Automatic Discontinuity Classification of Wind-turbine Blades Using A-scan-based Convolutional Neural Network

Jiyeon Choung, Sun Lim, Seung Hwan Lim, Su Chung Chi, and Mun Ho Nam

***Abstract*—Recent development trends in wind power generation have increased the importance of the safe operation of wind-turbine blades (WTBs). To realize this objective, it is essential to inspect WTBs for any defects before they are placed into operation. However, conventional methods of fault inspection in WTBs can be rather difficult to implement, since complex curvatures that characterize the WTB structures must ensure accurate and reliable inspection. Moreover, it is considered useful if inspection results can be objectively and consistently classified and analyzed by an automated system and not by the subjective judgment of an inspector. To address this concern, the construction of a pressure- and shape-adaptive phased-array ultrasonic testing platform, which is controlled by a nanoengine operation system to inspect WTBs for internal defects, has been presented in this paper. An automatic classifier has been designed to detect discontinuities in WTBs by using an A-scan-imaging-based convolutional neural network (CNN). The proposed CNN classifier design demonstrates a classification accuracy of nearly 99%. Results of the study demonstrate that the proposed CNN classifier is capable of automatically classifying the discontinuities of WTB with high accuracy, all of which could be considered as defect candidates.**

***Index Terms*—Wind-turbine blade (WTB), blade inspection platform, convolutional neural network (CNN), discontinuity, phased-array ultrasonic testing (PAUT), A-scan.**

I. Introduction

As a promising renewable energy resource for reducing greenhouse gas emissions and satisfying the stringent environmental regulations, wind power has witnessed remarkable technical advancement in recent years. Many countries of the world have direct and uninterrupted access to wind power. This feature further enhances the economic applicability of wind-based electricity production [1]. Consequently, the global cumulative installed wind capacity has witnessed a dramatic 20-fold increase in recent years, from 23.9 GW in 2001 to 486.8 GW in 2016 [2]. However, alongside this unprecedented development, the number of structural-failure occurrences at wind power stations has also steeply increased, from an average of 33 accidents between 1998 and 2002 to nearly 167 incidents each year between 2013 and 2017 [3]. Since most wind-power stations have been installed either in the mountains or near shorelines, they are subject to erratic winds and strong turbulence. Such severe operation environment is likely to cause the development of fatigue cracks in wind-turbine blades (WTBs), thereby resulting in failure during natural disasters. As the size of wind turbine increases, there will be more risks of structural failure. In particular, large wind turbines frequently experience structural failure of blades. This means that the risk posed by structural failures to wind-power plants is presently difficult to estimate. Wind turbines need to operate continuously to provide the most efficient power production. Accordingly, quality management and maintenance of wind power generation facilities assume critical importance. When WTBs experience structural issues, additional costs are incurred to facilitate their repair and reoperation, and at the same time, the maintenance and production schedules are also delayed. Apart from the loss of time and money, pieces of broken blades pose a safety concern. The risk associated with these problems can be alleviated via thorough inspection of WTBs before they are placed into operation.

In particular, WTBs made of composite materials such as glass-fiber reinforced polymer (GFRP) and carbon-fiber reinforced plastic, may result in the occurrence of internal defects, including debonding, delamination, and crack initiation, owing to difficulties encountered during the manufacture of blades. Initially, these defects can be so small that they may be considered to have little impact on the structural integrity of WTBs. However, since WTBs are subject to repeated impulsive loadings, these defects tend to propagate, thereby threatening WTB structural integrity and resulting in catastrophic failures. If these initial defects could be detected in their infancy during the manufacturing stage and duly repaired in accordance with established guidelines, many serious failures and the accompanying economic damage can be

Manuscript received: October 10, 2018; accepted: October 12, 2019. Date of CrossCheck: October 12, 2019. Date of online publication: March 26, 2020.

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) "Development of System for Damage Detection on the Outer of Fibrous Composite Blade for Wind Power Plants In-process and In-service Inspection" (No. 20153030024070), funded by the Ministry of Trade, Industry, and Energy (MOTIE), Korea.

J. Choung (corresponding author), S. Lim, and S. H. Lim are with Korea Electronics Technology Institute (KETI), 401-402, Bucheon Technopark, 655, Pyeongcheon-ro, Wonmi-gu, Bucheon-si, Gyeonggi-do 14502, Korea (e-mail: cjy2457@keti.re.kr; sunishot@keti.re.kr; shair@keti.re.kr).

S. C. Chi and M. H. Nam are with Sam Yong Inspection Engineering Co. Ltd., New Black Seok Building, 19, Dasan-ro 11-gil, Jung-gu, Seoul 04598, Korea (e-mail: suchung@samyong.co.kr; nmh4241@samyong.co.kr).

DOI: 10.35833/MPCE.2018.000672

prevented.

Internal defects can be accurately detected via the use of an appropriate method. The traditional method of defect detection is based on the judgment of an experienced inspector concerning results obtained via the use of nondestructive testing (NDT) techniques such as the phased-array ultrasonic testing (PAUT). However, the reliability and objectivity of this approach has often been a matter of dispute. In view of this, the development of automatic defect detection techniques has become a subject attracting intense research focus in recent years.

The proposed study has investigated an automatic detection technique by developing a PAUT platform to facilitate precise examination of WTBs. The proposed platform has been equipped with a pressure- and shape-adaptive mechanism to accommodate the WTB profile curvatures, and the same has been controlled by the nanoengine operation system (NEOS), a real-time operation system. Among the signals from NDT, discontinuity has been a detectable change in WTBs. This could be generated due to some undesirable quality or the specimen state including the normal state. Therefore, not all discontinuities could be attributed to defects. They would need to be classified based on the quality level to identify the ones corresponding to defects. Discontinuity classification has played a role in classifying discontinuities to detect defects causing unacceptable quality levels. The study has also established a convolutional neural network (CNN)-based classifier capable of identifying internal discontinuities through the use of A-scan image data provided by PAUT test results. Discontinuities in the shape and structure of WTBs have been considered a factor for detection during the proposed research. Recent studies [4], [5] have reported detection and classification of certain types of defects. The objective of the proposed study, however, is to detect possible defects in WTBs regardless of their type. The proposed study is the first to report discontinuity diagnosis to facilitate the detection of the all possible defect types automatically with high levels of precision and reliability.

