



Journal Homepage: -[www.journalijar.com](http://www.journalijar.com)  
**INTERNATIONAL JOURNAL OF  
 ADVANCED RESEARCH (IJAR)**

Article DOI:10.21474/IJAR01/5117  
 DOI URL: <http://dx.doi.org/10.21474/IJAR01/5117>



### RESEARCH ARTICLE

#### A NOVEL APPROACH TOWARDS ONLINE DEVNAGARI HANDWRITTEN WORD RECOGNITION BASED ON ROBUST FEATURE EXTRACTION METHOD AND FFNN CLASSIFIER.

Saniya Ansari<sup>1</sup>, Dr. Bhavani S<sup>2</sup> and Dr. Udaysingh Sutar<sup>3</sup>.

1. Research Scholar, Dept. of ECE, Karpagam Academy of Higher Education, Karpagam University, Coimbatore, Tamilnadu, State, India.
2. Professor&HoD, Dept Of ECE, Karpagam Academy of Higher Education, Karpagam University, Coimbatore, Tamilnadu, State, India.
3. Professor&HoD, Dept. Of E&TC, AISSMSCOE, Pune, Maharashtra State, India.

#### Manuscript Info

##### Manuscript History

Received: 09 June 2017  
 Final Accepted: 11 July 2017  
 Published: August 2017

##### Key words:-

Handwritten Recognition, Devnagari Script, Feature extraction, FFNN, SVM, KNN, Recall, Precision & accuracy

#### Abstract

In recent time the use of handwritten recognition tools are increasing in various fields of applications, therefore it becomes interesting area of research to many researchers. Different methods are presented with goal of more accuracy for character recognition with less time for different languages like English [Major], Hindi, Telagu, Gurumukhi, Kannada etc. Devnagari is also one of the most widely used scripts in India. This paper is focused to present accurate, robust and faster method for Devnagari handwritten recognition for Marathi language using efficient algorithms in each phase of handwritten recognition. There are three main phases of any handwritten character recognition methods such as Segmentation, Feature extraction and classification. In first phase, input image is preprocessed using Gaussian filter for smoothing and noise removal. Further using thresholding preprocessed image is segmented with additional morphological operations such as dilation, filling, erosion in order to get finalized segmented image. In second phase, hybrid efficient and optimized feature vector of length 91 is presented by using combination of geometrical features, regional features, distance transform and gradient features. In third phase, use of efficient, accurate classifier called Feed Forward Neural Network [FFNN] for online handwritten character recognition is presented. Also the performance of both SVM & K-NN classifiers is computed for same database. Here 200 commonly used handwritten words are collected from 50 users with different handwriting styles to create database of 10,000 words. For experimentation, 7500 word samples are used to create 20 dataset out of which 70% samples are used for training, 20% for testing & 10% for validation. Overall recognition accuracy obtained using SVM, K-NN & FFNN classifiers is 84.70%, 82.30% & 94.57% respectively.

Copy Right, IJAR, 2017,. All rights reserved.

#### Corresponding Author:-Saniya Ansari.

Address:-Research Scholar, Dept. of ECE, Karpagam Academy of Higher Education, Karpagam University, Coimbatore, Tamilnadu, State, India.

## Introduction:

Since from last two decades, various research studies are presented based on character recognition over English Language as it is universal language. However in India, there are several languages which are used to communicate in day to day life rather than English. Devnagari script has the most accurate scientific basis and it is used by Sanskrit, Hindi, Marathi and Nepali languages. Marathi is the widely spoken language in Maharashtra and it is well known official language and used as co-official language in Maharashtra and Goa states of Western India respectively. There is about more than 500 million people uses this language for day to day life communication. For Indian languages, so far any effective method for character recognition is not available as very less research work reported towards the problem. Other than English language, different research methods for the languages like Chinese, Arabic, Roman, and Japanese for handwritten character recognition are exists. The complexity of Devnagari script is more as compared to English due to variations in writing styles of different characters which are composed of various order, direction, shape, strokes etc. Like English, Devnagari script also consists of basic set of symbols of 34 consonants and 18 vowels those can be used to construct words. Therefore, the base of presented work is to target research of online handwritten word recognition for Devnagari script especially for Marathi language.

The recognition of handwritten Marathi words comes under the domain of document image analysis which deals with automatic reading of information from the input word image. This task is done by OCR which is called as Optical Character Recognition which is nothing but technique of reading scanned text document by machine. The existing OCR has implicit assumption that type of script that has to be processed is aware before processing. However, in automated applications and environment, this type of document processing techniques depends on human intervention in order to choose particular OCR type, and this becomes very inefficient, impractical and undesirable for end users. If the document itself contains different types of languages, then document analysis as well as recognition becomes more challenging and complex, as OCR needs to select any one type of language before processing the document. To overcome this problem, in recent years, few methods for automatic character recognition are presented. Online or offline handwritten analysis is one of next research domain of OCR. Handwritten character recognition is most interesting research domain since from last decade in pattern recognition. OCR can be used in applications such as digit recognition, bank check processing, vehicle plate recognition, postal address block detection as well as their recognition etc. For the OCR applications, accuracy and speed are most vital parameters to decide the efficiency of recognition system which heavily depends on feature extraction phase and methods used for it.

This paper presents the improved framework for automatic online Devnagari handwritten word recognition. The process of handwritten word recognition is broadly classified into four major parts such as 1) Preprocessing, 2) Segmentation, 3) Feature Extraction, 4) Classification and Recognition. The first important step of any image processing domain is preprocessing in which raw, noisy and varying sized images are processed to get smooth, noise free and fixed sized images for further research processing. Preprocessing helps next functions to work efficiently. In presented approach Gaussian filtering is used for removing noise and smoothing image with a 512\*512 standard size. After that segmentation is performed to extract the real information from input image in order to do further processing.

The third step of handwritten recognition is feature extraction. The selection of feature extraction technique is main factor for delivering highest handwritten recognition accuracy. The various methods for feature extractions are presented in literature like Gradient features, structural features, regional features, projection histograms, Zernike moments, zoning etc. Many methods requires more time for feature extraction delivering better accuracy. But in handwritten character recognition, both high accuracy and less time are desirable & preferred. In this paper, the new hybrid method of feature extraction is presented in which total 91 features are extracted and further used for recognition. These features are combination of structural or geometric features, regional feature, gradient features and distance transform features. The structural features are optimized by using Universe of discourse over input segmented image, which can speed up the tasks of 91 extracted features.

