Handwritten Bangla Word Recognition using Elliptical Features

Showmik Bhowmik[1], Samir Malakar [2]
[1]Department of Computer Sc. and Engg.,
Dumkal Institute of Engineering and Technology,
Dumkal, India.
[2] Department of Master of Computer Application,
MCKV Institute of Engineering,
Howrah, India.
{showmik.cse, malakarsamir}@gmail.com

Ram Sarkar, Mita Nasipuri
Department of Computer Sc. and Engg.,
Jadavpur University,
Kolkata, India.
{raamsarkar, mitanasipuri}@gmail.com

***Abstract*—In the present work, a holistic word recognition technique is proposed for the recognition of the handwritten Bangla words. Holistic word recognition technique assumes a word as a single and indivisible entity and extracts features from the entire word to recognize it. In this work, a set of elliptical features is extracted from handwritten word images to represent them in the feature space. Then, a comparison among 5 well known classifiers is carried out in terms of their accuracies to select the suitable classifier for evaluating the present work. Based on that, finally, a neural network based classifier is chosen for the recognition task. Using the elliptical features, the proposed system provides a satisfactory result on a small dataset.**

Keywords—Holistic word recognition; handwritten words; Elliptical features; Bangla script;

I. Introduction

Handwritten Word Recognition (HWR) is a process of automatic reading of handwritten words from document images. HWR always magnetizes the attention of the researchers due to its wide range of real life applications, such as Library / Office automation, Postal automation, Keyword extraction, Document categorization etc. Two major areas that come under the preview of HWR are 1) *Online* HWR and 2) *Offline* HWR.

In case of *online* HWR systems, words are written on an electronic device using some special pen. The recognition algorithm uses some temporal information like position, pressure and direction of the pen for the recognition of the handwritten word. On the other hand *Offline* HWR systems recognize the words at later time after it was inscribed. Words from the handwritten documents are much harder to recognize compare to the printed words due to its complex nature. More specifically, handwritten words may be cursive, discrete, touching or combination of these.

The approaches followed by the experts in the literature of HWR can broadly be classified into two major categories namely the *Analytical approach* [1] and the *Holistic approach* [2]. In *Analytical approach*, a word is considered as a collection of small units such as characters and/ or character subparts and a given word is first segmented into such small units, which are then recognized to form recognized words. On the other hand, in *Holistic approach*, a word is considered as a single and indivisible unit. So, in this approach the features of the entire word image are used for the recognition purpose. Here, the individual word is treated as an individual class. As it is a segmentation free approach, it evades most of the problems like *word segmentation ambiguity, variability of segment shape* which generally occurs in analytical approach [1, 2]. Despite of the above fact most of the holistic approaches in the offline HWR domain, as found in the literature, have been applied over limited vocabularies and static lexicon size. In this work, a holistic HWR system is proposed for the recognition of handwritten Bangla word images containing major city names of West Bengal, India.

II. PREVIOUS WORKS

The problem of HWR is one of the most challenging research areas in pattern recognition domain and has been studied for several decades. In [2], authors presented a survey on the holistic paradigm in human reading and its applications in handwritten word recognition. They show that holistic word recognition is a viable alternative to the popular analytical (segmentation-based) approach to handwriting recognition. In [3], a holistic approach for word recognition is described using right-left discrete HMM and Kohonen self-organizing feature map (SOFM) to recognize the Arabic words. The work described in [4] involves an accurate extraction and representation of three zones, *namely*, lower, upper and central zones from the off-line cursive word to obtain a descriptor which provides a coarse characterization of word shape. The recognition system is primarily based on the sequential combination of Hopfield model and Multi-Layer Perceptron (MLP) based classifiers for prototype recognition leading to handwritten word recognition. This application is used in processing of poor quality bank cheques. A system towards Indian postal automation based on the recognition of multilingual pin-code and city name of the postal document has been proposed in [5]. The application narrated in [6] is the recognition of the Portuguese handwritten names of the months. This paper brings a contribution to the problem of efficiently recognizing handwritten words from a limited size lexicon. For that, a multiple classifier system has been developed that analyzes the words from three different approximation levels, in order to get a computational approach inspired on the human reading process. In [7], authors proposed a longest-run based holistic feature set that has been used to classify word images belonging to different classes, using a neural network based classifier. To evaluate the

978-1-4799-6929-6/14 $31.00 © 2014 IEEE
DOI 10.1109/CICN.2014.66

technique, few words from the handwritten documents of the *CMATERdb*1.2.1 [8] dataset have been used. Frequently occurring English words are manually extracted from the handwritten pages and the accuracy of the technique is evaluated using a threefold cross-validation method.

