Efficient Signatures Verification System Based on Artificial Neural Networks

H. Said-Ahmed[1] and E. Natsheh[2]

[1]*College of Applied Studies, King Faisal University, Saudi Arabia.*
[2]*College of Engineering, AMA International University, Bahrain.*

dr_natsheh@hotmail.com

Abstract—**Biometrics refer to the system of authenticating identities of humans, using features like retina scans, thumb and fingerprint scanning, face recognition and also signature recognition. Signatures are a simple and natural method of verifying a person's identity. It can be saved as an image and verified by matching, using neural networks. Signature verification can be offline or online. In this work, we present a system for offline signature verification. The user has to submit a number of signatures that are used to extract two types of features, statistical features and structural features. A vector obtained from each of them is used to train propagation neural network in the verification stage. A test signature is then taken from the user, to compare it with those the network had been trained with. A test experiment was carried out with two sets of data. One set is used as a training set for the propagation neural network in its verification stage. This set with four signatures form each user is used for the training purpose. The second set consists of one sample of signature for each of the 20 persons is used as a test set for the system. A negative identification test was carried out using a signature of one person to test others' signatures. The experimental results for the accuracy showed excellent false reject rate and false acceptance rate.**

Keywords—Signature verification; neural network; false reject rate; false acceptance rate

I. INTRODUCTION

AUTHENTICATION of individual identities is a routine part of our daily lives. Signature verification is used for traditional transactions like bank cheques, credit and debit card transactions and legal documentation. Most financial transactions require signatures which have to be authenticated. The downside of signatures is that they can be forged. This brings into focus the need to have robust, foolproof and efficient automated systems for signature verification [1]. Offline signature verification systems consider the image of a signature and match it with the one submitted as a sample. Here, the geometrical dimensions are extremely important, since the accuracy of the result depends on a number of factors like the scanning device, the paper used to capture the signature and the pressure exerted while signing. Signatures are first scanned. Then it is preprocessed for extracting the features. Thereafter, this preprocessed image is used to extract accurate geometric patterns, which can be used to compare the features using Artificial Neural Networks (ANN) for possible forgeries. The main goal of signature verification systems is to convert the handwritten word into a digital form to compare it with a sample given earlier. The ANN uses algorithms, which are first trained by testing samples, Thereafter, the algorithms are set to work as per predefined norms, which defines the acceptable limits of errors in the signatures fed in for verification. Then the ANN checks these signatures with the original samples earlier fed in and returns values, which indicate whether a signature is genuine or forged. These systems are also called optical character recognition (OCR) systems.

Article history: Manuscript received 19 September 2019; received in revised form 27 March 2020; Accepted 27 March 2020.

We will examine in detail how the offline signature verification using artificial neural networks work in actual operations. In the following discussion, we will detail the entire workings, broken up into distinct sections, starting from the acquisition of the signature to the output derived from the artificial neural network.

II. RELATED WORK

In signature verifications, three types of forgeries are anticipated [1]. These are: Random forgery, where the person doing the forgery is taking a chance on his/her forgery passing authentication checks. Here, the forger is not actually aware of the name or signature pattern of the person whose signature he/she is forging. The second type is the "simple" forgery, where the person doing the forgery is aware of the name of the person whose signature he/she is impersonating, but not the signature pattern. The third is the "skilled" forgery, where the person forging the signature, knows the geometric pattern of the original signature and the signature generated by a skilled forger is akin to the original. Various papers have been printed outlining different methodologies for signature verification.

Kaur and Aujla [2], describe the other biometric means of personal identification. These are based on various personal identifiers like voice, hand structure, lip movement, odor, iris, retina, fingerprint and gait. Biometrics have the edge over other forms of authentication like passwords, PIN and personalized smart cards. They cannot be stolen or lost. Further, biometrics are unique to a person and cannot be hacked. It have to be mentioned here, that the signature needed to be verified may be genuine or forgery. The only input here is the image of the signature. This image is captured by a traditional scanner in most cases, and there are other types of scanners also like C-pen, which consists of a digital camera, a processor and a memory. The digital camera inside the pen captures the text and saves it in the memory, from where it can be transferred to a PC using a cable.

