Convolutional Neural Network Training with Artificial Pattern 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, a 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, CNN has been trained with the patterns plus a number of artificial patterns. A simple rotation based approach is employed to generate artificial patterns. The proposed CNN with artificial pattern 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 for handwritten Bangla numeral using two different feature sets. Feature sets they investigated are called shadow feature and centroid feature. 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. They used Support Vector Machine (SVM) classifier combined with extensive feature extractor 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 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, CNN has been trained with the patterns plus a number of artificial patterns. A simple rotation based approach is employed to generate artificial patterns. The proposed CNN with artificial pattern 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 CNN with artificial pattern in detail which contains preprocessing of handwritten scan images, artificial pattern generation and classification using CNN. 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. CNN Training with Artificial Pattern for Bangla Handwritten Numeral Recognition

This section explains proposed Convolutional Neural Network Training with Artificial Patterns (CNNAP) for Bangla Handwritten Numeral Recognition in detail which has three major steps: (i) preprocessing of handwritten raw images of numerals; (ii) generate artificial patterns; and (iii) training CNN. The following subsection gives brief description of each steps.

A. Benchmark Handwritten Numeral Image Data and Preprocessing

This 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.

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.

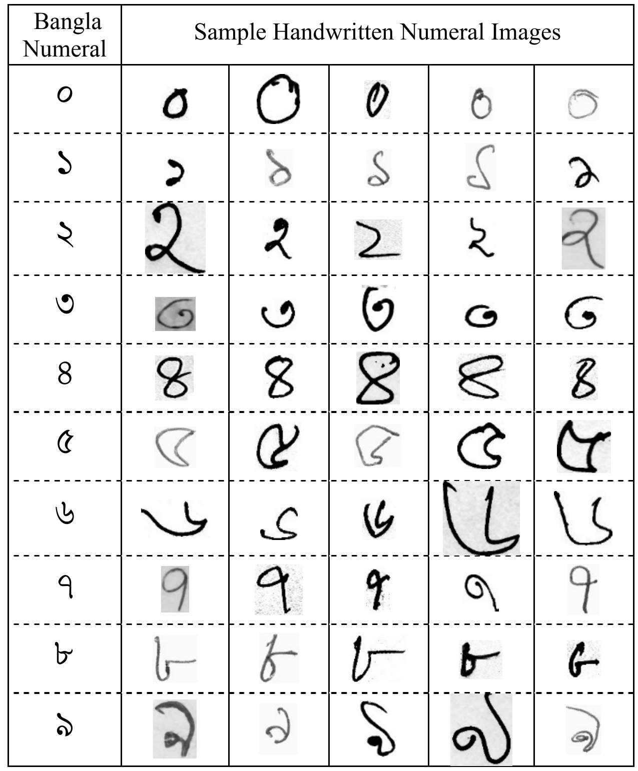


Fig.1. Samples of handwritten Bangla numerals from the dataset.

B. Artificial Pattern Generation

The use of artificial pattern to train CNN owing to improve recognition performance is the most important step of this study. Generated artificial pattern is commonly used in ensemble construction to promote diversity among individual classifiers (e.g., feed forward neural networks or decision trees) hence to improve recognition performance [14, 15]. Different techniques of pattern generation are investigated and such ensemble construction is found most effective for the small sized problems (i.e., problems having relatively small number of patterns) [14, 15]. In case of handwritten numeral recognition, Bhattacharya and Chaudhuri [11] used random

Scanned Handwritten digit	(1) Transform into binary image	(2) Foreground - background interchange	(3) Crop to actual portion	(4) Resized into 28×28 dimension	Pattern for classification

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

rotation and blurring to generate patterns. They trained MLPs with generated patterns plus original patterns; and training set size was 10 times of original set.

A simple rotation based technique has been followed in the present study to generate artificial patterns. The rotation is performed in between foreground-background interchange (Step 2) and crop operation (Step 3). It is observed that rotation before crop operation is most effective because it helps to remove some boundary line noise in scanned image. For a defined fixed rotational angle two different patterns are generated from an image for clock wise and anti-clock wise rotation. Therefore, the training set used in this study contains patterns following steps of Fig. 2 and patterns with rotation explained here. Since CNN has the ability to capture rotation, a small different rotation angle might enhance recognition capability of CNN.

C. CNN training for Classification

The most common CNN architecture has been considered in this study for classification purpose [12, 16, 17]. Figure 3 shows CNN structure that holds two convolutional layers with 5×5 sized local receptive 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 (i.e., 28-5+1) 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. This subsampling operation reduces both the spatial resolution of the feature map and sensitivity to shift and distortions.

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).

