(IJIEST) 2019, Vol. No. 5, Jan-Dec **e-ISSN: 2454-9584; p-ISSN: 2454-8111**

ENGLISH CHARACTER RECOGNITION SYSTEM USING HYBRID CLASSIFIER BASED ON MLP AND SVM

***Bashar Ibrahim Hameed, **Humam Khalid Yaseen, ***Raed Sami Sarhan**

**Assistant Lecturer, Master of Information Systems, Iraqi Sunni Affairs*

*** Assistant Lecturer, Master of Computer Science, Iraqi Sunni Affairs*

****Assistant Lecturer, Master of Information Systems, Ministry of Education, Iraqi-Baghdad Education Directorate Karkh-3*

ABSTRACT

The initial stage of recognizing individual characters is through the use of machine in analyzing handwritten documents. This research focuses on the recognition of isolated characters obtained from handwritten English documents. The focus is given to the study of preprocessing and feature extraction methods. To bring about size uniformity, the character image samples are normalized to 70X50 pixel sizes using nearest neighbor interpolation method. A normalized image is next transformed into a binary image utilizing Otsu's threshold selection technique. This is followed by thinning or contour tracing. An important concern in character recognition is the selection of best discriminative features. This study took an overview of various statistical and structural features and recommended a unique and new view on the basis of features. The effectiveness of these features on handwritten English character recognition is experimented using MultiLayer Perceptron (MLP) and Support Vector Machine (SVM) classifiers. The proposed modified view based feature is found to be very efficient. An extensive assessment experiments outcomes carried out with different features/combination of features and classifiers is presented. From the experiments, view based features are found to give high recognition accuracy. By combining different features in an optimal way, recognition accuracy of 97% is obtained.

Keywords—nearest neighbor interpolation; English character recognition; view based feature; support vector machine; multilayer Perceptron

INTRODUCTION

Off-line handwriting analysis as well as recognition served as an attractive area to the researchers since last many decades [1-2]. The world had witnessed the enhanced advancement of handwritting recognition methods.

Documents that are hand printed has a significant function in our lives despite the prevalence in the utilization of computers in document processing. In the current times that do not use papers, analyzing hand printed documents by using machine plays a major role. The machine analysis of documents prepared with paper and pen is termed as offline document recognition systems that are written by hand (HWR).

In comparison to the aforementioned offline document recognition system, on the other hand, the online system that recognizes characters that are handwritten (HCR) focuses on recognizing characters that are written on a sensor device like touch pad which can be directly be converted into digital form.

The identification of characters that are written by hand via the use of machines is an emerging pertinent technology in the modern world. Recognizing documents that are written by hand is the enablement of interpretation by a computer program on the received entry that is written by hand, where it can perform correct segmentation, extract optimal features and recognize the possible words. The evolution of automatic character recognition takes place in two modes (based on the data acquisition): offline recognition and online recognition. The online HCR system uses the digitizer that directly captures the written characters according to the strokes, speed, pen up and pen down information. However, Offline character recognition (OCR), also known as Optical Character Recognition, retains the image of the data on the paper via the use of optical scanner or cameras. It is also termed as optical character recognition (OCR) due to the conversion of the text image into a bit pattern via digitizing devices in an optical manner.

The major challenges in HCR amounts from the huge differences in handwriting forms, as they vary according to distinctive styles at the individual level, across varying instance of period, for example, varies according to the expanse, form, speed of writing, width of the characters and others[2]. The current state of the art printed character recognition systems perform quite well and provides recognition accuracy as good as 99.9% but HCR systems are restricted in their capacities because of issue of writing similarity between two characters and infinite shape possibility by different users.

The main research goal is in recognizing the HCR of English characters with acute precision. All pattern recognition system functionings are dependant upon the functioning of its constituents at the individual level. Thus the research is designed in focusing on the preprocessing feature extraction and categorization phases in the form of either suggesting the new one or modifying the available approaches in pattern recognition.

