Study and evaluation of Segmentation and feature Extraction techniques for Digit recognition

Dinesh Kumar Dewangan, Om Prakash Dewangan
Computer Science and Engineering
Rungta College of Engineering & Technology , Bhilai, India

***Abstract*— Image processing is wider area of research due to use of image processing in real time applications and devices which requires more accuracy and efficiency in recognition. Fundamental task of Image processing are Image segmentation, Feature extraction and classification of extracted feature. For Computer vision system, the basic purpose of image processing is to produce images which make the subsequent analysis more reliable and simpler. We are not at all the interested how well the image looks while we focuses in interesting and hidden information over there. The text or digit recognition by capturing images or scanned images is very crucial task because various preprocessing required before the segmentation and feature extraction. Purpose of this paper is to compares single or combination of methods for Image segmentation and similarly compares multiple Feature extraction methods. The main aim of this paper is to find out best method or combination or hybrid methods for digit recognition in real time basis. And proposed segmentation and Feature extraction method is to segment the object less distortion and extract the feature or information in maximum accuracy level. The purpose of developing this method is to recognize text or digit of captured photographs automatically in real time. Also the algorithm must work for worst cases efficiently.**

***Index Terms*— Image processing, Digit recognition, Hybrid segmentation, Multiple Feature extraction, Classification, real time recognition.**

I. INTRODUCTION

Extracting meaningful information from any kind of captured image is too challenging task, character detection, extraction and recognition have been an active field of research. It still remains an open problem in the field of digit recognition and image processing at the real time basis. There are mainly three phases of a character/text recognition system namely Preprocessing, Segmentation and Feature extraction. Preprocessing aims at eliminating the inconsistency that is inherent in captured images. The different preprocessing techniques applied by various researchers in regard to increase the performance of the Segmentation and Feature Extraction process. The objective of segmentation is to simplify the representation of an image into something which is more meaningful and easier to recognize and analyze. The result of segmentation is to properly separate the region of interest where the information is located and after that those segmented portions are used for information extraction like Text recognition, digit recognition [4] and classification.

Image processing is an important tool to analyze any raw image and fetch precious information of them. Image processing generally refers to digital image processing for recognition. Image processing method is closely related to computer graphics and computer vision. For computer graphics, images required manual intervention which is made through physical models of objects, environments, and lighting, instead of being acquired (via image capturing devices such as cameras) from natural scenes, as in most animated movies. In computer vision it is often considered high-level image processing out of which a machine / computer / software means to interpret the physical contents of an image or a sequence of images (e.g., Scan photographs, videos or 3D full-body magnetic resonance scans). The Segmentation and feature extraction of the image is very useful in medical applications to diagnose the abnormalities in the image [3], Satellite imaging and in computer vision as well as in Artificial Neural Network, for criminal record management based on various features of images also an another important use of image processing. The criteria for segmentation and feature extraction of image is very hard to decide as it varies from image to image and also varies significantly on the resource used to capture the image. There are various literatures available to understand and analyze the segmentation and feature extraction techniques which are discussed below:

Brhmadesam S. (2010) has proposed hybrid segmentation methods such as Gaussian Filtering, Sobel Edge detection and another fine tuning methods Dilation and Erosion for increasing the efficiency of segmentation of image. Use of mathematical morphological operators, edge detection and filtering techniques in a novel way without complex mathematical jargons it provide better results. After applying proposed hybrid method in 350 images and get better results of segmentation.

R.Radha et al (2014) has proposed Hybrid Classifier method for efficient classification or feature extraction method as per image the property of image element based on the feature of the input image, classifier is selected automatically:

- Hole Feature
- Fourier Descriptor Classifier
- Chain Code Feature

The proposed system was tested with 700 samples containing 70 set of each digit (0-9) of varying size and font style. The samples contain various fonts like Arial, Times New Roman, Courier New, etc. The digit images were Captured using canon digital camera. It was stored in Portable Network Graphics (PNG) format. Image segmentation was performed on the processed Input image. Feature extraction was performed to the segmented digits and recognized using hybrid classifiers. Accuracy in feature extraction of Hole classifier is 99%, Fourier descriptor classifier is 100 % and MLP-Back propagation classifier shows100%.

