

RESEARCH ARTICLE

A Survey on Synthetic Biometrics: Fingerprint, Face, Iris and Vascular Patterns

ANDREY MAKRUSHIN¹, ANDREAS UHL², AND JANA DITTMANN¹, (Member, IEEE)¹Advanced Multimedia and Security Laboratory (AMSL), Otto von Guericke University Magdeburg (OVGU), 39106 Magdeburg, Germany²Multimedia Signal Processing and Security Laboratory (Wavelab), University of Salzburg, 5020 Salzburg, Austria

Corresponding author: Andrey Makrushin (andrey.makrushin@ovgu.de)

The research in this paper is a part of the joint project GENSYNTH (Tools for the Generation of Synthetic Biometric Sample Data). AMSL is funded in part by the Deutsche Forschungsgemeinschaft (DFG) under project no. 421860227. Wavelab is funded in part by the Austrian Science Fund (FWF) under project no. I4272.

ABSTRACT Synthetic biometric samples are created with an ultimate goal of getting around privacy concerns, mitigating biases in biometric datasets, and reducing the sample acquisition effort to enable large-scale evaluations. The recent breakthrough in the development of neural generative models shifted the focus from image synthesis by mathematical modeling of biometric modalities to data-driven image generation. This paradigm shift on the one hand greatly improves the realism of synthetic biometric samples and therefore enables new use cases, but on the other hand new challenges and concerns arise. Despite their realism, synthetic samples have to be checked for appropriateness for the tasks they are intended which includes new quality metrics. Focusing on sample images of fingerprint, face, iris and vascular patterns, we highlight the benefits of using synthetic samples, review the use cases, and summarize and categorize the most prominent studies on synthetic biometrics aiming at showing recent progress and the direction of future research.

INDEX TERMS Biometrics, biometric modeling, face, fingerprint, iris, synthetic data, vascular patterns.

I. INTRODUCTION

The need for a new survey on synthetic biometrics is argued by the rush development and popularization of Deep Convolutional Neural Networks (DCNN) moving the focus in image generation from mathematical modeling to data-driven synthesis. This can be seen as a paradigm change because of an overall change in techniques to be used and requirements on auxiliary data. While in early studies the problem of generating synthetic biometric samples was addressed by (mathematical) modeling of physiological structures and acquisition procedures [1], modern studies rely on Generative Adversarial Networks (GAN) transferring the underlying image characteristics from training data to synthetic biometric samples. Moreover, the existing surveys on this topic are quite outdated having been published in 2004 [2] and 2006 [1].

The most powerful driver for the recent development of synthetic biometrics is the introduction of regulations on protection of private data such as the General Data

Protection Regulation (GDPR) in the EU [3]. Biometric data are a special kind of private data protected by Article 9 of the GDPR. Biometric data cannot even be anonymized due to its nature. This fact dramatically hinders usage of real biometric samples not only in industrial but also in academic research. Synthetic biometrics solves this problem by introducing virtual individuals. And biometric samples of virtual individuals can be made public without any legal concerns.

If done properly, replacement of datasets of real biometric samples by synthetic datasets is a step towards open research which requires the reproducibility of experimental results by any independent third party and hence public sharing of both algorithms and datasets used in studies [4].

Besides privacy protection, an important use case of synthetic data is augmentation of training and evaluation datasets. Synthetic samples may compensate for unavailable data aiming at covering all possible natural variations as well as unnatural malicious modifications. Taking control over sample variation and sample attributes as well as ensuring equal distribution of sample attributes (e.g. race, gender, age)

The associate editor coordinating the review of this manuscript and approving it for publication was Vincenzo Conti¹.

in datasets enables fair and unbiased application of machine learning.

The main advantage of neural generative models in comparison to traditional modeling is the realistic appearance of generated samples. In recent studies, GAN models have been already used for rendering photorealistic face, iris and fingerprint images. However, the number of publicly available datasets of synthetic biometric images is very limited, as well as the commonly agreed evaluation methodology for appropriateness of synthetic biometric datasets is still not presented.

Our main contribution is in reviewing the use cases of synthetic samples as well as summarizing and categorizing the most prominent studies on synthetic biometrics aiming at showing recent progress and the direction of future research.

The remainder of the paper is organized as follows: Section II elaborates on the nature of synthetic biometric samples and categorizes use cases, requirements on synthetic data and generation techniques. Section III outlines traditional modeling studies to create synthetic fingerprints, faces, irises and vascular patterns. Section IV focuses on data-driven image generation of the same modalities. The public synthetic biometric datasets and synthesis tools are listed in Section V. Our summary is drawn in Section VI.

II. FUNDAMENTALS

A. SYNTHESIS IN BIOMETRICS

Following the argumentation in [5], we notice that the “biometric data” is an ambiguous term, so is a synthesis of “biometric data”. This survey focuses on biometric samples as “biometric data” acquired by biometric sensors and represented in the form of images which is the case for several of the most established biometric modalities: face, fingerprint, iris, and vascular traits. Synthesis of other types of “biometric data” e.g. feature measurements, matching scores and decision data is beyond the scope of this work.

1) SYNTHESIS AS MODELING

In the encyclopedia of biometrics, Buettner [6] defines the ultimate goal of biometric sample synthesis as the application of computer-aided parametric modeling to creation of a synthetic corpus of biometric samples which is indistinguishable from a corpus of real biometric samples obtained from people. The parametric models are either mathematical equations describing physics of biometric sample acquisition or statistical models derived from empirical analysis of real biometric samples.

Buettner stresses that modeling helps on the one hand to fundamentally understand which factors affect the digitization process of biometric modalities for a specific sensor type and on the other hand to efficiently generate synthetic images that are visually similar to real biometric samples which in turn may improve testing of biometric algorithms. This perfectly fits into an analysis-by-synthesis paradigm

in which the models of real-world phenomena are learnt to predict perceptual observations [1].

2) SYNTHESIS AS TEMPLATE INVERSION

In fact, biometric analysis can be seen as a special case of representation learning, a process in which a raw object is translated to an abstract feature vector that is suitable for object recognition. From this simplistic perspective, synthesis of biometric samples has been seen as an inverse task to biometric analysis [1] in which a raw biometric signal is reconstructed from its abstract representation.

Let us refer to the result of biometric analysis as a biometric template, then biometric template inversion is a special case of parametric modeling in which the identity of a subject is given by the data stored in a template. In early stages of research it was believed that biometric samples cannot be reconstructed from templates. Meanwhile, this belief has been contradicted. In [7] one can find an overview of approaches to invert biometric templates.

For face, the identity is directly associated with a facial image and there is no commonly agreed convention on features to be stored apart from a frontal face image [8]. Hence, for face, the problem of template inversion exists only for deep learning based representations. The de facto standard for iris representation is a 2048-bit binary code called IrisCode [9]. How IrisCode can be turned into an iris image is demonstrated in [10]. For fingerprint, the template comprises a list of minutiae represented by type, x- and y-coordinates, and an orientation angle [11]. The most recent study on fingerprint reconstruction is in [12]. Note that for the template inversion task the visual realism of the resulting pattern is not always a primary objective.

3) PARADIGM'S CHANGE

Similar to how the focus of biometric research has been changing from the proof of the concept to applicability of biometrics, to security issues in biometric systems e.g. template protection, to intentional attacks mitigation like Presentation Attack Detection (PAD), the focus of synthetic biometrics has been changing from physics-based modeling to statistical modeling and finally to data-driven synthesis.

Physics-based modeling relies on studying physics of an object of interest and its interaction with an environment to predict possible projections onto the sensor space. Statistical modeling relies on statistical analysis of the object's representations to predict the representations of modified objects. Data-driven synthesis relies on collecting enough representations for learning of a deep neural network that is then capable of generating random representations as well as representations of modified or even unseen objects.