Unlike online HCR systems, which utilize factors like strokes sequence, speed, pen up and pen down information, the systems that are not online use scanned images of handwritten documents. Hence offline HCR system design is more challenging. An offline system has to address issues such as the distinctive styles at the individual level, across varying instances according to size, shape, thickness of characters and others. These issues limit the accuracy of offline HCR systems. Due to the challenges in design faced by offline HCR system, it is an emerging area of concern in research pertaining to image processing and pattern recognition.

A well-structured and standard database of isolated handwritten character images is developed in this research work. The constructed database has the potential to be used as a benchmarking resource for the handwritten character recognition. A detailed investigation of performance of a set of selected features used to recognize segregated English characters that were set apart for is carried out. A set of view based feature is proposed. The proposed features are found to be effective for English characters and are simple and require less feature extraction time. An efficient combination of classifier MLP and SVM is used in proposed study, that generates an average recognition precision of 97%, which is significant achievement in English HCR studies.

The remainder of the study are as the following. Section two presents HCR, different steps in HCR and the broad area of research. A review of selected literature related to HCR system in English scripts is also presented. Section three describes the different preprocessing steps used in handwritten character recognition systems, thinning algorithm, tracing algorithm, feature extraction method and classifiers. Section four contains the description of the hybrid classification experiment and their results. Finally, the work is concluded in Section five along with the direction for future research.

PRELIMINARIES

A. Handwritten Character Recognition

Handwritten Character Recognition (HCR) system is developed with an objective to recognize handwritten characters from a digital image of handwritten documents. An HCR system includes steps such as character segmentation, pre-processing the image representation of the character, feature extraction, in addition to the recognition of character class with the extracted features.

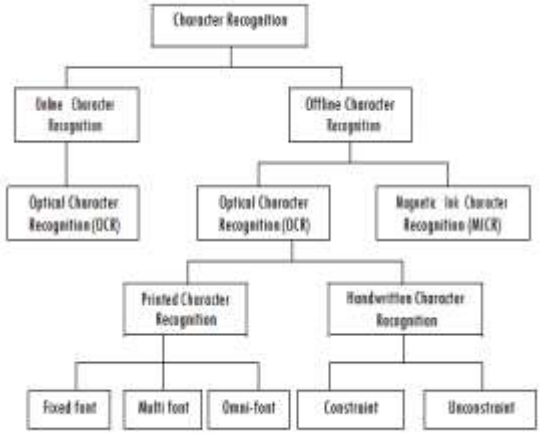

Fig. 1. Classification of handwritten character recognition

Character Recognition methods links the identity that is in the form of a symbol with the image of a certain character. It is categorized into two categories, according to the data acquisition and kind of

text : online and offline (Figure 1).

The online system that involves recognizing characters uses the digitizer that obtains in a direct way the writing according to the sequence of the strokes, speed, pen up and pen down data collected [3]. On the other hand, the offline character recognition (OCR) obtains the information from paper using the optical scanner or cameras. The perceived general acceptance is that the on-line handwritten recognition of text technique has attained more superior outcomes in comparison to its offline opponent. The Offline Character Recognition may be segregated further into optical character recognition (OCR) and magnetic Ink character recognition (MICR).

MICR used the property of magnetic field to recognize the written text. Different character provides unique magnetic field, which help in its recognition. On the contrary, OCR recognizes the character with the help of its optical feature: appearance and shape. OCR uses the same approach as the human used for character recognition. This approach helps in recognizing the character of different shape, size and orientation, which could either be hand written or printed. OCR is segregated into printed character recognition and HCR.

B. Different Steps in HCR Systems

Generally, a system that is used for recognizing comprises of the following steps: image acquisition, thinning, contour, segmentation, feature extraction and finally a classification stage to generate a digital output.

