Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 08, August - 2017

Conversion and Recognition of Handwritten Devnagari Character String into Printed Character String Using KNN

Sushma Pilawan,
SGGSIE & T,
India

Milind Bhalerao
SGGSIE & T,
India

Abhijeet Nandedkar
SGGSIE & T,
India

Sanjiv Bonde
SGGSIE & T,
India

***Abstract*— This paper presents a system for the conversion of handwritten string of Devnagari character to printed character string by using character segmentation approach. 11 different statistical features of segmented characters are extracted which are compared with features extracted from printed string of characters available in training data for cross validation purpose using K- nearest neighborhood (kNN) algorithm.**

Use of handwritten string of Devnagari characters written in different styles and converting it into printed string makes the system more prone to real life application. System mainly works on segmentation of characters using bounding box, after segmentation, features are extracted which is compared with training feature set. We have analyzed our system with existing Devnagari handwritten character recognition systems. In given framework, we have focused on a creating database in different styles and recognizing them as printed characters.

Keywords*— *K-Nearest Neighborhood Algorithm, Connected Components Labeling, Bounding Box, Statistical Feature, Feature Extraction Technique, Handwritten Devnagari string segmentation, Object extraction, Printed String of Characters.

I. INTRODUCTION

Devnagari character recognition system transforms a two dimensional image of text, containing either machine printed text or handwritten text, ideally present in any script, from its image form to machine readable form. This method helps to convert any type of documents such as historical document, newspaper, books even unrestricted documents etc to a comprehensible format.

Many business areas such as writer identification, bank check processing, mail sorting and postal automation uses character recognition system.

Beside that, the character recognition systems are concerned to the research field such as writer authentication and verification Writer authentication is the process of identifying the writer and the originality of the document. OCR technique is widely used technique in many other areas like mail sorting, education, finance, government or private offices. Automation of reading of addresses on letters and parcels is done by OCR technology.

One of the most important application of character recognition system is retrieval of machine editable information from handwritten character. As human writing style varies greatly the conversion of handwritten words into character is very important. This paper explains the system of conversion of handwritten string of characters into machine printed document form.

Large number of research is being carried out in character recognition system. [28] provides details survey of Indian script character recognition. They have provided detail studies scripts used in Indian language such as Hindi, Bangla, Devnagari etc. An intensive survey on properties of different Indian script character recognition is presented by U. Pal and B.B. Chaudhuri [12]. D. Joshi [34] have combined different features such as GLCM and Hu moment feature for the recognition of Marathi barakhadi. P. P. Roy [32] have used five different types of features and HMM classifier for the recognition of Bengali words Different researchers have applied different techniques over printed and handwritten characters as well as numerals. Kumar Singh [19] have used Zernike moments features and obtained accuracy of 80%. U. Pal, et al in [20] obtained the result up to 95.13%. The character recognition system is broadly classified into two categories which are printed and handwritten character recognition. [18]have used different type of text such as noiseless, scattered document and find out the most effective way for segmentation. U pal, P P Roy [17] proposed a novel method for the recognition of curved document. They used the concept of reservoir for the recognition of curved document text B. B. Chaudhuri and U. Pal [24] have proposed different methods for skew angle detection in the scanned documented Bangla script. They obtained mean and standard deviation of given skewed document and apply Hough transform, to find connected components in the binary document image . After that find cluster and angle between left most and right most white pixel and obtained skew angle. Utpal Garain and Bidyut B. Chaudhuri [23] used analysis of fuzzy multifactorial for effective segmentation and identification of touching printed Devnagari as well as Bangla character. They obtained recognition accuracy of 98% for good print and paper quality documents and recognition accuracy comes down to 85–90%.