B. USE CASES FOR SYNTHETIC SAMPLES

Since synthetic biometric samples can be generated with significantly less effort in terms of time and manpower than collecting real biometric samples, the general motivation of

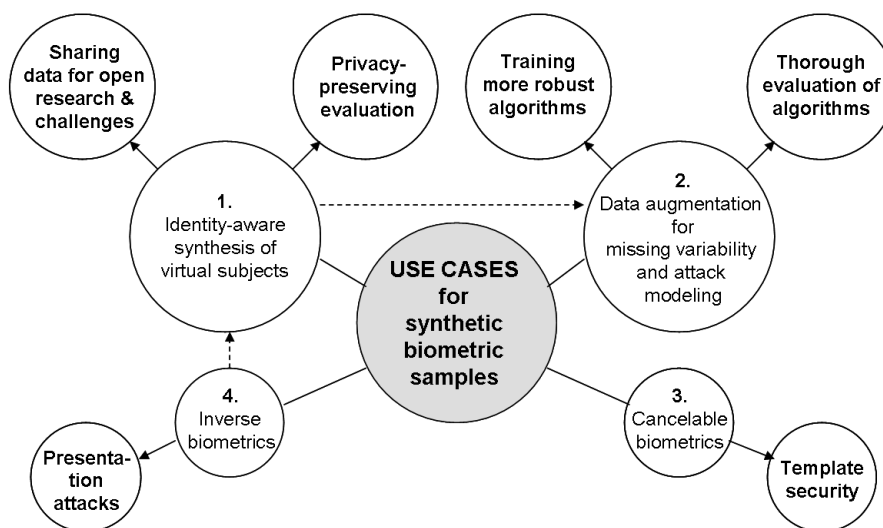


FIGURE 1. Use cases for synthetic biometrics; arrows for specification of use cases and dashed arrows for implication of one use case in another.

creating synthetic samples is a lack of collected samples of a specific kind [5]. To be more precise, synthesis helps to compensate for biases and simulate natural (e.g. aging, environmental influences) or malicious (e.g. presentation attacks) variations at less cost. Use cases for synthetic data in medical research are listed in [13].

We propose to assign use cases to one of the four categories: (1) Identity-aware synthesis of virtual subjects, (2) Data augmentation for missing variability and attack modeling, (3) Cancelable biometrics, and (4) Inverse biometrics; see Fig. 1. The first two use cases are common for all kinds of synthetic data without special focus on biometric samples.

1) IDENTITY-AWARE SYNTHESIS OF VIRTUAL SUBJECTS

The first use case can be interpreted as replacement of real data by synthetic data which helps getting around data privacy concerns [4]. Introduction of virtual subjects with the corresponding biometric modalities has two major goals: (i) Sharing data for open research and challenges, and (ii) Privacy-preserving development of biometric systems including training and evaluation.

2) DATA AUGMENTATION

Data augmentation is a common approach for bias mitigation. In fact, it is tremendously hard or almost impossible to collect data that cover most of the possible presentation variations of a biometric modality or even collect data from minority group people. At the same time, missing important variations in training samples lead, on the one hand, to biased models and, on the other hand, to high classification error rates. Moreover, biometric classification models are often subjected to intentional attacks that could be modeled in synthetic samples. All in all, there are two major goals of data augmentation: (i) Training more robust algorithms and (ii) Thorough evaluation of already trained classification

models (better testing of biometric algorithms). Data augmentation for advanced model training has received massive attention in recent studies in e.g. fingerprint classification [14] and face recognition [15]. GANs designed for face representation modeling are capable of building a robust individual face model from a single still image [16], [17]. In a trivial case, data augmentation can be performed by modest image deformations which do not destroy the initial structure of a biometric modality, or by identity-aware synthesis of new images.

3) CANCELABLE BIOMETRICS

Cancelable biometrics addresses the usage of synthetic samples in place of real biometrics in stored templates. Such a replacement grants a biometric authentication system flexibility towards the case of a compromised reference data storage. If an individual synthetic template has been stolen, it can be easily replaced by another synthetic template. This use case has been first mentioned in [1] and elaborated in [18] on example of fingerprints and in [19] on example of faces. Although this use case can be seen as an implication of the first use case (replacement of a real subject by a virtual one) with the same requirements on the synthetic samples, we single it out because cancelable biometrics targets security rather than privacy concerns. Note that, similar to the first use case, it is essential here that synthetic samples obfuscate individual characteristics i.e. do not match real reference samples.

4) INVERSE BIOMETRICS

Inverse biometrics addresses reconstruction of biometric samples from biometric templates. Inverse biometrics itself is not directly a use case of synthetic biometric samples but rather a tool for creating those and a proxy to performing presentation attacks on biometric systems. It is also an important tool for identity-aware synthesis of virtual subjects.

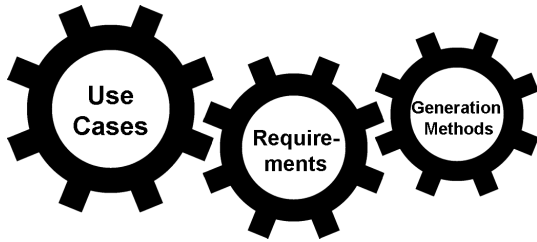


FIGURE 2. Different use cases imply different requirements on synthetic samples and the requirements in turn imply different generation methods.

Note that a strong generative model applied for inversion of biometric templates is a powerful instrument in the hands of a perpetrator. The capability of reconstructing biometric images from “irreversible” templates may lead to identity theft in the case of a compromised biometric reference dataset. Moreover, modeling of multi-identity biometric images such as face morphing [20] by e.g. blending GAN embeddings and a subsequent image reconstruction [21] or direct image blending [22] is a serious threat to identity verification systems. Even simple face swapping in e.g. deep-fakes [23] can be used as a means of mass consciousness manipulation.

C. GENERAL REQUIREMENTS ON SYNTHETIC SAMPLES

It is clear that different use cases imply different requirements on synthetic samples (see Fig. 2 for visualization). For the first use case the original identities must be disguised and several impressions of the same biometric entity are required. For the second use case the variability of samples must be increased by preserving the identity. The main objective in the third use case is concealing the original identity and stable generation of biometric substitutes. Nevertheless, general requirements on synthetic biometric samples can be formulated to assure the utility of samples in general and for a special use case.

Inappropriate synthetic samples, no matter for which purpose they are used, lead to non-meaningful results. To the best of our knowledge, the first attempt to formulate requirements on biometric generative models is done in [5]. The authors propose three criteria: flexibility, parsimony and consistency. Flexibility means that the generative model has enough parameters to model data under study. Parsimony means that the generative model is as simple as the data allows but not simpler. Consistency means that the distribution of synthetic samples is sufficiently close to that of real samples. Out of the three proposed criteria, the first two are rather the guidelines for model design, while the consistency can be practically validated in synthesized images. The proposed metric is goodness-of-fit criteria measured in a Chi-square or Kolmogorov-Smirnov test.

However, a commonly agreed methodology on how to practically assess the quality and utility of synthetic biometric samples still does not exist. In our previous study [24], we propose seven requirements on synthetic biometric

samples and/or generative models on example of synthetic fingerprint images:

- (R) Synthetic images should appear realistic which encompasses two aspects: real and synthetic samples cannot be told apart, neither by the naked eye nor by analyzing image statistics.
- (I) Synthetic images should be of sufficiently high resolution.
- (A) Synthetic images should preserve privacy, meaning that the virtual individuals cannot be linked to real individuals.
- (D) Synthetic images assembled into a dataset should be diverse enough to describe the broad variety of representations of a biometric modality caused by physical interactions with an environment.
- (C) The generative model should be capable of controllable generation of samples meaning that not only non-mated but also mated samples can be produced.
- (B) Synthetic images should reflect basic characteristics of the ground truth images.
- (E) In a synthetic dataset, the distributions of subject attributes such as gender, age or ethnicity should be controlled to avoid biases.

Although the requirements are formulated focusing on fingerprint images, we believe that they can be simply generalized for images of other biometric modalities and moreover for any kind of synthetic images generated by a neural generative model. We also believe that linking the proposed requirements to metrics specific for images under consideration can grow into an established evaluation methodology for synthetic data.

An evaluation methodology for synthetic data looking at realistic recognition accuracy values and the similarity of comparison score distributions to real data for the case of fingerprints is suggested in [25] and for the case of person re-identification data in [26].

Medical studies such as [27] have an alternative perspective on utility evaluation of synthetic data.

D. GENERATION METHODS

In general, there are three approaches for obtaining synthetic biometric samples: *adaptation*, *synthesis* and *reconstruction from a biometric template*.

In the case of adaptation, an existing biometric sample is modified to mimic certain acquisition conditions such as sensor characteristics, environmental variability, or physiology-based variations. For fingerprints, there is an adaptation tool called StirTrace [28]. Adaptation is a commonly used technique for data augmentation. An exhaustive review of face data augmentation approaches is presented in [15].

Synthesis aims at creating artificial samples from scratch according to a predefined sensor, physiological and environmental conditions making use of a pre-trained conditional generative model e.g. Glow [29] for faces or SFinGe [30] for fingerprints.

TABLE 1. Studies on model-based synthesis/reconstruction of biometric images.

