Segmentation-based ID preserving iris synthesis using generative adversarial networks

Vijay Kakani[1] · Cheng-Bin Jin[2] · Hakil Kim[3]

Abstract

This study proposes a method for generating ID preserving synthetic iris database. The proposed method can be applied in the generation of a synthetic iris database for various iris recognition tasks. This work successfully combines the main idea of generative adversarial learning, segmentation, and identification to solve real-world problems. The method produces synthetic iris images from the segmentation masks given ID information. The segmentation mask, iris pose, is devised from the input image by using a segmentation network. By doing this, the ID-preserving iris synthesis method generates an unlimited number of synthetic iris images by processing the provided input images. The accuracy of the generated iris images is validated by measuring top-1, top-5, and Area under the Curve (AUC). The SegNet and IDNet performance was evaluated using class accuracy in terms of precision, recall, and F1-score alongside the computation model complexity. This study exhibits ease of use, compatibility, and accuracy in preserving ID information for the generated synthetic images compared to the other baseline methods. Evaluation results prove the efficacy of this work by comparing the randomly generated iris images using the current study alongside existing methods.

✉ Hakil Kim
hikim@inha.ac.kr

Vijay Kakani
vjkakani@inha.ac.kr

Cheng-Bin Jin
sbkim0407@gmail.com

[1] Department of Integrated System Engineering, Inha University, 100 Inha-ro, Nam-gu 22212, Incheon, Korea

[2] HUYA Inc., 280 Hanxi Road, Panyu District 511446 Guangzhou, China

[3] Department of Electrical and Computer Engineering, Inha University, 100 Inha-ro, Nam-gu 22212, Incheon, Korea

1 Introduction

Personal identification is an essential issue in current security applications. Conventional methods include knowledge-based methods and token-based methods. These methods for personal identification are still used commonly considering their low cost and convenient maintainability. However, traditional methods are not reliable enough due to their susceptible nature in terms of security and user compatibility (knowledge may be forgotten, stolen, or lost). Therefore, biometrics-based methods have become a dominant research topic in recent years. Biometrics aims to recognize a person using physiological or behavioral characteristics such as fingerprints, face, iris, palmprint, gait, and voice.

The iris is the annular part between the pupil and the sclera. The purpose of the iris is to change pupil size by using the dilator and sphincter muscles to control the amount of light entering the pupil. Iris has a very complicated texture pattern because of the elastic fibrous tissue covering it such as freckles, coronas, crypts, stripes, and furrows. The most crucial point is that the pattern of the iris never changes over a person's lifetime. The initial studies on iris pattern recognition by Daugman et al. [10] revealed that an iris pattern has more than 250 degrees of freedom. It means that the probability of two different persons having the same iris structure is about 1 in 7 billion. Besides, even the two irises of an individual are different, and the iris pattern is also independent of the genetic constitution. Therefore, some family members, including the twins possess various iris patterns. Compared with other biometric features, personal identification based on iris pattern shows high accuracy due to the above-stated biological observations [53]. Iris identification has many potential applications, such as national border controls, network security, airports, and cellphone authentication [17].

1.1 Motivation for ID-preserving synthetic iris database

- The idea of ID-preserving is very crucial in biometric applications in terms of authentication and authorization. The current study aims to preserve the identity of the iris sample while synthesis variants of it. For instance, given the distribution of the iris samples, the proposed framework can synthesize various iris samples, yet the identity of the iris sample is intact (keeping the source ID unchanged). This will facilitate the growth in sample size with variants while preserving the ID information.
- A significant progress has been achieved in the vicinity of iris identification algorithms. Besides, most iris identification algorithms claim a very low false accept rate (FAR) and false reject rate (FRR). However, none of these algorithms were tested on a large database. Even though the most famous and prominent public iris database such as CASIA exists, it only consists of 756 images with 108 identities.
- Experiments on these relatively small databases suffer from the following drawbacks: first, the actual performance on real-world applications is unpredictable; second, the algorithms devised based on small databases lack generalization ability. Therefore, there is an obvious need for a larger public database to evaluate the performance of new iris identification algorithms and provide more knowledge on essential information about iris characteristics.
- It is difficult to collect numerous amounts of iris images to construct a sizeable real-world database. The difficulties include close relation with personal privacy concerns and the efficiency issue in terms of cost and time in assembling a massive database representing a variety of intra-class variations.

- One alternative way to cope with the problem is to design a synthetic iris generation system to construct large databases which can be used by researchers and developers to evaluate their performances.

The established premise of the problem statement has one significant aspect to consider in attaining efficiency during the generation of a synthetic iris database; that is "ID preservation". In other words, it is crucial that the ID is preserved during the generation of synthetic iris images.

1.2 Significance of employing collaborative framework

The main pipeline of this collaborative framework is to apply segmentation, identification, and generative adversarial network (GAN) for the ID-preserving iris synthesis. In order to generate the pool of iris images, this work utilized the use of synthetic databases in biometrics is not new and has been previously studied in the context of fingerprints and face biometrics [43]. The synthetic method should not only generate reliable and diverse irises but also should be able to preserve the class ID of the identity for the synthetic iris; providing conditional ID information of the identity. In order to signify the needed conditional ID-preserving iris generation, the following networks were employed:

- SegNet: estimates a mask from an input eye image, this mask serves as an input for the novel IrisGAN.
- IDNet: extracts iris features from the anchor image to feed novel IrisGAN to generate specific iris images, as the goal of iris generation has a condition for preserving ID information of a given target iris.
- IrisGAN: generates a synthetic eye image corresponding to the input mask fed by the SegNet and iris features fed by the novel IDNet.

In summary, the main idea of the method is that an input eye image is fed as a template to the segmentation network to estimate a label mask that consists of pupil, iris, sclera, and background. Then the mask is utilized as an input of a generator network that constructs the synthetic eye image while preserving given identity information from the designed iris identification network. Therefore, our method includes a collaborative framework of SegNet, IDNet, and IrisGAN. The schematic of these three networks is shown in Fig. 1. The overall framework of the iris segmentation and generation, shows that there are three parts included in the system: (1) the multi-class eye segmentation, (2) the iris identification, and (3) the generation of iris images with ID preserving feature. Each part is interdependent on the others and cannot work independently. Eye segmentation is the necessary process used by identification and generation, and iris generation utilizes both segmentation and identification. First, SegNet [25] estimates a mask from an input eye image; then, IrisGAN generates a synthetic eye image corresponding to the mask. The goal of iris generation has a condition for preserving the ID information of a given target iris.

To estimate the performance of the generated synthetic iris images, the outcomes are evaluated via various experiments: (a) by estimating the cumulative matching characteristic (CMC) with ID preserving and without ID preserving from IDNet, and (b) by examining the distributions of genuine and impostor matching score of the synthetic images. Additionally, global pixel accuracy, each class accuracy, precision, recall, and F1 score is included for analyzing the model accuracy. Moreover, the model complexity is approximated using the number of learned parameters in each stage using semantic segmentation challenge metrics.
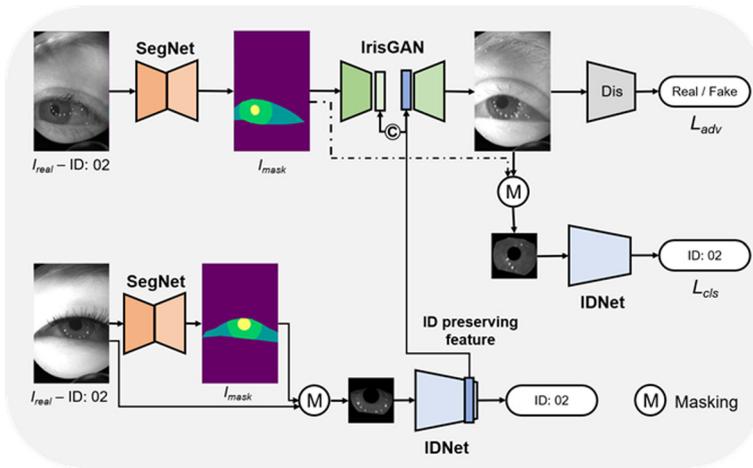
Fig. 1 Proposed framework of ID preserving iris synthesis with the combination of SegNet, IDNet, and IrisGAN

This paper first summarizes previous research on the synthesis of iris generation in Section 2. The proposed method is elaborated in Section 3 along with the network architectures of SegNet, IrisGAN, and IDNet while Section 4 describes the design of the objective loss function of the proposed framework followed by Section 5 reports the experimental results and discussion. The conclusion of the study is presented in Section 6.

2 Literature Review

2.1 Synthetic biometric datasets

Synthetic biometrics are relatively new and related research has been vigorously studied in the last few years. In the fingerprint literature, a software known as SFINGE proposed by Cappelli et al. [4] has been used to generate synthetic fingerprint datasets that have been included in many versions of the Fingerprint Verification Contest (FVC) [33–35]. The experimental results indicate that the performance of competing algorithms on these synthetic datasets and real datasets is comparable. Also, the software can model the intra-class and inter-class dynamics of the real dataset with acceptable precision. Similarly, Orlans et al. [44] utilized the FaceGen software [24] using Viisage FaceTools and 3D Studio Max to develop a face recognition algorithm.

2.2 Synthetic iris image datasets

There are very few literature precedents discussing synthetic iris image generation [3]. Lefohn et al. [28] created and rendered the prosthetic iris by stacking several transparent layers on top of each other modeling a particular iris component, such as stoma, collarette, and sphincter. Cui et al. [8] used principal component analysis (PCA) and super-resolution for iris synthesis. Learned eigenvectors from the training data are used to construct a coarse iris image. Then, the coarse iris images are refined through super-resolution. This work

evidently is the first research regarding the synthesis of artificial irises in the biometric domain. To avoid the sampling of a probability distribution, Makthal and Ross [36] used primitive textural patterns to guide the process of generating synthetic iris images based on deterministic Markov Random Field (MRF). Zuo et al. [64] developed a "model-based anatomy-based" approach, which includes generating the 3D fibers in a cylindrical shape and projecting them onto a 2D field, followed by adding blur effects and eyelids. Wei et al. [56] used the iris patch as a fundamental element to characterize the visual primitive of iris texture. Then, the patch-based sampling is applied to the visual primitive to create a synthetic iris prototype. However, the significant weaknesses of these standard statistical models lie in their unrealistic visual appearance, high complexity, and procedural heuristics.