for documents which were degraded. S. Arora [25] have used different set of features such as Shadow, Chain Code Histogram of Character Contour, View based features etc along with MLP to obtain accuracy of 89.58% for the recognition of handwritten Devnagari character. [11] have proposed a novel method for segmentation of unconstrained handwritten text-lines. Technique is used to remove foreground portion of the document image and sobel operator is applied for smoothing purpose. The recognition accuracy obtained up to 93% using Block- based Hough transform .Alireza Alaei, Umapada Pal, P. Nagabhushan [10] have

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 08, August - 2017

proposed a noval technique for recognition of Persian and Arabic handwritten character. [4] proposed HMM-based Indic handwritten word recognition using zone segmentation. They used Global histogram based OTSU method for image preprocessing and run length smoothing algorithm for line segmentation and used four different types of features which are PHOG feature, LGH feature, GABOR feature, G-PHOG feature, Marti– Bunke feature. Out of all features G-PHOG which is the combination of Gabor and FOG presents maximum results. For the preprocessing of Handwritten digit images Sandhya Arora, Debotosh Bhattacharjee, Mita Nasipuri [9] used Multi Layer Perceptron(MLP) based classifier on 4900 samples the overall recognition rate observed is 92.80%. The different features used are intersection point, shadow of image, chain code histogram and number of straight line. Firstly performing scaling of character and extracting three features after that. One of the most important application of character recognition system is conversion of handwritten text into printed text.[33] have used structural and stoke based features for character recognition. Very few research on conversion of handwritten text to printed text system is done till date. Kavallieratou, Ergina, and Stathis Stamatatos.[4] have proposed a novel approach to discriminate between handwritten and printed text, fir that purpose they used IAM-DB and GRUHD databases of English and Greek database of mixed i.e. printed and handwritten text lines.[5] Zheng, Yefeng, Huiping Li, and David Doermann have proposed an method for segmentation and identification of handwritten and machine printed text. The proposed system have achieved 72.19 % extraction for handwritten words. Bloomberg, Dan S [6] have used a morphological operator to separate the handwritten annotation from the text line segment. Chiang, Mike W [7] have provided a method to convert displayed text which was initially written in iragana and/or Katakana characters to Kanji characters. Keskar, Dhananjay, John Light, and Alan McConkie [8] have obtained a system for the conversion of speech to text document. Many people have used statistical feature, [30] have used these feature for maraathi number recognition. The six closest neighbour connected components (CCs) method rather than 4 or 8 connected components has been embraced to perceive written by hand numerals. The ANN classifier strategy has been utilized to perceive the numerals in the Stick code. In this paper, the advancement of correct scanner tag to each Stick code is proposed for sorting the postal letters consequently. [29]

This paper presents a system for offline handwritten character string segmentation into separate characters then conversion into printed text and reconstruction using Microsoft Office. For that purpose, 11 statistical features are extracted and classified by using K nearest neighbour (KNN) algorithm. In the classification stage, experiments are carried out on train and test features using KNN classifier. Section 2 explains the proposed methodology, in section 3 results are discussed. Section 4 gives conclusion and future work.

II. PROPOSED METHODOLOGY

For the conversion of handwritten string of characters into text string of characters, the system is developed which extracts statistical features of handwritten string available in test data and then compares it with features of printed characters available in training set. Before feature extraction the string of characters is preprocessed and segmented into individual characters. El Abed, Haikal, and Volker Margner [2] have compared different features and pre-processing techniques for offline Arabic handwritings. Several feature Extraction methods are available for Devnagari characters and digits recognition but the time required and accuracy is not optimized. Golait, Snehal S., and Latesh G. Malik [27] suggested faster efficient and optimized feature extraction method for character and digits recognition. In a further subsection, the features extraction and the classification is clarified in detail. Figure 2 show the proposed segmentation and recognition algorithm.

Methodology

The steps or techniques applied before processing by correcting images from various errors is called pre-processing the pre-processing is to be done before image enhancement. The input to the system can be handwritten string of characters i.e the handwritten characters written one after other.

