Notice of Violation of IEEE Publication Principles

“An Approach to Recognize Handwritten Bengali Numerals for Postal Automation"
by Md. Saidur Rahman, G. M. Atiqur Rahaman, Asif Ahmed and G. M. Salahuddin
in the Proceedings of 11th International Conference on Computer and Information Technology (ICCIT 2008, 25-27 December, 2008, Khulna, Bangladesh)

After careful and considered review of the content and authorship of this paper by a duly constituted expert committee, this paper has been found to be in violation of IEEE's Publication Principles.

This paper contains significant portions of original text from the paper cited below. The original text was copied without attribution (including appropriate references to the original author(s) and/or paper title) and without permission.

Due to the nature of this violation, reasonable effort should be made to remove all past references to this paper, and future references should be made to the following article:

“Handwritten Bangla Numeral Recognition System and Its Application to Postal Automation"
by YingWen,Yue Lu, Pengfei Shi
in *Pattern Recognition* (Jan 2007, vol. 40, pp.99–107) Elsevier

Proceedings of 11th International Conference on Computer and Information Technology (ICCIT 2008)
25-27 December, 2008, Khulna, Bangladesh

AN PPROACH TO RECOGNIZE HANDWRITTEN BENGALI NUMERALS FOR POSTAL AUTOMATION

Md. Saidur Rahman, G. M. Atiqur Rahaman, Asif Ahmed and G. M. Salahuddin

Computer Science and Engineering Discipline, Khulna University, Khulna-9208, Bangladesh

E-mail: msrku@yahoo.com, atiq@ku.ac.bd, asif_cseku@yahoo.com, gms_far@yahoo.com

***Abstract* — A recognition approach for handwritten Bengali numerals and its application for Bangladesh Postal system are presented in this paper. The approach consists of preprocessing, feature extraction and recognition. Using the canny edge detector the top left corner of the post-code box has been detected and the four handwritten numeral images have been segmented. Each image then goes through the steps of normalization, filtering and thinning. Kirsch mask has been used to get the edges through the horizontal, vertical, right and left diagonal. Principal component analysis (PCA) has been used for dimension reduction as the final feature vector consists of the four directional feature vector along with the normalized image. Then the output of the PCA is passed to a trained Support Vector Machine (SVM) to determine which class the input belongs to. Experiments demonstrate that the average recognition rate, error rate and reliability achieved by the proposed system are 92.5%, 7.5% and 92.5% respectively.**

***Index Terms* — Bengali Numeral, Feature Extraction, Kirsch Mask, Principal Component Analysis, Support Vector Machine.**

I. INTRODUCTION

Bengali is the second most popular language in South Asia. But unfortunately we have not yet got the appropriate application towards handwritten Bengali numeral character recognition system as most of the research works have been on retrieving content of printed document [1]-[3]. But for our postal automation a system is really necessary that can recognize Bengali numerals from the envelope. Most of the countries in the world already have been developed the system for their postal automation. Postal automation is a topic of research interest for last two decades and many pieces of published article are available towards postal automation of non-Bengali languages documents [4]-[7]. Several systems are also available for postal automation in USA, UK, France, Canada and Australia. But very few works have been done towards the automation of Bangladeshi postal system [8].

Handwritten character recognition is highly subjective as different people have different art of writing the same numeral character. That is why to get the actual numerals from handwriting is difficult in real time. So the main objective of our work is to recognize handwritten Bengali numerals from the envelope of Bangladesh Postal system to get the post code with high recognition rate, less error rate, high reliability and comparatively less response time.

II. LITERATURE SURVEY

Various approaches have been proposed by the researchers for numeral recognition [9] but most of them are used for non-Indian numeral recognition. One of the most widely used approaches is based on neural networks [10]. Here the network architecture is trained by a set of training data and then the input is classified by the trained networks. In structural approach each pattern class is defined by structural description and the recognition is performed according to structural similarities [11]. Statistical approach is also applied to numeral recognition. It is relatively insensitive to pattern noise and distortion but modeling of statistical information is a tedious task [9]. Among others, Hidden Markov Models, Fourier and Wavelet Descriptors [12], Fuzzy rules [13], tolerant rough set [14], are reported in the literature.

