Modern Education
and Computer Science
PRESS

Bangla Handwritten Character Recognition using Convolutional Neural Network

Md. Mahbubar Rahman, M. A. H. Akhand, Shahidul Islam, Pintu Chandra Shill
Dept. of Computer Science and Engineering
Khulna University of Engineering & Technology Khulna, Bangladesh
Email: {akhand, pintu} @cse.kuet.ac.bd

M. M. Hafizur Rahman
Dept. of Computer Science, KICT, International Islamic University Malaysia Selangor, Malaysia
Email: hafizur@iium.edu.my

Abstract—Handwritten character recognition complexity varies among different languages due to distinct shapes, strokes and number of characters. Numerous works in handwritten character recognition are available for English with respect to other major languages such as Bangla. Existing methods use distinct feature extraction techniques and various classification tools in their recognition schemes. Recently, Convolutional Neural Network (CNN) is found efficient for English handwritten character recognition. In this paper, a CNN based Bangla handwritten character recognition is investigated. The proposed method normalizes the written character images and then employ CNN to classify individual characters. It does not employ any feature extraction method like other related works. 20000 handwritten characters with different shapes and variations are used in this study. The proposed method is shown satisfactory recognition accuracy and outperformed some other prominent exiting methods.

Index Terms—Handwritten Character Recognition, Bangla, Convolutional Neural Network.

I. INTRODUCTION

In recent years, there has been much interest in automatic character recognition. Between handwritten and printed forms, handwritten character recognition is more challenging. Handwritten characters written by different persons is not identical but varies in both size and shape. Numerous variations in writing styles of individual character make the recognition task difficult. The similarities in distinct character shapes, the overlaps, and the interconnections of the neighboring characters further complicate the problem. A handwritten character recognition system consists of two major steps: feature extraction from the character set and then employ learning tool(s) to classify individual character [1-4].

With distinct feature extraction techniques, a number of methods based on artificial neural network are investigated for handwritten English character recognition. Morphological /Rank/Linear Neural Network (MRL-NN) [1], in which the combination of inputs in

every node is formed by hybrid linear and nonlinear (of the morphological/rank type) operations, is investigated for handwritten digit. A hybrid Multilayer Perceptron-Support Vector Machine (MLP-SVM) based method was used for English digit [2] and Chinese character [3] recognition. Support Vector Machine (SVM) with Radial Basis Function (RBF) network is used in Ref. [4] for both small and capital handwritten English character set.

Recently, Convolutional Neural Network (CNN) [5] is found efficient for handwritten character recognition due to its distinct features. CNNs add the new dimension for image classification systems and recognizing visual patterns directly from pixel images with minimal preprocessing. In addition, CNN automatically provides some degree of translation invariance. A CNN based model was tested on UNIPEN [6] English character dataset and found recognition rates of 93.7% and 90.2% for lowercase and uppercase characters, respectively [7]. Moreover, CNN committees are also found to enhance the recognition performance where the committees are formed by varies aspect ratios on characters images to same architecture [8].

Character recognition complexity varies among different languages due to distinct shapes, strokes and number of characters. A number of works, including the above discussed methods, is available for English with respect to other major languages such as Bangla. Bangla is one of the most spoken languages, ranked fifth in the world. It is also an important language with a rich heritage; 21st February is declared as the International Mother Language day by UNESCO to respect the language martyrs for the language in Bangladesh at the year of 1952. Bangla is the first language of Bangladesh and the second most popular language in India. About 220 million people use Bangla as their speaking and writing purpose. There are 50 characters in Bangla and some contains additional sign up and/or below. Moreover, Bangla contains many similar shaped characters; in some cases a character differ from its similar one with a single dot or mark. That makes difficult to achieve better performance with simple technique as well as hinders to work with Bangla handwritten character recognition.

