*International Journal of Engineering & Technology*, 7 (4.19) (2018) 1055-1059

International Journal of Engineering & Technology

*Website: www.sciencepubco.com/index.php/IJET*

***Research paper***

# Study of Different Features and Classification Techniques for Recognition of Handwritten Devanagari Text

**Vijay More[1*], M U Kharat[2*], S V Gumaste[3*]**

[1.] *Assistant Professor,*
[2.] *HOD,*
[3.] *Research Coordinator*
[*.] *Department of Computer Engineering, MET's Institute of Engineering, Bhujbal Knowledge City, Adgaon, Nashik, MS, India.*
**Corresponding author E-mail: vbmore2005@rediffmail.com*

**Abstract**

Devanagari script is most popular and an older script in India. Millions of people all over the globe are using Devanagri script for various purposes such as communication, understanding the history, record keeping, research, etc. Recognition of handwritten Devanagari word is one of the popular area of research from decades because of its wide scope of applications. Different features and techniques of classification are the most important steps in the process of recognizing Devanagari handwritten word, are described in this paper.

***Keywords***: *classification; Devanagari script recognition; feature extraction; feature selection; pattern recognition.*

## 1. Introduction

India is multilingual, regional and multiscript country. Indian Community using various regional-languages in practice like Sanskrit, Marathi, Urdu, Hindi, Bengali, Tamil, Telugu, Punjabi etc. Hindi is the national language of India. Marathi, Hindi, Konkani, Sanskrit are the languages comes under Devanagari script. Language recognition and identification of writer are most promising research area, which also covers very important area of applications in forensic science, banking, security, graphology etc.
In India, the state governments are using regional language (script) for recording the decisions, minutes of meeting notes, comments and suggestions given by higher authority to their subordinates, etc. The authorship of written decisions, comments, meeting notes, and suggestions given need to be recognized in disputed documents.
Handwritten Devanagari script Recognition is an intelligent computer application that receives and interprets intelligible handwritten input which is then passed through various automated process to recognize and generate equivalent script. Script recognition is a three step process which contains preprocessing, feature extraction and classification (recognition). Preprocessing is the process of removing noise from source images so that these noiseless images can be further used efficiently extracting unique features. Various features are extracted in feature extraction stage. Recognition is nothing but classification of extracted features with the available or stored features for those character/word sets. This paper concentrates only on the study of various (appropriate) features and classification techniques.
Handwritten character recognition is not an easy task because there are inherent variations of the style of writing among different writers. The differences and styles or patterns in writing character and their sizes will make the recognition task very difficult. It makes Feature extraction is one of the critical dimension in image processing and image recognition domain. Features are uniquely identified components of characters. The basic task or need of feature extraction and selection is to find a group of most prominent and effective features for classification process. This may also includes compressing the high-dimensional feature space into low-dimensional feature space maintaining classification accuracy high.

### 1.1 Characteristics and symbol set of Devanagari script

#### 1.1.1 Vowels

| | |
|---|---|
| Vowels | अ आ इ ई उ ऊ ए ऐ ओ औ अं अः |

**Fig 1:** Vowels of Devanagari Script

#### 1.1.2 Consonents

| | |
|---|---|
| Consonants | क ख ग घ ङ च छ ज झ ञ ट ठ ड<br>ढ ण त थ द ध न प फ ब भ म<br>य र ल व ह श स ष क्ष ज्ञ |

**Fig 2:** Consonants of Devanagari Script

#### 1.1.3 Numerals

| | |
|---|---|
| Numerals | ० १ २ ३ ४ ५ ६ ७ ८ ९ |

**Fig 3:** Numerals of Devanagari Script

#### 1.1.4 Modifiers

| | |
|---|---|
| Modifiers | ा ि ी ु ू ृ ॅ ॆ े ै ॉ<br>ॊ ं ः |

**Fig 4:** Modifiers for Devanagari Script

**Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.**

### 1.2 Shirorekha

Shirorekha or headline is the upper horizontal line of the character or word as in figure 5. It does not contain any useful information so it should detect and then discarded using Shirorekha detection algorithm.