Different steps the system used for recognizing are illustrated in Figure 2 block diagram. The images for HCR system can be obtained through writing straight into the computer using stylus capturing or by taking a document photo, in addition to the alternative of scanning handwritten document. It is termed as digitization process [4]. Preprocessing entails a sequence of functions executed to improve it, enabling it to be appropriated for segmentation. Preprocessing step entails removing noise, thinning and contour and others. Appropriate filter such as gaussian filter, mean filter and min-max filter can be utilized in noise removal from the target image. Binarization process transforms the gray scale or colored image to image in black-and-white format.

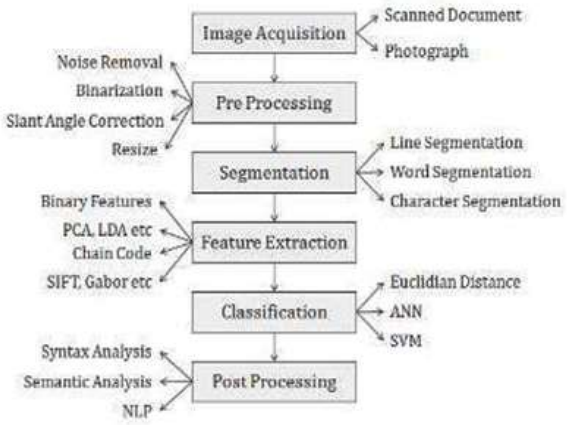

Fig. 2. Different steps of character recognition system

Furthermore, the size of the input document might need to be adjusted in the event that the size is extremely huge for the improvement of operation immediacy. If document is scanned then it might be tilt and needs to be aligned by performing slant angle correction. However tilt correction and size reduction might also lead to removal of some important features for the recognition so both should be done very carefully. Typically, there is an hierarchy in the processing of the document. At the first stage, the lines are segregated into segments utilizing row histogram. Out of every row, the extraction of the words is executed utilizing column histogram, and lastly the extraction of the characters is implemented. The preciseness of the accurate categorization of the input document is hugely dependent on the preciseness of suitable segmentation.

Feature extraction is the essential core of all character recognition system [5]. Feature extraction methods such as Linear Discriminant Analysis (LDA), Principle Component Analysis (PCA), Chain Code, Independent Component Analysis (ICA), zoning, gradient based features, Scale Invariant Feature Extraction (SIFT) and histogram are executed in the extraction of the characteristics of individual characters. The characteristics are utilized to train the classification system. At the instance of the introduction of a new input image to the HCR system, the extraction of the characteristics is implemented, and provided as an input to the trained classifier such as artificial neural network or support vector machine. The classifiers will make a comparison between the input feature with the stored pattern, and ascertain the optimum matching class for the input. The Semantic or Syntax analysis or likewise greater level notions could be implemented to assess the appropriateness of the recognized character in the post processing, though it is not a mandatory step but helps in the improvement of accuracy

LITERATURE REVIEW

Various classifiers comparative performance study that are implemented on handwritten digit

recognition are discussed in Liu et al. [6]. The characteristics producing the latest preciseness in handwritten character recognition tasks involves gradient [13], curvature [7] characteristics. Gabor transform [8] and statistical/structural features have been utilized in a successful way in several recognition studies[9] .

Recently[10] in a study on recognizing handwritten Devanagari and Bangla character, the wavelet transforms of the image of the character image inputted were put through the three layer approach.

Rajib et al. [11] recommended a handwritten English character recognition system according to the Hidden Markov Model (HMM). The said technique utilize a pair of varying feature extractions entailing global and local feature extraction. The global feature consists of numerous features such as gradient features, projection features and curvature features in the sequence of four, six and four accordingly. The estimation of the local features is realized through the division of the sample image into nine equal blocks. The calculation of the gradient feature of each block utilizing four feature vector, that produces thirty-six total number of local features . This produces fifty features (local + global) for individual sample image. After that, the features are inputted into HMM model for the purpose of training it.

The post processing information is used by this technique for the purpose of decreasing the cross classification of varying categorizations. This techniques is time consuming in the training and feature extraction. Additionally, its performance is inferior in such inputs instances, when numerous characters are mixed in an individual image.