Feature extraction plays a vital role for a recognition system. Principal component analysis is a well-known method for feature extraction. By calculating the eigenvectors of the covariance matrix of the original inputs, PCA linearly transforms a high-dimensional input vector into a low-dimensional one whose components are uncorrelated [15]-[17]. Recently, support vector machine (SVM) has become a popular tool for numeral recognition due to its remarkable characteristics such as good generalization performance, the absence of local minima. SVM implements the structural risk minimization principle which seeks to minimize an upper bound of the generalization error rather than minimizing the training error [18].

In this paper, we present a recognition system for handwritten Bengali numerals. The system consists of three main parts: preprocessing, feature extraction and recognition. The input of the system is the image of an envelope and the output is the recognized Bengali numerals of the post-code written upon the envelope. Here we propose a recognition approach based on the direction feature combined with PCA and SVM. The experimental

results show that the recognition performance of the proposed approach is superior to other conventional recognition approaches.

III. Materials and Methods

The overall system is represented in Fig.1. Here the preprocessing phase consists of post-code box detection, segmentation, normalization, filtering, thinning and dilation. In the dimension reduction phase Kirsch mask and the PCA method are applied. Then the SVM is trained up. And in the recognition phase the numeral characters are recognized.

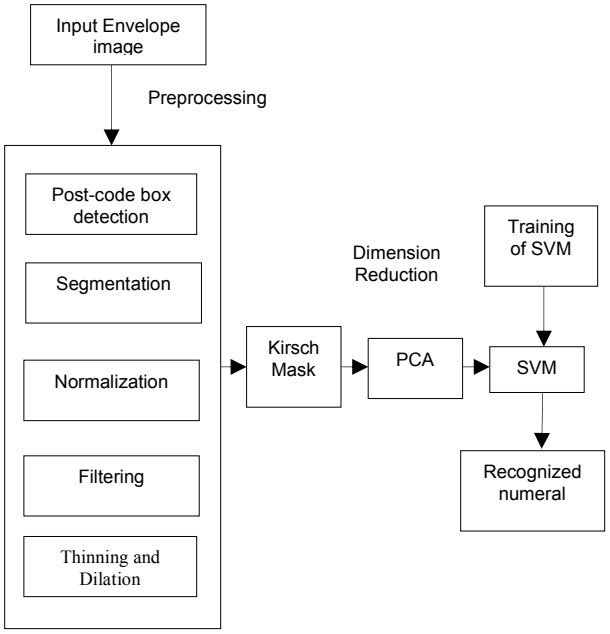
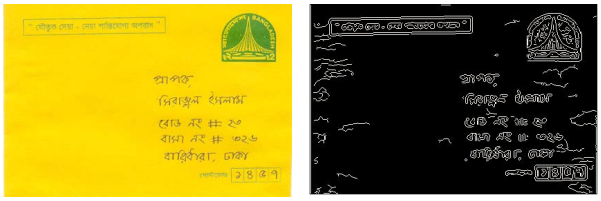


Fig. 1. Block diagram of the proposed system

A. Preprocessing

The goal of this part is to extract numeral images of postcode from an envelope image for the subsequent recognition. In Bangladeshi postal documents there are pre-printed boxes to write post-code. We call these boxes as post-code boxes. People generally write the destination post-code of four digits inside these boxes.

Detection of Post-code Box and Segmentation: Getting an image of envelope as input, the first task is to locate the post-code box in the envelope. To serve the purpose at first the true color image has been transformed into a binary image and then the edge detection procedure has been applied on it. For edge detection the canny method [19] has been used. The Canny method finds edges by looking for local maxima of the gradient of binary image. The gradient is calculated using the derivative of a Gaussian filter. The method uses two thresholds, to detect strong and weak edges, and includes the weak edges in the output only if they are connected to strong edges. This method is therefore less likely than the others to be "fooled" by noise, and more likely to detect true weak edges. The Fig.2 shows an example of the effect of applying Canny Method.



(a) (b)

Fig. 2. (a) The input image (b) After applying 'Canny' method for edge detection.

We have started searching for the post-code box from a pre-specified location which is nearby the searching location. During the search for an edge, whenever we have found the top horizontal edge of the post-code box, we also get the left most point. We have noticed that for a 72 dpi image of the envelope, each box is of size 18x20 pixels. So after getting the left-top point of the post-code box, we can easily crop the required 4 images of handwritten Bengali numerals.