भारतमाता की जय

**Fig 5:** Shirorekha example

### 1.3 Vertical Bars in the characters

1. End bar characters

अ ख घ च ज झ ञ त थ ध न प ब भ म य ल व श स

2. Non bar characters

इ उ ऊ ए छ ट ठ द ड ढ र ह

3. Middle bar characters

क ऋ फ

**Fig 6:** Vertical bars in the characters

## 2. Recognition of Devanagari Handwritten Word

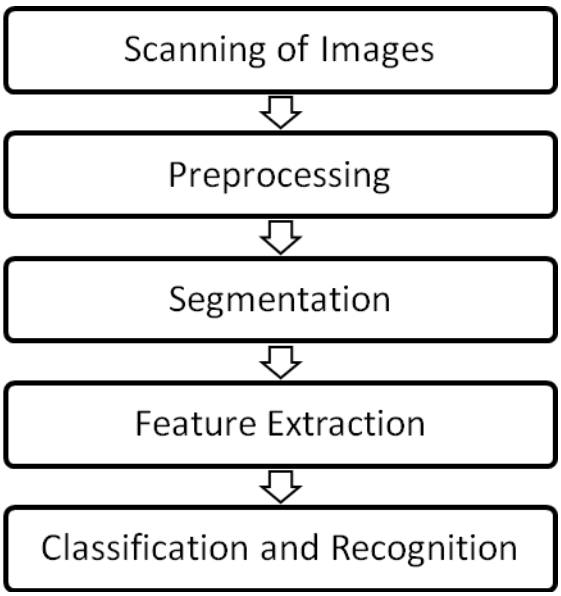



**Fig. 7:** Stages of Handwritten Script Recognition

### 2.1. Pre-processing

This pre-processing stage is very important stage in the process of text recognition in that the scanned digital image assumed to have noise. The noise is the main hurdle in the process of text recognition accurately. Examples of noise are quality of the paper, unwanted pen stokes, ink float mark due to water contact, tea-coffee floated on the written paper, etc. There are numerous techniques available to remove or to reduce the noise from the image. Noise removal should be handled carefully because this process may remove original data. Preprocessing stage is required because it is used to produce a noiseless image of character which can be further used for segmentation and then recognition.

#### 2.1.1 Binarization

Structural symmetry of strokes is used to estimate the local threshold by the structural symmetric pixels (SSP) of the region to suppress the pixels that do not contain text. This procedure is called Local threshold binarization method. The region surrounding the strokes whose gradient magnitudes are large enough and having opposite directions can be declared as SSP. Background surface is also extracted using SSP in the compensated image to deal with degradations of document images like non uniform illumination, stain and low contrast. It is nothing but transforming gray scale image into black and white image.

#### 2.1.2 Size Normalization

This is the step in which segmented character is normalized to fit within the matrix so that all characters have same data size.

#### 2.1.3 Thresholding

Thresholding is the process of reducing an original image by a designated factor into binary image. Means converting or aligning binary shape images obtained from edge detection or boundary detection or thresholding to 1-pixel wide lines. The threshold version of hand written or printed alphanumeric can be thinned for better representation and further processing.

#### 2.1.4 Noise Removal

It is necessary to eliminate noise like disconnected lines, gap of lines, incomplete matra / vilanti etc. Median Filtering, Wiener Filtering method and morphological operations can be performed to remove noise. Sobel technique can be used to detect edges in binarized Image.

### 2.2 Segmentation

In this stage Pre-processed image is segmented into paragraphs, lines, words and characters. Segmentation process deeply involves segmentation of paragraph (paragraph segmentation), segmentation of line (line segmentation), segmentation of word (word segmentation) and segmentation of character (character segmentation). Devanagari script words can be split into character by removing Shirorekha and then can be recognized. Paragraph segmentation can be skipped. Here, overlapped lines must be dealt carefully.

### 2.3 Feature Extraction

Features are the sense of information which was extracted from the input scanned and cleaned image. All this information must be similar for similar images but must not be same for other images. These features play major role in pattern recognition. Thus the selection of feature extraction technique becomes a key aspect in achieving high recognition performance. In other words, this stage is used to remove or reduce redundancies from data. Difference between feature extraction and feature selection must be clarified before processing. Feature extraction is a technique used to extract unique features from the input image. On the other hand, feature selection is the process of selection of most relevant features from available pool of features which helps to improve the classification accuracy of the extracted features.

