Handwritten Bangla Digit Recognition Employing Hybrid Neural Network Approach

Tahsina Hashem*, Mohammad Asif† and Md. Al-Amin Bhuiyan‡
*†Department of Computer Science and Engineering, Jahangirnagar University, Savar, Dhaka-1342
*Email: tahsinahashem@gmail.com
†Email: asif_pogo@yahoo.com
‡Department of Computer Engineering, King Faisal University, Alahssa 31982, Saudi Arabia
‡Email: mbhuiyan@kfu.edu.sa

***Abstract*—Handwritten Bangla digit recognition is one of the most attractive area for researchers who have interest in image processing and pattern recognition field. In our everyday activities like bank check identification, passport and document analysis, number plate identification and especially in our postal automation service, recognition of handwritten digits plays a significant role. That's why a rich body of literature already has been published in this area. But most traditional techniques are generally based on complex feature extraction approach that introduces a great overhead in recognition tasks. Recently a novel approach Back-propagation algorithm is used for recognition which simplifies the recognition process but the main drawback is, the network takes lots of iterations to converge. This paper addresses a faster and efficient Hybrid Neural Network Solution called (BAM+BPNN) which is a combination of Bidirectional Associative Memory and Back-Propagation Neural Network. BAM is used for dimensional reduction and BPNN is used to train the neural network with the set of input patterns for acquiring separate knowledge of each digit. This research can take a decision that Hybrid Neural Network algorithm (BPNN with BAM) takes less iteration to train and less time to recognize digits than Back-propagation algorithm (BPNN). Experimental study shows the effectiveness of our proposed technique.**

Keywords*—Hybrid Neural Network, Back-propagation, BAM, Handwritten Bangla Digit

I. Introduction

Digit Recognition refers to the process of translating the images of handwritten, typewritten or printed digits into a standard format. So that everyone can understand and use it for different purposes such as editing, searching/indexing and especially in postal automation service. Every literate human has his own manner of writing styles. Many peculiarities may also exist in different writing styles. Thus the recognition process becomes very much complicated by noisy inputs, distorted images and differences between type phases, sizes and fonts.

Bengali is one of the most spoken languages (ranked seventh [1]) in the world and is the native-tongue of 99% citizens of Bangladesh. A number of researches already has been conducted with feature extraction [2]–[7] and neural network [8]–[10] based handwritten Bangla digit recognition problem. Complex feature extraction approaches have showed good performance in digit recognition. But in this process, all the features for each digit are need to be collected, which introduces a great overhead in recognition tasks. Morshed et al. [9] and some other researchers [4], [5], [10] used the novel approach Back-propagation Neural Network algorithm [11], which successfully overcomes the limitations of feature extraction approach, but the major drawback is it requires a large number of iterations to converge.

Recently, Sanjit et al. presented a hybrid neural network solution [12] for Bangla character recognition which combines local image sampling and artificial neural network. We have utilized their idea and modified their algorithm to make our hybrid neural network solution (BAM + BPNN) faster and efficient.

In summary the contribution of this research are as follows:

- A Hybrid Neural Network Solution (***B**iderectional **A**ssociative* **M***emory and* ***B**ack-**P**ropagation **N**eural **N**etwork*) called **(BAM+BPNN)** has been proposed to recognize handwritten Bangla digits. Employing this algorithm we have achieved our main motive of reducing number of iterations as well as making the system robust and efficient.
- The processed image of Bangla digits have been directly taken rather than using problem dependent feature extraction process.
- To avoid loss of information of each digit characteristics, we do not input directly the smaller image to BPNN. Rather than we have used BAM to reduce dimension of the digit image. Then the output of BAM is used as input to BPNN.
- The proposed scheme has effectively used Multiple Training Encoding Strategy to overcome the initial limitations of BAM and later strengthen the training phase.
- A number of experiments have validated the performance of our algorithm.