Normalization, Filtering, Thinning and Dilation: Size normalization is implemented on each character image that has been collected from previous stage. The size is normalized to be 16x16 pixels. Then, we employ the Gaussian filter to remove the noise and then transform the image into binary image. We notice 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, we use the thinning and dilation algorithm to process the normalized numerals and get the same thickness of the stroke.

B. Dimension Reduction

Use of Kirsch Mask on the Normalized Image: The normalized image can pass through the SVM as input but there will be lot of relevant information missing in that case. That is why, in this experiment, we first obtain four directional feature maps of a numeral image by the Kirsch Mask. The size of each map is 16×16 pixels. The original image is considered as the fifth feature map. Then we arrange the five images to get the feature vectors (5×16×16=1280 dimensions). But this is huge dimension to be given as input to the SVM. To solve this problem we have used PCA to decrease the dimension and extract more significant features. In this way, recognition time and training time can be reduced. Fig.3 illustrates this step.

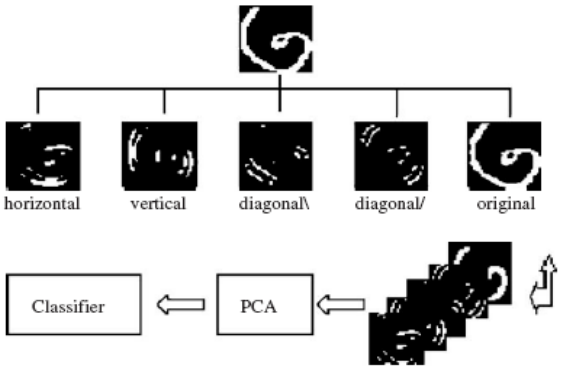


Fig. 3. Applying kirsch mask on the normalized image.

C. Training of SVM

We have collected hand written Bengali numerals of a number of people. A sample has been shown in Fig.4.

Following the same preprocessing procedure of an input image, we have collected the 10 Bengali numerals of different handwriting. After using Kirsch mask and PCA on these images, we have stored it in a .mat (a MATLAB file extension) file which is used to classify the SVM.

After getting input from PCA for each numeral image, the SVM tries to find the class where the input numeral image should belong.

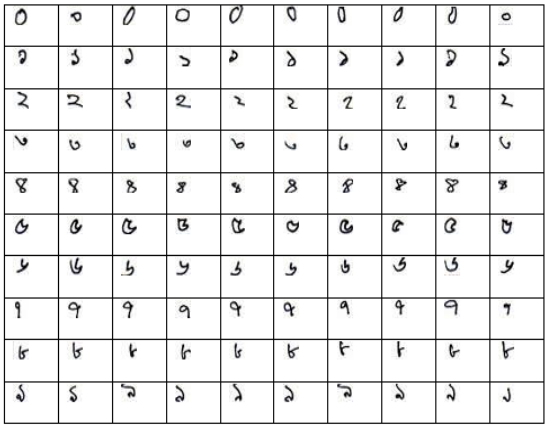


Fig. 4. Samples of handwritten Bengali numeral images (০ -৯)

D. Recognition

Based on PCA and SVM, we propose the recognition approach based on Kirsch mask combined PCA and SVM. Generally, Kirsch edge detector, Prewitt edge detector, Sobel edge detector and so on are the representative edge detectors. Among them the Kirsch edge detector has been known to detect four directional edges more accurately than others because the Kirsch edge detector considers all eight neighbors. Kirsch defined a nonlinear edge enhancement algorithm as follows [20]:

$$G(i,j) = \max_{k=0}^{7}\{1, \max[|5S_k - 3T_k|]\} \quad (1), \text{ where}$$

$$S_k = A_k + A_{k+1} + A_{k+2} \quad (2)$$

$$T_k = A_{k+3} + A_{k+4} + A_{k+5} + A_{k+6} + A_{k+7} \quad (3)$$

In Equations (1), (2) and (3) G (i, j) is the gradient of pixel (i, j), $A_k\,(k = 0,1,.....,7)$ is eight neighbors of pixel (i, j) as defined in Fig.5.

