, P. Dutta, Advancements in image classification using convolutional neural network, in: 2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), 2018, pp. 122-129.
[8] Alom, M.Z., Sidike, P., Taha, T.M., Asari, V.K. “,Handwrittenbangla digit recognition using deep learning”, May 2017.
[9] U Ravi Babu, Y V VSatyanarayana and S. MarthuPerumal , " Printed Telugu Numeral Recognition based on Structural, Skeleton and Water Reservoir Features", International Journal Of Computers & Technology, Vol 10, No 7. August, 2013
[10] R. Vijaya Kumar Reddy, Dr.U. Ravi Babu, “Digit Recognition Using Geometrical and Hotspot Features”, Journal of Advanced Research in Dynamical & Control Systems, Vol. 10, No. 3, 2018.
[11] Alkhateeb, J.H., Alseid, M. “ DBN - based learning for Arabic handwritten digit recognition using DCT features.”, In 6th International Conference on Computer Science and Information Technology (CSIT), pp.222–226 (2014).
[12] Akarshan Sarkar, Kritika Singh, Prof. AmitabhaMukerjee” Project Report Handwritten Hindi Numerals Recognition System”.
[13] U Ravi Babu, Y V VSatyanarayana and S. MarthuPerumal , " Printed Telugu Numeral Recognition based on Structural, Skeleton and Water Reservoir Features", International Journal Of Computers & Technology, Vol 10, No 7. August, 2013

[14] U. Pal, B. B. Chaudhuri, A. Belaid, A complete system for bangla hand-written numeral recognition, IETE Journal of Research 52 (1) (2006) 27{34.
[15] Dua, D. and Graff, C. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.
[16] He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition.In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
[17] Hochreiter, S., 1998. The vanishing gradient problem during learning recurrent neural nets and problem solutions.International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 6(02), pp.107-116.
[18] Weiyang Liu, YandongWen, Zhiding Yu, Meng Yang," Large-Margin Softmax Loss for Convolutional Neural Networks ", Proceedings of the 33 rd International Conference on Machine Learning, New York, NY, USA, 2016. JMLR: W&CP volume 48.
[19] Kingma, D.P. and Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
[20] Dongre, V.J. and Mankar, V.H., 2015, March. Devanagari offline handwritten numeral and character recognition using multiple features and neural network classifier.In 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 425-431). IEEE
[21] Singh, P.K., Das, S., Sarkar, R. and Nasipuri, M., 2016, December. Recognition of offline handwritenDevanagari numerals using regional weighted run length features. In 2016 International Conference on Computer, Electrical & Communication Engineering (ICCECE) (pp. 1-6). IEEE
[22] Basu, S., Sarkar, R., Das, N., Kundu, M., Nasipuri, M. and Basu, D.K., 2005, December. Handwritten Bangla digit recognition using classifier combination through DS technique.In International conference on pattern recognition and machine intelligence (pp. 236-241).Springer, Berlin, Heidelberg.
[23] U. Bhattacharya, B. B. Chaudhuri, Handwritten numeral databases of in-dian scripts and multistage recognition of mixed numerals, IEEE Transactions on Pattern Analysis and Machine Intelligence 31 (3) (2009) 444{457.doi:10.1109/TPAMI.2008.88.
[24] C. Vasantha Lakshmi, Ritu Jain, C. Patvardhan, "Handwritten Devanagari Numerals Recognition With Higher Accuracy", Proc. of IEEE Int. Conf. on Computational Intelligence and Multimidia Application, 2007, pp 255-259.
[25] G.G.Rajput, S.M.Mali, "Fourier Descriptor Based Isolated Marathi Handwritten Numeral Recognition", Int. Journal of Computer Application (0975 – 8887), Vol.3, No.4, 2010, pp.9-13.
[26] Bhowmik,T.K.; Parui, S.K.; Bhattacharya, U.; Shaw, B.: An HMM based recognition scheme for handwritten Oriya numerals.In: Proceedings of 9th International Conference on Information Technology (ICIT'06), Bhubaneswar, 18–21 December 2006, pp. 105–110.
[27] K. Roy, T. Pal, U. Pal and F. Kimura, "Oriya handwritten numeral recognition system," Eighth International Conference on Document Analysis and Recognition (ICDAR'05), Seoul, South Korea, 2005, pp. 770-774 Vol. 2, doi: 10.1109/ICDAR.2005.183.
[28] Sarangi, P.K., Sahoo, A.K. and Ahmed, P., 2012. Recognition of isolated handwritten oriya numerals using hopfield neural network. International Journal of Computer Applications, 40(8), pp.36-42.
[29] U. Bhattacharya and B. B. Chaudhuri, "Databases for research on recognition of handwritten characters of Indian scripts", In Proc. of the 8th Int. Conf. on Document Analysis and Recognition (ICDAR-2005), Seoul, Korea, vol. II, page: 789-793, 2005.