Many problems arise such as variations in genuine signatures, and noise introduced by the scanning device, difference in pen width, which make off-line signature verification a problem. Signature verification research works on random forgery detection and skilled forgery detection. Skilled forgery detection is a much more difficult task; since differentiation between a good forgery and a genuine signature is difficult. Random detection uses only the image based features to classify signatures as genuine or forgeries.

Biswas et. al. [3] write that to develop an offline signature system, work on some geometrical features, based on the shape and dimension of the signature image are used. Baseline Slant angle, which is the imaginary line around which the signature is assumed to rest. The features were extracted from a group of signature images .The mean values and standard deviation for all features are computed and used for verification and the Euclidian Distance given by the following equation:

$$d((x_1,y_1),(x_2,y_2)) = \sqrt{(x_1-x_2)^2 + (y_1-y_2)^2} \quad (1)$$

where distance is defined as $d(x,y) = |x-y|$.

If this distance is less than a defined value then the signature accepted as a genuine, otherwise it is forged. When the system tested it give a results of 8.5% False Reject Rate for original signature, for the forged it gives 13.3% False Acceptance Rate. This system give inaccurate results if large database is used, also it fails in case of skilled forgeries.

The artificial neural network (ANN) is modeled on the human brain and uses algorithms to emulate the brain. It tries to imitate the brain in developing the means to process data as quickly as the brain does. Neural networks are composed of inputs which carry weights on neural connections, which calculates the input signal entering the neuron and an activation function calculating the level of activity to modulate neuron generation, based on the incoming load. Finally, there is an output calculation, which is based on the signal reaching a predefined threshold. Different

algorithms are used in neural networks for handwriting recognition. They propose that an ANN algorithm is developed and trained to detect patterns in samples of different handwritten signatures; some examples are [4], [5], [6] and [7].

Kovari and Charaf [8] maintain that the biggest drawback of offline signature verification is the fact that it deals with a two dimensional image. Though other features of signature verification like velocity and pressure are available, due to the two dimensional limitation, there is a minimal image transformation, which impacts the output values. The rate of rejection of authentic signatures and the rate of acceptance of forgeries is minimized by fine tuning the neural network to reach a reach a point where both the rate are equal. This is known as Equal Error Rate (EER).

In [9], authors presented algorithm finds the best non-linear alignment of two vectors such that the overall distance between two corresponding vector elements is minimized in least square sense. The overall distance between two signatures calculated in linear time using specific equation, which is mainly work on distances on the signature, the test signature is compared with each reference signature, resulting in a number of distance to the closest, farthest, and the template signatures are all used to classify the test signature as genuine or forgery. Same method was used for offline and online signature verification. It gets low performance in off-line signature verification, because it is hardly to differentiate between forgery and genuine signature of a writer.

III. PROPOSED SYSTEM

Biometrics is the science of identifying or verifying the identity of person based on physiological or behavioral characteristic`s. The behavioral characteristic like signature and voice depends on physical characteristics since these are actions performed by persons. Biometrics, take the measurement from a person and make such a procedure to compare the given data with that collected before. This technique has two applications, identification and verification [4].

Over thousands of years, biometrics was used for identification. In Egypt's Nile Valley, traders were formally identified based on physical characteristics such as height, eye color and complexion. Biometrics serves multiple purposes. It offers an alternate method for identification other than the PIN and passwords, which can be lost or stolen. In biometrics the characteristics are used as identifier, so there is nothing dependent on human memory or theft. The internet and advances in technology have increased the geographical reach of biometrics.

When choosing a biometric for an application one of the important questions that must be asked is whether the application needs to verify or identify. If an application requires an identification of a subject from a large database, it needs a scalable and relatively more distinctive and foolproof biometric (e.g., fingerprint, iris, or DNA). A biometrics system is an automatic identification or verification of an individual by using a behavioral or physical characteristic of that person [2]. It may be called either a verification system or an identification system.

A verification system verifies a person's identity by comparing the captured biometric characteristic with that already saved in a data base. A verification system may accept or reject these entered data. An identification system recognizes a person by searching the data base for a match [10] [11] [12].

A simple biometric system can be designed in the following steps [8]:

- Acquisition
- Enhancement (enrolment)
- Feature extraction
- Matching

In evaluating a performance of a biometrics system there are two important factors, False Rejection Rate (FRR) occurs when the system rejects an authorized user, and the False Acceptance Rate (FAR) occurs, when the system accepts an unauthorized user. Also, where FAR=FRR, it is called Equal Error Rate (EER).

