Off-Line Handwritten Signature Recognition by Wavelet Entropy and Neural Network

Khaled Daqrouq [1,*,†], Husam Sweidan [2,†], Ahmad Balamesh [1,†] and Mohammed N. Ajour [1,†]

[1] Electrical and Computer Engineering Department, King Abdulaziz University, Jeddah 21589, Saudi Arabia; abalamesh@kau.edu.sa (A.B.); majour@kau.edu.sa (M.N.A.)

[2] Department of Electrical and Computer Engineering, Michigan Technological University, Houghton, MI 49931, USA; hisweida@mtu.edu

[*] Correspondence: haleddaq@gmail.com; Tel.: +966-566-980-400; Fax: +966-2-695-2686

[†] These authors contributed equally to this work.

Abstract: Handwritten signatures are widely utilized as a form of personal recognition. However, they have the unfortunate shortcoming of being easily abused by those who would fake the identification or intent of an individual which might be very harmful. Therefore, the need for an automatic signature recognition system is crucial. In this paper, a signature recognition approach based on a probabilistic neural network (PNN) and wavelet transform average framing entropy (AFE) is proposed. The system was tested with a wavelet packet (WP) entropy denoted as a WP entropy neural network system (WPENN) and with a discrete wavelet transform (DWT) entropy denoted as a DWT entropy neural network system (DWENN). Our investigation was conducted over several wavelet families and different entropy types. Identification tasks, as well as verification tasks, were investigated for a comprehensive signature system study. Several other methods used in the literature were considered for comparison. Two databases were used for algorithm testing. The best recognition rate result was achieved by WPENN whereby the threshold entropy reached 92%.

Keywords: wavelet; entropy; signature; threshold entropy; PNN

1. Introduction

Handwritten signatures have a very widely known tradition of use in commonly encountered recognition tasks such as financial transactions and document authentication. Moreover, signatures are easily used and well accepted by the public, and signatures are straightforward to produce with fairly cheap devices. Signature recognition is divided into verification and identification. In the verification, we detect whether or not a claimed signature is genuine and belongs to the claiming signer (accept or reject), in contrast to identification, in which an imposter signature is recognized and referred to the correct signer.

The advantages of signature recognition make it more preferable over other biometrics. Nevertheless, signature recognition also has some weaknesses: It has a puzzling pattern recognition problem due to the large variations between different signatures made by the same person. These variations may be caused by instabilities, emotions, environmental changes, etc., and are person-dependent. Moreover, signatures can be forged more easily than other biometrics [1]. User recognition by his or her signature can be divided into static (offline), where the signature written on paper is digitized, and dynamic, where users write their signature in a digitizing tablet, tablet-personal computer, personal digital assistant (PDA) stylus [2] or similar; the information acquired depends on the input device. Taking into consideration the highest security ranks that can be attained by dynamic systems, most of the efforts of the international scientific community are concentrated on

this type [3,4]. Static systems are limited to use in legal-related cases [5]. However, most academic and research tasks start from analyzing the signature recognition with an off-line procedure. Forensic document examiners, e.g., Robertson [6], Russell, et al. [7], Hilton [8], and Huber and Headrick [9], have cited in their textbooks the method of automatic writer verification and identification based on offline handwriting. A correspondence can be drawn between this writer identification process and offline signature verification and identification because the individuality analysis of words is similar to signature verification. The approaches that have been seen in the literature to extract significant information from signatures can be broadly split into the following [10]:

- Feature-based approaches, in which a holistic vector, consisting of global features such as signature duration, standard deviation, etc., is derived from the acquired signature trajectories.
- Function-based approaches, in which time sequences describing local properties, such as position trajectory, pressure, and azimuth, are used. A system for verifying handwritten signatures where various static and dynamic signature features are extracted and used as a pattern to train several network topologies is presented [11]. A signature verification system based on a Hidden Markov Model approach for verifying the hand signature data is presented in [12]. Instrumented data gloves furnished with sensors for detecting finger bend, hand position, and orientation for detecting hand signatures are used in handwritten verification [13]. A method for automatic handwritten signature verification that depends on global features that summarize different aspects of signature shape and dynamics of signature production is studied in [14]. A signature recognition algorithm, relying on a pixel-to-pixel relationship between signature images based on extensive statistical examination, standard deviation, variance, and theory of cross-correlation, is investigated in [15]. Online reference data acquired through a digitizing tablet is used with three different classification schemes to recognize handwritten signatures, as discussed in [16]. The influence of an incremental level of skill in the forgeries against signature verification systems is clarified in [17]. Principles for an improved writer enrollment based on an entropy measure for genuine signatures is presented in [18]. Dynamic signature verification systems, using a set of 49 normalized features that tolerate inconsistencies in genuine signatures while retaining the power to discriminate against forgeries, are studied in [19]. A statistical quantization mechanism to suppress the intra-class variation in signature features and statistical quantization mechanism, thus distinguishing the difference between genuine signature and its forgery, is emphasized in [20]. This method is not well-established in the field of signature recognition and needs to be analyzed based on several databases. Other signature verification systems based on extracting local information time functions of various dynamic properties of the signatures are used for comparison. The use of a discrete wavelet transform (DWT) in extracting features from handwritten signatures that achieved higher verification rates than that of a time domain verification system is reported in [21–23]. The limitation of 1D DWT is that it generates discrete wavelet sub-signals from the approximation first level sub-signal. Therefore, the DWT representation could lose the high-frequency components.

According to the classification stage, different proposals can be found in the literature to measure the similarity between the claimed identity model and the input features. In the Signature Verification Competition 2004 (SVC04), dynamic time warping (DTW) [24] and hidden Markov model (HMM)-based schemes [25] were shown to be the most popular; the system based on DTW was the first. A similar result can be seen in [26], where a vector quantization (VQ) pattern recognition algorithm is also tested. Recent proposals show some improvements, by combining systems (VQ-DTW) in [26], by fusing local and global information [12], or by combining different system outputs [25]. In [26] was seen that DTW had performed well in the problem of measuring similarity when the signatures to be compared have different lengths. On the other hand, if the signatures have the same length, i.e., the same number of points, the similarity of which can be computed in a more straightforward way, as a simple distance measurement between vectors can be applied. This distance calculation is

based on the use of the Euclidean one. However, due to the high dimensionality of feature vectors, the authors proposed the use of fractional distances to alleviate the concentration phenomena [26]. A system for handwriting signature verification using a two-level verification method, by extracting wavelet features and through neural network recognition, is presented [27]. Dynamic handwritten signature verification by the wavelet transform with verification using the backpropagation neural network (BPNN) is studied [28]. DTW, VQ, HMM, and BPNN are widely employed in different tasks. These methods are not spatial signature recognition methods. And many modifications and combinations of VQ, DWT, DTW or other methods have been suggested in the literature. That could have success in some places, but we still need original signature methods designed to meet the needs with less limitation related to complexity, time consumption, and accuracy.

For the offline signature verification and identification tasks, many approaches based on neural networks [29], Hidden Markov Models, and regional correlation [30], have been discussed in the literature [31,32]. The paper by Lee and Pan gave a summary of feature selection [33], broadly classifying them into global, statistical and topological features. Global features, as well as statistical features, were extracted and verified using statistical distance distribution, where a simple approach for offline signature verification was adopted [34].

In the literature, many methods have been used widely in recognition tasks, such as DTW, VQ, or HMM. Speaking, they have been used in recognition applications such as speaker identification, speech recognition, face recognition, and signature recognition. They could have good results in some places, but we are sure that the mentioned methods are not designed particularly for signature recognition. Therefore, we need to develop a particular way that possesses a structure inspired from the signature recognition task basis.

In this paper, we have proposed a new method for an off-line signature recognition system based on the use of wavelet AFE. Two proposals by DWT and WP are used for feature extraction. For classification, PNN is suggested. In the proposed system, the entropy wavelet feature vectors are fed to PNN classifier to be recognized. The essential motivations of such choice are as follows:

(1) For the wavelet, the crucial feature of a signature is the signature morphology, as well as the lines concentration [35]. Bearing in mind this fact, the use of wavelet transform entropy would benefit immensely in features tracking, because of the possibility of signal analysis over several passbands of frequency.

(2) For PNN, the feature vector is relatively not long, and that would not affect the PNN algorithm computational complexity. On the other hand, the possibility of working in an embedded training mode makes the system works online. This is easier for implementation, as well as giving PNN the ability to provide the confidence in its decision that follows directly from the Bayes' theorem [36,37]. Although this process doesn't affect the system's performance, it will offer a speedy process as well as perform in a very timely manner.