The fourth and final phase for online handwritten character recognition is classification. Various classifiers are presented in literature for recognition such as SVM, KNN, ANN statistical, structural and neural networks. Recently, significant contributions towards the improvement of recognition rates have been made by means of different combination strategies. To the best of our knowledge there are only a few research reports available on Devnagari offline handwritten character, word and numeral recognition. This paper presents the use of SVM, K-NN

& FFNN classifier for classification and handwritten word recognition for Marathi language. The performance metrics used for comparative analysis are false positive rate, false negative rate, f-score, accuracy, precision rate, recall rate etc.

In this paper, section II gives the details information about various Devnagari character recognition methods, related segmentation methods, different feature extraction techniques and various approaches adopted for recognition using classifiers are discussed. Section III, gives brief information about various issues related with online handwritten character and word recognition, Section IV introduces all algorithms and designs involved in complete proposed framework. Section V gives detailed results and comparative analysis. Section VI discusses conclusion and future work.

### **Related Works:**

In this section, different methods reported previously in literature by different authors are summarized based on main categories such as segmentation, feature extraction and classification method used.

### **Review of Devnagari Character Recognition Methods**

Anoop Namboodiri presented a method to classify words and a line in an online handwritten document based on 11 different spatial and temporal features extracted from the strokes of the words. In an online handwritten document words and lines are classified into one of the six major scripts like Arabic, Cyrillic, Devnagari, Han, Hebrew, or Roman. The proposed system achieves an overall classification accuracy of 87.1% for a data set of 13,379 words [1]. M. Hanmandlu et al., presented Fuzzy logic based method for recognition of handwritten Hindi numerals and characters with 92.67% and 90.65% over all accuracy for Handwritten Devnagari numerals and characters respectively [2]. Satish Kumar et.al.author presented Zenrike moment feature based method for recognition of handwritten Devnagari character using artificial neural network for classification purpose[3]. N.Sharma et.al, presented handwritten Devnagari characters recognition with the directional chain code features extraction and quadratic classifier which yields 80.36 % of overall recognition accuracy. Sharma et al. used the five preprocessing stages in sequential order i.e. size normalization and centering, interpolating missing points, smoothing, slant correction and resampling of points. The stroke written with high speed will have missing points which can be calculated using Bezier interpolation [4]. Prachi Mukherji et.al proposed method using basic structural features like endpoint, cross point, junction points. The segments of characters are coded using Average Compressed Direction Code algorithm. With top modifier 71.68 % and without top modifier 88.33% accuracy is achieved [5].

Sandhya Arora et.al presented a scheme for online Handwritten Devnagari Character Recognition, which uses different feature extraction and recognition algorithms. After preprocessing, Chain code histogram, four side views, shadow based features are extracted and fed to MLP. The proposed system was tested on 1500 samples yielding 98.16% and 89.58% recognition rates for top 5 and top 1 result respectively [6]-[7]. Sushama Shelke et.al elaborated a novel approach for recognition of handwritten Marathi compound characters using a multi-stage multi-feature classifier. The various features like pixel density features, Euclidean distance features and modified wavelet approximation features are extracted and then applied to three different neural networks which yielded recognition accuracy of 97.95% [8]. Mitrakshi B. Patil et.al developed method for recognition of offline handwritten Devnagari characters using artificial neural networks. The input characters are represented as N-vector and given as input to neural network. The size of the hidden layer selected for first and second layer is 8 and 16. The output values are represented using Binary vectors of size 4 with the learning rates ranging from 0.1 to 0.4 [9].

Tanuja K. et.al, implemented a novel technique for Handwritten Hindi Character Recognition System using Canny Edge Detection technique and artificial neural network. This approach provides 95% Accuracy with Feedback propagation Neural Network algorithm for single character. [10]. Ved Agnihotri et.al proposed a new method of classification by extracting diagonal features from zones of an image using neural network for Handwritten Devnagari script recognition system. Here the features of each character image are converted into chromosome bit string with length 378. Genetic Algorithm is used for the recognition and classification, which yields precision as 85.78% match and 13.35% mismatch [11]. Anilkumar N Holambe, presented combining statistical, structural Global transformation and moments features to form hybrid feature vector. The combination of SVM & KNN algorithm gives highest accuracy of 96% [12]. Muhammad et al., presented a review about HCR in general and specifically concentrated on feature extraction and selection and investigated metaheuristic search algorithm as an optimization tool [13]. Vijaya Pawar proposed an artificial neural network based classifier and statistical and structural method based feature extraction. Features are extracted in terms of various structural and statistical features like End points,

middle bar, loop, end bar, aspect ratio etc. Feature vector is applied to Self organizing map (SOM) which attains 95% accuracy [14].

### **Review of Different Preprocessing & Segmentation Techniques:**

Satish Kumar presented method for recognition of handwritten Devnagari compound character using Zenrike moment feature extraction and artificial neural network for classification. The proposed classification system preprocess and normalized the handwritten character images into 30x30 pixels images and divides them into zones. The pre-classification produces three classes depending on presence or absence of vertical bar [3]. Aparna et al. proposed a preprocessing technique that consists of interpolation, smoothing and normalization of strokes. The strokes are then converted onto curve length base and then smoothed independently along t-axis using a Gaussian filter [15]. Huang et al. proposed a new preprocessing technique for online handwriting recognition which first removes the hooks of the strokes by using changed-angle threshold with length threshold, and then filtered the noise by using a smoothing technique. The basic steps used for preprocessing are Removing duplicated Points, Elimination of Hooks, Interpolating points, Detection of sharp points, Removing Hooks, Smoothing Data, and Normalization [16].