Quin J. et al (2006) has proposed method for digit recognition modules including image pre-processing, image binarization, morphological filtering, segmentation, feature extraction and digit recognition using a graphic language. The approach was

examined and tested on 5000 banknotes with 4000 digits and the approach achieved s single digit recognition rate of more than 99.60%, a serial number recognition rate of 99.50%, and a recognition time of 157ms. The results show that our approach is effective and efficient and can clearly satisfy the system requirements. Our experiments also suggest that the current system is still quite sensitive to the noise of the image, but it could be improved by increasing the number of template digits in the training set.
Andrew J. (2011) has proposed method Histograms of Oriented Gradients (HOG) could be used in conjunction with a Nearest Neighbor classifier to produce better performance than all previous methods. Accuracy of this method will be verified using the chars74k dataset, our best performing method achieves an 8.1% increase in performance when tested using 5 training images per class and an increase of 6.8% when using 15 training images per class.
Gonzalez A. (2012) has proposed Direction Histogram (DH) and it is slightly inspired by Gradient oriented of Histogram. We propose to detect the edge pixels of the binarised objects and then to compute the direction of the gradient for each edge pixel. As it is a binarised image, there is only gradient on the edge pixels, so it is faster to compute. Later we quantize the direction of the gradients in the edge pixels into 8 bins: {−135◦, −90◦, −45◦, 0◦, 45◦, 90◦, 135◦, 180◦}, and we compute the histogram for each bin. Experimental results shows the 85.8% matched or classified images, 3.6% found miss-matched object and 10.6% not found over the given data set.
Choudhari A. et al (2013) has proposed threshold method of segmentation; it is difficult to find precise character boundaries without knowledge about characters. So to avoid the error prone process, segmentation approach that is based on selecting the objects or regions that meet the criterion of having area greater than some threshold value has been proposed as one of the segmentation strategies. The objects whose area is greater than 25 and less than 3000 are assumed to be the valid character objects. The objects of area less than 25 are assumed to be some noise and the objects having area greater than 3000 are assumed to be some picture or logo etc. Very promising results are achieved by using this technique to extract character images. Although this technique is having a few shortcomings in cases where a character image is not completely connected, yet it can be extensively used in a variety of applications where the extracted character images are used to train and test a character recognition system like a car number plate recognition system etc.
Khan A. et al (2013) has proposed combination of threshold based and clustering methods as well as Region-edge-based with morphological watershed segmentation method for better segmentation of raw image , this paper shows the comparative study of various hybrid method of segmentation on the basis of different properties of input or captured image. So a single approach to segment all variety of images may be practically unachievable. The prior knowledge on the image can give better results and gives user the choice to decide proper method to segment the image
Arora S. et al (2008) has proposed multiple classifier system. Shadow features are extracted from scaled bitmapped character image. Chain code histogram features are extracted by chain coding the contour points of the scaled character bitmapped image. Intersection features are extracted from scaled, thinned one pixel wide skeleton of character image. Experimental results of the training and test sets a sample set of 4900 handwritten devnagari characters are considered. A training set of 3332 samples and test set of 1568 samples are formed. Combination of individual classifier outputs overcomes deficiencies of features and trainability of single classifiers. Outputs from several classifiers can be combined to produce a more accurate result. Classifier combination takes two forms: combination of like classifiers trained on different data sets and combination of dissimilar classifiers, accuracy observed using this method is approx 92.80% which is much better to using individual method of classification.
This paper is organized as follows: Section I provides introduction about image processing techniques and various literature methods review. In Section II we have discus about problem identification of character or digit recognition which affect cost and quality of recognized character or digit. Section III provides the description of proposed algorithm by various literatures and their pros. and cons. Section IV provides Results and discussions. The conclusion is provided in section V.

