Privacy-Safe Iris Presentation Attack Detection

Mahsa Mitcheff, Patrick Tinsley, and Adam Czajka
384 Fitzpatrick Hall of Engineering, University of Notre Dame, IN 46556, USA
{mmitchef,ptinsley,aczajka}@nd.edu

2024 IEEE International Joint Conference on Biometrics (IJCB) | 979-8-3503-6413-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/IJCB62174.2024.10744455

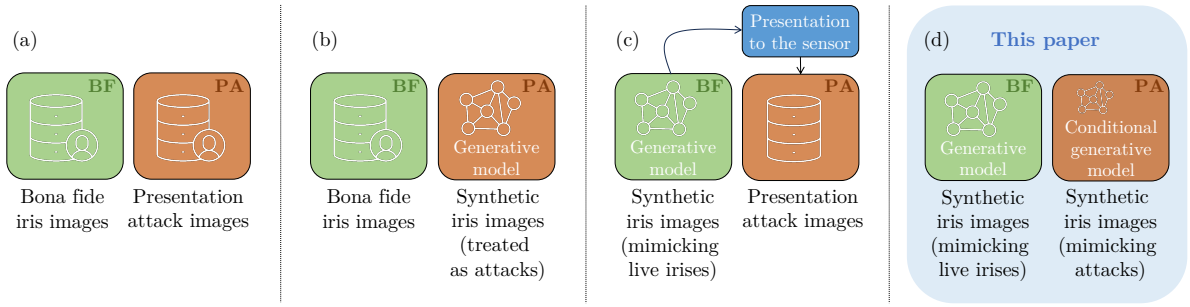


Figure 1. Approaches to source training samples to develop iris presentation attack detection methods: (a) a classical approach, in which authentic images of living individuals and physical attacks are used, (b) similar to (a) but targeting specifically detection of synthetically-generated iris images, or "deep fakes", (c) a newer approach, in which *bona fide* samples are synthesized, and then used to carry out physical presentation attacks captured by iris sensors, and (d) the approach proposed in this paper, in which both *bona fide* and attack samples are synthesized and used to train iris presentation attack detection models. **BF** and **PA** stand for "bona fide" and "presentation attack," respectively. A person icon next to the dataset icon denotes the presence of identity information in the data.

Abstract

This paper proposes a framework for a privacy-safe iris presentation attack detection (PAD) method, designed solely with synthetically-generated, identity-leakage-free iris images. Once trained, the method is evaluated in a classical way using state-of-the-art iris PAD benchmarks. We designed two generative models for the synthesis of ISO/IEC 19794-6-compliant iris images. The first model synthesizes bona fide-looking samples. To avoid "identity leakage," the generated samples that accidentally matched those used in the model's training were excluded. The second model synthesizes images of irises with textured contact lenses and is conditioned by a given contact lens brand to have better control over textured contact lens appearance when forming the training set. Our experiments demonstrate that models trained solely on synthetic data achieve a lower but still reasonable performance when compared to solutions trained with iris images collected from human subjects. This is the first-of-its-kind attempt to use solely synthetic data to train a fully-functional iris PAD solution, and despite the performance gap between regular and the proposed methods, this study demonstrates that with the increasing fidelity of generative models, creating such privacy-safe iris PAD methods may be possible. The source codes and generative models trained for this work are offered along with the paper.

1. Introduction

A textured contact lens (TCL), if worn, partially occludes the natural pattern of an iris, and thus almost certainly leads to a false non-match result if the iris covered by such a lens is being matched to the same, clear iris [10, 35, 7]. This, combined with the fact that the true intention of wearing contact lenses cannot be easily assessed (*e.g.*, cosmetic/medical reasons), has led to TCLs being one of the most popular concealer attacks on iris recognition systems in real operational scenarios. The quest for effective iris presentation attack detection (PAD) methods, specifically crafted for detecting TCLs, is thus unfading.

Effective iris PAD methods, especially those based on modern deep learning-based models, are data-hungry. Acquiring large sets of authentic iris images without and with TCLs (and of multiple and diverse brands) may become more challenging in the near future due to rising debate around Artificial Intelligence (biometrics included) and the ways of how our privacy is protected when our data is used to train AI methods [30, 31, 18].

However, these fears can be transformed into opportunities. With constantly increasing fidelity of image synthesis (including biometric samples) offered by modern generative models, we may eventually reach a point when authentic and synthetically-generated iris images will be indistinguishable. This may include synthetic iris images showing

rare and difficult-to-collect anomalies, such as deformations caused by post-mortem decomposition [3], but also irises with TCLs of various and unknown brands. This paper investigates this latter opportunity and makes an attempt to train a regular iris PAD model using *solely* synthetically-generated data. Fig. 1 illustrates the difference between the proposed approach and other works related to iris PAD and utilizing synthetically-generated samples. Additionally, we also make sure that none of the synthetic iris images "leak" identity information from the training samples (used to train the generative models) to the synthetic data employed in training our PAD models. This is done by excluding samples synthesized by the generative model that are "too close" (in a biometric matching sense) to samples used in training this generative model. Such "privacy-safe" PAD models are then tested as usual, with regular iris PAD benchmarks containing authentic iris images with and without TCLs, and, interestingly, achieve performance which is not significantly worse than the one obtained for models trained with authentic data.

Our experiments show that the PAD accuracy of models trained solely with synthetic data is statistically significantly worse than models trained with the same number of samples but originating from human subjects: the average Area Under the ROC Curve (AUC) across five independent training runs and across all models equal to 0.90 (lowest) and 0.93 (highest) in the former case, compared to 0.97 obtained in the latter case. However, this discrepancy in accuracy is relatively small and demonstrates the feasibility of training iris PAD methods using only synthetic data. This accuracy gap may disappear quickly with increasing fidelity of biometric data synthesizers.