II. Existing Inspection Platform and Classification Method

A. Inspection Platform

Owing to mathematically and structurally complex curvatures of WTB surfaces, the use of conventional testing equipment and platforms could hardly ensure reliable inspection. Another important consideration has been the size of inspection equipment. Pre- and in-service WTB inspection equipments used by specialty companies and designed to accommodate complex WTB geometries are typically so large that they need to be operated by cranes. Presently, limitations with regard to the platform size have been addressed by researchers through the development of various types of robots such as those employed during the inspection of wind-turbine-rotor blades [6], micro-aerial vehicle-type wall-climbing robots [7], and the inchworm-type blade-inspection robot system [8]. The paper attempts to address the issue of complex blade geometries via the use of a novel method to adapt the proposed platform to WTB surfaces.

B. Classification Method

With the advancement in wind power generation technology, many studies have been performed with an objective to replace human-based inspection techniques with an automatic system to detect and identify blade defects. Reference [9] proposed an optical inspection method for the detection of surface cracks on an installed blade and demonstrated crack detection by using the Sobel and Canny image-processing technique. One of the most recent trends in defect detection has involved automatic classification of inspection results by using machine learning (ML) and neural network (NN). Reference [10] used unmanned aerial vehicles to obtain blade surface images, from which they extracted Haar-like features and detected defects by implementing Logiboost along with an extended cascading classifier. In another study concerning automatic detection, [11] applied the short-time Fourier transform to A-scan signals generated during surface inspection to extract features, and developed a new potential real-defect miner algorithm based on the k-means and support vector machine classifiers for clustering and classification. Reference [12] captured the surface of a cutting tool using a CMOS camera, and the features from resulting images were extracted by performing edge detection and discrete Fourier transforms whilst identifying defects on the surface via NN. To facilitate WTB maintenance, [13] acquired ultrasonic signals and performed signal preprocessing to extract features. Subsequently, detection results obtained via artificial neural networks (ANNs) were compared against those obtained via the use of various ML methods such as the decision tree, quadratic discriminant analysis, and weighted k-nearest neighbors, only to find that the use of ANN demonstrated the highest detection accuracy. However, most studies concerning the use of ML or NN have been performed to determine the means to extract features prior to employing a detection technique. In this regard, there has been growing interest in the use of CNN, a deep-learning method widely employed in image analysis. CNN applications typically include the use of an NN that could identify features without the need for a separate feature-extraction step. Reference [14] employed a one-class classifier based on the CNN approach to facilitate the image-defect detection with regard to the small amount of electronic component present within the image dataset. Reference [15] employed both mathematical and CNN models to perform photometric extraction of stereo images to detect defects on the surface of a steel rail, thereby demonstrating that the CNN-based approach could provide higher detection accuracy than mathematical models.

III. Experimental Method

The procedure for automatic defect detection contains three steps: WTB inspection, data preprocessing, and CNN for data classification. In the first step, accurate data acquisition is very important in order to ensure precise blade to internal examination. To address this concern, a pressure- and shape-adaptive PAUT platform is developed in this study to facilitate the precise inspection of WTB-surface curvatures.

Data are first collected on a blade testbed by performing inspections using the proposed PAUT platform. Subsequently, in the data pre-processing stage, wherein image data are transformed to be interpretable by a deep-learning algorithm, datasets are grouped based on certain criteria, and concerned images are transformed into arrays. This is accomplished through the use of *numpy.asarray* function in Python. The CNN model is designed using the TensorFlow (TensorFlow™) and Keras packages within Eclipse IDE (Eclipse Foundation, Inc). Transformed data are subsequently identified via the training of the CNN model, which is subsequently used to detect discontinuities. Finally, the original dataset is successfully used to automatically identify internal discontinuities. Once a WTB is thoroughly inspected, the CNN model is used to automatically determine whether the collected data contain any discontinuities.

A. WTB Testbed

A large-sized WTB typically measures 40-50 m in length. In this study, an actual WTB structure measures 43.5 m in length, is fabricated using GFRP, and comprises two shear webs. However, such a large structure is difficult to test experimentally in view of the storage- and cost-related concerns. Since the objective of this study is to test the detection accuracy and performance of the proposed technique using the general shapes of the WTBs, the testbed is fabricated as a scaled-down 3 m long model of the actual WTB geometry described above with the corresponding scaling factor 12 in consultation with the blade manufacturer, as depicted in Fig. 1. The ply means a thickness of layer of a laminated material in Fig. 1, and 1 ply is 0.9 mm in the testbed. Edge joints, which typically forms main inspection targets, and internal shapes are designed to resemble those of a full-size WTB, as closely as possible. GFRP, identical to that used in full-sized WTBs, is used to manufacture the test blade. Candidate defects that develop and potentially occur are listed in Table I, which have been compiled in accordance with manufacture and operation conditions of typical WTBs. Candidate defects are described based on the experiences and knowledge of NDT inspection experts. Among the candidate defects that occur during manufacture and operation, the delamination is chosen as an artificial defect. The delamination is introduced along the leading and trailing edges of the blade in accordance with inputs from actual blade manufacturers.

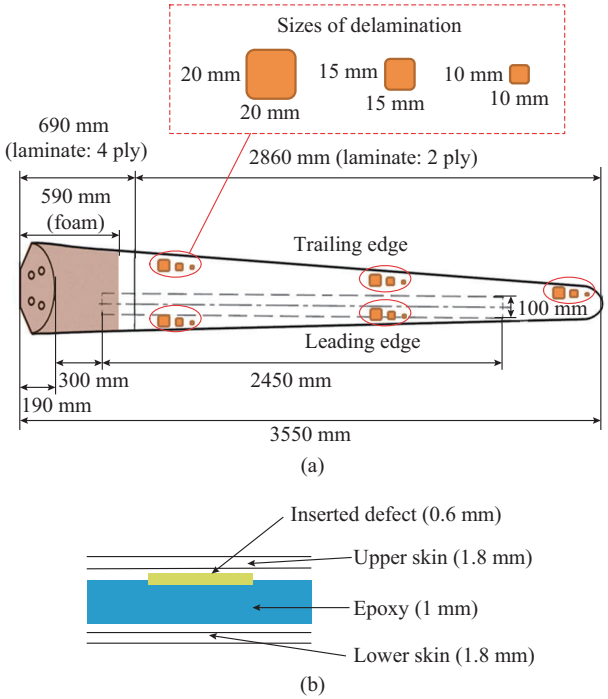

Fig. 1. Testbed with sizes and locations of artificial defects. (a) Overview. (b) Side view of testbed and delamination.

