Enhanced Iris Recognition Method by Generative Adversarial Network-Based Image Reconstruction

MIN BEOM LEE[ID], **JIN KYU KANG**[ID], **HYO SIK YOON**[ID],
AND KANG RYOUNG PARK[ID], (Member, IEEE)
Division of Electronics and Electrical Engineering, Dongguk University, Seoul 04620, South Korea

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

ABSTRACT Iris recognition is one of the non-contact biometric identification methods that are hygienic and highly accurate. Iris recognition involves using iris images obtained by a near-infrared (NIR) camera or a visible light camera. A clear image of iris can be obtained when an NIR camera is used, but it requires an NIR illuminator in addition to the NIR camera. Iris recognition can be performed with a built-in camera device when a visible light camera is used, which also has the advantage of obtaining a three-channel image containing the color information. Accordingly, studies are being conducted on iris recognition by obtaining iris images from the face images taken by a high-resolution visible light camera in smartphones. However, when iris images have unconstrained conditions or are obtained without the cooperation of the subjects, the quality of iris images are reduced by noises such as optical and motion blur, off-angle view, specular reflection (SR), and other artifacts, thus ultimately deteriorating the recognition performance. Therefore, in this study, a method has been proposed for enhancing the quality of iris images by blurring the iris region and deep-learning-based deblurring. In addition, we propose the method for improving the recognition performance by integrating the recognition score in periocular regions and support vector machine (SVM). The method proposed in this study, which was experimented with noisy iris challenge evaluation-part II training database and MICHE database, exhibited an improved performance compared to the state-of-the-art methods.

INDEX TERMS Biometrics, iris recognition, deep learning, generative adversarial network.

I. INTRODUCTION

Biometric technologies are essential in the fields that require personal identification. Biometrics can be categorized into a contact-based method in which a part of body such as palmprint, finger-vein, or palm-vein comes in contact with a device to obtain the data, and a non-contact-based method in which a body part such as face, iris, or retina does not come in contact with a device. A contact-based method can result in high-quality image data in a stable environment as a body part comes directly in contact with a device. In contrast, a non-contact-based method may result in poor data quality due to the movement of a person. However, a non-contact based method is being widely applied in devices used by many people as practicing personal hygiene has become quite

important during the recent COVID-19 pandemic. An iris recognition method involves using the iris pattern having a high degree of freedom (DOF) between the pupil and the sclera. Therefore, iris recognition is commonly employed in various fields as the iris pattern does not change with aging and is not easily damaged [1]. Iris recognition can be performed with grayscale images obtained using a near-infrared (NIR) camera and an NIR illuminator and with color images obtained using a visible light camera. Images obtained with an NIR camera have clear iris patterns, which result in a high recognition accuracy. However, the quality of iris image is reduced and intense lighting is required when the distance between the iris and device is increased. Additional lighting is not required when using a visible light camera; the obtained images contain the color information. Based on these characteristics, the device can be miniaturized or the iris recognition feature can be added to an existing device.

According to the visible light based iris recognition characteristics, studies are being conducted using a database consisting of eye images taken with a high-resolution visible camera at 4–8 m to include poor focus, off-angle view, motion blur, rotation, and low illumination [2] and a database consisting of face images taken with a visible light camera embedded in smartphones [3]. However, when iris images have unconstrained conditions or are obtained without the cooperation of the subjects, the quality of iris images is reduced by noises such as optical and motion blur, off-angle view, SR, and other artifacts, which ultimately deteriorates the recognition performance. Therefore, this study has been proposed for enhancing the quality of iris images by blurring the iris region and deep-learning-based deblurring. In addition, we propose the method for improving the recognition performance by integrating the recognition score in iris and periocular regions using an SVM. The method proposed in this study, which was experimented with two open databases, exhibited an improved accuracies of iris recognition compared to the state-of-the-art methods.

Our research is novel in the following three ways compared to previous works.

- Iris recognition performance is improved through proposed method of iris pattern blurring and GAN-based deblurring.
- We find that using the images generated by inputting a periocular region in the generator of GAN is not advantageous for performance improvement due to the limitations of GAN. Accordingly, only the iris region is used for blurring and as an input to GAN for deblurring.
- In order to make it possible to have the fair comparisons for other researchers, we make our trained CNN models public as shown in [48].

This paper is organized as follows. Section II presents previous studies in iris recognition, and Section III provides detailed explanations on the deep-learning-based deblurring and iris recognition methods proposed in this study. Section IV present the experimental results with analysis, and lastly Section V concludes this paper.

II. RELATED WORKS

Studies on iris recognition mostly cover a segmentation process for detecting iris and removing noises, feature extraction process for extracting information from iris images, and matching process for calculating the distance between the enrolled features and the features being recognized. In this section, previous studies on iris recognition are summarized. The methods for extracting features from iris images in previous studies on iris recognition can be roughly divided into handcrafted feature-based methods and deep-feature-based methods.

A. HANDCRAFTED FEATURE-BASED METHODS

As a handcrafted feature-based method, Kaur *et al.* proposed a method for extracting features in NIR or visible light environment using Zernike moments, polar complex exponential transform, polar cosine transform, and polar sine transform, grouped under polar harmonic transforms, and then classified them using the K-nearest neighbor (K-NN) classifier [4]. Hamd *et al.* proposed a method for extracting iris features in frequency domain region using Fourier descriptors consisting of low frequency descriptor and high frequency descriptor and classified them using a neural network [5]. Ahmed *et al.* suggested a score fusion method for the Chi-Square distance value calculated using the features obtained using multi-block transitional local binary patterns (MB-TLBP) improved with local binary patterns (LBP) and the hamming distance value between iris codes obtained using 1-D Log-Gabor filter [6]. Galdi *et al.* suggested a score fusion method for the three distance values calculated using the features obtained with color descriptor, texture descriptor, and cluster descriptor for enrolled and recognized images [7]. In the study conducted by Sajjad *et al.*, an iris recognition method was proposed using contrast-limited adaptive histogram equalization (CLAHE) for images in order to reinforce against low contrast and low illumination of the noisy iris challenge evaluation-part II (NICE.II) training database [8].

B. DEEP FEATURE-BASED METHODS

With a recent advancement in deep learning technology, supervised learning-based CNN is being used for extracting deep features in iris recognition. Lee at al. proposed a matching method by extracting features using a non-square convolution filter for normalized iris images for which segmentation has not been performed [9]. Zhao *et al.* proposed a model comprising two fully convolutional network (FCN) based sub-networks (FeatNet and MaskNet) for extracting small and local pattern in iris images, and an extended triplet loss (ETL) function that shortens the distance between iris images of the same person and increases the distance between iris images of different individuals for training this model [10]. Gangwar *et al.* proposed DeepIrisNet-A consisting of convolutional layer, batch normalization layer, and fully connected layer and DeepIrisNet-B having an inception architecture for extracting features for iris recognition, and a method for matching by measuring the Euclidean distance between 4096-dimensional feature vectors extracted from the network [11]. Minaee *et al.* suggested a recognition method in which features are extracted by inputting eye images in the visual geometry group (VGG)-Net 16 model and the extracted features are classified with SVM [12]. Nguyen *et al.* proposed a method for extracting features for recognition using five state-of-the-art and off-the-shelf CNNs (AlexNet, VGG, Google Inception, ResNet, and DenseNet) and classifying them using SVM [13]. Proença *et al.* distinguished between genuine and imposter by obtaining the central coordinates and size of iris using single shot multibox detector (SSD) and then inputting normalized iris images for which segmentation has not been performed into a VGG-19 based CNN model [14]. Al-Waisy *et al.* suggested a method for which both the right and left irises of an individual are used.
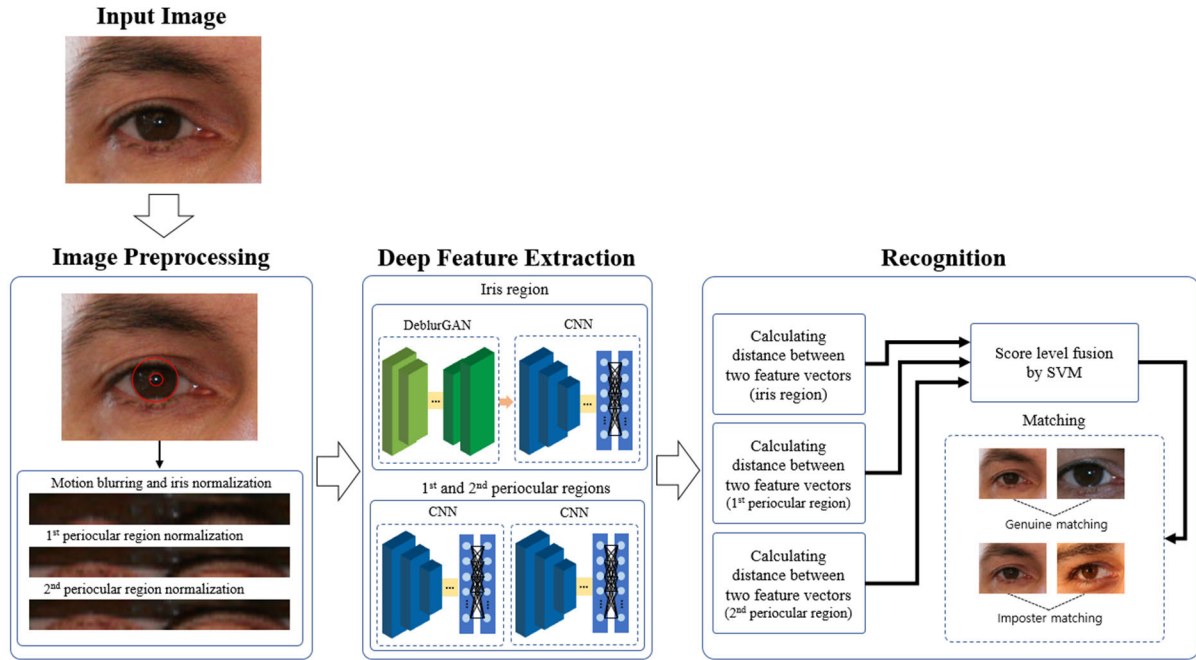
FIGURE 1. Overview of the proposed method.

