Original Article | Thoracic Imaging

eISSN 2005-8330
https://doi.org/10.3348/kjr.2021.0146
Korean J Radiol 2022;23(1):139-149

Korean Journal of Radiology
KJR

Bone Suppression on Chest Radiographs for Pulmonary Nodule Detection: Comparison between a Generative Adversarial Network and Dual-Energy Subtraction

Kyungsoo Bae[1, 2], Dong Yul Oh[3], Il Dong Yun[4], Kyung Nyeo Jeon[1, 2]

[1]Department of Radiology, Institute of Health Sciences, Gyeongsang National University School of Medicine, Jinju, Korea; [2]Department of Radiology, Gyeongsang National University Changwon Hospital, Changwon, Korea; [3]Monitor Corporation, Seoul, Korea; [4]Division of Computer and Electronic System Engineering, Hankuk University of Foreign Studies, Yongin, Korea

Objective: To compare the effects of bone suppression imaging using deep learning (BSp-DL) based on a generative adversarial network (GAN) and bone subtraction imaging using a dual energy technique (BSt-DE) on radiologists' performance for pulmonary nodule detection on chest radiographs (CXRs).

Materials and Methods: A total of 111 adults, including 49 patients with 83 pulmonary nodules, who underwent both CXR using the dual energy technique and chest CT, were enrolled. Using CT as a reference, two independent radiologists evaluated CXR images for the presence or absence of pulmonary nodules in three reading sessions (standard CXR, BSt-DE CXR, and BSp-DL CXR). Person-wise and nodule-wise performances were assessed using receiver-operating characteristic (ROC) and alternative free-response ROC (AFROC) curve analyses, respectively. Subgroup analyses based on nodule size, location, and the presence of overlapping bones were performed.

Results: BSt-DE with an area under the AFROC curve (AUAFROC) of 0.996 and 0.976 for readers 1 and 2, respectively, and BSp-DL with AUAFROC of 0.981 and 0.958, respectively, showed better nodule-wise performance than standard CXR (AUAFROC of 0.907 and 0.808, respectively; $p \leq 0.005$). In the person-wise analysis, BSp-DL with an area under the ROC curve (AUROC) of 0.984 and 0.931 for readers 1 and 2, respectively, showed better performance than standard CXR (AUROC of 0.915 and 0.798, respectively; $p \leq 0.011$) and comparable performance to BSt-DE (AUROC of 0.988 and 0.974; $p \geq 0.064$). BSt-DE and BSp-DL were superior to standard CXR for detecting nodules overlapping with bones ($p < 0.017$) or in the upper/middle lung zone ($p < 0.017$). BSt-DE was superior ($p < 0.017$) to BSp-DL in detecting peripheral and sub-centimeter nodules.

Conclusion: BSp-DL (GAN-based bone suppression) showed comparable performance to BSt-DE and can improve radiologists' performance in detecting pulmonary nodules on CXRs. Nevertheless, for better delineation of small and peripheral nodules, further technical improvements are required.

Keywords: *Chest radiography; Deep learning; Generative adversarial network; Pulmonary nodules; Bone suppression imaging*

INTRODUCTION

Chest radiography is the most frequently used diagnostic imaging modality in routine clinical practice. It is performed not only for patients with thoracic disease but also for others during a routine checkup [1]. Findings on chest radiographs (CXRs) often guide subsequent evaluations, such as laboratory tests, additional imaging studies, and tissue sampling. Detection of nodules on CXR is an important task for radiologists because many malignant and benign conditions can manifest as nodular lesions [2]. However, the risk of detecting errors on two-dimensional (2D) images, is relatively high because of the low contrast resolution and overlapping anatomical structures.

As per previous studies related to missed lung cancer, overlapping bony structures are major factors that cause

Received: February 19, 2021 **Revised:** August 4, 2021
Accepted: August 17, 2021
Corresponding author: Kyung Nyeo Jeon, MD, PhD, Department of Radiology, Institute of Health Sciences, Gyeongsang National University School of Medicine, 11 Samjeongja-ro, Seongsan-gu, Changwon 51472, Korea.
• E-mail: knjeon@gnu.ac.kr

errors in the diagnosis of CXRs [3-6]. To improve the visibility of parenchymal lesions on CXRs, bony structures need to be eliminated using advanced techniques. The dual-energy subtraction (DES) technique can remove bones from images using differences in the degree to which body tissues attenuate low- and high-energy photons. Despite its positive effect on the reader's performance, the DES technique has not been widely used owing to issues such as its limited portability, increased radiation exposure, and specific hardware requirements [7-12]. Additionally, cardiac and respiratory motion artifacts can degrade image quality procured using a dual-exposure technique [9]. With recent advances in deep learning, software-based bone suppression algorithms have been introduced [13,14]. Application of a software-based bone suppression technique on chest radiography can improve the diagnostic accuracy for detecting nodules by radiologists and computer-aided diagnosis systems without increasing radiation exposure or requiring specialized equipment [15-19]. However, further technical advances are required for their practical use in patients.