Velappa Ganapathy et al. [12] recommended technique used for recognizing according to the multi-scale neural network training. For the improvement of the precision, this technique utilized selective threshold, that used calculation according to minimum distance method.. It entails the creation and advancement of GUI, that is able to ascertain the character within the whole the scanned image. It offers an 85% precision with medium training level. It utilized huge resolution images (20 × 28 pixels) for training with a decrease in training duration.

Som et al. [13] utilized fuzzy membership function in improving the precision of hand written text recognition system. Here, the text images are normalized to 20 × 10 pixels, and fuzzy approach is next utilized for every individual class. The bonding box is constructed surrounding the character for the purpose of determining the text's vertical and horizontal projection.

When the image undergoes cropping to a bounding box, the image is adjusted in its scale to the size of 10×10 pixels. Next, the cropped images undergoes thinning process via the thinning operation. For the creation of the test matrix, the entire pre-processed images are put into a single matrix; in

sequential order. During the presentation of new (test) images is made by the user, it is assessed for the test matrix matching. The technique was rapid but it offers a low precision.

Rakesh kumar et al [14] recommended a technique for the reduction of the system training by using a one-layer neural network, however, this produced a reduction in the precision rate. The scaling of the segmented characters is to 80 X 80 pixels. The performance of the data normalization onto the input matrices is for the improvement of the training performance.

EXPERIMENTAL DETAILS

A sequence of operations are executed onto the input image at the time of the pre-processing stage, that comprises de-noising, binariation thinning and others. It assists in the improvement of the image rendering in addition to enabling the image to be appropriated for segmentation and feature extraction.

C. Modified Stentiford Thinning (MST)

The current research involves modifying through the utilization of the stentiford thinning algorithm , which contains three steps: pre-processing, thinning and post processing. Firstly, the hole is removed through the utilization of six patterns of bits, which is then followed the implementation of smoothing through the removal of the entire black elements which possess less than three black neighbors. Following the pre-processing, the Zhang Suen thinning algorithm is executed as according to the description in Ref [15]. Just four matrices M are utilized in stentiford thinning algorithm for the deletion of a central black pixel. However, for the recommended MST algorithm, we utilized two new matrices Bi and Ci additional to M_i where i= 1 2 3 and 4. Due to the fact that M_i scans matrices and only executes the removal of additional lines on the left, right upward and downward. For the purpose of additional slanting lines removal, the MST algorithm utilized B_i and C_i scan matrices as illustrated in Figure 3.

The scanning of the entire matrices in totality is carried out over the character and instances when a matrix fits, the central black pixel is pinpointed to be deleted. The summary of post processing of the recommended MST algorithm is as the following:

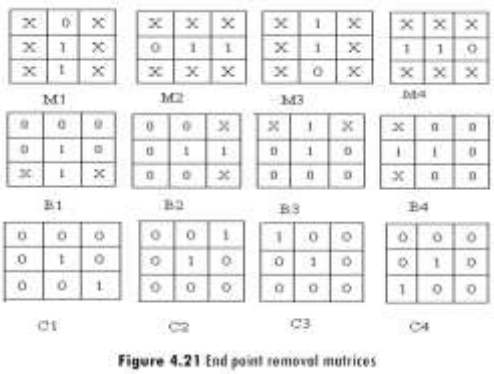

Figure 4.21 End point removal matrices

Fig. 3. End point removal matrices

Step1. Firstly, the matrix M_1 is scanned across the character in accordance to to Figure 4, and the identification of the next fit position is carried out.
Step2. In the event that the central element at a fit does not fall as an endpoint, and it possesses a connectivity of value one, it is then identified for deletion.
Step3. Steps 1 and 2 for the entire fit positions are repeated
Step4. Steps 1-3 for individually scanned of matrix M_2, M3, M4 are repeated.
Step5. The entire marked elements are deleted.
Step6. In the event that in Strep 4, one or more elements are deleted, then revert to Step1.
Step7. Steps 1-6 for the scanned matrices B_i and C are repeated.
Step8. Exit