This method performs recognition through the ranking-level fusion of the values resulted from inputting each image into a CNN that classifies it into N classes [15]. Wang *et al.* proposed DRFNet consisting of residual blocks including three scales of dilated convolutional kernels and extended triplet loss for training this model [16]. Alaslani *et al.* suggested a method for classifying features vectors extracted from AlexNet using SVM [17]. In [53], Kashihara proposed the method of super-resolution by super-resolution generative adversarial network (SRGAN) and iris recognition by deep convolutional neural network (DCNN). However, the SRGAN and DCNN could not enhance the accuracies with the restored image of low resolution compared to those with original image of high resolution. However, our proposed method of motion blurring and DeblurGAN-based deblurring shows the higher accuracies of iris recognition than those with original images as shown in Tables 7 and 11. In [54], Minaee *et al.* proposed Iris-GAN to generate realistic iris images, but they did not measure the recognition accuracies.

C. DEBLURRING OF IRIS IMAGES

There have been various researches on deblurring for iris images. In [56], authors proposed iris deblurring algorithm for iris capture based on the domain knowledge inherent in iris images and iris capture settings to improve the performance. Alaoui *et al.* proposed the method of motion deblurring of iris image to achieve a quality edge preserving image restoration using Total Variation (TV)-L1 regularization technique [57]. In [58], authors proposed iris image deblurring method to enhance the quality of blurred iris images based on refinement of point spread function. Liu *et al.* proposed deblurring method to automatically enhance the quality of

both defocused and motion blurred iris images [59]. In [60], authors proposed the method of deblurring of noisy images in iris recognition.

However, no previous study on iris recognition has investigated the removal of various noises in iris patterns through blurring and deblurring for improving the recognition performance.

Hence, we propose a method for enhancing the quality of iris images by blurring the iris region and GAN-based deblurring. In addition, we propose the method for improving the recognition performance by integrating the recognition score in iris and periocular regions using an SVM. Table 1 presents the comparison of strengths and weaknesses between the existing methods and the proposed method.

III. PROPOSED METHOD
A. OVERVIEW OF PROPOSED METHOD
Figure 1 shows the overall flow of the algorithm proposed in this study. Iris and pupil regions are extracted from the input image, and then normalization is performed for the iris region by applying a blur kernel. Subsequently, two periocular regions are defined based on the iris region, and then normalization is performed again. The iris region contains the information on iris patterns, while the periocular regions contain the information on the shape of eyelid and eyebrow in addition to skin color. Feature vectors are extracted from the three regions using a deep learning model. Then, the distance between these feature vectors and enrolled feature vectors is calculated. Score fusion is performed for three distance values based on SVM to create one score. Using the last score value, it is determined whether the matched value is genuine matching or imposter matching.

TABLE 1. Comparisons of existing and proposed method. (d', EER, and GAR mean d-prime value, equal error rate, and genuine acceptance rate (100 - false rejection rate (FRR)), respectively. The detailed descriptions are provided in Sections III.E and IV.E.)
(A: noisy iris challenge evaluation-part II (NICE.II) training database / B: Institute of automation of Chinese academy of sciences (CAISA)-Iris-Ver.1.0 /
C: Mobile iris challenge evaluation (MICHE) II competition database / D: CASIA-Iris-Distance / E: Notre Dame (ND)-Iris-0405 / F: CASIA-Iris-Thousand /
G: CASIA-Iris-V4-Lamp / H: CASIA-Iris-Interval / I: MICHE database / J: Indian Institute of Technology (IIT) Delhi database /
K: Subset of University of Beira iris (UBIRIS).v2 database / L: ICE 2005).

Category	Recognition method	Accuracy	Pros	Cons
Handcrafted feature-based method	Classifying with the K-NN classifier [4]	d' of 2.82, EER of 10% (K)	Less time for developing algorithm	Lower recognition performance than deep feature-based methods
	Extracting features using Fourier descriptors [5]	EER of 0.17% (B)		
	Using 1-D Log-Gabor filter and MB-TLBP [6]	EER of 1.22% (C)		
	Score fusion of the distance found with color, texture, and cluster descriptors [7]	EER of 29% (C)		
	CLAHE-based image enhancement [8]	EER of 18.82% (A)		
Deep-feature-based method	Three CNNs-based method [9]	d' of 2.62, EER of 10.36% (A) d' of 1.87–2.26, EER of 16.25–17.9% (I)	Higher recognition performance than handcrafted feature-based methods	Does not consider removing noise by motion blur and improving iris recognition accuracy accordingly
	Small and local pattern extraction using FCN-based model, and ETL [10]	EER of 3.85% (D)		
	DeepIrisNet [11]	EER of 2.19% (E)		
	Classifying features extracted with VGG-Net 16 using SVM [12]	GAR of 99.4% (J)		
	Five state-of-the-art and off-the-shelf CNNs [13]	GAR of 98.5% (F)		
	SSD detector-based segmentation less method [14]	d' of 6.61, EER of 0.006% (G) d' of 6.61, EER of 0.03% (F)		
	DRFNet [16]	EER of 1.33% (E), EER of 7.05% (D)		
	Classifying feature vectors extracted from AlexNet using SVM [17]	GAR of 89% (H) GAR of 98% (F) GAR of 98.3 (B)		
	Three CNNs and SVM-based score fusion [18]	d' of 3.31, EER of 8.58% (A) d' of 1.87–2.09, EER of 15.49–17.39% (I) d' of 3.29, EER of 2.96% (D)		
Deblurring only	Handcrafted feature-based deblurring of iris images [56-60]	d' of 5.02, EER of 0.63% (L) [59]	Fast processing time	- Performance is affected by database - Do not attempt the performance enhancement by deblurring after motion blurring
Deblurring after motion blurring	Iris recognition method using images restored with GAN (**Proposed method**)	d' of 3.4, EER of 7.49% (A) d' of 1.82–2.22, EER of 14.18–17.02% (I)	Motion blur noise removal and recognition accuracy improvement	Requires longer time to train algorithm

B. IMAGE PREPROCESSING

1) DETECTION IRIS AND PERIOCULAR REGION

Previous image preprocessing for iris recognition consisted of iris detection, segmentation, and normalization processes [1]. The iris is detected in the iris detection step, and then noise factors such as reflected light, eyebrow, and eyelid are removed in the segmentation step. Lastly, Cartesian coordinates are converted to polar coordinates and iris size is normalized in the normalization step.

The images captured with a visible light camera in this study have a significantly poorer image quality than regular iris images captured with a NIR camera and illuminator, thus causing difficulty in accurately performing segmentation of the iris region. Moreover, the proposed method involves training with periocular regions in addition to the iris region. Therefore, only detection and normalization are performed in this study and segmentation is excluded. Eyebrow, eyelid, and skin color are included for training the feature extraction

algorithm since segmentation was not performed. The iris region was extracted from the input image using the following two circular edge detector (CED) [9], [19].

$$\arg\max_{(x_0,y_0),r} \left[\frac{\partial}{\partial r} \left(\int_{-\frac{\pi}{4}}^{\frac{\pi}{6}} \frac{I(x,y)}{5\pi r/12} ds + \int_{\frac{5\pi}{6}}^{\frac{5\pi}{4}} \frac{I(x,y)}{5\pi r/12} ds \right) \right] \quad (1)$$

where r is the radius of iris region. The coordinates (x_0, y_0) shows the center position of the iris region. We do not perform the operations in the range $0–2\pi$ for the two integro-differential operations of Equation (1) because the regions of the other ranges can be occluded by eyelids, which can reduce the detection accuracy of iris outer boundary.

In addition to the iris region, the periocular regions were defined by increasing the central location of the iris radius (IRrad) obtained using Equation (1) by the size of $w_1 \times IRrad$ and $w_2 \times IRrad$ (w_1 and w_2 are 1.43 and 1.53, respectively) as shown in Figure 2(b) and (c), respectively [9].

2) NORMALIZATION OF IRIS AND PERIOCULAR REGIONS

In general, each individual has a different sized iris. Furthermore, even the iris size of the same individual varies because of expansion and contraction of the pupil according to the change in the brightness of lighting; the iris size also varies depending on the distance between the camera and the eyes. To solve the problem of performance deterioration due to these factors, size normalization shown in Figure 3 is performed for the iris and periocular regions defined in Figure 2 [9]. In this method, the iris images of polar coordinates are respectively divided into 8 tracks and 256 sectors as shown in Figures 3 (b). In each track, the pixel values are averaged in the vertical (ρ axis) direction by a one-dimensional (1-D) Gaussian kernel. Consequently, the normalized iris and periocular images of 256×8 pixels are generated.

FIGURE 2. Example of detected iris region and periocular region definition. (a) Iris region. (b) Periocular region based on $w_1 \times IRrad$. (c) Periocular region based on $w_2 \times IRrad$.
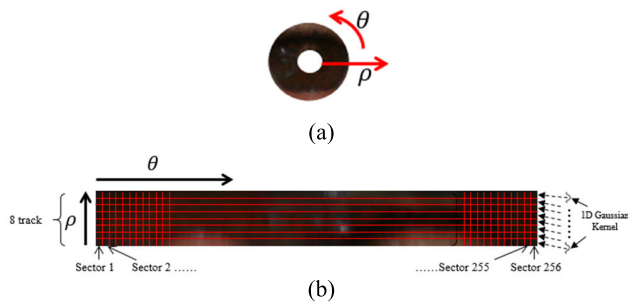

FIGURE 3. Example of the normalized iris image. (a) Iris region in cartesian coordinates. (b) Normalized iris image of (a) in polar coordinates.

C. DEEP LEARNING MODEL FOR DEBLURRING

Extracting excellent features in biometrics indicates that invariance is maintained in the data of the same classes, while the data of different classes have highly distinctive characteristics. Iris images must be sharp and the pattern should be clear for extracting these features from the iris images. In this study, deblurring is performed for iris images by combining the generator of DeblurGAN and CNN, and features are extracted from CNN.

