IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 26, 2019, accepted January 15, 2020, date of publication January 20, 2020, date of current version January 27, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2967771

Modified Conditional Generative Adversarial Network-Based Optical Blur Restoration for Finger-Vein Recognition

JIHO CHOI, KYOUNG JUN NOH, SE WOON CHO, SE HYUN NAM, MUHAMMAD OWAIS, AND KANG RYOUNG PARK

Division of Electronics and Electrical Engineering, Dongguk University, Seoul 04620, South Korea

Corresponding author: Kang Ryoung Park (parkgr@dongguk.edu)

This work was supported in part by the National Research Foundation of Korea (NRF) funded by the Ministry of Education through the Basic Science Research Program under Grant NRF-2018R1D1A1B07041921, in part by the NRF funded by the Ministry of Science and ICT through the Basic Science Research Program under Grant NRF-2019R1A2C1083813, and in part by the NRF funded by the Ministry of Science and ICT through the Basic Science Research Program under Grant NRF-2019R1F1A1041123.

ABSTRACT Among the existing biometrics methods, finger-vein recognition is beneficial because finger-veins patterns are locate under the skin and thus difficult to forge. Moreover, user convenience is high because non-invasive image capturing devices are used for recognition. In real environments, however, optical blur can occur while capturing finger-vein images du to both skin scattering blur caused by light scattering in the skin layer and lens focus mismatch caused by finger movement. The blurred images generated in this manner can cause severe performance degradation for finger-vein recognition. The majority of the previous studies addressed the restoration method o skin scattering blurred images; however, only limited studies have addressed the restoration of optically blurred images. Even the previous studies on the restoration of optical blur restoration have performed restoration based on the estimation of the accurate point spread function (PSF) for a specific image-capturing device. Thus, it is difficult to apply these methods to finger-vein images acquired by different devices. To address this problem, this paperproposes a new method for restoring optically blurred finger-vei images using a modified conditional generative adversarial network (conditional GAN) and recognizing the restored finger-vein images using a deep convolutional neural network (CNN). The results of the experiment performed using two open databases, the Shandong University homologous multimodal traits (SDUMLA-HMT) finger-vein database and Hong Kong Polytechnic University finger-image database (version 1) confirmed that the proposed method outperforms the existing methods.

INDEX TERMS Finger-vein recognition, optical blur image restoration, modified conditional GAN, CNN.

I. INTRODUCTION

Among biometric technologies such as face, iris, fingerprint, and finger-vein recognition, finger-vein recognition has the following benefits [1]. (1) As a finger-vein is hidden inside the body and is typically invisible to the human eye, it is difficult to be forged or stolen. (2) The non-invasive image capturing ensures both convenience and cleanliness and is more acceptable for the user. (3) Because a person has ten fingers, if something unexpected happens in one finger, the other fingers can be used for authentication. In real environments, however, blurred images can be generated while capturing finger-vein images due to the focus mismatch of the finger-vein acquisition camera lens and light scattering in the skin layer. That is, because of the nature of the near-infrared (NIR) light used to acquire the images of the finger-veins present under the skin, skin scattering blur that reduces the sharpness of the acquired finger-vein images caused by light scattering in the tissues and moisture in the skin frequently occurs. Numerous studies have been conducted to improve recognition accuracy by solving skin scattering blur [2]–[8]. Conversely, motion blurring rarely occurs in the input images because the images are captured while the fingers are fixed, to a certain extent, on the finger-vein image capturing devices. Differences in finger thickness among individuals, infants, and adults, and the depth difference from the surface of the

The associate editor coordinating the review of this manuscript and approving it for publication was Inês Domingues.

finger skin to the finger-vein, however, can cause differences in the distance from the camera lens to the finger-vein. This can result in optical blurring and reduces the sharpness of the finger-vein images. Consequently, these skin-scattering or optically blurred images can cause considerable performance degradation for finger-vein recognition. To address this problem, the restoration of blurred finger-vein images is essential.

The majority of the previous studies on finger-vein restoration, however, have focused on skin scattering restoratio; only limited studies have been conducted on optical blur restoration. Moreover, the previous studies on the restoration of optical blur [9] have performed restoration based on the estimation of the accurate point spread function (PSF) for a specific image-capturing device. Thus, it is difficult to apply the results of the study to finger-vein images acquired by different devices. The problems of the previous studies can be solved using a deep convolutional neural network (CNN). This is because it is not necessary for users to estimate PSF, as CNN determines the optimal filters from the weights learned from the training data. Moreover, CNN can be applied to finger-vein images acquired from different environments.

Considering these reasons, in this study, we propose a method of performing optical-blurred finger-vein image restoration using a conditional generative adversarial network (conditional GAN) [10]. Unlike the early generative adversarial network (GAN) [11] that generates images from random vector inputs, the input and target images are defined as paired cases in conditional GAN, and training is performed such that the input image is generated similar to the target image. Difference images are generated using the restored finger-vein images, and finger-vein recognition is performed with these images using deep CNN. The contents of this study are as follows. In Section II, previous studies and their differences from this study are described. In Section III, the contributions of this study are explained. In Section IV, the modified conditional GAN-based blurred-image restoration and finger-vein recognition method proposed in this study are described. In Sections V and VI, experimental results with analysis and conclusions are provided, respectively.

II. RELATED WORKS

Existing finger-vein recognition can be divided into two methods: finger-vein recognition without blur restoration and finger-vein recognition with blur restoration. These two methods can be further divided into the subcategories of non-training-based finger-vein recognition and training-based finger-vein recognition.

A. FINGER-VEIN RECOGNITION WITHOUT BLUR RESTORATION

For the non-training-based finger-vein recognition without blur restoration method, Lee *et al*. proposed a method of extracting and recognizing finger-vein features using a local binary pattern (LBP)-based method [12]. They first performed the alignment of images through affine transform using minutia points extracted from the finger-vein region, and then extracted finger-vein features using LBP. Peng *et al*. used an 8-way Gabor filter created by selecting the optimal parameters [13]. They extracted the finger-vein pattern through the fusion of images where the veins were emphasized among the Gabor filter-applied images; they then evaluated its performance using scale-invariant feature transform (SIFT) matching robust for rotation and shift. These studies [12], [13] have the advantage that the recognition performance is improved when the appropriate filter is applied to the image features; however, performance degradation can occur when the designed filters are applied to finger-vein images with different features. Moreover, when image quality is reduced due to factors such as positional variation of the finger, misalignment, uneven illumination, and shading that can occur during image acquisition, relying on pre-processing to address such problems is required.

To complement the drawbacks of this non-training-based methods, i.e., handcrafted feature-based methods, research on training-based methods were conducted. Wu *et al*. performed dimension reduction and feature extraction through principal component analysis (PCA) and linear discriminant analysis (LDA), and proposed a method of classifying finger-vein images using a support vector machine (SVM) [14]. This method learns various features of the images and is robust for various elements and environmental changes during finger-vein recognition; however, feature extraction and dimension reduction must be preceded for classification. In [15], [16], the performance of finger-vein verification by genuine matching (matching when input and enrolled images are in the same class) and imposter matching (matching when input and enrolled images are in different classes) was evaluated using the difference image of the enrolled and input images as the input to CNN. This method has the advantage of easiness to classify new finger-vein images that have not been trained. Qin *et al*. created vein-pattern maps that distinguished the finger-vein pattern area from the background using seven baseline algorithms [17]. They performed the labeling of the vein and background by calculating the finger-vein feature probability of each pixel on the created vein-pattern maps. After labeling, training was performed using a divided N $\times$ N size of the original image as the input to CNN, and the probability of the input image being a vein pattern was finally calculated. In this method, however, the seven baseline methods were likely to increase the labeling error if the image quality was not acceptable due to factors of illumination change, low resolution, and optical blur, which could lead to recognition performance degradation. In [18], a finger-vein recognition method based on densely connected network (DenseNet) and shift matching, was proposed; however, blurring that occurs during the image acquisition was not considered.

B. FINGER-VEIN RECOGNITION WITH BLUR RESTORATION

The methods explained in Section II.A did not consider blur that could occur during the acquisition of the finger-vein

images. In general, finger-vein image blur can be largely divided into skin scattering blur, optical blur and motion blur. These blurred images cause performance degradation for finger-vein recognition. To address this problem, studies on blur restoration-based finger-vein recognition have been conducted. As explained in Section I, the majority o the blur that occurs during the acquisition of finger-vein images is skin scattering blur and optical blur. Therefore, in this study, the previous studies were divided into skin scattering blur and optical blur, and analyzed. Among the studies on non-training-based skin scattering blur restoration, Lee *et al*. [2] measured the PSF of the skin scattering blur and then restored the skin scattering blur using the constrained least squares (CLS) filter based on fast Fourier transform. However, it is difficult to apply this method if the finger-vein image is optically blurred because this method considers only skin scattering blur. Yang *et al*. [3], [4] designed a biological optical model (BOM) considering light scattering components and proposed a scattering removal algorithm based on the designed model. This method, however, did not consider background tissue, and could degrade performance if the scattering components were not properly measured. Yang *et al*. [5] proposed an algorithm for removing the scattering effect from finger-vein images using the weighted biological optical model (WBOM), anisotropic diffusion and gamma correction (ADAGC), non-scattered transmission map (NSTM), Gabor wavelets, and inter-scale multiplication operation. This method, however, is not invariant for image rotation, scaling, or translation that can occur when it is applied to real environments. Shi *et al*. [6] proposed a scattering removal method for finger-vein images using haze removal techniques. In the study, the anisotropic diffusion method was applied to the original finger-vein image; an averaging filter mask was additionally applied to the generated diffused image to measure the scattering illumination. After the smoothed image generated in this manner, was defined as scattering illumination, finger-vein image restoration was performed based on Koschmieder's law. However, there is a drawback that the scattering factor was only measured approximately. Yang and Bai [7] proposed a blur restoration method by combining depth PSF and BOM. Moreover, [8] proposed a method of effectively removing the scattering effects by measuring the scattering components and transmission maps based on an optical model-based scattering algorithm. The scattering radiation and scattering components, however, were measured approximately, and the simplified optical model could reduce the performance of the finger-vein recognition.

Lee and Park [9] restored finger-vein images using PSF and the CLS filter that considered both optical blur and scattering blur. They effectively restored finger-vein images by measuring both the optical blur and scattering blur components and improved recognition performance. This method, however, requires the accurate prediction of parameters for measuring the two PSFs, and long processing time. To address these drawbacks of previous studies, this study proposes a new method for restoring optically blurred finger-vein images using modified conditional GAN and for recognizing the restored finger-vein images using deep CNN. Table 1 presents a comparison between the proposed method and previous methods.

III. CONTRIBUTIONS

This research is novel compared to the previous studies for the following five reasons.

- Whereas the majority of the previous studies on finger-vein image blur restoration used non-training-based image blur restoration methods, this study proposes finger-vein image blur restoration by applying modified conditional GAN, which is a training-based generative network.
- The previous studies on finger-vein image blur restoration were conducted with a focus on skin scattering blur restoration; optical blur restoration was rarely addressed. This study proposes an optical blur restoration method based on modified conditional GAN.
- The training complexity and time have been reduced and training convergence is improved by training modified conditional GAN for image blur restoration and deep CNN for finger-vein recognition separately.
- Conditional GAN applies random noise to the generator in a form of dropout to prevent deterministic output. The generation of various outputs due to dropout, however, can change the vein pattern of the finger-vein image to be restored, which leads to degradation of recognition performance. Therefore, dropout was removed in this study because the deterministic output is required rather than the various outputs.
- The trained modified conditional GAN and deep CNN developed in this study, and blurred image database according to blur intensity are published in [19] to allow other researchers to perform fair performance evaluations.

IV. PROPOSED METHOD

A. OVERVIEW OF PROPOSED METHOD

Figure 1 is the flowchart of the proposed method. A finger image is captured in Step (1). In Step (2), the captured image is binarized, the broken finger boundaries are restored, and in-plane rotation compensation is conducted. In Step (3), the upper and lower boundaries of the binarized image are detected using a 4×20 mask and noise is removed. After restoring the collapsed area, the final region of interest (ROI) image for CNN input is obtained. In Step (4), the optically blurred image is resized to 256×256 pixels and used as the input to the modified conditional GAN to obtain a restored image as output. In Step (5), the difference image to be the input to CNN is generated using the restored finger-vein image and enrolled image. Then, the output score is obtained by inputting the difference image to deep CNN. Finally,

TABLE 1. Summarized comparisons of previous and proposed methods.

