Asian Journal of Information Technology 14 (1): 11-22, 2015
ISSN: 1682-3915

Adaptive Fuzzy Membership Functions with AGA Techniques in Telugu Character Recognition

[1]P.K. Venkateswar Lal, [2]Ande Prasad and [1]M. Venkat Rao
[1]Department of CSE, Vignan University, Vadlamudi, Guntur, India
[2]Department of CSE, Vikrama Simhapuri University, Nellore, India

Abstract: In the proposed Telugu character recognition technique, the given Telugu handwritten document is processed by normalizing the document and removing the noise. Then, skew detection and correction process is carried out by using the bilinear interpolation method to get more accurate result. Thus, the de-skewed documents text lines and characters are segmented by exploiting the Adaptive Histogram Equalization (AHE). In the next stage, the features of the segmented characters are extracted with the help of the Zoning Method. In Zoning Method, an adaptive fuzzy membership function will be developed by the AGA (Adaptive Genetic Algorithm). By using AGA in Zoning Method the features are extracted from the segmented characters and that extracted features is given to the FFBNN (Feed Forward Back Propagation Neural Network) for accomplishing the training process. During testing, more number of handwritten segmented Telugu characters will be given to the well trained FFBNN to check whether the input character is recognized or not. Thus, our proposed method has given more accurate recognition results by using our proposed adaptive fuzzy membership function with AGA Method. The proposed method performance is evaluated by getting more number of handwritten Telugu documents and compared with the GA-FFBNN and FFBNN.

Key words: Adaptive Histogram Equalization (AHE), Feed Forward Back Propagation Neural Network (FFBNN), Adaptive Genetic Algorithm (AGA), zoning, bilinear interpolation

INTRODUCTION

Pattern recognition is the study of how machines can observe the environment, learn to distinguish patterns of interest from their background and make sound and reasonable decisions about the categories of the patterns (Basu *et al*., 2010). In spite of almost 50 years of research, design of a general pattern recognizer remains an elusive goal (Lakshmi *et al*., 2006). Pattern recognition can be defined as a process which leads to a decision. The quality of this decision can only be measured by statistic relating to the number of “good” and “bad” classifications (Abed *et al*., 2010). Pattern recognition encompasses two fundamental tasks: description and classification (Kaur and Kaur, 2013). The design of a pattern recognition system essentially involves the following three aspects: data acquisition and preprocessing, data representation and decision making (Basu *et al*., 2010). The four approaches for pattern recognition are: template matching, statistical classification, syntactic or structural matching and neural networks (Kaur and Kaur, 2013).

The essential problem of pattern recognition is to identify an object as belonging to a particular group (Cheng and Cheng, 2009). Pattern recognition is a subject researching object description and classification method, it is also a collection of mathematical, statistical, heuristic and inductive techniques of fundamental role in executing the tasks like human being on computers (Liu *et al*., 2006). Methods of Pattern Recognition are Statistical Model, Structural Model, Neural Network Based Model (Jomy *et al*., 2011), Fuzzy Based Model and Hybrid Model (Singh, 2013).

Character recognition is one of the most fundamental topics in the context of pattern recognition and is included as a key issue in the recognition (Swethalakshmi *et al*., 2006). Most character recognition procedures can be visualized as consisting of three steps which use: the pre-processor, feature extractor and recognizer. The following are some of the applications of character recognition Signature Verification Writer Identification in Examination Assessment as a Mark Sheet Reader (Sitamahalakshmi *et al*., 2010). To recognize characters we need to divide the document into

Corresponding Author: P.K. Venkateswar Lal, Department of CSE, Vignan University, Vadlamudi, Guntur, India

classifiable objects (Mithe *et al*., 2013). In a simple form, these classifiable objects can be directly inputted in the form of isolated characters to classify (Kumar and Rao, 2013a, b).

To recognize the sentences or paragraphs, consideration of segmentation at word, line and character level and up to the level to decompose into classifiable objects is required (Sharma and Singh, 2013). In context to most Indian scripts, many researchers have proposed the segmentation into three horizontal zones where upper and lower zones consist of modifiers and middle zone is framed by basic character set (Siddharth *et al*., 2011).

The main objective of character recognition is to interpret input as a sequence of characters from an already existing set of characters (Dongre and Mankar, 2010). Character recognition can be applied on type-written, printed or handwritten text. The character recognition for handwritten characters is more complex due to varying writing styles of people (Siddharth *et al*., 2011). Telugu is one of the popular languages of India that is spoken by >66 million people especially in South India (Sastry *et al*., 2010). Telugu character recognition a feature based approach where features of all the basic symbols in several different fonts and sizes are stored and symbols of same fonts but different sizes are recognized on the basis of these features (Rao *et al*., 2013). The features used by them are the local gradients at various pixels called the radial direction features (Lakshmi *et al*., 2006). Character recognition system can contribute tremendously to the advancement of the automation process (Sastry *et al*., 2010) and can improve the interaction between man machine in many applications including office automation, check verification and large variety of banking, business and data entry applications (Abed *et al*., 2010).

