AUTOMATIC SORTING OF MAILS BY RECOGNIZING HANDWRITTEN POSTAL CODES USING NEURAL NETWORK ARCHITECTURES

A. S. M. Mahbub Morshed[*], Md. Shahidul Islam[!], Md. Abul Faisal[#], Md. Abdus Sattar[+], Mitsuru Ishizuka[^]
[*]University of Tokyo, Tokyo, Japan, email: mmorshed@mi.ci.i.u-tokyo.ac.jp
[!]University of Ottawa, Canada, email: shah1d1698@hotmail.com
[#]Royal Melbourne Institute of Technology, Melbourne, Australia, email: fosi41@hotmail.com
[+]Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, email: masattar@cse.buet.ac.bd
[^]University of Tokyo, Tokyo, Japan, email: ishizuka@i.u-tokyo.ac.jp

***Abstract* - We develop a neural-network architecture for recognizing handwritten Bangla digits and its application to automatic mail sorting machine for Bangladesh Postal system is presented. The system consists of preprocessing, classification and implementation. Preprocessing is fully automated for envelops with a postcode frame. Otherwise, a human operator has to manually select the block if the numerals of postcode are in the destination address region. Neural network with one hidden layer is trained using back-propagation algorithm for classification. We also discuss implementing the system in a real-world mail sorting system. The performance of the network on single digits is 92.2% recognition, 6% substitution and 1.8% rejection. The system is a hybrid of image processing, segmentation and neural network classification.**

***Keywords* – Bangla Postal codes recognition, Neural Network, Back-propagation algorithm, Handwritten pattern recognition**

I. INTRODUCTION

The Government of Bangladesh approved acquisition of a mail sorting machine for its sluggish postal service almost a decade ago. However, due to financial constraints, it can not afford the excessively expensive foreign technology. We strongly feel that the same technology can be developed using home-grown expertise and research facilities. We adopt the approach used by Le Cun et al.[1] with some minor modification: our system is developed on Bangla digits only and the network has only one hidden layer instead of 4 hidden layers or convolution layers.

A widely accepted approach to handwritten character recognition is to divide the classification process into a feature extraction followed by a classification. The feature extractor contains most of the problem-dependent information and it requires most of the design effort and determines the performance to a large extent. The classifier, on the other hand, contains little *a priori* knowledge about the task. Our approach is to demonstrate that neural network trained with back-propagation can be applied to real image-recognition system without complex feature-extraction stage. As architecture of the network strongly influences recognition performances [2], we can achieve a good recognition by designing a network that contains a certain amount of a *priori* knowledge about the problem.

The remainder of the paper is structured as follows: Section 2 gives a brief overview of the related work on Bangla character recognition, Section 3 presents the preprocessing of the system, which mainly discusses postcode location, digit segmentation and normalization. Section 4 describes the neural network design and the training process. Section 5 presents the experimental results achieved by the proposed approach. Section 6 discusses a possible way of implementing and integrating the system. Finally, in Section 7 we draw out conclusions.

II. RELATED WORKS

Bangla is the second-most widely spoken language in the Indian subcontinent and is the native-tongue of 99% citizens of Bangladesh. Unfortunately, researches on handwritten Bangla character recognition are not sufficient. Some papers on printed Bangla digit recognition are available [3][4][5] but the most promising work on the recognition of handwritten Bangla digit is by Pal et al. using the concept of water reservoir [6][7][8].

On the other hand, postal automation is a topic of research interest for last two decades and many pieces of published article are available towards postal automation in languages other than Bangla [1][9][10][11]. Several systems are also available for postal automation in USA, UK, Japan, France, China, Australia and Canada. Other than Wen et al. [12], no work has been done towards the automation of Bangladesh postal system. Wen et al. use support vector machine (SVM) classifier combined with extensive feature extractor using principal component analysis (PCA) and Kernel principal component analysis (KPCA). Our system directly takes the processed image of Bangla digit and does not need problem dependent feature extraction.

Neural Network trained using back-propagation [13] is a popular tool for numeral recognition due to its good generalization performance. Neural network tries to mimic human brain in the sense that it provides an abstraction of parallel processing and can be implemented using digital signal processing toolkit.