2.3 Deep learning object detection models

Recent advancements in deep learning techniques based on neural networks such as CNN and GANs have paved the way for generating synthetic datasets in major fields such as autonomous vehicles [1, 21, 38], biometrics [29], video surveillance [12, 19], machine vision [20, 27, 30]. Additionally, the need for image-based forensics [31, 50] and biometric applications is increasing with the advent of technology. In literature, few studies leveraged deep learning object detection models in the context of self-driving applications. These studies focused on employing detection models for the scene classification approaches dealing with the similarities in the scene elements. [40]. Similarly, advanced networks such as SSD-like object detectors were employed to construct an object detection scenario in an indoor scene structure [41]. The parallels can be drawn between these object detection models and the plausible application of such models for the biometric object detection aspects. This could provide a systematic coarse-to-fine detection approach which is essential for the detection approach in the context of delicate structures such as iris, face, fingervein etc.

2.4 Deep learning iris detection and synthesis models

There were a few studies focused on gazing at the liveness detection of the iris using cascade deep learning networks [51] and domain-specific DenseNet features [7]. Besides, the iris synthetic dataset must also have a strong correlation towards the ID information that the sample holds as a unique signature. Minaee et al. [37] proposed a deep convolutional generative adversarial network (DCGA) [47] based synthetic iris method, which can generate iris images sampled from a prior distribution. Yadav et al.[59, 60] utilized the generative capability of a relativistic average standard generative adversarial network (RaSGAN) and Cyclic Image Translation Generative Adversarial Network (CIT-GAN) to synthesize high-quality images of the iris to analyze its potential for presentation attack detection (PAD). A novel algorithm, 4DCycle-GAN, was proposed by Zou et al. [63] for expanding the spoof iris datasets by synthesizing fake iris images in conjunction with textured contact lenses. The significant weakness of these methods is the same as the traditional approaches in that the methods cannot preserve the identity of the generated irises, and the quality of the results is still unrealistic. Tables 1 and 2 present a more detailed comparison between the above-mentioned methods.

Table 1 Comparisons of the related traditional methods for iris image synthesis

Algorithm	Lefohn et al. [28]	Cui et al. [8]	Makthal and Ross [36]	Wei et al. [56]
Method used	Domain knowledge provided by ocularists	Coarse image with PCA and super resolution refinement	Markov random field	Patch-based sampling
Variational control	Manually control	Template control	Template control, rotation, distortion	Template control, defocus, deformation, rotation
Preserving ID information	No	No	No	No
Generated pattern	Ocular prosthesis	Iris code	Iris code	Iris code
Application area	Image synthesis	Image synthesis	Image synthesis	Image synthesis
Dataset	None	CASIA	CASIA	CASIA, UPOL
Evaluations	None	FRR, FAR	Score distribution	Human visual test, score distribution, ROC (FRR, FAR, EER)

Table 2 Comparisons of the related GAN-based methods for iris image synthesis

Algorithm	Kohli et al. [26]	Minaee and Abdolrashidi [37]	Zou et al. [63]	Yadav et al. [59, 60]
Method used	iDCGAN	DCGAN	4DCycle-GAN	RaSGAN,CIT-GAN
Variational control	Random vector	Random vector	Random vector	Random vector
Preserving ID information	No	No	No	No (RaSGAN), Yes (CIT-GAN)
Generated pattern	Iris	Iris	Iris	Iris
Application area	PAD	Synthesis	PAD	PAD
Dataset	IITD Contact Lens, IIT Delhi Iris, Multi-Sensor Iris	CASIA, IIT Delhi	irisKing, IrisGuard, ND-Contact	CASIA-iris-fake
Evaluations	Pupil boundary circularity, Pupil contrast, concentricity Pupil-iris ratio, sharpness, overall quality	FID	None	FID, score distribution, ROC (FRR, FAR, EER)

3 Framework for ID-preserving iris synthesis

The overall framework of ID-preserving iris synthesis shows that there are three parts included in the system: (1) the multi-class eye segmentation, (2) the iris identification, and (3) the generation of iris images with ID-preserving feature.

3.1 SegNet for multi-class eye segmentation

Multi-class eye segmentation is the conditional task where a segmented output mask is generated from an input eye image, and the segmented mask should be conditioned with the ground-truth one. The discriminator of GAN learns the geometry information from real distribution, e.g., a pupil is surrounded by an iris, the iris is always within the sclera, and all three categories are grouped. SegNet incorporates the framework of adversarial learning in the training stage. Besides, considering memory-and-compute constraints, a Dense U-Net-Light is provided based on the landmark structure of the segmentation network, U-Net, and it is applied as the generator in semi-supervised adversarial learning. A discriminator network tries to distinguish ground-truth mask from the outputs generated by Dense U-Net-Light. Instead, the Dense U-Net-Light tries to generate as realistic masks as the discriminator cannot differentiate from the ground-truth masks. Besides, a recursive process is utilized by using the trained Dense U-Net-Light to translate unlabeled data into noisy labeled data. Then, the generator is re-trained and utilized to estimate the correct labeled data again. In this study, we discuss a new approach to multi-class eye segmentation with semi-supervised adversarial learning and massive noisy labeled data. SegNet estimates clear anatomical masks with high precision and recall compared to the traditional segmentation methods, and this method (CVLab-CLv2) achieved 4th place in the https://eval.ai/web/challenges/challenge-page/353/leaderboard/1002 OpenEDS Semantic Segmentation Challenging held by Facebook in the Workshop of ICCV. Semi-supervised adversarial learning can improve the quality of the segmentation by training the generator to estimate segmentation masks that are indistinguishable from masks annotated by human annotators. Moreover, the massive noisy labeled data by annotating unlabeled one can improve the performance of segmentation in advance. Similar observations are also found in other research [58]. The SegNet workflow and network architecture are illustrated in Fig. 2. The weight file of Dense U-Net-Light that includes trainable parameters with float32 data type is less than 1MB.

The upper part of the SegNet in Fig. 2(a) indicates Dense U-Net-Light with semi-supervised adversarial learning. The bottom part of the SegNet in Fig. 2(a) represents the trained Dense U-Net-Light applied in the dataset to convert unlabeled data into labeled data. Even though the labeled data is not entirely clear, the significant volume of labeled with noise data is helpful to improve the model performance. The network shown in Fig. 2(b) includes the encoder and the decoder network. For combining coarse features and refined features, convolutional layers of the encoder and decoder are connected by skip connections. The dashed red line indicates the addition-skip connection, and the black line represents the concatenation-skip connection.

For getting a segmentation mask from the multi-class probability map, the position of probability with maximal is required, and we usually use an argmax function. However, the problem is that this function cannot be backpropagated. It is impossible to calculate the gradient from the discriminator to update Dense U-Net-Light. To address this issue, we use an idea very similar to the soft version of the max function named soft-argmax trick in Fig. 3.
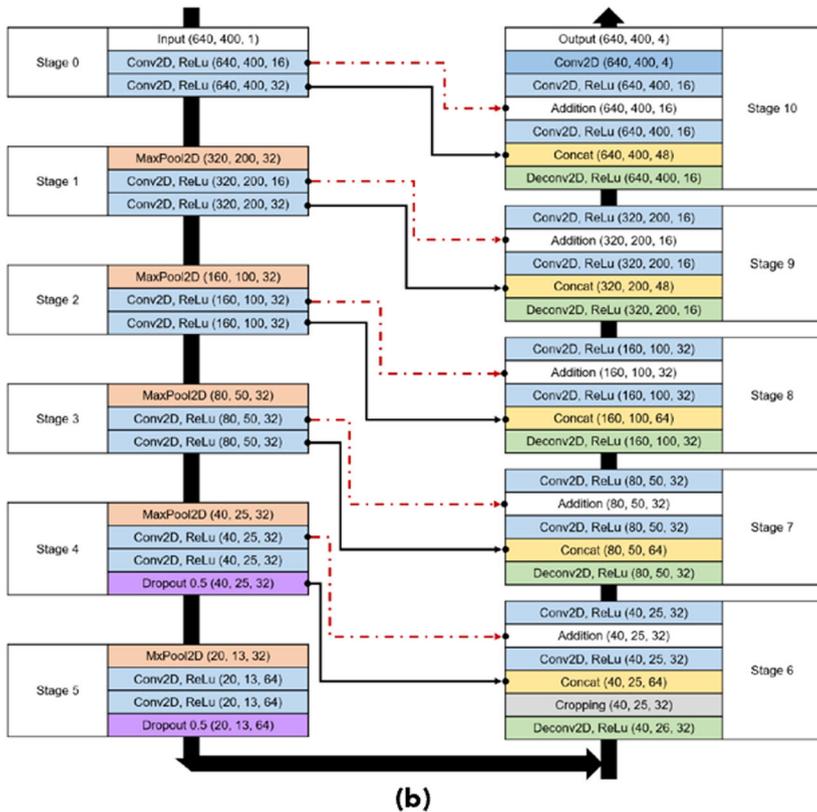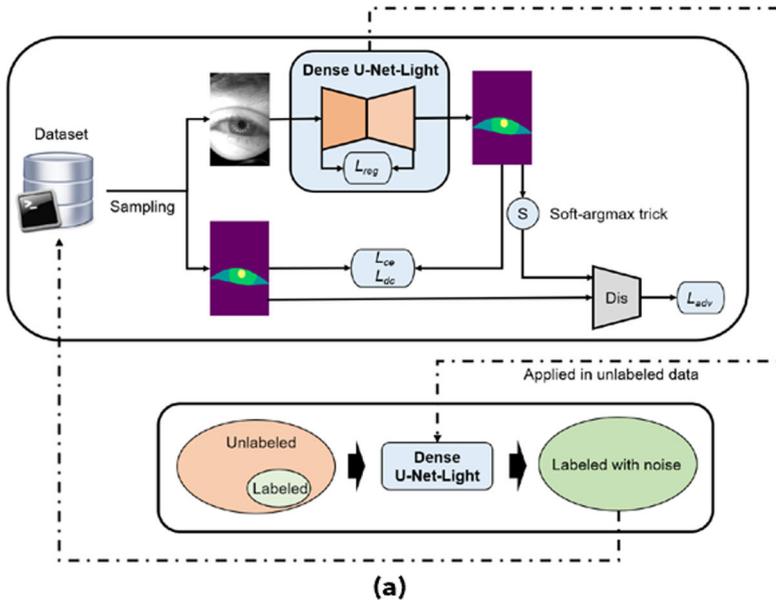
(a)

(b)