RECENT RELATED RESEARCHES: A REVIEW

Desai (2010) has proposed an Optical Character Recognition (OCR) System for handwritten Gujarati numbers. We may find so much of work on Indian languages like Hindi, Kannada, Malayalam, Tamil, Bangala, Gurumukhi, etc. but Gujarati was a language for which hardly any work was traceable especially for handwritten characters. In that work a neural network was proposed for Gujarati handwritten digits identification. Multi-layered feed forward neural network was suggested for classification of digits. The features of Gujarati digits were abstracted by four different profiles of digits. Thinning and skew-correction were also done for preprocessing of handwritten numerals before their classification.

Pirlo and Impedovo (2011) have presented a class of membership functions which were called Fuzzy-Membership Functions (FMFs) for zoning-based classification. Those FMFs can be easily adapted to the specific characteristics of a classification problem in order to maximize classification performance. In that research, a real-coded genetic algorithm was presented to find in a single optimization procedure, the optimal FMF, organized with the optimal zoning described by Voronoi tessellation. The experimental results which were carried out in the field of handwritten digit and character recognition, indicate that optimal FMF performs better than other membership functions built on abstract-level, ranked-level and measurement-level weighting models which can be found in the works.

John *et al*. (2012) have presented the application of wavelet processing in the domain of handwritten character recognition. To achieve high recognition rate, vigorous feature extractors and powerful classifiers that were invariant to degree of variability of human writing were needed. The proposed scheme consists of two stages: a feature extraction stage which was based on Haar wavelet transform and a classification stage that uses support vector machine classifier. Experimental results have been shown that the proposed method is effective.

Pirlo and Impedovo (2012) have proposed a new class of zone-based membership functions with adaptive capabilities is introduced and its effectiveness is shown. The basic idea was to select for each zone of the Zoning Method, the membership function greatest suited to exploit the characteristics of the feature distribution of that zone. Along with this, a Genetic algorithm was proposed to determine in a unique process the most favorable membership functions along with the optimal zoning topology, defined by Voronoi tessellation. The experimental tests have shown the superiority of the new technique with respect to traditional Zoning Methods.

Soman *et al*. (2013) have proposed a pattern analysis technique to develop a powerful and efficient system for handwritten character recognition. The four techniques were: Convolution Neural Networks (CNN), Principal Component Analysis (PCA), Support vector machines, Multi classifier systems. The proposed system that embodies the above-mentioned four techniques was used for recognition of offline handwritten Telugu characters. Telugu aksharas of Consonant-Vowel (CV) type with 36 consonant classes and 15 vowel modifier classes were used for the study. Telugu dataset consisted of 47428 CV images in the training set and 5156 CV images in the test set.

Choudharya *et al.* (2013) have presented a pattern classifier technique which is used to extract the features. The main focus of that work was to extract features obtained by binarization technique for recognition of handwritten characters of the English language. Recognition of handwritten character images have been done by using multi-layered feed forward artificial neural network as a classifier. Some preprocessing techniques such as thinning foreground and background noise removal, cropping and size normalization, etc. were also employed to preprocess the character images before their classification. Very promising results were achieved when binarization features and the multilayer feed forward neural network classifier was used to recognize the off-line cursive handwritten characters.

Shao *et al.* (2013) have proposed a fast Self-Generation Voting Method for further improving the performance in handwritten Chinese character recognition. In their method, firstly, a set of samples were generated by the proposed fast Self-Generation Method and then these samples were classified by the baseline classifier and the final recognition result was determined by voting from these classification results. Two methods that are normalization-cooperated feature extraction strategy and an approximated line density were used for speeding up the Self-Generation Method. They evaluated the proposed method on the CASIA and CASIA-HWDB1.1 databases. High recognition rate of 98.84% on the CASIA database and 91.17% on the CASIA-HWDB1.1 database are obtained. These results have demonstrated that the proposed method outperforms the state of the art methods and is useful for practical applications.

Wakahara and Yamashita (2014) have proposed a GAT Correlation Method to reduce the computational cost of matching in k-NN classification. They accelerate GAT Correlation Method by reformulating its computational model and adopting efficient lookup tables. Recognition experiments performed on the IPTP CDROM1B handwritten numerical database have shown that the matching techniques of the simple correlation, the tangent distance and the accelerated GAT correlation achieved recognition rates of 97.07, 97.50 and 98.70%, respectively. The computation time ratios of the tangent distance and the accelerated GAT correlation to the simple correlation are 26.3 and 36.5-1.0, respectively.

Tao *et al.* (2014) have proposed the kernel version of DLA, Kernel Discriminative Locality Alignment (KDLA) and carefully prove that learning KDLA is equal to conducting Kernel Principal Component Analysis (KPCA) followed by DLA. This theoretical investigation can be utilized to better understand KDLA, i.e., the subspace spanned by KDLA is essentially the subspace spanned by DLA on the principal components of KPCA. Experimental results demonstrated that DLA and KDLA are more effective than representative discriminative information extraction algorithms in terms of recognition accuracy.

PROPOSED TELUGU CHARACTER RECOGNITION SYSTEM

In the proposed technology, initially the text documents obtained from the database are preprocessed in order to remove the noise which in turn helps to improve the accuracy of the recognition. After that bilinear interpolation is applied on the preprocessed documents for skew detection and correction. Then, each character is segmented from the preprocessed documents by using AHE technique. Subsequently, features from the segmented characters are extracted by exploiting the voronoi based zoning approach. At the time of zoning, the adaptive fuzzy membership function is optimized with the utilization of Adaptive Genetic Algorithm (AGA) and thus optimal features are obtained from each character. These optimal features are then fed to FFBNN to attain the training process. At the testing time, more number of characters is utilized to analyze the performance of the FFBNN. Performance of the proposed technique is compared to other techniques to show the efficacy of the proposed technique. Architecture of the proposed Telugu Character Recognition System is given in Fig. 1.

