



Multi-task CNN for restoring corrupted fingerprint images¹

Wei Jing Wong^a, Shang-Hong Lai^{a,b,*}

^aDepartment of Computer Science, National Tsing Hua University, Hsinchu, Taiwan

^bMicrosoft AI R&D Center, Taipei, Taiwan



ARTICLE INFO

Article history:

Received 5 February 2019

Revised 11 November 2019

Accepted 9 January 2020

Available online 10 January 2020

Keywords:

Fingerprint image enhancement

Fingerprint recognition

Convolutional neural networks

Multi-task learning

ABSTRACT

Fingerprint image enhancement is one of the fundamental modules in an automated fingerprint recognition system (AFRS). While the performance of AFRS advances with sophisticated fingerprint matching algorithms, poor fingerprint image quality remains a major issue to achieve accurate fingerprint recognition. In this paper, we present a multi-task convolutional neural network (CNN) based method to recover fingerprint ridge structures from corrupted fingerprint images. By learning from the noises and corruptions caused by various undesirable conditions of finger and sensor, the proposed CNN model consists of two streams that reconstruct the fingerprint image and orientation field simultaneously. The enhanced fingerprint is further refined using the orientation field information. Moreover, we create a deliberately corrupted fingerprint image dataset associated with ground truth images to facilitate the supervised learning of the proposed CNN model. Experimental results show significant improvement on both image quality and fingerprint matching accuracy after applying the proposed fingerprint image enhancement technique to several well-known fingerprint datasets.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Fingerprint has been one of the most commonly used biometric traits in both security and forensic applications. Despite the fact that fingerprint offers high distinctiveness, the recognition accuracy of an automated fingerprint recognition system (AFRS) depends on several other factors, such as fingerprint image quality and the fingerprint matching algorithm.

A typical AFRS consists of five fundamental modules, including image acquisition, image enhancement, feature extraction, fingerprint matching and classification. During fingerprint image acquisition, the sensor type, condition of finger and contact between sensor and finger may affect the quality of the acquired fingerprint image. The purpose of image enhancement is to eliminate noises and corruptions caused by imperfect circumstances and to enhance fingerprint ridge structures through some image processing techniques. Fig. 1 depicts examples of fingerprint enhancement results for fingerprint images of good and poor quality by using the proposed method. Fingerprint image enhancement often includes contrast enhancement, image filtering and some other image processing techniques.

The CNN model has become a popular deep learning tool for solving image classification problems due to its capability to extract most appropriate features and perform classification simultaneously with sufficient training data. However, we aim at reconstructing the fingerprint image from its corrupted version in this work. The denoising auto-encoder [1] is an example of using artificial neural network, or specifically deep belief network, for noise removal. Some existing applications of CNNs for solving similar tasks are image super-resolution [2,3] and inpainting [4].

In this paper we propose a novel CNN model for low-quality fingerprint image enhancement coined as the orientation field-corrected fingerprint image enhancement network (OFFIENet). OFFIENet is a multi-task learning CNN model that recovers the fingerprint ridge structures assisted by the ridge orientation information. It reconstructs the fingerprint image and the orientation field simultaneously while utilizing the orientation information to further enhance the ridge structures. Although some works on end-to-end minutiae extraction using CNNs already existed in the literature [5,6], the proposed method only performs fingerprint image enhancement so that both minutiae-based and image-based matching methods can benefit from the enhanced images. Besides, fingerprint images with well-defined ridge structures provide more information, except minutiae, to be harvested.

Since CNN is a supervised learning approach, we need to provide both the input (corrupted fingerprint) and the expected output (clean fingerprint or ground truth) during the training of the

* Corresponding author at: Dept. of Computer Science, National Tsing Hua University, No. 101, Sec. 2, KuangFu Rd., Hsinchu 300, Taiwan.

E-mail address: lai@cs.nthu.edu.tw (S.-H. Lai).

¹ This work was supported in part by Ministry of Science and Technology, Taiwan (grant 105-2218-E-007-030) and Novatek Microelectronics Corp.

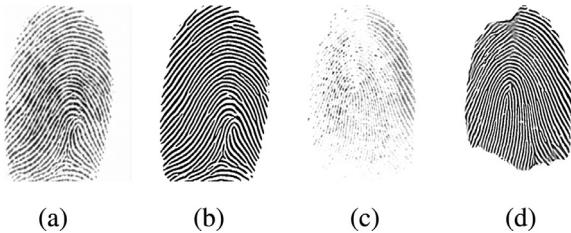


Fig. 1. Examples of (a) a good quality and (c) a poor quality fingerprint image. (b) and (d) are the corresponding enhancement results using the proposed CNN-based fingerprint ridge structure enhancement method.

CNN model so that it can learn the transformation between the images. One of the challenges in this work is the lack of existing public dataset with such pairs of fingerprint images. Therefore, we generate our own training dataset by using the Synthetic Fingerprint Generator (SFinGe) software [7] to generate ground-truth fingerprint images and apply various fingerprint corruption algorithms to simulate noises and corruptions that would appear in real fingerprint images.

The previous paper presented in [8] is most relevant to the method proposed in this paper, so we would like to discuss the major differences between our work and theirs. First, we use a deeper network model and most importantly, we introduce orientation field-correction into the proposed CNN model to improve the image quality of fingerprint reconstruction. In addition, the work in [8] directly used the corrupted fingerprints generated by SFinGe while we create our own training dataset that contains a wide variety of fingerprint corruptions. Lastly, we use image patches instead of the entire fingerprint images for training.

The main contributions of this paper are summarized as follows.

1. We propose the OFFIENet which include the reconstruction of fingerprint orientation field and fingerprint ridge structure simultaneously to enhance low-quality fingerprint images.
2. We create a fingerprint image dataset by synthetically generating corrupted fingerprints associated with the corresponding ground truths for the training of fingerprint enhancement CNN.

2. Related works

2.1. Fingerprint image enhancement techniques

Gabor filtering and its variants are one of the most popular filtering approaches utilized for fingerprint image enhancement. Hong et al. [9] pioneered this idea by taking advantage of the frequency-selective and direction-selective properties of Gabor filters to amplify and correct the local ridge patterns. The log-Gabor filters have also been proven [10] to outperform traditional Gabor filters. In addition, curved Gabor [11], orthogonal curved-line Gabor [12] and circular Gabor [13] were claimed to adapt better to the unique curved structures in fingerprints.

Contextual filtering in the Fourier domain [14] is another way of enhancing the ridge structures. Chikkerur et al. [15] used short-time Fourier analysis to compute the frequency image, orientation image and region-of-interest, which are used to construct the contextual filter bank. Fingerprint noises can be removed more effectively while preserving the ridge lines by processing in the Fourier domain. Furthermore, the combination of spatial domain and Fourier domain filtering is adopted [16–18] as an improvement to single domain filtering.

Fingerprint image filtering in the wavelet domain was demonstrated [19] by performing textural and directional filtering on the approximation image of wavelet decomposition. It has been shown [20] that the combination of Gabor filtering and wavelet decomposition produces more visually appealing outcome compared to mere Gabor filtering. Wang et al. [21] applied adaptive singular value decomposition on wavelet coefficients to improve the contrast of images.

An adaptive filtering approach [22,23] was proposed to dynamically select the filter size based upon the local ridge frequency. Yun and Cho [24] presented an adaptive preprocessing algorithm that categorizes fingerprint images into oily, dry and normal image. Moreover, Sutthiwichaiporn et al. [25] developed the adaptive boosted spectral filtering algorithm that iteratively assesses the quality of image blocks and performs spectral filtering on unfiltered blocks. The advantage of these adaptive methods is that the filters are manipulated to address certain fingerprint structures depending on the criteria.