Fig. 2 (a). SegNet for multi-class segmentation. (b).The model architecture of the Dense U-Net-Light
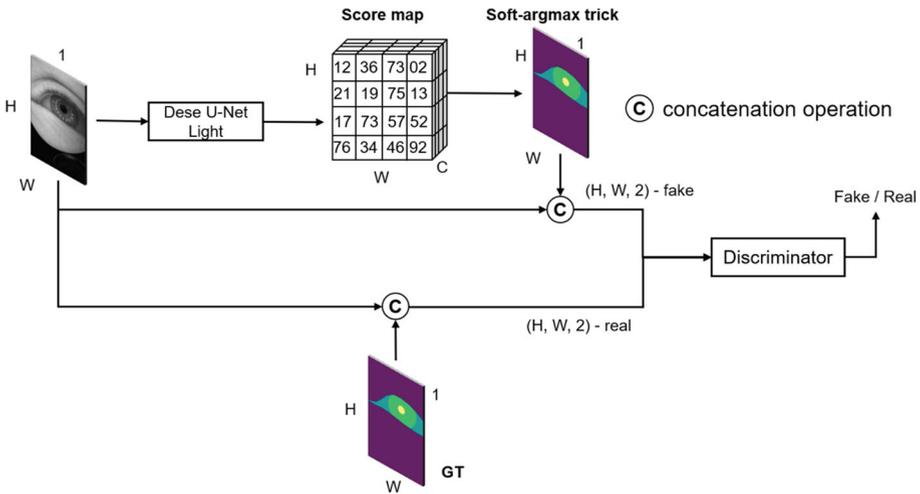
Fig. 3 Explanation of the soft-argmax trick in adversarial learning. In binary-class segmentation, there is no need for soft-argmax. However, this is critical for multi-class segmentation with adversarial learning

The soft-argmax function is expressed as follows:

$$E[x] = \sum_i \frac{e^{x_i}}{\sum_j e^{x_j}} \cdot i \tag{1}$$

where x is the unnormalized score from the network, and i is the index in a vector. Using the above function, we calculate the normalized probabilities for each x_i. The expectation of this is the sum of the indices multiplied by their respective probabilities. However, this expectation is weak if there are multiple modes. For raising the max and lowering the other values, x can be multiplied by an arbitrarily big β. Therefore, the above equation is redefined as follows :

$$soft - argmax(x) = \sum_i \frac{e^{\beta x_i}}{\sum_j e^{\beta x_j}} \cdot i \tag{2}$$

3.2 IDNet for iris identification

As a part of the proposed framework and considering the memory requirements, a relatively shallow but famous network, ResNet18v2 [14] has been employed for designing an identification network IDNet. The personal identification information is only considered from the located iris region of the eye. Therefore, for the input eye image, SegNet is utilized to estimate the segmentation mask that includes pupil, iris, sclera, and background. The iris area is extracted from the eye image using the estimated mask and mapped into the fixed 200×200 as the input of the ResNet18v2. Figure 4 illustrates the architecture diagram of IDNet and an example of the input image feed into the network. While mapping the iris region into the fixed size, the ratio between height and width is preserved and resized to be in the center of an input image.
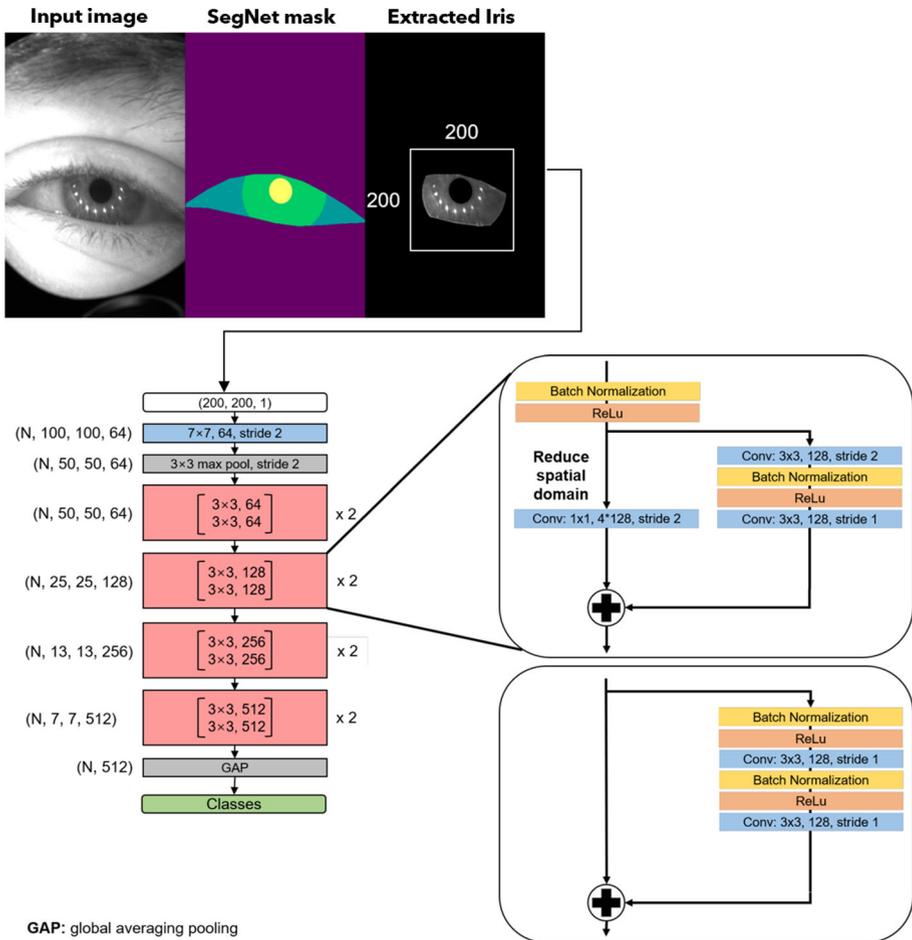
Fig. 4 Sample image for the input of IDNet. The iris area was extracted using mask information. In the network architecture of IDNet, the numbers in the block indicate filter size, the number of convolutional filters, and the stride of convolution. The shape of features after each operation block is also given on the left side of the block

The objective of the identification task involves a data loss term and a regularization term which are defined as follows:

$$L = \frac{1}{N} \sum_{i=1}^{N} L_i + \lambda_{reg} R(W) \tag{3}$$

$$L_i = -\log\left(\frac{e^{s_{y_i}}}{\sum_j e^{s_j}}\right) \tag{4}$$

$$R(W) = \sum_k \sum_l W_{k,l}^2 \tag{5}$$

where s is the unnormalized score of the network and λ_{reg} is a hyperparameter for balancing the data loss term and the regularization term. The data loss is defined using soft-max cross-entropy as usual, and the regularization term uses $L2$ norm constraint. The λ_{reg} is set as $1e^{-4}$ from the experimental observations.

3.3 IrisGAN for ID-preserving iris synthesis

The proposed framework consists of a forward cycle-consistent and a backward cycle-consistent network. Each cycle consistently includes aligned and unaligned learning. The forward cycle enforces the translation from the mask to the eye, while the backward cycle moves from the eye to the mask. As conditional ID-preservation is essential for the synthetic iris generation, IDNet extracts iris features from the anchor image to feed IrisGAN to generate specific iris images as shown in Fig. 5. The proposed IrisGAN also applies a cycle-consistent structure for its unsupervised learning setup. However, the proposed network has a dual cycle-consistent structure for the adoption of semisupervised learning: one cycle-consistent structure for supervised learning with aligned data and the other for unsupervised learning with unaligned data. Because the forward and backward cycle-consistent networks with aligned data or unaligned data are similar, we only illustrate a forward cycle-consistent adversarial network with unaligned learning in Fig. 5(a) and a backward cycle-consistent with aligned learning in Fig. 5(b).

In the forward cycle-consistent adversarial network with unaligned learning, the Syn_{Eye} network generates a synthetic eye image from a mask image, and this eye image is then used by the Syn_{Mask} network to generate the original mask image in order to learn the eye mask structures. The input to the eye discriminator network is either a sample eye image from the real eye data or a synthetic eye image. The objective function for unaligned learning includes both cycle-consistent and adversarial loss. In the backward cycle-consistent adversarial network with aligned learning, a synthetic mask image is generated from an eye
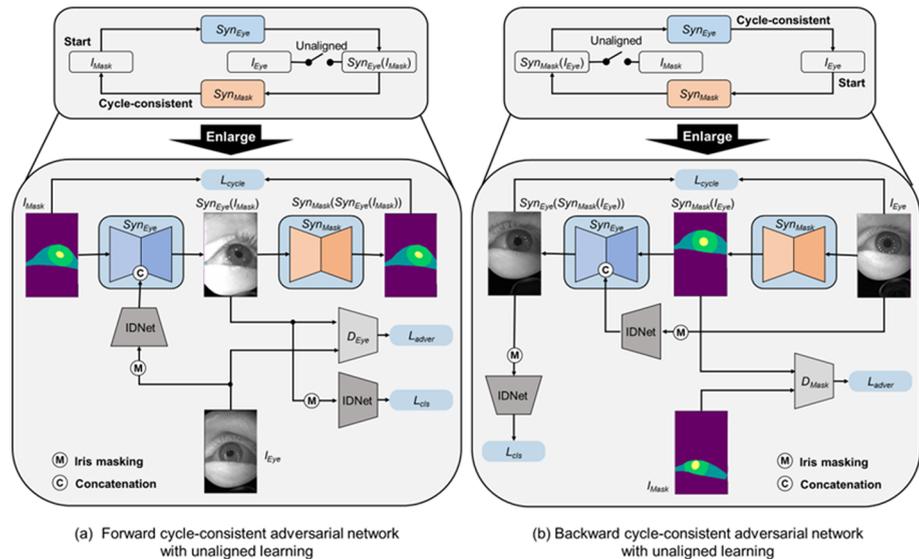

(a) Forward cycle-consistent adversarial network with unaligned learning

(b) Backward cycle-consistent adversarial network with unaligned learning

Fig. 5 IrisGAN: ID-Preserving iris synthesis pipeline

image, and this mask image is employed by the Syn_{Eye} network to generate the original eye image. The mask discriminator is used to distinguish between the synthetic mask and reference mask images. In aligned learning, a reference image is matched with the synthetic image to restrain the generated structure of the output. The four switches are simultaneously employed to control the data flow from the reference image, and these are connected in aligned learning but disconnected in unaligned learning. It is also important to note that the Syn_{Eye} and Syn_{Mask} networks utilized in the forward and backward cycles share the same weights.

