FingerGAN: A Constrained Fingerprint Generation Scheme for Latent Fingerprint Enhancement

Yanming Zhu, Xuefei Yin, and Jiankun Hu, Senior Member, IEEE

Abstract—Latent fingerprint enhancement is an essential preprocessing step for latent fingerprint identification. Most latent fingerprint enhancement methods try to restore corrupted gray ridges/valleys. In this paper, we propose a new method that formulates latent fingerprint enhancement as a constrained fingerprint generation problem within a generative adversarial network (GAN) framework. We name the proposed network FingerGAN. It can enforce its generated fingerprint (i.e., enhanced latent fingerprint) indistinguishable from the corresponding ground truth instance in terms of the fingerprint skeleton map weighted by minutia locations and the orientation field regularized by the FOMFE model. Because minutia is the primary feature for fingerprint recognition and minutia can be retrieved directly from the fingerprint skeleton map, we offer a holistic framework that can perform latent fingerprint enhancement in the context of directly optimizing minutia information. This will help improve latent fingerprint identification performance significantly. Experimental results on two public latent fingerprint databases demonstrate that our method outperforms the state of the arts significantly. The codes will be available for non-commercial purposes from https://github.com/HubYZ/LatentEnhancement.

Index Terms—Constrained fingerprint generation, deep convolutional generative adversarial network, latent fingerprint enhancement.

I INTRODUCTION

Fingerprint have been widely used for human verification and identification in many civil or criminal applications [1], [2]. Different from plain and rolled fingerprints that are acquired professionally, latent fingerprints refer to finger skin impressions unintentionally left at a crime scene and are generally used as important evidence to identify criminals by law enforcement and forensic agencies. Compared with plain and rolled fingerprints, latent fingerprints are usually smudgy and blurred, with incomplete regions, unclear ridge structures, and complex background noise. Due to these factors, the identification accuracy of latent fingerprints, which heavily relies on fingerprint quality, is much lower than that of plain and rolled fingerprints. Therefore, latent fingerprint enhancement, which aims to improve latent fingerprint quality, becomes one of the most necessary and important preprocessing steps for latent fingerprint identification.

Over the past few decades, many efforts have been made toward latent fingerprint enhancement [3], [4], [5], [6], [7]. In the early days, classical image processing techniques such as contextual filtering and directional filtering were introduced to enhance fingerprints. For example, Cappelli et al. [8] proposed tuning a Gabor filter to the local orientations and frequencies of fingerprints to suppress noise and improve the clarity of ridge structure. Chikkerur et al. [9] proposed performing contextual filtering in the Fourier domain to enhance fingerprints. However, these methods are mainly effective for bad-quality plain or rolled fingerprints, and tend to fail in latent fingerprint enhancement due to: 1) the corrupted ridge structures caused by structural noise in latent fingerprints; and 2) the unreliable orientation and frequency estimation caused by the low clarity of ridge structures of latent fingerprints. Therefore, varieties of smoothing and global modeling techniques were proposed to address the above problems and devoted to reliable orientation estimation to improve the latent fingerprint enhancement [10], [11], [12], [13]. For example, Yoon et al. [11] proposed using a polynomial model together with Gabor filters to estimate the fingerprint orientations to improve the latent enhancement. Feng et al. [12] proposed using an orientation patch dictionary to estimate orientations and then applying Gabor filtering to the orientations to achieve latent fingerprint enhancement. Yang et al. [13] proposed to further improve the above method by replacing its orientation dictionary with a set of localized orientation dictionaries. However, tuning of Gabor filters requires a fixed ridge frequency. This is problematic because the ridge frequency of fingerprints is not constant.

Later, to further improve the enhancement of latent fingerprints, various total variation (TV) image models, which minimize the total variation of an image and decompose the image into two components of texture and cartoon, were adopted to take advantage of ridge structures [14], [15], [16], [17]. For example, Zhang et al. firstly proposed an adaptive TV model [14] to remove the structural noise of latent fingerprints and then proposed an adaptive directional TV model [15] for latent fingerprint enhancement. These methods can restrain the
structural noise in the decomposed texture components of latent fingerprints by integrating local orientations and scales of fingerprints. However, estimating the local parameters of these models for poor-quality latent fingerprints is difficult and thus the extracted ridge structures by these models are usually weak. Therefore, in later research, TV decomposition is generally used as a preprocessing for latent fingerprint enhancement [16], [17].

After that, with the success of deep learning, deep neural networks were proposed for latent fingerprint enhancement [18], [19], [20]. For example, Svoboda et al. [21] proposed using a convolutional autoencoder to reconstruct latent fingerprints. Li et al. [22] proposed a deep convolutional network consisting of one convolution and two deconvolution parts for latent fingerprint enhancement. Qian et al. [23] proposed a latent fingerprint enhancement method based on DenseUnet. Horapong et al. [24] used a sparse autoencoder to boost the ridge/valley spectrum to enhance latent fingerprints. Liu et al. [25] proposed using deep nested UNets for latent fingerprint enhancement. These methods take advantage of the strong representation ability of deep neural networks and achieve remarkable results, but the corrupted ridge/valley structures of latent fingerprints are not well restored in most cases.

Recently, generative adversarial networks (GANs) have been used for latent fingerprint enhancement to enhance the restoration of ridge/valley structures. For example, Dabouei et al. [26] proposed a conditional GAN for partial latent fingerprint enhancement, which achieves an enhancement of rejecting seriously corrupted fingerprint regions while improving ridge structure clarity in relatively good-quality regions. Joshi et al. [27] proposed a GAN-based algorithm to amplify the ridge/valley structure of latent fingerprint for enhancement. Huang et al. [28] proposed using a progressive PatchGAN to achieve latent fingerprint enhancement. The enhancement ability of these methods mainly comes from the powerful feature representation and reconstruction ability of GANs.