A_0	A_1	A_2
A_7	(i, j)	A_3
A_6	A_5	A_4

Fig. 5. Definition of eight neighbors of pixel (i, j)

In this paper, we calculate directional feature vectors for horizontal (H), vertical (V), right-diagonal (R), and left-diagonal (L) directions as follows:

$$G_H(i,j) = \max(|5S_0 - 3T_0|, |5S_4 - 3T_4|) \quad (4)$$

$$G_V(i,j) = \max(|5S_2 - 3T_2|, |5S_6 - 3T_6|) \quad (5)$$

$$G_R(i,j) = \max(|5S_1 - 3T_1|, |5S_5 - 3T_5|) \quad (6)$$

$$G_L(i,j) = \max(|5S_3 - 3T_3|, |5S_7 - 3T_7|) \quad (7)$$

The Fig.6 shows the Kirsch masks used for calculating directional feature vectors. The normalized image is processed by the pairs of Kirsch masks (3 × 3 pixels), resulting in four graded feature maps, coding for the presence of horizontal, vertical or diagonal edges. In this way, we get four directional feature maps. We also take the original image as the fifth map in order to compensate the disadvantages of four directional feature maps. The final data representation corresponds to 5×16×16(=1280) dimensional inputs. The 1280-dimension is too much for input of a classifier so we employ PCA to extract feature and decrease the dimension. Finally, we obtain the new feature presentation as the input of SVM.

6	5	5	-3	-3	-3
-3	0	-3	-3	0	-3
-3	-3	-3	5	5	5

(a)

-3	-3	5	5	-3	-3
-3	0	5	5	0	-3
-3	-3	5	5	-3	-3

(b)

-3	5	5	-3	-3	-3
-3	0	5	5	0	-3
-3	-3	-3	5	5	-3

(c)

-3	-3	-3	5	5	-3
-3	0	5	5	0	-3
-3	5	5	-3	-3	-3

(d)

Fig. 6. Kirsch masks, (a) horizontal, (b) vertical, (c) right-diagonal and (d) left-diagonal

IV. Experimental Results

A. Data Collection

To meet the requirements of a practical usage of the letter sorting machine for Bangladesh Postal System, we should consider not only the recognition rate and the error rate but also the response time. There are various writing styles. For our experiments, 1600 numerals were obtained from the real letters written by Bangladeshis. We randomly select 900 numerals as the training set and the rest as the test set.

For this paper, our experiments were tested on Windows XP, Pentium T(R) D CPU 2.8G and 760 RAM. We employ MATLAB to realize the program.

B. Time Performance:

To evaluate the performance of our proposed approach, we have implemented both the proposed method as well as an existing method which takes normalized image as the input of SVM classifier. Table I shows the training and recognition time of our proposed system and the existing system.

TABLE I
THE RECOGNITION AND TRAINING TIME OF TWO APPROACHES

Approach	Recognition time (ms)	Training time (s)
Normalized image + SVM	21.44	67.94
Our Proposed System	31.24	146.47

C. Recognition Performance

The proposed method demonstrates a very high recognition rate. The experimental result shows that close similarity in geometric shape of some numerals led to the chance of false recognition. The Table II represents the recognition rate for different numerals with respect to all numerals.

To evaluate the recognition results, the following measures are employed: recognition rate (Recog.), error rate (Err.), rejection rate (Reject.), and reliability (Reliab.). They are defined as follows:

TABLE II
THE RATE OF RECOGNITION OF EACH NUMERAL WITH RESPECT TO ALL THE 10 NUMERALS

	Recognition Rate (%)									
	Numeral recognized as									
Input	০	১	২	৩	৪	৫	৬	৭	৮	৯
০	91.7	0	0	8.3	0	0	0	0	0	0
১	0	91.7	8.3	0	0	0	0	0	0	0
২	0	16.7	83.3	0	0	0	0	0	0	0
৩	16.7	0	0	75	0	0	8.3	0	0	0
৪	0	0	0	0	100	0	0	0	0	0
৫	0	0	0	0	0	100	0	0	0	0
৬	0	0	0	0	0	0	100	0	0	0
৭	0	0	0	0	0	0	0	100	0	0
৮	0	0	0	0	0	0	0	0	100	0
৯	0	16.7	0	0	0	0	0	0	0	83.3