A few notable works are available for Bangla

handwritten character recognition. Bhowmik et al. [9] proposed a fusion classifier using Multilayer Perceptron (MLP), RBF network and SVM. They used wavelet transform for feature extraction from character images. In classification, they considered some similar characters as a single pattern and trained the classifier for 45 classes. Basu et al. [10] proposed a hierarchical approach to segment characters from words and MLP is used for classification. In segmentation stage they used three different feature extraction techniques but they reduced character patterns into 36 classes merging similar characters in a single class. Recently, Battacharya et al. [11] considered a two-stage recognition scheme for 50 basic character classes. Feature vector for the first classifier is computed by overlaying a rectangular grid consisting of regularly spaced horizontal and vertical lines over the character bounding box. The response of this first classifier is analyzed to identify its confusion between a pair of similar shaped characters. Second stage of classification is used to resolve the confusion and feature vector is computed by overlaying another rectangular grid but consisting of irregularly spaced horizontal and vertical lines over the character bounding box. They used Modified Quadratic Discriminant Function (MQDF) classifier and MLP as classifiers in first and second stages, respectively.

In this paper, Convolutional Neural Network (CNN) [5] based Bangla handwritten character recognition is investigated. The proposed method first normalizes the written character images and then employ CNN to classify individual characters. It does not employ any feature extraction method like other related works. Handwritten characters of different persons having different shapes and variations are used in this study. 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 Bangla handwritten character recognition using CNN which contains dataset set preparation, preprocessing and classification using CNN. Section III presents experimental results of the proposed method and compares performance with other related works. Finally, a brief conclusion of the work is given in Section IV.

II. BANGLA HANDWRITTEN CHARACTER RECOGNITION USING CNN(BHCR-CNN)

This section explains proposed BHCR-CNN in detail which has two major steps: preprocessing of raw images of characters 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

We prepared a moderately large handwritten dataset for 50 Bangla isolated characters. Fig.1 shows the characters in printed form with English spelling, among them 39 are consonants and 11 are vowels. For handwritten character scripts, we have considered around 30 individuals from different ages and education levels. Our prepared dataset size is 20000 having 400 samples for each character. The dataset contains wide variation of distinct characters because of different peoples' writing styles. Some of these character images are very complex shaped and closely correlated with others. Fig. 2(a) shows several samples handwritten character images for first character "অ" that shows variations in size, shape and orientation.

B. Preprocessing of Raw Charatcer Images

Preprocessing cleans the arbitrary images into common shape or form that makes appropriate to feed into classifiers. At first handwritten characters 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 character 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 characters.

Since we considered black color for writing on white paper (background), the grayscale image files contain more white pixels than black for writing. To reduce computational overhead, images are converted through foreground character black to white and background changed to black. Fig. 2 shows sample figures of character "অ" in two different forms. It is observed in Fig. 2(a) that the raw images for same "অ" are quite different in styles besides sizes. Fig. 2(b) is resized and foreground-background interchanged images that is used as input of CNN.

C. Classification using CNN

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
অ	আ	ই	ঈ	উ	ঊ	ঋ	এ	ঐ	ও	ঔ	ক	থ	গ	ঘ	ঙ	চ
A	AA	I	II	U	UU	R	E	AI	O	AU	KA	KHA	GA	GHA	NGA	CA

18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
ছ	জ	ঝ	ঞ	ট	ঠ	ড	ড়	ণ	ত	থ	দ	ধ	ন	প	ফ	ব
CHA	JA	JHA	NYA	TTA	TTHA	DDA	RRA	NNA	TA	THA	DA	DHA	NA	PA	PHA	BA

35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
ভ	ম	য	র	ল	শ	ষ	স	হ	ড়	ঢ়	য়	ৎ	ং	ঃ	ঁ
BHA	MA	YY	RA	LA	SHA	SSA	SA	HA	DDHA	DHRA	YYA	KHAND	ANUS	VISARG	BINDU

Fig. 1. Bangla basic characters: first 11 characters are vowels and rest 39 are consonants. Name of each character is provide below the corresponding Bangla character symbol.