To evaluate the performance, we have to calculate each of (FRR) and (FAR) and since the relative lowering of one of them will increase the other's value.

A human characteristic, whether physical or behavioral, can be used to validate a person if it meets some criteria, as explained below:

Universality: indicates that each person should display biometric characteristics;

Distinctiveness: indicates, that any two persons should display separate biometric characteristics to identify them as separate entities.

Permanence: Meaning that biometric characteristics should remain constant for the same person over a period of time. .

Collectability: This indicates that the biometrics be procured for quantitative measurements.

Performance: This refers the achievement of accurate acceptable results, in a given span of time.

Acceptability: This refers to the people's acceptability of a specific biometric validation in their daily routine.

Circumvention: This indicates that despite the systems validations, it is not very difficult to make the system tag a forged signature as genuine and vice versa.

The four stages of offline signature verification will be detailed, starting with acquisition, image reprocessing and extraction of geometric features and further processing to detect forgeries, using the artificial neural networks.

A. Acquisition

In this stage a biometric data is collected using special biometric reader for each type, and stored in a data base. The performance of a biometric system is largely affected by the reliability of the reader used. Moreover if the biometric trait or being measured is noisy, the result given by the comparison tool is not reliable.

B. Preprocessing

The signature is first captured and then converted into a digital format which can be read by computers. Here, the colored scanned image is first converted into a grayscale and then further processed into a binary image. This is a stage where the signature is being ready for extraction of features. There are two steps here: Color Inversion, where the actual color is transformed into a grayscale image, by rejecting the hue and depth saturation and retaining the lightness or luminescence. The second step is the image filtration and conversion to binary image. Preprocessing is used to remove unwanted and distracting elements such as noise. Any ordinary scanner with enough resolution can be used as an image acquisition device. The scanned signature image is preprocessed to be suitable for feature extraction stage. Scanning may introduce noise to the signature image. Another source of the noise may be paper background on which the signature is given. A noise reduction filter is used with the binary image in order to put black pixels on white background. Gaussian smoothing filter is selected to do this function. The Gaussian smoothing filter is used to remove details and noise. It has the formulas shown below:

$$G(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}} \quad (2)$$

where σ is the standard deviation of the distribution which determines the width. Gaussian function is symmetric, so smoothing performed by it will be the same in the two directions (vertical and horizontal), thus the edges will be in its direction. This is shown in Figure 1.

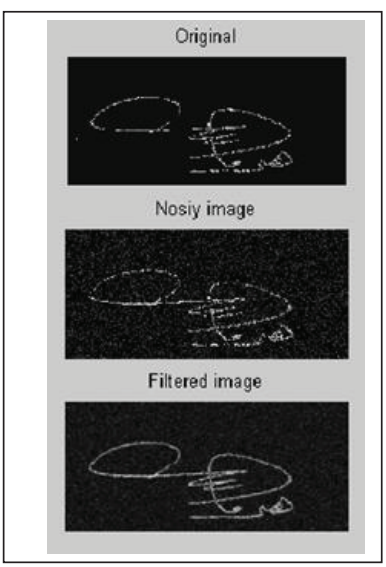

Fig. 1. Signature after filters are applied

C. Features Extraction

Feature extraction is the main component of offline signature verification systems. Global features that are extracted from every pixel that lies within a rectangle circumscribing the signature. These features do not reflect any local, geometrical, or topological properties of the signature, but include transformations, series expansions, image gradient analysis etc. Although global features are easily extractable and insensitive to noise, they are dependent upon the position alignment and highly sensitive to distortion and style variations [2].

Statistical features are derived from the distribution of pixels of a signature, e.g. statistics of high gray-level pixels to identify pseudo-dynamic characteristics of signatures. This technique includes the extraction of high pressure factors with respect to specific regions (for example, upper, middle and lower envelope).

Geometrical and topological features describe the characteristic geometry and topology of a signature and thereby preserve the signatures global and local properties, e.g. local correspondence of stroke segments to trace signature.

The selection of features is the most critical process, because selecting wrong features or not selecting useful features can return null or improper results. Features should contain enough properties to segregate between different classes, ignore irrelevance in input, and also limit the quantity of experimental data required for training the system.