In this paper, we propose a new method that formulates latent fingerprint enhancement as a constrained fingerprint generation problem within a GAN framework. The proposed network is named FingerGAN. It can enforce its generated fingerprint (i.e., enhanced latent fingerprint) indistinguishable from the corresponding ground truth instance in terms of the fingerprint skeleton map weighted by minutia locations and the orientation field regularized by the FOMFE model. Because minutia is the primary feature for recognition and minutia can be retrieved directly from the fingerprint skeleton map [29], we offer a holistic framework that can perform latent fingerprint enhancement in the context of directly optimizing minutia information. This will help improve latent fingerprint identification performance significantly. Experimental results on two public latent fingerprint databases demonstrate that our method significantly outperforms the state of the arts.

The main contributions of this paper are summarized as follows.

1) Unlike most latent fingerprint enhancement methods that try to restore corrupted gray ridges/valleys, we propose a new method that formulates latent fingerprint enhancement as a constrained fingerprint generation problem within a GAN framework.

2) We propose a FingerGAN which can generate enhanced latent fingerprints conditioned on a fingerprint-to-fingerprint translation and can enforce its generated enhanced latent fingerprints indistinguishable from the ground truth instances in terms of fingerprint skeleton map and orientation field.

3) The fingerprint skeleton map is proposed as a ground truth because minutia is the primary feature for recognition and minutia can be retrieved directly from the fingerprint skeleton map. Also, a Gaussian-based minutia weight map is proposed to apply to the reconstruction loss, which can accommodate a moderate loss of the accuracy of minutia locations.

4) The orientation field is proposed as a ground truth in a way of bringing in correspondence between the generated enhanced latent fingerprint and the ground truth orientation field. Also, the FOMFE model is adopted to regularize the orientation field so that the effects of spurious pixels and noise can be rectified.

5) A synthetic latent fingerprint generation method is proposed, which can address the issue of lacking high-volume latent fingerprints and their true mates required for deep learning.

The rest of this paper is organized as follows. Section II provides background information on related techniques. Section III describes the proposed method in detail. Section IV presents and discusses the experimental results. Finally, the paper is concluded in Section V.

II BACKGROUND

Since the proposed method involves GAN, U-Net, and the FOMFE fingerprint orientation model, relevant background knowledge is provided as follows.

A. Generative Adversarial Network

GAN is one of the most popular groups of generative networks, which learns to map an embedding space to a data distribution of interest, and has achieved great success in various image generation and processing tasks [26], [30], [31]. The underlying strategy of a GAN is emulating a competition, with a generative network, called generator $G$, which takes a random vector $z$ sampled from a noise distribution $\mathcal{Z}$ as input and tries to generate samples as ‘real’ as possible, and a discriminative network, called discriminator $D$, which performs binary classification to distinguish samples generated by $G$ from the real samples and acts as an adversary. The goal of $G$ is to maximize the misclassification error of $D$ while the goal of $D$ is to beat $G$ by learning to identify the generated samples. Through such a zero-sum game, the GANs have the ability to learn any kind of data distribution in an unsupervised manner. The networks of $G$ and $D$ are trained iteratively with two steps: 1) fixing the parameters of $G$ and optimizing $D$; and 2) fixing the parameters
of $D$ and optimizing $G$ by using a loss function formulated as [32]:

$$
\min_{G} \max_{D} L(G, D) = \mathbb{E}_{x \in \mathcal{X}} [\log(D(x))] + \mathbb{E}_{z \in \mathcal{Z}} [\log(1 - D(G(z)))] ,
$$

where $x$ is the real sample from the data distribution $\mathcal{X}$. $D(x)$ represents the binary classification score given input $x$. During the training, half of the samples are real and the rest $G(z)$ are samples generated by $G$ given $z$.

Although the superiority of GAN in unsupervised representation learning, it can not be directly used for latent fingerprint enhancement due to its high probability of generating unrelated fingerprints. A GAN conditioned on the given information or constrained by prior knowledge can address this issue [31], [33], which inspires us to propose the FingerGAN. It is elaborately designed by customizing a GAN to fit the latent fingerprint enhancement task.

### B. U-Net and Its Variations

U-Net [35] is a fully convolutional neural network (CNN) that was originally invented for biomedical image segmentation. It has a U-shaped encoder-decoder network architecture consisting of two main parts: a contracting path (encoder network) and an expansive path (decoder network). The encoder and decoder networks have four encoder blocks and four decoder blocks, respectively, and are connected via skip connections. The encoder network is responsible for feature extraction, which compresses the resolution of the input image and extracts target sensitive information. The decoder network is responsible for mixing the extracted features with the outputs of horizontally corresponding encoder blocks to generate a semantic segmentation mask. U-Net has been proven to be a powerful tool to learn efficient data presentation and semantically meaningful information. In fact, after this, the U-shaped network has been widely used in various tasks including image-to-image translation [31], [36].

Inspired by this, we propose embedding a U-shaped network in a GAN for latent fingerprint enhancement. It can leverage both the advantages of the U-shaped network and the advantages of GAN by jointly training the U-shaped network with an adversarial loss, as in [37].

### C. FOMFE Model

FOMFE model describes the global topology of fingerprint ridges and is for modeling fingerprint orientations [38]. It is a regularized orientation field that is more reliable against noise and works well for low-quality fingerprints. Therefore, in this paper, we introduce it as prior knowledge to guide the constrained fingerprint generation.

### III PROPOSED METHOD

#### A. Problem Formulation

We propose to formulate the latent fingerprint enhancement as a constrained fingerprint generation problem conditioned on a fingerprint-to-fingerprint translation. For this purpose, we propose a FingerGAN by embedding a U-shaped network in a GAN such that the U-shaped network acts as the generator of the GAN, as illustrated in Fig. 1.

