BanglaNet for Bangla Handwritten Characters and Numerals Recognition

Simanta Deb Turja*
simantaturja@gmail.com

Mehedi Hasan†
abir.aust.102@gmail.com

Md. Hasib Al Zadid‡
hzsakkhor@gmail.com

I. Introduction

BANGLANET is developed to recognize bangla handwritten basic characters and numerals. Bangla handwritten characters and numerals recognition is one the most interesting issues in present time due to its various applications and usage in digitization techniques. BanglaNet is a simple, lightweight convolutional Neural Network Model for classifying Bangla Handwritten Characters which contains 50 basic characters(11 vowels and 39 consonants) and 10 numerals.

Bengali also known by its endonym Bangla, is an Indo-Aryan language primarily spoken by the Bengalis in Bangladesh and the Indian states of West Bengal, Tripura, Assam's Barak Valley. It is the official and most widely spoken language of Bangladesh and second most widely spoken of the 22 scheduled languages of India. With approximately 250 million native speakers and another 37 million as second language speakers, Bengali is the fifth most-spoken native language and the seventh most spoken language by total number of speakers in the world.

The task of recognizing Bangla handwritten characters and numerals is simple yet challenging because of its alignment and similarity. One of the most common reason of facing challenge in handwritten recognition research is because of the unique style to writing, different shapes and sizes. Moreover, automatic handwritten recognition is one of the most important research fields in recent years for its different applications such as OCR which helps to recognize the characters from images.

Fig. 1 Bangla Handwritten Characters and Numerals

*4th Year 2nd Semester, Department of Computer Science and Engineering, Ahsanullah University of Science and Technology
†4th Year 2nd Semester, Department of Computer Science and Engineering, Ahsanullah University of Science and Technology
‡4th Year 2nd Semester, Department of Computer Science and Engineering, Ahsanullah University of Science and Technology

II. Related Works

A Bangla handwritten character recognition system using CNN was developed by Rahman et al. [1] in 2015. Here, the CNN method achieved 85.96% accuracy on 50 classes. A custom dataset was prepared with total samples: 20,000 and 400 sample images for each class. Image resolution was 28x28 in this experiment.

A deep learning approach was applied on Bangla numeric digits by Zahangir et al. [2]. This approach managed a 98.8accuracy on the 10 classes. The dataset used in this experiment was CMATERdb3.1.1 with 6000 images for Bangla Numerals [10 numeric digit classes]. Image resolution was 32x32 in this experiment.

A Deep convolutional Neural Network approach was applied on merged dataset of EkushDB, BanglaLekhaDB and CMaterDB by Rabbu et al.[3] and achieved 96.5% accuracy on the combined datasets which contains approximately 0.4 Million images of Bangla Characters, numerals and compound characters.

A research on offline handwritten numeral recognition was conducted by Bhattacharya et al. [4] and the approached used in this case was, an ensemble of MLPs (Multi Layer Perceptrons) [5] which were combined by Adaboost. This paper provides a detailed comparison between handwritten Bangla, Oria and Devnagari numeric digit recognition using the Multi-Layer Perceptron.

III. Proposed Objectives

The following research model has several sub-tasks including dataset preprocessing, augmentation and designing the model.

A. Dataset Preprocessing

The proposed method used Ekush[6] dataset for training and testing purpose. It consists of Bangla basic characters, numerals and compound characters. As the proposed scheme is only working on basic characters and numerals, we extracted the basic characters and numerals from the dataset.

The dataset containing 200000 images resized into 32x32 pixels. For handwritten characters recognition, data augmentation helps more because a single person can write a character in a different variation. For data augmentation, some of the affine transformations such as 20 degree random rotation, scaling has been applied on the images.

B. Designing the CNN Model

The proposed CNN Model is a 13-layer convolutional neural network. All the convolutional layer with a kernel size of 3, these layers use ReLU activation with same padding. The output of these layer connected with (2x2) max pooling layer.