In this work three different features were extracted from the signature. These features are based on some statistical calculations and structural features depending on the shape of the signature. After extraction, feature vectors are used in verification phase. To get these features we have:

1. Horizontal and vertical projection of the image of the signature.
2. Lower envelop of the signature which can be defined as the curve connecting lower most pixels of the signature trajectory.
3. Upper envelop of the signature which can be defined as: the curve connecting upper most pixels of the signature trajectory [13].

At the beginning, let's define features as follows:

1. Horizontal projection image:

$$X(i)=\sum_{j} im(i,j) \tag{3}$$

2. Vertical projection image:

$$Y(i)=\sum_{i} im(i,j) \tag{4}$$

where $im(i,j)$ is either 1 or 0 and i is the row index and j refers to the column, and the r^{th} order moment measure for projection is defined in:

$$\mu_r=\sum_{i}(xi-x^c)rG(xi) \tag{5}$$

where x^c is the centroid of the corresponding projection image and $G(xi)$ can be either X() or Y().The horizontal and vertical projection image can be defined as mean value of the vertical (horizontal) points on the x, y coordinates.

The Moment is defined as follows: If X is a random variable, the r^{th} moment of X, usually denoted by $\mu' r$, and is defined as:

$$\mu_r^{'}=\varepsilon[X^{r}] \tag{6}$$

Note that $\mu_r^{'}=\varepsilon[X^{'}]=\mu_x$, the mean of X.

Central moments are defined as: if X is a random variable, the r^{th} central moment of X about A is defined as:

$$\mu_r^{'}=\varepsilon[(X-\mu_x)^r] \tag{7}$$

Note that $\mu_1=\varepsilon[(X-\mu_x)]=0$ and $\mu_2=\varepsilon[(X-\mu_x)^2]$ the variance of X.

This feature is used to produce skewness and kurtosis measures in the following equations:

Kurtosis measures:

$$Kv=\frac{\mu 4v}{(\mu 2v)^2},\quad K_H=\frac{\mu 4H}{(\mu 2H)^2} \tag{8}$$

Skewness measures:

$$Kv=\frac{\mu_{3v}}{(\mu_{2v})^{1.5}},\quad K_H=\frac{\mu_{3H}}{(\mu_{2H})_{1.5}} \tag{9}$$

Skewness pertains to the degree of asymmetry of the distribution. Conversely, symmetry indicates that the data set looks the same from both the left and right sides of the center point. Kurtosis indicates the trajectory of distribution and shows whether the data set is flat or peaked in relation to normal distribution patterns [2]. *V* indicates vertical moments computed from the vertical image and *H* indicates horizontal moments.

Number the image of the signature after preprocessing can be used to get its contour in different forms; one of these forms is a skeleton (see Figure 2).The most used form of contour in the signature verification methods is the upper and lower contour. There are many ways to get the different points of the contour. The methods selected here to get the contour of the image is tracing. To do this, trace the signature upper and lower contour. To get the vector of the upper contour, take the first black pixel in each column of the image matrix, this represents the upper most pixels of the signature. Also for the lower contour, take the first black pixel in each column of the image matrix, which represents the lower most pixels of the signature. Take these values as features of ANN.

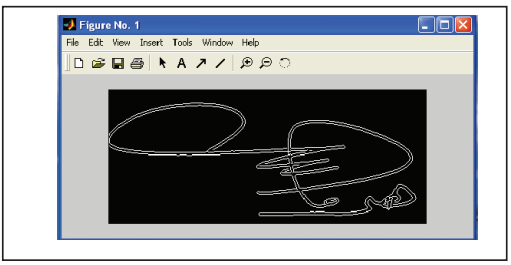

Fig. 2. Image Contour

D. Training, Recognition and Matches in ANN

Initialize setting, where the neural network architecture details are fixed in relation to the inputs and outputs. The number of patterns for training the network is also fixed. Here, they have used 900 units in the input. The formula for this looks like this:

$$h \geq (p-1)/(n+2) \quad (10)$$

where *p* is the number of training samples and n is the number of inputs to the neural network [4].

Generate Training Set: This is considered crucial to the efficient functioning of a neural network for signature verification. The algorithms are dependent of the training set for their outputs. The training set is constructed by saving all the sample signatures into an array, which is used to train the neural network.