The U-shaped network is responsible for generating enhanced latent fingerprints given input latent fingerprints. Because minutiae is the primary feature for recognition and minutia can be retrieved directly from the fingerprint skeleton map, we propose using the minutia location weighted fingerprint skeleton map as a ground truth to force the U-shaped network to perform latent fingerprint enhancement in the context of directly optimizing minutia information. The discriminator is used to force the U-shaped network to generate enhanced latent fingerprints indistinguishable from the ground truth instances in terms of both the fingerprint skeleton map and the FOMFE-based orientation field. For this purpose, its input is a concatenation of the fingerprint skeleton map and the FOMFE-based orientation field. Specifically, the U-shaped network generated enhanced latent fingerprint and the ground truth orientation field are concatenated to form a type of input. The ground truth skeleton map and the ground truth orientation field are concatenated to form another type of input. The discriminator tries to distinguish these two types of inputs to beat the U-shaped network. This design of concatenation brings in correspondence between the generated enhanced fingerprint and the ground truth orientation field. Therefore, the generation of the U-shaped network is constrained by prior knowledge of FOMFE-based orientation field and can address the problem of generating unrelated fingerprints. Details of the proposed FingerGAN are provided in the following Section III.B.

#### B. Details of the Proposed FingerGAN

Fig. 2 illustrates the details of the proposed FingerGAN.

1) **U-Shaped Network**: The U-shaped network consists of an encoder with five composite convolutional blocks (C1-C5) and a decoder with five deconvolutional blocks (DC1-DC5), where skip connection [39] is adopted for the first four composite deconvolutional blocks. This is proposed to keep the high-frequency details of the inputs and increase the quality of the reconstruction from the decoder. Each of the first four composite convolutional blocks consists of two convolutional layers, and each convolutional layer is followed by a batch-normalization layer and a leaky rectified linear unit (ReLU) [40]. The last composite convolutional block consists of one convolutional layer, which is followed by a batch-normalization layer and a leaky ReLU layer. Each of the first four composite deconvolutional blocks consists of two up-convolutional layers, and each up-convolutional layer is followed by a batch-normalization layer and a leaky ReLU layer. The last composite deconvolutional block consists of an up-convolutional layer, a batch-normalization layer, and a sigmoid layer. According to the study in [41], successive convolutions by a set of small kernels are equal to one convolution by a larger kernel. It can effectively enhance a network’s discriminative

1 A concatenation is a two-channel tensor where the fingerprint skeleton map is its first channel map and the FOMFE-based orientation field is its second channel map.
power and reduce the number of parameters required to be learned. In this paper, we use a set of small kernels and their details are reported in Table I. Also, we double or halve the kernel numbers when the size of feature maps halving or doubling.

The input of the U-shaped network is a latent fingerprint to be enhanced, and the output is the generated enhanced latent fingerprint. In the training stage, the input latent fingerprints are synthesized using rolled fingerprints by our proposed method described in the following Section III.C. The ground truths used to optimize the generation of the U-shaped network are the fingerprint skeleton maps of the rolled fingerprints. This way, by calculating a reconstruction loss between the generated enhanced latent fingerprint and the ground truth, the U-shaped network can learn to denoise the input latent fingerprint and reconstruct its fingerprint skeleton.

2) **Discriminator**: The architecture of the discriminator is a classical CNN. It has seven composite blocks, and each of the first six blocks consists of a convolutional layer, a batch-normalization layer, and a leaky ReLU layer. The last block consists of a convolutional layer, a batch-normalization layer, and a sigmoid layer. Similar to the parameter choice of the U-shaped network, we use small kernels for the discriminator. Details of the kernels are in Table I.

The discriminator takes a two-channel map as input and outputs a binary classification score. Specifically, the U-shaped network generated enhanced latent fingerprints and the ground truth skeleton map are respectively concatenated with the ground truth orientation field to form two types of two-channel inputs to the discriminator. The discriminator tries to distinguish them and thus can force the U-shaped network generated enhanced latent fingerprint indistinguishable from the ground truth in terms of the fingerprint skeleton map and the FOMFE-based orientation field. This way, it enables the U-shaped network to have an ability of deep semantic understanding, and thus to learn to restore the corrupted ridge structure of the latent fingerprint in addition to the denoising.

3) **Gaussian-Based Minutia Weight Map**: To force the U-shaped network to optimize minutia information, we propose a
TABLE I
DETAILS OF THE ARCHITECTURE OF THE FINGERGAN

| Block | Layer | Kernel Size | Stride | Kernel Number | Block | Layer | Kernel Size | Stride | Kernel Number |
|-------|-------|-------------|--------|---------------|-------|-------|-------------|--------|---------------|
| C1    | conv1 | 3 × 3       | 1      | 64            | DC1   | up-conv1 | 2 × 2       | 2      | 512           |
|       | conv2 | 3 × 3       | 1      | 64            |       | up-conv2 | 3 × 3       | 1      | 512           |
| C2    | conv1 | 2 × 2       | 2      | 128           | DC2   | up-conv1 | 2 × 2       | 2      | 256           |
|       | conv2 | 3 × 3       | 1      | 128           |       | up-conv2 | 3 × 3       | 1      | 256           |
| C3    | conv1 | 2 × 2       | 2      | 256           | DC3   | up-conv1 | 2 × 2       | 2      | 128           |
|       | conv2 | 3 × 3       | 1      | 256           |       | up-conv2 | 3 × 3       | 1      | 128           |
| C4    | conv1 | 2 × 2       | 2      | 512           | DC4   | up-conv1 | 2 × 2       | 2      | 64            |
|       | conv2 | 3 × 3       | 1      | 512           |       | up-conv2 | 3 × 3       | 1      | 64            |
| C5    | conv1 | 2 × 2       | 2      | 1024          | DC5   | up-conv1 | 3 × 3       | 1      | 1            |

conv: convolution, up-conv: up-convolution.

Fig. 3. Illustration of the proposed Gaussian-based minutia weight map. (a) a fingerprint skeleton map; (b) the minutia map M of (a); and (c) the Gaussian-based minutia weight map of (a).