We developed a new bone suppression model based on the wavelet transform and generative adversarial networks (GANs) [20]. Although the performance of this model was evaluated by comparing quality metrics with those of other convolutional neural network-based models, its potential for clinical application has not yet been assessed. In addition, the performance of software-based bone suppression images compared to that of bone subtraction images using the dual energy technique remains unclear. Therefore, the purpose of the present study was to evaluate the effect of bone suppression imaging using deep learning (BSp-DL) based on GAN compared to bone subtraction imaging using a dual energy technique (BSt-DE) on radiologists' performances for pulmonary nodule detection on CXR.

MATERIALS AND METHODS

This study was approved by our Institutional Review Board. The written informed consent was waived because of the retrospective study design (IRB No. 2019-08-018).

Patients and Image Selection

Between April 2016 and December 2019, a total of 1858 consecutive patients underwent chest radiography using the dual energy technique. Among them, 194 patients had chest CT images taken within 7 days of dual-energy radiographic imaging. With chest CT as a reference, two board-certified radiologists with 24 and 25 years of experience, respectively, reviewed CXRs and determined the presence or absence of non-calcified nodules. After excluding 83 cases, 111 patients were finally enrolled for the evaluation. The reasons for exclusion were as follows: presence of extensive infiltrative lesions (n = 36), severe parenchymal distortion (n = 15), lung atelectasis (n = 7), prior lung surgery (n = 4), large amount of pleural effusion or diffuse pleural calcification (n = 13), and nodules apparent on CT but not observed on CXRs (n = 8). The image selection process is illustrated in Figure 1. After reviewing axial and coronal CT scan images, annotations were made in the center of 83 nodules on CXRs to serve as standard reference to evaluate readers' performance. When a patient underwent multiple dual energy chest radiographic imaging, only one study that was performed at the nearest time to CT was used for the evaluation. The study did not include any image datasets used in the bone suppression model development process or internal validation.

Image Acquisition

Posteroanterior chest radiography was performed using a flat-panel detector digital radiographic imaging system (Discovery XR656, GE Healthcare) with dual energy capability. The detector had an image size of 41 x 41 cm and a pixel dimension of 0.2 x 0.2 mm. For dual-energy image acquisition, there were two exposures using 120 and 60 kVp pulses at intervals of 200 ms. Fixed tube currents of 194 mA and 252 mA were used at high and low voltage values, respectively. The exposure time was automatically controlled according to the patient size. Each acquisition resulted in three images: a composite image (equivalent standard CXR), a soft tissue selective image (BSt-DE), and a bone-selective image. Bone or soft-tissue images were produced by weighted subtraction of each absorption coefficient. CXR and BSt-DE were anonymized and saved for review.

Image Processing Using Bone Suppression Model

Our bone suppression model used adversarial training in the GAN framework to learn the conditional probability distribution of the output images according to the input images. The Haar wavelet decomposition was adopted as an input system. It pre-defined features that the network should learn via wavelet transformation of an image into four directional feature images (Fig. 2). This allowed the model

https://doi.org/10.3348/kjr.2021.0146

Korean Journal of Radiology
KJR

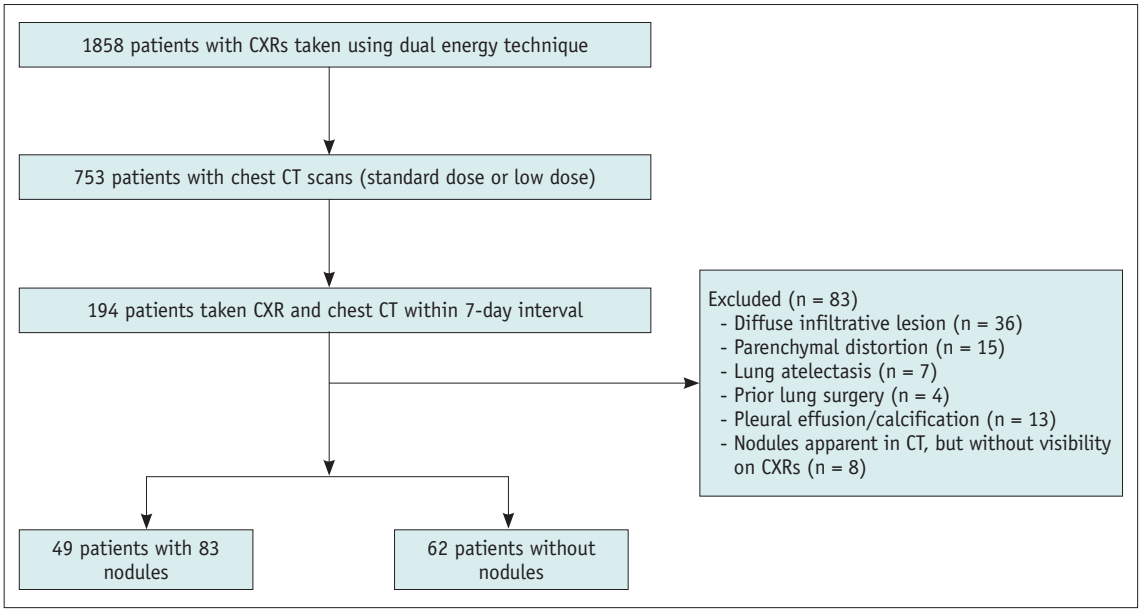

Fig. 1. Flow chart depicting the selection of the study population. CXR = chest radiograph

