International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 4 Issue: 6
ISSN: 2454-4248
70 – 76

Feature Extraction Techniques for Marathi Character Classification using Neural Networks Models

Ms. Smita Ashokrao Bhopi
Department of Computer Science & IT,
MGM's College of Computer Science & IT,
Nanded, Maharashtra (India)
smitabhopi@gmail.com

Mr. Manu Pratap Singh
Department Of Computer Science, IET
Dr. B. R. Ambedkar University
U.P.(India)Agra,
manu_p_singh@hotmail.com

***Abstract*—** Hand written Marathi Character Recognition is challenges to the researchers due to the complex structure. This paper presents a novel approach for recognition of unconstrained handwritten Marathi characters. The recognition is carried out using multiple feature extraction methods and classification scheme. The initial stages of feature extraction are based upon the pixel value features and the classification of the characters is done according to the structural parameters into 44 classes. The final stage of feature extraction makes use of the zoning features. First Pixel values are used as features and these values are further modified as another set of features. All these features are then applied to neural network for recognition. A separate neural network is built for each type of feature. The average recognition rate is found to be 67.96% , 82.67%,63,46% and 76.46% respectively for feed forward , radial basis , elman and pattern recognition neural networks for handwritten marathi characters.

Keywords- *Marathi handwritten characters recognition, Pixel Value Features, Histogram based features , Zoning based features Artificial neural networks.*

*****

I. INTRODUCTION

Character classification is a form of pattern recognition process. Presence of unwanted objects or disoriented patterns will affect the percentage accuracy. The most basic way of recognizing patterns is through using the probabilistic methods. It is very difficult to achieve 100% accuracy in recognition of handwritten characters. There are verities of writing styles because different peoples will write the same character differently. Handwritten character recognition has many possible application areas in various fields like postal automation, bank automation, form filling etc.

Handwritten character recognition for Indian scripts is quite a challenging task for the researchers because Indian languages are more complicated in terms of structure. Devnagari is the most popular script in India many languages like Hindi, Marathi, Nepali etc. are written in devnagari script .Devnagari script consists of 16 vowels and 36 consonants making 52 alphabets. Marathi is written from left to right. It has no upper and lower case characters. Every character has a horizontal line at the top called as the header line. The header line joins the characters in a word. Vowels are combined with consonants with the help of specific characteristic marks. These marks occur in line, at the top, or at the bottom of a character in a word. It also has complex compound characters in which two or more consonants are combined forming a new special symbol. Compound characters in Marathi script occur more frequently in the script as compared to other languages derived from Devanagari.

OCR work for printed and handwritten characters in various Indian scripts [1-3] is carried out by researchers but major work is found for Bangla [4, 5] and Devanagari. OCR work on printed Devanagari script started in early 1970s. Sinha [6-8] and Mahabala. Sethi and Chatterjee [5] also have done some earlier studies on Devanagari script and presented a Devanagari hand-printed numeral recognition system based on binary decision tree classifier. They used a similar technique for constrained hand-printed Devanagari character recognition. The first complete OCR system development of printed

Devanagari is perhaps due to Palit and Chaudhuri [10] , Pal and Chaudhuri [9]. The method proposed by Pal and Chaudhuri gives about 96% accuracy. Due to the complexities involved with Devanagari script, already existing methods cannot be applied directly with this script. A report on handwritten Devanagari characters was published in 1977 [11]. Kumar & Singh [13] proposed Zernike moments based approach for Devanagari character recognition. The other work on Devanagari character recognition is proposed 64 dimensional chain code features, but still no any standard OCR is available for the same. An excellent survey of the area is given in [15].For recognition of handwritten Devanagari numerals, Ramakrishnan et al. [16] used independent component analysis technique for feature extraction from numeral images. Bajaj et al [14] considered a strategy combining decisions of multiple classifiers. An extensive research on printed Devanagari text was carried out by Veena Bansal and R. M. K. Sinha [12, 13]. Multilayer perceptron was also used for classification by Sandhya Arora et al. [15] for handwritten Devanagari characters.

