Denoising Autoencoder for Iris Recognition in Noncooperative Environments

Eduardo Luz[1(✉)] and David Menotti[1,2]

[1] Computing Department, Federal University of Ouro Preto, Ouro Preto, MG, Brazil
{eduluz,menottid}@gmail.com
[2] Informatics Department, Federal University of Paraná, Curitiba, PR, Brazil

Abstract. The iris is considered as the most unique phenotype feature visible in a person's face and has been explored in the last three decades. Outstanding approaches are known for iris recognition task when the image is acquired in a well controlled environment. However, the problem is still challenging in a noncooperative environment. Having this context in mind, and from a learning representation perspective, in this paper, we propose the use of denoising autoencoders networks to create descriptors to iris recognition. We extract features from six regions of the iris and also use a specific scheme in the literature that employ a set of thresholds for iris acceptance / rejection. We perform experiments on two well-know databases, by comparing our descriptor with 2D Gabor and Wavelet representations of implementations of us. In both data sets, the proposed descriptor outperforms these features, and presents comparable results with the best performing method in a NICE contest.

Keywords: Learning representation · Denoising autoencoders · Neural networks · Iris recognition · Noncooperative environment

1 Introduction

Modalities acquired from the eye are considered as the most unique phenotype feature visible in a person's face [4]. It is composed of particular and random textures for each individual, which are even different for each eye of the same person, which increases its uniqueness (possibility of 1 among 10^{72} individuals), making even difficult to fraud. In constrained environments, in short distances, and smaller databases, there are currently efficient and effective methods for correct (or almost perfect) iris identification [3,1]. Nonetheless the main question is how to identify it on adverse conditions in natural environments.

Thus a major problem in using iris for person identification is in its recognition on noncooperative environment where the images could be acquired at an uncontrolled distance, moving people, and use of some accessories such as lenses and glasses, among others. Difficulties also increases when the system runs on a large dataset, with degraded images by light reflections and other noise, where a misidentification may cause great damages.

A. Pardo and J. Kittler (Eds.): CIARP 2015, LNCS 9423, pp. 200–207, 2015.
DOI: 10.1007/978-3-319-25751-8_25

In general, an iris recognition system is divided in six main steps such as: image acquisition; iris segmentation; normalization; feature extraction; representation of features; and classification. How to represent the information presented in the iris image (feature extraction step), specifically on degraded ones, are often discussed and investigated in works related to iris recognition in noncooperative environments.

Proença & Alexandre [12] divide the iris into two independent blocks, one composed of four sub-regions and the other composed of two sub-regions. Although the sub-regions of the same block are independent (non-overlapping), there is overlap among the sub-regions of distinct blocks. The rationale for this division is the robustness to noisy environments and loss of biometric signature. From these sub-regions, six dissimilarity values extracted using a 2D Gabor descriptor are obtained and then fused by means of a classification rule, which is also employed in our work and involves an ordered set of optimized thresholds obtained to minimize the false rejection (false negative) and false acceptance (false positive) rates.

Szewczyk et al. [14] initially segmented the iris image and then pre-processed, using the following steps: blue channel removal; conversion to monochrome images; histogram equalization; and removal of reflections, eyelashes, and occlusions caused by eyelid. They also analyze and choose the best Wavelet function for iris feature extraction.

Other works in the literature coping with iris images in noncooperative environment are focused on effectiveness comparison of different strategies. In [5], the features extracted from the LoG-Gabor filters, Haar wavelet, Discrete Cosine Transform (DCT), and Fast Fourier Transform (FFT) are compared. Marsico et. al. [7] combine two techniques, Linear Binary Patterns (LBP), which produces a local texture description, and Discriminable Textons (BLOBs), which highlights uniqueness of the texture (furrows, crypts and spots), and verify that the resulting method increases the final recognition performance.

Despite the large number of techniques presented in the literature for iris recognition in noncooperative environments, here we propose the use of Denoising Autoencoder Neural Network (DAeNN) to extract features following the classification scheme proposed in [12]. To the best of our knowledge the DAeNN technique for iris feature extraction has been used here for the first time. The rationale to use DAeNN in the noncooperative scenario is it capability to deal with noise, and it is the main contribution to the literature of our work and brings promising results as stated below. To validate our proposal the databases CASIA-IrisV4 database [2] and UBIRIS.v2 [11] are used. The CASIA-IrisV4 database is very popular among iris recognition methods and UBIRIS.v2 database includes images captured in unconstrained conditions. By comparing our descriptor with 2D Gabor [12,16] and Wavelet [14] representations of implementations of us, in both data sets, the proposed descriptor outperforms these features, and presents comparable results with the best performing method in a NICE contest [10].

