Generation of Privacy-friendly Datasets of Latent Fingerprint Images using Generative Adversarial Networks

Stefan Seidlitz, Kris Jürgens, Andrey Makrushin, Christian Kraetzer and Jana Dittmann
Otto-von-Guericke University Magdeburg, Universitaetsplatz 2, Magdeburg, Germany

Keywords: Digitized Forensics, Latent Fingerprint, Image Synthesis, Generative Adversarial Networks, GAN, Privacy.

Abstract: The restrictions posed by the recent trans-border regulations to the usage of biometric data force researchers in the fields of digitized forensics and biometrics to use synthetic data for development and evaluation of new algorithms. For digitized forensics, we introduce a technique for conversion of privacy-sensitive datasets of real latent fingerprints to "privacy-friendly" datasets of synthesized fingerprints. Privacy-friendly means in our context that the generated fingerprint images cannot be linked to a particular person who provided fingerprints to the original dataset. In contrast to the standard fingerprint generation approach that makes use of mathematical modeling for drawing ridge-line patterns, we propose applying a data-driven approach making use of generative adversarial neural networks (GAN). In our synthesis experiments the performance of three established GAN architectures is examined. The NIST Special Database 27 is exemplary used as a data source of real latent fingerprints. The set of training images is augmented by applying filters from the StirTrace benchmarking tool. The suitability of the generated fingerprint images is checked with the NIST fingerprint image quality tool (NFIQ2). The unlinkability to any original fingerprint is established by evaluating outcomes of the NIST fingerprint matching tool.

1 INTRODUCTION

Fingerprints are known to be directly linked to an individual and fingerprint as biometric modality is well accepted and widely established means of user authentication. The applications making use of fingerprints spread form biometric access control systems to forensic investigation of latent fingerprints. Empirical studies on forensic or biometric fingerprint processing and recognition require a large dataset of fingerprints for validation of results. Moreover, development of fingerprint detection and recognition algorithms based on machine learning is hardly possible without an abundant amount of training data. However, fingerprints as well as any other biometric data are seen as a special category of personal data which is prohibited to be processed by the recent trans-border regulations without exception for the purpose of non-commercial research. A prominent example of a fingerprint dataset valuable for digitized forensics which was removed from the public access after establishing such regulations is the NIST Special Database 27 (Garris et al., 2000). Note that some categories of personal data can be processed after anonymization which is not the case for biometric samples because they require no meta-data to be linked to individuals. An elegant solution to the privacy-caused processing restrictions is a generation of artificial fingerprints that have the same characteristics as real fingerprints, but cannot be linked to particular persons. For biometrics, there is a de facto standard synthesizing tool called SFinGe (Cappelli, 2009). Fingerprints are generated to fit a certain basic pattern and a predefined set of minutitae. In contrast, in the field of digitized forensics the focus is on mimicking substrate and environmental influences on a digitalized latent fingerprint. Since such influences can be hardly formalized, modern data-driven image generation approaches has to be adopted for fingerprints.

Recent achievements in development of deep convolutional neural networks and especially generative adversarial networks (GAN) allow for automated generation of artificial images that can be hardly told apart from real images. In this paper, we examine several GAN architectures in application to generation of fingerprint images and assess the quality of the generated images. The quality is two-fold. The generated fingerprints must appear natural

Seidlitz, S., Jürgens, K., Makrushin, A., Kraetzer, C. and Dittmann, J.
Generation of Privacy-friendly Datasets of Latent Fingerprint Images using Generative Adversarial Networks.
DOI: 10.5220/0010251603450352
In *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2021) - Volume 4: VISAPP*, pages 345-352
ISBN: 978-989-758-488-6

and be privacy-friendly. Natural appearance means that the naked eye cannot see difference to a typical original pattern and a fingerprint is applicable for the further investigations (appropriate basic pattern, sufficient number of minutiae, etc). Privacy-friendly means that a generated fingerprint does not match one particular original fingerprint. If a generated fingerprint matches k ($k \neq 1$) original instances, then k-anonymity will be indicated. As a metric for both criteria we use NIST tools (Ko, 2007): NFIQ2 for the former and the combination of the MINDTCT and Bozorth3 for the latter.