The proposed Telugu Character Recognition System consists of five stages namely, preprocessing, skew detection and correction, segmentation, feature extraction and recognition. Let the Database (D) consists of many text documents and let $x_{i,j}$ be one of the document x at location (i, j) taken from the database.

Preprocessing: Preprocessing process ensures to attain high segmentation accuracy. The input handwritten document image is named as d(mxn) and given to the preprocessing process. Let us consider the Database (D) which consists of document images {i = 1, 2, ..., N} where i represents the number of document images and each image has the size of M×N represents the row and column of the image. The RGB document image is first converted into grayscale image and it is named as gD_i. Next, this gD_i image is given to the binarization process. Binarization is a technique where the gray scale images are converted

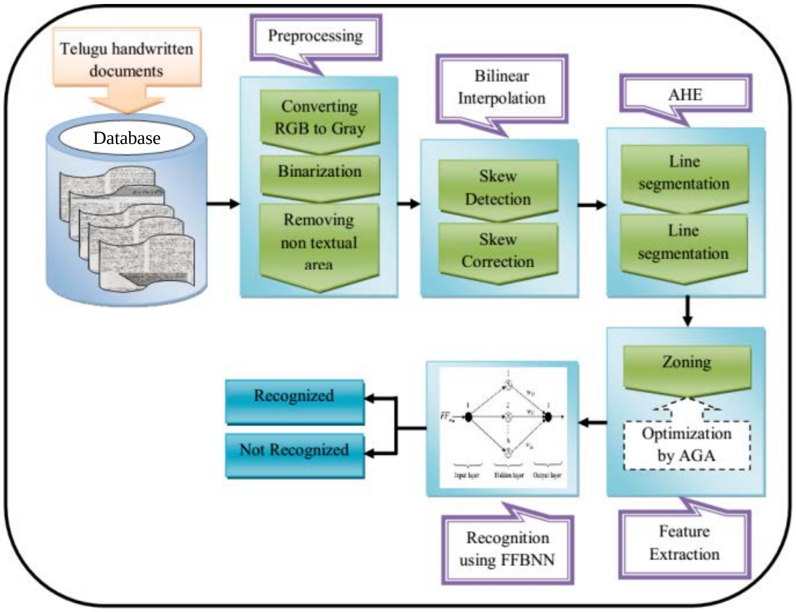

Fig. 1: Architecture of our proposed Telugu Character Recognition System

into binary images. Binarization separates the foreground (text) and background information. The process of binarization is stated as follows:

$$BD_i = \begin{cases} 1 & gD_i > \tau \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Where:
Bd_i = The binary image
τ = A threshold value

If the image gD_i intensity value is greater than the given threshold value means we change the pixels values into 1, otherwise we change into 0. Afterward, the non-textual area is removed from the binary image BD_i by traversing the image in four directions, top, bottom, left and right. During the traversal, the white pixel areas are removed in all directions. Finally, we get the textual area image TgD_i which having the textual part of the document image D_i. The textual image TgD_i is mapped to the original image D_i and the resultant image from the preprocessing process is denoted as:

$$\begin{gathered} PD_i \text{ where } i = 1, 2, ..., D: D \in M' \times N' \\ \text{where } M' < M \,\&\, N' < N \end{gathered} \quad (2)$$

The preprocessed image (PD_i) is then subjected to skew detection and correction process.

Bilinear interpolation: An important part of any document recognition system is detection of skew in the image of a page. Text lines in a document are generally parallel to one another in the horizontal direction and the space between every neighboring text lines is relatively constant. Since, scanning every pixel in the whole document image is time-consuming, it is appropriate to select a suitable sub-region in calculating the text line direction that corresponds to the skew angle of the image. After finding skew angle, the document image can be corrected according to the detected skew angle.

Here, Bilinear Interpolation (BI) is utilized to detect the skew angle (θ) of the document image and also correcting the document image according to the skew angle (θ). In mathematics, bilinear interpolation is an extension of linear interpolation for interpolating functions of two variables on a regular 2D grid. The key idea is to perform linear interpolation first in one direction and then again in the other direction. Although, each step is linear in the sampled values and in the position, the interpolation as a whole is not linear but rather quadratic in the sample location. The stepwise procedure of the bilinear interpolation technique (Tao *et al.*, 2014) is given:

Step 1: Get the preprocessed image (PD_i) as the input image.

Step 2: Resize the row size ($M' \rightarrow H$) and the column size ($N' \rightarrow H$) of the image (PD_i) as equal and denote the resized image as RD_i.

Step 3: Enlarge the resized image (RD_i) from ($H \times H$) to ($H' \times H'$) and denote the enlarged image as ED_i.

Step 4: Imagine scaling the $H' \times H'$ grid to fit over $H \times H$.

Step 5: For each and every pixel in ED_i do:

Step 6: Take a pixel (x', y') in enlarged image (ED_i).

