IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 4, 2020, accepted October 14, 2020, date of publication October 19, 2020, date of current version October 28, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3032030

Damage Detection With an Ultrasound Array and Deep Convolutional Neural Network Fusion

DONGGEUN KIM, SAN KIM, SIHEON JEONG, JI-WAN HAM, SEHO SON, AND KI-YONG OH

Department of Energy System Engineering, Chung-Ang University, Seoul 06974, South Korea

Corresponding author: Ki-Yong Oh (kiyongoh@cau.ac.kr)

This work was supported in part by the Korea Electric Power Corporation under Grant R20XO02-5, and in part by the Chung-Ang University Graduate Research Scholarship, in 2019.

ABSTRACT Diagnostic methods for power transmission facilities are important for energy security because the growth of defects in power facilities increases the risk of blackouts in an entire power grid. However, damage in power transmission facilities is difficult to detect because cracks or defects are minuscule and are challenging to determine. One interesting phenomenon caused by damage in power transmission facilities is ultrasound emissions on a damaged surface. However, measuring ultrasound emissions to detect defects is limited by the severity of the surrounding noise. To overcome this limitation, this study proposes a new method for damage detection by fusing ultrasound measurements with recorded optical images. The proposed method consists of two phases. The first phase preprocesses ultrasound measurements for ultrasound feature extraction. This phase aims to detect the location of ultrasound emissions by analyzing ultrasound characteristics including the intensity and density. The second phase detects and classifies a damaged object with optical images recorded using a deep convolutional neural network. This phase not only discards the noise from the ultrasound measurements but also classifies a damaged system among many components in power transmission facilities. The experiments validate the effectiveness of the proposed method using ultrasound measurements and recorded images and finally suggest scenarios for potential applications.

INDEX TERMS Health management, diagnostics, deep learning, damage detection, feature extraction, ultrasound camera.

I. INTRODUCTION

The reliability of power transmission facilities is important for securing the stability of power resources and the safety of power facilities. Specifically, the failure of transmission facilities causes serious problems, including blackouts [1] that threaten the livelihood of society. Diagnostics and prognostics must be considered for transmission facilities using novel sensors and smart inspection systems to ensure the safety and reliability of power transmission facilities.

Power transmission facilities suffer from a variety of failure modes [2]. Problems, such as bullet holes, strand failures, or insulator cracks, can damage the surfaces of transmission facilities. Vibrations in transmission lines cause wear, spacer fractures, and breakaways between transmission facilities. Hence, most damage originates from transmission lines, spacers, stock bridge (SB) dampers, insulators, and marker balls, which comprised 159 out of 205 cases of defects (78%) in 2016 in Korea [3], [4]. The failure mode analysis suggests that inspecting major failure modes and their components is crucial for the intelligent operation and maintenance of power facilities. Therefore, several direct and indirect diagnostic methods have been studied to inspect damage in transmission facilities [2].

The direct diagnostic method requires that electricians climb transmission towers and walk along transmission lines using simple tools to visually inspect whether damage has occurred in the current [2]. This intuitive inspection method can detect many failure modes directly. However, the performance of this traditional inspection method depends on the skills and concentration of the electrician. If the electrician is distracted from inspection, this traditional inspection approach does not efficiently detect damage in power facilities. Moreover, this method exposes electricians to potential

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

accidents, such as being hurt from a tower or electrical short, due to the geographical and overhanging nature of power transmission facilities.

In these dangerous environments, a nondestructive inspection approach is useful for mitigating concerns regarding potential accidents [1]. These methods measure corona discharge [5], ultrasonic emissions [6], and heat generation [7] from a damaged system. With the technological advances in the fourth industrial revolution, these methods have attracted attention. Specifically, smart mobilities have enabled these sensors to closely approach power facilities to measure informative data. Mobile robots [8] and crewless aerial vehicles [9] have been vigorously developed by equipping these nondestructive sensors to inspect power facilities. This suggests that proactive maintenance with these novel sensors equipped with smart mobilities reduces maintenance cost, decreases the potential for accidents, and avoids the catastrophic failure of power facilities.

Each nondestructive inspection method has benefits and drawbacks. Specifically, an infrared camera measures heat generation on a damaged surface, which was the type of defect in 75 out of 205 (37%) cases in 2016 in Korea [4]. This method is quite useful for detecting defects in connection parts because the wear or damage in connection parts results in overheating. However, the exact temperature in outside environments is difficult to estimate using infrared thermal cameras because many of the parameters used to convert thermal energy into temperature depend on the surrounding environment [7]. Specifically, parameter tuning for an infrared thermal camera depends on the weather and physical surroundings; thus, recording thermal images requires significant effort. Consequently, infrared thermal cameras are primarily used for controlled environments, such as the inside of buildings or underground tunnels [10].

A corona camera can detect partial discharges, called corona discharges, on damaged surfaces [11]. This approach is effective because the type of defect in 88 of 205 cases (43%) in 2016 in Korea [4] emitted corona discharge that can be detected using a corona camera. Corona discharge does not naturally occur on Earth because the corona discharge from the sun is absorbed by the ozone layer, suggesting that measuring the corona discharge is an effective metric. However, the main drawback of this approach is that corona cameras are too expensive to produce, especially considering the wide distribution of power transmission lines.

Interestingly, corona discharge accompanies ultrasonic emissions, suggesting that an ultrasound array can be used to detect most damaged surfaces [12]. Moreover, an ultrasound array is a much cheaper method than using a corona camera; therefore, an ultrasound array is economically feasible. However, a current ultrasound array measures the surrounding noise, including reflection and diffraction noises. Relatively weak sources may also be misjudged as noise if more than one ultrasound source exists. Moreover, this sensor only measures ultrasound intensity to deduce whether information for an object emitting ultrasound is missing.

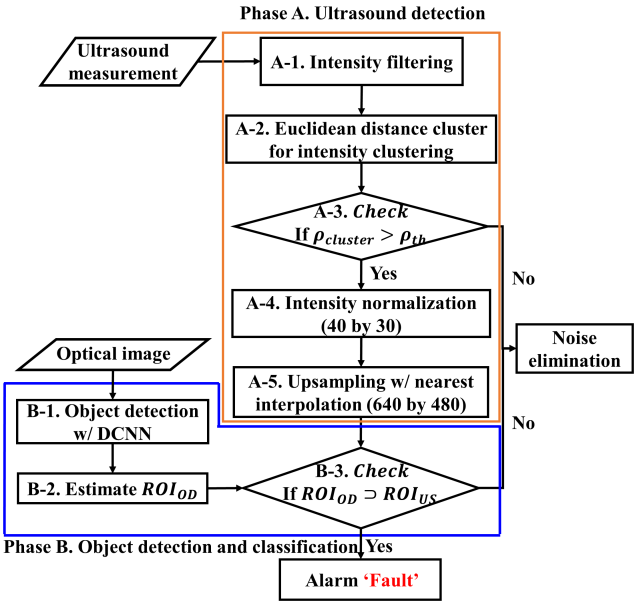

FIGURE 1. Flowchart for the proposed fault detection method using ultrasound measurements and optical image fusion.

To overcome these limitations, this study proposes a method for damage detection using ultrasound and optical images. The proposed method comprises two phases. The first phase preprocesses ultrasound measurements to eliminate noise using intensity and density filters. An intensity filter applies a statistical outlier detection method, and a density filter uses an unsupervised clustering method called Euclidean distance clustering to account for the circle-like shapes of ultrasound emissions. The second phase analyzes optical images through a deep convolutional neural network (DCNN) for object detection. Hence, this phase not only enhances the accuracy of distinguishing true ultrasound sources from noises but also classifies the damaged facilities. The experiments show that the proposed method detects ultrasounds in a variety of operational conditions even with the inherent characteristics of ultrasounds. Moreover, the proposed neural network effectively cognizes five power facilities with high accuracy. Finally, a potential scenario for damage detection is illustrated with the proposed method to suggest future applications using a crewless aerial vehicle.