First, confirm that you have the correct template for your paper size. This template has been tailored for output on the A4 paper size. If you are using US letter-sized paper, please close this file and download the file “MSW_USltr_format”.

A. Preprocessing

The input to the system can be presented in RGB format but it must be present in the gray scale for further processing. So the main purpose of pre-processing is to convert the given image into grey scale and make the image suitable for feature extraction for that purpose we perform process like binarization obtaining bounding box etc.Firstly the input image is resized to 256×256.

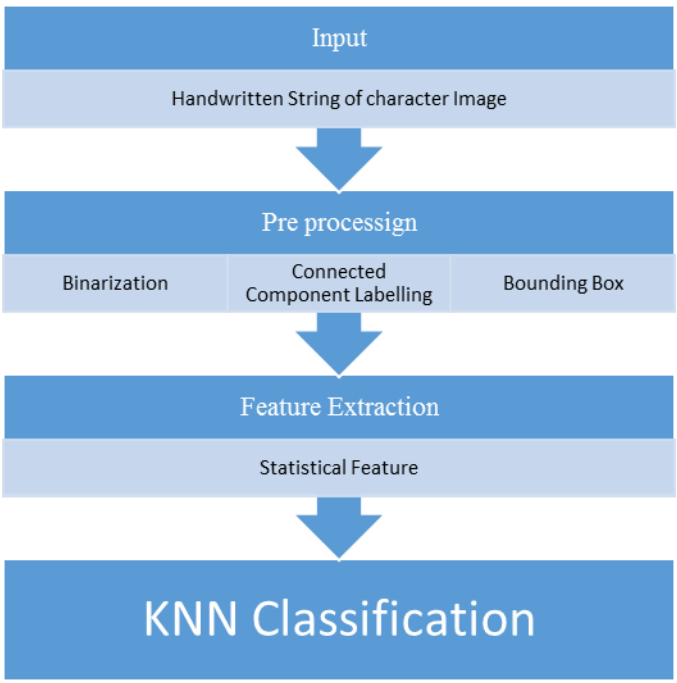

Figure 2: Flowchart

The input to the system can be presented in RGB format but it must be present in the gray scale for further processing.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 08, August - 2017

So the main purpose of pre-processing is to convert the given image into grey scale and make the image suitable for feature extraction for that purpose we perform process like binarization obtaining bounding box etc.Firstly the input image is resized to 256×256.

a. Character Separation :

Step 1: Image Binarization:

- The OSTU [26] method is employed to convert the gray scale image into binary image. The advantage of using this method is that it chooses threshold automatically to reduce the intra class difference between black and white pixels.

$$\sigma_w^{\,2}(t) = w_0(t)\sigma_0^2(t) + w_1(t)\sigma_1^2(t)$$

where w_0 and w_1 are the class probabilities separated by a threshold t and

σ_0^2 and σ_1^2 are the variances of these two classes.

The class probability w_0 and w_1 is computed from the L histogram, L histogram is the histogram of the image pixel intensity.

$$w_0(t) = \sum_{i=o}^{t-1} p(i)$$

$$w_1(t) = \sum_{i=t}^{L-1} p(i)$$

b. Connected Component labelling:

Connected components labelling is used to label the image pixels into different groups depending upon the 8 neighbourhood connectivity. It follows the following algorithm:

- Scan the image along a row up to point p, where p is the pixel to be labelled.
- Find the pixel in the image having pixel intensity V={1}. When this is true, it examines the four neighbours of p which have already been encountered in the scan.
- Based on this information, the labelling of p occurs as follows:

If all four neighbours are 0, assign a new label to p, else

if only one neighbour has V={1}, assign its label to p, else

if more than one of the neighbours have V={1}, assign one of the labels to p and make a note of the equivalences.

c. Finding the bounding box:

Bounding box is the most important process for character separation. It finds out each separated character and forms the box around each character. Bounding box is used to represent the region of interest i. e. the character. It returns the region of intrest in the form of pixel coordinates. The line joining the coordinate forms the bounding box.