Gaussian-based minutia weight map w which is defined as:

\[ w(x, y) = \begin{cases} w'(x, y), & \text{if } w'(x, y) \neq 0, \\ w_0, & \text{otherwise}, \end{cases} \]

with

\[ w'(x, y) = \frac{\sum_{u=-r}^{r} \sum_{v=-r}^{r} w_g(u, v) \cdot M(x + u, y + v)}{\sum_{u=-r}^{r} \sum_{v=-r}^{r} w_g(u, v)}, \]

\[ w_g(u, v) = \frac{1}{2\pi\sigma^2} e^{-\frac{u^2 + v^2}{2\sigma^2}}, \]

and

\[ w_0 = \frac{w_g(r, r)}{\sum_{u=-r}^{r} \sum_{v=-r}^{r} w_g(u, v)}, \]

where \((x, y)\) are coordinates of each pixel, \((u, v)\) are coordinates of pixels in a local window centered at \((x, y)\), \(r\) is half the size of the local window, \(M\) is the minutia map whose value is 1 at minutia and 0 otherwise (as shown in Fig. 3b), and \(\sigma\) is the standard deviation of Gaussian. In the experiment, \(\sigma\) is set to be 8 and \(r\) is set to be 17. An example of the proposed weight map is illustrated in Fig. 3.

4) Loss Functions: The FingerGAN has two losses: 1) an adversarial loss which is used to jointly train the discriminator and the U-shaped network, and force the U-shaped network to generate enhanced latent fingerprints indistinguishable from the ground truths in terms of fingerprint skeleton map and FOMFE-based orientation field; and 2) a reconstruction loss which is used to further force the U-shaped network to generate enhanced latent fingerprints in the context of optimizing minutiae information.

Denote the training latent fingerprint as \(l\) and its domain as \(\mathcal{L}\), the U-shaped network as \(G\), the generated enhanced latent fingerprint as \(G(l)\), the ground truth skeleton map as \(g\) and its domain as \(\mathcal{G}\), the ground truth FOMFE-based orientation field as \(G_F\) and its domain as \(\mathcal{G}_F\) [38], and the discriminator as \(D\). According to the loss function in (1), the adversarial loss \(L_a\) is formulated as:

\[
\min_{G} \max_{D} L_a(G, D) = \mathbb{E}_{g \in \mathcal{G}, g_F \in \mathcal{G}_F} [\log(D(g, g_F))] + \mathbb{E}_{l \in \mathcal{L}, g \in \mathcal{G}} [\log (1 - D(G(l), g_F))].
\]

We use the L1 loss as the reconstruction loss \(L_r\), and thus it is formulated as:

\[
L_r(G) = \mathbb{E}_{l \in \mathcal{L}} ||w \odot (g - G(l))||_1,
\]

where \(\odot\) denotes the element-wise multiplication. Overall, the total loss function is formulated as:

\[
\min_{G} \max_{D} L = L_a + \eta L_r,
\]

where \(\eta\) is a parameter that weights the contributions of the reconstruction loss and the adversarial loss. It is empirically set to be 0.001 in the experiments.

C. Proposed Training Data Generation

Applying deep learning to latent fingerprint applications is challenging because the current public databases either are short of the correspondence between latent fingerprints and their true mates or lack quantity. In this paper, we propose an effective procedure to generate the training data.

1) Overview: Fig. 4 illustrates the process of the proposed training data generation. First, a quality evaluation is performed on rolled fingerprints to select good-quality ones. This is important because it ensures obtaining reliable ground truth labels to provide meaningful supervision for the training. Then, the TV decomposition [34] is applied to those selected good-quality fingerprints to obtain their texture components, which are subsequently enhanced and thinned to generate ground truth skeleton.
maps. Also, ground truth FOMFE-based orientation fields are calculated based on the ground truth skeleton maps using the method in [38]. Meanwhile, noise is added to those selected good-quality fingerprints to obtain synthesized latent fingerprints, which are subsequently decomposed by the TV decomposition to obtain their texture components to be used as the training latent fingerprints. In the experiment, the rolled fingerprints are from the database NIST SD14 [42]. The quality evaluation is achieved using the method in [43] due to its effectiveness. The enhancement is achieved using the gradient-based method in [9] due to the good quality of those selected fingerprints. The noise is added by the proposed latent fingerprint synthesization method, which is described as follows.

2) Latent Fingerprint Synthesization: For better noise simulation, we propose adding complex and realistic noise instead of simple line or character noise adopted in previous works [18], [22]. This helps provide abundant training data that better mimics real latent fingerprint cases, and is important for the U-shaped network to learn more effective representations of fingerprints from tough situations.

Given a selected rolled fingerprint \( b \), firstly, a plastic distortion [44] is added by the following equation:

\[
b'_i = b_i + \Delta(b_i) \cdot g(h(b_i), k),
\]

where \( b_i = [x_i, y_i]^T \) is a point in \( b \) and \( b_i' \) is its distorted point. \( k \) is the skin plasticity coefficient. \( \Delta(b_i) \) is the torsion and traction amount computed on the basis of a rotation angle \( \theta \) and a displacement vector \( e = [e_x, e_y]^T \), and is given by

\[
\Delta(b_i) = (R_\theta \cdot (b_i - o_r) + o_t + e) - b_i,
\]

with

\[
R_\theta = \begin{bmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta 
\end{bmatrix},
\]

where \( o_r \) is the center of rotation. \( g(h(b_i), k) \) is the gradual transition defined as:

\[
g(h(b_i), k) = \begin{cases} 
0 & h(b_i) < 0 \\
\frac{1}{2} \left(1 - \cos \left(\frac{\pi h(b_i)}{k}\right)\right) & 0 < h(b_i) < k, \\
1 & \text{otherwise}
\end{cases}
\]

where function \( h(b_i) \) returns a measure proportional to the distance between the point and the border of an ellipse centered

Fig. 4. Schematic diagram of the proposed training data generation.

Fig. 5. Example of (a) various plastic distortions and (b) corresponding distorted fingerprints based on the same rolled fingerprint.

at \( o_r \) with semi-axes \( s_x \) and \( s_y \), and is formulated as:

\[
h(b_i) = \sqrt{(b_i - o_r)^T A^{-1} (b_i - o_r) - 1},
\]

with

\[
A = \begin{bmatrix}
s_x^2 & 0 \\
0 & s_y^2
\end{bmatrix}
\]

In the experiments, to generate reasonable distortions, the ranges of values for parameters \( k, \theta, e, \) and \( A \) are empirically set to \([0.5, 2], [0, 5], [-15, 15] \), and \([0.2s, 0.6s] \) (for \( s_x \)), \([s_x, 2s_x] \) (for \( s_y \)), respectively, where \( s \) is half of the size of the fingerprint image width. \( o_r \) and \( o_t \) are both set to be the center of the fingerprint image. Fig. 5 shows various plastic distortions and their corresponding distorted fingerprints based on the same rolled fingerprint.