The remainder of this paper is organized as follows:

Section II presents the proposed algorithm which is explained elaborately with adequate examples. Section III focuses on the performances of the algorithm with respect to a variety of parameters. Section IV draws the conclusions of the paper.

II. Overview of Proposed System

In this section we give a brief overview of our robust Hybrid Neural Network Approach to recognize Bangla digits. Our Hybrid Neural Network Approach can be applicable to any field for the purpose of recognition of handwritten Bangla digits. Here in this paper, we highlight it's application to postal automation service. Each of the phase is designed and implemented very carefully to make our recognition system more faster and efficient than previous research. The overall process of digit recognition is done into two major phases:

1) Image Processing
2) Hybrid Neural Network Approach

In the following subsections we will describe the phases listed above in a more details.

A. Image Processing

This is the first phase of our system. The goal of image processing is to extract the Bangla digits from the envelope and then convert these extracted digits into a binary format for recognition. This phase is very much essential for removing noise and for minimizing the differences between type phases, sizes and fonts of different handwriting styles. In our research all the digit samples are collected from Bangladesh mail envelope. Fig.1 shows some sample of handwritten digits.

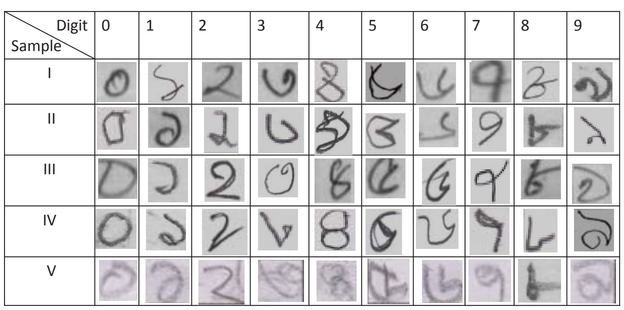

Fig. 1: Sample of Bangla Handwritten Digits

Image processing operations are performed through following steps:

1) **Scanning Image:** At first Bangladesh postal envelopes are scanned. Four digit postal codes are written in the box at lower right corner of the envelope. The digits were written by different people. So these digits vary in shape and size.
2) **Block Detection and Clipping:** The frame or block containing 4 digits postal codes is 15 pixels wide and 118 pixels long. Heuristic approach is used to clip the desired part. Clipping is done from fixed pixel position. The position is fixed by several trial and error methods. In this research, the position from which lower right portion is clipped: X=453, Y=256. So, 80% pixels are removed from envelope images. As a result, the probability of occurring error is reduced.
3) **Segmentation:** After detection of postcode frame, horizontal and vertical segmentation is done to finally extract each digit from the envelope.