As a summary, the **novel contributions** of this paper are:

C1: a framework to train a privacy-safe iris PAD method with solely synthetically-generated data,

C2: a proof of concept demonstrating current capabilities of the approach when tested on state-of-the-art iris PAD benchmarks,

C3: all elements needed to replicate our results and facilitate further research in this area: (a) source codes and weights of unconditional and conditional StyleGAN2-ADA models synthesizing ISO/IEC 19794-6-compliant iris images without and with textured contact lenses (mimicking seven distinct contact lens brands)[1], and (b) dataset of synthetic iris images used to train the privacy-safe PAD models[2].

[1] `https://github.com/CVRL/PrivacySafeIrisPAD`

[2] instructions of how to request a copy of the dataset can be found at `https://cvrl.nd.edu/projects/data/`; look for dataset named *Privacy-Safe Iris PAD (UND-2024-PSIPAD)*

2. Related Work

2.1. Iris Image Synthesis

Generative Adversarial Networks (GANs)[14] have revolutionized the field of image generation, enabling the creation of highly realistic synthetic images. Researchers have extensively used GAN-based models to also generate iris images. Such first attempts were made by Kohli *et al*. [24] and Minaee and Abdolrashidi [26], who introduced the iris synthesis models based on deep convolutional generative adversarial network, called iDCGAN and DC-GAN, respectively. Yadav *et al*. [36] employed a relativistic average standard GAN (RaSGAN), which – unlike iDCGAN – trains its generator to maximize the probability that randomly sampled synthetic iris images are more realistic than a set of real irises. The problem with these early, yet successful, approaches was that generated images had the iris filling out the image frame almost entirely (we could classify this type as a *cropped* iris image, according to ISO/IEC 19794-6), and were consequently lacking some of the fine iris details due to low image resolution.

Image-to-image translation approaches were also found to be useful in iris image synthesis. Zou *et al*. [42] proposed 4DCycle-GAN that adds two more discriminators to Cycle-GAN to increase diversity of features in synthesized irises. Their goal was to convert irises with textured contacts lenses into images of irises without texture contact lenses. Yadav and Ross [37] proposed a cyclic image translation generative adversarial network (CIT-GAN) for iris image manipulation. The model transfers stylistic elements from one domain to another domain, *e.g*. from cosmetic contact lenses to plain iris. More recently, Yadav and Ross [38] proposed the iWarpGAN model that can generate irises of new identities not seen during training, with the increased variety of image styles present in the generated irises.

The above works lack one important, from the biometrics point of view, property of the synthesized images: identity preservation. This means that generated samples, even exhibiting high visual realism, could not be considered as same-eye samples. This has been partially addressed by Khan *et al*. [21, 22], who proposed deep learning-based iris pattern deformation models to synthesize iris images with different pupil sizes, correctly modeling the anatomical nonlinear deformations of the iris tissue, and thus preserving identity. Bhuiyan and Czajka [3] also propose a conditional StyleGAN2-based model to generate both same-eye and different-eye post-mortem iris images, given the time elapsed since death (post-mortem interval).

2.2. Iris Presentation Attack Detection

Iris PAD has seen significant attention to date, with more than two hundred papers proposing iris PAD algorithms

published to date, if we base our estimates solely on the methods surveyed by Czajka and Bowyer [7] and Boyd *et al*. [5], as well as algorithms submitted to the LivDet-Iris competitions [40, 32]. Despite the increasing performance of iris PAD in closed set scenarios (when attack types used in testing are known), especially demonstrated by deep learning-based algorithms, iris PAD still lacks the ability to generalize well to unknown attack types and unknown properties of known attack types [6]. **This paper addresses the latter challenge, in which properties** (brand, texture, opaqueness) **of known attack type** (textured contact lenses) **are unknown**.

2.3. Synthetic Data in the Context of PAD

The fidelity of modern generative models has increased so far that instead of treating generated images as spoof examples, synthesized images started to replace authentic samples. Yadav *et al*. [36] proposed an approach in which synthetically-generated irises were either treated as authentic samples, or were added to enlarge the set of presentation attack types. In both scenarios, the authors found that adding synthetic samples increases the generalization capabilities of iris PAD. Yadav *et al*. [37] proposed a generative model to transfer the style of known presentation attacks (*e.g*., paper printouts or artificial eyes) onto a bona fide samples.

Both solutions above may considerably increase the size of iris PAD-related datasets. One potential aspect not yet discussed is the increase of privacy: whether generative models are appropriately used in the PAD context. Fang *et al*. explicitly included the discussion about privacy into synthesizing PAD datasets [12] and introduced *SynthASpoof*, the first privacy-preserving dataset for training face PAD algorithms [11]. This dataset consists of synthetically-generated faces (via StyleGAN2-ADA) that serve as "bona fide" samples. The researchers then captured real physical presentation attacks (printed and displayed images of synthetic faces), and demonstrated the feasibility of training face PAD models solely with synthetic data, eliminating the ethical and legal issues associated with collecting images of real faces. Later, *SynthASpoof* was used in the "Face Presentation Attack Detection Based on Privacy-Aware Synthetic Training Data" competition (SynFacePAD 2023) [12], demonstrating promising results achieved by the competition participants in training PAD models solely on synthetic face images from SynthASpoof.

Our paper differs from the above works in at least two important aspects. Firstly, both "bona fide" and "spoof" samples are synthesized by the proposed generative models. This eliminates the need to carry out physical presentation attacks, and thus naturally makes the creation of the PAD dataset much easier. Secondly, and more importantly, due to a potential data leakage observed in generative models, we have added a component to verify that synthesized samples do not biometrically match samples of authentic subjects, whose data was used in training the generators.