Bharath and Madhvanath gives relevance of stroke size and position information for the recognition by comparing three different preprocessing schemes. These schemes are Word-level preprocessing retaining the vertical positions of the strokes completely, Word level preprocessing followed by stroke level preprocessing retaining the coarse vertical positions of the strokes and Stroke-level preprocessing ignoring the positions of the strokes [17]. Zhao et al. uses four preprocessing techniques like De-hooking, smoothing, size normalization, Resampling [18]. N. Anupama, Ch. Rupa proposed an algorithm based on multiple histogram projections using morphological operators to extract features of the image. Horizontal projection is performed on the text image, and then line segments are identified by the peaks in the horizontal projection. Vertical histogram projections are used for the line segments and decomposed into words using threshold and further decomposed to characters [19].

### **Review of Feature Extraction Methods:**

Satish Kumar presented Zenrike moment feature based method for recognition of handwritten devnagari compound character using artificial neural network for classification purpose [3]. N. Sharma presented the directional chain code features with quadratic classifier based on the contour code and stroke direction. The contour code feature utilizes the rate of change of slope along the contour profile in addition to other properties such as the ascender and descender count, start point and end point [4]. Sandhya Arora, presented an OCR for Handwritten Devnagari Characters using neural classifier. They used four feature extraction techniques namely, intersection, shadow feature, chain code histogram and straight line fitting features[6]. Anilkumar Holambe, presented combination of statistical, structural Global transformation and moments features to form hybrid feature vector. The Classifiers are combined to achieve maximum accuracy for Devnagari Script. [12].

U. Pal et.al did a comparative study of four sets of different feature extracting methods and 12 different classifiers for handwritten character recognition. Projection distance, linear differentiates function, subspace method, modified quadratic discriminate function, support vector machine, Euclidian distance, image learning, nearest neighbour, modifies projection distance, compound projection distance and compound revised quadratic discriminate function were used as different classifiers[20]. Bikash Shaw et al. presented an offline handwritten Devnagari word recognition: a segmentation based approach for recognition handwritten Devnagari words. Stroke based features are used as feature vectors and hidden Markov model is used for recognition [21]. J. Pradeep, presented a handwritten character recognition system using multilayer Feed forward neural network. Three different orientations, namely, horizontal, vertical and diagonal directions are used to extract 54 features from each character. The diagonal orientation for feature extraction is identified to be the most suitable method as it yields higher recognition accuracy [22].

Brijmohan Singh, used two different methods for extracting features from handwritten Devnagari characters, the Curvelet Transform and the Character Geometry, and compares their recognition performances using two different classifiers, viz., the Support Vector Machine with Radial Basis Function and the k-Nearest Neighbour classifier[23]. Mahesh Jangid have proposed the method for feature extraction like Zonal density, Projection histogram, Distance Profiles, Background Directional Distribution and SVM for classification and they have got 98%, 99.1% and 99.2% of accuracy [24]. Gita Sinha, presented an overview of Feature Extraction techniques for off-line recognition of isolated Devnagari numeral recognition. Zone based approach presents the combination of

image centroid zone and centroid zone of numeral/character image to obtain 200 Feature Vector from both methods [25]. Pratibha Singh calculated the features based on three different zoning methods. Directional feature is obtained using chain code and gradient direction quantization of the orientations [26]. Rajiv Kumar presented offline handwritten character recognition for Devnagari. The evaluated feature extraction methods includes, direct pixel, image zoning, wavelet transformation and Gaussian image transformation techniques. These features were classified by using KNN and neural network classifier. [27].

### **Issues In Online Devnagari Handwritten Word Recognition System:**

In Handwriting recognition, the major problem is the huge variations in the handwriting styles of various writers or within the handwriting style of the same writer. Variations may occur due to several reasons extending from personal to material factors. Thus the performance of the handwriting recognition system depends heavily on various methods adopted to train these variations. A good recognition system should be able to recognize different but similar looking characters. Various issues are summarized in below section:

#### **Variations in Handwriting Styles**

A variation in handwriting style mostly occurs due to various writers or due to same writer. For handwriting capturing, input devices used gives the information about shape, size, order of stroke and speed of handwriting. Due to various handwriting styles the characters may look similar even though number of strokes, drawing order and direction of strokes may vary significantly.

#### **Personal or Background Factors**

The direction and position in handwriting may affect due to left-handed or right-handed writing, age and health. Rather than this, education, origin and profession of a person also lead to variation in handwriting.

#### **Material Factors**

Material factors are related with the hardware devices used for writing. Due to uncomfortable and inconvenient hardware used for writing may produce variation in writing. These factors may be type, position and size of the writing board.

#### **Similarity in Shape of Some Characters**

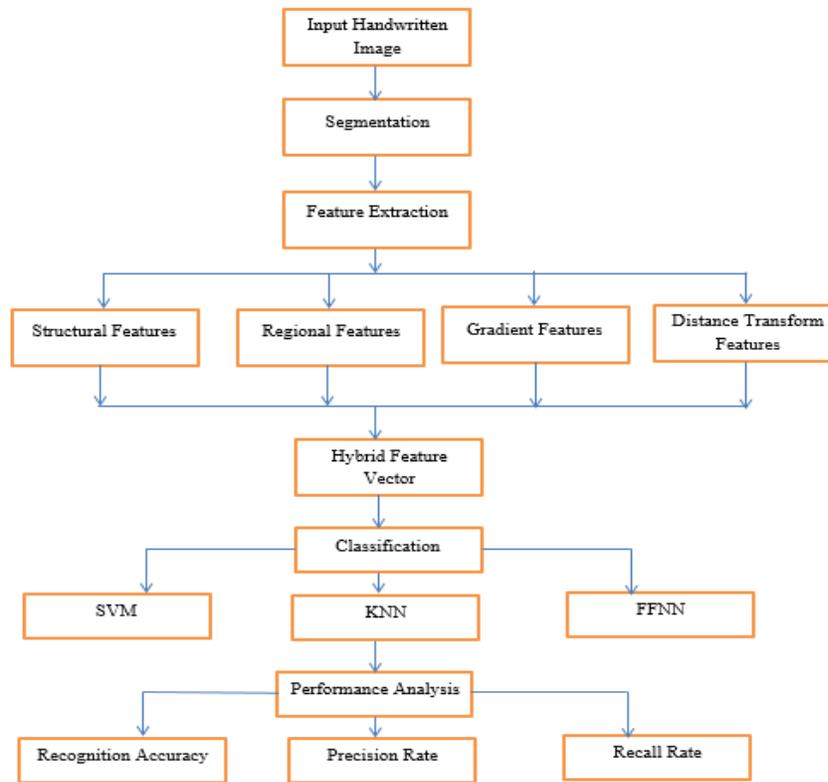
There are several characters in various scripts which have almost the similar shape. Due to similar shape of character, it becomes difficult to recognize them accurately.