C. Sample Input Output of the Model

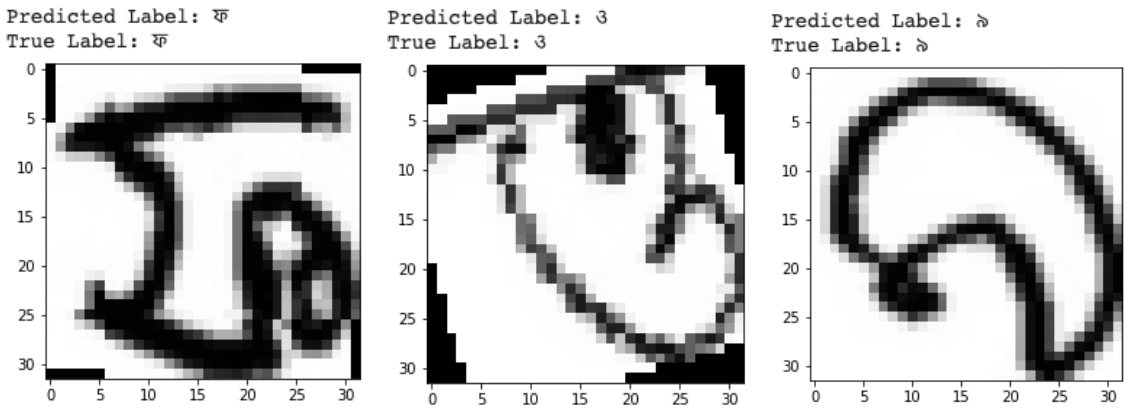

Fig. 2 Sample Input Output of the Model

IV. Methodology and System Architecture

Among different types of neural networks, CNN is one of the best methods to perform classification on image data. The proposed model BanglaNet that has been used to classify the Bangla handwritten characters and numerals was implemented using Pytorch as backend. It contains 6 convolution layers with kernel size of 3x3. The activation function used in these layers is ReLU activation function which introduces non-linearity. For the first layer, the input dimensions had to be specified, which, in this case, is 32x32x3, which means each image will be 3 channel 32x32 image. This layer is followed by a Max Pooling layer with pool size 2x2. There are 6 convolutional layers each followed by ReLU activation function and 2x2 Max pooling. The values obtained from the last pooling layer are then flattened as one dimensional tensors. There are three fully connected layers with 1024 nodes in the first layer, 512 nodes, 256 nodes in the second and third layer respectively. The optimization function used in this model is Adam optimizer with a learning rate of 0.001. Summary of the architecture is displayed on Table I.

Table I.

Model Architecture Summary		
Layer no (type)	Output Shape	Connected to
(Input Layer)	32, 32, 3	-
1 (Conv2D)	32, 32, 32	Input Layer
2 (Conv2D)	32, 32, 64	1
3 (Maxpool2D)	16, 16, 64	2
4 (Conv2D)	16, 16, 128	3
5 (Conv2D)	16, 16, 128	4
6 (Maxpool2D)	8, 8, 128	5
7 (Conv2D)	8, 8, 256	6
8 (Conv2D)	8, 8, 256	7
9 (Maxpool2D)	4, 4, 256	8
10 (Flatten)	4096	9
11 (Dense)	1024	10
12 (Dense)	512	11
Output(Dense)	60	12

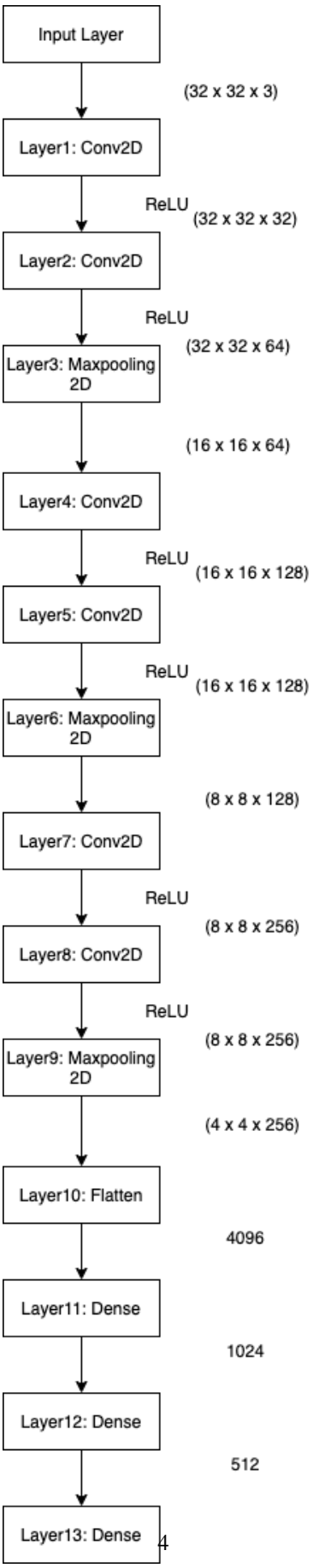

Fig. 3 BanglaNet Model Architecture

V. Experiments

A. Dataset

The dataset consists of 200000 images in total after deducting the compound characters from EkushDB. It contains sample images of 10 Bangla numeric digits and 50 Bangla basic characters, a total of 60 classes. The dataset was converted to a pytorch dataloader class for training. Then train_val_test is performed, to create training, validation and testing set. Validation set has been used during the training phase for validating the accuracy of the model on unknown data. 10% of the total dataset was used as validation set and 10% of the data as testing set.

B. Result

The training phase was executed for 15 epochs with a learning rate of 0.001. After 15 epochs proposed model gets 95.84% validation accuracy and 0.1 cross entropy validation loss for Ekush dataset.