Category			Methods	Advantages	Disadvantages
Finger-vein recognition without blur restoration	Non-training-based		Minutia-based alignment + LBP-based feature extraction [12]	Recognition performance is improved when an appropriate filter is applied to the image features.	Performance degradation occurs when the designed filter is applied to finger-vein images with different features. Recognition performance is influenced by the positional variation of a finger, misalignment, uneven illumination, and shading. High-quality images are required. A Blur that can occur during the acquisition of finger-vein images is not considered.
			Gabor filter + SIFT feature matching [13]		
	Training-based		PCA + LDA + SVM [14]	By learning the features of various images, it is possible to recognize new finger-vein images that have not been trained.	Increased end-to-end processing time compared to the non-training-based method. A Blur that can occur during the acquisition of finger-vein images is not considered.
			Difference image + CNN [15,16]		
			Vein-pattern maps + CNN [17]		
			DenseNet [18]		
Finger-vein recognition with blur restoration	Skin scattering blur restoration	Non-training-based	PSF + CLS filter [2]	Optimal filters and models suitable for finger-vein images are designed by measuring scattering components.	Optical blur restoration is not considered.
			Skin scattering removal based on optical model [3,4]		
			BOM + ADAGC + NSTM + Gabor wavelets [5]		
			Haze removal techniques + Koschmieder's law [6]		
			Depth PSF + BOM [7]		
			Optical model-based scattering removal algorithm [8]		
	Optical blur restoration	Non-training-based	Two PSFs for optical blur + CLS filter [9]	Blurred images are restored considering both optical blur and scattering blur.	Parameters must be accurately predicted for the estimation of two PSFs, and the long processing time is required.
		Training-based	Modified conditional GAN-based method **(proposed method)**	Optical blur restoration. Applicable to blurred images acquired from various devices.	Intensive training is necessary for conditional GAN and deep CNN

in Step (6), finger-vein verification is performed based on the output score.

B. PREPROCESSING, IN-PLANE ROTATION COMPENSATION, AND DETECTION OF FINGER ROI

To remove the background area of the captured image, the image is binarized and obtained as indicated in Figure 2(b). As the background adjacent to the finger region is not completely removed, the background is removed using the Sobel edge detector and area threshold method [20]. A difference image is generated using the edge map created by the Sobel edge detector and the binarized image; then, an image with the background removed is obtained by applying the area threshold method as indicated in Figure 2(c). To compensate for misalignment that degrades the recognition performance, the second-order moments for the binarized mask M (Figure 2(c)) are calculated using Equation (1).

$$s_{11} = \frac{\sum_{(x,y)\in M} (y - m_y)^2 \cdot f(x,y)}{\sum_{(x,y)\in M} I(x,y)}$$
$$s_{12} = \frac{\sum_{(x,y)\in M} (x - m_x)\left(y - m_y\right) \cdot f(x,y)}{\sum_{(x,y)\in M} I(x,y)}$$
$$s_{22} = \frac{\sum_{(x,y)\in M} (x - m_x)^2 \cdot f(x,y)}{\sum_{(x,y)\in M} I(x,y)}, \quad (1)$$

$f(x,y)$ and (m_x, m_y) represent the image pixel value and center coordinates, respectively. Based on these, the rotation angle θ of Equation (2) is calculated to compensate

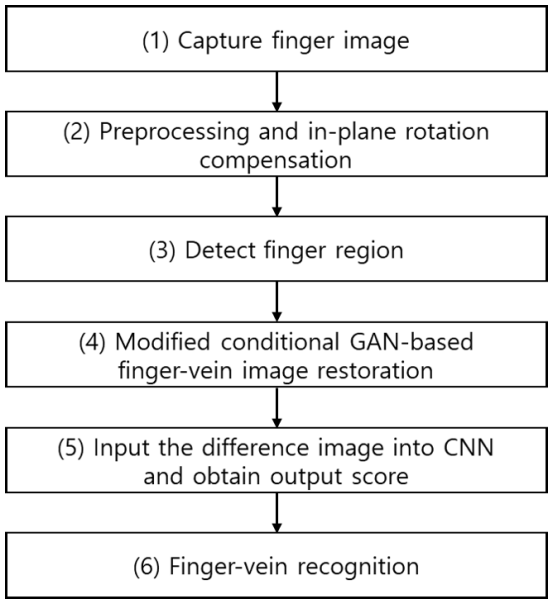

FIGURE 1. Flowchart of the proposed method.

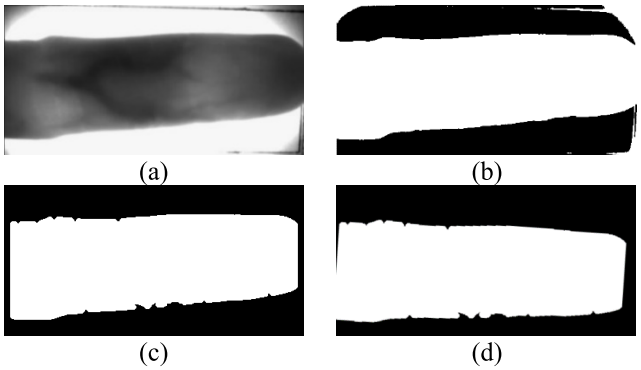

FIGURE 2. Example of the input image and in-plane rotation compensation: (a) original image, (b) binarized image, (c) image by background removal, and (d) image after in-plane rotation compensation.

for the in-plane rotation [21]. That is, a compensated image is obtained through image rotation based on θ and bilinear interpolation as indicated in Figure 2(d).

$$\theta = \begin{cases} tan^{-1}\left\{\dfrac{s_{11} - s_{22} + \sqrt{(s_{11} - s_{22})^2 + 4s_{12}^2}}{-2s_{12}}\right\} \\ \quad if s_{11} > s_{22} \\ tan^{-1}\left\{\dfrac{-2s_{12}}{s_{22} - s_{11} + \sqrt{(s_{22} - s_{11})^2 + 4s_{12}^2}}\right\} \\ \quad if s_{11} \leq s_{22} \end{cases} \quad (2)$$

The left and right ends of Figure 3(a) are thick regions of the finger or regions with a nail that are not well illuminated. These regions are not suitable to be used for recognition because the accurate vein pattern cannot be captured. Therefore, in this study, the left and right sections of an image are removed based on the predetermined size. The

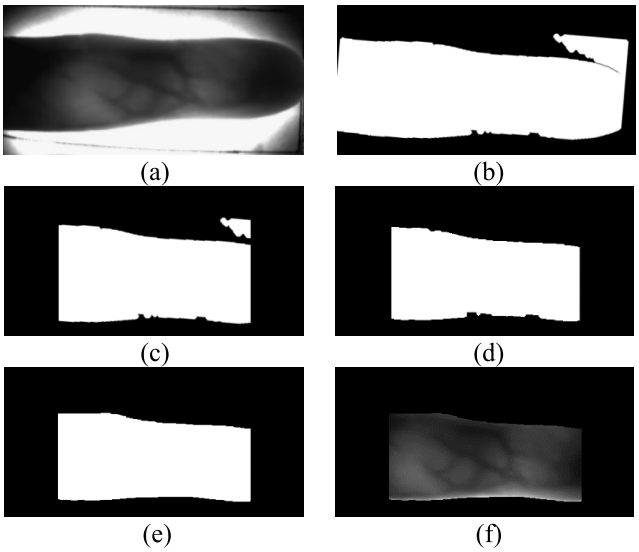

FIGURE 3. Procedure for detecting finger ROI for recognition: (a) original image, (b) image of rotated mask, (c) image after removal of left and right areas of (b), (d) resulting image by component labeling with (c), (e) ROI mask by filling black region inside finger area of (d), and (f) detected ROI image.

noise that has not been removed as shown in Figure 3(c) is then removed using the component labeling method, and an image is obtained as indicated in Figure 3(d). The black area in the finger region is not necessary for recognition because the vein pattern is not captured due to bright lighting. Hence, the black area is filled with the average value of the surrounding pixel positions using a 4 × 20 mask to create an ROI mask (Figure 3 (e)). The created ROI mask is used to obtain the ROI image as shown in Figure 3(f).

C. MODIFIED CONDITIONAL GAN-BASED BLUR RESTORATION OF FINGER-VEIN IMAGE

GAN is a generative model where training is performed to allow two different networks to create fake images similar to real images for their own purposes [11]. GAN consists of two networks, i.e., a generator and a discriminator. The generator is trained to create fake images similar to the real images when random noise variables are received as input, and the discriminator discriminates fake images from the real images. The generator is trained to allow it to create fake images that are more similar to the real images, and the discriminator is trained to allow it to discriminate fake images from the real images more accurately. It is called a generative adversarial network because it employs a structure where the generator and discriminator gradually improve their performances as adversaries to each other.

1) CONDITIONAL GAN ARCHITECTURE

The existing GAN models that use random noise variables as input [11], [2]–[24] do not have sufficient control elements to allow the input to be produced as the desired output within a reasonable period of time, and thus convergence to the

optimal point is difficult. To address this problem, conditional GAN sets a condition that an image with a similar form to the target image is given as input. Conditional GAN was designed to allow the network to create an image similar to the target image when the input image was received. That is, it was designed to produce the desired output from the beginning of training. As the vein-pattern information is extremely important in finger-vein recognition, this study uses the original vein-pattern image as a target and aims to restore blurred finger-vein images similar to the original finger-vein images. Considering these points, we decided that conditional GAN was suitable for image restoration, and the restoration of blurred finger-vein images proceeded using the pix2pix network [10], which applies the concept of conditional GAN.

The generator of the network is designed to have a structure similar to that of U-Net [25]. U-Net was designed to allow the features extracted from the encoder part to be shared with the decoder part by adding skip connections based on the encoder-decoder network [26]. If there is only encoder-decoder without skip connections, an image with different input and output identities could be generated when a blurred finger-vein image is restored. The generator can compensate for the loss of information that occurs during down-sampling by adding skip connections and can produce output with features similar to those of the input. For the generator, leaky rectified linear unit (ReLU) and batch normalization (BN) are used in the encoder par; ReLU, BN, and the hyperbolic tangent (tanh) function are used in the decoder part for structure optimization. Table 2 and Figure 4 show the structure of the generator of the conditional GAN.

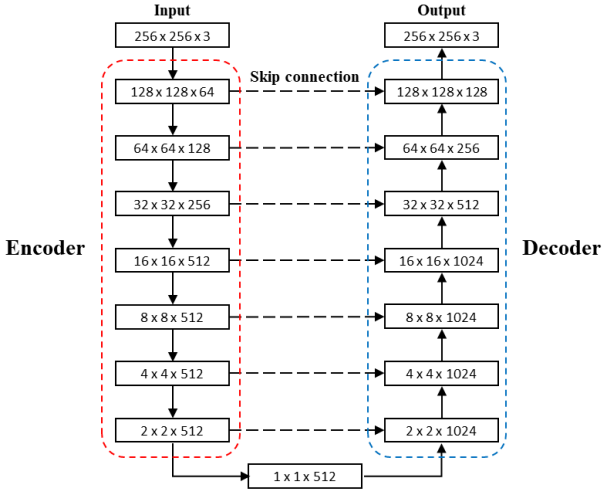

FIGURE 4. Architecture of generator with skip connections.

L1 loss, which is the difference in absolute value between the target image and output image, is used as the generator's loss. This is because the L1 loss can capture low-frequency information in many cases, even though it cannot capture the high-frequency information of an image. L1 loss alone, however, is not suitable to be used for restoration because it is extremely important to capture the vein pattern, which is the high-frequency information of the finger-vein. Conditional GAN allows the discriminator to concentrate on the local patch region to capture the high-frequency information [10], which is defined as patchGAN. The patch of this patchGAN moves within the entire image and determines whether the local region is real or fake. This discriminator is applied to restore the vein pattern in the local region of a finger-vein image more similar to the original image. In this study, this patch scale was set to 70× 70. For structure optimization, leaky ReLU, BN, and the sigmoid function are used in the discriminator. Table 3 and Figure 5 show the structure of the discriminator.

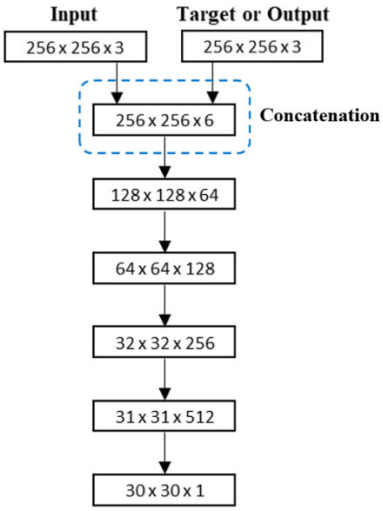

FIGURE 5. Architecture of discriminator.