Step 7: Find the scaled coordinates of the grid point using equation:

$$x = \frac{x'(H-1)}{(H'-1)} \quad (3)$$

$$y = \frac{\hat{y}(H-1)}{(H'-1)} \quad (4)$$

The range of the scaled coordinates (x, y) are lies between the size of the resized image (RD_i).

Step 8: Perform the floor and ceil operations on the obtained scaled coordinates (x, y) to get the four coordinates such as (x_1, y_2), (x_2, y_1) and (x_2, y_2).

Step 9: Find the intensity values (Intsty) of the four neighbours such as (x_1, y_1), (x_1, y_2), (x_2, y_1) and (x_2, y_2).

Step 10: Construct the matrix as:

$$\begin{bmatrix} x_1 & y_1 & x_1y_1 & 1 \\ x_1 & y_2 & x_1y_2 & 1 \\ x_2 & y_1 & x_2y_1 & 1 \\ x_2 & y_2 & x_2y_2 & 1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = \begin{bmatrix} Insty(x_1y_1) \\ Insty(x_1y_2) \\ Insty(x_2y_1) \\ Insty(x_2y_2) \end{bmatrix} \quad (5)$$

Step 11: Find out the values of A-D by solving step 10 and substitute it in the below equation:

$$BI(x, y) = Ax + By + Cxy + D \quad (6)$$

Step 12: Replace $(\hat{x}, \hat{y})$ by using the value BI(x, y) which is obtained from the above equation. Finally, the skew corrected image (SkD_i, $D \in M' \times N'$) is obtained and it is passed to carry out the segmentation process.

Character segmentation using adaptive histogram equalization: In order to recognize each character, the character should be segmented from the document image (SkD_i) and so AHE is used here to segment the character from the document image (SkD_i). AHE is an extension to traditional histogram equalization technique. It enhances the contrast of images by transforming the values in the intensity document. Unlike histogram equalization, it operates on small data regions, rather than the entire document. Each region's contrast is enhanced, so that the histogram of the output region approximately matches the specified histogram. The neighboring regions are then combined using bilinear interpolation in order to eliminate artificially induced boundaries. AHE is used to enhance the contrast of the document x' by modifying the pixel based on its surrounding pixels. AHE is routine, locally adaptive and habitually produces superior document. Let us consider a moving window w (w = OxO) and I is the intensity of pixel (i, j). Then, the modification of pixel (i, j) is given as:

$$\begin{aligned} Map(I) = {} & p[q \times map_{\_,\_}(I) + (1-q)\, map_{+,\_}(I)] + \\ & [1-p]\,[q \times map_{\_,+}(I) + (1-p)\, map_{+,+}(I)] \end{aligned} \quad (7)$$

Where:

$map_{+,\_}$ = Mapping of right upper $(i_+, \_)$
$map_{+,+}$ = Mapping of right lower $(i_+, {}_+)$
$map_{\_,+}$ = Mapping of left upper $(i_\_, {}_+)$
$map_{\_,\_}$ = Mapping of left lower $(i_\_, \_)$

$$p = \frac{(j - j_\_)}{(j_+ - j_\_)} \quad (8)$$

$$q = \frac{(i - i_\_)}{(i_+ - i_\_)} \quad (9)$$

This modification is done for all the pixels in the entire document (SkD_i) and finally the segmented characters (sc_i, i = 1, 2, ..., d where $sc_i \in SkD_i$) are obtained.

Optimal feature extraction with the help of voronoi based zoning: The segmented characters (sc) are then subjected to feature extraction process. To extract the features, the segmented characters (sc) are zoned by using Voronoi based technique. In order to get the optimal features (ff_i), the membership function is optimized by using AGA. Zoning is nothing but dividing the given input image into number of sub images that can give information related to a particular part of the pattern. Zoning is of two types:

- Static: obtained by superimposing the regular m×n grids on a pattern image
- Dynamic: the design of zoning is considered as an optimization problem

Steps to zone the image:

- Define the number of points $\{Z_i = Z_1, Z_2, ..., Z_n\}$ to be divide the image
- Obtain the size of the image (sc_i: u×v)
- Find the difference (d_i) between the number of points generated $\{Z_j = Z_1, Z_2, ..., Z_n\}$ and the size of the image (sc_i)
- Get the minimum difference for each number of points generated $\{Z_j = Z_1, Z_2, ..., Z_n\}$
- Sort the points $\{Z_j = Z_1, Z_2, ..., Z_n\}$ in descending order

By using these above steps, the image (sc_i) is partitioned into number of zones $\{Z_j = Z_1, Z_2, ..., Z_n\}$ based on the number of points generated already.

