2nd Int'l Conf. on Electrical Engineering and Information & Communication Technology (ICEEICT) 2015
Jahangirnagar University, Dhaka-1342, Bangladesh, 21-23 May 2015

Bangla Handwritten Numeral Recognition using Convolutional Neural Network

M. A. H. Akhand, Md. Mahbubar Rahman, P. C. Shill, Shahidul Islam
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. Although Bangla is a major language in Indian subcontinent and is the first language of Bangladesh study regarding Bangla handwritten numeral recognition (BHNR) is very few with respect to other major languages such Roman. The existing BHNR methods uses 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. It also automatically provides some degree of translation invariance. In this paper, a CNN based BHNR is investigated. The proposed BHNR-CNN normalizes the written numeral images and then employ CNN to classify individual numerals. It does not employ any feature extraction method like other related works. 17000 hand written numerals with different shapes, sizes and variations are used in this study. The proposed method is shown satisfactory recognition accuracy and outperformed other prominent exiting methods.**

Keywords—Bangla Numeral; Convolutional Neural Network 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]. Research on recognition of unconstrained handwritten numerals has made impressive progress in Roman, Chinese and Arabic script [1, 2, 3]. On the other hand, recognition of handwritten Bangla numeral is largely neglected 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. At first, they used two different methods of feature extraction: i) boundary extraction that does the extraction of a numeral in a single window by horizontal and vertical scanning; and ii) scaling is done to convert the image into fixed sized matrix. Then Multi-Layer Perceptron (MLP) are evolved for recognition. Basu et al. [3] 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. [2] introduced a new 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. [6] 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). Recently, Wen and He [7] proposed a kernel and Bayesian Discriminant based method to recognize handwritten bangla numeral. Most recently, Nasir and Uddin [8] 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.

In this paper, Convolutional Neural Network (CNN) [9] based Bangla handwritten numeral recognition (BHNR) is investigated. Recently, CNN is found efficient for image classification with its distinct features. CNN automatically provides some degree of translation invariance. Handwritten numeral images are normalized first and then employ CNN to classify individual numeral. It does not employ any feature extraction method like other related works. We used 17000 hand written numerals with different shapes and variations. Experimental studies reveal that the proposed CNN based method shows satisfactory classification accuracy and outperformed some other exiting methods.

The rest of the paper is organized as follows. Section II explains proposed BHNR using CNN which contains dataset set preparation, preprocessing and classification using CNN. Section III presents experimental results of the proposed method and compare performance with other related works. Finally, a brief conclusion of the work is given in Section IV.

978-1-4673-6676-2/15/$31.00 ©2015 IEEE

II. BANGLA HANDWRITTEN NUMERAL RECOGNITION USING CNN (BHNR-CNN)

This section explains proposed BHRR-CNN in detail which has two major steps: preprocessing of raw images of numerals and classification using CNN. The following subsection gives brief description of each steps. At first it explain Dataset preparation for better understanding.

A. Dataset Preparation

Here we prepared a moderately large handwritten dataset for Bangla isolated numerals. For handwritten scripts, we have considered around several individuals from different ages and education levels. The dataset contains 17000 images with wide variation of distinct numeral images because of different peoples writing styles. Fig.1 shows few sample images of each numeral.

B. Preprocessing of Raw Numeral Images

Preprocessing cleans the arbitrary images into common shape or form that makes appropriate to feed into classifiers. At first handwritten numerals are scanned and produces gray scale image files. In a grayscale image, each pixel value is a single integer number (from 0 to 255) that represents the brightness of the pixel. Typically white pixel has value 255 whereas black pixel has value 0. The image files even for a numeral are often found different sizes for different persons. The arbitrary images are resized into 28×28 dimension to maintain appropriate and equal inputs for all the numerals.

Since we considered black color for writing on white paper (background), the grayscale image files contains more white pixels than black for writing. To reduce computational overhead, images are converted through foreground numeral black to white and background changed to black.

C. Classification using CNN

Handwritten numeral classification is a high-dimensional complex task and traditional MLP require much computation to work with grayscale image. Therefore, a number of traditional methods [5] first extract features from the input image and then use MLP based methods for classification task. On the other hand, CNN itself extract features from the input image or speech signals through its convolution operation [10]. Moreover, CNN has ability to perform right operation on invariance to scaling, rotation and other distortions. Therefore, CNN is considered for classification for Bangla handwritten numeral in this study.

