Multiple Convolutional Neural Network Training for Bangla Handwritten Numeral Recognition

M. A. H. Akhand, Mahtab Ahmed
Dept. of Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh

M. M. Hafizur Rahman
Dept. of Computer Science, KICT
International Islamic University Malaysia
Selangor, Malaysia

***Abstract*—Recognition of handwritten numerals has gained much interest in recent years due to its various application potentials. The progress of handwritten Bangla numeral is well behind Roman, Chinese and Arabic scripts although it is a major language in Indian subcontinent and is the first language of Bangladesh. Handwritten numeral classification is a high-dimensional complex task and existing methods use distinct feature extraction techniques and various classification tools in their recognition schemes. Recently, convolutional neural network (CNN) is found efficient for image classification with its distinct features. In this study, three different CNNs with same architecture are trained with different training sets and combined their decisions for Bangla handwritten numeral recognition. One CNN is trained with ordinary training set prepared from handwritten scan images; and training sets for other two CNNs are prepared with fixed (positive and negative, respectively) rotational angles of original images. The proposed multiple CNN based approach is shown to outperform other existing methods while tested on a popular Bangla benchmark handwritten dataset.**

Keywords—Bangla Numeral; Convolutional Neural Network; Pattern Generation; Handwritten Numeral Recognition.

I. INTRODUCTION

Recognition of handwritten numerals has gained much interest in recent years due to its various application potentials in postal system automation, passports and document analysis, automatic bank cheque processing and even for number plate identification [1]. The progress of handwritten Bangla numeral is well behind Roman, Chinese and Arabic script [1, 2, 3] although it is a major language in Indian subcontinent and is the first language of Bangladesh.

A few notable works are available for Bangla handwritten numeral recognition. Bashar et al. [4] investigated a digit recognition system based on windowing and histogram techniques. Windowing technique is used to extract uniform features from scanned image files and then histogram is produced from the generated features. Finally, recognition of the digit is performed on the basis of generated histogram. Khan et al. [5] employed evolutionary approach to train artificial neural network for Bangla handwritten numeral. They used boundary extraction in order to extract the numeral in a single window by horizontal-vertical scanning and applied scaling to convert the image into fixed sized matrix. Then Multi-Layer Perceptron (MLP) are evolved for recognition. Basu et al. [6] used Dempster-Shafer (DS) technique for combination of classification decisions obtained from two MLP based classifiers trained with shadow feature and centroid feature for handwritten Bangla numeral. Pal et al. [3] introduced a technique based on the concept of water overflow from the reservoir for feature extraction and then employed binary tree classifier for Bangla handwritten numeral recognition.

Wen et al. [7] proposed handwritten Bangla numeral recognition system for automatic letter sorting machine. In their study Support Vector Machine (SVM) is combined with extensive feature extractors using Principal Component Analysis (PCA) and kernel Principal Component Analysis (KPCA). Das et al. [8] also used SVM for classification but used different techniques for feature selection. In their proposed system, a genetic algorithm (GA) based region sampling strategy has been employed to select an optimal subset of local regions which contains high discriminating information about the pattern shapes.

Wen and He [9] proposed a kernel and Bayesian Discriminant based method to recognize handwritten bangla numeral. Recently, Nasir and Uddin [10] introduced a hybrid system for recognition of handwritten Bangla numeral for the automated postal system, which performed feature extraction using k-means clustering, Baye's theorem and Maximum a Posteriori, then the recognition is performed using SVM. On the other hand, Bhattacharya and Chaudhuri [11] presented a multistage cascaded recognition scheme using wavelet-based multi-resolution representations and multilayer perceptron (MLP) classifiers. The scheme first computes features using wavelet-filtered image at different resolutions. The scheme has two recognition stages and the first stage involves a cascade of three MLP classifiers

Handwritten numeral classification is a high-dimensional complex task and traditional MLP require much computation to work with grayscale image. In general, traditional methods first extract features from the input image and then use MLP based methods for classification task [6, 11]. On the other hand, Convolutional Neural Network (CNN) is efficient for image classification with its distinct features. It automatically provides some degree of translation invariance. Most recently, Akhnad et al. [12] proposes a CNN based Bangla handwritten numeral recognition is investigated. It does not employ any feature extraction method like other existing works. Handwritten numeral images are normalized first and then employ CNN to classify individual numeral in the study. The

978-1-5090-2427-8/16 $31.00 © 2016 IEEE
DOI 10.1109/ICCCE.2016.73

CNN based method showed satisfactory classification accuracy and outperformed some other existing methods.