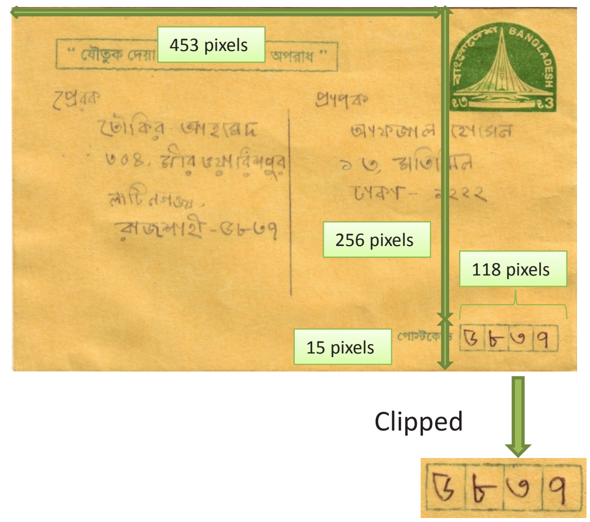

Fig. 2: Block Detection and Clipping Image

4) **Noise Reduction:** The image may contain noisy pixels. Such noises are removed for good performances even though neural networks has noise tolerance capability.
5) **Normalization and Scaling:** The segmented digits are then normalized to 16 by 16 pixels. Aspect Ratio Adaptive Normalization (ARAN) technique is used for normalization.
6) **Binarizing and Thinning:** Using Binarization the images are converted into matrices containing 0s for insignificant pixels and 1s for significant pixels. This research has applied Otsu's Method [13] for selecting the thresold value of the image. Based on this thresold value the image is converted to binary image. It is noticed that the thickness of the stroke is varying because different people use different writing tools. Since the thickness of the stroke influences the recognition result, thinning algorithm has been employed [14] to process the binary images of digits and get the same thickness of the stroke. So that all the thin images will have width 1 pixel.
At last the binary image is converted into binary digit matrix which consists of 0s and 1s. If the pixel is black, it becomes to 1 otherwise it is 0.

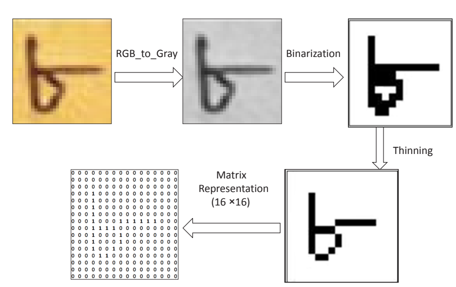

Fig. 3: Binarizing and Thinning

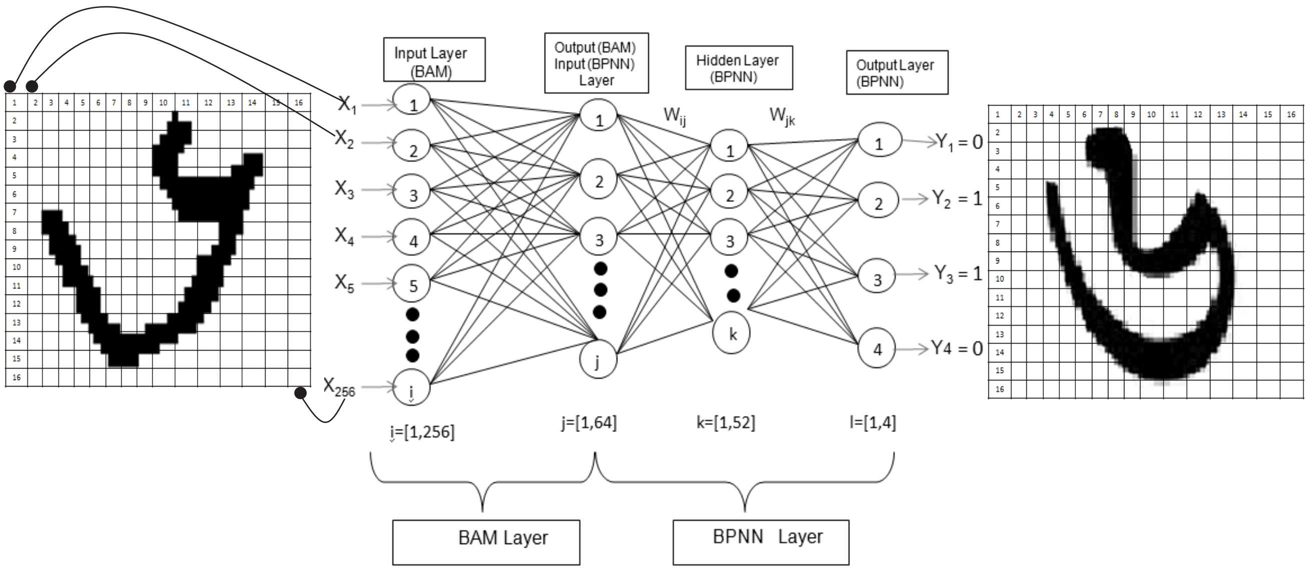

Fig. 4: Hybrid Neural Network Topology

B. Hybrid Neural Network Approach

The second stage of digit recognition uses the proposed hybrid neural network approach. A hybrid neural network is composed of two neural networks. One is **B**ack-**P**ropagation **N**eural **N**etwork (***BPNN***) and the other is **B**idirectional **A**ssociative **M**emory (***BAM***). The network architecture for the hybrid neural network is shown in Fig.4. The recognition process is divided into two phases:

1) Training Phase
   - Dimensional Reduction using BAM
   - Training Using BPNN
2) Recognition Phase

Each of the phase is discussed in details in the following subsections.