A conditional generator adversarial network (cGAN) has an improved architecture that helps in controlling the output of the generator by adding more information [21] to the GAN architecture proposed by Goodfellow *et al.* [20]. Isola *et al.* proposed a cGAN based pix2pix model, which has been successfully applied in various image transform fields [22]. DeblurGAN [23] used in this study was designed as a cGAN using the Wasserstein gradient penalty(WGAN-GP) [24]. Training GAN models needs the

process of finding a Nash equilibrium of a noncooperative two-player game [25]. In some cases, the gradient descent can successfully perform this, but in other cases, it cannot. Therefore, no good equilibrium-finding algorithm has been reported yet. Therefore WGAN was employed as it uses an alternative objective function based on the Wasserstein distance instead of the traditional Jensen–Shannon distance, which can help increase the training stability [26]. Gulrajani *et al.* [24] proposed WGAN-GP robust to the choice of generator architecture. Based on these, DeblurGAN adopts WGAN-GP as a critical function, which enables DeblurGAN to use a lightweight CNN architecture as a generator.

Figure 4 shows the training architecture of DeblurGAN. Blur image is input to the generator, and the generator generates a restored image. The generator is trained with adversarial loss and content loss, which is represented using Equation (2).

$$\mathcal{L} = \mathcal{L}_{GAN} + \lambda \cdot \mathcal{L}_X \tag{2}$$

where $\mathcal{L}$ is the total loss, which is the sum of adversarial loss $\mathcal{L}_{GAN}$ and content loss $\mathcal{L}_X$. λ was set to 100 in all experiments. Adversarial loss $\mathcal{L}_{GAN}$ is calculated using Equation (3).

$$\mathcal{L}_{GAN} = \sum_{n=1}^{N} -D_{\theta_D}(G_{\theta_G}(I^B)) \tag{3}$$

where D_{θ_D} and G_{θ_G} are the discriminator and generator, respectively; θ_D and θ_G are the trainable parameters of the discriminator and generator, respectively; and I^B is the blur image of Figure 4. Content loss was the perceptual loss [27], which is defined as:

$$\mathcal{L}_X = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}\left(I^S\right)_{x,y} - \phi_{i,j}(G_{\theta_G}(I^B))_{x,y})^2 \tag{4}$$

$\phi_{i,j}$ is the feature map obtained from the ith convolution layer of the VGG19 network pretrained with ImageNet database [28]. $W_{i,j}$ and $H_{i,j}$ are the width and height of the feature maps. I^S is the sharp image (Sharp image of Figure 4).
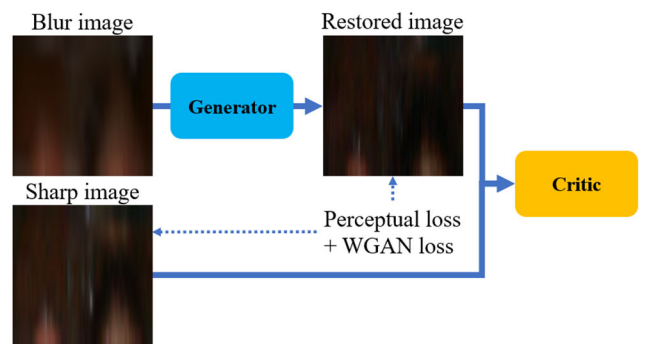

FIGURE 4. DeblurGAN training architecture.

Figure 5 and Table 2 show the generator architecture of DeblurGAN, while Table 3 lists different parameters of discriminator architecture. The generator consists of three convolution blocks, nine residual blocks (ResBlocks [29]
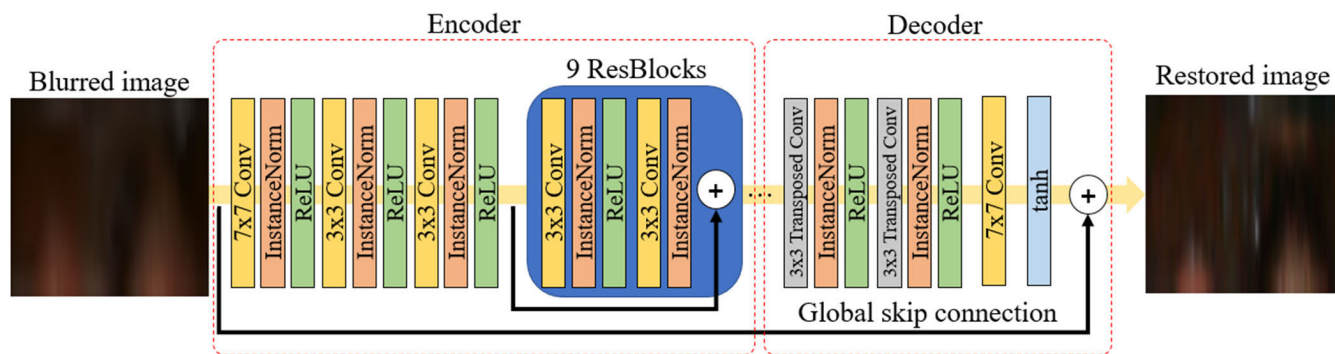
FIGURE 5. DeblurGAN generator architecture (Conv: Convolution layer and ReLU: Rectified linear unit).

TABLE 2. DeblurGAN generator architecture.

	Layer type	Number of filters	Size of feature map (height × width × channel)	Filter size (height × width)	Number of strides (height × width)	Number of padding (height × width)
Encoder	Image input layer		256 × 256 × 3			
	1st convolution layer	64	256 × 256 × 64	7 × 7	1 × 1	3 × 3
	2nd convolution layer	128	128 × 128 × 128	3 × 3	2 × 2	1 × 1
	3rd convolution layer	256	64 × 64 × 256	3 × 3	2 × 2	1 × 1
	ResBlock $\begin{bmatrix} 3 \times 3 \ conv \\ 3 \times 3 \ conv \end{bmatrix} \times 9$	256	64 × 64 × 256	3 × 3	1 × 1	1 × 1
Decoder	1st transposed layer	128	128 × 128 × 128	3 × 3	2 × 2	
	2nd transposed layer	64	256 × 256 × 64	3 × 3	2 × 2	
	4th convolution layer	3	256 × 256 × 3	7 × 7	1 × 1	3 × 3
	Global skip connection (input + 4th convolution layer)		256 × 256 × 3			

TABLE 3. DeblurGAN discriminator architecture.

Layer type	Number of filters	Size of feature map (height×width× channel)	Filter size (height× width)	Number of stride (height×width)	Number of padding (height× width)
Image input layer		256 × 256 × 3			
1st convolution layer	64	128 × 128 × 64	5 × 5	2 × 2	2 × 2
2nd convolution layer	128	64 × 64 × 128	5 × 5	2 × 2	2 × 2
3rd convolution layer	256	32 × 32 × 256	5 × 5	2 × 2	2 × 2
4th convolution layer	512	16 × 16 × 512	5 × 5	2 × 2	2 × 2
5th convolution layer	512	16 × 16 × 512	5 × 5	1 × 1	2 × 2
6th convolution layer	512	16 × 16 × 512	5 × 5	1 × 1	2 × 2

of Figure 5), two transposed convolution blocks, and global skip connection. Each ResBlock consists of convolution layer, instance normalization layer [30], and rectified linear unit (ReLU) [31] activation. The generator receives a 256 × 256 × 3 (height × width × channel) size image as an input, and calculates the feature map in the 1st–3rd convolution layer using a 7 × 7 and 3 × 3 sized filter. Without using a separate pooling layer, a stride is calculated to be 2 × 2 in the 2nd and 3rd convolution layers, and thus, the feature map is designed to be reduced to 64 × 64 × 256.

After the feature map is calculated in nine ResBlocks, the size of the feature map becomes 256 × 256 × 64 through the 1st and 2nd transposed convolution layer in which the width and height increase while the number of channels decreases. Finally, a restored image of 256 × 256 × 3 is generated through the 4th convolution layer and global skip connection.

D. FEATURE EXTRACTION AND MATCHING

For recognizing iris in the normalized iris and periocular images in this study, three CNN models having the same

TABLE 4. CNN model architecture for extracting iris features.

Layer type	Number of filters	Size of feature map (height × width × channel)	Filter size (height× width)	Number of strides (height × width)	Number of padding (height × width)
1st convolution layer	64	8 × 244 × 64	1 × 13	1 × 1	0 × 0
2nd convolution layer	64	8 × 232 × 64	1 × 13	1 × 1	0 × 0
Max pooling layer	1	8 × 116 × 64	1 × 2	1 × 2	0 × 0
3rd convolution layer	128	8 × 104 × 128	1 × 13	1 × 1	0 × 0
4th convolution layer	128	8 × 92 × 128	1 × 13	1 × 1	0 × 0
Max pooling layer	1	8 × 46 × 128	1 × 2	1 × 2	0 × 0
5th convolution layer	256	8 × 36 × 256	1 × 11	1 × 1	0 × 0
6th convolution layer	256	8 × 26 × 256	1 × 11	1 × 1	0 × 0
Max pooling layer	1	8 × 13 × 256	1 × 2	1 × 2	0 × 0
7th convolution layer	512	6 × 11 × 512	3 × 3	1 × 1	0 × 0
8th convolution layer	512	4 × 9 × 512	3 × 3	1 × 1	0 × 0
Max pooling layer	1	4 × 5 × 512	1 × 2	1 × 2	0 × 1
1st fully connected layer		4096			
2nd fully connected layer		4096			
3rd fully connected layer		#classes			
Softmax layer		#classes			
Classification layer (output layer)		#classes			

architecture were used as a feature extractor for extracting the features from each type of images (CNN parts of "Deep Feature Extraction" in Figure 1). This model uses a non-square filter, which is appropriate for the iris image environment, and each CNN model consists of eight convolution layers and three fully connected layers [9]. Table 4 summarizes the detailed architecture of the feature extractor.