Among the 22 official languages available in India, Bangla is second most popular language. With nearly 230 million native speakers, it is one of the most spoken languages (ranking 7th) in the world. This is the national and official language of Bangladesh. It is also the official language of the states like West Bengal and Tripura in India. Bangla is an Indo-Aryan language of the eastern Indian subcontinent, evolved from the Bramhi script. Modern Bangla script has 11 vowels and 39 consonant. Apart from vowels and consonants, there are around 280 compound characters [9] in Bangla script, which are formed by combining two or more consonants. Bangla script, with a few small modifications is also used for writing Assamese, Manipuri, Sylheti scripts. Despite of these facts, a very few works like [10-12] of HWR system of Bangla script are available in the literature compared to other languages like English, French and Arabic etc. The motivation behind this work is to draw a strong attention of the research community to this popular but less explored language.

III. PRESENT WORK

The proposed work describes a holistic approach to recognize handwritten city names using elliptical features and MLP based classifier. Here, first, handwritten word images are collected from various people in A4 size datasheets within a predefined rectangular box. Then these datasheets are scanned and words are cropped manually in order to prepare the database for the experiment in the present work. After that all word images are preprocessed to remove noises and also to smooth the contours of the same. Then from each of these word images a set of 65-element elliptical features are extracted. For the selection of the suitable classifier, first, five well known supervised classifiers like Naïve Bayes, Multilayer Perceptron (MLP), Bagging, Dagging and Support Vector Machine (SVM) are compared and then the final classifier (here, MLP) for the recognition task is selected based on their recognition accuracies on test dataset. Fig 1 shows the block diagram of the entire system.

A. Database Preparation

Database preparation is a vital task in any pattern recognition research. In the present work, handwritten word samples are collected from various people varying in age, educational background and profession. Pre-formatted blank A4 size datasheets were given to the people for writing the words in Bangla script. One such sample datasheet containing 5 different city names (as Jalpaiguri, coochbehar, Raigunge, Balurhat and Krishnanagar) of state West Bengal, India with 28 instances of each city name is depicted in Fig. 2. All datasheets are then scanned in a flatbed scanner with 300 dpi resolution and stored as bmp file format. Finally, from those scanned documents word images are cropped manually to prepare the database for the experiment. The current database contains 20 different word classes (i.e., city names) with 51 samples per class i.e., total 1020 words. At the preprocessing step, all word images are first smoothed using 'Gaussian' filter to remove the noise and then binarized using simple threshold based technique. After the binarization, two morphological operators 'erosion' and 'dilation' are applied on the word images to remove the isolated foreground pixels and to smooth the contours**.**

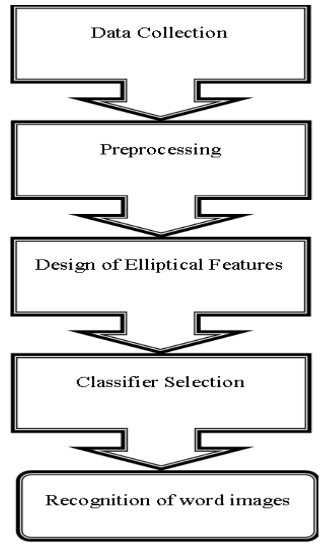

Fig 1: Block diagram of the present work

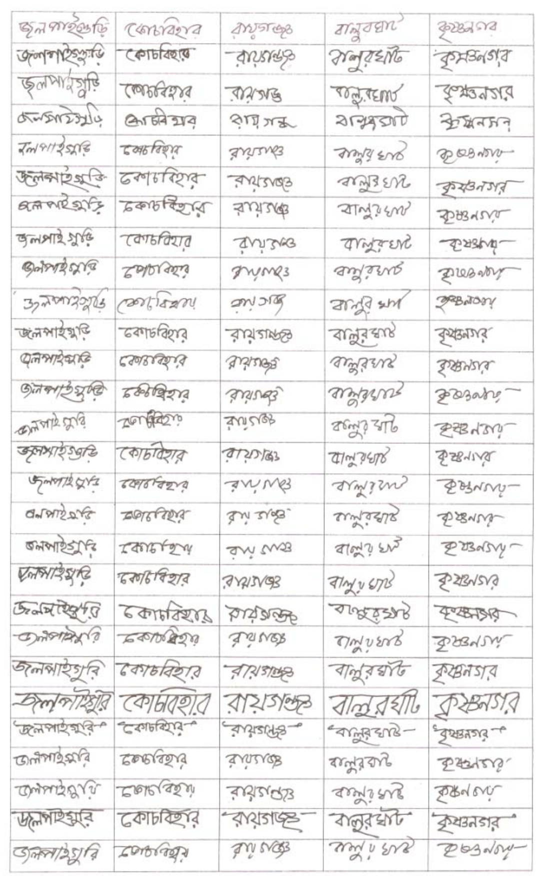

Fig 2: Sample of handwritten Bangla word images

B. Design of Elliptical Features

Feature extraction is another important step in any pattern recognition domain where each pattern is represented numerically. In this work, several elliptical features are estimated for this purpose.

