BENGALI HANDWRITTEN NUMERAL RECOGNITION USING ARTIFICIAL NEURAL NETWORK AND TRANSITION ELEMENTS

Zahidur Rahim Chowdhury[a,*] **Mohammad Abu Naser**[b] **Ashraf Bin Islam**[c]

[a] United International University, Dhaka.
[b] Islamic University of Technology, Gajipur.
[c] Bangladesh University of Engineering and Technology, Dhaka.

* Corresponding email address: zahidur@gmail.com

Abstract: *Bengali hand-writing recognition has potential application in document processing for one the widely used for language in the world. A method using Artificial Neural Network (ANN) is utilized primarily to identify numerals of the language using transition features. Maximum accuracy of 82% is reported in this article for an optimized network. The typical performance of the handwriting recognition system that uses a single recognition scheme is around 85%. The significance of local features in a character should be incorporated to enhance the overall performance of the network.*

Key words: *Bengali Hand-writing, Numeral, Pattern Recognition, Neural Network, Transition.*

INTRODUCTION

Hand-written Bengali character recognition is a process where techniques of pattern recognition are applied to analyze handwritings of Bengali language, one of the most popular languages in the world. Beside Bengali, researchers have studied the recognition process using different techniques for other popular languages like English [1–4], Chinese [5], Arabic [6], Japanese [7, 8], and Indic [9]. Potentials to accelerate the processing of bank cheques, postal addresses, ID forms and manuscripts have drawn researchers' attention toward this field. Numeral recognition is the primary step in order to build a complex hand-writing recognition system that can be utilized in applications like postal code detection, car tracking, price tagging, etc.

Word recognition in 'holistic' approaches is based on the information or image of a complete word [2, 10]. This algorithm performs better to provide auxiliary information but can not work as a stand alone recognizer. On the other hand, Segmentation based techniques generate optimized sections of a word to match and give better performance [11]. Bisnu et al [12] and Pal et al [13] proposed approaches to segment Bengali handwritten characters in order to be utilized in the segmentation based recognition approaches. Moreover, recognition techniques or networks like Modified Quadratic Discriminate functions [7, 8], Hidden Markov Model (HMM) [11, 14] and Artificial Neural Network (ANN) [4] are utilized for both segmentation-based and segmentation-free approaches.

Wen et al [15] proposed two approaches for recognizing handwritten Bengali numerals. One method was based on image reconstruction recognition approach and for the other one, direction feature extraction approach combined with Principal Component Analysis (PCA) and Support Vector Machine (SVM) were utilized. An integrated system utilizing outputs of different approaches (with average accuracy of 88%) was also reported in that article. Recognition results from different systems were compared to make the final decision. The average recognition rate, error rate and reliability achieved by the integrated system were 95.05%, 0.3% and 99.03%, respectively.

In this article, a recognition process for Bengali hand written numerals is presented using 'holistic' approaches due to limited number of possible outputs. Transition features of an image, instead of the complete image, were used as inputs of the ANN for an efficient and a compact recognition system. According to the authors' best knowledge, the transition features were not used to identify Bengali characters. The adaptive nature of the neural network en bled the recognition process to consider the variation of writings and gave the capability to learn from prior experiences. Network parameters were optimized and can be used in a segmentation-based approach for complete hand writing recognition system where words, instead of characters, need to be recognized. The recognition system described in this article may serve as a primary recognition unit for the integrated system proposed by Wen et al [15] in order to provide faster and/or more accurate output.

RECOGNITION PROCESS

A multilayer, feed-forward neural network of two hidden layers and trained with back-propagation were used to recognize Bengali handwritten numerals. Hand-writing samples, collected on a form in a tabular format, were utilized for the off-line recognition system. Transition points were calculated from these samples after necessary preprocessing. The neural network was trained, adapted and tested based on these transition elements.

