

HIGH FIDELITY FINGERPRINT GENERATION: QUALITY, UNIQUENESS, AND PRIVACY*Keivan Bahmani, Richard Plesh, Peter Johnson, Stephanie Schuckers**Timothy Swyka*Clarkson University
Potsdam, NY, USAPrecise Biometrics
Potsdam, NY, USA**ABSTRACT**

In this work, we utilize progressive growth-based Generative Adversarial Networks (GANs) to develop the Clarkson Fingerprint Generator (CFG). We demonstrate that the CFG is capable of generating realistic, high fidelity, 512×512 pixels, full, plain impression fingerprints. Our results suggest that the fingerprints generated by the CFG are unique, diverse, and resemble the training dataset in terms of minutiae configuration and quality, while not revealing the underlying identities of the training data. We make the pre-trained CFG model and the synthetically generated dataset publicly available at https://github.com/keivanB/Clarkson_Finger_Gen

Index Terms— Fingerprint Synthesis, Generative Adversarial Networks, Fingerprint Matching, Fingerprint Quality

1. INTRODUCTION

Fingerprint identification systems are widely accepted as inexpensive and secure, making them one of the most prevalent forms of biometric recognition. Current state-of-the-art fingerprint recognition systems heavily rely on Convolutional Neural Networks (CNNs) [1]. While these systems show exceptional performance, they require expensive large-scale fingerprint datasets for training and evaluation. However, collecting and sharing large-scale biometric datasets comes with inherent risks and privacy concerns. For instance, the National Institute of Standards and Technology (NIST) recently discontinued several publicly available datasets from their catalog due to privacy issues [2]. Generating synthetic fingerprints can help alleviate both expense and privacy issues. Synthetic datasets can be generated easily at scale while still representing the training dataset and shielding the identity of the individuals that were used during training.

The traditional approach to fingerprint synthesis involves sampling from independent statistical models for orientation field and minutiae with Gabor-filtering or other models to generate the final ridge structure [3, 4, 5]. The fingerprints generated using these approaches suffer from some shortcomings. The additive noise used in the ridge valley generation

process can give the generated fingerprints a distinct visual pattern. Additionally, fingerprints generated using the traditional approaches can be distinguished from real fingerprints using their minutiae distribution [6]. Finally, the independent modeling used in the traditional approach is not necessarily able to capture the correlation between the orientation field, ridge valley structure, and minutiae patterns.

More recently, GAN models have been employed for fingerprint synthesis [7, 8, 9, 10]. Contrary to the previous fingerprint synthetic approaches, GANs don't rely on independent statistical models for each aspect of fingerprints. They are capable of learning the high-dimensional probability distribution of the training data and generating samples from the learned distribution. Previous GAN-based fingerprint synthesis work mainly utilized the Improved Wasserstein GAN (IWGAN) architecture [11]. These models, however, are unstable and fail to generalize to high-resolution images [8, 9, 10]. Previous work has proposed different training regimens to address these issues. Finger-GAN utilizes additional total variation constraints to impose connectivity within the generated images [9]. Fahim et al. proposed a loss doping approach to stabilize the training process and prevent mode collapse [10]. The proposed methods improved the stability of the training process, but failed to produce full plain impression fingerprints with precise boundary and high fidelity. Cao et al. focused on fingerprint search at scale and utilized a large-scale dataset of 250,000 rolled fingerprints. Despite the large dataset, the authors still observed the same mode collapse issues. In response, they proposed an unsupervised pre-training step, adding additional computational overhead. To the best of our knowledge, this is the only GAN-based model in previous work that is capable of producing high fidelity, 512×512 pixels, 500 dpi, rolled fingerprints [8]. In this work, we introduce the Clarkson Fingerprint Generator (CFG), a GAN-based fingerprint synthesis model which uses progressive growth training to generate realistic 512×512 pixels, plain impression fingerprint images, Figure 1 illustrates examples of the synthetic fingerprints. Our contributions are as follows:

- Contrary to the previous IWGAN-based fingerprint synthesis models, the CFG utilizes a multi-resolution and progressive growth training approach [12, 13]. The

This material is based upon work supported by the Center for Identification Technology Research and the National Science Foundation under Grant No.650503.