The rest of the paper is organized as follows: Section II discusses the feature extraction techniques, while section III describes the classification methods. Section IV presents the proposed system. Section V discusses the results obtained iby the proposed system and section VI finally discusses the conclusion and future scope.

 ISSN: 2454-4248
70 – 76

II. FEATURE EXTRACTION TECHNIQUES

Feature extraction is the process of extracting different features from the matrices of digitized characters. The characters are recognized on the basis of these features. Features of a character can be classified into two classes: Global or statistical features and Structural or topological features.

A. Global or Statistical Features

Global features are obtained from the collection of points representing the character matrix. These features can be easily discovered as compared to topological features. Global features are not affected too much by noise or distortions as compared to topological features. A number of techniques are used for feature extraction; some of these are: moments, zoning, projection histograms, n-tuples, crossings and distances.

B. Structural or Topological features

These features are related to the geometry of the character set to be considered. Some of these features are concavities and convexities in the characters, number of end points, no of holes in the characters etc. A lot of research has been done by different researchers to find different structural features. This feature set includes information for a character like location and number of holes in the characters, concavities in the skeletal structure, crossings of strokes, vertical end points of the character and bounding box of the character.

III. CLASSIFICATION METHODS

Classification is another most important component of OCR system. It is basically decides the feature space to which the unknown pattern belongs. Classification is usually done by comparing the feature vectors corresponding to the input character with the representative of each character class. But before doing this the classifier should posses a number of training patterns. A number of classification methods were purposed by different researchers some of these are statistical methods, syntactic methods, template matching, artificial neural networks, kernel methods.

A. Statistical methods

The purpose of the statistical methods is to determine to which category the given pattern belongs. By making observations and measurement processes, a set of numbers is prepared, which is used to prepare a measurement vector. Statistical classifiers are automatically trainable. *k*-NN method compares an unknown pattern to the set of patterns that have been already labeled with class identities in the training stage. A pattern is identified to be of the class of pattern, to which it has the closest distance, where as a bayesian classifier assigns a pattern to a class with the maximum a posteriori probability.

B. Syntactic or structural methods

Syntactic methods are good for classifying hand written texts. This type of classifier, classifies the input patterns on the basis of components of the characters and the relationship among these components. Firstly the primitives of the character are identified and then strings of the primitives are checked on the basis of pre-decided rules. Generally a character is represented as a production rules structure, whose left-hand side represents character labels and whose right-hand side represents string of primitives. The right-hand side of rules is compared to the string of primitives extracted from a word. So classifying a character means finding a path to a leaf.

C. Template matching

This is one of the simplest approaches to patter recognition. In this approach a prototype of the pattern that is to be recognized is available. Now the given pattern that is to be recognized is compared with the stored patterns. The size and style of the patterns is ignored while matching.

D. Artificial neural networks

A neural networks composed of inter connected elements called neurons. A neural network can trained itself automatically on the basis of examples and efficient tools for learning large databases. This approach is non-algorithmic and is trainable. The most commonly used family of neural networks for pattern classification task is the feed-forward network, which includes multilayer perception and Radial Basis Function (RBF) networks. But the limitation of the systems based on neural networks is their poor capability for generalization.

E. Kernel methods

Some of the most important Kernel methods are Support Vector Machines , Kernel Principal Component Analysis (KPCA), Kernel Fisher Discriminant Analysis (KFDA) etc. Support vector machines (SVM) are a group of supervised learning methods that can be applied to classification. In a classification task usually data is divided into training and testing sets. The aim of SVM is to produce a model, which predicts the target values of the test data. Different types of kernel functions of SVM are: Linear kernel, Polynomial kernel, Gaussian Radial Basis Function (RBF) and Sigmoid.

IV. THE PROPOSED SYSTEM

The proposed system is designed to recognize 44 characters of Marathi language. The characters used in the proposed system are shown in Figure1. The recognition system consists of phases like segmentation, pre-possessing, feature extraction, training and testing phase. First of all the handwritten document is segmented into lines. Then the lines are segmented into words, after that words are segmented into characters then pre-processing and normalization is performed on the characters to remove noise and make them identical. We have used three different feature extraction techniques i.e. pixel value features, histogram based features and zoning or blocks based features to get the features of characters. These features are used for both the phases i.e. training and testing the neural network. In the training phase features are extracted from the handwritten normalized characters and used for

 ISSN: 2454-4248
70 – 76

training the neural network. After the training neural network with sufficient samples the weights and biases for each network are saved. In the testing phase characters are extracted from handwritten script and then similar features are extracted from the character again after pre-processing and normalization of the character. The features are applied as inputs to the neural network. The output of the neural network gives the final recognition result. The next section discusses the proposed system in detail. The flow chart of the proposed scheme is described in Figure 2.

