POLITECNICO DI TORINO
Repository ISTITUZIONALE

Authnet: Biometric Authentication Through Adversarial Learning

Original
Authnet: Biometric Authentication Through Adversarial Learning / Ali, Arslan; Testa, Matteo; Bianchi, Tiziano; Magli, Enrico. - (2019), pp. 1-6. (Intervento presentato al convegno 2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)) [10.1109/MLSP.2019.8918810].

Availability:
This version is available at: 11583/2779293 since: 2021-03-17T16:15:00Z

Publisher:
IEEE

Published
DOI:10.1109/MLSP.2019.8918810

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

25 July 2024

AUTHNET: BIOMETRIC AUTHENTICATION THROUGH ADVERSARIAL LEARNING

Arslan Ali, Matteo Testa, Tiziano Bianchi and Enrico Magli

Department of Electronics and Telecommunications
Politecnico di Torino, Italy
name.surname@polito.it

ABSTRACT

We present AuthNet: a generic framework for biometric authentication, based on adversarial neural networks. Differently from other methods, AuthNet maps input biometric traits onto a regularized space in which well-behaved regions, learned by means of an adversarial game, convey the semantic meaning of authorized and unauthorized users. This enables the use of simple boundaries in order to discriminate among the two classes. The novel approach of learning the mapping regularized by target distributions instead of the boundaries further avoids the problem encountered in typical classifiers for which the learnt boundaries may be complex and difficult to analyze. With extensive experiments on publicly available datasets, it is illustrated that the AuthNet performance in terms of security metrics such as accuracy, Equal Error Rate (EER), False Acceptance Rate (FAR) and Genuine Acceptance Rate (GAR) is superior compared to other methods which confirms the effectiveness of the proposed method.

Index Terms— Biometric authentication, Deep neural network, Latent mapping

1. INTRODUCTION

In recent years, there has been a growing interest in biometric authentication systems due to their convenience in providing access to sensitive data.

Ideally, a good biometric authentication system should maximize the probability of accepting authorized users while keeping a negligible probability of accepting wrong users. This problem has traditionally been addressed by means of handcrafted feature extraction and distance matching in the enrollment and verification phases, respectively. However, the rise of deep learning based models has showed the superiority of learned features, see e.g. [1, 2].

Nonetheless, deep learning based classification learns highly non-linear boundaries with complex shape in order to partition the feature space. As shown in [3], the geometry of the decision boundaries heavily affects the robustness of the classifier, leading to potential errors.

To overcome this, in this paper we propose to learn a compact and meaningful mapping of the input biometric traits onto the latent (feature) space. This mapping should yield a latent space that is shaped in a simple and well-behaved manner so that the region of the space corresponding to the authorized user is well-separated from that containing all the other users. More specifically, biometric traits similar to that of the authorized user should cluster in a region of the latent space and exhibit given statistical properties. In a similar way, all the biometric traits of every other possible user should gather in a region of the space which semantically represents the non-authorized users. This complex task requires to map two possibly very different input distributions onto target distributions having the same shape.

The resulting system, which we will refer to as AuthNet, employs an adversarial model to regularize the latent space such that simple threshold-based rules can be employed to discriminate between the authorized user and everyone else.

1.1. Related work

Even though AuthNet can be used in principle with any biometric trait, in this work we will focus on two of the most commonly used ones, i.e. faces and fingerprints.

Outstanding progress has been made in face recognition task with the advent of deep learning methods which, by learning the most discriminative facial features, reach state-of-art performance. Examples include Facenet [4], ArcFace [5], DeepFace [6] and others, see [7]. Other non deep learning based approaches which rely on low-dimensional representations of the faces include sparse representations [8] and manifold [9] representations.

Regarding the fingerprints, we mention [10, 11] in which the matching is performed on a global descriptor of the whole fingerprint and [12, 13] in which the matching is made on the local minutiae information. The performance can be improved when additional information such as shape context and orientation is included, see e.g. [14]. Deep learning models have also been proposed, examples include [1] in which convolutional neural networks (CNN) are used to extract minutiae from raw fingerprint images and [2] where a

stacked autoencoder is used to classify fingerprints into arch, left/right loop, and whorl.

To the best of our knowledge, this is the first approach in which learning a mapping regularized by target distributions is employed to achieve robust biometric authentication.

