IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 12, 2019, accepted December 1, 2019, date of publication December 4, 2019,
date of current version December 16, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2957532

A Study on the Elimination of Thermal Reflections

GANBAYAR BATCHULUUN, HYO SIK YOON, DAT TIEN NGUYEN, TUYEN DANH PHAM, AND KANG RYOUNG PARK
Division of Electronics and Electrical Engineering, Dongguk University, Seoul 04620, South Korea

Corresponding author: Tuyen Danh Pham (phamdanhtuyen@gmail.com)

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 Recently, thermal cameras have been used in various surveillance and monitoring systems. In particular, in camera-based surveillance systems, algorithms are being developed for detecting and recognizing objects from images acquired in dark environments. However, it is difficult to detect and recognize an object due to the thermal reflections generated in the image obtained from a thermal camera. For example, thermal reflection often occurs on a structure or the floor near an object, similar to shadows or mirror reflections. In this case, the object and the areas of thermal reflection overlap or are connected to each other and are difficult to separate. Thermal reflection also occurs on nearby walls, which can be detected as artifacts when an object is not associated with this phenomenon. In addition, the size and pixel value of the thermal reflection area vary greatly depending on the material of the area and the environmental temperature. In this case, the patterns and pixel values of the thermal reflection and the object are similar to each other and difficult to differentiate. These problems reduce the accuracy of object detection and recognition methods. In addition, no studies have been conducted on the elimination of thermal reflection of objects under different environmental conditions. Therefore, to address these challenges, we propose a method of detecting reflections in thermal images based on deep learning and their elimination via post-processing. Experiments using a self-collected database (Dongguk thermal image database (DTh-DB), Dongguk items and vehicles database (DI&V-DB)) and an open database showed that the performance of the proposed method is superior compared to that of other state-of-the-art approaches.

INDEX TERMS Thermal reflection, thermal imaging, deep learning, thermal reflection detection, thermal reflection removal.

I. INTRODUCTION

Recently, in various fields, data analysis, object detection, pattern recognition, etc. have been performed using a long-wavelength infrared (LWIR) camera. LWIR cameras can measure electromagnetic radiation (EMR) in the range of 8–12 μm [1]. Most of the thermal radiation emitted from living bodies and objects is infrared, and in many cases, LWIR cameras are used to measure heat information.

Therefore, the LWIR camera is also called a thermal camera. Figure 1 shows the example of visible light and thermal images with thermal camera. Thermal cameras can enhance the visibility of near and far objects in dark environments without additional illuminators. However, it is difficult to detect and recognize an object due to the thermal reflections generated in the image obtained using a thermal camera.

FIGURE 1. Example of a thermal camera, a visible light image, and a thermal image, respectively, from left to right.

The EMR emitted by object reflects off the floor or walls nearby, creating shadow-like areas around the object. This is called the thermal reflection of the object. The pattern, size, and pixel value of thermal reflections vary greatly depending on the material, the generated heat in the range of reflection, the temperature of the object and the surrounding environment. For example, in an 8-bit thermal image, the pixels in the area of thermal reflection can have any value between 0–255.

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

FIGURE 2. Examples of thermal reflections. Left and right figures show visible light images and the corresponding thermal images.

In addition, unlike a typical shadow, thermal reflection may include a spatial feature in addition to the body shape of the object. Figure 2 shows the thermal image and the corresponding visible-light image. The area of thermal reflection is represented by a red dashed line.

Therefore, in this study, we propose a technique for the detection and removal of thermal reflections using deep convolutional neural network (CNN). There are no studies in which the elimination of thermal reflection from near and distant objects in thermal images is studied using deep learning-based methods that consider the aforementioned issues.

The remainder of this paper is divided into following sections. Related studies are described in Section II, and the contribution of this study is described in Section III. A detailed description of the proposed method is presented in Section IV. Experimental results and comparative experiments are presented in Section V. Finally, the conclusions of this study are presented in Section VI.

II. RELATED WORKS

The existing thermal camera-based studies can be divided into categories that include with and without thermal reflection elimination. There are a variety of detection [2]–[11], identification [12], [13], and recognition [14]–[16] methods that have been executed using thermal imaging. In addition, a survey study on thermal image-based methods has been conducted [17].

In the previous studies [2] and [8], a method for the detection of subjects in images with the halo effect was proposed using a contour-based approach. The Halo effect in a thermal image is a circular region of high intensity pixels that surrounds an object [18]. However, this contour-based method is not a method for the removal of the halo effect, but rather, an approach for accurately detecting subjects in thermal images with halo effects. In a separate study [19], the authors proposed a method of suppressing thermal reflection. In this approach, they experimented with visible light reflection and also studied the reflection of heat. In addition, they used various polarizers and plates, and graphically represented the change in the thermal reflection according to the angle of the plates. In this way, they proposed a thermal reflection suppression method considering the angles according to the plates of various materials and polarizers in the experiment. However, in this method, given that the angle varies depending on the material of the plates, the suppression performance is reduced when there are nearby walls or floors made of different materials. Moreover, these previous studies did not consider the problem of thermal reflection of various objects in different environments. Furthermore, there have been no studies that investigated the use of deep learning to address these problems.

Therefore, in this study, we present a method using deep learning that considers various environments, subjects, and objects to resolve the thermal reflection problem. Table 1 summarizes examples of related works and compares existing methods to the proposed approach.

III. CONTRIBUTIONS

This research is novel in the following four ways compared to previous works:

- There is no previous research on the elimination of thermal reflections from floors and walls in different environments. Therefore, this study proposes a method of removing thermal reflections of various objects (humans, vehicles, and items) in thermal images acquired in different environments (various illumination environments, indoor and outdoor environments, seasons, distances and angles).
- In this study, we propose a method using mask region-based CNN (R-CNN) to detect thermal reflection in various environments and objects.
- In this study, thermal reflection, and thermal rays are removed using the surrounding information of the area detected using Mask R-CNN, and post-processing based on morphological dilation, and complement operation.
- The Mask R-CNN model, generated data and self-constructed Dongguk thermal image database (DTh-DB) and Dongguk items and vehicles database

TABLE 1. Summary of comparison between the proposed method and previous studies.

Category		Method	Advantage	Disadvantage
Without removing thermal reflections	Handcrafted feature-based	Detection [2–11], recognition [14–16],	Large data acquisition, processing, and training are not required	- Handcrafted features are not suitable for use under various environments and camera settings. - Performance is affected by thermal reflections
	Deep feature-based	Identification [12, 13]	- The method extracts suitable features in a variety of environments and camera settings	- Performance is affected by thermal reflections - Large data acquisition, processing, and training requires more time
Removing thermal reflections	Handcrafted feature-based	Thermal reflection suppression [19]	- Large data acquisition, processing, and training are not required - Performance is not affected by thermal reflections	- It cannot be adopted in various environments because it requires additional polarizers and plates
	Deep feature-based	Removing thermal reflection (**proposed method**)	- Performance is not affected by thermal reflections - Applicable to various environments and camera settings	- It requires intensive training of deep CNN

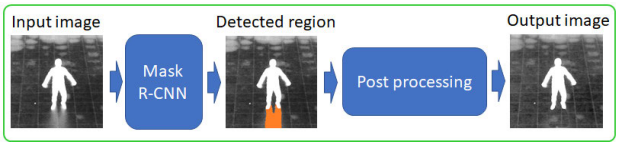

FIGURE 3. Example of a flow chart for the proposed method.

(DI&V-DB) that we used for training in this study, are published in [20] for performance evaluation by other researchers.

IV. PROPOSED METHOD

A. OVERALL PROCEDURE OF PROPOSED METHOD

In this section, we describe in detail the proposed method used in this study. The used thermal camera can acquire images (depth of 14 bits and size of 640 × 480 pixels) at a rate of 30 frames per second (fps) [21]. The object is visible in a dark or bright environment by measuring the temperature in the range from −40 °C to + 80 °C. However, there are many thermal reflections in the acquired image as previously described, which degrades the object detection and recognition performance. In this regard, this study proposes a method to remove thermal reflections from acquired images as shown in Figure 3. The Mask R-CNN is used to detect thermal reflection regions in the input thermal images, and the detected areas are removed during post-processing based on morphological dilation and complement operation. The following sections, IV.B and IV.C, describe thermal reflection detection and removal methods in detail.

B. THERMAL REFLECTION DETECTION USING MASK R-CNN

RetinaNet [22] was used in the Mask R-CNN-based [23] thermal reflection detection method used in this study.

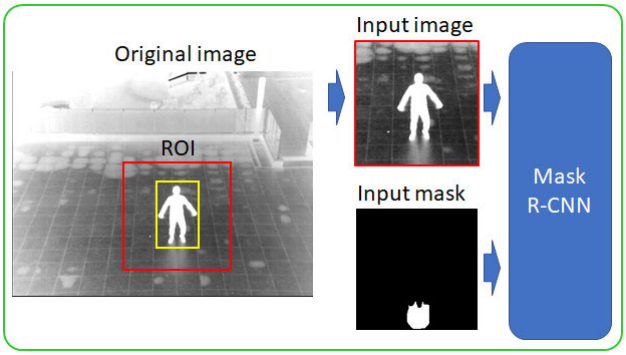

FIGURE 4. Example of cropping image and training mask R-CNN.