The zones $\{Z_j = Z_1, Z_2, ..., Z_n\}$ obtained in the above steps are utilized to calculate the features (ff_i) of the image (sc_i). After obtaining the zones $\{Z_j = Z_1, Z_2, ..., Z_n\}$, the center points of each zone are calculated and are represented as the Voronoi points ($\{V_j = v_1, v_2, ..., v_n\}$). After the calculation of the Voronoi points ($\{V_j = v_1, v_2, ..., v_n\}$), feature Instances ($fI_i$) are find out with the help of the Feature Set (FS_i) from each zone (Z_j). Then, the Euclidean Distance (ED_{ij}) between the location of the feature Instance (fI_i) and the Voronoi points are calculated and based on that distance ranked index sequence (R) is created. The ranked index sequence (R) consists the zones $\{Z_j = Z_1, Z_2, ..., Z_n\}$ which are arranged based in the ascending order of the Euclidean Distance (ED_{ij}) calculated. Finally, weight (WE = $\{\omega_1, \omega_2, ..., \omega_{ij}$) of each zone ($Z_j$) is assigned by using the adaptive membership function (ω_{ij}) (https://en.wikipedia.org/wiki/Sensitivity_and_specificity) where ω_{ij} is the weight of the zone (Z_j) at the corresponding feature Instance (fI_i). The membership function used here is:

$$\omega_{ij} = e^{-\gamma ED_{ij}} \tag{10}$$

where, γ is a positive parameter which denotes the falling rate. After the calculation the membership function, the range of the membership functions (ω_{ij}) are fuzzified between 0 and 1. Thus, the adaptive membership function becomes adaptive fuzzy membership function. Finally, features (ff_i) of each character image (sc_i) is obtained as follows:

$$FF_{sc_i} = \begin{bmatrix} FF_{sc_i}(1,1) & FF_{sc_i}(1,2) & ... & FF_{sc_i}(1,n) \\ FF_{sc_i}(2,1) & FF_{sc_i}(2,2) & ... & FF_{sc_i}(2,n) \\ .... & .... & ... & .... \\ FF_{sc_i}(S,1) & FF_{sc_i}(S,2) & ... & FF_{sc_i}(S,n) \end{bmatrix} \tag{11}$$

In Eq. 11, FF_{sci} represents the feature matrix of each character (sc_i) and FF_{sci} denotes the weight obtained using the adaptive fuzzy membership function of sth number of the feature in the feature set at nth zone. In order to get the optimal features, the falling rate (γ) in the adaptive fuzzy membership function is optimized using the well-known optimization algorithm called Adaptive Genetic algorithm.

As the convergence rate of the conventional GA is low, AGA is utilized in our proposed technique to speed up the convergence rate. The adaptive GA is employed with the help of Cauchy mutation in the mutation operator. Cauchy mutation is the mutation operator introduced in the Genetic algorithm to speed up GA process and also to enhance the GA performance. Here, the falling rate (γ) of the adaptive fuzzy membership function is optimized using AGA. Thus, the GA with Cauchy mutation operator will produce the optimal falling rate and thus the optimal features will be obtained. The GA with Cauchy mutation operation is shown below. The parameters utilized to control the interest point detection process are optimized using the well-known optimization algorithm called Adaptive Genetic Algorithm (GA) (Fig. 2) which is a metaheuristic algorithm that minimizes the natural evolution process. Most probably, it used to spawn elucidations to optimization and search problems.

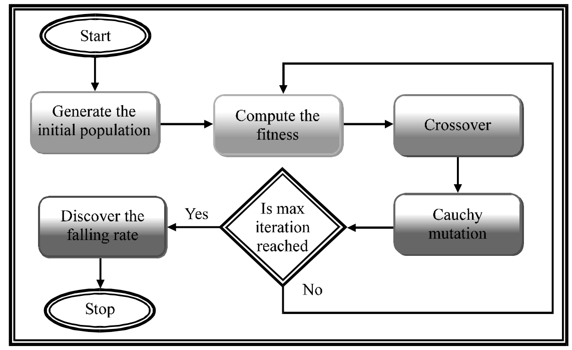

Fig. 2: Flowchart of the Adaptive Genetic algorithm

It spawns elucidations to optimization problems using techniques stirred by natural evolution such as inheritance, mutation, selection and crossover.

Initial phase: Initially the populations of the chromosomes x_i (i = 1, 2, ..., T) are generated randomly. T denotes the size of the population. The chromosome (x_i) contain the value of the falling rate (γ_i).

Fitness function: Fitness value of each parameter is calculated and the chromosome which has the highest fitness value is selected as the best chromosome:

$$F(i, j) = \min(er) \tag{12}$$

In Eq. 12, er(i, j) is the error rate of tth parameter.

Selection of chromosomes: One or more parent chromosomes are selected based on the 'T/2' best chromosomes which have minimum fitness and new solution is created.

Crossover: Single point crossover is performed at the crossover rate of (C_r) and hence ('T/2') offspring are obtained. In every crossover operation (TC_r) genes are exchanged between corresponding parents.

Mutation: Individuals are perturbed probabilistically to bring a change in the individuals. Using mutation operator, there is a probability that some new features might appear due to change in the chromosome. Cauchy mutation is used to mutate the individuals according to the equation given below. Mutation is performed on the basis of pre-determined mutating probability. In case the Cauchy mutation is applied, random variable 'r' is a Cauchy distribution. The Cauchy distribution function is defined as:

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\arctan(x) \tag{13}$$

Updation: In the 6th step initial chromosome replaced by new chromosome. Next to the mutation process the initial population chromosomes are replaced by ('T/2') elected and new 'T/2' offspring chromosomes.

Termination criteria: The process is continued until it meets the termination criteria. Thus the optimal falling rate obtained by using the AGA is substituted in Eq. 11 to get the optimal features.