(a)   (b)
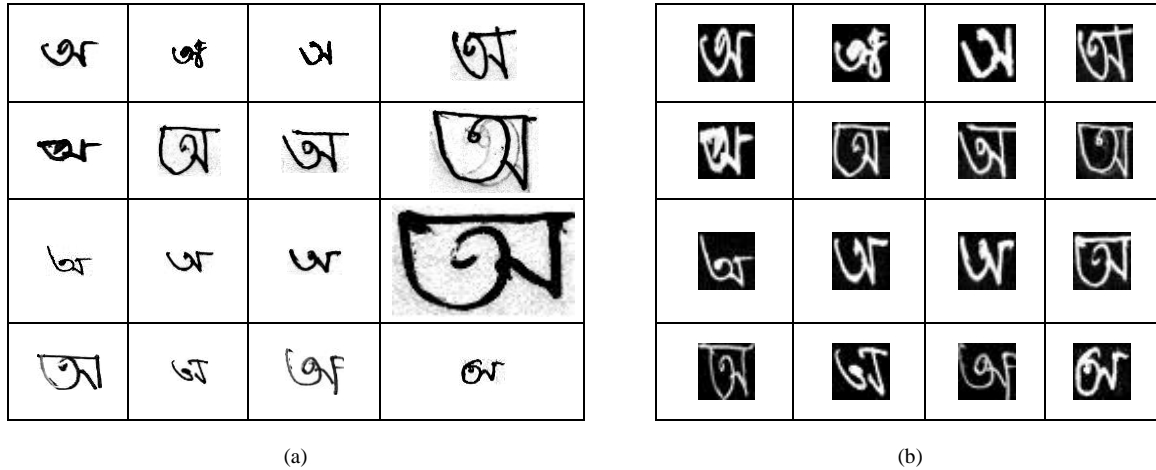
Fig. 2. Sample figures of character "অ": (a) raw images and (b) corresponding resized (28×28) and foreground-background interchanged images.

Handwritten character classification is a high-dimensional complex task and traditional MLP require much computation to work with grayscale image. Therefore, a number of traditional methods [9-11] 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 [12]. 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 character in this study as it is found to perform well for English.

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 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 3 shows CNN structure of this study for classification Bangla handwritten characters 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.
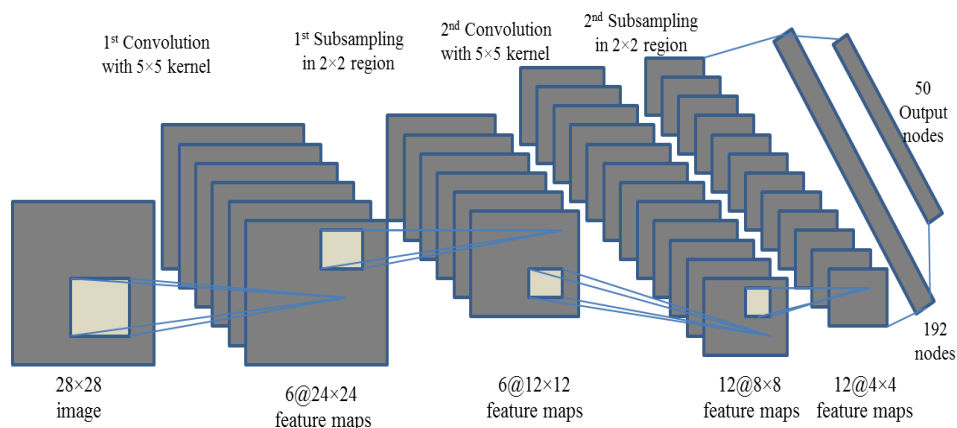
Fig. 3. Structure of CNN for proposed BHCR-CNN.

Convolution operation with kernel spatial dimension 5 converts 28 spatial dimension to 24 (i.e., 28-5+1) spatial dimension [13]. 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:

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

where down(.) represents a subsampling function through local averaging; β and b are multiplicative coefficient and additive bias, respectively. The output image becomes 2-times smaller in both spatial dimensions for 2×2 local area averaging in down(.) function. 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 12 distinct 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 50 feature maps (the output nodes) for character set. For total 50 character set, each output node represents a particular character and the desired value of the node was defined as 1 (and other 49 desired output nodes value as 0) for the input set of the pattern. Finally, the Error (E) that minimizes the CNN is

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

where P is the total number of pattern (here 17500); O is the total output nodes (i.e., 50); d_o and y_o are the desired and actual output of a node for a particular pattern p.

III. Results and Discussions

Experimental results of the proposed recognition scheme have been collected based on the samples of the prepared dataset discussed earlier. From 400 samples for a character, 350 samples have been randomly selected for training and rest 50 samples were reserved for testing. Therefore, training and test sets contained 17500 and 2500 samples, respectively. Training samples are evenly distributed over the underlying 50 classes. The recognition performance reported in this paper are based on the test set accuracies.

