

Recognition of an Indian Script using Multilayer Perceptrons and Fuzzy Features

Shamik Sural, Data-Core Systems, 3700 Science Center, Philadelphia, PA 19104, USA.
P.K.Das, Dept. of Comp. Science & Engineering, Jadavpur University, Calcutta 700032, India.

Abstract




We present a multi-stage character recognition system for an Indian script, namely, Bengali (also called Bangla) using fuzzy features and multilayer perceptrons (MLP). The fuzzy features are extracted from Hough transform of a character pattern pixels. We first define a number of fuzzy sets on the Hough transform accumulator cells. The fuzzy sets are then combined by t-norms to generate feature vectors from each character. A set of fuzzy linguistic vectors is next generated from these feature vectors. The MLPs used for classification have the fuzzy features as inputs. The MLP outputs also represent the belongingness of an input pattern to different fuzzy character pattern classes. To improve the recognition accuracy of Bengali characters, we divide all the patterns into three distinct sets. Each set of characters is once again divided into a number of mutually exclusive character pattern classes. During recognition, the class of each pattern is first determined, followed by recognition of the actual character within that class. Recognition accuracy of the system is more than 98%.

Keywords – Bengali character recognition, Multilayer perceptron, Hough transform, Fuzzy sets.

1. Introduction

A number of Indian languages use two major scripts, namely, Devnagari and Bengali (also called *Bangla*). We have worked on the recognition of Bengali characters. About 200 million people of eastern India and Bangladesh use the *Bangla* script. In recent years, Chaudhuri and Pal [2,9] have proposed an optical character recognition (OCR) system for recognition of the *Bangla* script. An OCR system to read both *Bangla* and Devnagari has also been proposed by them [3]. Other areas of work include processing of scripts containing text and non-text blocks as well as segmentation and recognition of handwritten text [1,8].

In *Bangla*, the number of characters (a character is called a *varna*) is large, although there is no upper and

lower case differentiation. Two or more characters often combine to form a new compound character. As a result, the total number of patterns to be recognized is about 300. There are 11 vowels (a vowel is called a *swara varna*) and 39 consonant characters (a consonant is called a *byanjan varna*) in modern *Bangla* alphabet. A vowel following a consonant takes a modified shape, which could be to the left, right (or both), top or bottom of the consonant. These are called vowel modifiers. In some situations, a consonant following or preceding a consonant is represented by a consonant modifier. In some cases, two or more consonant characters are combined to form a shape different from either of the participating basic characters. These are called *yuktakshars*. There are about 250 *yuktakshars* in the Bengali script. As an example, the compound character  is formed by combining the consonants  and . In present day *Bangla* text, the compound character occurrence is in the range of 4-6 %, the rest being basic characters and modifiers [9]. This distribution is considered while designing a classifier for the *Bangla* script.

We had earlier reported a neuro-fuzzy technique for the recognition of Bengali characters [11]. We have extended this work by further fine-graining the classification process during recognition. To improve recognition accuracy for the large character set, a multi-stage recognition strategy has now been formulated in which the characters are divided into three sets and those in each set are clustered into a number of classes. During recognition, the class of each character is first determined followed by identification of the character within that class. This new system for the recognition of characters by multilayer perceptrons [6] using fuzzy features is presented in this paper.

We describe our strategy for multi-stage character recognition in the next section. A method of fuzzy feature extraction and recognition of Bengali characters using these features and multilayer perceptrons is covered in section 3 of the paper. We present the results and draw conclusions in the last section.

2. Multi-stage recognition of Bengali characters

Since the number of characters to be recognized in the Bengali script is large, it is impractical to employ a single multilayer perceptron to classify all these patterns. We, therefore, use a hierarchy of MLPs to detect the characters from a segmented Bengali document image. All the character patterns are divided into three sets. The first set consists of the *swara* and the *byanjan varnas*, the vowel modifiers and two consonant modifiers. There are in all 53 patterns in the first set. The second set consists of 57 *yuktakshars* while the third set consists of another 71 *yuktakshars*. We employ three sets of MLP where the first set of MLP is used to recognize the first set of characters. Similarly, the second and the third sets of MLPs are trained with the second and the third sets of characters, respectively. Within each set, we divide the characters into a number of classes. Each class consists of similar looking characters belonging to that set.

Two stages of multilayer perceptrons are used to detect characters within each set. The first stage MLP detects the class of each character belonging to that set. Once the class of a character is identified, a second stage MLP trained with all the characters of this class is used to recognize the particular character within that class. Thus, the number of perceptrons in the second stage for each set of characters is the same as the number of classes in that set. All the multilayer perceptrons are trained with fuzzy features as explained in the next section.