In this paper, we divide the content into seven parts. The first part is an introduction of detailed information about the topic background and a literature survey. Parts 2 and 3 cover an elaboration on the presented feature extraction method. In the fourth part, we explained the mathematical intuition of the PNN. The results and discussion were managed in Section 5. Finally, conclusions are presented in Section 6.

2. Wavelet Transform Entropy for Feature Extraction

For a given orthogonal wavelet function, a library of wavelet packet bases is generated. Each of these bases offers a particular way of coding the data, preserving global energy and reconstructing exact features. The wavelet transform is used to extract additional features to guarantee a higher recognition rate. In this study, either WPT or DWT are applied at the stage of feature extraction [38], but the extracted data are not proper for the classifier due to their great length. Thus, we have to seek for a better representation of the signature features. Previous studies showed that the use of

entropy of WP, as features in recognition tasks, is efficient. A method was proposed in [39] to calculate the entropy value of the wavelet norm in digital modulation recognition. In the biomedical field, Ref. [40] presented a combination of genetic algorithm and wavelet packet transform used in the pathological evaluation was introduced [40], and the energy features were determined from a group of wavelet packet coefficients. A wavelet packet entropy adaptive network-based fuzzy inference system was developed to classify twenty 512×512 texture images obtained from the Brodatz image album [41]. Reference [42] proposed a features extraction method for speaker recognition based on a combination of three entropy types (i.e., sure, logarithmic energy, and norm) was proposed. The proposed scheme by Sengur [43] is composed of a wavelet domain feature extractor and an ANFIS classifier, where both entropy and energy features were are used in the wavelet domain. An effective rotation and scale invariant holistic handwritten word recognition system is proposed in [44] is introduced, such that the packet wavelet transform entropy was utilized to extract the feature vector of a Farsi word image. As seen above, the entropy of the particular sub-band signal may be employed as features for image recognition tasks. In this paper, the entropy obtained from the wavelet transform will be employed over the extracted data from either WPT or DWT, thus creating feature vectors with the appropriate length for the signature classification process. We should take into account the fact that signature recognition differs from speech/speaker recognition, as the length of the training set is rather short, and it is hard in this situation to estimate an accurate statistical model.

3. Wavelet Transform for Signature Feature Extraction

3.1. Wavelet Packet Transform

The wavelet packet method is considered to be the general case of wavelet decomposition that provides a richer analysis. In the following, the wavelet transform is defined as the inner product of the data $x(t)$ with the mother wavelet $\psi(t)$:

$$\psi_{a,b}(t) = \psi\left(\frac{t-b}{a}\right) \tag{1}$$

$$W_\psi x(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\psi * \left(\frac{t-b}{a}\right) dt \tag{2}$$

where a and b are the scale and shift parameters, respectively. The "*" symbol, denotes the conjugate version. Through modulating *a* and *b*, $\psi(t)$ may be dilated or translated.

The binarized signature image is recursively decomposed through the wavelet packet transform (WPT) using the recursive binary tree (see Figure 1). Essentially, the WPT is very similar to the discrete wavelet transform (DWT), but instead of only performing the decomposition process on approximations, WPT decomposes both details and approximations. In principle, wavelet packet (WP) uses a pair of low pass and high pass filters to generate two sequences, with the purpose of capturing different frequency sub-band features of the original signal. The two wavelet orthogonal basis generated from a previous node are defined as:

$$\psi_{j+1}^{2\,p}(k) = \sum_{n=-\infty}^{\infty} h[n]\psi_j^p(k - 2/n) \tag{3}$$

$$\psi_{j+1}^{2\,p+1}(k) = \sum_{n=-\infty}^{\infty} g[n]\psi_j^p(k - 2/n) \tag{4}$$

where *h*[*n*] and *g*[*n*] denote the low-pass and high-pass filters, respectively. $\psi(n)$ represents the wavelet function. Parameters *j* and *p* are the number of decomposition levels and nodes of the previous node, respectively [36].
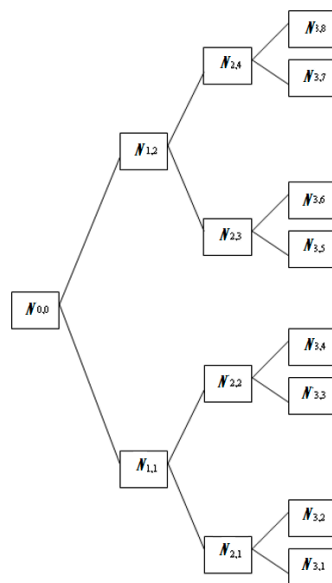
Figure 1. Wavelet packet at depth 3. $N_{level,sequence}$ is the WP node.

3.2. Discrete Wavelet Transform

The DWT indicates an arbitrary square integrable function as a superposition of a family of basic functions, namely, the wavelet functions. A family of wavelet basis functions can be generated by translating and dilating the mother wavelet [14]. Also, by taking the inner product between the original signal and the wavelet functions, the DWT coefficients can be produced. Given that wavelet functions are simply translated and dilated versions of each other, a simpler algorithm, known as Mallat's pyramid tree algorithm, has been proposed (Figure 2) [45–47].

It is possible to use DWT as the multi-resolution decomposition of a sequence. It takes a length N sequence $a(n)$ as the input and produces a length N sequence as the output. The output has $N/2$ values at the highest resolution (level 1) and $N/4$ values at the next resolution (level 2), and so on. Let $N = 2^m$, and let the number of frequencies, or resolutions, be m, while keeping in mind that $m = \log N$ octaves. Therefore, the frequency index k varies as $1, 2, \ldots, m$ corresponds to the scales $2^1, 2^2, \ldots, 2^m$. As depicted by the Mallat pyramid algorithm (Figure 2), the DWT coefficients of the previous stage are expressed as:

$$W_L(n,k) = \sum_i W_L(i, k-1)h(i - 2n), \tag{5a}$$

$$W_H(n,k) = \sum_i W_L(i, k-1)g(i - 2n), \tag{5b}$$

where $W_L(p,j)$ is the p-th scaling coefficient at the j-th stage, $W_H(p,j)$ is the p-th wavelet coefficient at the j-th stage, and $h(n)$, $g(n)$ are the dilation coefficients relating to the scaling and wavelet functions, respectively [47].

In the last decade, a tremendous increase in the applications of wavelets, in various scientific fields, has emerged [46]. The applications of wavelets include signal processing, image processing, security systems, biomedicine, statistics, etc. Wavelet transform offers a wide range of useful features, surpassing other transforms, such as Fourier transform or cosine transform. Some of these are as follows:

− Adaptive time-frequency windows;
− Lower aliasing data deformation for one and two-dimensional signal processing applications;
− Low computational complexity ($O(N)$), where N is the length of data; and
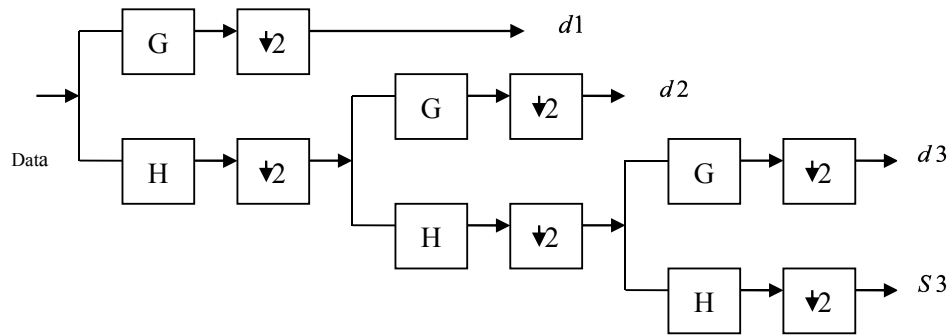
—   Inherent scalability.

Figure 2. DWT-tree by paramedic Mallat's algorithm. Where down-arrow with number 2 symbol denotes the decimation (dyaddown) operation.

The higher verification rate attained from utilizing DWT in extracting features from handwritten signatures rather than that of a time domain verification system is reported.

3.3. Feature Extraction Procedure

As discussed in the previous section, entropy is a common concept in many fields, mainly in image and signal processing. Classical entropy-based criterion describes information-related properties for the precise representation of a given image. Entropy is commonly used in image processing; it possesses information about the concentration of the image [46]. On the other hand, the method for measuring entropy appears as a supreme tool for quantifying the ordering of non-stationary processes. We will use it for signature feature extraction. This conclusion has been obtained empirically by interpreting the following criterion: The feature vector extracted should possess the following properties:

(1)   Varying widely from class to class;
(2)   Stable over a long period of time;
(3)   Should not have a correlation with other features.