II. DAMAGE DETECTION METHOD

The proposed damage detection method detects defects and classifies damaged systems by fusing information from ultrasound measurements and optical images. The proposed damage detection algorithm comprises two phases: the detection of ultrasound locations using ultrasound measurements (Phase A in Fig. 1) and object detection with optical images using a DCNN (Phase B in Fig. 1). The details for each phase are described in the following subsections.

A. ULTRASOUND DETECTION

This section describes the ultrasound detection method comprising five steps (Phase A in Fig. 1). This phase aims to eliminate noise from a) the surroundings of other sources,

b) the characteristics of a system, and c) the reflection and diffraction of ultrasound emissions in constrained environments. It is assumed that the intensity of these types of noise is relatively weaker than that of ultrasound emissions from the damaged surface of power transmission facilities.

In the first step (A-1 in Fig. 1), the measured intensity of ultrasound I_{US}, which is lower than the intensity threshold I_{th}, is filtered. Specifically, the ultrasound intensity I_{US} is measured as a two-dimensional (2D) vector map (40 by 30 pixels) from an ultrasound array. Each pixel contains information regarding the ultrasound intensity I_{US} calculated using the time difference of the arrival method [13] at an embedded digital-signal-processing board in an ultrasound array. In this step, the ultrasound intensity I_{US}, which is higher than the intensity threshold I_{th}, maintains the same value for each pixel, whereas the other cases are converted to the value of the intensity threshold I_{th} for each pixel. This study sets the intensity threshold I_{th} automatically by applying the three-sigma outlier detection method [14], where the measured intensity from faults is significantly stronger than that in normal conditions. This statistical filtering method automatically determines the intensity threshold I_{th}, implying that this method is easily applied to other future applications. Hence, this step effectively eliminates weak surrounding noise.

In the second step (A-2 in Fig. 1), the filtered intensity is clustered by addressing Euclidean distance clustering [15], which effectively calculates density data clustered in a circle shape, such as ultrasound emissions. The proposed method accounts for the characteristics (i.e., the strong intensity/high density) of ultrasound emissions from the damaged surface of power transmission facilities.

The third step (A-3 in Fig. 1) calculates the density of each cluster $\rho_{cluster}$ and then filters the clusters with the density threshold ρ_{th}, where the density threshold ρ_{th} is set to 0.5 times the maximum density ρ_{max} of each measurement because this is a problem of binary classification: true ultrasound vs. noise. Hence, the center point between 0 and 1 is an appropriate threshold. This threshold is also automatically determined from the maximum of each measurement. A cluster that is lower than the density threshold ρ_{th} is considered noise and is eliminated. The filtered clusters with a strong intensity/high density are defined as the region of interest for ultrasounds (ROI_{US}) which are candidates for ultrasounds emitted from a damaged system. To conduct this procedure, the Euclidean distance R is properly selected by considering the environment and the distance from the ultrasound array because ultrasound intensity depends on these factors.

The fourth step (A-4 in Fig. 1) transforms a 2D intensity vector map to a normalized 2D intensity map in grayscale as follows:

$$I_{nor} = \frac{255}{I_{max} - I_{th}} \times (I - I_{th}), \tag{1}$$

where I denotes the intensity of each pixel in each ROI_{US} and I_{nor} and I_{max} denote the normalized intensity in the range of 0 to 255 with a rounding off operation to make the result an integer and the maximum intensity in each ROI_{US}, respectively. An optical image is normally represented as red, green, and blue in integer form with 8 bits of data expressed for color, ranging from 0 to 255. To match the 2D ultrasound intensity map with this characteristic of optical images, a 2D intensity map is linearly normalized in the range of 0 to 255.

The final step upsamples the 2D intensity map using the nearest-neighbor interpolation method [16] to correct the resolution mismatch between the normalized ultrasound 2D intensity map (40 by 30 pixels) and the optical image (640 by 480 pixels). The nearest-neighbor interpolation method is introduced because its mean processing time is the fastest among interpolation methods.

B. OBJECT DETECTION AND CLASSIFICATION

This section describes Phase B of the proposed method comprising three steps. This phase increases the damage detection accuracy when distinguishing ultrasounds emitted from a damaged system with noisy ultrasounds. Moreover, object classification provides useful information to identify a damaged system in power transmission facilities.

The first step (B-1 in Fig. 1) executes object detection using an optimal image recorded through a DCNN because the DCNN outperforms other methods for object detection if sufficient training datasets are provided. The prediction through DCNN provides information for object detection with a confidence score, including the center of the object, the width of the object w_{obj}, and the height of the object h_{obj}.

The second step (B-2 in Fig. 1) calculates the region of interest for object detection ROI_{OD}, which requires a proper margin in terms of width and height to account for the different resolutions of ultrasound measurements and optical images. This is because one pixel in ultrasound measurement covers several pixels in optical images, as ultrasound measurements are upsampled at A-5 in Fig. 1. Hence, the mismatch between the location of an ultrasound emission and object detection should be considered, especially for a long-distance object. In consideration of these characteristics, ROI_{OD} is defined as follows:

$$ROI_{OD,*} = G_A \times G_{*,a} \times *_{obj} \tag{2}$$

where G_A and $G_{*,a}$ denote the distance gain and aspect ratio gain. The asterisk $*$ indicates the width w or height h of the object estimated from the previous step (i.e., w_{obj} or h_{obj}). The distance gain is calculated as follows:

$$G_A = \begin{cases} 1, & if\, A_{obj} > A_{th} \\ 1.2 - 0.2 \times \dfrac{A_{obj}}{A_{th}}, & else \end{cases} \tag{3}$$

where A_{obj} and A_{th} are the object area and area threshold, respectively. The object area is calculated by multiplying the width of an object w_{obj} and the height of an object h_{obj}. The threshold A_{th} is set to 5% of the overall area of an optical image. The distance gain G_A is addressed due to the resolution mismatch depending on the distance between the ultrasound array and the object of interest. The longer the distance between an ultrasound array and object of interest,

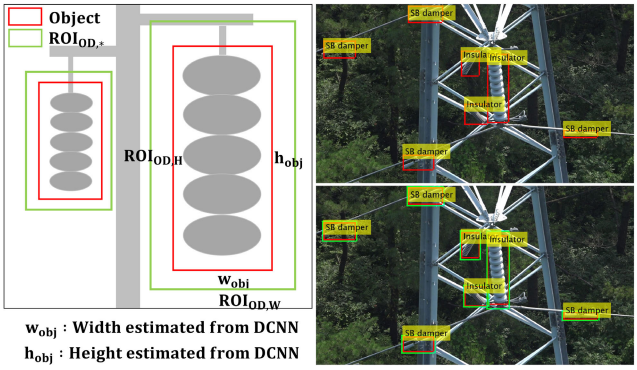

FIGURE 2. The region of interest for object detection $\mathbf{ROI_{OD}}$ with respect to the size and aspect ratio of an object.

the greater the distance mismatch. The aspect ratio gain $G_{*,a}$ is set as follows:

$$G_{*,a} = \begin{cases} 1.2 \mp 0.2 \times \gamma, & \text{if } \gamma \geq 1.2 \\ 1.2 \pm 0.2 \times \gamma, & \text{if } \gamma \leq 0.8 \\ 1, & \text{else} \end{cases} \quad (4)$$

where γ denotes the aspect ratio between the width of object w_{obj} and the height of object h_{obj} (i.e., w_{obj}/h_{obj}). This gain is addressed to maintain a proper margin for all objects with different aspect ratios. In Eq. (4), the upper sign in front of aspect ratio γ is used for the aspect ratio gain for width $G_{w,a}$, whereas the bottom sign in front of aspect ratio γ is used for the aspect ratio gain for height $G_{h,a}$. This study presents area threshold A_{th} and aspect ratio γ, considering the characteristics of power facilities. However, the main principle in determining A_{th} and γ is simple, suggesting that selecting these two thresholds is intuitive and easily implemented in other applications.

Fig. 2 depicts an example of executing the first and second steps in this phase. The first step detects objects, such as insulators or SB dampers (red boxes in Fig. 2), and provides the width of the object w_{obj}, and the height of the object h_{obj}. Then, the second step calculates $ROI_{OD,W}$ and $ROI_{OD,H}$ to compare ROI_{OD} to ROI_{US} in the following step.