As shown in Figures 1 and 5, the normalized iris image becomes a restored image through the DeblurGAN generator, which then becomes the input of the CNN model. Subsequently, two normalized periocular images are input in each CNN model. A total of 4096-dimensional features are extracted from the 1st fully connected layer of the three CNNs. This process is performed for the enrolled images and the recognized images; a total of three pairs of 4096-dimensional features are extracted. We do not use a triplet loss based approach with positive and negative anchor comparisons in our research, but adopts a conventional cross-entropy loss [52] with positive anchor. In detail, with training data of classes 1 ∼ N, our CNN model for recognition having N output nodes is trained to produce the output of classes 1 ∼ N. Then, in order to use our trained model for untrained testing data of classes N + 1 ∼ N + M, we extract the mentioned 4096-feature vector with each input image, and measure the distance between the two vectors from enrolled and input images for the final decision of positive and negative matching [9], [18].

Then, three distance values are obtained by calculating the Euclidean distance between the three pairs of 4096-dimensional features. Score level fusion is performed in this study using SVM [32] for the three calculated distance values.

This method sets the distance (or margin) between the classes to be maximum to configure an optimized hyper-plane. Low dimension space data are transformed to higher dimension space data in order to enable the hyper-plane to more easily classify using a kernel function in the case of a complicated problem such as non-linear environment, rather than a simple classification problem. SVM forms a hyper-plane by selecting several support vectors, as shown in Equation (5). x_i and y_i are the selected support vectors and their corresponding labels (−1 or 1). a_i and b are the parameters of the SVM model, and $K(\cdot)$ is the kernel function. In our experiments with the training data, we experimentally selected radial basis function of Equation (6) as optimal one.

$$f(\mathrm{x}) = sgn(\sum_{i=1}^{k} a_i y_i K(x, x_i) + b) \qquad (5)$$

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2} \quad (\gamma > 0) \qquad (6)$$

If the score (distance) output through SVM is greater than the threshold, it is determined as imposter matching; if the score is smaller than the threshold, it is determined as authentic (genuine) matching. Here, authentic matching refers to when the input image has the same class as the enrolled image, while imposter matching refers to when the input image has a different class from the enrolled image. The two types of errors which occur here, i.e., false acceptance rate (FAR) (the error of accepting imposter matching as authentic one) and false rejection rate (FRR) (the error of rejecting authentic matching as imposter one), typically have a tradeoff relationship. The error rate at the threshold point at which FAR and FRR become identical is called the equal error rate (EER).

IV. EXPERIMENTAL RESULTS

A. EXPERIMENTAL DATABASES AND ENVIRONMENTS

sectionTwo types of open databases were used in the experiment to evaluate the proposed method. The first database, NICE.II training database, was partially selected from UBIRIS.v2 database in order to evaluate the iris recognition performance in an unconstrained environment under a visible light with extensive noise. This database was used to evaluate the performance in the NICE.II contest [2]. NICE.II training database contains a total of 1,000 images consisting of 171 classes. The image resolution is 400 × 300 pixels, and iris images were captured using a high-resolution visible light camera which uses visible lighting of persons who were walking 4–8 m away from the camera [33]. The images in this database include various factors which lead to deterioration in recognition performance such as in-plane rotation, low-illumination, blurring, and off-angle view as shown in Figure 6(a).

The second database MICHE [3] was built to study iris recognition for smartphones. The images in each database were captured with a front-facing or rear-facing camera of a device in indoor and outdoor of visible light environment. As shown in Figure 6(b), they include the factors causing performance deterioration such as optical and motion blur, off-angle view, SR, and shadow.

For training the three CNN models for extracting features, the normalized images were augmented by normalizing the central coordinates of iris [9]. Moreover, for the CNN model that extract features from the iris region, the number of images was increased by generating iris images with a cGAN-based model [18]. The number of images in each database and the number of images after augmentation are shown in Table 5. For a fair comparison with the previous methods, the augmented data were used only during the training, and the original data that have not been augmented were used during the testing.

TABLE 5. Number of images of each database.

Database		Sub-dataset	Number of classes	Number of images	Training dataset
NICE.II training		A	86	492	39,852
		B	85	508	41,148
MICHE	Galaxy S4	A	35	330	26,730
		B	35	342	27,702
	iPhone5	A	35	312	25,272
		B	35	298	24,138

The proposed method was verified using a two-fold cross validation method. As shown in Table 5, this method involves forming sub-datasets group A and group B by dividing the number of database classes in half and then proceeding with training and testing of those sub-datasets. Specifically, when one sub-dataset is trained, the other sub-dataset is tested (1st fold cross validation). Subsequently, these two sub-datasets are interchanged for training and testing (2nd fold cross validation) to measure the average accuracy.

(a)

(b)

FIGURE 6. Example images from the databases used in the experiments. (a) NICE.II training database. (b) MICHE database.

All experiments, including both training and testing, and were performed on a desktop computer with an Intel® Core™ i7-7700 CPU @ 3.6 GHz (4 cores) with 32 GB of main memory, and NVIDIA GeForce GTX 1080 (2560 compute unified device architecture (CUDA) cores) [34] with graphics memory of 8 GB (NVIDIA, Santa Clara, CA, USA). The DeblurGAN used in the experiment was implemented with TensorFlow framework (version 1.14.0) [35]. The three CNN models used for extracting features were implemented with Caffe framework (version 1) [36]. The versions of the installed framework and library include Python 3.6. NVIDIA CUDA® toolkit [49] and NVIDIA CUDA® deep neural network library (CUDNN) [50] versions are 10.0 and 7.6, respectively.

B. NOISE REDUCTION BY MOTION BLURRING

As shown in "Image Preprocessing" in Figure 1, motion blurring is performed to remove the noise in the iris region.

FIGURE 7. Example of a motion blurred image.

To this end, random trajectories generation method proposed by Boracchi *et al.* [37] was used to generate a motion blurring kernel to create a motion blurred image as shown in Figure 7. Therefore, original images of NICE.II and MICHE databases were used as sharp image data (ground truth data) for the training of our model for deblurring.

As shown in Figure 1, motion blurring and image restoration by DeblurGAN were performed only for the iris region because it contains detailed iris texture in addition to other noises including SR and external light reflection (ghost), while the periocular regions have a relatively less amount of noises. The features that can be obtained from the periocular region such as the shape of eyelid or wrinkles can be damaged when blurring is performed followed by restoration using DeblurGAN or the skin color may change due to the change in lighting, which hinders predicting the original color. The inter-class distance may be reduced between the restored images due to these problems.

C. TRAINING OF DEBLURGAN AND CNN MODELS

In this study, one DeblurGAN and three CNN models were trained. For DeblurGAN, dropout regularization with the probability of 0.5 was used in the 1st convolution layer of each ResBlock to prevent overfitting [38]. For optimization, we followed the approach of [26] and performed 5 gradient descent steps on discriminator and one step on generator using adaptive moment estimation (Adam) [39] as a solver. The initial learning rate was set to 0.0001 for both generator and critic. After the first 150 epochs, we linearly decayed the rate to 0 over the next 150 epochs. The generator and discriminator were trained with a batch size of 1. The discriminator of DeblurGAN was trained so that the loss converges to 0 as shown in Figure 8. The training results of the generator were verified with a peak signal-to-noise ratio (PSNR). PSNR is used to evaluate based on the mean square error (MSE) between the ground truth image ($I^S(x, y)$) and deblurred image ($I^R(x, y)$) as shown in Equations (7) and (8) [40].

$$MSE = \frac{\left(\sqrt{\sum_{y=1}^{N} \sum_{x=1}^{M} (I^S(x, y) - I^R(x, y))^2}\right)^2}{MN} \quad (7)$$

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE}\right) \quad (8)$$

In the above equations, M and N represent the image width and height, respectively. Figure 9 shows the PSNR between the image generated by the generator through the training of DeblurGAN and the ground truth image. The image quality
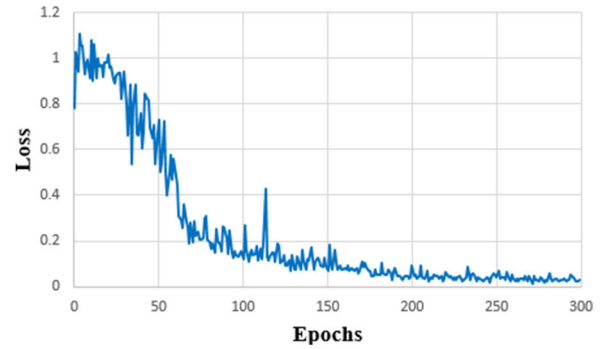

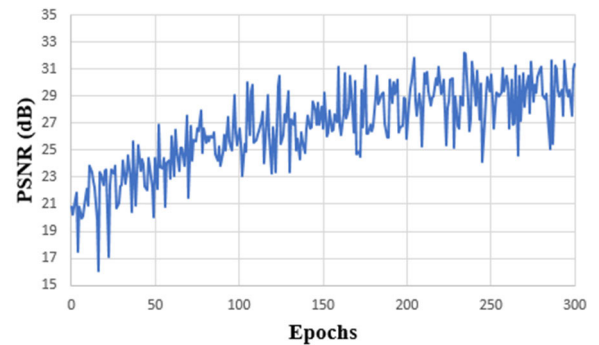**FIGURE 8.** Discriminator loss according to training epochs.

FIGURE 9. PSNR according to training epochs of the image generated with the generator.

was improved after training the DeblurGAN for 300 epochs, and the generator was also sufficiently trained.

The training of the three CNN models was performed as follows. We used the Adam optimizer as a solver. The initial parameters for this optimizer were learning rate of 0.001, momentum of 0.9, momentum2 of 0.999, epsilon of *1e-08*. The convolution filter was initialized by using a method suggested by He *et al.* [41] and the biases were initialized to 0. A batch size of 128 was used and learning was conducted in 50 epochs. Figure 10 shows the training loss and accuracy graphs when the A sub-dataset of NICE.II training database was trained with the three CNN models. The training results showed that all training losses converge to 0 and the training accuracy also converges to nearly 100%, thus proving that the three CNN models used in this study have been sufficiently trained.