3. Methodology

3.1. Generative Model Selection

Limited data per contact lens brand presents a challenge for choosing a suitable generative model. Conventional Generative Adversarial Networks (GAN)-based models typically require considerably large datasets. However, in the case of TCL iris images, we usually only have a few hundred to a few thousand images per lens brand, which may cause the discriminator to overfit. As such, we opted for the StyleGAN2-ADA architecture [20], which is designed for synthesizing images after training with limited number of samples. The StyleGAN2-ADA framework is designed to increase the number and variety of samples through differing augmentation techniques. The applied training data augmentations followed those recommended by the StyleGAN2-ADA authors.

It is important to note that the presented framework is not generative model-specific (for instance, StyleGAN-specific). In particular, the identity leakage mitigation component is a post-hoc operation once the synthesis is completed, and thus can be applied to any generator. In general, any generative model able to synthesize ISO-compliant iris images, such as diffusion models [28, 29], can be deployed.

3.2. Overview of the Proposed Framework

Fig. 2 outlines a framework consisting of four distinct steps to train an iris PAD system while preserving user privacy. To synthesize iris images, we trained two separate StyleGAN2-ADA-based models (**Step 1** in Fig. 2). The first model synthesizes irises with TCLs, and is conditioned by the contact lens brand, enabling image generation with contact lenses matching designs and patterns of seven distinct brands: Bausch & Lomb, FreshLook, CooperVision, Ciba Vision, United Contact Lens, Johnson & Johnson, and ClearLab. The second model synthesizes images mimicking live irises without any textured contact lenses.

Next, we use Neurotechnology's VeriEye SDK [27] to identify any matches between the generated irises without TCLs and irises used to train the StyleGAN2-ADA model, and remove all synthetic samples matching live irises (**Step 2** in Fig. 2). The proposed framework is independent of the iris matcher. We decided to use VeriEye in this work due to its availability and relatively high position in the NIST IREX program's leaderboard [1].

This curated, solely-synthetic data is then used to train regular iris PAD methods (**Step 3** in Fig. 2), which in this work are based on three popular architectures: DenseNet

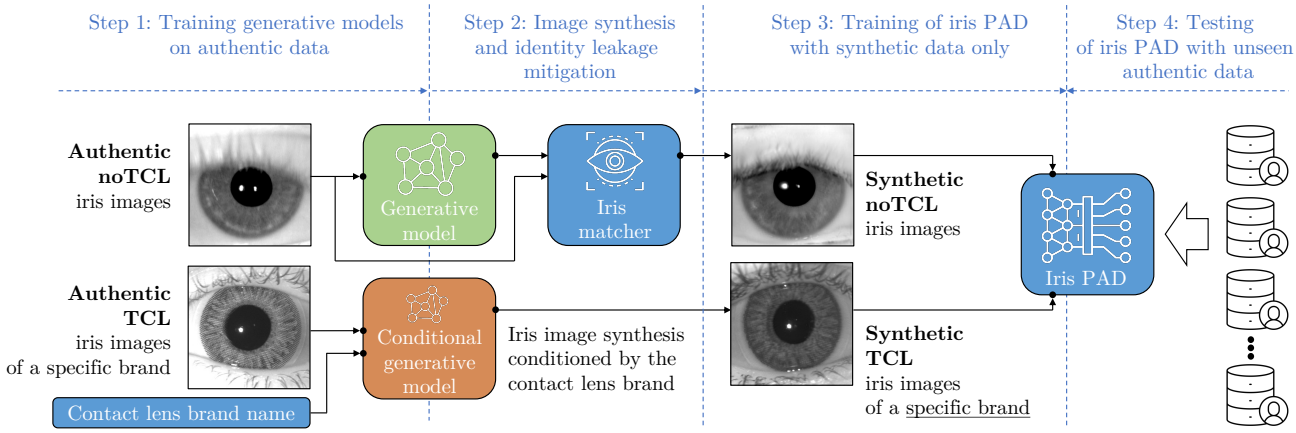


Figure 2. The pipeline of privacy-safe, synthetic data-only iris presentation attack detection (PAD) training and validation. **TCL** and **noTCL** denote images of irises with and without contact lenses, respectively. After training generative models (**Step 1**), we exclusively use synthetically-generated data (mimicking irises both with and without textured contact lenses) to train iris PAD as usual (**Step 3**). The iris matcher is used (in **Step 2**) to exclude synthetic samples that are "too close" to non-synthetic samples used for generative models training, which prevents the "leakage" of identity information from the training set into the generated samples. Resulting iris PAD methods are tested on regular (non-synthetic) data composed of *bona fide* and fake samples (**Step 4**).

[16], ResNet [15], and Vision Transformer [8].

Finally, the trained PAD models are tested with several state-of-the-art benchmarks composed of non-synthetic images of irises with and without TCLs (**Step 4** in Fig. 2). Sections 3.3 through 3.6 provide details related to all elements of the proposed framework. Since each step accomplishes an independent task, optimizing all tasks jointly is unnecessary and was not performed.

3.3. Training Generative Models on Authentic Data (Step 1 in Fig. 2)

3.3.1 Datasets Used

Authentic TCL: ND3D [13] and "ND Cosmetic Contact Lenses 2013" (further: ND-CCL) [9] datasets were used to train the StyleGAN2-ADA model to synthesize iris images with TCLs. ND3D [13] consists of 4,328 iris images captured with AD100 and LG4000 sensors, and represent three different contact lens brands: Bausch & Lomb, FreshLook, and Johnson & Johnson. ND-CCL consists of 16,925 iris images captured using the same sensors (AD100 and LG4000) and representing five different contact lens brands: CooperVision, Ciba Vision, United Contact Lens, Johnson & Johnson, and ClearLab. In total, ND3D and ND-CCL combined provides 21,253 iris images with textured contact lenses.

