A New Handwritten Signature Verification System Based on the Histogram of Templates Feature and the Joint Use of the Artificial Immune System with SVM

Yasmine Serdouk, Hassiba Nemmour[(✉)] [ID], and Youcef Chibani

Laboratoire d'Ingénierie des Systèmes Intelligents et Communicants (LISIC),
Faculty of Electronic and Computer Sciences, University of Sciences
and Technology Houari Boumediene (USTHB), 16111 Algiers, Algeria
{yserdouk,hnemmour,ychibani}@usthb.dz

Abstract. Verifying the authenticity of handwritten signatures is required in various current life domains, notably with official contracts, banking or financial transactions. Therefore, in this paper a novel histogram-based descriptor and an improved classification of the bio-inspired Artificial Immune Recognition System (AIRS) are proposed for handwritten signature verification. Precisely, the Histogram Of Templates (HOT) is introduced to characterize the most widespread orientations of local strokes in handwritten signatures, while the combination of AIRS and SVM is proposed to achieve the verification task. Usually, using the k Nearest Neighbor rule, a questioned signature is classified by computing dissimilarities with respect to all AIRS outputs. In this work, using these dissimilarities, a second round of training is achieved by the SVM classifier to further improve the discrimination power. In comparison with existing methods, the experiments on two widely-used datasets show the potential and the effectiveness of the proposed system.

Keywords: Artificial Immune Recognition System
Handwritten signature verification · Histogram Of Templates · SVM

1 Introduction

Handwritten signature is a biometric feature unique to each person. As it depends on physical and psychological conditions of the writer, researchers have to deal with the intra variability of the signer in order to develop robust systems for signature verification. One can mention two verification approaches: on-line and off-line. The on-line verification, in which signatures are acquired via an electronic device, considers dynamic information of signatures. In the off-line approach, signatures are written on a sheet of paper. In this case, features are

calculated from the signature shape. Furthermore, the verification can be carried out according to two strategies: writer-dependent or writer-independent [1]. To authenticate genuine and forged signatures, the writer-dependent strategy develops a specific system adapted to each person's style while only one generic system is developed for all persons in the writer-independent framework.

To characterize effectively signature images, several global and local descriptors were employed during the past years. For instance, typical global features are the mathematical transforms, such as Wavelets, Ridgelets and Contourlets [2–4]. Nevertheless, local features are much preferred since they describe specific parts of signature images, which makes them robust to global shape variations [5]. In this respect, we note topological features, such as pixel density and pixel distribution, curvature features, orientation features and gradient features [1,5,6].

Moreover, various methods were developed to achieve the verification task, such as dynamic time warping, neural networks, hidden Markov models and SVM [7]. Currently, SVM is the most commonly used classifier since it can significantly outperform the others [8]. Nevertheless, the scores reported in literature are not optimal and still need improvements. Recently, many interesting mechanisms inspired from the natural immune system allowed the development of Artificial Immune Recognition Systems (AIRS) that tackle with various pattern recognition applications, such as thyroid diagnosis [9] and fault detection [10]. AIRS classification adopts a supervised learning process to create new representative data for each class, called Memory Cells (MC). Then, the k Nearest Neighbor (kNN) rule is performed over the established MC to classify test data. In [6,11,12], the authors successfully employed the AIRS classifier for off-line signature verification. However, experiments showed that the user-defined parameters must be carefully tuned for each writer's characteristics in order to achieve a competitive performance. Also, since the kNN decision depends only on the pertinence of the produced MC, a more powerful decision is conceivable.

Presently, we propose a robust system for off-line signature verification. A novel descriptor using the Histogram Of Templates (HOT) is introduced to characterize stroke orientations in signatures. For the verification step, we jointly use AIRS and SVM to overcome the shortcomings of the conventional AIRS classifier. Precisely, after the AIRS training, a set of dissimilarities is calculated between original data and the evolved MC from the AIRS training. Then, these dissimilarities are used to train a SVM to develop an automatic decision about questioned signatures. The rest of this paper is organized as follows: Sect. 2 introduces the proposed signature verification system. Experiments are presented and discussed in Sect. 3, followed by the main conclusions in the last Section.

