IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 27, 2021, accepted August 2, 2021, date of publication August 6, 2021, date of current version August 19, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3102954

PGGAN-Based Anomaly Classification on Chest X-Ray Using Weighted Multi-Scale Similarity

HYE-JEONG KWON[1], DONG-HOON SHIN[1], AND KYUNGYONG CHUNG[2]
[1]Department of Computer Science, Kyonggi University, Suwon-si, Gyeonggi-do 16227, South Korea
[2]Division of AI Computer Science and Engineering, Kyonggi University, Suwon-si, Gyeonggi-do 16227, South Korea

Corresponding author: Kyungyong Chung (dragonhci@gmail.com)

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) (No. 2019R1F1A1058651).

ABSTRACT To use artificial intelligence to assist in diagnoses applications, a model to utilize quality data is required, which results in massive time and cost. In medical data, data imbalance occurs because the amount of data with lesions is less than that without lesions. To overcome this limitation, this study proposes a progressive growth of generative adversarial networks (PGGAN)-based anomaly classification on chest X-rays using weighted multi-scale similarity. An anomaly detection method is applied to learn the distribution of normal images to solve the problem of data imbalance. The use of PGGAN, which is a model that generates high-resolution images by gradually adding layers, enables to find image characteristics on a multi-scale and define the similarity between an original image and a generated image. The anomaly score is calculated by applying the weighted arithmetic mean to a resolution-by-resolution similarity. The threshold is defined after the analysis of the F1-score, and then the classification performance is evaluated. The accuracy of the proposed model was assessed using a confusion matrix and compared with that of a conventional classification model, and the efficiency was demonstrated through ablation studies. The classification accuracy of the test dataset was 0.8525. Compared to a U-net-based disease classifier with low-resolution which accuracy was 0.8410, the performance of the proposed model was 0.8507, exhibited an improvement.

INDEX TERMS Artificial intelligence, computer-aided diagnostics, deep learning, healthcare, PGGAN, unsupervised learning, X-ray data.

I. INTRODUCTION

Artificial intelligence (AI), one of the main technologies of the 4th Industrial Revolution, has been studied and applied in several industries. Accordingly, a variety of AI-based methodologies for computer-assisted diagnosis have been studied in the medical service area [1]–[3]. The development of an assisted diagnosis tool based on AI enables to solve issues of this area, such as monitoring healthcare data collected by massive amounts of sensors and devices [39]. Appropriate diagnosis criteria help to achieve consistent judgment. In this regard, high accuracy and fast judgment are influenced by data evaluated and algorithms applied. Therefore, the use of a good assisted-diagnosis technology based on AI is expected to improve this process [4], [5]. To apply AI to the diagnosis of a disease and the detection of lesions, which are abnormal changes in the tissue of an organism, a model is required to understand the disease [40]–[42]. Therefore, it is necessary to use and process massive data such that AI can appropriately learn the characteristics of a disease [4], [6]. Additionally, an image expert's accurate annotation process is essential in the training data such that AI can recognize that a disease is an anomaly. These preparation processes require time, labor, and cost [7]. In terms of the application of deep learning technology, a variety of models have been used to minimize the consumption of unnecessary resources. Conventionally, they have been researched in the categories of supervised and unsupervised learning.

Wang and Xia [8] proposed a model for diagnosing 14 diseases based on chest X-ray images in a convolutional neural network (CNN) and attention-based model. In the proposed method, the attention-based model enabled to detect diseases despite inefficient information of data (position, label, etc.). This model achieved the second highest area under

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

the curve (AUC) score, of 0.781, among the state-of-the-art (SOTA) models. Wang *et al.* [9] proposed a method for detecting chest diseases through weekly supervised multiple classification by using patients' frontal-view X-ray images. In other proposed methods, the network of the pre-trained classification model was changed, and the multiple-label deep CNN (DCNN) model was trained [10]–[12]. Thus, faster and better performance was achieved, such that the annotated data were automatically generated. In [13], a variety of pre-trained DCNN models were compared to evaluate their performance, from which ResNET-50 showed the best performance for seven out of eight diseases [13]. Because the proposed method used a large dataset with 112,120 data, it failed to guarantee its performance in the case of data shortage. Aviles-Rivero *et al.* [14] proposed an image classification of the lung with lesions through minimum supervision. By applying graph-based semi-supervised learning and using a multiple classification function, the proposed method resolved the lack of labeled data. Thus, with the use of 20% trained data in the lung image classification of ChestX-ray14 data, the dependency on annotated data was minimized. The AUC of the GraphX net using 20% of the data was 0.79.

Regarding supervised learning, the problem of data imbalance, which is caused by the quantity difference between normal data (without lesions) and abnormal data (with lesions), has been extensively researched [15]. To solve this problem in advance, unsupervised learning-based models perform learning using only normal data. Tang *et al.* [16] proposed a classifier for detecting chest diseases using a generative adversarial network (GAN). In the proposed model, the structures of the discriminator and encoder were added to the U-net structure, and chest X-ray images were regenerated. The U-Net auto-encoder architecture enables preserving high-resolution capabilities through connections in the upsampling process [17]. For classification, the result (fake image) from the model learning normal images was compared with that from the model learning abnormal images. The accuracy of the proposed method was 0.841, which was far higher than that of the U-net architecture, of 0.627. Nevertheless, the GAN-based learning model requires significant resources. The actual data had a high resolution of 1024×1024, which was reduced to 64×64 in the study. As a result, data information was lost. Madani *et al.* [18] proposed a classifier for detecting chest diseases using GAN. Compared to the CNN-based classifier [19], the proposed classifier showed a higher performance (85.10%) in terms of 2,000 annotated data. When the number of annotated data was 10, the CNN-based method and the proposed method had performances of 51.27% and 73.08%, respectively; thus, there was a large difference in performance. This means that the proposed method solves the problem of insufficient data in image classification using deep learning by classification with the use of small amount of annotated data [20]. Nevertheless, based on the training of adjusted 32×32 images, it is difficult to consider fine diseases.

For learning, a conventional lesion classifier changes a high-resolution image to a low-resolution image [21]. Given that medical images have high resolution, information on fine tissues can be lost, which negatively influences the classification [22]. Therefore, in terms of a GAN-based model for learning the distribution of normal data and discerning the distribution of abnormal data, this study proposes the use of progressive growth of GANs (PGGANs) as a base model. This model shows better performance for high-resolution image generation than the deep convolutional GAN (DCGAN) [23], [24]. An anomaly score (AS) is defined by analyzing the similarity between an abnormal image and a fake image generated using the PGGAN model. It is calculated as a weighted average according to the similarity in multiple resolutions of an image and the amount of information. This study makes the following contributions:

- AI-based assisted diagnosis technology solves the problem of insufficient human and physical resources in the medical service area;
- A segmented high-resolution image generation model overcomes information loss and failures in the detection of small lesions, which are common in conventional unsupervised learning-based models;
- Solve the class imbalance problem of medical data using an unsupervised learning model that learns only the normal data distribution;
- Analyze similarities considering the scale space of the data generated by PGGAN. It is possible to find the invariant features of lesions robustly, despite a scale change. By weighting the distortion of image regeneration through scale and similarity, small-sized lesions were prevented from being excluded from abnormality.