Sample collection

Samples were collected in custom made forms. Ten numerals were collected in different columns while different rows were allocated for different people. Figure 1

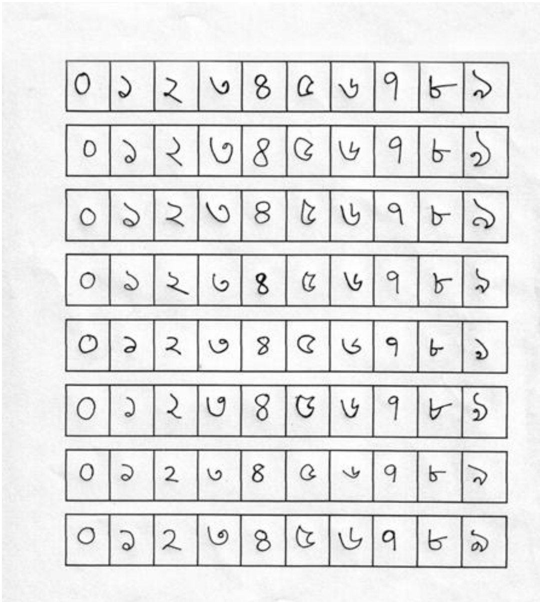

Figure 1. Sample collection form. Different rows are filled with different people.

shows one of those forms. Later these forms were scanned with a resolution of 300 dpi and further processing was performed to remove background noises and other defects.

Preprocessing

Scanned samples, saved in gray-scale mode, were aligned based on the tilt of the image found using Hough transformation [16]. Noise introduced during the scanning process was removed while gray-scale images were converted to binary matrices based on a threshold value obtained from the average gray-scale value of the total sample image. A linear spatial filter was also used to obtain the desired numeral from the background of irrelevant detail. The broken lines of the input characters that caused degradation of performance were avoided by a morphological ‘dilation’ operation in order to ‘grow’ objects in the broken areas. Image with continuous lines can be skeletonized to remove ambiguity and spurious features from the sample. ’bwmorph’ and ’medfilt2’ functions available in the Matlab were utilized to filling, producing the skeleton and filtering noises from the sample. The binary samples at the end of all the processing steps were used for further analysis.

Transition calculation

In transition features, location and number of transitions from background to foreground along vertical and horizontal lines are computed [4]. Transitions can represent a character more efficiently, i.e. with reduced information compared to that required to represent the total image of the character. Moreover, samples with different sizes and aspect ratios can be compared without normalization provided that transition-points are calculated maintaining particular ratios. Normalized values (Figure 2) of N number of transition points of a sample were calculated starting from left to right, bottom to top, right to left and top to bottom while scanning the sample top to bottom, right to left, bottom to top and left to right, respectively. In cases where number of transition points in a particular row, K, were more than N, extra points were ignored as shown in Figure 2. On the other hand, $N - K$ number of zeros were padded in the matrix where K was less than N. A local averaging technique was used to represent the sample transitions by R number of rows for a particular scanning direction that made the total dimension of the transition matrix as $4\times N\times R$ when all four directions were considered.

Transition matrix element calculation along different rows are demonstrated in Figure 2 for $N = 3$. Background of the image is represented by 0 (or white in the image) while 1 (or black in the image) represents the foreground. Normalized values of the transition for ith row are calculated using:

$$T_{ij} = \frac{I_{ij} + L_{col} - 1}{L_{col} - 1} \qquad (1)$$

$$\text{for } i = 1, 2, \mathrm{K}\ , L_{row};\ j = 1, 2, \mathrm{K}\ , L_{col}$$

where, I_{ij} represents the transition values in terms of column numbers. Moreover, L_{col} and L_{row} are the total column and row number, respectively. So transitions at the leftmost column of the matrix during left to right consideration will produce 1 and the rightmost column will produce 0 as normalized transition values. Transitions along different columns can be calculated in a similar manner.

Neural network structure

A four layer neural network with two hidden layers, one input layer and one output layer was used to optimize the N and R. The input layer consisted of a column matrix having $4NR$ elements. The numbers of neurons in the first and second layers were $4NR$ and $2NR$, respectively, in order to maintain a proportional network arrangement while the N and R were varied. Ten numerals, from 0 to 9, were the ten possible outputs giving the total number of elements at the output layer to be ten.