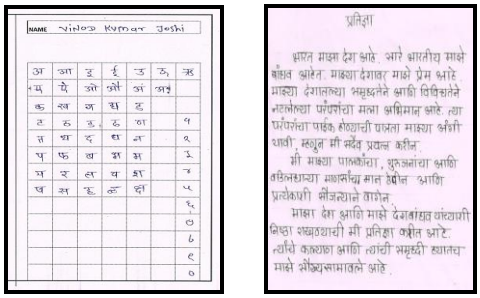

Figure.1 Sample of Handwritten Marathi Characters and Script

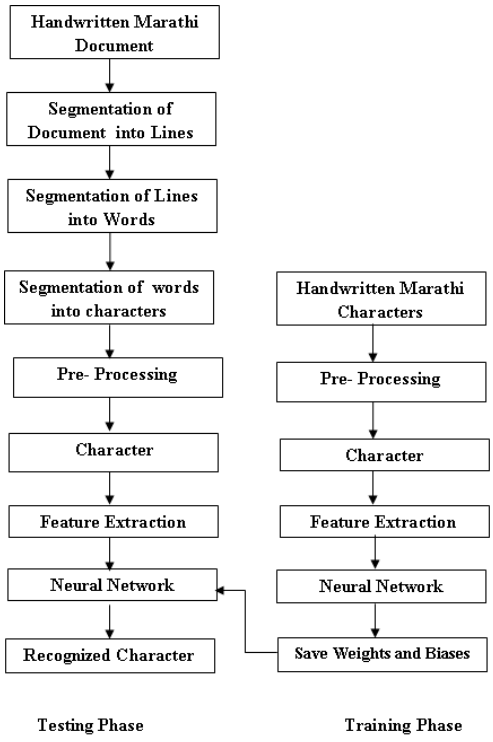

Figure 2. Proposed Character Recognition System

A. *Data collection*

We have collected the samples of handwritten documents from different peoples belonging to different categories in a separate sheet without any restrictions. The database for handwritten marathi documents is created by scanning the handwritten documents at 300 dpi using a flatbed scanner in color mode. The images are stored in jpeg file format. Here we assume that the documents may contain different styles of writing and the documents may be having skewed lines. Multiple, global or non uniform skew may present in the documents. Total 65 samples of handwritten documents having approximately 15 lines , 70 words and 150 characters in each document are scanned resulting into about 9750 different characters samples in the database.

B. *Pre-processing*

Before segmentation we have to perform some preprocessing on every document image. Preprocessing is used to reduce the noise and remove the unwanted data if any. We can also decrease the variation and transforms the data in specific format so that the data can be processed more easily and efficiently. First we need to convert the color document into gray image and then by using the threshold value we can convert it into to black and white. Remove all objects containing less than 30 pixels i.e. remove the unwanted objects or noise included in the image. Perform morphological operations like erosion and dilation of image as the writer may use different pen so thickness of character may be different .In normalization we will convert it into in to characters of single pixel width. We can also remove the skew if it is present in the document.

C. *Segmentation*

The text line segmentation methods can be normally classified into two types bottom-up and top-down. In the bottom-up approach, the neighboring components are grouped using some easy rules depending on the geometric relationship between neighboring blocks. The projection based methods are the top-down algorithms which are one of the most successful methods for machine printed text. The projection based methods are also successful for handwritten text where text lines are straight or easily separable. But due to different writing styles of the people, the text line segmentation is still very challenging. In general, text-line segmentation techniques are script independent. In the proposed system we have used projection based method for segmentation of lines and characters.