Mainly, in this section, the procedure for feature extraction method is to be demonstrated. The binary signature matrix Γ may be represented as:

$$\Gamma = \begin{bmatrix} \delta_{1,1} & \delta_{1,2} & \dots & \delta_{1,250} \\ \delta_{2,1} & \delta_{2,2} & \dots & \delta_{2,250} \\ & \vdots & \vdots & \vdots \\ \delta_{80,1} & \delta_{80,2} & \dots & \delta_{80,250} \end{bmatrix} \tag{6}$$

where $\delta_{a,b}$, represents the pixels of the signature image. Before applying wavelet transform over the matrix Γ, it is reshaped into the vector A, as follows:

$$A = \begin{bmatrix} \delta_{1,1} \\ \delta_{2,1} \\ \vdots \\ \delta_{79,250} \\ \delta_{80,250} \end{bmatrix} \tag{7}$$

When DWT is calculated for the column vector A at a depth (level) of 5 selected empirically for the best performance, the matrix D_{DWT} is obtained:

$$D_{DWT} = \begin{bmatrix} \boldsymbol{cd_1} & \boldsymbol{cd_2} & \ldots & \boldsymbol{cd_L} & \boldsymbol{ca_L} \end{bmatrix} \tag{8}$$

cd_l is the detailed *DWT* subsignal (a column in the matrix *D*), $l = 1, 2, \ldots, L$, and ca_L is the approximation *DWT* sub-signal. In our paper $L = 5$ (this was selected empirically, for the best performance). Originally, the resulting detailed and approximation subsignals are vectors with different lengths. Therefore, they are reshaped into equal size matrices to fit consistently in D_{DWT}. This is achieved by adding zeros to the shorter vectors.

Now, for the case where *WPT* is applied, the *WPT* sub-signals are as follows:

$$D_{WPT} = \begin{bmatrix} \boldsymbol{c_1} & \boldsymbol{c_2} & \ldots & \boldsymbol{c_H} \end{bmatrix} \tag{9}$$

c_h is the WPT node, $h = 1, 2, \ldots, H$ where H is the number of wavelet packet transform nodes; this depends on the level used in our paper, which is 5. The next step is to evaluate the entropy over the rows as well as the columns of the matrix *D* resulting from the wavelet transform (by means of *DWT* or *WPT*). In the proposed study, the focus will be on modifying entropy to reduce the size of feature vectors. In our work, we propose the use of AFE to extract features from the frames of each column or row:

$$U_q(n) = \{u_{q1}(n), \ u_{q1}(n), \ldots, u_{qZ}(n)\}, \tag{10}$$

where Z is the number of considered frames for q-th column or row donated $u_q(t)$. The average of entropy coefficients calculated for Z frames of $u_q(t)$ is utilized to extract a feature vector as follows:

$$afe_q = \sum_{z=1}^{Z} E(u_{qz}(n)) \frac{1}{Z} \tag{11}$$

With the purpose of investigating a better performance, different types of entropy are tested:
The non-normalized Shannon entropy:

$$E(s) = -\sum_{\tau} s_{\tau}^2 \log(s_{\tau}^2) \tag{12}$$

The log-energy entropy:

$$E(s) = \sum_{\tau} \log\left(s_{\tau}^2\right) \tag{13}$$

The Threshold entropy:

$E(s) = 1$ if $|s_i| > \mathbb{P}$ and 0 elsewhere so $E(s) = \#\{i$ such that $|s_i| > \mathbb{P}\}$ the number of times where $|s_i|$ is greater than the threshold $\mathbb{P}$ $\tag{14}$

The input signal, which is either a row vector or a column vector from the matrix *D*, and s_i are the coefficients of *s* in an orthonormal basis.

The obtained entropy results vector $\mathbb{F}$ is the feature vector output:

$$\mathbb{F} = \begin{bmatrix} afe_{q_1} & afe_{q_2} & \ldots & afe_{q_N} \end{bmatrix} \tag{15}$$

where N is the number of elements in the feature vector F, which equals the number of columns times the number of rows of the input matrix *D*.

The obtained entropy values are tackled as the feature vector is fed to the classifier. Figure 3 shows WPE calculated for WP at depth 5 with db5 for two writers. For each person, three different signatures were used. We can notice that the feature vector extracted by WPE is appropriate for signature recognition.
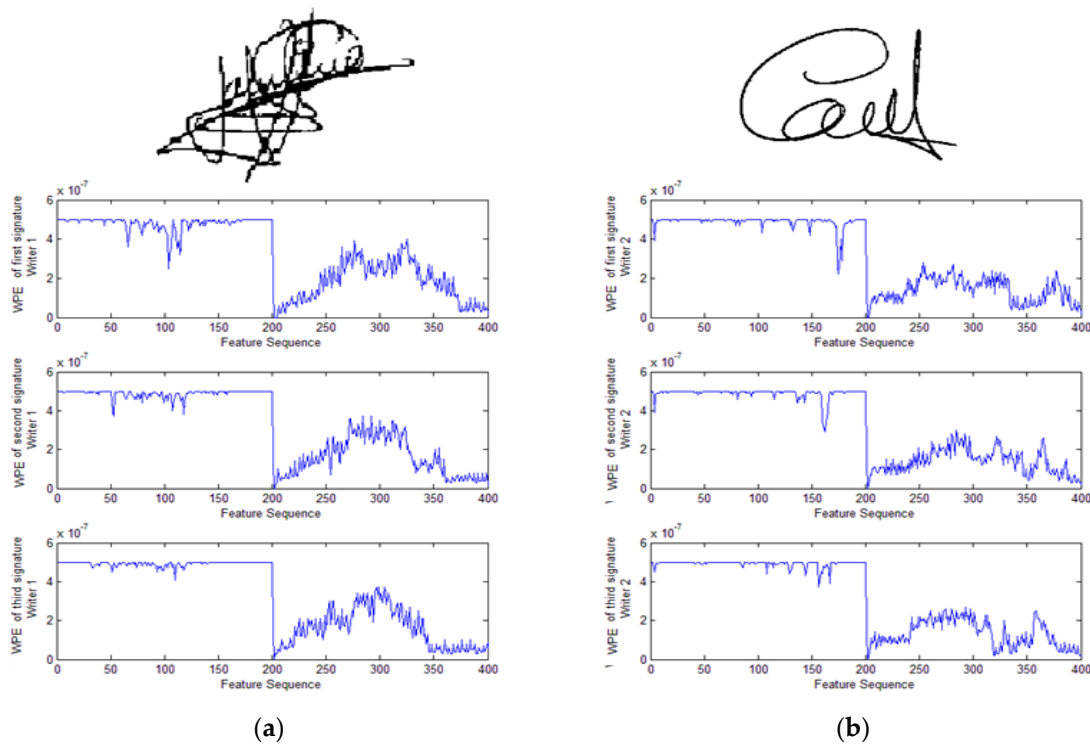
Figure 3. Shows WPE calculated for WP at depth 5 for two persons. (**a**)The feature extraction vector for three signatures of person 1; (**b**) The feature extraction vector for three signatures of person 2.

4. Classification

Along with the introduction of the original probabilistic neural network by Specht [48], several extensions, enhancements, and generalizations have been proposed. These attempts are intended to improve either the classification accuracy of PNNs or the learning capability; on the other hand, they optimize network size, which depresses the memory requirements and the resulting complexity of the model, as well as attaining lower operational times [49]. For the classification of the signature's feature vectors, we use the PNN as a classifier (seen in Figure 4). The primary stimulus of such choice is the potential of working in an embedded training mode, which lets the system operate online and makes implementation easier, as well as PNN; moreover, it can provide a credible decision that follows directly from Bayes' theorem. In addition to the harmless impact, this procedure has on the system performance; it will speed up the process, as well as operate in a real-time manner.