Then, speckle noise is added to the distorted fingerprint \( b' \) by the equation \( b'' = b' + n \cdot b' \), where \( n \) is uniformly distributed random noise with mean 0 and variance set to \((0, 0.02)\). Finally, a latent fingerprint \( c \) is synthesized by fusing \( b'' \) and a realistic noise image \( d \) according to the equation:

\[
c = (1 - \lambda) b'' + \lambda d,
\]

where \( d \) is randomly cropped from the background regions of latent fingerprints in the NIST SD27 database [42]. \( \lambda \) is a weight that measures the intensity degree of the realistic noise image. In the experiments, its value ranges from 0.2 to 0.8. Fig. 6 shows some synthesized latent fingerprints and their training latent
fingertips and ground truth skeleton maps generated by the proposed latent fingerprint synthesization method.

IV EXPERIMENTAL RESULTS

In this section, we evaluate our proposed method. Database and implementation details are firstly introduced in Section IV.A. Experimental results are then presented in Sections IV.B and IV.C. Finally, the proposed method is analyzed and discussed in Section IV.D.

A. Database and Implementation Details

1) Database: Training Database The database NIST SD14 [42] is used to generate the training data, which consists of 54,000 rolled fingerprint. According to the described method in Section III.C, a total of 13,000 good-quality fingerprints are selected from them. For each of the selected fingerprints, 10 latent fingerprints are synthesized, and thus the final training database consists of a total of 130,000 training latent fingerprints and 13,000 corresponding ground truth skeleton maps.

Test Databases Two challenging latent fingerprint databases NIST SD27 [45] and IIIT-Delhi MOLF [46] are used to evaluate the performance of the proposed method. Database NIST SD27 is provided by the National Institute of Standards and Technology in collaboration with the FBI. It contains 258 latent fingerprints collected from crime scenes, which are classified based on three different qualities, ‘good’, ‘bad’, and ‘ugly’, with numbers of images 88, 85, and 85, respectively. Latent fingerprints in this database contain complex noises and degradation of various types and levels, and therefore is a rigorous benchmark for evaluating the performance of the proposed method. Database IIIT-Delhi MOLF is provided by Sankaran et al. and is widely used in latent fingerprint tasks in recent years. It contains 4,400 latent fingerprints and three sets of live-scan fingerprints obtained by different acquisition sensors of ‘Crossmatch’, ‘Secugen’, and ‘Lumidigm’. Each set has 4,000 live-scan fingerprints and can be used as a reference database for latent fingerprint identification. These three reference databases are denoted as ‘C’, ‘S’, and ‘L’, respectively. The resolution of images in these two databases is 500 ppi.

2) Implementation Details: Enhancement. Details of the architecture of the FingerGAN are provided in Fig. 2 and Table I. It was implemented in PyTorch and its optimizations are solved by the SGD solver Adam [47] with a learning rate of 0.001. During the training, 192 × 192 patches are used to train the FingerGAN. During the testing, for a latent fingerprint to be enhanced, a sliding window of size 192 × 192 with a step size of 8 was adopted to generate the enhanced latent fingerprint using the trained U-shaped network. Implementation codes will be available for non-commercial purposes from https://github.com/HubYZ/LatentEnhancement.

Identification. Enhanced latent fingerprint identification experiments are conducted to quantitatively evaluate the performance of the proposed method. For experiments conducted on the NIST SD27 database, the manually marked regions of interest provided in [12] are used consistently for all compared methods. Also, to make the identification more challenging, the reference fingerprint database is extended by adding rolled fingerprints from the NIST SD14 database. This is reasonable because the NIST SD14 database has been only used for enhancement training and has not been used in any way for the identification task. Therefore, each enhanced latent fingerprint is compared with a total of 27,258 rolled fingerprints for the identification. For experiments conducted on the IIIT-Delhi MOLF database, each enhanced latent fingerprint is compared with the first and second fingerprint samples of each subject for each of the three reference databases according to the test protocol established by Sankaran et al. [46]. The commercial software Neurotechnology VeriFinger SDK12.1 3 is used for the identification. The Cumulative Match Characteristic (CMC) curve is employed to evaluate the performance of the latent fingerprint identification.

B. Minutia Recovery Accuracy

1) Quantitative Evaluation: To evaluate the performance of our proposed method, we investigate our minutia recovery accuracy and compare it with those of the state-of-the-art Tang’s method [19], Qian’s method [23], Cao’s method [48], and Huang’s method [28].

This experiment is conducted on the NIST SD27 database because it provides manually marked minutiae which can be used as genuine minutiae. Recovered minutiae of each compared method are extracted using the VeriFinger12.1 from its enhanced

Fig. 6. Examples illustrating the proposed training data generation by (a) selected good-quality rolled fingerprints, (b) synthesized latent fingerprints of (a), (c) training latent fingerprints (TV decomposed textures) of (a), and (d) ground truth skeleton maps of (a).
Fig. 7. Numbers of recovered genuine minutiae extracted from the enhanced latent fingerprints generated by our and the four compared methods, compared with the numbers of manually marked minutiae for each of the 258 latent fingerprints in the NSIT SD27 database.

Fig. 8. Comparison of numbers of introduced fake minutiae in the enhanced latent fingerprints generated by our and the four compared methods for each of the 258 latent fingerprints in the NSIT SD27 database.

Table II

| Method  | Recovered genuine minutiae | Introduced fake minutiae |
|---------|-----------------------------|--------------------------|
| Cao’s   | 1,887                       | 12,235                   |
| Qian’s  | 1,631                       | 13,985                   |
| Huang’s | 1,581                       | 12,358                   |
| Tang’s  | 1,919                       | 11,579                   |
| Ours    | 1,982                       | 11,152                   |

We compare the extracted minutiae with the manually marked minutiae and define the recovered genuine minutiae as those extracted minutiae with both correct location, orientation, and minutia type in accordance with the manually marked minutiae. All the other extracted minutiae are defined as introduced fake minutiae. As can be seen from the results in Table II, our method achieves the best result, recovering more genuine minutiae meanwhile introducing fewer fake minutiae than all the other methods.