1) Dimensional Reduction Using BAM: Before training with Back-propagation neural network the dimension of the digit images are reduced. Employing **KOSKO's** hetero-associative correlator BAM, stored pattern pair are recalled even in presence of noise. The basic idea behind the BAM is to store pattern pairs for each digit. So that when M-dimensional vector **X** from set **A** is presented as input, BAM recalls N-dimensional vector **Y** from set **B**, where $M \geq N$.

But the limitation of BAM is it does not give guarantee to retrieve pattern pairs, when the stored pairs are at local minima. So in addition **WANG's** Multiple Training Encoding Strategy has been employed (which is an enhancement of the encoding strategy proposed by KOSKO) to make this research robust that ensures correct recall of al types of pattern pairs.

We have taken binary digit matrix as input matrix and employed BAM strategy to reduce dimension of the input digit matrix.

At first we have defined $m \times m$ digit matrix as $1 \times M$ digit input matrix of BAM which act as input vector **X** (where, $M = m \times m$ = Dimension of Input Digit Matrix). And the desired output is defined as $1 \times N$ digit matrix which is the representative of $n \times n$ digit matrix which act as output vector **Y** (where, $N = n \times n$ = Dimension of Output Digit Matrix). Then BAM generalized correlation matrix $(M \times N)$ is constructed using WANG's Multiple Training Encoding Strategy. Finally this weight matrix is used to retrieve the stored pattern which is associated with the input pattern. As $m > n$, so dimension is reduced from $m \times m$ digit matrix to $n \times n$ digit matrix.

For example, consider the case m=16 and n=8. BAM recalls 8×8 Output Vector **Y** from 16×16 Input Vector **X** using 256×64 BAM correlation weight matrix. As a result, dimension is reduced from 16×16 digit matrix to 8×8 digit matrix. Fig.5 shows an example for sample digit THREE.

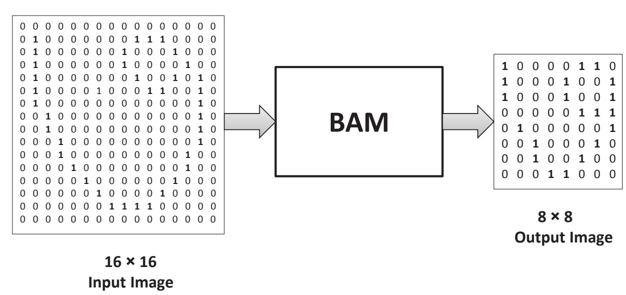

Fig. 5: Retrieving Associated Pattern Using BAM for Bangla Digit THREE

Algorithm 1 provides the pseudocode of the dimension reduction using BAM.

2) Training Using Back-Propagation Algorithm: The neural network is trained using back-propagation algorithm with the set of input patterns of the digits to learn the patterns corresponding to each Bangla digit.

Algorithm 1 Reduce Dimension Using BAM

Desired inputs ← binary matrix representation of each digit having larger dimension
Desired outputs ← binary matrix representation of each digit having smaller dimension
Calculate BAM weight matrix using Multiple Training Encoding Strategy for all associated pattern pairs
Save the BAM weight matrix for retrieving the associated pattern/digit pair

Algorithm 2 provides the pseudocode of the training process of hybrid neural network (***BAM+BPNN***). For each particular input pattern, algorithm 2 calls algorithm 3 to retrieve its associated output pattern from BAM. This associated pattern is defined as desired input of BPNN. The desired output of BPNN is a binary value of input decimal digit which consists of 4 neurons to classify 10 digits (0 to 9) in a binary format from 0000 to 1001. After initialization, training is continued until desired accuracy rate is achieved.