### 2.4 Types of Features

For recognition approaches, extracting a set of appropriate, unique and useful features with usage of efficient method/s of extraction are important factors in getting high recognition performance.

- Structural Features: Crossing points, loops, branching points, zigzags and strokes in different directions, dots and their positions, length of contour segment.
- Statistical Features: Centre of gravity (mean), Characteristics loci, zoning, crossing and distances, projections and profiles are statistical features.
- Global Transformations: Method of obtaining n-dimensional feature vector that represents the pattern of the contour of the images using suitable transform like Fourier transform.
- Local Features: Large number of small component of text, large number of sub-images which contains text pixels which are at fine grained level.
- Texture Features: mean, mean deviation and slope of the regression line are the features extracted by 2D DWT using normalized co-occurrence histograms.

There are various features that can be extracted are helpful in recognition of script which are Statistical and Structural Features, Grey Level Co-Occurrence Matrix, Chain Code Histogram Features, Gabor Filters, Stroke Features, Gradient Features, Fractal Features, Project Profile Features, Wavelet Features, Morphological Geometrical Features, Texture Features, Steerable Pyramid Transform, Features of Connected Component, Features based on Water Reservoir, Topological Features, Moment Features etc.

### 2.5 Classification and Recognition

There are various classification algorithms identified in the field of script identification. These are Support Vector Machines (SVM), Neural-Network (NN), K-Nearest Neighbour (KNN), and Linear Discriminent Classifier (LD). These are the most prominent classification algorithms used for script identification. Some of the classifiers illustrated below:

#### 2.5.1 SVM Classifiers (SVM)

The SVM is a classifier applied for two class problem which looks for the optimal hyper plane and maximizes the distance, margin between nearest examples of both classes, named support vectors (SVs). The linear SVM is further extended to various non-linear variants. Experiments with Gaussian kernel SVM shown better performance compared to non-linear SVM kernels. Gaussian kernel gives highest accuracy in most of the cases.

#### 2.5.2 Neural Network (NN)

Classification can also be done by a two-layer feed-forward neural network [3] with back-propagation. The topology of such NN classifier has an input and output layers with embedded neurons which are determined by the length of feature vector and number of classes available, respectively. To minimize the squared error between the network output values and the target values, this network can be used with the gradient descendent rule. The trained network is then used to test an input image entered by the user and the classifier maps input pattern to a number of classifications.

#### 2.5.3 K-Nearest Neighbour (KNN)

K-nearest Neighbour is non-parametric classifier, where posterior probability is estimated from the number of occurrences of nearest neighbours of the unknown pattern. The key idea behind k-nearest Neighbour classification [6] is that, similar observations always belong to same classes. The test image is classified with its k-nearest Neighbour belongs to certain pre-defined region. Feature vectors stored a-priori is always used to decide the nearest Neighbour of the given test image feature vector.

## 3. Detailed Study of Various Techniques

Sk Md Obaidullah et.al, Huge number of features were considered by authors:

1. Length of Perimeter of Connected Component
2. Maximum Width and Height
3. Area: Number of Pixels inside the boundary
4. Aspect Ratio: Width / Height
5. Rectangularity: A / (W x H) (A- Area, W- Width, H- Height)
6. Circularity: C = A/B, A is the area of the test shape and B is the area of the circle
7. Curvature
8. Energy of Curvature
9. Mean Distance
10. Mean, standard deviation, skews of vertical centroid, relative horizontal centroid, number of holes, sphericity, and aspect ratio of connected components
11. Variance Distance
12. Euler Number: Number of close contour Number of hole inside
13. Moments: Spatial moments, Central moments, HU invariant moments, third central moment measure the skew and forth central moment measure the kurtosis that describe the flatness
14. Angle of inclination with respect to major axis
15. Size of the object
16. Polygon Approximation
17. Gray level co-occurrence matrix, multi channel Gabor filter based features
18. Local energy feature, distribution of local energy;