Figure 5 demonstrates the basic configuration of a PNN for classification in K classes. As illustrated in the figure, the first layer of the PNN, indicated as an input layer receives the input vectors to be classified. The nodes in the second layer, namely the pattern layer, are aggregated in K groups according to the class they belong to. All nodes in the pattern layer, also referred to as pattern units or kernels, are connected to all inputs of the input layer. Even though there are various possible probability density function estimators, here, we assume that every pattern unit can be realized as having an activation function, namely the Gaussian basis function:

$$f_{ij}(x; c_{ij}, \sigma) = \frac{1}{(2\pi)^{d/2}\sigma^d} exp\left(-\frac{1}{2\sigma^2}(x - c_{ij})^T(x - c_{ij})\right) \tag{16}$$

where $i = 1, \ldots, K$, $j = 1, \ldots, M_i$ and M_i is the number of pattern units in a given class k_i. σ is the standard deviation, it is also referred to as the smoothing or spread factor. It regulates the receptive field of the kernel. The input vector x and the centers $c_{ij} \in R^d$ of the kernel have a dimensionality d.

exp stands for the exponential function, and the transpose of the vector is indicated by the superscript *T* [46].
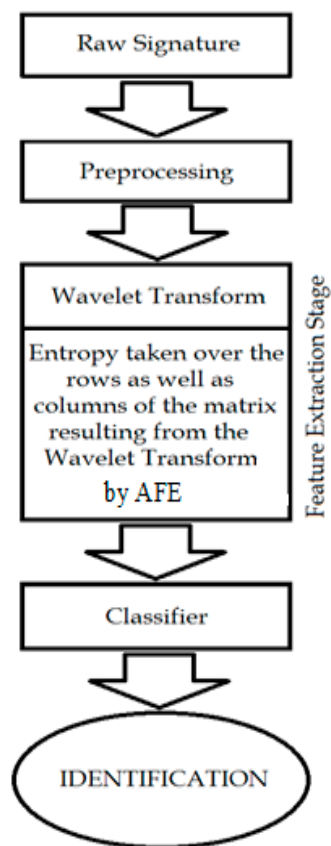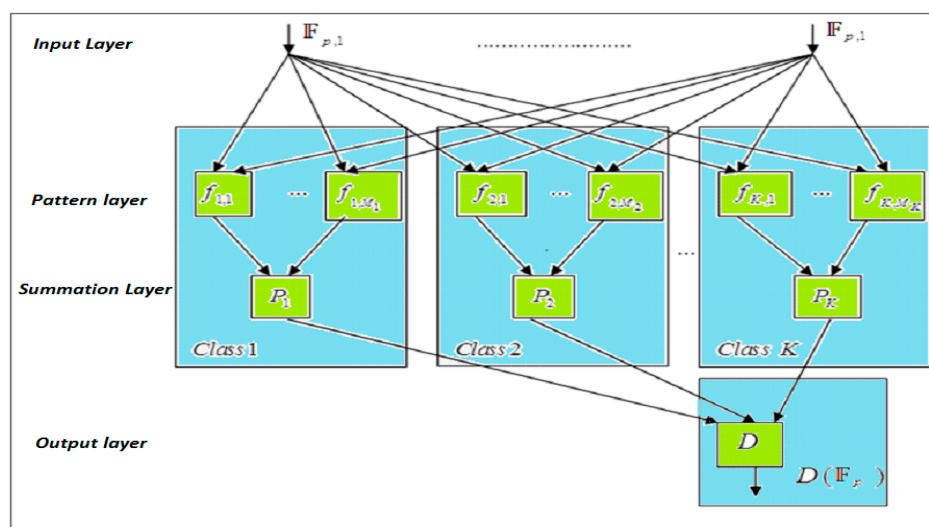

Figure 4. Flowchart of the proposed system.

Figure 5. The original probabilistic neural network Structure.

Obviously, the total number of the pattern layer nodes is represented as a sum of the pattern units for all classes:

$$M = \sum_{i=1}^{K} M_i \tag{17}$$

Next, the weighted outputs of the pattern units from the pattern layer that belong to the group k_i are connected to the third layer that is chosen as summation layer corresponding to that specific class k_i. The weights are resolved with the assistance of the decision cost process and the a priori class distribution. The positive weight coefficients ω_{ij} used for weighing the member functions of class k_i need to fulfill the following requirement:

$$\sum_{j=1}^{M_i} \omega_{ij} = 1 \text{ for every given class } k_i, \; i = 1, ..., K. \tag{18}$$

PNN for the classification task is proposed in [50,51]. Even though there exist many improved versions of the original PNN, that can be either more economical or show an appreciably better performance, for simplicity of exposition, we embrace the original PNN for classification process. The used algorithm is denoted by *PNN* (we used Matlab function "newpnn" to create a PNN) and relies on the following construction:

$$Net = [I, T, SP] \tag{19}$$

where **I** is the input writers feature vectors (pattern) resulting from the feature extraction method discussed previously, and it is used for net training:

$$I = \begin{bmatrix} \mathbb{F}_{11} \; \mathbb{F}_{12} \ldots \; \mathbb{F}_{1Tr} \\ \mathbb{F}_{21} \; \mathbb{F}_{22} \ldots \; \mathbb{F}_{2Tr} \\ \\ \cdots \\ \cdots \\ \cdots \\ \\ \mathbb{F}_{N1} \; \mathbb{F}_{N2} \ldots \; \mathbb{F}_{NTr} \end{bmatrix} \tag{20}$$

where, *Tr* is the training vectors' number. *P* is the target class vector

$$P = [1, 2, \ldots, Tr] \tag{21}$$

The *SP* parameter is known as the spread of radial basis functions. The value of one is used for the *SP*, since that is a typical distance among the input vectors. If the *SP* value approaches zero, the network will perform as the nearby neighbor classifier. The larger the *SP* becomes, the more the designed network will account for several nearby design vectors. A two-layer network is created. The first layer acquires radial basis transfer function (RB) neurons (as appears in Figure 6):

$$RB(n) = \exp(-n^2), \tag{22}$$

With the use of Euclidean distance (ED) the weighted inputs are calculated;

$$ED = \sum \sqrt{((x - y)^2)}, \tag{23}$$

and its net input with net product functions, which calculate a layer's net input by combining its weighted inputs and biases. The pattern layer has competitive transfer function (see Figure 7) neurons, and it utilizes dot product weight function to calculate its weighted input. Its weight function applies weights to an input to get weighted inputs. The proposed net calculates its net input functions (called NETSUM), which calculate a layer's net input by combining its weighted inputs and biases. Only the first layer has biases. PNN sets the first layer weights to $\mathbb{F}$ and the first layer biases are all set to $0.8326/SP$ resulting in radial basis functions that cross 0.5 at weighted inputs of $+/- SP$. The pattern layer weights are set to *P* [47].
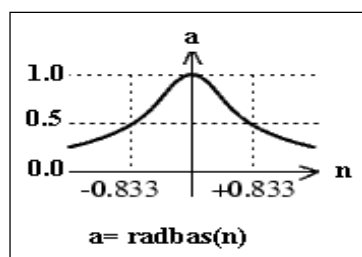
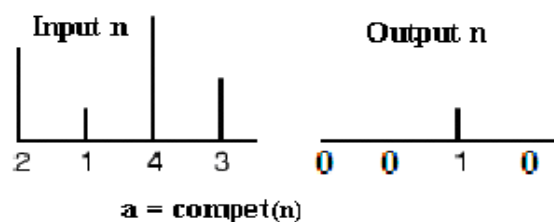**Figure 6.** Radial basis transfer function.

Figure 7. Competitive transfer function.

Now, we test the network on new feature vector (testing signatures) with our network to be classified. This process will be called simulation.