Algorithm 2 Training Hybrid Neural Network (***BAM+BPNN***)

for $each\ of\ the\ digit$ **do**
    Input Pattern ← the binary matrix representation of digit having larger dimension.
    Desired input of BPNN ← retrieve AssociatedPattern (Input Pattern)
    Desired output of BPNN ← the binary form of decimal digit (4 neurons)
    while $sumsq\_error > 0.001$ **do**
        Train the neural network using Back-propagation algorithm
    end while
    Save the finally trained weights and bias values for recognition
end for

Algorithm 3 retrieve AssociatedPattern (Input Pattern)

Output Pattern (smaller dimension) ← retrieve associated output of the Input Pattern (larger dimension) using saved BAM weight matrix
return Output Pattern

3) Recognition Phase: To finally recognize digit we input an unknown $m \times m$ digit matrix and retrieve it's associated stored pair $n \times n$ digit matrix using BAM (where $m > n$). This $n \times n$ digit matrix is defined as input of BPNN. During the recognition phase, no learning takes place i.e., weight matrices are not changed. There is no backward stage, only feed forward stage. Trained BPNN weight matrices are used to calculate the actual output in a binary format.

Algorithm 4 provides the pseudocode of the recognition process to identify the handwritten Bangla digits.

Algorithm 4 Recognize Handwritten Bangla Digit

Input Pattern ← binary matrix representation of an unknown pattern/digit having larger dimension
Desired input of BPNN (Output Pattern) ← retrieve AssociatedPattern (Input Pattern)
Saved trained weights and bias values are set to Back Propagation Neural Network to initialize
Actual output of the BPNN ← calculate using the back-propagation feed forward stage which is in a binary format.
Recognized digit ← Conversion (binary to decimal)

III. Experimental Results

This section presents the performance of the proposed system. In order to evaluate the effectiveness of the proposed system, various experiments were carried out for different real captured images of Bangladesh postal envelope. The experimental results will be discussed in brief with respect to execution time and memory usage.

In order to justify the performance, 1400 digits were obtained from the real letters written by Bangladeshis. The samples set is a mixture of good and indistinct, obscure hand written digits to evaluate the proposed scheme effectively. Randomly 800 digits are selected as the training set and the rest of 600 digits are used as the test set.

In this research, we consider different combination of matrix size i.e. 10×10, 12×12, 16×16. If each digit is 16×16 pixels image, then the digit has a feature matrix of 256 elements in it. Each element is nothing but binary values (0 and 1). Hence, in the experiment the number of neurons in input layer of BAM is 256 and the number of neurons in output layer of BAM is 64 which in turn act as input layer for BPNN. The number of neurons in hidden layer is 52 (80% of input layer). Finally the number of neurons in output layer is 4 that generate 16 distinct binary values. However, in order to uniquely recognize the 10 decimal digits correctly we used the first 10 binary patterns only.

The training algorithm depends on various network parameters which are selected by observing the performance of the hybrid network. The values of network parameters are listed in table I.

TABLE I: Values of Network Parameters (16×16 Digit Sample)

Parameters	Values
No. of layers (BAM)	2 (1 Input & 1 Output Layer)
No. of input neurons in input layer (BAM)	$16 \times 16 = 256$
No. of output neurons in output layer (BAM)	$8 \times 8 = 64$
Activation function (BAM)	Y^{Sign}=1 (if X ≥ 37); 0 (if X < 37)
Momentum factor (BPNN)	0.95
Learning rate (BPNN)	0.2
Weight and Bias Initialization range (BPNN)	[-2.4/64,2.4/64]
No. of layers (BPNN)	3 (1 Input, 1 Hidden & 1 Output Layer)
No. of input neurons in input layer (BPNN)	$8 \times 8 = 64$
No. of neurons in hidden layer (BPNN)	52 (80% of maximum no. of inputs)
No. of output neurons in output layer (BPNN)	4
Activation function (BPNN)	$Y^{Sigmoid}$=(1 / (1 + e^{-X}))
Error rate (Sum Squared Error)	0.001
Average number of epochs	1676

The sample output of the recognized digit 2 is shown in

Fig.6.