Study	Application	Approach
Fingerprint synthesis		
Sherlock and Monro '03 [31]	Orientation image from cores and deltas	Zero-pole model for fingerprint topology
Cappelli '09 [30]	Realistic fingerprint synthesis based on [31]	Statistical ridge pattern model
Ram et al. '10 [32]	Estimation of ridge orientation	Modeling ridges by Legendre polynomials
Johnson et al. '13 [33]	Texture mapping from real to master fingerprint	Modeling of ridge texture & valley noise features
Drahansky '19 [34]	Simulation of sensor, environment, user and finger conditions, skin diseases	Modeling of synthesis states based on Petri net
Priesnitz et al. '22 [35]	Synthesis of contactless fingerprints	Simulation of capturing process, subject characteristics, and environmental influences
Fingerprint reconstruction from minutiae		
Cappelli et al. '07 [36]	Fingerprint reconstruction	Orientation from zero-pole model
Ross et al. '07 [37]	Partial fingerprints reconstruction	Orientation from minutiae triplets
Feng and Jain '11 [38]	Fingerprint reconstruction	Orientation from AM-FM model
Li and Kot '12 [39]	Fingerprint reconstruction	Orientation from AM-FM model
Cao and Jain '15 [40]	Idealistic ridge patterns (no noise)	Patch dictionaries for orient. & continuous phase
Face synthesis		
Sirovich and Kirby '87 [41]	Face as a linear combination of eigenpictures Applied to face recognition in [42] - Eigenfaces Software tools: EVOFIT, E-Fit V	Principal component analysis (PCA) for faces
Cootes '01 [43]	Active Appearance Model	PCA on texture + PCA on facial landmarks
Blanz and Vetter '99 [44]	Morphable 3D Face Model Software tool: FaceGen	PCA on texture + PCA on 3D face model
Kähler et al. '02 [45]	Deformable head model with anatomical structure	Modeling musculature and skeletal structure
Iris synthesis		
Lefohn et al. '03 [46]	Ocularist's model	Stacking layers: stroma, collarette, sphincter, etc.
Cui et al. '04 [47]	Iris synthesis	PCA + super-resolution
Makthal and Ross '05 [48]	Iris texture synthesis	Markov random field (MRF) model
Shah and Ross '06 [49]	Realistic iris synthesis	MRF model for background texture + embedding radial and concentric furrows, collarette and crypts
Zuo et al. '07 [50]	Compilation of large-scale synthetic iris datasets	Anatomy based model
Iris reconstruction from binary IrisCode		
Venugopalan et al. '11 [51]	Iris reconstruction	Deterministic approach based on Gabor filtering
Galbally et al. '13 [52]	Iris reconstruction	Probabilistic approach based on genetic algorithms
Synthesis of vascular samples		
Crisan et al. '08 [53]	Realistic hand vein samples	Vein nodes are connected by randomly driven vessel growth and imaging variation modeling
Hillerström et al. '14 [54]	Realistic finger vein samples	Vein nodes are connected by randomly driven vessel growth and imaging variation modeling
Fiorini et al. '14 [55]	Realistic retina (fundus) samples	Patch-based approach to create texture
Bonaldi et al. '16 [56]		Active shapes and Kalman filtering for vessels
Das et al. '17 [57]	Realistic sclera patterns	Non-parametric texture synthesis

Reconstruction from a biometric template aims at interpreting the essential identity-specific information from a biometric template and generating artificial samples by parameterizing a previously learned generic model towards the identity.

III. TRADITIONAL MODELING

The most prominent and recent studies on model-based synthesis of biometric images and reconstruction from biometric templates are listed in Table 1.

A. FINGERPRINTS

Historically, first conclusive modeling efforts were made for fingerprints back in the first half of the 20th century [58].

A simple yet practical model for ridge pattern orientation based on singular points (cores and deltas) is proposed in [31] and applied in SFinGe [30] for algorithmic synthesis of realistic fingerprints. In [32], it is shown how ridge orientations can be modeled by fitting Legendre polynomials. Modeling of fingerprints in [30] goes far beyond modeling of master fingerprints i.e. iterative application of Gabor filters, tailored to local orientation and frequency. A crucial step is mimicking distortions to simulate traction and torsion forces applied to a finger during its placement on the sensor as well as rendering texture and adding noise. However, the applied texture and noise models are not well justified leading to inability of reproducing texture that is statistically representative for real fingerprints. In [33], texture and noise are derived from

real fingerprint images enabling realistic textures in synthetic images. Simulation of a fingerprint generation process by Petri net is proposed in [34] with the focus on mimicking skin diseases. The net depicts synthesis states and possible transitions between them. First, master fingerprint is created, then environmental influences are added, then user and finger conditions, then fingerprint damages, and finally sensor characteristics. The study in [35] focuses on modeling contactless fingerprints. Synthetic samples are designed to reflect properties of capturing process, subject characteristics, and environmental influences.

Inversion of fingerprint templates was first addressed by Hill back in 2001 [59]. However, the intensive research started in 2007 with the studies of Cappelli et al. [36] and Ross et al. [37]. In contrast to fingerprint drawing from scratch, the locations of singular points are derived from the given minutiae and further on the orientation map. Ridge frequencies are also adjusted to minutiae. Apart from the zero-pole model [31], ridge patterns can be estimated using the minutiae triplet model [37] or the AM-FM model [38], [39] which makes use of eight neighboring minutiae. Reconstruction of ridges based on patch dictionaries, proposed in [40], allows for idealistic ridge patterns.

B. FACE

Modeling of faces has always been an intensively studied topic of computer vision. Indeed, face recognition is only one domain in a broad variety of applications making use of highly realistic synthetic faces. The first important mathematical concept allowing for modeling face attributes and preserving individual characteristics was proposed by Sirovich and Kirby [41] way back in 1987. Later on, Turk and Pentland [42] called this concept “Eigenfaces” and demonstrated its application to face recognition. Eigenfaces is a domain specific rephrase of eigenvectors known as an integral part of Principal Component Analysis (PCA). The main idea is that a face image is decomposed in a linear combination of Eigenfaces and the coefficients of Eigenfaces define face appearance. Individual characteristics, environmental influences and other appearance variations are modeled by modifying one or several coefficients. The next step was made by Cootes [43] who split face into texture and a graph of facial landmarks and applied PCA to texture and geometry separately. Image rendering is then done by texture warping. This move allows for better modeling of face poses. The breakthrough in face modeling is achieved by Blanz and Vetter [44] with the concept of a morphable 3D face model. Practically, for geometry representation, facial landmarks are replaced by a 3D head shape and PCA is used, as before, for modeling variations in pose, illumination, expression, etc. This model has been actively used in most face modeling studies until the paradigm change in 2014 due to the success of DCNN in classification tasks and introduction of GAN [60]. A head model suitable for deformation in anthropometrically meaningful ways using the underlying muscle and bone structure

is proposed in [45]. The head growth from early childhood to adult can be simulated. In fact, complex 3D models can be perfectly combined with facial landmark localization and texture warping to rotate or frontalize faces in 2D images [61] which in turn greatly improves the performance of face recognition.

We intentionally omit “face reconstruction from a template” because of an absence of commonly agreed convention on “face features” so that the standard face template is a high-quality frontal face image. Hence, face reconstruction makes sense only for deep-face features addressed in Section IV-B.

C. IRIS

Modeling of iris images is a relatively young research field having its roots in the early 2000s. The ophthalmologist's approach to iris synthesis [46] overlays semi-transparent texture layers built from topological and optic models. The layers are designed to mimic stroma, collarette, limbus, pupil and sphincter muscle components. It has been stated that such fake iris patterns printed onto vanity contact lenses pose a threat to iris identification systems. Synthesis of iris images based on PCA and super-resolution is proposed in [47]. In [48], iris is modeled by Markov Random Fields (MRF) applied to stitching of texture primitives representing radial furrows, crypts and limbus. The study in [49] is a further development of [48] in which MRF modeling is applied to background texture, and radial furrows, concentric furrows, collarette and crypt are synthesized as texture patches and embedded into the texture. An anatomy-based method for synthesizing iris images with the focus on compiling a large database of synthetic irises is proposed in [50]. The authors put emphasis on “realism” stating that a synthetic iris should not only look like a real iris but also have statistical characteristics of a real iris.

Iris reconstruction from binary IrisCode relies either on a deterministic approach based on Gabor filtering [51] or on a probabilistic approach based on genetic algorithms [52].

D. VASCULAR BIOMETRICS

A review on several synthesis techniques for hand-related vascular biometric samples provided in [62] is not restricted to palm vein data as suggested by the title. The first work in this field is more recent as compared to other modalities - [53] proposes a method for creating hand vein sample data, employing (randomly positioned) “key points” which are interconnected by a random process and dilated to achieve vessel appearance. Texture is generated using a contrast variation model. As introduced in [54] similar methodology for finger veins, based on setting slightly disturbed vein nodes, which are interconnected by randomly steered vessel growth and NIR variability is also modeled to result in realistic appearance.

Eye-related traits have been considered in the context of creating synthetic data - in particular retina recognition,

as fundus imagery is an important diagnosis element in ophthalmology as well. In [55], a patch-based approach is used to create retinal backgrounds and foveae, model-based texture synthesis techniques have also been employed for the generation of realistic optic discs and vessel networks. The approach has been refined in [56], where the Active Shape model and Kalman filtering combined with an extension of the Multiresolution Hermite vascular cross-section model facilitates realistic vessel network modeling.