Also, we provide a detailed comparison of the minutiae recovery accuracy of our and the four compared methods on each of the 258 enhanced latent fingerprints in Figs. 7 and 8. As can be seen, compared with the second-best method (Tang’s method), there are 112 latent fingerprints where our enhanced latent fingerprints recover more genuine minutiae, while there are 98 latent fingerprints where Tang’s enhanced latent fingerprints recover more genuine minutiae than ours. Furthermore, there are 128 latent fingerprints where our enhanced latent fingerprints introduce fewer fake minutiae than Tang’s, while there are 111 latent fingerprints where Tang’s enhanced latent fingerprints introduce fewer fake minutiae than ours. These results demonstrate the superiority of our method in terms of minutia recovery accuracy, and support our claim that the FingerGAN can perform latent fingerprint enhancement in the context of directly optimizing minutia information.

2) Visual Inspection: We provide an illustrative example in Fig. 9 for visually comparing the enhanced latent fingerprint of our and the four compared methods. Recovered minutiae are also labeled to compare with the manually marked minutiae. By observing and comparing the bottom right areas (yellow...
Fig. 9. Example of the comparison of the enhanced latent fingerprints generated by different methods. (a) Latent fingerprint B176 from the NIST SD27 database with the manually marked minutiae labeled as red circles or crosses, (b) the enhanced latent fingerprint by Cao’s method, (c) the enhanced latent fingerprint by Qian’s method, (d) the enhanced latent fingerprint by Huang’s method, (e) the enhanced latent fingerprint by Tang’s method, and (f) the enhanced latent fingerprint by our method. Recovered genuine minutiae and introduced fake minutiae in (b-f) are labeled as red and blue circles or crosses, respectively. Some regions of interest are highlighted in red rectangles.

rectangles) of the fingerprints in Fig. 9(b-f), we can observe that our enhanced latent fingerprint (f) gets better ridge/valley structures than the enhanced latent fingerprints (b-e) of the four compared methods. The superiority of our method can also be proved by observing the recovered genuine minutiae in Fig. 9(b-f). As can be seen, a total of 16 minutiae is manually marked in the latent fingerprint (a), only three, four, five, and three recovered genuine minutiae are extracted from Cao’s, Qian’s, Huang’s, and Tang’s enhanced latent fingerprints (b-e), respectively. However, ten recovered genuine minutiae are extracted from our enhanced latent fingerprint (f).

C. Identification Performance

1) Evaluation on Database NIST SD27: To comprehensively evaluate our proposed method, we perform fingerprint identification using our enhanced latent fingerprints and compare its performance with those achieved using enhanced latent fingerprints of the state-of-the-art Tang’s method [19], Dabouei’s method [49], Qian’s method [23], Joshi’s method [27], Cao’s method [48], and Huang’s method [28]. We conduct comparison experiments on all latent fingerprints, the ‘good’ latent fingerprints, the ‘bad’ latent fingerprints, and the ‘ugly’ latent fingerprints, respectively. Comparison results are shown in Fig. 10. As can be seen, our method achieves significantly better results than all the other methods on the overall, the ‘good’, and the ‘bad’ latent fingerprints. For the identification on the ‘ugly’ latent fingerprints, our method achieves the tied best rank-1 result with Tang’s method, and outperforms Tang’s method in rank-2. These results demonstrate the superiority of our method in latent fingerprint enhancement.

2) Evaluation on Database IIIT-Delhi MOLF: We also compare the identification performance of using our enhanced latent fingerprints with those of using enhanced latent fingerprints of the six state-of-the-art methods (Tang’s [19], Dabouei’s [49], Qian’s [23], Joshi’s [27], Cao’s [48], and Huang’s [28] methods) on the IIIT-Delhi MOLF database. Fig. 11 shows the comparison of CMC curves achieved over the three reference databases ‘C’, ‘S’, and ‘L’, respectively. As can be seen, our method achieves consistently the best rank-1 performance over the three reference databases, which demonstrates the superiority and robustness of our method.

D. Ablation Study

To further analyze our method and justify the effectiveness of the FingerGAN design, we conduct the following ablation studies. These experiments are conducted on the NIST SD27 database using a reference database consisting of 258 corresponding rolled fingerprints of the NIST SD27 database. All
Fig. 10. Comparison of CMC curves achieved using the enhanced latent fingerprints generated by Cao’s, Qian’s Huang’s, Tang’s, Joshi’s, Dabouei’s, and our methods on the NIST SD27 database. (a) CMC curves achieved using all latent fingerprints, (b) CMC curves achieved using the ‘good’ latent fingerprints, (c) CMC curves achieved using the ‘bad’ latent fingerprints, and (d) CMC curves achieved using the ‘ugly’ latent fingerprints.

Fig. 11. Comparison of CMC curves achieved using the enhanced fingerprints generated by Cao’s, Qian’s Huang’s, Tang’s, Joshi’s, Dabouei’s, and our methods on the IIIT-Delhi MOLF database over the three reference databases (a) ‘C’, (b) ‘S’, and (c) ‘L’, respectively.

other experimental settings are the same as those described in Section IV.A.2, except stated otherwise.

1) The Advantage of Embedding the U-Shaped Network in a GAN: To demonstrate the effectiveness of embedding the U-shaped network in a GAN, we conduct the following ablation study. We use only the proposed U-shaped network for latent fingerprint enhancement without using the discriminator, and name this method FingerGAN-noDiscriminator. Specifically, we train the U-shaped network using only the reconstruction loss in (7) with only the fingerprint skeleton maps as the ground truths. Fig. 12 compares the CMC curves achieved using the proposed FingerGAN and the FingerGAN-noDiscriminator. As can be seen, the rank-1 accuracy achieved using the FingerGAN (76.36%) is significantly higher than that achieved using the FingerGAN-noDiscriminator (70.54%). This demonstrates the effectiveness of embedding the U-shaped network in a GAN and supports our claim that the proposed FingerGAN can force its generated enhanced latent fingerprints indistinguishable from the ground truths.