TABLE III
RECOGNITION RESULTS BY THE PROPOSED SYSTEM

Recognition Attributes	০	১	২	৩	৪	৫	৬	৭	৮	৯	Avg.
Recog(%)	91.67	91.67	83.3	75	100	100	100	100	100	83.3	92.5
Err.(%)	8.3	8.3	16.7	25	0	0	0	0	0	16.7	7.5
Reliab.(%)	91.67	91.67	83.3	75	100	100	100	100	100	83.3	92.5

TABLE IV
THE RECOGNITION RATES (%) OF TWO APPROACHES

Approach \ Numerals	০	১	২	৩	৪	৫	৬	৭	৮	৯
Normalized Image to SVM	58.3	66.7	83.3	58.3	100	83.3	83.3	100	100	58.3
Proposed System	91.7	91.7	83.3	75	100	100	100	100	100	83.3

TABLE V
COMPARISON OF AVERAGE RECOGNITION RATE WITH OTHER EXISTING APPROACHES

	Proposed System	Normalized Image to SVM	Neural Network Based method	
			With momentum factor	Without momentum factor
Avg. rate of recognition (%)	92.5	79.1	60.75	89.75

$$\text{Recog.} = \frac{\text{N1 (correctly recognized)}}{\text{N4 (tested)}} \times 100\% \quad (8)$$

$$\text{Err.} = \frac{\text{N2 (error recognized)}}{\text{N4 (tested)}} \times 100\% \quad (9)$$

$$\text{Reject.} = \frac{\text{N3 (rejected)}}{\text{N4 (tested)}} \times 100\% \quad (10)$$

$$\text{Reliab.} = \frac{\text{Recog.}}{\text{100\% - Reject.}} \times 100\% \quad (11)$$

where N1 (correctly recognized) is defined as the number of numerals correctly classified, N2 (error recognized) is defined as the number of numerals falsely classified, N3 (rejected) is defined as the number of numerals rejected by the classifier, and N4 (tested) is the number of input numerals for test.

Table III shows the recognition rate, the error rate and the reliability of different numerals for the proposed system. Finally, we get the average recognition rate; the average error rate and the average reliability are 92.5%, 7.5% and 92.5%, respectively. Experimental results demonstrate that our presented system performs very well. It meets the response time with the accuracy requirements.

V. PERFORMANCE COMPARISONS AND DISCUSSIONS

To compare the results of our proposed method we have implemented existing method applying normalized image to SVM and neural network based system using back-propagation algorithm [21]. Table IV presents the recognition results of our proposed approach and other existing method.

The Table V represents the average recognition rate of Bangla numerals. The third method which is based on neural network using back propagation algorithm [21] is tested during the training period with iteration number from 4000 to 7000 with momentum and without momentum.

The experimental results of Table IV and Table V show that the recognition rate of our method is better than other existing system. But from Table I it is found that the time performance of our proposed approach is not better than the other. But it is remarkable that the recognition rate of our method is much higher than the two other existing methods.

From Table II it has been found that the highest reliability is 100% obtained by Bengali numeral "৪", "৫", "৬", "৭", "৮". The false recognition or rejection

is due to variability of handwritings as well as bad writing. From the experiments we note that the most confusing numeral pair is "০" and "৩" as shown in Fig.7(a). Second confusing pair is "৯" and "৯" as shown in Fig.7(b). Third confusing pair is "৯" and "২" as shown in Fig.7(c). Therefore, error rate or rejection rate for these numerals is higher.

(a) (b) (c)

Fig. 7. .Examples of some confused handwritten Bengali numeral pairs: (a) "০" and "৩" (b) "৯" and "৯" (c) "৯" and "২".

VI. CONCLUSIONS

In a recognition system the single most important issue is to ensure a good recognition rate. We have used PCA and SVM to enhance the recognition rate. In order to maximize the performance of numeral recognition, our proposed approach extracts four directional local feature vectors with the Kirsch masks and one global feature vector from a normalized input image. Hence, the important features are retained and the non-important pixels are minimized. The recognition result of our proposed approach outperforms the others. However, our proposed approach has extracted more significant features than the method applies normalized image to SVM. So it needs more time to estimate than that specific approach. But our method also works better with respect to recognition performance than the method for numeral recognition by neural network using back-propagation algorithm with and without momentum factor. As far as the recognition rate is concerned, it is experimentally shown that the recognition approach that we have proposed in this paper is superior to other conventional approaches.