CNN automatically obtains the relevant features like invariance to translation, rotation by forcing the replication of weight configurations of one layer to a local receptive field in the previous layer. Thus, a feature map is obtained in the next layer. By reducing the spatial resolution of the feature map, a certain degree of shift and distortion invariance is also

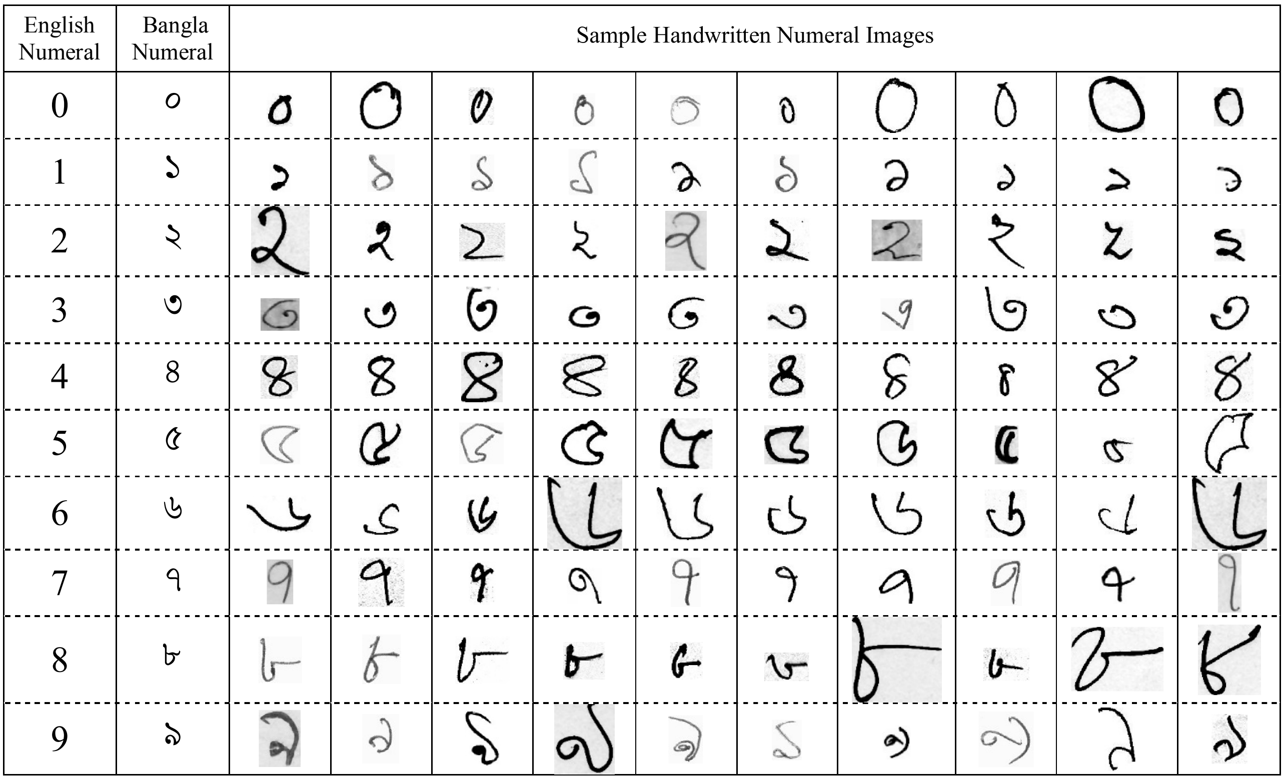


Fig.1. Samples of handwritten Bangla numerals from the dataset.