#### **Presence of vertical appendages of modifiers**

The horizontal and vertical extent of characters gets affected due to existence of consonant and vowel modifiers. Due to this a significant variation is observed in height in Devnagari characters which have vertical appendages of diacritic strokes.

### **Proposed Methodology:**

This section describes framework and algorithms used during each phase of proposed Devnagari word recognition for Marathi language. Figure-1 below shows the overall flow of system and then present's algorithms designed for each step.



**Figure 1:-** Architecture and flowchart of proposed system.

This paper is particularly focused on domain of handwritten word recognition for Devnagari script. There are different techniques presented for offline handwritten recognition, but very few are presented over online handwritten recognition. Online handwritten character recognition is dynamic and needs immediate and accurate recognition. The presented system is composed of four main phases such as Data Collection, Pre-processing & Segmentation, Feature Extraction and Classification. These phases are elaborated in following paragraph in detail.

In Data collection phase, the database for handwritten Marathi word is created with respect to various handwriting styles. In online recognition electronic tablets or digitizers are most commonly used devices. In the presented work, an android based i-ball Digital Tablet-4030 is used as an input device to prepare the database. Gesture class is used to capture gesture which user will draw on screen and then the gestures are transferred to Matlab tool for further processing. Preprocessing & Segmentation is the first vital step of any image processing and use of effective methods in preprocessing and segmentation defines the efficiency and accuracy of handwritten character recognition. Here first Gaussian filter is used on resized image and then applied thresholding method for binarization, these both approaches gives better outputs with less processing time. The image is segmented further using dilation, erosion and perimeter detection morphological operations. Characters from segmented image i.e. word samples are separated using vertical segmentation method and then horizontal segmentation method.

Feature extraction is a very important step as the success of a recognition system is always based on feature extraction method. The feature extractor determines which properties of the preprocessed data are most significant and useful in further phases. In presented research work, hybrid efficient and optimized feature vector is used which is combination of geometrical features, regional features, distance transform and gradient features. Total 91 features are extracted which is highest ever as compared to all existing methods for handwritten character recognition. In addition to this, in existing cases, the time required for extracting the geometrical features is very high; however Universe of discourse is used to speed up the retrieval.

Feed-forward neural network, is used which is the most commonly used family of neural networks for pattern classification tasks. The FFNN is nothing but biologically motivated approach of classification which is composed of large number of simple processing units organized in layers. Each unit in current layer is connected with all other

previous layer units. Every connection may have variable weight or strength; hence there is no possibility of similarity between all connections. Network knowledge is encoded into the weights on such connections. When FFNN acts as classifier, there is no feedback mechanism among layers. Therefore such classifier is known as feed forward neural network classifier. For learning three different objects together are needed to select: a FFNN (the *classifier*), a Pattern (the *inputs*) and Categories (the *correct outputs*). During the learning phase the weights in the FFNN will be modified. All weights are modified in such a way that when a pattern is presented, the output unit with the correct category, hopefully, will have the largest output value. The computing time for the learning phase depends on the size of the neural network, the number of patterns to be learned, the number of epochs, the tolerance of the minimizer and the speed of computer. FFNN is used with 91 numbers of input layers neurons, one hidden layer with 10 neurons and 10 neurons in output layer. The log sigmoid transfer function is used as activation function for hidden layer with maximum 1000 epochs.

For each phase, different methods and algorithms used are explained in detail in below section.

### **Pre-processing and Segmentation Algorithm:**

#### **Step 1: Preprocessing:**

- 1.1. Image acquisition: Browse input handwritten image
- 1.2. Image resizing: Resize input image to fix 512 \* 512 sizes.
- 1.3. Grayscale conversion: RGB image is required to be converted into grayscale.
- 1.4. Image Denoising and smoothing: Once image is resized and converted into grayscale format, further it is preprocessed for removal of noise and enhance its contrast by using Gaussian filter.
- 1.5. Final preprocessed image is generated as output of preprocessing.

#### **Step 2: Image Segmentation:**

- 2.1. Binarization: Grayscale preprocessed image is further segmented using thresholding method in which pixels those having intensity value less than 128 are kept as black and rest all kept as white pixels. Output of binarization is segmented binary image.
- 2.2. Edge Detection: Sobel Edge detection operator is used to detect the edges on binary image.
- 2.4. Dilation: The basic role of dilation morphological operator is that the value of the output pixel is the maximum value of all the pixels in the input pixel's neighborhood. In a binary Image, if any of the pixels is set to the value 1, the output pixel is set to 1. Dilation is applied on edge detected image by using two flat linear structuring elements with angle of 180 and 90.
- 2.5. Clear border: Used to suppress light structures connected to image border of dilated image.
- 2.6. Erosion: Used after dilation on image to erode the image using diamond-shaped structuring element.
- 2.7. Segmented Image: Finally segmented image is generated using finding perimeters objects from image.

#### **Step 3: Vertical Character Segmentation:**

- 3.1. Find white pixel from segmented image column wise
- 3.2. Check if column having number of white pixels less than or equal to 10, assign 0 value to entire column to make sure that it is represented as black, else keep as it is in output image.
- 3.3. Below For loop shows how it is done.
 

```

for i=1:c
    a = nnz (a2 (:,i)); // measure the number of white pixels from each column.
    if (a <= 10)
        opim (:,i) = 0;
    else
        opim (:,i) = a2 (:,i);
    end
end;

```

#### **Step 4: Horizontal Character Segmentation:**

- 4.1. Find white pixel from segmented image row wise
- 4.2. Check if row having white pixels and then start counting number of rows those having white pixel using count variable.
- 4.3. upper body segmentation  
If count == 10, then Insert two black rows after 10<sup>th</sup> row to represent upper part of word.
- 4.4. Lower body segmentation: Apply reverse process of above step 3 above on segmented image.