The remainder of this paper is organized as follows: Section 2 describes related work on GAN-based high-resolution image generation and multi-scale image structure similarity evaluation. Section 3 describes the proposed chest X-ray anomaly classification method using PGGAN-based weighted multi-sale similarity. Section 4 describes the experimental results and performance evaluation of the model. Section 5 presents the conclusions of this study.

II. RELATED WORK

A. GAN-BASED HIGH-RESOLUTION IMAGE GENERATION

Conventional GAN models have an unstable learning process and low-resolution images for learning. In the case of generators and discriminators competitively learning, the higher the image resolution, the more difficult a generator deceives a discriminator. In addition, high-resolution images entail limitations of storage space and learning speed limit. Accordingly, a reduction in the batch size leads to unstable learning and performance degradation [25]. The proposed model to solve the problem is the PGGAN. PGGAN learns by adding layers progressively to the generator and discriminator. With a progressive increase in resolution, it is possible to first find the large-scale structure of the image distribution and

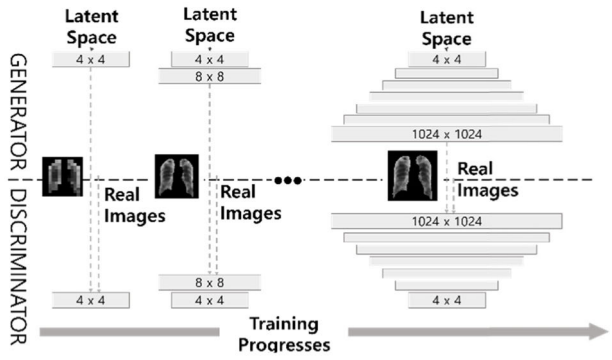


FIGURE 1. Learning process of PGGAN.

to gradually represent detailed features. Fig. 1 shows the learning process of PGGAN.

As shown in Fig. 1, layer addition starts from the smallest layer (4×4) and ends at the largest layer (1024×1024). At the beginning of learning, the large-scale structure, a feature found in low-resolution images, is discovered through image distribution. Through progressive learning, the overall features of the learning images are trained. Along with layer growth, detailed features are trained. Through the fade-in type in the process of increasing the resolution, the output of the previous resolution layer influences the high-resolution output. Considering the learning result of a previous layer, PGGAN progressively generates a high-resolution image from a large-scale low resolution.

Thus, PGGAN overcomes the limitation of conventional GAN-based models, which usually generate images with a low resolution of approximately 64×64, by creating images with a high resolution of 512×512. In addition, it achieves stable learning and reduces the learning time through the layer growth of the generator and discriminator. In the medical imaging field, studies have been conducted on anomaly classification and nodule detection using the excellent characteristics of PGGAN. Yoo *et al.* [26] applied PGGAN to conjunctival melanoma detection to improve performance. PGGAN was used to generate images with a high resolution of 256×256 to solve the problems of data imbalance and data shortage in ocular surface images. The detection performance of the data was improved using CycleGAN and PGGAN [44]. However, a GAN model was not applied after the data augmentation.

B. MULTI-RESOLUTION IMAGE STRUCTURE SIMILARITY METRIC

To detect image features, a variety of feature detection techniques have been applied on multiple scales. The image size is inversely proportional to the scale. Therefore, a scaled-down image shows the entire structure, and a scaled-up image provides a microscopic feature. A multi-scale image presented in such a way is called a scale space [27]. In scale space, it is possible to find image structures from different perspectives. The simplest method for obtaining a scale space is to size up or down an image. If a scale space is obtained through a simple size change, the detailed image features can be lost. Thus, a conventional image quality assessment method is used to maintain the image size and obtain the scale space of an image through blur or lower resolution [28]. A typical image similarity assessment method considering multiscale is the multi-scale structural similarity (MS-SSIM) [29]. Unlike the quality measurement method to discriminate the similarity of images through the calculation of the distance between pixels, the structural similarity index measure (SSIM) considers a cognitive structure and generates a result similar to that of human perception ability [30]. The luminance, contrast, and structure of an image are compared, and thereby one local quality map is calculated. Equation (1) is the formula for calculating the SSIM:

$$SSIM\ (x, y) = [l\ (x, y)]^{\alpha}[c(x, y)]^{\beta}[s(x, y)]^{\gamma} \tag{1}$$

x and y are the input images for comparison. Functions l, c, and s calculate the image feature maps of luminance, contrast, and structure, respectively. α, β, and γ are the variables that adjust the contribution of each evaluation element. Equation (2) presents a formula for calculating the luminance of each image:

$$l\ (x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \tag{2}$$

μ represents the mean of the two images. Equation (3) is used to calculate the contrast of each image:

$$c\ (x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \tag{3}$$

σ represents the standard deviation of each image. Equation (4) is used to calculate the structure of each image:

$$s\ (x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \tag{4}$$

In (2)–(4), $C_i (i = 1, 2, 3)$ is a positive constant to solve the instability when the operation value of the denominator converges to zero [30]. After the contribution of the three elements is adjusted, the final structural similarity of an image is drawn. Accordingly, visual image features are obtained by comparing three image features for the structural evaluation of an image, rather than a simple difference between pixels. MS-SSIM developed from SSIM can compare SSIM on multiple scales and discriminate similarity. After blurring an original image, MS-SSIM downsamples it and considers the K scales of an image in a top-down manner ($k = 1$; full resolution image). Equation (5) presents the formula for calculating the similarity through multi-scale SSIM:

$$\text{MS-SSIM} = [l_k(\mathrm{X}, \mathrm{Y})^{\alpha_K}] \prod_{k=1}^{K} [c_k(X, Y)^{\gamma_k}][s_k(X, Y)^{\gamma_k}] \tag{5}$$

In (5), α_k, β_k, and γ_k are the variables for adjusting the relative contributions of each element. With the use of PGGAN, this study considers the scale space of images in a bottom-up manner to generate an image (from low resolution