To specifically address the scars in fingerprint images, Khan et al. [26] proposed a spatial domain scar removal strategy. In their method, scars are detected using Gaussian derivative filters and filled by tracing the orientation of nearby ridges.

In the literature reviewed above, fingerprint ridge enhancement filters are constructed based on the characteristics of ridge structures and theoretical deduction of the filter response. Cao et al. [27] incorporated dictionary learning into the filter design process for fingerprint image enhancement. First, two sets of ridge pattern dictionaries (coarse-level and fine-level) are learned from the training data. Then, the parameters of the Gabor filter are tuned according to the matched pattern so that each block can be optimally filtered according to its local orientation and frequency.

Some neural network-based fingerprint image enhancement methods were also proposed such as cellular neural networks [28], convolutional neural networks (CNN) [29] and generative adversarial networks (GAN) [30]. In [28], the authors proposed to apply some pre-defined Gabor-type filters for denoising and enhancement of fingerprint images, but the experimental results were very limited. Note that Li et al. [29] also proposed a multi-task learning approach which is closely related to this work. In their work, they proposed the FingerNet model that contains the enhancement branch and the orientation branch. The enhancement branch is designed to remove structured noise and enhance fingerprint structure. The orientation branch performs the task of guiding enhancement through a multi-task learning strategy. However, the enhancement ground truth in [29] is generated by gradient-based fingerprint enhancement method, which is difficult to ensure the quality of the ground truth. In addition, the orientation branch and enhancement branch in their network architecture are in parallel, which means the corresponding deconvolution networks are independent. But in fact, the enhanced fingerprint image and the orientation map are highly related. We would like to draw the differences between this work and theirs for the following points. First, we produce ground truth and noisy fingerprint pairs for the model training by using image synthesis method. Secondly, in our model, the orientation field predicted is then fed into the enhancement branch to directly regularize the enhancement. In addition, instead of using the quantized orientation angles in [29], we use the sine and cosine of angles to represent the fingerprint orientation for better data contiguity.

More recently, Joishi et al. [30] proposed a Generative Adversarial Network model to enhance the ridge structure quality for latent fingerprint images. The enhancement network in the GAN model is an auto-encoder architecture. They also generated a synthetic dataset to synthesize a number of pairs of latent fingerprint images and the corresponding ground-truth fingerprint images for training the GAN model.

2.2. CNNs for other fingerprint recognition applications

CNNs have been used in various aspects of fingerprint recognition. Jiang et al. [5] used CNNs to perform direct minutiae extraction on gray-scale fingerprint images. The core of the framework includes a JudgeNet that determines if an image patch contains a minutia and a LocateNet that locates the position of the minutia in the patch. A similar work [6] then surfaced but used fully convolutional network for latent fingerprint minutiae extraction instead. Another notable work on CNN-based minutiae extraction is FingerNet [31]. Besides minutiae, detection of singular points (core and delta) using CNNs was also proposed [32].

Other applications of CNNs in fingerprint recognition include fingerprint classification [33], fingerprint segmentation [34] and latent orientation field estimation [35].

3. Proposed CNN model for fingerprint image enhancement

3.1. Single-task network

Fig. 2(a) depicts the proposed single-task network for fingerprint image enhancement. The network consists of 13 convolutional layers, each followed by a rectifier linear unit (ReLU). The first 12 layers are called the BaseNet and will be used as the base network for extension later. The filter size of the first convolutional layer is 11×11 and is reduced by 2 after every three layers. The number of feature maps is 64 for all layers. The 13th layer is called the reconstruction layer which output the enhanced fingerprint image.

Like CNNs used for super-resolution [2,3] and image completion [4], one consideration of designing the proposed network is that the image size should be preserved. Therefore, padding is required to ensure that the size of the feature maps remains the same throughout the network. Another common practice of preserving the image size is to use unpooling layers or deconvolutional layers to reverse the size reduction effect caused by pooling layers or strided convolutional layers, for example the encoder-decoder architecture employed in [4]. However, we deliberately avoid using pooling layers in the proposed network as they lead to spatial information loss, especially for fine local structures like fingerprint ridges.

For gray-scale image reconstruction, the L_2 norm squared function is used to measure the reconstruction loss of the network for optimization, and it is given by

$$L_{\text{recon}} = \|Y - Y'\|_2^2 \quad (1)$$

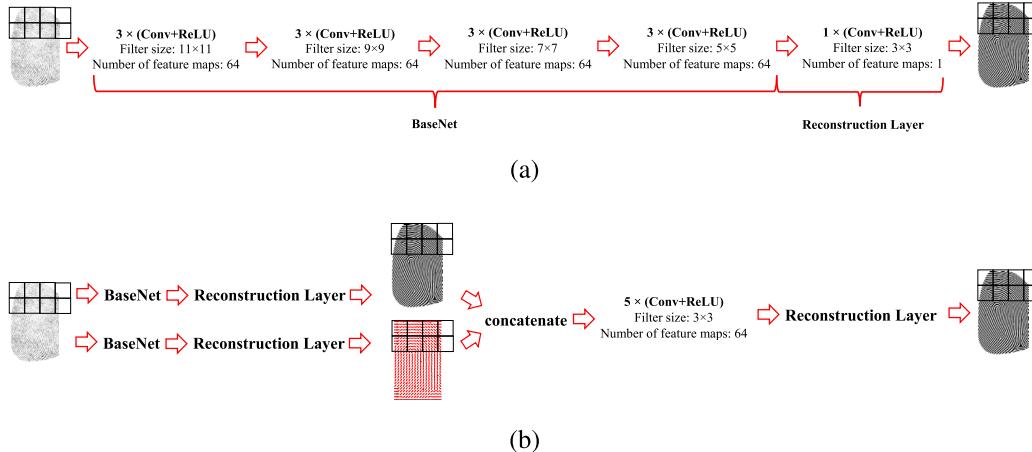


Fig. 2. The proposed CNN architecture for (a) single-task and (b) multi-task (OFFIENet) fingerprint image enhancement.

where Y and Y' represent the ground truth image and the reconstructed image, respectively.

Alternatively, the network can be tweaked to output a binary image. In this case, we modify the reconstruction layer to output two score maps of the same size. Instead of using the L_2 loss for binary image reconstruction as suggested in [8], we use the cross-entropy loss given by

$$L_{\text{recon-bin}} = -Y_0 \log(Y'_0) - Y_1 \log(Y'_1), \quad (2)$$

where Y_0 and Y_1 are the ground truth labels of 0's (black pixels) and 1's (white pixels), and Y'_0 and Y'_1 are the predicted probabilities for the corresponding binary values.

3.2. OFFIENet: multi-task network with orientation field correction

Fingerprints have unique ridge structures that flow in certain directions, known as the ridge orientation. In this work, we exploit the orientation feature of fingerprint to assist in the reconstruction of low-quality fingerprint images. Since ridge orientation is merely abstract feature of fingerprint, it is easier to recover from corrupted images, hence it can be used to refine the reconstruction. **Fig. 2(b)** shows the proposed OFFIENet to reconstruct the fingerprint image and orientation field simultaneously. First, the input image runs through two independent BaseNets to reconstruct the clean fingerprint and orientation field. Then the reconstructed images from the two BaseNets are concatenated and fed into another five convolutional layers as shown in the figure to obtain the final output.

Weight-Sharing. Since the input image is the same, the extracted lower level features should be similar regardless of the reconstruction target. Thus, the two BaseNets may share some of the filter weights even though they are trained to perform different task to reduce the number of parameters and the complexity while training the network. In this paper, we choose to apply weight-sharing between the first six layers of the two branches in **Fig. 2(b)**.