The synthesis networks Syn_{Eye} and Syn_{Mask} in IrisGAN adapt the same architecture as used in the study reported by Johnson et al. [18], which produced impressive results in real-time style transfer and single image super-resolution. The network contains two stride-1 convolutions at the beginning and the end, two stride-2 convolutions, nine residual blocks [14, 15], and two fractionally stridden convolutions with a stride of 0.5. Each residual block includes two convolutions with 256 filters of size 3×3 and a stride of 1. Instance normalization [52] and a rectified linear unit (ReLU) [32] activation function follows each convolution except in the final convolutional layer. The Hyperbolic Tangent (Tanh) activation function follows the final convolution to guarantee that the output is within $[-1, 1]$. For the discriminator networks Dis_{Eye} and Dis_{Mask}, we use a patch-based GAN (PatchGAN) [16] architecture, which aims to classify small overlapping image patches as either real or synthetic rather than whole images. This patch-level discriminator architecture has fewer parameters than a whole-image discriminator and can emphasize detailed information in local areas. The flow size of a volume of aligned data is $(N, H, W, 2)$. N is the batch size, H and W are the image height and width, respectively, and 2 represents a concatenation of the synthetic and input images. The number of discriminator layers is fixed at five. The network architecture of IrisGAN is shown in Table 3 where layers marked with IN indicate that the convolution layer is followed by the instance normalization layer. The ReLU activation layers are omitted in the illustration.

Table 3 IrisGAN Network Architecture

Layer	Output	[Kernel,Stride]	Parameters
Conv1, IN	(H x W x 64)	[7x7,1]	3200
Conv2, IN	(H/2 x W/2 x 128)	[3x3,2]	73856
Conv3, IN	(H/4 x W/4 x 256)	[3x3,2]	295168
Residual block1, IN	(H/4 x W/4 x 256)	[3x3,2]	590080
Residual block2, IN	(H/4 x W/4 x 256)	[3x3,2]	590080
Residual block3, IN	(H/4 x W/4 x 256)	[3x3,2]	590080
Residual block4, IN	(H/4 x W/4 x 256)	[3x3,2]	590080
Residual block5, IN	(H/4 x W/4 x 256)	[3x3,2]	590080
Residual block6, IN	(H/4 x W/4 x 256)	[3x3,2]	590080
Residual block7, IN	(H/4 x W/4 x 256)	[3x3,2]	590080
Residual block8, IN	(H/4 x W/4 x 256)	[3x3,2]	590080
Residual block9, IN	(H/4 x W/4 x 256)	[3x3,2]	590080
Fractional conv1, IN	(H/2 x W/2 x 128)	[3x3,0.5]	295040
Fractional conv2, IN	(H x W x 64)	[3x3,0.5]	73782
Conv4, Tanh	(H x W x 1)	[7x7,1]	3137

4 Objectives of ID-preserving iris synthesis

The proposed iris synthesis method can target a given personal identity to generate a synthetic iris image while preserving identity information. While the conventional methods construct some synthetic images by generating iris codes, our method is different in terms of preserving ID information. IrisGAN can generate a synthetic iris image using a segmentation mask from the SegNet and the ID-targeted feature from the IDNet. The objective of ID-targeted IrisGAN is defined as follows:

$$
\begin{aligned}
L_{unsup}\left(Syn_{Eye},\ Syn_{Mask},\ Dis_{Eye},\ Dis_{Mask}\right) = \\
L_{unsup\text{-}adver}\left(Syn_{Eye},\ Syn_{Mask},\ Dis_{Eye},\ Dis_{Mask}\right) \\
+\ \lambda_{cycle}\cdot L_{unsup\text{-}cycle}\left(Syn_{Eye},\ Syn_{Mask}\right) \\
+\ \lambda_{iden}\cdot L_{unsup\text{-}iden}\left(Syn_{Eye},\ Syn_{Mask}\right)
\end{aligned}
\tag{6}
$$

where λ_{cycle} and λ_{iden} are hyper-parameters that balance the relative importance of adversarial, cycle-consistent, and iris identification loss.

The adversarial loss has been applied in an unsupervised setup only for the ID-preserving iris synthesis. The forward and backward mapping are: $Syn_{Eye}\ :\ Mask\ \rightarrow\ Eye$ and $Syn_{Mask}\ :\ Eye\ \rightarrow\ Mask$ respectively and the discriminators Dis_{Mask} and Dis_{Eye} are expressed as follows:

$$
\begin{aligned}
L_{unsup\text{-}adver}\left(Syn_{Eye},\ Dis_{Eye},\ Syn_{Mask},\ Dis_{Mask}\right) = \\
E_{I_{Eye}\sim p_{data}(I_{Eye})}\left[\log\left(Dis_{Eye}\left(I_{Eye}\right)\right)\right] \\
+\ E_{I_{Eye}\sim p_{data}(I_{Eye})}\left[\log\left(1 - Dis_{Eye}\left(Syn_{Eye}\left(A,\ B\right)\right)\right)\right] \\
+\ E_{I_{Eye}\sim p_{data}(I_{Eye})}\left[\log\left(Dis_{Mask}\left(A\right)\right)\right] \\
+\ E_{I_{Eye}\sim p_{data}(I_{Eye})}\left[\log\left(1 - Dis_{Mask}\left(D\right)\right)\right]
\end{aligned}
\tag{7}
$$

where the first two terms are the forward adversarial loss, and the last two terms are the backward adversarial loss. The S^* and T^* are the segmentation and identification network. Also, $A = S^*\left(I_{Eye}^1\right)$, $B = T^*\left(I_{Eye}^2, S^*\left(I_{Eye}^2\right)\right)$ and $D = Syn_{Mask}\left(I_{Eye}^1\right)$. The parameters of the segmentation and identification network are fixed during the training of the ID-preserving networks.

The forward cycle-consistent network should be able to bring I_{Mask} back to the original image, i.e., $I_{Mask} \rightarrow Syn_{Eye}\left(I_{Mask}\right) \rightarrow Syn_{Mask}\left(Syn_{Eye}\left(I_{Mask}\right)\right) \approx I_{Mask}$. Similarly, the backward cycle-consistent network should extract an image I_{Eye} from the image domain to satisfy $I_{Eye} \rightarrow Syn_{Mask}\left(I_{Eye}\right) \rightarrow Syn_{Eye}\left(Syn_{Mask}\left(I_{Eye}\right)\right) \approx I_{Eye}$. The cycle-consistent losses are expressed as follows:

$$
\begin{aligned}
L_{unsup\text{-}cycle}\left(Syn_{Eye},\ Syn_{Mask}\right) = \\
E_{I_{Eye}\sim p_{data}(I_{Eye})}\left[\left\|Syn_{Mask}\left(Syn_{Eye}\left(A,\ B\right)\right) - A\right\|_1\right] \\
+\ E_{I_{Eye}\sim p_{data}(I_{Eye})}\left[\left\|Syn_{Eye}\left(D\right), C\right) - I_{Eye}^1\right\|_1\right]
\end{aligned}
\tag{8}
$$

where $L_{unsup\text{-}cycle}$ are the cycle-consistent structures for unsupervised learning. The S^* and T^* are the SegNet and IDNet, respectively. Also, $A = S^*\left(I_{Eye}^1\right)$, $B = T^*\left(I_{Eye}^2, S^*\left(I_{Eye}^2\right)\right)$, $C = T^*\left(I_{Eye}^1, S^*\left(I_{Eye}^1\right)\right)$ and $D = Syn_{Mask}\left(I_{Eye}^1\right)$.

The iris identification loss is constrained using soft-max cross-entropy function. The loss is defined as follow:

$$L_{iden}\left(Syn_{\text{Eye}}, Syn_{\text{Mask}}\right) =$$

$$E_{I_{Eye} \sim p_{\text{data}}\left(I_{Eye}\right)}\left[-\log\left(\frac{e^{T^*\left(Syn_{\text{Eye}}(A,B),A\right)_{y|I^2_{Eye}}}}{\sum_{j=1}^{c} e^{T^*\left(Syn_{Eye}(A,B),A\right)_j}}\right)\right]$$

$$+ E_{I_{\text{Eye}} \sim p_{\text{data}}\left(I_{Eye}\right)}\left[-\log\left(\frac{e^{T^*\left(Syn_{\text{Eye}}(D,C),D\right)_{y|I_{Eye}}}}{\sum_{j=1}^{c} e^{T^*\left(Syn_{Eye}(D,C),D\right)_j}}\right)\right]$$

(9)

where, the j is the class index of unnormalized score from $T^*(x)$, the c is the number of identities, and y indicate the ID of the sampled eye image. Also, $A = S^*\left(I^1_{Eye}\right)$, $B = T^*\left(I^2_{Eye}, S^*\left(I^2_{Eye}\right)\right)$, $C = T^*\left(I^1_{Eye}, S^*\left(I^1_{Eye}\right)\right)$ and $D = Syn_{Mask}\left(I^1_{Eye}\right)$. In summary, the training objective function can be expressed mathematically as

$$Syn^*_{Eye} = \arg\min_{Syn_{Eye}, Syn_{\text{Mask}}} \max_{Dis_{Eye}, Dis_{\text{Mask}}}$$

$$L_{\text{unsup}}\left(Syn_{\text{Eye}}, Syn_{\text{Mask}}, Dis_{Eye}, Dis_{\text{Mask}}\right)$$

(10)

where Syn_{Eye} and Syn_{Mask} minimize the objective function, while Dis_{Eye} and Dis_{Mask} maximize it.

5 Experimental results and discussions

5.1 Implementation details

The image pool technique was employed [16] that updates the discriminator networks Dis_{Eye} and Dis_{Mask} using a history of synthetic images rather than the ones generated by the latest synthesis networks ensuring the stability of the IrisGAN training process. An image pool buffer was maintained that stores the 50 previously synthesized images. For ID preserving iris synthesis, data augmentation was carried out using random horizontal flipping, random rotation (-5 to 5 degrees rotation), and the random translation of up to 15 pixels in each spatial dimension in the training images. For iris identification, data augmentation is a commonly used technique in the classification task. However, in the iris identification, the vertical-axis flip is not appropriate, and the random translation is also unnecessary for the extracted iris region. Random brightness and rotation are used for the illumination and rotation invariance. For the random brightness, the RGB domain of the iris region is firstly converted into the HSV domain, then the magnitude of the V channel is multiplied by the uniform distribution of [0.5, 1.5]. Rotation with the uniform distribution of (-15 to 15 degrees) is used for data augmentation.