Finally, also sclera data has been subject to synthesis attempts: [57] proposes to apply an earlier published non-parametric texture synthesis technique which relies on systematically selecting random patches from “primitive images” (i.e. seed or source images) to produce similar texture, using a part of the UBIRISv1 dataset as primitive images.

IV. DATA-DRIVEN SYNTHESIS

Neural generative models have been applied in biometrics for many purposes. The most prominent are biometric image completion [95], image quality enhancement [96], image style transfer [97], generation of random realistic biometric samples [78] and conditional generation for identity-aware image reconstruction or retaining/obfuscation of soft biometric attributes [98]. The majority of neural generative models are based on GAN. Since face as biometric modality has always been a pioneering one in computer vision research, the majority of feasibility studies on GAN-based synthesis deal with face images, but the developed GAN architectures can be applied to other biometric images as well. GAN models allow for creating realistic patterns and transferring domain characteristics from an existing dataset. These two aspects are very challenging for model-based generation and are natively supported in data-driven generation. The most prominent and recent studies on data-driven synthesis/reconstruction of biometric images are listed in Table 2.

The key feature of data-driven synthesis is that no knowledge about image semantics is required. However, the prerequisite for training of a generative model is a large set of diverse real biometric samples. Here, one can see a logical mistake. If real biometric samples are already presented, why do we use them for training and not for the initial task e.g. evaluation of a biometric system? The first reason is privacy protection. Indeed real biometric data can hardly be used in open research. Second, a synthetic dataset can be made of arbitrary size enabling training of DCNN from scratch.

A. FINGERPRINTS

In [63], a Wasserstein GAN is applied for generating master fingerprints that match multiple original fingerprints. The Finger-GAN proposed in [64] is a connectivity imposed fingerprint generation GAN that is introduced and applied to simulate images from two fingerprint datasets: FVC2006 and PolyU. The SynFi approach in [65] enables high-resolution realistic fingerprints generation by combining GAN with a super-resolution network. In [66], a combination of

a convolutional autoencoder and an improved Wasserstein GAN is used for synthesizing 512×512 pixel rolled and plain fingerprints. In [67], the feasibility of three established NVIDIA GAN architectures: progressive growing GAN, StyleGAN and StyleGAN2 is validated for realistic fingerprint pattern generation. The Clarkson Fingerprint Generator (CFG) [68] is trained using StyleGAN and proprietary dataset captured with a Crossmatch Guardian scanner. In [69], a CycleGAN is applied for texture transfer from real fingerprints to Anguli generated fingerprints with added sweat pores. Then a super-resolution network increases the image size. The PrintsGAN is proposed in [70] for fingerprint synthesis from DeepPrint representation with control over a fingerprint identity.

In fact, there are two ways of fingerprint synthesis with a pre-defined identity. The first is in combining model-based reconstruction from minutiae and further application of style-transfer (e.g. CycleGAN) to add realism and transfer domain characteristics, as in [69]). The second way is to use conditional GAN with identity as a condition. If identity is given by a deep representation, as in [70], the identity-aware synthesis is straightforward. If identity is not explicitly presented or given by minutiae, an additional encoding step is required.

Fingerprint reconstruction from a minutiae template is addressed in [71] using a conditional GAN (pix2pix) originally proposed in [72]. In [73], a convolutional minutiae-to-vector encoder is used in combination with StyleGAN2 for identity-preserving, attributes-aware fingerprint reconstruction from minutiae. In [12], fingerprint reconstructions from deep network embedding and from minutiae are compared qualitatively and quantitatively regarding the inversion attack performance.

B. FACE

Modeling of face as a vector in a latent space of a generator network is a current research trend. Making moves in the latent space of the generator both person identity and facial attributes can be modified on demand. Sophisticated GAN models allow for face learning from a single still images and therefore unconstrained face recognition [17]. The most successful face modeling networks are Glow [29] from OpenAI and StarGAN [75] from CLOVA AI Research. However, the best visual quality have face images randomly produced by NVIDIA GAN models: progressive growing GAN [77], StyleGAN [78], StyleGAN2 [79], StyleGAN2-ada [80] and StyleGAN3 [81]. The majority of recent face modeling approaches build on NVIDIA architectures trying to extend it by a conditional generation mechanism. This can be done either by statistical analysis of the latent space [99], [100], or by semi-supervised disentanglement learning [101], or by detecting “StyleSpace channels” for attribute control [102].

The problem of reconstructing face images from deep embeddings goes far beyond the field of inverse biometrics. The question is whether an image encoded by one deep

TABLE 2. Studies on data-driven synthesis/reconstruction of biometric images.

Study	Application	Approach
Fingerprint synthesis		
Bontrager et al. '18 [63]	Dictionary attacks via latent variable evolution	DeepMaster-Prints
Minaee et al. '18 [64]	Finger-GAN, realistic fingerprint synthesis	Connectivity Imposed GAN
Riazi et al. '20 [65]	SynFi, realistic fingerprint synthesis	GAN + super-resolution DCNN
Mistry et al. '20 [66]	Forensic fingerprint generation, NIST SD4	Conv. autoencoder + WassersteinGAN
Seidlitz et al. '21 [67]	Random fingerprint patch generation, NIST SD27	PGGAN, StyleGAN, StyleGAN2
Bahmani et al. '21 [68]	Random fingerprint generation, In-house CrossMatch	StyleGAN (Clarkson Finger Generator)
Wyzykowski et al. '22 [69]	Controlled fingerprint generation	Anguli + CycleGAN + super-resolution CNN
Engelsma et al. '22 [70]	Synthesis from DeepPrint representation with identity control	PrintsGAN
Fingerprint reconstruction from minutiae		
Kim et al. '18 [71]	Reconstruction from minutiae map 256x256x1	pix2pix from Isola et al. [72]
Bouzaglo and Keller '22 [73]	Reconstruction from minutiae map 512x512x3 & attributes adaption while preserving identity	Minutiae-to-vector encoder + StyleGAN2 Adv. Loss + Reconstruction Loss
Wijewardena et al. '22 [12]	Reconstruction from minutiae map 512x512x6 & deep templates	U-Net generator, Adv. Loss + Rec. Loss + Minutiae Loss + DeepPrint loss
Makrushin et al. '23 [74]	Reconstruction from minutiae map 512x512x1	adapted pix2pix + advanced minutiae encoding
Face synthesis		
Kingma et al. '18 [29]	Modeling face incl. identity & attributes	Glow
Choi et al. [75], [76]	Modeling face incl. identity & attributes	StarGAN, StarGAN v2
Karras et al. [77]–[81]	Random high-resolution face image generation	PGGAN, StyleGAN, StyleGAN2(-ada), StyleGAN3
Face reconstruction from deep face templates		
Zhmoginov et al. '16 [82]	Rec. from the final conv. layer of FaceNet	Iterative and feed-forward methods
Cole et al. '17 [83]	Rec. from the 1024-D avgpool layer of the FaceNet NN2 architecture	CNN-based autoencoder with differentiable image warping & spline interpolation
Mai et al. '19 [84]	Rec. from the final conv. layer of FaceNet	NbNet, Neighbourly de-conv. neural network
Iris synthesis		
Minaee et al. '18 [85]	Iris-GAN, iris synthesis	Vanilla DC-GAN
Lee et al. '19 [86]	Data augmentation for accurate iris recognition	Conditional GAN (pix2pix)
Yadav et al. '19 [87]	Synthesis of iris images for PAD	RaSGAN
Yadav and Ross '21 [88]	Iris style transfer for PAD	CIT-GAN
Tomasevic et al. '22 [89]	Bimodal (Vis/IR) ocular images synthesis + annotation	BiOcularGAN
Iris reconstruction from templates		
Ahmad and Fuller '20 [90]	RESIST, Reconstructing irises from 3 types of templates (Gabor, DenseNet, DenseNet norm.)	U-Net (pix2pix) extended by L1 perception loss and SSIM loss
Synthesis of vascular samples		
Zhang et al. '19 [91]	Data augmentation for enhanced training	FCGAN for CNN finger vein recognition
Yang et al. '20 [92]	Realistic finger vein samples	Model-based pattern, GAN-based texture
Salazar et al. '21 [93]	Realistic finger vein samples	Model-based pattern, GAN-based texture
Finger/hand veins reconstruction		
Kauba et al. '21 [94]	Reconstruction from binary templates	pix2pix, convolutional neural network

convolutional model can be successfully reconstructed by another deep convolutional model trained with previously unseen samples. For the case of facial images, the positive examples of such reconstruction are demonstrated in [82], [83], and [84]. All approaches mentioned reconstruct face images from FaceNet embeddings.