RetinaNet uses Resnet-50 [24] to extract features. In addition, we used a feature pyramid network (FPN) [25] and a small fully convolutional network (FCN) [26] instead of the region proposal network (RPN) [27] when detecting region of interest (ROI), and candidate object boxes.

In the thermal reflection detection method, box classification and box regression were simultaneously performed using small FCN subnets, and mask segmentation was performed using the detected final box as an input to another small FCN. Previous Mask R-CNN studies have reported that the Resnet-FPN backbone increases accuracy and speed when extracting features [23]. Therefore, in this study, thermal reflection was detected using Resnet-FPN/RetinaNet. The flowchart of Mask R-CNN used in this study is shown in Figure 5.

Tables 2 and 3 describe the structure of Resnet-50 and FPN included in Mask R-CNN. In Tables 4, 5, and 6, three FCN structures (classification subnet, box regression subnet, and mask segmentation network) included in Mask R-CNN

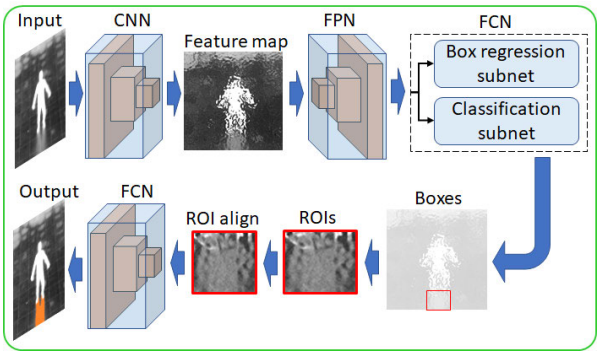

FIGURE 5. Example of mask R-CNN architecture.

TABLE 2. Description of the structure of Resnet-50.

Occurrence	Layer type	# of filters	Filter size (stride)
1 time	Input layer		
1 time	Conv1	64	7×7 (2)
1 time	Maxpool		3×3 (2)
3 times	Conv2 /C2	64	1×1 (4)
		64	3×3 (4)
		256	1×1 (4)
4 times	Conv3/C3	128	1×1 (8)
		128	3×3 (8)
		512	1×1 (8)
6 times	Conv4/C4	256	1×1 (16)
		256	3×3 (16)
		1024	1×1 (16)
3 times	Conv5/C5	512	1×1 (32)
		512	3×3 (32)
		2048	1×1 (32)

TABLE 3. Description of the structure of FPN.

Layer type	# of filters	Filter size (stride)
Stage 1		
C5 →Conv1 →P5	256	1×1 (1)
C5 →Conv2 → P6	256	3×3 (2)
P6→ Relu → Conv3 → P7	256	3×3 (2)
(C4 → Conv4) + (P5 → 2×Up) → P4	256	1×1 (1)
(C3 → Conv5) + (P4 → 2×Up) → P3	256	1×1 (1)
(C2 → Conv6) + (P3 → 2×Up) → P2	256	1×1 (1)
Stage 2		
P5 → Conv7 →*P5*	256	3×3 (1)
P6 → Conv8 →*P6*	256	3×3 (1)
P7 → Conv9 →*P7*	256	3×3 (1)
P4 → Conv10 →*P4*	256	3×3 (1)
P3 → Conv11 →*P3*	256	3×3 (1)
P2 → Conv12 →*P2*	256	3×3 (1)

are described. In the proposed method, a larger auto cropped image (red box) of the original image of Figure 4 with reference to the object region (yellow box) detected by the method of [12] is used as an input to the Mask R-CNN network. In the training phase, the cropped images and the manually created mask images are used as the input to the Mask R-CNN. In the test phase, thermal reflection is detected using only the cropped images. In Table 2, the dimension of the input image

TABLE 4. Description of the structure of classification subnet.

Layer type	# of filters	Filter size(stride)
Conv1	256	3×3 (1)
Relu1		
Conv2	256	3×3 (1)
Relu2		
Conv3	256	3×3 (1)
Relu3		
Conv4	256	3×3 (1)
Relu4		
Conv5	A	3×3 (1)
Sigmoid1		

TABLE 5. Description of the structure of box regression subnet.

Layer type	# of filters	Filter size(stride)
Conv1	256	3×3 (1)
Relu1		
Conv2	256	3×3 (1)
Relu2		
Conv3	256	3×3 (1)
Relu3		
Conv4	256	3×3 (1)
Relu4		
Conv5	4×A	3×3 (1)
Sigmoid1		

TABLE 6. Description of the structure of FCN.

Layer type	# of filters	Filter size (stride)
Conv1	256	3×3 (1)
Relu1		
Conv2	256	3×3 (1)
Relu2		
Conv3	256	3×3 (1)
Relu3		
Conv4	256	3×3 (1)
Relu4		
Trans-Conv1	256	2×2 (1)
Relu5		
Conv6	1	1×1 (1)
Relu6		

(input layer) and the class number (Softmax) are $224 \times 224 \times 1$ and 2×1, respectively.

In Tables 2 and 3, convolution sets (conv2, conv3, conv4, conv5) are expressed as (C2, C3, C4, C5). In Table 3, arrows, 2×Up, P1, and *P1* represent the next step, upsampling, the first feature map of FPN, and the final feature map of FPN, respectively. The first feature map of FPN (extracted in stage 1 of table 3) is extracted from the feature map extracted with Resnet-50. The final feature map of FPN (extracted in stage 2 of table 3) is a feature map to reduce the aliasing effect caused by upsampling. A detailed description of P1–P7 can be found in [22]. In Tables 4 and 5, A represents anchors, and in the proposed method, $A = 9$.

C. REMOVING THERMAL REFLECTION

Thermal reflections detected in the proposed method are removed during post-processing. The procedure for removing thermal reflections is shown in Figure 6. In the image where

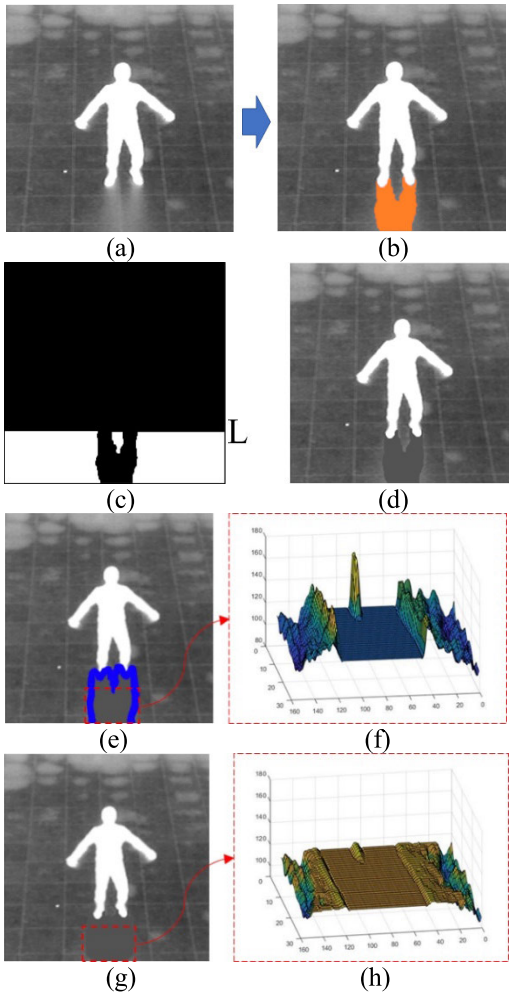

FIGURE 6. Example showing the removal of the thermal reflection area. (a) An original image; (b) a detection result with thermal reflection; (c) an ROI image for finding μ value; (d) the result after removing thermal reflection; (e) the resulting image with region of thermal rays; (f) a 3D review of a portion of the thermal rays; (g) the result after removing thermal rays; (h) a 3D view of a portion of the thermal rays after suppression.

thermal reflection is detected, as shown in Figure 6 (b), except the thermal reflection and the upper region (Figure 6 (c) black area) of the L line in Figure 6 (c), the mean μ of the pixel value of the original image in the remaining background (Figure 6 (c) white area) is calculated. Here, the L line is based on the top position of the detected thermal reflection area. By replacing the pixel value of the thermal reflection with the value of μ in the detected region, the thermal reflection is removed as shown in Figure 6 (d). However, as shown in Figure 6 (a), the value of the boundary of thermal reflection is not clearly distinguished from the background. Therefore, when performing Mask R-CNN-based detection proposed in this study, the thermal reflection is often detected to be slightly smaller than the actual thermal reflection region, in which case the thermal rays that surround the area of thermal reflection remain as shown in Figure 6 (d).