In addition, we provide an illustrative example in Fig. 13 for visually inspecting the enhanced latent fingerprints generated by the two models. As can be seen, the enhanced fingerprint generated by the FingerGAN is richer in ridge/valley details than that generated by the FingerGAN-noDiscriminator, as shown in the zoomed rectangles. This also results in more genuine minutiae being extracted from the enhanced latent fingerprint generated by the FingerGAN. This can be explained by the fact that the FingerGAN-noDiscriminator has only the $L_1$ loss, which makes it focus only on the overall error of the corresponding pixels of the enhanced latent fingerprint and the...
ground truth, while the FingerGAN has an additional $L_a$ loss to force the overall pattern of the enhanced latent fingerprint indistinguishable from that of the ground truth. It is well known that minutiae are salient features of fingerprints. They will affect fingerprint distinguishability significantly. Therefore, with the discriminator, the FingerGAN can facilitate better ridge/valley reconstruction, leading to better minutiae details.

2) The Advantage of Using the Skeleton Map: To demonstrate the effectiveness of using the fingerprint skeleton map as ground truth, we conduct the following ablation studies. We use fingerprint gray images and binary images instead of fingerprint skeleton maps as the ground truths respectively to train the FingerGAN, and name these two methods FingerGAN-gray and FingerGAN-binary respectively. Fig. 14 compares the CMC curves achieved using the proposed FingerGAN, the FingerGAN-gray, and the FingerGAN-binary. As can be seen, the rank-1 accuracy achieved using the proposed FingerGAN (76.36%) is significantly higher than those achieved using the FingerGAN-gray (70.15%) and FingerGAN-binary (71.32%). This demonstrates the effectiveness of using the fingerprint skeleton map as ground truth and supports our claim that the skeleton map facilitates ridge/valley reconstruction.

In addition, we provide an illustrative example in Fig. 15 for visually inspecting the enhanced latent fingerprints generated by these three models. As can be seen, compared with the enhanced fingerprint generated by the FingerGAN-gray and FingerGAN-binary, the one generated by FingerGAN is much clearer in ridges/valleys, as shown in the zoomed rectangles, resulting in more minutiae being identified from it. This can be explained by the fact that minutia is defined on the skeleton map [29], thus using the skeleton map as the ground truth can directly...
optimize the skeleton map of the enhanced latent fingerprint, leading to more accurate minutiae recovery.

3) The Advantage of the Gaussian Minutia Weight: To demonstrate the effectiveness of using the Gaussian-based minutia weight map, we conduct the following ablation study. We train the FingerGAN using a loss function without the Gaussian-based minutia weight map and name this method FingerGAN-noWeight. That is, we remove $w$ in the (7) and make the reconstruction loss as $L_r(G) = \mathbb{E}_{l \in L} [\|g - G(l)\|_1]$ to train the FingerGAN. Fig. 16 compares the CMC curves achieved using the proposed FingerGAN and the FingerGAN-noWeight. As can be seen, the rank-1 accuracy achieved using the proposed FingerGAN (76.36%) is significantly higher than that achieved using the FingerGAN-noWeight (62.02%). This demonstrates the effectiveness of using the Gaussian-based minutia weight map and supports our claim that the FingerGAN can perform latent fingerprint enhancement in the context of optimizing minutia information.

In addition, we provide an illustrative example in Fig. 17 for visually inspecting the enhanced latent fingerprints generated by the two models. As can be seen, the enhanced fingerprint generated by the FingerGAN obtains better ridge/valley reconstruction, especially around minutiae, as shown in the zoomed rectangles. This results in obviously more minutiae being identified from it. This can be explained by the fact that the weighted reconstruction loss $L_r$ forces the network to focus on the reconstruction of the weighted areas.

4) The Advantage of Using Orientation Field: To demonstrate the effectiveness of using the FOMFE-based orientation field, we conduct the following ablation study. We train the FingerGAN without using the FOMFE-based orientation field and name this method FingerGAN-noOF. Fig. 18 compares the CMC curves achieved using the proposed FingerGAN and the FingerGAN-noOF. As can be seen, the rank-1 accuracy achieved using the proposed FingerGAN (76.36%) is significantly higher than that achieved using the FingerGAN-noOF (71.72%). This demonstrates the effectiveness of using the FOMFE-based orientation field and supports our claim that the FOMFE-based orientation field acts as an additional constraint to guide the generation of enhanced latent fingerprints.

In addition, we provide an illustrative example in Fig. 19 for visually inspecting the enhanced latent fingerprints generated by the two models. As can be seen, the enhanced
This paper proposed a FingerGAN for latent fingerprint enhancement, which formulates latent fingerprint enhancement as a constrained fingerprint generation problem. It can enforce its generated enhanced latent fingerprint indistinguishable from the corresponding ground truth instance in terms of the fingerprint skeleton map weighted by minutia locations and the orientation field regularized by the FOMFE model. Because minutia is the primary feature for recognition and minutia can be retrieved directly from the fingerprint skeleton map, we offer a holistic framework that can perform latent fingerprint enhancement in the context of directly optimizing minutia information. This will help improve latent fingerprint identification performance significantly. Experimental results on two public latent fingerprint databases demonstrate that our method outperforms the state of the arts significantly.

