creative commons
COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

조선대학교
CHOSUN UNIVERSITY

[UCI]I804:24011-200000358618

February 2021
Master's Degree Thesis

A light-weight generative adversarial network for fingerprint patch generation

Graduate School of Chosun University
Department of Computer Engineering
Masud An-Nur Islam Fahim

조선대학교
CHOSUN UNIVERSITY

A light-weight generative adversarial network for fingerprint patch generation

경량 생성 대립 네트워크지문 패치 생성

February 25, 2021

Graduate School of Chosun University

Department of Computer Engineering

Masud An-Nur Islam Fahim

조선대학교
CHOSUN UNIVERSITY

A light-weight generative adversarial network for fingerprint patch generation

Advisor: Prof. Ho Yub Jung, PhD

A thesis submitted in partial fulfillment of the requirements for a master's degree

October, 2020

Graduate School of Chosun University

Department of Computer Engineering

Masud An-Nur Islam Fahim

조선대학교
CHOSUN UNIVERSITY

마수드 안 누르 이슬람 파힘
속사학위논문을 인준함

심사위원장 조선대학교 최우열 교수 (인)

위 원 조선대학교 정호엽 교수 (인)

위 원 조선대학교 강문수 교수 (인)

2020년 11월

조선대학교 대학원

조선대학교
CHOSUN UNIVERSITY

TABLE OF CONTENTS

조선대학교
CHOSUN UNIVERSITY

조선대학교
CHOSUN UNIVERSITY

LIST OF ABBREVIATIONS AND ACRONYMS

CNN	Convolutional Neural Network
GAN	Generative Adversarial Network
MS-SSIM	Multi-Scale Structural Similarity
JSD	Jensen–Shannon Divergence
KLD	Kullback–Leibler Divergence
DCGAN	Deep Convolutional Generative Adversarial Network
WGAN	Wasserstein Generative Adversarial Network
GP	Gradient Penalty
MSE	Mean Square Error
RMS	root mean square
SN	Spectral Normalization
cGAN	Conditional Generative Adversarial Network
Pix2Pix	Pixel 2 pixel
ReLU	Rectified Linear Unit
G	Generator
D	Discriminator
TanH	Tan Hyperbolic
IS	Inception Score
FID	Fréchet Inception Distance

조선대학교
CHOSUN UNIVERSITY

LIST OF FIGURES

조선대학교
CHOSUN UNIVERSITY

조선대학교
CHOSUN UNIVERSITY

조선대학교
CHOSUN UNIVERSITY

LIST OF TABLES

조선대학교
CHOSUN UNIVERSITY

ABSTRACT

A light-weight generative adversarial network for fingerprint patch generation

Masud An-Nur Islam Fahim
Advisor: Prof. Jung, Ho Yub, Ph.D.
Department of Computer Engineering
Graduate School of Chosun University

Generating fingerprint images for biometric purposes is both necessary and challenging. In this thesis, we presented a fingerprint generation approach based on a generative adversarial network [1]. To ensure GAN training stability, we have introduced conditional loss doping that allows a continuous flow of gradients. Our study utilizes a careful combination of a residual network and spectral normalization [2] to generate fingerprints. The proposed average residual connection shows more immunity against vanishing gradients than a simple residual connection. Spectral normalization allows our network to enjoy reduced variance in weight generation, which further stabilizes the training [2]. The proposed scheme uses spectral bounding only in the input and the fully connected layers. Our network synthesized fingerprints up to 256 by 256 in size. We used the multi-scale structural similarity (MS-SSIM) metric [3] for measuring the diversity of the generated samples. Our model has achieved 0.23 MS-SSIM scores for the generated fingerprints. The MS-SSIM score indicates that the proposed scheme is more likely to produce more diverse images and less likely to face mode collapse.

조선대학교
CHOSUN UNIVERSITY

조선대학교
CHOSUN UNIVERSITY

한 글 요 약

경량 생성 대립 네트워크
지문 패치 생성

마수드 안 누르 이슬람 파힘
지도 교수: 정 호엽
학과 : 컴퓨터공학과
조선대학교 대학원

생체 인식 목적으로 지문 이미지를 생성하는 것은 필수적이면서도 도전적입니다. 이 논문에서 우리는 생성 적 대립 네트워크 [1]에 기반한 지문 생성 접근법을 제시했습니다. GAN 훈련 안정성을 보장하기 위해 지속적으로 기울기 흐름을 허용하는 조건부 손실 도핑을 도입했습니다. 우리의 연구는 '레즈넷(ResNet, Residual Network)'과 스펙트럼 정규화 [2]의 신중한 조합을 사용하여 지문을 생성합니다. 제안 된 평균 잔류 연결은 단순한 잔류 연결보다 소실 기울기에 대한 내성이 더 큽니다. 스펙트럼 정규화를 사용하면 네트워크에서 가중치 생성의 분산을 줄일 수 있으므로 훈련 [2]를 더욱 안정화 할 수 있습니다. 제안 된 방식은 입력 계층과 완전 연결 계층에서만 스펙트럼 경계를 사용합니다. 네트워크는 최대 256 x 256 크기의 지문을 합성했습니다. 생성 된 샘플의 다양성을 측정하기 위해 다중 스케일 구조적 유사성 (MS-SSIM) 메트릭 [3]를 사용했습니다. 우리 모델은 생성 된 지문에 대해 0.23 MS-SSIM 점수를 달성했습니다. MS-SSIM 점수는 제안 된 방식이 더 다양한 이미지를 생성 할 가능성이 더 높고 모드 붕괴에 직면 할 가능성이 적음을 나타냅니다.

조선대학교
CHOSUN UNIVERSITY

I. INTRODUCTION

Fingerprints are essential to biometric systems. Many approaches have been proposed to create fingerprints for biometric purposes. Previous studies on fingerprint synthesis have been mostly associated with manual generation from a single or multiple fingerprint base structure. These studies show that using morphological or minutiae point manipulation, we can produce synthetic fingerprints[4]–[7].

A master fingerprint is capable of bypassing a small-scale biometric security system such as those used on smartphones. This MasterPrint invasion is possible due to the input data limitation of these devices. This scenario was introduced by [8], [9], who attempted to emulate the master fingerprint. Their study has shown that a master fingerprint can be obtained by manipulation of original images or through synthesis using the hill-climbing scheme. However, their master fingerprints are visually distinguishable from the original data. These studies [10]–[12] proposed a zero-pole model for synthetic fingerprint generation. However, visually indistinguishable fingerprints are still very difficult to produce. Compared to previous studies, fingerprint generative adversarial network-based studies [13]–[15] have shown great promise. A summary of these studies is presented in table 1.

The GAN [1] is a massive leap in the area of artificial intelligence since it popularizes information synthesis using a neural network. The GAN framework takes a sampled data distribution and produces a synthetic data distribution that closely reflects the input data. Despite introducing an enormous number of possibilities, the GAN framework also has challenges, such as mode collapse, training stability, and a large computational budget. To mitigate these challenges,

조선대학교
CHOSUN UNIVERSITY

Table 1: Previous fingerprint synthesis studies at a glance.

Algorithm	Approaches
Cappelli et al. [9]	Three models were used autonomously. However, the visual distinction between synthetic and real fingerprints is clear. No metric for measuring the similarity between the generated fingerprints was provided.
Johnson et al. [10]	Three models were used autonomously. However, the visual distinction between synthetic and real fingerprints is clear. No metric for measuring the similarity between the generated fingerprints was provided.
Johnson et al. [10]	Three models were used autonomously. However, the visual distinction between synthetic and real fingerprints is clear. No metric for measuring the similarity between the generated fingerprints was provided.
Bontrager et al. [13]	128×128 slices.
Kai et al. [12]	Produces up to 512× 512. Autoencoder-based image generation always approximates images close to the original image
Kai et al. [12]	Produces 128 × 128 patches; produces blurry ridges. Additionally, like all other previous studies, they did not show any diversity analysis

many schemes have been proposed, and all of them are based on the idea of the deep convolutional generative adversarial network (DCGAN) [16]. DCGAN uses deep convolutional layers instead of fully connected layers to generate fake images. It can successfully produce 64 by 64 images, somewhat successful with 128 by 128 size, but fails for larger scale.

Following DCGAN, three significant studies [17]–[19] were reported in 2017. Instead of typical cross-entropy-based loss, Arjovsky et al. [17] uses the Wasserstein distance as the loss function. This particular method enables the network to enjoy a continuous flow of loss values throughout the training in comparison to cross entropy-based loss functions. This method also improves Lipschitz constraint enforcement by proposing weight clipping. Later, Gulrajani et al. [18] improved weight clipping by introducing gradient penalty and it has been adopted by other GAN studies. An alternating gradient update-based scheme [19] has shown convergence when the distribution from the generator and the original distribution are absolutely continuous [20]. A much more stable training strategy was later introduced by Google [21] based on an encoder-decoder-based discriminator. This study employs the idea of the Wasserstein

조선대학교
CHOSUN UNIVERSITY

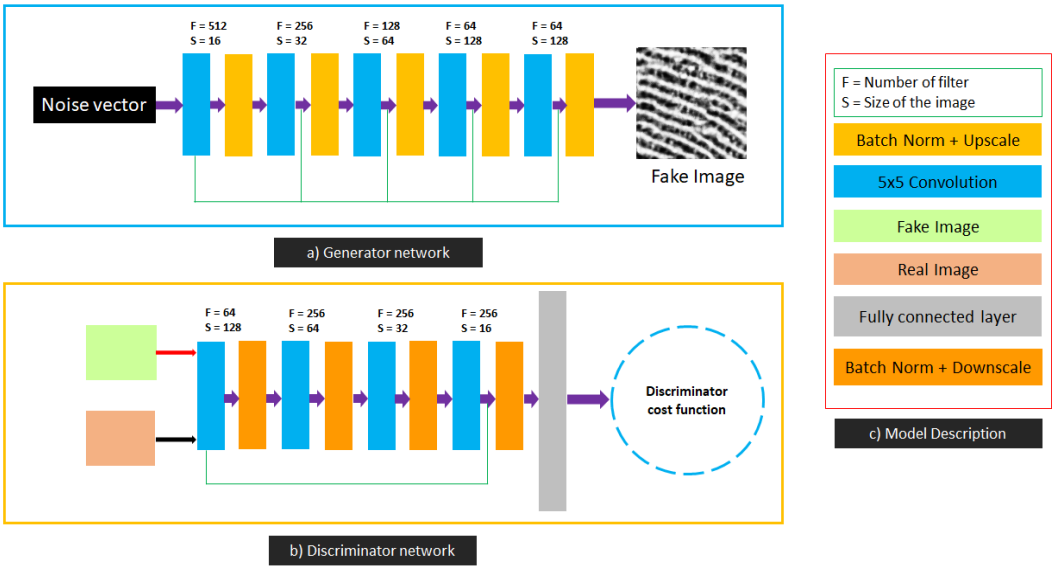

Figure 1: Proposed structure for the generator and the discriminator. a) This schematic diagram is indicating the details of the proposed generator. b) This schematic diagram is indicating the details of the proposed discriminator. c) This section is showing the meaning behind each block as presented in a and b.

distance for constructing their adaptive learning rate dependent loss function. BEGAN [21] studied the training stability issue and successfully produced 128 by 128 images.

Currently, we have a handful of practical advances in GAN training mechanisms. However, the underlying dynamics have yet to be established. From [20], the authors of Dirac-GAN show that if it is possible to examine the eigenvalues of the Jacobian of the associated gradient vector field, local convergence and stability properties can be analyzed. A far-fetched similarity can be shown in the spectral analysis of the generative adversarial network [2]. This study place emphasis on normalizing the weights of the neural network to

조선대학교
CHOSUN UNIVERSITY

stabilize the training procedure. Similar to BEGAN [21], these studies [2], [20] examined the means of stable training to generate 128 by 128 images. Enforcing the Lipschitz continuity may lead to stable training [2]. This intuition is also supported by Odena et al. [22] since they have shown that the stability of the generator is dependent on taming the Jacobian. Another stable training scheme from Google [23] has also supported spectral analysis for the further stabilization of GAN training.