D. TESTING OF PROPOSED METHOD
1) COMPARATIVE EVALUATION OF GENERATED IMAGES
Figure 11 shows examples of original image, motion blurred image, and motion deblurred image by DeblurGAN generator for the normalized iris region. More specifically, the blurred and normalized iris image in Figure 11(b) were used as an input for the trained DeblurGAN to generate the final image of the generator in Figure 11(c). As shown in Figure 11(c), the noise of motion blurring is removed by DeblurGAN, and then the image which resembles the original image (Figure 11(a)) is generated.
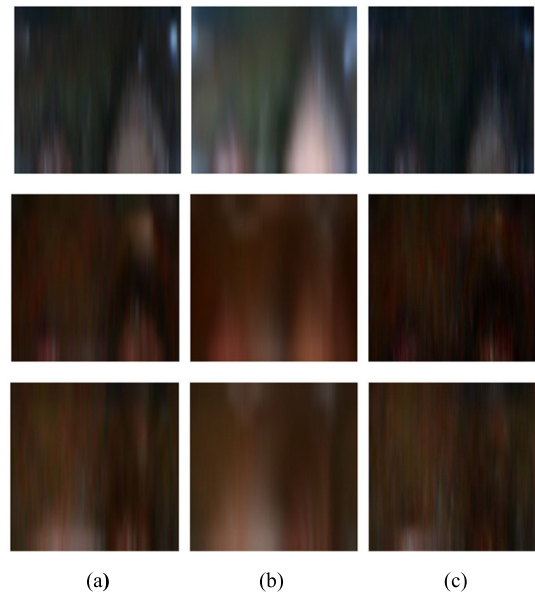
(a)

(b)

(c)
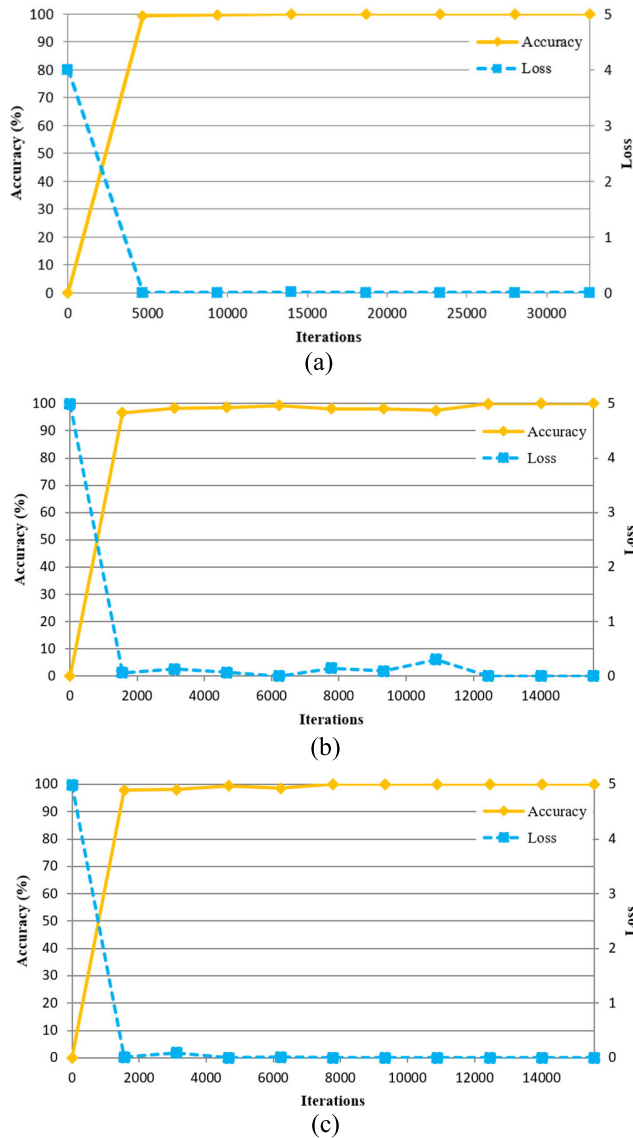
FIGURE 10. Loss and accuracy curves of CNN training. (a) 1st CNN training for iris region (b) 2nd CNN training for 1st periocular region (c) 3rd CNN training for 2nd periocular region.

(a)   (b)   (c)

FIGURE 11. Examples of images generated by DeblurGAN generator (a) Original image, (b) motion blur image, and (c) generated image.

and decidability value (d-prime value). The d-prime value (d') is used for objective performance evaluation of iris recognition in NICE.II contest [2], and is calculated using Equation (9). The d-prime value is calculated based on the mean (μ_A and μ_I) and standard deviation of (σ_A and σ_I) authentic and imposter matching distributions. FAR and FRR in biometrics are commonly caused by the overlap of authentic and imposter matching distributions; FAR, FRR, and EER tend to decrease as there are no overlaps between these two distributions with a large distance between them. The d-prime value in Equation (9) increases when the two distributions are far apart and decreases when the two distributions are closer with overlaps; thus, the biometric system is estimated to have a better performance when the value of d-prime increases.

$$d' = \frac{|\mu_A - \mu_I|}{\sqrt{\frac{\sigma_A^2 + \sigma_I^2}{2}}} \qquad (9)$$

TABLE 6. Comparisons of PSNR of generated image by proposed method and the state-of-the-art methods.

Model	PSNR
CycleGAN [51]	26.72
Pix2pix [22]	24.1
Proposed method	31.33

In Table 6, we quantitatively compare the PSNR of generated image by proposed method and the state-of-the-art methods [22], [51]. As shown in this table, our method outperforms the state-of-the-art methods for image generation.

In this study, iris recognition performance was evaluated using EER, receiver operating characteristic (ROC) curve,

2) ABLATION STUDY (NICE.II DATABASE)

Table 7 presents the comparison of the recognition rate of when motion blurring and DeblurGAN based deblurring proposed in this study were performed and when they were not performed for the iris region using NICE.II database. As shown in Table 7, the recognition performance is better when motion blurring and DeblurGAN based deblurring proposed in this study were performed, compared to when they were not performed. Moreover, the comparison of the recognition rate of when motion blurring and DeblurGAN based deblurring were performed and when they were not performed for the periocular region of NICE.II database are presented in Table 8. Unlike in Table 7, however, the recognition performance is poorer when motion blurring and DeblurGAN based deblurring were performed for the periocular region, compared to when they were not performed. As explained

TABLE 7. Propose method experimented for the iris region.

Method	Two-fold cross validation	EER (%)		d-prime value	
		Each fold	Average	Each fold	Average
Without motion blurring & deblurring	1st fold	12.07	12.79	2.23	2.25
	2nd fold	13.51		2.26	
With motion blurring & deblurring	1st fold	12.92	11.49	2.45	2.6
	2nd fold	10.05		2.74	

TABLE 8. Propose method experimented for the periocular region.

Method	Two-fold cross validation	EER (%)		d-prime value	
		Each fold	Average	Each fold	Average
Without motion blurring & deblurring	1st fold	10.37	11.87	2.54	2.42
	2nd fold	13.36		2.29	
With motion blurring & deblurring	1st fold	11.78	13.12	2.42	2.29
	2nd fold	14.45		2.15	

in Section IV.B, the periocular region has a relatively less amount of noises than the iris region, which contains detailed iris texture and an extensive amount of noises including SR and external light reflection (ghost) in cornea. In addition, the recognition performance was deteriorated partly due to the features that can be obtained from the periocular region such as eyelid shape or wrinkles being damaged after blurring and image restoration by DeblurGAN and the skin color being varied depending on the lighting, thus causing difficulty in predicting the original color.

In Table 9 and Figure 12, the recognition performance is compared between when recognition was performed for the iris region or the 1st or 2nd periocular region, and when the matching scores of all regions proposed in this study were fused with SVM. In Figure 12, genuine acceptance rate (GAR) is 100 − FRR (%). As shown in Table 9 and Figure 12, the recognition performance was found to be better when the matching scores of all regions proposed in this study were fused with SVM.

In addition, we compared our SVM-based fusion method to other famous score level fusion methods of weighted SUM and weighted PRODUCT [61], [62] methods. As shown in Table 9, proposed SVM-based fusion method outperforms the weighted SUM and weighted PRODUCT, from which we can estimate that SVM can obtain more sophisticated classifier plane on three dimensional spaces of matching distances compared to weighted SUM and weighted PRODUCT [61], [62].
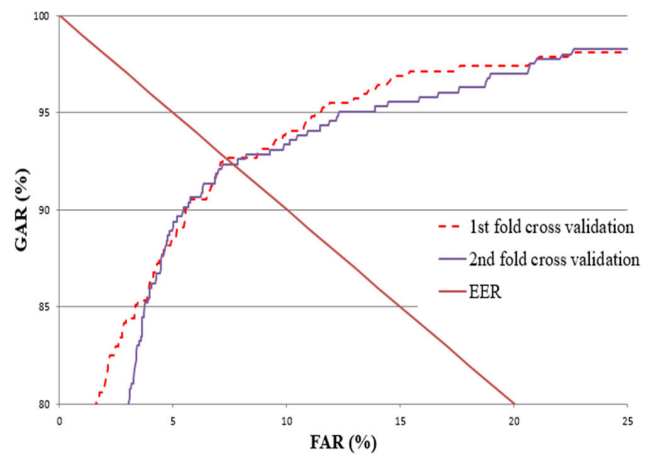
3) COMPARISONS WITH THE STATE-OF-THE-ART METHODS (NICE.II DATABASE)

Table 10 presents the comparison of the recognition performance between the proposed method and the state-of-the-art

TABLE 9. Recognition accuracies in case of using iris (or periocular) region compared to those using all the regions. (A: Only with iris region, B: Only with the 1st periocular region, C: Only with the 2nd periocular region, D: Score fusion by SVM with all the regions, E: Score fusion by Weighted SUM [61], [62] with all the regions, F: Score fusion by Weighted PRODUCT [61], [62] with all the regions).

Method	Two-fold cross validation	EER (%)		d-prime value	
		Each fold	Average	Each fold	Average
A	1st fold	12.92	11.49	2.45	2.6
	2nd fold	10.05		2.74	
B	1st fold	10.37	11.87	2.54	2.42
	2nd fold	13.36		2.29	
C	1st fold	10.36	11.63	2.59	2.46
	2nd fold	12.89		2.32	
D	1st fold	7.33	7.49	3.34	3.4
	2nd fold	7.64		3.46	
E	1st fold	8.63	8.75	2.94	2.96
	2nd fold	8.86		2.98	
F	1st fold	8.51	8.59	3.03	3.04
	2nd fold	8.67		3.05	

FIGURE 12. ROC curves of recognition accuracies by proposed method.