3. Fuzzy feature extraction and MLP based classification

We use Hough transform [4] to extract fuzzy features from each Bengali character both during training and recognition. An important observation on Hough transform is that it provides three characteristics of a straight line in an image. These are the values of ρ , θ and count of a $(\rho-\theta)$ accumulator cell used for Hough Transform implementation. Here ρ and θ specify the position and orientation of a straight line, while count specifies the length of the line. If an input character pattern is corrupted by noise, some of the features may be missed out due to the thresholding done on the accumulator cell counts. To overcome this problem, instead of thresholding, we define fuzzy sets to extract information from the Hough transform accumulators. These fuzzy set membership functions are listed in table I for θ values in the first quadrant. Similar membership functions are defined for θ values in the other quadrants.

Based on the basic fuzzy sets defined on Hough transform accumulator, a number of new fuzzy sets are next synthesized using t-norms to represent each line as a

combination of its length, position and orientation. We use the standard intersection: $i(p,q) = \min(p,q)$ as the t-norm [5]. Similar basic and synthesized fuzzy sets are defined on the (a,b,c) accumulator cells for circle extraction using the transform

$c = \sqrt{(x-a)^2 + (y-b)^2}$ where (a,b) denotes the centre of a circle and c , its radius. A non-null support of a synthesized fuzzy set implies the presence of the corresponding feature in a character pattern. We, therefore, choose the height of each synthesized fuzzy set to define a feature element and the set of 'n' such feature elements constitute an n-dimensional feature vector for the character. The feature vectors extracted from a set of Bengali characters form the input of an MLP trained with that set of characters.

We use fuzzy feature vectors as MLP inputs and fuzzy pattern classes as its outputs. When a feature vector, as defined above, is extracted from a degraded character pattern for recognition, the strength of the features in the vector may vary due to the presence of noise. To combat the effect of noise, we generate membership values in three linguistic fuzzy sets, namely, *weak*, *moderate* and *strong* from the individual feature elements to form the fuzzy MLP inputs. The linguistic set membership functions are derived from the Butterworth polynomials for low pass, band pass and high pass filter transfer functions [7] as shown below.

$$\begin{aligned} \mu_{\text{weak}}(x) &= [1+(x/a)^{2m}]^{-1/2}, \\ \mu_{\text{moderate}}(x) &= ([1+(x/a_1)^{2m}][1+(a_2/x)^{2m}])^{-1/2}, \\ \mu_{\text{strong}}(x) &= [1+(a/x)^{2m}]^{-1/2} \end{aligned} \quad (1)$$

A processing step before the actual MLP input layer, converts each feature element into membership values in these linguistic sets, generating a $3n$ -dimensional fuzzy feature vector from the n -dimensional feature vector. These $3n$ -dimensional vectors form the input pattern for the MLP. The advantage of using linguistic features is that, for small variations in the feature values, the linguistic set membership values remain unchanged. The system can thus recognize even highly degraded character patterns. There is an inherent dichotomy in the requirements of a character recognition system (i) To be able to distinguish between similar looking patterns and (ii) To correctly identify a character even in the presence of noise. By defining fuzzy features we effectively draw a balance between these two requirements.

We define the MLP outputs to represent fuzzy pattern classes and the MLP is trained to learn the degree by which an input linguistic feature vector belongs to each of these classes. When the MLP is trained with the sample Bengali character patterns, the expected outputs corresponding to each input pattern is computed based on a distance measure between the input feature vector and the feature vector of the character represented by the particular output unit. The membership function denoting

belongingness to the different character pattern classes for an input pattern is determined as explained below.

Consider a P-class problem domain with P nodes in the output layer of the MLP. Each character pattern, before converting to linguistic sets, is represented by an n-dimensional feature vector \bar{F}_i . The Euclidean distance between \bar{F}_i and other feature vectors is calculated as follows.

$$d_{ik} = \left[\sum_j (F_{ij} - F_{kj})^2 \right]^{1/2} \quad k = 1, 2, \dots, P \quad j = 1, 2, \dots, n \quad (2)$$

We use eq. (2) to calculate the distances of all the P patterns from the i^{th} input pattern where the summation is done over all the feature elements subscripted by j. The k^{th} expected output of the MLP for the input vector \bar{F}_i is equal to the membership degree of the i^{th} character pattern to the k^{th} fuzzy pattern class. These expected outputs are determined using the relation:

$$O_{k(\text{exp})}^i = \mu_k(\bar{F}_i) = 1 / [1 + (d_{ik}/f_{\text{den}})^{f_{\text{pow}}}] \quad (3)$$