Instead of dealing with a cost function, some research suggests looking at the differences in the number of networks [24]–[27]. They proposed that multiple generators or discriminators can improve GAN stability. One study [24] shows that using the KL divergence and the reverse KL divergence as dual discriminators for GAN can improve training. These studies [25], [26] employ more than two discriminators to produce gradients with low variance, consequently improves the generator. However, this idea increases the training cost dramatically for higher-dimensional images. AdaGAN [28] utilizes the boosting technique to improve training performance. This study [29] uses multiple generators to improve the GAN training scheme. Mixture GAN (MGAN) [30] proposes to reduce the mode collapse situation by training multiple generators to learn the statistics of different data modes. However, this increases training difficulty, and mode collapse remains unsolved.

The goal of this study is to design a GAN system to produce faithful synthetic fingerprint images. Synthetic fingerprints can aid in many research applications. For example, an extensive collection of usable synthetic fingerprint images can help in fingerprint detection, fingerprint classification, fingerprint liveness detection, or data augmentation for deep learning tasks.

Due to hardware, computational, and design complexities, it is hard to

조선대학교
CHOSUN UNIVERSITY

generate images with a higher dimension. We have designed the proposed network in a way that it can easily deal with the above concerns. To ensure stability in the fingerprint generation, we adopted spectral normalization [2]. The presence of skip connection in the generator and the discriminator helps our network to mitigate the vanishing gradient problem. We provided a more detailed analysis in the methodology section.

A. Contributions

Here, we have proposed a GAN scheme that can successfully produce fingerprint patches. The overall contributions of this study can be summarized as follows:

- The proposed lightweight network can generate fingerprint patches for up to 256 by 256 fingerprint images.
- This study introduced loss-doping for overall training stability. Loss doping allows our network to avoid training collapse, which is prevalent in previous GAN studies. Also, we have observed improved convergence for this technique.
- Mode collapse is still an open challenge in GAN research. The proposed network is fairly free from this drawback, with an average MS-SSIM score of 0.23. We have observed that produced images are less likely to be similar to each other. Additionally, the utilization of data augmentation has enabled it to cope with the same image with different appearances.
- Our minimalistic residual-spectral network for fingerprint generation enjoys good stability during the entire training procedure and is less likely to suffer from training collapse.

조선대학교
CHOSUN UNIVERSITY

B. The Research Objectives

There are few studies out there to deal with fingerprint generation. Due to the generative feature, GAN shows superior performance over the handcrafted algorithms. However, the GAN performance is constrained by the nature of loss function, network design, and gradient processing techniques. The proposed study aims to produce fingerprint patches by utilizing GAN. It is usual to face stability challenges with the generative adversarial network. Moreover, no GAN studies ensure free of mode collapsing. Our goal is to design a generative adversarial network that can successfully produce fingerprint patches without facing usual GAN training problems. We aim to provide a GAN training scheme that can aid other GAN studies to skip mode collapse as much as possible while maintaining desired stability.

C. Thesis Layout

The thesis follows this organization. Chapter II presents an overview of the GAN objective function, GAN loss function variants, GAN metrics, and several GAN alternatives. Then in chapter III, we describe the related works for GAN-based image generation and GAN tuning techniques. Next, in Chapter IV, we explain the network architecture and relevant analysis for the proposed loss doping employed to mitigate the training collapse during the training as much as possible. We present sufficient pictorial demonstration and comparison, tables, quantitative analysis for this study in Chapter V. We conclude our proposed approach in Chapter VI.

조선대학교
CHOSUN UNIVERSITY

II. Generative Adversarial Network

A. Generative Adversarial Network

After having information, X, but no proper label, Y, one will utilize unsupervised learning to extract value from it. Unlabeled information is the default state of all information found in nature; it is not until we need to extricate a few causal interface meanings from the data to relate labels to that information. People frequently make this identification in a dataset. To many, this procedure is costly and time expending. In any case, many valuable data can beget from a dataset without having a particular forecast objective in intellect. One exciting property can be to be able to create new tests from a dataset. The generative adversarial network makes it possible to develop new tests comparable to those found in a dataset. It comprises two convolutional neural networks, the generator G and the discriminator D. In together, they create an adversarial system. This adversarial phenomenon's underlying façade is a cyclic operation, where the generator tries to fools the discriminator, and the only priority of the discriminator is to see through the generator's output.

B. Theoretical Formulation

In the desired setting, we have a discriminator that can decide between real or fake data for a given input from a given dataset. The discriminator return the probability score for the input data, necessarily $\mathrm{D} : \mathrm{X} \rightarrow [0, 1]$. If the mapped output is 0, the discriminator is sure about its fake nature. On the other hand, for 1, the discriminator will consider it as genuine. For this discussion, we will denote the discriminator as ϑ_{d}.

조선대학교
CHOSUN UNIVERSITY

On the other hand, the genarator works with pure noise $\mathrm{z} \sim \mathscr{D}_{\mathrm{z}}$,. For the generator input, as an example $\mathscr{D}_{\mathrm{z}} = \mathscr{N}\left(0, \sigma^2 \mathrm{I}\right)$ is a possible choice and produces samples as output, $\mathrm{G} : \mathscr{D}_{\mathrm{z}} \rightarrow \mathscr{X}$.

The noise provides arbitrariness to the generator G. People call the input vector z as the latent vector. There are numerous ways to manipulate the latent vector space. A summary of the latent vector design is discussed later. We denote the generator's parameters as ϑ_{g} when necessary along with the discriminator. To avoid extra complexity, we represent the total parameter of the whole model as ϑ.

Let's set up a condition that will govern the GAN to work according to our objective. For this, it is obvious to design a generator that can convince the discriminator with its output. We seem to maximize the likelihood, $\mathrm{D}(\mathrm{G}(\mathrm{z}))$, that the discriminator guesses that produced tests are from the real distribution. This is often comparable to minimizing $1-\mathrm{D}(\mathrm{G}(\mathrm{z}))$, which can be helpful afterward. On the other hand, we need the discriminator to classify produced tests as fake, which implies maximizing $1-\mathrm{D}(\mathrm{G}(\mathrm{z}))$. So, the formulation of the generator objective is

$$\min_{\mathrm{G}} \max_{\mathrm{D}} \mathbb{E}_{\mathrm{z} \sim \mathscr{D}_*}[1-\mathrm{D}(\mathrm{G}(\mathrm{z}))]$$

This formulation does not work complexities. One typical case is to provide output with a min-max limit of 0 to 1. This bound is not helpful for the overall optimization procedure. A standard measurable approach would require the logarithm of the likelihood and mirrors the log probability definition from insights. From the inferential statistical view, here we are trying to fit a distribution to our data, and to do so, we should select the distribution with the maximum entropy. So, the reformulation of the generator objective function is as

조선대학교
CHOSUN UNIVERSITY

follows:

$$\mathscr{L}_{\mathrm{g}}(\mathrm{G},\mathrm{D}) = \mathbb{E}_{\mathrm{z}\sim\mathscr{D}_{\mathrm{z}}}[\log(1-\mathrm{D}(\mathrm{G}(\mathrm{z})))]$$

And the minimization task for the generator is:

$$\min_{\mathrm{G}} \mathscr{L}_{\mathrm{g}}(\mathrm{G},\mathrm{D})$$

Now, we will settle things for the discriminator. Before that, one may set up the perfect discriminator at the beginning of the training. This is not a desirable choice for the GAN since the generator will yield easily to the discriminator. Instead, setting the discriminator and the generator with poor initialization is more suitable. We let the discriminator learn to distinguish samples from the dataset and become strong at seeing through the generator's deception. We will define the discriminator's objective as

$$\mathscr{L}_{\mathrm{d}}(\mathrm{G},\mathrm{D}) = \mathbb{E}_{\mathrm{x}\sim\mathscr{D}_{\mathrm{dat}}}[\log \mathrm{D}(\mathrm{x})] + \mathbb{E}_{\mathrm{z}\sim\mathscr{D}_{\mathrm{z}}}[\log(1-\mathrm{D}(\mathrm{G}(\mathrm{z})))]$$

which will be maximized throughout the training procedure.

$$\max_{\mathrm{D}} \mathscr{L}_{\mathrm{d}}(\mathrm{G},\mathrm{D})$$

By combining the above objectives, we get the adversarial loss function as

$$\mathscr{L}_{\mathrm{JS}}(\mathrm{D},\mathrm{G}) = \mathbb{E}_{\mathrm{x}\sim\mathscr{D}_{\mathrm{dat}}}[\log \mathrm{D}(\mathrm{x})] + \mathbb{E}_{\mathrm{z}\sim\mathscr{D}_{\mathrm{x}}}[\log(1-\mathrm{D}(\mathrm{G}(\mathrm{z})))]$$

However, this choice of the loss function is not sufficient. Later we will see the relationship between maximizing this over D with the Jensen-Shannon divergence.

C. Loss function variants

There are many ways we can follow to design the loss function of the GAN. We incorporate a constrained outline for the loss function and the regularization

조선대학교
CHOSUN UNIVERSITY

techniques in the following tables. The Non-Immersing GAN (NS GAN) is comparable to the JS GAN, which differs where the generator tries to maximize its objective, not minimize. It is subsequently not a minimax diversion. To increase this distinction, JS GAN is additionally cited as MM (MiniMax) GAN. BEGAN utilizes the autoencoder and the decoder to improve the performance of GAN while compromising the diversity. GAN regularization is also as critical as the loss function. This necessarily includes discriminator supervision for the continuity of the training.

Table 2: Different types of loss functions for GAN.

Name	Discriminator Objective	Generator Objective
JS GAN	$\mathscr{L}_d = -\mathbb{E}_{x\sim\mathscr{D}_{dat}}[\log D(x)] - \mathbb{E}_{z\sim\mathscr{D}_z}[\log(1 - D(G(z))]$	$\mathscr{L}_g = \mathbb{E}_{z\sim\mathscr{D}_z}[\log(1 - D(G(z))]$
NS GAN	$\mathscr{L}_d = -\mathbb{E}_{x\sim\mathscr{D}_{dat}}[\log D(x)] - \mathbb{E}_{z\sim\mathscr{D}_z}[\log(1 - D(G(z))]$	$\mathscr{L}_g = -\mathbb{E}_{z\sim\mathscr{D}_z}[\log(1 - D(G(z))]$
LS GAN	$\mathscr{L}_d = -\mathbb{E}_{x\sim\mathscr{D}_{dat}}[(D(x)-1)^2] - \mathbb{E}_{z\sim\mathscr{D}_z}[(1-D(G(z))^2]$	$\mathscr{L}_g = -\mathbb{E}_{z\sim\mathscr{D}_z}[(1-D(G(z))^2]$
WGAN	$\mathscr{L}_d = \mathbb{E}_{x\sim\mathscr{D}_{dat}}[D(x)] - \mathbb{E}_{z\sim\mathscr{D}_z}[D(G(z))]$	$\mathscr{L}_g = -\mathbb{E}_{z\sim\mathscr{D}_z}[D(G(z))]$
BEGAN	$\mathscr{L}_d = \mathbb{E}_{x\sim\mathscr{D}_{dat}}[\lVert x-AE(x)\rVert_1] - k_t\mathbb{E}_{z\sim\mathscr{D}_z}[\lVert G(z)-AE(G(z))\rVert_1]$	$\mathscr{L}_g = \mathbb{E}_{z\sim\mathscr{D}_z}[\lVert G(z)-AE(G(z))\rVert_1]$

To analyze the overall performance of a given generative adversarial network, researchers usually use the Inception score and the Fréchet Inception Distance. Two primary intuition results in the Inception score. The conditional label distribution of samples containing meaningful objects should have a low entropy, and the samples should have a high amount of variability. To calculate this score, we need the Inception network, which was trained with the ImageNet data set. The inception score is as follows:

$$\mathrm{IS}(G) = \exp\left(\mathbb{E}_{x\sim\mathscr{D}_g}\left[D_{KL}(p(y \mid x)\|p(y))\right]\right)$$

조선대학교
CHOSUN UNIVERSITY

Table 3: Different varieties of GAN regularizers. For L1 and L2 we are in a supervised setting.