TABLE 10. Comparisons of recognition accuracy by proposed method with those by the state-of-the-art methods.

Method	EER (%)	d-prime value
Sajjad et al. [8]	18.82	*Not reported*
Proença et al. [42]	16 (approximate value)	2.42 (approximate value)
Zanlorensi et al. [43]	13.98	2.25
Tan et al. [44]	12 (approximate value)	2.57
Lee et al. [9]	10.36	2.62
Lee et al. [18]	8.58	3.31
Proposed method	7.49	3.4

methods. From the table it can be seen that the proposed method exhibited a better recognition performance than the state-of-the-art methods.

4) ABLATION STUDY (MICHE DATABASE)

Table 11 presents the comparison of the recognition rate of when motion blurring and DeblurGAN based deblurring

TABLE 11. Propose method experimented for the iris region.

Method		Two-fold cross validation	EER (%)		d-prime value	
			Each fold	Avg.	Each fold	Avg.
Without motion blurring & deblurring	Galaxy S4	1st fold	19.58	18.08	1.61	1.72
		2nd fold	16.57		1.83	
	iPhone5	1st fold	20.88	19.5	1.57	1.67
		2nd fold	18.12		1.77	
With motion blurring & deblurring	Galaxy S4	1st fold	17.67	17.07	1.75	1.83
		2nd fold	16.46		1.91	
	iPhone5	1st fold	18.67	18.14	1.75	1.76
		2nd fold	17.62		1.77	

TABLE 12. Recognition accuracies in case of using iris (or periocular) region compared to those using all the regions
(A: Only with iris region, B: Only with the 1st periocular region, C: Only with the 2nd periocular region, D: Score fusion by SVM with all the regions).

Method	Sub-dataset	Two-fold cross validation	EER (%)		d-prime value	
			Each fold	Average	Each fold	Average
A	Galaxy S4	1st fold	17.67	17.07	1.75	1.83
		2nd fold	16.46		1.91	
	iPhone 5	1st fold	18.67	18.14	1.75	1.76
		2nd fold	17.62		1.77	
B	Galaxy S4	1st fold	22.78	18.32	1.59	1.82
		2nd fold	13.85		2.04	
	iPhone 5	1st fold	21.7	20.91	1.56	1.66
		2nd fold	20.12		1.75	
C	Galaxy S4	1st fold	22.57	19.41	1.58	1.72
		2nd fold	16.25		1.85	
	iPhone 5	1st fold	23.4	21.19	1.64	1.67
		2nd fold	18.98		1.7	
D	Galaxy S4	1st fold	16.83	14.18	1.88	2.22
		2nd fold	11.52		2.56	
	iPhone 5	1st fold	16.36	17.02	1.87	1.82
		2nd fold	17.68		1.76	

FIGURE 13. ROC curves of recognition accuracy by proposed method.

TABLE 13. Comparisons of recognition accuracies by proposed method with those by the state-of-the-art methods.

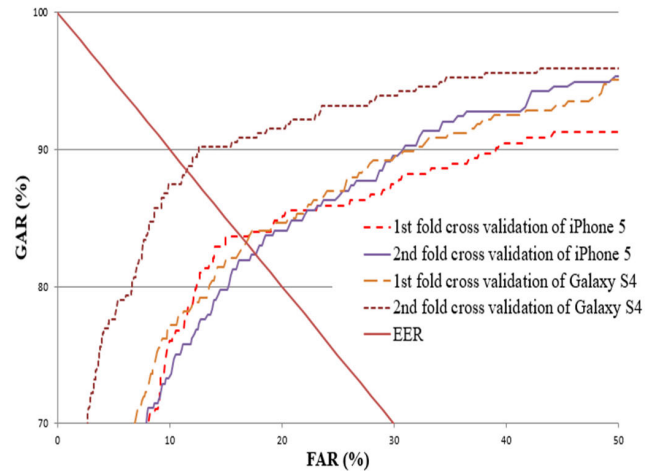| Method	Sub-dataset	EER (%)	d-prime value
Raja et al. [45]	Galaxy S4	38.8	6.49
	iPhone5	38.6	6.21
Santos et al. [46]	Galaxy S4	19.8	6.13
	iPhone5	22	5.44
Lee et al. [9]	Galaxy S4	17.9	1.87
	iPhone5	17.45	2.00
Lee et al. [18]	Galaxy S4	15.49	2.09
	iPhone5	17.39	1.87
Proposed method	Galaxy S4	14.18	2.22
	iPhone5	17.02	1.82

proposed in this study were performed and when they were not performed for the iris region of MICHE database.

As shown in Table 11, the recognition performance for MICHE database is also better when motion blurring and DeblurGAN based deblurring proposed in this study were performed, compared to when they were not performed.

In Table 12 and Figure 13, the recognition performance is compared between when recognition was performed for the iris region or for the 1st or 2nd periocular region, and when the matching scores of all regions proposed in this study were fused with SVM. As shown in Table 12 and Figure 13, the recognition performance was found to be better when the matching scores of all regions proposed in this study were fused with SVM.

5) COMPARISONS WITH THE STATE-OF-THE-ART METHODS (MICHE DATABASE)

Table 13 lists the comparison of the recognition performance between the proposed method and the state-of-the-art methods. From the table it can be seen that the proposed method

exhibited a better recognition performance than the state-of-the-art methods.

E. ANALYSIS AND DISCUSSION
1) ANALYSIS OF RECOGNITION ACCURACIES

Figure 14 shows the successful cases of genuine matching by the proposed method. Figures 14(a) and (c) show the successful case of genuine matching even under different lighting environment of the two irises being recognized. Figure 14(b) shows the successful case of when the size of the two irises being recognized is different. Figure 14(d) shows the successful case of genuine matching even when the reflected light is substantial, and the external light fluctuates.

Figure 15 shows the failed recognition cases. Figures 15(a) - (d) are false acceptance cases, while (e) - (h) are false rejection cases. In Figure 15(a), as the iris is located at the end of the image, features are not easily extracted from the periocular region and the iris pattern is unclear. Hence, it is a false acceptance case in which a different individual is recognized as the actual subject. Figures 15(b) and (c) are false acceptance cases because of rotation and noises such as reflected light in addition to the eyelid and skin color being very similar. Figure 15(d) is a false acceptance case because the reflect light in the iris is substantial while the makeup and eyebrow are very similar. Figure 15(e) is a false rejection case because of unclear iris pattern and severe
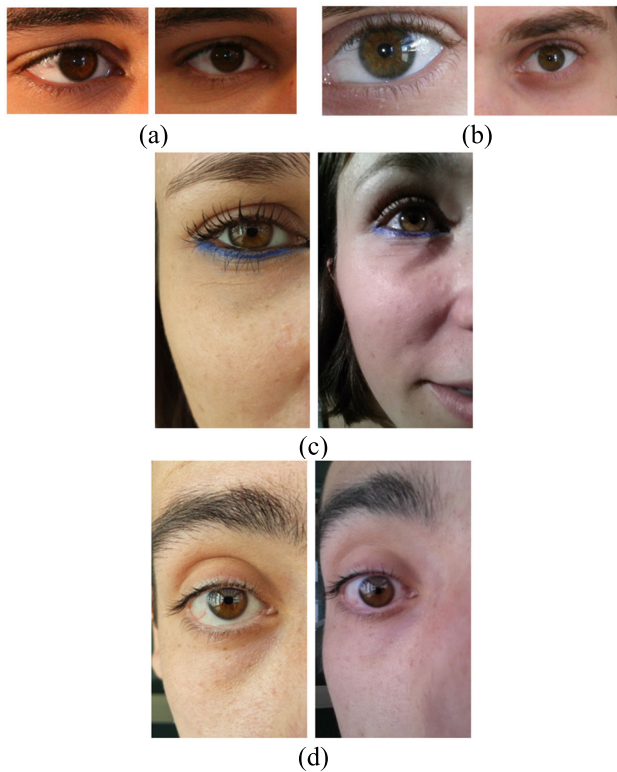
FIGURE 14. Correct recognition cases of genuine matching. (a) Case strong against lighting fluctuation, (b) Case strong against off-angle and changes in iris size, (c) Case strong against reflected light in iris and lighting fluctuation, (d) Case strong against reflected light and external light noise. (a) and (b) are from NICE.II database whereas (c) and (d) are from MICHE database.

reflected light even when the two images have the same class. Figures 15(f) and (g) are failed cases because of severe off-angle, changed eye shape, and changed eyelid shape due to external light fluctuation even if the images are of the same class. Figure 15(h) is a false rejection case due to severe reflected light and a changed skin color caused by lighting fluctuation.

Overall, the proposed method was proven to be effective against noises such as a change in iris size, fluctuation in visible lighting, and off-angle which occur frequently in an unconstrained environment. However, the recognition rate drops when the eyelid shape extracted as a periocular feature is distorted due to off-angle, when the periocular region is difficult to extract because the iris is positioned at the edge of an image, and when a periocular feature becomes similar due to makeup.