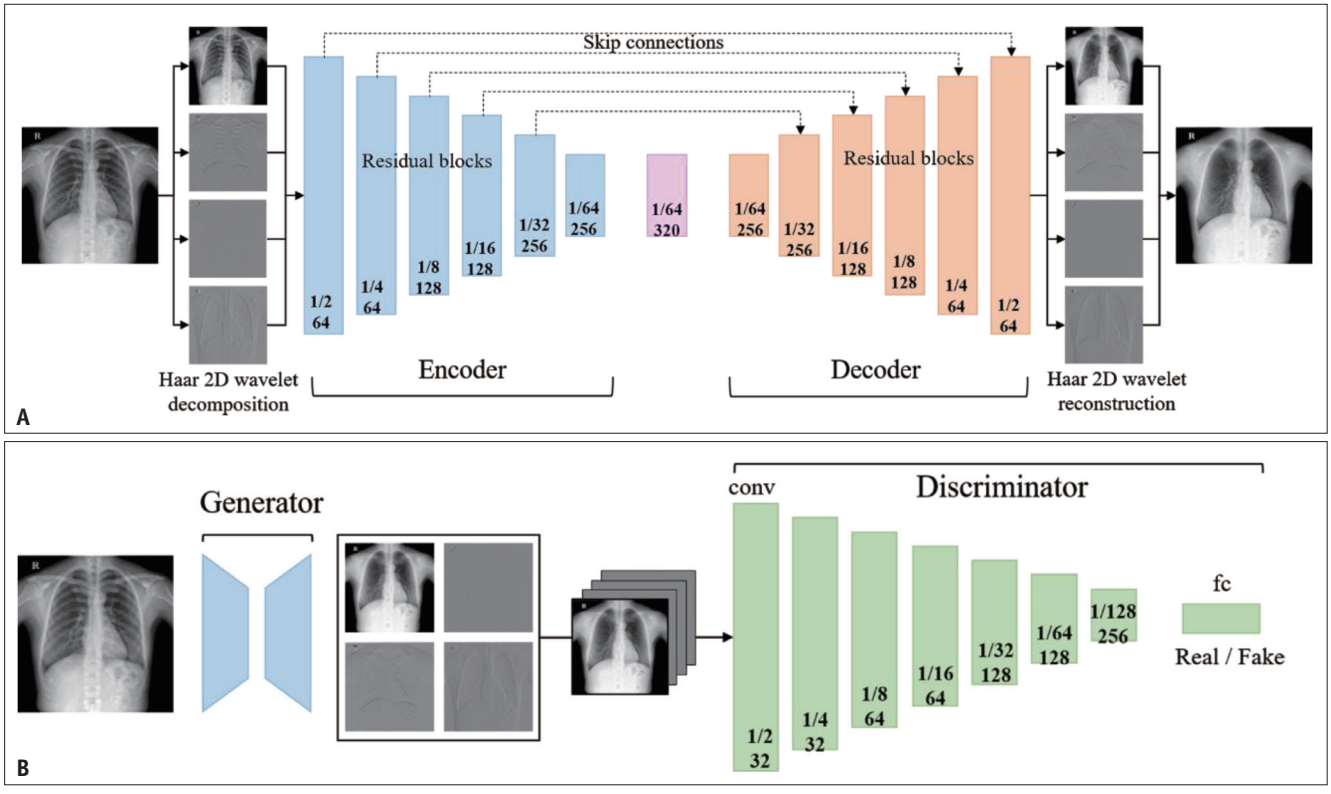

Fig. 2. The architecture of bone suppression algorithm based on wavelet transform and generative adversarial networks.
A. The architecture of generator that receives the original image and produces bone suppressed images. This system uses the frequency information obtained from Haar wavelet transformation which pre-defines features that the network should learn, allowing the network to converge more quickly and efficiently. The generator takes four channels of 512 x 512 source (original) images obtained through Haar 2D wavelet decomposition and tries to produce output (bone suppressed) image that can fool the discriminator by avoiding image blurring. The output image is finally reconstructed to 1024 x 1024 by Haar 2D wavelet reconstruction. Values below each conv block are image compression ratios with the number of channels. **B.** The architecture of the discriminator. The discriminator distinguishes whether the input is a fake image that comes from the generator, or a real one from the training set. In the training process, the two networks compete. The discriminator plays a critical role in preventing the generator from producing a blurred image and also considers the distribution of image batches s as the generator does. conv = convolution, fc = fully connected, 2D = two-dimensional