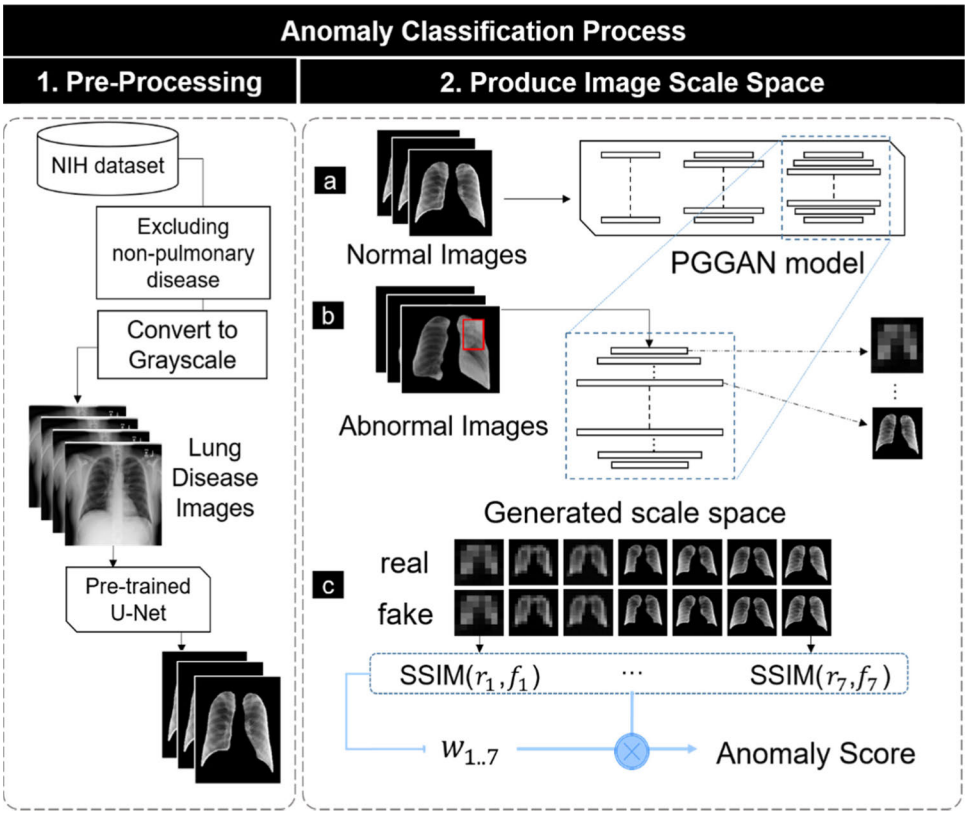


FIGURE 2. Structure of process for the chest X-ray anomaly classification using PGGAN-based weighted multiscale similarity.

to high resolution). The method used in this study identifies a feature through model learning in a random space, rather than down-sampling-after-blurring, and then increases the resolution. A generator that learns different resolutions is used to find the scale space of an image. Thus, the structural feature of a high-resolution image is found by considering the information amount of the image at multiple scales.

III. PGGAN-BASED ANOMALY CLASSIFICATION ON CHEST X-RAY USING WEIGHTED MULTI-SCALE SIMILARITY

To detect lesions in chest X-ray images, this study proposes a PGGAN-based model using weighted multi-scale similarity. The proposed model consists of three steps: pre-processing, model training, and estimation of the decision boundary. Fig. 2 shows the structure of the process for chest X-ray anomaly classification using PGGAN-based weighted multi-scale similarity.

In the first pre-processing step of Fig. 2, a pre-trained U-net is used to remove unnecessary features other than the lung from an image. In other words, this is the step of lung segmentation [31]. The NIH Chest X-ray dataset applied is image data that include a patient's chest, and a diversity of burned-in annotations and other organ structures [8], [9]. Therefore, to increase the efficiency of anomaly classification, regions other than the region of interest (lung) are removed. In the second step, the PGGAN is trained using normal data. The PGGAN model can generate images with high resolution. With the use of the model, an X-ray image is generated as health data. A normal lung image generation model is trained only with data that do not include lesions, and compared with a ground truth image [32]. At this time, as for the fake image generated in the PGGAN model, in which a random image is regenerated in the existing random vector, it is difficult to find a reference image clearly. Accordingly, the latent vector calculated through the feature extract layer of the discriminator is applied to generate a fake image that matches the ground truth image [33]. To find abnormal regions in images with high resolution, it is necessary to detect fine defects and abnormal regions in multiple scales, such as medically assisted diagnosis systems for visual inspection [34]. Therefore, an anomaly is defined by considering a scale space, and its image is compared with the original image. From the view of the multi-resolution detected based on the drawn similarity, an image information amount is weighted, and an AS is defined. In the last step, the decision boundary is estimated based on the AS of the entire dataset. With this estimation, the image is classified into normal or abnormal.

A. DATA PRE-PROCESSING

The NIH Chest X-ray dataset [9] was used in this study as the data for anomaly classification, which are frontal-view chest X-ray image data offered by the National Institutes of Health of the US (NIH).

The data consist of 112,120 images with a resolution of 1024 × 1024, with a total of 30,805 patient data.

In addition, they have the labels of 14 diseases (atelectasis, consolidation, infiltration, pneumothorax, edema, emphysema, fibrosis, effusion, pneumonia, pleural thickening, cardiomegaly, nodule, mass, and hernia). Because a GAN-based model learns the probability distribution of data, it is important to use a large amount of data. Only normal data were used as training data, totaling only 50,361 data, excluding test data and disease-existing data. Among the 14 diseases, cardiomegaly and hernia are disorders related to the heart and bowel. As this study focused on the classification of lung defects, disorders unrelated to the lung were not considered.

An anomaly is defined as one of the 12 diseases related to the lung. In terms of lung diseases, the lesion region is very small compared to the entire image size. For this reason, the presence of a lesion has no influence on a feature of the entire image. In addition, X-ray images in the NIH dataset include a patient's physique, organs, and various components. In addition, situational features at the time of shooting X-ray images, such as a patient's posture, are influential. These unnecessary data affect the classification of the abnormal images. Fig. 3 shows the preprocessing of the chest X-ray data for the removal of unnecessary features.

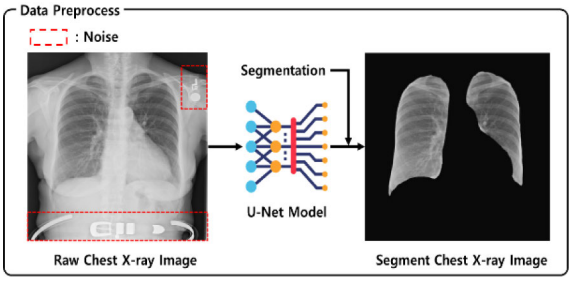


FIGURE 3. The preprocessing process of the chest X-ray data for the removal of unnecessary features.

In the chest X-ray, the original image of Fig. 3 shows text and unnecessary noise marked with a red dotted line, which includes an indicator presenting the direction (L; left, R; right), and objects such as the examinator's waist belt shown on the bottom, or medical tools [35]. These features are not on the region of interest of the model and are unnecessary information for lesion classification. Thus, to prevent incorrect classification of abnormalities due to other structures, it is necessary to preprocess unnecessary regions and extract regions of interest. In the medical service area, U-net [17] is applied to image segmentation for labeling particular regions in a medical image. To improve the accuracy of the model, this study used a pre-trained U-net to perform lung segmentation. The segmented data were used as the learning data for the PGGAN, which regenerates images. By using the segmented lung data of the chest X-ray image as learning data, it is possible to increase the learning accuracy and make a robust classifier.