II. PROBLEM STATEMENT

In any system, manual intervention decreases the quality of outcomes and increases the error rate. Various use of image processing in the day to day activities in the real time applications and devices, we are facing more challenges to recognition a better character outcomes in case of connected input and varying size of images. It is observed that no one single method applied for recognition of character or digits from raw images which is perfect, there are issues associated with them either quality related or cost related, quality related means outcome of final processed image is blurred or not clear /neat and cost related issue means it has taken much time for processing or recognition of images.
Problems identification

- Each captured image has varying properties of image like pixel intensity, size, dimension and clearness visibility so common automatic methodology for pre-processing not suite best.
- Quality of image also changes the pre-processing criteria so we need first use standard image capturing method and devices with proper specification.
- Main problem in segmentation is to find out ROI (region of Interest).
- Various segmentation algorithms are affected from over/under segmentation issue due to blurred, damaged and ambiguous raw images.
- And some segmentation algorithms are very costly in term of execution (segmentation) time those are not applicable in real time application.
- No one single method of extraction is perfect to extract features for all segments of images.
- Speed and time of extraction also differs as per input image and extraction on the basis of texture or shape.

- Accuracy in feature extraction is also a big challenge or research area for the real time environment.

III. METHODOLOGY

In previous papers authors illustrate about various methods of segmentation and Feature extraction and as well as fine tuning of segmented images, we have studied that various methods are proposed by literatures:-
As per above method proposed by various literatures some best suited Segmentation methods are mentioned here as follows:
1. Clustering Method
2. Region Growing Method
3. Watershed Transformation
4. Threshold
5. Hybrid Approach (combining two or more methods)

1. In clustering method, it is an unsupervised learning technique, where one needs to know the number of clusters in advance to classify pixels [12]. A similarity condition is defined between pixels, and then similar pixels are grouped together to form clusters. Four of the most used clustering algorithms are:
-K-means
-Fuzzy C-means
-Hierarchical clustering

2. The region based segmentation is partitioning of an image into similar/homogenous areas of connected pixels through the application of homogeneity/similarity criteria among candidate sets of pixels. Each of the pixels in a region is similar with respect to some characteristics or computed property such as color, intensity and/or texture. The goal of region growing Image Segmentation is to find regions that represent objects or meaningful parts of objects. Major.
problems of image segmentation are result of noise in the image.
An image domain X must be segmented in N different regions R (1)…… R (N). A simple approach to image segmentation is to start from some pixels which is called seeds representing distinct image regions and to grow them, until they cover the entire image.
For region growing it needs a rule describing a growth mechanism and a rule checking the homogeneity of the regions after each growth step.
The growth mechanism – at each stage k and for each region Ri(k), i = 1,…,N, we check if there are unclassified pixels in the 8-neighbourhood of each pixel of the region border Before assigning such a pixel x to a region Ri(k),we check if the region homogeneity:

P(Ri(k) U {x}) = TRUE , is valid

3. The watershed transformation considers the gradient magnitude of an image as a topographic surface. Pixels having the highest gradient magnitude intensities (GMIs) correspond to watershed lines, which represent the region boundaries. Water placed on any pixel enclosed by a common watershed line flows downhill to a common local intensity minimum (LIM). Pixels draining to a common minimum form a catch basin, which represents a segment.
Watershed Transform: A watershed region is defined as the region over which all points flow “downhill” to a common point. The watershed algorithm uses concepts from edge detection and mathematical morphology to partition images into homogeneous regions [22] and is applied to the gradient magnitude of the image. Steps involved in Watershed method:-

Step1: The boundary values of the pixels of g(x,y) is to be found and the minimum value is to be assigned to Mi. Start flooding by initializing n=min+1 (min corresponds to the minimum boundary value). Let $C_n(M_i)$ as the coordinates in the catchment basin associated with minimum Mi that are flooded at stage n.