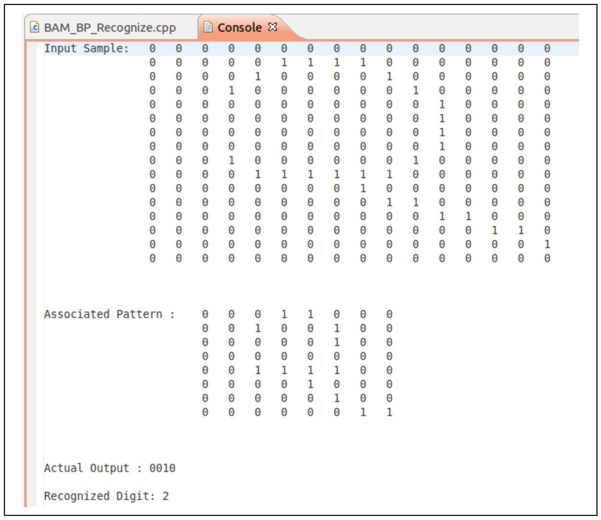

Fig. 6: Output of Test Image (Digit '2')

A. Error Minimization

Training is continued until the error comes to a minimum level. The error decreasing rate is graphically compared (%) between BPNN and BPNN with BAM with respect to iteration in the following Fig.7.

At first the task has been implemented using the BPNN algorithm and secondly, BAM and BPNN algorithm has been merged to train and recognize the Bangla Numeral. An error level is chosen (suppose 0.001 in percentage) to stop the training process. The error is calculated using sum of squared method. The number of iteration depends on the number of neurons in hidden layer and also on what algorithm has been used for the training process.

The performance of the network has been simulated by varying number of neurons in hidden layer from 50% to 90% with respect to input neurons in input layer. The training curve indicating the gradual reduction in error over several iterations due to our back-propagation (BPNN) learning algorithm. It is clear that the training curve converges more rapidly when 80% of input neurons (256) has been used in hidden layer. The number of iteration is 6360 when 50% of input neurons has been used in hidden layer but it decreases to 4070 when 80% of input neurons has been used in hidden layer. Finally when hybrid neural network (BPNN with BAM) approach has been used for the same percentage of hidden layer that is 80% of input neurons (64), the iteration decreases from 4070 to 1676. So, this research can take a decision that the considered hybrid neural network (BPNN with BAM) takes less iteration to train and less time to recognize digits than BPNN.

This research has been implemented in C++ on Linux platform with 3GB RAM. The performance of the algorithm has been tested using the following measures: recognition rate, error rate and rejection rate. Fig. 8 shows the bar diagram of the individual performance of different digits. From the figure it is seen that the proposed system achieves an average accuracy

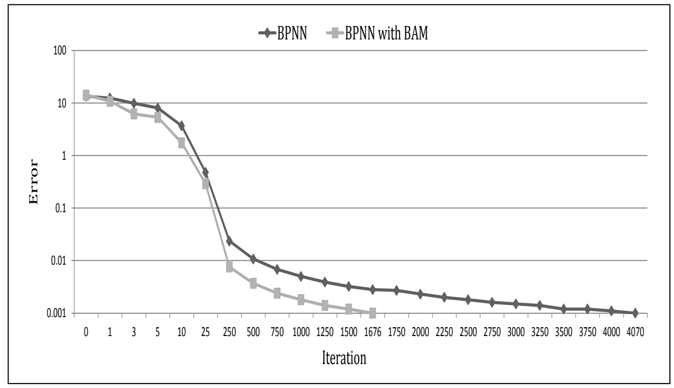

Fig. 7: BPNN vs BPNN with BAM (80% Hidden Layer)

rate (Recog.) of 91.1% and average error rate (Err.) and average rejection rate (Reject.) is 5.5% and 3.4% respectively. Highest accuracy rate is achieved for digit "eight" and "four". The mistaken recognition or rejection is due to variability of handwritings as well as bad writing.