IrisGAN is trained with mini-batch stochastic gradient descent (SGD) [49] with a mini-batch size of 1. All weights are initialized from a zero-centered truncated normal distribution with a standard deviation of 0.02. All networks are trained with a learning rate of 0.0002 for the first 100, 000 iterations and a linearly decaying rate that goes to zero over the next 100, 000 iterations. Adam optimizer is one of the most pervasive and robust optimizers used in various field. The model is also optimized using the Adam optimizer [23] with $\beta_1 = 0.5$ and $\beta_2 = 0.999$ as suggested in [47]. The following empirical values are used to train the synthesis

Table 4 Statistics of the OPENEDS used for multi-class eye segmentation and iris image synthesis

Category	No. of images	No. of identities
Train	7,658	122
Validation	1,220	122
Test	2,440	122
Total	11,318	122

networks: $\lambda_{cycle} = 10$, $\lambda_{iden} = 100$, and $\lambda_{reg} = 0.0001$ from the experimental observations. In the Leaky ReLU, the slope of the leak is set to 0.2. Reflection padding is used to reduce artifacts instead of zero padding in the convolution layers. The model takes about 5 days to train for 400, 000 iterations using a single GeForce GTX 1080Ti GPU.

5.2 OpenEDS Dataset

To validate the proposed ID preserving iris synthesis, we conduct extensive quantitative and qualitative evaluations over OpenEDS (Open Eye Dataset) [11]. The OpenEDS is collected using virtual reality (VR) head-mounted display mounted with two synchronized eye-facing cameras under controlled illumination. This dataset includes more than 150 individual participants, 12, 759 pixel-level labeled images, and 252, 690 unlabeled images. For the task of iris identification, the initial training and validation datasets were rearranged. In all cases, training, validation, and test sets are mutually disjoint. The new separated dataset can be found in Table 4. OpenEDS includes 122 identities. For each identity, 10 and 20 images are randomly sampled and used to construct the validation and test set respectively, and the remaining images are used as the training images. The number of images for each identity is very different. The statistical information regarding the data distribution is illustrated in Fig. 6. Three identities have more than 150 samples. However, the identity number 98 has only 37 images. Apart from 30 samples for the validation and test combined, the training data for this identity is only seven samples in total.
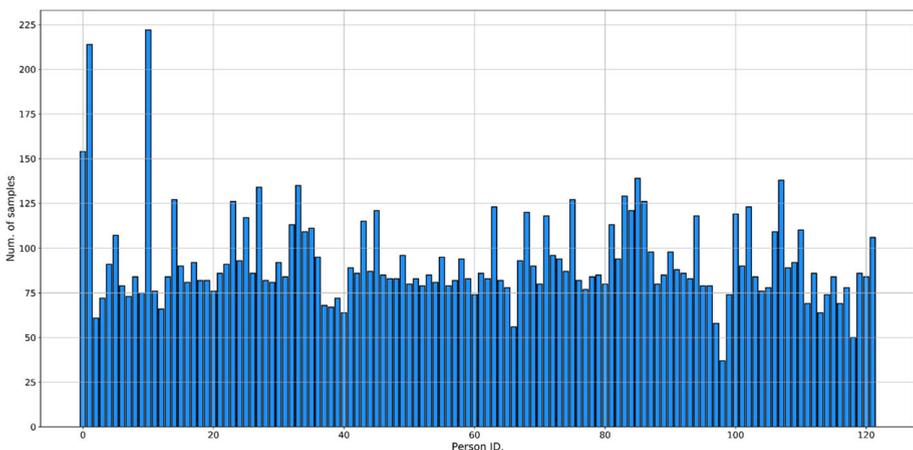
Fig. 6 The number of images for each identity. The number of images has high variance

Besides significant reflection and illumination changes on the small iris region, the limited amount of data for certain identities presents a huge challenge for accurate identification. Several sample images of OpenEDS are shown in Fig. 7. The dataset even includes some participants with eyeglasses. The reflections on the iris from the eyeglasses make it very challenging for the semantic segmentation and personal identification task.

5.3 Evaluation Metrics

The iris identification accuracy is evaluated by the top-1, top-5 accuracy, and the Area Under the Curve (AUC) over the Cumulative Matching Characteristic (CMC) curve [9]. The condition of ID preserving iris is also evaluated by the top-1 accuracy, the top-5 accuracy, and AUC over the CMC curve. Measurement of realism is also another essential issue. The generated image quality is measured using Frechet Inception Distance (FID) [13] score, the FID compares the statistics of the generated synthetic samples against the real samples:

$$FID = \left\| \boldsymbol{\mu}_r - \boldsymbol{\mu}_s \right\|^2 + \mathrm{Tr} \left(\sum_r + \sum_s - 2 \sqrt{\sum_r \sum_s} \right) \tag{11}$$

where $\boldsymbol{\mu}_s$, $\boldsymbol{\mu}_r$, $\sum_r$ and $\sum_s$ represent the statistics of two distributions and T_r is the trace of the co-variance matrix $\left(\sum_r + \sum_s - 2\sqrt{\sum_r \sum_s} \right)$.
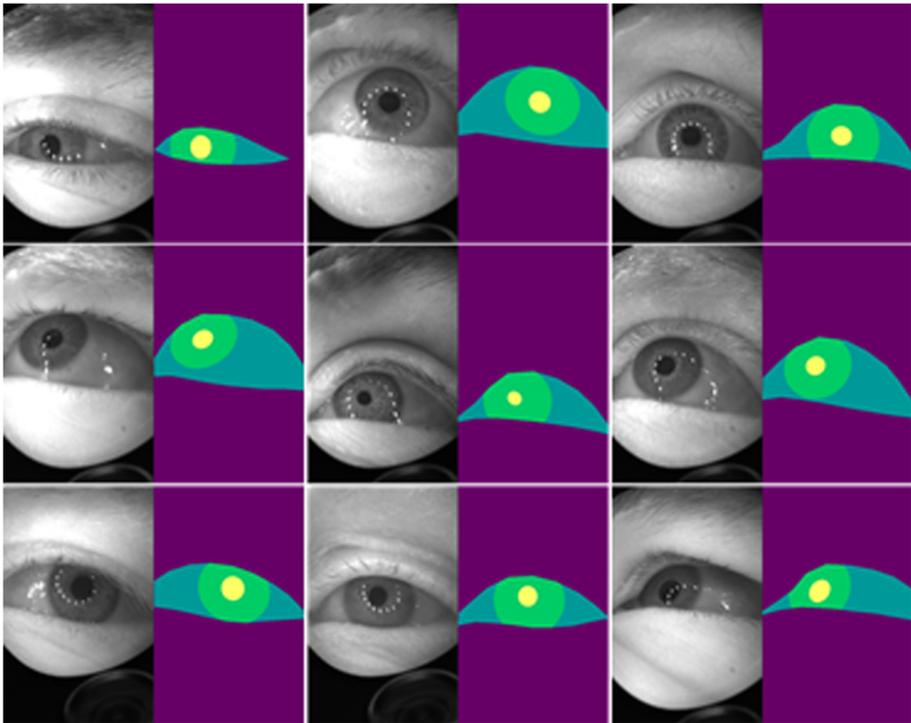

Fig. 7 Sample images of OpenEDS dataset. Annotation involves pupil (yellow), iris (green), and sclera (cyan). The background indicates by violet

Table 5 Top-1, Top-5 and AUC for iris identification on OpenEDS dataset

Category	Training accuracy	Validation accuracy	Test accuracy		
			Top-1	Top-5	AUC
IDNet without data augmentation	100.00	80.41	68.89	89.39	0.977
IDNet with data augmentation	99.88	89.34	80.04	93.07	0.982

The segmentation algorithm requires an accurate estimation of each class in 2D images, typically per-pixel segmentation into the key eye regions: the pupil, the iris, the sclera, and the background. The practical and ideal algorithm should be accurate, robust, extremely power efficient, and possible in real-time processing for the embedded system. Therefore, we evaluate both the accuracy of the model and approximate the complexity using the number of learned parameters in the stage. This metric follows the OpenEDS Semantic Segmentation Challenge; it is defined as follows:

$$M = 0.5 \left(P + \min \left(1, \frac{1}{S} \right) \right)$$ (12)

where $0 \leq P \leq 1$ measures model accuracy as defined by the un-weighted mean intersection-over union (mIoU) score over all classes, and $S > 0$ measures the model complexity, as defined by the number of learned model parameters in the test stage, measured in the unit of model-size in MB. More precisely, model size in MB will be computed as S = number of learned model parameters $\times$ 4(floating point)/(1024 $\times$ 1024).

Moreover, global pixel accuracy, each class accuracy, precision, recall, and F1 score are included for analyzing the model accuracy. The definition of the accuracy, precision, recall, and F1 score is given as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$ (13)

$$Precision = \frac{TP}{TP + FP}$$ (14)
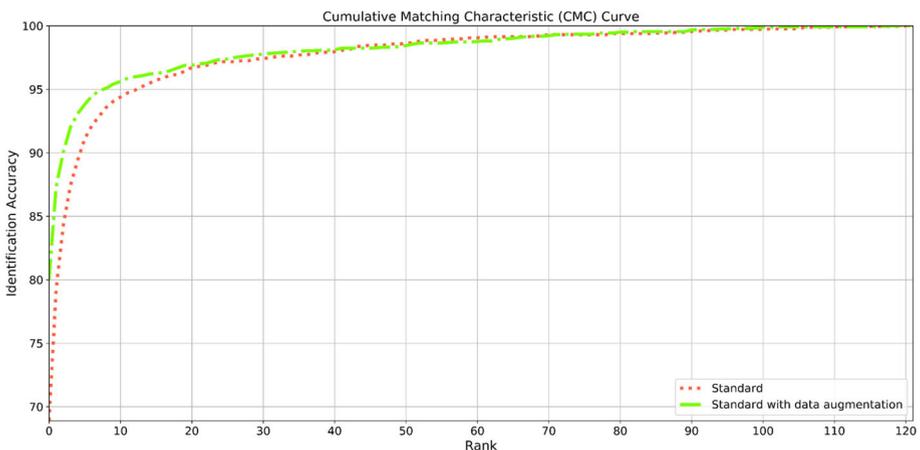
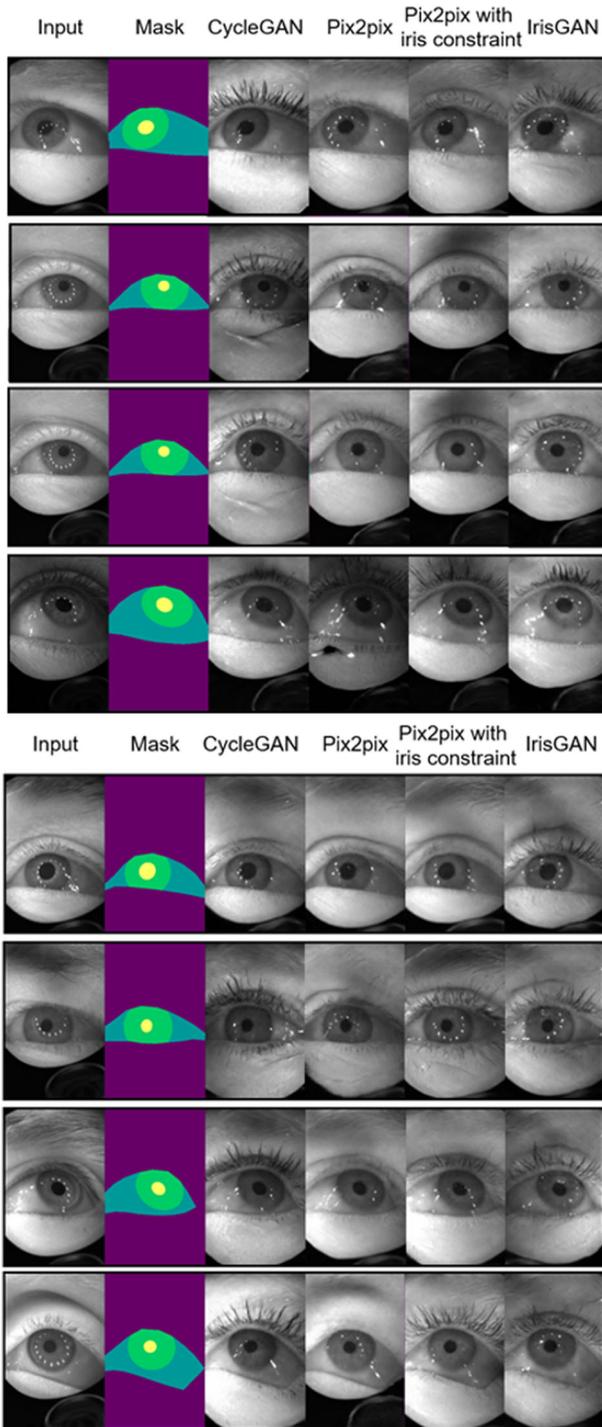**Fig. 8** CMC curve of iris identification on OpenEDS dataset