C. IRIS

There are only few studies on GAN-based iris synthesis. Iris-GAN is proposed in [85] to generate irises from random variables using a deep convolutional GAN. The synthetic irises lack realism and mostly suffer from unrealistic patterns around the iris boundary. In [86], a conditional GAN

(pix2pix) is applied to iris synthesis aiming at data augmentation to improve accuracy of iris recognition. In [87], the Relativistic Average Standard Generative Adversarial Network (RaSGAN) is applied to synthesize high-quality images to mimic iris presentation attacks and hence to support presentation attacks detection (PAD) against unseen attacks. Aiming at compensating for under-represented iris presentation attacks in the training set, the Cyclic Image Translation Generative Adversarial Network (CIT-GAN) is proposed in [88] to enable domain style transfer and improve PAD accuracy.

GAN-based iris reconstruction from three types of iris templates is introduced in [90] referred to as RESIST. Apart from a traditional Gabor filter template, two deep templates

TABLE 3. Public synthetic datasets / generation software.

Name, creator, tool	URL	# samples
Fingerprint		
Anguli: Synthetic Fingerprint Generator [104], Indian Institute of Science	https://dsl.cds.iisc.ac.in/projects/Anguli/	software only
FVC2000 DB4, University of Bologna, SFinGe [30]	http://bias.csr.unibo.it/fvc2000/databases.asp	110 fp, 8 imp.
FVC2002 DB4, University of Bologna, SFinGe v2.51	http://bias.csr.unibo.it/fvc2002/databases.asp	110 fp, 8 imp.
FVC2004 DB4, University of Bologna, SFinGe v3.0	http://bias.csr.unibo.it/fvc2004/databases.asp	110 fp, 8 imp.
FVC2006 DB4, University of Bologna, SFinGe v3.0	http://atvs.ii.uam.es/atvs/fvc2006.html	150 fp, 12 imp.
Sokoto Coventry Fingerprint Dataset (SOCOFing) [105]	https://www.kaggle.com/datasets/ruizgara/socofing	6k fp, 600 Africans
Clarkson Fingerprint Generator (CFG) [68]	https://github.com/keivanB/Clarkson_Finger_Gen	50K random fp
L3 Synthetic Fingerprint dataset (L3-SF) [106]	https://andrewzyy.github.io/L3-SF/	148 identities, 10 imp., 5 subsets
Multiresolution Synthetic Fingerprint Dataset (MR-SF) [69]	https://drive.google.com/file/d/1MsSCJZreekT9NVlj04Rsw5jry1LbOFid/view?usp=sharing	148 identities, 10 imp., 5 subsets, L2 and L3
Face		
Face Synthetics, Microsoft [107]	https://github.com/microsoft/FaceSynthetics	100,000 faces + 70 landmark annot.
iFakeFaceDB, UAM [108]	https://github.com/socialabubi/iFakeFaceDB	87000 faces
Real and Fake Face Detection, Yonsei Univ.	https://www.kaggle.com/datasets/ciplab/real-and-fake-face-detection	960 faces, manual creation
Syn-Multi-PIE [100], IDIAP	https://gitlab.idiap.ch/bob/bob.paper.ijcb2021_synthetic_dataset	software only
Notre Dame Synthetic Face Dataset [109]	https://cvrl.nd.edu/projects/data/	12338 faces, 21 views
Parametric-face-image-generator, Basel Univ. [110]	https://github.com/unibas-gravis/parametric-face-image-generator	software only
Univ. Chicago [111]	https://tlab.uchicago.edu/databases/	17000
Iris		
Textured-based synthesis [49], West Virginia University	https://biic.wvu.edu/data-sets/synthetic-iris-dataset	1000 irises, 7 imp.
Model-based synthesis [50], Clarkson University	https://citer.clarkson.edu/research-resources/biometric-dataset-collections-2/synthetic-iris-model-based/	10000 irises, 16 imp.
BiOcularGAN [89], University of Ljubljana	https://github.com/dariant/BiOcularGAN	software only
Vascular modalities		
Natural-based Synthetic Palm Vein Database, NS-PVDB [93], Universidad Católica del Maule	http://www.litrp.cl/repository	10000 palm vein ROIs, 6 imp.
Synthetic Finger-Vein Image Database Generator [54], Hong Kong Polytechnic University	https://web.comp.polyu.edu.hk/csajaykr/fvgen.htm	software only

are addressed: embedding of DenseNet with and without normalization. The RESIST architecture is a modified U-Net (pix2pix) extended by L1 perception loss and SSIM loss.

For both iris recognition and periocular recognition, GAN-based synthetic eye images are created using a segmentation mask as input and being gaze-preserving as described in [103].

D. VASCULAR TRAITS

In [91], GAN-based data augmentation is proposed for enhanced training in CNN-based finger vein recognition, where the suggested FCGAN approach outperforms other GAN architectures. Here, classically augmented finger vein sample images are used as training data. Contrasting to this approach, [92] generates the vascular network using a classical approach, while generating the texture for the image using a GAN approach. Similarly, a nature-inspired algorithm is used in [93] to form the vessel pattern, and the StyleGAN2 for texture generation.

A good overview on GAN-based synthetic retina (fundus) images generation methods (including an analysis of weaknesses) is contained in [112]. Among other techniques, [113] uses a multiple-channels-multiple-landmarks (MCML) approach in which vessel structure is combined with optical disc and optical cup images, which are then fed into pix2pix or CycleGAN architectures to generate realistic texture. A multitude of different CNN architectures is used and compared overall. A related approach is taken in [112], where a VAE is employed to create vessel trees, which are then used by a GAN architecture to create full fundus images. After training, this results in an end-to-end system. In [114], a pix2pixHD GAN is used to create synthetic retinal images, where ophthalmologists were not successful in discriminating real from artificial data.

V. DATASETS AND CHALLENGES WITH SYNTHETIC DATA

A. DATASETS

Due to recent data protection regulations, many biometric datasets were removed from the public domain. For instance,

NIST has taken down forensic fingerprint datasets NIST SD4, NIST SD14 and NIST SD27, and Microsoft has taken down MS-Celeb-1M database of 10 million faces scraped from the internet. This tendency explains the growing interest in synthetic biometric datasets that can replace the real ones. There is a number of studies on compilation of large synthetic datasets for almost all biometric modalities. In [66], [67], and [68] there are approaches for fingerprints, in [99], [100], and [115] for faces, and in [50] for irises. However, the amount of publicly available datasets is quite limited. Table 3 summarizes public synthetic datasets.

B. CHALLENGES

After it has been established that synthetic faces can be successfully applied in face recognition and even completely substitute for real faces [100], [107], [116], [117], the next step is organizing large-scale biometric campaigns in which recognition models are trained solely from synthetic samples addressing the whole variety of biometric modalities. To the best of our knowledge, the only such competition related to biometrics and focusing on face morphing attack detection took place within the International Joint Conference on Biometrics 2022.¹

VI. CONCLUSION

The abundance of recent studies on generation of synthetic biometric samples suggests that interest in experimenting with synthetic data is currently very high and is even growing. There is a belief that by 2024, 60% of the data used for the development of AI-based solutions will be synthetically generated,² and by 2030, synthetic data will completely overshadow real data in AI models.³ Although there is still controversy on whether the synthetic datasets can replace real biometric datasets, the number of published neural generative approaches for biometric images increases. While face biometric still enjoys a sufficient number of publicly available datasets, many fingerprint datasets have been removed from public domain due to privacy concerns. Synthetic biometric samples have a versatile use spreading from augmentation of training and test data to enabling privacy-friendly evaluation of biometric algorithms and sharing data for open research. Hence, many researchers seek for publicly available synthetic biometric datasets. Synthetic biometrics also have a close relation to several security-related domains such as inverse biometrics, cancelable biometrics and privacy-enhancing technologies implicating an improvement of security if synthetic samples are involved. Utility evaluation of synthetic biometric samples is the major challenge now. We believe that linking the requirements proposed

in this paper to metrics specific for images under consideration can grow into an established evaluation methodology for any kind of synthetic images generated by a neural generative model.