B. STRUCTURAL SIMILARITY METRIC OF MULTI-RESOLUTION IMAGE

The lung data segmented in the preprocessing step were used as learning data for the PGGAN. The DCGAN-based model applied for anomaly detection sometimes fails to generate high-quality images with high resolution. A model using a generator of low-quality images fails to perform anomaly classification of images with high resolution [36]. Although it does not consider detailed features of lesions necessary for the classification of medical images with high resolutions, it allows many noises. For learning, this study applied the PGGAN, which supports the stable learning of images with high resolutions. PGGAN learns by adding layers progressively to generator G and discriminator D. The layer addition starts with a small size of 44 and ends with a large size (512×512), sequentially. At the beginning of learning, a large-scale structure, a feature found in images with low resolution, is found through image distribution. With progressive learning, the overall features of the learning images are learned. Detailed features are trained with the layer addition. Thus, PGGAN progressively generates images with high resolutions, overcoming the limitations of conventional GAN-based models that generate images with a low resolution of approximately 64×64. In this study, by using PGGAN, images with progressively growing resolutions were generated and converted into multi-scale ones. Thus, it is possible to find the feature vector of an image in a random latent space. The distribution of normal lung images is only used as input data, and a model learns the data to represent lung images without diseases only. Fig. 4 shows the ground truth image used for learning (a) and the normal lung image generated by the PGGAN model that learned a normal distribution (b).

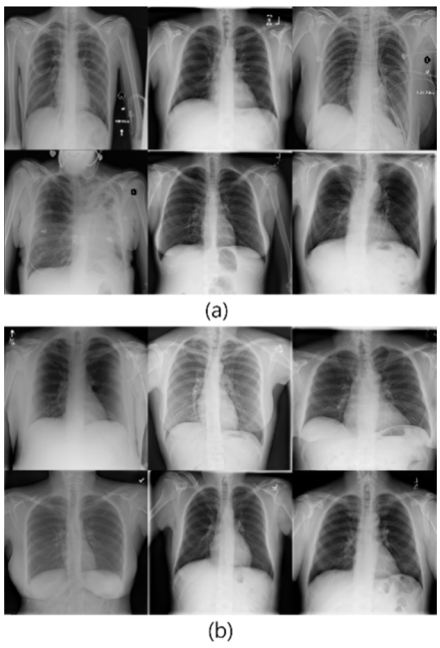


FIGURE 4. (a) Ground truth image used for learning and (b) normal lung image generated by the PGGAN model that learned the normal distribution.

As shown in Fig. 4, the structural features of the lung, such as its structure and size, are well represented, but the input image and the generated image fail to match each other.

To find a matching pair of an original image and a generated image, the feature extract layer of the discriminator is used to calculate the feature vector of an image. Fig. 5 shows the architecture to obtain a fake image matching the input image through the feature extractor.

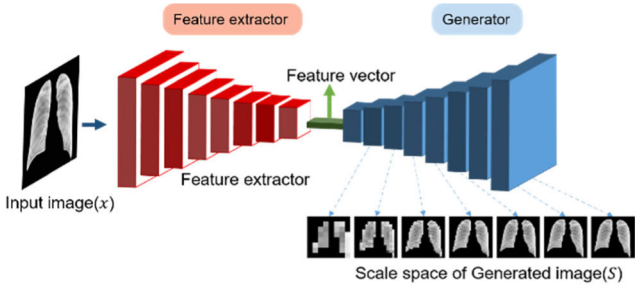


FIGURE 5. Architecture to obtain a fake image matching the input image through the feature extractor.

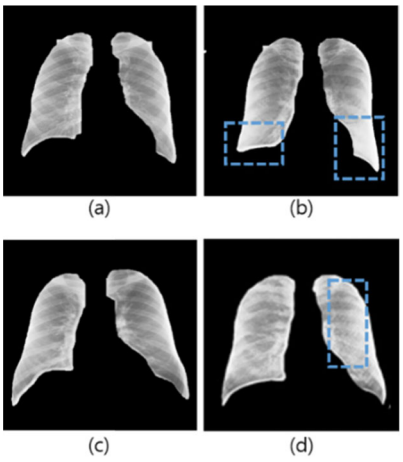


FIGURE 6. (a) and (c) show ground truth images of abnormal lungs; (b) and (d) present the images of abnormal lungs generated in the model.

Fig. 5 shows the seven-step progressive image generator used for classification of abnormal images. It is composed of the discriminator and generator network of the PGGAN trained with normal images. The generator creates images for seven types of resolutions from $2^3 \times 2^3$ (the smallest resolution in which the lung structure is identified) to $2^0 \times 2^0$. The latent feature vector of the input image(x) is obtained using the last feature extract layer of the discriminator. Based on the feature vector of x, the generator generates images with various resolutions, thereby obtaining the scale space of the image(x). With the application of such process, it is possible to analyze the similarities between images with multiple resolutions.

C. THRESHOLD DEFINITION USING WEIGHTED MULTI-SCALE SIMILARITY

To classify chest X-ray images with lesions, it is necessary to apply the model learned only with normal data. The trained PGGAN generates a false image using a query image as input data. Fig. 6 shows the results of the generation of abnormal lung images in the trained model.

Fig. 6 (a) and (c) show ground truth images of abnormal lungs; (b) and (d) present the images of abnormal lungs generated in the model. Fig. 6 shows the failure to generate differences in the chest structural deformation and contrast of the abnormal lung. In the case of such a disease diagnosis, the human perception ability was used to determine the structure and level of turbidity. Therefore, to detect lung defects, this study used an image similarity calculation technique based on human perception ability to define an anomaly. The image quality assessment method used in this study was the full-reference (FR) method, which calculates the similarity with the image generated in a GAN model using the entire information of an original image. Because a metric similar to the human perception ability was applied [37], the structural and perceptual features of the image generated by GAN were considered. Based on a scale space, the deformation of a fine region was also considered. Abnormal regions for detection, such as fine lesion tissues or medical equipment in the lung region, have diverse sizes. Accordingly, if any defect that is clearly different from that in an original image is found, it should be considered a significant region in anomaly classification regardless of size. In addition, regardless of whether the fine lung defect is outsized by other anomaly regions, its image should be clearly classified as an abnormal image. To do this, similarity analysis was conducted in multiscale. Each of the different scales appropriately presents the features of different objects in an image. The abnormal regions to detect include the deformation of organ structures, such as noticeable atelectasis and relatively small pulmonary nodules. Considering different image features depending on the scales, it is possible to find detailed representation features of an image. Appropriate features and objects are detected faster from low-resolution images than from high-resolution images. In addition, some features that are not found in a low-resolution image are obtained in a high-resolution image. After the images with progressively growing resolutions were obtained, defects in the images with low resolutions were detected. Subsequently, the coarse-to-fine method of verifying the defects progressively and accurately was applied. This method analyzes image features in diverse resolution solutions and provides a weight to consider detailed features of images. For the dataset that includes normal and abnormal images, the AS values of all images were drawn, and then the threshold for the classification of abnormal images was found through min-max scaling [38]. Equation (6) is the structural similarity assessment formula in weighted arithmetic mean-based multi-resolution:

$$\mathrm{AS} = 1 - \frac{\sum_{i=1}^{n} w_i \cdot S(x_i, \hat{x}_i)}{n} \tag{6}$$

x represents the ground-truth image of the dataset with normal and abnormal images. $\hat{x}$ the image generated based on v (latent vector) generated by the trained D (discriminator of PGGAN). n is the number of resolutions with the overall contour of the lung in the entire image(x) generated by trained G