TABLE I
CANDIDATE DEFECTS THAT OCCUR DURING MANUFACTURE AND OPERATION OF BLADE

Defect occuring time	Defect	Detail
Defect that occurs during manufacture	Delamination	Gap between stiffener (glass fiber) and base material (resin)
	Bubble and blow hole	Occurrence of bubbles between plies
	White area	Some plies incompletely filled with resin during reaction injection molding process
	Crack	Ply fiber severed
	Wrinkle	Excessive strengthening work causing wrinkles to appear on surface
Defect that occurs during operation	Debonding	For cases wherein plies are incompletely bonded, pressure loads during operation result in debonding
	Lightning	When all parts of the blade are wet, lightning cannot be conducted to ground
	Fatigue defect	Invisible small defects that continue to grow during operation
	Corrosion	Wind gradient and foreign substances tending to corrode blade surfaces

A WTB testbed with intentional defects is constructed with three different delamination sizes. As depicted in Fig. 1, the delamination, in the form of artificial square-shaped defect measuring 20 mm, 15 mm, 10 mm, and 0.6 mm in depth, respectively, is induced along the leading and trailing edges of the WTB. The 20 mm defect corresponds to the largest size that could be inserted into the small blade joint. The 10-mm defect is selected based on data obtained from a plane testbed experiment, wherein the defects of this size are either detectable or independent of the thickness of the composite material. The 15 mm defect is selected to represent the mean value between the 20 mm and 10 mm extremes.

B. Pressure- and Shape-adaptive PAUT Platform

Phased-array ultrasonic testing is an NDT technique for identifying internal defects within structures. The PAUT ap-

proach provides a 2D thermogram in real time by transmitting ultrasonic waves of diverse amplitudes into an object. Compared with other ultrasonic testing techniques, the implementation of this method could improve detection reliability and reduce testing time by simultaneously generating ultrasonic signals at various angles. In this study, a Dynaray Lite tester and Zetec probes are used to perform PAUT. Since the testbed blade geometry is rather small, the dual-probe technique is considered suitable for performing edge inspection. An efficient dual-probe method is adopted as the pitch-catch technique [16], which enables Zetec probes to possess the advantage of high resolution.

WTB structures are highly non-linear and irregularly shaped, which possess significant non-linearity in the form of large surface curvatures. Furthermore, since WTBs could be fabricated in different shapes in accordance with their intended service environment, if a conventional PAUT platform is used for inspection, the accuracy and reliability of inspection signals collected across a wide range of testbed geometries would tend to degrade. Accordingly, to facilitate the inspection of different types of blade-joint shapes with high curvatures, a flexible active-passive constant pressure mechanism platform capable of performing various roll, pitch, and *Z*-axis motions is fabricated as part of this study. The platform described is completely attached to the blade surface, thereby maintaining constant pressure against it and facilitating accurate inspection of the internal structure of the leading and trailing edges of blade even if the inspection tools are moved. This approach enables the platform to adjust to the flexible range of motion over surfaces of diverse shapes. When the inspection by conventional method is repeated three times for each size of the artificial defect on the leading edge, the average measurement error is 1.59 mm with a standard deviation of 1.92 mm. The average measurement error is 1.06 mm with standard deviation of 2.24 mm when using the pressure- and shape-adaptive PAUT platform under equivalent conditions. By comparing the average measurement errors between the conventional and proposed platform inspection methods on the trailing edges, errors of 1.93 mm and 1.01 mm are obtained, respectively. The standard deviations are 3.4 mm and 1.22 mm, respectively. The proposed PAUT platform thus reduces measurement error by 33.34% on the leading edge and 47.67% on the trailing edge. The NEOS software is used to control platform operation. NEOS is a real-time operation system with DO-178B safety certification. It is a special-purpose embedded system that facilitates accurate and precise processing [17].

Figure 2 depicts the experimental setup for defect inspection with the PAUT platform applied to the testbed. Three supporting legs are attached to the testbed along with the control part for the movement of axes of the PAUT device. The PAUT device consists of a wedge, probe, and encoder attached to the instrument unit of the testing platform for defect inspection. The two Zetec probes serve as the pulser and receiver and are moved using the 23 mm flat wedge. The entire experimental setup is placed inside an indoor pool to facilitate the use of water as a medium for defect inspection. A Windows notebook running the X86 NEOS software is used to control the platform. The collected data are transmitted to the notebook PC, wherein the Zetec program, UltraVision 3.8R30, is installed.

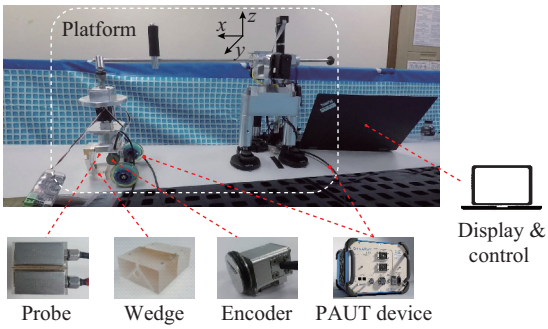

Fig. 2. Defect inspection system with PAUT platform.

C. Discontinuity Classification Criteria

The data obtained via the PAUT device consist of the A-scan signals, C-scan image, and D-scan image. S-scan images could also be obtained from PAUT instead of D-scan images. Among them, the basic signal output obtained from the PAUT platform is an A-scan signal. C-scan and D-scan represent a mixture of A-scan and distance information. The A-scan signal is obtained by transmitting a single ultrasonic beam from top to bottom along the length of the blade. This signal is expressed as the ultrasonic path travelled by an ultrasonic wave against ultrasonic-pulse amplitude (along the *y*-axis). The corresponding graph depicts locations and intensities of reflected ultrasonic waves. The index axis represents the inspection direction along which the probe and wedges traverse, whereas the scan axis denotes the inspection range. C-scan images are captured to depict the top view of the blade as well as to illustrate the relationship between the index and scan. Likewise, D-scan and S-scan images depict the relationships between the ultrasonic path and index. They are obtained by multiple moving ultrasonic beams, causing multiple A-scan signals to overlap along and perpendicular to the index and inspection directions, respectively. D-scan and S-scan images are obtained using linear and sectorial beams, respectively. This study mainly focuses on the use of A-scan signals for differentiating between internal discontinuities and defect-free detections, since A-scan signals facilitate the sorting of defects. A discontinuity may not necessarily imply the presence of a defect, but it could be used as a basic criterion to determine candidate defect probabilities. Therefore, A-scan signals form the basic information necessary to determine the presence of internal discontinuities, whereas C-scan and D-scan images could be used to measure the lengths of the defects and facilitate the visualization of discontinuity.