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

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

Experimental results of the proposed recognition scheme have been collected based on the samples of the prepared dataset discussed earlier. We applied CNN on the resized and normalized grayscale image files without any feature extraction technique. The method is implemented in Matlab R2015a. The experiment has been conducted 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

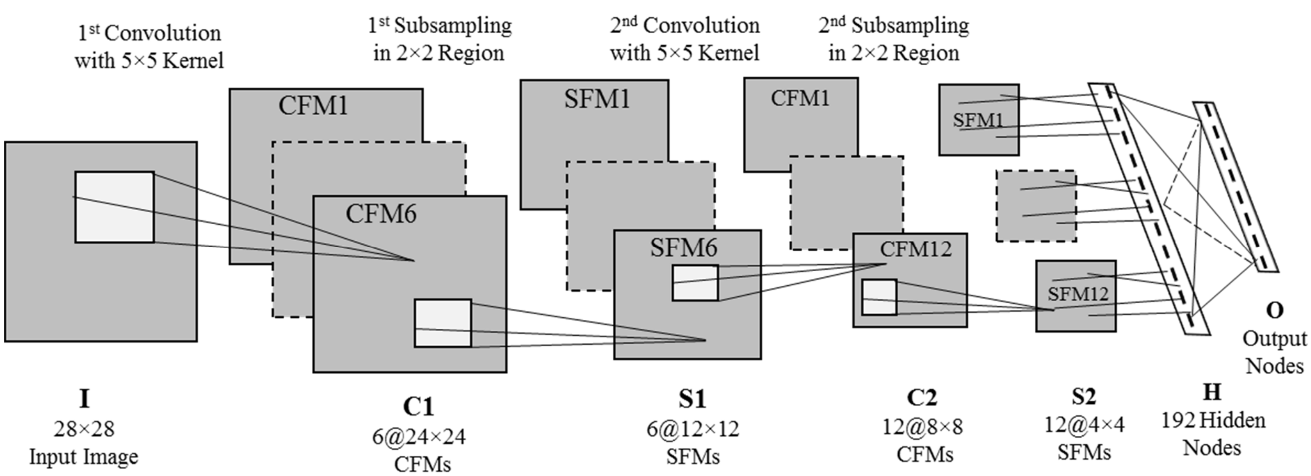
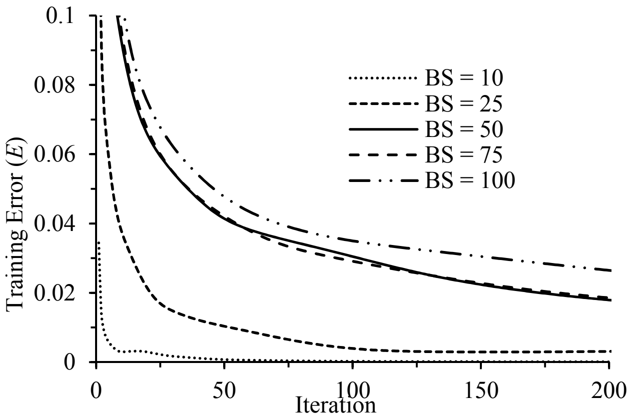
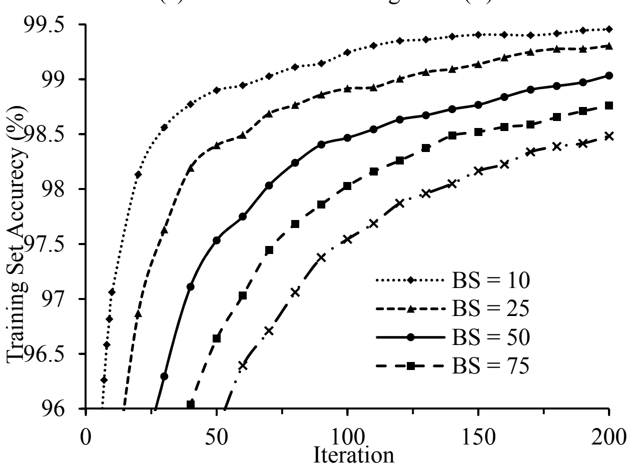


Fig. 3. Structure of CNN for proposed CNNAP.