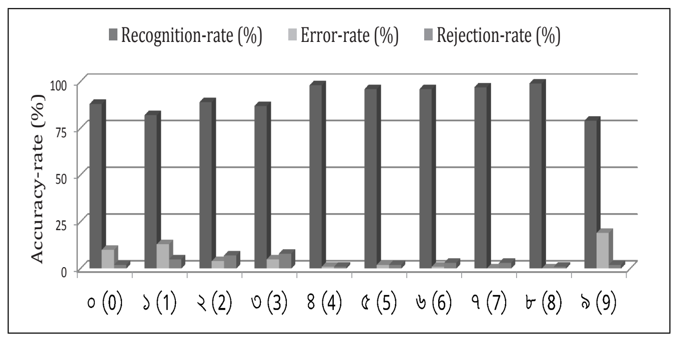

Fig. 8: Accuracy Rate of Each Digit

From the experiments it is noted that, our proposed scheme gives satisfactory outputs on recognition of indistinct handwritten digits. But sometimes it fails to recognize some special samples according to writers intension as because those samples are very much confusing. Fig.9 reflects some of such scenarios. The most confusing numeral pair is zero and three. Second confusing pair is one and two. Third confusing pair is one and nine. Therefore, error rate or rejection rate for these digits is higher. So the neural network needs to be trained using more samples especially for these digits.

We have also taken same noisy Bangla digits. Then we have observed the performance of our hybrid network on these noisy samples. Fig.10 shows the rate of error with different noise level. From the figure, it is noted that the recognition rate decreases with the increase of noise level.

Experimental results demonstrate that the presented system performs very well. It meets the response time as well as the accuracy requirements.

B. Cost Analysis

The performance of the hybrid network (BAM+BPNN) has been measured with respect to memory usage and execution

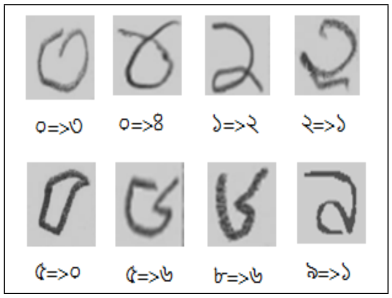

Fig. 9: Examples of some confused and indistinct handwritten Bangla digits

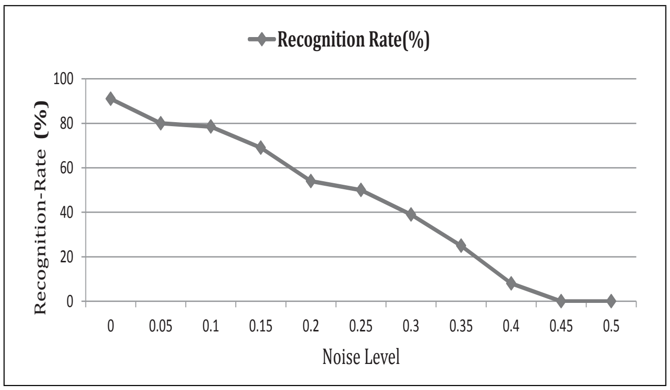

Fig. 10: Performance Evaluation on Noisy Digit Recognition

time. For the context of execution time, it has been observed that the training time is 21.14s and response time is 0.23s. Most of the time has been needed for training using back-propagation learning algorithm. But this training time is comparatively much shorter than previous research.

Now come to the memory consumption area. BAM is used in addition with BPNN for dimension reduction to make the recognition system faster. Although it requires some extra space to store the BAM correlation weight matrix for associating pattern pairs but it is the main reason that reduces the number of neurons in input layer of BPNN which in turn makes the training process faster.