https://doi.org/10.3348/kjr.2021.0146

to exploit high-frequency details of CXR and to converge more quickly and efficiently [21]. In the GAN framework, the generator attempted to deceive the discriminator by creating an image similar to the training set. This bone suppression model was trained and validated using a total of 348 pairs of composite and soft tissue selective images obtained by dual-energy radiography that were available in public domain [20]. Technical details of the model, including network architecture, quantitative performance metrics, and training process, have been described previously [20]. No additional training was imparted for this study. A total of 111 CXRs obtained using the dual energy technique were postprocessed. The generated bone-suppressed images (BSp-DL) were anonymized and stored for evaluation.

Evaluation of Image Quality

Two radiologists (with 6 and 8 years of experience, respectively) independently assessed the image quality of BSp-DL and BSt-DE in two reading sessions. In each session, the images were randomly re-ordered. With reference to a previous study [22], the evaluation criteria included visibility of pulmonary vessels in the lung field, retrocardiac and diaphragm area, visibility of the central airway, and degree of bone suppression. The evaluation criteria and their scales are listed in Supplementary Table 1.

Nodule Detection

Before an observer study, a preliminary study was conducted with images taken using an anthropomorphic chest phantom with synthetic nodules for readers' training (Supplementary Fig. 1). Two board-certified radiologists (with 5 and 7 years of experience, respectively) who were blinded to the diagnosis independently evaluated 333 image sets for detecting pulmonary nodules in three reading sessions at two-week intervals.

In the first session, standard CXRs were displayed on a 12-megapixel wide monitor (Coronis Uniti, Barco Healthcare) for evaluation. In the second and third sessions, standard CXR and corresponding BSt-DE or BSp-DL were displayed side by side on the same monitor. In each reading session, the images were randomly re-ordered. Readers were asked to mark nodules if any, and record a score of 1–4 (very low to very high confidence). A score of 0 was assigned to unmarked lesions and normal images without false positive marks. Calcified nodules and masses > 30 mm were ignored. When there were multiple nodules in one lung zone, up to three nodules in order of their sizes (from the largest to the smallest) were evaluated.

Lesion localizations determined by readers were considered correct if they were marked within 10 mm of the reference standard. The location of a nodule was determined according to both the longitudinal and transverse divisions of the lung fields on CXRs. The vertical distance of each lung from the apex to the costophrenic angle was divided into three equal parts: right upper (RU), right middle (RM), right lower (RL), left upper (LU), left middle (LM), and left lower (LL) zones. Horizontally, the outer half of the lung parenchyma from the chest wall was defined as the peripheral zone, and the remaining inner portion was designated as the central zone. Nodules were considered to overlap with bony structures when 50% or more of the nodular area overlapped with the ribs and/or clavicles.

Statistical Analysis

Alternative free-response receiver operating characteristic (AFROC) curve analysis was performed to compare nodule-wise localization performances of CXR, BSt-DE with CXR, and BSp-DL with CXR, while ROC analyses were employed to compare person-wise diagnostic performances. In the person-wise analysis, the highest diagnostic confidence score for each image set was used for calculation [23].

In addition, sensitivity, specificity for person-wise detection, and sensitivity for nodule-wise detection were evaluated by assigning scores of 2–4 as positive. Subgroup analyses for nodule-wise detection according to the size, location, and presence or absence of overlapping bones were also undertaken. The McNemar test was used to compare the sensitivity and specificity.

Inter-reader agreements on person-wise and nodule-wise diagnoses were determined based on the weighted kappa coefficient, which was interpreted as follows: a kappa value of < 0.20, poor; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; and < 0.81, almost perfect.

Following Bonferroni correction for multiple comparisons, a p value of < 0.017 was considered statistically significant. AFROC analysis was performed using RJafroc (Artificial Intelligence Systems and Observer Performance; R package version 2.0.1, R Foundation of Statistical Imaging). Other statistical analyses were undertaken using commercially available software (MedCalc, version 19.5, MedCalc Software).

RESULTS

Among the 111 subjects, there were 66 males and 45

https://doi.org/10.3348/kjr.2021.0146

Korean Journal of Radiology
KJR

females. The mean age was 54.7 years (range, 18–94 years). According to standard reference, 49 patients had 83 lung nodules with a mean diameter of 13.24 ± 6.80 mm (range, 5–30 mm). However, 62 patients had no lung nodules. The location of nodules in the lung zone was RU for 20, RM for 15, RL for 18, LU for 10, LM for 7, and LL for 13. While 31 nodules were located in the central area, 52 were in the peripheral area. Fifty-five nodules overlapped with bones, whereas 28 nodules did not. Patient demographics are shown in Table 1.