III. PREPROCESSING

The main aim of preprocessing is to extract bangla digits from the images of envelopes. It is designed to separate the signal from the noise, making the subsequent recognition more robust. Noise in the case of handwriting is not only dirt and interference in the image capture. It is also variability in handwriting styles: the size, the slant, thickness of the strokes and other variables.

Bangla Digits for this present work is collected from postal codes that appeared on Bangladesh Mail envelopes. We used a flatbed scanner for digitization. The images are in gray tone and digitized at 200dpi and stored as Joint Photographic Expert Group (jpg) format. We have used the following stages to convert the grey tone image into two-tone (0 and 1) image….

1) *Block Detection:* If the digits are located in the postcode frame (Figure 1), then a heuristic method is used to select the frame initially. The frame or block containing 4 digit postal codes is 25 pixels wide and 95 pixels long. This heuristic approach reduces the response time of subsequent recognition. If it fails to detect the frame or if the frame length is not more than 3.5 times its width, then Run Length Smoothing Algorithm (RLSA) is applied at the bottom quarter of the envelop, which contains the least written image. Using component labeling, each component is then selected and mapped to the original image. A component is selected if the length of the component is greater than 5 times the width of the component and the length of the component is less than seven times the width of the component. Since a postcode frame contains 6 square boxes, the length of the desired framed component will be about 6 times the width of the component. In the case of wrong detection – i.e. when the numerals of postcode are in the destination address region – the frame is selected manually.
2) *Segmentation:* After detection of postcode frame, horizontal lines are detected and deleted using line segmentation. Through vertical pixel analysis, vertical lines and numeric digits are segmented. Depending on the positions of the vertical lines the postal code digits are extracted from left to right to preserve the order of the occurrence of the digits. In most cases, our method is able to extract those digits that touched/crossed the border of postcode box. The segmentation process is illustrated in Figure 2.
3) *Binarization:* Binarization involves the representation of images into matrices containing 0s for insignificant pixels and 1s for significant pixels. Histogram based global binarizing algorithm [3] is used in this process. We further used Hough transform to de-skew the image.
4) *Noise Reduction:* The binary image may contain spurious noise pixels and irregularities on the boundary. Such noises are removed for better performances even though neural networks can tolerate some noises.
5) *Boundary Extraction:* Boundary extraction involves finding the boundary position of each digit matrix. We observe that this preprocessing step has a profound effect in reducing some of the spurious noises mentioned in the previous step.
6) *Normalization and Scaling*: The segmented digits vary in size. The size of the digits is then normalized to 12 by 12 pixels. We use an Aspect Ratio Adaptive Normalization (ARAN) technique for normalization. This technique is better than linear mapping as ARAN preserves the shape and aspect ratio of digit image. A normalized image of Bangla digit '7' is given in Figure 3.

Finally, our database consists of 10400 segmented Bangla digits. The training set consists of 5500 digits and the remaining 4900 digits are used for testing. We ensure that both the training set and the testing set contain numerous examples that are ambiguous or unclassifiable.

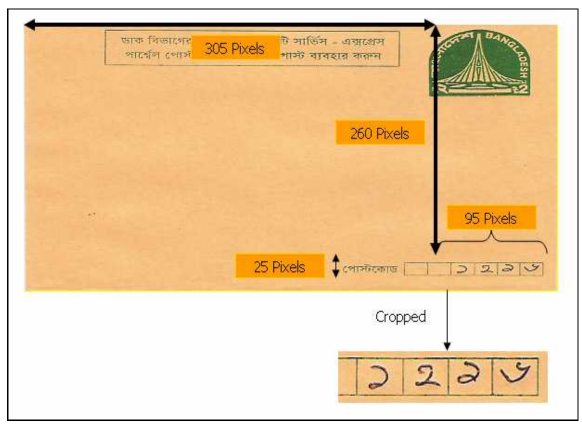


FIGURE 1:
HEURISTIC APPROACH TO BLOCK DETECTION.