2) OBJECTIVE FUNCTION AND MODIFIED CONDITIONAL GAN

The ultimate purpose of this study is to train the generating function G that estimates the corresponding deblurred original finger-vein image $\boldsymbol{F}^{ori}$ from the given optically blurred finger-vein image $\boldsymbol{F}^{blur}$. The objective function of the proposed method can be expressed as follows.

$$\mathcal{L}_{cGAN}(G,D) = \mathbb{E}_{\boldsymbol{F}^{blur},\boldsymbol{F}^{ori}}\left[logD\left(\boldsymbol{F}^{blur},\boldsymbol{F}^{ori}\right)\right] + \mathbb{E}_{\boldsymbol{F}^{blur}}\left[log(1-D(\boldsymbol{F}^{blur},G\left(\boldsymbol{F}^{blur}\right))\right] \quad (3)$$

$$\mathcal{L}_{L1}(G) = E_{\boldsymbol{F}^{blur},\boldsymbol{F}^{ori}}\left[\parallel \boldsymbol{F}^{ori} - G\left(\boldsymbol{F}^{blur}\right) \parallel_1\right] \quad (4)$$

The final objective function is as follows.

$$G^* = arg\min_G \max_D \mathcal{L}_{cGAN}(G,D) + \lambda\mathcal{L}_{L1}(G) \quad (5)$$

Conditional GAN applies random noise to the generator in a dropout form to prevent deterministic output. The generation of various outputs due to dropout, however, can transform the vein patterns of the image to be restored, which can cause degradation of the recognition performance. Therefore, modified conditional GAN without dropout is proposed in this study because the deterministic output is required rather

TABLE 2. Discriminator of conditional GAN used for our research.

	Layer	Number of filters	Size of feature map (height × width × channel)	Size of kernel (height × width × channel)	Number of Strides (height × width)	Number of Paddings (height × width)
	Image input layer		256 × 256 × 3			
Encoder	1st convolutional layer	64	128 × 128 × 64	4 × 4 × 3	2 × 2	1 × 1
	Leaky ReLU		128 × 128 × 64			
	2nd convolutional layer	128	64 × 64 × 128	4 × 4 × 64	2 × 2	1 × 1
	Batch normalization		64 × 64 × 128			
	Leaky ReLU		64 × 64 × 128			
	3rd convolutional layer	256	32 × 32 × 256	4 × 4 × 128	2 × 2	1 × 1
	Batch normalization		32 × 32 × 256			
	Leaky ReLU		32 × 32 × 256			
	4th convolutional layer	512	16 × 16 × 512	4 × 4 × 256	2 × 2	1 × 1
	Batch normalization		16 × 16 × 512			
	Leaky ReLU		16 × 16 × 512			
	5th convolutional layer	512	8 × 8 × 512	4 × 4 × 512	2 × 2	1 × 1
	Batch normalization		8 × 8 × 512			
	Leaky ReLU		8 × 8 × 512			
	6th convolutional layer	512	4 × 4 × 512	4 × 4 × 512	2 × 2	1 × 1
	Batch normalization		4 × 4 × 512			
	Leaky ReLU		4 × 4 × 512			
	7th convolutional layer	512	2 × 2 × 512	4 × 4 × 512	2 × 2	1 × 1
	Batch normalization		2 × 2 × 512			
	Leaky ReLU		2 × 2 × 512			
	8th convolutional layer	512	1 × 1 × 512	4 × 4 × 512	2 × 2	1 × 1
	Batch normalization		1 × 1 × 512			
	ReLU		1 × 1 × 512			
Decoder	1st deconvolutional layer	512	2 × 2 × 512	4 × 4 × 512	2 × 2	1 × 1
	Batch normalization		2 × 2 × 512			
	Concatenation (+7th conv)		2 × 2 × 1024			
	ReLU		2 × 2 × 1024			
	2nd deconvolutional layer	512	4 × 4 × 512	4 × 4 × 1024	2 × 2	1 × 1
	Batch normalization		4 × 4 × 512			
	Concatenation (+6th conv)		4 × 4 × 1024			
	ReLU		4 × 4 × 1024			
	3rd deconvolutional layer	512	8 × 8 × 512	4 × 4 × 1024	2 × 2	1 × 1
	Batch normalization		8 × 8 × 512			
	Concatenation (+5th conv)		8 × 8 × 1024			
	ReLU		8 × 8 × 1024			
	4th deconvolutional layer	512	16 × 16 × 512	4 × 4 × 1024	2 × 2	1 × 1
	Batch normalization		16 × 16 × 512			
	Concatenation (+4th conv)		16 × 16 × 1024			
	ReLU		16 × 16 × 1024			
	5th deconvolutional layer	256	32 × 32 × 256	4 × 4 × 1024	2 × 2	1 × 1
	Batch normalization		32 × 32 × 256			
	Concatenation (+3th conv)		32 × 32 × 512			
	ReLU		32 × 32 × 512			
	6th deconvolutional layer	128	64 × 64 × 128	4 × 4 × 512	2 × 2	1 × 1
	Batch normalization		64 × 64 × 128			
	Concatenation (+2th conv)		64 × 64 × 256			
	ReLU		64 × 64 × 256			
	7th deconvolutional layer	64	128 × 128 × 64	4 × 4 × 256	2 × 2	1 × 1
	Batch normalization		128 × 128 × 64			
	Concatenation (+1th conv)		128 × 128 × 128			
	ReLU		128 × 128 × 128			
	8th deconvolutional layer	3	256 × 256 × 3	4 × 4 × 128	2 × 2	1 × 1
	Tanh (output layer)		256 × 256 × 3			

than the various outputs. Finally, the restored image $\boldsymbol{F}^{res}$ from $\boldsymbol{F}^{blur}$ is generated based on the objective function of Equation (5).

D. CNN-BASED FINGER-VEIN RECOGNITION

In this study, the difference image of the restored finger-vein image is used for CNN-based finger-vein recognition.

TABLE 3. Discriminator of conditional GAN used for our research.

Layer	Number of filters	Size of feature map (height × width × channel)	Size of kernel (height × width × channel)	Number of Strides (height × width)	Number of Paddings (height × width)
Image input layer (input + target) or (input + output)		256 × 256 × 6 (concatenation)			
1st convolutional layer Leaky ReLU	64	128 × 128 × 64 128 × 128 × 64	4 × 4 × 6	2 × 2	1 × 1
2nd convolutional layer Batch normalization Leaky ReLU	128	64 × 64 × 128 64 × 64 × 128 64 × 64 × 128	4 × 4 × 64	2 × 2	1 × 1
3rd convolutional layer Batch normalization Leaky ReLU	256	32 × 32 × 256 32 × 32 × 256 32 × 32 × 256	4 × 4 × 128	2 × 2	1 × 1
4th convolutional layer Batch normalization Leaky ReLU	512	31 × 31 × 512 31 × 31 × 512 31 × 31 × 512	4 × 4 × 256	1 × 1	1 × 1
5th convolutional layer	1	30 × 30 × 1	4 × 4 × 512	1 × 1	1 × 1
Sigmoid layer (output layer)		30 × 30 × 1			

The difference image is obtained from the absolute difference between the pixels of the enrolled and input images. As image differencing sensitively responds to changes in two different images, a typically dark image is produced for the same classes due to the small difference between the pixels of the two images and a bright image for different classes due to the relatively large difference between the pixels of the two images. The difference image can express the characteristics of genuine and imposter matching with a single generated image. Here, genuine matching refers to matching when the enrolled image and input image are data of the same class, whereas imposter matching refers to matching when the images are data of different classes. In the case of a finger-vein, it is possible to significantly improve the finger-vein recognition performance through the difference image because the vein patterns between images of the same class has a high similarity whereas the vein patterns between images of a different class has a low similarity. Figures 6 (c) and (f) show examples of generated difference images.

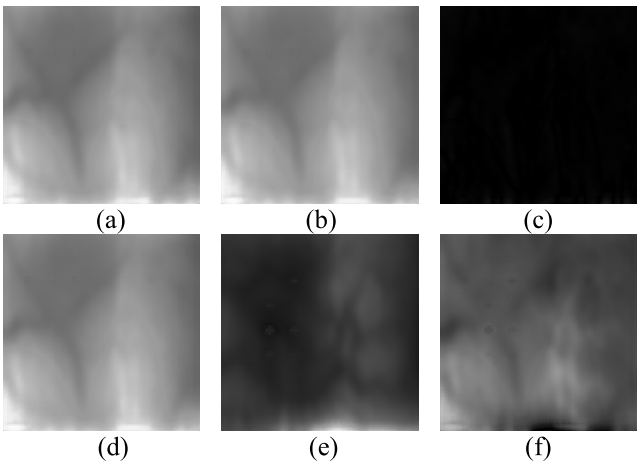

FIGURE 6. Examples of difference images for enrolled and input images: (a) enrolled image and (b) input image from same class, (c) difference image of (a) and (b), (d) enrolled image and (e) input image from different class, and (f) difference image of (d) and (e).

The generated difference image is used as the input to the deep CNN. In this study, the performances of visual geometry group (VGG) Net-16 [27], deep residual network (ResNet)-101 [28], and DenseNet-161 [29] are compared for the recognition of finger-vein images. Each pre-trained model is fine-tuned with the training data of this study. The data used in the training and testing are difference images obtained after resizing the images restored bymodified conditional GAN into 224 × 224, and the output classes of each model are set to two classes: genuine matching and imposter matching.

VGG Net-16 consists of 13 convolutional layers, 5 pooling layers, and 3 fully connected layers (FCL). In the 1st convolutional layer, 64 filters of size 3 × 3 are used. Based on this, the size of the feature map is 224 × 224 × 64 in the 1st convolutional layer, where 224 and 224 are the height and width of the feature map, respectively. They are calculated based on (output height (or width) = (input height (or width) – filter height (or width) + 2 × the number of padding) / the number of strides + 1) [30].

The ResNet model is characterized by applying the identity mapping using shortcut connections. This is effective in preventing the accuracy-saturated and degradation problems that occur as the depth of the network increases. It is also possible to reduce information loss by placing a shortcut connection with the previous layer for vein-pattern feature information that can disappear through convolution and pooling operations. In this study, it is important to reduce the loss of the vein pattern for reasons similar to those for removing the dropout of conditional GAN. Therefore, the shortcut connections of ResNet are important for this study.

In the case of ResNet-101, the depth of the network is high because there are as many as 101 layers. As the

depth increases, the dimension increases, and the subsequent parameters increase. To address this problem, ResNet added a bottleneck layer that includes 1 × 1 convolution operations. The bottleneck layer was designed to reduce the feature map dimension using 1 × 1 size convolution operations, applying 3 × 3 convolution operations. for the feature extraction and increasing the dimension again by applying 1 × 1 convolution operations. This can effectively reduce the feature map dimension, and consequently reduce the computational cost compared to a structure where two 3 × 3 convolution operations are directly connected.

DenseNet uses dense connectivity, which improves the skip-connection structure of ResNet. Dense connectivity is a method for concatenating the feature map of the l^{th} layer and those of previous layers in a dense block while adding a channel. Therefore, the input of the l^{th} layer consists of the feature maps $(x_0, x_1, \ldots, x_{l-1})$ of previous layers. Equation (6) indicates how nonlinear transformation is performed in the l^{th} layer of DenseNet [29].

$$x_l = H_l([x_0, x_1, \ldots, x_{l-1}]) \tag{6}$$

In Equation (6), $[x_0, x_1, \ldots, x_{l-1}]$ is the concatenation of the feature maps from Layer 0 to Layer l. The dense block delivers the features of the previous layer to the subsequent layer through the concatenation of the previous and subsequent layers. This can improve performance while preventing signal attenuation that occurs as the depth of layers increases. As the depth of the network increases, however, the size of the network becomes extremely large because there is an excessive number of channels for the concatenated feature maps. To prevent this, DenseNet added a bottleneck layer between the layers in the dense block. Consequently, the use of the bottleneck structure reduces the computational cost while preventing an increase in the size of the feature map. Nevertheless, the output of the dense block concatenates all the layers in the block. Thus, the size of the feature map becomes extremely large as the depth of layers increases or the number of layers increases in the dense block. To address this problem, the size of the feature map was reduced by adding a transition layer between the dense blocks. The transition layer reduces the number of channels in the feature map by half through a 1 × 1 convolution operation, and decreases the width and height by half through 2 × 2 average pooling. Moreover, DenseNet adjusts the output feature map size by specifying a growth rate. Each layer in the dense block produces a feature map in the size of the growth rate. The growth rate of DenseNet-161 used in this study was 48. The dense block is an advantage in that it can reduce the loss of vein pattern information that can disappear as the depth of the network increases.