Once the bounding box is obtained, each separate box represents one character. All these characters are stored in separate folder.

III. FEATURE EXTRACTION

For accomplishing high recognition rate, the determination of features extraction strategy is essential. After pre-processing and character separation the feature extraction process is applied over each separated characters.

Different features which are extracted are:

1. Mean: Mean value gives the contribution of individual pixel intensity for the entire image. This feature can be used for noise removal and low pass filtering. Mathematically given by:

$$\mu = \frac{x_1 + x_2 + x_3 + ...x_N}{N}$$

Where,

$x_1, x_2, \ldots\ldots x_N$ represents the image pixel intensity,

μ being the mean of image.

And N is the total number of pixel.

2. Variance: Variance is normally used to find how each pixel varies from the neighbouring pixel (or centre pixel) and is used in classify into different regions

$$v = \frac{1}{N}[(x_1 - \mu)^2 + (x_2 - \mu)^2 + ...+ (x_N - \mu)^2]$$

3. Standard Deviation: It is calculated by taking the under root of variance given by:

$$\sigma = \sqrt{\frac{1}{N}[(x_1 - \mu)^2 + (x_2 - \mu)^2 + ...+ (x_N - \mu)^2]}$$

4. Regional Area: This is the ratio between number of the pixels in the skeleton to the total number of pixels in the image.

$$\text{RegionalArea} = \frac{\text{No. of pixel in skeleton}}{\text{Total no. of pixel}}$$

5. Centroid: Represents the centre of image. The two elements gives the x and y co- ordinates of the centre of image.
6. Eccentricity: It is the eccentricity of the smallest ellipse which fits the skeleton of the image. Its value ranges between 0 and 1.

$$\text{Eccentricity} = \frac{\text{Distance between the foci of the ellipse}}{\text{major axis length}}$$

7. Equi diameter: Specifies the diameter of a circle with same area as region. Where the area of the region is given by,

$$Area = \sqrt{\frac{4 * area}{\Pi}}$$

8. Extent: It is obtained by dividing the area of whole image by the area of the bounding box.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 08, August - 2017

9. Major Axis Length: It is calculated as the length given by obtaining the major axis of the ellipse that has the moment same as that of the region moment

   Major Axis= a+b

   Where,
   a,b are the distances from each focus to any point on the ellipse
10. Minor Axis Length: It is calculated as the length given by obtaining the minor axis of the ellipse that has the moment same as that of the region moment

    $$\text{Minor Axis} = \sqrt{(a+b)^2 + f^2}$$

    where
    f is the distance between foci.
    a,b are the distances from each focus to any point on the ellipse.
11. Orientation: The angle obtained between the major axis and the x-axis of the ellipse that has the comparable second moments as the area.

Perimeter: It represents the distance which is calculated around the boundary. It is given by distance between adjacent pair of image pixel.

IV. CLASSIFICATION

Statistical classifier along with artificial neural network are most widely used classifier [13], [14]. In case of statistical classifiers, the features obtained are presented in the form n-tuples or vectors.[31] have used multi class svm for handwritten character recognition. The objective of such classifiers is to calculate the character probability of a character and evaluating the classes which are possible. The method used for obtaining classes can be categories into parametric and nonparametric classification. Parametric classifiers consists of Linear Discriminant Function (LDF) [15] and Quadratic Discriminant Function (QDF) [15]. Non parametric classifier mainly include K-Nearest Neighbour classifier [16]. The kNN classification is a statistical method of classification without training phase. The main objective of kNN is to estimate a training data with the X pattern which is nothing but the test data, the kth pattern which is closest to the pattern X is selected. The most frequently occurred class is the class of pattern X in all the k patterns.

Steps in KNN

Step 1. Store every input training samples with its label.