In this study, an updated and different CNN based method has been investigated for Bangla handwritten numeral recognition. A moderated pre-processing has been adopted to produce patterns from handwritten scan images. On the other hand, three CNNs have been trained with different training sets and combined their decisions for final decision. One CNN is trained with ordinary training set prepared from the handwritten images. A simple rotation based approach is employed to generate training sets for other two CNNs. The proposed multiple CNN based approach is shown to outperform other existing methods while tested on a popular Bangla benchmark handwritten dataset.

The rest of the paper is organized as follows. Section II explains proposed multiple CNN based approach in detail which also contains preprocessing of handwritten images and training set preparation. Section III presents experimental results of the proposed method and performance comparison with other related works. Finally, a brief conclusion of the work is given in Section IV.

II. Multiple CNN Training for Bangla Handwritten Numeral Recognition

This section explains proposed Multiple Convolutional Neural Network (MCNN) for Bangla Handwritten Numeral Recognition in detail which has two major steps: (i) preprocessing of handwritten raw images of numerals and generate three training sets; and (ii) training three CNNs and combine their decisions. The following subsection gives brief description of each steps.

A. Training Set Generation from Benchmark Handwritten Numeral Image

Benchmark Bengali handwritten numerals image dataset maintained by CVPR unit, ISI, Kolkata [13] is considered in this study. Several recent studies used the dataset or in a modified form [13, 20]. The samples of CVPR dataset are the scanned images from pin codes used on postal mail pieces. The digits are from people of different age and sex groups as well as having different levels of education. The dataset has been provided in training and test images. The test set contains total 4000 images having 400 samples for each of 10 digits. On the other hand, the training set contains total 19392 images having around 1900 images of each individual digit. In this study, all 4000 test images and 18000 training images (1800 images from each digit) are considered and pre-processed. Fig.1 shows few sample images of each numeral.

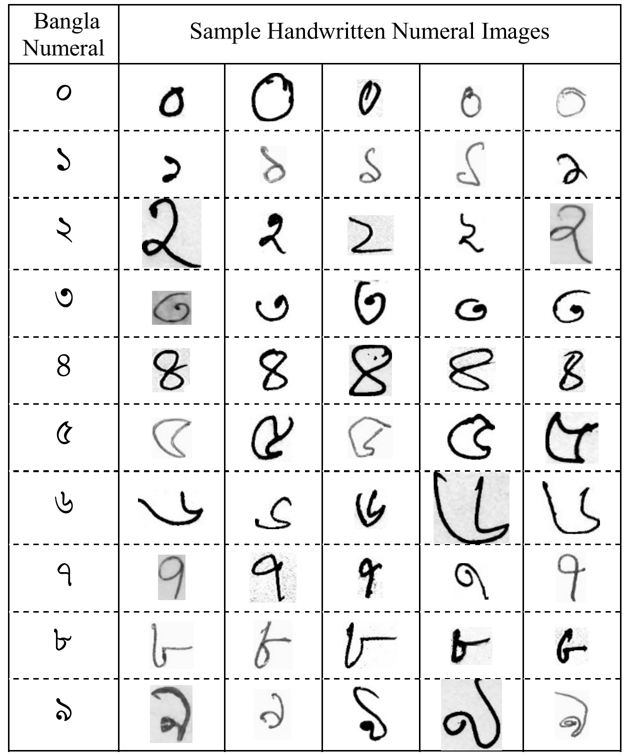


Fig.1. Samples of handwritten Bangla numerals from the dataset.