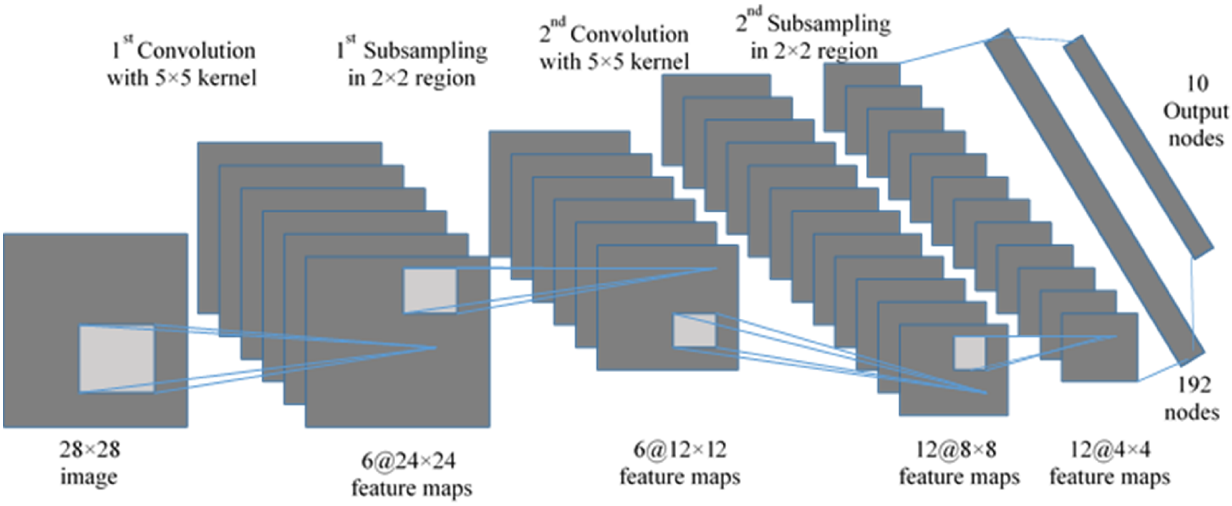


Fig. 2. Structure of CNN for proposed BHNR-CNN.

achieved. Also, the number of free parameters is significantly decreased by using the same set of weights for all features in the feature map.

Figure 2 shows CNN structure of this study for classification Bangla handwritten numeral that holds two convolutional layers with 5×5 receptive fields (i.e., kernel) and two subsampling layers with 2×2 averaging area with input and output layers. Input layer contains 784 nodes for 28×28 pixels image. 1st convolutional operation produces first level six feature maps. Distinct kernel having different weights and biases from other kernels are used to produce a 1st level feature map so that it can extract different types of local features. Convolution operation with kernel spatial dimension 5 converts 28 spatial dimension to 24 (i.e., 28-5+1) spatial dimension [11]. Therefore, each 1st level feature map size is 24×24. This local receptive field can extract the visual features such as oriented edges, end-points, corners of the images. Higher order features are obtained by combining those extracted features.

In 1st subsampling layer, the 1st level feature maps are down-sampled from 24×24 into 12×12 feature maps by applying a local averaging with 2×2 area, multiplying by a coefficient, adding a bias and passing through an activation function. More formally it can be shown as follows:

$$x_j^l = f(\beta_j^l down(x_j^{l-1}) + b_j^l), \qquad (1)$$

where *f(.)* is the activation function; *down(.)* represents a subsampling function through local averaging; β and b are multiplicative coefficient and additive bias, respectively. Sigmoid is commonly used as activation function and is considered in this study. For 2×2 local area averaging in *down(.)* function, the output image becomes 2-times smaller in both spatial dimensions. This subsampling operation reduces both the spatial resolution of the feature map and sensitivity to shift and distortions.

Second convolution and 2nd subsampling operations are similar to 1st convolution and 1st subsampling operations, respectively. 2nd convolutional operation produces distinct 12 feature maps; a receptive field size of 5×5 produces a feature map size of 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 (=12×4×4) distinct nodes those are fully connected to 10 feature maps (the output nodes) for numeral set.

III. Results and Discussions

Experimental results of the proposed recognition scheme have been collected based on the samples of the prepared dataset discussed earlier. 17000 samples divided into 13000 and 4000 samples, for training and testing, respectively. Training samples are evenly distributed over the underlying 10 classes. The recognition performance reported in this paper are based on the test set accuracies. In the test set, equal number of samples (i.e., 400) for each numeral were considered.

We applied CNN on the resized and normalized grayscale image files without any feature extraction technique. The method is implemented in Matlab2013a. The experiment has been conducted on HP pro desktop machine (CPU: Intel Core i5 @ 3.20 GHz and RAM: 4.00 GB) in Window 7 environment.