2. PROPOSED METHOD

2.1. Adversarial Learning

Adversarial models gained popularity in the context of generative models, with the introduction of the Generative Adversarial Networks (GANs) [15]. The main goal of a GAN is to implicitly learn the probability distribution of the input data in such a way that the network is then able to *generate* samples similar to the input data.

The main idea behind adversarial models is to reach the minimum of a functional defined as a minimax game where two entities have adversarial (opposite) goals. Within the deep learning framework, the two entities called generator and discriminator are modeled as neural networks and the minimax game is introduced in the loss function in order to make the two networks compete against each other during the training process. In more detail, the discriminator should be able to correctly discriminate between generated and real samples, while the generator should be able to generate samples which are realistic enough to fool the discriminator.

In AuthNet, as described in detail in the following, samples of the data distribution are mapped onto a latent representation which follows a target distribution. This can be considered as the inverse mapping of a conventional GAN, in which samples of a fixed distribution are mapped onto the captured distribution of the data.

2.2. Latent Mapping

In the following we formalize the main concept of AuthNet.

Let $\mathcal{B} = \{\mathcal{B}_{a=0}, \mathcal{B}_{a=1}\}$ denote the set of all possible biometric traits and $a \in \{0, 1\}$ an indicator variable such that $a = 1$ represents the authorized user and $a = 0$ represents all other unauthorized users. Moreover, let us define as $\mathbf{x} \in \mathbb{R}^n$ a generic biometric trait in $\mathcal{B}$ and as $\mathbf{z} \in \mathbb{R}^d$ its latent representation with $d < n$. The goal is to learn an encoding function of the input biometric trait $\mathbf{z} = H(\mathbf{x})$ such that $\mathbf{z} \sim \mathbb{P}_1$ if $\mathbf{x} \in \mathcal{B}_{a=1}$ and $\mathbf{z} \sim \mathbb{P}_0$ if $\mathbf{x} \in \mathcal{B}_{a=0}$, with $\mathbb{P}_1$ and $\mathbb{P}_0$ the target distributions in the latent space. If the distributions $\mathbb{P}_1$ and $\mathbb{P}_0$ are well-behaved, a simple distance-based thresholding approach can be employed to determine whether the biometric trait $\mathbf{x}$ corresponds to authorized user or not.

Let us set $\mathbb{P}_1 = \mathcal{N}(\boldsymbol{\mu}_1, \sigma_1\mathbf{I})$ and $\mathbb{P}_0 = \mathcal{N}(\boldsymbol{\mu}_0, \sigma_0\mathbf{I})$ to be Gaussian, this amounts to enclosing the energy of the latent representation of authorized and unauthorized users within hyperspheres whose radius depends on both d and the distribution parameters. For the sake of simplicity and without loss of generality, we set $\mathbb{E}[||\mathbf{z}_1||_2] < \mathbb{E}[||\mathbf{z}_0||_2]$ with $\mathbf{z}_1 \sim \mathbb{P}_1$ and $\mathbf{z}_0 \sim \mathbb{P}_0$. Then, the authentication is performed on the basis of the decision rule: if $||H(\mathbf{x})||_2 \leq \tau$ the user is authenticated; otherwise the user is rejected.

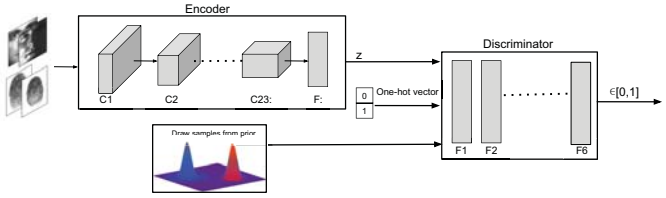


Fig. 1: AuthNet architecture at enrollment phase. Training biometric traits are given as input to the encoder which consists of 23 convolutional layers followed by a fully connected layer. The output of the encoder, together with a one-hot vector and samples of the target distributions, is given as input to the discriminator which is made of 6 fully connected layers.

During the enrollment phase, the network is trained to learn the mapping for the specific user to be enrolled. In the verification phase, a threshold decision rule is applied on the latent representation of the input biometric trait computed through the trained encoding function $H(\mathbf{x})$, in order to output a decision.