Fig. 9 Qualitative results of ID preserving IrisGAN

$$Recall = \frac{TP}{TP + FN} \tag{15}$$

$$F1score = \frac{2 \times (Recall \times Precision)}{Recall + Precision} \tag{16}$$

where TP, FP, FN and TN indicate true positive, false positive, false negative, and true negative, respectively. Iris identification is the closed recognition system.

5.4 Experimental results

The purpose of IDNet is to extract ID-preserving features from the iris to provide to IrisGAN. An ablation study was put forward to understand the effect of data augmentation during the training stage. Two sets of samples were separated and one set is subjected to random illumination and rotation during the training stage and another set is just used as it is for the training. For the random brightness, the RGB domain of the iris region is firstly converted into the HSV domain, then the magnitude of the V channel is multiplied by the uniform distribution of [0.5, 1.5]. Rotation with the uniform distribution of (−15 to 15 degrees) is used for data augmentation. Table 5 and Fig. 8 demonstrate the performance of the IDNet. Top-1 and top-5 accuracy of IDNet without or with data augmentation are 80.04 and 93.07 respectively. The AUC score improves from 0.9771 to 0.9816 after doing random illumination and rotation during the training stage.

A segmentation mask estimated from SegNet becomes an input to the ID preserving IrisGAN, and the IrisGAN tries to generate a new iris image while keeping a given ID from another input iris. The input images, the estimated masks, and generated iris images are represented in Fig. 9. The given ID is randomly selected from the test set, and all the generated iris images in Fig. 9 are constructed using the same ID of the identity. We observe that the generated iris images correspond to the input mask states appropriately, and the proposed ID preserving IrisGAN can generate a massive number of iris images at no cost. It will be of significant benefit in both developing and evaluating the iris recognition algorithms.
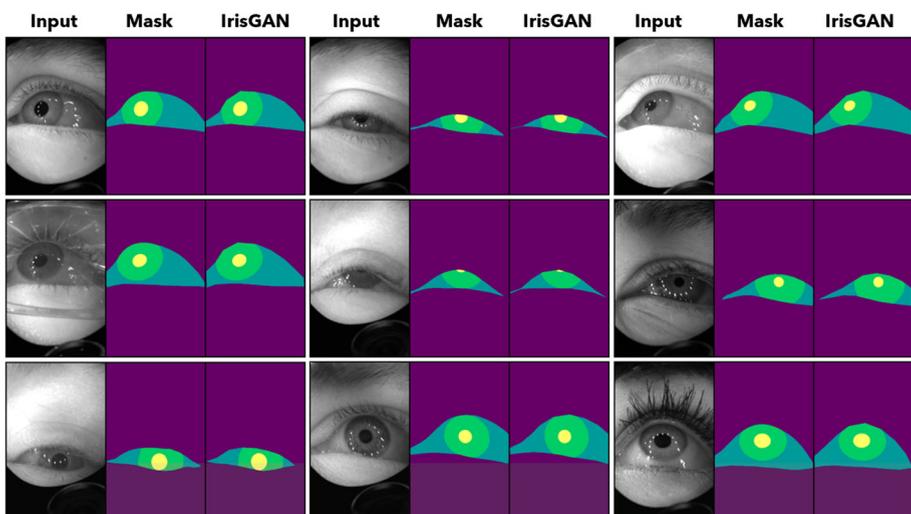

Fig. 10 Segmentation results based on semi-supervised adversarial learning on validation of OpenEDS dataset
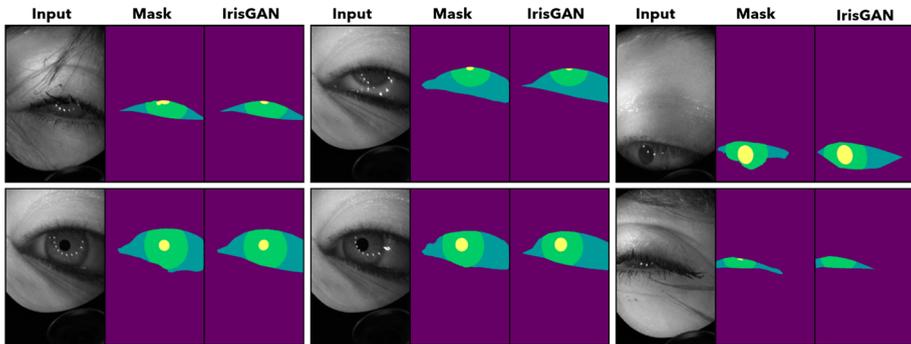
Fig. 11 Bad examples using our method on validation of OpenEDS dataset

To examine the multi-class eye segmentation capability of the proposed framework, a set of iris samples with diverse levels of difficulties is employed. The results using semi-supervised adversarial learning for the multi-class eye segmentation are given in Fig. 10. Also, Fig. 11 shows some bad cases for the segmentation. Even though there are various reflection and illumination situations, our method can robustly detect the pupil, iris, and sclera. The failure cases mostly happen when occlusion with eyelid and eyelash.

Table 6 demonstrates the comparative results for conditional ID preserving generation of synthetic irises. The top-1, top-5, and AUC are measured by using 1, 440 synthetic iris images while selecting random ID from the data set. CycleGAN [62], Pix2pix [16], and Pix2pix [16] with iris constrain are utilized to compare with the proposed method. The mean and standard deviation are included for five trails. The proposed method has a higher ID preserving characteristic compared with baseline methods. The top-1 and top-5 accuracy achieve 21.842 and 79.234, respectively. Fig. 12 gives comparison results with baseline methods for ID preserving iris synthesis using CMC curve.

Distributions of matching distance are shown following figures, which are evaluated using the features extracted from the last layer of the identification network. All the possible combinations are illustrated in Table 7. The distribution of matching distance for genuine matching between the live and synthetic iris generation using the same live iris image is shown in Fig. 13(A). The distribution of matching distance for genuine matching between live and synthetic iris generation using the different live iris images is shown in Fig. 13(B). The distribution of matching distance for genuine matching between the synthetic iris and synthetic iris is shown in Fig. 13(C). Figure 13(D) shows the distribution of matching distance

Table 6 Comparison results for conditional ID preserving generation of synthetic irises

Model	FID	Test accuracy		
		Top-1	Top-5	AUC
Zhu et al. [62]	42.58 ± 3.23	5.737 ± 0.652	18.182 ± 1.571	0.6891 ± 0.0158
Isola et al. [16]	54.94 ± 4.18	8.146 ± 0.735	23.382 ± 0.972	0.7303 ± 0.0092
Isola et al. [16] with iris constrain	45.11 ± 3.18	6.558 ± 0.442	19.974 ± 0.510	0.7073 ± 0.0069
Proposed method	48.43 ± 4.42	21.842 ± 0.693	49.234 ± 1.205	0.8503 ± 0.0082

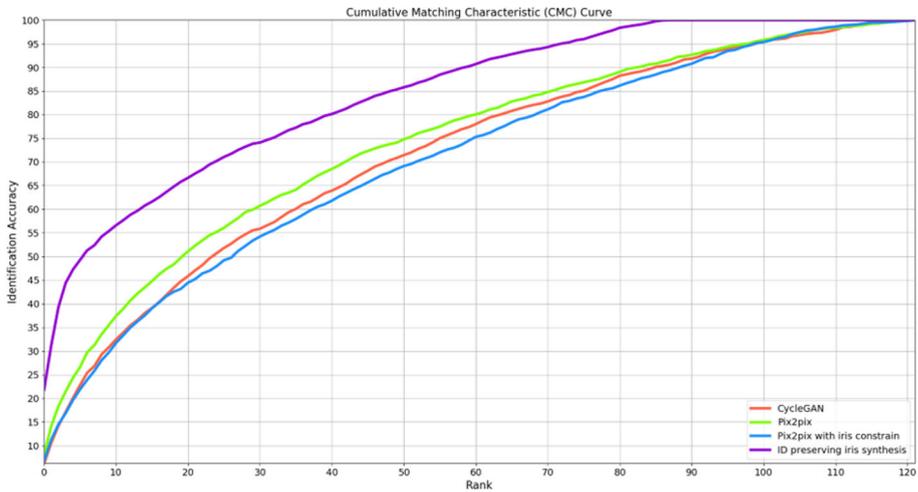The means and standard deviations are presented for the five trials

Fig. 12 Comparison results with baseline methods for ID preserving iris synthesis using CMC curve

for imposter matching between live and synthetic iris. Fig. 13(E) illustrates the distribution of matching distance for imposter matching between the synthetic iris and synthetic iris.