Based on the output score obtained after inputting the difference image into CNN, the matching, genuine or imposter, is determined. The following method is used. Genuine matching is concluded if the output score is less than the threshold determined based on the equal error rate (EER) of the genuine matching and imposter matching distribution, which was obtained in advance from the training data. Otherwise, imposter matching is determined. EER is the error rate at the point where the false acceptance rate (FAR), which is the error rate of incorrectly accepting imposter data as genuine data, becomes equal to the false rejection rate (FRR), which is the error rate of incorrectly rejecting genuine data as imposter data. The recognition error rate is measured by the receiver operating characteristic (ROC) curve for FAR and the genuine acceptance rate (GAR). In this case, GAR is calculated as (100 − FRR) (%).

V. EXPERIMENTAL RESULTS

A. TWO OPEN DATABASES FOR EXPERIMENTS

In this study, two types of finger-vein databases were used. The first database was the Hong Kong olytechnic University finger-image database (Version 1), which consisted of two sessions [20]. In Session 1, index and middle-finger images (six images per finger) were acquired from 156 people, amounting to 1,872 images (156 people × 2 fingers × 6 images). In Session 2, index and middle-finger images (six images per finger) were obtained from 105 people from the 156 people of Session 1, amounting to 1,260 images (105 people × 2 fingers × 6 images). In this study, only the images of Session 1 were used. The second database was the Shandong University homologous multimodal traits (SDUMLA-HMT) finger-vein database [31]. The SDUMLA-HMT finger-vein database acquired the index, middle, and ring-finger images (six images per finger) of both hands from 106 people, which produced 3,816 images (106 people × 2 hands × 3 fingers × 6 images). In this study, the Hong Kong olytechnic University finger-image database (Version 1) is referred to as PolyU-DB and the SDUMLA-HMT finger-vein database as SDU-DB.

In the experiment, two-fold cross-validation was performed. The PolyU-DB and SDU-DB databases were composed of 312 and 636 classes, respectively. In the 1st fold validation, the images of 156 classes were used for training and the remaining images of 156 classes were used for testing in the case of PolyU-DB. In the case of SDU-DB, the images of 318 classes were used for training and the remaining images of 318 classes were used for testing. For blurred images, a database was constructed in this study by applying Gaussian blur to PolyU-DB and SDU-DB because there was no open database for optically blurred finger-vein images. Gaussian blur was generated with filter sizes of 11× 11 (standard deviation of 11), 15 × 15 (standard deviation of 15), and 19 × 19 (standard deviation of 19). As mentioned, the experiment was performed using two-fold cross-validation. Therefore, in the 2nd fold validation, training and testing were performed again by switching the training and testing data used in the 1st fold validation. This prevented the data of the same classes from being used again in training and testing (open-world setting). The average of the accuracy measured from these two experiments was used as the final recognition accuracy.

TABLE 4. Descriptions of experimental databases with augmented data.

			PolyU-DB	SDU-DB
Original images		# of images	1,872	3,816
		# of people	156	106
		# of hands	1	2
		# of fingers	2 (index and middle fingers)	3 (index, middle, and ring fingers)
		# of classes (# of images per class)	312 (6)	636 (6)
Training for 1st or 2nd fold cross validation	Training of conditional GAN	# of images (original + augmented data)	13,104 (6 images × 14 times × 156 classes)	26,712 (6 images × 14 times × 318 classes)
	Training of CNN for finger-vein recognition	# of images for genuine matching	12,012 ((6 images × 13 times - 1) × 156 classes)	24,486 ((6 images × 13 times - 1) × 318 classes)
		# of images for imposter matching	12,012	24,486

B. DATA AUGMENTATION AND EXPERIMENTAL SETUP

The two databases (PolyU-DB and SDU-DB) used in this study had a small number of images, which were not sufficient for training a large number of weights in the deep CNN structure. Therefore, an overfitting problem could occur. To address this problem, a data augmentation method was performed to increase the amount of training data. By using the data augmentation method, the number of images increased by a factor of 14 times, including the original images, through a 1-3 pixel translation of the images in the up, down, left, right, and overall directions. Table 4 describes the original and training images of PolyU-DB and SDU-DB. As explained in Section V.A, half of the total classes of each database were used as training data and the other half as testing data. The data augmentation method generated 78 images, which was 13 times as many as the original 6 images per class. One of these 78 images was selected to be the enrolled image; the remainder became input images. Using the data obtained from these, training based on genuine and imposter matching was performed. As the number of imposter matches is typically greater than the number of genuine matches, there is a problem in that the finger-vein recognition CNN is biased to imposter matching data. Therefore, in this study, imposter matching images were randomly selected to match the number of genuine matching images and were used as the training data. This augmentation method was applied to both databases in the same manner. The method was applied only to the training data; the original images without augmentation were used for testing data. To restore optical blur for the finger-vein images, optically blurred finger-vein images were required in this study. As there was no optically blurred data acquired from the real environments, it was necessary to generate blurred images. In this study, optically blurred images were generated by applying Gaussian blur, which is similar to the optical-blur effect, to the original finger-vein images.

Gaussian filtering is commonly used for the removal of noise in images and it generates smoothing and blurry output [32]. As the blur intensity does not occur constantly when images are acquired in real environments, images were generated by applying differing blur intensities.

The generated optically blurred image $\boldsymbol{F^{blur}}$ and original image $\boldsymbol{F^{ori}}$ were used as the input and target of the conditional GAN, respectively, as indicated in equations (3) and (4).

In this study, training and testing were performed on a desktop computer environment including an Intel®Core™ i7-3770K CPU @ 3.5 GHz (4 cores) with 16 GB RAM and NVIDIA GeForce GTX 1070 (1920 compute unified device architecture (CUDA) cores) graphic processing unit (GPU) card with graphics memory of 8 GB [33]. An algorithm was implemented using the Caffe framework [34], TensorFlow framework (version 1.9.0) [35], Python 3.6.8 Version [36], and compute unified device architecture (CUDA) (Version 9.0) [37] with CUDA deep neural network library (CUDNN) (version 7.1.4) [38].

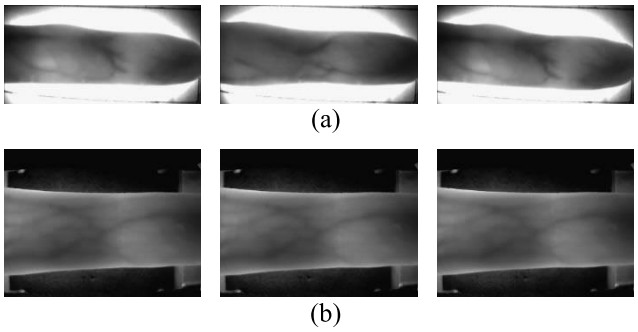

FIGURE 7. Example images of different trials from the same finger of one individual from each database: (a) PolyU-DB and (b) SDU-DB.

C. TRAINING OF MODIFIED CONDITIONAL GAN MODEL FOR BLUR RESTORATION

The number of maximum epochs, mini-batch size, and learning rate, which were used as the training parameters of themodified conditional GAN, were set to 40, 4, and 0.002, respectively. During training, random jittering was applied to the scale [10]. That is, the input image of the size of 256× 256 pixels was resized to 286× 286 pixels and then randomly cropped to 256× 256 pixels to be used as the input to the network. Moreover, in this study, the adaptive moment estimation (Adam) optimization [39] method was used for the training of the modified conditional GAN model. Figure 8 shows the discriminator loss and generator L1 loss, which are the training loss of the modified conditional GAN based on the epoch. As indicated in Figure 8, the loss values converged as the number of epochs increased. This indicates that the modified conditional GAN used in this study was fully trained.

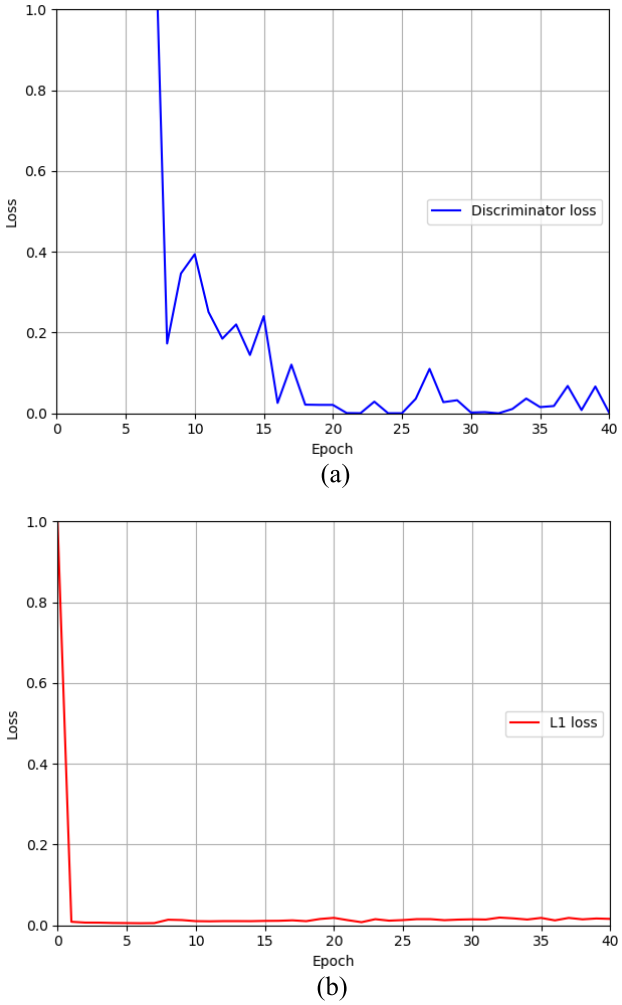

FIGURE 8. Examples of loss curves of modified conditional GAN with training data of SDU-DB: (a) discriminator loss and (b) generator L1 loss.

D. TRAINING OF CNN MODEL FOR FINGER-VEIN RECOGNITION

For the training of the CNN model for finger-vein recognition, the stochastic gradient descent (SGD) method [40] was used. The SGD method reduces the learning rate by multiplying the learning rate by the gamma value for each step size in mini-batch units and achieves the rapid convergence of the training accuracy and loss. The value obtained by dividing the number of total training data by the mini-batch size is referred to as the number of iterations, and the moment when the training has completed the number of iterations is referred to as 1 epoch.

Therefore, the number of maximum iterations is identical to the number obtained by multiplying the number of iterations by the epoch. As explained in Section IV.D, the performances of VGG Net-16, ResNet-101, and DenseNet-161 were compared in this study for the CNN model for finger-vein recognition. For the training parameters of VGG Net-16, the number of output classes was 2 (authentic, imposter), number of maximum epochs was 25, mini-batch size was 24, learning rate was 0.001, step size was 16 epochs, momentum was 0.9, and gamma value was 0.1. For the training parameters of ResNet-101, the number of output classes was 2 (authentic, imposter), number of maximum epochs was 25, mini-batch size was 4, learning rate was 0.001, step size was 16 epochs, momentum was 0.9, and gamma value was 0.1. Lastly, for the training parameters of DenseNet-161, the number of output classes was 2 (authentic, imposter), number of maximum epochs was 25, mini-batch size was 4, learning rate was 0.001, step size was 16 epochs, momentum was 0.9, and gamma value was 0.1. Figures 9 and 10 show the graphs of the training loss and accuracy in the VGG Net-16, ResNet-101, and DenseNet-161 models, which used the difference images of the images restored by the modified conditional GANs obtained from PolyU-DB and SDU-DB as input. As can be observed from these training graphs, the training loss virtually converged to zero and the accuracy essentially converged to 100.