2.3. Enrollment

During the enrollment phase we want to learn an encoding function $H(\mathbf{x})$ which maps the users onto the target distributions. This optimal $H(\mathbf{x})$ should be the one for which a distance metric between $H(\mathbf{x}) : \mathbf{x} \in \mathcal{B}_{a=1}$ and $\mathbb{P}_1$, and between $H(\mathbf{x}) : \mathbf{x} \in \mathcal{B}_{a=0}$ and $\mathbb{P}_0$ is minimized. To achieve this goal we employ an adversarial model whose optimum is reached when the Jensen Shannon (JS) divergence between the latent mapping and target distribution is minimized [15].

The general AuthNet architecture at enrollment phase is depicted in Fig. 1. It consists of two competing neural networks: an encoding function $H(\mathbf{x}, \boldsymbol{\theta}_h)$ having parameters $\boldsymbol{\theta}_h$ and a discriminator $D(\mathbf{p}, \boldsymbol{\theta}_d)$ with parameters $\boldsymbol{\theta}_d$. For the sake of readability, unless needed, we will drop the parameters in the notation of the encoding and discriminator networks. The biometric traits $\mathbf{x}$ are given as an input to $H(\cdot)$ which encodes them to their latent representation $\mathbf{z}$. The discriminator $D(\mathbf{p})$ takes as input the vector $\mathbf{p} \in \{\mathbf{s}, \mathbf{z}\}$, namely it is given in an alternate fashion, either a sample from one of the target distributions $\mathbf{s}$ or the encoded latent representation $\mathbf{z}$. The vector $\mathbf{s} \in \mathbb{R}^d$ contains samples randomly drawn from the target distributions $\mathbb{P}_1$ or $\mathbb{P}_0$ if $\mathbf{x} \in \mathcal{B}_{a=1}$ or $\mathbf{x} \in \mathcal{B}_{a=0}$ respectively. The labelled information a of the user (as one-hot vector) is also given to the discriminator; this acts as a switch to select a "sub-discriminator" function for either authorized or unauthorized users. The output of $D(\mathbf{p})$ is a scalar representing the probability that the given input is coming either from the encoding function or the target distribution.

d	GAR@10^{-2}FAR%
1	99.998
2	89.884
3	53.351

Table 1: GAR comparison of a randomly selected user of Yale DB2 when considering different dimensionality of the latent space d. The best case is obtained for $d = 1$.

2.3.1. Loss Function

The loss function we consider to address the above-defined adversarial setting is given by

$$V(H, D) = \mathop{\mathbb{E}}_{\mathbf{s}\sim\mathbb{P}} [\log(D(\mathbf{s}, a))] + \mathop{\mathbb{E}}_{\mathbf{x}\sim\mathcal{B}} [\log(1 - D(H(\mathbf{x}), a))] , \quad (1)$$

which is optimized as a minimax two-player game according to $\min_{\boldsymbol{\theta}_h} \max_{\boldsymbol{\theta}_d} V(H, D)$, where the optimization is carried over the networks parameters $\boldsymbol{\theta}_h$ and $\boldsymbol{\theta}_d$ in an alternate fashion.

2.4. Authentication

The authentication phase, which directly translates to the testing phase of the neural network, only requires the use of the encoder network. Indeed, given a trained encoder network it is possible to compute the latent representation related to the input biometric trait. Then, a threshold is applied on the ℓ_2 norm of the latent representation in order to output a decision. The decision step can be formalized as follows:

$$\begin{cases} \text{accept if } ||\mathbf{z}||_2 \leq \tau, \\ \text{reject if } \ ||\mathbf{z}||_2 > \tau \, . \end{cases}$$

Selecting τ: τ is an adjustable threshold that can be varied to obtain the desired trade-off between false acceptance rate (FAR) and false rejection rate (FRR). The results presented in the following section are obtained by fixing a metric value FAR, obtaining the corresponding τ and computing the other metrics i.e. Genuine Acceptance Rate (GAR) at that threshold. The same applies to the Equal Error Rate (EER): we compute the value of τ at the intersection of FAR and FRR values. Additionally, we also compute the maximum accuracy achieved by considering different values of τ.

