Segmentation Technique for the Learning in Recognition of the Two Handwritten Bangla Digits Using Counterpropagation Neural Network

Abul L. Haque, Mohammed T. Siddique, Tanvir M. Khan, Imtiaz Ahmed

Dept. of Computer Science & Engineering, North South University

Dhaka, Bangladesh

ahaque@northsouth.edu tanvirnazrul@yahoo.com, n_tanvir@hotmail.com, Imtiaz_81@hotmail.com,

Afsaneh Minaie,

Engineering Department,

Utah Valley State College

minaieaf@uvsc.edu,

Abstract

We are proposing a segmentation technique for the learning of the two Bangla digits incorporating the grid method, the regional search method, the feature extraction method, and the counterpropagation neural network. We emphasized the recognition of the two Bangla digits one and nine since these two digits look similar. The experimental result shows an overall increased rate of recognition for Bangla one and nine.

Keywords: Bangla numeral, Bangla digit, Neural Network, Counterpropagation Neural Network, segmentation

1. Introduction

Bangla is a language spoken by the people of Bangladesh and the people of the state of West Bengal in India. Bangla character recognition using different approaches has become a field of interest in the last few years. In spite of its interests, the progress in this field has not been very fast. So far, only a limited amount of work has been done successfully with Bangle character recognition due to its complexity. Literature survey on the Bangla digit recognition techniques revealed that the simulation results of all the approaches were suffering from the low recognition rates of Bangla one and nine since both of them look similar. Therefore, we paid special attention on the recognition of Bangla digit one and nine. Earlier work of Rahman et al.[9], Syeed et al.[10], Mollah and Talukder[11], and Rahman et al.[12] inspired us to initiate the work in this area. The technique that was formulated here for the improvement of the recognition incorporated the concepts of grid method [6], regional search method [8], and most prominently, the segmentation and feature extraction method [7], and the Counterpropagation Neural Network (CPN) [1, 2, 3, 4, 5] to recognize the Bangla digits one and nine. The simulation result suggested a significant improvement in the recognition rate of Bangla digits one and nine using our technique.

2. Methodology

2.1 Segmentation

The most important part of a character recognition system is the segmentation of the character. We shall consider each Bangla digit as an ordered set of straight line segments each with a certain length, direction, probable place and area. We can consider the most complex digit as the reference digit, which closely resembles the various digits written by different people. For our example, we consider Bangla nine as our reference digit. It can be represented by ten ordered set of straight-line segments which can be denoted by CP1, CP2, CP3, CP4, CP5, CP6, CP7, CP8, CP9, and CP10 respectively. The representation is shown in Figure 1.

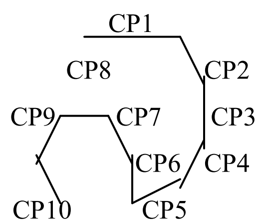

Figure 1. Bangla ”nine”

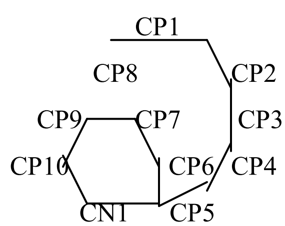

Figure 2. Bangla “one”

We can generalize the rule by representing a digit consisting of M straight line segments by the $CP = (CP_1, CP_2, CP_3 \ldots CP_M)$.

Segments can be defined using the following seven criteria:

1. Direction (d): this is the direction in which a segment may align. We used 10 directions to define all the digits in Bangla alphabet. All these directions can be derived from Figure 1.
2. Weight (w): this can be called as intelligent weighting factor, which is being used for reward or punishment. The weighting factor is positive if the segment is present in the specific digit, and negative otherwise.

3. Length (l): this is the normal length of a segment, derived from the starting and ending coordinates of the segment.
4. Start X (X_1): this is the starting x coordinate of the segment.
5. Start Y (Y_1): this is the starting y coordinate of the segment.
6. End X (X_2): this is the ending x coordinate of the segment.
7. End Y (Y_2): this is the ending y coordinate of the segment.