RESULTS

Most samples were collected from an age group of 20 to 25 years. In total, there were 66 sets of sample handwritings and 10 and 15 sets out of those were used for training and adaptation purposes, respectively. The rest of the samples were used to test the performance of the designed network. Separate blocks in a row of the sample collection forms were allocated for different numerals and separate rows were for different sets of handwriting. The primary samples were extracted from a region consisted of 80% of height and width of a cell leaving 10% of the dimensions in four sides of the cell to ensure enough space for large hand-writings. Image representing a particular set of numerals were vertically aligned since tilted images produced distorted sample for training or testing the network. Top margin was difficult to determine for such cases and the margin became a part of the sample along with all the insignificant blank areas in between the sample and the margin. The collected samples were scanned at 300 dpi. Size of the extracted portion of an image for numerals, after removing blank region between the sample and the margin, after all necessary processing varied from 20 to 80 pixels (height or width) based on the handwritings. Initial sample, the sample after background noise reduction and skeleton of the numerals in order to find the transitions are shown in Figure 3.

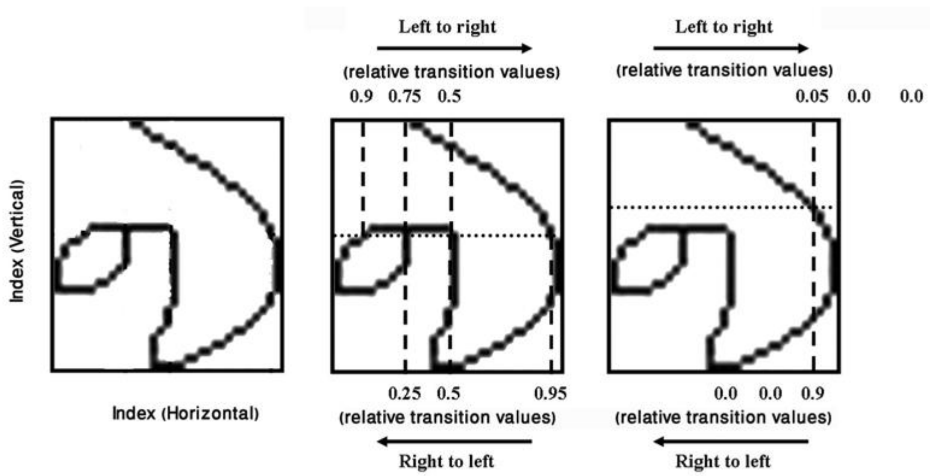

Figure 2. (a) Original image for transition calculation (b and c) transition matrix calculation from left to right and right to left when three transitions were considered. Moreover for (b) the number of transitions, N, considered is less than the number of transitions available in that row, K (c) K < N. The samples were scanned from top to bottom while transitions were calculated from left to right and vice versa.

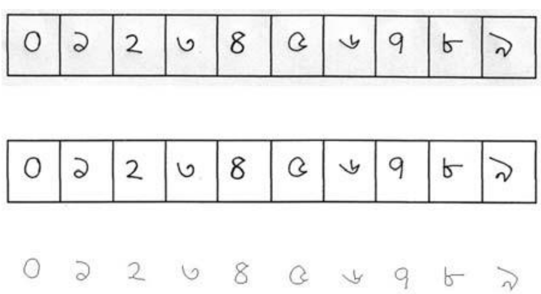

Figure 3. Samples at different stages of the processing. (a) Primary sample (b) background noise reduced (c) skeleton of numerals

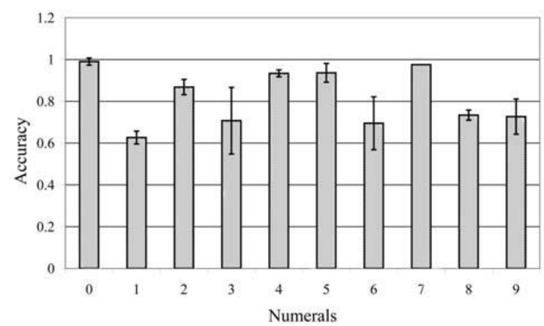

Figure 5. Identification accuracy for different numerals. Error bars represent the standard deviations for different runs.