Image Quality

The total image quality score by two independent readers did not differ between BSt-DE and BSp-DL (Table 2). While the visibility of pulmonary vessels in the lung field was superior ($p < 0.001$) in BSt-DE as compared to BSp-DL, the visibility of pulmonary vessels in the retrocardiac and diaphragm areas was superior ($p < 0.001$) in BSp-DL than in BSt-DE (Fig. 3). The visibility of the central airway and degree of bone suppression were not significantly different between the two image datasets.

Table 1. Characteristics of Study of Patients and Nodules

Male:female	66:45
Mean age, years (range)	54.7 (18–94)
Patient with vs. without nodules	49:62
Location of nodules (n = 83)	
RU, RM, RL, LU, LM, LL	20, 15, 18, 10, 7, 13
Central vs. peripheral	31:52
Overlapping vs. non-overlapping with bones	55:28
Size of nodules, mean ± standard deviation (range)	13.2 ± 6.8 (5–30)
5–10 mm	30
10–20 mm	39
20–30 mm	14

Data are number of patients or nodules, unless specified otherwise. LL = left lower, LM = left middle, LU = left upper, RL = right lower, RM = right middle, RU = right upper

Nodule Detection

The use of BSt-DE {reader 1: area under the AFROC curve (AUAFROC) = 0.996 (95% confidence interval [CI]: 0.989, 1.000), $p < 0.001$; reader 2: AUAFROC = 0.976 (95% CI: 0.954, 0.997), $p < 0.001$} or BSp-DL (reader 1: AUAFROC = 0.981 [95% CI: 0.964, 0.997], $p = 0.005$; reader 2: AUAFROC = 0.958 [95% CI: 0.926, 0.990], $p < 0.001$) showed better nodule-by-nodule diagnostic performance than the use of CXR alone (reader 1: AUAFROC = 0.907 [95% CI: 0.858, 0.955]; reader 2: AUAFROC = 0.808 [95% CI: 0.733, 0.883]) (Table 3, Fig. 4). There were no differences in AUAFROCs between BSt-DE and BSp-DL. In the person-wise analysis, the diagnostic performance was significantly improved when BSt-DE was used (reader 1: area under the ROC curve [AUC] = 0.988 [95% CI: 0.946, 0.999], $p = 0.011$); reader 2: AUC = 0.974 [95% CI: 0.925, 0.995], $p < 0.001$) or when BSp-DL was used (reader 1: AUC = 0.984 [95% CI: 0.940, 0.998], $p = 0.005$; reader 2: AUC = 0.931 [95% CI: 0.867, 0.971], $p < 0.001$) than when CXR was used alone (reader 1: AUC = 0.915 [95% CI: 0.846, 0.959]; reader 2: AUC = 0.798 [95% CI: 0.711, 0.868]) (Table 4). There was no significant difference in diagnostic performance between the use of BSt-DE and BSp-DL.

In subgroup analysis according to nodule size, radiologists' detection sensitivity for nodules > 1 cm was significantly ($p < 0.017$) improved when either BSt-DE or BSp-DL was used compared to when CXR alone was used for both readers. There was no significant difference in

Table 2. Image Quality Scores of BSt-DE and BSp-DL by Two Independent Readers

Criteria	Reader	BSt-DE	BSp-DL	*P*
Visibility of pulmonary vessels in the lung fields	1	9.93 ± 0.96	8.49 ± 0.71	< 0.001
	2	10.22 ± 0.96	9.07 ± 1.03	< 0.001
Visibility of pulmonary vessels in the retrocardiac/ diaphragmatic area	1	3.90 ± 0.86	5.02 ± 0.69	< 0.001
	2	3.95 ± 0.81	4.93 ± 0.58	< 0.001
Visibility of central airway	2	2.85 ± 0.36	2.73 ± 0.45	0.096
	2	2.90 ± 0.30	2.80 ± 0.40	0.102
Bone suppression	1	2.56 ± 0.55	2.63 ± 0.54	0.180
	2	2.17 ± 0.50	2.17 ± 0.44	1.000
Total score	1	19.24 ± 1.46	18.88 ± 1.33	0.077
	2	19.24 ± 1.61	18.95 ± 1.40	0.105

Data are mean ± standard deviation. A *p* value < 0.017 was considered statistically significant (Bonferroni correction for multiple comparisons). BSp-DL = bone suppression imaging using deep learning, BSt-DE = bone subtraction imaging using a dual energy technique

https://doi.org/10.3348/kjr.2021.0146

sensitivity for detecting nodules > 1 cm between the use of BSt-DE and the use of BSp-DL. Meanwhile, radiologists detected more sub-centimeter nodules with the use of BSt-DE than with the use of BSp-DL. The difference in detection sensitivity between the use of BSt-DL and the use of BSp-BL for sub-centimeter nodules was significant (p = 0.002) for reader 1 (Table 3). There was no significant difference in detecting sub-centimeter nodules between the use of BSp-DL and the use of CXR alone for both readers.