These thermal rays form discontinuous boundaries with the background cause object detection errors. To solve

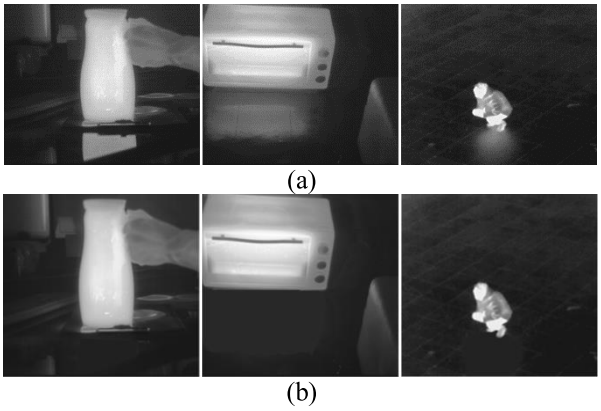

FIGURE 7. Example showing the removal of thermal reflection. (a) original images;(b) processed images.

this problem, in this study, the detected region X is made wider by a structuring element M using morphological dilation [28] and complement operation [29], and region Y where $(X \oplus M)$ and the detected region (X) do not overlap, is set as the thermal ray region (Figure 6 (e) blue area). Then, using Equation (4), the pixel values in the thermal ray region as shown in Figure 6 (f) are reduced as shown in Figure 6 (h). Equation (4) further reduces the higher values of the pixels corresponding to the thermal rays and causes a smaller reduction of the lower values. Figure 6 (g) shows the final result without the thermal reflection, and Figure 7 (b) shows the images with the thermal reflection removed for various objects.

$$A \oplus B = \bigcup_{b \epsilon B} A_b \tag{1}$$

$$B \backslash A = B \bigcap A^c \tag{2}$$

$$Y = (X \oplus M) \backslash X \tag{3}$$

$$Y' = Y - r \cdot (Y - \mu) \tag{4}$$

where Y and Y' represent the area of the original and changed thermal rays, respectively. Y and r represent the pixel value of thermal rays and a regulation factor that determines the extent to which the thermal rays will be suppressed, respectively.

V. EXPERIMENTAL RESULTS

A. DESCRIPTION OF EXPERIMENTAL SETUP AND DATABASES

The DTh-DB and DI&V-DB [20] collected in this study consist of thermal images of distant objects acquired in a dark or bright environment (including dawn, day and night) and near objects acquired in a dark or bright indoor environment. Images were acquired using different camera settings in various weather and seasonal environments as shown in Tables 7 and 8. The DTh-DB contains both visible light and thermal images captured by visible light and thermal cameras simultaneously, and a laptop computer [9]. However, only thermal images were used in this study. In addition, when acquiring DTh-DB, the height of the camera installed is

TABLE 7. Capturing conditions and the detail of the descriptions of DTh-DB.

Dataset	Condition	Detail description
I (see in Figure 8(a))	Humidity 62.6 %, wind speed of 1.3 m/s, 21.9 °C, afternoon, cloudy, autumn	- Object is distinctive in the current position. However, it is not distinctive in the left-upper area which is brighter in the thermal image - Thermal reflection is shown below the human area
II (see in Figure 8(b))	6.0 °C, afternoon, cloudy, humidity 39.6 %, wind speed of 1.9 m/s	- Object is distinctive in the current position. However, it is not distinctive in the left-lower area which is brighter in the thermal image - Thermal reflection is shown below the human area
III (see in Figure 8(c))	14.0 °C, afternoon, sunny, humidity 43.4 %, wind speed of 3.1 m/s	- The air heating system of the building increases the temperature outside the building in the thermal image
IV (see in Figure 8(d))	1.2 °C, morning, humidity 73.0 %, wind speed of 1.6 m/s	- The temperature of the window of the building is not kept over time, which causes objects to be indistinctive in the thermal image
V (see in Figure 8(e))	1.0 °C, afternoon, humidity 50.6 %, wind speed of 1.7 m/s	- The pixel values of trees and leaves are higher in the thermal image when it is sunny in cold environments
VI (see in Figure 8(f))	31.3 °C, noon, humidity 43.4 %, wind speed of 3.1 m/s	- The temperature of the human body and background is similar in the lower area of the thermal image, which increases the difficulty of segmentation of the human area from the background
VII (see in Figure 8(g))	10.2 °C, afternoon, cloudy and rainy, humidity 60.6 %, wind speed of 1.7 m/s	- Object is clearly observed in the thermal image - Thermal reflection is shown below the human area
VIII (see in Figure 8(h))	38.5 °C, noon, humidity 35.3 %, wind speed of 1.2 m/s	- The temperature of the human area is lower than that of the background in the thermal image
IX (see in Figure 8(i))	18.9 °C, night, humidity 62.6 %, wind speed of 1.3 m/s	- Small-sized objects are included in the thermal image - Thermal reflection is shown below the human area
X (see in Figure 8(j))	10.9 °C, dark night, humidity 48.3 %, wind speed of 2.0 m/s	- The dataset was collected at night in winter - Thermal reflection is shown below the human area
XI (see in Figure 8(k))	10.9 °C, dark night, humidity 48.3 %, wind speed of 2.0 m/s	- The temperature of the window of the building is not kept over time, which causes objects to be indistinctive in the thermal image
XII (see in Figure 8(l))	20.2 °C, dark night, humidity 58.6 %, wind speed of 1.2 m/s	- The air heating system of the building increases the temperature outside the building in the thermal image
XIII (see in Figure 8(m))	-2.0 °C, dark night, humidity 50.6 %, wind speed of 1.8 m/s	- The pixel value of trees and leaves is high in the thermal image due to sunlight during daytime
XIV (see in Figure 8(n))	12.0 °C, dark night, humidity 63.1 %, wind speed of 1.5 m/s	- The dataset was collected at night in winter - Thermal reflection is shown below the human area
XV (see in Figure 8(o))	28.0 °C, night, humidity 45.1 %, wind speed of 1.6 m/s	- The temperature of the background is higher than that of the human bodies in the thermal image due to sunlight during daytime
XVI (see in Figure 8(p))	10.0 °C, dark night, humidity 63.1 %, wind speed of 1.5 m/s	- The temperature of the human body and the background is similar in the left-upper area of the thermal image - Thermal reflection is shown below the human area
XVII (see in Figure 8(q))	19 °C, 62 lx.	- A man is walking in the corridor - Thermal reflection occurs below the feet in the thermal image
XVIII (see in Figure 8(r))	17 °C, 84 lx.	- A man is walking in the corridor - Thermal reflection occurs below the feet in the thermal image
XIX (see in Figure 8(s))	20 °C, 125 lx.	- A man is walking in the corridor - Thermal reflection occurs below the feet in the thermal image
XX (see in Figure 8(t))	18 °C, 130 lx.	- Pixel values below human feet are higher than other areas - Thermal reflection occurs below the feet in the thermal image
XXI (see in Figure 8(u))	17 °C, 1 lx.	- Two men are walking in a dark corridor - Thermal reflection occurs below the feet in the thermal image
XXII (see in Figure 8(v))	18 °C, 130 lx.	- Two men are in a corridor - Thermal reflection occurs below the feet in the thermal image

2.6–10 meters and the distance between an object and the cameras is 3–30 meters.

DTh-DB and DI&V-DB contain a total of 278,261 images. Detailed information and examples of databases can be found in Tables 7 and 8, and Figures 8 and 9. The experiment was conducted with two-fold cross-validation.

That is, half of the total data was used for training, the other half was used for testing, and the training and testing data

TABLE 8. Capturing conditions and detail descriptions of DI&V-DB.

Dataset	Condition	Detail description
I (see in Figure 9(a))	29 °C, humidity 53 %	- A plastic noodle cup with hot water on a table in a room - Thermal reflection is shown below the cup
II (see in Figure 9(b))	29 °C, humidity 53 %	- A water bottle with hot water on a bookshelf in a room - Thermal reflections are shown below the bottle and at both sides
III (see in Figure 9(c))	31 °C, humidity 75 %	- A hot oven on the floor in a room - Thermal reflection is shown below the oven
IV (see in Figure 9(d))	29 °C, humidity 53 %	- A water pot with hot water on the floor in a room - Thermal reflection is shown below the pot
V (see in Figure 9(e))	29 °C, humidity 53 %	- Thermal reflection of a human hand is shown on a table in a room
VI (see in Figure 9(f))	35 °C, humidity 80 %	- A car is moving on the road - Thermal reflection is shown below the car
VII (see in Figure 9(g))	35 °C, humidity 80 %	- A truck is moving on the road - Thermal reflection is shown below the truck
VIII (see in Figure 9(h))	35 °C, humidity 80 %	- A car is moving on the road - Thermal reflection is shown below the car
IX (see in Figure 9(i))	35 °C, humidity 80 %	- Back view of a car moving on the road - Thermal reflection is shown below the car
X (see in Figure 9(j))	35 °C, humidity 80 %	- Back view of a car moving on the road - Thermal reflection is shown below the car
XI (see in Figure 9(k))	35 °C, humidity 80 %	- Side view of a car moving on the road - Thermal reflection is shown below the car

were swapped, and the same process was repeated to determine the average value of the two testing accuracies to be the final accuracy. The model training does not contain the testing set.

The training and testing of the algorithm proposed in this study were performed using a desktop computer. The specification of the laptop computer used in the acquisition of the database is intel CPU (core i5-2520M CPU @ 2.50 GHz) and RAM (4 GB) and the specifications of the desktop computer are Nvidia graphic card (Nvidia GeForce GTX TITAN X [30]), intel CPU (core i7-6700 CPU @ 3.40 GHz (8 CPUs)) and RAM (32 GB). The proposed method is implemented and run using python-based Keras application programming interface (API) with Tensorflow backend engine [31] and OpenCV library [32].