2   Proposed Signature Verification System

The proposed Signature Verification System (SVS) is composed of a feature generation module that is based on the Histogram Of Templates and a verification module, which combines AIRS with SVM. The verification task is achieved according to the writer-dependent strategy. So, for each writer, an SVS

is developed to discriminate between genuine signatures and skilled forgeries. In this section, a detailed explanation of the feature extractor "Histogram Of Templates" is carried out. Then, a brief overview of the AIRS theory is made, followed by the details of how the combination of AIRS and SVM is performed.

2.1 Histogram Of Templates

The Histogram Of Templates (HOT) is proposed for highlighting local stroke orientations by using a set of templates. As shown in Fig. 1, sliding windows covering (3×3) pixels are applied on a signature image to count the number of pixels that fit each template [13]. The resulting counts constitute the histogram of templates. So, if we consider twenty templates, the histogram will have 20 bins. Each bin corresponds to the number of pixels P matching a template k. Presently, HOT feature is computed by considering both pixel information and gradient information. This leads to a histogram of 40 bins that combines pixel and gradient information vectors.

Pixel Information-Based HOT. For each template, if the gray value $I(P)$ of a pixel P is greater than the gray value of the two adjacent pixels, P matches the template.

$$I(P) > I(P1) \quad \&\& \quad I(P) > I(P2). \tag{1}$$

Gradient Information-Based HOT. For each template, if the gradient magnitude $Mag(P)$ of a pixel P is greater than the gradient magnitudes of the two adjacent pixels, P matches the template.

$$Mag(P) > Mag(P1) \quad \&\& \quad Mag(P) > Mag(P2). \tag{2}$$

2.2 Artificial Immune Recognition System

The Artificial Immune Recognition System (AIRS) is a bio-inspired classifier that was introduced by Watkins in [14]. Through mutations and resources competition processes, the training of AIRS generates new data that are called antibodies (or Memory Cells, MC) in order to represent variability within the classes of interest. The training algorithm considers each training signature as an antigen. Also, each generated antibody with its associated class label (genuine class or forged class) is called Artificial Recognition Ball (ARB). Note that ARBs are provisional MC that will be used during the training process to produce the final established MC. Before the beginning of the AIRS training, an initialization of the MC set is carried out by selecting randomly one training sample from each class.
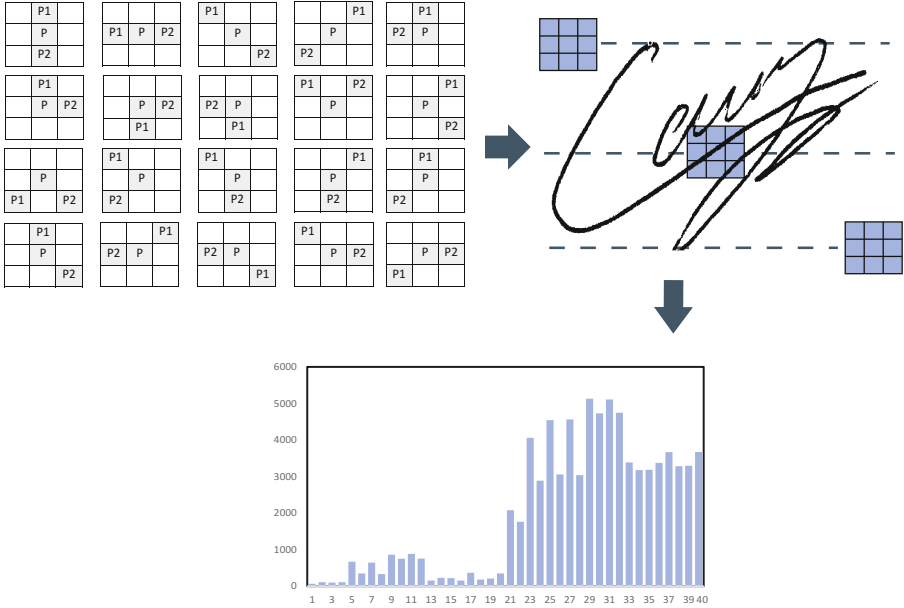
Fig. 1. HOT calculation for a signature sample.

Training Process. The training of each antigen (i.e. a training signature image) is a one-shot process that is described in what follows:

– **MC-match selection:** MC-match represents the highest stimulated MC. The stimulation ST is calculated between each MC and the actual antigen as:

$$ST(ag_i, MC) = 1 - affinity(ag_i, MC). \tag{3}$$

While *affinity* is the Euclidian distance and ag_i is the ith antigen. Then, using the selected MC-match a set of randomly mutated clones (ARBs) is generated.

