ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

Biometric system based on off-line signatures

Bhanu Priya Taneja[1], Navdeep Kaur[2]

Student, Department of CSE, (Rayat)Bahra Group of Institutes, Patiala, Punjab, India [1]

Assistant Professor, Department of CSE, (Rayat)Bahra Group of Institutes, Patiala, Punjab, India [2]

Abstract: This paper discusses the process of off line signature recognition and verification system using neural networks. The paper consists of two stages 1) Evaluating feature vectors of off line-signature images. 2) Training and Validation using back-propagation artificial neural network. Three layer feed forward MLP classifier have been used for pattern recognition process. The main advantage of feed forward neural network is that they are easy to use, and that they can approximate any input/output mapping with weights and thresholds (biases) of the model. Different stages of the process are explained in the paper by making use of proper algorithms. The overall accuracy achieved is 96.3%.

Keywords: Confusion matrix, Feed forward neural network, Kurtosis, Receiver operating characteristics (ROC).

I. INTRODUCTION

Offline signature recognition and verification system (SRVS) is the most widely acceptable form because handwriting of an individual differentiates him from other individual. It not only authenticates a person but also validates the documents. One does not need to use any traditional form of protection from forgeries because an individual's signature acts as a password in every authentication mode. The signature is a behavioural biometrics, and is therefore inherently dependent on the changing activity pattern of the signer and the signing process [1].

Signatures are composed of special character; therefore, usually they are not readable. The objective of signature recognition is to recognize the writer [2]. According to the available input, signature recognition and verification process is classified in two parts: Static signature recognition technique which is helpful in automatic signature recognition found on documents and Dynamic signature recognition technique that lays emphasis on dynamic properties of signature.

Offline signature recognition systems are more difficult than online recognition systems because the information like duration, flow, velocity is lost, in case of offline signatures [2]. But, on the other hand, the advantage is that they do not need any special gadgets to traverse them. Other advantages are: collection of signatures does not require special instruments, No special training is required for data collection.

Generally, the signature is a word with which a person identifies him/her, and as such will have a greater personal significance than any other words he /she writes [4]. Thus, it is an ideal method to secure your legal as well as personal documents.

This paper aims at giving the brief idea about the process to be carried out to achieve the desired aim. The paper discusses the entire implementation in detail covering the algorithm used, process of confusion matrix, receiver operating characteristics (ROC) and neural networks.

II. PROPOSED METHODOLOGY

We have divided the proposed work into two sections-

Section 1: Evaluating feature vectors of off line-signature images.

Section 2: Training and Validation using feed forward artificial neural network.

A. *Evaluating feature vectors of off line-signature images*

1) Process 1: Binarization and Thinning of signature image:

The graphic data taken from an image scanner and processed by the computer need to be pre-processed first. We used im2bw command in mat lab to achieve binarization.

Algorithm used for thinning process:

1. In the first sub iteration, delete pixel p if and only if the conditions G_1, G_2, and G_3 are all satisfied.

2. In the second sub iteration, delete pixel p if and only if the conditions G_1, G_2, and G_3' are all satisfied.

Condition G1:

$X_H(p)=1$

where

$$X_H(p) = \sum_{i=1}^{4} b_i$$

$$b_i = \begin{cases} 1, \text{ if } x_{2i-1} = 0 \text{ and } (x_{2i} = 1 \text{ or } x_{2i+1} = 1) \\ 0 \text{, otherwise} \end{cases}$$

x_1, x_2, ..., x_8 are the values of the eight neighbors of p, starting with the east neighbor and numbered in counter-clockwise order.

Condition G2:

$2 \leq \min\{n_1(p), n_2(p)\} \leq 3$

Where,

The image is now processed for the next level of conversion. These conversions are made to make effective verification.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

$$n_1(p) = \sum_{k=1}^{4} x_{2k-1} \vee x_{2k}$$

$$n_2(p) = \sum_{k=1}^{4} x_{2k} \vee x_{2k+1}$$

Condition G3:

$$(x_2 \vee x_3 \vee \bar{x}_8) \wedge x_1 = 0$$

Condition G3':

$$(x_6 \vee x_7 \vee \bar{x}_4) \wedge x_5 = 0$$

Step 1: Collect the signatures from different people in different directions on a white sheet.

Step 2: Scan them in the format which is compatible with the software. Here, the image is scanned in png format. One can use jpg or jpeg as well.

Step 3: After collecting all the images, create a database that could be used further.

