

# Radial Basis Function For Handwritten Devanagari Numeral Recognition

Prerna Singh  
M.Tech Student  
Shobhit University,  
Meerut, India

Nidhi Tyagi  
Assistant Professor  
ISCEIT, Shobhit University,  
Meerut, India

**Abstract**—The task of recognizing handwritten numerals, using a classifier, has great importance. This paper applies the technique of Radial Basis Function for handwritten numeral recognition of Devanagari Script. Lot of work has been done on Devanagari numeral recognition using different techniques for increasing the accuracy of recognition. Since the database is not globally created, firstly we created the database by implementing pre-processing on the set of training data. Then by the use of Principal Component Analysis we have extracted the features of each image, some researchers have also used density feature extraction. Since different people have different writing style, so here we are trying to form a system where recognition of numeral becomes easy. Then at the hidden layer centers are determined and the weights between the hidden layer and the output layer of each neuron are determined to calculate the output, where output is the summing value of each neuron. In this paper we have proposed an algorithm for determining Devanagari numeral recognition using the above mentioned system.

**Keywords**-Radial Basis Function; Devanagari Numeral Recognition; K-means clustering; Principal Component Analysis (PCA).

## I. INTRODUCTION

Handwritten character recognition has been around since 1980's. Till date many researches have been done. Automatic reading of numerical fields has been attempted in several application areas such as online handwritten recognition on computer tablets, recognize zip codes on mail for postal address sorting, processing bank check amounts, numeric entries in forms filled up by hand (for eg. Tax forms) and so on. While solving this domain of handwritten recognition many challenges are faced. As the handwritten digits are not always of same size, thickness, or orientation and position relative to the margins, many handwritten versions are even hard to recognize.

Handwritten recognition [1] is the ability of a computer to receive and interpret intelligible handwritten input from sources such as document, photographs, touch-screens & other devices. But in the past various schemes or algorithms are proposed in

different languages like English [2], Chinese[3], Arabian, Persian, Bengali, Gujarati. Different classifiers have been used for handwritten digit recognition using neural network. Recently, significant contribution towards the improvement of recognition rates have been made by means of different combination strategies [4] and by the use of support vector machines [5].

Machine learning is an artificial system[6] capable of autonomous acquisition & integration of knowledge. This capacity to learn from experience, analytical observation other means, results in system that can improve its own speed or performance and allows computer to "learn". It can be categorized into supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning [7].

In this paper we proposed a method for recognition of Handwritten Devanagari numerals using Radial basis function. It is a feed forward neural network that computes activation at the hidden neuron in a way that is different from product between input vector and the weight vector. Hidden neuron activation in RBF are computed using an exponential of distance measure (usually the Euclidean distance or weighted norm) between the input vector and a prototype vector that characterizes the signal function at a hidden neuron.

Devanagari script was originally developed to write Sanskrit and after some time it was adapted by many other languages such as Rajasthani, Marathi [8], Nepali, Bhojpur, Konkani and Hindi [9], the mother tongue of majority of Indians.

The paper is organized in various sections. Further sections are like section 2, contains the related work on pattern recognition, section 3 gives the overview of system, how recognition takes place using RBF, section 4 elaborates the proposed work and section 5 describes certain examples on the proposed system, section 6 gives the conclusion and future work.

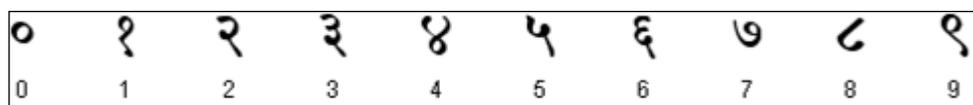


Fig.1 Devanagari Numerals

## II. RELATED WORK

Connell, R.M.K. Sinha and Anil K. Jain [10] describes the combination of classifiers capturing both online and offline features using markov model, three nearest neighbour and clustering. A classification accuracy of 86.5% is obtained. Anita Pal & Dayashankar Singh has performed a work to recognize handwritten English character [2] using a multilayer perceptron with one hidden layer. Fourier descriptors with back propagation network yields good recognition accuracy of 94%. The skeletonized and normalized binary pixels of English characters were used as the inputs of the . M.Hanmandlu, J.Grover, S.Vasikarla, V.K.Madasu [9] used fuzzy modelling for recognition and produces 95% recognition rate. The criterion function of the relationship between entropy and energy of the fuzzy sets is used. G.G.Rajput, S. M. Mali [8] presented a method for recognition of isolated Marathi handwritten numerals using Fourier descriptors, invariant translation, scaling and rotation, and used as features of the images representing handwritten numerals. Support Vector Machine is used as pattern recognition and gives the accuracy rate of 97.85%. Shailedra Kumar Shrivastava and Sanjay S.Gharde [5] applies support vector machine for recognition and produces the average recognition rate of 99.85% using moment invariant and affine moment invariant.