Step 2. To do a prediction for a test sample, compute its distance from every training
example.

Step 3. Then, keep the k training samples which are having minimum distance from
the test sample, where k 1 and an integer.

Step 4. Check the most common label from this k minimum distance training samples.

This common label is the prediction for the test example. There are many distance metrics used in KNN. Most commonly used comparing parameters are :Euclidian distance (1), Manhattan distance (2) or Minkwoski distance (3). The constant k is a number which is small in magnitude. These distances are given by:

$$d(x,y) = \sqrt{\sum_{i=}^{n}(x_i - y_i)^2} \quad (1)$$

$$d(x,y) = \sum_{i=1}^{n}|x_i - y_i| \quad (2)$$

$$d(x,y) = \sqrt[p]{\sum_{i=1}^{n}(x_i - y_i)^p} \quad (3)$$

Where x is the feature vector of pattern which is to be recognized and y is a pattern of training data, n is feature vector size and p is a constant which is defined experimentally.

V. RESULT AND CLASSIFICATION

B. *Database*

For the recognition of the handwritten Devnagari character string we collected Devnagari character string handwritten by different student of our institute. These collected samples are scanned by Hp Scanjet Enterprise flow 7000s2 with resolution of 150 dpi. We have collected string of 3 or 4 consecutive Devnagari characters picked up at random. It can be extended to any number of characters depending upon system requirement. Some of the sample images taken from the database are shown in figure 4. These string of character are needed to be segmented. The characters after segmentation are shown in figure 5.

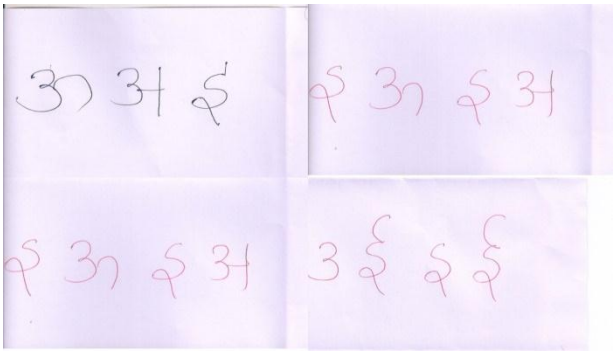

Fig. 4.Sample character string images in database

These images are preprocessed and bounding box are obtained as shown in fig. 5

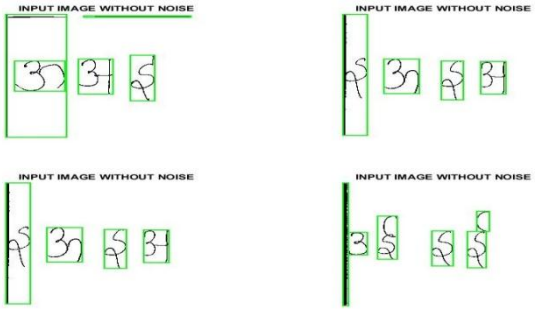

Fig. 5.Bounding box of character string.

After forming a bounding box these characters are needed to be separated. Fig. 6 shows segmented characters.


Segmented1


Segmented2


Segmented3

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 08, August - 2017

Fig. 6.Segmented Characters.

Once separated characters are available in binary form, different features of test image are extracted. Below table shows the feature vector extracted for the first image. As shown in above figure, first image after segmentation consists of three characters, so all the features are extracted for each image. The statistical features combines the three i.e mean, variance, standard deviation of the each character, and other features are represented separately. Combination of all these feature gives a feature vector of 1×10 dimension.