E. TESTING RESULT WITH POLYU-DB

Figure 11 shows examples of finger-vein images restored through modified conditional GA. As indicated in Figure 11,

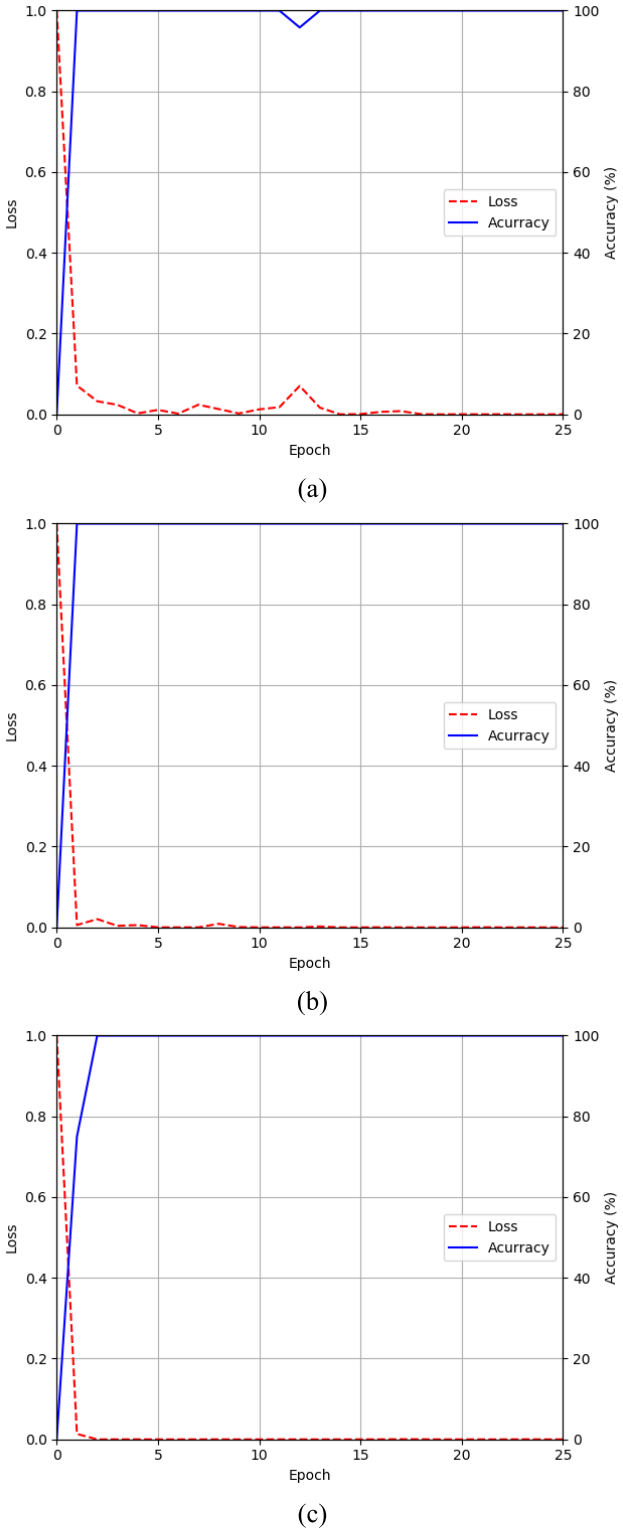

FIGURE 9. Examples of loss and accuracy curves of (a) VGG Net-16, (b) ResNet-101, and (c) DenseNet-161 with training data of PolyU-DB restored by modified conditional GAN.

the finger-vein images optically blurred by different blurring kernels were well restored by the modified conditional GAN proposed in this study, and clearer finger-vein patterns closer to the original images were observed.

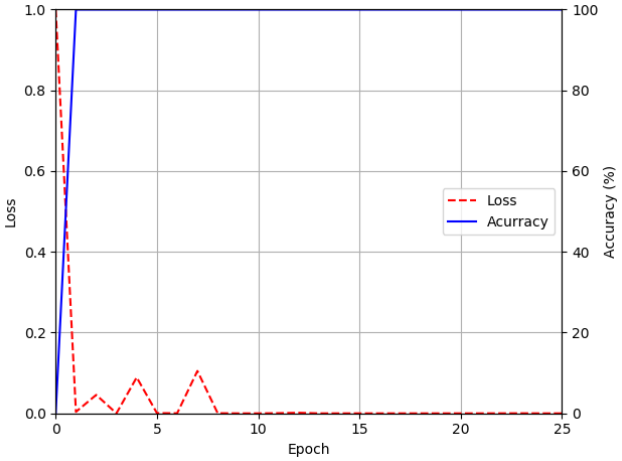

FIGURE 10. Example of loss and accuracy curves of ResNet-101 with training data of SDU-DB restored by modified conditional GAN.

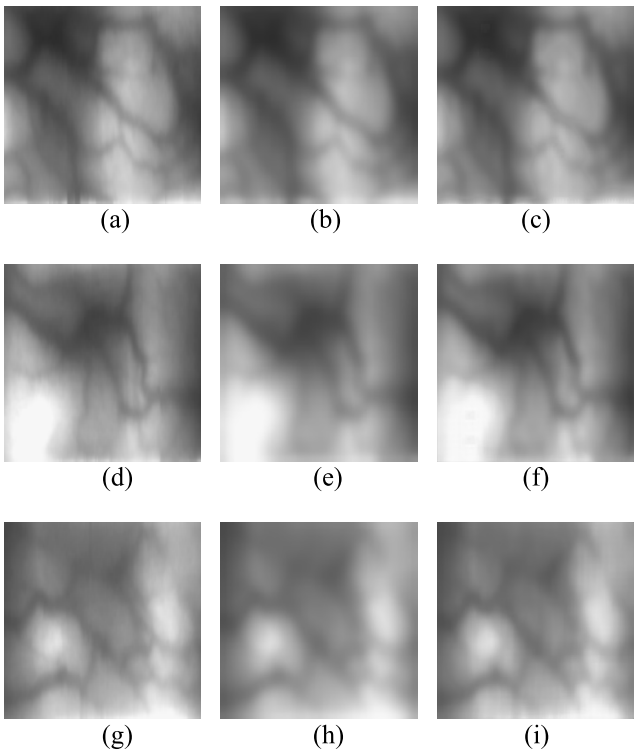

FIGURE 11. Results of image blur restoration by modified conditional GAN: (a), (d), and (g) original images. (b), (e), and (h) blurred images by Gaussian filter size 11 × 11 (standard deviation of 11), 15 × 15 (standard deviation of 15), and 19 × 19 (standard deviation of 19), respectively. (c), (f), and (i) restored images of (b), (e), and (h), respectively.

In the next experiment, the EERs of the finger-vein recognition based on without blurring and with different blur intensities were measured as indicated in Table 5. The experiments were performed using four methods. In scheme 1, the training was performed using the original training data without blurring and the EER of the finger-vein recognition was measured using the remaining testing data. In scheme 2, the EER of finger-vein recognition was measured with blurred testing

data using CNN classifier trained with the original training data without blurring. In scheme 3, training was performed using blurred training data and the EER of the finger-vein recognition was measured using the remaining blurred testing data. In scheme 4, training was performed using the training data restored by the proposed modified conditional GAN and the EER of the finger-vein recognition was measured using the remaining testing data restored by the modified conditional GAN. As indicated in Table 5, when the input image was blurred, the EER of the finger-vein recognition increased compared to no blurring. As the blur intensity increased (as the Gaussian filter size and standard deviation increased), the EER of the finger-vein recognition increased. It would appear that blurring reduced the distinctiveness between the vein pattern and other regions, and thus reduced the recognition accuracy as indicated in Figure 11. Moreover, in all cases, when the training was performed using the training data restored by the proposed modified conditional GAN and recognition was performed using the remaining testing data restored by the modified conditional GAN (scheme 4), the recognition accuracy was greater compared to schemes 2 and 3.

Figure 12 shows the recognition performances of schemes 2, 3, and 4 based on the blur intensities for PolyU-DB with receiver operating characteristics (ROC) curves. As indicated in Figures 12 (a), (b), and (c), when training and recognition were performed with the training data and testing data that were restored using the modified conditional GAN, scheme 4 is greater recognition accuracy than schemes 2 and 3 in all cases.

In Table 6, the performance of the proposed modified conditional GAN-based restoration was compared with the existing state-of-the-art restoration methods using different CNN recognition models. The state-of-the-art restoration methods included DeblurGAN [41], cycle-consistent adversarial networks (CycleGAN) [42], and original conditional GAN [10]. For fair performance evaluation, the same CNN models for finger-vein recognition were used in all cases, and the recognition accuracy was measured using the aforementioned scheme 4 method. As indicated in Table 6, when the proposed modified conditional GAN-based restoration, proposed in this study, was performed, the finger-vein recognition accuracy was greater compared to the existing state-of-the-art restoration methods.

In Table 7, the recognition accuracy was compared for the aforementioned schemes 1 through 4 using different CNN models for finger-vein recognition. The compared CNN models for finger-vein recognition included VGG Net-16 [27], ResNet-101 [28], andDenseNet-161 [29]. As indicated in Table 7, in all cases and all CNN models, the proposed cheme 4 method is greater recognition accuracy than chemes 2 and 3.

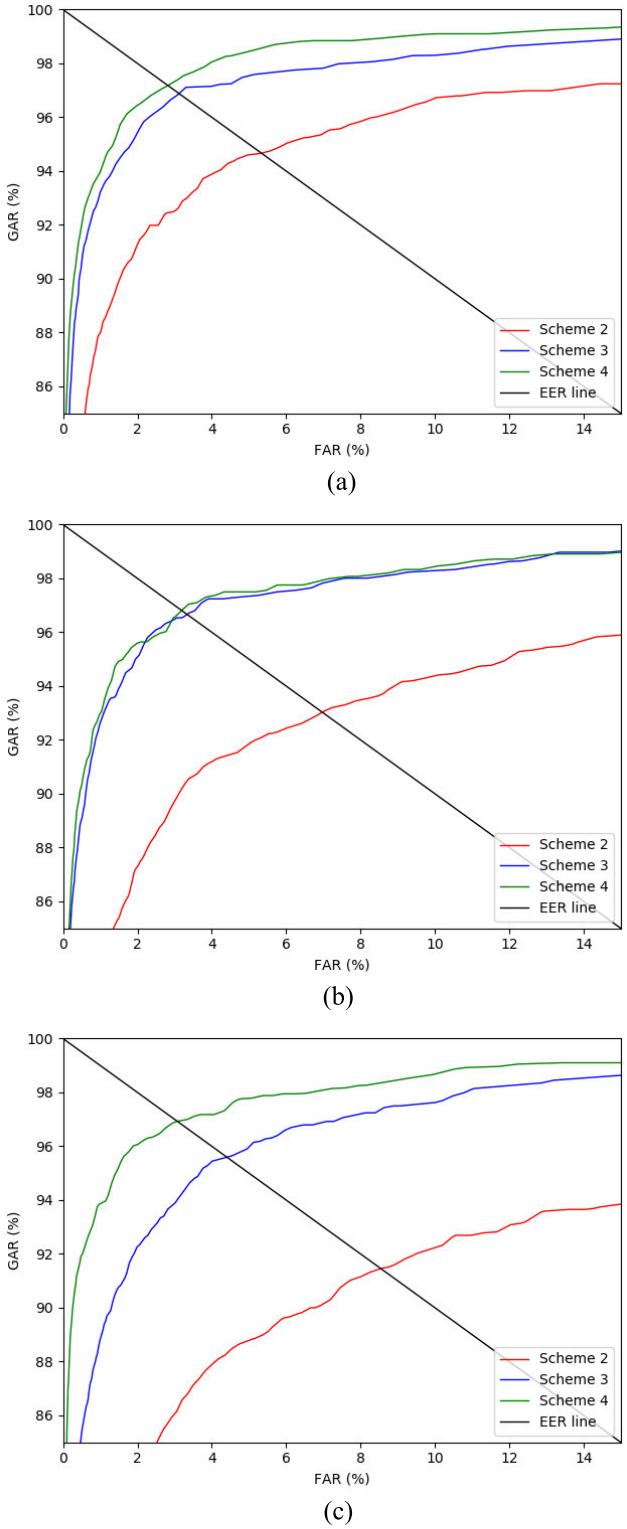

FIGURE 12. ROC curves of finger-vein recognition of schemes 2–4 based on different blur intensities in PolyU-DB for Gaussian filter size: (a) 11 × 11 (standard deviation of 11), (b) 15 × 15 (standard deviation of 15), and (c) 19 × 19 (standard deviation of 19).

F. TESTING RESULT WITH SDU-DB

To verify the performance of the proposed method in different finger-vein database environments, the testing performance was measured using SDU-DB, the second open database. As shown in Table 8, the EERs of the finger-vein recognition based on without blurring and with different blur

TABLE 5. Comparisons of finger-vein recognition error (EER) with PolyU-DB based on no blurring and different blur intensities (unit: %).