In the last step (B-3 in Fig. 1), an upsampled 2D intensity map (from A-5 in Fig. 1) is compared to an object detection result (from B-2 in Fig. 1). If the center of ROI_{US} is in ROI_{OD}, a fault alarm is triggered with the information regarding the class and location of the damaged system. Hence, the proposed method detects a damaged system and informs its class.

III. DCNN FOR OBJECT DETECTION

The accuracy of object detection in the B-1 phase is important to accurately predict a defect in the proposed method, suggesting that selecting the appropriate DCNN determines the overall accuracy of the proposed method. Hence, this study compares the architectures of three state-of-the-art one-stage DCNNs, which simultaneously extract the location and class of objects with parallel neural network structures because the two-stage model limits real-time applications. Specifically, this study examines the single-shot multi-box detector (SSD),

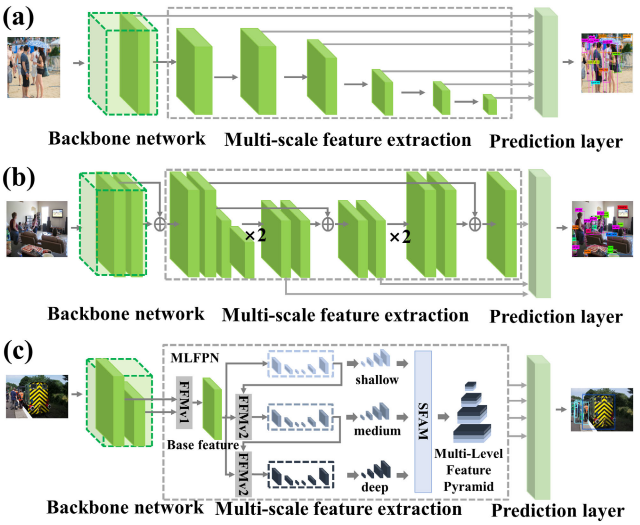

FIGURE 3. Architecture of deep convolutional neural networks (DCNNs) for object detection. (a) Single-shot multi-box detector (SSD). (b) You only look once version 3 (YOLOv3), and (c) M2Det.

you only look once (YOLOv3), and a multi-level feature pyramid network (MLFPN) called M2Det [17]–[19].

A. SINGLE-SHOT MULTI-BOX DETECTOR

The SSD was proposed in 2016 and has two key characteristics in comparison to the previous one-stage object detection models, including YOLOv1 [20]. One key characteristic is defining the centers, widths, and heights of the default bounding boxes called prior boxes, whereas only the centers of the prior boxes are defined in YOLOv1. The other characteristic is feature resampling. Feature resampling retains more meaningful information and is used to construct multi-scale feature maps. These characteristics improve the accuracy and calculation time.

Specifically, the SSD is a feed-forward convolutional network (Fig. 3(a)). An optical image is first inserted into a backbone network, which is the basis of the feature extraction layers. The SSD uses VGG-16, which is pretrained using the ImageNet Large Scale Visual Recognition Challenge dataset [21], among many backbone networks [22]–[25]. The features extracted through a backbone network pass through multi-scale convolutional layers to construct multi-scale feature maps, which is meaningful because of their respective field sizes. The location of an object, including the center, width, and height relative to the default box, is extracted, and the classification is simultaneously executed for each object location at each feature map layer. Data augmentation with selective sampling is also implemented for effective training. Hard negative mining is used to deal with class imbalance problems.

B. YOU ONLY LOOK ONCE (v3)

The YOLOv3 was proposed in 2018 as part of the YOLO series [18],[20],[26]. Fig. 3(b) illustrates the overall architecture of YOLOv3, which uses a custom network called the darknet as the backbone network. A skip connection between downsampling and upsampling is used to effectively extract

features containing semantic information for localization and classification. High-feature layers strengthen detailed information, but meaningful semantic information from an object can be lost. By concatenating the upsampled and previous features of the two-step layers on a channel basis, meaningful semantic information from low layers is conserved as features for high-feature layers. This neural network uses multi-scale training, data augmentation, and batch normalization to improve the performance of object detection.

C. THE M2Det NETWORK

The M2Det network includes multi-scale and multi-level characteristics [19]. Fig. 3(c) illustrates the overall architecture of M2Det. Specifically, M2Det extracts multi-level features by addressing an MLFPN, whereas the previous deep neural networks in object detection use single-level features that are extracted using only a serial structure with single-feature layers.

The MLFPN comprises a feature fusion module (FFM), a thin U-shaped module (TUM), and a scale-wise feature aggregation module (SFAM) to improve object detection performance. The FFM aims to avoid losing low-layer information by inserting lower features into a higher layer to extract higher features. The concatenated feature from the FFM is inserted into the TUMs. The TUMs consist of multi-level structures with multi-TUM layers. Each TUM is a multi-scale autoencoder-formatted convolutional layer to extract multi-scale features. The features in the encoding parts are concatenated to be featured in decoding parts to become distinct features. Hence, multi-level features from TUM layers are inserted into the SFAM to generate the final multi-scale feature maps.

Sparse features representing low scales are used to detect large objects, whereas dense features are for small objects. These feature maps are reweighted through the channel-wise attention submodule. This submodule executes global averaging pooling for each channel, and the resulting features are inserted into two fully connected layers. Finally, the features that pass through all these modules are inserted into a prediction layer to evaluate the object location and classes.

IV. EXPERIMENTS

This section describes experiments for ultrasound measurements and image data acquisition. Experiments were conducted to verify the detectability of ultrasound emissions using the proposed method in a variety of operational conditions. Image data acquisition presents procedures and methods to acquire image datasets using a crewless aerial vehicle. These image datasets were used for the training and validation of object detection with the DCNN.

A. EXPERIMENTS FOR ULTRASOUND DETECTION

The attenuation, reflection, and diffraction of ultrasounds are affected by the position, intensity, and surrounding environment of an ultrasound source, resulting in mirror or ghost ultrasounds [27]. Mirror or ghost ultrasound emissions interfere with detecting the true ultrasound in a damaged system.

TABLE 1. Experimental cases for ultrasound detection.

Case	Source	Experimental condition	Total measurements
1	1	Single strong ultrasound at different distances (5 m, 10 m, and 15 m)	15
2	2	Dual same-intensity ultrasounds at the same distance (5 m)	25
3	2	Dual different-intensity ultrasounds at the same distance (5 m)	25
4	2	Dual different-intensity ultrasounds at different distances (weak: 5 m, strong: 10 m)	25

The direction of ultrasound emissions is also important, considering the straightness of the ultrasound, because an ultrasound source might not directly face an ultrasound array in most cases. Hence, experiments were conducted to evaluate the detectability of the proposed method with respect to the distance between an ultrasound array and ultrasound source, emission angles of ultrasounds, single- or multiple- sources, and the intensity of the ultrasound source (Table 1).

The ultrasound array BATCAM 1.0 (SM Instruments, Republic of Korea), JETSON TX2 (NVIDIA, USA) and two ultrasound generators (Wildlife, USA) were used in the experiments (Fig. 4(a)). BATCAM 1.0 measures ultrasounds and provides a preprocessed 2D intensity vector map (40 by 30 pixels, 25 frame per second (FPS)) with a measurement angle of 90°. The signal bandpass of BATCAM 1.0 is set to the range of 38 to 42 kHz because the ultrasound frequency emitted from a damaged surface is around 40 kHz.

Four cases with different ultrasound sources were explored using a single source vs. multiple sources (Fig. 4(b)). Sources in the intensity range of 36 to 48 dB were classified as weak sources, and an intensity range of more than 100 dB was classified as strong. Each case was conducted for different distances and emission angles between the BATCAM 1.0 ultrasound array and the ultrasound generator. The ultrasound generator and BATCAM 1.0 were installed at the heights of 0.5 m and 0.3 m from the ground, and the measurement time was set to 4 s for all cases.