Step 4: Load the image into MAT lab from specified location.

FIGURE 1: ORIGINAL IMAGE

Step 5: Convert the original (RGB) image into gray scale image.

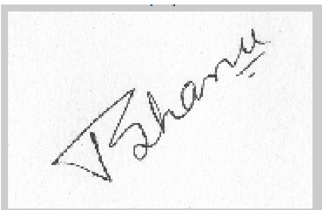

FIGURE 2: GRAY SCALE IMAGE

Step 6: Now, the gray scale image is converted into binary image.

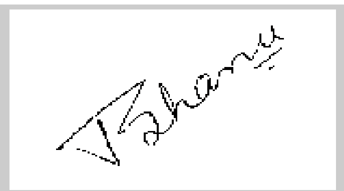

FIGURE 3: BINARY IMAGE

Step 7: Morphological operations like thinning and thickening are applied after the above step.

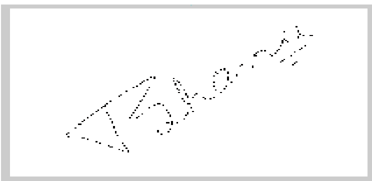

FIGURE 4: THINNED IMAGE

2) Process 2: Calculating center of mass of the thinned signature pixels in x and y direction:

Step 1: Centroid of image according to signature pixel on x and y direction is calculated.

Step 2: Eigen values in x-direction are calculated.

Step 3: Evaluate rotation angle of signature to horizontal center of mass and rotate image using this rotation angle to align the signature in horizontal direction. However, Theta is calculated and the binary image is rotated towards horizontal axis according to the theta value.

Step 4: Crop the signature area.

Step 5: The image is aligned and ready for the next stage.

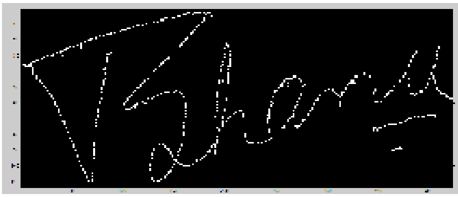

FIGURE 5: ALIGNED IMAGE

3) Process 3: Computation of signature features from the cropped image:

- FEATURE EXTRACTION: Following features are extracted so that we could we could put them into artificial neural network.
- *SKEWNESS:* Skewness is a measure of the asymmetry of the data around the sample mean. If skewness is negative, the data spread more to the left of the mean. If skewness is positive, the data spread is more to the right. The skewness of the normal distribution (or any perfectly symmetric distribution) is zero.

  The skewness of a distribution is defined as

  $$s = \frac{E(x-\mu)^3}{\sigma^3}$$

  where μ is the mean of x, σ is the standard deviation of x, and $E(t)$ represents the expected value of the quantity t. skewness computes a sample version of this population value.
- *KURTOSIS:* Kurtosis is a measure of whether the data are peaked or flattened, relative to a normal distribution. The kurtosis of the normal distribution is 3.
  The kurtosis of a distribution is defined as

  $$k = \frac{E(x-\mu)^4}{\sigma^4}$$

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

where μ is the mean of x, σ is the standard deviation of x, and $E(t)$ represents the expected value of the quantity t. Kurtosis computes a sample version of this population value of kurtosis these are primarily peakiness (width of peak), tail weight etc.

- *MEAN:* For a random variable vector A made up of N scalar observations, the mean is defined as

$$\mu = \frac{1}{N}\sum_{i=1}^{N} A_i .$$

- *STANDARD DEVIATION*: For a random variable vector A made up of N scalar observations, the standard deviation is defined as

$$S = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} |A_i - \mu|^2} ,$$

where μ is the mean of A:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} A_i .$$

The standard deviation is the square root of the variance.

B. *Training and Validation using back-propagation artificial neural network*

The feed-forward neural network (FFNN) is the most popular classifier. In this work MLP-FFNN classifier is being used. Their main advantage is that they are easy to use, and that they can approximate any input/output mapping with weights and thresholds (biases) of the model.

1) *Selecting number of input layer neurons:* The input layer to the neural network is the conduit through which the external environment presents a pattern to the neural network. Here the number of input layer neurons was taken 12.

2) *Selecting layer of output layer neurons:* As each neuron provides only a single output, the number of neurons required in the output layer will be equal to the number of scalars in the output vector. So the number of neurons in the output layer was fixed same as number of classes to be discriminated i.e. 12 to distinguish 1 types of signatures.