We applied CNN on the resized grayscale image files without any feature extraction technique. The CNN algorithm 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 OS environment.

The batch wise training performed in this study due to large sized training set. Each resized image that used in training consists of 784 (28 × 28) pixel values; and therefore, the training set with 17500 images cause memory overflow when consider whole set at a time. Therefore, number of Batch Size (BS) is considered as a user defined parameter. On the other hand, Learning Rate (LR) is also an element that influences learning. Figure 4 presents training error (E) for different LR values for a fixed BS = 10. From the figure it is observed that at the error reduces rapidly up to 50 iteration and after 150 iteration curses goes to saturation for any LR value. However lower E is shown to achieve for LR = 1.0. Table 1 presents the summary of the experimental results for 200 iterations. It is observed from the table that better training set and test set classification accuracy are achieved for LR =1.0. For LR=10, after 200 iteration the recognition accuracy for training and test sets were 93.93 % and 85.36%, respectively. On the basis of observation from the Fig. 4, the LR value 1.0 seems effective for training.
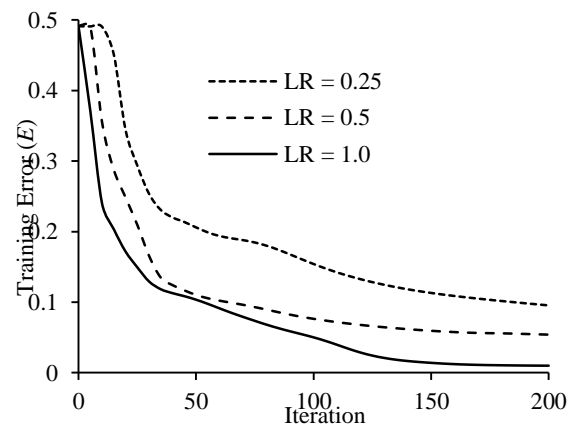
Fig. 4. Effect of Learning Rate (LR) on Training Error (E).

Table 1. Performance evaluation for different Learning Rate (LR) with BS = 10 after 200 iterations.

Learning Rate	Training Error (E)	Training Set Rec. Accuracy	Test Set Rec. Accuracy
0.25	0.095406	91.40%	85.28%
0.50	0.053969	92.73%	84.20%
1.0	0.009958	**93.93%**	**85.36%**

An experiment conducted to observe the effect of batch size on the performance of BHCR-CNN. Fig. 5 presents required time in minutes and training error (E) for three different batch sizes (BS) of 5, 10 and 20. For the experiment the LR values was considered as 1.0. It is observed from the Fig. 5(a) that batch size has a remarkable effect on training time. In general, an iteration required for BS = 5 is near about double of BS = 10 and three times of BS = 20. Table 2 presents the summary of the experimental results for 200 iterations. To complete 200 iterations, BHCR-CNN required training time of

455.01, 267.70 and 175.38 minutes for batch sizes of 5, 10 and 20, respectively. For larger batch size, operation performed on a large number of training patterns at a time and took relatively smaller time to train.

Figure 5(b) shows the effect of batch size on training error (E). For any batch size error reduced rapidly in few initial iterations (e. g., up to 50) and after that error minimization with iteration was not significant. However lower number of BS seems to minimize error rapidly although it required larger training time as of Fig. 5(a). It is also observed for the figure that for $BS = 5$ the error E is lower than that of BS values 10 or 20.