CFG can generate high fidelity plain fingerprint with realistic shape and boundaries at 512×512 pixels and does not suffer the mode collapse and quality issues associated with previously proposed IWGAN-based fingerprint generators [9, 10]. When compared to the model used in [8], the multi-resolution fingerprint synthesis model can be trained using a smaller dataset and without computationally expensive pre-training steps.

- We compare quality metric distributions to assess diversity of the synthetic fingerprints and their similarity to bonafide fingerprints. We also match every synthetic fingerprint to every bonafide fingerprint to ensure that the synthetic fingerprints do not reveal the real identities.
- We utilized a CNN-based Presentation Attack Detection (PAD) model to evaluate fingerprints generated using the CFG. This process reaffirms the high fidelity of the samples generated using the CFG.
- We make the pre-trained CFG model and the synthetically generated fingerprints publicly available. To the best of our knowledge, the CFG is the first publicly available GAN-based fingerprint synthesis model.

2. CLARKSON FINGERPRINT GENERATOR

In this work, we utilize multi-resolution training for fingerprint synthesis [13]. Multi-resolution GAN models start the training process by training both the Generator (G) and Discriminator (D) at lower spatial resolutions and progressively increasing (growing) the spatial resolution throughout the training. Progressive growth-based GANs are capable of effectively capturing high-frequency components of the training data and producing high-fidelity and realistic human faces [13]. The main known limitation of the progressive growth approach is the generator’s strong location preference for details. This issue has led to artifacts in generating high-resolution faces across different poses [14]. However, fingerprint recognition systems operate at a relatively fixed scale and do not suffer from the pose, illumination, and expression variations associated with face images. As a result, we believe this architecture alleviates the problems observed in the previous IWGAN-based fingerprint synthesis models while introducing a minimal amount of artifacts to the synthesized fingerprints. To the best of our knowledge, the CFG is the first fingerprint synthesis model that leverages multi-resolution and progressive growth training.

3. TRAINING AND EVALUATION

The CFG utilizes StyleGAN architecture [13]. The model is trained from scratch using 72,000, 512×512 pixels bonafide fingerprints from 250 unique identities, captured using a

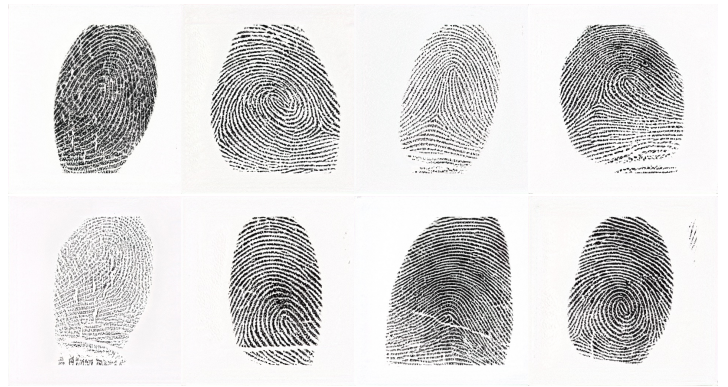


Fig. 1. Examples of generated fingerprint images by the proposed approach (512×512 pixels at 500 dpi)

Crossmatch Guardian scanner (DB-1). The CFG is trained in an unsupervised manner, i.e. we did not provide the model with unique identity labels. The bonafide fingerprints are processed at 8 resolutions from 4×4 up to 512×512 pixels during progressive training. Subsequently, we utilized the CFG to generate 50,000 synthetically generated 512×512 plain impression fingerprints (DB-2). Note that the synthetic fingerprints are generated without any truncation to represent the full range of the CFG [13].