The first case (Case 1 in Table 1) was conducted with a single strong ultrasound source. The ultrasound generator was installed at distances of 5, 10, and 15 m from the BATCAM 1.0 with an emission angle in the range of $-60°$ to $60°$ with an angular resolution of 30°. The emission angle is defined as 0° when the ultrasound generator faces the BATCAM 1.0. The counterclockwise direction is positive, and the clockwise direction is negative. Hence, datasets of 15 were measured, replicating a situation in which a single, strong ultrasound was emitted from a damaged system at different distances without external interference.

The second case (Case 2 in Table 1) assumes that two small instances of damage occur in the transmission facilities. Two ultrasound sources, which emit the same weak ultrasound intensity, were installed to mimic this situation at a distance of 5 m with an emission angle in the range of $-60°$ to $60°$ with an angular resolution of 30°. The distance between the two

FIGURE 4. (a) Hardware configuration and (b) experimental conditions.

sources was set to 3 m. Hence, this case records datasets of 25 measurements to test the interference of two same-intensity sources to check the detectability of the proposed method in multiple source situations.

The third case (Case 3 in Table 1) assumed that two damaged surfaces emit different-intensity ultrasounds at the same distance of 5 m from BATCAM 1.0 because of the different degrees of damage. The surface to the left features minor damage and emits an ultrasound of weak intensity, whereas the surface on the right features major damage and emits an ultrasound of strong intensity. These experiments were also conducted with the same emission angle range and angular resolution for the two ultrasound generations so that datasets of 25 measurements were recorded.

The fourth case (Case 4 in Table 1) was conducted to locate two different-intensity ultrasounds at different distances. The weak ultrasound generator was located at a distance of 5 m, and the strong ultrasound generator was located at a distance of 10 m from the BATCAM 1.0 with the same emission angle range and angular resolution as the other cases; thus, datasets of 25 measurements were recorded. This case replicated the case in which two instances of damage occurred in different systems at different distances from an ultrasound array where a heavily damaged system was farther from an ultrasound array.

TABLE 2. Detailed information on image data measurements.

Date	Site name	# of TL inspected	# of image datasets
17.05.23-17.05.23	AY	5	1463
17.06.14-17.06.15	AH	5	2523
17.06.16-17.07.27	SS	17	9727
17.07.14-17.07.27	ST	6	1538

B. DATA ACQUISITION OF OPTICAL IMAGES

The acquisition of optical image datasets is important for object detection through DCNNs because a large number of high-quality image datasets increases the accuracy of object detection. A specialized crewless aerial vehicle was employed by the Korea Electric Power Corporation (KEPCO) Research Institute to record images of core components in transmission facilities. This crewless aerial vehicle was equipped with SPMY FDR-AX100 (SONY, Japan) and acquired optical image datasets with a resolution of 1920 by 1080 pixels. The crewless aerial vehicle flew via autopilot controlled by a KEPCO ground control station [9] in actual fields.

This method is effective because transmission facilities are located in mountainous environments, which are difficult to reach. The ground control station has a virtual safety wall from the transmission facilities to mitigate concerns regarding the interference of an inertial measurement unit sensor from magnetic fields energized by high-voltage transmission lines. Hence, the data acquisition of optical images proceeded at 12X magnification using SPMY FDR-AX100 equipped in the crewless aerial vehicle while maintaining a safe distance from transmission lines.

Four sites where 154, 345, and 765 kV transmission lines (TLs) are located were navigated with a crewless aerial vehicle to acquire image datasets at a variety of transmission facilities (Table 2). The total recorded image datasets for the transmission facilities is 15 251. Image datasets of 13 726 (90%) were used to train the DCNN, whereas the rest of the image datasets (1525, 10%) were used to validate the performance of the DCNN. All transmission facilities installed on the towers were recorded and classified into five categories, including a tower, insulator, SB damper, spacer, and marker ball (Fig. 5), because these are potentially damageable facilities.

V. RESULTS AND DISCUSSION

This section analyzes two experimental results. Section 5.1 analyzes the results of the source detectability in various situations, such as using single or multiple sources with a variety of distances, emission directions, and intensities. Section 5.2 compares the performance of the three DCNN models for object detection to select a proper DCNN with the Microsoft Common Objects in Context (MS COCO) dataset [28]. Then, a selected object detection model was trained using the training datasets for the transmission facilities, and

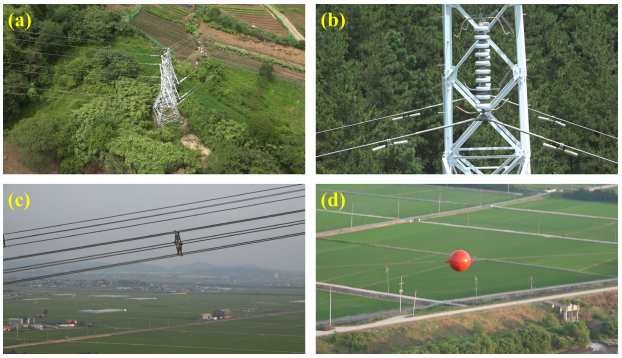

FIGURE 5. Core components in transmission facilities including (a) a tower, (b) insulator and stock bridge damper, (c) spacer, and (d) marker ball.

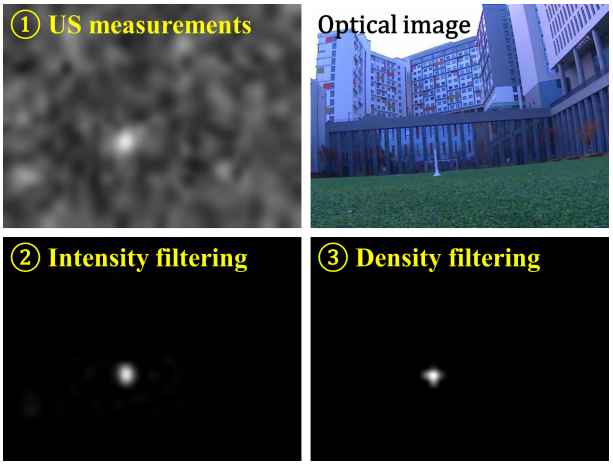

FIGURE 6. Results of single ultrasound detection at a distance of 5 m and an emission angle of 60° for Case 1.

the performance of the object detection step was validated using test datasets. Finally, Section 5.3 presents a scenario for damage detection using the proposed system.

A. ULTRASOUND SOURCE DETECTION

Case 1 tested the detectability of a single ultrasound at different distances. Fig. 6 depicts one experimental result through Phase A of the proposed method, where the distance between the ultrasound source and ultrasound array was 5 m, and the emission angle of the ultrasound source was 60°. The density threshold ρ_{th} was set to one pixel. Noise reflected from the surrounding buildings and the ground remained, even after intensity filtering (② in Fig. 6). However, the density filter eliminated most noise that was misjudged to be ultrasound emissions from a real source (③ in Fig. 6). Other results were similar and were omitted for the sake of brevity.

Fig. 7(a) shows the maximum measured intensity and detectability for Case 1. The maximum intensity decreased over the measurement distance. Moreover, for the same distance, the maximum intensity was the strongest when an ultrasound source directly faced an ultrasound array (0°) whereas the ultrasound intensity was attenuated when a source obliquely faced an ultrasound array. These results suggest that detection accuracy depends on the distance

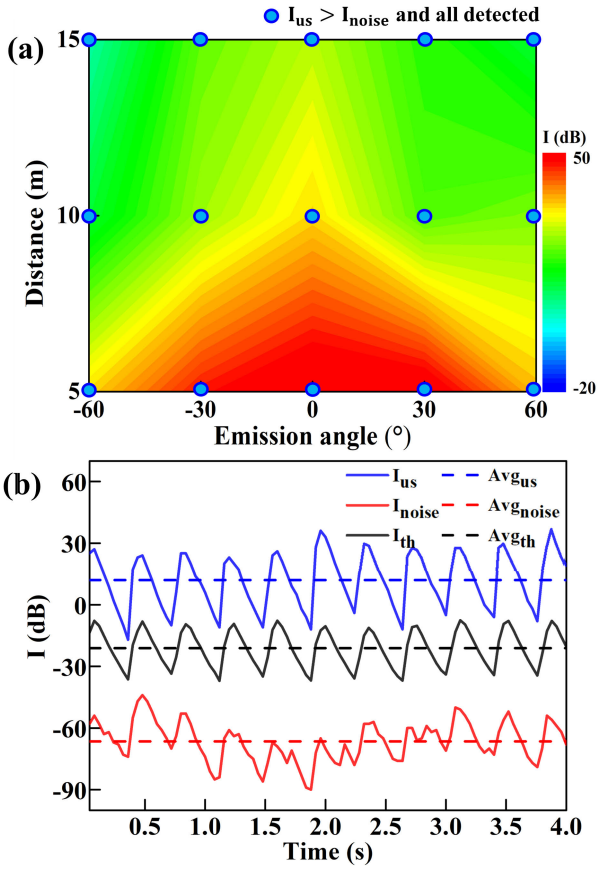

FIGURE 7. Results for Case 1. (a) Maximum intensity map. (b) Measured intensity of a single source and noise w.r.t time at 5 m from an ultrasound array with 60° emission angle.

and emission angle, implying that facilities should face an ultrasound array within a measurable angle.