Here the parameters f_{den} and f_{pow} control the degree of membership by which the pattern \bar{F}_i belongs to the different output fuzzy sets. It is seen that, $\mu_k(\bar{F}_i) \in [0, 1]$, $\mu_k(\bar{F}_i) = \mu_i(\bar{F}_k)$, $\mu_k(\bar{F}_k) = 1$, and $d_{ik} \geq d_{il} \Rightarrow \mu_k(\bar{F}_i) \leq \mu_l(\bar{F}_i)$. Further, for $f_{\text{den}} \rightarrow 0$ and $f_{\text{pow}} \rightarrow \infty$, the fuzzy MLP output reduces to a conventional MLP output with $O_{k(\text{exp})}^i = 1$ for $i = k$, and 0 otherwise. The MLP is trained with the input fuzzy feature vectors and fuzzy expected outputs by the back propagation algorithm [10].

The recognition decision is based on the α -cuts of the output fuzzy sets for $\alpha = \text{th_opt}$, a threshold value for the MLP outputs. The α -cuts are determined by the parameters f_{den} and f_{pow} of eq. (3). For $f_{\text{den}} > d_{ik}$, the number of elements in the α -cuts increases with higher values of f_{pow} . For a fixed value of the parameters f_{den} and f_{pow} , if th_opt is low, α -cuts of the output fuzzy sets may contain more than one element while a high value of th_opt results in null α -cuts for some or all of the outputs. Currently in our system, we are using only the highest value of output for each MLP for the recognition of Bengali scripts. We, therefore, choose the MLP output by making $f_{\text{den}} \rightarrow 0$ and $f_{\text{pow}} \rightarrow \infty$. However, if a dictionary search step is built to correct errors in the MLP based classification, fuzzy membership values can be used in the MLP output also. If each of the α -cuts contain one element, then during recognition only one of the MLP outputs goes above the threshold. The highest value output is then considered as the detected Bengali character. If, however, the membership value is above the threshold for more than one output, it indicates a

misclassification and the character is marked for correction by user intervention. All the high value outputs may also be considered for a dictionary search to uniquely identify the character. The advantage of using fuzzy set membership functions at the MLP output is that, only outputs with high membership values need to be considered for the search. Since the MLP outputs denote their similarities to the different pattern classes, this decision making process is justified.

We classify characters based on the fact that an MLP output shows very high values for the patterns learnt by it and much less values for unknown characters. Taking advantage of this MLP behavior and considering the distribution statistics of simple and compound Bengali characters, we submit the same input pattern to all the three sets of MLPs and each set returns its guess along with a measure of goodness of the guess. The measure of goodness used is:

$G = \text{Max}(\text{First stage class detector MLP output} + \text{Second stage character detector MLP output})$.

After results from all the three sets of MLPs have been obtained, the results are compared and the set which has the highest measure of goodness, is considered to be the most reliable. The character detected by this set of MLPs is considered as the recognized character. It should be noted, however, that the feature extraction process is done only once and the same set of features is used for all the perceptrons. Also, the selection of a new set of MLPs is a simple process of reading in a new set of weights for the inter-connecting links. Thus, resource requirement in our approach of multi-stage recognition is the same as that for a single MLP. The fuzzy MLPs are structurally similar to a crisp perceptron with the number of output units same as the number of *distinct* patterns to be classified in each stage.

4. Results and conclusions

In the implemented system, we use a 300 dpi flatbed HP scanner to scan the Bengali documents from different sources and a 233 MHz Intel Pentium machine for character recognition. There is a large variation in the printing quality of the scanned documents since the materials are composed manually as well as using monotype and linotype methods. Typical printed Bengali documents have uneven line spacing, varying thickness and they often contain broken fonts.

Our system has the advantage that the time required for recognition of the vowels, consonants, modifiers and compound characters is almost equal, being the propagation time of the MLPs after feature extraction. Thus, the speed of recognition in this system can be very accurately predicted for any combination of simple and compound characters in a Bengali document. As the fuzzy sets perform an implicit normalization of the character

dimensions, the system is robust enough to handle large point size variations for the same font. Table II lists the performance of the system in terms of recognition accuracy for different groups of Bengali characters. Since some of the *yuktakshars* are similar looking, the MLPs are unable to distinguish them correctly in a few cases. However, as the *yuktakshars* occur only sparsely in Bengali script, the overall accuracy of the system is about 98.5%. The average speed of recognition is 175 characters per minute. The results show a better performance compared to other such systems that have about 96% accuracy without a dictionary based error correction stage as well as a lower speed of recognition.