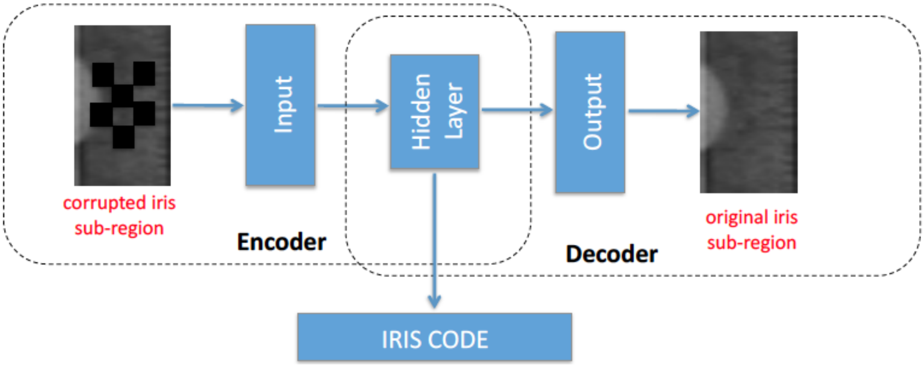

Fig. 1. Denoising Autoencoder for Iris Recognition problem.

2 Proposed Descriptor

In this section, we present the proposed descriptor for iris recognition based on denoising autoencoder representations. The steps of segmentation, normalization, and partition are straightforward and identical to the ones proposed in [12] and due to space constraints are omitted. Although the classification step used here is also similar to [12], a small modification in it is made and then it presented.

2.1 Feature Extraction

A denoising autoencoder network is employed for each iris partitions as Fig. 1 illustrates. Then, six independent biometric signatures are created, each one corresponding to a specific iris region. It is worth mentioning that we used the same autoencoder network topology to all partitions.

Autoencoder Neural Network. The Autoencoder is an unsupervised feature learning technique. Autoencoder can be considered as a neural network aiming to learn an identity function by setting the output equal to the input, with the less possible amount of error. During this process, the inputs are encoded (mapped) to a new representation, i.e., the hidden layer

$$h(x) = f(Wx + b), \tag{1}$$

where $h(x)$ is the hidden layer (the computation result), f is called activation function, x the input values, and W and b parameters are the weight matrix and the bias vector, respectively. The decoder portion of the network maps back the output of $h(x)$ to Y, i.e.,

$$Y(h(x)) = f(W_y(h(x)) + b_y), \tag{2}$$

where Y is the autoencoder output, and it should be considered as a prediction of X. While conceptually simple, the autoencoder technique has been shown to achieve state-of-the-art results in several classification problems [15,8,9], specially when deep/stacked architectures are employed.

The autoencoder architecture used here consists in one input layer, one output layer and one hidden layer. The layers are composed of artificial neurons in which each layer is fully-connected to each other. Each connection from the hidden layer is a weighted output of input neurons (an iris image) and its connections to the output layer should reproduce its own input.

The learning process of a neural network consists of obtaining the suitable weights for the connections by means of a training algorithm. Here we used a well-know backpropagation algorithm. The inner hidden layer outputs are the new feature vector (descriptor), and it can be calculated according to

$$h_{w,b}(x) = f(W^T x) = f(\gamma W_i x_i + b). \tag{3}$$

An approach to make the autoencoder more robust is to use a denoising algorithm [15]. The denoising algorithm consists in training the autoencoder with a corrupted version of the training images as input. According to Vincent et al. [15], the denoising phase is specially important to increase generalisation power of the network, by making it less vulnerable to input noise and other artefacts. Thus, this is the rationale to use autoencoder as feature extraction in noncooperative environment iris recognition, where the iris is often noisy or have some parts obfuscated.

2.2 Classification

Initially, dissimilarities between two images, for all six sub-regions $D_i = HD(I_i^1, I_i^2)$, are calculated in which $i = 1, ..., N$ ($N = 6$ is the number of sub-regions), I_i^1 the sub-region i of the image 1, I_i^2 the sub-region i of the image 2.

The Euclidean Distance (ED) is used to compute the dissimilarity between images. Given two feature vectors sets $A = \{a_1, \ldots, a_N\}$ and $B = \{b_1, \ldots, b_N\}$ the Euclidean Distance is given by

$$ED(A, B) = \sqrt{\sum_{i=1}^{N} (a_i - b_i)^2} \tag{4}$$

Given the sets of dissimilarities $D_i = [D_1, ..., D_N]$ and the thresholds $T_i = [T_1, ..., T_N]$, explained later, the next step is to count the number of $D_j \in D$ that are less or equal to T_i:

$$C(D, T_i) = \sum_{j=1}^{N} \Pi_{\{D_j \leq T_i\}}, \tag{5}$$

where, $\Pi_{\{.\}}$ is the indicator function.