3. TRAINING AND IMPLEMENTATION DETAILS

As previously discussed, the AuthNet architecture is made of two neural sub-networks: the encoder and the discriminator. Since the encoder takes as input images and it is supposed to learn discriminative features, we employ a CNN. Starting from the input layer we use 23 convolutional layers with ReLU activations followed by a final fully connected layer. The first convolutional layer is made of 32 filters with a kernel size of 7×7. At layers $1, 3$ and 7 we progressively increase the number of filters and decrease the size of the kernels; at the same time we perform downsampling by using a stride factor of 2. The last convolutional layer is made of 256 filters of size of 3×3. This layer is then followed by a fully connected layer having output of size corresponding to the latent space dimensionality d. It is important to notice that in this last layer we do not use any non-linear activation as the output should behave as a sample coming from one of the target distributions. In our tests we fixed the hyperparameter $d = 1$, since in our experiments this choice gave us better results as can be seen in Table 1. As d increases, under the same FAR, the GAR decreases. The highest GAR is obtained for $d = 1$. Intuitively, as the latent space grows in dimensionality, a larger number of training samples is required to avoid overfitting. The datasets we consider have a rather small number of samples per user, thus it is not surprising that a smaller d achieves better results. The discriminator is a fully connected network consisting of 6 layers. The first layer has the input of size $d+2$ since it is the concatenation of the latent representation vector with a one hot vector indicating the class the corresponding user belongs to. The size progressively increases to a maximum of 1000 which is decreased in the further layers to the final output of size equal to 1 with sigmoid activation.

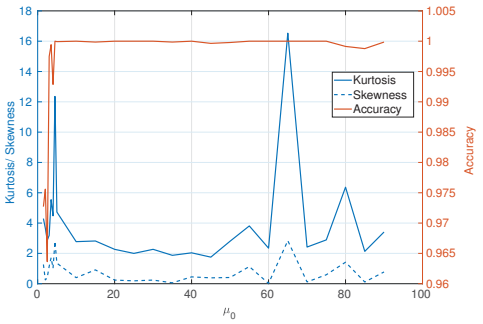


Fig. 2: Accuracy, kurtosis and skewness comparison of a randomly selected user from CMU-MultiPIE having $\mathbb{P}_0 = \mathcal{N}(k, 1)$ where $k = [0.5, 90]$, and $\mathbb{P}_1 = \mathcal{N}(0, 1)$. If the means of the two distribution are too far apart the training process gets unstable, hence it effects the accuracy, kurtosis and skewness of the imposed distributions.

Preprocessing and training parameters The network is trained using Adam optimizer using an iterative algorithm as discussed in [15]; the optimization is carried out one step for encoder and one for the discriminator. Weight decay is set to be 0.0004 and a dropout of 0.7 is used. Initially, the learning rate is set to be 0.01 for first 5000 iterations, and is then decreased by a factor of 10 after every 5000 iterations. In total, the network is trained for 30000 iterations. The only pre-processing we perform on all the considered datasets is an energy normalization of the input images.

4. EXPERIMENTS

In biometric authentication systems it is common to assume that the user puts him/herself in a controlled condition for the

Biometry	Method	EER%	GAR@10^{-1}FAR%	GAR@10^{-2}FAR%	Max accuracy
Face - Yale B	**AuthNet**	**0.426**	**99.903**	**99.890**	**99.615**
	AuthNet enc. classifier	0.961	99.310	99.310	99.252
	FaceNet	1.286	98.819	98.712	98.782
	ArcFace	0.893	99.159	99.108	99.229
Face - Multi-PIE	**AuthNet**	**0.044**	**100.0**	**100.0**	**99.983**
	AuthNet enc. classifier	0.279	100.0	100.0	99.818
	FaceNet	0.930	99.368	99.201	99.163
	ArcFace	1.811	98.811	98.125	98.738
Fingerprint FVC 2006	**AuthNet**	**0.414**	**100.0**	**99.908**	**99.630**
	AuthNet enc. classifier	1.454	99.497	99.108	98.603
	Verifinger	0.758	100.0	99.796	99.398
	Hybrid approach [14]	3.200	98.182	94.854	96.906

Table 2: Performance comparison of AuthNet with respect to other biometric authentication schemes.

biometric traits acquisition. In this regard, the datasets we consider are among the biggest ones acquired in such conditions.

4.1. Datasets

For the **face** authentication task we employ two commonly used datasets. The first dataset we employ is the CMU Multi-PIE dataset [16]. It consists of samples with different poses illumination and expressions. In total it has 337 subjects acquired in 4 different sessions. We consider the frontal poses of 129 subjects which are common in all 4 sessions. For each user enrollment 75% of the samples are used for training and the remaining 25% are left for testing. For unauthorized users out of 128 users, 96 users samples are used for the training and remaining 32 users samples are left for the testing. Keeping the aspect ratio, we resize the images to 144x192x3. Further to create more diverse samples, we employ the mixup strategy as described in [17]: positive and negative *training* samples are mixed through convex combination.