Order of Application	Direction of Single scan line	Direction of successive scan line
M1,B1,C1	Left to Right	Downwards
M2,B2,C2	Upwards	Left to Right
M3,B3,C3	Right to Left	Upwards
M4,B4,C4	Downwards	Right to Left

Fig. 4. Matrix ordering and scanning dierection of modified thining algorithm

D. Contour

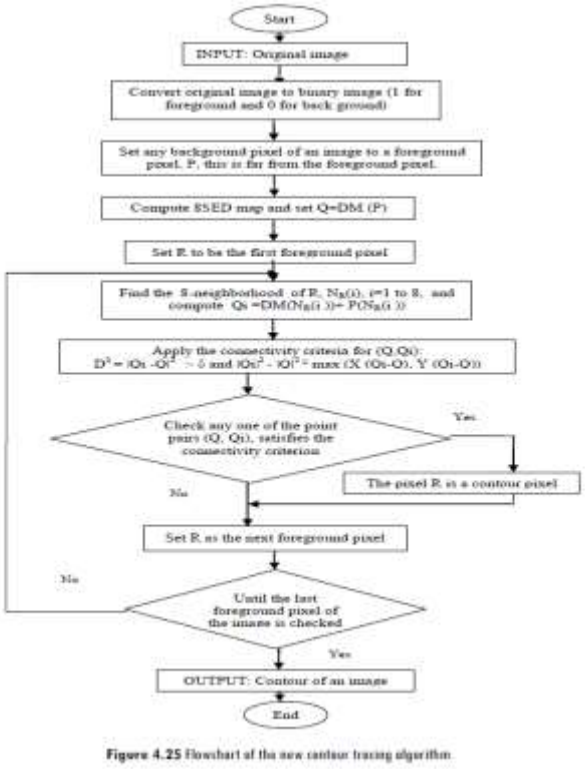

Figure 4.25 Flowchart of the new contour tracing algorithm

Fig. 5. Flow chart of countor traching algorithm

The key component of this algorithm is to apply eight-point sequential Euclidean distance 8SED mapping and connectivity criteria to achieve the correct contour point of an image. The proposed algorithm is illustrated in Figure 5. This algorithm assumes the value of background points as 0 and is represented by black color, the value of object point as 1 and is represented by white color. To find a contour of an image one needs to traverse the entire image. First we select a background pixel at random and set its value to 1(set it as foreground pixel, say P). This pixel can be far from the foreground pixel. Next we compute the 8SED mapping and set Q=DM (P). For the determination in ascertaining that a pixel point is a contour point, the corresponding points for every eight neighboring points are identified, and the connectivity criterion is next implemented onto the set of eight point pairs. The relative proximations of the closest points for individual pixels may be attained through the utilization of 8SED map. The eight neighboring points of the point taken into account may be attained through the addition of relative position of its closest point to the relative location of its neighborhoods.

E. Segmentation

The segmentation is executed based on the edge detection and the gap that exists between the varying

characters. Following the segmentation, the portions that are sub-divided are marked and then undergo further processing in succession, individually. The labeling is implemented for the determination of the number of characters in the whole image. The image that is scanned, is determined for valid segmentation points through the assistance of pixels value change (zero and one) or min-max detection among the characters, that is extremely simple to ascertain in texts that are not connected. In addition, the segmentation points are assessed for any error point inclusion through the inspection of every points against the average distance between two segmentation points in a whole image.

a. Line Segmentation Algorithm:

1. Choose a pre-processed binary image.

2. Calculate the size of image. Suppose it is (m, n)

3. Calculate the frequency of black and while pixels in each row using horizontal-vertical scanning and save the values in 2D array.

4. The row having 'n' white pixels are designated as blank row.

5. The row having mix of white and black pixels are designated as character row.

6. The distance between last character row and first blank row is calculated for each occurrence.

7. The average of all these distances are calculated and designated as line space.

8. The midpoint row of line space is used to provide a segmentation row between each of the last character row and first blank row.