In subgroup analysis according to nodule location, radiologists' detection sensitivity for nodules in the upper/mid lung zone was improved (p < 0.017) with the use of

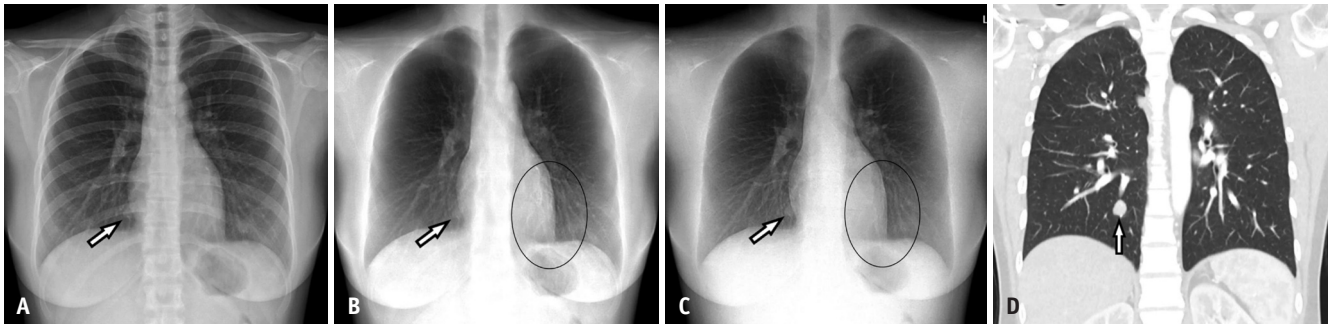

Fig. 3. A 27-year-old female who presented with cough.
A. A nodule in the right pericardiac area (arrow) is obscured by rib shadow on the CXR. **B, C.** BSp-DL **(B)** and BSt-DE **(C)** reveal the nodule (arrows) clearly. In the evaluation of image quality, readers scored visibility of pulmonary vessels in the lung fields higher in BSt-DE while visibility of pulmonary vessels in cardiac and diaphragmatic areas was higher in BSp-DL (circles). **D.** A coronal image of chest CT in lung window setting showing a 16 mm nodule (arrow) in the right lower lobe. BSp-DL = bone suppression imaging using deep learning, BSt-DE = bone subtraction imaging using dual energy technique, CXR = chest radiograph

Table 3. Nodule-Wise Analysis of Radiologists' Performance and Sensitivity for Detecting Lung Nodules

	Reader	CXR[a]	BSt-DE[b]	BSp-DL[c]	P^{a-b}	P^{a-c}	P^{b-c}
AUAFROC*	1	0.907 [0.858, 0.955]	0.996 [0.989, 1.000]	0.981 [0.964, 0.997]	< 0.001	0.005	0.095
	2	0.808 [0.733, 0.883]	0.976 [0.954, 0.997]	0.958 [0.926, 0.990]	< 0.001	< 0.001	0.350
Sensitivity							
For all nodules† (n = 83)	1	60.2 (50/83)	89.2 (74/83)	78.3 (65/83)	< 0.001	0.001	0.022
	2	53.0 (44/83)	83.1 (69/83)	71.1 (59/83)	< 0.001	0.001	0.031
< 1 cm† (n = 30)	1	36.7 (11/30)	80.0 (24/30)	46.7 (14/30)	< 0.001	0.375	0.002
	2	26.7 (8/30)	66.7 (20/30)	36.7 (11/30)	0.002	0.453	0.022
≥ 1 cm† (n = 53)	1	73.6 (39/53)	94.3 (50/53)	96.2 (51/53)	0.003	0.002	1.000
	2	67.9 (36/53)	92.5 (49/53)	90.6 (48/53)	0.001	< 0.001	1.000
Central† (n = 31)	1	67.7 (21/31)	87.1 (27/31)	87.1 (27/31)	0.070	0.070	1.000
	2	58.1 (18/31)	74.2 (23/31)	77.4 (24/31)	0.180	0.070	1.000
Peripheral† (n = 52)	1	55.8 (29/52)	90.4 (47/52)	73.1 (38/52)	< 0.001	0.012	0.012
	2	50.0 (26/52)	88.5 (46/52)	63.7 (35/52)	< 0.001	0.012	0.001
Upper/middle† (n = 52)	1	55.8 (29/52)	90.4 (47/52)	75.0 (39/52)	< 0.001	0.006	0.021
	2	48.1 (25/52)	82.7 (43/52)	69.2 (36/52)	< 0.001	0.003	0.065
Lower† (n = 31)	1	57.7 (21/31)	87.1 (27/31)	83.9 (26/31)	0.070	0.125	1.000
	2	61.3 (19/31)	83.9 (26/31)	74.2 (23/31)	0.039	0.219	0.453
Overlap with bones† (n = 55)	1	56.4 (31/55)	92.7 (51/55)	80.0 (44/55)	< 0.001	0.001	0.039
	2	47.3 (26/55)	85.5 (47/55)	70.9 (39/55)	< 0.001	0.001	0.039
No overlap with bones† (n = 28)	1	67.9 (19/28)	82.1 (23/28)	75.0 (21/28)	0.219	0.625	0.625
	2	64.3 (18/28)	78.6 (22/28)	71.4 (20/28)	0.219	0.625	0.687