The optimal features achieved in the above process are given to train the Feed Forward Neural Network (FFBNN) to recognize the character. The neural network consists of n number of input units, h hidden units and one output unit. Back propagation is used as the training algorithm. The structure of the FFBNN is given in Fig. 3:

- For all the neurons, allot weights arbitrarily except for input neurons
- The bias function and activation function for the neural network is described

$$x(t) = \beta + \sum_{n=1}^{H}\left(w_{tn} FF_{sc_{tn}}\right) \tag{14}$$

$$x(a) = \frac{1}{1+e^{-x(t)}} \tag{15}$$

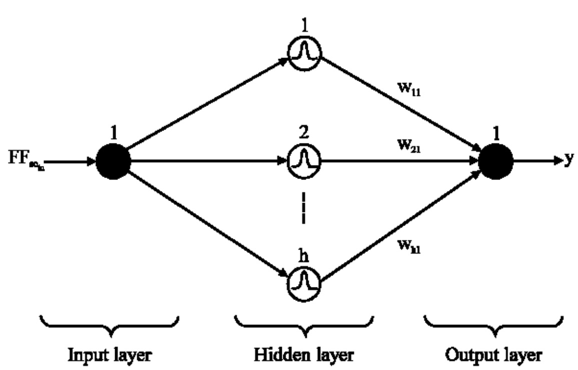

Fig. 3: Diagram of the FFBNN

In bias function, $FF_{sc_{tn}}$ is the input extracted from the character. The activation function for the output layer is given in Eq. 15.

- Get the learning error:

$$E_r = \frac{1}{h}\sum_{n=0}^{h-1} de_n - ac_n \tag{16}$$

Where:
E_r = The FFBNN network output
de_n and ac_n = The desired and actual outputs
h = Total number of neurons in the hidden layer

Weights are allocated to the hidden layer and output layer neurons by randomly chosen weights. The input layer neurons have a constant weight:

1. Determine the bias function and the activation function
2. Calculate BP error for each node and update the weights as follows:

$$w_{(tn)} = w_{(tn)} + \Delta w_{(tn)} \tag{17}$$

$\Delta w_{(tn)}$ is obtained as:

$$\Delta w_{(tn)} = \delta \cdot X_{(tn)} \cdot Be \tag{18}$$

Where:
δ = The learning rate which normally ranges from 0.2-0.5
Be = The back propagation error

3. Then, repeat the steps 2 and 3 until the BP error gets minimized. The process is repeated until it satisfies Be<0.1
4. The error gets minimized to a minimum value the FFBNN is well trained for performing the testing phase

Then, the result of the neural network (y) is compared with the threshold value (τ). If it satisfies the threshold value it denotes that the given character is recognized:

$$\text{Result} = \begin{cases} \text{Recognized} & y \geq \tau \\ \text{Not recognized} & y < \tau \end{cases}$$

EXPERIMENTAL RESULTS

Our proposed Telugu Character Recognition System is done with the help of FFBNN. In order to get the optimal features, the falling rate of the adaptive fuzzy membership function is optimized by using AGA. Then, the obtained features are given to the FFBNN to achieve the training process. In the testing process, more handwritten characters are given to the well trained FFBNN to validate the performance of the proposed technique.

The performance of the proposed Telugu Character Recognition System is evaluated by using more number of handwritten document images and the proposed technique's performance is compared with the GA-FFBNN and FFBNN. The proposed technique is implemented in Matlab Platform.

By applying the statistical measures which is specified by (http://csserver.evansville.edu/~richardson/courses/EE499_Image_Processing/resources/lectures/l05/BilinearInterpolationExample.pdf), the concert of our suggested Telugu Character Recognition System is examined. The database consists of 12 handwritten documents which are collected manually. The statistical measures TP, TN, FP and FN values of our suggested system are specified in Table 1. Figure 4-8 illustrate the sample of input handwritten document, preprocessed image, skew corrected image, line segmented image and segmented characters, respectively.

Accuracy, sensitivity, specificity, FAR (False Rejection Rate) and FRR (False Acceptance Ratio) measures are specified in Table 1. At this point, performance of our suggested Telugu Character Recognition System is compared with the GA-FFBNN and FFBNN. The accuracy of the suggested method is 94.08%. While the other techniques such as GA-FFBNN and FFBNN offers 89.3491 and 89.9408 accuracy correspondingly. When compared to the existing systems, the suggested AGA-FFBNN provides higher rate of accuracy. It symbolizes the suggested method identify the character more precisely than the other methods. The suggested method has 95.6522% of sensitivity. It is about 2.5-10% higher than the other methods. High percentage of sensitivity points out the good presentation of the

Table 1: Performance measures of proposed technique and other techniques

Measures	Proposed technique	GA-FFBNN	FFBNN
Accuracy	94.08	89.340	89.94
Sensitivity	95.65	92.500	85.71
Specificity	93.49	88.370	91.66
FAR	0.065	0.116	0.083
FRR	0.043	0.075	0.142

Fig. 4: Input handwritten document

Preprocessing image

Fig. 5: Preprocessed document

Interpolation image

Fig. 6: Skew corrected document

method. When seeing the other performance measurements such as FAR and FRR also the proposed technique yields our proposed method provides better performance results that is the rate of FAR and FRR are less than the FAR and FRR rate of GA-FFBNN and FFBNN techniques. Thus, by seeing the results, it is proved that the proposed technique recognize the Telugu characters efficiently.