Gaussian blur intensity (filter size, standard deviation)	Method			
	Training & testing with original unblurred images (scheme 1)	Testing blurred images without training (scheme 2)	Training & testing with blurred images (scheme 3)	Training & testing with restored images (scheme 4) (proposed method)
11 × 11, 11	1.082	5.422	3.088	2.746
15 × 15, 15		6.880	3.294	3.023
19 × 19, 19		8.587	4.254	3.103

TABLE 6. Comparisons of finger-vein recognition error (EER) in the case of training and testing with restored images (scheme 4) by four different restoration methods using different CNNs with PolyU-DB blurred images by Gaussian blurring of 19 × 19 (standard deviation of 19) (unit: %).

CNN model	DeblurGAN [41]	CycleGAN [42]	Original conditional GAN [10]	Modified conditional GAN (proposed method)
VGG Net-16 [27]	5.068	5.188	5.163	3.189
ResNet-101 [28]	3.255	4.827	3.475	3.103
DenseNet-161 [29]	1.939	3.203	1.996	1.814

TABLE 7. Comparisons of finger-vein recognition error (EER) for three different CNNs for finger-vein recognition with PolyU-DB blurred images by Gaussian blurring of 19 × 19 (standard deviation of 19) (unit: %).

CNN model	Method			
	Training & testing with original unblurred images (scheme 1)	Testing blurred images without training (scheme 2)	Training & testing with blurred images (scheme 3)	Training & testing with restored images (scheme 4) (proposed method)
VGG Net-16 [27]	3.425	7.913	5.137	3.189
ResNet-101 [28]	1.08	8.587	4.254	3.103
DenseNet-161 [29]	1.616	7.238	1.999	1.814

TABLE 8. Comparisons of finger-vein recognition error (EER) with SDU-DB based on no blurring and different blur intensities (unit: %).

Gaussian blur intensity (filter size, standard deviation)	Method			
	Training & testing with original unblurred images (scheme 1)	Testing blurred images without training (scheme 2)	Training & testing with blurred images (scheme 3)	Training & testing with restored images (scheme 4) (proposed method)
11 × 11, 11	3.338	5.816	4.275	3.886
15 × 15, 15		6.275	4.099	3.891
19 × 19, 19		6.633	4.413	3.934

intensities were measured. The experiments were performed in schemes 1 through 4 as mentioned above. As indicated in Table 8, when the input image was blurred, the EER of the finger-vein recognition increased compared without blurring.

Moreover, as the blur intensity increased (as the Gaussian filter size and standard deviation increased), the EER of the finger-vein recognition increased. It would appear that blurring reduced the distinctiveness between the vein pattern

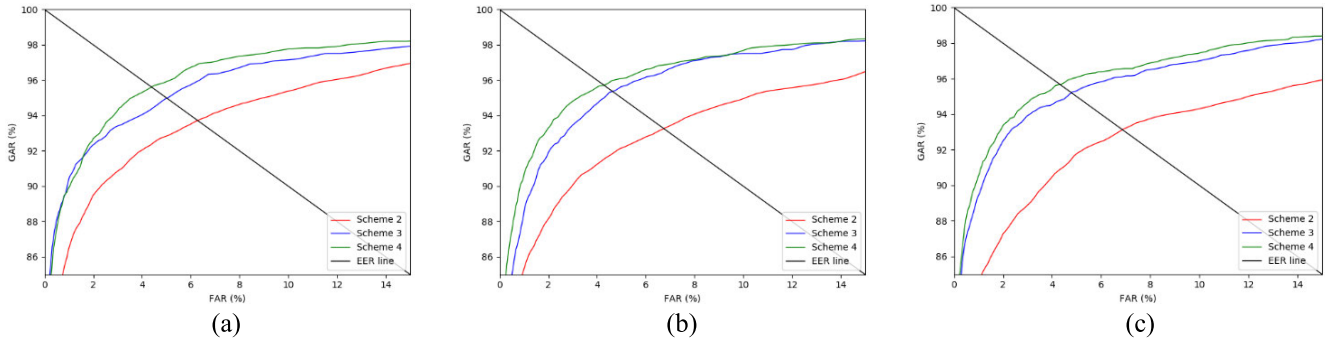

FIGURE 13. ROC curves of finger-vein recognition of schemes 2 4 according to various blur intensities in SDU-DB. In the case of Gaussian filter size (a) 11×11 (standard deviation of 11), (b) 15×15 (standard deviation of 15), and (c) 19×19 (standard deviation of 19).

TABLE 9. Comparisons of finger-vein recognition error (EER) in the case of training and testing with restored images (scheme 4) by four different restoration methods with SDU-DB blurred images by Gaussian blurring of19 × 19 (standard deviation of 19) (unit: %).

DeblurGAN [41]	CycleGAN [42]	Original conditional GAN [10]	Modified conditional GAN (proposed method)
3.985	8.831	4.198	3.934

TABLE 10. Comparisons of finger-vein recognition error (EER) for finger-vein recognition with SDU-DB blurred images by Gaussian blurring of 19 × 19 (standard deviation of 19) (unit: %).

CNN model	Method			
	Training & testing with original unblurred images (scheme 1)	Testing blurred images without training (scheme 2)	Training & testing with blurred images (scheme 3)	Training & testing with restored images (scheme 4) (proposed method)
ResNet-101 [28]	3.338	6.633	4.413	3.934

and other regions, and thus reduced the recognition accuracy. Moreover, in all cases, when training was performed using the training data restored by the proposed modified conditional GAN and recognition was performed using the remaining testing data restored by the modified conditional GAN (scheme 4), the recognition accuracy was greater compared to schemes 2 and 3.

Figure 13 shows the recognition performances of schemes 2, 3, and 4 based on the blur intensities for SDU-DB with ROC curves. As indicated in Figures 13 (a), (b), and (c), when training and recognition were performed with the training data and testing data that were restored as modified conditional GAN, scheme 4is greater recognition accuracy than schemes 2 and 3 in all cases.

In Table 9, the performance of the proposed modified conditional GAN-based restoration was compared with the existing state-of-the-art restoration methods [10], [41], [42]. For fair performance evaluation, the same CNN for finger-vein recognition was used in all cases, and the recognition accuracy was measured using the scheme 4 method mentioned above. As indicated in Table 9, when the proposed modified conditional GAN-based restoration was performed, the finger-vein recognition accuracy was greater compared to the existing state-of-the-art restoration methods.

In Table 10, the recognition accuracy is compared for the aforementioned schemes 1 through 4 for the finger-vein recognition [28]. As indicated in Table 10, in all cases and all CNN models, the proposed cheme 4 method is greater recognition accuracy compared to chemes 2 and 3.

Figures 14 (a) and (c) show authentic and imposter matching before restoration, respectively, and the vein pattern is hardly distinguishable due to optical blurring. These represent the case of incorrect recognition. Figures 14 (b) and (d) show authentic and imposter matching after restoration, respectively, and represent the case of correct recognition by solving the incorrect recognition problem that occurred in Figures 14 (a) and (c) with the proposed method.

Figure 15 shows an example of false rejection cases and false acceptance cases despite the optical blur restoration method. In the case of false rejection such as Figure 15 (a) and (b), the vein pattern is not clear because of the darkness of the vein region of the input image, and the degree of misalignment is extremely large. In this case, the problems could not be solved by the modified conditional

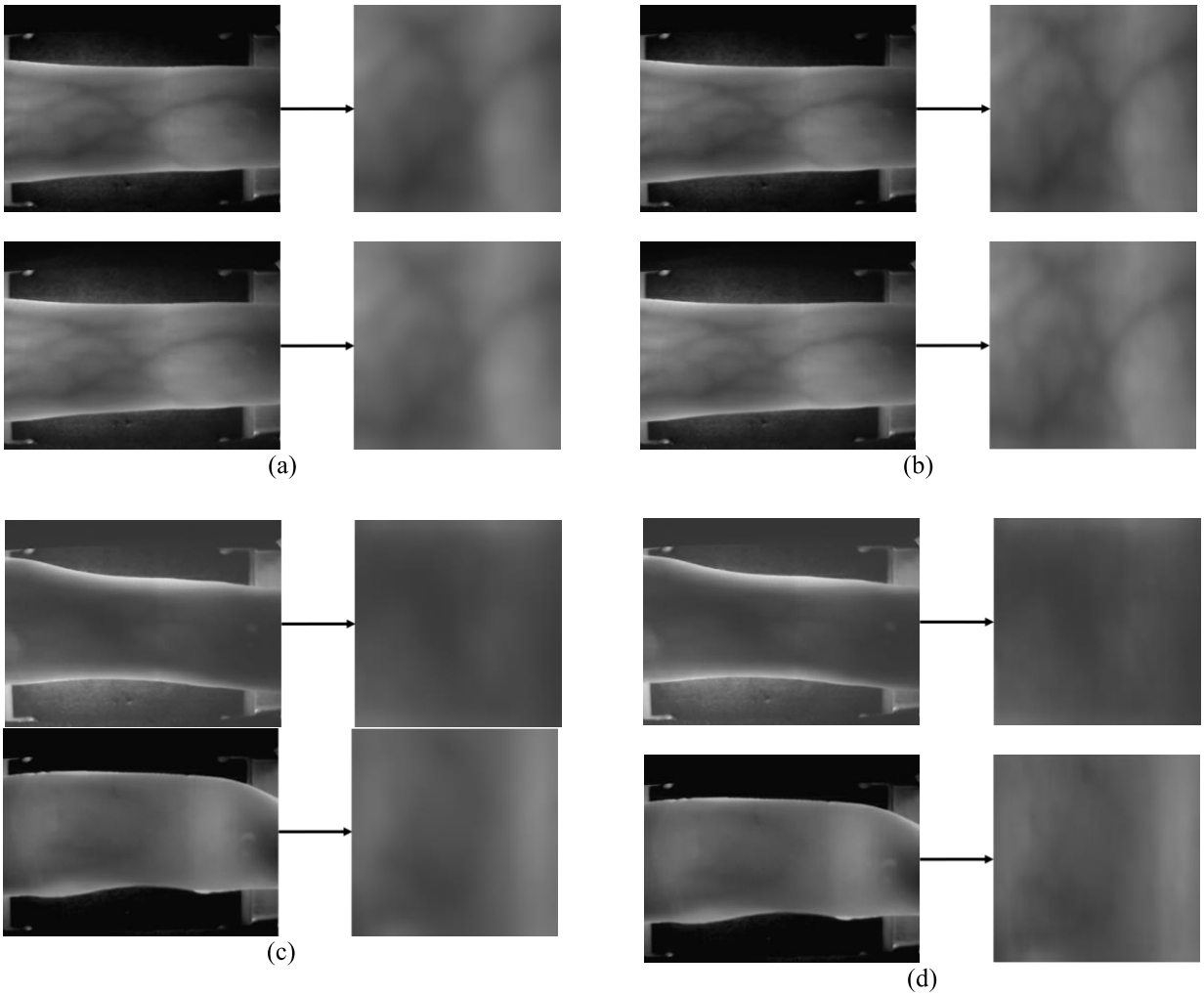

FIGURE 14. Examples of the correct recognition cases before optical blur restoration and correct recognition cases after optical blur restoration. (a) Incorrect authentic matching before restoration, (b) correct authentic matching after restoration, (c) incorrect imposter matching before restoration and (d) correct imposter matching after restoration. Upper and lower images in (a), (b), (c) and (d) show the enrolled and input images, respectively. Images are original and ROI images from the left.

GAN-based restoration proposed in this study. In the case of false acceptance such as Figure 15 (c) and (d), the vein patterns were not clear because of the extreme darkness of the vein region of the input image. Moreover, the shades of the enrolled and input images were similar. Therefore, it would appear that the incorrectly recognized results were a consequence of similar shades.

G. PROCESSING TIME OF PROPOSED METHOD

In the next experiment, the processing speed of the modified conditional GAN + CNN for the finger-vein recognition method proposed in this study was measured. The measurement was performed on the desktop computer described in Section V.B and the Jetson TX2 embedded system [43] shown in Figure 16. The reason that the processing speed was also measured in an embedded system was to verify whether on-board computing was possible for the proposed system. The majority of access-control finger-vein recognition systems are used in on-board computing (edge computing) environments, which operate in an embedded system attached to an entrance, rather than in a desktop computer-based server-client computing (cloud computing) environment. The Jetson TX2 includes NVIDIA Pascal^{TM}-family GPU (256 CUDA cores) with 8 GB of memory shared between the central processing unit (CPU) and GPU, and 59.7 GB/s of memory bandwidth; it requires less than 7.5 watts of power. As indicated in Table 11, for the proposed method, the recognition of one image required 24 ms in the desktop computer and 397 ms in the Jetson TX2 embedded system. The Jetson TX2 embedded system required longer processing time than the desktop computer because the computing resources were significantly limited; however, it was confirmed that the proposed method could be applied to embedded systems with limited computing resources.

