Offline Handwritten Signature Verification through Network Radial Basis Functions optimized by Differential Evolution

Saulo Henrique Leôncio de Medeiros Nápoles
Informatic Center
Federal University of Pernambuco
Recife, Brazil
shlmn@cin.ufpe.br

Cleber Zanchettin
Informatic Center
Federal University of Pernambuco
Recife, Brasil
cz@cin.ufpe.br

Abstract—**The handwritten signature is present in all important documents. In law, if the signature on a document is false, this document is also considered a fraud. This paper uses a neural network of radial basis function optimized by Differential Evolution Algorithm with features that best discriminates between a genuine signature of a simulated forgery. The experiments with this promising technique were made with a GPDS-300gray images base and the results subjected to statistical tests with the performance of technical literature.**

Keywords-component; Handwritten Signature; Off-line Verification; Radial Basis Function; Differential Evolution

I. INTRODUCTION

Throughout the history, the documents and signatures on them ever needed the guarantee of its authenticity. In the Middle Ages, the royal documentation came with your stamp ensuring its authenticity. In Visigothic Legal system, the most intellectual branch of the Germanic law, there was the confirmation document by witnesses who have touched it, signed and superscripted. Private documents were, on occasions, confirmed by royal documents [1].

The signature is brand present in all important documents, since payments for checks or credit cards, lawsuits, contracts and commitments to the business of the most varied. In a document, the signature proves that the subscriber is in agreement with what is written above, so the identification and authentication of digital documents is a main aspect to be considered to ensure the security and authenticity of their information, because the current rise of the information in the digital world has demanded more secure methods to protect and the falsity of documents and larceny, in all its aspects, constitute a delict, and if the signature on a document is false, this document is also considered invalid [2].

A way to validate signature is using Verification System, that aim examine the authenticity of the handwritten signature through of methods that can discriminate it from a forgery [3]. These systems can examine the signature so online or offline. The online approach requires a specific hardware, which can be a digitizing tablet, instrumented pen or other specialized hardware for the individual sign, so the dynamic characteristics are also observed as velocity, force and pressure. In the offline approach, the signature is done on paper and then scanned into the static characteristics are extracted, and then make a similar check to check images [4].

The forgeries are classified into three subgroups: random, simple and simulated. A random forgery is usually a genuine sample of another author. The simple forgery occurs when the simple forger knows the author's name, but does not have an example of the signature which he plans to forge. Finally, the simulated forgery occurs when the forger has an example of the signature and makes an imitation of the genuine signature [5].

The signature verification consists of, from a copy of the signature, check whether it belongs to the author. In other words, if it is write or not by the supposed author. This type of scan encounters two difficulties: The variation interpersonal and intrapersonal. The first case part of the two people do not have the same writing, already the second believes that nobody writes exactly twice, as can be seen in Figure 1 with two genuine signatures overlapped [1].

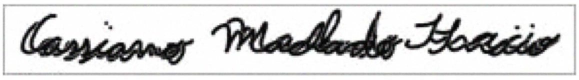

Figure 1. Variation intrapersonal of genuines signatures

The objective of this work is to develop an artificial intelligence technique, applied to signature verification, which goes beyond the traditional techniques through the use of a hybrid model that combines the capabilities of the RBF classifier and is perfected through the Differential Evolution optimization algorithm, as well will be easier to add new users to the system. The remaining material is organized as follows: section 2 is based on a literature of signature verification systems and classifiers related to this work. Section 3 is described which features the signature that will be adopted as the classification model was built and how it's performed the verification of signatures. Section 4 presents the results of experiments and a discussion about them. Finally, conclusions and directions for future work are presented in Section 5.

II. RELATED WORKS

This section will be a literature review in published studies about verification of signatures and radial basis functions neural networks.

978-1-4673-1490-9/12/$31.00 ©2012 IEEE

A. *Radial Basis Functions Neural Networks*

Radial basis function appeared in the '80s and is a neural network approached by viewing the design as a curve-fitting (approximation) problem in a high dimensional space. Learning is equivalent to finding a multidimensional function that provides a best fit to the training data, with the criterion for "best fit" being measured in some statistical sense. It was first introduced by Powell to solve the real multivariate interpolation problem. This problem is currently one of the principal fields of research in numerical analysis. In the field of neural networks, radial basis functions were first used by Broomhead and Lowe [6].