The multi-class eye segmentation results are evaluated on the test server of the OpenEDS Semantic Segmentation Challenge. Therefore, the test data is not approachable. In semantic segmentation, the model accuracy and model complexity were considered at the same time. The original U-Net architecture [48] has been analyzed, revised, and trained as the Dense U-Net-Light with semi-supervised adversarial learning.

Table 8 demonstrates the comparison results for the improvement of our method. The original U-Net structure is too cumbersome for the task, and it is not suitable for the embedded system. Therefore, the number of convolutional filters is reduced by half each time, and we represent C1/K. The compressed versions of the U-Net from C1/2, C1/4, and C1/8, the mIoU is not changed too much. However, the accuracy of U-Net with C1/16 drops sharply. The Dense U-Net-Light is constructed based on the U-Net C1/8. Because of the dense connections, the accuracy of the Dense U-Net-Light with supervised learning improves a lot. Based on adversarial learning, the network studies anatomical structure from the ground truth to improve from 94.380 (mIoU) to 94.597 (mIoU) by reducing false positives. The

Table 7 Combinations of matching distance

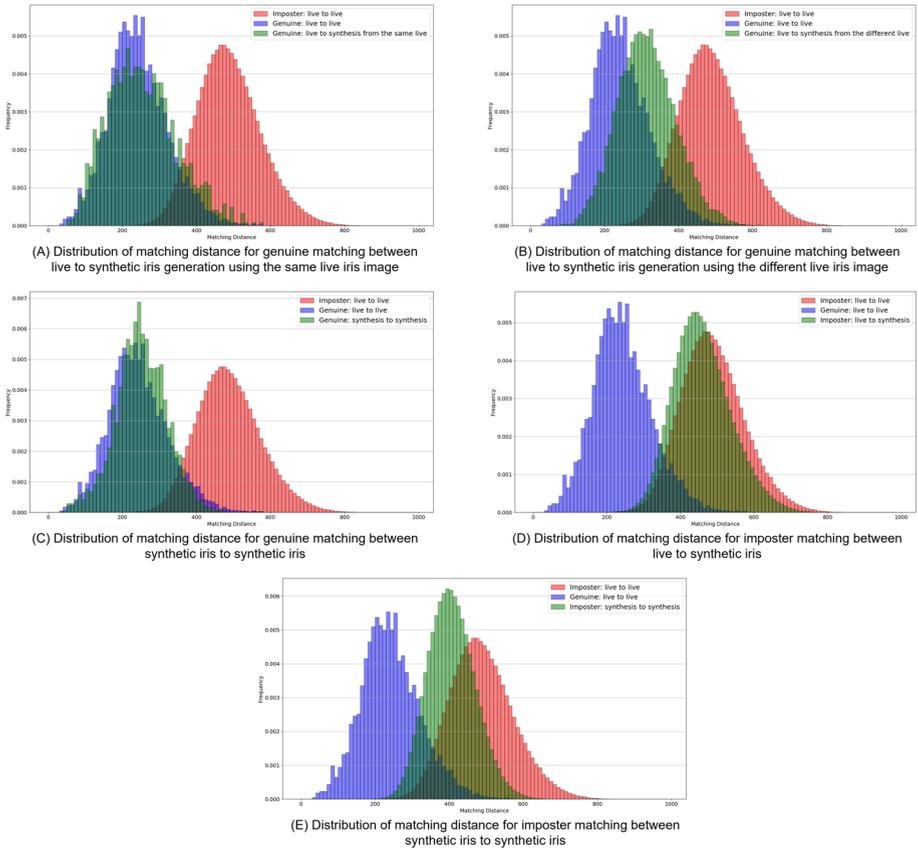| Category	Reference	Test	Number of combinations
	Live	Live	$C_{122}^1 \cdot C_{10}^2 = 5,490$
Genuine	Live	Synthesis from same live	$C_{122}^1 \cdot C_{10}^1 = 1,220$
	Live	Synthesis from different live	$C_{122}^1 \cdot C_{10}^1 \cdot C_9^1 = 10,980$
	Synthesis	Synthesis	$C_{122}^1 \cdot C_{10}^2 = 5,490$
Imposter	Live	Live	$C_{122}^2 \cdot C_{10}^1 \cdot C_{10}^1 = 738,100$
	Live	Synthesis	$C_{122}^2 \cdot C_{10}^1 \cdot C_{10}^1 = 738,100$
	Synthesis	Synthesis	$C_{122}^2 \cdot C_{10}^1 \cdot C_{10}^1 = 738,100$

Fig. 13 Distributions of matching distances

mIoU of network training with semi-supervised adversarial learning increases to 94.900. We speculate that the improvement is caused by the discriminator, observing more generated results of unlabeled images to promote capability. Then the discriminator delivers more precise feedback to the generator, Dense U-Net-Light. Besides, training with massive labels with noise data can improve performance in advance. However, the effect is limited. The comparison result with other methods is also given in Table 9.

6 Conclusion

In this paper, a method of ID-preserving iris synthesis is proposed to prove that it can be applied in the generation of synthetic iris databases for various iris recognition applications. This work successfully combines the main idea of generative adversarial learning, segmentation, and identification to approach the problem statement of ID-preserving synthetic iris generation. IrisGAN produces synthetic iris images from the segmentation masks given ID information. The segmentation mask, iris pose, is decided from the input image by using SegNet. By doing this, ID preserving IrisGAN generates an unlimited number of synthetic

Table 8 Evaluation results of SegNet on the OpenEDS test set

Model	Mean IoU	Model complexity	Global pixel accuracy	Class accuracy				F1 score
				Back	Sclera	Iris	pupil	
U-Net supervised	92.949	31030788.000	0.990	0.996	0.951	0.959	0.929	0.990
U-Net CI/2 supervised	93.042	7759620.000	0.990	0.996	0.951	0.962	0.945	0.990
U-Net CI/4 supervised	92.873	1940868.000	0.989	0.996	0.947	0.965	0.945	0.990
U-Net CI/8 supervised	92.924	485700.000	0.990	0.996	0.951	0.967	0.933	0.990
U-Net CI/16 supervised	92.230	121668.000	0.989	0.996	0.938	0.959	0.931	0.990
Dense U-Net-Light supervised	94.380	261860.000	0.992	0.997	0.961	0.968	0.954	0.990
Dense U-Net-Light supervised adversarial	94.597	261860.000	0.992	0.997	0.959	0.969	0.955	0.990
Dense U-Net-Light semi-supervised adversarial	94.900	261860.000	0.992	0.997	0.960	0.975	0.962	0.990
+ massive noisy label	94.970	261860.000	0.992	0.997	0.951	0.961	0.974	0.990

Table 9 Comparison results with baseline methods on the test set of OpenEDS Semantic Segmentation Challenge

Model	Mean IoU	Model complexity	Global pixel accuracy	Class accuracy				F1 score
				Back	Sclera	Iris	pupil	
U-Net [48]	92.949	31 030,788	0.99	0.996	0.951	0.959	0.929	0.99
SegNet [2]	89.478	416,088	0.976	Unknown	Unknown	Unknown	Unknown	0.974
EyeNet [22]	95.112	258,021	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown
Ours	94.97	261,860	0.992	0.997	0.951	0.961	0.974	0.99

iris images by changing input images. The generated iris images are validated by measuring top-1, top-5, and AUC. The SegNet and IDNet performance was evaluated using class accuracy in terms of precision, recall, and F1-score alongside the computation model complexity. This study exhibits ease of use, compatibility, and accuracy in preserving ID information for the generated synthetic images compared to the other baseline methods. Additionally, the shortcomings of the current study were reported with possible future directions toward a better iris synthesis approach.

7 Limitations

The proposed technique for ID preserving iris synthesis has the following limitations:

1. The experiments corresponding to the ID preserving iris synthesis were carried out specifically on the OpenEDS dataset; however, further datasets such as CASIA-iris-fake [61], ND-IRIS-0405 [45], and UBIRIS.v2 Database [46] are required to be investigated.
2. The performed experiments are targeting a single iris domain in terms of ID-preserving iris synthesis; however, various domains such as cosmetic contact lenses, and printed eyes are yet to be explored in the context of this study.
3. The proposed method validated the iris recognition and ID preserving sample synthesis with limited metrics pertaining to a specific dataset; however, the need for employing multiple metrics [3] on diverse datasets is still in progress for future work.
4. The proposed study used a single network architecture to perform the iris recognition; however using various neural network architectures and refining the better-performed architecture may produce the best results [60]
5. The interdependency of the IDNet and IrisGAN must be explored in a deeper sense to put forward a better approach toward identification.

8 Future directions

1. In the future, usage of different domains for iris synthesis may assist the improvement in overall results on the extended datasets.
2. Various datasets for iris recognition can be used to improve the diversity and aid in the critical validation of the proposed framework.
3. The future study aims to employ the synthetic iris generation system on various iris types based on age, quality, and disorientation; thereby, evaluating the performance of the proposed framework w.r.t several metrics in the context of iris recognition task [3].
4. A comparative study on ID preserving iris generation [39] and presentation attack detection [60] w.r.t diverse metrics can be carried out as a future study which may reveal critical findings in the research domain.
5. Also, the identification networks must be refined so as to consider the state-of-the-art physical adversarial and presentation attacks [5, 6, 54, 55] that pose a threat to the identification accuracy and also extend the IDNet to cater the multi-modalities such as the face [42, 57], iris, and finger-vein with single ID preserving information.
6. Additionally, alternative approaches such as vision-based transformer attention networks to the SegNet component must be explored in detail.

Data Availability The data used in the study is available upon request and directed to the corresponding author.

Declarations

Conflict of interest/Competing interests The authors have no relevant financial or non-financial interests to disclose.