(generator of PGGAN). Accordingly, if the input image size is 8×8, then $n = 1$. In this study, because an image with a size 512×512 was generated, $n = 7$. w adjusts the degree of contribution of each resolution to the similarity. Equation (7) is used to calculate the weight of each resolution:

$$S(x) = \frac{1}{1 + e^{x}}$$
$$w = S((r - \mu_r + C_r) \times \frac{S - \mu_s + C_s}{\delta_s}) \tag{7}$$

r means the resolution, μ_r represents the median of all resolutions, s represents $S(x_i, \hat{x_i})$, the SSIM operation result of a resolution, μ_s is the median of s, σ_s is the standard deviation of s. C is a positive integer number that prevents one of the two terms from converging to zero. In the image's scale space, row resolution loses pixels, so small anomalies such as nodules may not exist in the image [39], [40]. However, high-resolution images have little pixel loss, so nodules exist. The higher the resolution image, the subtler dissimilarity can be found. In this case, the influence of nodule dissimilarity on the arithmetic mean of SSIM in the entire scale space is minimal.

ALGORITHM 1. Function for calculating an AS of multi-scale image.

Algorithm 1 Anomaly Score Calculation Function in Multi Scale Image

Input: *Input image(x)*
Output: *Anomaly Score*
CALCULATE Anomaly Score
Create a list *RS* scale space of input x
Create a list *FS* scale space of Fake image $\hat{x}$ with regenerated image using $G(v)$
$v \leftarrow Extractfromlastto\ secondlayerinD(x)$
$S = []$
$R = []$
for res $= 8$ to *Res*
R $\leftarrow$ *Append* log_2 res
end for
for each *rs, fs* in the *RS, FS* **do**
$S \leftarrow Appendtheresultof$ (SSIM(rs, fs))
end for
for each *r,s* in the *R,S* **do**
$\alpha \leftarrow (r - \mu_r + C_1) \times (\frac{s - \mu_s + C_2}{\sigma_2})$
$w \leftarrow \frac{1}{1 + exp(-\alpha)}$
$Anomaly\_Score \leftarrow w * s + Anomaly\_Score$
end for
$Anomaly\_Score \leftarrow Anomaly\_Score/len(R)$
RETURN *Anomaly_Score*

This indicates that the smaller the size of the lesion, the less it is reflected in the Anomaly score. Therefore, it can be classified as normal compared to diseases in which dissimilarity is relatively large at all resolutions (e.g., organ deformation). The weights control the contribution of small anomalies to the anomaly score. By normalizing the image resolution and the SSIM score of the scale space, the influence of small lesions found at high resolution is increased. Since each resolution must be considered, the similarity analysis value cannot be zero. Therefore, the calculated value gets the final weight through the reverse sigmoid function. Algorithm 1 presents the function of calculating the anomaly score (AS) of a multiscale image. Algorithm 1 shows the function of drawing an anomaly of one input image using weighting. The resolutions considered for the input image range from a $2^3 \times 2^3$ resolution (in which a lung contour appears) to *Res* (the resolution of the input image). It is judged that as there is considerable fine distortion (texture loss, turbidity) for each resolution, there is more information for lesion image classification.

The importance of the resolution that has more information is reflected, and a weighted arithmetic mean is applied to increase the influence of a scale with a large distortion. In other words, the information amount of each resolution is reflected, and their contribution is adjusted.

IV. RESULT AND PERFORMANCE EVALUATION

The experimental environment for implementing the proposed anomaly classification was as follows: Ubuntu, Intel Skylake Xeon, NVIDIA Tesla V100 X 2(20RFLOPS), and RAM 128 GB. Tensorflow back-end engine was used as the software for the implementation. The training data for the model generating an image were only normal data. The model was trained using 50,361 normal image data. Two test datasets were generated and used for the threshold definition and validation. Each dataset consisted of 3,000 normal data and 3,000 abnormal data preprocessed by the proposed method in this study. Abnormal data include at least one disease. Based on all data ASs calculated using the aforementioned AS calculation method, a threshold was defined using the F1-score. Binary classification was performed based on the individual distribution of abnormal scores considering the threshold. Thus, images of the test dataset were regenerated, and a comparison with a conventional model for detecting lung defects was conducted [41]–[44]. As performance evaluation indexes, confusion matrix-based F1-score and AUC index were applied. The performance was evaluated using three methods. The first method consisted of using the weighted multi-scale similarity based on the decision boundary set through the F1-score. The second was to compare the proposed model with a classification model that includes a conventional generator. The third was to evaluate the excellence of the proposed algorithm in the ablation test based on the anomaly calculation method.

A. THRESHOLD DEFINITION USING WEIGHTED MULTI-SCALE SIMILARITY

For the binary classification of medical images in normal and abnormal categories, a threshold was defined. A randomly specified threshold would be inaccurate, and the

allowable range of normality could be significantly changed along with a shift in the threshold. Therefore, the F1-score was used to find an appropriate threshold. Fig. 7 shows the results from the AS distribution with the first test dataset.

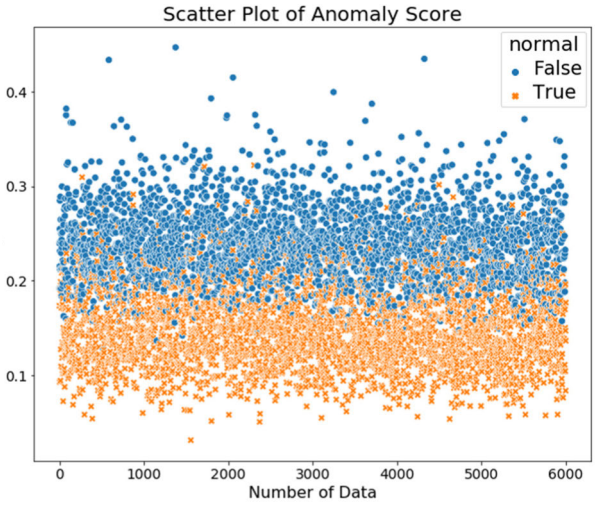


FIGURE 7. AS distribution of the first test dataset.

The orange points (×) in Fig. 7 represent normal images without any defect. Blue points (•) represent abnormal images. Through this distribution, it was possible to find the boundary between normal and abnormal images. To find an accurate threshold, the F1-score values were compared. Fig. 8 shows the F1-score as a result of abnormal image classification based on the threshold, such as to identify an effective threshold for abnormal image classification.