REFERENCES

- [1] S. N. Yanushkevich, "Synthetic biometrics: A survey," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jul. 2006, pp. 676–683.
- [2] N. Orlans, D. Buettner, and J. Marques, "A survey of synthetic biometrics: Capabilities and benefits," in *Proc. Int. Conf. Artif. Intell.*, vol. 1, Jan. 2004, pp. 499–505.
- [3] *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation)*, European Commission, Brussels, Belgium, 2019.
- [4] D. S. Quintana, "A synthetic dataset primer for the biobehavioural sciences to promote reproducibility and hypothesis generation," *eLife*, vol. 9, Mar. 2020, Art. no. e53275.
- [5] Y. Ma, M. Schuckers, and B. Cukic, "Guidelines for appropriate use of simulated data for bio-authentication research," in *Proc. 4th IEEE Workshop Autom. Identificat. Adv. Technol.*, Oct. 2005, pp. 251–256.
- [6] D. J. Buettner, "Biometric sample synthesis," in *Encyclopedia of Biometrics*, S. Z. Li and A. K. Jain, Eds., 2nd ed. Cham, Switzerland: Springer, 2015, pp. 211–217.
- [7] M. Gomez-Barrero and J. Galbally, "Reversing the irreversible: A survey on inverse biometrics," *Comput. Secur.*, vol. 90, Mar. 2020, Art. no. 101700.
- [8] P. Grother and E. Tabassi, "Face image data interchange formats, standardization," in *Encyclopedia of Biometrics*, S. Z. Li and A. Jain, Eds. Boston, MA, USA: Springer, 2009, pp. 314–321.
- [9] J. Daugman, "Information theory and the iriscodes," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 2, pp. 400–409, Feb. 2016.
- [10] J. Galbally, M. Savvides, S. Venugopalan, and A. A. Ross, "Iris image reconstruction from binary templates," in *Handbook of Iris Recognition*, K. W. Bowyer and M. J. Burge, Eds. London, U.K.: Springer, 2016, pp. 469–496.
- [11] *Information Technology—Biometric Data Interchange Formats—Part 2: Finger Minutiae Data*, Standard ISO/IEC 19794-2:2011, 2011.
- [12] K. P. Wijewardena, S. A. Grosz, K. Cao, and A. K. Jain, "Fingerprint template invertibility: Minutiae vs. deep templates," 2022, *arXiv:2205.03809*.
- [13] S. James, C. Harbron, J. Branson, and M. Sundler, "Synthetic data use: Exploring use cases to optimise data utility," *Discover Artif. Intell.*, vol. 1, no. 1, pp. 1–5, Dec. 2021.
- [14] P. Irtem, E. Irtem, and N. Erdogmus, "Impact of variations in synthetic training data on fingerprint classification," in *Proc. Int. Conf. Biometrics Special Interest Group*, A. Brömme, C. Busch, A. Dantcheva, C. Rathgeb, and A. Uhl, Eds., vol. 296, 2019, pp. 189–196.
- [15] X. Wang, K. Wang, and S. Lian, "A survey on face data augmentation for the training of deep neural networks," *Neural Comput. Appl.*, vol. 32, no. 19, pp. 15503–15531, Oct. 2020.
- [16] B. Gecer, S. Ploumpis, I. Kotsia, and S. Zafeiriou, "Fast-GANFIT: Generative adversarial network for high fidelity 3D face reconstruction," 2021, *arXiv:2105.07474*.
- [17] Y. Ding, Z. Tang, and F. Wang, "Single-sample face recognition based on shared generative adversarial network," *Mathematics*, vol. 10, no. 5, p. 752, Feb. 2022.
- [18] A. Othman and A. Ross, "On mixing fingerprints," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 1, pp. 260–267, Jan. 2013.
- [19] Z. Yuan, Z. You, S. Li, Z. Qian, X. Zhang, and A. Kot, "On generating identifiable virtual faces," in *Proc. 30th ACM Int. Conf. Multimedia*, Oct. 2022, pp. 1465–1473.
- [20] M. Ferrara, A. Franco, and D. Maltoni, "The magic passport," in *Proc. IEEE Int. Joint Conf. Biometrics*, Sep. 2014, pp. 1–7.
- [21] A. Makrushin, M. Trebeljahr, S. Seidlitz, and J. Dittmann, "On feasibility of GAN-based fingerprint morphing," in *Proc. IEEE 23rd Int. Workshop Multimedia Signal Process. (MMSP)*, Oct. 2021, pp. 1–6.
- [22] A. Makrushin, T. Neubert, and J. Dittmann, "Automatic generation and detection of visually flawless facial morphs," in *Proc. 12th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, 2017, pp. 39–50.
- [23] Y. Mirsky and W. Lee, "The creation and detection of deepfakes: A survey," *ACM Comput. Surv.*, vol. 54, no. 1, p. 7, Jan. 2021.

¹<https://sites.google.com/view/ijcb-syn-mad-2022>, accessed 26.01.2023

²Andrew White, Gartner Blog, https://blogs.gartner.com/andrew_white/2021/07/24/by-2024-60-of-the-data-used-for-the-development-of-ai-and-analytics-projects-will-be-synthetically-generated, accessed 26.01.2023

³Alexander Linden, Gartner Q&A, <https://www.gartner.com/en/newsroom/press-releases/2022-06-22-is-synthetic-data-the-future-of-ai>, accessed 26.01.2023