5. Results and Discussion

In this work, we deal with two databases. Database A is the GPDS960 offline signature database [45], which contains data from 960 individuals: there are 24 genuine signatures per writer, and for each genuine signature there are 30 skilled forgeries made by 10 forgers from 10 different genuine specimens. The contained signatures are in "bmp" format, in greyscale and 300 dpi resolution. Database B is also an offline database that contains data from 20 individuals, with 15 signatures per writer. This database was constructed by giving each writer a 5 min session to complete a form that asks for 15 signatures. The signatures were later digitized using a handheld scanner, namely the LG LSM-100 mouse scanner with a 300 dpi resolution. The signatures in database A are in black and white "PNG" format and a 300 dpi resolution. The signatures images were turned into the binary format using fixed thresholding [45]. Moreover, the signatures in database B are processed by cropping the background margins that have no data, followed by the conversion into a black and white "PNG" format. The signature images were turned into the binary format by thresholding using Otsu's method and eliminating any possible noise originating from the background. It is worthwhile mentioning that signatures, in both databases, were unified to a fixed size of (80 × 250) matrix, to prepare them for the feature extraction process. The conducted experiments will tackle two main recognition tasks: the identification task (for academic purposes) and the verification task. The verification task is more popular in practice. For the all conducted experiments, database A is considered, 900 writers called class (24 signatures per class), where 10 or 15 signatures for the training of each writer and 14 or nine signatures for testing of each writer were used. For verification, 15 signatures for the training of each writer and nine signatures for testing of each writer were used. This was selected after trying different numbers of training signatures. The choice was taken to the tradeoff between degreasing the dimensionality of the input matrix and the performance. For example, ten signatures for the training of each writer choice decreased the dimensionality of the input matrix but stepped down the identification rate of about 7%. At each time, 50 classes (750 signatures) were trained. Table 1 shows the typical signature identification results of the proposed system using DWENN concerning the different wavelet functions and the different entropies (Shannon, Log energy, and Threshold). The comparison of the scores with five sub wavelet functions for each wavelet family is also shown. From these tabulated results, we conclude that the system performance of DWENN, by threshold

entropy (in our experiments, we have set Threshold over the rows is 250, and Threshold over the columns is 15, this was selected empirically for the best performance), and db10 wavelet function, with recognition rate reaching 89.99%, was the best. The second database, B, had a slightly worse rate, 87%. Besides, the performance concerning the proposed system by means of WPENN was studied. Similar to the previous experiment, the comparison of the identification scores with five subwavelet functions (i.e., Daubechies (Db), Coiflet (Coif), Symlet (Sym), and Biorthanol (Bior)) for each wavelet family is shown in Table 2.

Table 1. The identification rates results by DWENN for several entropies and wavelet functions.

DWENN by Shannon Entropy				
Db1	**Db2**	**Db3**	**Db6**	**Db10**
79.44	77.44	77.88	79.88	80.08
Sym1	**Sym2**	**Sym3**	**Sym6**	**Sym10**
79.44	79.44	79.44	80.08	81.66
Bior1.1	**Bior1.3**	**Bior1.5**	**Bior2.2**	**Bior3.5**
79.44	78.23	81.11	83.88	81.11
Coif1	**Coif2**	**Coif3**	**Coif4**	**Coif5**
80	81.66	80.55	82.77	81.11
DWENN by Log Energy Entropy				
Db1	**Db2**	**Db3**	**Db6**	**Db10**
66.44	72.77	67.22	81.05	80.11
Sym1	**Sym2**	**Sym3**	**Sym6**	**Sym10**
67.22	72.77	67.22	72.77	81.88
Bior1.1	**Bior1.3**	**Bior1.5**	**Bior2.2**	**Bior3.5**
67.22	70	68.44	62.22	70
Coif1	**Coif2**	**Coif3**	**Coif4**	**Coif5**
69.44	80.55	80	76.11	78.33
DWENN by Threshold ($\mathbb{P}1 = 250$, $\mathbb{P}2 = 15$) Entropy				
Db1	**Db2**	**Db3**	**Db6**	**Db10**
86.11	84.88	84.22	86.66	89.99
Sym1	**Sym2**	**Sym3**	**Sym6**	**Sym10**
86.11	84.88	84.88	86.11	86.11
Bior1.1	**Bior1.3**	**Bior1.5**	**Bior2.2**	**Bior3.5**
86.11	83.44	82.44	86.11	85
Coif1	**Coif2**	**Coif3**	**Coif4**	**Coif5**
85.55	85.55	83.33	86.66	86.88

Table 2. The identification rates' results by WPENN for several entropies and wavelet functions.

WPENN by Shannon Entropy				
Db1	**Db2**	**Db3**	**Db6**	**Db10**
79.88	79.44	81.44	80.22	78.88
Sym1	**Sym2**	**Sym3**	**Sym6**	**Sym10**
79.88	79.44	81.11	81.44	82.77
Bior1.1	**Bior1.3**	**Bior1.5**	**Bior2.2**	**Bior3.5**
79.88	80.55	82.22	81.11	80
Coif1	**Coif2**	**Coif3**	**Coif4**	**Coif5**
82.33	80.11	80.11	80.11	81.88
WPENN by Log Energy				
Db1	**Db2**	**Db3**	**Db6**	**Db10**
43.33	80.55	77.22	83.33	81.11
Sym1	**Sym2**	**Sym3**	**Sym6**	**Sym10**
43.33	80.55	77.22	83.33	84.44
Bior1.1	**Bior1.3**	**Bior1.5**	**Bior2.2**	**Bior3.5**
43.33	65	73.33	61.66	83.33
Coif1	**Coif2**	**Coif3**	**Coif4**	**Coif5**
81.11	83.88	86.66	81.11	79.44

Table 2. *Cont.*

WPENN by Threshold ($\mathbb{P}1 = 175$, $\mathbb{P}2 = 10$) Entropy				
Db1	**Db2**	**Db3**	**Db6**	**Db10**
86.1	87.77	88.88	88.33	86.10
Sym1	**Sym2**	**Sym3**	**Sym6**	**Sym10**
86.1	87.03	88.23	85.55	89.88
Bior1.1	**Bior1.3**	**Bior1.5**	**Bior2.2**	**Bior3.5**
86.1	85.22	86.11	87.88	85.55
Coif1	**Coif2**	**Coif3**	**Coif4**	**Coif5**
87.88	83.88	85.55	87.77	92.06

From these results, we conclude that the system performance of WPENN, by Threshold (in our experiments, we have set Threshold over the rows is 175, and Threshold over the columns is 10) entropy and coif5 wavelet function, with a recognition rate that reaches 92.06%, is the best. Concerning second database B, it had the slightly better rate that reached 92.20%. The conclusions from those results tabulated in Tables 1 and 2 are as follows:

- The performance of identification rates WPENN is slightly better than that achieved with DWENN. That why it will be chosen for the final system.
- Threshold entropy results outperformed other used entropies.

Concerning the wavelet function, the best results are achieved with coif5. Wavelet functions should be selected empirically, for a given database, method or other circumstances. Table 3 shows the signature identification results of the proposed system using WPENN by threshold entropy about the different wavelet functions for the fraud-testing database. Looking at Table 3, we investigate the system's capability to detect the fraud-testing signatures. Since the results of identification rate of fraudulent signatures are worse to those of genuine signatures, we conclude that the proposed system is capable of detecting the fraudulent ones. It means the fraudulent ones were very bad as testing signals when genuine ones were trained.

To test our method with regards to the number of training/testing signatures, an experiment based on 15/9, 10/14 and first 5/random 5 was conducted. The results are tabulated in Table 4. Based on the results shown in the table we can conclude that the 15/9 training/testing system is the best choice.

Table 3. The identification rates of fraud signatures results by WPENN for several wavelet functions.

WPENN by Threshold (P1 = 175, P2 = 10) Entropy/Fraud				
Db1	**Db2**	**Db3**	**Db6**	**Db10**
46.66	48.33	46.66	46.66	46.11
Sym1	**Sym2**	**Sym3**	**Sym6**	**Sym10**
46.66	48.33	46.66	52.22	51.11
Bior1.1	**Bior1.3**	**Bior1.5**	**Bior2.2**	**Bior3.5**
46.66	50	49.44	47.22	50.55
Coif1	**Coif2**	**Coif3**	**Coif4**	**Coif5**
48.33	52.22	51.66	51.11	52.22

Table 4. The identification rates of different training/testing systems signature result by WPENN.