**4.2: Feature Extraction Algorithm:**

Input: Segmented Handwritten Image

Step 1: Extract Statistical/Geometric Features from Input and form feature vector GeF

Step 2: Extract Regional/Structural Features from Input and form feature vector ReF

Step 3: Extract Gradient Features: Gradient and Direction

Step 3.1: Apply mean and standard deviation on gradient

Step 3.2: Apply mean and standard deviation on direction

Step 3.3: Form final 4 features gradient vector called GrF

Step 4: Extract Distance Transform from Input.

Step 4.1: Apply mean and standard deviation on distance transform

Step 4.2: Form final 2 features distance transform vector called DiF.

Step 5: Combine GeF, ReF, GrF and DiF to form 91 feature vectors called CHF.

Output: Feature Vector CHF.

**4.2.1: Statistical Feature Extraction Algorithm:**

Input: Segmented handwritten image.

Step 1: Skeletonize input image.

Step 2: Apply Universe of discourse on skeletonized image.

Step 3: Zoning: An input image is divided into 9 equal size zones.

Step 4: From each zone, extract starters, intersections, and minor starters. Store them in one vector.

Step 5: line segments extraction from image and store them into one vector for each zone.

Step 6: line type detection from line segments such as horizontal, vertical, right diagonal and left diagonal etc.

Step 7: find total number of each line type.

Step 8: find normalized length of each line type.

Step 9: Final features collection: below are listed

- Number of horizontal lines.
- Number of vertical lines.
- Number of Right diagonal lines.
- Number of Left diagonal lines.
- Normalized Length of all horizontal lines.
- Normalized Length of all vertical lines.
- Normalized Length of all right diagonal lines.
- Normalized Length of all left diagonal lines.
- Normalized Area of the Skeleton

Step 10: Store all 81 features into vector GeF.

In above algorithm, Skeletonization and universe of discourse are used to reduce the features extraction time.

**4.2.2: Regional Feature Extraction Algorithm:**

Input: Segmented handwritten image.

Step 1: Euler Number Extraction

Step 2: Regional Area Extraction

Step 3: Eccentricity Extraction

Step 4: Orientation Extraction

Step 5: Formation of final feature vector called ReF using the four features i.e. Euler Number, Regional Area, Eccentricity and Orientation.

**4.3 Proposed FFNN Recognition Algorithm:****4.3.1: Training Algorithm:**

Step 1: Input feature matrix reading at layer 1  $F_i$ .

Step 2: Computation of activation value for every neuron  $[AN_i]$ .

Step 3: Search neuron with maximum  $AN_i$  value.

Step 4: Extract the step 3 results with its input\_id and max\_  $AN_i$  index.

Step 5: Output  $Ok$  is set to 1 for  $k$ th neuron who's having maximum  $AN_i$  value.

Step 6: Else set output to 0.

Step 7: Feed the input of previous layer to next layer still to the output layer.

Step 8: Repeat above steps for all input layers.

**4.3.2: Recognition Algorithm:**

Step 1: Read test pattern to be recognized or classified

Step 2: Compute activation value ANi during layer 2

Step 3: Select neuron with max ANi.

Step 4: Extract Neuron with max ANi index and save it as input\_id and max\_ANi\_index for purpose of matching.

Step5: If match is successful, then input\_id of max ANi is returned as output.

Step6: Stop

**Results and Discussion:**

**5.1. Dataset Preparation:**

For the research study, large dataset is prepared as standard database is not available. It contains total 200 Devnagari words, written by 50 different candidates using i-Ball digital tablet with different handwriting styles to create database of 10,000 words. Out of 10,000 handwritten word samples, 7500 samples are used to create 15 dataset. The training of 15 dataset is done using neural network for 91 input with one hidden layer of 10 neurons. Each Dataset consists of 10 words of 50 samples each. For experimentation, 7500 input samples are used out of which 70% samples are used for training, 20% for testing & 10% for validation. Based on this database, neural network is trained and performed classification using three different classifiers such as SVM, KNN and proposed FFNN to get three main performance metrics such as precision rate, recall rate and recognition accuracy.

**5.2. Performance Metrics Used:**

- True positive (TP) = correctly identified
  - False positive (FP) = incorrectly identified
  - True negative (TN) = correctly rejected
  - False negative (FN) = incorrectly rejected
- Precision Rate (%) =  $(TP / TP+FP) * 100$   
 Recall Rate (%) =  $(TP / TP+FN) * 100$   
 Accuracy (%) =  $(TP+TN / TP+TN+FN+FP) * 100$

**Results:**

Figure 2 below, shows the main GUI for handwritten Devnagari word recognition which shows the output result of all steps used with average performance of three classifiers. This figure shows that FFNN classifier outperforming the existing methods of classification on available dataset.