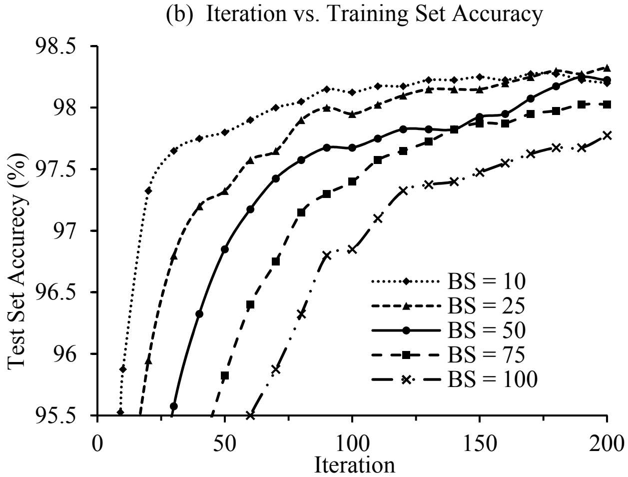
considered as a user defined parameter. At first we observed the effect of BS on the performance of CNN with the original training set. Then the effect of artificial patterns on CNN performance is measured while training with original patterns plus generated artificial patterns.



(a) Iteration vs. Training Error (E)



(b) Iteration vs. Training Set Accuracy

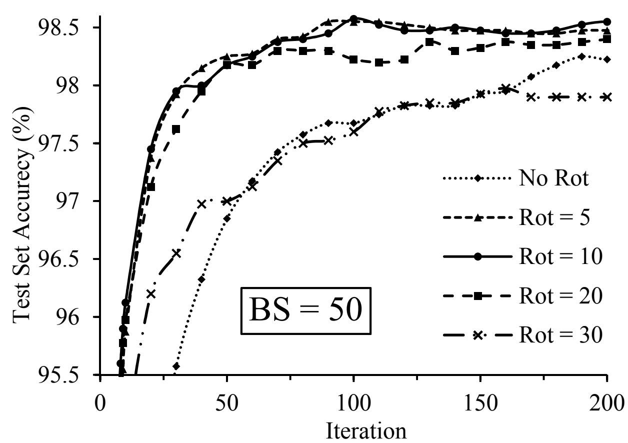


(c) Iteration vs. Test Set Accuracy

Fig.4. Effect of BS on CNN performance while training with original training set.

Figure 4 shows the effect of different BS values (i.e., 10, 25, 50, 75 and 100) on the performance of CNN while training with original training set only. It is observed from Fig. 4(a) that the training error (E) rapidly decreases for lower BS values (e.g., 10 or 25) and at the same fashion training set classification accuracy also improved. According to the Fig. 4(b) training accuracy for a lower BS value (e.g., 10 or 25) is always better than the accuracy for a larger BS value (e.g., 75 or 100) at any iteration. For lower BS value CNN are adjusted for small number of patterns at a time and as a result rapid error minimization is logical but it requires longer time to feed all the training patterns to CNN [16]. It is notable that accuracy on test set (the patterns of it is unseen during training) is more desirable than training set accuracy. The test set accuracy refers as the generalization ability of a system and indicates how good it will be for new or unseen cases. It is interesting to observe from Fig. 4(c) that test set accuracy for a larger BS value (e.g., 100) is compatible to the lower BS value (e.g., 25). The test set accuracy is observed 98.33 at 200 iteration for BS=50. On the other hand the test set accuracy for BS=10 at 200 iteration was 98.20.

Figure 5 shows test set accuracy of proposed CNNAP for two BS values (i.e., 50 and 100) with different rotational angle for artificial pattern generation. The rotational angle varied



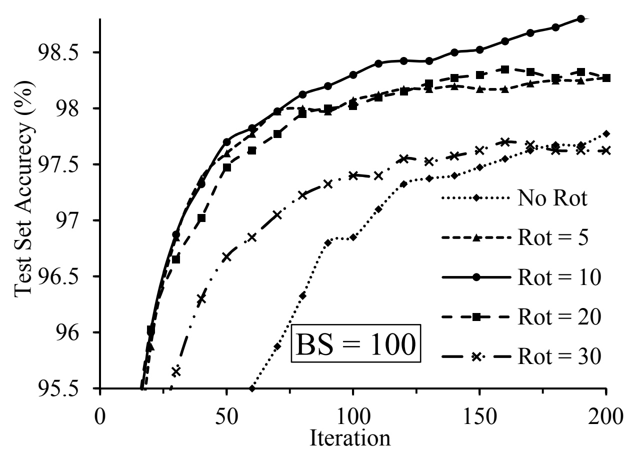


Fig.5. Test set accuracy of proposed CNNAP with different rotational angle for artificial pattern generation.

TABLE I. CONFUSION MATRIX PRODUCED FOR TEST SAMPLES OF BANGLA HANDWRITTEN NUMERALS. TOTAL SAMPLES ARE 4000 HAVING 400 FOR EACH NUMERAL.