9. The set of rows are segmented into different parts are designated as lines.

b. Character Segmentation Algorithm:

1. Choose a Line segmented image

2. Calculate the number of lines and assigned it value L.

3. Repeat for line = 1 to L

4. Calculate the frequency of black and while pixels in each column using vertical- horizontal scanning and save the values in 2D array.

5. The column having all white pixels are designated as blank column.

6. The column having mix of white and black pixels are designated as character column.

7. The distance between last character column and first blank column is calculated for each occurrence.

8. The average of all these distances are calculated and designated as character space.

9. The midpoint row of character space is used to provide a segmentation column between each of the last character column and first blank column.

10. The set of columns are segmented into different parts are designated as individual character.

11. The individual character is further processed using edge detection in order to remove any extra blank columns or rows.

12. Save the individual character for further processing.
13. End

F. Feature Extraction

In proposed feature extraction method, we used thinned binary images. There are two methods utilized for feature extraction, namely: features that are based on viewing capabilities, and bit map feature. The feature extracted from both the approaches are then combined and feed to classification stage.

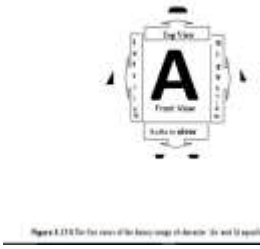

Fig. 6. Five views based features of English character "A"

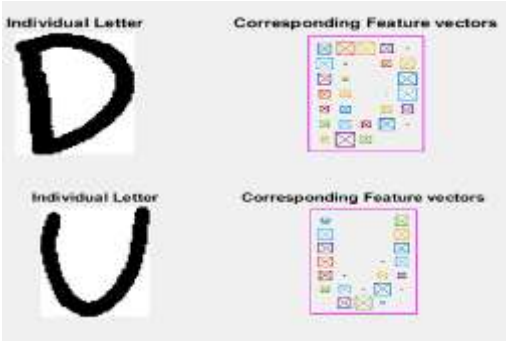

Fig. 7. Bit map features of different character

The view based approach is based on the fact that a human usually needs only partial information for correct character recognition, which is its shape. The first approach is to find the five views viz, top, bottom, left, right and front of original and thinned images and then each view is divided into 16 equal blocks. Thus, in each block, p, the Sum of Background Pixels (SBP) is calculated as SBP_p. In the second approach, the original and thinned images undergo a division that set them into 16 equal blocks, and from every individual block, five views are computed. Then for each view, the sum of back ground pixels is calculated as SBPv. The five view based features are shown in the Figure 6.

On the other hand, the calculation of the bit map feature is realized through the conversion of the pre-processed image and transforming them into bit mapped version of size 7×5. Several examples

of bit map format of varying characters are depicted in Figure 7. The preservation of the core features of input image with rather smaller amount of space/ data length are realized by the bit map form.

G. Classifier

The classifiers that are based on MLP and SVM are singly trained on the exact handwritten English database utilizing equal features set. The MLP classifier rate of recognition is 94.15%, and the recognition rate is 95.87% for the SVM classifier derived from the exact test data itself. Given that the two classifiers produce a precision greater than 94%, we merged the two through the utilization of varying combination techniques as explicated in Section Four. The outcome of the combined classifiers utilizing the sum rule, product rule and Borda count are illustrated in Table Two. The application of the weighted sum rule is implemented for the combination of the classifiers. The allocation of varying weights β and 1- β (where β varies from 0 to 1) are applied to both of the classifiers, and the rates of recognition are determined. The varying of the accuracy influenced by the weights that are given by both of the classifiers are shown in Figure 8.

RESULTS AND DISCUSSIONS

The recognition accuracies of the recommended system that possess the varying feature sets and varying classifiers are shown in Table I . It can be clearly seen from the result that the hybrid features and hybrid classifier combination provides the most accurate result. Table II is showing the variation of accuracy of hybrid classifier with different hybridization rule.