Our contribution is in demonstrating that GAN can be successfully applied for conversion of privacy-sensitive datasets of fingerprint images to privacy-friendly datasets and in comparing three currently very prominent GAN architectures: ProgressiveGAN, StyleGAN and StyleGAN2 for this purpose. While comparing generation perfor-mances of the networks, we consider the following characteristics: how many training iterations are required to obtain high-quality fingerprint images, the time of image generation, and the proportion of high-quality fingerprints in the whole number of generated fingerprints. The better network would be able to generate more high-quality fingerprints in a shorter time frame. In order to improve the diversity of generated images and make the process of image generation more stable, we augment the training set by applying filters from the StirTrace benchmarking tool (Hildebrandt et al., 2015). Note that here we are not focused on generation of one particular dataset of artificial fingerprints, but rather propose a technique for compilation of such privacy-friendly datasets out of existing data.

In Section 2, we overview related works. In Section 3, we introduce our concept of generation and assessment of synthetic fingerprint images. In Section 4, we elaborate on important aspects of our implementation. In Section 5, we evaluate the GAN generated fingerprint images. Section 6 concludes the paper with the summary of results.

2 RELATED WORKS

Early works on generation of synthetic fingerprint images were concerned rather with reconstruction of ridge-line patterns from biometric templates considering this act as a potential attack on a biometric system (Galbally et al., 2008). Starting from the set of minutiae, a fingerprint area, an orientation map and a frequency map are estimated. Then, the iterative pattern growing approach draws ridges along the orientation lines by applying Gabor filters (Cappelli et al., 2007). This approach has been implemented in the software called SFinGe (Cappelli, 2009). The most critical step in this workflow is the estimation of an orientation map based on the set of minutiae or, to be more precise, based on singular points of a basic pattern (core, deltas). It is shown in (Ram et al., 2010) that singular points can be modeled by the zero-poles of Legendre polynomials, resulting in a discontinuous orientation field. An exhaustive study on the possibility of modeling fingerprints by the phase portraits of differential equations is conducted in (Zinoun, 2018) and the limitations are outlined.

In contrast to mathematical modeling, the modern trend is a data-driven generation of realistic images by means of generative adversarial networks (GAN). Recently, very impressive results have been demonstrated with images of human faces. Based on a huge amount of face images, researchers from NVIDIA successfully created high-quality, high-resolution synthetic faces, which can be hardly told apart from the real ones (Karras et al., 2018). The first effort to synthesize fingerprints using a Wasserstein GAN is made in (Bontrager et al, 2017) aiming at generating so-called *master* fingerprints that match multiple original fingerprints. Later on, in (Minaee et al., 2018), a connectivity imposed GAN is introduced and applied to two datasets: FVC-2006 and PolyU. The size of generated images in both publications is rather insufficient. In (Attia et al, 2019) fingerprints are synthesized by a variational autoencoder. In (Cao et al., 2018), a combination of an autoencoder and an adapted Wasserstein GAN is used for synthesizing 512x512 pixel fingerprint images. A CycleGAN is applied in (Wyzykowski et al., 2020) to transfer texture from real fingerprints to conventionally synthesized ridge-line patterns with added sweet pores which dramatically improves their realistic appearance. An alternative approach for generating high-resolution realistic fingerprints is in combining GAN with a super-resolution network proposed in (Riazi et al. 2020). In (Fahim, et al., 2020), a lightweight GAN is proposed for creating 128x128 pixel images and compared with five established GAN architectures based on 64x64 pixel patches. The next breakthrough is done in (Mistry et al., 2020) by incorporating identity information into the fingerprint synthesis network which is based once again on combining auto-encoder and Wasserstein GAN.

Here, we look for a suitable GAN architecture to generate high-resolution (512x512 pixel) gray-scale fingerprint images. This target size corresponds to the high-quality fingerprint image format (1000 ppi) used in dactyloscopic analyses (Orandi et al., 2014). Images with this size and resolution are capable of representing a significant part of a scanned latent fingerprint incl. the level 1 feature (basic pattern) as well as a sufficient number of level 2 features (minutiae) and potentially, depending on the prints quality, also level 3 features (sweat pores). The vast majority of dated GAN architectures are designed to generate images with a resolution less or equal to 256x256 pixels (see e.g. (Karras et al., 2017)) and therefore omitted in our considerations. Generating images of higher resolution is possible by adding several up-sampling layers (Zhang et al., 2016), but it would increase instability of the training process. Some GAN architectures such as ConditionalGAN (Wang et al., 2018) suffice the resolution criterion but are not capable of generating plausible images from fully random latent vectors. Such GANs are also not addressed. To the best of our knowledge, currently only three GAN architectures fit to our requirements: ProgressiveGAN (Karras et al., 2017), StyleGAN (Karras et al., 2018) and StyleGAN2 (Karras et al., 2019). These networks are able to generate plausible high-resolution images from fully random latent vectors.