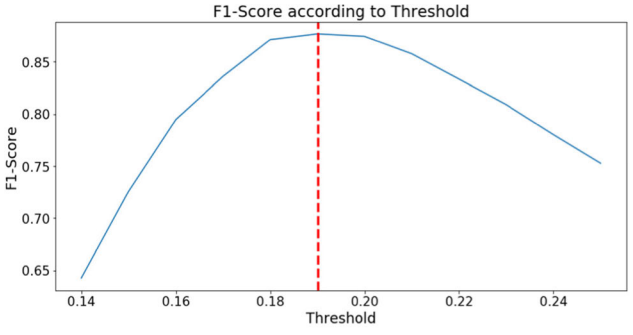


FIGURE 8. F1-score as a result of abnormal image classification based on the threshold.

When the threshold was 0.75, the image classification performance was the lowest; when it was 0.19, images were more effectively classified. Therefore, a threshold of 0.19, was applied to the classification of the second test dataset. Table 1 lists the confusion matrix through threshold classification.

TABLE 1. Confusion matrix through threshold classification.

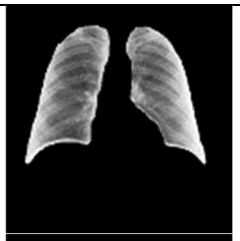
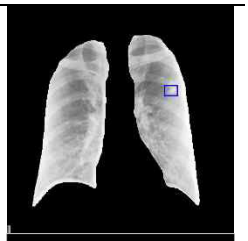
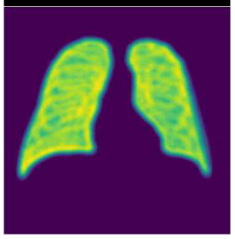
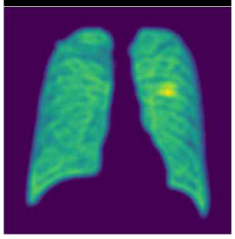
Real/Predicted	Positive	Negative
True	2,366	634
False	251	2,749

Based on the confusion matrix of Table 1, the following values were obtained: classification accuracy of 0.8525, recall of 0.9163, precision of 0.8125, and F1-score of 0.8524. In this study, data (True) whose values were higher than the threshold were considered to be normal data, and data (False) whose values were lower than the threshold were classified as data with lesions. Regarding incorrect classification, 634 normal data (10.56%) were classified as abnormal data, and 251 (4.18%) abnormal data were classified as normal data. In healthcare systems that require correct diagnosis and timely treatment, incorrect diagnosis is fatal to patients. Thus, the low rate of misdiagnosis in the proposed method is significant.

B. ANOMALY CLASSIFICATION RESULT ON CHEST X-RAY

The AS of a high-resolution image drawn using the proposed method was calculated. The anomaly classification was performed according to the threshold defined earlier. Table 2 presents the results of the proposed PGGAN-based anomaly classification. Table 2 presents the ground-truth and generated images. Low ASs calculated using the proposed technique indicate that the image is abnormal. Based on the threshold ($\theta = 0.81$) defined earlier, the images were classified as normal and abnormal. In the case of normal images, the organ structure and detailed tissues were well presented. In the case of abnormal images (nodule), the nodule in the boundary box of the ground truth image was not presented well. The resolution-by-resolution generator group progressively generating fake image learns the entire data distribution and generates an image using the latent vector of the input image. It retains the main features of the query image and regenerates an image based on the distribution of different image data. Therefore, distortion can be found between the images generated by the generator and the ground truth image. As for normal images, as the generator was trained with normal images, it effectively represents structural similarities. In contrast, for abnormal images, the generator fails to represent a structural shape and causes considerable distortion. In such a case, distortion (or the reflection of a different image feature), rather than the lesion region to detect, can influence the AS. For this reason, through the generator trained along with progressively growing resolutions, it is necessary to find scale-invariant features and give a different weight depending on the information amount of the image. Along with the learning process of PGGAN, the distortion influence of the fake image shape is reduced, and the influence of the fine lesion region not generated is improved. In addition, through the GAN model considering progressively growing resolutions, it is possible to consider the features of medical images with high resolutions efficiently.

TABLE 2. Results of the proposed PGGAN based anomaly classification.

Category	Normal	Abnormal (nodule)
Ground Truth Image (x)		
Fake Image ($\hat{x}$)		
Lesion between x and $\hat{x}$		
Anomaly Score	0.1096	0.2361

C. EVALUATION OF MODEL FIT THROUGH COMPARISON WITH EXISTING STUDIES

In the first goodness-of-fit assessment, the abnormal-chest X-ray classifier of conventional models was compared with the proposed model based on the AUC score in terms of accuracy. The models compared were a U-net Auto Encoder-based model (one-classification SOTA model for NIH dataset [9]), an encoder-based model, and a CNN-based model with VGG16. These models were used to classify 64 × 64 and 1024 × 1024 images [13], [43]. Table 3 shows the classification AUC results of the conventional one-classification models and the proposed model.

TABLE 3. Classification AUC results according to the resolution of the existing model and the proposed model.

Image size/Method	64×64
U-net Auto Encoder	0.6870
U-net Auto Encoder + Discriminator + Second Encoder	0.8410
Our proposal	0.8507

The models in Table 3 downsampled the chest X-ray dataset of the NIH and then classified the images. In the experimental environment of the NVIDIA TITAN Xp GPU, the PyTorch framework was used for classification. To prevent information loss at the time of down-sampling, the aim was to maintain the features of images with high resolutions through the skin-connection of the Auto Encoder structure. However, because features were trained in a DCGAN-based model, no high performance was achieved at high resolutions. On the contrary, the proposed model calculated the AS through multiple scales to find detailed features of lesions at high resolutions. In the comparative experiment, an image with a low resolution of 64 × 64 was used as the input image. In short, a smaller scale space than that with 512 × 512 was considered. Therefore, the classifier in the proposed method showed relatively lower accuracy owing to image feature loss, but produced high performance at both low and high resolutions.

For comparison of classification results of higher resolution images (1024 × 1024), proposed classifier was compared with a CNN-based classifier using VGG16. Fig. 9 shows the classification results of the CNN-based model with VGG16 and the proposed model.

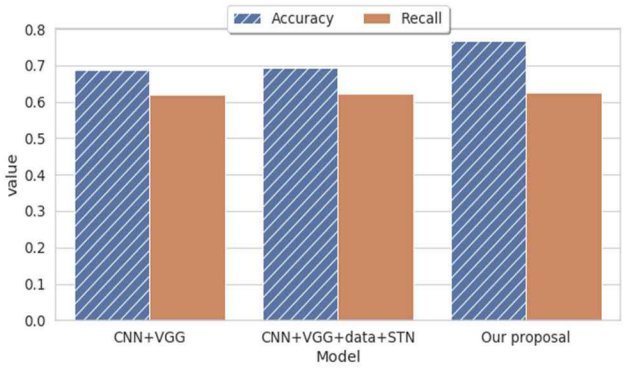


FIGURE 9. Classification results of the CNN-based model with VGG16 and the proposed model.

The CNN based models in Fig 9 was used without changing the resolution of the original data. Therefore, it is possible to classify images without losing the high-resolution features of the data. As a result of comparing the proposed model and the CNN-based model, the performance of recall was similar at about 0.6. However, there was a meaningful difference in accuracy. This means that our model can classify the anomalies properly more than the compared model. In the healthcare system, the timely judgment of abnormalities is effective for preventing disease and treatment of patients. Thus, our model is relatively suitable.