*Numbers in parentheses show 95% confidence intervals of AUAFROC values, †Numbers indicate percentages with raw data in parentheses. p values refer to differences in sensitivities among the use of CXR[a] alone, the use of BSt-DE[b], and BSp-DL[c]. A p value < 0.017 was considered statistically significant (Bonferroni correction for multiple comparisons). AUAFROC = area under the alternative free-response receiver-operating characteristic curve, BSp-DL = bone suppression imaging using deep learning, BSt-DE = bone subtraction imaging using a dual energy technique, CXR = chest radiograph

https://doi.org/10.3348/kjr.2021.0146

Korean Journal of Radiology
KJR

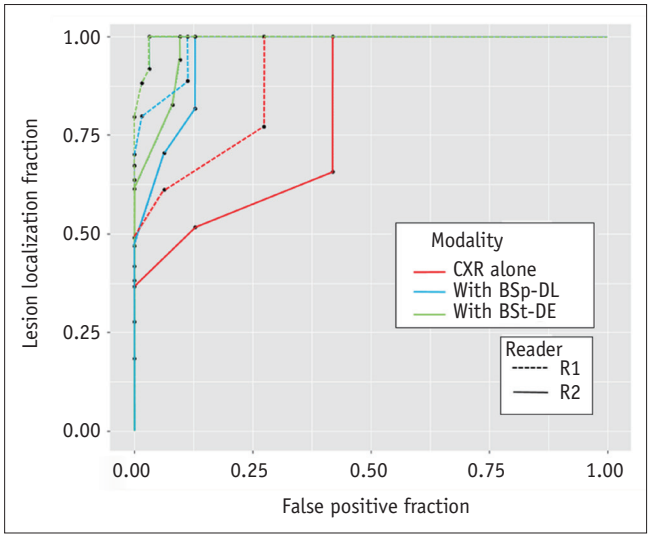

Fig. 4. The AFROC curves showing nodule-wise localization performance of radiologists. Area under the AFROC curve is improved significantly ($p < 0.017$) using BSp-DL or using BSt-DE than that using CXR alone. AFROC = alternative free-response receiver operating characteristic, BSp-DL= bone suppression imaging using deep learning, BSt-DE = bone subtraction imaging using dual energy technique, CXR = chest radiograph, R1 = reader 1, R2 = reader 2

BSt-DE or BSp-DL compared to that with the use of CXR alone for both readers (Fig. 5). However, there were no significant differences in the sensitivity for detecting nodules located in the lower lung zone among the three groups. For peripheral nodules (n = 52), the use of BSt-DE or BSp-DL showed significantly higher ($p < 0.017$) sensitivity than the use of CXR alone for both readers. In addition, the use of BSt-DE showed significantly (both $p < 0.017$) higher sensitivity than the use of BSp-DL for both readers. For central nodules (n = 31), there was no difference in detection sensitivity among the three groups.

Radiologists' detection sensitivity for nodules overlapping with bones (n = 55) was significantly ($p < 0.017$) improved with the use of BSt-DE or BSp-DL than with the use of CXR alone for both readers. However, there were no significant differences in the detection sensitivity between the use of BSt-DE and BSp-DL. Additionally, there were no differences in detection sensitivity among the three groups for nodules without overlapping bones (n = 28).

Inter-observer agreements on nodule-wise diagnosis were

Table 4. Person-Wise Diagnostic Performance of Nodule Detection in 111 Patients by Two Independent Readers

	Reader	CXR[a]	BSt-DE[b]	BSp-DL[c]	P^{a-b}	P^{a-c}	P^{b-c}
AUROC*	1	0.915 [0.846, 0.959]	0.988 [0.946, 0.999]	0.984 [0.940, 0.998]	0.011	0.005	0.772
	2	0.798 [0.711, 0.868]	0.974 [0.925, 0.995]	0.931 [0.867, 0.971]	< 0.001	< 0.001	0.064
Sensitivity†	1	71.4 (35/49)	95.9 (47/49)	91.8 (45/49)	< 0.001	0.002	0.625
	2	65.3 (32/49)	89.8 (44/49)	83.7(41/49)	0.002	0.012	0.453
Specificity†	1	93.5 (58/62)	98.4 (61/62)	98.4 (61/62)	0.250	0.250	1.000
	2	87.1 (54/62)	91.9 (57/62)	93.5 (58/62)	0.375	0.344	1.000