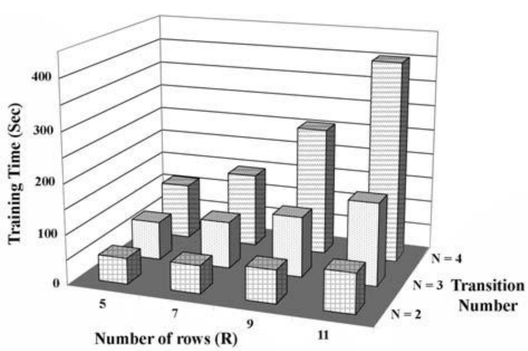

Figure 4. Time required to train the network for different combination of N and R.

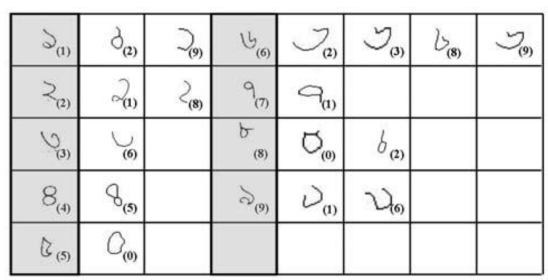

Figure 6. The samples that have full features of a numeral and have been properly recognized are inside gray regions. Moreover, samples that were difficult to identify are in white regions. Output of the neural network is given in printed numbers when the images were fed for recognition.

The number of transition, N, was varied from 2 to 4 during the optimization process. On the other hand, the number of row, R, was varied from 5 to 11 with a step of 2. The network did not converge for $N = 1$ and it also took several attempts for the network to converge for $R = 11$ and $N = 4$. It implies that one transition was not adequate to represent the numerals and the true transition features obtained from the sample became less dominant when more zeros were padded in cases where more number of transitions was considered than required to efficiently represent the sample.

Network performances, i.e. accuracy to recognize the numerals varied insignificantly for the range of parameters chosen for the optimization of N and R as shown in Table 1. There were 10 runs to train, adapt and test the network for each variation. The best performance, 82%, was close to that (85%) reported by Gader et al [4] for English handwriting recognition using transition elements. Moreover, the typical performances of handwriting recognition system, even for Bengali numerals, are 85±5% for approaches that use single scheme [15]. Moreover, as

N	Accuracy			
	***R* = 5**	***R* = 7**	***R* = 9**	***R* = 11**
2	0.78 (± 0.02)	0.80 (± 0.02)	0.82 (± 0.01)	0.80 (± 0.02)
3	0.80 (± 0.02)	0.80 (± 0.01)	0.81 (± 0.01)	0.80 (± 0.01)
4	0.78 (± 0.03)	0.80 (± 0.02)	0.81 (± 0.02)	0.76 (± 0.04)

Table 1. Variation of accuracy with number of transition, *N*, and number of rows to be considered, *R*. Standard deviations are also provided for different runs.

reported by Wen et al [15] the recognition accuracy can be improved to 95% using an integrated system combining three different approaches.

The integration strategy applied in that article is majority voting, i.e., if at least two out of three recognition results are identical. Time required to train, adapt and test the samples were also measured in order to estimate computational resources. The adaptation time for 15 samples and the testing time for around 40 samples were almost constant for different combinations of N and R. The adaptation time was around 1.2 seconds per sample while that for testing was around 0.75 seconds per sample for a Intel (R) Celeron (R) M 1.4 GHz processor with 256 MB of RAM. The training time that varies obviously for the computational load variation for different dimensions of the input matrix are shown in Figure 3. The network with $N = 2$ and $R = 9$ was considered to be the optimum network due to relatively higher accuracy, lower standard deviation and reduced training time.

The recognition accuracy of a particular network also varied with the numerals (Figure 3). Numeral 0, 4, 5 and 7 were recognized with more than 90% accuracy and with greater consistency. Similarity between Bengali numerals like (1, 9), (3, 6) and (5, 6) for made the identification method more difficult. Moreover, as the method described in this article emphasized on the overall shape of numerals instead of local fine features modifications should be made to distinguish numerals with such characteristics. The samples causing erroneous recognition for the approach followed for this article are given in Figure 6.