2) ANALYSIS OF FEATURE MAPS

In this section, the difference in the feature map is analyzed for the cases in which the recognition failed for the original image in genuine matching (false rejection); however, the recognition was successful when motion blurring and DeblurGAN based deblurring proposed in this study were used. Figures 16(a)–(c) shows the first, second, and third examples. The feature map is the image of visualizing the
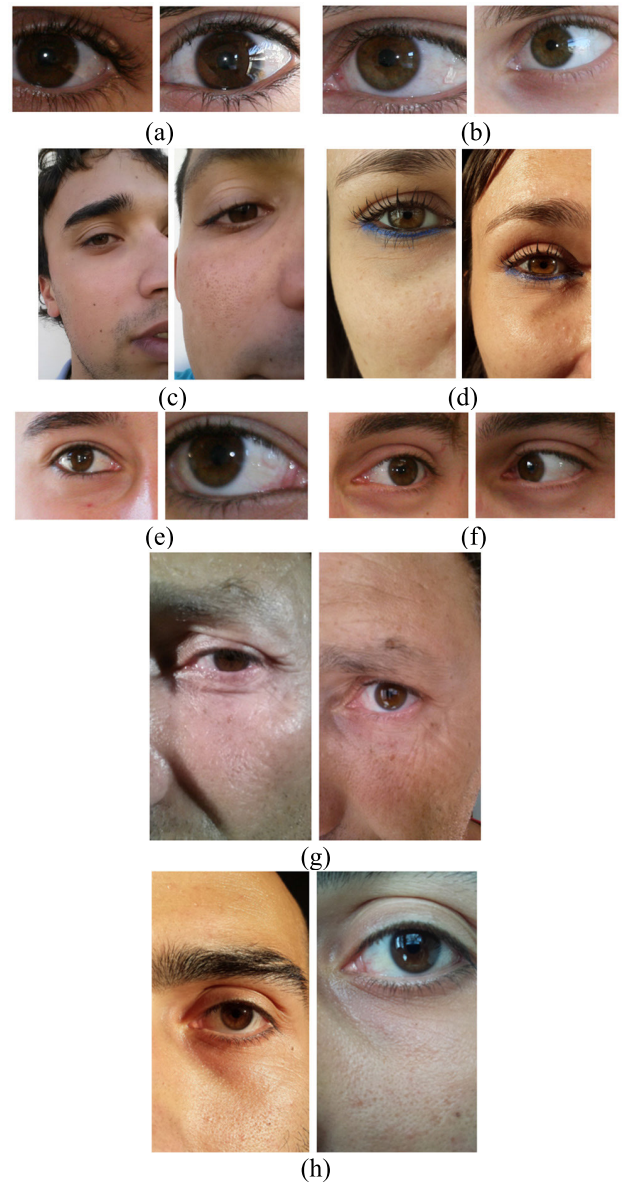
FIGURE 15. Failed recognition cases. In (a) - (d), false acceptance cases are shown whereas false rejection cases are shown in (e) - (h). (a) Case of false acceptance due to periocular region being difficult to obtain and unclear iris pattern, (b) Case of false acceptance due to rotation, reflected light, and similar wrinkle pattern and skin color, (c) Case of false acceptance due to unclear iris pattern, (d) Case of false acceptance due to substantial reflected light in iris and similar makeup, (e) Case of false rejection due to unclear iris pattern, reflected light, and severe off-angle, (f) Case of false rejection due to a change in eyelid shape caused by severe off-angle, (g) Case of false rejection due to a change in eyelid shape caused by external light and changed facial expression, (h) Case of false rejection due to substantial reflected light and changed skin color (a), (b), (e), and (f) are from NICE.II database whereas (c), (d), (g), and (h) are from MICHE database.

output of the 8th convolution layer in the CNN model for recognition which is presented in Table 4.

In Figures 16(a)–(c), the left-top, right-top, left-bottom, and right-bottom images show the feature maps of enrolled original image, corresponding deblurred image by our method, recognized original image, and corresponding deblurred image by our method, respectively.
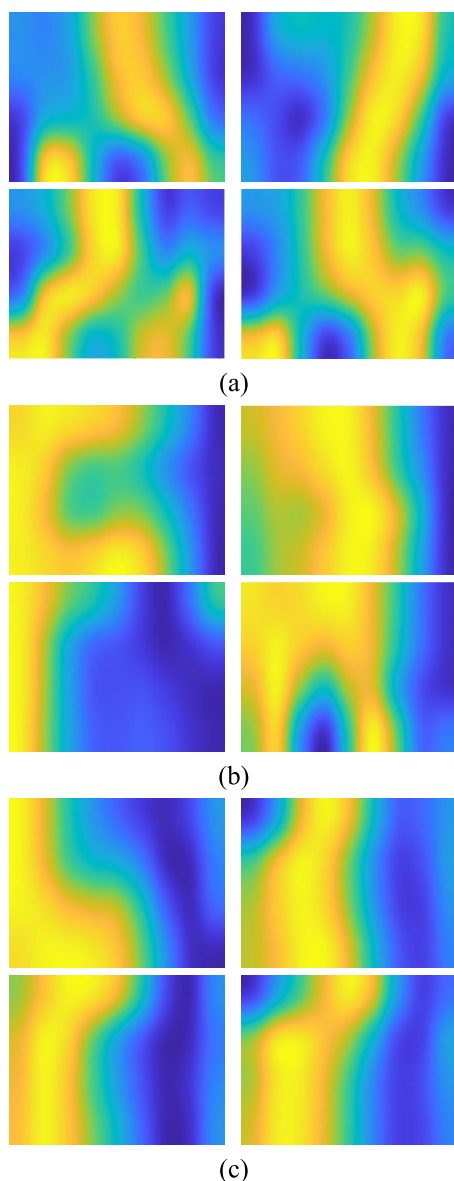
FIGURE 16. Comparison of the feature map between the original image and restored image. (a) - (c) respectively show the 1st–3rd examples. In (a) - (c), the left-top, right-top, left-bottom, and right-bottom images show the feature maps of enrolled original and corresponding deblurred images, the feature map of recognized original and corresponding deblurred images, respectively. In (a) - (c), upper and lower images show the enrolled and recognized images, respectively.

In Figures 16(a)–(c), upper and lower images show the enrolled and recognized images, respectively. In the feature maps, the area of important features are shown in brighter and yellow colors whereas the regions of insignificant features are shown in darker and bluish colors [55]. Therefore, the two images which have the similar areas of yellow and bluish colors usually have similar features, which causes the better similarity in recognition.

Figure 16 shows that feature maps of the original images (the left-top and left-bottom (the 1st column) images in Figures 16(a)-(c)) have a significant difference in the areas of yellow and bluish colors between the enrolled and recognized images. However, the feature maps of the images generated by the proposed DeblurGAN method (the right-top and right-bottom (the 2nd column) images in Figures 16(a)-(c)) have more similar areas of yellow and bluish colors between the enrolled and recognized images. Accordingly, when motion blurring and DeblurGAN-based deblurring are used, the recognition performance can be improved by making the feature maps between the enrolled and recognized images similar during genuine iris matching. These results were also confirmed in Tables 7 and 11.

V. CONCLUSION

This study proposed a method for enhancing the quality of iris images by blurring the iris region and deblurring using DeblurGAN-based method, thus improving the recognition performance. DeblurGAN was trained using the images to which a random motion blur kernel is applied and the original images as input. Using the normalized iris image to which a random motion blur kernel is applied as an input for the proposed model, the restored iris image is obtained from the generator. Then, the feature vector was extracted using the CNN for feature extractor from this image. The recognition performance was enhanced through score level fusion of three distance values for three pairs of feature vectors between enrolled image and recognized image using SVM. When the performance was measured based on a two-fold cross validation method using NICE.II training database and MICHE database, the proposed method resulted in a better recognition performance than when the proposed method was not used or the state-of-the-art methods were used. In case of iris recognition with near-infrared light image, iris patterns are distinctive, from which the high accuracy of iris recognition can be obtained. However, in case of iris recognition with visible light image like our research, there are lots of cases that iris patterns are not visible as shown in Figure 6, which degrades the accuracies of iris recognition. In this case, if the iris recognition is supported from periocular area, the accuracies can be improved, and these were already confirmed in previous researches [6], [9], [18], [46]. Therefore, we used three information from iris and two periocular areas in our research as shown in Figure 1.

In the future, various approaches for improving the recognition performance will be studied by automatically generating the iris region occluded by reflected light, eyelash, eyelid, and glasses frame using GAN methods. Furthermore, another method will be examined for reducing the errors of genuine matching or imposter matching caused by the dependence on the periocular region when the eyelid shape is changed due to off-angle or a change in facial expression.

REFERENCES

[1] J. Daugman, "How iris recognition works," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 1, pp. 21–30, Jan. 2004.

[2] NICE. (2020). *NICE Training Dataset*. Accessed: Feb. 11, 2020. [Online]. Available: http://nice2.di.ubi.pt/

[3] MICHE (2020). *MICHE Database*. Accessed: Feb. 25, 2020. [Online]. Available: http://biplab.unisa.it/MICHE/database/

[4] B. Kaur, S. Singh, and J. Kumar, "Iris recognition using Zernike moments and polar harmonic transforms," *Arabian J. Sci. Eng.*, vol. 43, no. 12, pp. 7209–7218, Dec. 2018.

[5] M. Hamd and S. Ahmed, "Fourier descriptors for iris recognition," *Int. J. Comput. Digit. Syst.*, vol. 6, no. 5, pp. 285–291, Sep. 2017.

[6] N. U. Ahmed, S. Cvetkovic, E. H. Siddiqi, A. Nikiforov, and I. Nikiforov, "Combining iris and periocular biometric for matching visible spectrum eye images," *Pattern Recognit. Lett.*, vol. 91, pp. 11–16, May 2017.

[7] C. Galdi and J.-L. Dugelay, "FIRE: Fast iris RE cognition on mobile phones by combining colour and texture features," *Pattern Recognit. Lett.*, vol. 91, pp. 44–51, May 2017.

[8] M. Sajjad, C.-W. Ahn, and J.-W. Jung, "Iris image enhancement for the recognition of non-ideal iris images," *KSII Trans. Internet Inf. Syst.*, vol. 10, no. 4, pp. 1904–1926, 2016.

[9] M. B. Lee, H. G. Hong, and K. R. Park, "Noisy ocular recognition based on three convolutional neural networks," *Sensors*, vol. 17, no. 12, pp. 1–26, 2017.

[10] Z. Zhao and A. Kumar, "Towards more accurate iris recognition using deeply learned spatially corresponding features," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Venice, Italy, Oct. 2017, pp. 3809–3818.

[11] A. Gangwar and A. Joshi, "DeepIrisNet: Deep iris representation with applications in iris recognition and cross-sensor iris recognition," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Phoenix, AZ, USA, Sep. 2016, pp. 2301–2305.