3) *Selecting number of hidden layers:* In MLFF networks one of the most important configuration issues is to select an optimal number of hidden layers. To solve this, repeated experiments were performed to determine the size of the hidden layer. The final ANN was made with 1 hidden layer with each node of the layer implementing a sigmoid activation function.

The artificial neural networks work mainly in three phases, which are

a) Training Phase
b) Validation Phase
c) Testing Phase

III. RESULTS AND DISCUSSIONS

C. *Confusion Matrix (Classification Matrix)*

The performance of classifier is analyzed using confusion matrix [32] which is also known as table of confusion. It displays the number of correct and incorrect predictions made by the model compared with the actual classifications in the test data. The confusion matrix lists the correct classification against the predicted classification for each class.

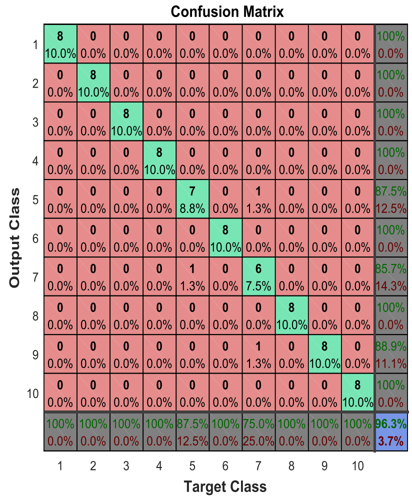

FIGURE 6: OVERALL CONFUSION MATRIX

OVERALL ACCURACY ACHIEVED IS 96.3%.

D. *RECEIVER OPERATING CHARACTERISTICS*

The Receiver Operating Characteristic is a metric used to check the quality of classifiers.

1) *TRAINING PHASE:*

In the below figure we can see that in training phase, the values of R is 0.95759 which means that data has been fitted and classified in a correct manner. Moreover, the validation phase lies on the random guess line (the diagonal line).

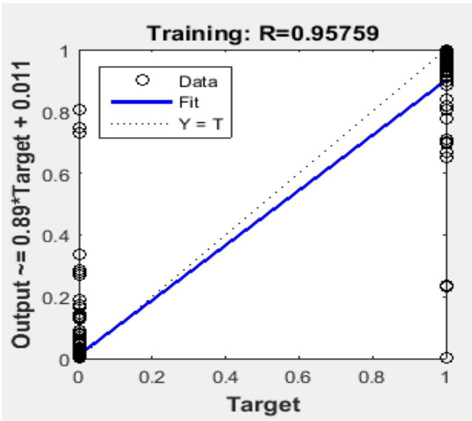

FIGURE 7: TRAINING PHASE ROC CURVE

2) *VALIDATION PHASE:*

In the below figure we can see that in validation phase the value of R is 0.99024 which means data has been fitted

 DOI 10.17148/IJARCCE.2015.41022

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

and classified in a correct manner. We can see that result of validation phase lies on the random guess line (the diagonal line).So, the accuracy is 100%.

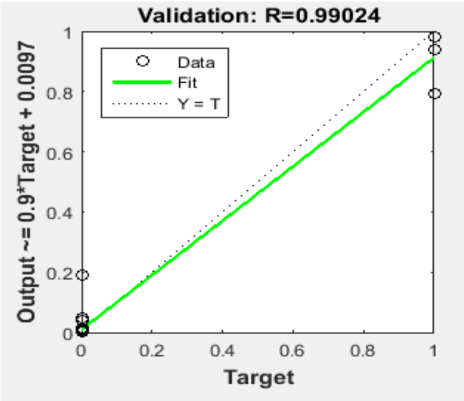

FIGURE 8: VALIDATION PHASE ROC CURVE

3) TESTING PHASE:

Value of R is 0.99567 and the coefficient of goodness i.e. R^2 is not 100%.

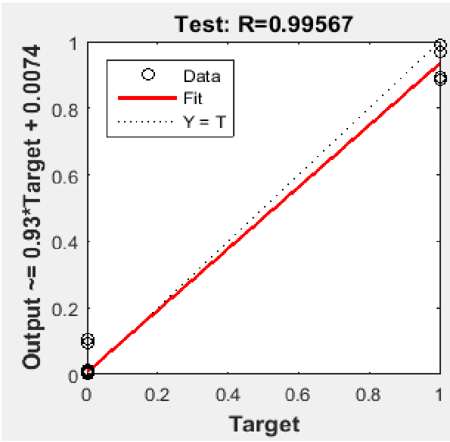

FIGURE 9: TESTING PHASE ROC CURVE

On the basis of above discussed phases (training, validation and testing) we can generate the final ROC curve according to which the values of R is 0.9611 which means data has been fitted and classified in a correct manner which is the ideal condition for goodness of classifier.