(a) Elapsed Time vs. Iteration

(b) Training Error vs. Iteration
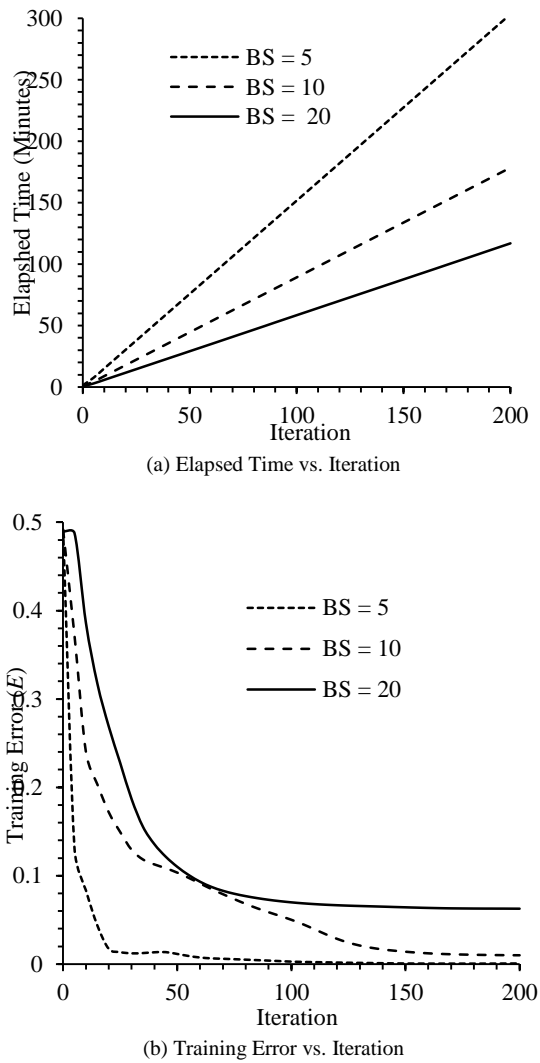
Fig. 5. Effect of Batch Size (BS) on Required Time and Training Error (E).

Table 2 presents the summary of the experiment results for 200 iterations. It is observed from the table that the best training set and test set classification accuracy is observed for $BS = 10$ and the values are 93.93% and 85.36%, respectively. It is notable from the figure that lower error might not give better classification accuracy. Therefore, it is require to check with different iteration to achieve better classification accuracy.

Table 2. Performance evaluation of different Batch Size (BS) with LR=1.0 after 200 iterations.

Batch Size	Time in Minutes	Training Error (E)	Training Set Rec. Accuracy	Test Set Rec. Accuracy
5	455.01	0.000678	93.43 %	84.08 %
10	267.70	0.009958	**93.93 %**	**85.36 %**
20	175.38	0.062655	92.98 %	85.16 %

Figure 6 shows the training set and test set classification (i.e., recognition) accuracy for different learning rates for batch size 10. For the experiment iteration was varied from 50 to 300. It is observed from the figure that training set classification is increased with iteration for any LR value. On the other hand, testing set accuracy did not improve with training set accuracy after certain iteration. The accuracy on test is more desirable that indicates the generalization ability of a system. The best test set recognition accuracy 85.96 % is achieved for $LR = 1.0$ at 225 iteration.

(a)

(b)
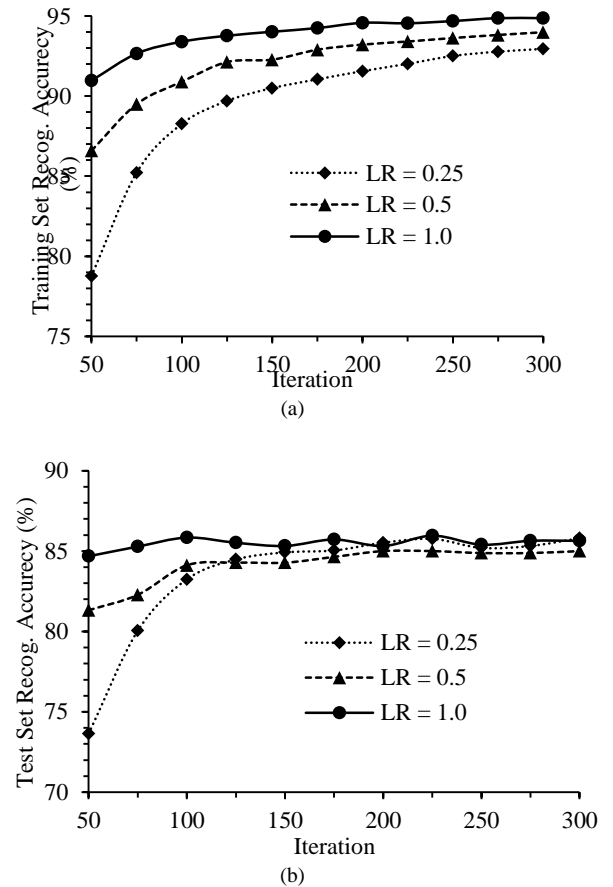
Fig. 6. Training and Test Recognition Accuracy for different Iteration with BS=10.