TABLE I. RECOGNITION ACCURACY WITH DIFFERENT FEATURES AND CLASSIFIER

Method	Recognition Accuracy		
	MLP Classifier	SVM Classifier	Hybrid MLP and SVM
View Based Features	89.25	90.12	92.21
Bit Map Features	93.25	94.84	95.34
Hybrid Features	94.15	95.87	97.04

TABLE II. RECOGNITION ACCURACY OF HYBRID CLASSIFER USING DIFFERENT RULES

Method	Recognition Accuracy (Hybrid MLP and SVM)
Sum Rule	95.87

Product Rule	95.98
Borda Rule	96.35
Threshold Rule (T=0.5, N=6)	97.04

The classifiers that are based on MLP and SVM are combined by the method shown in the Figure 8. Suppose that there are K classes denoted by $C_1, C_1,...C_k$. Hence, for a set pattern x, every single classifier (MLP and SVM) will generate K potential probabilities ($p_1^j, \quad p_2^j,....p_k^j$) that it x belongs to that particular class. Here j=1 represents SVM and j=0 represents MLP. In the initial phase, we chose top N ranked classes from the MLP confidence scores. We selected the value of N dependent upon the classifier's recognition accuracy. Accordingly let { $t_1, t_2,...t_n$} signify the top N ranked classes by the MLP classifier.

In stage two, the confidence estimates (corresponding to the N choices) from SVM P^1 are analyzed. The one that possesses the highest rank or score is selected. Let the t selected to signify this chosen class and s_{chosen}^1 to signify estimated probability for this class by the SVM classifier. The confidence measure s_{chosen}^1 might be low in the event that the SVM classifier is unable to accurately recognize the pattern as the result of certain confusion. Thus, a back up alternative to the MLP classifier is presented, which is dependent upon the confidence measure of the SVM classifier. In the event that s_{chosen}^1 is less than a particular threshold T, the top ranked class from the HMM classifier is forecast as the final output. Alternatively, t_{chosen} from SVM is the final output of the combined classifier.

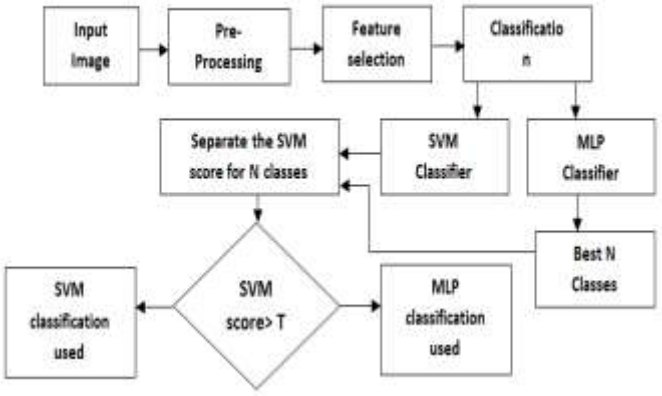

Fig. 8. Block diagram for the threshold hybridization of classifier

The classification accuracy is greatly depends on the two parameter N and T. The modification of character recognition accuracy (with the adjustment in the values of these two parameters) is shown in the Table III. Figure 9 is showing the sample output of the developed system.

TABLE III. VARIATION OF CHARACTER RECOGNITION WITH DIFFERENT THRESHOLD AND BEST CLASSES

Best Class (N)	Threshold (T)		
	T = 0.4	T = 0.5	T = 0.6
3	95.87	96.15	95.19
6	96.09	97.04	96.78
9	95.96	97.04	96.12