Step2: Compute $C_n(M_i) = C(M_i) \cap T[n]$

$$Cn(Mi) = \begin{cases} 1 & \text{for}(x,y)\epsilon Cn(Mi) \text{ and to } T[n] \\ 0 & \text{otherwise} \end{cases}$$

Then let C[n] denote the union of the flooded catchment basins at stage n:

$$C[n] = \bigcup_{i=1}^{R} Cn(Mi)$$

Set $n = n + 1$.
Step3. Derive the set of connected components in T[n] denoting as Q. For each connected component q Q[n], there are three conditions:

- If q∩C[n - 1], is empty, connected component q is incorporated into C[n - 1] to form C[n] because it represents a new minimum is encountered.
- If q∩C[n - 1] contains one connected component of C[n - 1], connected component q is incorporated into C[n - 1] to form C[n] because it means q lies within the catchment basin of some regional minimum.

- If q∩C[n - 1] contains more than one connected component of C[n- 1], it represents all or part of a ridge separating two or more catchment basins is encountered so that we have to find the points of ridge(s) and set them as "dam".

Step4. Construct C[n] using the values obtained for $C_n(M_i)$ and c[n] Set n = n + 1.

Step5. Repeat Step 3 and 4 until n reaches max + 1.

4. Thresholding is the simplest method of image segmentation. From a grayscale image, thresholding [11] can be used to create binary images. It is simplest approaches to segment an image is based on the intensity levels and is called as threshold based approach. The pixels that pass the threshold test are considered as object pixel and are assigned the binary value "1" and other pixels are assigned binary value "0" and treated as background pixels.

$$g(x,y) = \begin{cases} 1 & \text{for } i(x,y) \geq t \\ 0 & \text{for } i(x,y) < t \end{cases}$$

Where g(x,y) is the output image; i(x,y) is the input image and t is the threshold value. Steps involved in thresholding algorithm:-

- Divide the image into sub-image.
- Choose a local threshold for sub-image considered.
- Compare the pixels in that sub-image and segment the region.
- Consider all sub-images individually and choose corresponding threshold values.
- Stop segmentation when all the sub-images are processed.

5. Hybrid methods combine two or more of the basic segmentation methods. Here it has used the combination of several methods compare to its parent approaches and then applied the best combinations of methods to get the best result. For example- The combination of threshold based and clustering methods [10] are used in medical image segmentations and these combinations are selected as per the raw image properties means a perfect combination of methods which easily solve all those issues which could not be taken care by a single method.

Similarly literatures proposed various feature extraction methods are mentioned here as follows:

1. Hole Feature
2. Fourier Descriptor
3. Chain code Feature
4. Direction Histogram

1. Hole Feature: The digit image is processed to identify the presence of holes using the object region boundary extraction method. Initially the starting pixel is located and the boundary of the region is sketched until it reaches the initial co-ordinates again. This gives the actual co-ordinate values of the hole Boundary [13] (inner boundary) and region boundary (outer boundary). The digits 0, 4, 6, 8 and 9 contain the holes. The co-ordinates of the hole boundary was extracted. The feature vectors (1 FV) contains the number of holes (H_{cnt}), starting coordinates of the hole boundary ($start_x$, $start_y$), hole boundary length (H_{len}) and centroid (C_x C_y). Maximum x co-ordinate value (OB_{maxx}) for the outer boundary (OB) and (IB_{maxx}) for inner boundary (IB).

$FV_1=\{H_{cnt},Start_x,Start_y,H_{len},C_x,C_y,OB_{maxx},IB_{maxx}\}$

2. Fourier Descriptors: Fourier descriptors [13] represent the boundary co-ordinates values as complex numbers derived by the Chain code algorithm for the complete shape of the image. The digits like 1, 2, 3, 5 and 7 were implemented with FD. The feature vectors extracted contains centroid (C_x, C_y), minimum coordinates (Min_x, Min_y), maximum co-ordinates (Max_x, Max_y) and x_{val} for the Max_y.