3 CONCEPT

The usage of GAN is driven by privacy concerns. The synthesized fingerprints in a privacy-friendly dataset should exactly reproduce the characteristics of a reference dataset, but must leak no information in terms of reproducing the same minutiae.

The classic fingerprint synthesizing approaches start with minutiae or singular points to generate ridge-lines. If a set of minutia is taken from a reference fingerprint, the synthesized fingerprint would perfectly match it. Random selection of singular points or of a set of minutiae may lead to generation of implausible ridge-line patterns. Hence, tools like SFinGe can be successfully applied for generation of fingerprints with parameterizable characteristics reaching the diversity by randomizing parameters. However, SFinGe is hardly applicable to mimic characteristics that are presented in some reference data but cannot be formally described such as an appearance of a substrate. This is often a case for latent fingerprints which include not only characteristic of the fingerprint itself, but also characteristics of the environment in which the fingerprints were left behind (including topological characteristics of the surface). Table 1 summarizes the differences between the both aforementioned concepts of synthesizing fingerprint images.

Table 1: Comparison of SFinGe and GAN for generation of fingerprint images.

	SFinGe	GAN
Generation approach	mathematical modeling	data-driven
Fingerprint type	exemplar	latent
Reproduction of	characteristics of a fingerprint (basic pattern, minutiae)	characteristics of environment (incl. substrate, digitalization process etc.)

Our concept for the generation of privacy-friendly fingerprint image datasets is illustrated in Figures 1 and 2. The basic idea is that, for a given proprietary dataset of fingerprint images, we create a dataset of anonymous fingerprint images that are not linked to individuals but preserve all characteristics of the images in the initial dataset incl. background noise, image quality, frequency of ridge lines, one of the standard basic patterns, plausible number and locations of minutiae, etc. The number of samples in a new dataset is optional and depends only on time spent to the generation process. Note that images in the original dataset may include more than one fingerprint and may also vary in size. In contrast, the GAN training images must all have a specific size. Hence, a dataset specific pre-processing of images is required.

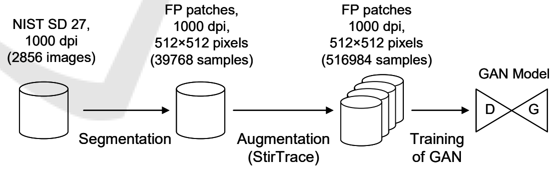


Figure 1: Training of a GAN model.

Figure 1 schematically describes training of a GAN model. In our further considerations, we take NIST Special Database (SD) 27 (Garris, 2000) with 2856 latent and matching tenprint fingerprint images as an example source of fingerprint data. There are two pre-processing steps performed before images are fed into the network: First, the fingerprints are segmented from the images. The patches are cut around fingerprint core points so that each resulting image contains exactly one partial or full fingerprint. The core points are located using the NIST tool MINDTCT (Ko, 2010). Note that the MINDTCT sometimes falsely highlight some artifacts like letters on the scanned trace cards in NIST SD 27 images as core points leading to non-fingerprint patches. However, we decided not to remove these outliers from the training data because they should be automatically sorted out in the later steps of the generation process. The number of extracted fingerprint patches is 39768. Second, the resulting set of fingerprint patches is augmented by their filtered versions (see Section 3.1) increasing the number from 39768 to 516984 for 12 StirTrace filters applied. This is done to improve the diversity of the generated images and to make the process of image generation more stable.