TABLE I:FEATURE VECTOR OF IMAGE 1

Feature	Segmented 1	Segmented 2	Segmented 3
Statistical	3193	66	1038
Area	32.61	7.15	33.17
Centroid	1.62	44.93	56.40
Equi Diameter	63.76	9.16	36.35
Eccentricity	0.27	0.93	0.94
Extent	0.74	0.94	0.60
Orientation	75.96	1.40	87.58
Perimeter	5.31	31.57	2.57
Major Axis Length	75.63	15.57	75.08
Minor Axis Length	72.64	5.58	24.07

These feature are compared with train feature using kNN classifier. Train data consists of printed characters whose statistical features are calculated and compared with test data which consists of handwritten characters As the shape of handwritten and printed character differs with very small distance. So the kNN classifier will obtain minimum distance between the handwritten testing and printed training data, which results in the conversion of handwritten to printed character. The overall recognition accuracy of 91% is obtained using this classifier. Fig.7 shows the final recognition result.

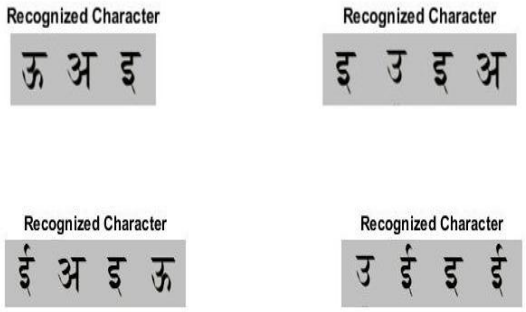

Fig.7. Recognition Result

C. *Comparision of results*

For handwritten Devnagari string of characters which is converted into printed character string, very few research work is done. So we have compared obtained results with the results of work done in the field of handwritten character recognition system. As a lot of work done previously in this area for Devnagari characters and according to the comparison, we have obtained a promising stand as our system have correctly recognised printed characters over handwritten database. Recognition results obtained on our own database are quite satisfactory. We hope results reported in this paper will be useful to the researchers for future work. Previous work comparison for the handwritten Devnagari character recognition is shown in table 2.

TABLE II: THE RECOGNITION RATE FOR HANDWRITTEN DEVNAGARI CHARACTERS

Reference	Features	Classifier	Results
[3]	PHOG	HVM	82.11
	LGH	SVM	79.29
	Gabor		48.47
	G- PHOG		66.27
[9]	1.Intersection	MLP	1.36.71
	2.Shadow		2.60.59
	3.Chain Code Histogram		3.64.90
[1]	Directional chain code	Quadratic	80.36
[16]	normalized distances	coarse	90.65
[Proposed]	Statistical	kNN	91

D. *Results Analysis*

The conversion of handwritten string of characters into printed character string is very important field as any computer application require data to be present in printed form. So the system presented in the paper is very important field. The system presented here is limited to simple characters only. We tried to extend the system for compound character it resulted in the recognition result as show in fig. 8

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 08, August - 2017

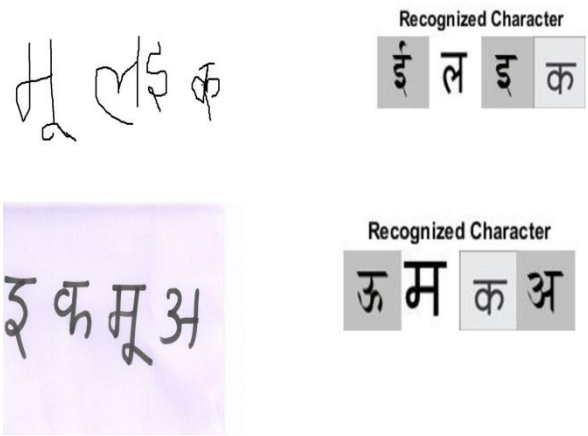

Fig.8. Error in Recognition Result

IV. CONCLUAION

In this paper, a system for Devnagari character string segmentation and recognition into printed character string using statistical features followed by kNN classifier is developed. The effectiveness of the proposed system is evaluated by observing the result over different string of characters developed by different individuals in different style. Misclassification is observed by adding compound characters in the given test database

In future, there is chance to improve the system by considering compound characters as well. Also the database can be developed to convert handwritten words into printed text which can be useful for many real life application.