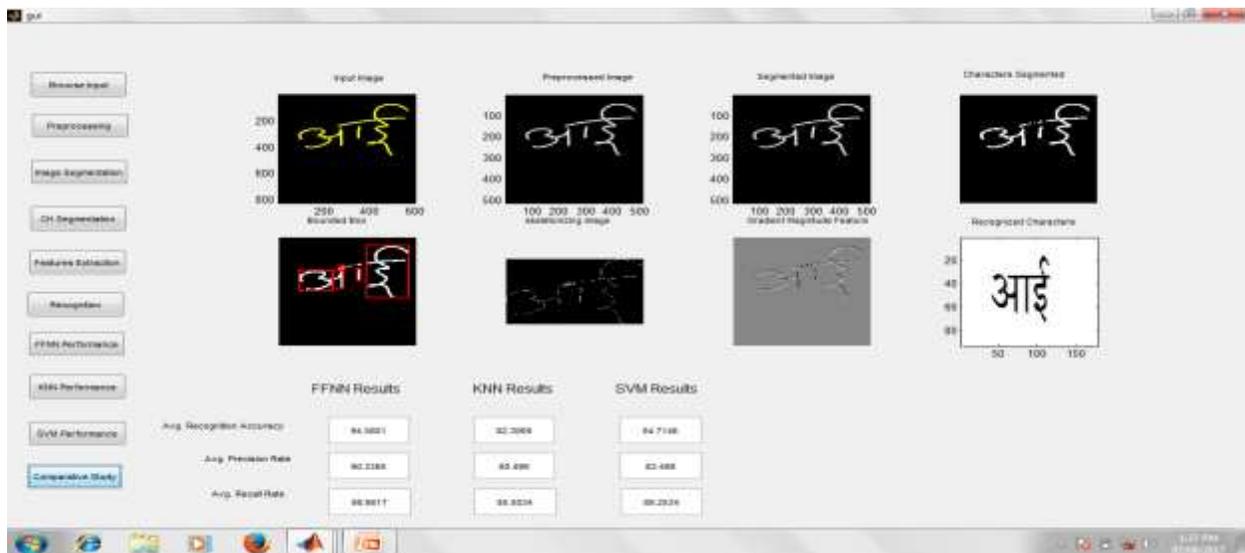


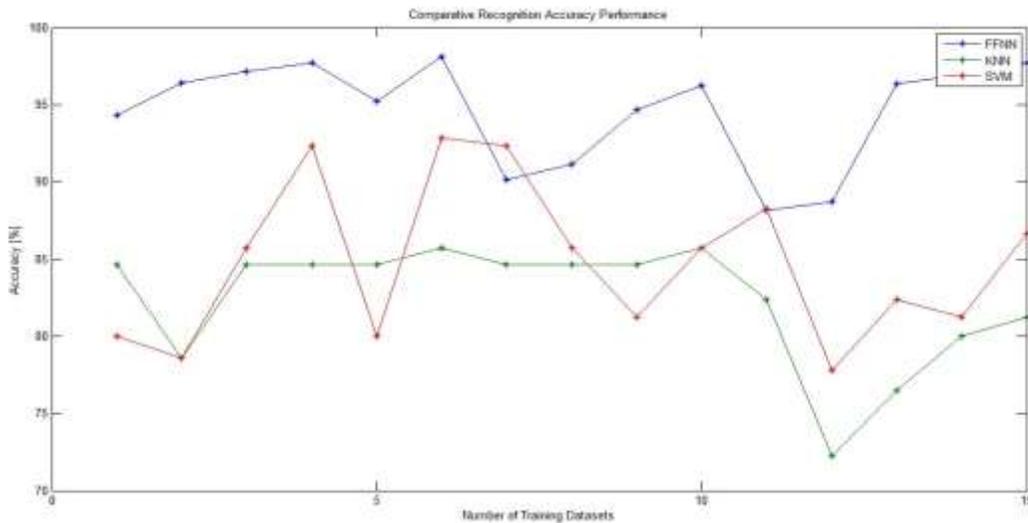
Fig 2:-Devnagari handwritten Word recognition using proposed approach

Any word sample that is to be recognized is preprocessed and extracted features from this word sample are sent to the Classifier. Three classifiers namely, FFNN, k-nearest neighbour (KNN) and SVM classifier are used to study the recognition accuracy on available database. The results of all the above classifiers are summarized in next section. The presented method performs well and appears promising compared to other methods in the literature. Table 1 below shows the output results & recognition accuracy achieved using FFNN, SVM & KNN classifier in terms of Precision, Recall and Accuracy.

**Table 1:-** Output results of FFNN, SVM & KNN classifiers

Dataset	FFNN			SVM			KNN		
	Precision	Recall	Accuracy	Precision	Recall	Accuracy	Precision	Recall	Accuracy
Dataset 1	89.76	98.69	94.28	78.50	93.75	80.00	78.57	81.25	84.61
Dataset 2	93.56	99.20	96.40	85.71	87.50	78.50	85.71	87.50	78.57
Dataset 3	94.81	99.39	97.11	80.00	87.50	85.71	73.33	81.25	84.61
Dataset 4	95.58	99.49	97.68	78.57	81.25	92.30	73.33	81.25	84.61
Dataset 5	91.49	98.82	95.20	86.66	93.75	80.00	91.66	81.25	84.61
Dataset 6	96.60	99.59	98.10	85.50	87.50	92.85	80.00	87.50	85.71
Dataset 7	82.43	97.34	90.09	78.57	81.25	92.30	73.33	81.25	84.61
Dataset 8	84.30	97.70	91.15	85.71	87.50	85.71	78.57	81.25	84.61
Dataset 9	90.51	98.66	94.64	93.33	100.0	81.25	73.33	81.25	84.61
Dataset 10	93.26	99.11	96.21	80.00	87.50	85.71	80.00	87.50	85.71
Dataset 11	78.89	96.77	88.13	83.33	89.47	88.23	83.33	89.47	82.35
Dataset 12	79.68	97.17	88.68	84.21	90.00	77.77	84.21	90.00	72.22
Dataset 13	93.56	99.12	96.39	83.33	89.47	82.35	83.33	89.47	76.47
Dataset 14	94.43	99.31	96.89	87.50	94.11	81.25	81.25	88.23	80.00
Dataset 15	95.88	99.50	97.70	81.25	88.23	86.66	87.50	94.11	81.25

Below graphs are showing the comparative study among proposed and existing methods of classification for precision, recall and recognition accuracy performances.