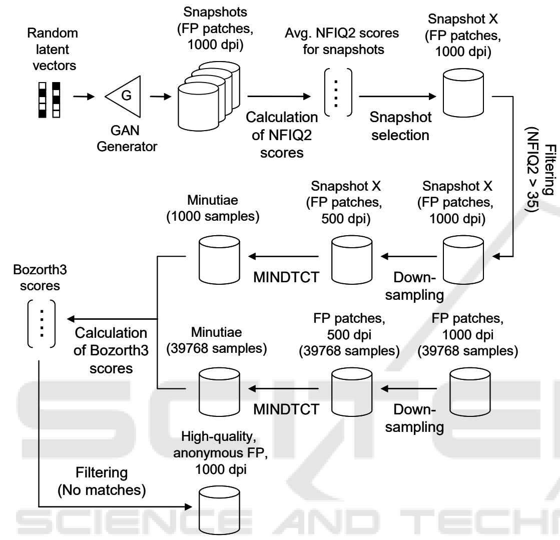

Random latent vectors
G
GAN Generator
Snapshots (FP patches, 1000 dpi)
Calculation of NFIQ2 scores
Avg. NFIQ2 scores for snapshots
Snapshot selection
Snapshot X (FP patches, 1000 dpi)
Filtering (NFIQ2 > 35)
Snapshot X (FP patches, 1000 dpi)
Down-sampling
Snapshot X (FP patches, 500 dpi)
MINDTCT
Minutiae (1000 samples)
Bozorth3 scores
Calculation of Bozorth3 scores
FP patches, 1000 dpi (39768 samples)
Down-sampling
FP patches, 500 dpi (39768 samples)
MINDTCT
Minutiae (39768 samples)
Filtering (No matches)
High-quality, anonymous FP, 1000 dpi

Figure 2: Generation of synthetic fingerprint images.

Figure 2 demonstrates the generation process of fingerprints and assessment of their quality. During GAN training, random latent vectors are fed into the generator resulting in synthetic fingerprint images. The snapshots of generated fingerprints are automatically stored at some training iterations. From each snapshot, we randomly pick 1000 fingerprint images for further analysis. For every image in every snapshot, we obtain a quality score using the NIST Fingerprint Image Quality estimator - NFIQ2 (Elham et al., 2013). From each snapshot, we calculate the mean value of the NFIQ2 scores and the number of images with NFIQ2 scores higher than 35 to determine at which iteration of the training process the generation model works best. Images with such scores correspond to the two highest quality classes in the 5-class NFIQ scale (Galbally et al., 2019) and, therefore, are referred to as high-quality fingerprints. The best snapshot has the highest number of high-quality 512x512 pixel fingerprints. Fingerprints with NFIQ2 scores lower or equal to 35 are filtered out. Each remaining fingerprint is biometrically matched to all training fingerprints using NIST tools (see Section 3.3).

3.1 Data Augmentation with StirTrace

StirTrace (Hildebrandt et al., 2015) is designed for benchmarking pattern recognition tasks in the context of digitized forensics. This tool can be seen as a set of filters to mimic typical artifacts that usually arise in the process of digitizing fingerprints, especially addition of noise. The most relevant filters used here for data augmentation are additive noise (strengths: 3, 5 and 9), additive Gaussian noise (strengths: 3, 5 and 9), median cut (strengths 3, 5 and 9) and salt and pepper noise (strengths: 3, 5 and 9). Note that filtering does not change the number and location of minutiae in a fingerprint. Hence, it is sufficient that the generated synthetic fingerprints are matched only against original patches and not against all training samples.

3.2 GAN Architectures

The term Generative Adversarial Network (GAN) was proposed in 2014 by Ian Godfellow for the architecture containing two neural networks Generator (G) and Discriminator (D), which are trained interchangeably. Generator produces synthetic data from random vectors and discriminator tries to distinguish these data from genuine data. The basic idea is that the generator improves with training while the discriminator's performance gets worse. However, after the point where the discrimi-nator is unable to tell apart genuine and synthesized data, the generator cannot be improved further, making the training process rather unstable. When operating with images, D is represented by a convolutional neural network and G by a de-convolutional neural network.

Progressive growing GAN (ProgressiveGAN) is an elegant solution to the convergence issue of the GAN training (Karras et al., 2017). Training begins with low-resolution images e.g. 4x4 pixels and the input images are re-scaled to this resolution. After the training process converges at the selected resolution, the resolution is increased and the training is repeated. This is done until the target resolution is reached. This is how the generator learns rough characteristics of training images first and then gradually fine characteristics. Technically, switching of resolution happens by gradual addition of intermediate layers into the network. Progressive-GAN was the first architecture that enabled gene-ration of high-resolution naturally looking fake faces which can be hardly told apart from real ones.

Based on ProgressiveGAN, an improved architecture called StyleGAN (Karras et al., 2018) was developed to take apart aggregated characteristics of images also referred to as styles and to move from one style to another. For face images the styles are e.g. a haircut or a skin color. Technically, the analysis and clustering of the latent space is done by introduction of adaptive instance normalization (AdaIN) layers and addition of noises to control the intensity of a particular style. However, StyleGAN often produces characteristic imperfection e.g. droplet or phase artifacts which are attributed to the network architecture. Droplet artifacts arise due to independent normalization of means and variations of different style feature maps in AdaIN layers and phase artifacts arise due to progressive growing. A variation of the StyleGAN architecture called StyleGAN2 (Karras et al., 2019) was proposed to avoid the aforementioned imperfections. Adaptive instance normalization is replaced by weight demodulation making normalization of means and variations of the different styles not independent anymore. The progressive growing issue is solved by generating images with only target resolution by adding up the weighted outcomes of all layers of the generator. It does not change the idea of progressive learning of image characteristics (from rough to fine), but helps to get rid of phase artifacts.