In Projection-based methods if an image A is of height H and width W i.e. of H * W size, the projection profile of the image is defined as follows:

$$P(i) = \sum_{j=1,2,\ldots..W} A\,(i,j), i = 1,2,\ldots..H, \quad (1)$$

Vertical Projection: For a binary image of size H * W where H is the height of the image and W is the width of the image, the vertical projection has been defined as

$$VP(j), j = 1,2,\ldots,W \quad (2)$$

This operation counts the total number of black pixels in each vertical column.

Horizontal Projection: For a binary image of size H * W where H is the height of the image and W is the width of the image, the horizontal projection is defined as

$$HP(i), i = 1,2,\ldots,H \quad (3)$$

This operation counts the total number of black pixels in each horizontal row.

 ISSN: 2454-4248
70 – 76

The projection is a one-dimensional signal that denotes the amount of text pixels per row as a result; the lobes (valleys) of the projection correspond to foreground (background) areas of the image. If the text lines have the same skew angle, the amplitude and the frequency of the projection are maximized when the skew of the text is zero. By calculating the projection for each angle and we can estimate the global skew angle according to a selected criterion. A common feature is the overlapping of successive text lines due to the ascenders and/or descenders of some characters. Hence, the formulation of a horizontal line as a separator is often not feasible to overcome this complexity, we can use the projections in order to locate the areas (i.e. the areas between two successive maxima) in which the separators should be allocated [24]. We have to find a path from the left to the right edge in each area, by attempting to move the obstacles.

In the proposed work text is written on plain paper. The lines and words are written in such a way that they do not overlap. We have segmented lines and words using horizontal and vertical projection profiles [21]. Peaks of the horizontal projection profiles separate the lines and in the vertical projection profiles separate the words in the document.
The following five steps are used in Segmentation.

- The line segmentation is based on horizontal histograms of the document. Those rows, for which HP[j] is zero; j = 1, 2, . . . ,H; serve as delimiters between successive text lines.
- The segmentation of the text line into words is based on the vertical projection of the text line. A vertical histogram of the text line is made and white spaces are used as word delimiter. Figure 3.a shows the horizontal and vertical projection profiles of the handwritten script.
- After extracting the sub-images corresponding to words for a text line, we locate the position of the header line of each word. In the horizontal projection of the word image the row containing maximum number of black pixels is considered as the header line.
- The header line can be removed by replacing all black pixels by white pixels.
- To separate character we make vertical projection of the image starting from the header line position to the bottom row of the word image box. The columns that have no black pixels are treated as boundaries for extracting corresponding characters.
- To separate symbols of the top strip we compute the vertical projection of the image starting from the top row of the image to the header line position. The columns that have no black pixels are used as delimiters for extracting top modifier symbol boxes. Figure 3.b shows the segmented words and characters.

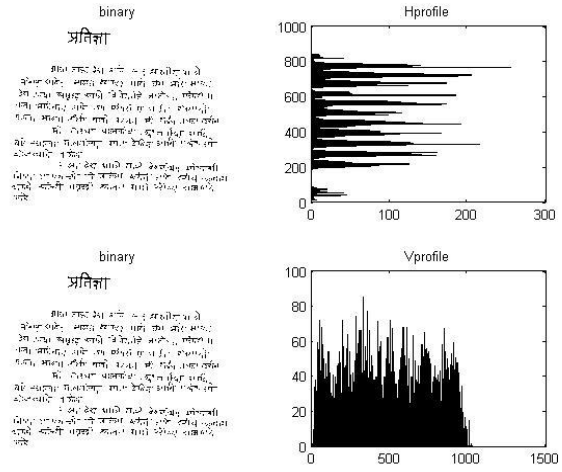

Figure 3.a

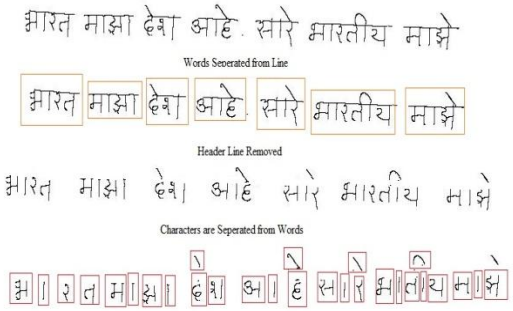

Figure 3.b

Figure 3. (a) Horizontal and Vertical Projection Profile
(b) segmented lines, words and characters