Fig. 7(b) presents the ultrasound intensity from the source I_{US} (blue line), background noises I_{noise} (red line), and ultrasound threshold I_{th}(black line), which is automatically determined using the three-sigma outlier detection method over time. The ultrasound threshold I_{th} has the same trend as that of the source because the threshold is set statistically. Measurement reveals that the ultrasound intensity increases significantly around the source and exceeds 3σ of the measurement. Specifically, the mean of the ultrasound intensity at the source is 3 dB and that of the ultrasound threshold is -27 dB, whereas the noise intensity is -67 dB, suggesting that background noises are effectively eliminated. Hence, the proposed method successfully detects ultrasound emissions for all experiments, although the measured ultrasounds depend on the distance and emission angle, suggesting that ultrasounds emitted from a damaged system can be easily detected if no other noise source exists in the surrounding environment.

Case 2 describes the condition in which two instances of damage occur at the same distance and emit a similar intensity of ultrasound. However, the emission angle could be different because of the complex shape of power facilities. This is a more complicated condition than Case 1 because of the interference of the two ultrasounds.

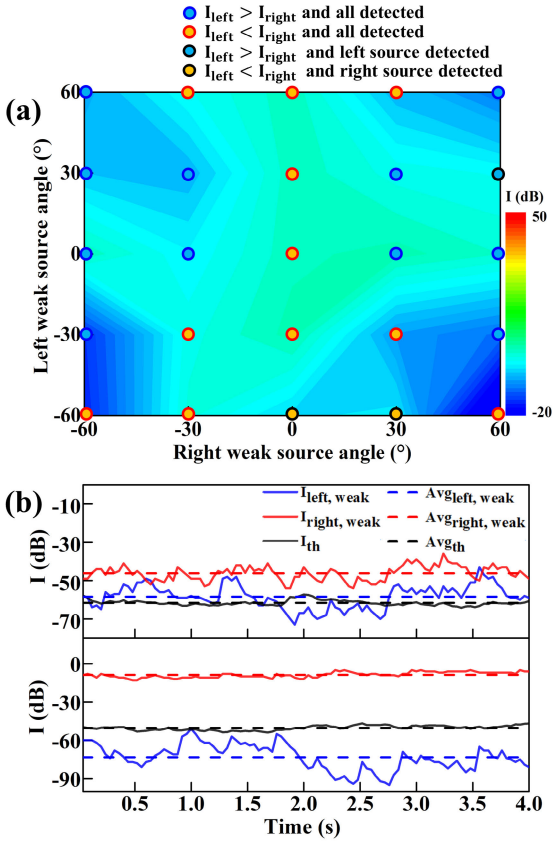

FIGURE 8. Results for Case 2. (a) Maximum intensity map. (b) Measured intensity of a weak source on the left side, a weak source on the right side, and intensity threshold with an emission angle of (top) −60° (left) /60°(right) and (bottom) −60°(left)/30°(right).

Fig. 8(a) presents the results for the maximum intensity and detectability. The experimental results reveal that the maximum intensity was low compared to the results of Case 1 in which both ultrasounds were emitted from weak sources. Moreover, detectability depends on the emission angle. Both ultrasounds were detected using an ultrasound array when both ultrasounds had a small emission angle (i.e., −30° to 30°; Fig. 9(a)), whereas ultrasounds headed with a larger emission angle were not detected (Fig. 9(b)). These results suggest that the relative emission angle of two ultrasounds affects the detectability due to interference. However, the noise decreased or was eliminated using intensity and density filters. Hence, the proposed method discriminates true ultrasounds from noise. Fig. 8(a) also reveals that the maximum intensity was affected by the emission angle of both sources because the intensity of the two sources was similar. In this case, the two sources were detected from 22 angles in 25 situations. Moreover, the source on the right side was detected when an emission angle of the source to the left was −60°, whereas the left source was detected when the emission angle of the right source was 60°. This phenomenon can be explained by the superposition of the forced linear wave equation [27].

Fig. 8(b) presents the ultrasound intensity of the left source (blue line), right source (red line), and intensity threshold

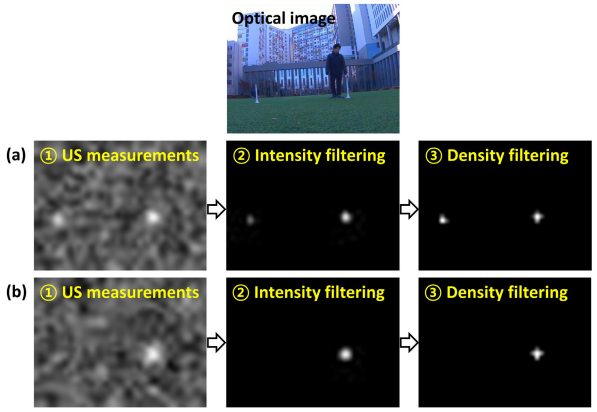

FIGURE 9. Results of ultrasound source detection with an emission angle of (a) −30°(left)/30°(right) and (b) −60°(left)/30°(right) for Case 2; US denotes ultrasound.

I_{th} (black line) over time with an emission angle of (top) −60°(left)/60°(right) and (bottom) −60°(left)/30°(right). With an emission angle of −60° and 60° for the left and right sources, respectively, the mean of both sources exceeds the threshold. However, the ultrasound intensity of the left source is lower than the ultrasound threshold I_{th} during a period of 1.75 to 2.75 seconds, suggesting that small defects may not be detected during entire period of measurements, and the detected frequency also should be accounted for. The left source is not detected with an emission angle of −60° (left) and 30° (right) because the ultrasound intensity of the left source is smaller than the ultrasound intensity threshold.

In Case 3, two sources emit ultrasounds with different intensities at 5 m from an ultrasound array to verify the detectability of multiple ultrasound sources. This case replicates two defects that are differently damaged in power facilities.

Fig. 10 illustrates the results of applying the proposed method for Case 3. Noise still occurs after the ultrasound measurement passes through the intensity filter, and the intensity of some noise was similar to that of a weak ultrasound, suggesting that the intensity filter limits the detection of two true sources. However, the two sources were detected after passing through the density filter (③ in Fig. 10(a)). Hence, the classification process through the DCNN might determine which component is damaged. However, a weak source is not detected, as depicted in Fig. 10(b) where two sources are installed at −60° and 30°, again confirming that the straightness and superposition of ultrasounds are important characteristics.

The detailed results for the maximum intensity and detectability are presented in Fig. 11(a). Both ultrasounds were detected for most cases; however, some cases could not detect the weak ultrasound, where the emission angle of the weak ultrasound θ_w on the left side was significantly oblique. Fig. 11(b) depicts the ultrasound intensity of the left source (blue line), right source (red line), and intensity threshold I_{th} (black line) over time with an emission angle of (top) −30°(left)/−30°(right) and (bottom) −60°(left)/−30° (right) to elucidate the reason for this observation.