Table 3 shows the confusion matrix of test set samples for $LR = 1.0$ at 225 iteration. In the figure 1-50 are the character index as of Fig. 1. From the table it has been observed that the proposed method worst performed for the character "ষ" and only 27 cases it classified truly out of 50 test cases. 14 cases of this character is classified as "শ" that looks similar even in printed form and more

Table 3. Confusion matrix produced for test samples of Bangla handwritten characters.

		1 অ	2 আ	3 ই	4 ঈ	5 উ	6 ঊ	7 ঋ	8 এ	9 ঐ	10 ও	11 ঔ	12 ক	13 খ	14 গ	15 ঘ	16 ঙ	17 চ	18 ছ	19 জ	20 ঝ	21 ঞ	22 ট	23 ঠ	24 ড	25 ঢ	26 ণ	27 ত	28 থ	29 দ	30 ধ	31 ন	32 প	33 ফ	34 ব	35 ভ	36 ম	37 য	38 র	39 ল	40 শ	41 ষ	42 স	43 হ	44 ড়	45 ঢ়	46 য়	47 ৎ	48 ং	49 ঃ	50 ঁ
1	অ	41	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
2	আ	0	44	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0
3	ই	0	0	41	1	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
4	ঈ	0	0	2	42	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	উ	0	0	0	0	38	4	0	0	0	2	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	2	0	0	0	0	0	1	0	0	0	1	0
6	ঊ	0	0	0	0	2	48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	ঋ	0	1	0	0	0	0	44	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	এ	0	0	0	0	0	0	0	44	1	1	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	ঐ	1	0	0	1	0	0	0	0	45	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	ও	0	0	0	0	0	0	0	0	0	44	0	0	0	0	0	3	0	0	0	0	0	0	0	0	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	ঔ	0	0	1	0	1	0	0	0	3	1	41	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	ক	0	0	0	0	0	0	0	0	0	0	0	48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	খ	1	0	0	0	0	0	0	0	0	0	0	0	47	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	গ	0	0	0	0	0	0	0	0	0	0	0	0	0	47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	ঘ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	ঙ	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
17	চ	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	45	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
18	ছ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
19	জ	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	44	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	ঝ	0	0	0	0	0	0	1	0	0	0	3	0	0	0	0	0	0	0	0	45	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	ঞ	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	ট	0	0	0	2	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	42	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
23	ঠ	0	0	1	0	1	0	0	0	1	0	1	0	0	0	0	1	0	1	0	1	0	39	0	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	ড	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	45	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	ড়	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	46	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
26	ণ	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	41	0	1	0	0	4	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
27	ত	2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
28	থ	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	39	0	0	1	2	1	0	0	3	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
29	দ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	1	0	0	0	0	45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	ধ	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	38	0	0	0	1	0	0	1	0	0	4	0	0	0	1	0	0	1	0	0	0	0	0	0	0
31	ন	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1	0	0	0	0	5	0	0	0	39	0	0	1	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
32	প	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	1	1	0	0	0	1	0	2	0	40	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	ফ	0	0	0	0	0	0	0	1	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	ব	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	47	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	ভ	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	1	0	0	0	0	0	0	3	0	0	0	0	0	0	0	41	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
36	ম	0	0	0	0	0	0	0	0	2	2	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	41	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	য	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	14	1	0	0	0	1	0	0	1	0	1	27	0	0	1	0	2	0	0	2	0	0	0	0	0
38	র	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	47	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
39	ল	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	3	2	0	0	0	0	0	0	41	1	0	0	0	0	0	0	0	0	0	0	0	0
40	শ	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	45	0	0	0	0	0	0	0	0	0	0	0
41	ষ	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	1	0	0	42	0	0	0	0	1	0	0	0	0	0
42	স	2	1	0	0	0	0	1	0	2	0	0	1	0	6	4	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	3	0	0	0	0	0	0	0	0	0	26	0	0	0	0	0	0	0	0	0
43	হ	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	42	0	0	0	0	1	1	1
44	ড়	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	5	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	38	1	0	0	0	0	0	0
45	ঢ়	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	43	0	0	0	0	0	1
46	য়	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	49	0	0	0	0
47	ৎ	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45	1	0	0
48	ং	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	1	0	0	46	0	0	0
49	ঃ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	48	0
50	ঁ	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	44

Table 4. Sample handwritten characters that misclassified by BHCR-CNN.