REFERENCES

[1] Sharma, Nabin, et al. "Recognition of off-line handwritten devnagari characters using quadratic classifier." Computer Vision, Graphics and Image Processing (2006): 805-816.

[2] El Abed, Haikal, and Volker Margner. "Comparison of different preprocessing and feature extraction methods for offline recognition of handwritten arabicwords." Document Analysis and Recognition, 2007. ICDAR 2007. Ninth International Conference on. Vol. 2. IEEE, 2007.

[3] Roy, ParthaPratim, et al. "HMM-based Indic handwritten word recognition using zone segmentation." Pattern Recognition 60 (2016): 1057-1075.

[4] Kavallieratou, Ergina, and Stathis Stamatatos. "Discrimination of machine-printed from handwritten text using simple structural characteristics." Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on. Vol. 1. IEEE, 2004.

[5] Zheng, Yefeng, Huiping Li, and David Doermann. "Machine printed text and handwriting identification in noisy document images." IEEE transactions on pattern analysis and machine intelligence 26.3 (2004): 337-353.H. Fujisawa, Y. Nakano, and K. Kurino, "Segmentation Methods for Character Recognition: From Segmentation to Document Structure Analysis,"Proc. IEEE, vol. 80, no. 7, pp. 1079-1092, 1992.

[6] Bloomberg, Dan S. "Segmentation of handwriting and machine printed text." U.S. Patent No. 5,181,255. 19 Jan. 1993.

[7] Chiang, Mike W. "Text conversion method for computer systems." U.S. Patent No. 6,154,758. 28 Nov. 2000.

[8] Keskar, Dhananjay, John Light, and Alan McConkie. "Correlating handwritten annotations to a document." U.S. Patent Application No. 09/896,123.

[9] Arora, Sandhya, et al. "Combining multiple feature extraction techniques for handwritten devnagari character recognition." Industrial and Information Systems, 2008. ICIIS 2008. IEEE Region 10 and the Third international Conference on. IEEE, 2008.

[10] Alaei, Alireza, Umapada Pal, and P. Nagabhushan. "A comparative study of persian/arabic handwritten character recognition." Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on. IEEE, 2012.

[11] Alaei, Alireza, Umapada Pal, and P. Nagabhushan. "A new scheme for unconstrained handwritten text-line segmentation." Pattern Recognition 44.4 (2011): 917-928.

[12] Pal, U., and B. B. Chaudhuri. "Indian script character recognition: a survey." pattern Recognition 37.9 (2004): 1887-1899.

[13] Liu, C.-L. et al. Handwritten digit recognition: Benchmarking of state-of-the-art techniques. Pattern Recognition, v. 36, p. 2271-2285, 2003.

[14] Liu, C.-L; Sako, H.; Fujisawa, H. "Performance evaluation of pattern classifiers for handwritten character recognition", Ineternacional Journal on Document Analysis and Recognition - IJDAR, v. 4, p. 191-204, 2002.

[15] Duda, Hart, and Stork, "Pattern classification" Chapter 5, Wiley, 2000.

[16] Hanmandlu, Madasu, OV Ramana Murthy, and Vamsi Krishna Madasu. "Fuzzy Model based recognition of handwritten Hindi characters." Digital Image Computing Techniques and Applications, 9th Biennial Conference of the Australian Pattern Recognition Society on. IEEE, 2007.

[17] Pal, Umapada, and ParthaPratim Roy. "Multioriented and curved text lines extraction from Indian documents." IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 34.4 (2004): 1676-1684

[18] GUPTA, DEEPA, and LeemaMadhu Nair. "Improving Ocr By Effective Pre-Processing And Segmentation For Devanagiri Script: A Quantified Study." Journal of Theoretical

[19] Kumar S, Singh C. A study of zernike moments and its use in devnagari handwritten character recognition. In Intl. Conf. on Cognition and Recognition 2005 Dec 22 (pp. 514-520).