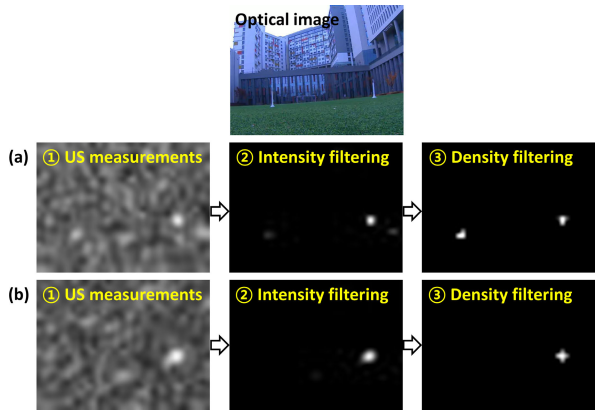

FIGURE 10. Results of ultrasound source detection with an emission angle of (a) −60°(left)/60°(right) and (b) −60°(left)/30°(right) for Case 3; US denotes ultrasound.

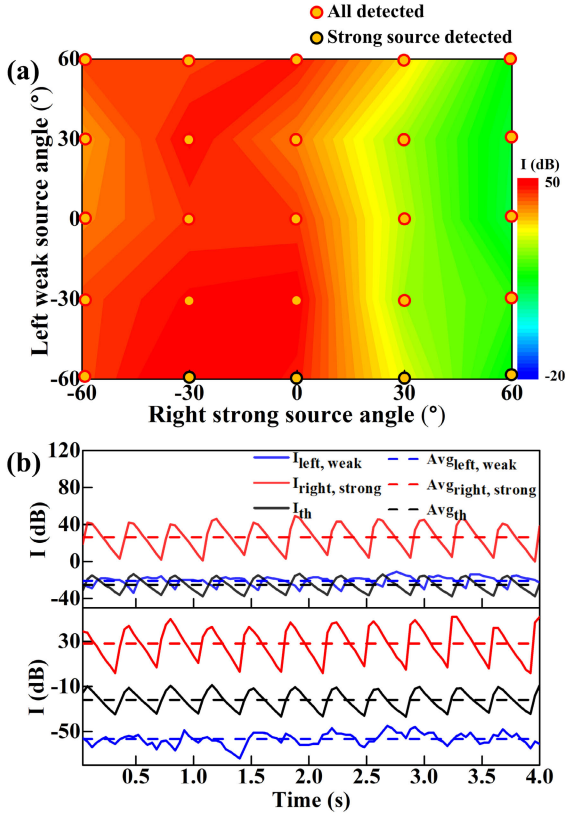

FIGURE 11. Results for Case 3. (a) Maximum intensity map. (b) Measured intensity of a weak source, strong source, and intensity threshold with an emission angle of (top)−30°(left)/−30°(right) and (bottom) −60°(left)/−30°(right).

This analysis demonstrates that the ultrasound intensity of the weak source is significantly different, although the maximum intensity from the strong source is similar, suggesting that ultrasound intensity depends on the emission direction.

Case 4 is the condition in which a weak and strong source were both installed at distances of 5 and 10 m from an ultrasound array. This condition replicated a situation in which a seriously damaged component was farther from an ultrasound array than a less-damaged component. This condition is plausible for our application because many transmission

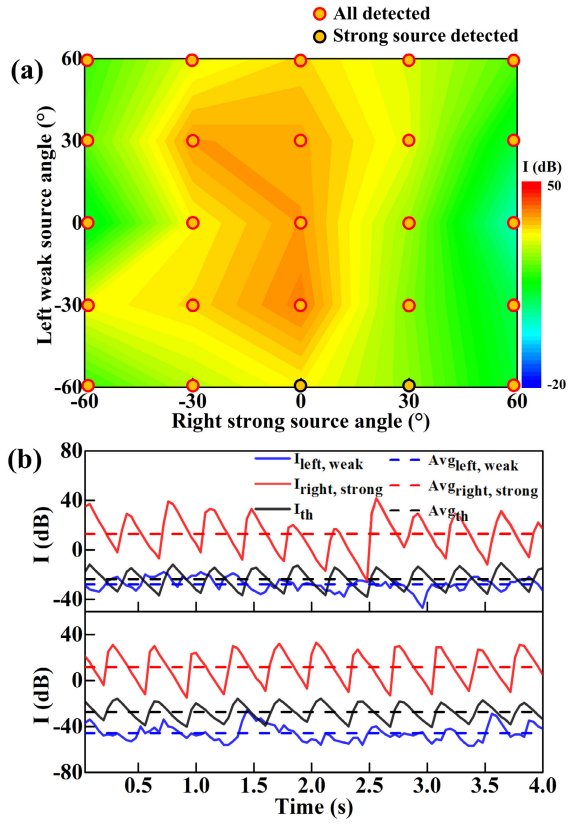

FIGURE 12. Results for Case 4. (a) Maximum intensity map. (b) Measured intensity of a weak source, strong source, and intensity threshold with an emission angle of (top) −30°(left)/0°(right) and (bottom) −60°(left)/0°(right).

facilities are located at different distances. Similar to the results of Cases 2 and 3, the intensity and density filters eliminated noise; therefore, ultrasound emissions from the true source were more clearly detected. However, a strong source was only detected for some cases, including those with an emission angle of weak and strong sources at −60°/0° and −60°/30° (Fig. 12(a)).

Detailed results for the maximum intensity and detectability are listed in Fig. 12(a). The maximum intensity from the strong ultrasound decreases in the overall experiments compared to Cases 1 and 3 because of the long distance. Specifically, the maximum intensity was 35 dB when the emission angles of the weak and strong sources were 60° and 0°, respectively, whereas the maximum intensity was 48 dB for the same angles in Case 3. Fig. 12(b) shows the reason for the undetected cases. The ultrasound intensity of the weak source is low, although the maximum intensity from the strong source is similar, suggesting that the relative ultrasound difference between two sources is important because the threshold is determined statistically. The proposed method detected both ultrasounds for 23 cases. The other two cases also detected strong sources, suggesting that seriously damaged components can always be detected, and less-damaged components can be detected if an inspection is conducted appropriately using the proposed method.

TABLE 3. Performance comparison with COCO data.

MS COCO data		Input size	FPS	mAP (IoU)			mAP (Area)		
				0.5:0.95	0.5	0.75	S (-32)	M (32-96)	L (-96)
SSD	[17],[19]	512	-	28.8	48.5	30.3	10.9	31.8	43.5
	Ours	× 512	15.2	28.4	47.9	29.4	10.4	32.6	44.5
YOLOv3	[18]	608	-	33.0	57.9	34.4	18.3	35.4	41.9
	Ours	× 608	27.3	30.7	52.8	32.4	15.6	32.7	39.6
M2Det	[19]	320	-	33.5	52.4	35.6	14.4	37.6	47.6
	Ours	× 320	19.5	34.1	51.9	36.9	13.6	39.1	49.1

The experimental results for Cases 1 to 4 suggest that the emission angle between the damaged surface and ultrasound array is not important for a single damaged condition. However, the emission angle between the damaged surface and ultrasound array and the angle between two ultrasounds are important, considering the convolution effects of the ultrasounds. All cases detected the ultrasounds at the emission angles of -30° to 30° using the proposed method, suggesting that damaged components can be detected if the ultrasound array faces the power facilities at the proper angle.

B. OBJECT DETECTION WITH DCNN

This study evaluated three state-of-the-art DCNN models for object detection using the MS COCO dataset [28]. The MS COCO dataset is one of the most popular datasets used to analyze the performance of DCNN models designed for object detection. In total, training/validation datasets of 2014 with 80 classification categories in the MS COCO dataset were used. The training of the DCNN model was executed on a graphics processing unit (GPU) server with two Intel Xeon 5218 (2.30GHz) CPUs with 32 GB memory and four RTX-2080Ti GPUs.