H. ANALYSIS OF FEATURE MAP

In this section, the feature map in DenseNet-161, which uses the images restored by the modified conditional GAN used in this study as input, was analyzed based on the depth of layers; output was produced for each channel of

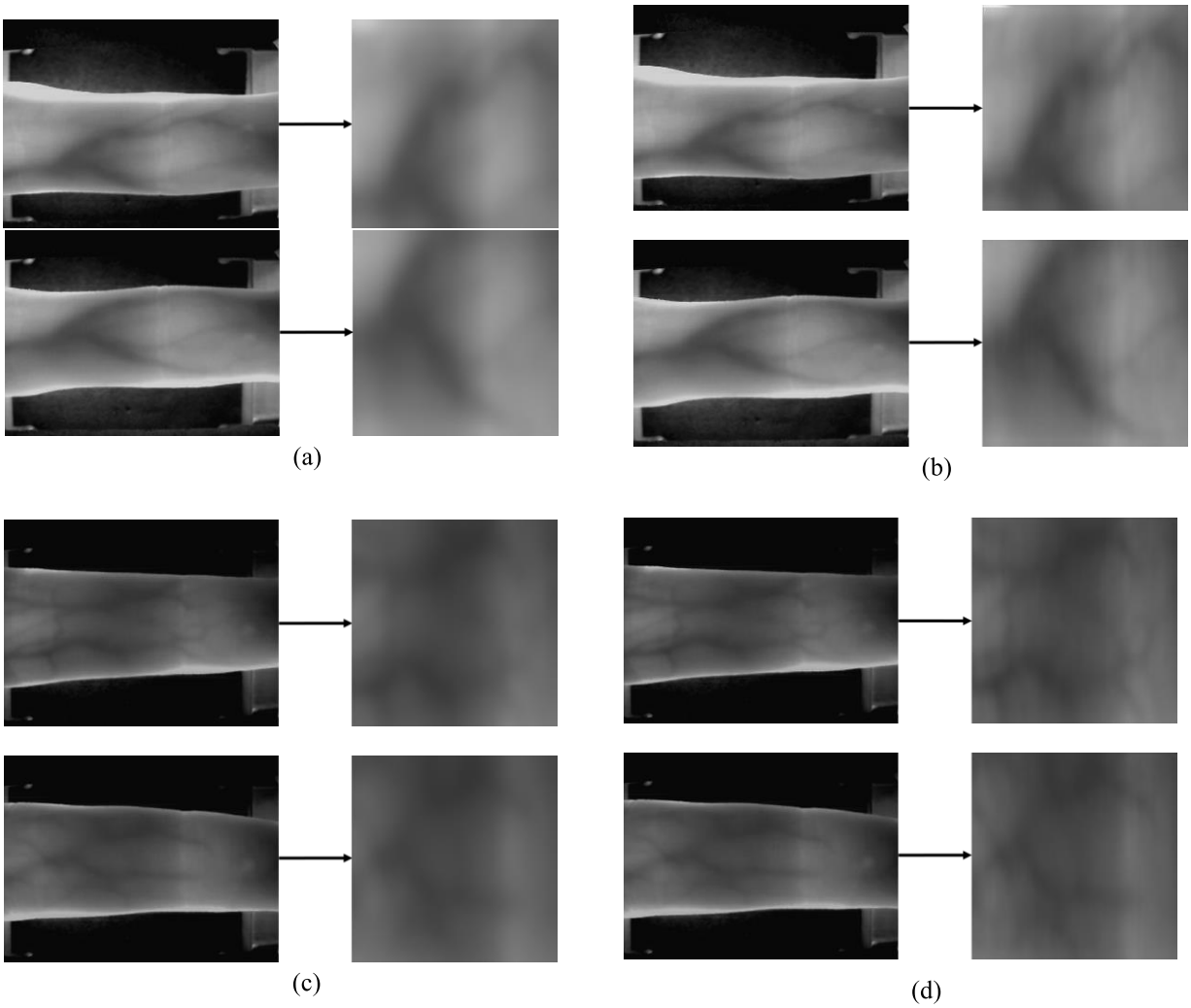

FIGURE 15. Examples of incorrect recognition cases: (a) and (b) show false rejection cases before restoration and after restoration, respectively. (c) and (d) show false acceptance cases before restoration and after restoration, respectively. Upper and lower images in (a), (b), (c) and (d) show the enrolled and input images, respectively. Images are original and ROI images from the left.

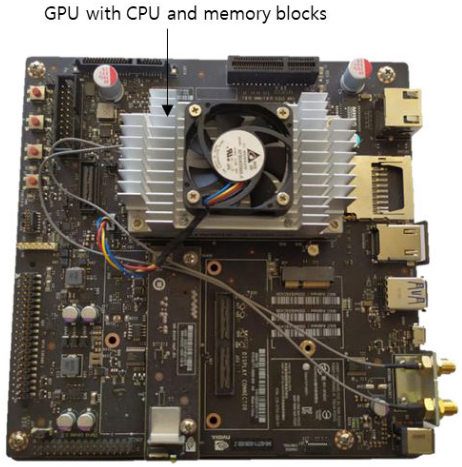

FIGURE 16. Jetson TX2 embedded system.

the feature map because the dimension of the feature map was overly large. Figure 17 shows examples of the feature maps extracted as authentic matching images and imposter matching images from each layer of DenseNet. Figure 17 (a) shows the feature maps extracted from the 1st convolutional layer; Figure 17 (b) shows the output feature maps of the 1st transition layer. Figure 17 (c) shows the output feature maps of the 2nd transition layer. Figure 17 (d) shows the output feature maps of the 3rd transition layer. Figure 17 (e) shows the output feature maps of the last dense block. Moreover, Figure 17 (f) shows the three-dimensional (3D) feature map images drawn by calculating the averages of the feature map values in Figure 17 (e).

TABLE 11. Comparisons of processing speed by proposed method on the desktop computer and embedded system (unit: ms).

	Modified conditional GAN for restoration	DenseNet-161 for finger-vein recognition	Total
Desktop computer	11	12.9	23.9
Jetson TX2	171	226.4	397.4

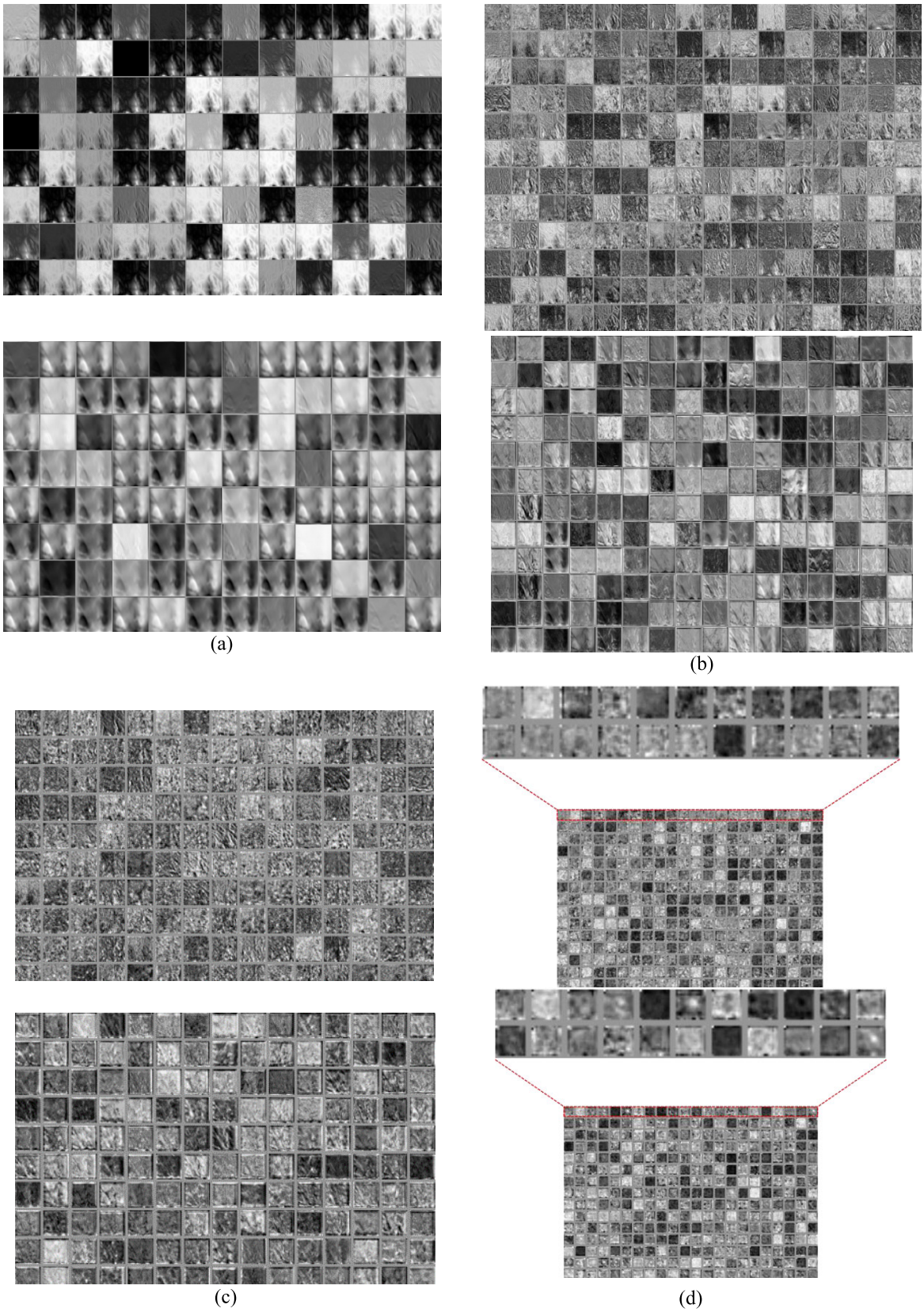

FIGURE 17. Examples of feature maps extracted from an authentic matching image and an imposter matching image from each layer of the DenseNet. Upper and lower images in (a), (b), (c), (d), (e) and (f) show authentic matching and imposter matching, respectively. (a) Feature map extracted from the 1st convolutional layer, (b) output feature map of the 1st transition layer, (c) output feature map of the 2nd transition layer, (d) output feature map of the 3rd transition layer, (e) output feature map of the last dense block, (f) 3D feature map image based on the average feature map values of (e).

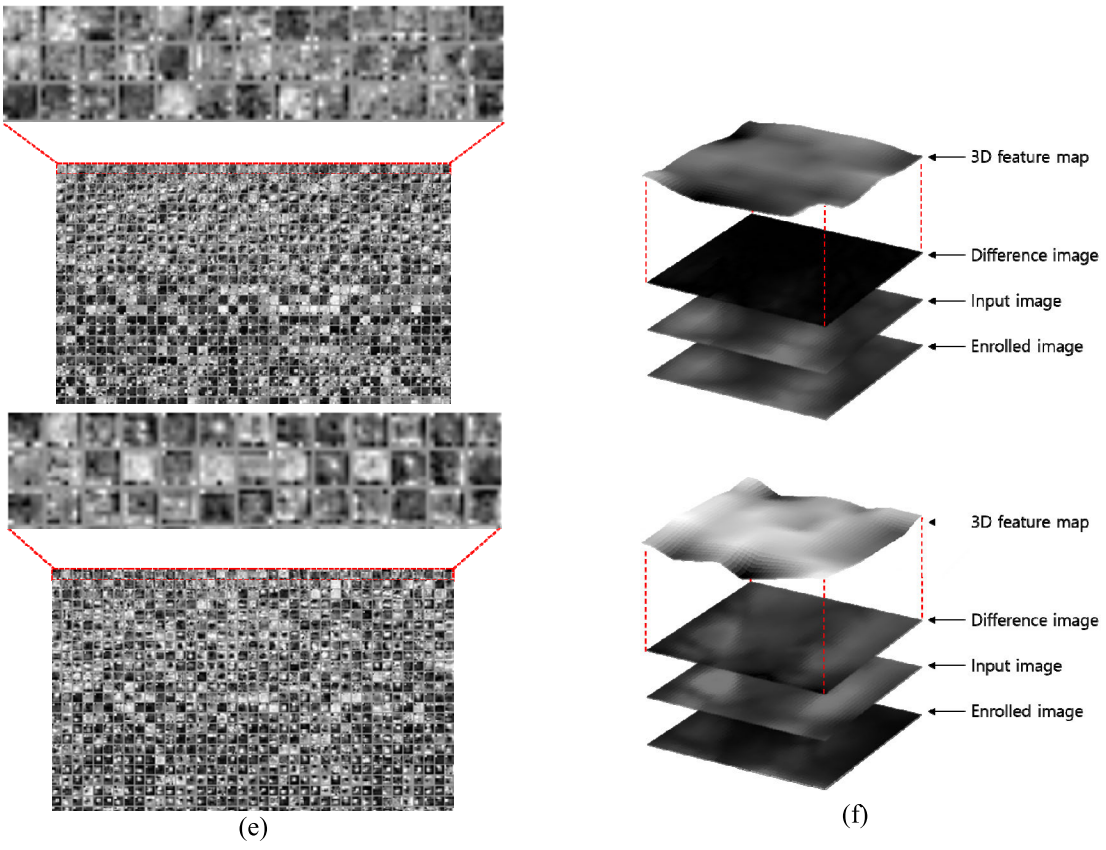

FIGURE 17. *(Continued.)* **Examples of feature maps extracted from an authentic matching image and an imposter matching image from each layer of the DenseNet. Upper and lower images in (a), (b), (c), (d), (e) and (f) show authentic matching and imposter matching, respectively. (a) Feature map extracted from the 1st convolutional layer, (b) output feature map of the 1st transition layer, (c) output feature map of the 2nd transition layer, (d) output feature map of the 3rd transition layer, (e) output feature map of the last dense block, (f) 3D feature map image based on the average feature map values of (e).**