REFERENCES

[1] B. B. Chaudhuri, and U. Pal, "A complete Printed Bangla OCR System", *Pattern Recognition* 31(5): 531–549, 1998.

[2] U. Pal, "On the development of an optical character recognition (OCR) system for printed Bangla script", *Ph.D. Thesis*, Indian Statistical Institute, 1997.

[3] U. Pal, and B. B. Chaudhuri, "Indian Script Character Recognition: a Survey", *Pattern Recognition*, 37(9): 1887–1899, 2004.

[4] Y. L. Cun , O. Matan, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, and H. S. Baird, "Handwritten Zip Code Recognition with Multilayer Networks", *In Proceedings of the 10th International Conference on Pattern Recognition (Atlantic City, NJ). IEEE Computer Society Press*, vol 2, pp. 35-40, 1990.

[5] U. Mahadevan, and S. N. Srihari, "Parsing and Recognition of City, State, and ZIP Codes in Handwritten Addresses", *In Proc. of Fifth ICDAR*, pp. 325-328, 1999.

[6] X. Wang, and T. Tsutsumida, "A New Method of Character Line Extraction from Mixed unformatted Document Image for Japanese Mail Address Recognition", *In Proc. of Fifth ICDAR*, pp. 769-772, 1999.

[7] S. N. Srihari, and E. J. Keubert, "Integration of Hand-Written Address Interpretation Technology into the United States Postal Service Remote Computer Reader System", *In Proc. Of Forth ICDAR*, pp. 892-896, 1997.

[8] A. S. M. M Morshed, M. S. Islam, M. A. Faisal, M. A. Sattar, and M. Ishizuka, "Automatic Sorting of Mails by Recognizing Handwritten Postal Codes using Neural Network Architectures", *Proc.(CD-ROM) Int'l Conf. on Computer and Information Technology (ICCIT2007),* session D3, 5 pages, Dhaka, Bangladesh , Dec. 2007.

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

[10] K. Kim, J. H. Kim and C. Y. Suen, "Recognition of Unconstrained handwritten Numeral Strings by Composite Segmentation method", *Proc.15th ICPR*, pp. 594-597, 2000.

[11] J. Cai and Z. Q. Liu, "Integration of structural and statistical information for unconstrained handwritten numeral recognition", *IEEE Trans. On PAMI*, 21(3): 263-270, 1999.

[12] P. Wunsch and A. F. Laine, "Wavelet Descriptors for Multi-resolution Recognition of Hand-printed Digits", *Pattern Recognition*, 28(8):1237-1249, 1995.

[13] Z. Chi and H. Yan, "Handwritten numeral recognition using self-organizing maps and fuzzy rules", *Pattern Recognition*, 28(1):59-66, 1995.

[14] K. Kim and S. Y. Bang, "A handwritten numeral character classification using tolerant Rough set", *IEEE Trans. on PAMI*, 22(9): 923-937, 2000.

[15] M. Turk and A. Pentland, "Eigenfaces for Recognition", *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991.

[16] K. Ohba and K. Ikeuchi, "Detectability, Uniqueness, and Reliability of Eigen Windows for Stable Verification of Partially Occluded Objects",*IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 9, pp. 1043-1048, 1997.

[17] H. Murase and S. Nayar, "Visual Learning and Recognition of 3D Objects from Appearance", *Interantional Journal of Computer Vision*, vol 14, pp. 5-24, 1995.

[18] V. N. Vapnik, "The Nature of Statistical Learning Theory", Springer, New York, 1995.

[19] http://www.mathworks.com, November 19, 2007.

[20] S. W. Lee, "Off-line recognition of totally unconstrained handwritten numerals using multilayer cluster neural network", *IEEE Trans. Pattern Anal. Mach. Intell.* 18 (6): 648–652, 1996.

[21] M. W. Islam, M A. M. Hasan, R. C. Debnath, "Handwritten Bangla Numeral Recognition using Back-propagation Algorithm With and Without Momentum factor ", *Proceedings of National Conference on Computer Processing of Bangla (NCCPB 2005)*, Independent University, Bangladesh, pp. 177-182, 2005.