The recognition accuracy of our system can be even higher if a post-processing step using dictionary search is introduced to resolve the ambiguous characters. Fuzzy feature extraction using Hough transform as suggested here, is a generic method. It can be used in many other pattern recognition problems besides Bengali character recognition. We have successfully used it in Roman character recognition where dictionary based error correction has been implemented to fully utilize the advantages of fuzzy MLP outputs. We plan to combine the two systems to design a bilingual OCR capable of reading Roman and Bengali scripts in the same document.

Table I Fuzzy set membership functions for a character pattern of height X and width Y.

Fuzzy Set	Membership Function
Long line	$\text{Count}/[(X^2+Y^2)^{1/2}]$
Short line	$2LL$ if $\text{count} \leq [(X^2+Y^2)^{1/2}]/2$ $2(1-LL)$ if $\text{count} > [(X^2+Y^2)^{1/2}]/2$
Left inclined line	1 if $\theta \leq 90.0$ 0 if $\theta > 90.0$
Right inclined line	1-LI
Nearly horizontal line	$\theta/90.0$
Nearly vertical line	1-HL
Slant line	$2HL$ if $\theta \leq 45.0$ $2(1-HL)$ if $\theta > 45.0$
Line near origin	ρ/X if $HL \geq VL$ ρ/Y if $VL > HL$
Line far from origin	1-NO
Line near top border	ρ/X if $HL > VL$ 0 otherwise
Line near bottom border	1-NT if $HL > VL$ 0 otherwise

Fuzzy Set	Membership Function
Line near vertical center	$2NT$ if $(HL > VL \text{ and } \rho \leq X/2)$ $2(1-NT)$ if $(HL > VL \text{ and } \rho > X/2)$ 0 otherwise
Line near right border	ρ/Y if $VL > HL$ 0 otherwise
Line near left border	1-NR if $VL > HL$ 0 otherwise
Line near horizontal center	$2NR$ if $(VL > HL \text{ and } \rho \leq Y/2)$ $2(1-NR)$ if $(VL > HL \text{ and } \rho > Y/2)$ 0 otherwise

Table II Recognition accuracy of the fuzzy OCR system for Bengali characters.

Group of characters	Recognition accuracy (%)
Vowels, consonants and modifiers	99.2
First set of compound characters	92.6
Second set of compound characters	86.2
Overall	98.5

References

1. Bishnu and B. B. Chaudhuri, "Segmentation of Bangla Handwritten Text into Characters by Recursive Contour Following", *Proc. Fifth International Conference on Document Analysis and Recognition*, pp. 402-405, 1999.
2. B.B. Chaudhuri and U. Pal, "A Complete Printed Bangla OCR System," *Pattern Recognition*, vol. 30, pp. 531-549, 1997.
3. Chaudhuri and U. Pal, "An OCR System for Two Indian Language Scripts: Bangla and Devnagari (Hindi)", *Proc. Fourth International Conference on Document Analysis and Recognition*, pp. 1011-1015, 1997.
4. J. Illingworth and J. Kittler, "A Survey of the Hough Transform," *Computer Vision, Graphics and Image Processing*, vol. 44, pp. 87-116, 1988.
5. G.J. Klir and B. Yuan, B, *Fuzzy sets and fuzzy logic - Theory and applications*, Prentice Hall Inc., Englewood Cliffs, NJ, USA, 1995.
6. R. P. Lippmann, "An Introduction to Computing with Neural Nets," *IEEE ASSP magazine*, pp. 4-22, 1987.

7. J. Millman and C. C. Halkias, *Integrated Electronics : Analog and Digital Circuits and Systems*, McGrawhill, Singapore, 1972.
8. U. Pal and B. B. Chaudhuri, "Automatic Separation of Machine-printed and Handwritten Text Lines", *Proc. Fifth International Conference on Document Analysis and Recognition*, pp. 645-648, 1999.
9. U. Pal and B. B. Chaudhuri, "Computer Recognition of Printed Bangla Script," *International Journal of System Science*, vol. 26, pp. 2107-2123, 1995.
10. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning Internal Representation by Error Propagation," in: D. E. Rumelhart and J. L. McClelland, Eds., *Parallel Distributed Processing : Explorations in the microstructure of cognition*, Vol. 1 : Foundations, Chapter 8, MIT Press, 1986.
11. S. Sural and P. K. Das, "An MLP using Hough Transform based Fuzzy Feature Extraction for Bengali Script Recognition," *Pattern Recognition Letters*, vol. 20, pp. 771-782, 1999.