Authentic noTCL: A subset of images presenting irises without textured contact lenses, selected from all datasets published by the University of Notre Dame [2], was used to train the generative model synthesizing iris images without contact lenses (noTCL). This subset contains 161,549 iris samples of an ISO-compliant resolution (640 $\times$ 480 pixels). All images were captured with the LG2200 sensor, representing 2,237 irises from 1,120 subjects.

Detailed information about the distribution of samples is presented in Table 1. Example images from ND3D and ND-CCL are shown in the upper row in Fig. 3. The bottom row in Fig. 3 presents example images synthesized by the models trained in Step 1, mimicking both images of irises with and without texture contact lenses.

3.3.2 Training

"Authentic TCL" and "Authentic noTCL" collections were used to train generative models synthesizing **Synthetic TCL** and **Synthetic noTCL** samples (used later in Step 3), respectively. We employed a class-conditional StyleGAN2 model to generate **Synthetic TCL** iris samples. For each textured contact lens brand, we defined a separate class. **Synthetic noTCL** iris images were generated using an unconditional StyleGAN2 model. The training code was adopted from the NVIDIA repository[3], specifically using the StyleGAN2 and StyleGAN2-ADA configurations. To ensure consistent input across models, all authentic noTCL and TCL iris images were re-scaled to a standard resolution of 512 $\times$ 512 pixels (this allows the GAN model to focus on learning features independent of the original image sizes). Synthesized images were then re-scaled back to ISO-compliant resolution of 640 $\times$ 480 pixels. StyleGAN-specific training parameters[4] were the following: batch size

[3] https://github.com/NVlabs/stylegan3/

[4] for further references, see https://github.com/NVlabs/stylegan3/blob/main/docs/configs.md

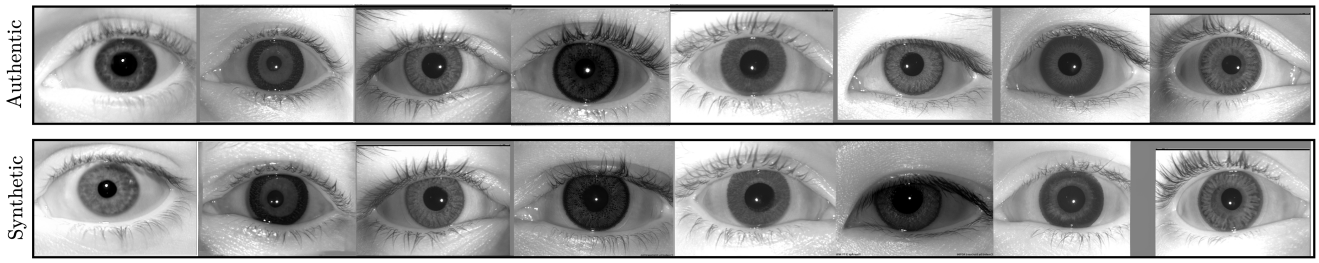


Figure 3. Examples of authentic (upper row) and synthetically-generated by a conditional StyleGAN2-ADA (bottom row) samples without textured contact lens (noTCL) and with textured contact lens (TCL) of a given brand.

= 32, gamma = 1.6384, mapping depth = 2, generator learning rate = 0.0025, and discriminator learning rate = 0.0025. Both StyleGAN-ADA models were trained from scratch using the selected iris image datasets.

3.4. Image Synthesis and Identity Leakage Mitigation (Step 2 in Fig. 2)

3.4.1 Synthesis

An identified issue with generative models is the potential for "identity leakage," where synthetic images tend to match authentic images used in the training set. To address this issue and make our solution privacy-safe, we first generate $N >> K$ iris images, where K is the number of images we eventually use in iris PAD training (Step 3). Any image that matches a sample used in StyleGAN2-ADA training is excluded.

3.4.2 Identity Leakage Mitigation

Neurotechnology's VeriEye SDK [27] was used for the leakage-related comparison experiments. This approach, however, is not matcher-specific. In our experiments $N = 10,000$. Additionally, any synthetic image that failed to enroll properly with the VeriEye was discarded. Given these exclusion rules, the remaining $K = 4,167$ synthetic samples, labeled as **Synthetic noTCL** in this paper, were used in Step 3.

3.4.3 Quality Evaluation

We evaluate the quality of synthesized iris images using the ISO/IEC 29794-6 quality metrics [4] to ensure that synthetic TCL and noTCL images present comparable quality as authentic TCL and noTCL irises. Such agreement in quality metrics is achieved, as shown in Fig. 4.

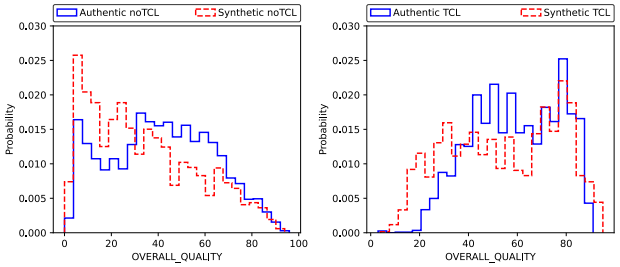


Figure 4. Distributions of the ISO/IEC 29794-6 overall quality score for iris images without contact lenses (noTCL) shown on the left plot, and for iris images with contact lenses (TCL) shown on the right plot. The quality score ranges from 0 to 100, with higher scores indicating better image quality.

3.5. Training of Iris PAD Algorithms with Synthetic Data Only (Step 3 in Fig. 2)

3.5.1 Presentation Attack Detection Models

To investigate the performance of PAD algorithms, we adopted the DenseNet [16], ResNet [15], and Vision Transformer (ViT) [8] architectures, which achieved a good accuracy in the most recent LivDet-Iris 2023 competition [33].

DenseNet employs a dense connectivity pattern, where each layer receives feature maps from all preceding layers as input, and its own feature maps become inputs to all subsequent layers. This dense connectivity strengthens feature propagation, improves feature re-usability, and significantly reduces the number of parameters compared to traditional convolutional neural networks. We used a pre-trained DenseNet with 121 layers for this study.