The images I^1 and I^2 are classified as corresponding to the same iris if:

$$\exists_i : C(D, T_i) \geq i, i = 1, \ldots, N. \quad (6)$$

3 Experiments

In this section, we describe the experiments performed in order to validate the proposed method. Initially, we describe details of the iris images and their databases and then the results of the experiments performed are presented. The evaluation of the proposed descriptor employs the use of the *Equal Error Rate* (ERR) which in turn is defined as the point in which the *False Acceptance Rate* (FAR) is equal to the *False Rejection Rate* (FRR) [12,3]. Finally, a brief discussion of the results is made.

3.1 Databases

The CASIA-IrisV4 database [2] is one of the most popular database for iris recognition evaluation. CASIA-Iris.v4 contains a total of 54,601 iris images from more than 1,800 natural individuals and 1,000 virtual individuals. The iris images of this database have low noise influence due to its acquisition protocol, where images are acquired in a very constrained and controlled environment.

In the UBIRIS.v2 [11] database, the images were captured in unconstrained conditions, such as at different distances, in motion and in the visible wavelength, which allow images with more realistic noise factors. The UBIRIS.v2 contains 11,102 iris images from more than 261 individuals.

In our experiments, 800 images of 80 subjects (10 images per individual) of UBIRIS database are used. Same amount of images are considered for CASIA database. Worth mentioning that, for both databases, we randomly select the individuals as done in previous works [12,14,13], which difficults a fair comparison between published works. However, we propose to let available our selection of individuals and also all the source code used in a webpage once the paper is *accepted to publication* to allow further comparisons.

For both databases, images from 40 individuals were used in the training phase for estimating the classifying thresholds, lefting 400 images of the other 40 individuals for testing/evaluation.

3.2 Results

In the followings, we report the results of our experiments. Initially, we evaluate the results obtained by using the Denoising Autoencoder descriptors and, by varying some parameters of the network and the increments used to obtain the classification thresholds. Finally, we took the parameters that obtained the best results for each descriptors and evaluated it on smaller thresholds values.

We stressed that we follows the same evaluation protocol used in previous works on literature to report the results here [12,14,13]. In that sense, we estimate the classification thresholds in the training set, and in the evaluation/testing set all images of each individuals are taken as query image in a scheme similar to leave-one-out.

As several other works in the literature, we use the *equal error rate* (EER) as our effectiveness measure of analysis.

Autoencoder. In single layer denoising autoencoder neural network, several parameters should be adjusted such as: the number of neurons in the layer; the learning rate; the number of training epochs; and the percentage of corruption noise applied on input of network. For this experiment we keep the normalized input image with a fixed sized of 16×128 pixels, and only one hidden layer, as an initial guess. Thus, a gridsearch is applied in order to find the remaining parameters that will produce the best results for the training phase. We vary the number of neurons in each hidden layer in 6 possibilites [15 20 30 40 50 100]. The network learning rate and the percent of corruption noise both vary in the range [0.3-0.1] with a step of 0.1. The number of training epochs used during trainig was [3000 6000]. The gridsearch figures was proposed based on empirical experimentation previously conducted. From the EER values we have obtained, we observed that the higher the number of epochs for training, the better values DAeNN produced. Also, we observed that for CASIA-IrisV4, less input noise level produced best results while UBIRIS.v2 database required higher noise level.

We selected the parameters that have achieved the lowest EER, 0.42% for CASIA V4 and 14.94% for UBIRIS V2. It is worth noticing that best NICE:II participant report EER of approximately 12% on a database similar to UBIRIS V2. We have also run the same methodology on two widely used

Table 1. Results comparison in CASIA

Method	Input	EER (%) incr. 0.025
Proposed descritpor	16×128	0.42
2D Gabor based [12]	16×128	2.55
Wavelet based [14]	16×128	3.84

Table 2. Results comparison in UBIRIS

Method	Input	EER (%) incr. 0.025
Proposed descriptor	16×128	14.94
2D Gabor based [12]	16×128	23.14
Wavelet [14]	16×128	18.48

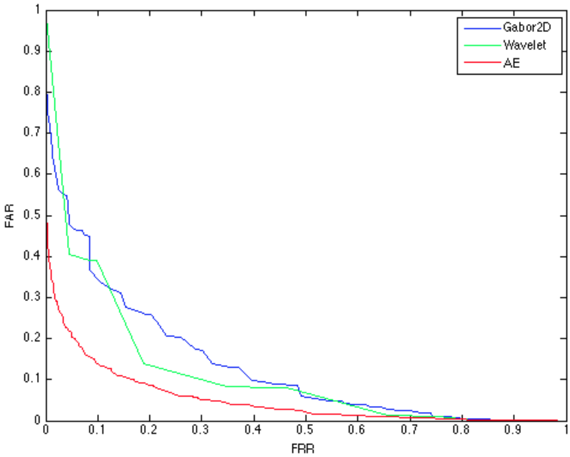

Fig. 2. The ROC curve obtained by the proposed method in the UBIRIS database considering an increment value of 0.025 in the classification threshold step. Autoencoder (AE) parameters: 30 neurons on hidden layer; noise level of 30%; learning rate of 0.1. Mother wavelet used: rbio3.1.