It is evident that in our work, a segment can be fully defined by (d, w, l, X_1, Y_1, X_2, Y_2). The Bangla nine can be defined by the set of ten segments (CP_1, CP_2, CP_3, CP_4, CP_5, CP_6, CP_7, CP_8, CP_9, and CP_{10}). The CP will stand for all segments having positive weighting factors (w). Any line segment with negative weighting factor (w) will be denoted by CN. CN refers that this line segment is not desired in a specific digit. For example, we can show that an additional CN in the digit Bangla nine will make the digit Bangla one, which is completely undesired in the representation of Bangla nine. Figure 1 shows the layout of Bangla nine having the required segments. Here, all the inputs with ten segments of the format in Figure 1 will be candidates for Bangla nine, whereas those containing eleven line segments with CP1 twice instead of CN1 will be candidates for Bangla one.

3. Parameters for the Counterpropagation Neural Network

In our proposed model, we require a few neurons or processing elements in the input layer. These neurons will be used to provide the input parameters of the straight line segments that we shall use to represent the digits in Bangla numerals. We also need ten neurons in the output layer to represent which digit has been detected by the system. Only one among the ten output neurons will be fired each time a digit is recognized. In our Kohonen layer or hidden layer, we shall use the same number of neurons. Each neuron in this layer will be fired based upon the clusters of the input pattern. When the inputs in the input layer will be processed and will be categorized as a member of one of the ten input clusters, the corresponding Kohonen layer will be fired to declare the specific cluster as ‘winner’. The following two training equations are used to train the Kohonen layer and Grossberg layer respectively.

Equation (2.1) is used to train the Kohonen Layer:

$$W_{new} = W_{old} + \alpha\,(X - W_{old}) \qquad (2.1)$$

Here,

W_{new} = new value of a weight connecting an input component x to the winner neuron.

W_{old} = previous value of this weight.

X = Input vector.

α = a training rate coefficient that may vary during the training process.

The following equation is used to train the Grossberg layer:

$$V_{ij\ new} = V_{ij\ old} + \beta\,(Y_j - V_{ij\ old})\,k_i \qquad (2.2)$$

Here,

K_i = the output of the Kohonen neuron i, which is the only nonzero output neuron.

Y_j = component j of the vector of desired outputs.

These two equations are used for the simulation.

4. Digital Recognition Technique

The input digit is processed for digitization and segmentation following the method described above. Once the preprocessing for segmentation is finished, we start working with the segment components. All the segments of a digit are used as input parameters to the input neurons. The system will automatically train the input sets for the Kohonen layer and then the similar set of inputs will be grouped together to fire a specific Kohonen neuron. The system will fire the same Kohonen neuron for a group of similar inputs.

Once a specific Kohonen neuron is fired for an input, the input for the Grossberg layer will be nonzero for only that specific fired neuron. All other inputs to the Grossberg layer will be zero. So the Grossberg layer adjusts weights only for the connection lines with the winning neuron of the Kohonen layer. Since Grossberg layer is trained with supervised training approach, it will identify the digit involved with that fired Kohonen neuron.

5. Simulation and Experimental Results

The handwritten Bangla numerals that we used in the simulation are shown in Figure 3. During the course of the experiments, we had to make some modifications of our system. Instead of representing the segments using those seven parameters, we had used matrices of size 7×5 each. In the grid method discussed previously [6], we found that to represent a segment, 3×3 matrix is sufficient. In order to increase the efficiency rate, we tried to increase the matrix size. We first started with a 10×10 matrix, which was found to overflow during training session. So we continued with reduced sized matrices. After some continual experiment, we came to the conclusion that 7×5 was the best-fit matrix for our purpose. Here a 1 in the corresponding matrix component represents occurrence of each segment and the rest of the components are represented by 0's. thus we had to use only 35 neurons in the input layer to represent a sample input. The Kohonen layer is trained in unsupervised training method, which is done by machine-generated noises involves standard deviation and mean calculation of input matrices. The noise function we used here is a built in function of MATLAB. The Grossberg layer training is done by taking the noise inputs and then with the target values set to the output layer.