– **Resources competition:** a competition between the generated clones (ARBs) is carried according to their stimulation level in order to ensure the development of more representative cells (i.e. the most stimulated ARBs that allow the recognition of antigens).

– **MC-candidate selection and MC pool update:** the ARB having the highest stimulation is selected as being the MC-candidate. Then, based on a comparison with MC-match, MC-candidate will replace MC-match or will be added to the memory cells population.

Classification. A questioned signature is classified as being genuine or forged according to its k Nearest Neighbors within the MC population.

AIRS Shortcomings. Because of the writer-dependent protocol, the AIRS training employs several set-up parameters that must be tuned according each writer's characteristics. This leads to several tests to find the optimal combination of parameters [6]. Mainly, AIRS parameters are described as follows:

- **Mutation rate:** a real taken between [0.002–0.01] that represents the mutation probability of an ARB.
- **Clonal rate:** an integer in the range [10–200] that controls the number of generated mutated clones.
- **Affinity threshold scalar:** a real taken in the range [0.1–1] used within the MC-candidate and MC-match comparison.
- **Stimulation threshold:** a real ranged between [0.1–1] used as a stopping criterion in the training routine of an antigen.
- **Resources number:** an integer between [100–700] that limits the number of mutated clones (ARBs) allowed in the system.

Furthermore, as the decision of the classical AIRS depends only on the pertinence of produced MC, we propose a hybrid verification system, in which the kNN classification is substituted by a support vector decision. This implementation allows us to globally tune the AIRS parameters for all writers while improving greatly the verification performance.

2.3   The Joint Use AIRS-SVM

The joint use AIRS-SVM as a verification system is achieved according to the following steps (see Fig. 2).

- Train AIRS according to the steps reported in Subsect. 2.2.
- Develop new training and testing sets by substituting signature features by their dissimilarities with respect to all memory cells in the MC Pool.
- Train SVM with the new training dissimilarity set to separate genuine dissimilarities from forged dissimilarities.
- Incorporate the dissimilarity vector of each questioned signature into the support vector decision to decide if it is a genuine or a forged signature.

3   Experimental Results

Our experimental study is conducted on two widely-used datasets: MCYT-75 and GPDS-300. MCYT-75[1] contains off-line signatures of 75 writers with 15 genuine and 15 skilled forgeries each. While the GPDS-300 corpus[2] contains off-line signatures of 300 writers represented by 24 genuine signatures and 30 skilled forgeries for each. Performance evaluation is based on the False Rejection Rate (FRR), the False Acceptance Rate (FAR) and the Average Error Rate (AER). This latter represents the average value between FAR and FRR.

---

[1] MCYT dataset is available on: http://atvs.ii.uam.es/databases.jsp.
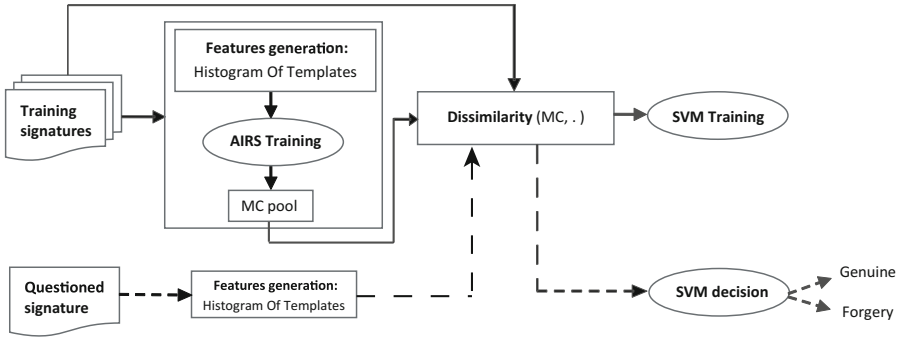[2] GPDS dataset is available on: http://www.gpds.ulpgc.es/download/.

Fig. 2. Proposed signature verification system (continuous lines indicate the training flowchart while dashed lines indicate the verification flowchart).