Handwritten Image	Character Classified as	True Character
	ত	অ
	ঙ	ঔ
	হ	দ
	থ	ঘ
	প	ণ
	ট	ঊ
	য	য়

difficult in handwritten form. Similarly the character "ন" classified as "ণ" in five cases. Due to such ambiguity, some exiting methods excluded such characters from recognition [9]. But the proposed method is shown best performance for "য়" only misclassifying one case as "য" that is excluded in some previous study. Table 4 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 BHCR-CNN misclassified 351 cases out of 2500 test cases and achieved accuracy 85.96%. On other hand, the method misclassified 954 characters out of 17500 training characters showing accuracy rate 94.55%.

Table 5 compares the outcome of the proposed method with other prominent works of Bangla handwritten character 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 single or two stages feature

selections. Moreover, the methods in Refs. [9] and [10] considers 45 and 36 classes merging or excluding some confusing character. Without feature selection and with 50 classes, proposed BHCR-CNN method is shown to outperform the MLP based methods of Ref. [9] and Ref. [10]. According to the table, BHCR-CNN is found inferior to the work of Ref. [11]. The recognition techniques that uses Ref. [11] is much complex than others; it uses two recognition stages each one consists of individual feature selection and classification techniques. Besides this, the proposed single stage method without feature selection is very simple. Moreover, since reported performance are for different datasets, the performance of proposed simple method is quite interesting and identified the ability of CNN based classifier for Bangla handwritten character.

Table 5. Recognition performance and structure comparison of proposed BHCR-CNN with some contemporary methods of handwritten Bangla character recognition.

The work reference	Total Classes	Feature Selection	Classification	Recog. Accuracy
Bhowmick et al. [9]	45	Wavelet Transformation	MLP	84.33 %
Basu et al. [10]	36	Longest run, Modified shadow, Octant-centroid	MLP	80.58 %
Bhattacharya et al. [11]	50	Regular and Irregular Grid based Selection	MQDF, MLP	95.84 %
Proposed BHCR-CNN	50	No	CNN	85.96 %

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 pattern classification in this study. The method has been tested on a large handwritten character dataset and outcome compared with existing prominent methods for Bangla. The proposed method is shown competitive performance with the exiting methods on the basis of test set accuracy but the proposed scheme seems efficient in size and computation. The presented result indicates that training of CNN might be improved and hence get better performance.

REFERENCES

[1] L. F. C. Pessoa and P. Maragos, *Neural networks with hybrid morphological/rank/linear nodes: a unifying framework with applications to handwritten character recognition*, Pattern Recognition, vol. 33, no. 6, pp. 945-960, June 2000.

[2] Bellili, M. Gilloux and P. Gallinari, *An MLP-SVM combination architecture for offline handwritten digit recognition,* Document Analysis and Recognition, Springer-Verlag, vol. 5, no. 4, pp. 244-252, 2003.

[3] J. Dong, A. Krzyżak and C. Y. Suen, *An improved handwritten Chinese character recognition system using support vector machine*, Pattern Recognition Letters, vol. 26, no. 12, pp. 1849-1856, September 2005.

[4] G. Vamvakas, B. Gatos and S. J. Perantonis, *Handwritten character recognition through two-stage foreground sub-sampling*, Pattern Recognition, vol. 43, no. 8, pp. 2807-2816, August 2010.

[5] 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.

[6] Guyon, L. Schomaker, R. Plamondon, M. Liberman, and S. Janet, *Unipen project of on-line data exchange and recognizer benchmarks,* in proc. of 12th International. Conference on Pattern Recognition (ICPR), vol. 2, pp. 29–33, IEEE, 1994.