$FV_2=\{C_x,C_y,Min_x,Min_y,C_x,X_{val}\}$

3. Chain code features: The Chain code features were extracted by using the proposed CTCC algorithm for digits 2, 3 and 5. In general the top left pixel was located as starting pixel on the first row. The Chain code was derived from the located starting pixel by considering the 8 connected boundaries. The Chain code [13] values were derived in clockwise or anticlockwise direction till the starting pixel is reached again by covering the complete shape of the object. The above Figure represents the boundary of the segmented image. The digits 2, 3 and 5 share similar characteristics in their shape. The Chain code for the complete shape resulted in duplication of values that increased the length and complexity of Chain code values.

$FV_3=\{ CH_{len}, DV_0, DV_1, DV_2, DV_3, DV_4, DV_5, DV_6, DV_7, DV_8, AVG_{dir} \}$

4. Direction Histogram Feature: The feature used in this paper as Direction Histogram (DH) and it is slightly inspired by histogram feature extraction method [9]. We propose to detect the edge pixels of the binarized objects and then to compute the direction of the gradient for each edge pixel. As it is a binarized image, there is only gradient on the edge pixels, so it is faster to compute. Later we quantize the direction of the gradients in the edge pixels into 8 bins: {−135◦,−90◦,−45◦, 0◦, 45◦, 90◦, 135◦, 180◦}, and we compute the histogram for each bin. The image is divided into 16 blocks in order to have spatial information, and the histograms

for each block are concatenated into a 128-dimensional vector. As this method is based exclusively on the direction of the edge pixels, it is not affected by color neither intensity.

For classification of extracted images: Classification is also a very crucial task after proper feature extraction it is an important portion of decision making in recognition system. The objective of image classification procedure is to categorize the pixels in an image which previously mentioned
classes or identify new classes. Here we have discussed about MLP-Back propagation method and K-NN methods of classification.
1. MLP-Back-propagation: Multilayer perceptron & BP (Back-propagation) model [14] Standard multilayer perceptron (MLP) architecture consists more than 2 layers; A MLP can have any number of layers, units per layer, network inputs, and network outputs. This network has 3 Layers; first layer is called input layer and last layer is called output layer; in between first and last layers which are called hidden layers. This model is the most popular in the supervised learning architecture because of the weight error correct rules. It is considered a generalization of the delta rule for nonlinear activation functions and multilayer networks.