3.3 Fingerprint Anonymity Assessment

Biometric matching of fingerprints is done by using NIST tools: MINDTCT and Bozorth3 (Ko, 2007). MINDTCT extracts the list of minutia from a fingerprint while Bozorth3 compares two such lists and produces a similarity score representing the number of matched minutiae. Note that MINDTCT requires 500 dpi images to properly detect minutiae. Since the target resolution of our synthesized fingerprints is 1000 dpi, the images are downscaled with a factor 2 before applying MINDTCT. The documentation of Bozorth3 suggests values over 40 for the perfect match. However, we use here a value of 30 as a decision threshold to guarantee better anonymity.

4 IMPLEMENTATION

We use the original NVIDIA implementations of the addressed GAN architectures from GitHub repositories: http://github.com/NVlabs/stylegan and http://github.com/NVlabs/stylegan2. The implementation of ProgressiveGAN is a part of the StyleGAN repository. The GANs are used in their default configuration, we only set the target image size to 512x512 pixel and kImages to 7500. The parameter kImages refers to the amount of real images which the GAN discriminator has seen during training. We switch off the estimation of the perceptual path length and linear separability of all GANs because these metrics do not work with gray scale images (for reasons see (Karras et al., 2018)). The amounts of layers used and the trainable parameters of each GAN are summarized in Table 2. For the training of the GANs we used a workstation with two NVIDIA Titan RTX graphic cards with 24 GB VRAM each.

The StirTrace tool is used in version 4 as provided at https://sourceforge.net/projects/stirtrace. The applied filters as well as the composition of the training set after this data augmentation step are summarized in Table 3.

The NFIQ2, MINDTCT and Bozorth3 are the parts of the NIST Biometric Image Software (NBIS) which is available at https://www.nist.gov/services-resources/software/nistbiometric-image-software-nbis.

Table 2: Configurations of the addressed GANs.

GAN type	network type	amount of layers	trainable parameters
Progressive GAN	Generator	43	23.067.048
	Discriminator	44	23.075.169
StyleGAN	Generator	76	26.174.696
	Discriminator	44	23.075.169
StyleGAN2	Generator	67	30.270.551
	Discriminator	29	28.982.721

Table 3: Composition of the training dataset.

Filter	Kernel size	Number of samples
without filter	-	39768
+ additive noise	3, 5, 9	3 x 39768 = 119304
+ additive Gaussian noise	3, 5, 9	3 x 39768 = 119304
+ median cut	3, 5, 9	3 x 39768 = 119304
Salt & pepper noise	3, 5, 9	3 x 39768 = 119304
		Total: 516984

5 EVALUATION

We compare GAN architectures regarding the following aspects: how many training iterations are required to obtain high-quality fingerprint images, the speed of image generation, and the proportion of high-quality fingerprints in the whole number of generated fingerprints. The better network is able to generate more high-quality fingerprint images in a shorter time frame. Additionally, we validate the "anonymity" of the created datasets.

5.1 Training and Generation Time

The training time strongly varies from one GAN to another. Due to the time constraints, we initially limited the training parameter kImages to 7500. Nonetheless, only StyleGAN reached the image size of 512x512 pixels with 7500 kImages on one GPU. The training of ProgressiveGAN with 7500 kImages resulted in images of 256x256 pixels. Hence, the training was continued with 12000 kImages on two GPUs. For the proper comparison the 256x256 images are upscaled to 512x512 pixels. The training time of StyleGAN2 is extremely long (see Table 4), so that we first switched from one GPU to two GPUs after 3429 kImages and stopped training of after 4452 kImages (ca. 11 days). Nevertheless, even at this point, the generated fingerprint images already reached a good subjective quality (to be attributed to the residual structure of generator and discriminator). Note that StyleGAN2 has an advantage over the two other architectures because training already starts with 512x512 pixel images while the other two start with very small images und gradually upscale them.

Table 4: Training time of the GANs on our reference PC.