D. Feature extraction

A feature vector is preferred for high recognition results in optical character recognition. The structural feature and gradient based feature perform well in similar shape characters recognition in Devnagari. We considered the performance with features ranging from the simple most features in which each feature vector element was the direct pixel value to more computationally expensive features obtained from projection profiles and zoning features. In further study we may combine the direct pixel, profile feature with the zoning features. In this paper we have used three feature extraction methods i.e. Pixel values features, Histogram based features and zoning/block based features. A brief description of feature extraction is given in this section.

1) Pixel Value Features

In our experiments we have started with a simple feature definition i.e. the pixel value. In Pixel values method the character image is converted to black & white then the noise is removed. The hole filling operation is performed to obtain the uniform character. We formed feature vectors by storing the size normalized two dimensional digit images into one dimensional column feature vectors where each feature element is the pixel value. The image is cropped and resized into 15 X 15 matrixes. Then the character matrix is reshaped into column vector of size 225 X 1.

 ISSN: 2454-4248
70 – 76

2) Histogram Based Features

We have used simple histogram based features, which are easy to extract. Characters can be represented by projecting the pixel gray values. This representation creates one-dimensional signal from a two dimensional image, which can be used to represent the character image Histogram values can be obtained by drawing the histogram of gray character image and store the histogram count values in a column matrix of 256x1.Each feature element represents the histogram value. The histogram feature values range from 0 to 470 which is the gray value of each pixel in the character image. Figure.4 shows histogram of handwritten a character "EE"

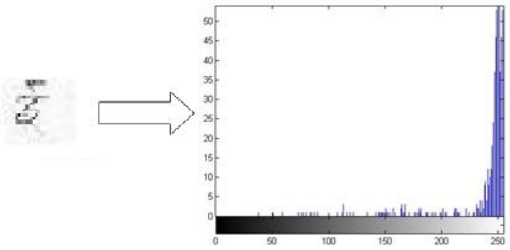

Figure 4. Histogram of character "EE"

3) Zoning /Block Based Features

The frame containing the character is divided into several overlapping or non-overlapping zones. The densities of the points or some features in different regions are analyzed. The image zoning value feature is the average pixel intensity value in a specified region or zone [23]. We defined these zones by dividing the size normalized image into equal number of rows and columns. We divide the character image into the zone size of 3x3 i.e. from the size normalized 15 x 15 character image we have extracted 25 features from 25 zones. We can calculate the mean values as follows. First we have converted the character image into black and white i.e. binary image. Then we have removed the blank space from left, right, upper and lower side of image. After that the images was resized into matrix of size 15x15 and divided into the blocks of 3x3.We have calculated the mean value for each block and then the mean values were stored in to a column matrix of size 25x1. The figure 5 shows the character matrix divide into 3x3 zone size.

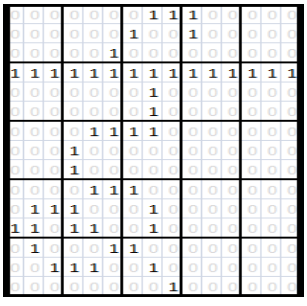

Figure5. Character EE divided into Zone size 3x3

E. Training and Testing of the Neural Network

A neural network is a set of connected input and output units in which each connection has a weight associated with it. The network will adjust the weights during the training phase so it will be able to predict the correct class for the input values. Feed forward neural networks, including multilayer perceptron (MLP), radial basis function (RBF) network, Back Propagation neural network (BPN) etc are used as classifiers in OCR system.

In feed forward architecture, the activations of the input units are set and then propagated through the network until the values of the output units are determined. The network act as a vector valued function it accepts one vector as the input and produce another vector at output.

Radial basis function can be used for approximating functions and recognizing patterns It uses Gaussian activation function. The response of such function is non-negative for all value of x. The function is defined as

$$f(x) = \exp(-x)2 \quad (4)$$

An Elman network is a three-layer network with the addition of a set of "context units". The middle (hidden) layer is connected to these context units fixed with a weight of one. At each time step, the input is feed-forward and a learning rule is applied. The fixed back-connections save a copy of the previous values of the hidden units in the context units.