**Figure 3:-** Recognition accuracy performance varying number of datasets

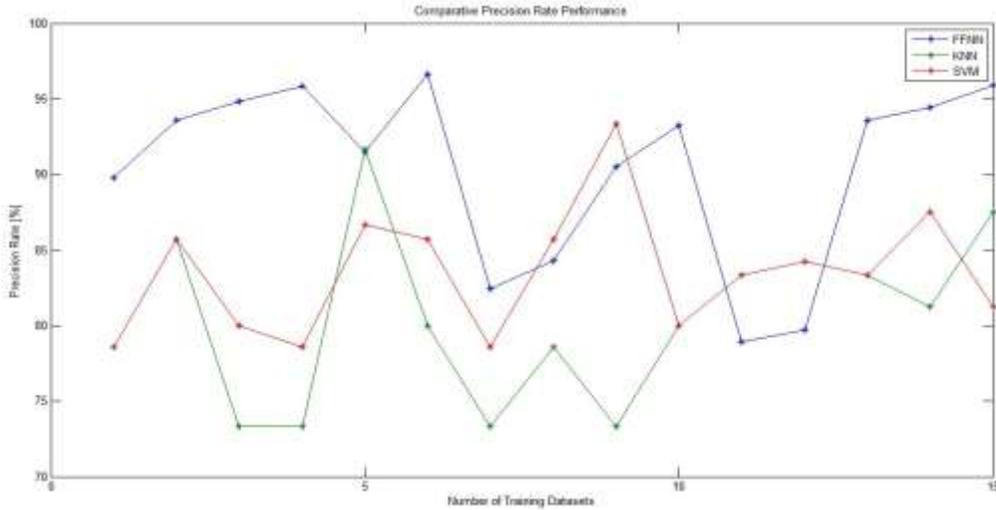


Figure 4:- Precision rate performance according to varying datasets

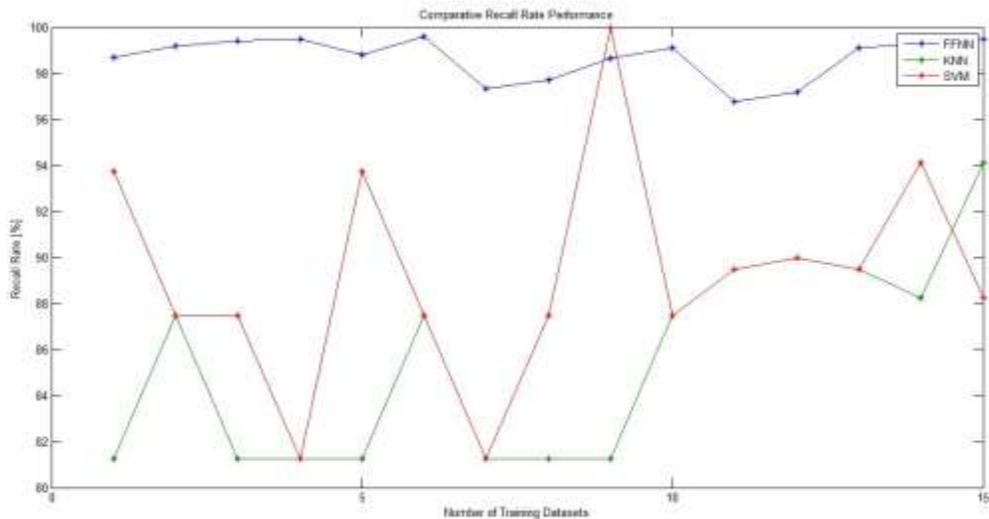


Figure 5:- Recall rate performance according to varying datasets

Figures 3, 4 and 5, shows the performance for accuracy, precision and recall rate analysis for each dataset using three different classifiers. From these results, it's observed that presented work using modified FFNN classifier shows more precision, recall and recognition accuracy as compared to SVM and KNN classifier results. Table 1 below shows the average performance results for presented and existing methods.

Table 2:- Average Performance of KNN, SVM & FFNN classifier for Proposed Methods

	KNN	SVM	Proposed-FFNN
<b>Recognition Accuracy (%)</b>	82.3069	84.7146	94.5801
<b>Precision Rate</b>	80.499	83.499	90.3385
<b>Recall Rate</b>	85.5034	89.2534	98.6617

From above table, performance of proposed approach shows great improvement which is nothing but approximate 11 % recognition accuracy as compared to existing methods on real dataset. Table 3 shows comparative analysis of proposed method with existing methods in terms of size of dataset, Features extracted, classifier used & recognition rate achieved.

**Table 3:-**Comparison of presented method with existing methods in literature

Sr. No.	Method Proposed by	Dataset Size	Features Extracted	Classifiers used	Recognition rate
1.	M.Hanmandlu, et al. [2]	4750 (Hindi Characters)	exponential membership function estimated by entropy and error function	Fuzzy model based recognition	90.65%
2.	S. Arora et al [6]	1500 (Hindi Characters)	Shadow, view based, Chain code histogram features	SVM ,ANN and with MLP neural network combination	98.61%
3.	SushmaShlke et al.[7]	24,000 (Compound Characters)	Pixel density, Euclidian distance, modified wavelet approximation	MLP neural network	97.95%
4.	A.N.Holambe et al. [10]	Unknown Marathi Characters	Statistical, Structural & Global Transformation moments	SVM & K-NN	96.00%
5.	Vijaya R. Pawaret al.[11]	2,000 (Marathi Characters)	<b>21 features</b> Structural features like End Points, Junctions, loops and Statistical Features like aspect ratio	Self-organizing Maps Neural Network	93.00%
6.	Presented Method	7500 (Marathi Words)	Statistical, Structural, Gradient & Distance transform	KNN	82.30
				SVM	84.70
				FFNN	94.57%

### Conclusion and Future Work:

This paper presents the hybrid and optimized framework for Devnagari handwritten character recognition. Section III shows detailed procedure of presented methodology which explains all new algorithms proposed for improving the recognition accuracy. The key contribution of the work is the design of hybrid feature extraction technique which contains optimized feature vector set of length 91. Another contribution is the use of modified FFNN algorithm for classification purpose which shows more accuracy than other methods of classification. Overall recognition accuracy obtained using SVM, K-NN & FFNN classifiers is 84.70%, 82.30% & 94.57% respectively. The main advantage of the work is that, it is scalable and works on diverse sets of datasets which is not achieved in past so far. In future, work on different languages like Hindi or Sanskrit can be suggested. It can be extended for the recognition of words, sentence and documents. This approach can be used in multilingual character recognition as well. Writer adaptation and spelling and semantic checks can be incorporated to correct errors at stroke and character levels. No OCR in this world is 100% accurate till date. Minimization of extraction time required so that overall process will become faster with maximum accuracy.