The critical look at the available literature reveals that lot of efforts have been made to get higher accuracy but still tremendous amount of work need to be done for improving recognition accuracy by developing new feature extraction techniques or modifying the existing feature extraction technique. This paper proposed a technique for Devanagari numeral recognition using Radial basis function by forming the cluster of training data at the hidden neuron and a set of test data are input to recognize the exact digit.

## III. RADIAL BASIS FUNCTION (RBF)

RBF was originally introduced into the literature for the purpose of interpolation of data points on a finite training dataset [11]. This network typically have three layers input layers, hidden layers with a non-linear RBF [12] activation function and a linear output layer. Hidden layer performs mapping from the input space into a higher dimensional space in which the patterns become linearly separable.

The output  $\phi$  of the network is thus for N number of neurons in the hidden layer is

$$\phi(x) = \sum a_i \sigma(\|x - C_i\|) \dots \dots \dots (1)$$

where  $C_i$  is the center vector for neurons  $i$  and  $a_i$  are the weight of the linear output neurons.

### SYSTEM OVERVIEW

The problem of recognition of handwritten numerals can be broadly classified into three blocks:-

- I. Pre-processing
- II. Feature extraction
- III. RBF classifier

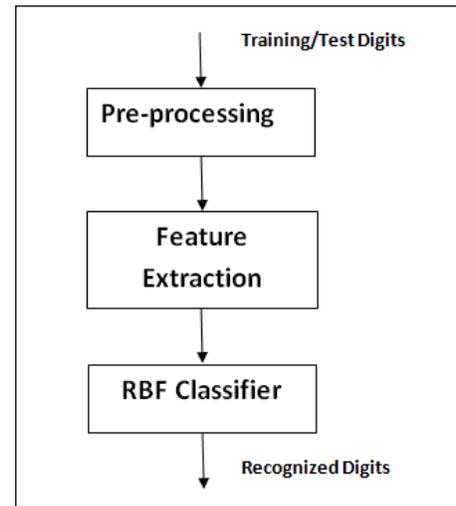


Fig 2. Block Diagram of System

### A. Creation of Data

Our objective is to obtain a set of handwritten samples of Devanagari numerals that capture variation in hand writing between and within writers. Therefore we need numeral sample from multiple writers as well as multiples sample from each writer.

**PRE-PROCESSING:-**Pre-processing is the process to removing noise and distortion present in input text due to hardware and software limitations. Steps are performed in pre-processing are size normalization and noise removal.

**Normalization-**General problem which occurs is “how to standardize [13] the size of each image”.

Following steps took place:-

- Start the image with  $xsize0 * ysize0$  pixels which fit the isolated numeral, by removing blank rows and columns.
- Rescale the size of the image to  $(Xsize * Ysize)$  pixels which is the maximum size according to  $xsize0$  or  $ysize0$  i.e  
 $Xsize = \max(xsize0, ysize0)$   
 $Ysize = \max(xsize0, ysize0)$
- Find the center of gravity co-ordinates.
- Now put the character at the center of frame.
- Finally rescale the image to the size of  $36 * 36$  pixels.

**Noise Removal:-** It may be, while scanning some noise is intercepted. That noise could be because of some particles of dust on the scanner or because of the poor quality of paper on which numerals are written. Remove the noise from the image file.

### B. Feature Extraction

When the input data to an algorithm is too large to be processed and is suspected to be notoriously redundant then the input data will be transformed into a reduced representation set

of features. We can also consider density feature extraction [1,14] to encode style characteristics without being affected by variations in minute details of similarly formed characters.

Using Principal Component Analysis:- It is an effective technique for dimensionality reduction and extracting important information used in several areas such as machine learning, neural networks and signal processing. It maximizes the rate of decrease of variance and transforms a number of correlated variables into (smaller) number of uncorrelated variables called principal components [15]. The first principal component accounts for as much variability in the data as possible and each succeeding component accounts for as much of the remaining variability as possible. PCA can be done by eigenvalues decomposition of covariance matrix of a data matrix, usually after mean centering the data for each attribute.

#### IV. PROPOSED RBF NETWORK

In addition to feature extraction, we need to calculate the centers of the basis function and the network is given in fig.3. The network divide its work in two phase, firstly it calculate the centers and then the weights between the hidden and the output neuron,

**K-means clustering**:-Now we want to calculate centers [16, 17] of basis function, the main idea is to define k centroids, one for each cluster. Steps for calculating centers are:-

1. Place K points in each cluster represent initial group center.
2. Assign each object to the group that has the closest centers
3. Re-calculate the position of the K centers , when all the objects are assigned
4. Repeat step 2,3 until the centers no longer move.

With the centers identified, now we need to determine the weights [18,19,20] using LMS procedure from the hidden neurons to output summing neurons.