An ellipse is a curved line forming a closed loop, where the sum of the distances from two points (foci) to every other points on the curve is constant. Ellipse can also be defined parametrically as well as non-parametrically.

The parametric definition of ellipse is

$$X = C_x + a\cos(t) \quad (1)$$

$$Y = C_y + b\sin(t) \quad (2)$$

,where (C_x, C_y) is the coordinate of the center of the ellipse.

'a' is the length of radius along X-axis.

'b' is the length of radius along Y-axis.

't' is the parameter such that $0 < t < 2\pi$.

The equation of an ellipse in non-parametric or Cartesian form is,

$$\frac{(X - C_x)^2}{a^2} + \frac{(Y - C_y)^2}{b^2} = 1 \quad (3)$$

In this work, hypothetical ellipses are fit on a word image. To decide the center of these ellipses, the center of gravity for a word image is computed as,

$$C_x = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (4)$$

$$C_y = \frac{1}{N}\sum_{i=1}^{N} y_i \quad (5)$$

,where (x_i, y_i) is the coordinate of the i^{th} foreground pixel.

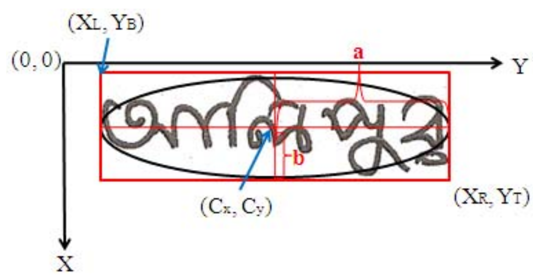

Fig 3: Illustration of outermost hypothetical ellipse of a sample image and the related parameters depiction therein

The parameters 'a' and 'b' are computed as follows

$$a = \min\{(C_y - Y_B), (Y_T - C_y)\} \quad (6)$$

$$b = \min\{(C_x - X_L), (X_R - C_x)\} \quad (7)$$

,where (X_L, Y_B) and (X_R, Y_T) are the coordinates of the top-left corner and bottom-right corner of the minimum boundary box of the considered word image. Here 'a' is considered as the length of the radius parallel to Y-axis and 'b' is the length of the radius parallel to the X-axis (see Fig. 3).

Elliptical features, considering an entire word image are computed following two ways:

- *Concentric ellipses:* These are four feature values designed by considering three concentric ellipses drawn on a word image as shown in Fig. 4. Let us assume that the outermost ellipse is marked as 1^{st}, the next ellipse is marked as 2^{nd} and the innermost ellipse is marked as 3^{rd} ellipses. The radii of the 1^{st} ellipse are computed using Eqs (6, 7). For rest of the ellipses radii are computed as follows ,

  $$a_{i+1} = \frac{a}{2^i} \quad (8)$$

  $$b_{i+1} = \frac{b}{2^i} \quad (9)$$

  ,where $i = 1, 2$.

  Here a_{i+1} and b_{i+1} represent the parameters 'a' and 'b' of $(i+1)^{th}$ ellipse.

  After that, first, *number of foreground pixels* outside the 1^{st} ellipse (but within the minimum boundary box of the word image) is computed and then the same is computed between 1^{st} and 2^{nd}, 2^{nd} and 3^{rd} and inside the 3^{rd} ellipse, which are considered as the feature values.

- *Outermost ellipse:* From the outermost ellipse, drawn on the word image, first 5 features are estimated. These are (i) *number of foreground pixels on boundary of the ellipse,* (ii) *number of foreground pixels along the axis parallel to X-axis,* (iii) *number of foreground pixels along the axis parallel to Y-axis,* (iv) *ratio of foreground pixels and background pixels inside the ellipse* and (v) *ratio of foreground pixels inside and outside of the ellipse* (not exceeding the minimum boundary box of the word). Then the region inside the outermost ellipse is divided into four sub-regions depending on the center and two foci points as shown in Fig. 5. Finally, *number of foreground pixels* inside these 4 sub-regions is computed. Thus, in total, 9 feature values are computed from the outermost ellipse.