CONCLUSION

A holistic approach to recognize handwritten Bengali numerals is presented in this article. Neural network was utilized and optimized for such application and transition elements were used to represent the features of different numerals. Around 82% overall accuracy was achieved for the optimized network. Local features made identification process of some of the numerals more difficult. Segmented portion of numerals along with the method presented in this paper should overcome the limitations of the current method where local features had limited significance.

ACKNOWLEDGEMENT

Authors like to thank Engr. Nurul Haque Prodhan for his support and inspiration.

REFERENCES

1. Gillies, A., 1992, "Cursive word recognition using hidden Markov models," *Proceedings of U.S. Postal Service Advanced Technology Conference*, Washington, USA, pp. 557-563.

2. Madvanath, S.,1992, "Using holistic features in hand-written word recognition", *Proceedings of U.S. Postal Service Advanced Technology Conference*, Washington, USA, pp. 51–62.

3. Plessis, B., 1992, Isolated handwritten word recognition for contextual analysis reading, *Proceedings of U.S. Postal Service Advanced Technology Conference*, Washington, USA, pp. 579-593.

4. Gader, P. D., Mohamed, M., and Chiang, J. -H., 1997, "Handwritten Word Recognition with Character and Inter-Character Neural Networks", *IEEE transactions on systems, man, and cybernetics* **27**, pp. 158–164.

5. Liu, H. and Ding, X., 2005, "Handwritten character recognition using gradient features and quadratic classifier and multiple discrimination scheme", *Proceedings of 8th International Conference on Document Analysis and Recognition*, Seoul, Korea, pp. 19–23.

6. Amin, A., 1998, "Off-line Arabic character recog-nition: The state of the art", *Pattern Recognition* **31**, pp. 517–530.

7. Kimura, F., Tsuruoka, S., Miyake, Y. and Shridhar, M., 1994, "A lexicon directed algorithm for Recognition of Unconstrained Handwritten words", *ICICE Transactions On Information and Systems*, **77**, pp. 785–793.

8. Kimura, F., Takashina, K., Tsuruoka, S., and Miyake, Y., 1987, "Modified quadratic discriminant functions and the application to Chinese character recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **9**, pp. 149–153.

9. Rahman, A. F. R., Rahman, R., and Fairhurst, M. C., 2002, "Recognition of handwritten bengali characters: A novel multi-stage approach", *Pattern Recognition*, **26**, pp. 997–1006.

10. Madvanath, S. and Govindaraju, V., 1993, "Holistic lexicon reduction", 3rd International Workshop on Frontiers in Handwriting Recognition, pp. 132–141.

11. Mohamed, M. A., and Gader, P. D., 1996, "Hand-written word rcognition using segmentation-free hidden Markov modeling and segmentation-based dynamic programming techniques", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **18**, pp. 548–554.

12. Bisnu, A., and Chaudhuri, B. B., 1999, "Segmentation of Bangla handwritten text into characters by recursive contour following", *Proceedings of 5th International Conference on Document Analysis and Recognition*, Bangalore, India, pp. 402–405.

13. Pal, U., and Datta, Sagorika, 2003, "Segmentation of Bangla unconstrained handwritten text", *Proceedings of 7th International Conference on Document Analysis and Recognition*, Edinburgh, Scotland, pp. 1128–1132.

14. Chen, M. Y., Kundu, A., and Zhou, J., 1994, "Off-line Handwritten Word Recognition using a Hidden Markov Model Type Stochastic Network", *IEEE transactions on Pattern Analysis and Machine Intelligence*, **16**, pp. 481–496.

15. Wen, Y., Lub, Y., and Shia, P., 2007, "Handwritten Bangla numeral recognition system and its application to postal automation", *Pattern Recognition*, **40**, pp. 99-107.

16. Rosenthal, A. S., Hu, J., Brown, M. K., 1997, "Size and orientation normalization of on-line handwriting using Hough transform", *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Munich, Germany, **4,** pp. 3077–3080.