Pre-processing is performed on the images into common form that makes it appropriate to feed into classifiers. The original images are in different sizes, resolutions and shapes. Matlab R2015a is used to pre-process the images to same dimension and format. For Bengali numeral, ISI images are transformed into binary image. At first, an image is transformed into binary image with automatic thresholding of Matlab. This step removes background as well as improves intensity of written black color. Since black color is used for writing on white paper (background), the binary image files contains more white point (having value 1) than black (having value 0). To reduce computational overhead, images are converted through foreground numeral black to white and background changed to black. Written digit may be a portion in the scan image that is easily visible from the foreground-background interchanged image. An image has been cropped to the actual writing portion removing black lines from all four sides (i.e., left, right, top and bottom). Finally, images are resized into 28×28 dimension to maintain appropriate and equal format for all the numerals. To capture pattern values of resized images, the double type matrix is considered (instead of binary in the previous stages) so that best possible quality in the resized images is retained. Fig. 2 shows the basic steps of pattern generation from scanned handwritten digit.

The steps of Fig. 2 is common to prepare three different training sets to train three CNNs. One CNN is trained with ordinary training set prepared using the steps of Fig. 2. A simple rotation based technique has been followed to prepare other two different training sets for two CNNs. The rotation is performed in between foreground-background interchange (Step 2) and crop operation (Step 3). A rotational angle is defined and two different training sets are generated for clock wise and anti-clock wise rotation of images.

Scanned Handwritten digit	(Step 1) Binary image transformation	(Step 2) Foreground and background interchange	(Step 3) Cropped to actual portion	(Step 4) Resized into 28×28	Pattern for classification

Fig. 2. Steps of pattern generation from scanned handwritten digit.

C. Multiple CNN training and decision combination for Classification

Figure 3 shows the structure of multiple CNN (MCNN) for handwritten digit classification for sample image. In the figure CNN2 is trained with ordinary training set. On the other hand, CNN1 and CNN3 are trained with the training sets prepared anti-clock and clock wise rotation of defined θ degree, respectively. Training perform independently with training set and combined outputs for MCNN output. Simple average is considered for decision combination. It is notable that structures of individual CNNs are same and the most common CNN architecture has been considered [12, 16, 17].

Figure 4 shows individual CNN structure of MCNN that holds two convolutional layers with 5×5 sized local receptive

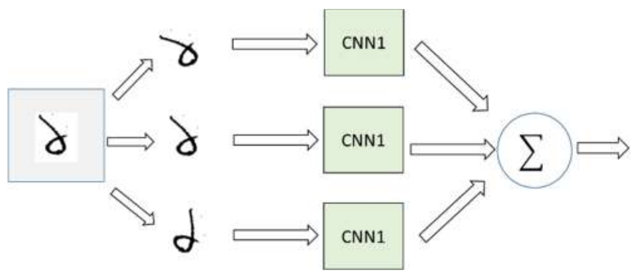


Fig.3. Structure of MCNN for handwritten digit classification.

fields (i.e., kernel) and two subsampling layers with 2×2 sized local averaging area along with input and output layers. Input layer contains 784 nodes for 28×28 pixels image. 1st convolutional operation produces first level six feature maps. Convolution operation with kernel spatial dimension of 5 reduces 28 spatial dimension to 24 spatial dimension [16, 17]. Therefore, each 1st level feature map size is 24×24. In 1st subsampling layer, the 1st level 24×24 feature maps are down-sampled into 12×12 feature maps by applying a local averaging with 2×2 area.

Second convolution and subsampling operations are similar to 1st convolution and subsampling operations, respectively. 2nd convolutional operation produces distinct 12 feature maps; a local receptive field of size 5×5 transforms a feature map of size 12×12 into 8×8. Then 2nd subsampling operation resizes each feature map to size of 4×4. These 12 features map values are considered as 192 ($=12 \times 4 \times 4$) distinct nodes those are fully connected to 10 feature maps (the output nodes) for numeral set. In the output layer, errors are measured by comparing desired output with the actual output. The training of CNN is performed to minimize the error (E).

$$E = \frac{1}{2}\frac{1}{PO}\sum_{p=1}^{P}\sum_{o=1}^{O}\left(d_o(p) - y_o(p)\right)^2, \qquad (1)$$

where P is the total number of patterns; O is the total output nodes of the problem; d_o and y_o are the desired and actual output of a node for a particular pattern p. In training, the kernel values with bias in different convolution layers and weights of hidden-output layers are updated. The description of the architecture is also available in the previous studies [12, 16].