Structural Feature Set:

1. Character Optical Density
2. Vertical distribution of the upward concavity distribution
3. Variances of the upward concavity distribution
4. Character height distribution
5. Upper and lower profile of the character bounding boxes
6. Distribution of Bounding box size and Characters density, Horizontal projections
7. Height distribution of connected component
8. Enclosing structure of connected component
9. Characters complexity, stroke, presence of circle, ellipse, vertical stroke, horizontal stroke, left and right diagonal stroke
10. Morphological opening and erosion in all directions (horizontal, vertical, right and left diagonal) using different types of structuring elements
11. Fourier coefficient of the horizontal projection profile,
12. Connected component based: mean and standard deviation of the outputs for a n-channel Gabor filter, Distribution of the width to height ratio
13. Aggregate distance of the pixels in the top most and bottom most profiles of the connected components of a character / image
14. Projection profile,
15. Shirorekha based features,
16. Busy zone based: distribution of the lower most points of the characters. For example in English it lies along the base line and bottom line. In Urdu it is randomly distributed
17. Water reservoir based feature for measure of the cavity region of the character
18. Jump discontinuity based feature: it refers to the maximum horizontal distance between two consecutive border pixels in a character pattern
19. Distribution of peaks and moments in horizontal projection profile
20. Histogram statistics: describe spatial distribution of gray level
21. Zone based feature calculation: Method of dividing the image into several zones for analyzing the distribution of the pixels in each zone
22. Contour, Convexity, Convexity defects
23. Fractal based features
24. Steerable pyramid based features

Script Dependent Feature Set:

1. Textual symbol: character fragments, discrete characters, adjoined characters, whole word,
2. Number of small, medium, and large components,
3. Number of dots,
4. Script specific principal stroke: horizontal, vertical, diagonal stroke,
5. Inclination feature of characters,
6. Tick feature: presence of tick like structure in characters (telakattu in tamil).

Main Classifier used for this work is MLP, Neural Network and Languages used for testing are Bangla, Devnagari, Malayalam, Urdu, Oriya and Roman scripts with accuracy of 92.8%. Authors used data set for this work from collected data by heterogeneous sources like University, Post office etc. From outside state they brought some data by their friends and different connections. They used a total of 152 documents written in Bangla, Devnagari, Malayalam, Urdu, Oriya and Roman script. Along with handwritten documents, authors also evaluated printed documents for printed environment. Originally the images are in gray tone and digitized at 300 dpi. A two stage based approach is used to convert the im-

ages into two-tone (0 and 1). Pre-binarization process is done in the first stage using a local window based algorithm in order to get an idea of different regions of interest. Run Length Smoothing Approach (RLSA) is applied on pre-binarized image, to overcome the limitations of the local binarized. Using component labeling, they select each component to map in the original gray image to get respective zones of the original image. Final binarized image is obtained using histogram based global binarization algorithm on these regions of the original image [1].

Mallikarjun Hangarge et.al, features of this work is Directional edge information and classifiers used mainly are KNN, D-DCT, linear discriminant analysis (LDA). A dataset of 9000 handwritten text words of six scripts languages namely Roman, Devanagari, Kannada, Telugu, Tamil and Malayalam were used for testing with an accuracy of 96. Dataset is created by collection of selected a paragraph of printed text from the various documents of History, Medical, Arts, Science and Religious. So they have collection of more number of unlike words in dataset as it is a collection from different sector / interest. They distributed the text so selected among 20 writers of various age groups and professions. Each script was written by these writers were digitized using a scanner with 300 dpi resolution[2].

Deepti Khanduja, et.al, mainly worked on Structural features of the character with Neural Network over Devanagari script. Authors received 93% accuracy over Thinning, zoning, raster scan, curve fitting [3].

K. Roy et.al, worked on features of Fractal dimension, Component Based Feature, circularity of the component, Freeman chain code using MLP classifier over handwritten Bangla, Devnagari, Malayalam, Urdu, Oriya, Roman script recived 89.4% accuracy over self generated dataset [4].