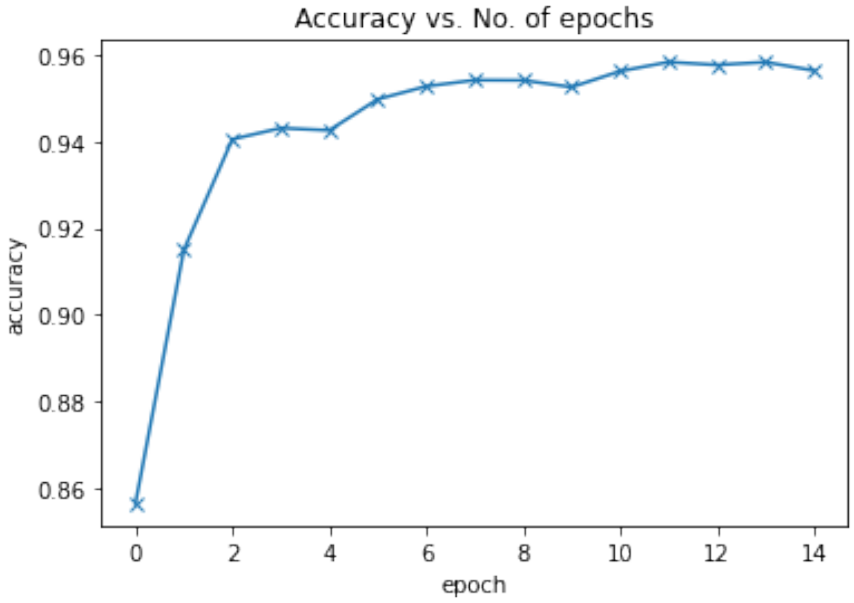

Fig. 4 Epoch vs. Validation accuracy

Table II.

Different Learning Rate Comparison		
Learning Rate	Accuracy	Loss
0.001	95.84%	0.2341
0.002	95.01%	0.2349
0.003	94.12 %	0.2402
0.004	93.97%	0.2435
0.005	93.77%	0.2468