III. Results and Discussions

MCNN is implemented in Matlab R2015a on HP pro desktop machine (CPU: Intel Core i7 @ 3.60 GHz and RAM: 8.00 GB) in Window 7 (64bit) environment. The batch wise training has been performed in this study due to large sized training set; and experiments have been conducted with different batch sizes. Weights of the CNN are updated once for a batch of image patterns and number of batch size (BS) is considered as a user defined parameter. At first we observed

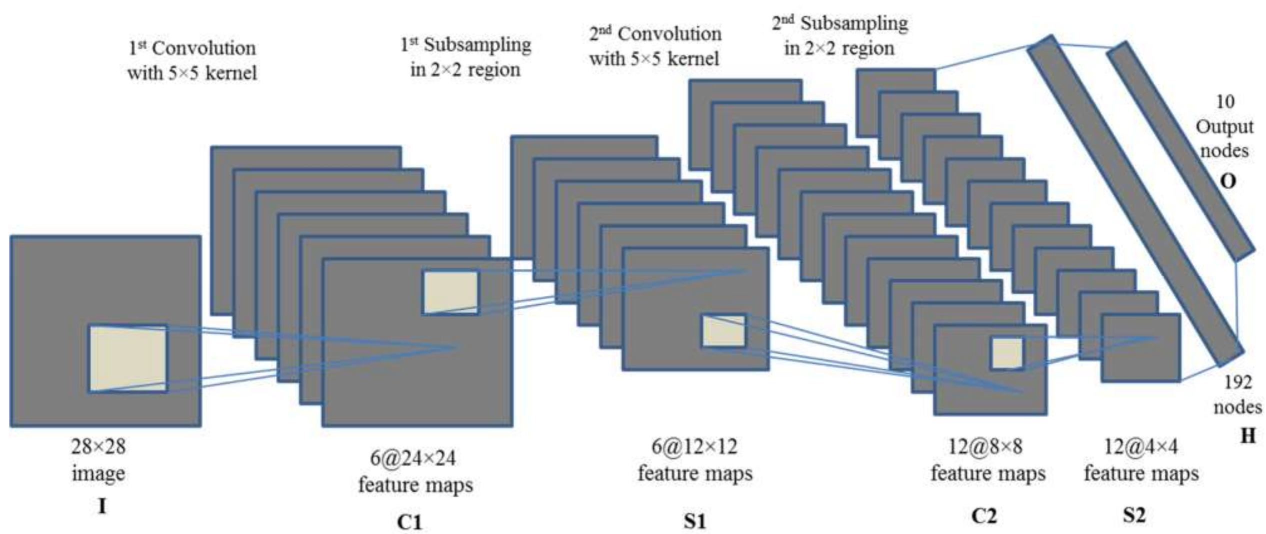
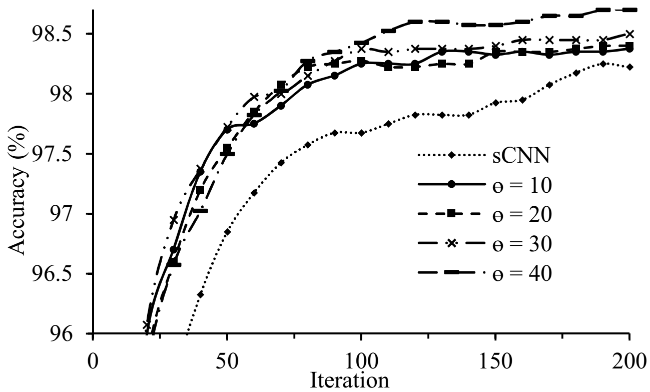
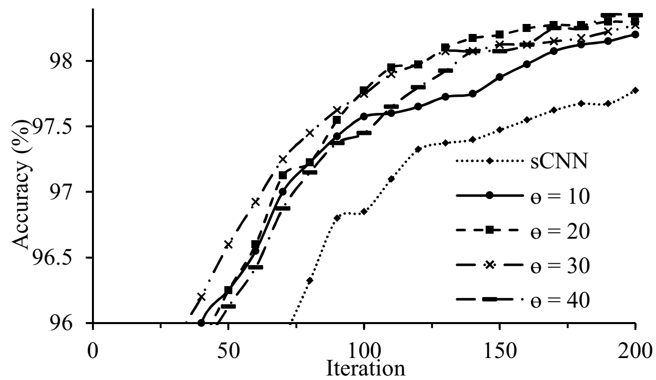


Fig. 4. Structure of the individual CNN in proposed MCNN



(a) Batch Size (BS) = 50



(b) Batch Size (BS) = 100

Fig.5. Test set accuracy of proposed MCNN for BS= 50 and 100 varying rotational angle (ɵ) 10 to 40.

the effect of BS on the performance of MCNN and outcome compared with related methods.