Elman networks are also known as "simple recurrent networks" (SRN).

we have implemented four types of neural networks i.e. Feed forward , Radial basis, Elman back propagation and Pattern recognition neural network using three different types of feature extraction methods i.e. pixel value features, histogram based features and zone/block based features for each network[20]. The transfer functions TF{i} can be any differentiable transfer function such as TANSIG, LOGSIG, or PURELIN. The training function BTF can be any of the backprop training functions such as TRAINLM, TRAINBFG, TRAINRP, TRAINGD, etc. We have used TANSIG as transfer function and TEAINLM and TRAINGDM as training function.

We have trained totally twelve different types of networks using character samples of handwritten marathi alphabets collected from different persons. The weights and biases for each network are saved after training the neural network with sufficient samples. These already trained networks are then used for classification of the characters which we have separated from the handwritten script in the segmentation step.

First we have trained the network using different features and then we have used the same networks for recognition of the characters which we have separated from the hand written script using segmentation. The network has to simulate to get the output values and after that we have plotted the regression and confusion matrix. The following figures (Fig.7 and Fig.8) shows the performance plots for all the four types of neural networks i.e. feed forward, Radial Basis, Elman, pattern recognition. This indicates that the performance of Radial basis neural network works is more effective. It has been seen from the analysis that the histogram and zoning feature is giving better result compared to other features.

 ISSN: 2454-4248
70 – 76

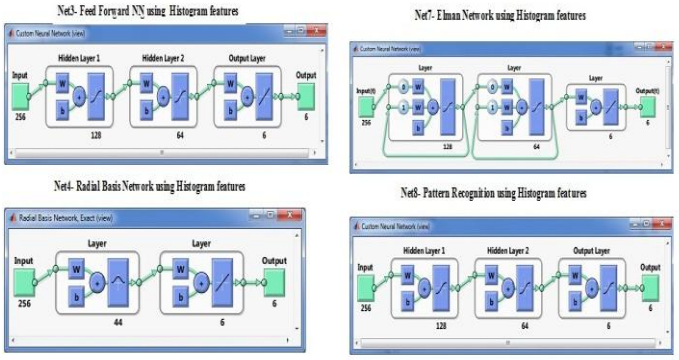

Figure6. Network structure of Feed Forward, Radial Basis, Elman and Pattern Recognition Neural Network

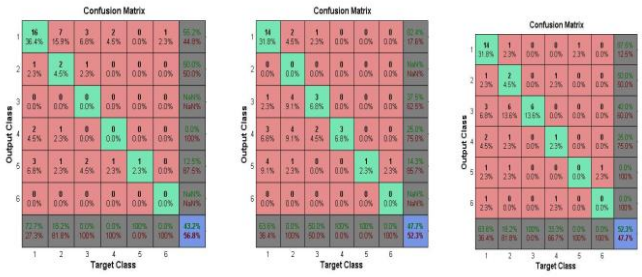

Figure 7. Confusion matrices for different feature extraction techniques

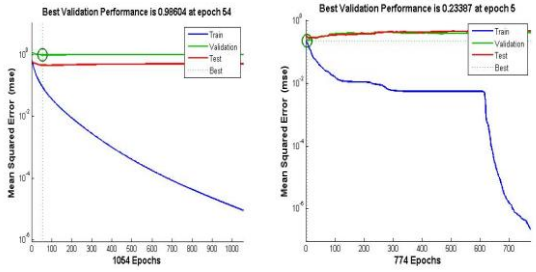

Figure 8. Performance plots for Elman and Pattern Recognition forward Neural Network