The second dataset we consider is the *extended Yale Face Database B* [18] (the cropped version). In total it has 38 subjects, with varying illumination conditions. The dataset is further split into train and test for each enrollment. For each authorized user enrollment 75% of the samples are used for training and the remaining 25% are left for testing. For unauthorized users, 31 users samples are used for training and 6 users samples are left for the testing. Further, the dataset is augmented by employing the crops of size 184×160 by an augmentation factor of $F = 81$ and $F_1 = 25$ for authorized and unauthorized users respectively. Further to create more diverse samples we employ mixup strategy as explained for CMU Multi-PIE dataset.

For **fingerprint** authentication, we evaluate our method on *Fingerprint Verification Competition (FVC 2006) DB2* [19] dataset that consists of 150 users. Maintaining the aspect ratio the images are resized to 202×149. For each authorized user enrollment, 75% of the samples are used for the training and the remaining 25%are left for the testing. For the case of unauthorized users, 124 users samples are used for the training and 25 are left for the testing. The dataset is augmented to a factor of $F = 289$ and $F_1 = 25$ by cropping the images of sizes 186×133 pixels. Lastly, mixup data augmentation is employed as explained for the face datasets.

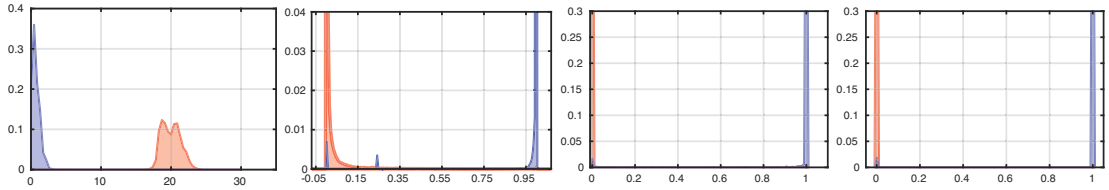
4.2. Parameters of authorized and unauthorized users distributions

In AuthNet the authorized and unauthorized target distributions are set to be Gaussian. This choice comes from the fact that the output of a (large enough) fully connected layer, by the central limit theorem, will naturally tend to a Gaussian distributed output [20]. We set the distributions to be $\mathbb{P}_1 = \mathcal{N}(0, 1)$ and $\mathbb{P}_0 = \mathcal{N}(20, 1)$. We choose $\mu_1 = 0$ and $\mu_0 = 20$ to be different enough to keep the distributions far apart from each other. In more detail, in Fig. 2 we show the maximum accuracy obtained by AuthNet, together with skewness and kurtosis of the latent representation as a function of of μ_0 for a randomly selected CMU-MultiPIE user. It can be seen that the region for which the accuracy is maximum, corresponds roughly to $15 \leq \mu_0 \leq 45$; in this region, skewness is close to 0 and kurtosis is approximately equal to 3, showing that the training indeed converges to Gaussian distributions. Further, if the difference between μ_0 and μ_1 is too large (e.g. $\mu_0 > 50$), the training process becomes unstable and the distributions are far from Gaussian.

4.3. Results

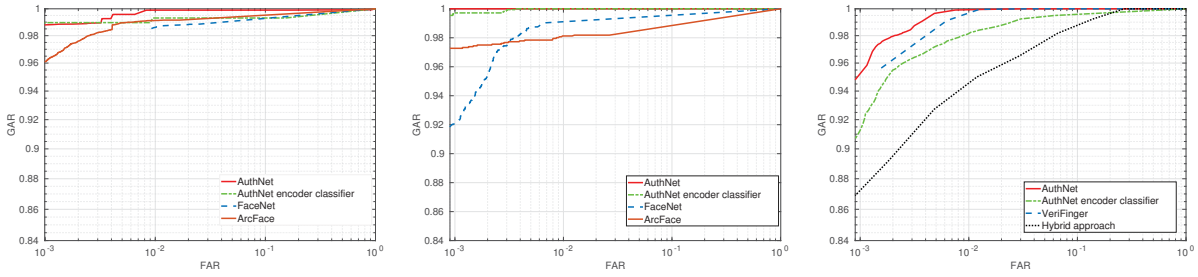
Face authentication. In the case of face authentication for benchmarking with state-of-the-art deep learning techniques we compare AuthNet with FaceNet [4] and ArcFace [5]. Regarding FaceNet and ArcFace, since it is not possible to train them from scratch because of the data scarcity, we compute the 512-dimensional embeddings of the input images given a pre-trained network on the CASIA WebFace dataset [22]. Then, a classifier is independently trained on the embeddings of each user. In order to assess the benefits of the adversarial scheme employed in AuthNet, we also include a comparison with a classifier based on the same network architecture as the AuthNet encoder, but trained with sigmoid cross entropy loss instead of the adversarial loss. In the following, this will be denoted as *AuthNet encoder classifier*.