The experiment we performed was specially for improving the recognition rates of Bangla one and nine. Along with the segmentation method mentioned above, we incorporated the concept of grid method [6], and the regional search method [8] too. Here the concept of matrix that we used came from the grid method. In the grid method the entire character was divided into some regions represented by a 3×3 matrix, whereas we have used 7×5 matrix for our purpose. Moreover, to improve the success rate, we have used the concept of down edge segment mentioned in the regional search method. To differentiate between one and nine we checked for the presence of down edge segment in the sample. Since the difference can be recognized from the presence of CN1 segment, we represented this segment by down edge segment, and fired the neuron for one when the segment was absent and the neuron for nine when it was present. This significantly improved the recognition rate of one and nine. The results are shown in the following table.

Table

Sample Bangla Digit	Success With Noise	Success Without Noise
0	80	86
1	94	95
2	90	88
3	80	84
4	95	96
5	90	91
6	88	89
7	88	87
8	80	82
9	94.5	93

6. Comparison of Our Results

The simulation result has been very impressive. Using the counterpropagation neural network algorithm for training, the performance of recognition has been significantly improved. The overall success rate of recognition of digits has been improved to 88%, which was below 85% for the previously proposed method of segmentation. Our objective of improving the recognition rate of Bangla one and nine has also been accomplished. Our new technique has achieved the approximate success rate of 94% for Bangla one and 93% for Bangla nine compared to the segmentation method [7] that showed the success rate of 74% and 71% respectively.

7. Conclusions and Future Work

Our goal was to improve the success rate of recognition of handwritten Bangla digits one and nine. We gradually progressed through applying counterpropagation neural network algorithm to train the system. While simulating, we found that the technique we used showed much superior performance showing significant improvements. This work can be extended to Bangla alphabets and other special character recognition and also for the recognition of other language characters too.

Bibliographic Information

[1] Wasserman, P.D., *Neural Computing*, Van Nostrand Reinhold, 1989

[2] Hecht-Nielsen, R., *Neurocomputing*, Addison Wesley, 1990

[3] Freeman, J.A., and D. M. Skapura, *Neural Networks,* Addison Wesley, 1991

[4] Principe, J.C., N. R. Euliano, and W. C. Lefebvre, *Neural and Adaptive Systems*. John Wiley & Sons Inc., 1999

[5] Mehrotra, K., C. K. Mohan, and S. Ranka, *Elements of Artificial Neural Networks,* Penram International Publishing (India), 1997

[6] Islam, M. M., S. Rahman, A. N. M. E. Rafiq, and M. M. Islam, "On-line Handwritten Bangla Numeral Recognition by Grid Method", *Proceedings of Second International Conference on Electrical and Computer Engineering (ICECE 2002)*, Dhaka, Bangladesh, pp 269-272, 2002

[7] Rahman, A. K. M. A., and C. M. Rahman. "Image Compression With a Counter Propagation Network", *Proceedings of International Conference on Computer and Information Technology (ICCIT2000)*, Dhaka, Bangladesh, pp 215-220, 2001

[8] Moshad, M. A. A., and M. M. Ali, "Recognition of Handwritten Bangla Digits by Intelligent Regional Search Method", *Proceedings of International Conference on Computer and Information Technology (ICCIT 2001)*, Dhaka, Bangladesh, pp 297-302, 2001

[9] Rahman, C. M., A. Akhter, and A. Wasif, "Handwritten Bangla Digit Recognition using a Neural Network", *Proceedings of International Conference on Computer and Information Technology (ICCIT 2000)*, Dhaka, Bangladesh, pp 221-225, 2001