Name	Discriminator Objective
L1	$\mathbb{E}_{(\mathrm{x},\mathrm{y})\sim\mathscr{D}_{\text{data}},\mathrm{z}\sim\mathscr{D}_{\mathrm{z}}}\left[\|\mathrm{y}-\mathrm{G}(\mathrm{x},\mathrm{z})\|_1\right]$
L2	$\mathbb{E}_{(\mathrm{x},\mathrm{y})\sim\mathscr{D}_{\text{data}},\mathrm{z}\sim\mathscr{D}_{\mathrm{z}}}\left[\|\mathrm{y}-\mathrm{G}(\mathrm{x},\mathrm{z})\|_2\right]$
WGAN-GP	$\mathbb{E}_{\tilde{\mathrm{x}}\sim\mathscr{D}_{\mathrm{g}},\mathrm{x}\sim\mathscr{D}_{\text{data}}}\left[(\|\nabla \mathrm{D}(\alpha \mathrm{x}+(1-\alpha)\tilde{\mathrm{x}})\|_2-1)^2\right]$
DRAGAN	$\mathbb{E}_{\tilde{\mathrm{x}}\sim\mathscr{D}_{\text{data}}+\mathcal{N}(0,\mathrm{c})}\left[(\|\nabla \mathrm{D}(\tilde{\mathrm{x}})\|_2-1)^2\right]$
InfoGAN	$\mathbb{E}_{\mathrm{x}\sim\mathscr{D}_{\text{data}},\mathrm{c}'\sim\mathrm{p}(\mathrm{c})}\left[\log \mathrm{Q}\left(\mathrm{c}' \mid \mathrm{x}\right)\right]+\mathrm{H}(\mathrm{c})$

However, this formula is not satisfactory enough to fulfill the criteria of the metric. This score shows the record to correlate with human evaluations. The Fréchet Inception Distance (FID) implants created outputs into the n-dimensional feature space given by a layer of Inception Net. This score assumes the feature space as a multivariate Gaussian distribution. For the given dataset, we can compute the mean and the variance of the generated outputs from the given network. At this point, we have two distributions for one generated output. Hence, the formulation of the Fréchet Inception Distance (FID) is as follows:

$$\mathrm{FID}(\mathrm{x},\mathrm{g})=\|\mu_{\mathrm{x}}-\mu_{\mathrm{g}}\|_2^2+\mathrm{tr}\left(\Sigma_{\mathrm{x}}+\Sigma_{\mathrm{g}}-2\left(\Sigma_{\mathrm{x}}\Sigma_{\mathrm{g}}\right)^{\frac{1}{2}}\right)$$

The FID score is also handy in detecting the intra-class mode dropping. This situation occurs when the generator shows limited distribution production performance for each class. In this scenario, one might get a good Inception distance score but will attain poor FID. However, both of them do not show any glimpse of accountability with the over-fitting issue. When a GAN can embed the

조선대학교
CHOSUN UNIVERSITY

training data into its space entirely but fails to produce new samples, GAN never achieves perfect ID score and FID for that dataset.

D. GAN's representative variants

Here, we have tried to note some of the alternatives to the original objective function. In the original GAN, the total optimization procedure tries to minimize the criteria of the min-max game. Since that reduces the overall training flexibility and leads to more complexities, several studies develop new training strategies. The following subsections contain a brief discussion upon some variants of the GAN objective function.

1. InfoGAN

InfoGAN utilizes a single unstructured noise vector z to decompose it into two parts: an incompressible noise z and the latent code c and features the semantic structure for real data distribution. InfoGAN wishes to solve the following equation :

$$\min_{\mathrm{G}} \max_{\mathrm{D}} \mathrm{V_I}(\mathrm{D},\mathrm{G}) = \mathrm{V}(\mathrm{D},\mathrm{G}) - \lambda \mathrm{I}(\mathrm{c};\mathrm{G}(\mathrm{z},\mathrm{c}))$$

Here, $\mathrm{V}(\mathrm{D},\mathrm{G})$ is represented as the objective function of original GAN, $\mathrm{G}(\mathrm{z},\mathrm{c})$, and λ is defined as the generated sample and tunable regularization parameter respectively. I is mutual information. InfoGAN wishes to maximize $\mathrm{I}(\mathrm{c};\mathrm{G}(\mathrm{z},\mathrm{c}))$ by maximizing the mutual information I between c and $\mathrm{G}(\mathrm{z},\mathrm{c})$ to make c contain more reliable and meaningful as much as possible according to the real sample features.

However, the GAN faces some difficulties in direct optimization at the time

조선대학교
CHOSUN UNIVERSITY

of its access to the posterior $\mathrm{P}(\mathrm{c}|\mathrm{x})$. InfoGAN finalizes its objective function as :

$$\min_{\mathrm{G}} \max_{\mathrm{D}} \mathrm{V_I}(\mathrm{D},\mathrm{G}) = \mathrm{V}(\mathrm{D},\mathrm{G}) - \lambda \mathrm{L_I}(\mathrm{c};\mathrm{Q})$$

where $\mathrm{L_I}(\mathrm{c};\mathrm{Q})$ is the lower bound of $\mathrm{I}(\mathrm{c};\mathrm{G}(\mathrm{z},\mathrm{c}))$ Lately, several variants of InfoGAN have been proposed for improvements such as semi-supervised InfoGAN and causal InfoGAN.

2. Conditional GANs (cGANs)

CGANs are the extended version of the GANs where both discriminator and genrator are conditioned on some additional information. Conditional GANs represents the objective function as follows:

$$\min_{\mathrm{G}} \max_{\mathrm{D}} \mathrm{V}(\mathrm{D},\mathrm{G}) = \mathrm{E}_{\mathrm{x}\sim \mathrm{p_{data}(x)}}[\log \mathrm{D}(\mathrm{x} \mid \mathrm{y})]$$

$$+\mathrm{E}_{\mathrm{z}\sim \mathrm{p_z(z)}}[\log(1-\mathrm{D}(\mathrm{G}(\mathrm{z} \mid \mathrm{y})))]$$

If we compare with the two objective functions of the InfoGAN and Conditional-GANs, the two generators are similar but the latent code c of the InfoGAN is unknown and that can be discovered through training. Moreover, InfoGAN has an extra network Q.

cGANs can generate samples conditioning on class labels, boundary box, tests, and key points. Based on the cGANs, stacked generative adversarial network (SGAN), a multi-task perspective composition of two cGANs can conduct text to photo-realistic image synthesis. Different applications of cGANs such as convolutional face generation, image translation, face aging, and outdoor image synthesis contain specific scenery attributes, natural image descriptions, and 3-D-aware scene manipulation. Robust cGANs augments the generator

조선대학교
CHOSUN UNIVERSITY

through an unsupervised pathway to promote the outputs o the generator to span the target, which is full of intense noise. Thekumparampil et al.[16] also maintained the robustness of conditional GANs to noisy labels. Conditional CycleGAN [16] maintains cyclic consistency. Considering the mode collapse issue of cans, Mao et al. proposed Mode Seeking GANs (MSGANs) with a simple, effective regularization term to fix it.

The discriminator equation of the original GANs can be likely as follows:

$$\mathrm{L} = \mathrm{E}\left[\log \mathrm{P}\left(\mathrm{S} = \text{real} \mid \mathrm{X}_{\text{real}}\right)\right] + \mathrm{E}\left[\log\left(\mathrm{P}\left(\mathrm{S} = \text{fake} \mid \mathrm{X}_{\text{fake}}\right)\right)\right]$$

The auxiliary classifier GAN (AC-GAN) consists of two different parts; such as the loglikelihood of the correct source $\mathrm{L_s}$ and the loglikelihood of the correct class label, $\mathrm{L_c}$. $\mathrm{L_s}$ is equivalent to L and $\mathrm{L_c}$ can be defined as follows:

$$\mathrm{L_C} = \mathrm{E}\left[\log \mathrm{P}\left(\mathrm{C} = \mathrm{c} \mid \mathrm{X}_{\text{real}}\right)\right] + \mathrm{E}\left[\log\left(\mathrm{P}\left(\mathrm{C} = \mathrm{c} \mid \mathrm{X}_{\text{fake}}\right)\right)\right]$$

AC-GAN is the first variant of GANs, where the discriminator and the generator maximize $\mathrm{L_c} + \mathrm{L_s}$ and $\mathrm{L_c} - \mathrm{L_s}$ respectively to produce recognizable examples of all the ImageNet classes.

Another special type of cGANs related software called pix2pix ,which performs sparse regularization for image-to-image translation. The generator of pix2pix learns mapping from the input image y to the output image $\mathrm{G(y)}$. The objective function of cGANs can be expressed as follows:

$$\mathrm{L_{cGANs}(D,G)} = \mathrm{E_{x,y}}[\log \mathrm{D(x,y)}] + \mathrm{E_y}[\log(1 - \mathrm{D(y,G(y))})]$$

조선대학교
CHOSUN UNIVERSITY

Furthermore, l_1 distance is used:

$$L_{l_1}(G) = E_{x,y}\left[\|x - G(y)\|_1\right]$$

The final objective of C-GAN is

$$L_{cGANs}(D,G) + \lambda L_{l_1}(G)$$

Following pix2pix, pix2pix-HD features matching loss for high-resolution images and semantic manipulation using cGANs . The learning problem of the discriminators of this pix2pix-HD can be formulated as follows:

$$\min_{G} \max_{D_1,D_2,D_3} \sum_{k=1,2,3} L_{GAN}(G, D_k)$$

The training set is given as a set of pairs of corresponding images $(s_i; x_i)$, where x_i is a natural photo and s_i is a corresponding semantic label map. (G, D_k) is the feature extractor of the discriminator D_k. If $L_{FM}(G, D_k)$ is the feature matching loss of the ith layer of the discriminator D_k can be expressed as follows:

$$L_{FM}(G, D_k) =$$

$$E_{(s,x)} \sum_{i=1}^{T} \frac{1}{N_i} \left[\left\| D_k^{(i)}(s,x) - D_k^{(i)}(s, G(s)) \right\|_1 \right]$$

where N_i is the number of elements in each layer and T denotes the total number of layers. The final objective function of is

$$\min_{G} \max_{D_1,D_2,D_3} \sum_{k=1,2,3} \left(L_{GAN}(G, D_k) + \lambda L_{FM}(G, D_k) \right)$$

3. CycleGAN

Computer vision problems require the image to image translation where the motto is to learn the mapping between the input and output images through

조선대학교
CHOSUN UNIVERSITY

training over a set of the dataset. Test data are also required for this purpose for proper prediction, which was effectively solved by the Cycle-consistent GANs (CycleGAN). CycleGAN trained the unpaired dataset and proved the cycle-consistency in an upper bound of the conditional entropy. Lately, DiscoGan has been proposed with a similar idea. After that, DualGAN and Wassesterian GAN use the loss format rather than the sigmoid cross-entropy loss, which was used in CycleGAN.