Method	15/9	10/14	First 5/Random 5
WPENN	92.06	87.12	87.89

For verification, we construct, at each time for each signer, a training matrix from 10 signatures (five genuine and five fraudulent ones). Then we test the remaining signatures (genuine and fraudulent)

to see which are rejected or accepted. If the genuine testing signature is recognized as genuine, it is signified as a true positive (TP) result, but if it is recognized as a fraud, it is signified as a false negative (FN) result. On the other hand, if the fraudulent testing signature is recognized as a fraud it is signified as a true negative (TN) result, but if it is recognized as genuine, it is signified as a false negative (FP) result. This means, the system has Jbeen tested for its accuracy and effectiveness on the 900 classes, which contains both their genuine and skilled fraudulent signature sample counterparts' (24 genuine signatures and 30 fraudulent signatures per class). At each time, five genuine signatures and five fraudulent signatures of a particular class were trained. Figure 8 illustrates the verification rate results obtained with regard to the genuine testing signatures (GTS) and the fraudulent testing signatures (FTS) and for three types of entropy (threshold, log energy and Shannon). WPENN (Figure 8a) and DWENN (Figure 8b) were involved in the experiment. We can notice that the result of WPENN at threshold entropy is the best.
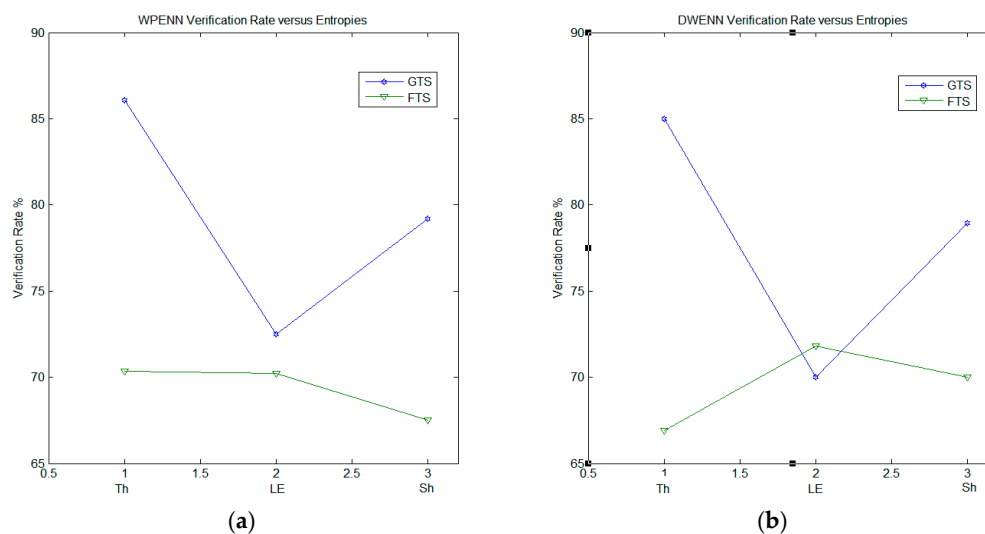

Figure 8. The verification rate results with regard to GTS and FTS for Threshold, Log energy, and Shannon. (**a**) WPENN; (**b**) DWENN. WP and DWT at level five and bior2.2 wavelet function were used.

For comparison purposes, a support vector machine (SVM) [52] classifier is utilized instead of PNN. Table 5 contains the verification rate results with regards to GTS and FTS for WPENN, DWENN, WP with SVM (WPESVM) and DWT with SVM (DWESVM). The experiments were conducted over different wavelet functions selected for the best recognition rate. The best recognition rate was achieved for WPENN, at 87.14% for GTS and 74.22% for FTS. In general, Bior2.2 wavelet (its results are bold in the table) function had the highest average calculated for all the methods. SVM had very good results for FTS, particularly in DWESVM.

Table 5. The verification rates [%] for comparison with SVM results by several wavelet functions.

Method	WPENN		WPESVM		DWENN		DWESVM		Average
Wavelet Function	GTS	FTS	GTS	FTS	GTS	FTS	GTS	FTS	
Db1	80.00	72.86	74.12	72.85	85.14	71.43	77.14	71.43	75.62
Sym6	85.71	71.43	82.00	65.70	82.80	65.71	80.00	65.71	74.88
Bior2.2	**87.14**	**74.29**	**82.72**	**74.29**	**88.71**	**68.57**	**74.29**	**72.86**	**77.85**
Coif1	85.73	65.70	78.57	70.00	84.29	71.48	74.28	75.78	75.72

Table 6 contains the verification rate results with regards to GTS and FTS for WPENN, DWENN, WPESVM and SVM (DWESVM), with a radial basis SVM, with margin parameter C and a Gaussian kernel with parameter G. C and G were optimized through a grid-search with $1 < C < 10^4$ and $1 < G < 10^3$.

Table 6. The verification rates [%] for comparison with SVM results by several wavelet transform levels.

Method	WPENN		WPESVM		DWENN		DWESVM		Average
Wavelet Function	GTS	FTS	GTS	FTS	GTS	FTS	GTS	FTS	
Level2	77.14	74.29	72.87	75.29	80.00	71.43	68.00	68.00	73.37
Level3	82.86	75.71	84.71	74.29	81.43	71.43	71.43	71.43	76.66
Level4	87.14	71.43	82.86	71.43	87.14	70.00	74.29	75.71	77.50
Level5	**87.14**	**74.29**	**82.72**	**74.29**	**88.71**	**68.57**	**74.29**	**72.86**	**77.85**
Level6	84.29	68.57	78.57	74.29	87.14	70.00	84.29	68.57	76.96

The experiments were conducted over different wavelet transform levels (2, 3, 4 and 5). The best recognition rate achieved for WPENN reached 87.14% for GTS and 74.22% for FTS at level 5 (its results are in bold). In general, level 5 had the highest average calculated for all the methods. We can see that SVM is competitive for PNN particularly for FTS.

In the case of systems that are tested under dissimilar, either unpredicted or predicted conditions (it is to extract objective conclusions), it is very difficult to judge quality. Bearing in mind these restrictions, several systems that are known in the literature methods were used for comparison. Taking into consideration these limitations the same database (GPDS960), some testing samples and some training samples were used. Table 7 shows the performance of several system proposal results.

Table 7. The identification rate results for comparison.

Rec. Method	Testing/Training	Length of the Feature Vector	Recognition Rate (%)	Confidence Interval 95%	
				Lower Bound	Upper Bound
RGMM (5 GMMs) [50]	9/15	250	87.94	84.12	90.08
RCGMM (5 GMMs) [50]	9/15	330	88.64	84.57	92.12
FFTGMM [53]	9/15	256	56.51	50.89	60.7
RPNN [54]	9/15	250	59.7	57.67	60.56
RCPNN [54,55]	9/15	330	60.5	52.23	62.25
FFTPNN [47]	9/15	256	60.6	55.34	68.78
DWEGMM (7 GMMs) [53,56]	9/15	450	65.56	60.4	70.54
WPEGMM (2 GMMs) [55,56]	9/15	2500	68.84	64.59	74.23
DWENN	9/15	450	89.99	85.23	92.55
WPENN	9/15	2500	92.06	90.34	94.24

The Gaussian Mixture Model (GMM) applied on the sums of the direct signatures matrix rows (without using any extra transformation of the raw data) denoted by RGMM, the GMM applied on the sums of the direct signatures matrix rows and columns denoted by row column GMM (RCGMM), and the GMM applied on the Fast Fourier Transform (FFT) of the vector of the sums of the direct signatures matrix rows and columns denoted by FFTGMM were utilized for comparison. Besides, for these above-mentioned methods, the PNN was investigated, providing that the same feature vectors were considered: the PNN with sums of the direct signatures matrix rows (RPNN), the PNN with sums of the direct signatures matrix rows and columns (RCPNN), the PNN [51] with the sums of the direct signatures matrix rows (RPNN), and the PNN with FFT of the vector of the sums of the direct signatures rows and columns (FFTPNN) [57]. The results are tabulated in Table 7. In this table, the best performance achieved was for our proposed method. We claim that our proposal results would also progress if combined with other systems or by adding original or derived features. In Table 7, the confidence interval was introduced for each method. The confidence interval states that 95% of the determined recognition rates for different signers are contained in this interval. A wider confidence interval is a bad database sign or not a good feature extraction method. In conclusion, all intervals calculated for presented methods in Table 7 are sensible. Concerning the second database B, it had a slightly better rate, reaching 92.87%.