The AP is given as the ratio of true positives (TPs) to the total predicted positives, which is the sum of TPs and false positives (FPs) (i.e., AP = TP/(TP + FP)). The mean average precision (mAP) is mean value of AP for all objects, which was used as a metric of prediction performance in two different ways. One is the mAP in terms of the intersection of union (IoU), which is the intersecting area over the overlapping area between the ground truth and prediction, indicating the performance of both the object localization and classification. The IoU threshold was set to eliminate the effects of well-classified and inaccurately localized predictions. A high IoU threshold increases the FPs for objects. Hence, the AP is inversely proportional to the IoU threshold. The mAP were calculated, such as 0.5 or 0.75 in Table 3. Moreover, the mAP, 0.5:0.95 in Table 3, is the averaged mAP over different IoU thresholds from 0.5 to 0.95 with 0.05 increments.

This is a primary metric in many studies [17]–[19] because this metric evaluates the overall performance of the neural network. The mAP with the following three pixel areas also is calculated to check the object detectability: small (area $< 32^2$), medium ($32^2 <$ area $< 96^2$), and large (area $> 96^2$). The mAP for different areas is also an important metric to evaluate the capability of multi-scale and multi-level networks. The input images for the three different sizes are predefined in the test datasets for quantitative comparison.

All conditions, including input size, backbone network, and hyperparameters, were set to the same values as the SSD, YOLOv3, and M2Det methods described in [17]–[19] to conduct quantitative comparisons, as presented in Table 3. These results reveal that the mAPs estimated from the three models in this study have similar accuracy to the references, confirming that the three models were built properly. Moreover, M2Det exhibits the highest mAP in both the IoU and area, whereas YOLOv3 is the fastest in terms of calculation time. The SSD is inferior to the others in terms of both accuracy and calculation time. The FPS in the reference literature is not displayed in this paper because quantitative comparisons between references and this study cannot provide useful information due to the different hardware configurations. Hence, this paper only compares the FPS of the three models from our GPU server. This result suggests that M2Det should be introduced as a damage detection method for power facilities because the calculation time and accuracy are both important when attempting to detect defeats in transmission facilities. However, accuracy is more important than calculation time for our study.

As the next step, M2Det was trained with the datasets from the transmission facilities. The number of objects classified in the datasets is 33 555, which includes 3253 towers (9.7%), 6628 spacers (19.7%), 7653 SB dampers (22.8%), 1044 marker balls (3.1%), and 14 977 insulators (44.6%). All images are downsampled from 1920 × 1080 pixels to 800 × 800 pixels to reduce the calculation time.

Data augmentation was randomly conducted before training. Data augmentation included cropping, expansion, distortion, and mirroring to supplement a lack of data diversity and increase the accuracy of object detection. When training M2Det with images of transmission facilities, a batch size of 2, a momentum of 0.9, and a decay weight of 0.0005 were used as hyperparameters.

The structure of M2Det for the feature extraction parts comprises four multi-level feature layers and four multi-scale features. The total training epochs were 200, and the initial value of the learning rate was 4×10^{-4}. Every 20 epochs, the learning rate decreased by a factor of 1/2 until 60 epochs. After 60 epochs, the learning rate was set as 1/10 of the previous learning rate for every 20 epochs until 140 epochs. At 140 epochs, the learning rate was fixed at 1×10^{-8} for efficient training.

Table 4 lists the AP of all categories through M2Det concerning the IoU. The AP for five objects and the mAP was around 0.90 for the IoU threshold of 0.5, suggesting that datasets over 15 000 are effective for training M2Det. The AP for five objects and the mAP are almost the same, implying that all objects were evenly trained, although several input images are different. The AP for the tower and marker ball is higher than that of the others for the IoU thresholds even though the number of images for towers and marker balls

TABLE 4. Prediction accuracy of M2Det for transmission facility data.

Class object	AP (IoU)					
	0.5	0.6	0.7	0.8	0.9	0.5:0.95
Tower	89.6	89.5	89.5	77.1	38.3	71.4
Spacer	89.4	89.2	87.4	70.3	11.7	65.7
Marker ball	90.4	90.4	89.8	86.2	25.2	70.8
Insulator	88.0	79.2	74.7	45.6	4.0	54.2
SB damper	89.8	89.3	77.0	57.4	11.0	60.8
mAP	89.4	87.5	83.7	67.3	18.0	64.6

is smaller than the rest of the objects. These observations can be explained a factor that the unique shape of marker balls and towers discriminated from the surroundings results in distinct features. The AP for the insulator is the lowest because insulators are located near a tower and overlap with the structure of a tower in an image. This spatial characteristic of insulators interferes with feature extraction and results in low accuracy compared with other objects.

The AP and mAP decrease when the IoU threshold increases. The same phenomena are shown in mAP in terms of different IoUs in Table 3 due to the formula for the AP. A smaller size significantly decreases the IoU considering the effect of the same pixel difference in the calculation of small and large objects. The same origin also explains with mAP (area) in Table 3. The mAP (area) is high for large objects over 40% and low for small objects less than 20% because the low resolution of small objects significantly decreases the IoU.

The IoU thresholds of 0.5 and 0.6 are generally used in DCNNs for object detection [29]. These thresholds are also reasonable for our application. The flight protocol with 12X magnification of the optical zoom measures medium and large images for five components. The mAP can be increased by omitting the downsampling process of the original optical images. The tradeoff between the prediction accuracy and calculation time can be optimized in future work for real-time applications.

C. SCENARIO FOR FAULT DETECTION IN POWER FACILITIES

This section describes the damage detection scenario in transmission facilities with the proposed method and system equipped in a crewless aerial vehicle. This scenario demonstrates how the proposed method and system could be used in the actual field. A defect is not frequently found through field inspection; therefore, it is difficult to show an example of detecting damaged facilities. Validating the proposed method in field applications requires significant effort and time.

To illustrate the potential scenario, the measured ultrasound was overlapped with optical images. It assumed that the damage of an insulator, SB damper, and spacer emitted ultrasounds because the main damage of these facilities is correlated to the partial discharge of TLs, whereas the towers are insulated for safety. In addition, the main defect in a maker ball occurs on the surface.

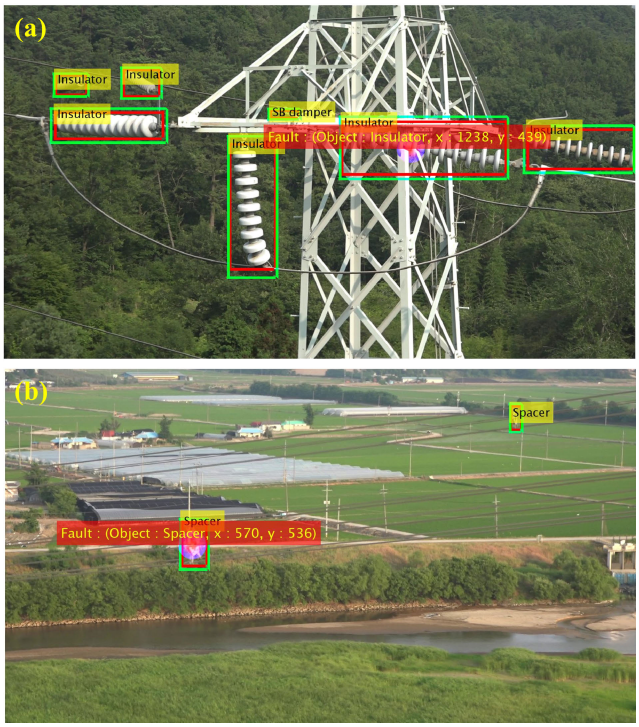

FIGURE 13. Scenario results using the object detection method. (a) Insulators and stock bridge (SB) dampers. (b) Spacers.

Considering the random characteristics in ultrasound diffusion, virtual ultrasound sources were randomly generated within an optimal image using the 2D Gaussian distribution. If the center of an ultrasound source was included in ROI_{OD}, a faulty alarm was generated with information on the defect location and class. Fig. 13 illustrates the examples of a scenario in which the defects in the transmission facilities were detected. This scenario demonstrates that Phase B provides information on the damaged components when ultrasounds are detected in Phase A using the proposed method.