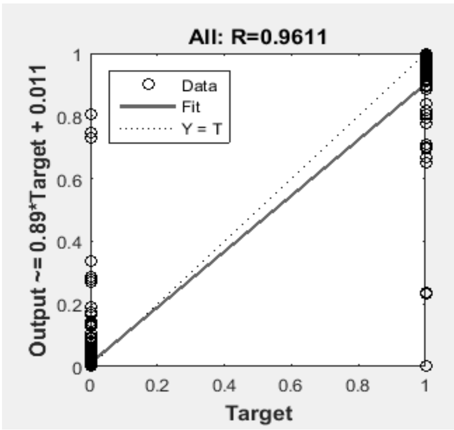

FIGURE 10: OVERALL ROC CURVE

IV. CONCLUSION

The paper presents a method of signature recognition and verification process. The extracted features like skewness, kurtosis, mean, standard deviation are fed to the feed forward neural network.

We have taken the samples of signature images from 10 different persons. Each person has given 8 different views of his/her signatures to create the database. From overall dataset 90% of dataset or a set of object used to design a classifier has been used for training phase.

According to table 3.1, the confusion or classification matrix is showing the overall accuracy achieved. The overall accuracy achieved is 96.3%.

Figure 3.2 shows the value of R in training phase which is 0.95759. The value of R in validation phase is 0.99024 and the value of R in testing phase is 0.99567.

The overall ROC achieved is 0.9611 which is the ideal condition for goodness of classifier.

The approximate overall classifier accuracy achieved is 96.3% which is a good amount.

ACKNOWLEDGMENT

Thanks to my family, my guide **Er. Navdeep Kaur** (AP, CSE Department), **Er. Harpreet Kaur** (AP & HOD, CSE Department), Bahra Group of Institute, Bhedpura, Patiala for having faith in me and allowing me to work with all terms and conditions and advising me time to time about my work.

REFERENCES

[1] M.Nasiri, A. Javaheri, "A Fuzzy Approach for the Automatic Off-line Persian Signature Verification Problem", 978-1-4577-1535-8/11 2011 IEEE
[2] Rahul Dubey, Dheeraj K Agrawal, "Comparative Analysis of Off-line Signature Recognition", 2012 International Conference on Communication, Information & Computing Technology (ICCICT) IEEE
[3] Vu Nguyen, Michael Blumenstein, "A Compact Size Feature Set for the Off-line Signature Verification Problem", 2012 10th IAPR International Workshop on Document Analysis Systems IEEE
[4] R.K.Bharathi, B.H.Shekar, "Discriminative DCT: An Efficient and Accurate Approach for Off-line Signature Verification", 978-0-7695-5100-5/13 2013 IEEE
[5] S. A. Angadi, Smita Gour, Gayatri Bajantri, "Offline Signature Recognition System Using Radon Transform", 978-0-7695-5100-5/13 2013 IEEE
[6] Srikanta Pal, Umapada Pal, Michael Blumenstein, "Off-line verification technique for Hindi signatures" , IET Biom., October 2013, Vol. 2, Iss. 4, pp. 182–190
[7] Moises Diaz-Cabrer, Marta Gomez-Barrero, Aythami Morales, Miguel A. Ferrer and Javier Galbally, "Generation of Enhanced Synthetic Off-line Signatures Based on Real On-line Data", 2014 14th International Conference on Frontiers in Handwriting Recognition IEEE
[8] R.K.Bharathi, B.H.Shekar, "Discriminative DCT-MLP based Approach for Off-line Signature Verification", 978-1-4799-3080-7/14 2014 IEEE
[9] Rapanjot Kaur, Gagangeet Singh Aujla, "Review on: Enhanced offline signature recognition using neural network and SVM", International Journal of Computer Science and Information Technologies", Volume5 (3), 2014
[10] Bhnau Priya Taneja, Navdeep Kaur, "A Literature Survey of Signature Recognition and Verification System", National conference, IIMEDII-2015, March 2015
[11] Bhnau Priya Taneja, Navdeep Kaur, "Biometric System Based on Off-Line Signatures", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue 5, May 2015