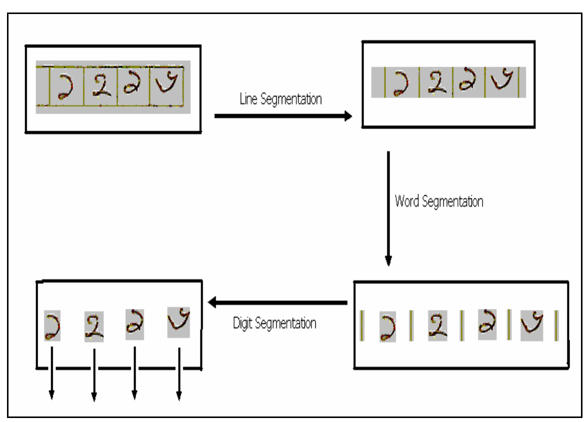
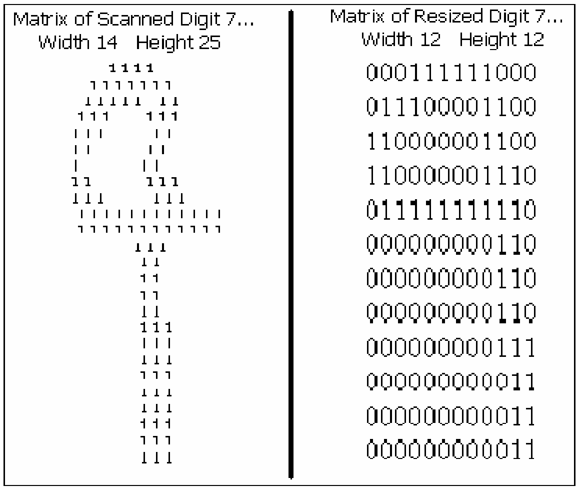


FIGURE 2:
DIFFERENT STEPS IN SEGMENTATION.



(a) (b)

FIGURE 3:
MATRIX REPRESENTATION OF (a)SCANNED AND (b)NORMALIZED BANGLA DIGIT SEVEN '৭'

IV. NEURAL NETWORK

The remainder of the recognition is entirely performed by 3 layered Neural Network [14]. Neural network, in general, is a layered feed-forward network that can be represented by a directed acyclic graph. Each node in the graph stands for an artificial neuron and the labels in each directed arc denote the strength of connection between the neurons and the direction of signal flow. All the connections are adaptive.

The number of neurons in the input layer is 144 as the size of the normalized image is 12 by 12 (144). The number of neurons in the hidden layer was varied during the experiment and best performance was achieved by 100 neurons. The output layer composed of 10 neurons: one per class. When a pattern belonging to class *j* is presented, the desired out is '1' for the *j*th output neuron and '0' for the other output neurons.

Training process using back-propagation algorithm [13] involves teaching the network a predefined set of input-output example pairs by using a two-phase *propagate-adapt* cycle. After an input pattern has been applied as a stimulus to the first layer of network, it is propagated through each upper layer until an output is generated. This output pattern is then compared to the desired output, and an error signal is computed for each output neuron.

The error signals are then transmitted backward from the output layer to each neuron in the intermediate layer that contributes directly to the output. However, each neuron in the intermediate layer receives only a portion of the total error signal, based roughly on the relative contribution the neuron made to the original output. This process repeats, layer by layer, until each neuron in the network has received an error signal that describes its relative contribution to the total error. Based on the error signal received, connection weights are then updated by each neuron to cause the network to converge toward a state that allows all the training patterns to be recognized correctly.

The significance of this process is that, as the network is being trained, the neurons in the intermediate layers organize themselves such that different neurons learn to recognize different features of the total input space. After training, when presented with an arbitrary input pattern that is noisy or incomplete, the neurons in the hidden layers of the network will respond with an active output if the new input contains a pattern that resembles the feature the individual neurons learned to recognize during training. Conversely, hidden-layer neurons have a tendency to inhibit their outputs if the input pattern does not contain the feature that they were trained to recognize.

The *stopping criteria* of BP algorithm selected for the present work is when the *mean squared error* between the actual output of the network and the desired output is less than 1.5%. 5500 different patterns were used to train the network. It took 7182 iterations for the network to converge. However, the number of iterations may vary depending upon the initial weight of the network, which was determined randomly.

V. RESULTS

Testing phase involves presenting the network with an arbitrary input pattern that is noisy or incomplete and it is expected that the neurons in the hidden layers of the network will respond with an active output if the new input contains a pattern that resembles the feature the individual neurons learned to recognize during training.