IV. Conclusion

In this paper, we have proposed a hybrid neural network solution (BAM+BPNN) for the recognition of Bangla digits. Our main objective was to make the recognition system faster and efficient simultaneously. The back-propagation neural network continues to iterate until desired output is achieved. Hence we have merged the back-propagation and BAM algorithm. Moreover, to ensure correct recall of all types of pattern pairs multiple training encoding strategy has been employed in BAM. The performance is at satisfactory level since recognition time (0.23s) is minimized drastically and it meets the accuracy requirements (90%-95%) as well even on recognition of noisy digits. This is because both BAM and BPNN has noise tolerance capability. So, it can be concluded that the proposed algorithm (BAM+BPNN) exhibits a very fast and robust approach in recognition of Bangla digits.

In future, we have a plan to train our network with more samples to improve the accuracy rate. Because the recognition accuracy of a hybrid neural network increases with the increase of training accuracy.

Acknowledgment

We would like to express our deep gratitude to the anonymous reviewers of this paper. Their useful comments have played a significant role in improving the quality of this work.

References

[1] "Bengali language." [Online]. Available: http://en.wikipedia.org/wiki/Bengali_language

[2] V. S. P. U. Roy, K. and B. Chaudhuri, Eds., *A system towards Indian postal Automation*, 2004.

[3] Y. Wen, Y. Lu, and P. Shi, "Handwritten Bangla numeral recognition system and its application to postal automation," *Pattern Recognition*, vol. 40, no. 1, pp. 99–107, Jan. 2007. [Online]. Available: http://dx.doi.org/10.1016/j.patcog.2006.07.001

[4] D. Singh1 and B. S. Khehra, "Digit recognition system using back propagation neural network," *International Journal of Computer Science and Communication*, vol. 2, no. 1, pp. 197–205, January-June 2011.

[5] R. B. Ashoka H. N., Manjaiah D. H., "Feature extraction technique for neural network based pattern recognition," *International Journal on Computer Science and Engineering (IJCSE)*, vol. 4, no. 3, pp. 331–339, March 2012.

[6] C.-L. Liu and C. Y. Suen, "A new benchmark on the recognition of handwritten bangla and farsi numeral characters," *Pattern Recogn.*, vol. 42, no. 12, pp. 3287–3295, Dec. 2009. [Online]. Available: http://dx.doi.org/10.1016/j.patcog.2008.10.007

[7] M. M. Hoque and M. O. Karim, "Extraction of meaningful fuzzy features for recognizing of bangla handwritten numerals," in *Computer and Information Technology (ICCIT), 2006 9th International Conference on*, 2006, pp. 61–66.

[8] N. Das, B. Das, R. Sarkar, S. Basu, M. Kundu, and M. Nasipuri, "Handwritten Bangla Basic and Compound character recognition using MLP and SVM classifier," *ArXiv e-prints*, Feb. 2010.

[9] M. A. F. M. A. S. M. I. A. S. M. Mahbub Morshed, Md. Shahidul Islam, "Automatic sorting of mails by recognizing handwritten postal codes using neural network architectures," in *Proceedings of the ICCIT (International Conference on Computer and Information Technology)*, 2007.

[10] M. S. S. Shariful Hasan Shaikot, Fahim Kawsar, "Bangali digit recognition system using 3-layer backpropagation neural network," in *Computer and Information Technology (ICCIT), 2003 6th International Conference on*, 2003, pp. 327–331.

[11] H. G. E. Rumelhart, D. E. and R. J. Williams, *Learning internal representations by error propagation. In Parallel distributed processing: Explorations in the microstructure of cognition*. Bradford Books, Cambridge, MA, vol. 1.

[12] P. D. M. A.-A. B. Sanjit Kumar Saha, Md. Shamsuzzaman, Ed., *On Bangla Character Recognition*. Dhaka, Bangladesh: IEEE Computer Society Press., 23-25 December 2010, proceedings of 13th International Conference on Computer and Information Technology (ICCIT 2010).

[13] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[14] S.-W. L. Lam, L. and C. Y. Suen, "Thinning methodologies-a comprehensive survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 9, p. 879, September 1992.