- [24] A. Makrushin, C. Kauba, S. Kirchgasser, S. Seidlitz, C. Kraetzer, A. Uhl, and J. Dittmann, "General requirements on synthetic fingerprint images for biometric authentication and forensic investigations," in *Proc. ACM Workshop Inf. Hiding Multimedia Secur.*, Jun. 2021, pp. 93–104.
- [25] S. Kirchgasser, C. Kauba, and A. Uhl, "Assessment of synthetically generated mated samples from single fingerprint samples instances," in *Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Montpellier, France, Dec. 2021, pp. 1–6.
- [26] C. Kang, "Is synthetic dataset reliable for benchmarking generalizable person re-identification?" in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Oct. 2022, pp. 1–8.
- [27] Z. Azizi, C. Zheng, L. Mosquera, L. Pilote, and K. E. Emam, "Can synthetic data be a proxy for real clinical trial data? A validation study," *BMJ Open*, vol. 11, no. 4, Apr. 2021, Art. no. e043497.
- [28] M. Hildebrandt and J. Dittmann, "StirTraceV2.0: Enhanced benchmarking and tuning of printed fingerprint detection," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 4, pp. 833–848, Apr. 2015.
- [29] D. P. Kingma and P. Dhariwal, "Glow: Generative flow with invertible 1×1 convolutions," in *Advances in neural information processing systems*, vol. 31, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2018.
- [30] R. Cappelli, "SFinGe," in *Encyclopedia of Biometrics*, S. Z. Li and A. Jain, Eds. Boston, MA, USA: Springer, 2009, pp. 1169–1176.
- [31] B. G. Sherlock and D. M. Monro, "A model for interpreting fingerprint topology," *Pattern Recognit.*, vol. 26, no. 7, pp. 1047–1055, 1993.
- [32] S. Ram, H. Bischof, and J. Birchbauer, "Modelling fingerprint ridge orientation using Legendre polynomials," *Pattern Recognit.*, vol. 43, no. 1, pp. 342–357, 2010.
- [33] P. Johnson, F. Hua, and S. Schuckers, "Texture modeling for synthetic fingerprint generation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2013, pp. 154–159.
- [34] M. Draňanský and O. Kanich, "Influence of skin diseases on fingerprints," in *Biometrics Under Biomedical Considerations*, A. Nait-Ali, Ed. Singapore: Springer, 2019, pp. 1–39.
- [35] J. Priesnitz, C. Rathgeb, N. Buchmann, and C. Busch, "SynCoLFInGer: Synthetic contactless fingerprint generator," *Pattern Recognit. Lett.*, vol. 157, pp. 127–134, May 2022.
- [36] R. Cappelli, A. Lumini, D. Maio, and D. Maltoni, "Fingerprint image reconstruction from standard templates," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 9, pp. 1489–1503, Sep. 2007.
- [37] A. Ross, J. Shah, and A. K. Jain, "From template to image: Reconstructing fingerprints from minutiae points," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 544–560, Apr. 2007.
- [38] J. Feng and A. K. Jain, "Fingerprint reconstruction: From minutiae to phase," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 2, pp. 209–223, Feb. 2011.
- [39] S. Li and A. C. Kot, "An improved scheme for full fingerprint reconstruction," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 6, pp. 1906–1912, Dec. 2012.
- [40] K. Cao and A. K. Jain, "Learning fingerprint reconstruction: From minutiae to image," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 1, pp. 104–117, Jan. 2015.
- [41] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," *J. Opt. Soc. Amer. A, Opt. Image Sci. Vis.*, vol. 4, no. 3, pp. 519–524, 1987.
- [42] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognit. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.
- [43] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, Jun. 2001.
- [44] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3D faces," in *Proc. 26th Annu. Conf. Comput. Graph. Interact. Techn.*, 1999, pp. 187–194.
- [45] K. Kähler, J. Haber, H. Yamauchi, and H.-P. Seidel, "Head shop: Generating animated head models with anatomical structure," in *Proc. ACM SIGGRAPH/Eurograph. Symp. Comput. Animation*, J. K. Hodgins, M. van de Panne, M. F. Cohen, and N. S. Pollard, Eds., Jul. 2002, pp. 55–63.
- [46] A. Lefohn, B. Budge, P. Shirley, R. Caruso, and E. Reinhard, "An ocularist's approach to human iris synthesis," *IEEE Comput. Graph. Appl.*, vol. 23, no. 6, pp. 70–75, Nov. 2003.
- [47] J. Cui, Y. Wang, J. Huang, T. Tan, and Z. Sun, "An iris image synthesis method based on PCA and super-resolution," in *Proc. 17th Int. Conf. Pattern Recognit.*, 2004, pp. 471–474.
- [48] S. Makthal and A. Ross, "Synthesis of iris images using Markov random fields," in *Proc. 13th Eur. Signal Process. Conf.*, 2005, pp. 1–4.
- [49] S. Shah and A. Ross, "Generating synthetic irises by feature agglomeration," in *Proc. Int. Conf. Image Process.*, Oct. 2006, pp. 317–320.
- [50] J. Zuo, N. A. Schmid, and X. Chen, "On generation and analysis of synthetic iris images," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 1, pp. 77–90, Mar. 2007.
- [51] S. Venugopalan and M. Savvides, "How to generate spoofed irises from an iris code template," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 2, pp. 385–395, Jun. 2011.
- [52] J. Galbally, A. Ross, M. Gomez-Barrero, J. Fierrez, and J. Ortega-Garcia, "Iris image reconstruction from binary templates: An efficient probabilistic approach based on genetic algorithms," *Comput. Vis. Image Understand.*, vol. 117, no. 10, pp. 1512–1525, Oct. 2013.
- [53] S. Crisan, I. Tärnovan, and T. Crisan, "A hand vein structure simulation platform for algorithm testing and biometric identification," in *Proc. 16th 6th IMEKO TC4 Symp.*, 2008, pp. 906–910.
- [54] F. Hillerström, A. Kumar, and R. Veldhuis, "Generating and analyzing synthetic finger vein images," in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, 2014, pp. 1–9.
- [55] S. Fiorini, M. D. Biasi, L. Ballerini, E. Trucco, and A. Ruggeri, "Automatic generation of synthetic retinal fundus images," in *Proc. Smart Tools Apps Graph.-Eurograph. Italian Chapter Conf.*, A. Giachetti, Ed., 2014, pp. 1–5.
- [56] L. Bonaldi, E. Menti, L. Ballerini, A. Ruggeri, and E. Trucco, "Automatic generation of synthetic retinal fundus images: Vascular network," *Proc. Comput. Sci.*, vol. 90, pp. 54–60, Jan. 2016.
- [57] A. Das, P. Mondal, U. Pal, M. Blumenstein, and M. A. Ferrer, "Sclera vessel pattern synthesis based on a non-parametric texture synthesis technique," in *Proc. Int. Conf. Comput. Vis. Image Process.*, B. Raman, S. Kumar, P. P. Roy, and D. Sen, Eds. Singapore: Springer, 2017, pp. 241–250.
- [58] M. Kücken, "Models for fingerprint pattern formation," *Forensic Sci. Int.*, vol. 171, nos. 2–3, pp. 85–96, Sep. 2007.
- [59] C. J. Hill, "Risk of masquerade arising from the storage of biometrics," B.S. thesis, Dept. Comput. Sci., Austral. Nat. Univ., Canberra, ACT, Australia, 2001. [Online]. Available: https://cyberatlantis.com/library/crypto_n_security/biometrics/fingerprints/risk_masquerade_storage_biometrics.pdf
- [60] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst. (NIPS)*, vol. 2. Cambridge, MA, USA: MIT Press, 2014, pp. 2672–2680.
- [61] T. Hassner, S. Harel, E. Paz, and R. Enbar, "Effective face frontalization in unconstrained images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 4295–4304.
- [62] E. H. Salazar-Jurado, R. Hernández-García, K. Vilches-Ponce, R. J. Barrientos, M. Mora, and G. Jaswal, "Towards the generation of synthetic images of palm vein patterns: A review," 2022, *arXiv:1710.10196*.
- [63] P. Bontrager, A. Roy, J. Togelius, N. Memon, and A. Ross, "DeepMasterPrints: Generating MasterPrints for dictionary attacks via latent variable evolution," in *Proc. IEEE 9th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Oct. 2018, pp. 1–9.
- [64] S. Minaee and A. Abdolrashidi, "Finger-GAN: Generating realistic fingerprint images using connectivity imposed GAN," 2018, *arXiv:1812.10482*.
- [65] M. S. Riazi, S. M. Chavoshian, and F. Koushanfar, "SynFi: Automatic synthetic fingerprint generation," *Cryptol. ePrint Arch.*, Paper 2020/217, 2020. [Online]. Available: <https://eprint.iacr.org/2020/217>
- [66] V. Mistry, J. J. Engelsma, and A. K. Jain, "Fingerprint synthesis: Search with 100 million prints," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Sep. 2020, pp. 1–10.
- [67] S. Seidlitz, K. Jürgens, A. Makrushin, C. Kraetzer, and J. Dittmann, "Generation of privacy-friendly datasets of latent fingerprint images using generative adversarial networks," in *Proc. 16th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, vol. 4, G. M. Fariella, Ed., 2021, pp. 345–352.