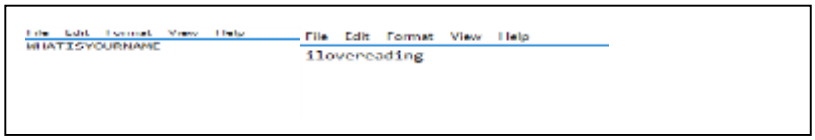

Fig. 9. Sample output of the system

CONCLUSIONS

This study was mainly focused on the efficient steps development for preprocessing, feature extraction and classification of HCR system. Extensive study was carried out by selecting a set of prominent preprocessing and feature extraction methods. Feature extraction algorithms used skeleton of character images (thinned images). Hence, efficient character thinning algorithms are developed after a detailed study of popular thinning algorithms. In feature extraction strategies, character contour is employed. A novel contour tracing algorithm and view point base feature extraction technique is utilized in this study. The proposed algorithm is implemented and applied on English handwritten characters. Extensive experiments are carried out with the different feature extraction methods in isolation and in combination. Classification is realized with SVM and MLP. Results show that the proposed view based features outperform the other approaches. Further, combination of features found to improve the accuracy, and a recognition rate of 97% could be achieved. In future, feature reduction methods will be incorporated with features combination approaches as well as better approaches in feature extraction and classification will be tried to develop for better performance of HCR.

REFERENCES

[1] N. Arica and F. Yarman-Vural, An Overview of Character Recognition Focused on Off-Line Handwriting, IEEE Trans. Systems, Man, and Cybernetics Part C: Applications and Rev., vol. 31, pp. 216-233, 2001.

[2] R. Plamondon and S. N. Srihari, On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey, IEEE Trans. Pattern Analysis and Machine Intelligence, vo1.22, pp.63-84, 2000.

[3] Connell, Scott D., and Anil K. Jain. "Template-based online character recognition." Pattern Recognition 34.1 (2001): 1-14.

[4] Von Ahn, Luis, et al. "recaptcha: Human-based character recognition via web security measures." Science 321.5895 (2008): 1465-1468.

[5] Trier, Øivind Due, Anil K. Jain, and Torfinn Taxt. "Feature extraction methods for character recognition-a survey." Pattern recognition 29.4 (1996): 641-662.

[6] H. Liu and X. Ding, Handwritten Character Recognition using Gradient Feature and Quadratic Classifier with Multiple Discrimination Schemes, Proc. Eighth Inti Conf. on Document Analysis and Recognition, pp. 19-25, 2005

[7] M. Shi, Y. Fujisawa, T. Wakabayashi, and F. Kimura, Handwritten Numeral Recognition Using Gradient and Curvature of Gray Scale Image, Pattern Recognition, vol. 35(10), pp. 2051-2059, 2002.

[8] Singh, Sukhpreet, Ashutosh Aggarwal, and Renu Dhir. "Use of Gabor Filters for recognition of Handwritten Gurmukhi character." International Journal of Advanced Research in Computer Science and Software Engineering 2, no. 5 (2012).

[9] Wu X-Q, Wang K-Q, Zhang D (2005) Wavelet energy feature extraction and matching for palmprint recognition. J Comput Sci Technol 20(3):411–418

[10] U. Bhattacharya and B. B. Chaudhuri, Handwritten Numeral Databases of Indian Scripts and Multistage Recognition of Mixed Numerals, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 31(3), pp. 444- 457, 2009.

[11] Rajib Lochan Das, Binod Kumar Prasad, Goutam Sanyal, "HMM based Offline Handwritten Writer Independent English Character Recognition using Global and Local Feature Extraction", International Journal of Computer Applications (0975 – 8887), Volume 46– No.10, pp. 45-50, May 2012.

[12] Velappa Ganapathy, Kok Leong Liew, "Handwritten Character Recognition Using Multiscale Neural Network Training Technique", World Academy of Science, Engineering and Technology, pp. 32-37, 2008.

[13] T.Som, Sumit Saha, "Handwritten Character Recognition Using Fuzzy Membership Function", International Journal of Emerging Technologies in Sciences and Engineering, Vol.5, No.2, pp. 11-15, Dec 2011.

[14] Parvez, Mohammad Tanvir, and Sabri A. Mahmoud. "Arabic handwriting recognition using structural and syntactic pattern attributes." Pattern Recognition 46, no. 1 (2013): 141-154.

[15] Zhang, T. Y., and Ching Y. Suen. "A fast parallel algorithm for thinning digital patterns." Communications of the ACM 27.3 (1984): 236-239.