4. f-GAN

f-GAN is a special area of the original GAN, uses the Kullback-Leibler divergence, which can measure the difference between two given probability distributions. For a large class of assorted divergence, this can be called as f– divergence. If P and Q are the two given probabilities, p and q are the absolute continuous density function, then f– divergence can be defined as follows:

$$\mathrm{D_f(P\|Q)} = \int_{\mathrm{X}} \mathrm{q(x)f}\left(\frac{\mathrm{p(x)}}{\mathrm{q(x)}}\right)\mathrm{dx}$$

Lately, f-GAN has been extended to quantitatively evaluated GANs with divergence, which is more efficient for training. For accuracy upgrade, f– divergence has been directly minimized in the generator step. As a result, this extension to f-GAN can predict the distributions' ratio between the realistic and generalized data in the discriminator step.

조선대학교
CHOSUN UNIVERSITY

III. Related Works

The generation of high-resolution images has attracted much attention in recent years. Even though the DCGAN architecture enables the production of higher resolution images, it needs several modifications; it is unable to ensure fidelity and stability in image generation with higher dimensionality. These problems were later addressed in other studies, and many remedies have been proposed. StackGAN [31] and StackGAN++ [32] have tried to solve these problems gradually. In the earlier version [31], they used a two-stage GAN strategy to achieve higher resolution images. Using the attention mechanism, self-attention GAN [33] has been proposed.

ID-GAN [34] uses variational auto-encoder (VAE) to distill latent distribution for GAN training and produced high dimensional images with the help of 3 networks. RAGAN [34] produces high-quality images by decreasing the probability of fake data to be recognized as real for the generator. MSG-GAN [35] allows gradient flow between the generator and the discriminator, which results in 1k resolution images. This study [36] utilizes domain translation from semantic label maps to produce crisp HD cityscape images. AE-GAN[3] combines WGAN and VAE to create stable, high-resolution photos. COCO-GAN [37] generates state-of-the-art images by utilizing spatial information as the constraint for the generator. The semantic bottleneck network combines progressive semantic generation network and segmentation-to-image synthesis network to produce 5k images. BigGAN [23], Progressive GAN [38], StyleGAN [39] these state-of-the-art methods provide means for large scale($\geq$ 512 by 512) image generation.

On the contrary, super-resolution GAN [40] can provide computational-friendly support to produce a high-resolution image. G-GANISR [41] improves

조선대학교
CHOSUN UNIVERSITY

the GAN performance for super-resolution by utilizing the least square loss function. Dual generative adversarial network [42] uses two generators to enhance the robustness of the network and successfully produces super-resolution images.

Very little has been done regarding studies on producing fingerprints compared to other GAN studies. [14] This study proposed Wasserstein distance-based GAN for fingerprint generation. Usually, GAN-based systems have advantages in producing sharper images than those produced by autoencoder-based schemes. Despite this trend, this scheme [14] seems to produce blurrier images. Additionally, compared to the original samples, the ridges presented in the images are more likely to be noisy. FingerGAN [15] proposed a DCGAN-based scheme that emphasizes the TV loss as an extension to the traditional loss for the DCGAN. They have produced 512 by 512 images. Even with this large-scale synthesis, images seem to be fuzzy in comparison with the original images. The stability power of an autoencoder combined with WGAN [13] provides a good way to implement fingerprint generation compared to [14], [15]. They [13] have produced 512 by 512 images and have claimed to produce millions of samples in one day. In contrast, they have presented very little pictorial representations. Moreover, none of these studies [13]–[15] have presented a diversity analysis. We can summarize the remaining challenges in previous studies as follows:

1) Their schemes rely on Gabor filtering, and AM-FM modeling produces a visually different image that appears synthetic when juxtaposed with real fingerprints.

2) Minutiae-based modeling is independent of minutiae formation. This results in poor pattern generation for fingerprint synthesis.

조선대학교
CHOSUN UNIVERSITY

3) Due to independent minutiae sampling, the generated ridges are very unrealistic.

4) Additionally, they cannot produce random realistic looking patterns, and the gaps between ridges are constant in the generated images [13]–[15].

In this study, we aim to produce fingerprint images that are free from the above traits. Additionally, we hope to produce faithful fingerprint images and least likely to face usual GAN training challenges.

조선대학교
CHOSUN UNIVERSITY

IV. Methodology

This section will cover our reasoning regarding the proposed training scheme for fingerprint generation. A theoretical framework for spectral normalization-based training, training without spectral normalization, and the intuition behind the proposed loss function will form the bulk of this section.

In the adversarial training setup, we need a generator G and a discriminator D. The generator produces fake images to outperform the classification performance of the discriminator. Intuitively, we can consider a value function V for both of the networks to express the GAN formulation theory [1]. The original formulation of the generative adversarial network can be given by:

$$\min_{\mathrm{G}} \max_{\mathrm{D}} \mathrm{V}(\mathrm{G},\mathrm{D}) \tag{1}$$

Throughout the entire training, the generator G is trained to minimize its cost, while the discriminator D is trained to maximize its cost. For this, let x denote the sample data and z denote the noise data for the generator. Now, $\mathrm{p_G}$ is the distribution from the generator , and $\mathrm{q_{data}}$ is the distribution over x. Then, the conventional formulation of equation 1 can be stated as follows:

$$\mathrm{V}(\mathrm{G},\mathrm{D}) = \mathrm{E}_{\mathrm{x}\sim\mathrm{q_{data}}}[\log(\mathrm{D}(\mathrm{x}))] + \mathrm{E}_{\mathrm{z}\sim\mathrm{p_G}}[\log(1-\mathrm{D}(\mathrm{G}(\mathrm{z})))] \tag{2}$$

In equation 2, we ensure that the discriminator is accurate over the real data by maximizing $\mathrm{E}_{\mathrm{x}\sim\mathrm{q_{data}}}[\log(\mathrm{D}(\mathrm{x}))]$. On the other hand, the generator G is producing data $\mathrm{G}(\mathrm{z})$ from the noise data z. The goal of the generator is to minimize $\mathrm{E}_{\mathrm{z}\sim\mathrm{p_G}}[\log(1-\mathrm{D}(\mathrm{G}(\mathrm{z})))]$ by producing high quality fake data $\mathrm{G}(\mathrm{z})$.

Let us define the training parameters for the discriminator and the generator as $\Omega_{\mathrm{D}}, \Omega_{\mathrm{G}}$. The discriminator maps the incoming data distribution in the network

조선대학교
CHOSUN UNIVERSITY

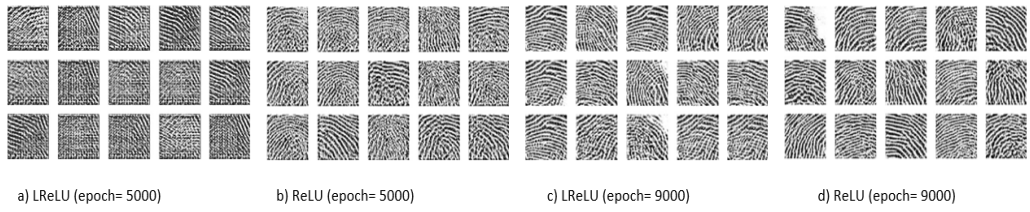

Figure 2: a) 128 by 128 patch generation using Leaky ReLU at epoch=5000 . b) 128 by 128 patch generation using ReLU at epoch=5000. c) 128 by 128 patch generation using Leaky ReLU at epoch=9000. d) 128 by 128 patch generation using ReLU at epoch=9000. From this figure, sharpness and clarity achieved with ReLU activation is clearly visible.

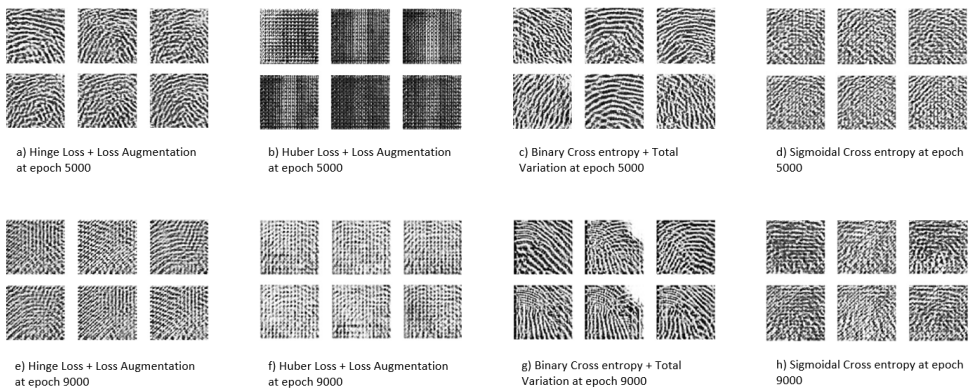

Figure 3: The performance of the loss functions, as mentioned in table 2, is present here. Our whole training period is consists of 9000 epochs. Binary Cross entropy + Total Variation loss function seems superior compared to other loss functions but shows inconsistency in training. From figure (b, d, f, h), Sigmoidal cross-entropy, and Huber Loss + loss augmentation was unable to produce meaningful structure over the whole training period. Hinge Loss + loss augmentation is somewhat successful in the first half of the training, and later it degrades the performance of the GAN.

조선대학교
CHOSUN UNIVERSITY

into a set of probabilities. These probabilities indicate which sample comes from the fake and real distributions. If $\mathrm{P}(\mathrm{x_i};\Omega_\mathrm{D})$ denotes the probability for the real data distribution that the discriminator classifies as real data and $\mathrm{P}(\mathrm{G}(\mathrm{z_i};\Omega_\mathrm{D})$ denotes the probability for the generated distributions that the discriminator classifies as real data, then we can set up the cost function for our discriminator and generator as stated in equations 3 and 4.

$$\mathscr{L}_\mathscr{D} = -\frac{1}{\mathrm{m}}\sum_{\mathrm{i}=1}^{\mathrm{m}}[\log \mathrm{P}(\mathrm{x_i};\Omega_\mathrm{D})] + [\log(1-\mathrm{P}(\mathrm{G}(\mathrm{z_i};\Omega_\mathrm{G});\Omega_\mathrm{D}))] \tag{3}$$

$$\mathscr{L}_\mathscr{G} = -\frac{1}{\mathrm{m}}\sum_{\mathrm{i}=1}^{\mathrm{m}}[\log(1-\mathrm{P}(\mathrm{G}(\mathrm{z_i};\Omega_\mathrm{G});\Omega_\mathrm{D}))] \tag{4}$$

Here, $\mathscr{L}_\mathscr{D}$ and $\mathscr{L}_\mathscr{G}$ denote the cost function for the discriminator D and the generator G, respectively. The data distribution for the discriminator is represented by $\mathrm{x_i}$. $\mathrm{G}(\mathrm{z_i})$ represents the data distribution generated by the generator throughout the training, and $\mathrm{z_i}$ stands for the noise vector.

One important concept we should keep in mind is that the generator acts as a black box. Thus, weights generated for the generator are not independent of the discriminator. This means that if the discriminator is not strong enough to classify fake images, we will see that the generated distributions are poor in terms of fidelity. This concept brings about the idea of an optimal discriminator [2], [17], [18]. We can find out the theoretical representation for the optimal discriminator $\mathscr{D}^*_\mathscr{G}(\mathrm{x})$ by fixing the generator. If $\mathscr{S}$ denotes the sigmoid function, then the optimal discriminator can be formulated as follows:

$$\mathscr{D}^*_\mathscr{G}(\mathrm{x}) = \frac{\mathrm{q_{data}}(\mathrm{x})}{\mathrm{q_{data}}(\mathrm{x}) + \mathrm{p_G}(\mathrm{x})} = \mathscr{S}(\lambda(\mathrm{x})) \tag{5}$$

Before explaining the further, first we will inspect the relationship between GAN and the KullbackLeibler (KL) divergence and Jensen-Shannon (JS) divergence.