The classification criteria for identifying A-scan signals are defined based on the AASHTO/AWS D1.5M/D1.5: 2015, a standard approved by the American National Standards Institute [18]. Following the K10.2 acceptance criterion prescribed in this standard, discontinuities are confirmed to be identified when the size and length of an indication exceed specific values. Figure 3 depicts the discontinuity classification indication levels for Classes A, B, C, and D. Here, full-

screen height (FSH) corresponds to the maximum signal strength provided by the equipment set for a 1.5 mm reflector, as specified in the standard under equipment calibration. If FSH is set to 100%, signals exceeding 100% could not be identified. In the proposed study, the value of FSH is set to be 80%. The standard sensitivity level (SSL) indicates correspondence to a reference value, and in this study, it is set to be 50±5% of the FSH value. The automatic reject level (ARL) is set higher compared to SSL by 5 dB and at 89% of FSH. The disregard level (DRL) is set lower compared to SSL by 6 dB and at 25% of FSH. When indication peak exceeds ARL, it is classified as Class A. Class B refers to the indication greater than SSL and less than or equal to ARL, whereas Class C refers to the indication greater than DRL and less than or equal to SSL. Indication peak less than or equal to DRL is categorized as Class D. Amongst the four classified defect classes, Class D is considered to possess negligible discontinuity. Therefore, the proposed study adopts 20% signal amplitude as the criterion for indicating the presence of a defect. When this criterion is applied to inspection images of the WTB leading edge, Fig. 4(a) is observed to contain a discontinuity, and the corresponding D- and C-scan images clearly demonstrate the presence of defects. On the other hand, Fig. 4(b) depicts negligible discontinuity, and the corresponding D- and C-scan images demonstrate the presence of non-apparent defects.

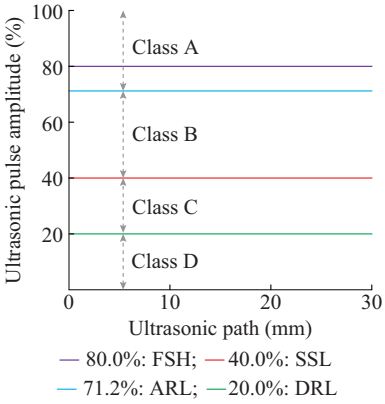

Fig. 3. Discontinuity classification indication levels.

D. Data Pre-processing

The collected dataset contains 12487 images, 9361 of which are observed to depict discontinuities, whereas the remaining 3126 images depict defect-free detection. Since the platform-based internal inspection of blade mainly focuses on defects determined from raw-image data, discontinuity data account for nearly 75% of the total data. With such an unbalanced dataset used to train the model, the model classifies all images containing discontinuity, thereby constituting an overfitting problem. Data augmentation is subsequently performed to efficiently solve this problem by increasing data volume artificially [19], [20], wherein horizontal flips of images containing indications are introduced to increase the presence of indications, thereby resulting in improved data balance.

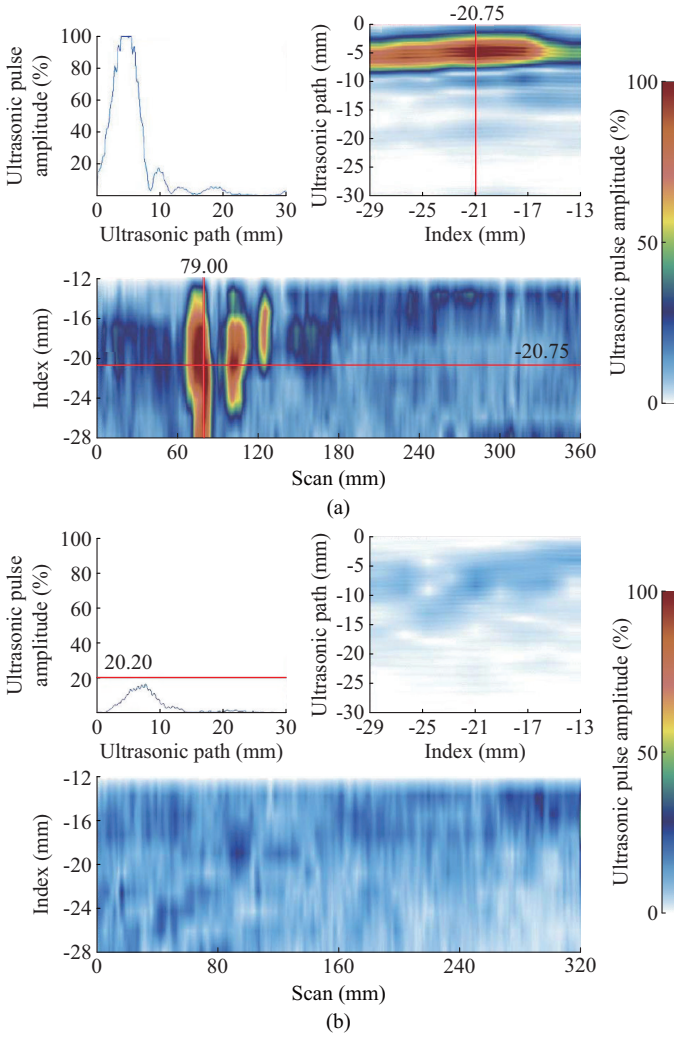

Fig. 4. Example of PAUT result image. (a) Example of observed defects. (b) Example of defect-free detections.

Horizontal flip is a technique of creating mirror images. Therefore, the adjusted dataset contains 15613 images, 9361 of which demonstrate discontinuities while the remaining 6252 contained defect-free detections. Each image initially possesses an approximate resolution of 440×400 pixels, but once the model is trained, the accuracy for smaller image sizes is also increased. Consequently, the size of collected images is reduced by half to approximately 220×200 pixels during data preparation. These half-sized images correspond to the smallest reduction rate, which in turn, demonstrate the increase of operation accuracy. RGB pixels of smaller images provide a similar data distribution by means of a normalization process, wherein the range is converted from [0, 255] to [0, 1], and the training and test datasets were divided into a ratio of 8:2 learning from inference data.