Following the protocol reported in [15], for each dataset, the training stage utilizes 10 genuine and 10 forged signatures that are randomly selected while the remaining signatures are used to test the verification performance. In this work, AIRS parameters take the same values for all writers to facilitate its implementation. So, as a first experiment, we tried to select the optimal k^{th} neighbor allowing the best AER on training data. As shown in Fig. 3 for both datasets, the best accuracy is obtained when considering one neighbor with an AER about 31% and 18% for MCYT-75 and GPDS-300, respectively. From these outcomes, we deduce that the kNN classification cannot deal with the variability of information offered by the evolved memory cells. Consequently, to improve the AIRS classifier, the proposed system performs a second round of training to take more advantage from both training set and memory cells to achieve a more robust verification. Table 1 reports error rates as well as the verification time using AIRS,
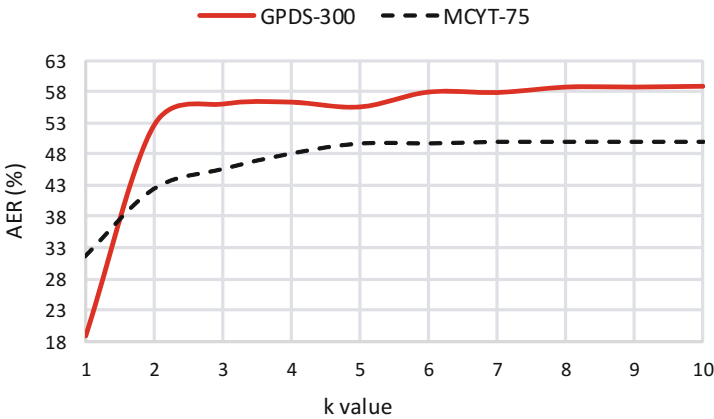

Fig. 3. AER variations according to the k values for kNN-based AIRS verification.

SVM and the proposed joint use of AIRS-SVM. For the proposed system, the verification time includes the HOT calculation, the dissimilarity computation, the SVM training and the decision time. In the conventional AIRS, the verification time corresponds to the calculation of HOT features and the computation of the KNN decision for a questioned signature while in SVM, it corresponds to the HOT computation plus the support vector decision time.

The results show that the proposed joint use AIRS-SVM outperforms both the classical AIRS and SVM performances. Indeed, the AIRS-SVM combination allows a significant improvement in AER values with at least a gain 6% for MCYT-75 and 0.7% for GPDS-300. Moreover, thanks to the support vector decision, the proposed combined verification system provides lower FAR than FRR, which reflects its ability to favor the reduction of false accepted signatures. The comparison of the verification time required to treat a questioned signature reveals that the proposed combination requires approximately the same duration as AIRS or as SVM. In addition, compared to the state-of-the-art results reported in Tables 2 and 3, the joint use of AIRS with SVM provides competitive outcomes.

Table 1. Signature verification results obtained for AIRS and the proposed implementation.

Dataset	Classifier	Verification time (s)	FAR (%)	FRR (%)	AER (%)
MCYT-75	AIRS	3.13	50.40	12.53	31.47
	SVM	3.13	17.06	17.06	17.07
	AIRS-SVM	3.15	**08.80**	**13.33**	**11.07**
GPDS-300	AIRS	1.42	22.43	13.76	18.87
	SVM	1.42	12.85	10.80	12.04
	AIRS-SVM	1.44	**07.61**	**16.66**	**11.35**

Table 2. MCYT-75 state-of-the-art.

References	Features	Classifier	#Genuine signatures	AER (%)
[16]	Slant measure	Variability measure	10	22.13
[17]	Geometric centroids	Degree of authenticity	9	21.61
[18]	LDP	LS-SVM	10	11.54
Proposed system	**HOT**	**AIRS-SVM**	**10**	**11.07**

Table 3. GPDS-300 state-of-the-art.

References	Features	Classifier	#Genuine signatures	AER (%)
[15]	LDP	SVM	10	15.35
[19]	Surroundedness	MLP	24	13.76
[20]	Gradient + equimass pyramid	Adaptive feature thresholding	16	14.01
Proposed system	**HOT**	**AIRS-SVM**	**10**	**11.35**

4 Conclusion

This paper aimed to introduce a novel histogram-based descriptor to characterize off-line signatures and proposes to verify the authenticity of these signatures using a joint use of Artificial Immune Recognition System with SVM. Specifically the kNN decision, which is commonly associated with the conventional AIRS is substituted by a support vector decision. Experiments conducted on two public datasets demonstrated the effectiveness of the proposed algorithm despite of using the same parameters selection for all writers. Precisely, a gain of 20.4% for MCYT-75 and of 7.52% for GPDS-300 in the AER is achieved over the classical AIRS performance. In order to further improve the verification accuracies, the histogram of templates descriptor could be implemented in different local parts of signature images for extracting more accurate information.