The results in Table 2 show that AuthNet achieves high performance in all the considered metrics. We highlight that, even though the AuthNet encoder classifier shares the same



(a) AuthNet (b) AuthNet enc class. (c) FaceNet + classifier (d) ArcFace + classifier

Fig. 3: Face authentication scores for authorized users (blue) and unauthorized users (red) for Multi-PIE. (a) Histogram of $||\mathbf{z}||_2$ decision statistics of AuthNet; (b) Histogram of the sigmoid outputs of AuthNet encoder classifier; (c) Histogram of the sigmoid outputs of FaceNet embeddings classifier; (d) Histogram of the sigmoid outputs of ArcFace embeddings classifier. The plots in (b)-(c) depict a detailed view to better appreciate the leakage effects.



(a) Face dataset - Yale B (b) Face dataset - MultiPIE (c) Fingerprint dataset

Fig. 4: ROC comparison on overall results of 32 users for faces (a)-(b) and fingerprint (c) datasets. AuthNet is compared with the AuthNet encoder classifier, FaceNet [4] and ArcFace [5] in (a)-(b); with AuthNet encoder classifier, VeriFinger [21] and the hybrid approach [14] in (c). AuthNet (red) achieves highest GAR at different values of FAR.

architecture as AuthNet, it achieves an EER which is an order of magnitude larger than the one of AuthNet, especially at low FAR, (see Fig. 4(a)-(b)). This suggests that the chosen target distributions impose well-defined decision regions in the latent space yielding a more robust classification scheme.

Indeed, as can be seen in Fig. 3(a), AuthNet effectively separates authorized and unauthorized users. A good separation is also achieved by other methods, see Fig. 3(b)-(d); however they fail to assign to all the unauthorized users a correct score, yielding some "leakage" into the wrong distribution. This behavior can indeed be more clearly noticed in the Receiver Operating Characteristic (ROC) comparison in Fig. 4(a)-(b). The well-separated distributions obtained from AuthNet lead to highest GAR even at small values of FAR compared to other benchmarking methods. Indeed, it can also be noticed that the proposed approach consistently outperforms other methods over the full range of FAR values.

At this point it is also interesting to notice that AuthNet is able to achieve higher performance when tested on the Multi-PIE dataset. Even though this dataset is more complex with respect to the Yale B dataset because of the less constrained acquisition, it has more samples. For this reason, complex mappings are easier to learn.

Fingerprint authentication. For fingerprint authentication we compare AuthNet with AuthNet encoder classifier, Verifinger [21], and the hybrid approach described in [14] in which a minutiae based template synthesis and matching is employed. Note that Verifinger achieves state-of-the-art performance without using deep learning approach. AuthNet outperforms all the other approaches in terms of EER by achieving an EER of 0.414%. At fixed small values of FAR, the resulting GAR of the proposed method is higher with respect to AuthNet encoder classifier and the hybrid approach, while achieving comparable performance with Verifinger. Lastly, in Fig. 4(c) the ROC comparison of AuthNet with respect to other fingerprint authentication schemes is depicted. It can be seen that AuthNet outperforms all other methods by achieving highest values of GAR at different FAR.

5. CONCLUSIONS

We presented a novel biometric authentication scheme based on deep learning which learns a mapping onto target distributions by means of an adversarial game. All the above results support the idea of AuthNet to move from learning the classification boundaries to learning a mapping onto a space regularized by target distributions. Our intuition behind this behavior is that the non-linear boundaries learnt by standard deep learning classifiers indeed become very complex as they try to closely fit the training data, leaving room for misclassification. Conversely, the adversarial learning of AuthNet enables much simpler boundaries to be used as it does not learn how to partition the space but rather how to map the input space into the latent one.