[10] Syeed, M.M.M., F.H. Siddiqui, A.S.A. Al-Mamun, S.K. Tanbeer, and M. A. Mottalib, "Bengali Character Recognition using Bidirectional Associative Memories (BAM) Neural Network", *Proceedings of International Conference on Computer and Information Technology (ICCIT 2002)*, Dhaka, Bangladesh, pp 247-251, 2002

[11] Molla, M.K.I., and K.H. Talukder, "Bangla Number Extraction and Recognition from Document Image", *Proceedings of International Conference on Computer and Information Technology (ICCIT 2002)*, Dhaka, Bangladesh, pp 200-206, 2002

[12] Rahman, M.O., M.A. Azim, M.S. Chowdhury, and M.N. Islam, "Different Segment Displays for Bangla, English and Arabic Digits", *Proceedings of International Conference on Computer and Information Technology (ICCIT 2003)*, Bangladesh, pp 299-302, 2003

Biographical Information

ABUL L. HAQUE is an associate professor of Computer Science at North South University, Dhaka, Bangladesh. He received his M.Sc. and Ph.D. degrees in Computer Science from the University of Western Ontario, London, Canada, and University of Oklahoma, Norman, Oklahoma, USA. His research interests are in the areas of Neural Networks, Artificial Intelligence, Analysis of Algorithms, and Computer Architecture.

MOHAMMED TANVIR SIDDIQUE received the BS in Computer Science from North South University, Dhaka, Bangladesh, in 2004. His research interests include Neural Network, Network Security, and Algorithm Design.

TANVIR MAHMOOD KHAN is currently pursuing a MS degree in Computer Science at University of Regina, Canada. He received the BS in Computer Science from North South University, Dhaka, Bangladesh in 2004.

IMTIAZ AHMED received the BS in Computer Science from North South University, Dhaka, Bangladesh in 2004.

AFSANEH MINAIE is an associate professor in the Engineering Department at Utah Valley State College. She received a B.S., M.S. and Ph.D. all in Electrical Engineering from University of Oklahoma in 1981, 1984 and 1989 respectively. Her current interests are in Computer Architecture, Embedded Systems, Digital Signal Processing, Digital Design, and Computer Interfacing.

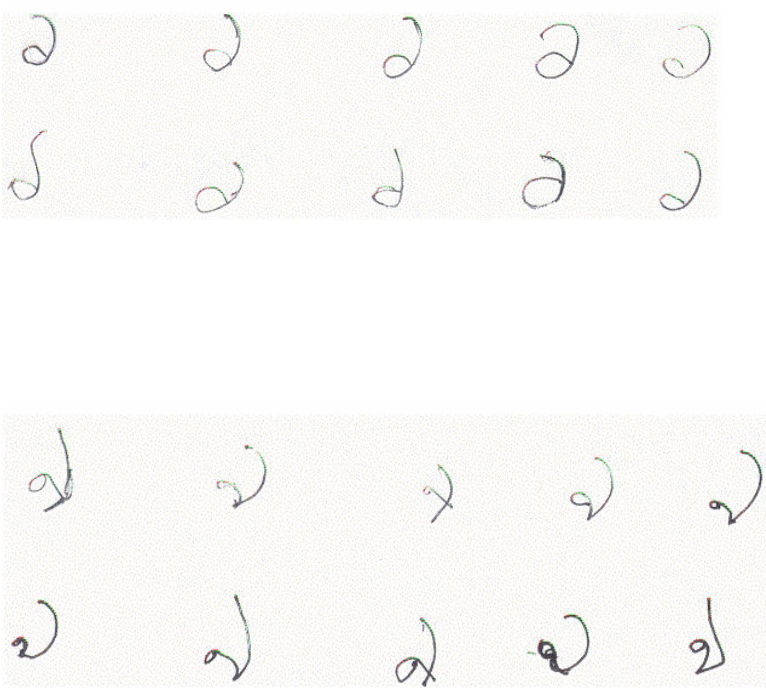

Figure 3. A sample of handwritten Bangla "one" (top) and "nine" (bottom)