Guezouri [7] created a Temporal-RBF (T-RBF) to address patterns which vary over time. The model integrates the time parameter on the network to solve some forgotten features in the standard model, as the state of memory, dynamic measurements, recalling phases. For this, two models are proposed: the first is the delay time introduced only to the input neurons in the second time delay is inserted in the entrance and in the neurons of the hidden layer.

Rovira *et. al.* [8] developed an RBF network to be used when the descriptors of the patterns are given by their orders of magnitude. The qualitative distance is built on the discrete structure of the absolute order of magnitude spaces. This distance is dependent on a metric structure defined in Rn. The aim is to capture the distance between the components of the patterns, locating labels with respect to the extreme magnitude.

Suresh *et. al.* [9] presents a new sequential multi-category classifier (SMC-RBF) para um problema real de classificação. The classification algorithm processes the training data one by one and builds the RBF network starting with zero hidden neuron. The growth criterion uses the is classification error, the approximation error to the true decision boundary and a distance measure between the current sample and the nearest neuron belonging to the same class. SMC-RBF uses the hinge loss function for a more accurate estimate of the posterior probability. For network parameter updates, a decoupled extended Kalman filter is used to reduce the computational overhead.

B. *Signature Verification*

A considerable number of related works to offline signature verification has been achieved recently [11] [12] [14] [15]. According to these studies, the first work in the area of off-line signature verification using neural networks was proposed by Mighell et al. [10], using 80 genuine signatures and 66 forgeries simulated produced by one individual.

Currently jobs involving Signature Verification System using features as signatures discrimination the image area, height and width ratio of the size, characteristics of direction, orientation and slope, contours, textures [11]. Can also be used as symbolic data rate continuous, interval or discrete multivalued, multivalued size, quantitative, categorical [12].

The classifiers more used in verification are Artificial Neural Networks [13], Dynamic Time Warping (DTW) [14], Hidden Markov Model (HMM) and Support Vector Machines (SVM) [15]. Some approaches use a combination of these techniques to form a new hybrid method. Can be used a genetic algorithm for choice of features to be used in a neural network [16], a combination of SVMs using Genetic Algorithm [5], among other forms and obtain satisfactory results.

Vargas *et. al.*[17] performs a classification by support vector machine with least squares using the texture image and the gray levels. The representation of the signature is made by means of an array of gray level co-occurrence matrix (GLCM), and after training, is made a model that simulates the author. Kovari *et. al.* [18] did a check through the features matching. Having the dissimilarity of the features obtained from a Dynamic Time Warping (DWT) the classification is done by two nested statistical thresholds.

Justino *et. al.* [15] use a Hidden Markov Model to generate a quantized vector for each author. Having created a lot of patterns is done using a cluster distance Euclidian and minimization of the sums of square errors in k-means algorithm. Biswas *et. al.* [19] also uses the cluster by cluster, but the algorithm used is the k-nearest neighbor. The features used were the ratio, the density of pixels, calculating the distance to the limit, relation of adjacent columns and the number of symbols within the image space.

Bajaj and Chaudhury [13] used a neural network for each feature extracted from the signature image. Upper envelope, lower envelope, vertical and horizontal projection moments were features used. The output of the classifiers are combined through the Adaline algorithm, and for each author, there is a node in the readout layer of the network. Hanmandlu *et. al.* [20] extracts the angles of certain points of the signature, then an Fuzzy CMeans create clusters of angles. These groupings are used as input to an MLP-BP network. The output layer has 3 neurons to represent the standard to which the entry belongs, because the base was composed of four writers signatures.

Arya e Inamdar [21] make a survey of techniques used in signature verification. They compare the methods with respect to the various levels of counterfeiting. They concluded that the neural network classifiers are the most used and, when it is necessary to add a new user, you will need to retrain the network. Statistical methods are used to identify random and simple forgeries, the reason these methods are suitable for describing the shape of the signature based on data graphometrics.