In the second goodness-of-fit assessment, the accuracy of the proposed anomaly classification was evaluated in an ablation study. In the AS calculation process, each factor influences classification performance. Table 4 presents the performance of the classifier in the proposed anomaly classification (ACM, anomaly classification model; S, structural similarity index metric; M, multi-scale image; W, weighted score). In Table 4, ACM_S classified abnormal data using the structural similarity of images with high resolutions. As it compared images in one resolution, it failed to consider invariant features. Accordingly, ACM_S showed the degree of distortion of the GAN-generated image and its original image, rather than the regions that discriminate legions.

TABLE 4. Performance of classifiers in the proposed anomaly classification.

Calculation method	Accuracy
ACM_S	0.593 ± 0.062
ACM_MS	0.8392 ± 0.035
ACM_WMS (ours)	**0.8592 ± 0.012**

ACM_MS can find invariant features of images at diverse resolutions using multi-scale similarity. However, because it does not consider the amount of information in each resolution for anomaly classification, it failed to find non-similarities caused by fine lesions. Therefore, although it was effective at detecting a relatively wide abnormal region, such as the deformation of organ structures, it did not discriminate a fine lesion. To discriminate fine lesions, the model proposed in this study uses weighted multi-scale similarity in consideration of resolutions and information amount. The proposed ACM_WMS model through repeated experiments showed the highest classification performance with 0.8592. Accordingly, we prove the effectiveness of anomaly classification using the weighted multi-similarity in this study.

V. CONCLUSION

Data imbalance is a common problem in medical images, caused by the quantity difference between normal data (without lesions) and abnormal data (with lesions), and requires excess resource consumption due to annotation. In addition, high-resolution data must be considered to detect the features of fine lesions. To overcome these problems with chest X-ray images and classify anomalies, this study proposed a chest X-ray anomaly classification model using PGGAN-based weighted multi-scale similarity. The proposed method uses the PGGAN-based anomaly classification model trained with normal data. The anomaly classification model using semi-supervised learning is unstable and cannot generate images with high resolution. PGGAN enables to achieve stable learning and generate high-resolution images through progressive layer learning. Additionally, it does not require the creation of a scale space for the analysis of image similarity. As the PGGAN model is trained along with growing resolutions in a latent space, the multi-resolution information of images is provided. Accordingly, it is possible to obtain a scale space and compare diverse features. By segmenting the lung in a chest X-ray image and using the segmented data as training data, it is possible to minimize the influence of data noise on the learning result. Failure to learn the distribution of abnormal data leads to a failure to learn the features of images with lesions and a failure to represent the structural features of abnormal images. Therefore, the multi-scale structure similarity of drawing a similarity in consideration of human perception ability was multiplied by the information amount of each resolution, and thereby an anomaly was defined. A threshold used for classifying abnormal lung images through the defined AS distribution was chosen based on the F1-score. When the test dataset was classified based on the defined threshold ($\theta = 0.81$), the accuracy rate was 85.2%. In addition, the proposed classifier drew the AUC score (0.8507), which was improved compared to that of the semi-supervised anomaly classifier. In other words, it obtained stable learning results for high-resolution images. In an ablation study, the anomaly calculation method considering the image information amount in multi-scale enabled to consider the detailed information of high-resolution images. Thus, the proposed semi-supervised learning-based model, as an automated assisted diagnosis tool, contributes to the development of healthcare systems. In future work, the model will be expanded to a multi-classification model that can judge not only the existence of diseases but also the types of various diseases and locations. Accordingly, it can provide an improved auxiliary opinion for disease diagnosis of healthcare providers.

REFERENCES

[1] W. Yang, Y. Chen, Y. Liu, L. Zhong, G. Qin, Z. Lu, Q. Feng, and W. Chen, "Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain," *Med. Image Anal.*, vol. 35, pp. 421–433, Jan. 2017.

[2] G. Wang, W. Li, M. A. Zuluaga, R. Pratt, P. A. Patel, M. Aertsen, T. Doel, A. L. David, J. Deprest, S. Ourselin, and T. Vercauteren, "Interactive medical image segmentation using deep learning with image-specific fine tuning," *IEEE Trans. Med. Imag.*, vol. 37, no. 7, pp. 1562–1573, Jul. 2018.

[3] S.-E. Ryu, D.-H. Shin, and K. Chung, "Prediction model of dementia risk based on XGBoost using derived variable extraction and hyper parameter optimization," *IEEE Access*, vol. 8, pp. 177708–177720, 2020.

[4] H. Yoo, S. Han, and K. Chung, "Diagnosis support model of cardiomegaly based on CNN using ResNet and explainable feature map," *IEEE Access*, vol. 9, pp. 55802–55813, Mar. 2021.

[5] J.-W. Baek and K. Chung, "Context deep neural network model for predicting depression risk using multiple regression," *IEEE Access*, vol. 8, pp. 18171–18181, 2020.

[6] J.-C. Kim and K. Chung, "Multi-modal stacked denoising autoencoder for handling missing data in healthcare big data," *IEEE Access*, vol. 8, pp. 104933–104943, 2020.

[7] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," Jun. 2017, *arXiv:1702.05747*. [Online]. Available: http://arxiv.org/abs/1702.05747

[8] H. Wang and Y. Xia, "ChestNet: A deep neural network for classification of thoracic diseases on chest radiography," 2018, *arXiv:1807.03058*. [Online]. Available: http://arxiv.org/abs/1807.03058

[9] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," Dec. 2017, *arXiv:1705.02315*. [Online]. Available: http://arxiv.org/abs/1705.02315

[10] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[11] 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.

[12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015, pp. 1–14.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015, *arXiv:1512.03385*. [Online]. Available: http://arxiv.org/abs/1512.03385

[14] A. I. Aviles-Rivero, N. Papadakis, R. Li, P. Sellars, Q. Fan, R. T. Tan, and C. B. Schönlieb, "GraphX-Net chest X-ray classification under extreme minimal supervision," 2019, *arXiv:1907.10085*. [Online]. Available: https://arxiv.org/abs/1907.10085

[15] Y. Rao, Y. Wang, F. Meng, J. Pu, J. Sun, and Q. Wang, "A symmetric fully convolutional residual network with DCRF for accurate tooth segmentation," *IEEE Access*, vol. 8, pp. 92028–92038, 2020.

[16] Y.-X. Tang, Y.-B. Tang, M. Han, J. Xiao, and R. M. Summers, "Abnormal chest X-ray identification with generative adversarial one-class classifier," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1358–1361.

[17] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2015, pp. 234–241.

[18] A. Madani, M. Moradi, A. Karargyris, and T. Syeda-Mahmood, "Semi-supervised learning with generative adversarial networks for chest X-ray classification with ability of data domain adaptation," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 1038–1042.

[19] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015.