Previous work on fingerprint synthesis relied on Frechet Inception Distance (FID) [9] and Structural SIMilarity (SSIM) [10] measures to respectively evaluate the quality and diversity of the generated samples. In the test, the CFG achieved FID of 24 which is considerably better than the FID of 70 reported by Finger-GAN [9]. We calculated FID using the 50,000 fingerprints. Unfortunately, the use of in-house datasets and proprietary code limits our ability to compare FID with previous fingerprint synthesis models. Also, since FID utilizes an inception network trained on ImageNet [15], it is best suited for evaluating generators of natural images rather than biometric images. Consequently, we utilize the BOZORTH3 minutiae-based fingerprint matcher [16] to evaluate the uniqueness of the synthetically generated fingerprints through their imposter distribution [8] and expand upon previous works by evaluating the quality and diversity of the synthetic fingerprints through fingerprint metrics. We evaluated the quality of the fingerprints using NIST NFIQ 2.0 [17] and utilized the NIST NBIS software [16] to evaluate and compare the minutiae configuration of the training (DB-1) and synthetic (DB-2) fingerprints. Additionally, we leveraged the work of Olsen et al. to extract features based on ridge-valley signature [18]. To accurately estimate the ridge-valley features, each fingerprint is decomposed into overlapping blocks of 32×32 pixels and we averaged the results over the 15 patches with lowest standard deviation in terms of ridge valley uniformity (highest quality).

Additionally we evaluate the synthetic fingerprints through

the use of a PAD model trained on a separate set of 50,000 (representing 250 people) and 35,000 PA images (representing 11 PA types) using Mobile NASNet architecture [19]. We replace the fully connected layers with four layers, three RELU activated layers of size 500, 50, 10, and a 2-unit softmax layer representing bonafide vs PA fingerprints. The PAD model is initialized using weights trained on ImageNet and subsequently trained using full fingerprint images. We utilized the Adam optimizer [20], a learning rate of 0.00001, and batch size of 128. The NASNet model achieves 13.5% Attack Presentation Classification Error Rate (APCER) at 0.5% Bonafide Presentation Classification Error Rate (BPCER).

The training of CFG, the CNN-based PAD model, and fingerprint synthesis process is carried out using 4 NVIDIA P100 GPUs provided by the National Science Foundation (NSF) Chameleon testbed [21].

4. RESULTS AND DISCUSSIONS

Figure 1 depicts several synthetic fingerprints generated with the CFG. Upon visual investigation, the CFG generates diverse, high-fidelity full plain impression fingerprints at 512×512 pixels without noisy boundaries and artifacts observed in the IWGAN-based fingerprint synthesis models [10, 9].

4.1. Matching Performance

To construct the empirical, genuine, and imposter distributions for the bonafide samples, we randomly select 1000 non-mated pairs for each fingerprint, resulting in 72 million non-mated scores, and we utilized all of the 3.8 million mated pairs. Figure 2 (a) depicts the imposter distribution of the bonafide fingerprints (DB-1). We select a conservative False Accept Rate (FAR) of 0.001% to evaluate the uniqueness of the synthetic fingerprints. Within the bonafide fingerprints (DB-1), this FPR translates into the True Positive Rate (TPR) of 0.873 at a score threshold of 41. Note that this threshold is in line with the rule of thumb threshold (40) suggested for the BOZORTH3 matcher [16]. We use the same threshold (41) to evaluate the uniqueness of the synthetically generated fingerprints (DB-2). The synthetically generated fingerprints in DB-2 are compared with every bonafide fingerprint used for training the CFG (DB-1). This process produces more than 3 billion match scores. We randomly select 40 million of those to construct an empirical imposter match score distribution between synthetic fingerprints (DB-2) and bonafide fingerprints (DB-1). Figure 2 (b) illustrates this imposter distribution.