TABLE I. SAMPLE HANDWRITTEN IMAGES THAT ARE MISCLASSIFIED BY MCNN.

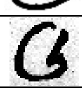
Handwritten Numeral Image	True Numeral	Image Classified as
	১	৯
	৫	০
	৬	৫
	৬	৫
	৯	১
	৯	১

Figure 5 shows test set accuracy of proposed MCNN for two BS values (i.e., 50 and 100) varying rotational angle (ɵ) 10 to 40 degree. The single CNN (sCNN) is also considered for better understanding and is marked as 'sCNN' in the figure. It clearly observed from the figure that MCNN with any rotational angle is better than sCNN. It is observed from the figure that accuracy improves with rotational angle up to a certain level and accuracy with larger rotational angle is found worse than small angle in some cases such as in Fig. 5(b) for BS=100. According to the results presented in the Fig. 5, the test set recognition accuracy at iteration 200 for BS = 50 was 98.7% (misclassifying 52 cases out of total 4000 test patterns). At that point, the method misclassified 89 cases out of 18000 training patterns showing accuracy rate of 99.51%. Finally, the best test set accuracy of MCNN varying BS values 10, 50, 75 and 100 and iteration up to 500 is shown 98.80% for BS=50. Table I shows some handwritten character images those are misclassified. Due to large variation in writing styles, such character images are difficult to classify even by human.

Table II compares the test set recognition accuracy of the proposed MCNN method with other prominent works of Bangla handwritten numeral recognition. It also presents distinct features of individual methods. It is notable that proposed method did not employ any feature selection technique whereas an existing method uses one or two feature selection techniques. Without feature selection, proposed MCNN method is shown to outperform the existing methods. According to the table, MCNN achieved testing accuracy of 98.80%, on the other hand, the testing accuracy are 97.70% and 98.20% for the works of [8] and [11], respectively. It is notable that the best performed existing method, i.e., Ref. 11, trained the classifier with 10 times larger training set with artificially generated patterns. Moreover, effectiveness of multiple CNN is clearly understandable comparing the accuracy of the proposed method with the work of [12] which

TABLE II. A COMPARATIVE DESCRIPTION OF PROPOSED MCNN WITH EXISTING METHODS FOR BANGLA HANDWRITTEN NUMERAL RECOGNITION.

Ref. Work, Year	Feature Selection	Classification	Test Set Recog. Accuracy
Bhattacharya and Chaudhuri [11], 2009	Wavelet filter at different resolutions	Four MLPs in two stages (three + one)	98.20%
Wen and He [9], 2012	Eigenvalues and eigenvectors	Kernel and Bayesian discriminant	96.91%
Das et al. [8], 2012	GA	SVM	97.70%
Nasir and Uddin [10], 2013	K-means clustering and Bayes' theorem	SVM	96.80%
Akhnad et al. [12], 2015	No	CNN	97.93%
Proposed MCNN	No	CNN	98.80%

used only original patterns to train CNN. Finally, the result compared in the table clearly revealed the efficacy of the proposed MCNN for Bangla handwritten numeral.

IV. Conclusions

Handwritten numeral classification is a high-dimensional complex task and multiple CNN based method has been investigated in this study for handwritten Bangla numeral recognition. A moderated pre-processing and simple rotation based technique has been adopted in proposed MCNN. The proposed method is shown to outperform the existing methods on the basis of test set accuracy on benchmark Bangla handwritten numeral dataset.

References

[1] R. Plamondon and S. N. Srihari, "On-line and off-line handwritten recognition: A comprehensive survey," *IEEE Trans. on PAMI,* vol. 22, pp. 62-84, 2000.