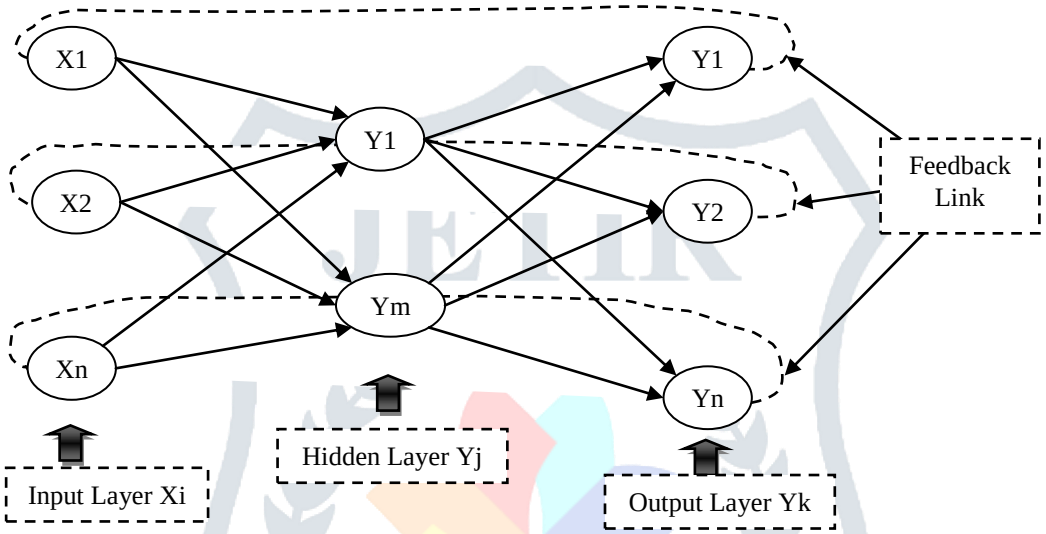

Fig: 3.1 Back propagation Classifier

2. The K-nearest neighbors (KNN) classifier is one of the easiest rather trivial & effective classifiers. KNN classifier finds a group of K samples from the training sets that are close to the test sample, and then the test sample is classified by the majority category of K-nearest neighbors. In other words, to classify an unlabeled sample, the distance of this sample to the entire training data is computed, its KNNs are identified, and the class labels of these KNNs are then used to determine the class label of the test sample. The Euclidean distance is the most widely used similarity (or dissimilarity) metric for KNN classifier. The Euclidean distance:

$$d(x,u) = \sqrt{\sum_{i=1}^{n}(x_i - u_i)^2}$$

- Determine the parameter K i.e., number of nearest neighbors beforehand.
- Distance between the query-instance and all the training samples is calculated using any distance measure algorithm.
- Distances for all the training samples are sorted and nearest neighbor based on the K^{th} minimum distance is determined.
- Since the K-NN is supervised learning, get all the Categories of your training data for the sorted value which fall under K.
- The prediction value is measured by using the majority of nearest neighbors.

