Offline Handwriting Identification Using Adaptive Neural Fuzzy Inference System

Devi T, Swathi Priya N, Prabaharan J, Manigandan J
Department of Computer Science and Engineering, Rajarajeswari College of Engineering, Bangalore-74
deviyabe@gmail.com, swathinatarajan@gmail.com, prabaharan.jothi@gmail.com, jmaniannauniv@gmail.com

ABSTRACT - A new method for handwriting identification was presented. The chosen word from the text image (handwritten) which is given as input to the system is separated as individual characters. Then each of the separated characters are converted into 625 values from column vectors, and then fed into the adaptive neural fuzzy inference system (ANFIS) for membership functions (MF) calculation and normalized firing strength. In our paper we were used triangular membership function and compare with others MF.The networks has been designed with single layered neural network corresponding to a character from a-z, the outputs of all the column vector is fed into network the which has been developed using the concepts of correlation. The output is compared with genetic algorithm and results are approximate to the scale, so ANFIS network is efficiency.

Keywords—*Handwriting Identification, Adaptive Neuro Fuzzy Inference System (ANFIS), Column vector*

I INTRODUCTION

The identification of written characters can be done using Handwriting recognition. The problem can be viewed in such a way to identify the most appropriate character the given figure matches to. Recognition technique is used for offline character recognition refers to the where the final figure is given to us. Online character recognition systems are contrary to the where the data can be sampled while the character is being written. Operating in offline mode gives input in the complete picture character that we need to recognize. The recognition is usually in complexity which is related with the size of the language being considered. If the language contains more number of characters; difficulty arises in identification than the case when the language contains lesser number of characters. They always have an effect on handwriting recognition system. In this paper we propose the use of Column vectors for solving this problem. The basic idea of column vector comes from the fact that it can be used as an excellent means of combining various styles of writing a character and make way to new styles. Closely observing the capability of human mind in the recognition of handwriting, we find that humans are able to recognize characters even though they might be seeing that style for the first time. This is possible because of their power to visualize parts of the known styles into the unknown character. In this paper we try to depict the same power into the machines.

The problem increases when it operates offline mode. We see a lot of work has been done in this area in the past few years. The solutions being proposed mainly use Adaptive Neural Fuzzy Inference System (ANFIS) and Hidden Markov Models (HMM) for solving the problem. Column vectors have not been applied much. They have been applied for feature selection optimization. ANFIS Neural system involve training of the system with all the characters [3].Then when an unknown input is given to the system, the ANFIS Neural Network is able search and find probable character by generalization. Hence once trained, the system would be ready to recognize input which is unknown. Hidden Markov Model is a complete statistical model that tries to predict the unknown sequence. It also tries to recognize the unknown character which is given as input.

II COLUMN VECTOR

Column vectors are a family of computational models inspired by evolution. These algorithms encode a potential solution to a specific problem on a simple chromosome-like data structure and apply recombination operators to these structures as to preserve critical information.

Column vectors are a very good means of optimizations in such problems. They optimize the desired property by generating approximate solutions from the presently existing solutions. Column vectors are optimization heuristics. It is used to search for good solution to complex problem and good heuristics for combinatorial problems. Mutation and crossover may only applicable for bit string and integer representation. It generates hybrid solution from the existing solution. This solution may be better than the already existing solution.

A. General Procedure

Handwriting recognition is a famous problem like the recognition of whatever input which is given in the form of image, scanned paper, etc. The handwriting recognition generally involves the following steps [5].

B. Segmentation

This step deals with the breaking of the lines, words and finally getting all the separated characters and the identification of the boundaries of the character, separating them for further processing. We assume that this step is already done using this algorithm. Hence the input to our system is a single character.

C. Preprocessing

It involves the initial processing of the image, so that it can be used as an input for the recognition system. In this algorithm a part of preprocessing step has been done and (thin) character segment is made to a unit pixel thickness. Various algorithms may be used for this purpose. The further processing is done by our algorithm.