After training, the network was presented with 490 patterns of each class and the test result is presented in Table 1.

TABLE 1:
TEST RESULT OF 4900 TEST PATTERNS

Bangla Digit	**Correct Output**	**Accuracy (%)**
০ (0)	483	98.6
১ (1)	482	98.4
২ (2)	447	91.2
৩ (3)	371	75.7
৪ (4)	484	98.8
৫ (5)	444	90.6
৬ (6)	460	93.9
৭ (7)	480	98.0
৮ (8)	490	100
৯ (9)	377	76.9

Table 1 shows that all 'eights' (৮) are perfectly recognized by the network. The performance of the network in recognizing the remaining digits, except for digits 'three' (৩) and 'nine' (৯), is quite acceptable, since their accuracy rate is above 90 percent – higher ones being 98.6% for 'zero' (০), 98.4% for 'one' (১), 98.8% for 'four' (৪) and 98% for 'seven' (৭). However, the network fails to recognize digits 'three' (৩) and 'nine' (৯) with high accuracy. It is observed the trained network is confusing 'nine' (৯) with 'one' (১) and 'three' (৩) with 'seven' (৭). In most cases, nine is wrongly recognized as one and three is misclassified as seven. It is also interesting to note that the confusion is unidirectional – i.e. nine is misclassified as one in 15.7% cases whereas one is misclassified as nine only in 1.2% cases (see Table 2). Out of 4900 test patterns, the network recognized 4518 patterns correctly. Thus, the network has an average accuracy rate of 92.2 percent.

The rejection rate of the network is 1.8%. The rejection criterion was that the difference between the activity levels of the two most-active output neurons should be larger than a threshold.

The average response time of the system between the taking of an envelop image and recognizing the postal code is 1.5 seconds.

TABLE 2:
CONFUSION MATRIX OF BANGLA DIGIT CLASSIFIER

	Classified as →									
Digits	**০**	**১**	**২**	**৩**	**৪**	**৫**	**৬**	**৭**	**৮**	**৯**
০ (0)	**483**	0	0	2	0	3	0	2	0	0
১ (1)	0	**482**	1	0	0	0	1	0	0	6
২ (2)	0	3	**447**	4	11	6	2	3	0	6
৩ (3)	2	0	0	**371**	0	3	3	67	5	3
৪ (4)	1	0	0	0	**484**	1	0	0	0	1
৫ (5)	12	4	0	6	0	**444**	11	0	0	3
৬ (6)	0	2	2	5	0	6	**460**	0	4	5
৭ (7)	2	0	4	1	0	2	0	**480**	0	0
৮ (8)	0	0	0	0	0	0	0	0	**490**	0
৯ (9)	0	77	0	5	0	4	0	3	0	**377**

VI. IMPLEMENTATION DISCUSSION

The entire mail sorting system (Figure 4) can be integrated into a mechanical machine controlled by a general purpose computer. The machine has a conveyor belt that transfers all the letters from one place to another. The control panel is attached on one side of the belt and contains a series of "push levers" separated by an equal distance. The optical sensors underneath the belt can detect the arrival of a mail. When activated by the control panel, the lever pushes a mail into a container on the other side of the conveyor belt. A 2D representation of the "letter sorting machine" is given in Figure 5.

First of all, all the mails – arranged in a predefined orientation – are loaded into the machine. The machine takes one letter at a time and places it under a digital camera. The image from the camera is then transferred to the recognition system running in the computer. After successful identification of postal code in the letter, the computer uses a lookup table for cities (for local mail) or countries (for oversea mail) to determine the lever that has to be activated to place the mail in its proper container.