[7] Yuan, G. Bai, L. Jiao and Y. Liu, *Offline handwritten English character recognition based on convolutional neural network*, in 10th IAPR International Workshop on Document Analysis Systems (DAS), pp. 125-129, doi: 10.1109/DAS.2012.61, 2012.

[8] D. C. Ciresan, U. Meier, L. M. Gambardella, and J. Schmidhuber, *Convolutional Neural Network Committees for Handwritten Character Classification,* International Conference on Document Analysis and Recognition (ICDAR), pp. 1135-1139, doi: 10.1109/ICDAR.2011.229, 2011.

[9] T. K. Bhowmik, P. Ghanty, A. Roy and S. K. Parui, *SVM-based hierarchical architec-tures for handwritten Bangla character recognition*, International Journal on Document Analysis and Recognition, vol. 12, no. 2, pp. 97-108, 2009.

[10] S. Basu, N. Das, R. Sarkar, M. Kundu, M. Nasipuri and D. K. Basu, *A hierarchicalapproach to recognition of handwritten Bangla characters*, Pattern Recognition, vol. 42, pp. 1467–1484, 2009.

[11] U. Bhattacharya, M. Shridhar, S. K. Parui, P. K. Sen and B. B. Chaudhuri, *Offline recognition of handwritten Bangla characters: an efficient two-stage approach*, Pattern Analysis and Applications, vol. 15, no. 4 , pp. 445-458, 2012.

[12] 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.

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

Authors' Profiles

Md. Mahbubar Rahman received the B.Sc. degree in Computer Science Engineering (CSE) from Khulna University of Engineering and Technology (KUET), Bangladesh in 2014. At present he is serving as a Lecturer at Bangladesh Army University of Engineering and Technology (BAUET), Qadirabad Cantonment, Natore, Bangladesh.

M. A. H. Akhand received the B.Sc. degree in Electrical and Electronic Engineering from Khulna University of Engineering and Technology (KUET), Bangladesh in 1999, the M.E. degree in Human and Artificial Intelligent Systems in 2006, and the Doctoral degree in System Design Engineering in 2009 from University of Fukui, Japan. He joined as a lecturer at the Department of Computer Science and Engineering at KUET in 2001, and is now a Professor. He is also head of Computational Intelligence Research Group of this department. He is a member of Institution of Engineers, Bangladesh (IEB). His research interest includes artificial neural networks, evolutionary computation, bioinformatics, swarm intelligence and other bio-inspired computing techniques.

Shahidul Islam received the B.Sc. degree in Computer Science Engineering (CSE) from Khulna University of Engineering and Technology (KUET), Bangladesh in 2014. At present he is serving as a Software Engineer in Kona SL in Dhaka, Bangladesh.

Pintu Chandra Shill received the B.Sc. degree in Computer Science Engineering (CSE) from Khulna University of Engineering and Technology (KUET), Bangladesh in 2003, the M.Sc. degree in Computer Engineering from Politecnico di Milano, Italy in 2008 and the Doctoral degree in System Design Engineering in 2009 from University of Fukui, Japan. He joined as a lecturer at the Department of CSE, KUET in 2004 and currently he is serving as an Assistant Professor. He has published several research papers in some reputed Journal and Conference. His research interest includes evolutionary computation, fuzzy logic, bioinformatics and artificial neural networks.

M. M. Hafizur Rahman received his B.Sc. degree in Electrical and Electronic Engineering from Khulna University of Engineering and Technology (KUET), Khulna, Bangladesh, in 1996. He received his M.Sc. and Ph.D. degree in Information Science from the Japan Advanced Institute of Science and Technology (JAIST) in 2003 and 2006, respectively. Dr. Rahman is now an assistant professor in the Dept. of Computer Science, KICT, IIUM, Malaysia. Prior to join in the IIUM, he was an associate professor in the Dept. of CSE, KUET, Khulna, Bangladesh. He was also a visiting researcher in the School of Information Science at JAIST and a JSPS postdoctoral research fellow at Research Center for Advanced Computing Infrastructure, JAIST \& Graduate School of Information Science (GSIS), Tohoku University, Japan in 2008 and 2009 \& 2010-2011, respectively. His current research include parallel and distributed computer architecture, hierarchical interconnection networks and optical switching networks. Dr. Rahman is member of IEICE of Japan and IEB of Bangladesh.