Create Neural Network: The training set is merged with a network comprising of multiple layers which are hidden. Here the data flows from input to the output layer.

Initialize Network: Here the weights for each input connection are determined. It is so configured using a weighing equation, so as to keep the range of output values within a time interval of 0.01 to 0.06.

Training the Network: The neural network has to be trained to give the desired output, based on the input values. The algorithms used can be either supervised or unsupervised. The network is trained using the training samples. The algorithm known as *back propagation*, is a common training algorithm, and is a supervised algorithm. In this, the error calculation is done using the simple formula: $O_{error} = D_{output} - N_{output}$, where, O_{error} is the output error, D_{output} is the desired output and N_{output} is the network output. Error calculations are compared with the predefined error value. If the error value touches the maximum predefined value the algorithm stops.

Matching: The signature to be verified is fed into the neural network in the form of a matrix. To determine the accuracy of the signature, the highest percentage value of the output layer is taken. If this is 10%, then it indicates that the match is only 10% and 90% is a not match, which points to forgery.

E. Artificial Neural Networks

Neural networks are constructed by interconnecting neurons. Artificial neural networks are constructed by using interconnected artificial neurons. These are constructed to imitate the actions of actual live neurons present in the human brain [2].

The algorithms built into artificial neural networks attempt to simplify problems to facilitate information processing. The learning stage is made up of a neural network known as multiplayer perceptron or MLR. MLR uses the back propagation algorithm to train the network.

Three feed-forward neural networks (NN1),(NN2),(NN3) were used in this system, one for each set of features. Each of these consists of sub-networks. The sub networks are trained such that it gives only two possible combination output neurons. If the output is (1, 0), this means that the input is recognized. Output of (0, 1) means that the input is not recognized. The neural networks of the system are Multi-layer perceptron.

In summary, the image of the signature is taken, in preprocessing phase it is prepared for feature extraction, and the extracted feature vectors are used for the training of the neural net.

IV. EXPERIMENT AND RESULTS

It is difficult to compare the results of this system with other systems because different systems use different sets of signatures.

For testing the performance of the system, a set of 80 signatures is used; different writers with different styles; (20 user with 4 samples of signatures for each one). First the system checked for the obvious case, in which a signature of a user is used as a replacement signature for another user. This means two different signatures here and no signature is accepted as a genuine signature to the original one.

In the second experiment, skilled forge signatures were needed, which is a subjective issue; The results of checking the system using a (semi-skilled) signature gave 12.31% error rate; while totally replaced signatures gave 9.7% error. Results are shown in Table I.

An experiment done on a small sub set of signatures with those complications and gave high rate of error table shown in Table II. Smaller sets throw up more errors, as reflected by the high FRR.

TABLE I. RESULT OF SECOND EXPERIMENT

Signature type	Result of the experiment		
	Number of signatures	*FAR*	*FRR*
Genuine	80	…….	9.73%
Forged	20	12.31%	…….

TABLE II. ERROR RATE FOR MORE COMPLEX SIGNATURES

Signature type	Result of the experiment		
	Number of signatures	*FAR*	*FRR*
Genuine multiple components	8	25%	37.3%

A small set of continuous handwritten signatures, such as signatures written in English language see (fig. 5) is used to check the system, It giver results of 15% FAR In general, good comments cannot be given with a test done on a very small set. A large set is time consuming because the training needs are bigger. One of them is from different person, which means that it takes a different width, as if it is a different user. Results are shown in table III. The table shows that due to the difference in geometry, the FRR percentage has increased.

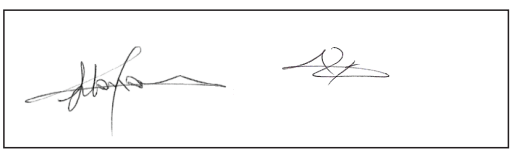

Fig. 3. Example of a figure caption. (figure caption)

TABLE III. RESULT OF SECOND EXPERIMENT

Signature type	Result of the experiment		
	Number of signatures	*FAR*	*FRR*
Genuine	28	33.2%	49.73%

Another experiment, using a combination of genuine, forged, skilled and random forged. In this test we get a 35% (FRR) error rate rejecting skilled forgeries while none of the genuine were rejected, we get a result of 26.4% (FAR) error for the random forge. This experiment was done using all the features, upper and lower envelopes and also horizontal and vertical projection.