D. Recognition

Recognition which identifies the character, once the input image is available in good condition. It may be processed for recognition which identifies the character. This algorithm uses an image as an input for the same.

i= imread('d:/filename.bmp')

To store the input image in bitmap file format that input image then again stored into variable i. Then calculate the column and row values for corresponding image. These vector values converted into the single row, similarly calculate covariance of same single row vector values. Then calculate intensity value for that corresponding image. These two inputs are given to ANFIS neural networks.

i(:)

s=cov (i(:))

intensityValue = i(row,column)

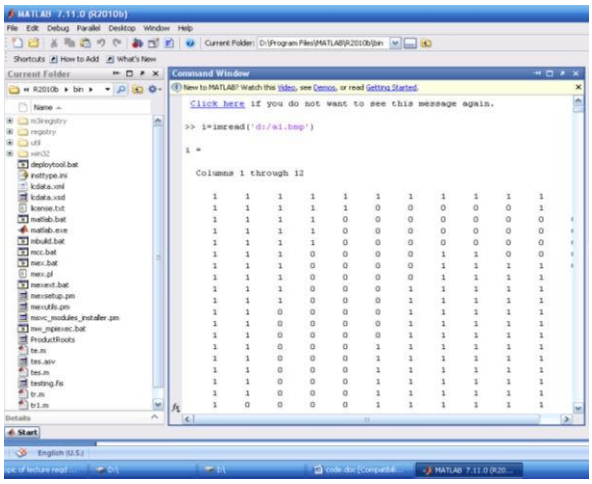

Fig. 1: Column Vector

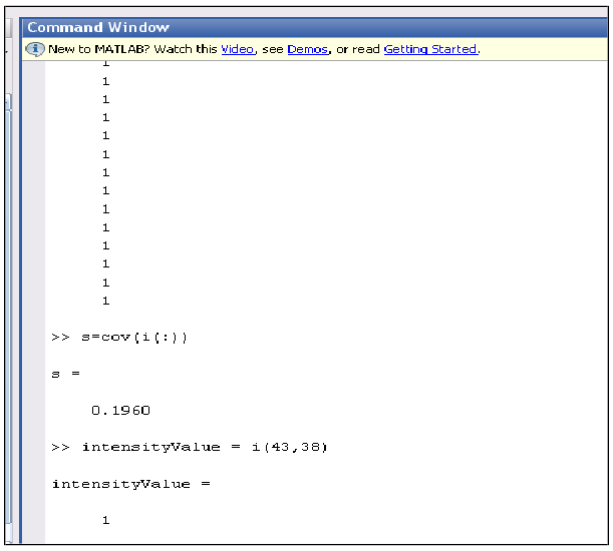

Fig. 2: Intensity Value

III ANFIS

ANFIS stands for Adaptive Neuro Fuzzy Inference System. ANFIS are a class of adaptive networks that are functionally equivalent to fuzzy inference systems. It integrates both neural network and fuzzy logic principles. ANFIS is a universal approximator. It's trained by a hybrid learning algorithm. Using a given input/output data set, the toolbox function ANFIS constructs a fuzzy inference system (FIS) whose membership function parameters are tuned.

The ANFIS algorithm combines the best features of both artificial neural network (ANN) and fuzzy logic.

The input data is trained here and the ANFIS toolbox is a five layered network. There are rules through which the nodes combine and give the output in a better way and compare with genetic algorithm. Fig1 shows the various layers of ANFIS and the combination of nodes.

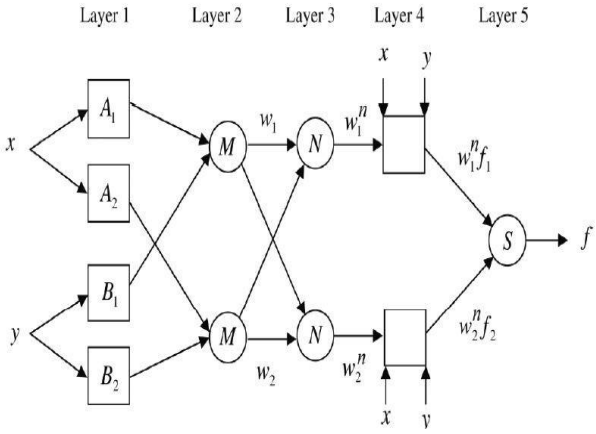

Fig. 3: ANFIS Architecture

In layer 2, we calculate membership function. Determine the function of F by using triangular fuzzy membership. The general expression of triangular fuzzy member function is,

$$f(x,a,b,c)=\max\left(\min\left(\frac{x-a}{b-a},\frac{c-x}{c-b}\right),0\right)$$ (equation 1)

Where a,b,c are parameter set.