Figure 9 illustrates the comparison results between several feature extraction methods based on False reject rate (FRR) = FN/(TP + FN) and False accept rate (FAR) = FP/(FP + TN). Some genuine signatures are correctly classified as positive (TP), while others are classified as negative (FN); likewise,

some fraud signatures are correctly classified as negative (TN), while some are classified as positive (FP). In our experiments, six feature extraction methods were involved in the experimental investigation by testing the first 900 signers from GPDS960 offline signature database. These methods are: the proposed WPE method; DWT with threshold entropy (DE); FFT applied on the rows and columns [47]; Shannon entropy applied directly on the rows and columns (DS); mix of Threshold, Log energy and Shannon applied on WP (WPME) and mix of Threshold, Log energy and Shannon applied on DWT (DME) [18,56]; mix of Threshold, Log energy and Shannon applied on WP (WPME); and a mix of Threshold, Log energy and Shannon applied on DWT (DME). For FRR, the best results were achieved by WPE and DS; WPE was better. For FAR, the best results were achieved by WPE, FFT and DME; DME was the best. The averages of the FAR and FRR results were calculated for the six methods, and the ones for WPE was the best.
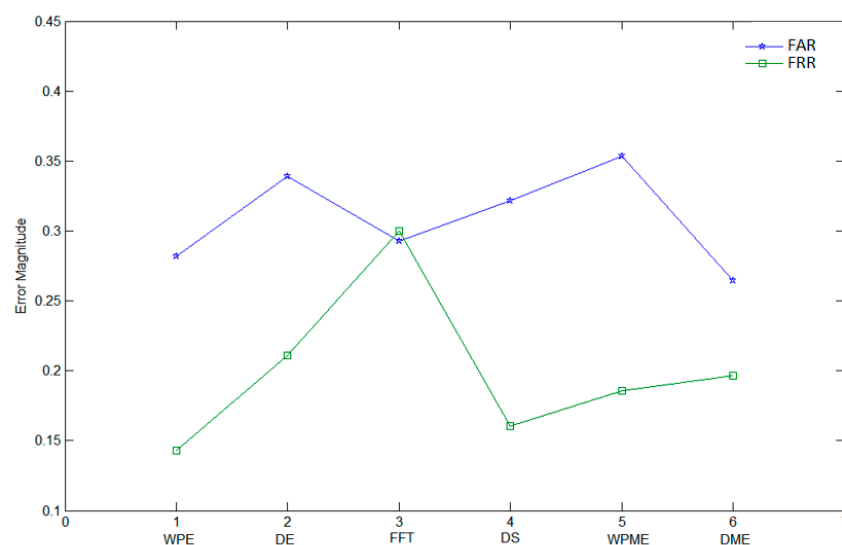

Figure 9. The verification rate results for comparison between several feature extraction methods.

In order to assess the performance of the suggested system, we utilize FAR, FRR, and Equal Error Rate (EER). The first 40 signers of GPDS960 offline signature database were utilized for the verification task [58]. Two other state-of-the-art methods were used for comparison [59]. EER is determined as the value where FAR and FRR are equal. The EER is the best and most popular single explanation of the error rate of a verification algorithm, and the lower the EER, the lower the error rate of the algorithm. The method with the lowest ERR is considered as the most precise method. Therefore, the results presented in Table 8 show that our technique proved to be the most efficient method for the accurate signature feature verification of offline handwritten signatures.

Table 8. The results for the first 40 signatures of the GPDS960 offline signature database utilized for the verification task.

Method	FAR (%)	FRR (%)	EER (%)
WPENN	16.1	16.2	16.15
Global Features for Offline Systems [59]	17.25	17.26	17.25
Graph Matching [59]	16.3	16.6	16.4

One of the most common systems of cross-validation is the leave-one-out cross-validation (LOOCV) system, where the model is continually refit, leaving out a single signature, which is then used to derive recognition for the left-out sample. LOOCV is a case of cross-validation in which the learning procedure is implemented once for each signature, utilizing all other signatures as a

training matrix and having the given signature as a single-instance test set [60]. We applied LOOCV for (40 signers) × (24 signatures) verification so that 960-1 signatures (50% genuine and 50% fraud) was used for training, and one signature was used for testing; we repeatedly changed the testing signature. The final result was calculated as a total result of the 960 testing cases achieved as an 83.37% verification rate.

6. Conclusions

In this paper, we have proposed a new method for an off-line signature recognition (identification and verification) system based on the use of wavelet average framing entropy. This method supposes a better recognition rate concerning the existing literature. An investigation over five wavelet sub-functions for four wavelet families has been done in the same conditions. The research of signature recognition by wavelet entropy is still in the beginning; moreover, we can state that our work is one of the first works in the academic community. In this work, we have shown that it is possible to use the probabilistic neural network for feature vector classification. The wavelet packet has shown that the best identification and verification results reached an average about (taken from database A) and were slightly better than a discrete wavelet transform. Concerning the second database B, it had a slightly better rate. Threshold entropy surpasses the other entropy types regarding the best recognition rate. In future work our algorithm will be further developed by adding or deriving new modifications of the proposed system to get better results.

Author Contributions: The contribution of the authors is similar presented by authors' order, as follows: Khaled Daqrouq: Ideas, Experiments design, Writing the paper; Husam Sweidan: Experiments simulation and results collection, Writing a part of the paper; Ahmad Balamesh: Experiments design and simulation; Mohammed N. Ajour: Reviewers' Comments interpretation and answering.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pascual-Gaspar, J.M.; Faundez-Zanuy, M.; Vivaracho, M. Fast on-line signature recognition based on VQ with time modeling. *Eng. Appl. Artif. Intell.* **2011**, *24*, 368–377. [CrossRef]
2. Alonso-Fernandez, F.; Fierrez-Aguilar, J.; Ortega-Garcia, J.; Gonzalez-Rodriguez, J. Secure access system using signature verification over tablet pc. *IEEE Aerosp. Electron. Syst. Mag.* **2007**, *22*, 3–8. [CrossRef]
3. Jain, A.K.; Griess, F.D.; Connell, S.D. On-line signature verification. *Pattern Recognit.* **2002**, *35*, 2963–2972. [CrossRef]
4. Lei, H.; Govindaraju, H. A comparative study on the consistency of features in on-line signature verification. *Pattern Recognit. Lett.* **2005**, *26*, 2483–2489. [CrossRef]
5. Zhang, D. Automated Biometrics. In *Technologies and Systems*; Kluwer Academic Publisher: Dordrecht, The Netherlands, 2000.
6. Robertson, E.W. *Fundamentals of Document Examination*; Nelson-Hall Publishers: Chicago, IL, USA, 1991.
7. Bradford, R.R.; Bradford, R.B. *Introduction to Handwriting Examination and Identification*; Nelson-Hall Publishers: Chicago, IL, USA, 1992.
8. Hilton, O. *Scientific Examination of Questioned Documents*; CRC Press: Boca Raton, FL, USA, 1993.
9. Huber, R.A.; Headrick, A.M. *Handwriting Identification: Facts and Fundamentals*; CRC Press: Boca Roton, FL, USA, 1999.
10. Fierrez-Aguilar, J.; Nanni, L.; Lopez-Peñalba, J.; Ortega-Garcia, J.; Maltoni, D. An on-line signature verification system based on fusion of local and global information, Audio- and Video-Based Biometric Person Authentication. In *Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3546/2005, pp. 523–532.
11. Trevathan, J.; Read, W.; McCabe, A. Neural network based handwritten signature verification. *J. Comput.* **2008**, *3*, 9–22.