Orientation Field Representation. The direct way of representing the orientation field is by angles in the range of $[0^\circ, 180^\circ]$. However since the orientations angle is periodic and circular, linear loss functions such as the L_2 loss in (1) cannot accurately measure the distance between them. For instance a 0° -oriented ridge is only 1° different from a 179° -oriented ridge but direct subtraction of the orientation values results in a distance of 179° . Such an issue is reflected as high gradient at the edge of 0° -to- 179° transition in the orientation map as highlighted in **Fig. 3(b)**.

To overcome the above orientation discontinuity problem, we use the sine and cosine of twice the orientation angles to obtain a smoothed orientation field as shown in **Fig. 3(c)** and (d). In

this case, the orientation field is supervised by two trigonometric function maps and hence the corresponding reconstruction layer is modified to output two images.

Joint Training. The proposed multi-task network is supervised at three different stages. Summing up the three losses with loss weights gives the total loss of the network:

$$L_{\text{total}} = \lambda_1 L_{\text{recon}} + \lambda_2 L_{\text{orient}} + \lambda_3 L_{\text{final}}, \quad (3)$$

where L_{recon} , L_{orient} and L_{final} are the squared L_2 loss functions associated with the initial image reconstruction, orientation field reconstruction and refined image reconstruction, respectively, as described in (1).

During training, the gradients derived from L_{total} are back-propagated through the entire network. In addition, the task-specific losses (L_{recon} , L_{orient} and L_{final}) are used to update the weights in the corresponding stream to preserve the independence among individual tasks. The initial image reconstruction stream and orientation field stream are first trained for 40K iterations separately before performing end-to-end training of the entire multi-task network.

4. Generation of corrupted fingerprint images

In this section, we discuss the noise models used to synthetically corrupt clean fingerprints in order to simulate noises and corruptions that appear in real fingerprints. They can be categorized into the following:

1. Sensor noise (SN): Poorly maintained sensing device may inflict noise to the resulting scanned image in the process of analog-to-digital conversion and data transmission. This type of noise can be approximated by the Perlin noise [36]. It is a coherent noise that affects only certain regions of the fingerprint and is created by using the procedure described in [36]. There is a parameter called the base noise probability p_L in sensor noise generation used to control the level of the noise.
2. Scars/Wrinkles (Sc): Scars or wrinkles are concave patterns on fingers. They cause broken ridges in the fingerprint image when scanned due to uneven contact between finger and sensor. The procedure of creating these broken ridges on fingerprints is given as follows:
 - (a) Define scars or wrinkles as elliptical occlusions. A set of random parameters $\{x_0, y_0, E_w, E_l\}$ is first generated for the ellipse, where (x_0, y_0) is the center of the ellipse which can be any point in the fingerprint area and E_w and E_l are the width and length of the ellipse, respectively.
 - (b) Draw a closed elliptical mask using the parameters obtained from the previous step based on the equation:

$$E(x, y) = \begin{cases} x = x_0 + \frac{E_w}{2} \cos \theta, \\ y = y_0 + \frac{E_l}{2} \sin \theta. \end{cases} \quad (4)$$

- (c) Remove the ridges enclosed by the mask.
3. Wetness (W): When the moist on the finger makes contact with the sensor, the wet regions will have darker pixels on the scanned image regardless whether they are ridges or valleys. Although these wet regions may appear randomly on any part of the finger surface, they usually come in patches. Therefore, we use the Perlin noise distribution as the base function to simulate the appearance of finger wetness. Pixels with probability (as defined in Eq. (4) of Cappelli et al. [36]) higher than a threshold τ_w are changed to value 0 to represent wetness.
4. Dryness (D): When the finger is too dry, it does not have full contact with the sensor, resulting in thinner ridges. This can be simulated by performing dilation on the fingerprint image so that the darker pixels (ridges) are reduced.

5. Over-pressurization (OP): As opposed to dryness, the ridges thicken when excessive pressure is applied onto the sensor. Thus, the effect can be simulated by performing erosion so that the ridges are expanded.
6. Valley noise (VN): This is the background noise which is also caused by the transmission of data from sensing device. It appears as isolated black pixels in the valleys. For this, we follow the exact same algorithm for sensor noise explained above except that black noise blobs are added instead of white ones.

5. Experimental results

5.1. Preparing for CNN training

In this section, we discuss the details of preparing training data for the proposed CNN. First of all, 300 distinct clean fingerprints from five fingerprint classes (whorl, left loop, right loop, arch and tented arch) are randomly created using SFinGe [7] demo version. The ridge frequency is also randomized to ensure the variability of the training data. These clean fingerprints are used as the ground truths of the reconstruction. After that, 12 different corrupted fingerprint samples, such as the examples shown in Fig. 4, are generated from each clean fingerprint based on the combinations of different corruption methods described in Section 4. The base noise probability p_L for sensor noise is taken randomly from 0.04 to 0.15 to create fingerprints from high to low quality. The threshold that controls the wet fingerprint coverage area τ_w ranges from 0.3 to 0.4. A 3×3 Gaussian filter with parameters $\mu = 0$ and $\sigma^2 = 3$ is applied to all corrupted fingerprints for image smoothing. Since the uncorrupted fingerprints have strong ridge lines, a simple gradient-based method [37] is sufficient to extract orientation field information as labels.

We extract image patches of 91×91 from the 3600 training fingerprints with step size of 80 as the input for CNN training. Patches containing no fingerprint area are discarded. There are totally 73,658 patches. The loss weights for joint training are set to $\lambda_1 = 1$, $\lambda_2 = 0.8$ and $\lambda_3 = 1$.

5.2. Datasets and experiment settings

In this paper, we use the portion of FVC datasets created by optical sensors and SFinGe to evaluate the proposed method as our corruption methods specifically address noises that would appear in these images. Each of the FVC2002 [38] and FVC2004 [39] datasets consists of 100 subjects with eight samples per subject, whereas FVC2006 [40] datasets are made out of 140 subjects with 12 samples each. In addition to the public datasets, we create 60 synthetic fingerprints with ground truth labels according to the procedures discussed in Sections 4 and 5.1. These fingerprints are different from the ones used for CNN training and are used to evaluate the performance of the proposed fingerprint image enhancement technique.

For fingerprint matching, the minutiae are extracted from fingerprint images by using the MINDTCT extractor provided by NIST biometric image software (NBIS) [41] after the enhancement method is applied. As there may be unrecoverable and low-quality regions in the fingerprints after the enhancement, we perform minutiae pruning to eliminate unreliable minutiae. For each 16×16 non-overlapping block on the image, we observe the distribution of the pixel intensities quantized into 16 bins. A block is marked as unrecoverable and the minutiae within the block are removed if (i) there are at least two bins with more than 40 counts in the distribution; and (ii) the two bins that have the highest counts are at least 10 bins apart. Note that the minutiae pruning process is not applicable to binary images. Finally, the minutia cylinder code (MCC) [42,43] is adopted for fingerprint matching.