V. RESULTS AND DISCUSSION

The dataset of handwritten marathi document is created by collecting the samples from different individuals. About 65 samples of handwritten document are collected containing about more than 9000 character samples. Out of these, two third of the samples i.e. 40 documents were used for training and remaining were used for testing. The handwritten characters may have different shapes as per the writing style of the writer. This may result in classification of the same character to different classes. The characters are then normalized to a size of 15x15 for feature extraction. 225 pixel value features, 256 histogram based features and 25 zone/block based features are used to train the neural network. The inputs to the neural networks are equal to the number of features derived, but the number of hidden neurons depends upon the characters in that class. The performance goal is kept to 0.001 and 5000 epochs are enough to train the network using feed forward algorithm. The feed forward algorithm is the fastest back propagation algorithm. The time required for training the neural networks in each of the 44 classes depends upon the number of characters and the number of samples per character in that class. The testing time is the same irrespective of the class. The proposed system is implemented using MATLAB. The time required to test a character is approximately 0.045 seconds. Table 1 indicates the recognition results for all the three feature extraction techniques. The results show that the zoning based features improve the results considerably for both the training as well as testing samples. We can also improve the results by combining the features and by using different values in zoning based features like sum or mean of the pixels .

Feature Extraction Method/ Neural Network	Feed Forward NN	Radial Basis NN	Elman NN	Pattern Recognition NN
Pixel Value Features	71%	82%	71%	79%
Histogram Based Features	61%	89%	46%	86%
Zoning or Blocks Based Features	72%	77%	74%	64%

Table 1. Results of Three Different Feature Extraction techniques

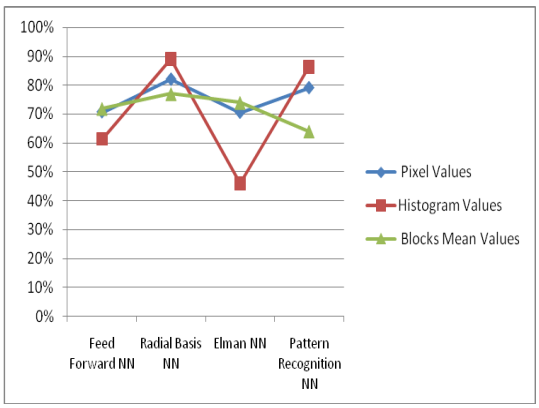

Figure 8. Recognition Results of all four types of Neural Networks

VI. CONCLUSION AND FUTURE SCOPE

In this paper we have tested the four different neural networks which clearly mention that the performance of Radial basis neural networks for recognition of characters separated from handwritten script which is better as compared to other networks. Here we have used three different types of segmentation methods i.e. projection profile, run length smearing and bounding box for. We can improve the results of segmentation by combining these projection profile method with bounding box method. From analysis we can say that zoning or blocks mean are more effective. We have not

 ISSN: 2454-4248
70 – 76

considered the modifiers and compound characters here so in future we can extend the same work for compound characters.

REFERENCES

[1] K. Roy, T. Pal, U. Pal, and F. Kimura, "Oriya Handwritten Numeral Recognition System", Proc. 8th Int. Conf. Document Analysis and Recognition, vol. 2, Seoul, Korea, Aug. 31-Sep. 1, 2005, pp. 770 – 774.

[2] Arun Pujari, C. Dhanunjaya Naidu, M. Sreenivasa Rao, and B. C. Jinaga, "An Intelligent Character Recognizer for Telugu scripts using Multi resolution Analysis and Associative Memory", Image and Vision Computing, vol. 22, 2004, pp. 1221-1227.

[3] U. Bhattacharya, S. K. Ghosh, and S. K. Parui, "A Two Stage Recognition Scheme for Handwritten Tamil Characters", Proc. 9th Int. Conf. Document Analysis and Recognition", Parana, Sept. 23-26, 2007, pp. 511 – 515.

[4] B. B. Chaudhuri, and U. Pal, "A Complete Printed Bangla OCR System", Pattern Recognition, vol. 31, no. 5, 1998, pp. 531-549.

[5] A. F. R. Rahman, R. Rahman, and M. Fairhurst, "Recognition of Handwritten Bengali Characters: A Novel Approach", Pattern Recognition, vol. 35, 2002, pp. 997-1006.

[6] Sinha R.M.K., Mahabala H. (1979) IEEE Trans. System, Man Cybern., 9, 435-441.

[7] Sinha R.M..K. (1973) Ph.D. Thesis ,Electrical Engineering Department, Indian Institute of Technology, India.