Komal et.al discussed Several techniques SVM, MLRMLL, SVRFMCetc Accordingly, plan of implementation process was finalized to spatial contextual, integrate spectral and spatial location cues of images with high resolution using conditional random fields (CRFSS) [5].

Hiremath P. S et.al, worked for Texture features using co-occurrence histogram. Connected component analysis (CCA) does not require as they adopted texture based approach using KNN classifier over languages like Tamil, Urdu, Kannada, Telugu, Hindi, Bengali, Malayalam. Digitized script documents were taken for experiments over four thousand images. Accuracy decreases as writer increases. (For one writer:97.5%, two writers: 91.8%, three writers: 79.75%) [6].

G. G. Rajput et.al, feature vector computed using DCT and DWT using KNN classifier over languages Kannada, English, Hindi, Malayalam, Punjabi, Tamil, Telugu, Gujarathi, Hindi. Documents of various languages are scanned with 300 DPI with accuracy of 96.4% [7].

Mallikarjun Hangarge, et.al, Thirteen spatial features were extracted using morphological filters, (segmentation: Devnagari: sirorekha; Roman: vertical strokes, left right diagonal stroke; Urdu: strong baseline, right and left diagonal stroke; stroke density: vertical and horizontal stroke density), pixel density of an image after fill holes using KNN classifier over languages Devnagari, English, Urdu. Data collection: Sample of 150 writers from various professionals, Colleges and schools were selected. They were asked to write 10 lines of text matter in Devnagari, English and Urdu scripts. A total of 300 handwritten document images were created from these writers. The collected documents were scanned at 300 DPI. Accuracy varies from language to language. English 86.00%, Devnagari 83.00%, Urdu 97.00% [8]

Sukalpa Chanda, et.al, Authors worked in two stages. In the first stage, they considered sixty four dimensional chain-code-histogram feature and in second stage of their work, four hundred dimensional gradient features were used with SVM classifier over languages English (Roman), Bengali (Bangla) and Devnagari. The digitized images (300 DPI) were in gray tone and were converted to two-tone images (0 ,1 ) using histogram-based thresholding approach. Considerably good accuracy of 98.5% on more than eleven thousand test words was achieved [9].

K. Roy, et.al Fractal-based features, Busy-zone based features and Topological features were considered using Neural Network over languages Bangla, English and Devnagari to achive 96.97% accuracy. Documents of Postal Letters written in Bangla, English, Devanagari were used [10].

Ramesh M. Kagalkar et.al, Authors implemented gradient based key frame extraction technique to solve the problem. These key frames were useful for splitting continuous gestures into sequence of signs by removing uninformative frames. Experiments were performed on their own continuous ISL dataset which is created using EOS camera in PG Research Laboratory (SPPU, Pune). Probes are tested exploitation varied forms of classifiers like, City block distance, Euclidian distance, Manhattan distance, Correlation, Manhattan distance etc. [11]

Sandhya Arora et.al, "Combining Multiple Feature Extraction Techniques for Handwritten Devnagari Character Recognition", They used MLP classifier over Handwritten Devnagari characters to achieve 92.8% accuracy using straight line fitting features, shadow feature, intersection and chain code histogram [12].

B.V.Dhandra, et.al, considered seven global features along with other three dominant local features were computed based on connected components, large number of text word images and handwritten alphanumeric were used with KNN classifier over Kannada, Roman and Devnagari scripts and they got variable accuracy Kannada 93.1%, Hindi 92.5%, English 89.5% [13].

Komal V. Rayate et.al, Hyper-spectral images display strong dependencies across spatial and spectral neighbours, which were very helpful for hyper-spectral image classification. Images with high- resolution have the type of plentiful geometric and detail information, which were useful for a detailed classification. A proposed classification algorithm is based on conditional random fields (CRFs), to integrate spectral, spatial contextual and spatial location cues by modeling the probabilistic potentials. An average accuracy of 98% was obtained [14].