We have observed classification accuracy of the system for various fixed number of iterations and it is observed that minimum 50 iterations are required for suitable performance. Table I shows the confusion matrix of test set samples after fixed 80 iterations. From the table it observed that the proposed method worst performed for the numeral “৩” and 382 cases it classified truly out of 400 test cases. In five cases this numeral classified as “০”. Both “০” and “৩” have circle

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

English Numeral	Bangla Numeral	Total samples of a particular numeral classified as									
		০	১	২	৩	৪	৫	৬	৭	৮	৯
0	০	397	0	1	0	0	0	1	0	0	1
1	১	0	394	1	1	3	1	0	0	0	0
2	২	0	0	400	0	0	0	0	0	0	0
3	৩	5	1	1	382	2	1	6	0	2	0
4	৪	0	0	0	0	397	2	0	0	1	0
5	৫	3	0	4	1	3	386	1	2	0	0
6	৬	0	2	0	7	0	8	383	0	0	0
7	৭	0	0	1	0	0	1	0	395	1	2
8	৮	0	0	0	0	0	1	0	0	399	0
9	৯	1	3	5	4	1	0	0	2	0	384

like ambiguity in handwritten form. In the Bangla handwritten numeral script, "৫" and "৬" looks very similar; therefore in eight cases "৬" classified as "৫". Similarly, the numeral "৯" classified as "১" and "২" in three and five cases, respectively; it is clearly observed confusion in several handwritten images of the three numerals due to diverse writing styles of individuals. But the proposed method is shown best performance for "২" and correctly classifying all 400 test samples.

TABLE II. SAMPLE HANDWRITTEN CHARACTERS THAT MISCLASSIFIED BY BHNR-CNN.

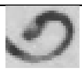
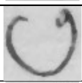
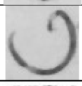
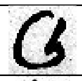
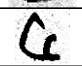
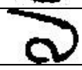
Handwritten Numeral Image	True Numeral	Image Classified as
	৩	০
	৩	০
	৩	০
	৫	০
	৬	৫
	৬	৫
	৯	১

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. Finally, the proposed BHNR-CNN misclassified 83 test samples out of 4000 test cases and achieved accuracy 97.93%. On other hand, the method misclassified only 78 cases out of 13000 training samples showing accuracy rate 99.40%. This indicates that there is a chance to improve CNN training and get better performance with the proposed method.

Table III compares the outcome of the proposed 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 methods. Without feature selection, proposed BHNR-CNN method is shown to

TABLE III. A COMPARATIVE DESCRIPTION OF PROPOSED BHNR-CNN WITH SOME CONTEMPORARY METHODS.

The work reference	Feature Selection	Classification	Recognition Accuracy
Pal et al. [2]	Water overflow from the reservoir based feature selection	Binary decision tree	92.80%
Wen et al. [6]	Principal component analysis (PCA) and Kernel PCA	SVM	95.05 %
Basu et al. [3]	Shadow feature and Centroid feature	MLPs with Dempster-Shafer technique	95.10%
Proposed BHNR-CNN	No	CNN	97.93 %

outperform the exiting methods. According to the table, BHNR-CNN achieved testing accuracy 97.93%, on the other hand, the testing accuracy are 92.80% and 95.05% for the works of [2] and [6], respectively. Although performance compared here are for different datasets, the efficacy of the proposed BHNR-CNN is quite interesting and identified the ability of CNN based classifier for Bangla handwritten numeral.

IV. Conclusions

Convolutional neural network (CNN) has ability to recognize visual patterns directly from pixel images with minimal preprocessing. Therefore, a CNN structure is investigated without any feature selection for Bangla handwritten numeral pattern classification in this study. The method has been tested on a large hand written numeral dataset and outcome compared with existing prominent methods for Bangla. The proposed method is shown to outperform the exiting methods on the basis of test set accuracy. Moreover, the proposed scheme seems efficient in size and computation.

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] 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.

[3] 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

[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] 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.

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

[8] 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.

[9] Y. Lecun and Y. Bengio, "Pattern Recognition and Neural Networks", in Arbib, M. A. (Eds), The Handbook of BrainTheory and Neural Networks, MIT Press 1995.

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

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

[12] Table II. A Comparative Description of Proposed BHNR-CNN with Some Contemporary Methods.