Bangla Numeral (Corresponding English Numeral)	Total samples of a particular numeral classified as									
	০	১	২	৩	৪	৫	৬	৭	৮	৯
০ (0)	398	1	0	0	0	0	0	1	0	0
১ (1)	0	386	1	1	2	0	3	1	1	5
২ (2)	0	0	400	0	0	0	0	0	0	0
৩ (3)	0	0	0	399	0	0	0	0	0	1
৪ (4)	0	0	0	0	400	0	0	0	0	0
৫ (5)	1	0	0	0	2	392	2	3	0	0
৬ (6)	0	0	0	1	0	2	397	0	0	0
৭ (7)	0	0	0	0	1	0	0	399	0	0
৮ (8)	0	1	0	0	0	0	1	0	398	0
৯ (9)	0	8	0	0	0	0	0	2	0	390

from 5 to 30 degree. Training with only original pattern was also presented (mark as 'No Rot') as of Fig.4 for better understanding. It is observed from the figure that accuracy improves with rotational angle up to a certain level and accuracy with larger rotational angle based artificial pattern is found to be worse than training without artificial pattern. For BS=100, the test set accuracy with artificial pattern by Rot=10 is much better than accuracy without artificial pattern. For the same BS=100, accuracy for Rot=30 is worse than accuracy without artificial pattern. Worse performance for larger rotational angle is logical because in that cases the generated patterns are more dislike with original and coincide with patterns of other numerals. According to the results presented in the Fig. 5, the best test set recognition accuracy was 98.98% (misclassifying 41 cases out of total 4000 test patterns) at iteration 310 for BS = 100. At that point, the method misclassified 128 cases out of 18000 training patterns showing accuracy rate of 99.29%.

TABLE II. SAMPLE HANDWRITTEN IMAGES THAT ARE MISCLASSIFIED BY CNNAP.

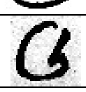
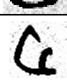
Handwritten Numeral Image	True Numeral	Image Classified as
	১	৯
	৫	০
	৬	৫
	৬	৫
	৯	১
	৯	১

Table I shows the confusion matrix of test set samples at best test set recognition point. From the table it is observed that the proposed method worst performed for the numerals "১" and "৯"; system truly classified 386 and 390 cases for "১" and "৯", respectively out of 400 test cases. Among the Bengali numerals, these two numerals seem to be most similar even in printed form. Numeral "১" recognized as "৯" in five cases; on the other hand "৯" recognized as "১" in eight cases. Similarly, in the Bengali handwritten numeral script, "৫" and "৬" looks similar; therefore in four cases system misclassified them as one another. It is notable that diverse writing styles enhance confusion in several numerals. But the proposed method shows best performance for "২" and "৪" numerals truly classifying all 400 test samples of each numeral. Table II 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 III compares the outcome of the proposed CNNAP 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 CNNAP method is shown to outperform the existing methods. According to the table, CNNAP achieved testing accuracy of 98.98%, 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 artificial pattern in training CNN is clearly understandable comparing the

TABLE III. A COMPARATIVE DESCRIPTION OF PROPOSED CNNAP WITH EXISTING METHODS FOR BANGLA HANDWRITTEN NUMERAL RECOGNITION.

Ref. work and year	Feature Selection	Classification	Recog. Accuracy
Pal et al. [2], 2006	Water overflow from the reservoir	Binary decision tree	92.80%
Basu et al. [6], 2005	Shadow feature and Centroid feature	MLPs with Dempster-Shafer	95.10%
Wen et al. [7], 2007	Principal comp. analysis (PCA) and Kernel PCA	SVM	95.05%
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 CNNAP	No	CNN	98.98%

accuracy of the proposed method with the work of [12] which used only original patterns to train CNN. Finally, the result compared in the table clearly revealed the efficacy of the proposed CNNAP and identified the effectiveness of rotation based artificially generated pattern to improve CNN performance for Bangla handwritten numeral.

IV. CONCLUSIONS

Handwritten numeral classification is a high-dimensional complex task and Convolutional neural network (CNN) is efficient for image classification with its distinct features. Artificial pattern based CNN method has been investigated in this study for handwritten Bangla numeral recognition. A moderated pre-processing and simple rotation has been adopted to produce artificial patterns. CNN is trained with the original patterns plus generated artificial patterns. 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. Hassan and H. A. Khan, "Handwritten Bangla numeral recognition using Local Binary Pattern," in Proc. of International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT2015), Dhaka, Bangladesh, pp. 1-4, May 21-23, 2015.

[19] H. A. Khan, A. A. Helal and K. I. Ahmed, "Handwritten Bangla digit recognition using Sparse Representation Classifier," in Proc. of International Conference on Informatics, Electronics & Vision (ICIEV2014), Dhaka, Bangladesh, pp. 1-6, May 23-24, 2014.