As indicated in Figure 17, abstract features were extracted as the depth of layers increased. For example, the vein-lines and high-frequency edge components of the original images were maintained in Figure 17 (a); however, the shapes of the vein-lines disappeared, and only abstracted low-frequency features were left in Figure 17 (e). Figure 17 (a) through (e), it appears that authentic matching and imposter matching did not have a significant difference in the feature map. Although the change in the 3D feature map values drawn by calculating the average of the feature map values was not significant in upper images of Figure 17 (f), which was the result of authentic matching in the step immediately before the classification layer, the change in the feature map value was found to be relatively larger in lower images of Figure 17 (f), which was the result of imposter matching, compared to authentic matching images of Figure 17 (f). This confirms that the difference in the CNN feature maps between authentic and imposter matching can be represented by the proposed method.

VI. CONCLUSION

To address the problem of finger-vein recognition performance degradation due to optical blur, this study proposed a method for restoring optically blurred finger-vein images using modified conditional GAN and recognizing the restored images using deep CNN. Existing conditional GAN applies random noise to the generator in a dropout form to prevent deterministic output. The generation of different outputs due to the dropout, however, can transform the vein pattern of the finger-vein image to be restored, which leads to the degradation of the recognition performance. Therefore, modified conditional GAN without dropout was proposed in this study because the deterministic output is required, rather than different outputs. When recognition was performed for two open databases using the proposed restoration method, the recognition error rate was found to be less compared to the rate in the case where no restoration was performed. Moreover, when experiments were performed using different state-of-the-art restoration methods and CNN models for finger-vein recognition, it was confirmed that effective optical blur restoration was achieved in all CNN models upon the application of the proposed method. Moreover, it was confirmed that the proposed modified conditional GAN + CNN for the finger-vein recognition method can be applied to embedded systems as well as typical desktop computers.

In future research, the possibility of using the proposed blur restoration method with different biometric modalities

such as face, iris, and palmprint recognition, will be examined. Furthermore, research will be conducted into methods to improve the restoration and recognition performances by training the modified conditional GAN model and CNN model for finger-vein recognition through reinforcement learning.

REFERENCES

[1] Z. Liu, Y. Yin, H. Wang, S. Song, and Q. Li, "Finger vein recognition with manifold learning," *J. Netw. Comput. Appl.*, vol. 33, no. 3, pp. 275–282, May 2010.
[2] E. Lee and K. Park, "Restoration method of skin scattering blurred vein image for finger vein recognition," *Electron. Lett.*, vol. 45, no. 21, pp. 1074–1076, 2009.
[3] J. Yang, B. Zhang, and Y. Shi, "Scattering removal for finger-vein image restoration," *Sensors*, vol. 12, no. 3, pp. 3627–3640, Mar. 2012.
[4] J. Yang and B. Zhang, "Scattering removal for finger-vein image enhancement," in *Proc. Int. Conf. Hand-Based Biometrics*, Hong Kong, Nov. 2011 pp. 17–18.
[5] J. Yang and Y. Shi, "Towards finger-vein image restoration and enhancement for finger-vein recognition," *Inf. Sci.*, vol. 268, pp. 33–52, Jun. 2014.
[6] Y. Shi, J. Yang, and J. Yang, "A new algorithm for finger-vein image enhancement and segmentation," *Inf. Sci Ind. Appl*, vol. 4, no. 22, pp. 139–144, 2012.
[7] J. Yang and G. Bai, "Finger-vein image restoration based on skin optical property," in *Proc. IEEE 11th Int. Conf. Signal Process.*, Beijing, China, Oct. 2012., pp. 749–752.
[8] J. Yang, Y. Shi, and J. Yang, "Finger-vein image restoration based on a biological optical model," in *New Trends Developments Biometrics*, vol. 3. London, U.K.: IntechOpen, 2012, pp. 59–76.
[9] E. C. Lee and K. R. Park, "Image restoration of skin scattering and optical blurring for finger vein recognition," *Opt. Lasers Eng.*, vol. 49, no. 7, pp. 816–828, Jul. 2011.
[10] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 5967–5976.
[11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, Montreal, QC, Canada, Dec. 2014, pp. 2672–2680.
[12] E. C. Lee, H. C. Lee, and K. R. Park, "Finger vein recognition using minutia-based alignment and local binary pattern-based feature extraction," *Int. J. Imaging Syst. Technol.*, vol. 19, no. 3, pp. 179–186, Sep. 2009.
[13] J. Peng, N. Wang, A. A. A. El-Latif, Q. Li, and X. Niu, "Finger-vein verification using Gabor filter and SIFT feature matching," in *Proc. 8th Int. Conf. Intell. Inf. Hiding Multimedia Signal Process.*, Piraeus, Greece, Jul. 2012, pp. 45–48.
[14] J.-D. Wu and C.-T. Liu, "Finger-vein pattern identification using SVM and neural network technique," *Expert Syst. Appl.*, vol. 38, pp. 14284–14289, Oct. 2011.
[15] H. G. Hong, M. B. Lee, and K. R. Park, "Convolutional neural network-based finger-vein recognition using NIR image sensors," *Sensors*, vol. 17, no. 6, p. 1297, Jun. 2017.
[16] W. Kim, J. M. Song, and K. R. Park, "Multimodal biometric recognition based on convolutional neural network by the fusion of finger-vein and finger shape using near-infrared (NIR) camera sensor," *Sensors*, vol. 18, no. 7, p. 2296, Jul. 2018.
[17] H. Qin and M. A. El-Yacoubi, "Deep representation-based feature extraction and recovering for finger-vein verification," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 8, pp. 1816–1829, Aug. 2017.
[18] J. M. Song, W. Kim, and K. R. Park, "Finger-vein recognition based on deep densenet using composite image," *IEEE Access*, vol. 7, pp. 66845–66863, 2019.
[19] *Dongguk GAN and CNN for Recognition of Blurred Finger-vein Image with Blurred Image Database*. Accessed: Aug. 20, 2019. [Online]. Available: http://dm.dgu.edu/link.html
[20] A. Kumar and Y. Zhou, "Human identification using finger images," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2228–2244, Apr. 2012.
[21] A. Kumar and D. Zhang, "Personal recognition using hand shape and texture," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2454–2461, Aug. 2006.
[22] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," Jan. 2017, *arXiv:1701.07875*. [Online]. Available: https://arxiv.org/abs/1701.07875
[23] X. Mao, Q. Li, H. Xie, R. Y. Lau, Z. Wang, and S. P. Smolley, "Least Squares Generative Adversarial Networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Venice, Italy, Oct. 2017, pp. 2813–2821.
[24] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," Nov. 2015, *arXiv:1511.06434*. [Online]. Available: https://arxiv.org/abs/1511.06434
[25] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Munich, Germany, Oct. 2015, pp. 234–241.
[26] G. E. Hinton, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.
[27] K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learn. Represent.*, San Diego, CA, USA, May 2015, pp. 1–14.
[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, pp. 770–778.
[29] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 2261–2269.
[30] *CS231n Convolutional Neural Networks for Visual Recognition*. Accessed: Mar. 7, 2019. [Online]. Available: http://cs231n.github.io/convolutional-networks/#overview
[31] *SDUMLA-HMT Finger Vein Database*. Accessed: May 7, 2019. [Online]. Available: http://mla.sdu.edu.cn/info/1006/1195.htm
[32] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2010.
[33] *NVIDIA GeForce GTX 1070*. Accessed: Apr. 21, 2019. [Online]. Available: https://www.nvidia.com/en-in/geforce/products/10series/geforce-gtx-1070/
[34] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Proc. ACM Int. Conf. Multimedia*, Orlando, FL, USA, Nov. 2014, pp. 675–678.
[35] *Tensorflow*. Accessed: Nov. 2, 2019. [Online]. Available: https://www.tensorflow.org/
[36] *Python*. Accessed: Apr. 21, 2019. [Online]. Available: https://www.python.org/downloads/release/python-368/
[37] *CUDA*. Accessed: Apr. 21, 2019. [Online]. Available: https://en.wikipedia.org/wiki/CUDA
[38] *CUDNN*. Accessed: Apr. 21, 2019. [Online]. Available: https://developer.nvidia.com/cudnn
[39] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, San Diego, CA, USA, May 2015, pp. 1–15.
[40] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proc. 19th Int. Conf. Comput. Statist.*, Paris, France, Aug. 2010, pp. 177–186.
[41] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas, "DeblurGAN: Blind motion deblurring using conditional adversarial networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, Jun. 2018, pp. 8183–8192.
[42] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Venice, Italy, Oct. 2017, pp. 2242–2251.
[43] *Jetson TX2 Module*. Accessed: Sep. 19, 2019. [Online]. Available: https://www.nvidia.com/en-us/autonomous-machines/embedded-systems-dev-kits-modules/

JIHO CHOI received the B.S. degree in business administration from Dongguk University, Seoul, South Korea, in 2016, where he is currently pursuing the combined M.S. and Ph.D. degrees in electronics and electrical engineering. His research interests include biometrics and pattern recognition. He designed the finger-vein restoration system based on modified conditional GAN and finger-vein recognition systems based on CNN, analyzed the result of experiments, and wrote the original paper.

KYOUNG JUN NOH received the B.S. degree in electronics and electrical engineering from Dongguk University, Seoul, South Korea, in 2019, where he is currently pursuing the M.S. degree in electronics and electrical engineering. He helped to construct the finger-vein recognition systems. His research interests include biometrics and deep learning. He helped the experiments and analysis.

SE WOON CHO received the B.S. degree in electronics and electrical engineering from Dongguk University, Seoul, South Korea, in 2017, where he is currently pursuing the combined M.S. and Ph.D. degrees in electronics and electrical engineering. He helped to port the algorithm on embedded systems and perform the experiments. His research interests include biometrics and pattern recognition. He helped the experiments and analysis.

SE HYUN NAM received the B.S. degree in electronics and electrical engineering from Seoul University, Seoul, South Korea, in 2015. He is currently pursuing the combined M.S. and Ph.D. degrees in electronics and electrical engineering with Dongguk University. He helped to construct the finger-vein restoration systems. His research interests include biometrics and pattern recognition. He helped in the implementation of modified conditional GAN.

MUHAMMAD OWAIS received the B.S. and M.S. degrees in computer engineering from the University of Engineering and Technology Taxila, Pakistan, in 2014 and 2016, respectively. He is currently pursuing the Ph.D. degree in electronics and electrical engineering with Dongguk University, Seoul, South Korea. He helped to collect databases. His research interests include medical image analysis and deep learning. He helped in the implementation of modified conditional GAN.

KANG RYOUNG PARK received the B.S. and M.S. degrees in electronic engineering from Yonsei University, Seoul, South Korea, in 1994 and 1996, respectively, and the Ph.D. degree in electrical and computer engineering from Yonsei University, in 2000. He has been a Professor with the Division of Electronics and Electrical Engineering, Dongguk University, since March 2013. His research interests include image processing and biometrics. He supervised this research and helped the revision of the original paper.