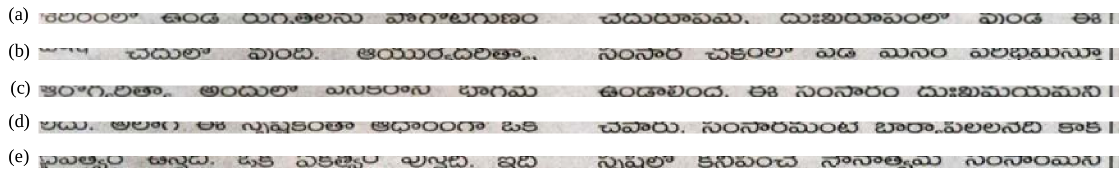

Fig. 7: a-e) Line segmented image

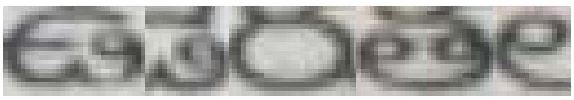

Fig. 8: Segmented characters

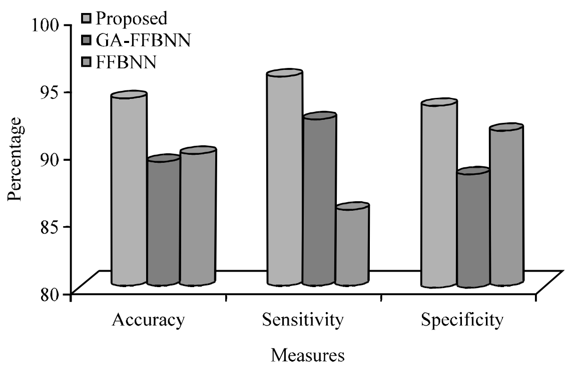

Fig. 9: Comparing the accuracy, sensitivity and specificity of our proposed technique (AGA-FFBNN) with other techniques such as GA-FFBNN and FFBNN

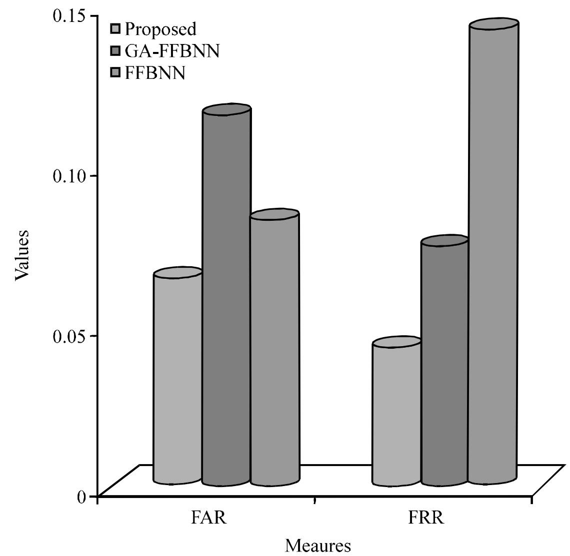

Fig. 10: Comparing the FAR and FRR of our proposed technique (AGA-FFBNN) with other techniques such as GA-FFBNN and FFBNN

Accuracy and sensitivity measures of the suggested method are compared with the other methods in Fig. 9. By seeing the graph, the precision of the suggested method is considerably higher than the other systems such as GA-FFBNN and FFBNN. Likewise, the sensitivity measure of the suggested AGA-FFBNN Method is moreover particularly higher than the other methods. It points out the presentation and the precision of the suggested method is higher than the other methods. Similarly, the specificity of the proposed technique is notably higher than the other techniques. Our suggested Telugu Character Recognition System based AGA-FFBNN has reached 94.08% of accuracy, 95.65% of sensitivity and 93.49% of specificity correspondingly. It points out that our suggested recognition system carried out better presentation when compared to the other techniques mentioned above. FAR and FRR measures of the suggested method are compared with the other methods in Fig. 10. FAR is obtained by subtracting the specificity value from 1. When looking the FAR value of the proposed technique (0.0650407), it is considerably smaller than the other techniques such as GA-FFBNN and FFBNN. But the GA-FFBNN and FFBNN yields the FAR rate 0.116279 and 0.0833333, respectively. And the FRR rate is attained by subtracting the sensitivity value from 1. On looking at the FRR value also, the proposed technique offers less rate of FRR than the other techniques. Thus, it indicates the proposed Telugu Character recognition technique performs efficiently with less error rate.

CONCLUSION

We have proposed a Telugu Character Recognition System based on FFBNN in this study. The suggested system was executed and more number of characters is utilized to examine the results of the suggested system. The presentation study confirmed that the suggested Telugu Character Recognition System in Telugu character recognition process presents an incredible rate of accuracy (94.08%), sensitivity (95.65%), specificity (93.49%), FAR (0.0650407) and FRR (0.0434783). The high value of these measures illustrates that our suggested technique more precisely identifies the character images from the specified test images. Based on FFBNN, the comparison result illustrates that our suggested Telugu Character Recognition System has specified high accuracy than existing methods. Hence our suggested Telugu Character Recognition System competently identifies the handwritten Telugu character imaged by applying the FFBNN technique.