E. Automatic Discontinuity Detection

This study utilizes a deep learning algorithm to design a classification system capable of automatically differentiating discontinuities in A-scan signals. Two NVIDIA GPU GeForce GTX 1080 Ti cards with 11 GB of memory are used to process the learning and inference datasets. Since A-scan

signals obtained from the proposed PAUT platform mainly contain images, CNN is used to distinguish between images. CNN is a deep feed-forward ANN used in the extraction and classification of image features. Basic CNN contains three key structural layers: a convolution layer to extract data features; a pooling layer that plays a key role in subsampling to reduce feature mapping; and a fully-connected layer to facilitate data classification.

An activation function is used within each layer to create a feature map and add non-linearity to the NN. Thus, complex data patterns could be learned [21]. Activation functions generally employed include the threshold, linear, Gaussian, and sigmoid functions [22]. Generally, the sigmoid function is used as an ANN activation function. However, when only the sigmoid function is applied to each layer in this study, a vanishing gradient problem is observed during the course of back-propagation. The vanishing gradient problem refers to a case where the gradient magnitude approaches zero as more gradients are calculated through the layer [23], which makes it difficult to calculate the error rate and reduce calculation accuracy. This problem is solved by using the rectified linear unit (ReLU) proposed in [24]. The ReLU function is based on the principle that a value of zero could be assigned to all negative parameters. Further, for any value assigned to positive parameters, the function returns the input value to the output [25].

The loss function could calculate the difference between the calculated values for a model and the corresponding intended values. The mean squared error (MSE) represents one possible example of a loss function. MSE measures the average of squared error values. The binary cross-entropy, also known as the negative log loss, denotes a probabilistic interpretation loss function in classifier problems [26]. The output of binary cross-entropy is a probabilistic result with an intermediate value between 0 and 1. This result could be interpreted as a probability belonging to the positive class. It is essential to determine the point where minimum losses are incurred. If a CNN model is trained to be excessively fit to a particular training dataset, it fails to possess sufficient generality, thereby, once again, representing an overfitting problem. Regularization [27], dropout [28], and batch normalization could be applied to prevent the occurrence of such overfitting. Regularization is a penalty on model complexity to improve generality and could be divided into L1 and L2 types. The L1 regularization technique adds an absolute magnitude to the cost function, while the L2 type adds a squared magnitude to the cost function. In this study, L2 regularization, also referred to as "weighty decay" regularization, is employed. Weighty decay regularization reduces large weights by adding a parameter that imposes a penalty upon the cost function [29], [30]. If a model is trained to reduce weights in this manner, it tends to be less affected by outliers corresponding to local noise, thereby preventing overfitting and providing acceptable generalization. Dropout, as a complementary method to regularization, randomly removes units (nodes) from a network [31]. Batch normalization is performed on a mini-batch basis to facilitate the uniform learning of layers, each of which could possess different distributions owing to an internal covariate shift within deep networks [32]. Theoretically, prior to nonlinearity evaluation, batch normalization requires a network that must be designed before the activation function.

F. Model Estimation

A model is typically evaluated based on its accuracy in terms of inferring test data. However, the demonstrated accuracy in this study is considered insufficient to perform detailed evaluations, thereby resulting in wrong conclusions for cases where datasets include an unbalanced number of data points between the classes. To address these issues, this study adopts a confusion matrix, as described in Table II, which could distinguish the polarity between actuality and prediction to evaluate the classification system model [33], [34].

TABLE II
PREDICTION CONFUSION MATRIX

Actual polarity	Predicted polarity	
	Predicted negative	Predicted positive
Actual negative	True negative (TN)	False positive (FP)
Actual positive	False negative (FN)	True positive (TP)

Estimation scores are derived in (1)-(6) based on the confusion matrix.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$REC = \frac{TP}{TP + FN} \tag{2}$$

$$SPE = \frac{TN}{TN + FP} \tag{3}$$

$$PRE = \frac{TP}{TP + FP} \tag{4}$$

$$FPR = \frac{FP}{TN + FP} \tag{5}$$

$$FNP = \frac{FN}{FN + TP} \tag{6}$$

where *TN*, *TP*, *FN*, and *FP* are the values of TN, TP, FN, and FP, respectively. Accuracy *ACC* is the number of images correctly predicted from the whole set of images. Recall *REC* is the number of positive predictions over the actually positive images. It is also called sensitivity or true positive rate. Specificity *SPE* is the number of negative predictions over the actually negative images. Precision *PRE* is the number of actually positive images among the positively predicted images. False positive rate *FPR* refers to the number of positively predicted images from actually negative images. Further, false negative rate *FNR* refers to the number of positively predicted images from the actually negative images. When the designed model could predict the results with high accuracy and reliability based on a new dataset, it is considered to be a good model. *ACC*, *REC*, *SP*, and *PREC* are all considered optimum as they approach 1, whereas *FRP* and *FNP* are considered optimum as their corresponding values approach zero.

However, in the case of an unbalanced binary classifica-

tion dataset, the F1 score, defined in (7) as the harmonic mean of precision and recall, could be used as an indicator superior to any of the previously defined elements [35]. As the F1 score approaches a value of unity, the evaluated model could be considered better, since the F1 score depends on the average of precision and recall, which are considered optimum when their value is closer to unity.

$$F1\_Score = 2\frac{PRE \cdot REC}{PRE + REC} \tag{7}$$

IV. Results

The CNN classification model contains 11 layers: 1 input layer, 6 convolution layers, 3 fully-connected layers and 1 output layer. Figure 5 depicts the final architecture of the CNN discontinuity classifier where Input, Conv-32, Conv-64, MaxPool, BN, Flatten, FC-512, FC-128, FC-64 and Output represent the input layer, 2D convolution layer with 32 filters, 2D convolution layer with 64 filters, max pooling layer, batch normalization, layer to change the 2D shape into 1D shape, fully-connected layer with 512 output neurons, fully-connected layer with 128 output neurons, fully-connected layer with 64 output neurons and output layer, respectively. Based on the observed variation in accuracy, batch normalization in this study is performed at the end of each layer to normalize input data on the basis of each layer. In terms of hyper parameters, the batch size considered is 50, the drop rate equals 0.25 at each layer, and the learning rate is set to be 0.01. To facilitate the evaluation of loss function, a binary cross-entropy is applied. Two types of active functions are utilized within each layer except the input. Additionally, within the convolution layer, the ReLU function is applied as an active function. Since a binary classification model is employed, the sigmoid function is applied in the fully-connected layers and output layer. Test accuracy is determined to be approximately 99% with regard to the classification of discontinuities and defect-free detection.