Gopal Datt Joshi, et.al, First level features (Coarse features): The energy profile, E(θ), Second-level features (Finer features): Ratios of normalized energies KNN and Parzenn for 10 Indian (Devanagari, , Kannada, Tamil, Gurumukhi, Malayalam, Bangla Roman, English, Oriya, Gujarati, Urdu) and 13 world languages:(Ethiopic, Arabic, Cyrillic, Korean, Chinese, Hebrew, Devanagari, Japanese, Armenian, Thai, Latin, Greek, Burmese). The scanned images were acquired from heterogeneous sources like books, magazines, newspapers and computer printouts. The images and a summary table are made available by the authors at http://www.c3.lanl.gov/kelly/LIFI/ and they achieved 94% accuracy [15].

B.V.Dhandra, et.al, Global and local features were extracted by morphological filters and regional descriptors of three major Indian languages/scripts: Kannada, Roman and Devnagari. They considered global features like vertical and horizontal Stroke Density, four different features like top-hat and bottom-hat transformations. Local Features like aspect ratio, eccentricity, extent etc were considered. KNN classifier was applied with different ranges of K (K= 3, 5, 7, 9) for languages Kannada, Roman and Devnagari. Experiment was conducted by selecting unconstrained handwritten text documents around two hundred and fifty writers of different ages and fields. The collected documents were scanned at 300 DPI, so they got low noise, good quality document image. The gray tone digitized images were converted to two-tone images into 0-1 labels using Otsu's global thresholding approach achieved accuracy of 96.05% [16].

K.-Roy, et.al, authors considered features of Shirorekha / Matra, Water reservoir and Presence of small component using Tree classifier over postal letters written in Bangla, Devnagari, English and received 88.58% of accuracy [17].

Jinjie Huang, et.al, features selection methods are generally categorized into filter approach, wrapper approach and embedded method. In this paper, authors suggested hybrid genetic algorithm

for feature selection using wrapper approach. The outer and inner optimization stages are proposed for feature selection. The outer stage completes global search. The inner optimization performs local search which will be an independent measure for feature ranking using SVM, HGA-p (Wrapper framework for feature selection). Data set is available at ics & microarray website[18].
Rong-En Fan, et.al, described Study of tools for classification using SVM, Linear Classifier. Dataset is available at csie website[19].
Kekre H B, et.al, Image Halftoning is applied over Handwritten Document images. Image Halftoning technique yields best compression factor of 12.5%. This compressed image is recompressed with RLE to obtain excellent compressed image data. [20].
Borse S B, et.al, Zoning is able to handle variation in handwritten pattern, which is the result of users' different physical and psychological conditions. Verona diagram is used for optimal zoning using GA, Voronoi tessellation over Handwritten Document Images [21].

## 4. Discussion and Future work

In the process of handwritten Devanagari script recognition, the preprocessing phase dominates further stages of recognition in that precisely cleaned image will lead to produce high level of recognition accuracy. Feature extraction stage leads to extract unique features of the text written by writer where accurately extracting desired feature that must be capable to rebuild / recognize the original character. The set of local and global features extracted will help to recognize the script properly when passed through recognition phase i.e. classification phase. Before being sent to classification phase, out of these extracted features, only selected features are useful to classify the source text. The different classification techniques are described with varying recognition accuracy depending on different features. Researchers can choose appropriate technique to proceed for recognition.

## 5. Conclusion

This paper mostly focuses on the study of different features, feature selection and different classification or recognition techniques to recognize handwritten Devanagari script. This study may help researchers and developers to understand various techniques which were implemented for recognition.

## References


[1] Sk Md Obaidullah, Supratik Kundu Das , Kaushik Roy , "A System for Handwritten Script Identification from Indian Document", Researchgate: Journal of Pattern Recognition Research 8 (2013) 1-12, Researchgate, 2013, pp.1-12.

[2] Mallikarjun Hangarge, Santosh K.C., Rajmohan Pardeshi, "Directional Discrete Cosine Transform for Handwritten Script Identification", 12th International Conference on Document Analysis and Recognition (ICDAR), 2013, IEEE, 2013, pp.344-348.

[3] DEEPTI KHANDUJA, NEETA NAIN, and SUBHASH PANWAR, "A Hybrid Feature Extraction Algorithm for Devanagari Script", ACM Trans. Asian Low-Resour. Lang. Inf. Process., Vol. 15, No. 1, Article 2, ACM, 2015, pp.1-10.