III. Proposed Method

The system is divided into two phases: the addition of a new author and verification of a signature. To achieve these tasks the system has modules for feature extraction, pattern storage and signature verification.

A. *Feature Extraction*

From the signature images, it is necessary to extract features that will discriminate the genuine signature of forgeries. A research in literature was made to find which characteristics distinguish a genuine from a fake signature. Some characteristics were investigated and experiments were performed with the database proposed in the competition of signature verification the International Conference on Document Analysis and Recognition (ICDAR) 2011 [22].

The characteristics were subjected to tests: proportion and distribution of pixels, tilt, pressure, and centroid. The proportion is the ratio between the height and width of the image, the distribution of pixels is made through a grid which subdivide the image and it is counted how many pixels are expressed in the current subimage. This feature incorporates a static descriptor, which provides an insensitivity to intrapersonal variations. For this feature, the image was subdivided into 12 subimages through a 3x4 grid so that this analysis is done locally in the image. A local analysis in the inclination on the most significant segments is performed and returns the integer, based on chain-code with eight neighbors, which have higher correlation with this segment. The pressure is a pseudo-dynamic feature, because pressure can only be purchased online approach, you can do the simulation using the equation:

$$\theta_{hpr} = g_{min} + 0.75(g_{max} - g_{min}) \qquad (1)$$

Where θ_{hpr} is the threshold pressure and g_{min} and g_{max} represent the minimum and maximum gray level in the image [23]. The result of this operation can be seen in Figure 2. It was subdivided into a 3x2 grid to obtain its distribution. Finally, the centroid is the gravitational center of the signature, consisting of a central point in space of the signature. These variables were chosen based on studies of Justin *et. al.* [24], which would cover the global features, statistics and dynamics of a signature.

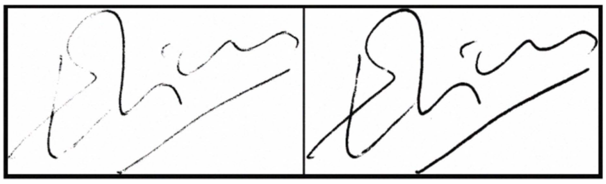

Figure 2: Signature Pressure

To verify which features are most discriminating, a Reservoir Computing Network was built to classify genuine signatures and forgeries using as input the features individually and combined. From the results, we selected the features that have lower error rate and shorter processing times. This type of network is improved by a recurrent neural network. By adding recurrent connections, the network training becomes more difficult because it is transformed into a very complex dynamical system [25]. To minimize this complexity, three independent studies have proposed a solution that had much in common: avoid the problem of a recurrent neural network formation, while still being able to use its powerful processing time. Reservoir Computing uses the network as a reservoir, whose weights are not changed in the training and are chosen randomly. The response of the reservoir is seen from the output layer by a simple classification using a linear evaluation function [26]. This technique was chosen because it is a powerful and fast, able to discriminate the figures as similar as simulated forgeries and genuine signatures.

Reservoir computing tests was made using 5-fold cross-validation. For comparison purposes, we will use the Equal Error Rate (EER), and two classic errors: Type I error or False Rejection Rate (FRR), when an authentic signature is rejected and the type II error or False Acceptance Rate (FAR) is accepted as a forgery. The approach presented the best result was using the variable distribution of pixels, pressure and proportion, like can be seen in the table below.

TABLE 1. RESULTS OF THE VARIABLE PERFORMANCE

Distribution + Pressure + Proportion	**Mean**	**Standard desviation**	**Median**
EER	0.04567	0.017123	0.04285
FRR	0.9762	0.023513	0.971
FAR	0.0236	0.027354	0.0115
Tempo	0.292	0.016	0.295

Each image is binarized. Then, the chosen characteristics are extracted and stored in an array of real values.

B. Pattern Storage

After extracting the characteristics of all signature images, we need to define a template that represent the author. For this, the vectors are subjected to an RBF network which will return an approximation of these features through the centers resulting from the training that will be used as the profile of the author.