Therefore, from each word image, 13 (i.e., 4+9) global feature values are extracted.

To get the local information, the a word image is divided into 4 small subparts depending on the center of the ellipse and from each subpart same feature values, as mentioned earlier, are computed. Therefore, in total 65 (i.e. 5 (4 sub-sectional and 1 whole word)×13) elliptical features are computed from a particular word image. All the feature values are suitably normalized before feeding them into the classifier.

Fig 4: Three concentric ellipses fit on a word image

Fig 5: Illustration of four sub-regions planned by considering the outermost ellipse

Elliptical features are computed from different local regions of a word image thus it will be helpful to discriminate a particular handwritten word from other as different words have dissimilar shapes and pixel distribution in different sub-regions.

C. Suitable classifier selection

After feature extraction, the next most important step is recognition of the handwritten word images and for that a suitable classifier is required. In the present work, 5 well known classifiers such as Naïve Bayes, Bagging, Dagging, MLP and SVM are considered. These classifiers are compared in terms of their accuracies in recognizing the words. For the comparison task a small dataset containing 680 number of word images of city names is created from the current database and a tool called *weka (Waikato Environment for Knowledge Analysis)*, a popular machine learning tool developed at the University of Waikato, New Zealand [13], is used. Table II shows the word recognition performances of these said classifiers.

TABLE I. ELLIPTICAL FEATURES

Feature #	***Feature Name***	***Count***
1	Concentric ellipses	4
2	Number of foreground pixels on outermost ellipse boundary	1
3	Number of foreground pixels along the axis parallel to X-axis in outermost ellipse	1
4	Number of foreground pixels along the axis parallel to Y-axis in outermost ellipse	1
5	Ratio of foreground pixels and background pixels inside the outermost ellipse.	1
6	Ratio of foreground pixels inside the ellipse and outside the outermost ellipse	1
7	Number of foreground pixels in the sub-region inside the outermost ellipse	4
	TOTAL	13

TABLE II. PERFORMANCE COMPARISON OF THE CLASSIFIERS IN RECOGNIZING THE WORD IMAGES

Sl #	***Classifier Name***	***Accuracy (in %)***
1	Naïve Bayes	74.41
2	Bagging	60.00
3	Dagging	69.41
4	SVM	77.35
5	MLP	**77.94**

From the Table II, it is found that, though the accuracies of MLP and SVM are very close but still MLP outperforms SVM. Thus, in the present work, final word recognition task is carried out using MLP classifier.

D. Multilayer Perceptron (MLP)

MLP is a neural network based classifier, consisting of three different layers *viz*., input, hidden and output layers. Each layer contains certain number of neurons or nodes. The number of neurons in the input layer depends upon the number of attributes or features and the number of neurons in the output layer depends upon the number of different classes considered in a certain task. The number of hidden layers and the number of neurons in each hidden layer depend upon the nature or the complexity of the considered problem. Each node in a layer is connected to all other nodes to the next layer, thus it is also known as feed forward network. During the training process, it basically learns the appropriate weight for each connection between nodes of different layers. For learning, MLP uses Back propagation, which is a supervised learning technique.

IV. EXPERIMENTAL RESULTS

As stated earlier, for the present work a database of 1020 handwritten words are prepared with 20 different word classes and in each class contains 51 samples.

For the recognition task a 3-fold cross validation method is used. Each fold contains 680 word images for training and 340 for testing For each fold, an MLP classifier is trained with learning rate (η) = 0.3, momentum term (α) = 0.2 with different number of neurons in the hidden layer. The detail result of the present word recognition system is shown in Table III.

A. Error case analysis

In the proposed work, in most of the cases word images are correctly classified still there are some word samples which are not classified properly. As, in the present work, only elliptical features are computed from the word images, it is found that words belonging to different classes have the almost same pixel distribution. Therefore, they are become indistinguishable in the feature space. Another problem is skewness in the word images (see Fig. 6). A word with large skew angle has different pixel distribution at various local regions, compare to other words of the same class and thus may be misclassified.

B. Conclusion

Handwritten word recognition is one of the most important and hardest challenges that have been fascinating the researchers from long time ago due to its large area of applicability. Innumerable research works are found in the domain of HWR considering Arabic, English etc. languages. But unfortunately there is no such amount of work found in case of Bangla script. In the proposed work, a holistic Bangla word recognition system is developed.