[4] K. Roy, S. Kundu Das, Sk Md Obaidullah, "Script Identification from Handwritten Document", Third National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics, 2011, Researchgate, 2011, pp.1-5.

[5] Komal V. Rayate, Shyamrao V. Gumaste, "Image Classification-Review" International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Volume 5, Issue 10, October 2016

[6] Hiremath P. S., Shivashankar S., Jagdeesh D. Pujari, V. Mouneswara, "Script identification in a handwritten document image using texture features ", IEEE Explore, IEEE, 2010, pp.110-114.

[7] G. G. Rajput, Anita H. B, "Handwritten Script Recognition using DCT and Wavelet Features at Block Level", RTIPPR 2010, IJCA, 2010, pp.158-163.

[8] Mallikarjun Hangarge, B.V.Dhandra, "Offline handwritten script identification in document images", IJCA 2010 Vol.4,No.6, IJCA, 2010, pp.6-10.

[9] Sukalpa Chanda, Srikanta Pal, Katrin Franke, Umapada Pal, "Two-stage Approach for Word-wise Script Identification", 2009 10th ICDAR, IEEE, 2009, pp.926-930.

[10] K. Roy, K. Majumder, "Trilingual Script Separation of Handwritten Postal Document", Sixth Indian Conference on Computer Vision, Graphics & Image Processing, 2008. ICVGIP '08. , IEEE, 2008, pp.693-700.

[11] Ramesh M. Kagalkar, S.V Gumaste, "Gradient Based Key Frame Extraction for Continuous Indian Sign Language Gesture Recognition and Sentence Formation in Kannada Language: A Comparative Study of Classifiers", International Journal of Computer Sciences and Engineering, Vol.-4(9), Sep 2016

[12] Sandhya Arora, Debotosh Bhattacharjee, Mita Nasipuri, Dipak Kumar Basu, Mahantapas Kundu , "Combining Multiple Feature Extraction Techniques for Handwritten Devnagari Character Recognition", 2008 IEEE Region 10 Colloquium and the Third ICIIS, Kharagpur, INDIA December 8-10., IEEE, 2008, pp.342-1-342-6.

[13] B.V.Dhandra, Mallikarjun Hangarge, "Global and Local Features Based Handwritten Text Words and Numerals Script Identification", International Conference on Computational Intelligence and Multimedia Applications 2007, IEEE, 2007, pp.471-475.

[14] Komal V. Rayate, Shyamrao V. Gumaste "Classification of High Resolution Images with Different Cues by using CRF Model", Sandip Foundation's International Journal on Emerging Trends in Technology (IJETT), Volume 4, issue 1 April 2017

[15] Gopal Datt Joshi, Saurabh Garg, Jayanthi Sivaswamy, "A generalised framework for script identification", IJDAR (2007), Springer, 2007, pp.55-68.

[16] B.V.Dhandra, Mallikarjun Hangarge, "Global and Local Features Based Handwritten Text Words and Numerals Script Identification", International Conference on Computational Intelligence and Multimedia Applications 2007, IEEE, 2007, pp.471-475.

[17] K.-Roy, A. Banerjee and U. Pal, "A System for Word-wise Handwritten Script Identification for Indian Postal Automation", IEEE INDIA ANNUAL CONFERENCE 2004, INDICON 2004, IEEE, 2004, pp.266-271.

[18] Jinjie Huang, Yunze Cai, Xiaoming Xu, "A hybrid genetic algorithm for feature selection wrapper based on mutual information", Pattern Recognition Letters 28 (2007), Elsevier, 2007, pp.1825-1844.

[19] Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, Chih-Jen Lin, "LIBLINEAR: A Library for Large Linear Classification", Journal of Machine Learning Research 9 (2008), JMLR, 2008, pp.1871-1874.

[20] Kekre H B, Kharat M U, Sange S R, "Image data compression using new Halftoning operators and Run Length Encoding", Thinkquest~2010, Springer India, 2011, pp.208-213.

[21] Borse S B, Bhalekar M, Kharat M U, "HANDWRITTEN CHARACTER RECOGNITION WITH OPTIMAL ZONING USING GA".