The choice of an RBF network is justified because, as reported by Arya and Inamdar [17], the training of new authors make the use of an MLP network, this can result in increasing the size and response time of the network. because it does not have this problem, a network of radial basis function (RBF) can be used, since this network is used to approach, clustering, interpolation and mixing models. The network has one hidden layers with 40 neurons where each one implements multiquadric function. The width of the receptive field and distance will be adjusted to improve the network. The output layer is the weighted sum of outputs of the hidden layers.

Network training is made with the feature vectors of each author. The database is divided into three to training, testing and tuning. The network input is the feature vector of each author. Each instance is submitted to the network and, at the end, the readout is the function that defines the user. The result of the training error is used as the fitness function in algorithm of the optimization Differential Evolution (DE). This approach aims to improve network performance, since, with the particularities of each author, a dynamic that would be recommended for a system with many authors.

The Differential Evolution (DE) is an improved version of Genetic Algorithm for fast optimization. The individual configuration simulates a network composed of a double representing the actual width of the receptive field and distance. Two individuals are stochastically generated is made to differentiation and a third target mutation. In possession of the difference between the first two, is multiply by a weighting factor, the result is summed with the individual mutation and then the crossing is made and submitted to the junction with the existing one. The selection is made with elitism of fitness of all the new population. This time is used the second part specifies the database for the adjustment of the network, thus avoiding the overfitting.

The Differential Evolution (DE) is an improved version of Genetic Algorithm for fast optimization of the population. The idea is to generate stochastically vectors of experimental parameters to disturb the vector population. The individual simulates a network configuration and is composed of a two number representing the actual width of the receptive field and far away. The error rate is the fitness of the optimizer.

From the result of performance in training the classifier, two individuals are randomly chosen to be made to a third differentiation and targeted mutation. The difference between the first two is made by multiplying a weighting factor, the result is summed with the individual mutation and then the crossing is done. Installation of a new individual, a sample of feature vectors is chosen randomly to calculate the fitness. The individual is chosen to present the best fitness, i.e., the lowest error rate. The algorithm terminates when we have 50 executions or when the result did not show better solution. The Figure 3 illustrates the functioning of this module.

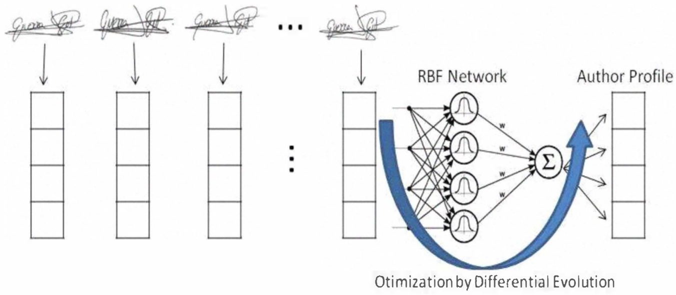

Figure 3: Pattern Storage Module

Having the best configuration for the user, the standard of the author are the centers of the resulting RBF network, which are stored in a hash table as the key is a code to identify it and the search to be optimized.

C. *Verification*

Unlike the task of identifying where you want to know the author is required and a comparison of 1: N, the verification wish to check if the signature really belongs to the author, so we have a 1:1 comparison. Having the signature test image and the user to which it belongs, the system will compare the input with the stored user template. Initially the features are extracted from the signature, then calculate the difference between the test and stored pattern, if the multiplication of differences is below the threshold of 0.2, defined empirically, the input signature is genuine. Otherwise, we cannot say that it belongs to the user in question.

IV. Experimental Results

To validate the system, the database-GPDS 300gray signature database [17] was used to contain a greater number of users. The base has 24 genuine signatures of 253 users totaling 6072 images and 7495 forgeries images. The signatures are in PNG format and were scanned at 600 dpi. In the tests were only random forgeries and went with 10-fold cross-validation, which is a selected sample of data to conduct the training, and the others are used in the test. Also collected were the results of an traditional RBF network without optimization the differential evolution to measure the impact of this approach. The results were collected and the average was taken between them and can be seen in Table 2 which also includes the performance of the method of Vargas *et al.* [17] published in Pattern Recognition.