Acknowledgment: This work results from the research cooperation with the Sony Technology Center Stuttgart (Sony EuTEC). We would like to thank Sony EuTEC for their feedback and the fruitful discussions.

References

[1] Lu Jiang, Tong Zhao, Chaochao Bai, A Yong, and Min Wu, "A direct fingerprint minutiae extraction approach based on convolutional neural networks," in *Neural Networks (IJCNN), 2016 International Joint Conference on*. IEEE, 2016, pp. 571–578.

[2] Ruxin Wang, Congying Han, and Tiande Guo, "A novel fingerprint classification method based on deep learning," in *Pattern Recognition (ICPR), 2016 23rd International Conference on*. IEEE, 2016, pp. 931–936.

[3] Alhussein Fawzi, Seyed-Mohsen Moosavi-Dezfooli, Pascal Frossard, and Stefano Soatto, "Classification regions of deep neural networks," *arXiv preprint arXiv:1705.09552*, 2017.

[4] Florian Schroff, Dmitry Kalenichenko, and James Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 815–823.

[5] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition," *arXiv preprint arXiv:1801.07698*, 2018.

[6] Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1701–1708.

[7] Omkar M Parkhi, Andrea Vedaldi, Andrew Zisserman, et al., "Deep face recognition," in *bmvc*, 2015, vol. 1, p. 6.

[8] Weihong Deng, Jiani Hu, and Jun Guo, "Extended src: Undersampled face recognition via intraclass variant dictionary," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 9, pp. 1864–1870, 2012.

[9] Xiaofei He, Shuicheng Yan, Yuxiao Hu, Partha Niyogi, and Hong-Jiang Zhang, "Face recognition using laplacianfaces," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 3, pp. 328–340, 2005.

[10] Chaoqiang Liu, Tao Xia, and Hui Li, "A hierarchical hough transform for fingerprint matching," in *International Conference on Biometric Authentication*. Springer, 2004, pp. 373–379.

[11] Nalini K Ratha, Kalle Karu, Shaoyun Chen, and Anil K Jain, "A real-time matching system for large fingerprint databases," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, , no. 8, pp. 799–813, 1996.

[12] Yuliang He, Jie Tian, Xiping Luo, and Tanghui Zhang, "Image enhancement and minutiae matching in fingerprint verification," *Pattern recognition letters*, vol. 24, no. 9-10, pp. 1349–1360, 2003.

[13] Dongjae Lee, Kyoungtaek Choi, and Jaihie Kim, "A robust fingerprint matching algorithm using local alignment," in *Pattern Recognition, 2002. Proceedings. 16th International Conference on*. IEEE, 2002, vol. 3, pp. 803–806.

[14] Joshua Abraham, Paul Kwan, and Junbin Gao, "Fingerprint matching using a hybrid shape and orientation descriptor," in *State of the art in Biometrics*. InTech, 2011.

[15] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[16] Ralph Gross, Iain Matthews, Jeffrey Cohn, Takeo Kanade, and Simon Baker, "Multi-pie," *Image and Vision Computing*, vol. 28, no. 5, pp. 807–813, 2010.

[17] Hongyi Zhang, Moustapha Cisse, Yann N Dauphin, and David Lopez-Paz, "mixup: Beyond empirical risk minimization," *arXiv preprint arXiv:1710.09412*, 2017.

[18] Athinodoros S. Georghiades, Peter N. Belhumeur, and David J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE transactions on pattern analysis and machine intelligence*, vol. 23, no. 6, pp. 643–660, 2001.

[19] Raffaele Cappelli, Matteo Ferrara, Annalisa Franco, and Davide Maltoni, "Fingerprint verification competition 2006," *Biometric Technology Today*, vol. 15, no. 7-8, pp. 7–9, 2007.

[20] Radford M Neal, *Bayesian learning for neural networks*, vol. 118, Springer Science & Business Media, 2012.

[21] SDK VeriFinger, "Neuro technology (2010)," *VeriFinger, SDK Neuro Technology*.

[22] Dong Yi, Zhen Lei, Shengcai Liao, and Stan Z Li, "Learning face representation from scratch," *arXiv preprint arXiv:1411.7923*, 2014.