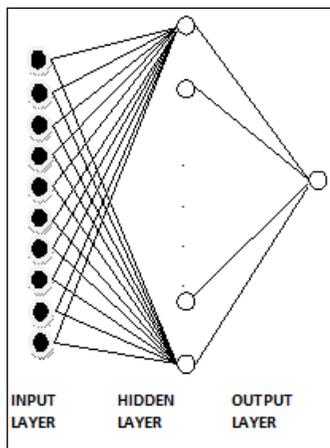


Fig.3 Radial Basis Function Network

Algorithm for the proposed network is given below:

#### ALGORITHM FOR PATTERN RECOGNITION USING RBF

**Input:** image x of Devanagari handwritten digit from test set.

**Initialize:**

1. the training set patterns after normalizing the size and removal of noise.
2. Extract important information for each data to be stored in cluster.
3. Remove the redundant data items using PCA

- Output:**
1. Calculate the centers of basis function.
  2. Determine weights between the hidden neuron
  3. Input data matches with any of the cluster, for that particular cluster weight =1 otherwise 0.
  4. Once values are calculated then we can determine the output by summing values of each neuron.

#### V. EXAMPLE:

As per the proposed system, let's consider an example when we insert an input, then what will be the values of each of the cluster represented by C.

I/P/	C0	C1	C2	C3	C4	C5	C6	C7	C8	C9
०	1	0	0	0	0	0	0	0	0	0
१	0	1	0	0	0	0	0	0	0	0
२	0	0	1	0	0	0	0	0	0	0
३	0	0	0	1	0	0	0	0	0	0
४	0	0	0	0	1	0	0	0	0	0
५	0	0	0	0	0	1	0	0	0	0
६	0	0	0	0	0	0	1	0	0	0
७	0	0	0	0	0	0	0	1	0	0
८	0	0	0	0	0	0	0	0	1	0
९	0	0	0	0	0	0	0	0	0	1

C0...C9 represents the cluster having information regarding ०, १ numeral and so on. Suppose if input get matched with the data of the cluster then for that cluster value will be 1 and for rest of the cluster value will be 0.

#### VI. CONCLUSION & FUTURE WORK

In this proposed work, we tried to recognize Devanagari numerals using RBF which have its own advantages and disadvantages over other systems. One of the advantages of RBF, is that the hidden layer is easier to interpret in RBF than the hidden layer in MLP. Training a RBF network is faster but once training is finished and it is being used, it is slower than a

MLP, so where speed is a factor a MLP may be more appropriate.

In future, we would like to implement this proposed work using MATLAB.

#### REFERENCES

- [1] "Handwritten Character Recognition using Neural Network", Sunith Bandaru.
- [2] "Handwritten English Character Recognition Using Neural Networks", Anita Pal and Dayashanker Singh, vol 1, no 2, July-December 2010.
- [3] "Printed Chinese Character Recognition", thesis by Yuan Lui.
- [4] "Devanagari numeral recognition by combining decision of multiple connectionist classifier" Reena Bajaj, Lipika Dey and Santanu Chaudhary, Sadhana Vol-27, Part-I, February 2002.
- [5] "SVM for handwritten Devnagri Numeral Recognition", Shailendra Kr. Srivastava and Sanjay GHarde, vol 07.
- [6] .Christopher M. Bishop, "Pattern Recognition and Machine learning", Springer Publication, Singapore, 2006.
- [7] "Introduction to Machine Learning" by Nils. J. Nilsson, 1997.
- [8] G.G. Rajput, S.M. Mali "Fourier Descriptor based Isolated Marathi Handwritten Numeral Recognition", vol 3-No.4, June 2010
- [9] "Input Fuzzy Modeling for Recognition of Handwritten Hindi Numerals", M. Hanmadhu, J. Grover, V.K. Madasu, S. Vasikarla.
- [10] Recognition of unconstrained on-line Devanagari characters by "Scott D. Connell, R.M.K. Sinha and Anil.K Jain.
- [11] "Neural Networks", Simon Haykins.
- [12] "Off-line handwritten characters recognition using Radial Basis Function", J. Ashok and Dr. E.G. Rajan, vol 2, issue 04.
- [13] "Handwritten Digits Recognition" by Gaurav Jain and Jason Ko.
- [14] "Statistical Pattern Recognition: A Review", Anil K. Jain, IEEE, Robert and Jianchang, vol 22 no.1, 2000.
- [15] "Fast and robust scheme for recognition of handwritten Devanagari Numerals", Vasantha, Ritu Jain and Patvardhan, NSC 2008.
- [16] Improving the performance of Radial Basis Function by learning center location", Dietrich and Thomas.
- [17] Pattern Recognition using K-nearest neighbours", Seiji, Kiyasu and Miyahara.
- [18] Weighted Radial Basis Function for improved Pattern Recognition and signal processing", Leonardo M. Reyneri.
- [19] Adrian G. Bors, "Introduction of Radial Basis Function".
- [20] R. M. K. Sinha and H. N. Mahabala, "Machine recognition of Devanagari script", IEEE Transactions on Systems, Man and Cybernetics.
- [21] Mohanty, S., Nandini, H., & Bebartta, D. (2011). Performance Comparison of SVM and K-NN for Oriya Character Recognition. International Journal of Advanced Computer Science and Applications - IJACSA, (Special Issue), 112-116.
- [22] Jipeng, T. (2011). Skew correction for Chinese character using Hough transform. International Journal of Advanced Computer Science and Applications - IJACSA, (Special Issue), 45-48.