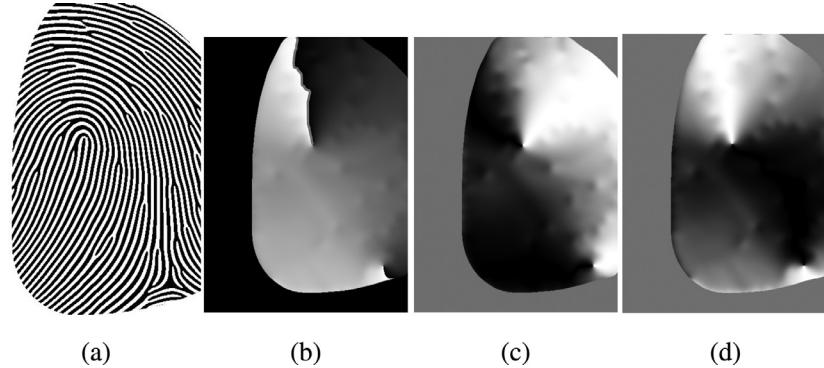


Fig. 3. Comparison between (b) the orientation map (θ) and the trigonometric function maps of twice the orientation angles, (c) $\sin 2\theta$ and (d) $\cos 2\theta$, of the same fingerprint, i.e. (a). All values are scaled to the range of [0,255] for better gray-scale visualization. The red line in (b) indicate the transition between 0° and 179° . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

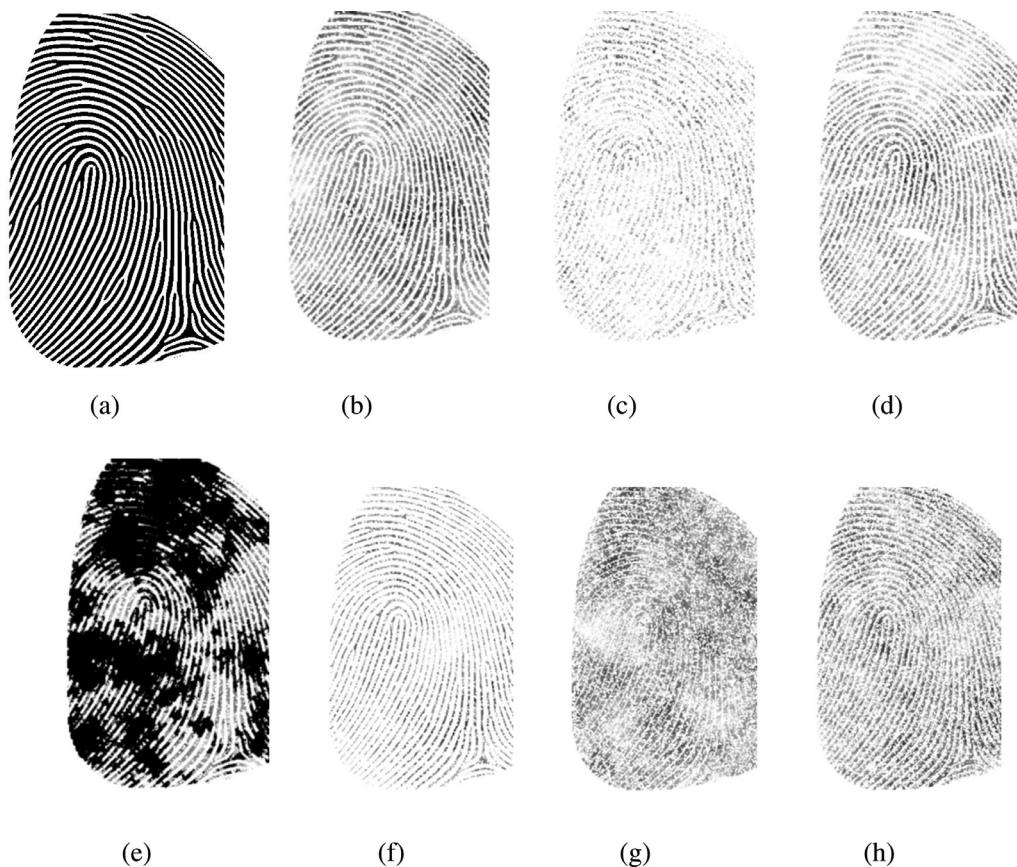


Fig. 4. Examples of corrupted fingerprints originated from fingerprint in Fig. 3(a) with combinations of different corruption method including (b) SN with $p_L = 0.04$, (c) SN with $p_L = 0.15$, (d) SN+Sc, (e) SN+W, (f) SN+D, (g) SN+OP and (h) SN+VN.

Note that the proposed fingerprint image enhancement method is used as a pre-processing step in a fingerprint recognition system. The system that uses MINDTCT for minutiae extraction and MCC for fingerprint matching has been widely used as baseline fingerprint recognition. They are selected in our evaluation because they provide reasonably accurate performance and the associated implementation code is available in public for our experiments. In addition, we also employ the state-of-the-art VeriFinger [44] fingerprint recognition system in conjunction with the proposed fingerprint enhancement method in our experiments. Note that our fingerprint enhancement method can be used in conjunction with any other fingerprint recognition systems.

We perform the experiments by using several different options of applying the enhancement methods discussed in this paper. They are

1. No enhancement: minutiae are extracted directly from the raw fingerprint images without any enhancement
2. Gabor filtering: traditional Gabor filtering [9] is applied as fingerprint image enhancement
3. BaseNet: CNN-based fingerprint enhancement based on network in Fig. 2(a)
4. BaseNet-bin: CNN-based fingerprint enhancement based on network in Fig. 2(a) with binary outputs

Table 1
Average PSNR (in dB) of the fingerprint images enhanced by the proposed method.

Method	Corruption methods						
	SN only	SN+Sc	SN+W	SN+D	SN+OP	SN+VN	All
No enhancement	8.32	8.07	9.66	7.38	8.08	8.59	8.33
Gabor filtering	7.04	7.12	6.99	7.14	6.24	6.87	6.91
BaseNet	19.26	18.93	17.95	16.74	10.55	16.17	16.86
BaseNet-bin	20.38	19.61	18.21	15.74	9.87	17.25	17.11
OFFIENet-ang	20.20	19.63	19.21	16.92	10.83	16.94	17.56
OFFIENet-tri	21.33	21.67	20.87	17.25	10.96	19.01	18.71
OFFIENet-tri-shared	22.06	21.28	20.45	17.24	11.63	18.91	18.86

Table 2
Average TMRR/FMRR (in %) of the fingerprint images enhanced by different methods.

Method	Corruption methods						
	SN only	SN+Sc	SN+W	SN+D	SN+OP	SN+VN	All
No enhancement	0.4/0.2	0.2/0.6	12.2/14.6	0.0/0.3	0.0/2.0	0.9/0.3	2.0/0.8
Gabor filtering	65.2/36.0	62.5/34.6	65.4/48.8	53.8/42.5	5.4/450.5	54.0/53.5	52.0/109.5
BaseNet	78.6/17.2	78.5/21.6	73.9/21.1	79.5/16.0	3.0/27.4	63.5/31.0	64.1/21.2
BaseNet-bin	83.6/17.6	85.8/16.3	75.2/34.8	79.6/21.9	16.5/231.1	76.7/21.8	70.2/56.9
OFFIENet-ang	83.6/17.0	78.6/16.2	74.1/22.0	79.0/15.6	5.4/29.9	70.2/21.8	66.2/20.0
OFFIENet-tri	85.1/14.0	88.1/13.7	79.4/20.4	83.2/15.6	11.0/44.6	77.6/18.1	71.4/20.7
OFFIENet-tri-shared	84.9/14.5	86.4/12.9	79.6/20.1	81.7/15.8	11.4/42.4	77.1/18.1	70.9/20.4

5. OFFIENet-ang: CNN-based fingerprint enhancement based on network in Fig. 2(b) and use angles as the orientation field labels
6. OFFIENet-tri: CNN-based fingerprint enhancement based on network in Fig. 2(b) and use $\sin 2\theta$ and $\cos 2\theta$ as the orientation field labels
7. OFFIENet-tri-shared: OFFIENet-tri with weight sharing

5.3. Fingerprint image enhancement results

In image restoration, it is common to use the peak signal-to-noise ratio (PSNR) to measure how close the restored image is to the clean image. We perform the same analysis on the enhanced fingerprint image to quantitatively evaluate how well the fingerprints are reconstructed from their corrupted versions. However, in the context of fingerprint enhancement, the ability to preserve features for fingerprint matching is crucial and it does not necessarily concur with high PSNR between the reconstructed image and the clean image. Therefore, we also compute the minutiae recovery rates of the enhanced fingerprints to compare the minutiae extracted from the enhanced image (recovered minutiae) and the original image (true minutiae). Here, we define true minutiae recovery rate (TMRR) as

$$\text{TMRR} = \frac{\text{number of correctly recovered minutiae}}{\text{number of true minutiae}} \quad (5)$$

and false minutiae recovery rate (FMRR) as

$$\text{FMRR} = \frac{\text{number of falsely recovered minutiae}}{\text{number of true minutiae}} \quad (6)$$

Tables 1 and 2 compares the average PSNR and minutiae recovery rates of the methods under evaluation. OFFIENet-tri-shared yields the best PSNR among the methods. However as discussed above, higher PSNR does not guarantee higher fingerprint recognition rate. The minutiae recovery rates measure how much valuable information is extracted from the enhanced fingerprints and is more relevant to minutiae-based fingerprint matching such as MCC. OFFIENet-tri has the highest TMRR compared to other methods, which leads to the prediction that it also has the highest recognition rate later on. Fig. 5 depicts some enhanced fingerprint images imposed with the extracted minutiae for different enhancement methods. From the zoomed-in region in Fig. 6, we can clearly

see that results of Gabor filtering and BaseNet contain falsely detected minutiae while OFFIENet-tri has the outcome almost identical to the ground truth.