GAN type	Used GPUs	kImages	Training time	Reached image size
Progressive GAN	1	< 7500	2d 1h 44m	256x256
Progressive GAN	2	7501-12000	4d 22h 57m	512x512
StyleGAN	1	< 7500	3d 3h 36m	512x512
StyleGAN2	1	< 3429	8d 5h 12m	512x512
StyleGAN2	2	3430-4452	2d 20h 46m	512x512

5.2 Quality of Generated Images

Figures 3 and 4 demonstrate the development of the average NFIQ2 score and the ratio of high-quality images of 1000 images randomly selected from GAN snapshots over the course of training, respectively. We see that StyleGAN2 starts genera-ting high-quality fingerprints after only few training iterations and the fingerprint quality even degrades with the higher kImages parameter. The diagrams also demonstrate that StyleGAN clearly outperforms ProgressiveGAN regarding both the average NFIQ2 score and the number of fingerprints with NFIQ2 scores higher than 35. For StyleGAN and ProgressiveGAN, the snapshot at which the targeted resolution of 512x512 pixels is reached is marked. For all three GANs, the average NFIQ2 scores stabilize between 25 and 30.

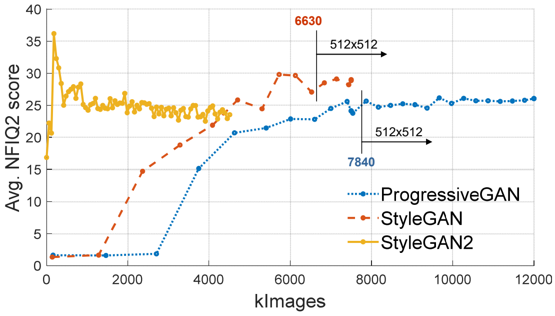


Figure 3: GAN training: average NFIQ2 scores.

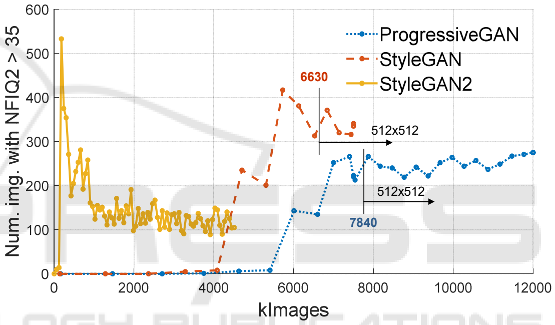


Figure 4: GAN training: ratio of high-quality fingerprints.

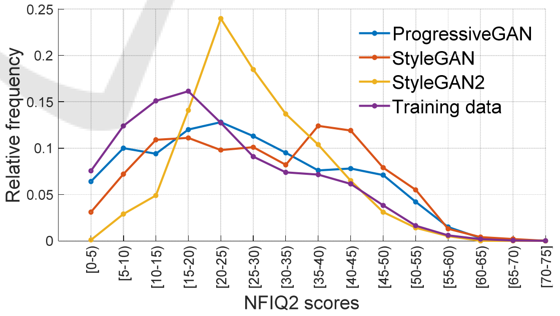
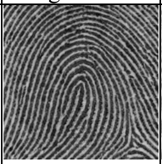
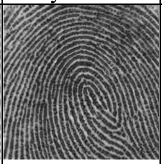
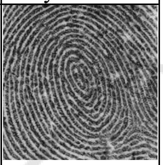
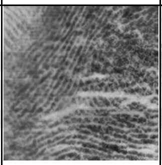
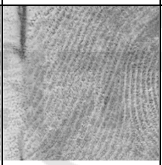
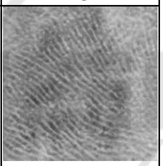


Figure 5: Distributions of NFIQ2 scores of 1000 random GAN fingerprint images from the selected snapshots.

We compare the performances of GANs by comparing the snapshots at which GAN reaches the highest average NFIQ2 score with the possibly low kImages parameter. The optimal value of kImages is 7870, 6840 and 1925 for ProgressiveGAN, StyleGAN and StyleGAN2 respectively. The distributions of NFIQ2 scores for the selected snapshots are depicted in Figure 5. The histograms shows that fingerprint images with the highest quality are generated by StyleGAN followed by Progressive GAN and StyleGAN2. Out of 1000 images in a corresponding snapshot, StyleGAN generated 371 images with NFIQ2 higher than 35, Progressive GAN 266 images and StyleGAN2 191 images. Note that all GAN generated images have on average higher NFIQ2 scores than original training images.

Table 5 visualizes the difference between GAN-generated fingerprint images of high (NFIQ2 > 35) and of relatively low quality (25 ≤ NFIQ2 < 35). The perceptual quality of the fingerprint images is extremely high, with plausible level 1 and level 2 features and, in the case of StyleGAN, even something that already looks like some sweat pores.