Interestingly the empirical distribution of synthetic fingerprints do not have the long tail that we observed in the imposter score distribution of the bonafide fingerprints (DB-1). Enforcing the same matching threshold (41) resulted in 190 false matches (maximum score of 53) as opposed to 330 false matches (maximum score of 286) observed in DB-1. Since

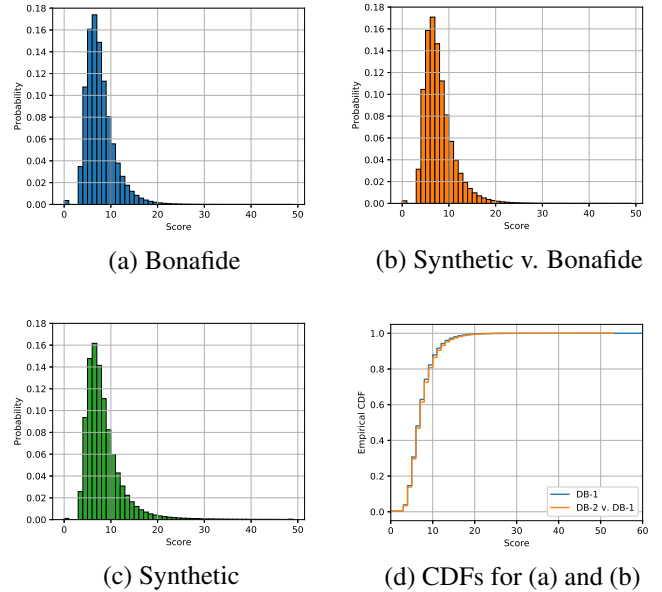


Fig. 2. Imposter score distributions for: (a) bonafide (DB-1), (B) synthetic (DB-2) versus bonafide (DB-1), (c) synthetic only (DB-2), and (d) empirical CDFs for (a) and (b).

the imposter distributions are not normally distributed, we conduct a non-parametric independent samples Kolmogorov-Smirnov test between the empirical imposter distributions to confirm the distinctness of the synthetic fingerprints. We perform the test between the empirical imposter distribution observed by matching bonafide fingerprints (DB-1) and the empirical imposter distribution observed by matching synthetic to bonafide fingerprints (DB-2 v. DB-1) with an alternative hypothesis that the imposter distribution of DB-2-v-DB-1 is *less* than DB-1. The Kolmogorov-Smirnov test ($D(40M) = 0.0158, p > 0$) indicates rejecting the null hypothesis in favor of the alternative hypotheses. As a result, we can confirm that the synthetic fingerprints generated by the CFG (DB-2) are distinct from bonafide samples in the training dataset (DB-1) and do not reveal the identities presented in the training dataset (DB-1).

Finally, we also compared every synthetic fingerprint to every other synthetic fingerprint in DB-2 and randomly selected 40M comparisons. Figure 2 (C) shows the empirical match score distribution for DB-2. We can observe that the distribution of match scores for synthetic fingerprints from DB-2 resembles that of bonafide in DB-1. However, the imposter distribution of DB-2 has a much longer tail that reveals an overlap of identities in the generated fingerprints. Enforcing the same threshold as DB-1 (41) results in 302,285 false matches out of 40M comparisons, compared to 330 in DB-1. Clearly there are multiple samples from the same identity in DB-2. This is expected as the CFG is trained in an unsupervised manner (without identity labels) and we did not force

Measure	DB-1 (Bonafide)				DB-2 (Synthetic)			
	Mean	STD	Skewness	Kurtosis	Mean	STD	Skewness	Kurtosis
Ridge Ending Minutiae Count [16]	47.378	15.019	1.040	2.323	45.606	12.371	0.662	0.602
Bifurcation Minutiae Count [16]	19.723	11.831	1.706	5.787	21.11	11.247	1.087	0.994
Reliability of Ridge Minutiae [16]	0.458	0.098	0.149	-0.048	0.453	0.081	-0.148	-0.326
Reliability of Bifurcation Minutiae [16]	0.592	0.142	-0.547	0.005	0.626	0.145	-0.668	-0.032
Percentage of Bifurcation Minutiae [16]	0.285	0.120	0.518	0.192	0.304	0.097	0.267	-0.043
Ridge Count [18]	3.272	0.375	0.420	0.66	3.331	0.334	0.326	-0.36
White Lines Count [18]	3.271	0.376	0.423	0.623	3.323	0.337	0.353	-0.356
RTVTR [18]	0.921	0.205	0.657	0.446	0.910	0.148	0.387	-0.205
Area of the Fingerprint	95.392	24.224	0.349	-0.317	94.019	21.232	0.688	0.307
NFIQ2 Score [17]	57.406	18.482	-0.365	2.622	54.687	14.616	-0.499	1.058