We see that Gabor filtering gives even lower PSNR than using the raw images but it is able to recover more minutiae. This is because Gabor filtering results in gray background as shown in Fig. 5 which differs from the original white background, thus increasing the mean-squared error in PSNR computation. This is one case where PSNR can be misleading in the context of fingerprint image enhancement.

Furthermore, both Tables 1 and 2 agree that fingerprints with SN alone are easier to recover than the others as they consist of only one type of noise. On the other hand, SN+OP is the most difficult case. This is because thickened ridges may merge with adjacent ridges and to a certain extent, the ridge structures become unrecoverable. An example of the reconstruction of an over-pressed fingerprint is shown in the first row of Fig. 5.

Apart from the synthetic data, we also perform analysis on the FVC datasets with real fingerprints. Since no ground truth image is available, PSNR, TMRR and FMRR cannot be computed. Based on the visual results in Fig. 7, we can see that OFFIENet-tri produces the most desirable outputs. Comparing the results of BaseNet and OFFIENet-tri, some minutiae exist in the former but not in the latter. These minutiae are spurious minutiae caused by scars or moist on the fingers and should be removed for robust fingerprint matching. In another case, more valuable minutiae are extractable if the contrast between the ridges and valleys is able to be recovered as shown in Fig. 8.

Fig. 9 depicts some of the fingerprints that the proposed method failed to reconstruct. These fingerprints contain severe and large defective areas which are difficult to recover even with human eyes. Although the well-known powerful GAN [45] may be able to make up the missing fingerprint regions, the recovered ridge patterns and minutiae might not represent the original fingerprint.

5.4. Fingerprint recognition accuracy

Since the objective of this work is to improve the performance of AFRS, evaluating the fingerprint recognition accuracy by using the proposed method is very crucial. In Table 3, we compare the

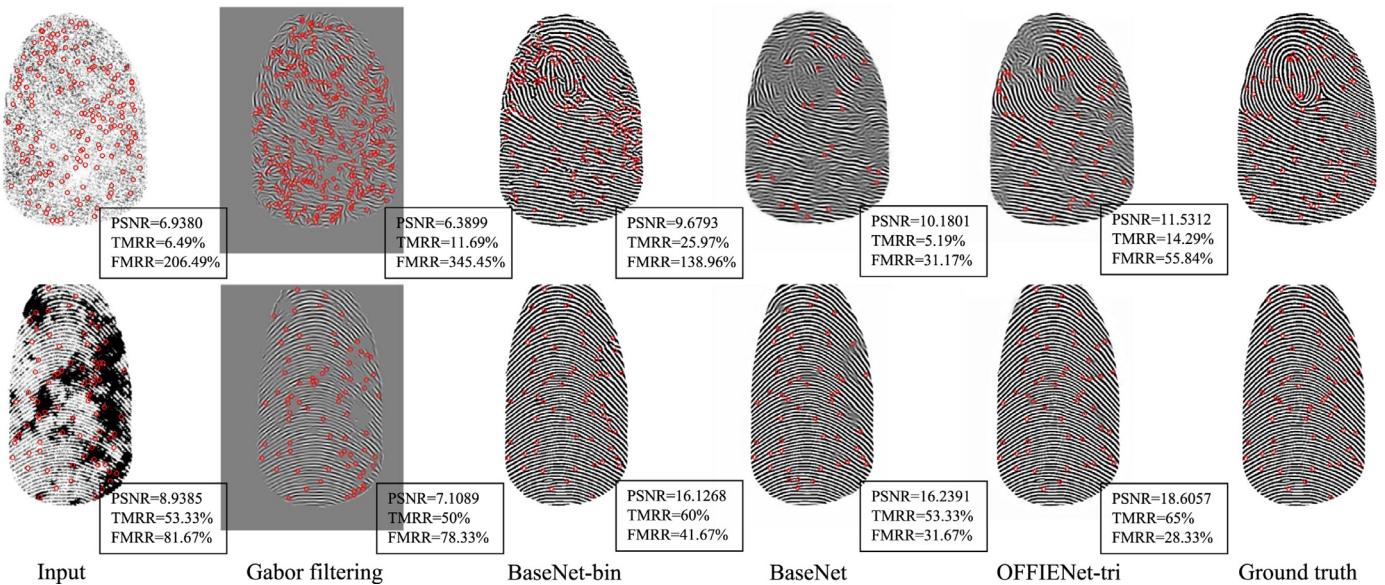


Fig. 5. Comparison between the results of using different enhancement methods on the synthetic testing dataset created using algorithms described in Section 4. The minutiae extracted are marked with red circles. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

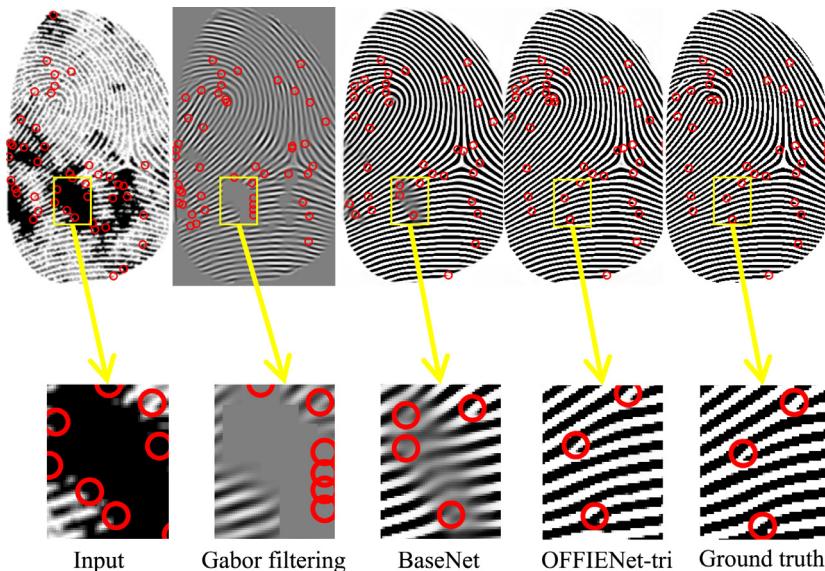


Fig. 6. Comparison between the zoomed-in view of the minutiae extracted from different enhanced synthetic fingerprints.

Table 3

EER (in %) of the proposed two-stage CNN-based fingerprint image enhancement schemes in conjunction with MCC or VeriFinger matching method compared with other existing methods. The values in the first three rows are directly taken from the respective papers.