- [68] K. Bahmani, R. Plesh, P. Johnson, S. Schuckers, and T. Swyka, "High fidelity fingerprint generation: Quality, uniqueness, and privacy," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2021, pp. 3018–3022.
- [69] A. B. V. Wyzkowsky, M. P. Segundo, and R. de Paula Lemes, "Multiresolution synthetic fingerprint generation," *IET Biometrics*, vol. 11, no. 4, pp. 314–332, Jul. 2022.
- [70] J. J. Engelsma, S. A. Grosz, and A. K. Jain, "PrintsGAN: Synthetic fingerprint generator," 2022, *arXiv:2201.03674*.
- [71] H. Kim, X. Cui, M.-G. Kim, and T. H. B. Nguyen, "Reconstruction of fingerprints from minutiae using conditional adversarial networks," in *Proc. 17th Int. Workshop Digit. Watermarking, Digit. Forensics Watermarking*, vol. 11378, C. D. Yoo, Ed. Cham, Switzerland: Springer, 2019, pp. 353–362.
- [72] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5967–5976.
- [73] R. Bouzaglo and Y. Keller, "Synthesis and reconstruction of fingerprints using generative adversarial networks," 2022, *arXiv:2201.06164*.
- [74] A. Makrushin, V. Mannam, and J. Dittmann, "Data-driven fingerprint reconstruction from minutiae based on real and synthetic training data," in *Proc. 18th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, 2023, pp. 229–237.
- [75] Y. Choi, M. Choi, M. Kim, J.-W. Ha, S. Kim, and J. Choo, "StarGAN: Unified generative adversarial networks for multi-domain image-to-image translation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8789–8797.
- [76] Y. Choi, Y. Uh, J. Yoo, and J.-W. Ha, "StarGAN V2: Diverse image synthesis for multiple domains," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 8185–8194.
- [77] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," 2017, *arXiv:1710.10196*.
- [78] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," 2018, *arXiv:1812.04948*.
- [79] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of StyleGAN," 2019, *arXiv:1912.04958*.
- [80] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila, "Training generative adversarial networks with limited data," 2020, *arXiv:2006.06676*.
- [81] T. Karras, M. Aittala, S. Laine, E. Härkönen, J. Hellsten, J. Lehtinen, and T. Aila, "Alias-free generative adversarial networks," 2021, *arXiv:2106.12423*.
- [82] A. Zhmoginov and M. Sandler, "Inverting face embeddings with convolutional neural networks," 2016, *arXiv:1606.04189*.
- [83] F. Cole, D. Belanger, D. Krishnan, A. Sarma, I. Mosseri, and W. T. Freeman, "Synthesizing normalized faces from facial identity features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3386–3395.
- [84] G. Mai, K. Cao, P. C. Yuen, and A. K. Jain, "On the reconstruction of face images from deep face templates," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 5, pp. 1188–1202, May 2019.
- [85] S. Minaee and A. Abdolrashidi, "Iris-GAN: Learning to generate realistic iris images using convolutional GAN," 2018, *arXiv:1812.04822*.
- [86] M. B. Lee, Y. H. Kim, and K. R. Park, "Conditional generative adversarial network-based data augmentation for enhancement of iris recognition accuracy," *IEEE Access*, vol. 7, pp. 122134–122152, 2019.
- [87] S. Yadav, C. Chen, and A. Ross, "Synthesizing iris images using RaSGAN with application in presentation attack detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2019, pp. 2422–2430.
- [88] S. Yadav and A. Ross, "CIT-GAN: Cyclic image translation generative adversarial network with application in iris presentation attack detection," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 2411–2420.
- [89] D. Tomasevic, P. Peer, and V. Struc, "BiOcularGAN: Bimodal synthesis and annotation of ocular images," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Oct. 2022, pp. 1–10.
- [90] S. Ahmad and B. Fuller, "Resist: Reconstruction of irises from templates," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Sep. 2020, pp. 1–10.
- [91] J. Zhang, Z. Lu, M. Li, and H. Wu, "GAN-based image augmentation for finger-vein biometric recognition," *IEEE Access*, vol. 7, pp. 183118–183132, 2019.
- [92] H. Yang, P. Fang, and Z. Hao, "A GAN-based method for generating finger vein dataset," in *Proc. 3rd Int. Conf. Algorithms, Comput. Artif. Intell.*, Dec. 2020, pp. 1–6.
- [93] E. Salazar, R. Hernández-García, R. J. Barrientos, K. Vilches, M. Mora, and A. Vásquez, "Automatic generation of synthetic palm vein images: A nature-based approach," in *Proc. 11th Int. Conf. Pattern Recognit. Syst. (ICPRS)*, 2021, pp. 38–43.
- [94] C. Kauba, S. Kirchgasser, V. Mirjalili, A. Uhl, and A. Ross, "Inverse biometrics: Generating vascular images from binary templates," *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 3, no. 4, pp. 464–478, Oct. 2021.
- [95] J. Mathai, I. Masi, and W. Abdalmageed, "Does generative face completion help face recognition?" in *Proc. Int. Conf. Biometrics (ICB)*, Jun. 2019, pp. 1–8.
- [96] A. S. Joshi, A. Dabouei, J. Dawson, and N. M. Nasrabadi, "FDeblurGAN: Fingerprint deblurring using generative adversarial network," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Aug. 2021, pp. 1–8.
- [97] R. Durall, J. Jam, D. Strassel, M. H. Yap, and J. Keuper, "FacialGAN: Style transfer and attribute manipulation on synthetic faces," 2021, *arXiv:2110.09425*.
- [98] V. Mirjalili, S. Raschka, and A. Ross, "PrivacyNet: Semi-adversarial networks for multi-attribute face privacy," *IEEE Trans. Image Process.*, vol. 29, pp. 9400–9412, 2020.
- [99] M. Grimmer, H. Zhang, R. Ramachandra, K. B. Raja, and C. Busch, "Generation of non-deterministic synthetic face datasets guided by identity priors," 2021, *arXiv:2112.03632*.
- [100] L. Colbois, T. D. Freitas Pereira, and S. Marcel, "On the use of automatically generated synthetic image datasets for benchmarking face recognition," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Aug. 2021, pp. 1–8.
- [101] W. Nie, T. Karras, A. Garg, S. Debnath, A. Patney, A. B. Patel, and A. Anandkumar, "Semi-supervised StyleGAN for disentanglement learning," in *Proc. 37th Int. Conf. Mach. Learn.*, 2020, pp. 1–4.
- [102] Z. Wu, D. Lischinski, and E. Shechtman, "StyleSpace analysis: Disentangled controls for StyleGAN image generation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 12863–12872.
- [103] H. Kaur and R. Manduchi, "EyeGAN: Gaze-preserving, mask-mediated eye image synthesis," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Oct. 2020, pp. 299–308.
- [104] A. H. Ansari, "Generation and storage of large synthetic fingerprint database," M.E. thesis, Indian Inst. Sci. Bangalore, India, Jul. 2011.
- [105] Y. I. Shehu, A. Ruiz-García, V. Palade, and A. E. James, "Sokoto coventry fingerprint dataset," 2018, *arXiv:1807.10609*.
- [106] A. B. V. Wyzkowsky, M. P. Segundo, and R. de Paula Lemes, "Level three synthetic fingerprint generation," in *Proc. 25th Int. Conf. Pattern Recognit. (ICPR)*, Jan. 2021, pp. 9250–9257.
- [107] E. Wood, T. Baltrusaitis, C. Hewitt, S. Dziadzio, T. J. Cashman, and J. Shotton, "Fake it till you make it: Face analysis in the wild using synthetic data alone," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 3681–3691.
- [108] J. C. Neves, R. Tolosana, R. Vera-Rodriguez, V. Lopes, H. Proenca, and J. Fierrez, "GANprintR: Improved fakes and evaluation of the state of the art in face manipulation detection," *IEEE J. Sel. Topics Signal Process.*, vol. 14, no. 5, pp. 1038–1048, Aug. 2020.
- [109] S. Banerjee, W. J. Scheirer, K. W. Bowyer, and P. J. Flynn, "Fast face image synthesis with minimal training," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2019, pp. 2126–2136.
- [110] A. Kortylewski, B. Egger, A. Schneider, T. Gerig, A. Morel-Forster, and T. Vetter, "Analyzing and reducing the damage of dataset bias to face recognition with synthetic data," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2019, pp. 2261–2268.
- [111] A. Todorov and D. Oh, "Chapter four—The structure and perceptual basis of social judgments from faces," in *Advances in Experimental Social Psychology*, vol. 63, B. Gawronski, Ed. New York, NY, USA: Academic, 2021, pp. 189–245.
- [112] P. Costa, A. Galdran, M. I. Meyer, M. Niemeijer, M. Abramoff, A. M. Mendonça, and A. Campilho, "End-to-end adversarial retinal image synthesis," *IEEE Trans. Med. Imag.*, vol. 37, no. 3, pp. 781–791, Mar. 2017.

- [113] Z. Yu, Q. Xiang, J. Meng, C. Kou, Q. Ren, and Y. Lu, "Retinal image synthesis from multiple-landmarks input with generative adversarial networks," *Biomed. Eng. OnLine*, vol. 18, no. 1 p. 62, Dec. 2019.
- [114] J. S. Chen, A. S. Coyner, R. P. Chan, M. E. Hartnett, D. M. Moshfeghi, L. A. Owen, J. Kalpathy-Cramer, M. F. Chiang, and J. P. Campbell, "Deepfakes in ophthalmology: Applications and realism of synthetic retinal images from generative adversarial networks," *Ophthalmol. Sci.*, vol. 1, no. 4, 2021, Art. no. 100079.
- [115] F. Boutros, M. Huber, P. Siebke, T. Rieber, and N. Damer, "SFace: Privacy-friendly and accurate face recognition using synthetic data," 2022, *arXiv:2206.10520*.
- [116] H. Qiu, B. Yu, D. Gong, Z. Li, W. Liu, and D. Tao, "SynFace: Face recognition with synthetic data," 2021, *arXiv:2108.07960*.
- [117] H. Zhang, M. Grimmer, R. Ramachandra, K. B. Raja, and C. Busch, "On the applicability of synthetic data for face recognition," 2021, *arXiv:2104.02815*.



ANDREAS UHL is currently a Professor of computer science with the Department of Artificial Intelligence and Human Interfaces, University of Salzburg, Austria. He is the Head of the Multimedia Signal Processing and Security Laboratory consisting of currently ten full-time researchers. The research emphasis of the group is on the analysis of biomedical signals and media security topics, dealing with biometric systems, media forensics, recognition of flora and fauna in advanced agriculture systems, medical image and video analysis, multimedia security, and image/signal processing in digital humanities. In biometrics, his main research interests include vascular biometrics, fingerprint recognition, face recognition, and associated questions of template security and privacy and the generation and assessment of synthetic biometric data.



ANDREY MAKRUSHIN received the degree in applied mathematics and computer science from Lomonosov Moscow State University and the M.Sc. degree in data and knowledge engineering and the Ph.D. degree in computer science from the Advanced Multimedia Security Laboratory, Otto von Guericke University Magdeburg (OVGU).

Since 2007, he has been with the Advanced Multimedia Security Laboratory, OVGU. He has participated in quite a few scientific projects joining the two disciplines: biometrics and forensics, as a Senior Research Assistant. He is currently with the generation of synthetic fingerprints within the project GENSYNTH. He authored or coauthored more than 30 research articles. His research interests include biometrics, pattern recognition, image processing, and digital image forensics.

Dr. Makrushin is a member of ACM, EAB, and INSTICC. He is a reviewer for several prominent scientific journals on multimedia forensics and biometrics.



JANA DITTMANN (Member, IEEE) has been the Leader of the Advanced Multimedia and Security Laboratory (AMSL), Otto von Guericke University Magdeburg (OVGU), since 2002. She has numerous experience in security and privacy by design issues. She has a special interest in hidden communication, media and crime scene forensics, and biometrics for data and user authentication as well as for automotive, ICS, and the IoT security and its testing. She has more than 20 years of multimedia and security expertise in research and teaching. She has successfully supervised and is supervising internal and external Ph.D. candidates, the master's, and bachelor's thesis in the field of security. In several projects, her laboratory researched the field of sensory investigations and its related experiments, industry control systems, biometrics, forensics, and crime scene trace investigations. Besides practical system design, theoretical analysis and theoretical modeling is a substantial research focus.

...