B. THERMAL REFLECTIONS FROM DIFFERENT MATERIALS

In general, if the background heat is higher than the object heat as shown in Figure 8 (o), thermal reflection does not occur at the bottom area of the object. This section shows how thermal reflection occurs with different materials. Figure 10 shows an experiment using a glass pot with hot water and a metal can with cold water. Due to the auto exposure of the thermal camera, the background became dark in Figures 10 (a)–(g) and the background was bright in Figure 10 (h)–(m). The shape of the can on the plastic sheet in Figure 10 (k) is not the thermal reflection of the can, but the shadow of the can. Figure 11 shows different cases of thermal reflection. As shown in Figure 10 (a), the hot object causes thermal reflection on the near surface and the cold object causes shadow. The heat tends to reflect well on smooth surfaces as in the case of visible light. However, the reflection of heat is not affected by the color of the surface, unlike the case of visible light.

C. TRAINING OF VARIOUS CNN MODELS FOR THERMAL REFLECTION DETECTION

We compared various CNN models including cycle-consistent generative adversarial network (CycleGAN) [33], perceptual loss network (PLN) [34], and SegNet [35] to the performance of Mask R-CNN for thermal reflection detection. This section describes the training process and the results for the four methods that were compared. In the experiment, thermal images whose size was 224 × 224 pixels were used for training and testing. In the training of CycleGAN, cycle-consistency loss, identity loss, training epoch, learning rate, and mini-batch were set to 9.0, 1.0, 4,500, 0.0001, and 1, respectively. In addition, the mean squared error (MSE)-based loss [36] and adaptive moment estimation methods (Adam)-based optimizer [37] were used. In PLN, the learning rate and training epoch were set to 0.001 and 3,000, respectively. In the SegNet method, the learning rate, learning rate drop period, learning rate drop factor, momentum, mini-batch, and optimizer were set to 0.001, 10, 0.3, 0.9, 2, and the stochastic gradient descent with momentum (SGDM), respectively. In the Mask R-CNN method, batch-size, training epoch, step-size, and optimizer were set to 1, 100, 10000, and Adam, respectively. The learning rate was determined according to the epoch number. The Resnet-50 structure was used as the backbone of Mask R-CNN and additional fine-tuning was performed using this experimental data with pre-trained weights from the ImageNet database. Figure 12 shows the training loss for each method with the increase in the number of epochs. For all the methods, the training loss converged with the increase of the number of epochs, and it is evident that all the methods used in the experiment sufficiently converged with this experimental data.

When training each method, we used paired or unpaired dataset as shown in Figure 13. As such, PLN, SegNet, and

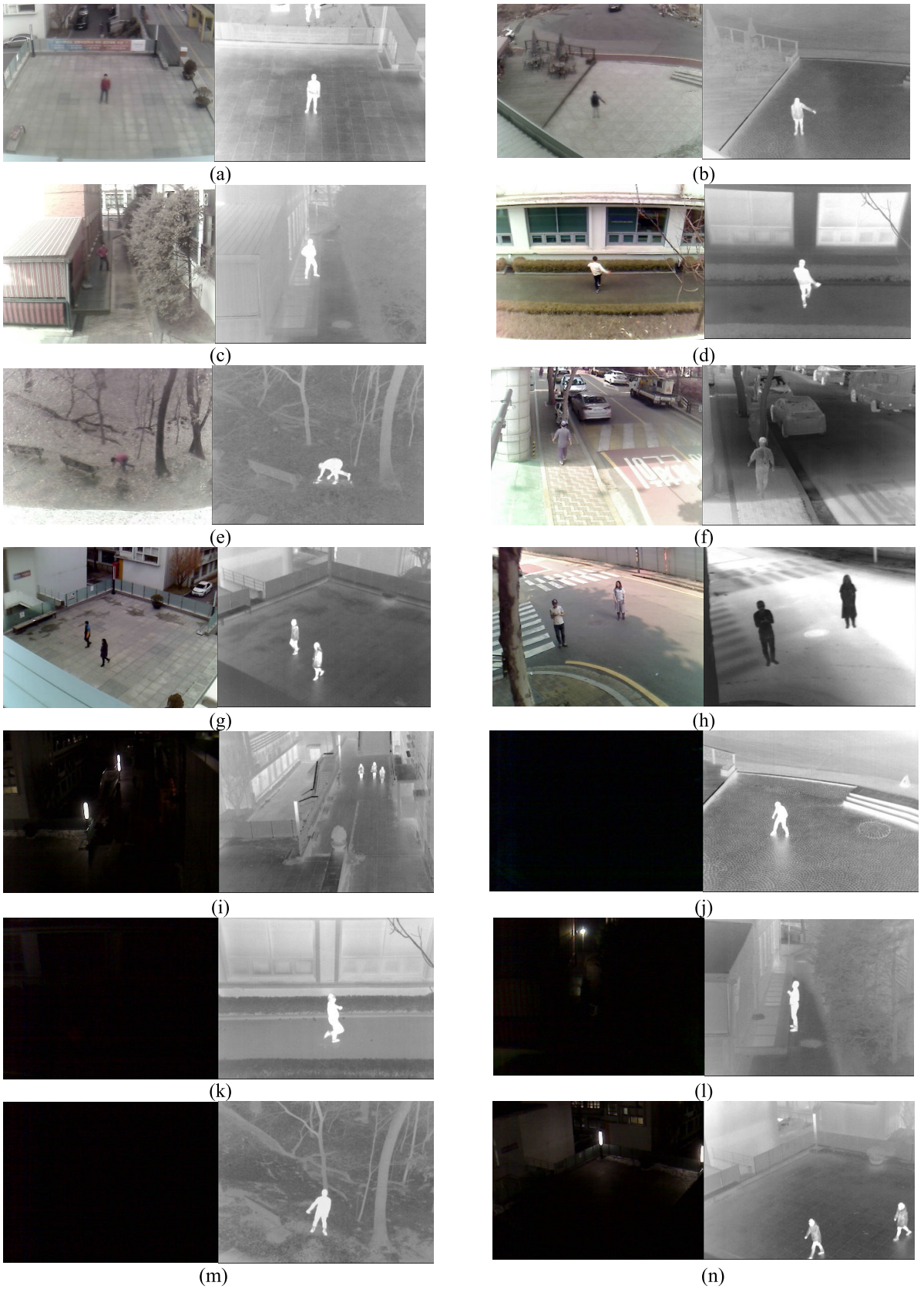

FIGURE 8. Examples of images acquired from each dataset. (a) ~ (v) show datasets I ~ XXII, respectively.

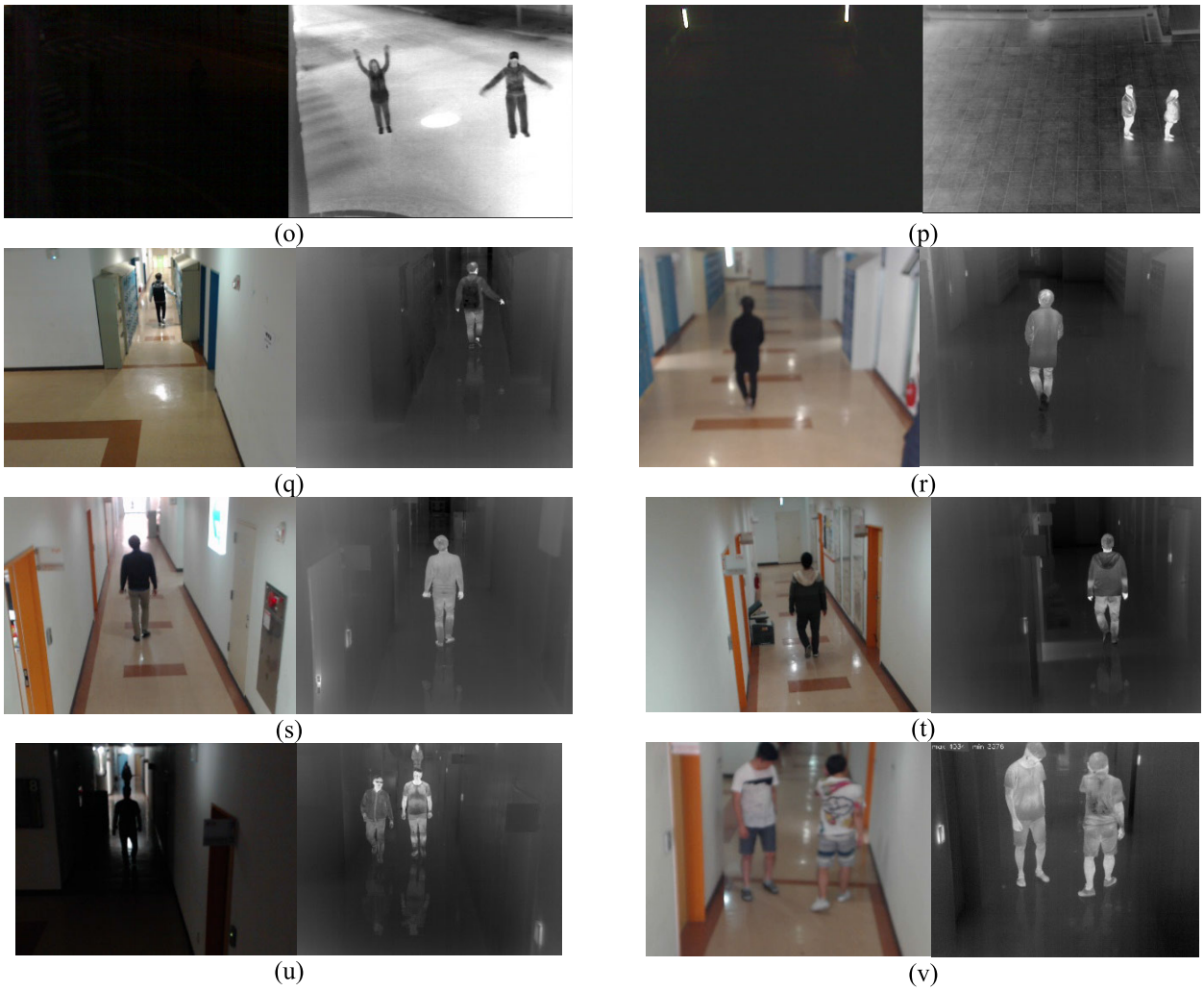

FIGURE 8. (*Continued:*) Examples of images acquired from each dataset. (a) ∼ (v) show datasets I ∼ XXII, respectively.