ResNet, instead of directly feeding information from layer to layer, introduces skip connections that bypass intermediate layers, allowing information to flow directly from earlier layer (L) to later layer (L+2). This enables the model to learn the difference between the two representations, rather than starting from scratch at each layer. This approach allows the model to build upon existing knowledge, facilitating the learning of complex and fine-grained

Table 1. Sources of data used in training and testing of the proposed framework.

Step as in Fig. 2	Image Type	Contributing Datasets	# of Samples
Step 1 (training generative models)	Authentic TCL	Sourced from ND3D [13] and ND-CCL [9]:	
		Bausch & Lomb	488
		FreshLook	1,901
		CooperVision	2,095
		Ciba Vision	2,716
		United Contact Lens	3,156
		Johnson & Johnson	5,015
		ClearLab	5,863
	Authentic noTCL	Sampled from publicly-available datasets [2]	161,549
Step 2 (privacy-safe synthesis), and Step 3 (PAD training with synthetic data)	Synthetic TCL	Generated and offered with this paper	4,167
	Synthetic noTCL	Generated and offered with this paper	4,167
Step 3 (PAD training with authentic data)	Authentic TCL	Sampled from publicly-available datasets [2]	4,167
	Authentic noTCL	Sampled from publicly-available datasets [2]	4,167
Step 4 (testing with iris PAD benchmarks)	Authentic TCL	BERC_IRIS_FAKE [25]	140
		IIITD Contact Lens Iris [23]	2,256
		LivDet-Iris Clarkson 2015 [41]	2,533
		LivDet-Iris Clarkson 2017 [39]	1,886
		LivDet-Iris IIITD-WVU 2017 [39]	957
	Authentic noTCL	BERC_IRIS_FAKE [25]	2,733
		IIITD Contact Lens Iris [23]	2,140
		LivDet-Iris Clarkson 2015 [41]	1,884
		LivDet-Iris Clarkson 2017 [39]	3,868
		LivDet-Iris IIITD-WVU 2017 [39]	2,216

features, and ultimately leading to improved performance in both shallow and deep networks. Both DenseNet and ResNet models tackle the problem of vanishing gradients, a fundamental challenge in training deep networks. We used a pre-trained ResNet with 101 layers for this study.

Vision Transformer (ViT) instead of feeding the entire image at once, it breaks the input down into smaller, fixed-size patches. These patches are then linearly embedded along with their position and passed to the transformer encoder, which analyzes the relationships between individual patches and the entire image, and classifies the image based on these learned representations. We used a pre-trained ViT-base model with a patch size of 16 for this study.

3.5.2 Training

In this step, the PAD models are trained solely on synthetic data. Although the ISO quality metrics are similar to those obtained for authentic iris images (cf. 4), to improve the model's robustness to image quality variations and enhance the diversity of the iris data during training, we employed various image augmentation techniques inspired by the `imgaug` library [19]. These techniques included basic transformations (*e.g.*, flipping, rotation), additive noise, and more advanced distortions like sharpening, blurring, and brightening. To enhance the similarity between synthetic and authentic irises, we augmented the synthetic irises used for training process of experiment 1 with a more rigorous brightness and contrast adjustment strategy and a broader range of blurring techniques.

Since there are 4,167 synthetic noTCL privacy-safe samples (curated in Step 2) available in this step, we included an equal number of synthetic TCL samples (4,167) representing seven contact lens types, with 595 or 596 samples per lens brand to achieve well-balanced set across all brands. 80% of the data was used for training, with the remaining 20% reserved for validation. Each model was trained for a maximum of 50 epochs, and the weights corresponding to the best accuracy on the validation set were picked for the final model. A batch size of 32 was used with stochastic gradient descent (SGD) as the optimizer and cross-entropy as the loss function. The learning rate was 0.005, the weight decay was 1e-6, and the momentum was 0.9.

3.6. Testing of Iris PAD Algorithms with Unseen Authentic Data (Step 4 in Fig. 2)

To evaluate the effectiveness of the PAD models trained on synthetic data and compare them to PAD models trained traditionally (with authentic data), we utilized several standard iris PAD benchmarks containing both noTCL and TCL authentic samples: BERC_IRIS_FAKE [25], IIITD Contact Lens Iris [23], LivDet-Iris Clarkson 2015 [41], LivDet-Iris Clarkson 2017 [39], and LivDet-Iris IIITD-WVU 2017 [39]. To our knowledge, this composition of benchmarks is the largest one can presently collect from publicly-available research datasets containing images with textured contact lenses. In the interest of fair evaluation, we have excluded textured contact lens datasets acquired by the University of Notre Dame, since data from that group was used at various stages of the method's design. For consistency, all images were center-cropped (using circular iris boundary approximations obtained by software offered with [34]) and resized to a uniform resolution of 256×256 pixels, which allows the models to be more focused on the iris texture. Details of the datasets used in various steps when designing the proposed framework are provided in Table 1.

4. Experiments and Results

4.1. Cross-validation and Metrics

Following ISO/IEC 30107-3:2017 [17], we evaluate PAD models by reporting Attack Presentation Classification Error Rate (APCER: the proportion of "attack" samples called "bonafide") and Bonafide Presentation Classification Error Rate (BPCER: the proportion of "bonafide" samples called "attack"). APCER and BPCER are used to (a) plot Detection Error Tradeoff (DET) curve, (b) calculate Area Under the Receiver Operating Characteristic curve (AUROC), and (c) calculate the decidability score:

$$d' = \frac{|\mu_{\text{BF}} - \mu_{\text{PA}}|}{\sqrt{0.5\left(\sigma_{\text{BF}}^2 - \sigma_{\text{PA}}^2\right)}}$$

where μ and σ are means and standard deviations, respectively, of the PAD scores obtained for bona fide (BF) and attack (PA) samples.