[20] Kumar, Manish, R. K. Sharma, and G. S. Lehal. Degraded Text Recognition of Gurmukhi Script. Diss. 2008.

[21] Pasha, Saleem, and M. C. Padma. "Handwritten Kannada character recognition using wavelet transform and structural features." Emerging Research in Electronics, Computer Science and Technology (ICERECT), 2015 International Conference on. IEEE, 2015.

[22] Liu, Cheng-Lin, and Hiromichi Fujisawa. "Classification and learning methods for character recognition: Advances and remaining problems." Machine learning in document analysis and recognition. Springer Berlin Heidelberg, 2008. 139-161.

[23] GUPTA, DEEPA, and LeemaMadhu Nair. "Improving Ocr By Effective Pre-Processing And Segmentation For Devanagiri Script: A Quantified Study." Journal of Theoretical and Applied Information Technology 52.2 (2013): 142-153. And segmentation for devnagiri script :A quantified study- 2013-ISSN

[24] Chen, Jin, et al. "Gabor features for offline Arabic handwriting recognition." Proceedings of the 9th IAPRInternational Workshop on Document Analysis Systems. ACM, 2010.

[25] Garain, Utpal, and Bidyut B. Chaudhuri. "Segmentation of touching characters in printed Devnagari and Bangla scripts using fuzzy multifactorial analysis." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 32.4 (2002): 449-459.

[26] Chaudhuri, B. B., and U. Pal. "Skew angle detection of digitized Indian script documents." IEEE Transactions on pattern analysis and machine intelligence 19.2 (1997): 182-186.

[27] Arora, Sandhya, et al. "Study of different features on handwritten Devnagari character." Emerging Trends in Engineering and Technology (ICETET), 2009 2nd International Conference on. IEEE, 2009.

[28] Level Otsu N. A threshold selection method from gray-level histogram. IEEE Transactions on Systems, Man and Cybernetics. 1979;9(1):62-6.

[29] Golait, Snehal S., and Latesh G. Malik. "Review on Feature Extraction Technique for Handwritten Marathi Compound Character Recognition." Emerging Trends in Engineering and Technology (ICETET), 2013 6th International Conference on. IEEE, 2013.

[30] Pal, U., and B. B. Chaudhuri. "Indian script character recognition: a survey." pattern Recognition 37.9 (2004): 1887- 1899.

[31] Velu, C. M. "Automatic letter sorting for Indian Postal Address Recognition System based on PIN codes." Journal of Internet and Information Systems 1.1 (2010): 6-15.

[32] Vaidya MV, Joshi YV, "Marathi numeral recognition using statistical distribution features". In Information Processing (ICIP), 2015 International Conference on 2015 Dec 16 (pp. 586-591). IEEE

[33] Naeem Ayyaz, Imran Javed and Waqar Mahmood, "Handwritten Character Recognition Using Multiclass SVM Classification with Hybrid Feature Extraction", Pak. J. Engg. & Appl. Sci, Vol. 10, pp. 57-67, 2012.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 08, August - 2017

[34] Roy PP, Bhunia AK, Das A, Dey P, Pal U. HMM based Indic handwritten word recognition using zone segmentation. Pattern Recognition 2016 Dec 31; 60: 1057-75.

[35] Shi CZ, Gao S, Liu MT, Qi CZ, Wang CH, Xiao BH. Stroke Detector and Structure Based Models for Character Recognition: A Comparative Study. IEEE Transactions on Image Processing. 2015 Dec;24(12):4952-64.

[36] Joshi D, Pansare S. Combination of Multiple Image Features along with KNN Classifier for Classification of Marathi Barakhadi. In Computing Communication Control and Automation (ICCUBEA), 2015 International Conference on 2015 Feb 26 (pp. 607-610) IEEE.