Mask R-CNN used ground truth data for the input due to their characteristics, and as shown in Figure 13, training was performed using a paired dataset of the input and ground truth. In contrast, CycleGAN uses unpaired reference data for training rather than ground truth data for input [33], and the unpaired dataset was used as shown in Figure 13.

D. TESTING OF THE PROPOSED METHOD

1) TESTING RESULTS OF THERMAL REFLECTION DETECTION

In this section, we present testing results for the detection of thermal reflections using SegNet and Mask R-CNN. As shown in Figure 13, CycleGAN and PLN are not the approaches for the detection of thermal reflections in the input, but directly generate images with thermal reflections that are eliminated from the input image. When comparing the detection accuracy, we measured the following five types of accuracies.

Accuracy (ACC) is the percentage of correctly classified pixels for each class as shown in Equation (5) [38]. Here, #TP, #TN, #FP, and #FN represent the number of true positive data, true negative data, false-positive data, and false-negative data, respectively. The positive and negative data represent the pixels of the thermal reflection and the background, respectively. TP represents the data that are positive and correctly classified as positive data whereas TN means data that are negative and correctly classified as negative data. FP represents data that are negative but incorrectly classified as positive data, whereas FN represents data that are positive, but incorrectly classified as negative data.

$$\text{Accuracy (ACC)} = \frac{\#\text{TP} + \#\text{TN}}{\#\text{TP} + \#\text{TN} + \#\text{FP} + \#\text{FN}} \tag{5}$$

The global accuracy (GlobalACC) is defined as the ratio of correctly classified pixels to the total number of pixels.

The F1 score is calculated based on precision and recall as shown in Equation (6) [39]. In this case, precision is calculated as #TP/(#TP+#FP), whereas recall is calculated as #TP/(#TP+#FN).

$$\text{F1} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{6}$$

For a class, the intersection over union (IoU) [40] is the ratio of the correctly classified pixels to the total number of ground

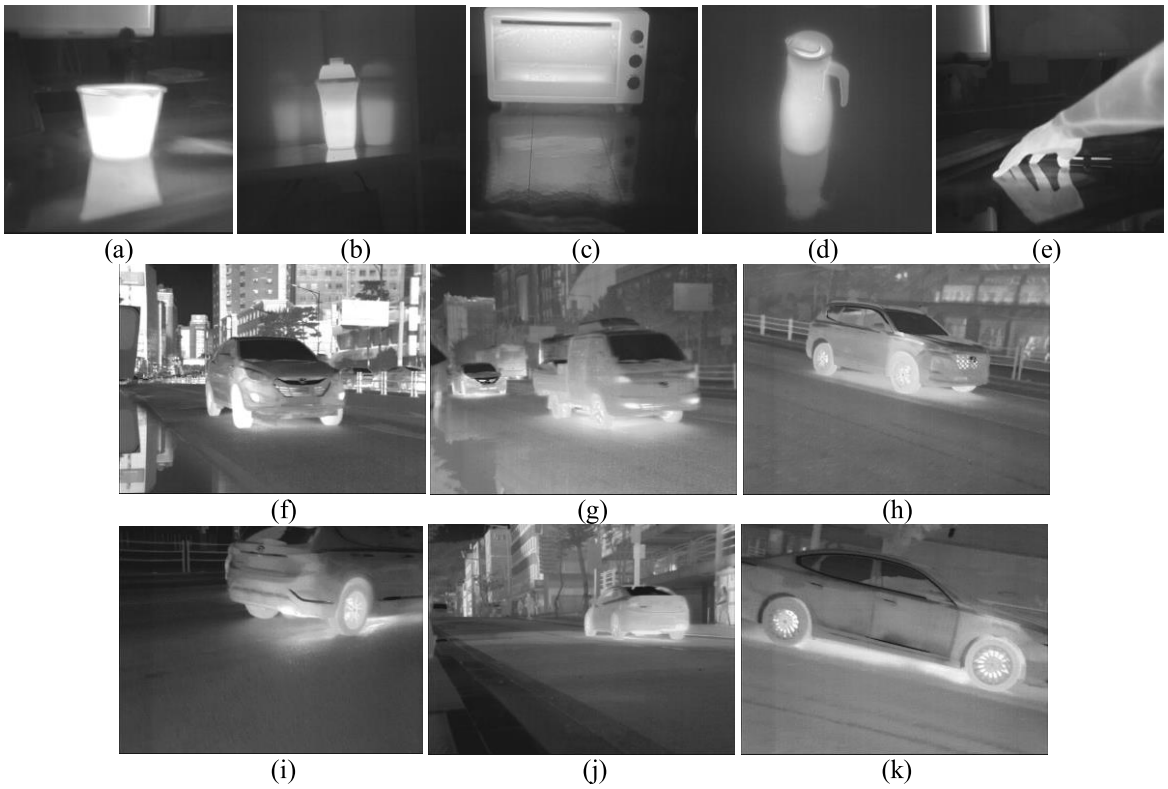

FIGURE 9. Examples of images from each dataset. (a) ∼ (v) show datasets I ∼ XI, respectively.

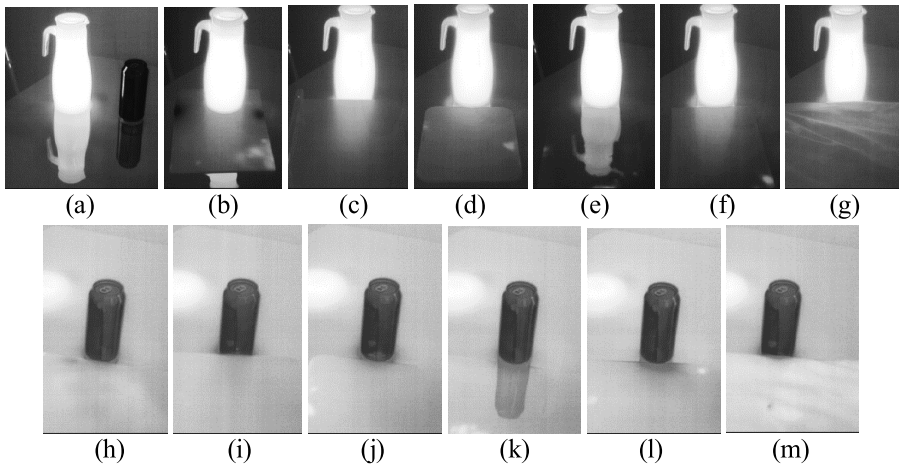

FIGURE 10. Example of thermal reflection using a glass pot with hot water and a metal can with cold water. (a) A pot and a can on glass; (b) a pot on a wooden surface; (c) a pot on a polyvinyl chloride (PVC) surface; (d) a pot on a rubber composite pad; (e) a pot on a plastic sheet; (f) a pot on a paper sheet; (g) a pot on a fabric; (h) a can on a wooden surface; (i) a can on a PVC surface; (j) a can on a rubber composite pad; (k) a can on a plastic sheet; (l) a can on a paper sheet; (m) a can on a fabric.

truth and predicted pixels in that class. The IoU score is also known as the Jaccard similarity, and it can be calculated with two sets X and Y as shown in Equation (7). In this case, X is the ground truth pixel of the thermal reflection whereas Y is the detected pixel of thermal reflection.

$$\text{Jaccard}\,(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} = \frac{TP}{TP + FP + FN} \tag{7}$$

As shown in Table 9, the method used in this study, Mask R-CNN, showed a higher detection accuracy compared to SegNet for all five metrics.

In Figure 14, the results of the SegNet and Mask R-CNN corresponding to the original images are compared. As shown in this figure, it is evident that the detection accuracy of the Mask R-CNN is higher than that of SegNet. In the training of SegNet, the loss is decreased until epoch 10 and does

FIGURE 11. Example of thermal reflections. (a) A car on an asphalt road; (b) a person on an asphalt road; (c) a person on a tile floor; (d) a person on the ground with leaves; (e) an oven on a PVC floor.

not decrease anymore. Therefore, the resulting images for the tests with the model obtained at epoch 10 are compared in Figure 14 (c).