On this set, the same experiment was done using only the horizontal and vertical projection which gives a different result, 23% (FRR) error rate rejecting skilled forgeries while none of the random forges was rejected. The results are shown in Table IV.

TABLE IV. COMBINATION SET

Signature type	Result of the experiment		
	Number of signatures	*FAR*	*FRR*
Genuine	28	33%	…….
Skilled forges	4	…….	23%
Random forges	10	…….	0%

Using the back propagation algorithm, the output values are given in the table V [13]. This shows the total rejection rate based on false rejection rate and false acceptance rate.

TABLE V. THE OUTPUT VALUES

No. of Iteration	Result of the experiment		
	FAR	*FRR*	*TER*
100	15%	12%	27.0%
102	12.50%	15%	27.5%
103	12.50%	10%	22.5%
104	14.50%	12%	26.5%
105	17.50%	15%	32.5%

V. CONCLUSION

Handwritten signatures are the simplest means of personal identification, and in this study, we have explored the different aspects of offline signature verification systems. Artificial neural networks are useful in automated signature recognition systems. We have also examined as to how crucial it is, to train the algorithms with sufficient number of sample signatures for this system to work accurately There is a lot of scope for further research in neural networks and with the advances in computer technology, it is expected that the current system of evaluating the authenticity of signatures will be further refined to reduce the FAR and FRR percentages and return highly accurate results.

REFERENCES

[1] Karouni, A., Daya, B. and Bahlak, S. (2011). Offline Signature Recognition Using Neural Networks Approach. *World Conference on Information Technology–Bhacesehir University. Science Direct*, 2 -4.

[2] Kaur, R. and Aujla, G.S. (2014). Review on: Enhanced Offline Signature Recognition Using Neural Network and SVM. *International Journal of Computer Science and Information Technologies*, 5 (3) 1 -4.

[3] Biswas, S., Kim, T.K. and Bhattacharyya, D. (2010). Features Extraction and Verification of Signature Image using Clustering Technique. *International Journal of Smart Home*. 4 (3), 2 – 10.

[4] Jarad, M., Al-Najdawi, N. and Tedmori, S. (2014). Offline Handwritten Signature Verification System Using a Supervised Neural Network Approach. *014 6th International Conference on CSIT*. 2 -6

[5] Choudhary, N.Y., Patil, R. Bhadade, U. and Chaudhari, B.M. (2013). Signature Recognition & Verification System Using Back Propagation Neural Network. *International Journal of IT, Engineering and Applied Sciences Research*. 1 – 7.

[6] Malekar, M.D. and Patel, S. (2013). Off-line Signature Verification Using Artificial Neural Network. *International Journal of Emerging Technology and Advanced Engineering*. 3 (9). 1 – 4.

[7] Sthapak, S., Khopade, M. and Kashid, C. (2013). Artificial Neural Network Based Signature Recognition & Verification. International *Journal of Emerging Technology and Advanced Engineering*, 3(8). 2 – 7.

[8] Kovari, B. and Charaf, H. (n.d.). Feature matching in off-line signature verification. *Recent Researches in Communications and Computers.* 2 - 4.

[9] Jambi, K. (2001). "Different Approaches for Arabic Signature recognition", Al-Azhar University Engineering Journal, vol. 12, 45-53.

[10] Nguyen, Vu., Blumenstein, M. and Leedham, G. (2009). Global Features for the Off-Line Signature Verification Problem. *2009 10th International Conference on Document Analysis and Recognition.* 1 – 4.

[11] Natsheh, E. (2013), Personalized Web Documents Filtering by Analyzing User Browsing Behaviors, *International Journal of Information Studies (IJIS)*, 5(2), 57-65.

[12] Natsheh, E. (2012). "Taxonomy of Clustering Methods Used in Fuzzy Logic Systems", Journal of Telecommunication, Electronic and Computer Engineering (JTEC), vol.4, no.1, pp. 65-72.

[13] Shikha, P. and Shailja, S. (2013). Neural Network Based Offline Signature Recognition and Verification System. *Research Journal of Engineering Sciences*, 2 (2) 1 – 4