[12] S. Minaee, A. Abdolrashidiy, and Y. Wang, "An experimental study of deep convolutional features for iris recognition," in *Proc. IEEE Signal Process. Med. Biol. Symp. (SPMB)*, Philadelphia, PA, USA, Dec. 2016, pp. 1–6.

[13] K. Nguyen, C. Fookes, A. Ross, and S. Sridharan, "Iris recognition with Off-the-Shelf CNN features: A deep learning perspective," *IEEE Access*, vol. 6, pp. 18848–18855, 2018.

[14] H. Proenca and J. C. Neves, "Segmentation-less and non-holistic deep-learning frameworks for iris recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Long Beach, CA, USA, Jun. 2019, pp. 1–10.

[15] A. S. Al-Waisy, R. Qahwaji, S. Ipson, S. Al-Fahdawi, and T. A. M. Nagem, "A multi-biometric iris recognition system based on a deep learning approach," *Pattern Anal. Appl.*, vol. 21, no. 3, pp. 783–802, Aug. 2018.

[16] K. Wang and A. Kumar, "Toward more accurate iris recognition using dilated residual features," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 12, pp. 3233–3245, Dec. 2019.

[17] M. G Alaslani and L. A. Elrefaei, "Convolutional neural network based feature extraction for IRIS recognition," *Int. J. Comput. Sci. Inf. Technol.*, vol. 10, no. 2, pp. 65–78, Apr. 2018.

[18] M. B. Lee, Y. H. Kim, and K. R. Park, "Conditional generative adversarial Network- based data augmentation for enhancement of iris recognition accuracy," *IEEE Access*, vol. 7, pp. 122134–122152, 2019.

[19] K. Y. Shin, G. P. Nam, D. S. Jeong, D. H. Cho, B. J. Kang, K. R. Park, and J. Kim, "New iris recognition method for noisy iris images," *Pattern Recognit. Lett.*, vol. 33, no. 8, pp. 991–999, Jun. 2012.

[20] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, Montréal, QC, Canada, Dec. 2014, pp. 2672–2680.

[21] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*. [Online]. Available: http://arxiv.org/abs/1411.1784

[22] 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. 1125–1134.

[23] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas, "Deblur-GAN: 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.

[24] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of Wasserstein GANs," in *Proc. 31st Conf. Neural Inf. Process. Syst.*, Long Beach, CA, USA, 2017, pp. 5767–5777.

[25] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training GANs," in *Proc. 30th Conf. Neural Inf. Process. Syst.*, Barcelona, Spain, Dec. 2016, pp. 2234–2242.

[26] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," 2017, *arXiv:1701.07875*. [Online]. Available: http://arxiv.org/abs/1701.07875

[27] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, Amsterdam, The Netherlands, Oct. 2016, pp. 694–711.

[28] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Miami, FL, USA, Jun. 2009, pp. 248–255.

[29] 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, Jun. 2016, pp. 770–778.

[30] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," 2016, *arXiv:1607.08022*. [Online]. Available: http://arxiv.org/abs/1607.08022

[31] V. Nair and G. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. 27th Int. Conf. Mach. Learn. (ICML)*, Haifa, Israel, Jun. 2010, pp. 807–814.

[32] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, Apr. 2011.

[33] H. Proenca, S. Filipe, R. Santos, J. Oliveira, and L. A. Alexandre, "The UBIRIS.v2: A database of visible wavelength iris images captured On-the-Move and At-a-Distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1529–1535, Aug. 2010.

[34] NVIDIA (2020). *NVIDIA GeForce GTX 1080*. Accessed: Mar. 10, 2020. [Online]. Available: https://www.geforce.com/hardware/desktop-gpus/geforce-gtx-1080/specifications

[35] TensorFlow (2020). *TensorFlow*. Accessed: Mar. 12, 2020. [Online]. Available: https://www.tensorflow.org/

[36] Y. Jia. *Caffe: An Open Source Convolutional Architecture for Fast Feature Embedding*. Accessed: Mar. 20, 2020. [Online]. Available: http://caffe.berkeleyvision.org/

[37] G. Boracchi and A. Foi, "Modeling the performance of image restoration from motion blur," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3502–3517, Aug. 2012.

[38] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.

[39] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, San Diego, CA, USA, Dec. 2014, pp. 1–15.

[40] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2010.

[41] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Santiago, Chile, Dec. 2015, pp. 1026–1034.

[42] H. Proença and G. Santos, "Fusing color and shape descriptors in the recognition of degraded iris images acquired at visible wavelengths," *Comput. Vis. Image Understand.*, vol. 116, no. 2, pp. 167–178, Feb. 2012.

[43] L. A. Zanlorensi, E. Luz, R. Laroca, A. S. Britto, L. S. Oliveira, and D. Menotti, "The impact of preprocessing on deep representations for iris recognition on unconstrained environments," in *Proc. 31st SIBGRAPI Conf. Graph., Patterns Images (SIBGRAPI)*, Parana, Brazil, Oct. 2018, pp. 289–296.

[44] T. Tan, X. Zhang, Z. Sun, and H. Zhang, "Noisy iris image matching by using multiple cues," *Pattern Recognit. Lett.*, vol. 33, no. 8, pp. 970–977, Jun. 2012.

[45] K. B. Raja, R. Raghavendra, V. K. Vemuri, and C. Busch, "Smartphone based visible iris recognition using deep sparse filtering," *Pattern Recognit. Lett.*, vol. 57, pp. 33–42, May 2015.

[46] G. Santos, E. Grancho, M. V. Bernardo, and P. T. Fiadeiro, "Fusing iris and periocular information for cross-sensor recognition," *Pattern Recognit. Lett.*, vol. 57, pp. 52–59, May 2015.

[47] M. De Marsico, M. Nappi, F. Narducci, and H. Proença, "Insights into the results of MICHE I—Mobile iris challenge evaluation," *Pattern Recognit.*, vol. 74, pp. 286–304, Feb. 2018.

[48] DM Lab. (2020). *Dongguk DeblurGAN and CNN for Iris Recognition*. Accessed: Apr. 10 2020. [Online]. Available: http://dm.dgu.edu/link.html

[49] NVIDIA. (2020). *CUDA*. Accessed: Mar. 11, 2020. [Online]. Available: https://developer.nvidia.com/cuda-90-download-archive

[50] NVIDIA. (2020). *CUDNN*. Accessed: Mar. 11, 2020. [Online]. Available: https://developer.nvidia.com/cudnn

[51] 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. 2223–2232.

[52] Caffe. (2020). *Cross-Entropy Loss Function*. Accessed: Dec. 11, 2020. [Online]. Available: http://caffe.berkeleyvision.org/doxygen/classcaffe_1_1SoftmaxWithLossLayer.html

[53] K. Kashihara, "Iris recognition for biometrics based on CNN with super-resolution GAN," in *Proc. IEEE Conf. Evol. Adapt. Intell. Syst.*, Bari, Italy, May 2020, pp. 1–6.

[54] S. Minaee and A. Abdolrashidi, "Iris-GAN: Learning to generate realistic iris images using convolutional GAN," 2018, *arXiv:1812.04822*. [Online]. Available: http://arxiv.org/abs/1812.04822

[55] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," 2016, *arXiv:1610.02391*. [Online]. Available: http://arxiv.org/abs/1610.02391

[56] X. Huang, L. Ren, and R. Yang, "Image deblurring for less intrusive iris capture," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Miami, FL, USA, Jun. 2009, pp. 1558–1565.

[57] F. Alaoui, K. Assid, V. Dembele, and A. Nassim, "Application of blind deblurring algorithm for iris biometric," *Int. J. Comput. Appl.*, vol. 79, no. 3, pp. 11–15, Oct. 2013.

[58] J. Liu, Z. Sun, and T. Tan, "Iris image deblurring based on refinement of point spread function," in *Proc. Chin. Conf. Biometric Recognit.*, Guangzhou, China, Dec. 2012, pp. 184–192.

[59] J. Liu, Z. Sun, and T. Tan, "A novel image deblurring method to improve iris recognition accuracy," in *Proc. IEEE Int. Joint Conf. Biometrics*, Washington, DC, USA, Oct. 2011, pp. 1–8.

[60] S. Jamaludin, N. Zainal, and W. M. D. W Zaki, "Deblurring of noisy iris images in iris recognition," *Bull. Electr. Eng. Informat.*, vol. 10, no. 1, pp. 156–159, Feb. 2021.

[61] A. Aboshosha, K. A. El Dahshan, E. A. Karam, and E. A. Ebeid, "Score level fusion for fingerprint, iris and face biometrics," *Int. J. Comput. Appl.*, vol. 111, no. 4, pp. 47–55, Feb. 2015.

[62] M. Singh, R. Singh, and A. Ross, "A comprehensive overview of biometric fusion," *Inf. Fusion*, vol. 52, pp. 187–205, Dec. 2019.

JIN KYU KANG received the B.S. degree in electronics and electrical engineering from Dongguk University, Seoul, South Korea, in 2016. He is currently pursuing the combined course of M.S. and Ph.D. degrees in electronics and electrical engineering with Dongguk University. His research interests include biometrics and deep learning. He helped the experiments and analysis of results.

HYO SIK YOON received the B.S. degree in electronics engineering from Kangwon National University, Chuncheon, South Korea, in 2015. He is currently pursuing the combined degree of master's and Ph.D. degrees in electronics and electrical engineering with Dongguk University. His research interests include pattern recognition and deep learning. He helped the experiments and analysis of results.
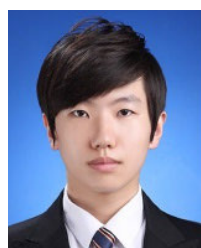
MIN BEOM LEE received the B.S. degree in information and telecommunication engineering from Dongyang Mirae University, Seoul, South Korea, in 2016. He is currently pursuing the combined course of M.S. and Ph.D. degrees in electronics and electrical engineering with Dongguk University. His research interests include biometrics and deep learning. He implemented the overall system and wrote the draft of original article.

KANG RYOUNG PARK (Member, IEEE) 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 deep learning. He supervised this research and helped the revision of draft of original article.

● ● ●