2) TESTING RESULTS FOR THE REMOVAL OF THERMAL REFLECTION

In this section, the testing results of the proposed thermal reflection removal method and other state-of-the-art methods are presented. Four methods based on CycleGAN [33], PLN [34], Mask R-CNN + CycleGAN [33], and SegNet [35] were compared to the proposed method in terms of accuracy. In the Mask R-CNN + CycleGAN methods, the region corresponding to the thermal reflection detected by the Mask R-CNN was filled with the corresponding pixel value in the resulting image obtained by CycleGAN. In the SegNet-based removal method, the thermal reflection region detected by SegNet was filled using the thermal reflection removal method of Section IV.C that is proposed in this study. Accuracy measurements were performed using the following three metrics based on the similarity between the original ground truth image (Im(i,j)) with manual removal of thermal reflection and the (Res(i,j)) with automatic removal of thermal reflection using each algorithm.

The signal-to-noise ratio (SNR) [41] and the peak signal-to-noise ratio (PSNR) [42] are mathematical measurements of image quality based on the pixel difference (mean squared error (MSE)) between two images as represented by Equations (8)–(10). The structural similarity (SSIM) index [43] is a method that predicts the perceived quality of images,

$$\mathrm{MSE} = \frac{\left(\sqrt{\sum_{j=1}^{M}\sum_{i=1}^{N}(\mathrm{Im}(i,j)-\mathrm{Res}(i,\mathrm{j}))^2}\right)^2}{MN} \tag{8}$$

$$\mathrm{SNR} = 10log_{10}\left(\frac{\left(\frac{\sum_{j=1}^{M}\sum_{i=1}^{N}(\mathrm{Im}(i,j))^2}{MN}\right)}{MSE}\right) \tag{9}$$

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

where M and N represent the image width and height, respectively. Table 10 shows the accuracies associated with these three metrics. As previously described, given that these metrics are based on similarity concepts, the larger the value, the higher the accuracy. As shown in Table 10, in all cases, the accuracy of thermal reflection removal of the proposed method is the highest.

TABLE 9. Comparison of detection results by SegNet and Mask R-CNN methods.

Metrics	SegNet [35]	Mask R-CNN
Recall	0.96	0.92
Precision	0.40	0.98
GlobalACC	0.93	0.99
F1	0.56	0.95
Jaccard	0.39	0.91

Figure 15 shows the result for thermal reflection removal using the proposed method and the state-of-the-art methods. Figures 15 (a) and (b) show the original input image with thermal reflection, and the ground truth image with the reflection removed manually, respectively. The resulting images obtained for the proposed method are shown in Figure 15 (c), and the resulting images for the SegNet, CycleGAN, PLN, and Mask R-CNN + CycleGAN are shown in Figures 15 (d)–(g), respectively. When CycleGAN and PLN were used as shown in Figures 15 (e) and (f), the pixel values of the resultant images were generally lower than those of the original input image, and the ground truth image, indicating low accuracy as shown in Table 10. In the case of Mask R-CNN + CycleGAN as shown in Figure 15 (g), the thermal reflection area was readily identified, however, as the pixel value was filled with the value from the CycleGAN conversion image, the reflection area was darker than that of ground truth image. In addition, when SegNet is used as shown in Figure 15 (d), the thermal reflection area is generally detected to be wider than that of the proposed method in Figure 15 (c), and a portion of the human body is also incorrectly detected as a reflection, resulting in inaccuracies in the reflection removal process.

3) TESTING RESULTS WITH AN OPEN DATABASE

To confirm that the proposed algorithm is applicable to various environments, additional experiments were performed using the existing thermal image open database. Although there are many different thermal image open databases available [2], [44]–[54], there are few open databases with thermal reflections in thermal imaging. Therefore, in this study, comparative experiments were performed using a thermal soccer dataset [44], an open database with thermal reflection, as shown in Figure 16. As shown

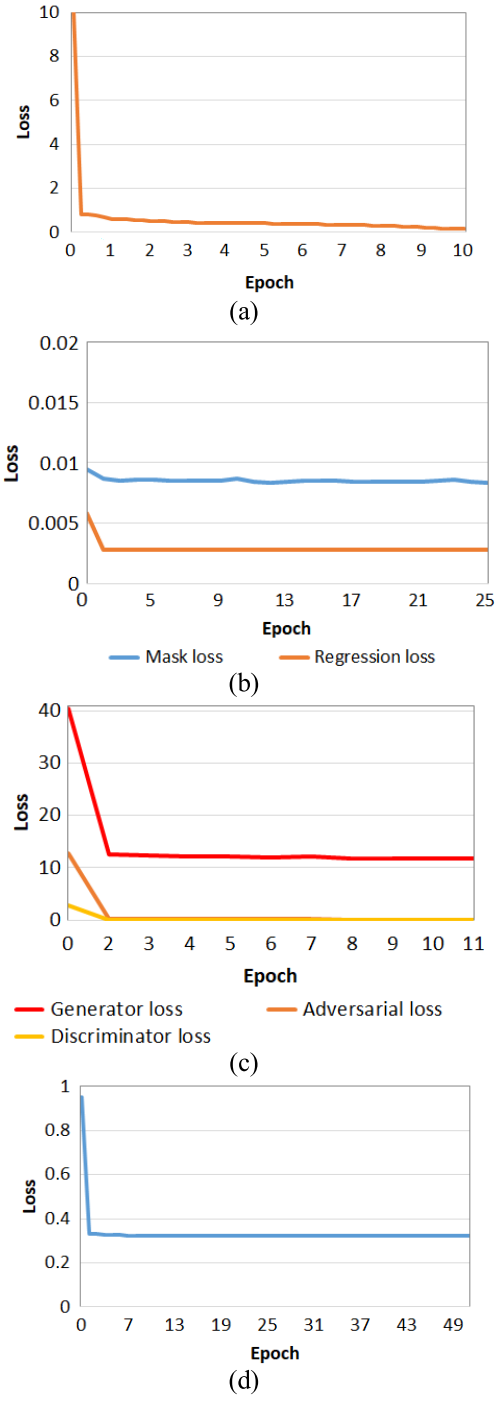

FIGURE 12. Examples of training loss curves. Loss curves for (a) SegNet, (b) Mask R-CNN, (c) CycleGAN, and (d) PLN.

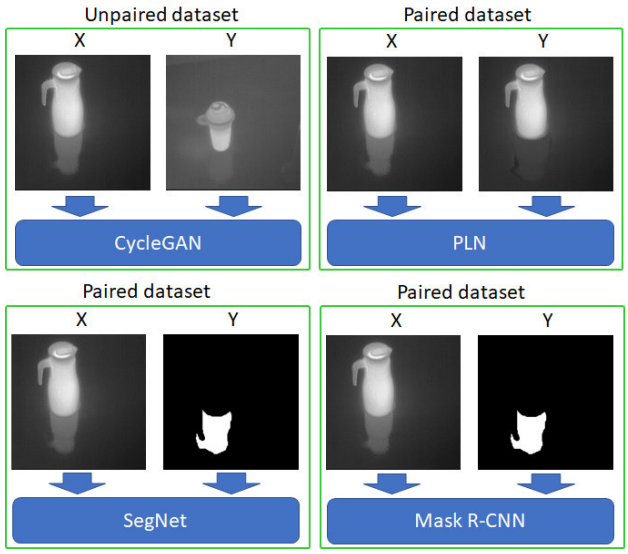

FIGURE 13. Examples of training methods using paired and unpaired datasets.

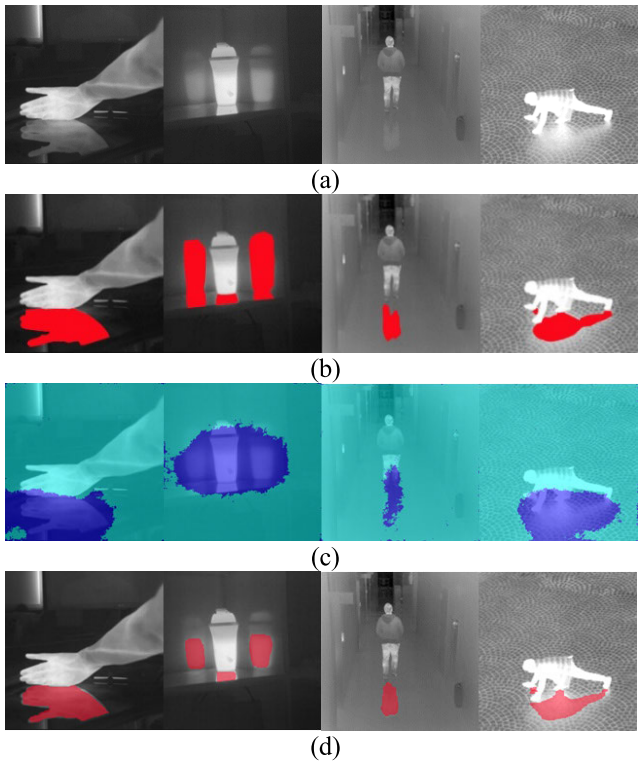

FIGURE 14. Example of detection results by SegNet and Mask R-CNN methods. (a) Original images; (b) ground truth images; (c) results obtained for SegNet;(d) results obtained for Mask-RCNN.

in Table 11, it is evident that in all cases, the accuracy of thermal reflection removal in the proposed method is higher than that of state-of-the-art methods. Figures 17 (a)–(c) show the original input image with thermal reflection, the ground truth image with the reflection removed manually, and the resulting image in which the thermal reflection is removed using the proposed method, respectively. As shown in Figure 17 (c), the thermal reflection is effectively removed using the proposed method compared to the ground truth image.