[20] A. Madani, R. Arnaout, M. Mofrad, and R. Arnaout, "Fast and accurate classification of echocardiograms using deep learning," 2017, *arXiv:1706.08658*. [Online]. Available: http://arxiv.org/abs/1706.08658

[21] M. Islam, B. Siddique, S. Rahman, and T. Jabid, "Image recognition with deep learning," in *Proc. Int. Conf. Intell. Inform. Biomed. Sci. (ICIIBMS)*, 2018, pp. 585–590.

[22] H. Liu, J. Xu, Y. Wu, Q. Guo, B. Ibragimov, and L. Xing, "Learning deconvolutional deep neural network for high resolution medical image reconstruction," *Inf. Sci.*, vol. 468, pp. 142–154, Nov. 2018.

[23] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*. [Online]. Available: http://arxiv.org/abs/1511.06434

[24] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," 2017, *arXiv:1710.10196*. [Online]. Available: http://arxiv.org/abs/1710.10196

[25] B. Zhou, S. Liu, B. Hooi, X. Cheng, and J. Ye, "BeatGAN: Anomalous rhythm detection using adversarially generated time series," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, Aug. 2019, pp. 4433–4439.

[26] T. K. Yoo, J. Y. Choi, H. K. Kim, I. H. Ryu, and J. K. Kim, "Adopting low-shot deep learning for the detection of conjunctival melanoma using ocular surface images," *Comput. Methods Programs Biomed.*, vol. 205, Jun. 2021, Art. no. 106086, doi: 10.1016/j.cmpb.2021.106086.

[27] T. Lindeberg, "Scale-space theory: A basic tool for analysing structures at different scales," *J. Appl. Statist.*, vol. 21, nos. 1–2, pp. 224–270, Jun. 1994.

[28] P. E. Robinson and Y. Roodt, "Blind deconvolution of Gaussian blurred images containing additive white Gaussian noise," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Feb. 2013, pp. 1092–1097.

[29] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, vol. 2, Nov. 2003, pp. 1398–1402.

[30] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[31] S. Reza, O. B. Amin, and M. M. A. Hashem, "TransResUNet: Improving U-Net architecture for robust lungs segmentation in chest X-rays," in *Proc. IEEE Region Symp. (TENSYMP)*, Jun. 2020, pp. 1592–1595.

[32] F. D. Mattia, P. Galeone, M. D. Simoni, and E. Ghelfi, "A survey on gans for anomaly detection," Jun. 2019, *arXiv:1906.11632*. [Online]. Available: https://arxiv.org/abs/1906.11632

[33] S. Akcay, A. Atapour-Abarghouei, and T. P. Breckon, "GANomaly: Semi-supervised anomaly detection via adversarial training," 2018, *arXiv:1805.06725*. [Online]. Available: http://arxiv.org/abs/1805.06725

[34] Z.-Q. Wang and X.-D. Yan, "Multi-scale feature extraction algorithm of ear image," in *Proc. Int. Conf. Electr. Inf. Control Eng. (ICEICE)*, Apr. 2011, pp. 528–531.

[35] M. S. Artemyev, A. R. Serazetdinov, and A. A. Smirnov, "X-ray image segmentation with the use of machine learning algorithms," *AIP Conf.*, vol. 2313, no. 1, 2020, Art. no. 080006.

[36] T. Schlegl, P. Seeböck, S. M. Waldstein, U. Schmidt-Erfurth, and G. Langs, "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery," in *Proc. Int. Conf. Inf. Process. Med. Imag.* Cham, Switzerland: Springer, 2017, pp. 146–157.

[37] G. Zhai and X. Min, "Perceptual image quality assessment: A survey," *Sci. China Inf. Sci.*, vol. 63, no. 11, pp. 211301–211352, Nov. 2020.

[38] D.-H. Shin, R. C. Park, and K. Chung, "Decision boundary-based anomaly detection model using improved AnoGAN from ECG data," *IEEE Access*, vol. 8, pp. 108664–108674, 2020.

[39] K. Shankar, E. Perumal, P. Tiwari, M. Shorfuzzaman, and D. Gupta, "Deep learning and evolutionary intelligence with fusion-based feature extraction for detection of COVID-19 from chest X-ray images," *Multimedia Syst.*, pp. 1–13, 2021.

[40] V. Chouhan, S. K. Singh, A. Khamparia, D. Gupta, P. Tiwari, C. Moreira, R. Damaševičius, and V. H. C. de Albuquerque, "A novel transfer learning based approach for pneumonia detection in chest X-ray images," *Appl. Sci.*, vol. 10, no. 2, p. 559, Jan. 2020.

[41] Y. Chaudhary, M. Mehta, R. Sharma, D. Gupta, A. Khanna, and J. J. P. C. Rodrigues, "Efficient-CovidNet: Deep learning based COVID-19 detection from chest X-ray images," in *Proc. IEEE Int. Conf. E-Health Netw., Appl. Services (HEALTHCOM)*, Shenzhen, China, Mar. 2021, pp. 1–6.

[42] A. K. Jaiswal, T. Prayag, K. Sachin, G. Deepak, K. Ashish, and J. R. Joel, "Identifying pneumonia in chest X-rays: A deep learning approach," *Measurement*, vol. 145, pp. 511–518, Oct. 2019.

[43] S. Tripathi, S. Shetty, S. Jain, and V. Sharma, "Lung disease detection using deep learning," *Int. J. Innov. Technol. Exploring Eng.*, vol. 10, no. 8, 2021, doi: 10.35940/ijitee.H9259.0610821.

[44] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2223–2232.

HYE-JEONG KWON received the B.S. degree from the Division of Computer Science and Engineering, Kyonggi University, Suwon, South Korea, in 2021. She is currently pursuing the master's degree with the Department of Computer Science, Kyonggi University. She worked as a Researcher with the Data Mining Laboratory, Kyonggi University. Her research interests include data mining, artificial intelligence, mobile healthcare, biomedical and health informatics, information systems, deep learning, machine learning, emerging health risk mining, and medical data analysis.

DONG-HOON SHIN received the B.S. degree from the Department of Computer Engineering, Dongseo University, South Korea, in 2019, and the M.S. degree from the Department of Computer Science, Kyonggi University, Suwon, South Korea, in 2021. He has been a Researcher with the Data Mining Laboratory, Kyonggi University. His research interests include data mining, artificial intelligence, mobile healthcare, biomedical and health informatics, knowledge systems, VR/AR, emerging health risk mining, and deep learning.

KYUNGYONG CHUNG received the B.S., M.S., and Ph.D. degrees from the Department of Computer Information Engineering, Inha University, South Korea, in 2000, 2002, and 2005, respectively. He worked with the Software Technology Leading Department, Korea IT Industry Promotion Agency (KIPA). From 2006 to 2016, he was a Professor with the School of Computer Information Engineering, Sangji University, South Korea. Since 2017, he has been a Professor with the Division of AI Computer Science and Engineering, Kyonggi University, South Korea. His research interests include data mining, artificial intelligence, healthcare, knowledge systems, HCI, and recommendation systems. He was named as a Highly Cited Researcher by Clarivate Analytics, in 2017.