feature extraction techniques for the iris recognition problem: 2D Gabor [12,16] and Wavelet [14,6]. Methods based on 2D Gabor descriptor embedded commercial of-the-shelf systems for iris recognition and according to [14] wavelet transform is one of the most relevant tool for extract features for iris recognition. Fig. 2 illustrates ROC curves for our proposed descriptor and for the others in comparison. The superiority of our proposal is remarkable.

3.3 Discussion

Results presented on Tables 1 and 2 show that DAeNN can be a promising tool for the iris recognition problem, specially in noncooperative environments. Experiments show that for UBIRIS V2 database, 30% of corruption on training images during the DAeNN training phase produces the best results.

4 Conclusion

In this work we addressed the problem of iris image representation in noncooperative environment using the denoising autoencoder technique. We performed experiments on two databases (UBIRIS and CASIA), and the proposed descriptor outperforms popular feature extraction methods, such as 2D Gabor [12,16] and Wavelet transform [14]. The proposed descriptor also showed comparable results, in terms of EER, with the best performing method in a NICE.II contest [10], even on UBIRIS database. Our result suggest that denoising autoencoder can be a promising tool for iris representation. As future work we plan to investigate stacked (deep) autoencoders architectures aiming to improve iris recognition effectiveness in noncooperative environments.

Acknowledgment. The authors would like to thank to UFOP, CAPES, and CNPq (grant #307010/2014-7).

References

1. Bowyer, K.W., Hollingsworth, K., Flynn, P.J.: Image understanding for iris biometrics: A survey. Comp. Vision and Image Understanding **110**(2), 281–307 (2008)
2. Chinese Academy of Sciences (CASIA): CASIA Iris Image Database. http://biometrics.idealtest.org/findTotalDbByMode.do?mode=Iris (accessed April 2015)
3. Daugman, J.: New methods in iris recognition. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics **37**(5), 1167–1175 (2007)
4. Daugman, J.: High confidence visual recognition of persons by a test of statistical independence. IEEE TPAMI (1993)
5. Kumar, A., Passi, A.: Comparison and combination of iris matchers for reliable personal authentication. Pattern Recognition **43**, 1016–1026 (2009)
6. Li, P., Ma, H.: Iris recognition in non-ideal imaging conditions. Pattern Recognition Letters (2011)
7. Marsico, M.D., Nappi, M., Riccio, D.: Noisy iris recognition integrated scheme. Pattern Recognition Letters (2011)
8. Mohamed, A.R., Sainath, T.N., Dahl, G., Ramabhadran, B., Hinton, G.E., Picheny, M.A.: Deep belief networks using discriminative features for phone recognition. In: IEEE ICASSP, pp. 5060–5063 (2011)
9. Netzer, Y., Wang, T., Coates, A., Bissacco, A., Wu, B., Ng, A.Y.: Reading digits in natural images with unsupervised feature learning. In: NIPS Workshop on Deep Learning and Unsupervised Feature Learning, vol. 2, p. 5 (2011)
10. Proenca, H., Alexandre, L.: Toward covert iris biometric recognition: Experimental results from the nice contests. IEEE TIFS **7**(2), 798–808 (2012)
11. Proença, Hugo, Alexandre, Luís A.: UBIRIS: a noisy iris image database. In: Roli, Fabio, Vitulano, Sergio (eds.) ICIAP 2005. LNCS, vol. 3617, pp. 970–977. Springer, Heidelberg (2005)
12. Proença, H., Alexandre, L.A.: Toward noncooperative iris recognition: a classification approach using multiple signatures. IEEE TPAMI **29**(4), 607–612 (2007)
13. Song, Y., Cao, W., He, Z.: Robust iris recognition using sparse error correction model and discriminative dictionary learning. Neurocomputing (2014)
14. Szewczyk, R., Grabowski, K., Napieralska, M., Sankowski, W., Zubert, M., Napieralski, A.: A reliable iris recognition algorithm based on reverse biorthogonal wavelet transform. Pattern Recognition Letters (2011)
15. Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P.A.: Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. The J. Machine Learning Research **11**, 3371–3408 (2010)
16. Wang, Q., Zhang, X., Li, M., Dong, X., Zhou, Q., Yin, Y.: Adaboost and multi-orientation 2D Gabor-based noisy iris recognition. Pat. Rec. Letters (2011)