In this work, elliptical features are computed from the entire word images as well as from the local regions of a word image. Then a suitable classifier is chosen from 5 well known classifiers for the classification task. The final classification is carried out using MLP. Considering the complex nature and the richness of the Bangla script, it can be said that the performance of the proposed system is quite satisfactory. In future, this system will be tested with more word classes and having more samples. During experimentation, it is found that due to skewness some of the word images are misclassified. Therefore, a skew correction module needs to be added at the preprocessing step.

TABLE III. DETAIL RESULT OF THE PRESENT BANGLA HANDWRITTEN WORD RECOGNITION SYSTEM

Fold#	*Number of Training Samples*	*Number of Test Samples*	*Accuracy*
1	680	340	77.94%
2			80.29%
3			**85.88%**

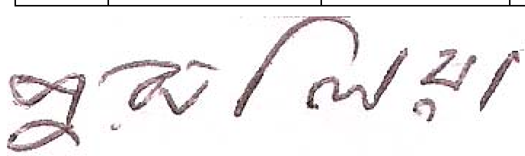

Fig 6: A sample word image with a large skew angle is misclassified by the present technique

ACKNOWLEDGMENT

The authors are thankful to the Center for Microprocessor Application for Training Education and Research (CMATER) and Project on Storage Retrieval and Understanding of Video for Multimedia (SRUVM) of Computer Science and Engineering Department, Jadavpur University, for providing infrastructure facilities during progress of the work. The current work reported here, has been partially funded by University with Potential for Excellence (UPE), Phase-II, UGC, Government of India.

REFERENCES

[1] U. Pal, B. Chaudhuri, "Indian script character recognition: a survey", *Pattern Recognition,* 37(9), 1887–1899 (2004).

[2] S. Madhvanath and V. Govindaraju, "The role of holistic paradigms in handwritten word recognition", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(2), 149-164 (2001).

[3] M. Dehghan, K. Faez, M. Ahmadi and M. Shridhar, "Handwritten Farsi (Arabic) word recognition: a holistic approach using discrete HMM",Pattern Recognition, 34, pp. 1057-1065, (2001).

[4] A. Namane, A. Guessoum and P. Meyrueis, "New holistic handwritten word recognition and its application to French legal amount", *Springer-Verlag Berlin Heidelberg*, pp. 654 – 663(2005).

[5] K. Roy, S. Vajda, U. Pal, B. B. Chaudhuri and A. Belaid, "A System for Indian Postal Automation", *International Conference on Document Analysis and Recognition,* pp. 1060-1064 (2005).

[6] J. J. de Oliveira Jr, COA Freitas, JM de Carvalho and R. Sabourin, "Handwritten Word Recognition Using Multi-view Analysis", *Springer-Verlag Berlin Heidelberg*, pp. 371–378 (2009).

[7] A. Acharyya, S. Rakshit, R. Sarkar, S. Basu and M. Nasipuri, "Handwritten Word Recognition Using MLP based Classifier: A Holistic Approach", *International Journal of Computer Science Issues*, Vol. 10, Issue 2, No 2, pp. 422-427, (2013).

[8] R. Sarkar, N. Das, S. Basu, M. Kundu, M. Nasipuri and D. K. Basu, "CMATERdb1: a database of unconstrained handwritten Bangla and Bangla–English mixed script document image", *International Journal on Document Analysis and Recognition*, pp. 71–83, (2012).

[9] N. Das, K. Acharya, R. Sarkar, S. Basu, M. Kundu, & M. Nasipuri, "A benchmark image database of isolated Bangla handwritten compound characters", *International Journal on Document Analysis and Recognition (IJDAR)*, 1-19 (2014).

[10] N. Das, S. Pramanik, S. Basu, P. K. Saha, R. Sarkar, M. Kundu, & M. Nasipuri, "Recognition of handwritten Bangla basic characters and digits using convex hull based feature set", *International Conference on Artificial Intelligence and Pattern Recognition* (2009).

[11] A. Roy,T. K. Bhowmik, S. K. Parui and U. Roy, "A Novel Approach to Skew Detection and Character Segmentation for Handwritten Bangla Words", *Digital Imaging Computing: Techniques and Applications,* pp. 30-30 (2005).

[12] U. Bhattacharya, A. Nigam, Y. S. Rawat, and S. K. Parui, "An analytic scheme for online handwritten Bangla cursive word recognition." *Proc. of the 11th ICFHR* ,320-325 (2008).

[13] Weka 3: Data Mining Software in Java, Machine Learning Group at the University of Waikato, http://www.cs.waikato.ac.nz/ml/weka/