TABLE 2. Results of the Proposed Method Performance

Measure	Proposed Method	Vargas *et. al.* [17]	RBF Standard
EER	0.68	0.74	15.48
FRR	3.42	5.56	6.78
FAR	0.4	0.04	10.62

The improvements of the proposed method can be seen when we compare its performance with the performance of an RBF network and the normal results of Vargas *et al.* [17] which uses the same image database. And, statistically, will be a test of hypotheses.

Hypotheses tests were performed of the three paired measures evaluated with a significance level of 95%, taking as the null hypothesis: H_0: $\mu_P = \mu_C$ and as alternative hypothesis H_1: $\mu_P < \mu_C$ where μ_P is a measure of the proposed method and μ_C as a competitor. Table 3 shows the results of tests performed Wilcox. In the test for the true positive rate, the alternative hypothesis had changed the sign of the inequality.

TABLE 3. Results of Statistical Tests

	Vargas *et al*[13]		RBF Standard	
Measure	**Pvalue**	**Decision**	**Pvalue**	**Decision**
EER	0,0341	Reject H_0	0,0057	Reject H_0
FRR	$932{*}10^{-6}$	Reject H_0	$176{*}10^{-7}$	Reject H_0
FAR	0,2376	Not Reject H_0	0,0021	Reject H_0

Having the values of statistical tests, we can state that the proposed method is very promising in signature verification systems because their performance was superior in all aspects when compared to other techniques for signature verification using the same database.

This performance was achieved by the right choices and analyzed the characteristics of the classifier. Another favorable factor was the optimization of the RBF network which will return a value close to the optimum. This technique shows promising because also it is easy to add a new user.

V. Conclusion

During this work it was shown how to perform the offline handwritten signature verification using a approach to the optimization of an RBF neural network.

For this aim, we extracted the features most commonly used in literature, which showed better ability to discriminate when tested with simulated forgery, using a network Reservoir Computing. Then the patterns were generated for each individual using a RBF network that fit to get the best results through an algorithm of ED. To compare the results of this methodology, we used a traditional RBF network and the balance of the experiments published by Vargas *et al.* [17]. These results were compared with statistical test and the proposed method was superior to others because of features selected and the optimization made in the RBF network.

Finally, as a suggestion for future work to further improve performance, we can use other types of variables, such as distributions or histograms and Symbolic Data.

ACKNOWLEDGMENT

This work received financial support from the Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE).

REFERENCES

[1] Oliveira, L. E. S., Justino, E., Freitas, C. O. A., Sabourin, R. "The Graphology Applied to Signature Verification". In: 12th Conference of the Internatioal Graphonomics Society, 2005, Salerno.

[2] Huang K., Yan, Hong. "Off-line Signature Verification based on Geometric Feature Extration and Neural Network Classification". Pattern Recognition, Vol 30, No. 1, 1997 pp. 9-17

[3] Heinen, M. R.; Osorio, F. S. "Handwritten Signature Autentication using Artificial Neural Network". In: IJCNN - IEEE Intenational Joint Conference on Neural Networks, 2006, Vancouver. Proceedings of the WCCI (World Congress on Computational Intelligence) - IJCNN. Vancouver - Canadá : IEEE Press, 2006. v. 1. p. 10111-10118.

[4] Sisodia, K. Anand, M. "Off-line Handwritten Signature Verification using Artificial Neural Network Classifier". International Journal of Recent Trends in Engineering, Vol 2, no. 2, November 2009.

[5] Bertolini, D, Oliveira, L. S. Justino, E. Sabourin, R. "Reducing Forgeries in Writer-independent Off-line Signature Verification Through Ensemble of Cassifier". Pattern Recognition, Vol 43, 2010 pp. 387-396.

[6] Bors, A. G. Introduction of the Radial Basis Function (RBF) Networks. http://www-users.cs.york.ac.uk/adrian/Papers/Others/OSEE01.pdf

[7] Guezouri M. "A new Approach Using Temporal Radial Basis Function in Chronological Series". Turkish Journal of Electrical Engineering and Computer Sciences, VOL.16, NO.2 2008

[8] Rovira, X. Agell, N. Sánchez, M. Prats, F. Parra, X. "Na Approach to Qualitative Radial Basis Function Networks over Orders of Magnitude" Proceedings of the 18th International Workshop on Qualitative Reasoning. Evanston, Illinois, USA (2004).