Table 5: Examples of GAN-generated fingerprint images.

	Progres.GAN	StyleGAN	StyleGAN2
NFIQ2 > 35	 42	 39	 43
NFIQ2 < 35	 26	 26	 25

5.3 Anonymity of Generated Fingerprints

After filtering out the low-quality fingerprint images with NFIQ2 score below 35, we compare each high-quality fingerprint image with all training images. Low Bozorth3 scores indicate that the generated fingerprints cannot be linked to any person who provided fingerprints to the training dataset. The average Bozorth3 scores range between two and six depending on the fingerprint and the GAN type. However, there are few generated fingerprints with exactly one Bozorth3 score higher than 30 indicating the match between a generated fingerprint and one of the fingerprints in the training dataset. Table 6 shows the ratio of synthesized fingerprints with Bozorth3 scores lower than 30, between 30 and 39, and higher than 40 for each considered GAN type. For StyleGAN for instance, 59 images have a score above 40 meaning 15.9% of high-quality non-anonymous images in the selected snapshot.

Table 6: Number of GAN fingerprints with Bozorth3 scores (s) in a certain range in the selected snapshots.

	Progres.GAN	StyleGAN	StyleGAN2
s < 30	169/266 ~ **63.53%**	182/371 ~ **49.06%**	106/191 ~ **55.50%**
30 ≤ s < 40	65/266 ~ 24.44%	130/371 ~ 35.04%	66/191 ~ 34.55%
s ≥ 40	32/266 ~ 12.03%	59/371 ~ 15.90%	19/191 ~ 9.95%

The experimental results suggest that Progressive GAN with 63.53% of Bozorth3 scores lower than 30 generates on average the highest number of anonymous fingerprints and therefore can be seen as the most privacy-friendly generation approach. ProgressiveGAN is followed by StyleGAN2 (55.5% anonymous fingerprints) and then StyleGAN (49.06% anonymous fingerprints). Considering the absolute number of the high-quality anonymous fingerprints in the selected snapshot, StyleGAN has clearly the best generation performance with 182 images followed by ProgressiveGAN (169 images) and then StyleGAN2 (106 images).

In our case study, we conducted experiments only with the NIST SD 27 database. The proportions of anonymous high-quality fingerprints within the whole number of generated fingerprints as well as the generation time cannot be generalized for any reference database taken as training data for a GAN. However, the experimental results clearly show that GAN is a suitable technique for "anonymization" of privacy-sensitive fingerprint datasets.

6 CONCLUSION

We demonstrate that a GAN is in general a suitable technique for generation of high-quality anonymous fingerprint images. As a data-driven approach a GAN takes a privacy-sensitive dataset and converts it to privacy-friendly dataset without loss of dataset characteristics. The resulting datasets can be used for research on fingerprints without privacy-caused limitations. In a case study with the NIST SD27 dataset, we show that all three addressed GAN architectures (ProgressiveGAN, StyleGAN and Style GAN2) are capable of converting original privacy-sensitive fingerprint images to privacy-friendly ones. StyleGAN2 has an advantage that the fingerprint images with high NFIQ2 scores are generated after only a few iterations of training. In contrast, ProgressiveGAN and StyleGAN require many training iterations to reach the target image resolution. However, StyleGAN2 is the worst approach regarding the absolute number of high-quality anonymous fingerprints generated. From the perspective of fast generation, StyleGAN is clearly superior. ProgressiveGAN is preferable regarding the better anonymity. Our future work will address the training of GAN models based on multifarious fingerprint images from many independent sources and conditional generation of fingerprint patterns such as predefined locations of minutia or substrate characteristics.

ACKNOWLEDGEMENTS

This research has been funded in part by the Deutsche Forschungsgemeinschaft (DFG) through the research project GENSYNTH under the number 421860227. We thank Philip Wiegratz, Alexander Heck and Mark Trebeljahr for participating in this work at an early stage and establishing the feasibility of GAN-based generation of fingerprint images.

REFERENCES

Attia, M., Attia, M. H., Iskander, J., Saleh, K., Nahavandi, D., Abobakr, A., Hossny, M., Nahavandi, S., 2019. Fingerprint Synthesis via Latent Space Representation, In *Proc. IEEE Int. Conf. on Systems, Man and Cybernetics*, pp. 1855-1861.

Bontrager, P., Roy, A., Togelius, J., Memon, N.D., Ross, A., 2018. DeepMasterPrints: Generating MasterPrints for Dictionary Attacks via Latent Variable Evolution, In *Proc. 9th IEEE Int. Conf. on Biometrics: Theory, Applications and Systems (BTAS)*, pp. 1-9.