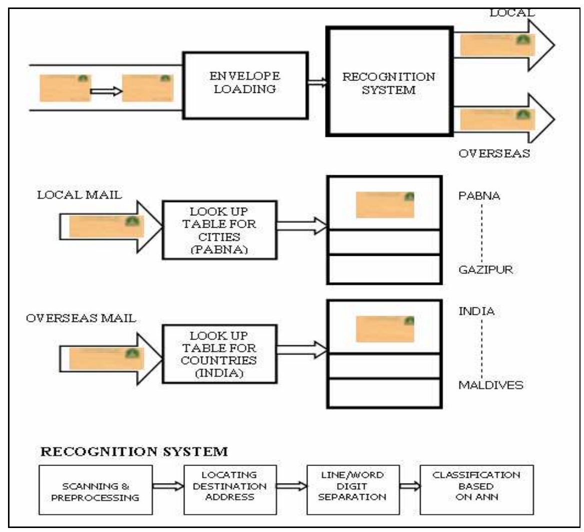


FIGURE 4:
A DESIGN OF MAIL SORTING SYSTEM

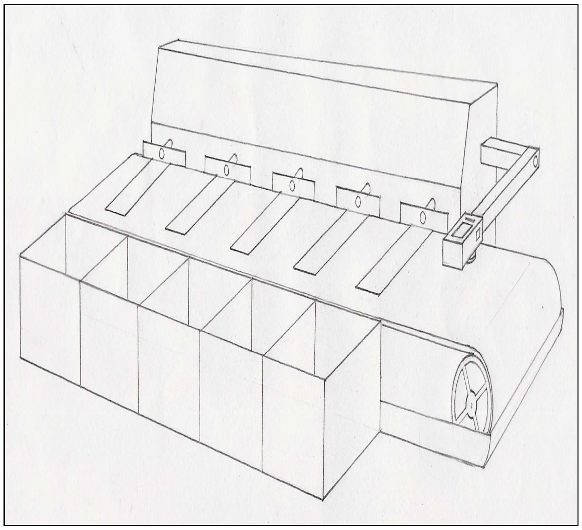


FIGURE 5:
COMPUTER CONTROLLED MECHANICAL MACHINE FOR REGION-WISE SORTING OF MAILS.

VII. CONCLUSION

An efficient recognition system for handwritten Bangla numerals has been developed. In the proposed system, we discuss a classifier based on neural network and experiment results confirm the relative effectiveness of the proposed approach. Then we discuss a simple implementation technique for sorting mails using postal codes. In our future work, we plan to include two or more recognition approaches for achieving better recognition.

VIII. ACKNOWLEDGEMENT

The research was supported by IUT Student Project Grant (FY 2004-2005). We would like to thank the National University Post Office, Board Bazar for allowing us to scan some sample real-life envelopes. We would also like to acknowledge the time and effort given by Mr. Nazam Shahpar for sketching our concept as shown in Figure 5.

REFERENCES

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

[2] Le Cun, Y. (1989). Generalization and Network Design Strategies. In R. Pfeifer, Z. Schreter, F. Fogelman, and L. Steels, editors, *Connectionism in Perspective*, Zurich, Switzerland, 1989. Elsevier.

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

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

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

[6] Pal, U., and Chaudhuri, B.B. (2000). Automatic Recognition of Unconstrained Off-line Bangla Hand-written Numerals. In T. Tan, Y. Shi and W. Gao, Editors, *Advances in Multimodal Interfaces, Lecture Notes in Computer Science* **vol. 1948**, Springer, Berlin (2000), pp. 371–378.

[7] Roy, K., Vajda, S., Pal, U., and Chaudhuri, B.B. (2004). A system towards Indian postal Automation. *Proceedings of the Ninth International Workshop on Frontiers in Handwritten Recognition (IWFHR-9 2004)* (October 2004), pp. 580–585.

[8] Pal, U., Belad, A., and Choisy, Ch. (2003). Touching Numeral Segmentation using Water Reservoir Concept. *Pattern Recognition Lett.* **24** (2003), pp. 261–272.

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

[10] Wang, X., and Tsutsumida, T. (1999). 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.

[11] Srihari, S. N., and Keubert, E.J. (1997). 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.

[12] Wen, Y., Lu, Y., and Shi, P. (2007). Handwritten Bangla Numeral Recognition System and its Application to Postal Automation. *Pattern Recogn.* 40, 1 (Jan. 2007), 99-107.

[13] Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning internal representations by error propagation. In *Parallel distributed processing: Explorations in the microstructure of cognition*, volume I, pages 318-362. Bradford Books, Cambridge, MA.

[14] Freeman, J. A., and Skapura, D. M. (2002). *Neural Networks – Algorithms, Applications, and Programming Techniques*, Pearson Education, 2002.