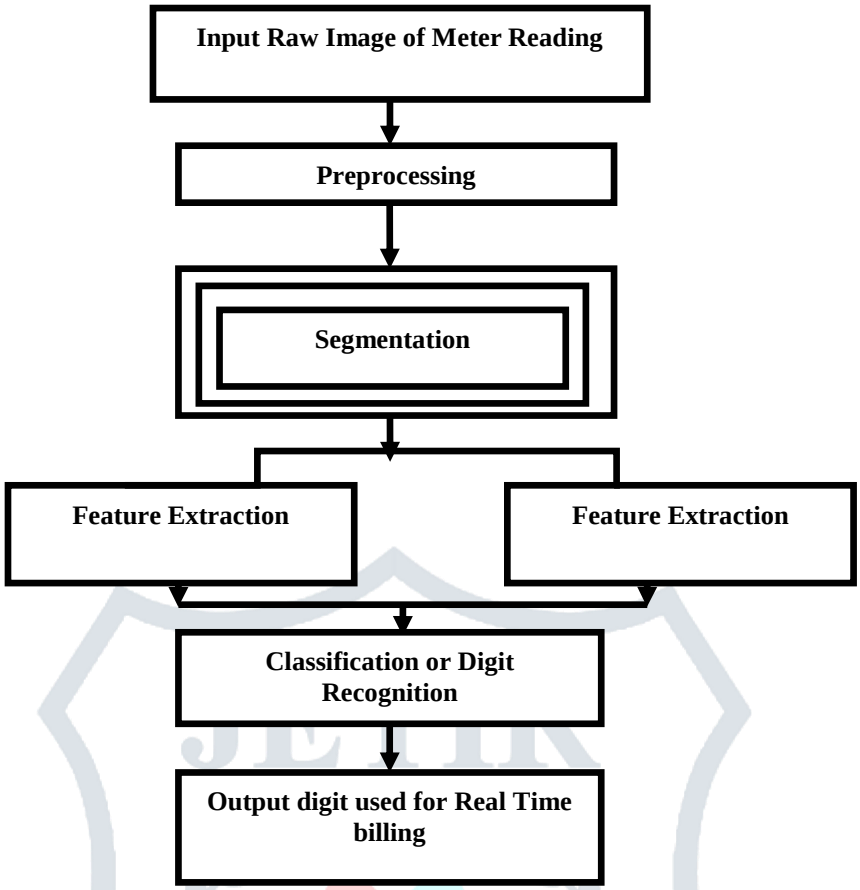

Fig: 3.2 Flow of Proposed Methodology

IV. Survey Result

As per proposed work, combination of selected methods for segmentation and feature extraction , gives good results and their proposed approach is a simple to be implemented in real time embedded system as it is less computational cost, a hybrid image segmentation algorithm with proper segmentation of region of interest without complex approaches, counter the smoothness constraints for segmentation, noise minimized, simple region filling logic, easy implementation using the available algorithms and Morphological operators in a novel way and a better segmented region of interest.

The results show that work has reduced processing time of segmentation and feature extraction. And also increase the accuracy of extracted feature.

- Increase Accuracy of digit recognition
- Reduce Error in Feature extraction
- Improve the performance of classification
- Real time digit recognition

V. Conclusion

The objective of this paper is to find out best suited methods of Segmentation and Feature Extraction while during the study of literatures proposed methods that show the no one single method is perfect for different type of images which are having different properties. So here, we evaluate the properties of multiple-segmentation multiple-feature extraction method to increase the efficiency of recognition of digit and also reduce the over segmentation with less computational time. Using hybrid method of segmentation and multiple feature extraction techniques we can increase the quality of image recognition. Purpose of this paper to accumulates various methods of segmentation and feature extraction for digits.

References

[1] Andrew J. et al (2011)" Natural Image Character Recognition Using Oriented Basic Image Features" International Conference on Digital Image Computing: Techniques and Applications.

[2] Arora S. et al (2008) "Combining Multiple Feature Extraction Technique for Handwritten Devnagari Character Recognition" IEEE Region 10 Colloquium and the Third ICIIS, Kharagpur, INDIA December 8-10;pp 342.

[3] Brhmadesam S.(2010) "A Simple Hybrid Image Segmentation Method for Embedded Visual Display Devices " 3rd International Congress on Image and Signal Processing.

[4] Babu B.(2009) "Character recognition in natural images," in Proceedings of the International Conference on Computer Vision Theory and Applications, Lisbon, Portugal.

[5] Camastra F. (2007). "A SVM-based cursive character recognizer", Pattern Recognition; , (40) 3721 3727.

[6] Chang C. et al (2011) "LIBSVM : a library for support vector machines ", ACM Trans actions on Intelligent Systems and Technology, 2 :27:1–27:27, 2011.

[7] Chaudhari A. et al(2013) "A New Approach to Detect and Extract Characters from Off-Line Printed Images and Text" Information Technology and Quantitative Management (ITQM2013) ;pp 434 – 440.

[8] Crosier M. et al (2010) "Using basic image features for texture classification," International Journal of Computer Vision, vol. 88, pp. 447– 460.

[9] Gonzalez A. et al (2012) "A Character Recognition Method in Natural Scene Images" 21st International Conference on Pattern Recognition. Tsukuba, Japan.

[10] Khan A. et al (2013) "Image Segmentation Methods: A Comparative Study" International Journal of Soft Computing and Engineering (IJSCE). ISSN: 2231-2307, Volume-3, Issue-4, September 2013.

[11] Mello, C.A.B. et al (2005) "Image Thresholding of Historical Documents Using Entropy and ROC Curves". Lecture Notes in Computer Science, v. 3773, p. 905-916.

[12] Narang V. et al (2013) "Devanagari Character Recognition in Scene Images" 12th International Conference on Document Analysis and Recognition.

[13] R. Radha et al (2014) "Digit Recognition Using Hybrid Classifier" 2014 World Congress on Computing and Communication Technologies ,978-1-4799-2877-4/14 2014 IEEE.

[14] Oliveira A. et al (2006) "Optical Digit Recognition for Images of Handwritten Historical Documents" Proceedings of the Ninth Brazilian Symposium on Neural Networks (SBRN'06) ;pp 0-7695-2680-2/06