For estimating the uncertainty of the results associated with stochastic optimization during neural network models training, we independently trained each model five times with the same datasets, but with five different training seeds. We thus further report the average performance metrics along with their standard deviations.

4.2. Experiments

To evaluate a potential gap between the performance of the iris PAD methods trained traditionally (with authentic samples) and trained solely with synthetic samples, we designed two experiments:

E1: the iris PAD models were **trained exclusively on synthetic data**, as described in Sec. 3.5, and **tested on iris PAD benchmarks**,

E2: the iris PAD models were **trained solely on authentic data** (including images of irises with and without textured contact lenses), and **tested on the iris PAD benchmarks**; to ensure fair comparisons, we sampled 4,167 authentic iris images without textured contact lenses (noTCL), and 4,167 authentic iris images with textured contact (TCL) for training, representing all identities whose data was used also for training generative models in Step 1.

4.3. Results

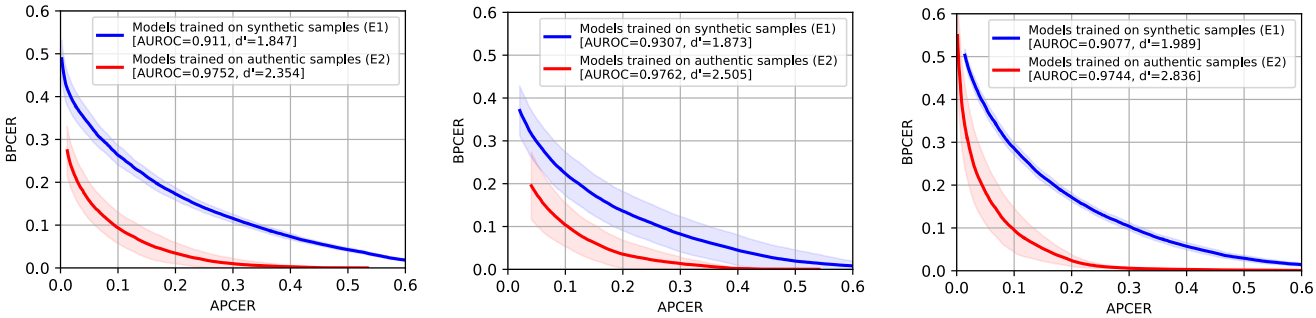
Figure 5 shows detection error tradeoff (DET) curve averaged across five train-test runs for all model architectures. These results are obtained on the iris PAD benchmarks (as described in Table 1) in both experiments (E1 and E2). Additionally, Table 2 shows a few selected operational points on the DET curves, specifically the average BPCER at various APCER values. It can be clearly seen that models trained exclusively with authentic samples (E1) outperformed models trained traditionally with synthetic iris data (E2), although the gap between two is not significant. For instance, the average AUROC for models trained with authentic data reached 97%, while the average AUROC values for models trained with synthetic data ranged from 90% (lowest) to 93% (highest), with the lowest and highest values obtained by the DenseNet and ResNet-based models, respectively. This performance gap, however, is not significant, and the entire experiment brings an interesting observation that iris PAD methods can be trained solely on relatively small, synthetically-generated datasets and still achieve satisfactory performance. Note that experiment E1 assumed using only privacy-safe synthetic data. The number of training samples (two sets of only 4,167 images each) was selected intentionally to run comparisons with models trained on authentic data, which were also limited to two sets of 4,167 images.

One potential reason for the performance gap is that synthetic samples may lack the diversity in term of brightness, contrast, blurriness, and other factors related to image quality, which are present in authentic samples. Additionally, the use of data from a single sensor in training the generative model in Step 1 might have further limited the diversity of synthetic images. Figure 6 illustrates examples of wrong classifications, likely caused by differences in image quality factors between the train and test sets. For instance, models trained with synthetic samples (on which we are focusing in this paper) failed to classify samples that are too dark (upper left part of Fig. 6) or those that were either too blurry or too sharp (lower left part of Fig. 6). This may suggest what additional augmentations could be made to increase the diversity of synthetic training samples. Since the tests presented in Step 4 are made on a sequestered set of benchmarks, we certainly did not make any adjustments in our augmentation strategies after seeing these mis-classifications.

To assess whether there are performance differences between model architectures (DenseNet, ResNet and ViT) we conducted paired samples t-test, comparing mean AUROC values separately for (a) models trained with authentic (red curves in Fig. 5) and (b) synthetic (blue curves in Fig. 5) iris images. In each (a) and (b) scenario, we conducted three

Table 2. BPCER (in % and averaged across five train-test runs) at different APCER levels obtained for iris PAD methods utilizing three architectures. **E1:** models trained solely with synthetic samples. **E2:** models trained solely with authentic samples.

	DenseNet		**ResNet**		**ViT**	
APCER	E1	E2	E1	E2	E1	E2
0.1%	53.17	33.08	44.85	34.79	53.93	56.86
1.0%	42.72	25.03	39.34	24.66	47.30	35.57
5.0%	33.60	14.20	28.68	14.04	35.50	16.56
10.0%	26.41	8.52	21.6	8.41	27.27	8.52



(a) DenseNet (b) ResNet (c) ViT

Figure 5. Average DET curves (thick lines), with shaded areas representing one standard deviation from five train-test runs, obtained for both experiments (E1 in blue and E2 in red) and for all three model backbones.