[2] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document Recognition,"*Proceedings of the IEEE,* vol. 86, no. 11, pp. 2278–2324, November 1998.

[3] U. Pal, C. B. B. Chaudhuri and A. Belaid, "A System for Bangla Handwritten Numeral Recognition," *IETE Journal of Research, Institution of Electronics and Telecommunication Engineers*, vol. 52, no. 1, pp. 27-34, 2006.

[4] M. R. Bashar, M. A. F. M. R. Hasan, M. A. Hossain and D. Das, "Handwritten Bangla Numerical Digit Recognition using Histogram Technique," *Asian Journal of Information Technology*, vol. 3, pp. 611-615, 2004.

[5] M. M. R. Khan, S. M. A. Rahman and M. M. Alam, "Bangla Handwritten Digits Recognition using Evolutionary Artificial Neural Networks*" in Proc. of the 7th International Conference on Computer and Information Technology (ICCIT 2004),* 26-28 December, 2004, Dhaka, Bangladesh.

[6] S. Basu, R. Sarkar, N. Das, M. Kundu, M. Nasipuri and D. K. Basu, "Handwritten BanglaDigit Recognition Using Classifier Combination Through DS Technique," *LNCS*, vol. 3776, pp. 236–241, 2005

[7] Y. Wen, Y. Lu and P. Shi, "Handwritten Bangla numeral recognition system and its application to postal automation," *Pattern Recognition*, vol. 40, pp. 99-107, 2007.

[8] N. Das, R.Sarkar, S. Basu, M. Kundu, M. Nasipuri and D. K. Basu, "A genetic algorithm based region sampling for selection of local features in handwritten digit recognition application," *Applied Soft Computing*, vol. 12, pp. 1592-1606, 2012.

[9] Y. Wen and L. He, "A classifier for Bangla handwritten numeral recognition," *Expert Systems with Applications*, vol. 39, pp. 948-953, 2012.

[10] M. K. Nasir and M. S. Uddin, "Hand Written Bangla Numerals Recognition for Automated Postal System," *IOSR Journal of Computer Engineering (IOSR-JCE)*, vol. 8, no. 6, pp. 43-48, 2013.

[11] U. Bhattacharya and B. B. Chaudhuri, Handwritten numeral databases of Indian scripts and multistage recognition of mixed numerals, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 31, no. 3, pp. 444-457, 2009.

[12] M. A. H. Akhand, Md. Mahbubar Rahman, P. C. Shill, Shahidul Islam and M. M. Hafizur Rahman, "Bangla Handwritten Numeral Recognition using Convolutional Neural Network," *in Proc. of International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT2015)*, Dhaka, Bangladesh, pp. 1-5, May 21-23, 2015.

[13] Off-Line Handwritten Bangla Numeral Database, Available: http://www.isical.ac.in/~ujjwal/, accessed July 12, 2015.

[14] M. A. H. Akhand, and K. Murase, "Ensembles of Neural Networks based on the Alteration of Input Feature Values*," International Journal of Neural Systems*, vol. 22, issue 1, pp. 77-87, 2012.

[15] M. A. H. Akhand, M. M. Hafizur Rahman, and K. Murase, "Pattern Generation through Feature Values Modification and Decision Tree Ensemble Construction," *International Journal of Machine Learning and Computing (IJMLC), vol. 3, no. 4, pp. 322-331, 2013.*

[16] Md. Mahbubar Rahman, M. A. H. Akhand, Shahidul Islam, Pintu Chandra Shill and M. M. Hafizur Rahman, "Bangla Handwritten Character Recognition using Convolutional Neural Network," *I.J. Image, Graphics and Signal Processing(IJIGSP),* vol. 7, no. 3, pp. 42-49, 2015.

[17] Feature extraction using convolution. Available: http://deeplearning.stanford.edu/wiki/index.php/

[18] T. Liu et al., "Implementation of Training Convolutional Neural Networks," *arXiv preprint arXiv:1506.01195*, 2015.

[19] Feature extraction using convolution, UFLDL Tutorial. Available: http://deeplearning.stanford.edu/, accessed November 12, 2015.

[20] CMATERdb 3.1.1: Handwritten Bangla Numeral Database, Available:http://code.google.com/p/cmaterdb/, accessed July 12, 2015.