4) COMPARISONS OF PROCESSING SPEED OF THE PROPOSED METHOD AND THE STATE-OF-THE-ART METHODS

In the next experiment, we compared the processing speed between the proposed method and state-of-the-art methods

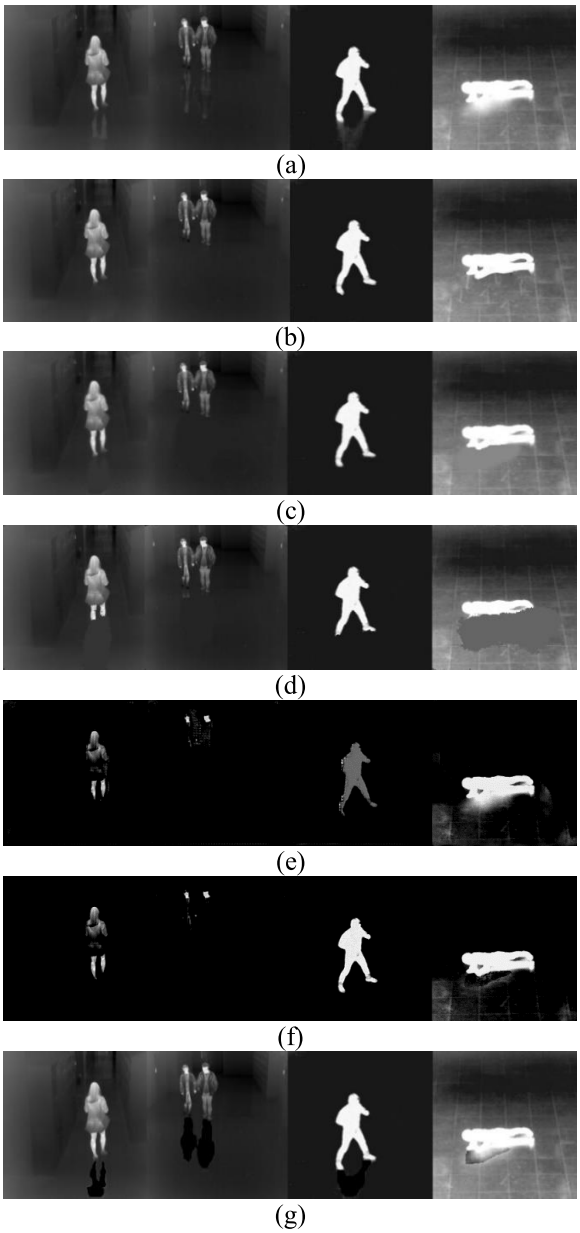

FIGURE 15. Examples of results obtained for the removal of thermal reflection.(a) Original images; (b) ground truth images; (c) results for the proposed method, (d) SegNet-based removal method, (e) CycleGAN, (f)PLN, and (g) Mask R-CNN + CycleGAN.

TABLE 10. Comparison of the accuracies of the proposed method to state-of-the-art methods.

Method	PSNR	SNR	SSIM
PLN [34]	11.40	2.31	0.092
CycleGAN [33]	11.42	2.34	0.096
SegNet-based removal [35]	24.71	15.63	0.939
Mask R-CNN + CycleGAN [33]	24.45	15.36	0.949
Proposed method	**31.43**	**22.35**	**0.973**

for one input image. The experiment was performed on a desktop computer described in Section V.A. As shown in Table 12, the proposed method required a processing time of 101.7 ms and was faster than the state-of-the-art methods.

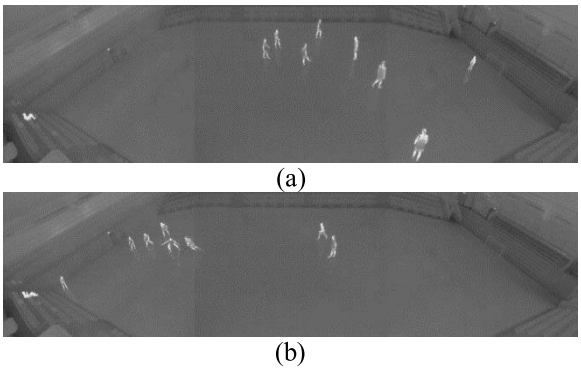

FIGURE 16. Sample images from the thermal soccer dataset.

TABLE 11. Comparison of the accuracies of the proposed method to state-of-the-art methods for the thermal soccer dataset.

Method	PSNR	SNR	SSIM
PLN [34]	14.34	1.99	0.055
CycleGAN [33]	14.03	1.67	0.052
SegNet-based removal [35]	28.18	15.82	0.960
Mask R-CNN + CycleGAN [33]	26.57	14.22	0.959
Proposed method	**36.58**	**24.23**	**0.983**

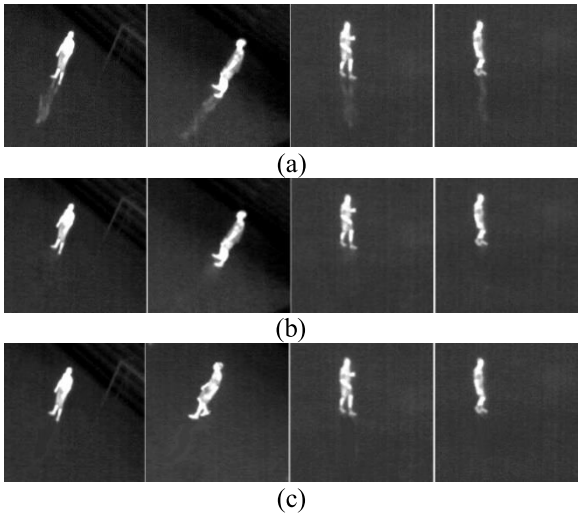

FIGURE 17. Examples of resulting images for the thermal soccer dataset. (a) Original images; (b) ground truth images; (c) results obtained using the proposed method.

TABLE 12. Comparison of processing speed for the proposed method and state-of-the art methods (unit: ms).

Method	Processing time
SegNet-based removal [35]	111.7
Mask R-CNN + CycleGAN [33]	107.9
Proposed method	**101.7**

VI. CONCLUSION

In this study, we proposed a method to detect thermal reflections produced by several objects, human subjects, and vehicles in thermal images using Mask R-CNN, and their subsequent removal during post-processing. In the experiments, we compared the performance of this approach to

several state-of-the-art methods such as SegNet, CycleGAN, PLN, and Mask R-CNN + CycleGAN. The results obtained using self-constructed DTh-DB and DI&V-DB databases, in addition to the thermal soccer open dataset showed that the proposed method showed a higher accuracy for thermal reflection detection and removal compared to the state-of-the art methods.

In future research, we will first detect the approximate ROI of thermal reflection for the entire input image, and then study the method of detecting and removing this artifact within the ROI more accurately. In addition, we plan to conduct research using methods that can detect all the objects, thermal reflections, and halo effects in a thermal image, followed by simultaneous removal of thermal reflections and halo effects.