Method	FVC2002 DB1	FVC2002 DB2	FVC2002 DB4	FVC2004 DB1	FVC2004 DB2	FVC2004 DB4	FVC2006 DB2	FVC2006 DB4
MCC [43]	1.00	0.49	-	-	-	-	0.12	-
Yang et al. [17]	-	-	-	3.12	2.50	4.19	-	-
Sutthiwichaiporn et al. [25]	2.07	0.88	1.53	5.65	5.46	2.59	0.25	1.94
VeriFinger [44]	0.25	0.31	0.35	1.80	0.82	1.06	0.02	8.76
No enhancement + MCC	0.67	0.75	4.57	6.57	8.56	9.52	1.43	9.02
Gabor filtering + MCC	0.71	0.47	2.86	3.97	5.54	5.71	1.35	5.13
BaseNet + MCC	0.57	0.14	1.29	3.43	4.43	1.57	1.04	1.43
BaseNet-bin + MCC	0.71	0.14	1.04	3.88	4.57	1.43	1.82	2.01
OFFIENet-ang + MCC	0.59	0.29	0.86	3.27	4.20	1.60	1.04	1.43
OFFIENet-tri + MCC	0.43	0.14	0.43	2.93	3.43	1.14	0.12	1.09
OFFIENet-tri-shared + MCC	0.47	0.17	0.50	2.85	4.18	1.14	0.12	1.16
OFFIENet-tri + VeriFinger	0.18	0.14	0.25	1.83	3.66	0.32	0.83	0.61

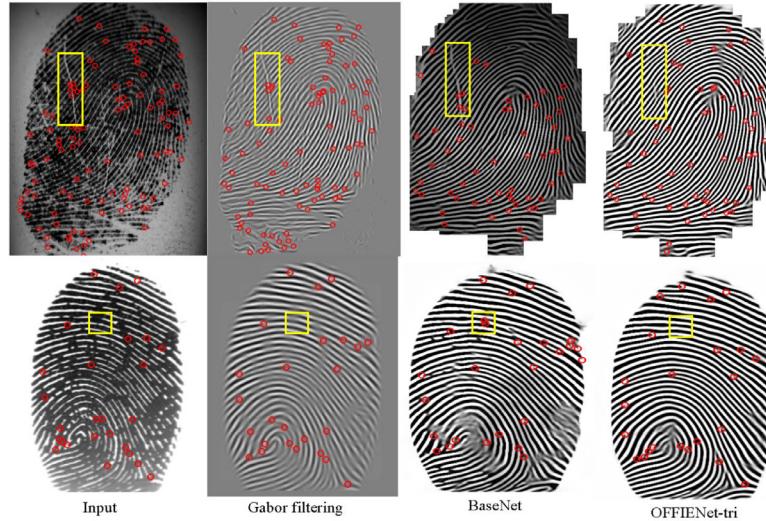


Fig. 7. Comparison between the results of using different enhancement methods on FVC datasets. The minutiae extracted are marked with red circles. Yellow boxes highlight the effect of orientation field correction on removing spurious minutiae. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

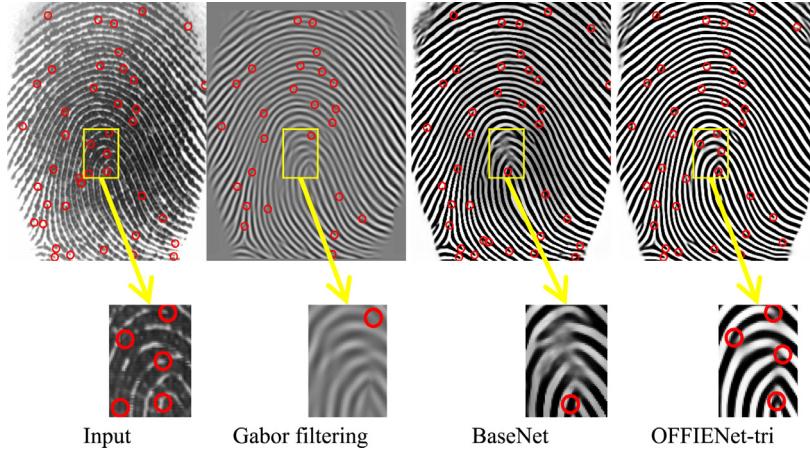


Fig. 8. Comparison between the zoomed-in view of the minutiae extracted from different enhanced real fingerprints.

equal-error rates (EER) of different methods and settings. The combination of using OFFIENet-tri for pre-processing and then applying Verifinger[44] for fingerprint classification outperforms the other methods in most datasets. From the detection error trade-off (DET) curves in Fig. 10, it is evident that OFFIENet-tri yields the best DET. Therefore, ridge orientation information does contribute to better reconstruction of the fingerprint structures. Besides, the degradation in OFFIENet-tri-shared when compared to OFFIENet-tri is marginal, thus showing that weight sharing is able to reduce the number of parameters without compromising much performance. To be specific, the number of parameters is reduced from 6.2M to 4.2M through weight sharing.

The proposed fingerprint enhancement method combined with VeriFinger does not provide the best accuracy for the experiment on FVC2004DB2 and FVC2006DB2 due to the unique background noise and large areas of dark regions that exist in the datasets as depicted in Fig. 11. This kind of background noises and dark regions were not modeled in the synthesis of our training data. Inclusion of such background noise and fingerprint corruption in the training data should improve the performance.

Our results show that gray-scale reconstruction performs better than binary reconstruction by comparing BaseNet and BaseNet-bin. In the case where a fingerprint region fails to be recovered,

BaseNet may output different intensity of gray pixels between black and white while BaseNet-bin forces the pixels to be either black or white. Such hard classification of BaseNet-bin results in unrealistic reconstructed fingerprint and many spurious minutiae.

In order to compare with the results of Schuch's work [8], we compute the area under DET curves (AUC) for OFFIENet-tri on datasets {FVC2002DB1, FVC2002DB2, FVC2004DB1, FVC2004DB2, FVC2006DB2}. We deduce the AUCs after enhancement from the baseline AUCs and the relative AUCs reported in [8] to be {0.1040, 0.1368, 0.38, 0.3174, 0.0680} as opposed to our AUCs {0.0019, 1.3544e-5, 0.0064, 0.0097, 4.5537e-5}, in respective order. Therefore, we prove that the proposed OFFIENet-tri outperforms the referred work with significantly lower AUC of DET curves. The reason might be that the proposed CNN is much deeper than theirs. A deeper network with larger receptive field allows a pixel to be reconstructed based on wider surrounding region, and thus it is more accurate.

5.5. The effect of number of epochs and training samples

We also investigate the effect of using different number of training epochs and sizes of training dataset on the fingerprint recognition performance. We train the proposed CNN with full



Fig. 9. Examples of low-quality fingerprints whose ridge lines cannot be fully recovered by the proposed method.

dataset as described in Section 5.1 as well as partial dataset (7/8, 3/4 and 1/2 of the full dataset) for 10 epochs. From Fig. 12, we can see that the performance converges at six to nine epochs. After convergence, the EER slightly increases. As the training data is synthetically created, it does not completely imitate the samples in testing data. Therefore, “over-training” the CNN might cause the resulting model to overly fit to the training data and lose its adaptability to real fingerprints. Furthermore, the performance is maintained when the training dataset is reduced to 7/8 of its full size, and starts to deteriorate when a quarter or more of the training dataset is discarded.

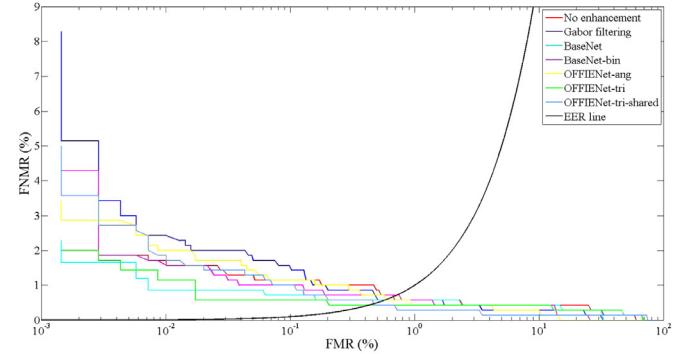


Fig. 11. Examples of low-quality fingerprints in FVC2004DB2. The proposed fingerprint enhancement method failed to recover the fingerprints well for accurate matching.