조선대학교
CHOSUN UNIVERSITY

From the equation (1) we can see min-max game as follows:

$$C(G) = \max_D V(D,G)$$

$$= E_{x\sim p_{data}}[\log D^*_G(x)] + E_{z\sim p_z}[\log(1 - D^*_G(G(z)))]$$

$$= E_{x\sim p_{data}}[\log D^*_G(x)] + E_{x\sim p_g}[\log(1 - D^*_G(x))]$$

$$= E_{x\sim p_{data}}\left[\log \frac{p_{data}(x)}{\frac{1}{2}(p_{data}(x)+p_g(x))}\right] + E_{x\sim p_g}\left[\frac{p_g(x)}{\frac{1}{2}(p_{data}(x)+p_g(x))}\right] - 2\log 2$$

For two given probabilistic distributions $p(x)$ and $q(x)$, we can write the KullbackLeibler (KL) divergence and Jensen-Shannon (JS) divergence as follows:

$$KL(p\|q) = \int p(x)\log\frac{p(x)}{q(x)}dx$$

$$JS(p\|q) = \frac{1}{2}KL\left(p\|\frac{p+q}{2}\right) + \frac{1}{2}KL\left(q\|\frac{p+q}{2}\right)$$

Therefore, above mentioned min-max formulation is:

$$C(G) = KL\left(p_{data}\|\frac{p_{data}+p_g}{2}\right) + KL\left(p_g\|\frac{p_{data}+p_g}{2}\right) - 2\log 2$$

$$= 2JS\left(p_{data}\|p_g\right) - 2\log 2$$

From this, we can infer that, the objective function of GAN is related to the KL. divergence and JS divergence.

From the min-max game formulation, its possible face a situation where the generator lacks in gradient severely. Informally, in the early stage of the training, G is poor in suitable weights and the generated samples are easily distinguishable from the input data. This enables the discriminator D to reject the generated samples with profound confidence. This scenario leads to saturate the $\log(1-D(G(z)))$. it is possible for us to train the generator in a way, where it

조선대학교
CHOSUN UNIVERSITY

will maximize $\log(D(G(z)))$ instead minimizing $\log(1-D(G(z)))$. Then the cost function for the generator will be as follows:

$$J^{(G)} = E_{z\sim p_{\mathbf{z}}(z)}[-\log(D(G(z)))]$$

$$= E_{x\sim p_g}[-\log(D(x))]$$

Above mentioned objective function may results in fixed point of dynamics but will supply larger gradients in comparison in the early stage of the training. The non-saturating game is not strongly rooted by theory. Additionally, the non-saturating game faces other problems such as unstable numerical gradient for training G. With optimal D^*_G, we have

$$E_{x\sim p_y}\left[-\log\left(D^*_G(x)\right)\right] + E_{x\sim p_q}\left[\log\left(1-D^*_G(x)\right)\right]$$

$$= E_{x\sim p_g}\left[\log\frac{\left(1-D^*_g(x)\right)}{D^{(x)}_G}\right]$$

$$= E_{x\sim p_y}\left[\log\frac{P_g(x)}{p_{data}(x)}\right] = KL\left(p_g \| p_{data}\right)$$

Therefore, $E_{x\sim p_g}\left[-\log\left(D^*_G(x)\right)\right]$ can be written as like noted below:

$$E_{x\sim p_g}\left[-\log\left(D^*_G(x)\right)\right]$$

$$= KL\left(p_g \| p_{data}\right) - E_{x\sim p_g}\left[\log\left(1-D^*_G(x)\right)\right]$$

From the previous derivation of the min-max game, we can bring the JS divergence here:

$$E_{x\sim p_{data}}\left[\log D^*_G(x)\right] + E_{x\sim p_g}\left[\log\left(1-D^*_G(x)\right)\right]$$

$$= 2JS\left(p_{data} \| p_g\right) - 2\log 2$$

조선대학교
CHOSUN UNIVERSITY

Therefore, $E_{x\sim p_g}[\log(1-D_G^*(x))]$ equals

$$E_{x\sim p_g}[\log(1-D_G^*(x))]$$

$$= 2JS\left(p_{data}\|p_g\right) - 2\log 2 - E_{x\sim p_{data}}[\log D_G^*(x)]$$

By substituting $E_{x\sim p_g}[\log(1-D_G^*(x))]$ KL divergence equation, we can reduce it to this extent:

$$E_{x\sim p_g}[-\log(D_G^*(x))]$$

$$= KL\left(p_g\|p_{data}\right) - 2JS\left(p_{data}\|p_g\right) + E_{x\sim p_{data}}[\log D_G^*(x)] + 2\log 2$$

From above equation, we can see that, due to negative sign, the optimization of the alternative generator in the non-saturating game is contradictory in nature. Because, the first term in the equation tries to make the divergence between the generated distribution and the real distribution as small as possible and the second term tries the opposite; make the divergence between these two distributions as large as possible. This phenomena brings the instability in the training dimension. Moreover, KL divergence is not a symmetrical quantity, which can be better understood from the following two examples:

- If $p_{data}(x) \to 0$ and $p_g(x) \to 1$, we have

$$KL\left(p_g\|p_{data}\right) \to +\infty$$

- If $p_{data}(x) \to 1$ and $p_g(x) \to 0$, we have

$$KL\left(p_g\|p_{data}\right) \to 0$$

These errors are the sole contribution of the generator. For them, we need to apply penalization. The first error leads the generator to produce absurd samples,

조선대학교
CHOSUN UNIVERSITY

and the penalization is rather significant. In the second error, the generator produces real distribution; in other words, it returns the training data as the output. For this error, penalization is relatively small. In essence, for the first error, we get the undesirable samples from the generator, and for the second error, we face the limited diversity issue. Due to this, the generator a safe gamble. Instead of making a risky decision like producing a diverse sample, which might lead to undesirable samples, it shows redundancy; in other words, it repeatedly produces safe samples. This situation is known as then mode collapse. However, this is still an open challenge. Going back to the equation (5), we can derive another unsolved situation from here: convergence complexity.

Here, $\lambda(x)$ is as follows:

$$\lambda(x) = \log_{q_{data}}(x) - \log_{p_G}(x) \tag{6}$$

Its derivative can be written as follows:

$$\delta_x \lambda(x) = \frac{1}{q_{data}(x)}(\delta_x q_{data}(x)) - \frac{1}{p_G(x)}(\delta_x p_G(x)) \tag{7}$$

The above equation is unbounded and may even be incomputable [2]. To mitigate this, some bounding conditions are necessary for convergence. This is the reason why researchers [2], [17], [18] have tried to bind the discriminator by K-Lipschitz, which is,

$$\max_{\|f\|_{lip} \leq k} V(G, D) \tag{8}$$

In the above equation, $\| f \|_{lip}$ denotes the smallest value y, where $\| f(m) - f(m') \| / \| m - m' \| \leq y$ for any m, m'. This situation can be addressed by introducing spectral normalization[2]. Spectral normalization stabilizes the training by normalizing the weights in any layer ξ. For a matrix A, the spectral

조선대학교
CHOSUN UNIVERSITY

norm $\sigma(\mathrm{A})$ is defined as follows:

$$\sigma(\mathrm{A}) = \max_{\xi \in \mathbb{R}^{\mathrm{n}}, \xi \neq 0} \frac{\| \mathrm{A}\xi \|_2}{\| \xi \|_2} \tag{9}$$

This equation resembles the maximum among the singular values of the matrix A. For a given vector s and weight matrix w,

$$\sigma(\mathrm{w}_\Omega, \mathrm{s}) \leq \prod_{\mathrm{n}=1}^{\mathrm{N}} \sigma(\mathrm{w}^{\mathrm{n}}) \tag{10}$$

which means that to bound the spectral norm, its enough to normalize the spectral norm of w^{n} in each layer. The theoretical guarantee for the above equation can be obtained from [2].

From the fundamental assumptions of convex optimization, to ensure convexity, a multidimensional linear function has to be Lipschitz continuous. Spectral normalization controls the Lipschitz constant of the discriminator by constraining spectral norm in every single layer [2]. If the previous statement holds, then Lipschitz constant is the largest singular value of the linear function. In other words, it is the spectral norm. If any multidimensional linear function M is K-Lipschitz at 0, then it is K-Lipschitz at any other point. This property simplifies Lipschitz continuity as follows:

$$\| \mathrm{M}\xi \| \leq \mathrm{K} \| \xi \|; \forall \xi \in \mathrm{I} \tag{11}$$

Here, I is the distribution domain. We can write the above equation as follows:

$$\langle \mathrm{M}\xi, \mathrm{M}\xi \rangle \leq \mathrm{K}^2 \langle \xi, \xi \rangle \tag{12}$$

This can be rewritten as follows:

조선대학교
CHOSUN UNIVERSITY

$$\langle (\mathrm{M}^{\mathrm{T}}\mathrm{M} - \mathrm{k}^2)\xi, \xi \rangle \leq 0 \tag{13}$$

Expanding ξ on the basis of eigenvector's orthonormality,

$$\begin{aligned} \langle (\mathrm{M}^{\mathrm{T}}\mathrm{M} - \mathrm{k}^2)\xi, \xi \rangle &= \langle (\mathrm{M}^{\mathrm{T}}\mathrm{M} - \mathrm{k}^2) \sum_{\mathrm{i}} \xi_{\mathrm{i}}, \mathrm{v}_{\mathrm{i}}, \sum_{\mathrm{j}} \xi_{\mathrm{j}}, \mathrm{v}_{\mathrm{j}} \rangle \\ = \sum_{\mathrm{i}} \sum_{\mathrm{j}} \xi_{\mathrm{i}} \xi_{\mathrm{j}} \langle (\mathrm{M}^{\mathrm{T}}\mathrm{M} - \mathrm{k}^2)\mathrm{v}_{\mathrm{i}}, \mathrm{v}_{\mathrm{j}} \rangle &\Rightarrow \sum_{\mathrm{i}} (\mathrm{k}^2 - \lambda_{\mathrm{i}}) \xi^2{}_{\mathrm{i}} \geq 0 \end{aligned} \tag{14}$$

From the above equations, $\mathrm{M}^{\mathrm{T}}\mathrm{M}$ is positive semi-definite, which means all the values of λ_{i} must be non-negative. To ensure this, each of $(\mathrm{k}^2 - \lambda_{\mathrm{i}}) \geq 0$. Since the value of K is the minimum to satisfy the above constraint, then it is obvious from the above relationship that K is the square root of the largest eigenvalue of $\mathrm{M}^{\mathrm{T}}\mathrm{M}$. Hence, the Lipschitz constant of any linear function is its spectral norm. This inherent property justifies the utilization of spectral normalization to ensure convergence [2]. We can only speculate that these properties also carry over to more complex non-linear models.

In our network setup, we used spectral bounding only in the dense and input layers of the discriminator. This makes our spectral norm-dependent setup different from [2], [23], where the authors have used it for every layer. We observed that this bounding has contributed to fingerprint generation by introducing more diversity. Our network achieved the best diversity score with the help of spectral normalization.