References

1. Abdigapporov S, Miraliev S, Kakani V, Kim H (2023) Joint multiclass object detection and semantic segmentation for autonomous driving. IEEE Access
2. Badrinarayanan V, Kendall A, Cipolla R (2017) Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence 39(12):2481–2495
3. Boutros F, Damer N, Raja K, Ramachandra R, Kirchbuchner F, Kuijper A (2020) Iris and periocular biometrics for head mounted displays: Segmentation, recognition, and synthetic data generation. Image and Vision Computing 104:104007
4. Cappelli R, Erol A, Maio D, Maltoni D (2000) Synthetic fingerprint-image generation. In: Proceedings 15th International Conference on Pattern Recognition. ICPR-2000, vol. 3, pp 471–474. IEEE
5. Cheng Z, Liang J, Choi H, Tao G, Cao Z, Liu D, Zhang X (2022) Physical attack on monocular depth estimation with optimal adversarial patches. In: Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXXVIII, pp 514–532. Springer
6. Cheng Z, Liang J, Tao G, Liu D, Zhang X (2023) Adversarial training of self-supervised monocular depth estimation against physical-world attacks. arXiv preprint arXiv:2301.13487
7. Choudhary M, Tiwari V, Venkanna U (2020) Iris liveness detection using fusion of domain-specific multiple bsif and densenet features. IEEE transactions on cybernetics 52(4):2370–2381
8. Cui J, Wang Y, Huang J, Tan T, Sun Z (2004) An iris image synthesis method based on pca and super-resolution. In: Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., vol 4. pp 471–474. IEEE
9. Dabouei A, Kazemi H, Iranmanesh SM, Dawson J, Nasrabadi NM et al (2018) Id preserving generative adversarial network for partial latent fingerprint reconstruction. In: 2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS), pp 1–10. IEEE
10. Daugman JG (1993) High confidence visual recognition of persons by a test of statistical independence. IEEE transactions on pattern analysis and machine intelligence 15(11):1148–1161
11. Garbin SJ, Shen Y, Schuetz I, Cavin R, Hughes G, Talathi SS (2019) Openeds: Open eye dataset. arXiv preprint arXiv:1905.03702
12. Ghimire A, Kakani V, Kim H (2023) Ssrt: A sequential skeleton rgb transformer to recognize fine-grained human-object interactions and action recognition. IEEE Access, p 1–1 https://doi.org/10.1109/ACCESS.2023.3278974
13. Heusel M, Ramsauer H, Unterthiner T, Nessler B, Hochreiter S (2017) Gans trained by a two time-scale update rule converge to a local nash equilibrium. Advances in neural information processing systems 30
14. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 770–778
15. He K, Zhang X, Ren S, Sun J (2016) Identity mappings in deep residual networks. In: European Conference on Computer Vision, pp 630–645. Springer
16. Isola P, Zhu J-Y, Zhou T, Efros AA (2017) Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 1125–1134
17. Jayanthi J, Lydia EL, Krishnaraj N, Jayasankar T, Babu RL, Suji RA (2021) An effective deep learning features based integrated framework for iris detection and recognition. Journal of Ambient Intelligence and Humanized Computing 12:3271–3281
18. Johnson J, Alahi A, Fei-Fei L (2016) Perceptual losses for real-time style transfer and super-resolution. In: European Conference on Computer Vision, pp 694–711. Springer

19. Juraev S, Ghimire A, Alikhanov J, Kakani V, Kim H (2022) Exploring human pose estimation and the usage of synthetic data for elderly fall detection in real-world surveillance. IEEE Access 10:94249–94261

20. Kakani V, Cui X, Ma M, Kim H (2021) Vision-based tactile sensor mechanism for the estimation of contact position and force distribution using deep learning. Sensors 21(5):1920

21. Kakani V, Lee S, Cui X, Kim H (2022) Performance analysis of spiking neural network using temporal spike-based backpropagation on field programmable gate array (fpga) platform. In: 2022 IEEE Region 10 Symposium (TENSYMP), pp 1–6. IEEE

22. Kansal P, Devanathan S (2019) Eyenet: Attention based convolutional encoder-decoder network for eye region segmentation. In: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), pp 3688–3693. IEEE

23. Keskar NS, Socher R (2017) Improving generalization performance by switching from adam to sgd. arXiv preprint arXiv:1712.07628

24. Ketchum KE (2009) Facegen and the technovisual politics of embodied surfaces. Women's Studies Quarterly 37(1/2):183–199

25. Kim T, Cha M, Kim H, Lee JK, Kim J (2017) Learning to discover cross-domain relations with generative adversarial networks. In: International Conference on Machine Learning, pp 1857–1865. PMLR

26. Kohli N, Yadav D, Vatsa M, Singh R, Noore A (2017) Synthetic iris presentation attack using idcgan. In: 2017 IEEE International Joint Conference on Biometrics (IJCB), pp 674–680. IEEE

27. Lee H, Kim S (2021) Sspnet: Learning spatiotemporal saliency prediction networks for visual tracking. Information Sciences 575:399–416

28. Lefohn A, Budge B, Shirley P, Caruso R, Reinhard E (2003) An ocularist's approach to human iris synthesis. IEEE Computer Graphics and Applications 23(6):70–75

29. Liao X, Huang Z, Peng L, Qiao T (2021) First step towards parameters estimation of image operator chain. Information Sciences 575:231–247

30. Liu Y, Dai W, Fang F, Chen Y, Huang R, Wang R, Wan B (2021) Dynamic multi-channel metric network for joint pose-aware and identity-invariant facial expression recognition. Information Sciences 578:195–213

31. Li Y, You J, Zhou J, Wang W, Liao X, Li X (2022) Image operation chain detection with machine translation framework. IEEE Transactions on Multimedia

32. Maas AL, Hannun AY, Ng AY et al (2013) Rectifier nonlinearities improve neural network acoustic models. In: Proc. Icml, vol 30. p 3 Citeseer

33. Maio D, Maltoni D, Cappelli R, Wayman JL, Jain AK (2002) Fvc 2000: Fingerprint verification competition. IEEE transactions on pattern analysis and machine intelligence 24(3):402–412

34. Maio D, Maltoni D, Cappelli R, Wayman JL, Jain AK (2002b) Fvc2002: Second fingerprint verification competition. In: Object Recognition Supported by User Interaction for Service Robots, vol 3. pp 811–814. IEEE

35. Maio D, Maltoni D, Cappelli R, Wayman JL, Jain AK (2004) Fvc2004: Third fingerprint verification competition. In: International Conference on Biometric Authentication, pp 1–7. Springer

36. Makthal S, Ross A (2005) Synthesis of iris images using markov random fields. In: 2005 13th European Signal Processing Conference, pp 1–4. IEEE

37. Minaee S, Abdolrashidi A (2018) Iris-gan: Learning to generate realistic iris images using convolutional gan. arXiv preprint arXiv:1812.04822

38. Miraliev S, Abdigapporov S, Kakani V, Kim H (2023) Real-time memory efficient multitask learning model for autonomous driving. IEEE Transactions on Intelligent Vehicles

39. Morampudi MK, Prasad MV, Raju U (2020) Privacy-preserving iris authentication using fully homomorphic encryption. Multimedia Tools and Applications 79(27):19215–19237

40. Ni J, Shen K, Chen Y, Cao W, Yang SX (2022) An improved deep network-based scene classification method for self-driving cars. IEEE Transactions on Instrumentation and Measurement 71:1–14

41. Ni J, Shen K, Chen Y, Yang SX (2023) An improved ssd-like deep network-based object detection method for indoor scenes. IEEE Transactions on Instrumentation and Measurement 72:1–15

42. Ning X, Gou D, Dong X, Tian W, Yu L, Wang C (2022) Conditional generative adversarial networks based on the principle of homologycontinuity for face aging. Concurrency and Computation: Practice and Experience 34(12):5792

43. Orlans NM, Buettner DJ, Marques J (2004) A survey of synthetic biometrics: Capabilities and benefits. In: IC-AI, pp 499–505

44. Orlans NM, Piszcz AT, Chavez RJ (2003) Parametrically controlled synthetic imagery experiment for face recognition testing. In: Proceedings of the 2003 ACM SIGMM Workshop on Biometrics Methods and Applications, pp 58–64

45. Phillips PJ, Scruggs WT, O'Toole AJ, Flynn PJ, Bowyer KW, Schott CL, Sharpe M (2009) Frvt 2006 and ice 2006 large-scale experimental results. IEEE transactions on pattern analysis and machine intelligence 32(5):831–846

46. Proença H, Filipe S, Santos R, Oliveira J, Alexandre LA (2009) The ubiris. v2: A database of visible wavelength iris images captured on-the-move and at-a-distance. IEEE Transactions on Pattern Analysis and Machine Intelligence 32(8):1529–1535

47. Radford A, Metz L, Chintala S (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434

48. Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, pp 234–241. Springer

49. Shrivastava A, Pfister T, Tuzel O, Susskind J, Wang W, Webb R (2017) Learning from simulated and unsupervised images through adversarial training. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 2107–2116

50. Tan J, Liao X, Liu J, Cao Y, Jiang H (2021) Channel attention image steganography with generative adversarial networks. IEEE Transactions on Network Science and Engineering 9(2):888–903

51. Tapia JE, Gonzalez S, Busch C (2021) Iris liveness detection using a cascade of dedicated deep learning networks. IEEE Transactions on Information Forensics and Security 17:42–52

52. Ulyanov D, Vedaldi A, Lempitsky V (2016) Instance normalization: The missing ingredient for fast stylization. arXiv preprint arXiv:1607.08022

53. Vijaykumar V (2021) A qualitative analysis on iris localization techniques and approaches of machine learning in iris recognition systems. Design Engineering, p 10921–10944

54. Wang W, Han C, Zhou T, Liu D (2022) Visual recognition with deep nearest centroids. arXiv preprint arXiv:2209.07383

55. Wei H, Tang H, Jia X, Yu H, Li Z, Wang Z, Satoh S, Wang Z (2022) Physical adversarial attack meets computer vision: A decade survey. arXiv preprint arXiv:2209.15179

56. Wei Z, Tan T, Sun Z (2008) Synthesis of large realistic iris databases using patch-based sampling. In: 2008 19th International Conference on Pattern Recognition, pp 1–4. IEEE

57. Wu G, He F, Zhou Y, Jing Y, Ning X, Wang C, Jin B (2022) Acgan: Age-compensated makeup transfer based on homologous continuity generative adversarial network model. IET Computer Vision

58. Xiao T, Xia T, Yang Y, Huang C, Wang X (2015) Learning from massive noisy labeled data for image classification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 2691–2699

59. Yadav S, Chen C, Ross A (2019) 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, pp 0–0

60. Yadav S, Ross A (2021) 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, pp 2412–2421

61. Zhang H, Sun Z, Tan T, Wang J (2011) Learning hierarchical visual codebook for iris liveness detection. In: International Joint Conference on Biometrics, vol 1

62. Zhu J-Y, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp 2223–2232

63. Zou H, Zhang H, Li X, Liu J, He Z (2021) Generation textured contact lenses iris images based on 4dcycle-gan. In: 2018 24th International Conference on Pattern Recognition (ICPR), pp 3561–3566. IEEE

64. Zuo J, Schmid NA, Chen X (2007) On generation and analysis of synthetic iris images. IEEE Transactions on Information Forensics and Security 2(1):77–90