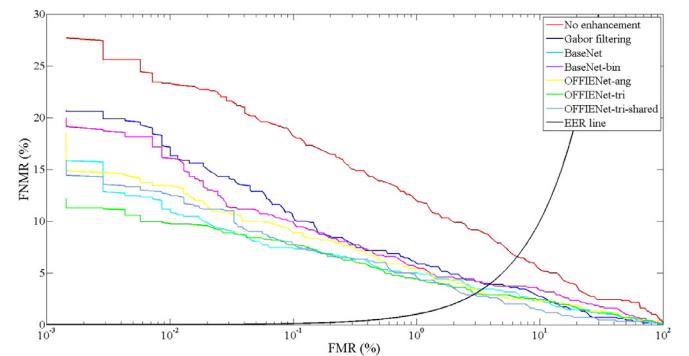
Table 4

Latent-to-Sensor fingerprint matching results (Rank-50 accuracy) on MOLF database by using different fingerprint pre-processing and matching methods. The results for Dabouei et al. are taken directly from [47], but their experiment setting was not completely the same as that used in our setting.

Method	Gallery	Probe	Accuracy (%)
Dabouei et al. [47]	DB1 (Lumidigm)	DB4 (Latent)	70.89
VeriFinger	DB1 (Lumidigm)	DB4 (Latent)	83.96
OFFIENET-tri + VeriFinger	DB1 (Lumidigm)	DB4 (Latent)	89.89
Dabouei et al. [47]	DB2 (Secugen)	DB4 (Latent)	66.11
VeriFinger	DB2 (Secugen)	DB4 (Latent)	76.67
OFFIENET-tri + VeriFinger	DB2 (Secugen)	DB4 (Latent)	84.10



(a)



(b)

Fig. 10. DET curves of the proposed methods when tested on (a) FVC2002DB1 and (b) FVC2004DB1.

5.6. Enhancement on latent fingerprints

Besides fingerprints obtained from optical sensors, we also apply the proposed method on latent fingerprints. We used the IIIT-MOLF database [46] for this experiment. We performed the fingerprint matching for the latent fingerprints by using VeriFinger

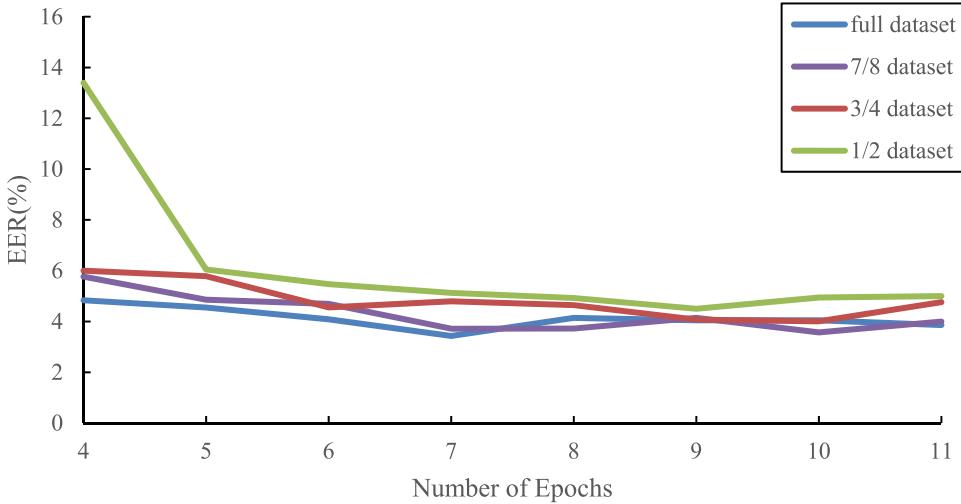


Fig. 12. Performance of the proposed CNN-based fingerprint image enhancement method on FVC2004DB2 with different number of training epochs and training samples.



Fig. 13. Enhancement results of some (a) good quality and (b) low quality latent fingerprints.

matching method [44] with and without applying the proposed OFFIENET enhancement method. We selected the first two fingerprint images for each person in DB1/DB2 dataset as the gallery, and randomly selected 10 images for each person from DB4, the latent image dataset, as the probe set. The matching accuracy is calculated as the percentage of the probe images with the corresponding ground truth within the top rank-50 matching results. Table 4 shows that the proposed fingerprint enhancement method in conjunction with VeriFinger matching considerably improves the matching accuracy for latent fingerprint images. Fig. 13 depicts some latent fingerprint enhancement results. Even though the exact noise patterns that exist in latent fingerprints were not simulated in our training data, the proposed enhancement CNN is still able to recover some parts of the fingerprint area, while it fails in other extremely low contrast and noisy images. This problem can be alleviated by adding more varieties, especially latent fingerprint-like samples, into the training data.

6. Conclusion

In this paper, we presented a multi-task CNN model specifically designed for fingerprint image enhancement, coined as OFFIENET. It embeds the orientation field information into the network for better reconstruction of the fingerprint image. We also created our own synthetic training data for the supervised learning of CNNs. Experimental results showed that the proposed method outperforms existing methods in a number of FVC fingerprint datasets.

We have proven that the proposed CNN is beneficial to the extraction of fingerprint minutiae, thus improving the overall fingerprint recognition EERs. The proposed scheme has demonstrated the feasibility and superiority of using the proposed OFFIENet model for fingerprint image enhancement on rolled and plain fingerprints. The future research direction is to extend this work for application to latent fingerprint images by simulating the noises and corruptions that would appear in such extremely low-quality fingerprints for the model training.

References

- [1] P. Vincent, H. Larochelle, Y. Bengio, P.-A. Manzagol, Extracting and composing robust features with denoising autoencoders, in: International Conference on Machine Learning, ACM, 2008, pp. 1096–1103.
- [2] C. Dong, C.C. Loy, K. He, X. Tang, Image super-resolution using deep convolutional networks, IEEE Trans. Pattern Anal. Mach. Intell. 38 (2) (2016) 295–307.
- [3] J. Kim, J.K. Lee, K.M. Lee, Accurate image super-resolution using very deep convolutional networks, in: IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [4] D. Pathak, P. Krähenbühl, J. Donahue, T. Darrell, A. Efros, Context encoders: Feature learning by inpainting, in: IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [5] L. Jiang, T. Zhao, C. Bai, A. Yong, M. Wu, A direct fingerprint minutiae extraction approach based on convolutional neural networks, in: International Joint Conference on Neural Networks, IEEE, 2016, pp. 571–578.
- [6] Y. Tang, F. Gao, J. Feng, Latent fingerprint minutiae extraction using fully convolutional network, in: International Joint Conference on Biometrics, IEEE, 2017.
- [7] D. Maltoni, Generation of synthetic fingerprint image databases, in: Automatic Fingerprint Recognition Systems, Springer, 2004, pp. 361–384.