REFERENCES

[1] M. Hawthorne, Fingerprints: Analysis and Understanding, Boca Raton, FL, USA: CRC, 2008.
[2] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, Handbook of Fingerprint Recognition, 2nd ed., Berlin, Germany: Springer, 2009.
[3] C. Wallace-Kunkel, C. Roux, C. Lennard, and M. Stoilovic, “The detection and enhancement of latent fingerprints on porous surfaces—a survey,” J. Forensic Identification, vol. 54, no. 6, 2004, Art. no. 687.
[4] A. Sankaran, M. Vatsa, and R. Singh, “Latent fingerprint matching: A survey,” IEEE Access, vol. 2, pp. 982–1004, 2014.
[5] A. V. Malwade, R. D. Raut, and V. Thakare, “Localized dictionaries based orientation detection and enhancement of latent fingermarks on porous surfaces—a survey,” IET Biometrics, vol. 54, no. 6, 2004, Art. no. 687.
[6] P. Schuch, S. Schulz, and C. Busch, “Survey on the impact of fingerprint image enhancement,” JET Biometrics, vol. 7, no. 2, pp. 102–115, 2018.
[7] B. Abebe, H. C. A. Murthy, E. Amare Zeretla, and Y. Desse, “Latent fingerprint enhancement techniques: A review,” J. Chem. Rev., vol. 2, no. 1, pp. 40–56, 2020.
[8] R. Cappelli, D. Maio, and D. Maltoni, “Semi-automatic enhancement of very low quality fingerprints,” in Proc. Int. Symp. Image Signal Process. Anal., 2009, pp. 678–683.
[9] S. Chikkerur, A. N. Cartwright, and V. Govindaraju, “Fingerprint enhancement using STFT analysis,” Pattern Recognit., vol. 40, no. 1, pp. 198–211, 2007.
[10] S. Yoon, J. Feng, and A. K. Jain, “On latent fingerprint enhancement,” in Biometric Technology for Human Identification VII, Bellingham, WA, USA: SPIE, 2010.
[11] S. Yoon, J. Feng, and A. K. Jain, “Latent fingerprint enhancement via robust orientation field estimation,” in Proc. Int. Joint Conf. Biometrics, 2011, pp. 1–8.
[12] J. Feng, J. Zhou, and A. K. Jain, “Orientation field estimation for latent fingerprint enhancement,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 4, pp. 925–940, Apr. 2013.
[13] Y. Yang, J. Feng, and J. Zhou, “Localized dictionaries based orientation field estimation for latent fingerprints,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 36, no. 5, pp. 955–969, May 2014.
[40] A. Sankaran, M. Vatsa, and R. Singh, “Multisensor optical and latent fingerprint database,” IEEE Access, vol. 3, pp. 653–665, 2015.

[41] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2014, arXiv:1409.1556.

[42] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2009, pp. 248–255.

NIST and FBI, “NIST special database SD14,” 2010. [Online]. Available: https://www.nist.gov/srd/nist-special-database-14

[43] NIST, NBIS (NIST biometric image software). [Online]. Available: https://www.nist.gov/itl/iad/ig/nbis.cfm

[44] R. Cappelli, D. Maio, and D. Maltoni, “Modelling plastic distortion in fingerprint images,” in Proc. Int. Conf. Adv. Pattern Recognit., 2001, pp. 371–378.

[45] NIST and FBI, “NIST special database SD27,” 2007. [Online]. Available: http://www.nist.gov/itl/iad/ig/sd27a.cfm

[46] A. Dabouei, S. Soleymani, H. Kazemi, S. M. Iranmanesh, J. Dawson, and D. P. Kingma, “Adam: A method for stochastic optimization,” in Proc. Int. Conf. Learn. Representations, 2015.

[47] K. Cao, D. L. Nguyen, C. Tymoszek, and A. K. Jain, “End-to-end latent fingerprint search,” IEEE Trans. Inf. Forensics Secur., vol. 15, pp. 880–894, 2020.

[48] A. Dabouei, S. Soleymani, H. Kazemi, S. M. Iranmanesh, J. Dawson, and N. M. Nasrabadi, “ID preserving generative adversarial network for partial latent fingerprint reconstruction,” in Proc. IEEE 9th Int. Conf. Biometrics Theory, Appl. Syst., 2018, pp. 1–10.

Xuefei Yin received the BS degree from Liaoning University, Liaoning, China, the ME degree from Tianjin University, Tianjin, China, and the PhD degree from the University of New South Wales, Canberra, Australia. He is currently a research associate with the University of New South Wales at Canberra, Australia. His research interests include biometrics, pattern recognition, privacy-preserving, information forensics, Internet of Things, and intrusion detection. He has published articles in top journals including IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Information Forensics and Security, ACM Computing Surveys, IEEE Transactions on Information Informatics, and IEEE Internet of Things Journal.

Jiankun Hu (Senior Member, IEEE) is currently a full professor with the School of Engineering and Information Technology, University of New South Wales, Canberra, Australia. He is an invited expert of the Australia Attorney-Generals Office assisting the draft of the Australia National Identity Management Policy. He has received nine Australian Research Council (ARC) Grants and has served at the Panel on Mathematics, Information and Computing Sciences, Australian Research Council ERA (The Excellence in Research for Australia) Evaluation Committee 2012. His research interest include the field of cybersecurity covering intrusion detection, sensor key management, and biometrics authentication. He has published many articles in top venues including IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Computers, IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Information Forensics and Security, Pattern Recognition, and IEEE Transactions on Industrial Informatics. He has served on the editorial board of up to seven international journals, including serving as senior area editor for IEEE Transactions on Information Forensics and Security. He received ten Australian research council (ARC) grants and has also served for the prestigious Panel of Mathematics, Information and Computing Sciences (MIC), ARC ERA (The Excellence in Research for Australia) Evaluation Committee.

Yanming Zhu received the BE degree from Shandong Agricultural University, China; the ME degree from Tianjin University, China, and the PhD degree from the University of New South Wales, Australia. She is currently a research fellow with the University of New South Wales, Sydney, Australia. Her research interests include focuses on developing computer vision and machine and deep learning methodologies for fingerprint biometrics, biomedical image analysis, and other applications such as false data injection attack detection. She has published articles in top journals including IEEE Transactions on Pattern Analysis and Machine Intelligence, Pattern Recognition, IEEE Transactions on Information Forensics and Security, ACM Computing Surveys, and Bioinformatics.

Authorized licensed use limited to: The University of Auckland. Downloaded on January 12, 2023 at 20:53:39 UTC from IEEE Xplore. Restrictions apply.