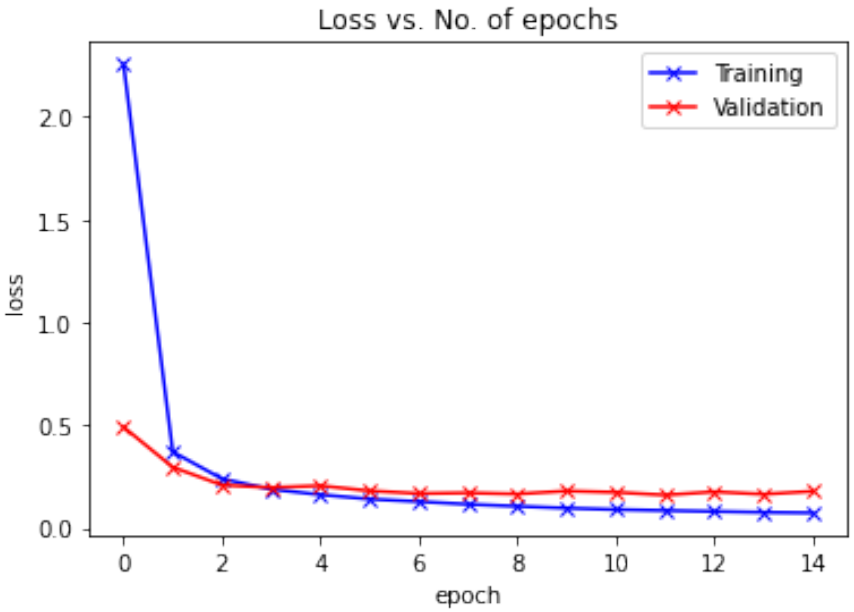

Fig. 5 Training and Validation Loss

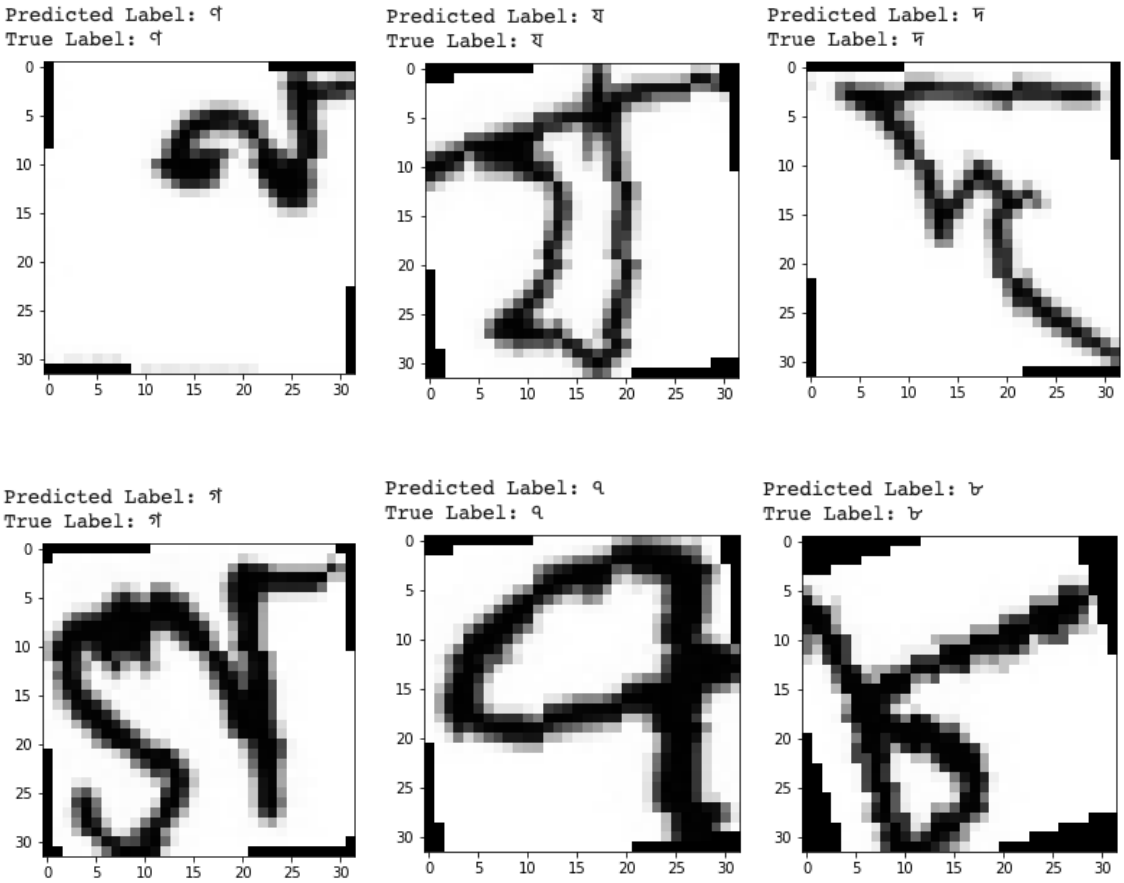

Fig. 6 Prediction and True Label of the model BanglaNet

Table III.

Comparison between ANN and CNN		
Model	Accuracy	Loss
ANN	92.67%	0.2520
CNN (BanglaNet)	95.84%	0.2341

VI. Conclusion

The research explores the opportunities and methods that are required to classify Bangla handwritten characters. The research proves that the CNN method is more efficient than ANN approaches. Although the system is providing admirable results in classifying individual letters of the Bangla alphabet, there is no scope to detect a sequence of characters or compound characters. Recognizing the compound characters and a sequence of characters or words as a whole, can be considered as future scope which requires more complex methods to implement.

References

[1] M. M. Rahman, S. I. P. C. S. M. R. e. a., M. Akhand, "Bangla handwritten character recognition using convolutional neural networ," *International Journal of Image, Graphics and Signal Processing (IJIGSP*, Vol. 7, No. 8, 2015, pp. 42, 49.

[2] M. Z. Alom, T. M. T., P. Sidike, and Asari, V. K., "Handwritten bangla digit recognition using deep learning," *Aviation Week & Space Technology*, Vol. arXiv preprint arXiv:1705.02680, 2017.

[3] A. K. M. S. Azad saha, S. A., S. Haque, and Hossain, S. A., "Ekushnet: Using convolutional neural network for Bangla handwritten recognition," *Procedia Comput. Sci.*, Vol. 143, No. 2, 2018, pp. 603–610.

[4] U. Bhattacharya, S. K. P. P. S., M. Shridhar, and Chaudhuri, B., "Offline recognition of handwritten bangla characters: an efficient twostage approach," *Pattern Analysis and Applications*, Vol. 15, No. 4, 2012, pp. 445–458.

[5] Jindal, T., and Bhattacharya, U., "Recognition of offline handwritten numerals using an ensemble of mlps combined by adaboost," *in Proceedings of the 4th International Workshop on Multilingual OCR*, Vol. ACM, No. 18, 2013.

[6] A. S. A. Rabby, M. S. I. S. A., S. Haque, and Hossain, S. A., "Ekush: A Multipurpose and Multitype Comprehensive Database for Online Off-Line Bangla Handwritten Characters," *in Recent Trends in Image Processing and Pattern Recognition, K. C. Santosh and R. S. Hegadi Eds. Singapore: Springer Singapore*, 2019, pp. 149–158.