### References:

1. Anoop Namboodiri, "Online Handwritten Script Recognition", *IEEE* Vol. 26 No. 1, January 2004.
2. Hanmandlu, O.V. Ramana Murthy, Vamsi Krishna Madasu, "Fuzzy Model based recognition of Handwritten Hindi characters", *IEEE Computer society, Digital Image Computing Techniques and Applications*, 2007.
3. Satish Kumar and Chandan Singh, "A Study of Zernike Moments and its use in Devnagari Handwritten Character Recognition", *Intl.Conf. on Cognition and Recognition*, pp. 514- 520, 2005.
4. N. Sharma, U. Pal, F. Kimura, and S. Pal, "Recognition of Off-Line Handwritten Devnagari Characters Using Quadratic Classifier", *ICVGIP 2006, LNCS 4338*, pp. 805 – 816, 2006. *Proc. of ICVGIP 2006, LNCS 4338*, springer , pp. 805{816, 2006.
5. Prachi Mukherji, Priti P. Rege, "Shape Feature and Fuzzy Logic Based Offline Devnagari Handwritten Optical Character Recognition", *Journal of Pattern Recognition Research (2009)*,pp 52-68.
6. Sandhya Arora, Debotosh Bhattacharjee "Combining Multiple Feature Extraction Techniques for Handwritten Devnagari Character Recognition", 2008 IEEE Region 10 Colloquium and the Third ICIIS, Kharagpur, INDIA December 8-10. 978-1-4244-2806-9/08, *IEEE* © 2008, vol., no., pp.1-6, 8-10 Dec. 2008 doi: 10.1109/ICIINFS.2008.4798415
7. Sandhya Arora et al., "Performance Comparison of SVM and ANN for Handwritten Devnagari Character Recognition", *IJCSI International Journal of Computer Science Issues*, Vol. 7, Issue 3, May 2010.

8. Sushama Shelke, Shaila Apte, "A Multistage Handwritten Marathi Compound Character Recognition Scheme using Neural Networks and Wavelet Features", International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 4, No. 1, March 2011
9. Mitrakshi B. Patil, Vaibhav Narawade, "Recognition of Handwritten Devnagari Characters through Segmentation and Artificial neural networks", International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Vol. 1 Issue 6, August – 2012
10. Tanuja K, UshaKumari V and Sushma T M, "Handwritten Hindi Character Recognition System Using Edge detection & Neural Network", International Journal of Advanced Technology and Engineering Exploration, ISSN (Print): 2394-5443 ISSN (Online): 2394-7454 Volume-2 Issue-6 May-2015
11. Ved Agnihotri, "Offline Handwritten Devnagari Script Recognition", IJCSI International Journal of Computer Science Issues, 2012.
12. Anilkumar N Holambe, Ravindra C Thool, "Combining Multiple Feature Extraction Technique and Classifiers for Increasing Accuracy for Devnagari OCR", IJSCE, ISSN: 2231-2307, Volume-3, Issue-4, September 2013
13. Muhammad, Arif Mohamad, Dewi Nasien, "A Review on Feature Extraction and Feature Selection for Handwritten Character Recognition", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 6, No. 2, 2015
14. Vijaya Rahul Pawar and Arun Gaiwad, "Multistage Recognition Approach for Offline Handwritten Marathi Script Recognition", International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.7, No.1, 2014, pp.365-378 <http://dx.doi.org/10.14257/ijcip.2014.7.1.34>, ISSN: 2005-4254 IJSIP
15. Aparna, K.H., Subramanian, V., Kasirajan, M., Prakash, G. V., Chakravarthy, V.S., and Madhvanath, S., 2004. "Online handwriting recognition for Tamil", Ninth International Workshop on Frontiers in Handwriting Recognition, pp. 438-443.
16. Huang, B. Q., Zhang, Y. B. and Kechadi, M. T., 2007. "Preprocessing techniques for online handwriting recognition", *Seventh International Conference on Intelligent Systems Design and Applications*, pp. 793-800.
17. Bharath, A. and Sriganesh M., 2010. "On the significance of stroke size and position for online handwritten Devnagari word recognition: An empirical study", *20th International Conference on Pattern Recognition (ICPR)*, pp. 2033-2036.
18. ZHAO Bin, LIU Yong and XIA Shao- Wei "Support Vector Machine and its Application in Handwritten Numeral Recognition" 0-7695-0750-6/00 2000 IEEE.
19. N. Anupama, Ch. Rupa & Prof. E. Sreenivasa Reddy "Character Segmentation for Telugu Image Document using Multiple Histogram Projections" Global Journal of Computer Science and Technology Graphics & Vision Volume 13 Issue 5 Version 1.0 Year 2013, Online ISSN: 0975-4172 & Print ISSN: 0975-4350
20. U. Pal and B. B. Chaudhuri, "Indian script character recognition: A survey", *Pattern Recognition*, vol. 37, pp. 1887–1899, 2004.
21. Bikash Shaw, Swapan Kumar Parui, "Offline Handwritten Devnagari Word Recognition: A holistic approach based on directional chain code feature and HMM", 978-0-7695-3513-5/08 © 2008 IEEE DOI 10.1109/ICIT.2008.33
22. J.Pradeep, E.Srinivasan, "Diagonal Feature Extraction Based Handwritten Character System Using Neural Network", *IJCA*s (0975 – 8887) Volume 8– No.9, October 2010, PP.17-22
23. Brijmohan Singh, Ankush Mittal, "An Evaluation of Different Feature Extractors and Classifiers for Offline Handwritten Devnagari Character Recognition", *Journal of Pattern Recognition Research* 2 (2011) 269-277
24. Mahesh Jangid, "Devnagari Isolated Character Recognition by using Statistical features" International Journal on Computer Science and Engineering (IJCSE) ISSN : 0975-3397 Vol. 3 No. 6 June 2011
25. Gita Sinha, Mrs. Rajneesh Rani, Prof. Renu Dhir, "Handwritten Devnagari Numeral Recognition Using Zonal Based Feature Extraction Method and SVM Classifier", © 2012, *IJARCSSE*, Volume 2, Issue 6, June 2012 ISSN: 2277 128X
26. Pratibha Singh, "Reliable Devanagri Handwritten Numeral Recognition using Multiple Classifier and Flexible Zoning Approach", *I.J. Image, Graphics and Signal Processing*, 2014, 9, 61-68 Published Online August 2014 in *MECS* (<http://www.mecspress.org/>) DOI: 10.5815/ijigsp.2014.09.08, *I.J. Image, Graphics and Signal Processing*, 2014, 9, 61-68
27. Rajiv Kumar, Kiran Kumar Ravulakollu, "On the Performance of Devnagari Handwritten Character Recognition", *World Applied Sciences Journal* 31 (6): 1012-1019, 2014 ISSN 1818-4952, © IDOSI Publications, 2014, DOI: 10.5829/idosi.wasj.2014.31.06.2062.