- [8] P. Schuch, S. Schulz, C. Busch, De-convolutional auto-encoder for enhancement of fingerprint samples, in: International Conference on Image Processing Theory Tools and Applications, IEEE, 2016, pp. 1–7.
- [9] L. Hong, Y. Wan, A. Jain, Fingerprint image enhancement: algorithm and performance evaluation, *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (8) (1998) 777–789.
- [10] W. Wang, J. Li, F. Huang, H. Feng, Design and implementation of Log-Gabor filter in fingerprint image enhancement, *Pattern Recognit. Lett.* 29 (3) (2008) 301–308.
- [11] C. Gottschlich, Curved-region-based ridge frequency estimation and curved Gabor filters for fingerprint image enhancement, *IEEE Trans. Image Process.* 21 (4) (2012) 2220–2227.
- [12] Y. Mei, B. Zhao, Y. Zhou, S. Chen, Orthogonal curved-line Gabor filter for fast fingerprint enhancement, *Electron. Lett.* 50 (3) (2014) 175–177.
- [13] E. Zhu, J. Yin, G. Zhang, Fingerprint enhancement using circular Gabor filter, in: International Conference on Image Analysis and Recognition, Springer, 2004, pp. 750–758.
- [14] B. Sherlock, D. Monro, K. Millard, Fingerprint enhancement by directional fourier filtering, *IEE Proc.-Vis. Image Signal Process.* 141 (2) (1994) 87–94.
- [15] S. Chikkerur, A.N. Cartwright, V. Govindaraju, Fingerprint enhancement using STFT analysis, *Pattern Recognit.* 40 (1) (2007) 198–211.
- [16] M. Ghafoor, I.A. Taj, W. Ahmad, N.M. Jafri, Efficient 2-fold contextual filtering approach for fingerprint enhancement, *IET Image Process.* 8 (7) (2014) 417–425.
- [17] J. Yang, N. Xiong, A.V. Vasilakos, Two-stage enhancement scheme for low-quality fingerprint images by learning from the images, *IEEE Trans. Hum.-Mach. Syst.* 43 (2) (2013) 235–248.
- [18] M. Zahedi, O.R. Ghadi, Combining Gabor filter and FFT for fingerprint enhancement based on a regional adaption method and automatic segmentation, *Signal Image Video Process.* 9 (2) (2015) 267–275.
- [19] C.-T. Hsieh, E. Lai, Y.-C. Wang, An effective algorithm for fingerprint image enhancement based on wavelet transform, *Pattern Recognit.* 36 (2) (2003) 303–312.
- [20] M.-I. Wen, Y. Liang, Q. Pan, H.-c. Zhang, A Gabor filter based fingerprint enhancement algorithm in wavelet domain, in: IEEE International Symposium on Communications and Information Technology, vol. 2, IEEE, 2005, pp. 1468–1471.
- [21] J.-W. Wang, N.T. Le, C.-C. Wang, J.-S. Lee, Enhanced ridge structure for improving fingerprint image quality based on a wavelet domain, *IEEE Signal Process. Lett.* 22 (4) (2015) 390–394.
- [22] J.S. Bartunek, M. Nilsson, J. Nordberg, I. Claesson, Adaptive fingerprint binarization by frequency domain analysis, in: Asilomar Conference on Signals, Systems and Computers, IEEE, 2006, pp. 598–602.
- [23] J.S. Bartunek, M. Nilsson, B. Sallberg, I. Claesson, Adaptive fingerprint image enhancement with emphasis on preprocessing of data, *IEEE Trans. Image Process.* 22 (2) (2013) 644–656.
- [24] E.-K. Yun, S.-B. Cho, Adaptive fingerprint image enhancement with fingerprint image quality analysis, *Image Vis. Comput.* 24 (1) (2006) 101–110.
- [25] P. Sutthiwichaiorn, V. Areekul, Adaptive boosted spectral filtering for progressive fingerprint enhancement, *Pattern Recognit.* 46 (9) (2013) 2465–2486.
- [26] M.A. Khan, T.M. Khan, D. Bailey, Y. Kong, A spatial domain scar removal strategy for fingerprint image enhancement, *Pattern Recognit.* 60 (2016) 258–274.
- [27] K. Cao, E. Liu, A.K. Jain, Segmentation and enhancement of latent fingerprints: a coarse to fine ridgestructure dictionary, *IEEE Trans. Pattern Anal. Mach. Intell.* 36 (9) (2014) 1847–1859.
- [28] E. Saatci, V. Tavsanoglu, Fingerprint image enhancement using CNN Gabor-type filters, in: IEEE International Workshop on Cellular Neural Networks and Their Applications, IEEE, 2002, pp. 377–382.
- [29] J. Li, J. Feng, C.-C.J. Kuo, Deep convolutional neural network for latent finger-print enhancement, *Signal Process.* 60 (2018) 52–63.
- [30] I. Joshi, A. Anand, M. Vatsa, R. Singh, S.D. Roy, P. Kalra, Latent fingerprint enhancement using generative adversarial networks, in: IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, 2019, pp. 895–903.
- [31] Y. Tang, F. Gao, J. Feng, Y. Liu, FingerNet: an unified deep network for finger-print minutiae extraction, in: IEEE International Joint Conference on Biometrics, 2017.
- [32] H.H. Le, N.H. Nguyen, T.T. Nguyen, Automatic detection of singular points in fingerprint images using convolution neural networks, in: Asian Conference on Intelligent Information and Database Systems, 2017.
- [33] R. Wang, C. Han, Y. Wu, T. Guo, Fingerprint classification based on depth neural network, arXiv:1409.5188, (2014).
- [34] B. Stojanović, O. Marques, A. Nešković, S. Puzović, Fingerprint ROI segmentation based on deep learning, in: Telecommunications Forum (TELFOR), 2016 24th, IEEE, 2016, pp. 1–4.
- [35] K. Cao, A.K. Jain, Latent orientation field estimation via convolutional neural network, in: International Conference on Biometrics, IEEE, 2015, pp. 349–356.
- [36] R. Cappelli, D. Maio, D. Maltoni, An improved noise model for the generation of synthetic fingerprints, in: Control, Automation, Robotics and Vision Conference, 2004. ICARCV 2004 8th, vol. 2, IEEE, 2004, pp. 1250–1255.
- [37] A.M. Bazen, S.H. Gerez, Systematic methods for the computation of the directional fields and singular points of fingerprints, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (7) (2002) 905–919.
- [38] D. Maio, D. Maltoni, R. Cappelli, J.L. Wayman, A.K. Jain, Fvc2002: second finger-print verification competition, in: International Conference on Pattern Recognition, vol. 3, IEEE, 2002, pp. 811–814.
- [39] D. Maio, D. Maltoni, R. Cappelli, J.L. Wayman, A.K. Jain, Fvc2004: third finger-print verification competition, in: Biometric Authentication, Springer, 2004, pp. 1–7.
- [40] J. Fierrez, J. Ortega-Garcia, D.T. Toledano, J. Gonzalez-Rodriguez, Biosec baseline corpus: a multimodal biometric database, *Pattern Recognit.* 40 (4) (2007) 1389–1392.
- [41] C.I. Watson, M.D. Garris, E. Tabassi, C.L. Wilson, R.M. McCabe, S. Janet, K. Ko, User's guide to NIST biometric image software (NBIS), 2007.
- [42] R. Cappelli, M. Ferrara, D. Maltoni, Minutia cylinder-code: a new representation and matching technique for fingerprint recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* 32 (12) (2010) 2128–2141.
- [43] M. Ferrara, D. Maltoni, R. Cappelli, Noninvertible minutia cylinder-code representation, *IEEE Trans. Inf. Forensics Secur.* 7 (6) (2012) 1727–1737.
- [44] Neurotechnology, Verifinger SDK, (<http://www.neurotechnology.com/verifinger.html>).
- [45] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, in: Advances in Neural Information Processing Systems, 2014, pp. 2672–2680.
- [46] A. Sankaran, M. Vatsa, R. Singh, Multisensor optical and latent fingerprint database, *IEEE Access* 3 (2015) 653–665.
- [47] A. Dabouei, S. Soleymani, H. Kazemi, S.M. Iranmanesh, J. Dawson, N.M. Nasrabadi, ID preserving generative adversarial network for partial latent fingerprint reconstruction, in: IEEE 9th International Conference on Biometrics Theory, Applications and Systems, IEEE, 2018.