VI. CONCLUSION

This paper presents a method for damage detection by fusing ultrasounds emitted from damaged surfaces and optical images. The proposed method comprises two phases for discriminating true ultrasounds from noise. The first phase preprocesses ultrasound measurements through intensity and density filters. The second phase aims to not only increase the accuracy of ultrasound detection but also classify a damaged system by comparing the location of an ultrasound with an optical image, which detects the object using object detection through a DCNN. The experiments demonstrated that single and multiple areas of damage could be effectively detected. However, proper angles that consider the straightness and superposition characteristics of ultrasounds are important for enhancing accuracy. Moreover, M2Det successfully recognizes the power facilities with up to 15 000 image datasets, suggesting that the DCNN is useful not only for enhancing accuracy but also for classifying damaged components. Finally, a scenario for damage detection with a crewless

aerial vehicle was demonstrated for potential applications. Future work includes many field experiments equipping the proposed method and system in a crewless aerial vehicle and verifying the quantitative accuracy in actual fields.

ACKNOWLEDGMENT

The authors would like to especially thank Dr. Joon-Young Park and Mr. Seok-Tae Kim for providing the image datasets for power transmission line facilities. Finally, the authors gratefully acknowledge the support of SM Instruments with the donation of the BATCAM 1.0 used for this research.

REFERENCES

[1] G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa, P. Kundur, N. Martins, J. Paserba, P. Pourbeik, J. Sanchez-Gasca, R. Schulz, A. Stankovic, C. Taylor, and V. Vittal, "Causes of the 2003 major grid blackouts in north america and europe, and recommended means to improve system dynamic performance," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1922–1928, Nov. 2005.

[2] R. Ferraro, *Field Gudie: Inspection of Conductors for Overhead Transmission Lines*. Washington, DC, USA: Electric Power Research Institute, 2015.

[3] S. Kim, D.-G. Kim, S.-H. Jung, and K.-Y. Oh, "Diagnostic methods for power transmission lines using a corona camera deployed on a mobility," in *Proc. PHMAP*, Beijing, China, 2019, pp. 285–287.

[4] *Annual Operation and Maintenance Report for Power Transmission Facilities*, Korea Plant Service & Engineering, Republic of Korea, South Korea, 2016.

[5] L. Wu, W. He, J. Hu, F. Zhou, B. Song, X. Shen, and T. Xu, "Study on daytime camera for corona discharge detection," *Semicond. Optoelectron.*, vol. 31, pp. 443–446, Jun. 2010.

[6] H. Ha, S. Han, and J. Lee, "Fault detection on transmission lines using a microphone array and an infrared thermal imaging camera," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 1, pp. 267–275, Jan. 2012.

[7] M. Voller and K. Mollmann, *Infrared Thermal Imaging Fundamentals, Research and Applications*. Hoboken, NJ, USA: Wiley, 2014.

[8] N. Pouliot, P.-L. Richard, and S. Montambault, "LineScout power line robot: Characterization of a UTM-30LX LIDAR system for obstacle detection," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2012, pp. 4327–4334.

[9] J. Park, S. Kim, J. Lee, J. Ham, and K. Oh, "Method of operating a GIS-based autopilot drone to inspect ultrahigh voltage power lines and its field tests," *J. Field Robot.*, vol. 37, no. 3, pp. 345–361, Apr. 2020.

[10] J. Szrek, J. Wodecki, R. Błażej, and R. Zimroz, "An inspection robot for belt conveyor maintenance in underground mine—Infrared thermography for overheated idlers detection," *Appl. Sci.*, vol. 10, no. 14, p. 4984, Jul. 2020.

[11] S. Kim, D. Kim, S. Jeong, J.-W. Ham, J.-K. Lee, and K.-Y. Oh, "Fault diagnosis of power transmission lines using a UAV-mounted smart inspection system," *IEEE Access*, vol. 8, pp. 149999–150009, 2020.

[12] Y. Wang, X. Li, Y. Gao, H. Zhang, D. Wang, and B. Jin, "Partial discharge ultrasound detection using the sagnac interferometer system," *Sensors*, vol. 18, no. 5, p. 1425, May 2018.

[13] J. Tuma, P. Janecka, M. Vala, and L. Richter, "Sound source localization," in *Proc. 13th Int. Carpathian Control Conf. (ICCC)*, High Tatras, Slovakia, May 2012, pp. 740–743.

[14] R. S. Figliola and D. E. Beasley, *Theory and Design for Mechanical Measurements*, 6th ed. Hoboken, NJ, USA: Wiley, 2016.

[15] J. Yadav and M. Sharma, "A review of K-mean algorithm," *IJETT*, vol. 4, no. 7, pp. 2972–2976, Jul. 2013.

[16] D. Han, "Comparison of commonly used image interpolation methods," in *Proc. 2nd Int. Conf. Comput. Sci. Electron. Eng. (ICCSEE)*, 2013, pp. 1–4.

[17] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. Berg, "SSD: Single shot MultiBox detector," in *Proc. ECCV*, Sep. 2016, pp. 21–37.

[18] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," in *Proc. CVPR*, Apr. 2018, pp. 1–6.

[19] Q.-J. Zhao, T. Sheng, Y.-T. Wang, Z. Tang, Y. Chen, L. Cai, and H.-B. Ling, "M2Det: A single-shot object detector based on multi-level feature pyramid network," in *Proc. CVPR* Jan. 2019, pp. 9259–9266.

[20] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788.

[21] O. Russakovsky, H. S. JiaDeng, and J. Krause, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Apr. 2015.

[22] A. Krizhevsky and I. Sutskever, "ImageNet classification with deep convolutional neural networks," in *Proc. NIPS*, Jan. 2012, pp. 1097–1105.

[23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *ICLR*, Apr. 2015, pp. 1–14.

[24] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[26] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 7263–7271.

[27] L. E. Kinsler, A. R. Frey, A. B. Coppens, J. V. Sanders, *Fundamentals of Acoustics*, 4th ed. Hoboken, NJ, USA: Wiley, 1999, pp. 128–139.

[28] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," in *Proc. ECCV*, Mar. 2014, pp. 740–755.

[29] M. Everingham, L. V. Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The PASCAL visual object classes (VOC) challenge," *Int. J. Comput. Vis.*, vol. 88, no. 2, pp. 303–338, Sep. 2009.

DONGGEUN KIM received the B.S. degree in energy system engineering from Chung-Ang University, Seoul, South Korea, in 2019, where he is currently pursuing the M.S. degree.

His research interests include diagnostic robots for extreme environments, artificial intelligence for diagnostics and prognostics, and design optimization.

SAN KIM received the B.S. degree in energy system engineering from Chung-Ang University, Seoul, South Korea, in 2019, where he is currently pursuing the M.S. degree.

His research interests include smart mobility, artificial intelligence for prognostics, and health management.

SIHEON JEONG received the B.S. degree in energy system engineering from Chung-Ang University, Seoul, South Korea, in 2019, where he is currently pursuing the Ph.D. degree.

His research interests include diagnostics and prognostics with smart mobility, artificial intelligent, simultaneous localization and mapping, and object cognition with novel sensors for autonomous driving and flight of smart mobility.

JI-WAN HAM received the B.S. degree in mechatronics engineering from Chungnam National University, Daejeon, South Korea, in 2010. He is currently pursuing the M.S. degree in energy system engineering with Chung-Ang University.

His research interests include artificial intelligence for prognostics and health management, simultaneous localization and mapping, and object cognition for autonomous driving and flight for smart mobility.

SEHO SON received the B.S. degree in energy system engineering from Chung-Ang University, Seoul, South Korea, in 2020, where he is currently pursuing the Ph.D. degree.

His research interests include deep learning, machine learning, and its applications in prognostics and health management.

KI-YONG OH received the B.S. degree in mechanical engineering from Hanyang University, Seoul, South Korea, in 2005, the M.S. degree in mechanical engineering from KAIST, in 2006, and the Ph.D. degree in mechanical engineering from the University of Michigan, Ann Arbor, in 2016. He joined the School of Energy System Engineering, Chung-Ang University, in 2017, where he is currently an Assistant Professor. His teaching and research interests include applied dynamics, prognostics, and health management in complex energy systems.