A vanishing gradient is a common challenge in GAN-based networks. To mitigate this, we can simply use residual connections. This residual formation has also been utilized by other studies [2], [21], [23] for this purpose. Let us say that the initial layer for our network is X^0. After applying the common activation function $\Psi()$, we can write this as α^1=$\Psi(\mathrm{X}^0)$. The nth layer of the proposed

조선대학교
CHOSUN UNIVERSITY

generator is stated as follows:

$$\alpha^{n} = \Psi^{n}(X^{(n-1)}; \Omega^{n}) + \Psi^{n}(X^{0}; \Omega^{1}) = \alpha^{(n-1)} + \alpha^{1} \tag{15}$$

We experimented with the skip connections using different algebraic operations. We found that simply adding the distributions from α^{1} creates gradient explosion and convergence difficulty. We performed an averaging operation in skip connections instead of simple addition. This ensures the mitigation of the gradient-related problems that were typical of previous GANs. From figure 1, we used the skip connection in every layer for the generator. For the discriminator, we used it only in the last layer. We maintained this structure for training with and without spectral normalization.

For the objective function, we applied some of the modifications to obtain a better result. In GAN theory, there is no incentive for the GAN training scheme to reach a minimal point [20]. Even though it reaches a point where it can successfully produce high-quality images, it does not have the motivation to stop the training. Additionally, it can propagate towards an unstable point. We found this through some experiments with the DCGAN. Similar findings were also observed by researchers from Google [23].

During the training procedure, we observed that a minute random modification in the generator loss can introduce dramatic changes. These changes rely on the degree of loss modification. This observation motivated us to take a different approach to mitigate the static loss scenario. In the traditional setup, when the discriminator wins over the generator, the discriminator and generator loss remain the same for some epochs. This static loss generation continues until the generator produces better images to fool the discriminator. To avoid

조선대학교
CHOSUN UNIVERSITY

Training Scheme: Conditional Loss Doping

1	**Data:** Noise Vector, Real images
2	**Output:** Synthetic fingerprint images
1	**For** e=1 to k do,
2	$\mathrm{G}'_{\mathrm{loss}} = \mathrm{E}_{\mathrm{x}'\sim \mathrm{P_G}}[\log(1-\mathrm{D}(\mathrm{x}'))]$
3	**If** $\mathrm{G}'_{\mathrm{loss}}[\mathrm{e}] = \mathrm{G}'_{\mathrm{loss}}[\mathrm{e}-1]$
4	$\mathrm{G}_{\mathrm{loss}} = (1+\beta) * \mathrm{G}'_{\mathrm{loss}}$
5	**else** $\mathrm{G}_{\mathrm{loss}} = \mathrm{G}'_{\mathrm{loss}}$
6	$\mathrm{D}_{\mathrm{loss}} = \mathrm{E}_{\mathrm{x}\sim \mathrm{P_D}}[\log(\mathrm{D}(\mathrm{x}))] + \mathrm{G}_{\mathrm{loss}}$
7	**End**

this situation, we introduced loss doping. Loss doping implements a minute loss augmentation in the generator loss instead of returning its original loss value. We implement this only if the generator produces the same loss value for two consecutive epochs. The total procedure is presented in the training scheme above. In this way, our network has enjoyed a non-freezing epochs during the entire training time.

Table 4: Summary of loss augmentation in our network architecture.

Loss augmentation choice	Comments
Hinge Loss + Loss Augmentation	With spectral normalization, it works well at the beginning. However, later, it shows performance degradation
Huber Loss + Loss Augmentation	It fails to produce any sort of meaningful image structure
Binary Cross entropy + Total Variation	Severely unstable performance due to rapid total variation
Sigmoidal Cross entropy	Produces very poor-quality images but is better than the abovementioned cases

Intuitively, loss doping acts more like a 'conditional momentum' scenario.

조선대학교
CHOSUN UNIVERSITY

Table 5: Summary of the noise vectors results for proposed GAN.

Noise vector choice	Comments
Uniform	Helps the network implement good training
Normal	As good as uniform distribution
Binomial	As good as uniform distribution
Binomial + Sigmoid	Struggles to generate necessary latent values to continue the training
TanH + Normal	Similar to binomial+sigmoid

Table 6: Summary of the proposed network's properties.

Inclusion	Inclusion	Detail Comments
Activations	ReLU, LeakyReLU, TanH, Sigmoid	Sigmoid replacement of ReLU. LeakyReLU performs worst regarding the generator and discriminator.
Batch Normalization	True	For both the generator and discriminator, it helps the network perform well.
Spectral normalization	True	Only for the discriminator. We have tried it for both in every layer. Least design choice is appreciated in our network.
Kernel size	5	Between 3 and 5, we have seen better results with 3.
Layer Normalization	None	We have tried it for every layer and observed desirable results without it.
Number of filters (generator)	64, 128, 256,512	
Number of filters (generator)	64,64, 128, 256,512	For 256 by 256 size, we have increased one layer.
Number of filters (discriminator)	64,256,256,256	
Number of filters (discriminator)	64,256,256,256	For 256 by 256 size, we did not increase the number of layers

Momentum allows the gradient descent algorithm to escape the saddle point and push the optimization procedure towards the convergence. In this study, loss doping does the same by changing the loss values minutely if and only if consecutive loss value is the same. This conditional doping supposedly changes the trainable weights and helps the optimization to converge faster. Empirically, we have observed faster convergence with the help of loss doping. Figure 10 and figure 11 show the effect of loss doping.

To ensure the best means of loss doping, we experimented with piecewise loss difference, random numbers, and percentile loss. With percentile loss augmentation, we observed stable training without facing training collapse. In

조선대학교
CHOSUN UNIVERSITY

our work, we updated the generator loss G_{loss} using loss doping as follows:

$$G_{loss} = (1+\beta) * G_{loss} \tag{16}$$

Here, β is the doping amount for our network. We maintained the value of $\beta = G_{loss}/10000$ for the training. Due to this, our network enjoyed a variable learning rate during the entire training time. In table 2, the summary of the loss doping is noted.

We did not limit our experiments only to networks for fingerprint generation. For maintaining stable training, we experimented with weight initialization techniques. In terms of convergence speed, we observed that Xavier initialization helps the network converge faster.

We focused on finding the learning capability of the network. To find this, we experimented with the selection of the latent vector. We started with the normal distribution, which ranges from zero to one, and this distribution aids the overall GAN training. We used the spherical uniform distribution [-1,1], whose performance is the same as that of the normal distribution. We also found that the Bernoulli distribution (0,1) aids in stable training. One interesting case for the normal distribution is that when it occurs with a nonzero formation, it is superior to the typical normal distribution.

We also applied the activation function ReLU [43], [44]. and TanH to the latent vectors. Compared to other choices, ReLU with a spherical uniform distribution shows mixed performance. The hyperbolic tangent version of the Bernoulli distribution showed identical performance as the spherical uniform distribution. Although these experiments show good results, the overall performance depends on the network and the objective function. Our choice of network parameter in the end governs the final outcome. For that reason, we

조선대학교
CHOSUN UNIVERSITY

have presented tables 3 and 4 summarizing the compact picture of our total scheme. Figure 2 and 3 is showing the necessary diagrams for the table 2 and 4. The findings in table 2,3, and 4 is obtained through experiments with the proposed network and our input dataset. Among all the evaluated architectures, we observed the best fingerprint generation training with this architecture. Additionally, we produced full-sized fingerprint images and patch images from this structure. Our discriminator has two versions: with and without spectral normalization. For each model choice, we observed the meaningful generation of synthetic images. The results of this study are presented in the next section.

조선대학교
CHOSUN UNIVERSITY

V. Performance Analysis

Our study utilizes two different architectures for fingerprint generation. In our network, spectral normalization was used for 128x128 and 256x256 images. We also produced these images without spectral normalization. For this purpose, we used the LivDet fingerprint dataset [27], [45]–[48]. Images in this dataset come with five different scanners. We used images from the Greenbit scanner, which contains 1000 real fingerprints. We applied rotation, translation, and flipping for data augmentation.

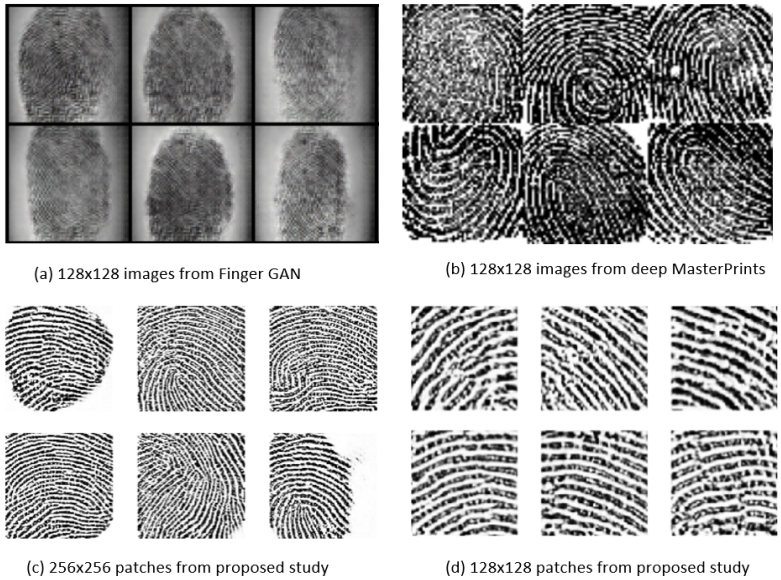

Figure 4: a) Images from Finger GAN [15]. b) Images from deep MasterPrints [14]. c) 256x256 patches from the proposed study. d) 128x128 patches from the proposed study. By visual inspection, ridges are clearer and sharper in the images from this study.

Figure 4 shows the output from [14], [15], and the proposed study. Images produced by [15] are blurry compared to those from the other two studies.

조선대학교
CHOSUN UNIVERSITY

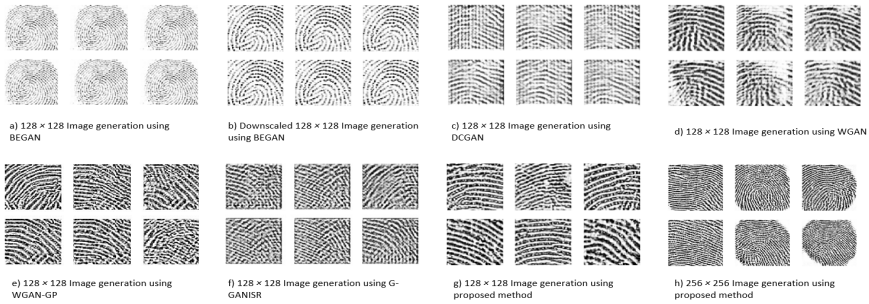

Figure 5: a) 128 × 128 images with BEGAN [21]. b)128 × 128 images with BEGAN. Here, input images are 128 × 128 patches downscaled from 256 × 256 patches. For both of these cases, patches are highly similar to each other. BEGAN produces different images with different initializations. However, the amount of diversity is explicitly negligible in all cases for BEGAN. c) Images from the DCGAN where images are somewhat diverse and not fully developed [16]. d) 128 × 128 Images from WGAN where the images are diverse and not fully developed [17]. e) 128 × 128 Images from WGAN-GP where the images are diverse, not fully developed, and comparatively sharper than WGAN [18]. f) 128 × 128 images from G-GANISR are somewhat recognizable as fingerprints [41]. g)128 × 128 patches from the proposed study. h) 256×256 patches from the proposed study. By visual inspection, ridges are more precise and sharper in the images from the proposed study. Even though BEGAN produced a very stable structure, the amount of diversity is nearly negligible.

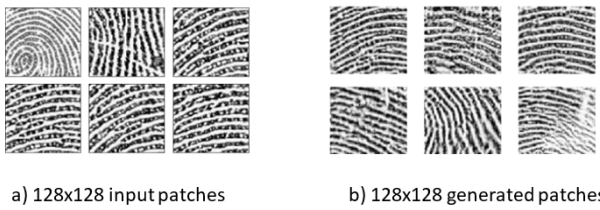

Figure 6: 128 by 128 patch from the input images and the corresponding fake fingerprints.