REFERENCES

[1] L. St-Laurent, D. Prévost, and X. Maldague, "Thermal imaging for enhanced foreground-background segmentation," in *Proc. Int. Conf. Quant. Infr. Thermogr.*, Padova, Italy, Jun. 2006, pp. 1–10.
[2] J. W. Davis and V. Sharma, "Background-subtraction using contour-based fusion of thermal and visible imagery," *Comput. Vis. Image Underst.*, vol. 106, pp. 162–182, May/Jun. 2007.
[3] W. K. Wong, H. L. Lim, C. K. Loo, and W. S. Lim, "Home alone faint detection surveillance system using thermal camera," in *Proc. Int. Conf. Comput. Res. Develop.*, Kuala Lumpur, Malaysia, May 2010, pp. 747–751.
[4] C. Beyan, "Object tracking for surveillance applications using thermal and visible band video data fusion," M.S. thesis, Dept. Inf. Syst., Middle East Tech. Univ., Ankara, Turkey, Dec. 2010.
[5] P. Kumar, A. Mittal, and P. Kumar, "Fusion of thermal infrared and visible spectrum video for robust surveillance," in *Proc. Int. Conf. Comput. Vis., Graph. Image Process.*, Madurai, India, Dec. 2006, pp. 528–539.
[6] P. Kumar, A. Mittal, and P. Kumar, "Study of robust and intelligent surveillance in visible and multi-modal framework," *Informatica*, vol. 31, no. 4, pp. 447–461, 2007.
[7] D. Gangodkar, P. Kumar, and A. Mittal, "Segmentation of moving objects in visible and thermal videos," in *Proc. Int. Conf. Comput. Commun. Inform.*, Coimbatore, India, Jan. 2012, pp. 1–5.
[8] J. W. Davis and V. Sharma, "Robust detection of people in thermal imagery," in *Proc. 17th Int. Conf. Pattern Recognit.*, Cambridge, U.K., Aug. 2004, pp. 713–716.
[9] J. H. Lee, J.-S. Choi, E. S. Jeon, Y. G. Kim, T. T. Le, K. Y. Shin, H. C. Lee, and K. R. Park, "Robust pedestrian detection by combining visible and thermal infrared cameras," *Sensors*, vol. 15, pp. 10580–10615, May 2015.
[10] E. S. Jeon, J.-S. Choi, J. H. Lee, K. Y. Shin, Y. G. Kim, T. T. Le, and K. R. Park, "Human detection based on the generation of a background image by using a far-infrared light camera," *Sensors*, vol. 15, no. 3, pp. 6763–6788, 2015.
[11] E. S. Jeon, J. H. Kim, H. G. Hong, G. Batchuluun, and K. R. Park, "Human detection based on the generation of a background image and fuzzy system by using a thermal camera," *Sensors*, vol. 16, no. 4, p. 453, 2016.
[12] G. Batchuluun, H. S. Yoon, J. K. Kang, and K. R. Park, "Gait-based human identification by combining shallow convolutional neural network-stacked long short-term memory and deep convolutional neural network," *IEEE Access*, vol. 6, pp. 63164–63186, 2018.
[13] G. Batchuluun, R. A. Naqvi, W. Kim, and K. R. Park, "Body-movement-based human identification using convolutional neural network," *Expert Syst. Appl.*, vol. 101, pp. 56–77, Jul. 2018.
[14] G. Batchuluun, J. H. Kim, H. G. Hong, J. K. Kang, and K. R. Park, "Fuzzy system based human behavior recognition by combining behavior prediction and recognition," *Expert Syst. Appl.*, vol. 81, pp. 108–133, Sep. 2017.
[15] G. Batchuluun, Y. G. Kim, J. H. Kim, H. G. Hong, and K. R. Park, "Robust behavior recognition in intelligent surveillance environments," *Sensors*, vol. 16, no. 7, p. 1010, 2016.
[16] H. Eum, J. Lee, C. Yoon, and M. Park, "Human action recognition for night vision using temporal templates with infrared thermal camera," in *Proc. Int. Conf. Ubiquitous Robots Ambient Intell.*, Jeju Island, South Korea, Oct./Nov. 2013, pp. 617–621.
[17] R. Gade and T. B. Moeslund, "Thermal cameras and applications: A survey," *Mach. Vis. Appl.*, vol. 25, no. 1, pp. 245–262, 2014.
[18] FLIR Systems. (2019). *Uncooled Detectors for Thermal Imaging Cameras*. [Online]. Available: http://www.flirmedia.com/MMC/CVS/Appl_Stories/AS_0015_EN.pdf
[19] S. Henke, D. Karstädt, K.-P. Möllmann, F. Pinno, and M. Vollmer, "Identification and suppression of thermal reflections in infrared thermal imaging," in *Proc. InfraMation*, Las Vegas, CA, USA, Oct. 2004, pp. 287–298.
[20] Digital Media Lab. (2019). *Mask R-CNN Model and Generated Data With Dongguk Thermal Image Database (DTh-DB) and Dongguk Items & Vehicles Database (DI&V-DB)*. [Online]. Available: http://dm.dgu.edu/link.html
[21] FLIR Systems. (2019). *FLIR TauTM 2*. [Online]. Available: https://www.flir.com/products/tau-2/
[22] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, Venice, Italy, Oct. 2017, pp. 2999–3007.
[23] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," Mar. 2017, *arXiv:1703.06870*. [Online]. Available: https://arxiv.org/abs/1703.06870
[24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.
[25] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Honolulu, HI, USA, Jul. 2017, pp. 936–944.
[26] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," Mar. 2015, *arXiv:1411.4038v2*. [Online]. Available: https://arxiv.org/abs/1411.4038
[27] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," Jan. 2016, *arXiv:1506.01497v3*. [Online]. Available: https://arxiv.org/abs/1506.01497
[28] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2010.
[29] K. J. Devlin, *Fundamentals of Contemporary Set Theory*. New York, NY, USA: Springer-Verlag, 1979.
[30] NVIDIA Corporation. (2019). *NVIDIA Titan X*. [Online]. Available: https://www.nvidia.com/en-us/geforce/products/10series/titan-x-pascal/
[31] Keras. (2019). *Keras: The Python Deep Learning Library*. [Online]. Available: https://keras.io/
[32] OpenCV. (2019). *OpenCV: Open Source Computer Vision*. [Online]. Available: http://opencv.org/.
[33] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," Mar. 2017, *arXiv:1703.10593*. [Online]. Available: https://arxiv.org/abs/1703.10593
[34] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," Mar. 2016, *arXiv:1603.08155*. [Online]. Available: https://arxiv.org/abs/1603.08155
[35] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.
[36] E. L. Lehmann and G. Casella, *Theory of Point Estimation*. New York, NY, USA: Springer-Verlag, 1998.
[37] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," Dec. 2014, *arXiv:1412.6980*. [Online]. Available: https://arxiv.org/abs/1412.6980
[38] Wikipedia Foundation. (2019). *Precision and Recall*. [Online]. Available: https://en.wikipedia.org/wiki/Precision_and_recall
[39] L. Derczynski, "Complementarity, f-score, and NLP evaluation," in *Proc. Int. Conf. Lang. Resour. Eval.*, Portorož, Slovenia, May 2016, pp. 261–266.
[40] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Boston, MA, USA: Addison-Wesley, 2005.
[41] T. Stathaki, *Image Fusion: Algorithms and Applications*. Cambridge, MA, USA: Academic, 2008.
[42] D. Salomon, *Data Compression: The Complete Reference*, 4th ed. New York, NY, USA: Springer-Verlag, 2006.
[43] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[44] Kaggle. (2019). *Thermal Soccer Dataset*. [Online]. Available: https://www.kaggle.com/aalborguniversity/thermal-soccer-dataset#Public%20dataset.zip
[45] (2019). *OSU Thermal Pedestrian Database*. [Online]. Available: http://vcipl-okstate.org/pbvs/bench/Data/01/download.html

[46] (2019). *IRIS Thermal/Visible Face Database*. [Online]. Available: http://vcipl-okstate.org/pbvs/bench/Data/02/download.html
[47] (2019). *Terravic Facial Infrared Database*. [Online]. Available: http://vcipl-okstate.org/pbvs/bench/Data/04/download.html
[48] (2019). *Terravic Motion Infrared Database*. [Online]. Available: http://vcipl-okstate.org/pbvs/bench/Data/05/download.html
[49] (2019). *Terravic Weapon Infrared Database*. [Online]. Available: http://vcipl-okstate.org/pbvs/bench/Data/06/download.html
[50] A. Akula, N. Khanna, R. Ghosh, S. Kumar, A. Das, and H. K. Sardana, "Adaptive contour-based statistical background subtraction method for moving target detection in infrared video sequences," *Infr. Phys. Technol.*, vol. 63, pp. 103–109, Mar. 2014.
[51] G.-A. Bilodeau, A. Torabi, P.-L. St-Charles, and D. Riahi, "Thermal–visible registration of human silhouettes: A similarity measure performance evaluation," *Infr. Phys. Technol.*, vol. 64, pp. 79–86, May 2014.
[52] Z. Wu, N. Fuller, D. Theriault, and M. Betke, "A thermal infrared video benchmark for visual analysis," in *Proc. 10th IEEE Workshop Perception Beyond Visible Spectr.*, Columbus, OH, USA, Jun. 2014, pp. 1–8.
[53] M. M. Zhang, J. Choi, K. Daniilidis, M. T. Wolf, and C. Kanan, "VAIS: A dataset for recognizing maritime imagery in the visible and infrared spectrums," in *Proc. 11th IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Boston, MA, USA, Jun. 2015, pp. 10–16.
[54] S. Brahmbhatt, C. Ham, C. C. Kemp, and J. Hays, "ContactDB: Analyzing and predicting grasp contact via thermal imaging," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Long Beach, CA, USA, Jun. 2019, pp. 8709–8719.

GANBAYAR BATCHULUUN received the B.S. degree in electronic engineering from Huree University, Ulaanbaatar, Mongolia, in 2011, the M.S. degree in electronic engineering from Pai Chai University, Daejeon, South Korea, in 2014, and the Ph.D. degree in electronics and electrical engineering from Dongguk University, Seoul, South Korea, in 2019. He has been a Professor with the Division of Electronics and Electrical Engineering, Dongguk University, since March 2019. His research interests include biometrics and pattern recognition. He designed the entire system and wrote the original draft of this article.

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 joint master's and Ph.D. degrees in electronics and electrical engineering with Dongguk University. His research interests include biometrics and pattern recognition. He helped experiments and analysis.

DAT TIEN NGUYEN received the B.S. degree in electronics and telecommunications from the Hanoi University of Science and Technology (HUST), Hanoi, Vietnam, in 2009, and the Ph.D. degree in electronics and electrical engineering from Dongguk University, in 2015. He has been a Professor with the Division of Electronics and Electrical Engineering, Dongguk University, since March 2015. His research interests include image processing, biometrics, and deep learning. He helped experiments and analysis.

TUYEN DANH PHAM received the B.S. degree in electronics and telecommunications from the Hanoi University of Science and Technology (HUST), Hanoi, Vietnam, in 2010, and the M.S. and Ph.D. degrees in electronics and electrical engineering from Dongguk University, in 2013 and 2017, respectively. He has been a Professor with the Division of Electronics and Electrical Engineering, Dongguk University, since March 2017. His research interests include image processing, biometrics, and deep learning. He supervised this research and revised this article.

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 helped experiments and analysis.