[8] BansalV., Sinha R.M.K. (1999) Proceedings of the Fifth International Conference on Document Analysis and Recognition , 653-656.

[9] Pal U., Chaudhuri B.B. (1997) Vivek, 10, 12-24.

[10] Palit S., Chaudhuri B.B., Das P.P., Chatterjee B.N. (1995) (Eda.) Pattern Recognition, Image Processing and Computer Vision, Narosa Publishing House: New Delhi, India, 163-168.

[11] Arora S., Bhattacharya D., Nasipuri M., Malik L. (2006) IEEE –International Conference on Signal And Image Processing, Hubli, Karnataka.

[12]]Veena Bansal, and R. M. K. Sinha, "On How to Describe Shapes of Devanagari Characters and Use them for Recognition", Proc. 5th Int. Conf. Document Analysis and Recognition, Bangalore, India, Sept. 20-22, 1999,410-413.

[13] Veena Bansal, "Integrating Knowledge Sources in Devanagari Text Recognition", Ph.D. Thesis, IIT Kharagpur, India, 1999.

[14] R. Bajaj, L. Dey, and S. Chaudhary, "Devanagari Numeral Recognition by Combining Decision of Multiple Connectionist Classifiers", Sadhana, 27, 2002, pp. 59-72.

[15] Sandhya Arora, Debotosh Bhattacharjee, Mita Nasipuri, Deepak Basu, and Mahantapa Kunda, "Combining Multiple Feature Extraction Techniques for Handwritten Devanagari Character Recognition", IEEE Region Colloquium and the Third ICIIS, Kharagpur, India, 2008, pp. 342-347.

[16] U. Pal, S. Chanda, T. Wakabayashi, and F. Kimura, "Accuracy Improvement of Devanagari Character Recognition Combining SVM and MQDF", 11th Int. Conf. Frontiers in Handwriting Recognition, Québec, Canada, Aug. 19-21, 2008, pp. 367-372.

[17] U. Pal, T. Wakabayashi, and F. Kimura, "Comparative Study of Devanagari Handwritten Character Recognition using Different Feature and Classifiers", Proc. 10th Int. Conf. Document Analysis and Recognition, Barcelona, Spain, Jul. 26-29, 2009, pp. 1111-1115.

[18] U. Pal, T. Wakabayashi, and F. Kimura, "Handwritten Bangla Compound Character Recognition using Gradient Feature", Proc. 10th Int. Conf. Information Technology, Orissa, India, Dec. 17-20,2007, pp. 208-213.

[19] U. Pal, N. Sharma, T. Wakabayashi, and F. Kimura, "Off-line Handwritten Character Recognition of Devanagari Script", Proc. 9th Int. Conf. Document Analysis and Recognition, vol. 1, Parana, Sept. 23-26, 2007, pp. 496-500.

[20] Ms. S.A. Bhopi, Mr. M. P. Singh, "Performance analysis of Handwritten Devnagari Character Recognition using Feed Forward , Radial Basis , Elman Back Propagation, and Pattern Recognition Neural Network Model Using Different Feature Extraction Methods", May 18 Volume 4 Issue 5 , International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRSCE), PP: 152 – 158

[21] Ms. S.A. Bhopi, Mr. M. P. Singh , "A Comparative Study of Segmentation Techniques used in Handwritten Documents " , Vol.6(4), Apr 2018, International Journal of Computer Sciences and Engineering(IJCSE),E-ISSN: 2347-2693

[22] R. Kumar and K. K. Ravulakollu, "Handwritten Devnagari Digit Recognition : Benchmarking On New Dataset," vol. 60, no. 3, pp. 543–555, 2014

[23] R. Kumar and K. K. Ravulakollu, "On the Performance of Devnagari Handwritten Character Recognition," vol. 31, no. 6, pp. 1012–1019, 2014.

[24] V. Katsouros and V. Papavassiliou, "Segmentation of Handwritten Document Images into Text Lines," 2005.

[25] V. Bansal and R. M. K. Sinha, "Segmentation of touching and fused Devanagari characters," Pattern Recognit., vol. 35, no. 4, pp. 875–893, Apr. 2002.