조선대학교
CHOSUN UNIVERSITY

Figure 7: 256 by 256 patch from the input images and the corresponding fake fingerprints.

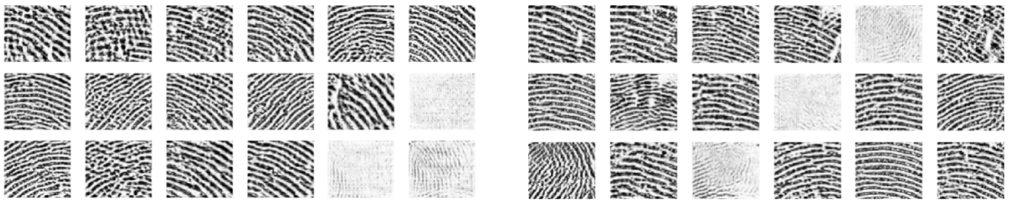

Figure 8: a) 128 by 128 patch generation. Patches from the proposed network without spectral normalization. b) 128 by 128 patch generation. Patches from the proposed network with spectral normalization.

From figure 4, we can easily observe the ridge difference in the fingerprints. Our research can successfully emulate sharper ridges than the other two studies. Figures 6 and 7 show the input patch images and the respective patches of 128 by 128 and 256 by 256 sizes produced in this study. Figures 8 and 9 contain the images generated using the proposed method. These images show 128 by 128 patches with and without spectral normalization and 256 by 256 patches with and without spectral normalization. Using spectral normalization, we enjoyed a similar quality in the output compared to images without spectral normalization.

For comparative purposes, we have used the DCGAN [16], WGAN [17], WGAN-GP[18], BEGAN [21], and G-GANISR [41] for image generation. We stacked the output from these in figure 5. BEGAN produces images

조선대학교
CHOSUN UNIVERSITY

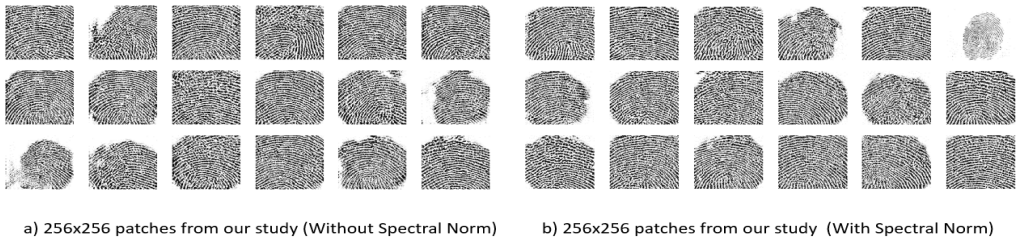

Figure 9: a) 256 by 256 patch generation. Patches from the proposed network without spectral normalization. b) 256 by 256 patch generation. Patches from the proposed network with spectral normalization.

without any deformity. However, the diversity between generated images is minimal. WGAN and DCGAN produce distorted fingerprint images compared to BEGAN. Additionally, these networks provide greater variety in image generation compared to BEGAN. WGAN-GP is somewhat successful in fingerprint generation compared to DCGAN and WGAN. Even though WGAN-GP produces sharper images compared to WGAN, it also struggles to produce desirable fingerprint patches. Photos from the G-GANISR performed similarly to the DCGAN. All of these methods are viable for producing images with a 128 by 128 size. If we increase the dimension, it is tough for these methods to create any meaningful structure. Compared to these methods, our network can produce up to 256 by 256 patches with desirable fidelity and diversity, as shown in figure 5(g-h).

A common way of evaluating the performance of the GAN is to measure the inception distance [18]. Lately, researchers have used the FID score[49] more frequently than the inception distance. The inception score gives us a way to measure the quality of the generated images. This score can be calculated using a large number of generated images. The FID score is an improvement on the

조선대학교
CHOSUN UNIVERSITY

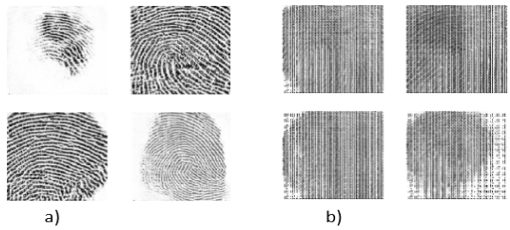

Figure 10: a) Images with loss doping. b) a) Images without loss doping. Left images are from the proposed network at epoch 3244 with the help of loss doping. The pictures on the right were produced using the same system at epoch 3244 without the aid of loss doping. From these figures, the effect of loss doping on the proposed network is apparent.

Figure 11: a) Images with loss doping. b) Images without loss doping. Images from both sides are at epoch 5975. Here, the proposed network is producing fake fingerprints.

조선대학교
CHOSUN UNIVERSITY

previously mentioned inception distance. This performance metric compares the statistics of the synthesized images according to the original images. The MS-SSIM [3] score can help us measure the diversity of the generated images. Likewise, this metric also utilizes a large number of generated images. This metric returns a score between 0.0 and 1.0. The higher the score per batch, the lower the amount of diversity among the generated images.

However, we did not use the inception distance or FID score to measure our network performance. These two metrics use weights from the inception network, and these weights are valid for images similar to those in the ImageNet dataset. Since our fingerprint dataset is absent in the ImageNet dataset, it is futile to use these metrics. Hence, we used the MS-SSIM score [3]. This score is entirely different from the other metrics in terms of its application. Since this score does not require an inception network, we can easily use it to measure our network's performance.

To quantitatively measure our network performance, we used the MS-SSIM metric[3]. Other studies [13]–[15]were not subject to diversity analysis. Table 5 shows the differences between the proposed method and other studies. In this table, [R] stands for the model with a residual connection, and [S] stands for the model with spectral-residual connections. From table 5, BEGAN shows the lowest diversity performance. Our fingerprint generator achieved better MS-SSIM scores when it came to patches with larger sizes. Since larger size patches contain shapes, ridges, and orientation, it is easier to introduce more diversity.

The MS-SSIM score is empirically lower for patches from the spectral-residual discriminator. This observation is consistent with both small- and large-scale patches. However, the DCGAN and WGAN achieve somewhat good scores even though they seem to produce irregular patches. For BEGAN, the MS-SSIM

조선대학교
CHOSUN UNIVERSITY

Table 7: MS-SSIM score comparison.

Methods	MS-SSIM [2]
BEGAN (patch images,128 by 128) [20]	0.944
DCGAN (patch images,128 by 128) [15]	0.5-0.613
WGAN (patch images,128 by 128) [16]	0.43-0.476
WGAN-GP (patch images,128 by 128) [17]	0.393-0.441
G-GANISR (patch images,128 by 128) [41]	0.47-0.558
Proposed method (patch images[S], 128 by 128)	0.31 -0.374
Proposed method (patch images[R], 128 by 128)	0.351-0.425
Proposed method (patch images[S], 256 by 256)	0.23-0.34
Proposed method (patch images[R], 256 by 256)	0.258 -0.32

score is the highest. This result justifies the figure 5. We inserted 128 by 128 patches and downscaled 256 by 256 patches to 128 by 128 patches in the BEGAN network. BEGAN seems to produce the same image every time with very slight ridge variation for both of these cases. Moreover, blurred structures are prominent in figure 5 for BEGAN. WGAN-GP produces deformed patches and shows better diversity performance than other methods. G-GANISR shows better diversity than BEGAN and DCGAN.

Our network achieves better diversity performance for both 128 by 128 and

조선대학교
CHOSUN UNIVERSITY

Table 8: MS-SSIM and average SSIM score comparison.

Proposed method (MS-SSIM, 128 by 128)	0.31 -0.374
Proposed method (average SSIM, 128 by 128)	0.257-0.324

Table 9: Number of weights.

Methods	(Generator, Discriminator)
BEGAN [20]	(19.7 M , 3.9 M)
DCGAN [15]	(45.3 M , 7.5 M)
WGAN [16]	(50.3 M , 8.4 M)
WGAN-GP [17]	(50.3 M , 8.4 M)
G-GANISR [41]	(232.4 M , 38.7 M)
Proposed method	(17.8 M , 3.3 M

256 by 256 patches compared to all of them. Our study has performed at best 0.23 MS-SSIM for 256x256 images with spectral normalization. Without spectral normalization, we achieved a slightly lower score of .258 for 256x256 images. We also measured SSIM scores between 1000 counterfeit and real photos. Thus, there are 1000 SSIM scores from the real photos for one generated image, and we averaged them. We have performed the same for the rest of the fake pictures. Table 6 contains the average SSIM score information for the 1000 fake

조선대학교
CHOSUN UNIVERSITY

fingerprints. This table justifies the MS-SSIM score since the mean SSIM score of all those GAN generated images did not exceed our achieved MS-SSIM score.

The presented GAN scheme is lighter in terms of trainable weights. For comparison, we have counted number of trainable weights for our network and other studies [16]–[18], [21], [41]. From table 7, we can see that the weight count for our architecture is significantly lower than other state-of-the-art studies.

Usually, the training time required for the GAN is higher than other deep learning networks. Our model's training time varies from 30 hours to several days, depending upon the size of the generated images and the dataset. Our trained model can produce a batch of 36 fake fingerprints in between 4 to 7 seconds.

조선대학교
CHOSUN UNIVERSITY

VI. CONCLUSION

In this study, we have presented a new GAN scheme to generate fingerprints. The proposed method can successfully produce whole and cropped fingerprint patches with 128 by 128 and 256 by 256 sizes. We have experimentally shown that our network can converge faster with the help of proposed loss doping. Additionally, to generate these fingerprints, our scheme utilizes comparatively a fewer number of weights. Furthermore, our network has demonstrated better divergence performance compared to other state-of-the-art studies. We have also presented experimental results to justify our selection of activation function, noise vector, and network design.

We can easily extend the proposed study for different lines of fingerprint scanners. The proposed doping allows our models to converge faster, although this paper does not cover the generalization of loss doping. We hope this work can provide general insight into the designing of the GAN networks. However, like the current GAN studies, our GAN scheme is not free from a redundant distribution and does not guarantee the deformity free image generation.

In our future work, we would like to extend our research for stable 512 by 512 fingerprint patch generation. Additionally, we are hoping to design a stable GAN architecture that can produce fingerprints with a precise boundary line and high fidelity.

조선대학교
CHOSUN UNIVERSITY

PUBLICATIONS

A. Journals

1. M. A.-N. I. Fahim and H. Y. Jung, “A lightweight gan network for large scale fingerprint generation”, *IEEE Access*, vol. 8, pp. 92 918–92 928, 2020.

2. M. A.-N. I. Fahim and H. Y. Jung, “Fast single-image hdr tone-mapping by avoiding base layer extraction”, *Sensors*, vol. 20, no. 16, p. 4378, 2020.

B. Conferences

1. S. Hossain and M. A. N. I. Fahim, “A simple way of image encryption using pixel shuffling and pixel manipulation”, in *2017 20th International Conference of Computer and Information Technology (ICCIT)*, IEEE, 2017, pp. 1–4.

2. M. A. N. I. Fahim, S. Mostafa, J. Tasnim, *et al.*, “Alignment of 3-d scanning data for polygonal mesh based on modified triangulation”, in *2017 6th International Conference on Informatics, Electronics and Vision & 2017 7th International Symposium in Computational Medical and Health Technology (ICIEV-ISCMHT)*, IEEE, 2017, pp. 1–5.

3. S. Mostafa, M. A. N. I. Fahim, and A. A. Hossain, “A new chaos based medical image encryption scheme”, in *2017 6th International Conference on Informatics, Electronics and Vision & 2017 7th International Symposium in Computational Medical and Health Technology (ICIEV-ISCMHT)*, IEEE, 2017, pp. 1–6.