Table 1. Fingerprint metrics for bonafide (DB-1) and synthetic (DB-2) fingerprints. NFIQ2 score has a range of $[0, 100]$ and The Ridge to Valley Thickness Ratio (RTVTR) has a range of $[0, 1]$.

the generator to generate fingerprints by varying the disentangled identity portion of the learned latent variable (W). However, the tests also show that out of 50K samples 11.5K samples have no comparisons with match score above 41 and can be considered as synthetically generated unique identities. In our future work, we aim to train the next generation of the model using identity, finger, and quality labels. This would allow researchers to have granular control over those factors in the generated fingerprints.

4.2. Quality, Diversity, and Fingerprint Metrics

Table 1 presents fingerprint metrics for the synthetic and bonafide fingerprints. We observe high standard deviation in the evaluated metrics of the synthetic fingerprints. As an example, Figure 3 illustrates the distribution of Ridge Ending Minutiae in the training (DB-1) and synthetic (DB-2) fingerprints. This suggests that the CFG is generating fingerprints with diverse minutiae configurations and is not suffering from the mode collapse issues associated with other IWGAN models [10, 9]. Additionally, we observe that the mean and standard deviation of the fingerprint metrics from the synthetically generated fingerprints are close to that of the training dataset (DB-1). Consequently, we can confirm the diversity and quality of the synthetically generated fingerprints. However, our results indicate differences in the skewness and kurtosis between synthetic and bonafide fingerprint metrics. For future work, we believe integrating such metrics into the GAN-based fingerprint synthesis model could improve the model’s performance even further.

Additionally, we evaluated our synthetically generated fingerprint (DB-2) using the trained PAD model. The PAD model classified 95.2% of the synthetic fingerprint as bonafide fingerprints. The PAD test reaffirms the quality and fidelity of the synthetically generated fingerprints.

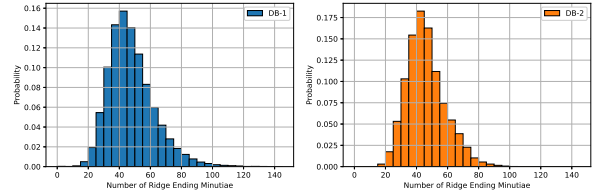


Fig. 3. Number of ridge ending minutiae for DB-1 (Left) and DB-2 (Right)

5. CONCLUSIONS

In this work, we developed a progressive growth-based fingerprint synthesis model capable of generating 512×512 pixels, synthetic, plain, impression fingerprints which are diverse, unique, and high-fidelity. Our results show that progressive growth-based GAN models do not suffer from the shortcomings associated with the previously proposed IWGAN-based fingerprint synthesis models. Our results confirm that the synthetic fingerprints generated by the CFG closely resemble the fingerprints in the training dataset in terms of fingerprint minutiae configurations, ridge-valley structure, and quality while not revealing the identities presented in the training dataset. Finally, we make both the CFG and a dataset of synthetically generated samples publicly available to allow other researchers to continue this work.

6. FUTURE WORK

In future work, we aim to improve the CFG by integrating quality metrics into the loss function. Additionally, we aim to provide our model with identity, finger, and quality labels. This would allow us to control such factors during the synthesis process. Finally, we aim to further investigate the usefulness of synthetic fingerprints in training and evaluating PAD models.