IV ANFISGA

This method brings the technique of ANFIS and GA together. The dataset which is obtained by the segment characters for training and testing. A set of data is trained by ANFIS and this is done in ANFIS Graphical User Interface of MATLAB. Here we use hybrid optimization method and error tolerance is set to zero. The membership function used is triangular membership function. The data to be evaluated is fed into the optimization toolbox. The various operations of the column vector are performed and the fitness function is evaluated until the stopping criterion is met. Fitness of each chromosome is computed by minimum optimization method. The mean absolute percentage error (MAPE), also known as mean absolute percentage deviation (MAPD), is a measure of accuracy of a method for constructing fitted time series values in statistics, and is defined by the formula:

$$M=\frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{A_t-F_t}{A_t}\right|$$ (equation2)

Where A_t is the actual value and F_t is the forecast value.

The difference between A_t and F_t is divided by the actual value A_t again. The absolute value in this calculation is summed for every fitted or forecasted point in time and divided again by the number of fitted points n. multiplying by 100 makes it a percentage error.

Fig. 3 shows the operation of ANFIS and the Column vector operations and the output is the dosage level which is produced as per the given data.

The membership function used is triangular membership function. It is calculated as in the following equation1.

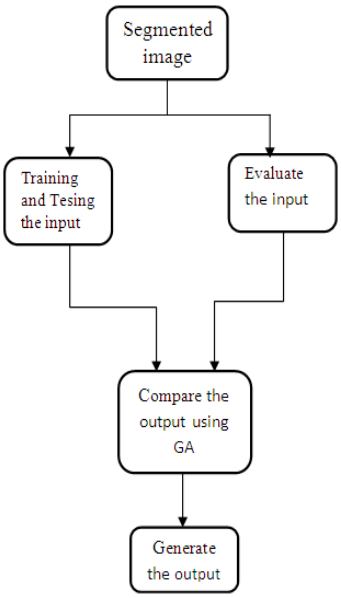

Fig. 4: ANFISGA data flow diagram

V RESULTS

The training and testing is done in ANFIS toolbox of MATLAB. The results show that this technique is much better than other statistical techniques and artificial intelligence techniques such as artificial neural network and fuzzy logic where the accuracy is low and a lot of risk is associated.

This model is implemented till the ANFIS and an accuracy of 93% is achieved which is much accurate when compared to the previous models. These results are stored and then evaluated by Column vector.

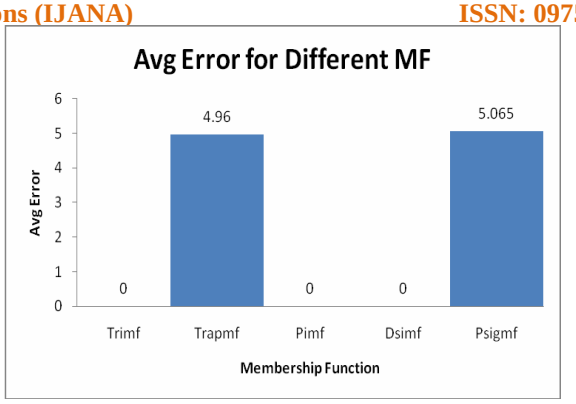

Fig. 5: Average error for different MF

VI CONCLUSION

In our project improve efficiency for handwriting recognition has been proposed. The proposed system is expected to produce better accuracy than the existing system. The result is efficient while comparing with genetic algorithm. We proposed the use of column vector for solving the problem of offline handwriting recognition. Input is taken in the form of images. Training data is taken to train the algorithm and it was already present in the database.

REFERENCES

[1] Rahul KALA1, Harsh VAZIRANI2, Anupam SHUKLA3 and Ritu TIWARI,‘ Offline Handwriting Recognition using Genetic Algorithm‘ IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 2, No 1, March 2010

[2] FLink, Gernot A and Plotz, Thomas, _Unsupervised Estimation of Writing Style Models for Improved Unconstrained Off-line Handwriting Recognition‘, International Workshop on Frontiers in Handwriting Recognition, Suvisoft, France, 2006

[3] Tripathy N. and Pal U,‘Handwriting segmentation of constrained Oriya text‘, Sadhna,2006.

[4] Som, Tanmoy and Saha, Sumit, _Handwritten character recognition by using Neural-network and Euclidean distance metric‘, Social Science Research Network, 2008.

[5] Liwicki, Marcus and Bunke, Horst, _Int. Journal of Pattern Recognition and Artificial Intelligence, 2007, 83-98.

[6] Graves, Alex; Fernandez, Santiago; Liwicki, Marcus; Bunke,Horst and Schmidhuber, Jurgen, _Unconstrained Online Handwriting Recognition with Recurrent Neural Networks‘,Advances in Neural Information Processing Systems 20,2008.