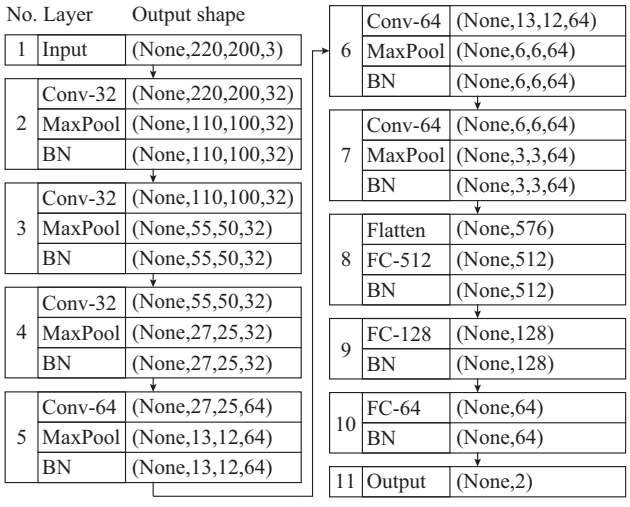

Fig. 5. Architecture of a discontinuity classifier.

The confusion matrix of the proposed classifier, which is used as an additional evaluation criterion, has been presented in Table III. The total number of successfully classified images within the test dataset equals 3094, which represents the sum of 1826 and 1268 images containing discontinuities and defect-free detections, respectively. The sum of classified failure images equals 29 with 23 images constituting the false-positive section, while the false-negative section contains 6 images. As can be observed in Table IV, the values of REC, SPE, and PRE are 0.995, 0.988, and 0.982, respectively, while the observed F1 score is 0.989. The FPR is also considered important, because if blade discontinuities remain unidentified or are classified as normal instead, a serious failure could occur. The resulting FRP is observed to be 0.012, which is considered to be acceptably close to zero.

TABLE III
Confusion Matrix for Discontinuity Classifier

Actual population	Predicted population	
	Negative (discontinuity)	Positive (defect-free)
Negative (discontinuity)	1826	23
Positive (defect-free)	6	1268

TABLE IV
Result Concerning Elements Derived Using Confusion Matrix

Element	Value		
ACC	0.991	FPR	0.012
REC	0.995	FNR	0.004
SPE	0.988	F1 score	0.989
PRE	0.982		

Failure cases are classified into two types. In the first type, the discontinuity classifier identifies six images as containing a discontinuity, whereas in actuality, the images are defect-free detections. In the second type, the classifier identifies 23 images to demonstrate defect-free detections. However, the images contain discontinuities. With regards to the first type, as shown in Fig. 6, peaks of all signals within images attain the 20% signal amplitude set as the criterion. With regards to the second type, Fig. 7(a) depicts classification failure for small-amplitude defects, whereas Fig. 7(b) depicts corresponding failures for large-amplitude defects. For cases depicted in Figs. 6 and Fig. 7(a), signal amplitudes demonstrate the existence of an ambiguous peak near the 20% boundary. As shown in Fig. 7(b), for a majority of cases, signal peaks are located on the extreme right of images. Therefore, the classifier would recognize the right side of these peaks as the end or boundary of the image. If more images are provided with their respective signal peaks lying on the extreme right of the domain, the use of the CNN classifier could effectively reduce the number of failure cases.

V. Conclusion

This paper presents the design of a CNN-discontinuity classifier capable of identifying discontinuities within WTBs through the use of A-scan image data. To obtain accurate A-scan images of raw data collected from a scaled-down WTB testbed, a pressure- and shape-adaptive PAUT platform using NEOS is fabricated.

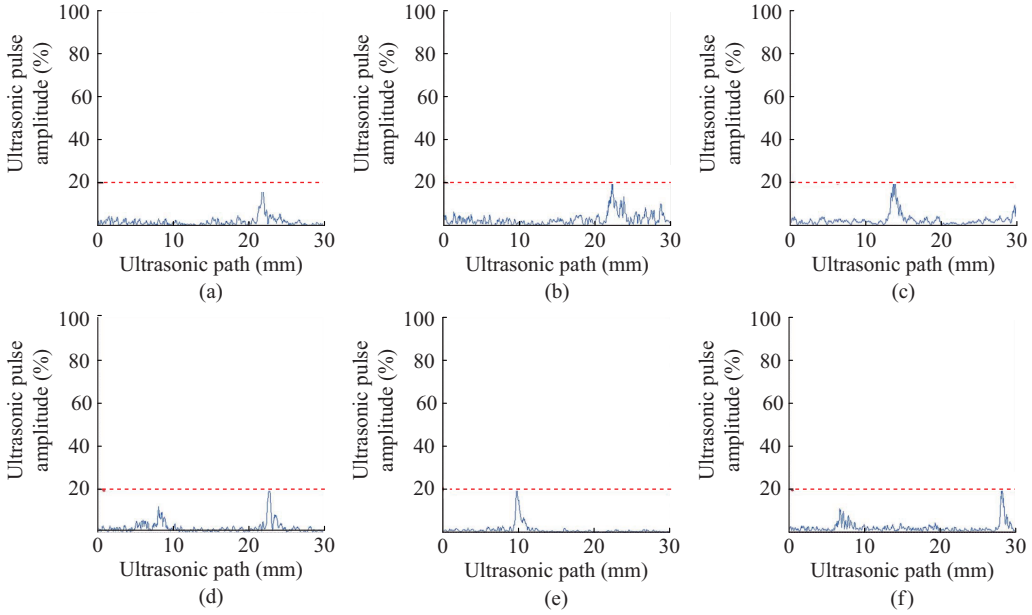

Fig. 6. Actual defect-free detection results predicted as discontinuities. (a) Result image 1. (b) Result image 2. (c) Result image 3. (d) Result image 4. (e) Result image 5. (f) Result image 6.