7. REFERENCES

- [1] Kalaivani Sundararajan and Damon L. Woodard, “Deep learning for biometrics: A survey,” *ACM Computing Surveys (CSUR)*, vol. 51, no. 3, pp. 1–34, 2018, Publisher: ACM New York, NY, USA.
- [2] “NIST Special Database Catalog,” Available at <https://www.nist.gov/srd/shop/special-database-catalog> (2021-01-05).
- [3] Raffaele Cappelli, D. Maio, and D. Maltoni, “SFinGe: an approach to synthetic fingerprint generation,” in *International Workshop on Biometric Technologies (BT2004)*, 2004, pp. 147–154.
- [4] Peter Johnson, Fang Hua, and Stephanie Schuckers, “Texture modeling for synthetic fingerprint generation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2013, pp. 154–159.
- [5] Qijun Zhao, Anil K. Jain, Nicholas G. Paulter, and Melissa Taylor, “Fingerprint image synthesis based on statistical feature models,” in *2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. 2012, pp. 23–30, IEEE.
- [6] Carsten Gottschlich and Stephan Huckemann, “Separating the real from the synthetic: minutiae histograms as fingerprints of fingerprints,” *IET Biometrics*, vol. 3, no. 4, pp. 291–301, 2014, Publisher: IET.
- [7] Philip Bontrager, Aditi Roy, Julian Togelius, and Nasir Memon, “DeepMasterPrint: Fingerprint Spoofing via Latent Variable Evolution,” *arXiv preprint arXiv:1705.07386*, 2017.
- [8] Kai Cao and Anil Jain, “Fingerprint synthesis: Evaluating fingerprint search at scale,” in *2018 International Conference on Biometrics (ICB)*. 2018, pp. 31–38, IEEE.
- [9] Shervin Minaee and Amirali Abdolrashidi, “FingerGAN: Generating realistic fingerprint images using connectivity imposed GAN,” *arXiv preprint arXiv:1812.10482*, 2018.
- [10] Masud An-Nur Islam Fahim and Ho Yub Jung, “A Lightweight GAN Network for Large Scale Fingerprint Generation,” *IEEE Access*, vol. 8, pp. 92918–92928, 2020, Publisher: IEEE.
- [11] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C. Courville, “Improved training of wasserstein gans,” in *Advances in Neural Information Processing Systems*, 2017, pp. 5767–5777.
- [12] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen, “Progressive growing of gans for improved quality, stability, and variation,” *arXiv preprint arXiv:1710.10196*, 2017.
- [13] Tero Karras, Samuli Laine, and Timo Aila, “A style-based generator architecture for generative adversarial networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2019, pp. 4401–4410.
- [14] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila, “Analyzing and improving the image quality of stylegan,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 8110–8119.
- [15] Deng Huang, Theodore T. Allen, William I. Notz, and Ning Zeng, “Global optimization of stochastic black-box systems via sequential kriging meta-models,” *Journal of global optimization*, vol. 34, no. 3, pp. 441–466, 2006.
- [16] Kenneth Ko, “Users Guide to Export Controlled Distribution of NIST Biometric Image Software (NBIS-EC),” July 2007, Last Modified: 2018-11-10T10:11:05:00.
- [17] E. Tabassi, “NFIQ 2.0: NIST fingerprint image quality,” in *NIS-TIR 8034*, 2016.
- [18] Martin Aastrup Olsen, Vladimír Šmida, and Christoph Busch, “Finger image quality assessment features—definitions and evaluation,” *IET Biometrics*, vol. 5, no. 2, pp. 47–64, 2016, Publisher: IET.
- [19] Barret Zoph, Vijay Vasudevan, Jonathon Shlens, and Quoc V. Le, “Learning transferable architectures for scalable image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8697–8710.
- [20] Diederik P. Kingma and Jimmy Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [21] Kate Keahey, Jason Anderson, Zhuo Zhen, Pierre Riteau, Paul Ruth, Dan Stanzione, Mert Cevik, Jacob Colleran, Haryadi S. Gunawi, and Cody Hammock, “Lessons learned from the chameleon testbed,” in *Proceedings of the 2020 USENIX Annual Technical Conference (USENIX ATC ’20)*. 2020, pp. 219–233, USENIX Association.