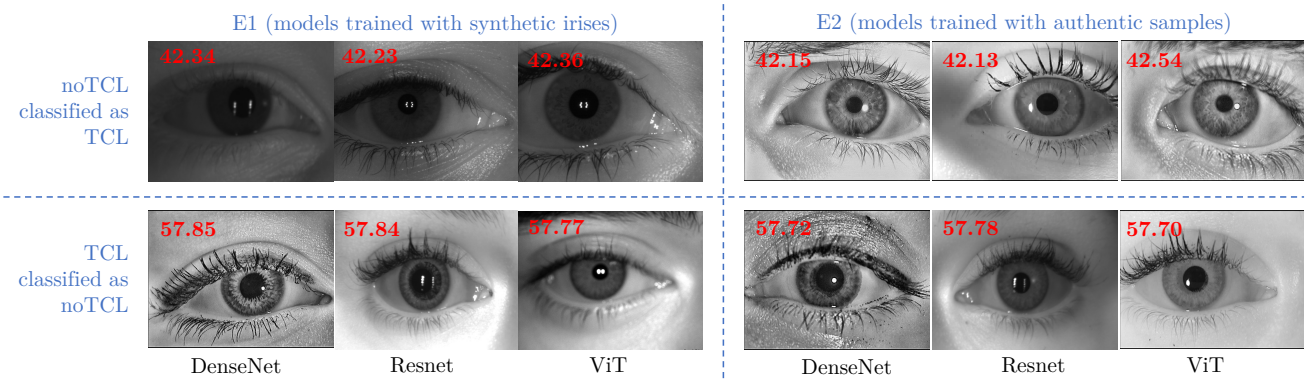


Figure 6. Examples of noTCL samples classified as TCL, and TCL samples classified as noTCL for each model. Numbers over each image is the liveness detection score. The score closer to 100.0 indicates clean iris (noTCL sample), and the score closer to 0.0 indicates an iris with textured contact lens (TCL). A threshold on 50.0 was used in classification into TCL and noTCL.

tests comparing AUROC values of two models, namely: DenseNet vs ResNet, DenseNet vs ViT, and ResNet vs ViT. The null hypothesis was that there are no differences between mean values. The p-values in scenario (a) ranged from 0.75 to 0.82, while those in scenario (b) ranged from 0.17 to 0.61. Thus, there are no reasons, at the assumed significance level $\alpha = 0.05$, to reject the null hypothesis amd hence we conclude that there are no statistically significant differences in the performance across model architectures.

5. Conclusions

This paper proposed the framework in which exclusively synthetically-generated iris images were used to build the entire iris PAD method detecting textured contact lenses. This study demonstrates that it is possible to train effective iris PAD models without using any authentic data, collected from human subjects. To achieve this goal we trained unconditional generative models synthesizing iris images without contact lenses, and conditional generative models synthesizing images of irises wearing contact lenses offered by seven different manufacturers. By applying an "identity leakage" mitigation mechanism in the pipeline, the proposed framework offers an advantage of reducing privacy concerns associated with using iris data from authentic subjects. As a result, we obtained privacy-safe iris PAD methods that perform comparably well when tested on all the existing benchmarks offering iris images with and without textured contact lenses (benchmarks used in models training were excluded from testing to avoid bias).

Possible extensions of this work: One obvious extension of this work is to keep generating synthetic data and see when the performance gap is filled out (if at all). The second extension of this work is to mix the existing authentic datasets used to train iris PAD to date with synthetically-generated samples (both mimicking irises not covered by textured contact lenses, and images of irises wearing contact lenses).

Acknowledgments: This material is based upon work partially supported by the National Science Foundation under Grant No. 2237880. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] IREX 10: Identification Track. https://pages.nist.gov/IREX10/. Accessed: April 6, 2024.

[2] ND Computer Vision Research Lab. https://cvrl.nd.edu/projects/data/. Accessed: July 8, 2021.

[3] R. A. Bhuiyan and A. Czajka. Forensic iris image synthesis. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 1015–1023, 2024.

[4] I. J. S. . Biometrics. Iso/iec 29794-6: 2015 information technology–biometric sample quality–part 6: Iris image data. 2015.

[5] A. Boyd, Z. Fang, A. Czajka, and K. W. Bowyer. Iris presentation attack detection: Where are we now? *Pattern Recognition Letters*, 138:483–489, 2020.

[6] A. Boyd, J. Speth, L. Parzianello, K. W. Bowyer, and A. Czajka. Comprehensive Study in Open-Set Iris Presentation Attack Detection. *IEEE Transactions on Information Forensics and Security*, 18:3238–3250, 2023.

[7] A. Czajka and K. W. Bowyer. Presentation attack detection for iris recognition: An assessment of the state of the art. *ACM Computing Surveys*, 54(4):86:1–86:35, 2018.

[8] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

[9] J. Doyle and K. Bowyer. Notre dame image dataset for contact lens detection in iris recognition. *Iris and periocular biometric recognition*, pages 265–290, 2014.

[10] J. S. Doyle, P. J. Flynn, and K. W. Bowyer. Automated classification of contact lens type in iris images. In *International Conference on Biometrics (ICB)*, pages 1–6, 2013.

[11] M. Fang, M. Huber, and N. Damer. Synthaspoof: Developing face presentation attack detection based on privacy-friendly synthetic data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1061–1070, 2023.

[12] M. Fang, M. Huber, J. Fierrez, R. Ramachandra, N. Damer, A. Alkhaddour, M. Kasantcev, V. Pryadchenko, Z. Yang, H. Huangfu, et al. Synfacepad 2023: Competition on face presentation attack detection based on privacy-aware synthetic training data. In *2023 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–11. IEEE, 2023.

[13] Z. Fang, A. Czajka, and K. W. Bowyer. Robust iris presentation attack detection fusing 2d and 3d information. *IEEE Transactions on Information Forensics and Security*, 16:510–520, 2021.

[14] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[15] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning. *Image Recognition*, 7, 2015.

[16] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[17] ISO/IEC 30107-3. Information technology – Biometric presentation attack detection – Part 3: Testing and reporting, 2016.

[18] A. K. Jain, D. Deb, and J. J. Engelsma. Biometrics: Trust, but verify. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 4(3):303–323, 2021.