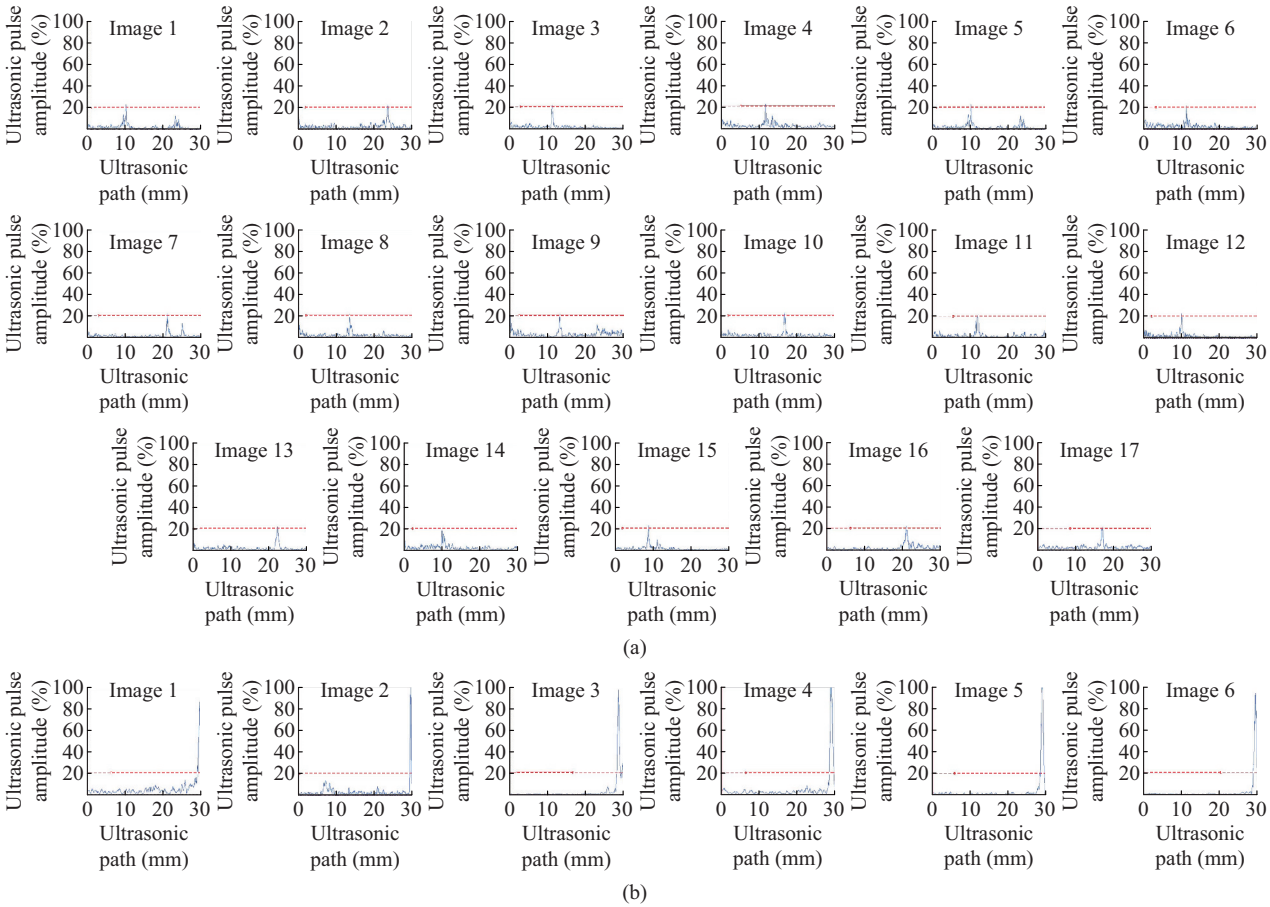

Fig. 7. Actual discontinuity results predicted as defect-free detections. (a) With small-amplitude defects predicted as defect-free detections. (b) With large-amplitude defects predicted as defect-free detections.

Scan data are successfully obtained by using a real-time control system, and an accurate CNN classifier is designed to exhibit a classification accuracy of nearly 99% along with an F1 score of 0.989. The novelty of the proposed classifier design could be stated in terms of its ability to automatically detect all candidate defects via discontinuity classification. The proposed CNN classifier could be used for a variety of WTB sizes, whilst it is also applicable in a wide range of

scenarios concerning internal-defect detection using the equipment capable of generating A-scan signals such as PAUT or other UT devices. As a future endeavor, the authors would intend to apply the proposed CNN classifier to actual full-scale WTBs to verify its real-time detection accuracy.

REFERENCES

[1] T. Traber and C. Kemfert, "Gone with the wind? - Electricity market prices and incentives to invest in thermal power plants under increasing wind energy supply," *Energy Economics*, vol. 33, no. 2, pp. 249-256, Mar. 2011.

[2] Global Wind Energy Council. (2018, Mar.). Global wind report 2016. [Online]. Available: https://gwec. net/publications/global-wind-report-2/global-wind-report-2016/

[3] Caithness Windfarm. (2019, Mar.). Summary of wind turbine accident data. [Online]. Available: http://www. caithnesswindfarms. co. uk/AccidentStatistics.htm

[4] M. Ruiz, L. E. Mujica, S. Alferez *et al*., "Wind turbine fault detection and classification by means of image texture analysis," *Mechanical Systems and Signal Processing*, vol. 107, pp. 149-167, Jul. 2018.

[5] S. Zhou, Y. Chen, D. Zhang *et al*., "Classification of surface defects on steel sheet using convolutional neural networks," *Materiali in Tehnologije*, vol. 51, no. 1, pp. 123-131, Feb. 2017.

[6] N. Elkmann, T. Felsch, and T. Forster, "Robot for rotor blade inspection," in *Proceedings of 2010 1st International Conference on Applied Robotics for the Power Industry (CARPI 2010)*, Montreal, Canada, Oct. 2010, pp. 1-5.

[7] S. Jung, J. Shin, W. Myeong *et al*., "Mechanism and system design of MAV (micro aerial vehicle)-type wall-climbing robot for inspection of wind blades and non-flat surfaces," in *Proceedings of 2015 15th International Conference on Control, Automation and Systems (ICCAS)*, Busan, South Korea, Oct. 2015, pp. 1757-1761.

[8] S. Lim, C. Park, J. Hwang *et al*., "The inchworm type blade inspection robot system," in *Proceedings of 2012 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, Daejeon, South Korea, Nov. 2012, pp. 1-4.