REFERENCES

Abed, M.A., A.N. Ismail and Z.M. Hazi, 2010. Pattern recognition using genetic algorithm. Int. J. Comput. Electr. Eng., 2: 583-588.

Basu, J.K., D. Bhattacharyya and T.H. Kim, 2010. Use of artificial neural network in pattern recognition. Int. J. Software Eng. Applic., 4: 23-34.

Cheng, H.P. and C.S. Cheng, 2009. Control chart pattern recognition using wavelet analysis and neural networks. J. Qual., 16: 311-321.

Choudharya, A., R. Rishib and S. Ahlawat, 2013. Off-line handwritten character recognition using features extracted from binarization technique. AASRI Procedia, 4: 306-312.

Desai, A.A., 2010. Gujarati handwritten numeral optical character reorganization through neural network. Pattern Recogn., 43: 2582-2589.

Dongre, V.J. and V.H. Mankar, 2010. A review of research on devnagari character recognition. Int. J. Comput. Applic., 12: 8-15.

John, J., K.V. Pramod and K. Balakrishnan, 2012. Unconstrained handwritten malayalam character recognition using wavelet transform and support vector machine classifier. Procedia Eng., 30: 598-605.

Jomy, J., K.V. Pramod and B. Kannan, 2011. Handwritten character recognition of South Indian scripts: A review. Proceedings of the National Conference on Indian Language Computing, February 19-20, 2011, Kochi, pp: 1-6.

Kaur, N. and U. Kaur, 2013. Survey of pattern recognition methods. Int. J. Adv. Res. Comput. Sci. Software Eng., 3: 317-319.

Kumar, K.V. and R.R. Rao, 2013a. Improvement in efficiency of recognition of handwritten telugu script. Int. J. Inventive Eng. Sci., 2: 1-4.

Kumar, K.V. and R.R. Rao, 2013b. Online handwritten character recognition for telugu language using support vector machines. Int. J. Eng. Adv. Technol., 3: 189-192.

Lakshmi, C.V., R. Jain and C. Patvardhan, 2006. OCR of Printed Telugu Text with High Recognition Accuracies. In: Computer Vision, Graphics and Image Processing, Kalra, P.K. and S. Peleg (Eds.)., LNCS. Vol. 4338, Springer-Verlag, Berlin, Heidelberg, pp: 786-795.

Liu, J., J. Sun and S. Wang, 2006. Pattern recognition: An overview. Int. J. Comput. Sci. Network Secur., 6: 57-61.

Mithe, R., S. Indalkar and N. Divekar, 2013. Optical character recognition. Int. J. Recent Technol. Eng., 2: 72-75.

Pirlo, G. and D. Impedovo, 2011. Fuzzy-zoning-based classification for handwritten characters. IEEE Trans. Fuzzy Syst., 19: 780-785.

Pirlo, G. and D. Impedovo, 2012. Adaptive membership functions for handwritten character recognition by voronoi-based image zoning. IEEE Trans. Image Process., 21: 3827-3837.

Rao, K.R.M., B. Ramesh and G.I. Reddy, 2013. Font and size identification in telugu printed document. Int. J. Eng. Res. Dev., 6: 92-110.

Sastry, P.N., R. Krishnan and B.V.S. Ram, 2010. Classification and identification of Telugu handwritten characters extracted from palm leaves using decision tree approach. ARPN J. Eng. Applied Sci., 5: 22-32.

Shao, Y., C. Wang and B. Xiao, 2013. Fast self-generation voting for handwritten Chinese character recognition. Int. J. Document Anal. Recogn., 16: 413-424.

Sharma, P. and R. Singh, 2013. Survey and classification of character recognition system. Int. J. Eng. Trends Technol., 4: 316-318.

Siddharth, K.S., M. Jangid, R. Dhir and R. Rani, 2011. Handwritten gurmukhi character recognition using statistical and background directional distribution features. Int. J. Comput. Sci. Eng., 3: 2332-2345.

Singh, S., 2013. Optical character recognition techniques: A survey. J. Emerg. Trends Comput. Inform. Sci., 4: 545-550.

Sitamahalakshmi, T., A.V. Babu and M. Jagadeesh, 2010. Character recognition using dempster-shafer theory-combining different distance measurement methods. Int. J. Eng. Sci. Technol., 2: 1177-1184.

Soman, S.T., A. Nandigam and C.S. Chakravarthy, 2013. An efficient multiclassifier system based on convolutional neural network for offline handwritten Telugu character recognition. Proceedings of the National Conference on Communications, February 15-17, 2013, Indian Institute of Technology Delhi, New Delhi, India, pp: 1-5.

Swethalakshmi, H., A. Jayaraman, V.S. Chakravarthy and C.C. Sekhar, 2006. Online handwritten character recognition of devanagari and telugu characters using support vector machines. Proceedings of the 10th International Workshop on Frontiers in Handwriting Recognition, October 2006, La Baule, France, pp: 1-8.

Tao, D., L. Liang, L. Jin and Y. Gao, 2014. Similar handwritten Chinese character recognition by kernel discriminative locality alignment. Pattern Recogn. Lett., 35: 186-194.

Wakahara, T. and Y. Yamashita, 2014. k-NN classification of handwritten characters via accelerated GAT correlation. Pattern Recogn., 47: 994-1001.