조선대학교
CHOSUN UNIVERSITY

4. M. A. N. I. Fahim and S. Hossain, “A simple way to create pointillistic art from natural images”, in *2017 3rd IEEE International Conference on Cybernetics (CYBCONF)*, IEEE, 2017, pp. 1–5.

5. S. Mostafa, M. A. N. I. Fahim, and J. Tasnim, “An approach to effective 3d reconstruction based on point cloud merging”, in *2016 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, IEEE, 2016, pp. 262–264.

6. S. Hossain, M. A. N. Fahim, N. N. Jui, *et al.*, “A simple nonlinearity upgradation based method of enhancing low light images”, in *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)*, IEEE, 2019, pp. 1–6.

조선대학교
CHOSUN UNIVERSITY

REFERENCES

[1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets", in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[2] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida, "Spectral normalization for generative adversarial networks", *arXiv preprint arXiv:1802.05957*, 2018.

[3] Y. Guo, Q. Chen, J. Chen, Q. Wu, Q. Shi, and M. Tan, "Auto-embedding generative adversarial networks for high resolution image synthesis", *IEEE Transactions on Multimedia*, 2019.

[4] H. Jung and Y. Heo, "Fingerprint liveness map construction using convolutional neural network", *Electronics Letters*, vol. 54, no. 9, pp. 564–566, 2018.

[5] G. L. Marcialis, A. Lewicke, B. Tan, P. Coli, D. Grimberg, A. Congiu, A. Tidu, F. Roli, and S. Schuckers, "First international fingerprint liveness detection competition—livdet 2009", in *International Conference on Image Analysis and Processing*, Springer, 2009, pp. 12–23.

[6] D. Yambay, L. Ghiani, P. Denti, G. L. Marcialis, F. Roli, and S Schuckers, "Livdet 2011—fingerprint liveness detection competition 2011", in *2012 5th IAPR international conference on biometrics (ICB)*, IEEE, 2012, pp. 208–215.

[7] H. Y. Jung, Y. S. Heo, and S. Lee, "Fingerprint liveness detection by a template-probe convolutional neural network", *IEEE Access*, vol. 7, pp. 118 986–118 993, 2019.

조선대학교
CHOSUN UNIVERSITY

[8] A. Roy, N. Memon, J. Togelius, and A. Ross, "Evolutionary methods for generating synthetic masterprint templates: Dictionary attack in fingerprint recognition", in *2018 International Conference on Biometrics (ICB)*, IEEE, 2018, pp. 39–46.

[9] A. Roy, N. Memon, and A. Ross, "Masterprint: Exploring the vulnerability of partial fingerprint-based authentication systems", *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 9, pp. 2013–2025, 2017.

[10] R. Cappelli, D. Maio, and D. Maltoni, "Synthetic fingerprint-database generation", in *Object recognition supported by user interaction for service robots*, IEEE, vol. 3, 2002, pp. 744–747.

[11] P. Johnson, F. Hua, and S. Schuckers, "Texture modeling for synthetic fingerprint generation", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2013, pp. 154–159.

[12] Q. Zhao, A. K. Jain, N. G. Paulter, and M. Taylor, "Fingerprint image synthesis based on statistical feature models", in *2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, IEEE, 2012, pp. 23–30.

[13] K. Cao and A. Jain, "Fingerprint synthesis: Evaluating fingerprint search at scale", in *2018 International Conference on Biometrics (ICB)*, IEEE, 2018, pp. 31–38.

[14] P. Bontrager, A. Roy, J. Togelius, N. Memon, and A. Ross, "Deepmasterprints: Generating masterprints for dictionary attacks via latent variable evolution", in *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, IEEE, 2018, pp. 1–9.

조선대학교
CHOSUN UNIVERSITY

[15] S. Minaee and A. Abdolrashidi, “Finger-gan: Generating realistic fingerprint images using connectivity imposed gan”, *arXiv preprint arXiv:1812.10482*, 2018.

[16] A. Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks”, *arXiv preprint arXiv:1511.06434*, 2015.

[17] M. Arjovsky, S. Chintala, and L. Bottou, “Wasserstein gan”, *arXiv preprint arXiv:1701.07875*, 2017.

[18] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, “Improved training of wasserstein gans”, in *Advances in neural information processing systems*, 2017, pp. 5767–5777.

[19] N. Kodali, J. Abernethy, J. Hays, and Z. Kira, “How to train your dragan”, *arXiv preprint arXiv:1705.07215*, vol. 2, no. 4, 2017.

[20] L. Mescheder, A. Geiger, and S. Nowozin, “Which training methods for gans do actually converge?”, *arXiv preprint arXiv:1801.04406*, 2018.

[21] D. Berthelot, T. Schumm, and L. Metz, “Began: Boundary equilibrium generative adversarial networks”, *arXiv preprint arXiv:1703.10717*, 2017.

[22] A. Odena, C. Olah, and J. Shlens, “Conditional image synthesis with auxiliary classifier gans”, in *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, JMLR. org, 2017, pp. 2642–2651.

[23] A. Brock, J. Donahue, and K. Simonyan, “Large scale gan training for high fidelity natural image synthesis”, *arXiv preprint arXiv:1809.11096*, 2018.

조선대학교
CHOSUN UNIVERSITY

[24] T. Nguyen, T. Le, H. Vu, and D. Phung, “Dual discriminator generative adversarial nets”, in *Advances in Neural Information Processing Systems*, 2017, pp. 2670–2680.

[25] I. Durugkar, I. Gemp, and S. Mahadevan, “Generative multi-adversarial networks”, *arXiv preprint arXiv:1611.01673*, 2016.

[26] B. Neyshabur, S. Bhojanapalli, and A. Chakrabarti, “Stabilizing gan training with multiple random projections”, *arXiv preprint arXiv:1705.07831*, 2017.

[27] B. Zhao, W. Li, and W. Gong, “Deep pyramid generative adversarial network with local and nonlocal similarity features for natural motion image deblurring”, *IEEE Access*, 2019.

[28] I. O. Tolstikhin, S. Gelly, O. Bousquet, C.-J. Simon-Gabriel, and B. Schölkopf, “Adagan: Boosting generative models”, in *Advances in Neural Information Processing Systems*, 2017, pp. 5424–5433.

[29] A. Ghosh, V. Kulharia, V. P. Namboodiri, P. H. Torr, and P. K. Dokania, “Multi-agent diverse generative adversarial networks”, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8513–8521.

[30] Q. Hoang, T. D. Nguyen, T. Le, and D. Phung, “Mgan: Training generative adversarial nets with multiple generators”, 2018.

[31] H. Zhang, T. Xu, H. Li, S. Zhang, X. Wang, X. Huang, and D. N. Metaxas, “Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks”, in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 5907–5915.

조선대학교
CHOSUN UNIVERSITY

[32] ——, “Stackgan++: Realistic image synthesis with stacked generative adversarial networks”, *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 8, pp. 1947–1962, 2018.

[33] H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena, “Self-attention generative adversarial networks”, *arXiv preprint arXiv:1805.08318*, 2018.

[34] W. Lee, D. Kim, S. Hong, and H. Lee, “High-fidelity synthesis with disentangled representation”, *arXiv preprint arXiv:2001.04296*, 2020.

[35] A. Karnewar, O. Wang, and R. S. Iyengar, “Msg-gan: Multi-scale gradient gan for stable image synthesis”, *CoRR, abs/1903.06048*, vol. 6, 2019.

[36] Y. Miron and Y. Coscas, “S-flow gan”, *arXiv preprint arXiv:1905.08474*, 2019.

[37] C. H. Lin, C.-C. Chang, Y.-S. Chen, D.-C. Juan, W. Wei, and H.-T. Chen, “Coco-gan: Generation by parts via conditional coordinating”, in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 4512–4521.

[38] T. Karras, T. Aila, S. Laine, and J. Lehtinen, “Progressive growing of gans for improved quality, stability, and variation”, *arXiv preprint arXiv:1710.10196*, 2017.

[39] T. Karras, S. Laine, and T. 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.

[40] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, *et al.*, “Photo-realistic single image super-resolution using a generative adversarial network”, in *Proceedings*

조선대학교
CHOSUN UNIVERSITY

of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4681–4690.

[41] P. Shamsolmoali, M. Zareapoor, R. Wang, D. K. Jain, and J. Yang, "G-ganisr: Gradual generative adversarial network for image super resolution", *Neurocomputing*, vol. 366, pp. 140–153, 2019.

[42] M. Zareapoor, H. Zhou, and J. Yang, "Perceptual image quality using dual generative adversarial network", *Neural Computing and Applications*, pp. 1–11, 2019.

[43] W. Zhou, C. Wu, Y. Yi, and G. Luo, "Structure preserving non-negative feature self-representation for unsupervised feature selection", *IEEE Access*, vol. 5, pp. 8792–8803, 2017.

[44] Z. Qiumei, T. Dan, and W. Fenghua, "Improved convolutional neural network based on fast exponentially linear unit activation function", *IEEE Access*, vol. 7, pp. 151 359–151 367, 2019.

[45] *Http://livdet.org*.

[46] L. Ghiani, D. A. Yambay, V. Mura, G. L. Marcialis, F. Roli, and S. A. Schuckers, "Review of the fingerprint liveness detection (livdet) competition series: 2009 to 2015", *Image and Vision Computing*, vol. 58, pp. 110–128, 2017.

[47] V. Mura, G. Orrù, R. Casula, A. Sibiriu, G. Loi, P. Tuveri, L. Ghiani, and G. L. Marcialis, "Livdet 2017 fingerprint liveness detection competition 2017", in *2018 International Conference on Biometrics (ICB)*, IEEE, 2018, pp. 297–302.

조선대학교
CHOSUN UNIVERSITY

[48] L. Ghiani, D. Yambay, V. Mura, S. Tocco, G. L. Marcialis, F. Roli, and S. Schuckcrs, “Livdet 2013 fingerprint liveness detection competition 2013”, in *2013 International Conference on Biometrics (ICB)*, IEEE, 2013, pp. 1–6.

[49] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, “Gans trained by a two time-scale update rule converge to a local nash equilibrium”, in *Advances in Neural Information Processing Systems*, 2017, pp. 6626–6637.

조선대학교
CHOSUN UNIVERSITY

ACKNOWLEDGEMENTS

In the beginning, there was light. I am grateful to the creator of the light who created me, ALLAH. Then, I have no words to explain my gratitude towards my venerable advisor, my advisor, Professor Ho Yub Jung. I am incredibly grateful that I have enjoyed advanced hardware resources, absolute creative freedom, limitless overlooking of my foibles, continuous support, and great individualism under his supervision. Moreover, I would like to mention his profound fastidiousness, which drives me to be more precise in my life.

Here, I have met two extraordinary people, Md. Atikul Islam and Abol Basher vai. Maybe I can write another thesis about their support in my life. Anyone would want selfless Abol Basher in his/her life to stay for the rest. I would also like to say special thanks to beloved Arafat Habib, aka AH!, definitely brings a special theme in my life here. I am also indebted to Yasir Arafat vat for his never-ending official guidance and informative support. I am grateful to Vidal, aka N Saqib, for his friendly support, Naim, Soumen, Saif aka Guti, Proshen, Sarkar, aka Soka, Fuad, Zedi. Their colorful presence in my life is beyond my words. A special thanks to my lovely guffy brother TushyKing Asif, who never misses bringing sunshine in my face. I mean it, Never. Founding People: A B M Aowlad Sir, Md. Jahirujjaman, Sakib Mostafa, Jarin Tasnim; lifetime thanks towards them.

My parents, I guess it is impossible to say anything about their contribution to my life. Arigato Mahmoud, Arigato.

Lastly, I am grateful to Almighty ALLAH for all the blessings, luck, and guidance. Toda, Elohim.