Cao, K., Jain, A.K. 2018. Fingerprint Synthesis: Evaluating Fingerprint Search at Scale. In *Proc. Int. Conf. on Biometrics (ICB)*, pp. 31–38.

Cappelli, R., 2009. SFinGe, In *Li, S.Z., Jain, A. (eds) Encyclopedia of Biometrics*. Springer, Boston, MA

Cappelli, R., Lumini, A., Maio, D., Maltoni, D., 2007. *IEEE Trans. on Pattern Analysis and Machine Intelligence 29(9):1489-1503*.

Elham, T., Olsen, M. A., Makarov, A., Busch, C., 2013. Towards NFIQ II Lite: Self-Organizing Maps for Fingerprint Image Quality Assessment, *NIST Interagency Report 7973*, December 13, 2013.

Fahim M. A. I., Jung, H. Y., 2020. A Lightweight GAN Network for Large Scale Fingerprint Generation, *IEEE Access, vol. 8,* pp. 92918-92928.

Galbally, J., Cappelli, R., Lumini, A., Maltoni, D., Fierrez, J., 2008. Fake Fingertip Generation from a Minutiae Template, In *Proc. 19th Int. Conf. on Pattern Recognition (ICPR 2008)*, pp. 1-4.

Galbally, J., Haraksim, R., Ferrara, P., Beslay, L., Tabassi, E., 2019. Fingerprint Quality: Mapping NFIQ1 Classes and NFIQ2 Values, In *Proc. Int. Conf. on Biometrics (ICB 2019)*, pp. 1-8.

Garris, M., Mccabe, R., 2000. NIST Special Database 27 Fingerprint Minutiae From Latent and Matching Tenprint Images, *NIST Interagency Report 6534*, June 1, 2000.

Hildebrandt, M., Dittmann, J., 2015. StirTrace V2.0: Enhanced Benchmarking and Tuning of Printed Fingerprint Detection, *IEEE Trans. on Information Forensics and Security 10 (4):833-848*.

Karras, T., Aila, T., Laine, S., Lehtinen, J., 2017. Progressive Growing of GANs for Improved Quality, Stability, and Variation, *CoRR abs/1710.10196*.

Karras, T., Laine, S., Aila, T., 2018. A Style-Based Generator Architecture for Generative Adversarial Networks, *CoRR, vol. abs/1812.04948*.

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., Aila, T., 2019. Analyzing and Improving the Image Quality of StyleGAN, *CoRR, vol. abs/1912.04958*.

Ko, K., 2007. User's Guide to NIST Biometric Image Software (NBIS), *NIST Interagency Report 7392*, January 21, 2007.

Minaee, S., Abdolrashidi, A., 2018. Finger-GAN: Generating Realistic Fingerprint Images Using Connectivity Imposed GAN. *CoRR, vol. abs/ 1812.10482*.

Mistry, V., Engelsma, J.J., Jain, A.K., 2020. Fingerprint Synthesis: Search with 100 Million Prints. *CoRR, vol. abs/1912.07195*.

Orandi, S., Libert, J. M., Grantham, J. D., Ko, K., Wood, S. S., Byers, F. R., Bandini, B., Harvey, S. G., Garris, M. D., 2014. Compression Guidance for 1000 ppi Friction Ridge Imagery, *NIST Special Publication 500-289*, February 24, 2014

Ram, S., Bischof, H., Birchbauer, J., 2010. Modelling fingerprint ridge orientation using Legendre polynomials, *Pattern Recognition 43(1):342-357*.

Riazi, M.S., Chavoshian, S.M., Koushanfar, F., 2020. SynFi: Automatic Synthetic Fingerprint Generation, *CoRR, vol. abs/ 2002.08900.*

Wang, T., Liu, M., Zhu, J., Tao, A., Kautz, J., Catanzaro, B., 2018. High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs, In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2018)*, pp. 8798-8807.

Wyzykowski, A., Segundo, M., Lemes, R., 2020. Level Three Synthetic Fingerprint Generation, *CoRR, vol. abs/ 2002.03809.*

Zhang, H. Xu, T., Li, H., Zhang, S., Huang, X., Wang, X., Metaxas, D. N., 2016. StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks, *CoRR, vol. abs/1612.03242*.

Zinoun, F., 2018. Can a Fingerprint be Modelled by a Differential Equation?, *CoRR, vol. abs/1802.05671*.