[9] Suresh S., Sundararajan N., Saratchandran P. "A Sequential multi-category classifier using radial basis function networks" Neurocomputing 71 (2008) 1345–1358. June 2007.

[10] Mighell, D. A.; Wilkinson, T. S.; Goodman, J. W. "Backpropagation and Its Application to Handwritten Signature Verification". Advances in Neural Information Processing Systems 1, 1989 pp. 340–347.

[11] Dimauro, G. Impedovo, S. Lucchese, M. G. Modugno, R. Pirlo, G. "Recent Advancement in Automatic Signature Verification". Proceedings of 9th International Workshop on Frontiers in Handwritting Recognition (IWFHR), 2004 pp. 179-184.

[12] Prakash, H. N. e Guru, D. S. "Relative Orientations of Geometrics Centroids for Off-line Signature Verification". International Conference on Advanced Pattern Recognition (ICAPR-2009), ISI, Kolkata, India 2009 pp. 201-204.

[13] Bajaj, R. e Chaudhary, S. "Signature Verification using mutiple neural classifiers". Pattern Recognition, Vol 30, 1997 pp. 1-87.

[14] Shankar, A. P. e Rajagopalan "Off-line Signature Verification using DWT". Pattern Recognition Letters, Vol 28, 2007 pp. 1407-1414.

[15] Justino, E. J. R. Bortolazzi, F. e Sabourin, R. "A Comparasion of SVM and HMM classifiers in the Off-line Signature Verification". Pattern Recognition Letters, Vol 26, Issue 9, 2005 pp. 1377-1385.

[16] Xuhua, Y., Furuhashi, T., Obata, K., Uchikawa, Y. (1997) Selection of Features for Signature Verification Using the Genetic Algorithm. Computers & Industrial Engineering Vol. 30, No. 4, pp. 1037-1045.

[17] Vargas, J. F., Ferrer, M. A. Travieso, C. M. Alonso J. B. Off-line Signature Verification based on grey level information using texture features. Pattern Recognition (2010), in: Pattern Recognition, ISSN:0031-320, vol 44, no. 2, pp 375-385.

[18] Kovari, B. Kertész, Z. Major, A. Off-line Signature Verification Based on Feature Matching. In 11th International Conference on Intelligent Engineering Systems, 2007. Budapest, Hungary.

[19] Biswas, S. Kim, T. Bhattacharyya, D. "Features Extraction and Verification os Signature Image using Clustering Technique. In: International Joournal of Smart Home. 2010. Vol 4, no. 3. July, 2010.

[20] Hanmandlu, M. Madasu, V.K. Madasu, S. "Neuro-fuzzy Aproaches to Signature Verification" Proceedings of the Second National Conference on Document Analysis and Recognition. *The Second National Conference on Document Analysis and Recognition,* 2003. Mandya, India..

[21] Arya, M. S. Inamdar, V. S. "A Preliminary Study on Various Off-line Hand Written Signature Verification Approaches. In: International Journal of computer Applications (2010) Vol 1, No. 9.

[22] Sig Comp 2011 Disclaimer. Available at: <http://www.dfki.de/~liwicki/SigComp2011/disclaimer.pdf>

[23] Vargas, J. F. Ferrer, M. A. Travieso, C. M. Alonso, J. B. "Off-line Signature Verification Based on High Pressure Polar Distribution". In: ICFHR08, Montreal, 2008.

[24] Edson J. R. Justino, Flávio Bortolozzi, Robert Sabourin, "Off-line Signature Verification Using HMM for Random, Simple and Skilled Forgeries," icdar, pp.1031, Sixth International Conference on Document Analysis and Recognition (ICDAR'01), 2001

[25] Holzmann G. "Reservoir Computing: A Powerful Black-Box Framework for Nonlinear Audio Processing". Proceedings of the 12th International Conference on Digital Audio Effects (DAFx-09), Como, Italy, 2009

[26] Schrauwen, B. Verstraeten, D. Campenhout, J. V. "An overview of reservoir computing: theory, applications and implementations". Proceedings of the 15th European Symposium on Artificial Neural Networks (2007) pp. 471-482.