References

1. Bertolini, D., Oliveira, L.S., Justino, E., Sabourin, R.: Reducing forgeries in writer-independent off-line signature verification through ensemble of classifiers. Pattern Recogn. **43**, 387–396 (2010)
2. Deng, P.S., Liao, H.M., Ho, C.W., Tyan, H.: Wavelet-based off-line handwritten signature verification. Comput. Vis. Image Underst. **76**, 173–190 (1999)
3. Nemmour, H., Chibani, Y.: Off-line signature verification using artificial immune recognition system. In: 10th International Conference on Electronics Computer and Computation (ICECCO), pp. 164–167 (2013)
4. Pourshahabi, M.R., Sigari, M.H., Pourreza, H.R.: Offline handwritten signature identification and verification using contourlet transform. In: International Conference of Soft Computing and Pattern Recognition (SoCPaR), pp. 670–673 (2009)
5. Yilmaz, M.B., Yanikoglu, B., Tirkaz, C., Kholmatov, A.: Offline signature verification using classifier combination of HOG and LBP features. In: International Joint Conference on Biometrics (IJCB), pp. 1–7 (2011)
6. Serdouk, Y., Nemmour, H., Chibani, Y.: New off-line handwritten signature verification method based on artificial immune recognition system. Expert Syst. Appl. **51**, 186–194 (2016)
7. Impedovo, D., Modugno, R., Pirlo, G., Stasolla, E.: Handwritten signature verification by multiple reference sets. In: 11th International Conference on Frontiers in Handwriting Recognition (ICFHR), pp. 19–21 (2008)

8. Justino, E.J.R., Bortolozzi, F., Sabourin, R.: A comparison of SVM and HMM classifiers in off-line signature verification. Pattern Recogn. Lett. **26**, 1377–1385 (2005)
9. Kodaz, H., Ozsen, S., Arslan, A., Gunes, S.: Medical application of information gain based artificial immune recognition system (AIRS). Expert Syst. Appl. **36**, 3086–3092 (2009)
10. Laurentys, C.A., Palhares, R.M., Caminhas, W.M.: A novel artificial immune system for fault behavior detection. Expert Syst. Appl. **38**, 6957–6966 (2011)
11. Serdouk, Y., Nemmour, H., Chibani, Y.: Topological and textural features for off-line signature verification based on artificial immune algorithm. In: 6th International Conference on Soft Computing and Pattern Recognition (SoCPaR), pp. 118–122 (2014)
12. Serdouk, Y., Nemmour, H., Chibani, Y.: An improved artificial immune recognition system for off-line handwritten signature verification. In: 13th International Conference on Document Analysis and Recognition (ICDAR), pp. 196–200 (2015)
13. Tang, S., Goto, S.: Histogram of template for human detection. In: International Conference on Acoustics Speech and Signal Processing (ICASSP), pp. 2186–2189 (2010)
14. Watkins, A.B.: AIRS, A resource limited artificial immune classifier. Master's thesis, Faculty of Mississippi state university, USA (2001)
15. Ferrer, M.A., Vargas, J.F., Morales, A., Ordonez, A.: Robustness of off-line signature verification based on gray level features. IEEE Trans. Inf. Forensics Secur. **7**, 966–977 (2012)
16. Alonso-Fernandez, F., Fairhurst, M.C., Fierrez, J., Ortega-Garcia, J.: Automatic measures for predicting performance in off-line signature. In: International Conference on Image Processing (ICIP), pp. 369–372 (2007)
17. Prakash, H.N., Guru, D.S.: Geometric centroids and their relative distances for off-line signature verification. In: 10th International Conference on Document Analysis and Recognition (ICDAR), pp. 121–125 (2009)
18. Grupo de Procesado Digital de Senales. http://www.gpds.ulpgc.es/download/index.htm
19. Kumar, R., Sharma, J.D., Chanda, B.: Writer-independent off-line signature verification using surroundedness feature. Pattern Recogn. Lett. **33**, 301–308 (2012)
20. Larkins, R., Mayo, M.: Adaptive feature thresholding for off-line signature verification. In: 23rd International Conference on Image and Vision Computing, pp. 1–6 (2008)