[19] A. B. Jung, K. Wada, J. Crall, S. Tanaka, J. Graving, C. Reinders, S. Yadav, J. Banerjee, G. Vecsei, A. Kraft, Z. Rui, J. Borovec, C. Vallentin, S. Zhydenko, K. Pfeiffer, B. Cook, I. Fernández, F.-M. De Rainville, C.-H. Weng, A. Ayala-Acevedo, R. Meudec, M. Laporte, et al. imgaug. https://github.com/aleju/imgaug, 2020. Online; accessed 01-Feb-2020.

[20] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila. Training generative adversarial networks with limited data. In *Proc. NeurIPS*, 2020.

[21] S. K. Khan, P. Tinsley, and A. Czajka. Deformirisnet: An identity-preserving model of iris texture deformation. In *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 900–908, 2023.

[22] S. K. Khan, P. Tinsley, M. Mitcheff, P. Flynn, K. W. Bowyer, and A. Czajka. Eyepreserve: Identity-preserving iris synthesis. *arXiv preprint arXiv:2312.12028*, 2023.

[23] N. Kohli, D. Yadav, M. Vatsa, and R. Singh. Revisiting iris recognition with color cosmetic contact lenses. In *2013 International Conference on Biometrics (ICB)*, pages 1–7. IEEE, 2013.

[24] N. Kohli, D. Yadav, M. Vatsa, R. Singh, and A. Noore. Synthetic iris presentation attack using idcgan. In *2017 IEEE International Joint Conference on Biometrics (IJCB)*, pages 674–680. IEEE, 2017.

[25] S. J. Lee, K. R. Park, Y. J. Lee, K. Bae, and J. H. Kim. Multifeature-based fake iris detection method. *Optical Engineering*, 46(12):1 – 10, 2007.

[26] S. Minaee and A. Abdolrashidi. Iris-gan: Learning to generate realistic iris images using convolutional gan. *arXiv preprint arXiv:1812.04822*, 2018.

[27] Neurotechnology. Verieye sdk. https://www.neurotechnology.com/verieye.html, 2023. Accessed: July 12, 2024.

[28] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer. High-resolution image synthesis with latent diffusion models. In *IEEE/CVF International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10674–10685, 2022.

[29] J. Sohl-Dickstein, E. A. Weiss, N. Maheswaranathan, and S. Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International Conference on Machine Learning (ICML)*, volume 37 of *ICML'15*, page 2256–2265. JMLR.org, 2015.

[30] P. Tinsley, A. Czajka, and P. Flynn. This face does not exist... but it might be yours! identity leakage in generative models. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 1320–1328, 2021.

[31] P. Tinsley, A. Czajka, and P. J. Flynn. Haven't i seen you before? assessing identity leakage in synthetic irises. In *2022 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–9. IEEE, 2022.

[32] P. Tinsley, S. Purnapatra, M. Mitcheff, A. Boyd, C. Crum, K. Bowyer, P. Flynn, S. Schuckers, A. Czajka, M. Fang, N. Damer, X. Liu, C. Wang, X. Sun, Z. Chang, X. Li, G. Zhao, J. Tapia, C. Busch, C. Aravena, and D. Schulz. Iris Liveness Detection Competition (LivDet-Iris) – The 2023 Edition. In *2023 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–10, 2023.

[33] P. Tinsley, S. Purnapatra, M. Mitcheff, A. Boyd, C. Crum, K. Bowyer, P. Flynn, S. Schuckers, A. Czajka, M. Fang, et al. Iris Liveness Detection Competition (LivDet-Iris)–The 2023 Edition. In *IEEE International Joint Conference on Biometrics, Ljubljana, Slovenia, 25-28 September*, 2023.

[34] M. Trokielewicz, A. Czajka, and P. Maciejewicz. Post-mortem iris recognition with deep-learning-based image segmentation. *Image and Vision Computing*, 94:103866, 2020.

[35] D. Yadav, N. Kohli, J. S. Doyle, R. Singh, M. Vatsa, and K. W. Bowyer. Unraveling the effect of textured contact lenses on iris recognition. *IEEE Transactions on Information Forensics and Security*, 9(5):851–862, 2014.

[36] S. Yadav, C. Chen, and A. Ross. Synthesizing iris images using rasgan with application in presentation attack detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pages 0–0, 2019.

[37] S. Yadav and A. Ross. Cit-gan: Cyclic image translation generative adversarial network with application in iris presentation attack detection. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 2412–2421, 2021.

[38] S. Yadav and A. Ross. iwarpgan: Disentangling identity and style to generate synthetic iris images. *arXiv preprint arXiv:2305.12596*, 2023.

[39] D. Yambay, B. Becker, N. Kohli, D. Yadav, A. Czajka, K. W. Bowyer, S. Schuckers, R. Singh, M. Vatsa, A. Noore, et al. Livdet iris 2017—iris liveness detection competition 2017. In *2017 IEEE International Joint Conference on Biometrics (IJCB)*, pages 733–741. IEEE, 2017.

[40] D. Yambay, P. Das, A. Boyd, J. McGrath, Z. Fang, A. Czajka, S. Schuckers, K. Bowyer, M. Vatsa, R. Singh, et al. Review of iris presentation attack detection competitions. In *Handbook of Biometric Anti-Spoofing: Presentation Attack Detection and Vulnerability Assessment*, pages 149–169. Springer, 2023.

[41] D. Yambay, B. Walczak, S. Schuckers, and A. Czajka. Livdet-iris 2015 - iris liveness detection competition 2015. In *2017 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*, pages 1–6, 2017.

[42] H. Zou, H. Zhang, X. Li, J. Liu, and Z. He. Generation textured contact lenses iris images based on 4dcycle-gan. In *2018 24th International Conference on Pattern Recognition (ICPR)*, pages 3561–3566. IEEE, 2018.