[9] H. Zhang and J. Jackman, "Feasibility of automatic detection of surface cracks in wind turbine blades," *Wind Engineering*, vol. 38, no. 6, pp. 575-586, Dec. 2014.

[10] L. Wang and Z. Zhang, "Automatic detection of wind turbine blade surface cracks based on UAV-taken images," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 9, pp. 7293-7303, Mar. 2017.

[11] G. Rui and A. Gachagan, "Advanced defect detection algorithm using clustering in ultrasonic NDE," *AIP Conference Proceedings*, vol. 1706, no. 1, pp. 1-8, Feb. 2016.

[12] S. Inoue, M. Konishi, and J. Imai, "Surface defect inspection of a cutting tool by image processing with neural networks," *Memoirs of the Faculty of Engineering, Okayama University*, vol. 43, pp. 55-60, Jan. 2009.

[13] A. A. Jiménez, C. G. Muñoz, and F. G. Márquez, "Machine learning for wind turbine blades maintenance management," *Energies*, vol. 11, no. 1, pp. 1-13, Dec. 2017.

[14] M. Zhang, J. Wu, and H. Lin, "The application of one-class classifier based on CNN in image defect detection," *Procedia Computer Science*, vol. 114, pp. 341-348, Oct. 2017.

[15] D. Soukup and R. Huber-Mörk, "Convolutional neural networks for steel surface defect detection from photometric stereo images," *Advances in Visual Computing*, vol. 2014, no. 1, pp. 668-677, Dec. 2014.

[16] M. Moore, B. Phares, G. Washer *et al*. (2014, Jul.). Guidelines for ultrasonic inspection of hanger pins. [Online]. Available: https://rosap.ntl.bts.gov/view/dot/40559

[17] National Defense/Aerial RTOS (NEOS). (2018, May). Defense/aviation RTOS (NEOS). [Online]. Available: http://www. mdstec. com/solutions/?no=376

[18] Bridge Welding Code, AASHTO/AWS D1.5M/D1.5: 2015-AMD1. (2016, Nov.). An American National Standard. [Online]. Available: https://pubs.aws.org/Download_PDFS/D1.5M-D1.5-2015-PV.pdf

[19] S. Wong, A. Gatt, V. Stamatescu *et al*., "Understanding data augmentation for classification: when to warp?" in *Proceedings of 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, Gold Coast, Australia, Nov.-Dec. 2016, pp. 1-6.

[20] L. Perez and J. Wang. (2017, Dec.). The effectiveness of data augmentation in image classification using deep learning. [Online]. Available: https://arxiv.org/abs/1712.04621

[21] S. Abrahams, D. Hafner, E. Erwitt *et al.*, *TensorFlow for Machine Intelligence*. Seoul: Hongneung Gwahak (in Korean), 2016, p. 171.

[22] A. Jain, J. Mao, and K. Mohiuddin, "Artificial neural networks: a tutorial," *Computer*, vol. 29, no. 3, pp. 31-44, Mar. 1996.

[23] T. Mikolov, A. Joulin, S. Chopra *et al.* (2015, Apr.). Learning longer memory in recurrent neural networks. [Online]. Available: https://arxiv.org/abs/1412.7753

[24] V. Nairand and G. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, Haifa, Israel, Jun. 2010, pp. 807-814.

[25] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks", in *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS)*, Ft. Lauderdale, USA, Apr. 2010, pp. 315-323.

[26] M. Bernico, *Deep Learning Quick Reference*. Birmingham: Packt Publishing, 2018, p. 57.

[27] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge: MIT Press, 2016, pp.116-117.

[28] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS)*, Siem Reap, Cambodia, Dec. 2012, pp. 1097-1105.

[29] A. Ng, "Feature selection, L1 vs. L2 regularization, and rotational invariance," in *Proceedings of the 21st International Conference on Machine Learning (ICML)*, Banff, Canada, Apr. 2004, p. 78.

[30] A. Krogh and J. Hertz, "A simple weight decay can improve generalization," *Neural Information Processing Systems*, vol. 1992, pp. 950-957, Jan. 1992.

[31] N. Srivastava, G. Hinton, A. Krizhevsky *et al*., "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929-1958. Jan. 2014.

[32] S. Ioffe and C. Szegedy. (2015, Mar.). Batch normalization: accelerating deep network training by reducing internal covariate shift. [Online]. Available: https://arxiv.org/abs/1502.03167

[33] R. Kohavi and F. Provost, "Glossary of terms," *Machine Learning*, vol. 30, no. 23, pp. 271-274, Feb. 1998.

[34] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," *PLoS One*, vol. 10, no. 3, pp. 1-21, Mar. 2015.

[35] A. C. Müller, S. Guido, and H. Bak, *Introduction to Machine Learning with Python*. Seoul: Hnipanbit Media (in Korean), 2016, pp. 344-345.

Jiyeon Choung received the M.Sc. degree in advanced engineering robotics from University of Bristol, Bristol, U.K. in 2012. She is currently an Assistant Researcher with the Intelligent Robotics Research Center in Korea Electronics Technology Institute, Bucheon-si, South Korea. Her research interests include data analysis, autonomous robot system, deep learning and artificial intelligence.

Sun Lim received the Ph. D. degree in electrical engineering from Korea University, Seoul, South Korea, in 2014. He is currently a Researcher with the Intelligent Robotics Research Center in Korea Electronics Technology Institute, Bucheon-si, South Korea. His research interests include control theory, robot control, adaptive control and synchronization control.

Seung Hwan Lim received the M. Eng. degree in mechanical engineering from Korea University, Seoul, South Korea, in 2014. He is currently an Assistant Researcher with the Intelligent Robotics Research Center in Korea Electronics Technology Institute, Bucheon-si, South Korea. His research interests include mechatronics, computer aided engineering (CAE) and mechanism design.

Su Chung Chi received the M. Sc. degree in electrical engineering from University of Soongsil, Seoul, South Korea, in 2013. He is currently a Researcher with an attached research institute in Sam Yong Inspection Engineering Co. Ltd., Seoul, South Korea. His research interests include nondestructive testing, inspection method and control.

Mun Ho Nam received the Doctor of Philosophy Candidate in electrical engineering from University of Soongsil, Seoul, South Korea, in 2012. He is currently a Senior Researcher with an attached research institute in Sam Yong Inspection Engineering Co. Ltd., Seoul, South Korea. His research interests include nondestructive testing, inspection method, renewable energy systems and control systems.