12. McCabe, A.; Trevathan, J. Markov model-based handwritten signature verification. In Proceedings of the 2008 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, Shanghai, China, 17–20 December 2008.
13. Tolba, A.S. GloveSignature: A virtual-reality-based system for dynamic signature verification. *Digit. Signal Process.* **1999**, *9*, 241–266. [CrossRef]
14. Gu¨ler, I.; Meghdadi, M. A different approach to off-line handwritten signature verification using the optimal dynamic time warping algorithm. *Digit. Signal Process.* **2008**, *18*, 940–950. [CrossRef]
15. Bandyopadhyay, S.K.; Bhattacharyya, D.; Das, P. Handwritten signature recognition using departure of images from independence. In Proceedings of the 2008 3rd IEEE Conference on Industrial Electronics and Applications, Singapore, 3–5 June 2008.
16. Zimmer, A.; Ling, L.L. Offline signature verification system based on the online data. *EURASIP J. Adv. Signal Process.* **2008**, *2008*, 492910. [CrossRef]
17. Fernandez, F.A.; Fierrez, J.; Gilperez, A.; Galbally, J.; Ortega-Garcia, J. Spainrobustness of signature verification systems to imitators with increasing skills. In Proceedings of the 2009 10th International Conference on Document Analysis and Recognition, Barcelona, Spain, 26–29 July 2009.
18. Garcia-Salicetti, S.; Houmani, N.; Dorizzi, B. A novel criterion for writer enrolment based on a time-normalized signature sample entropy measure. *EURASIP J. Adv. Signal Process.* **2009**, *2009*, 9. [CrossRef]
19. Lee, L.L.; Berger, T.; Aviczer, E. Reliable on-line human signature verification systems. *IEEE Trans. Pattern Anal. Mach. Intell.* **1996**, *18*, 643–647. [CrossRef]
20. Ong, T.S.; Khoh, W.H.; Teoh, A. Dynamic handwritten signature verification based on statistical quantization mechanism. In Proceedings of the 2009 International Conference on Computer Engineering and Technology, Singapore, 22–24 January 2009; pp. 312–316.
21. Nanni, L.; Lumini, A. A novel local on-line signature verification system. *Elsevier* **2007**, *29*, 559–568. [CrossRef]
22. Nakanishi, I.; Nishiguchi, N.; Itoh, Y.; Fukui, Y. On-line signature verification based on subband decomposition by DWT and adaptive signal processing. *Fundam. Electron. Sci.* **2005**, *88*, 1–11. [CrossRef]
23. Zhan, E.; Guo, J.; Zheng, J.; Ma, C.; Wang, L. On-line handwritten signature verification based on two levels back propagation neural network. In Proceedings of the 2009 International Symposium on Intelligent Ubiquitous Computing and Education, Chengdu, China, 15–16 May 2009.
24. Kholmatov, A.; Yanikoglu, B. Identity authentication using improved online signature verification method. *Pattern Recognit. Lett.* **2005**, *26*, 2400–2408. [CrossRef]
25. Nanni, L.; Lumini, A. Advanced methods for two-class problem formulation for on-line signature verification. *Neurocomputing* **2006**, *69*, 854–857. [CrossRef]
26. Faundez-Zanuy, M. On-line signature recognition based on vq-dtw. *Pattern Recognit.* **2007**, *40*, 981–992. [CrossRef]
27. Nakanishi, I.; Sakamoto, H.; Nishiguchi, N.; Itoh, Y.; Fukui, Y. Multi-matcher on-line signature verification system in DWT domain. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Sapporo, Japan, 26–29 October 2006.
28. Lejtman Dariusz, Z.; George Susan, E. On-line handwritten signature verification using wavelets and back-propagation neural networks. In Proceedings of the Sixth International Conference on Document Analysis and Recognition, Seattle, WA, USA, 13 September 2001; p. 0992.
29. Bajaj, M.; Chaudhury, S. Signature verification using multiple neural classifiers. *Pattern Recognit.* **1997**, *30*, 1–7. [CrossRef]
30. Plamondon, R.; Parizeau, M. A comparative analysis of regional correlation, dynamic time warping and skeletal tree matching for signature verification. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 710–717.
31. Wu, Q.; Lee, S.; Jou, Z. On-line signature verification based on split-and-merge matching mechanism. *Pattern Recognit. Lett.* **1997**, *18*, 665–673.
32. Ammar, M.; Yoshido, Y.; Fukumura, T. Structural description and classification of signature images. *Pattern Recognit.* **1990**, *23*, 697–710. [CrossRef]
33. Lee, S.; Pan, J.C. Offline tracing and representation of signatures. *IEEE Trans. Syst. Man Cybern.* **1992**, *22*, 755–771. [CrossRef]
34. Kalera, M.K.; Srihari, S.; Xu, A. Off-line signature verification and identification using distance statistics. *Int. J. Pattern Recognit. Artif. Intell.* **2004**, *18*, 1339–1360. [CrossRef]

35. Daqrouq, K. Wavelet entropy and neural network for text-independent speaker identification. *Eng. Appl. Artif. Intell.* **2011**, *24*, 796–802. [CrossRef]

36. Ganchev, T.; Tasoulis, D.; Vrahatis, M.; Fakotakis, D. Generalized locally recurrent probabilistic neural networks with application to text-independent speaker verification. *Neurocomputing* **2007**, *70*, 1424–1438. [CrossRef]

37. Bayes, T. An essay towards solving a problem in the doctrine of chances. *Philos. Trans. R. Soc. Lond.* **1963**, *53*, 370–418. [CrossRef]

38. Mallat, S. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **1989**, *11*, 674–693. [CrossRef]

39. Souani, C.; Mohamed, A.; Kholdoun, T.; Rached, T. VLSI design of 1-D DWT architecture with parallel filters. *Integration* **2000**, *29*, 181–207. [CrossRef]

40. Ghaffari, A.; Golbayani, H.; Ghasemi, M. A new mathematical based QRS detector using continuous wavelet transform. *Comput. Electr. Eng.* **2008**, *34*, 81–91. [CrossRef]

41. Tareq, M.; Hamza, T.; Radwan, E. Off-line Handwritten Signature Recognition Using Wavelet Neural Network. *Int. J. Comput. Sci. Inf. Secur.* **2010**, *8*, 1–9.

42. Vivaracho-Pascuala, M.; Faundez-Zanuyb, M.; Pascualc, J. An efficient low cost approach for on-line signature recognition based on length normalization and fractional distances. *Pattern Recognit.* **2009**, *42*, 183–193. [CrossRef]

43. Avci, D. An expert system for speaker identification using adaptive wavelet sure entropy. *Expert Syst. Appl.* **2009**, *36*, 6295–6300. [CrossRef]

44. Avci, E. A new optimum feature extraction and classification method for speaker recognition: GWPNN. *Expert Syst. Appl.* **2007**, *32*, 485–498. [CrossRef]

45. Broumandnia, A.; Shanbehzadeh, J.; Varnoosfaderani, M. Persian/Arabic Handwritten Word Recognition Using M-band Packet Wavelet Transform. *Image Vis. Comput.* **2008**, *26*, 829–842. [CrossRef]

46. Daqrouq, K.; Azzawi, A.K.Y. Average framing linear prediction coding with wavelet transform for text independent speaker identification system. *Comput. Electr. Eng.* **2012**, *38*, 1467–1479. [CrossRef]

47. Daqrouq, K.; Abu Sbeih, I.; Daoud, O.; Khalaf, E. An investigation of speech enhancement using wavelet filtering method. *Int. J. Speech Technol.* **2010**, *13*, 101–115. [CrossRef]

48. Specht, D.F. Probabilistic neural networks. *Neural Netw.* **1990**, *3*, 109–118. [CrossRef]

49. Specht, D.F. Enhancements to probabilistic neural networks. In Proceedings of the IEEE International Joint Conference on Neural Networks, Baltimore, MD, USA, 7–11 June 1992.

50. Behroozmand, R.; Almasganj, F. Optimal selection of waveletpacket-based features using genetic algorithm in pathological assessment of patients' speech signal with unilateral vocal fold paralysis. *Comput. Biol. Med.* **2007**, *37*, 474–485. [CrossRef] [PubMed]

51. Yang, Z.R.; Chen, S. Robust maximum likelihood training of heteroscedastic probabilistic neural networks. *Neural Netw.* **1998**, *11*, 739–748. [CrossRef]

52. Dong, Z.C. Preclinical Diagnosis of Magnetic Resonance (MR) Brain Images via Discrete Wavelet Packet Transform with Tsallis Entropy and Generalized Eigenvalue Proximal Support Vector Machine (GEPSVM). *Entropy* **2015**, *17*, 1795–1813.

53. Ketabdar, H.; Richiardi, J.; Drygajlo, A. Global feature selection for on-line signature verification. In Proceedings of the 12th International Graphonomics Society Conference, Salerno, Italy, 26–29 June 2005.

54. Quan, Z.H.; Huang, D.S.; Xia, X.L.; Lyu, M.R.; Lok, T.-M. Spectrum analysis based on windows with variable widths for online signature verification. In Proceedings of the 18th International Conference on Pattern Recognition, Hong Kong, China, 20–24 August 2006.

55. Sengur, A. Wavelet transform and adaptive neuro-fuzzy inference system for color texture classification. *Expert Syst. Appl.* **2008**, *34*, 2120–2128. [CrossRef]

56. Avci, E.; Hanbay, D.; Varol, A. An expert discrete wavelet adaptive network based fuzzy inference system for digital modulation recognition. *Expert Syst. Appl.* **2006**, *33*, 582–589. [CrossRef]

57. Ji, G.L. Fruit Classification by Wavelet-Entropy and Feedforward Neural Network trained by Fitness-scaled Chaotic ABC and Biogeography-based Optimization. *Entropy* **2015**, *17*, 5711–5728.

58. Blumenstein, M.; Miguel, A.F.; Vargas, J.F. The 4NSigComp2010 off-line signature verification competition: Scenario 2. In Proceedings of the 12th International Conference on Frontiers in Handwriting Recognition, Kolkata, India, 16–18 November 2010; pp. 721–726.

59. Sulong, G.H.; Ebrahim, A.Y.; Jehanzeb, M. Offline handwritten signature identification using adaptive window positioning techniques. *SIPIJ* **2014**, *5*. [CrossRef]
60. Sammut, C.; Webb, G. Encyclopedia of Machine Learning. In *Leave-One-Out Cross-Validation 2010*; Springer: Berlin, Germany, 2010; pp. 600–601.