*Numbers in parentheses show 95% confidence intervals of AUROC values, †Numbers are shown as percentages with the raw data in parentheses. *p* values indicate differences in the diagnostic performance between the use of CXR[a] alone, the use of BSt-DE[b], and BSp-DL[c]. A *p* value < 0.017 was considered statistically significant (Bonferroni correction for multiple comparisons). AUROC = area under the receiver operating characteristic, BSp-DL = bone suppression imaging using deep learning, BSt-DE = bone subtraction imaging using a dual energy technique, CXR = chest radiograph

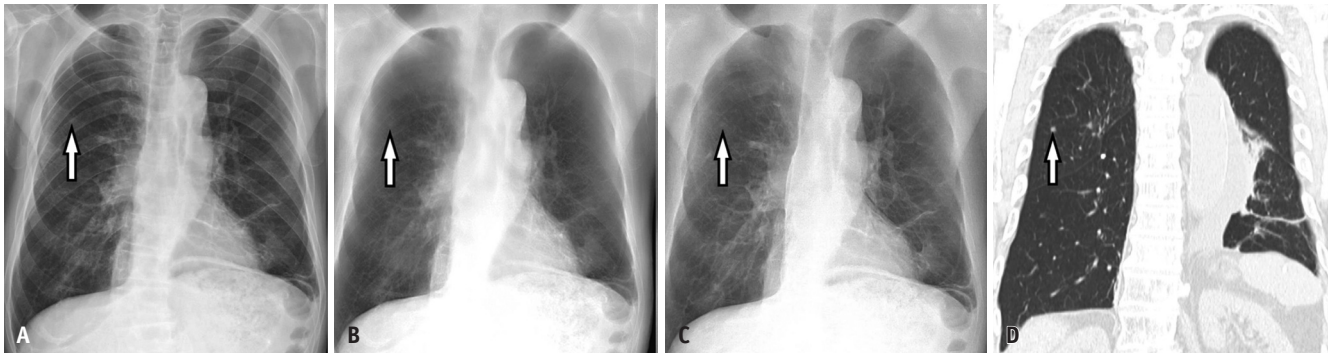

Fig. 5. An 85-year-old male who underwent CXR as a part of routine follow-up after gastrectomy for stomach cancer.
A. A tiny nodule overlapping with anterior arc of right 3rd rib (arrow) is suspicious in the right upper lobe on CXR. **B, C.** BSp-DL **(B)** and BSt-DE **(C)** confirm persistence of the nodule (arrows) after eliminating ribs. **D.** A coronal image of chest CT in lung window setting showing the presence of a 7 mm nodule (arrow) in the right upper lobe. BSp-DL= bone suppression imaging using deep learning, BSt-DE = bone subtraction imaging using dual energy technique, CXR = chest radiograph

https://doi.org/10.3348/kjr.2021.0146

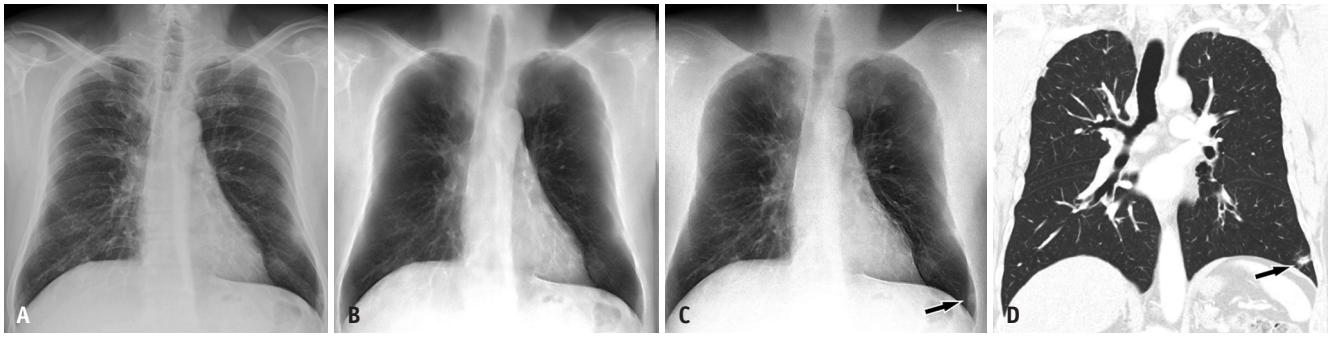

Fig. 6. A 61-year-old male who presented with cough and sputum.
A. Readers missed a subpleural nodule in the left lower lobe near the costophrenic angle on chest radiograph. **B.** BSp-DL cannot define the nodule clearly. **C.** BSt-DE delineates the nodule (arrow) in the left lower lobe. **D.** A coronal image of chest CT in lung window setting showing the presence of a 9 mm nodule (arrow) in the left lower lobe. BSp-DL= bone suppression imaging using deep learning, BSt-DE = bone subtraction imaging using dual energy technique

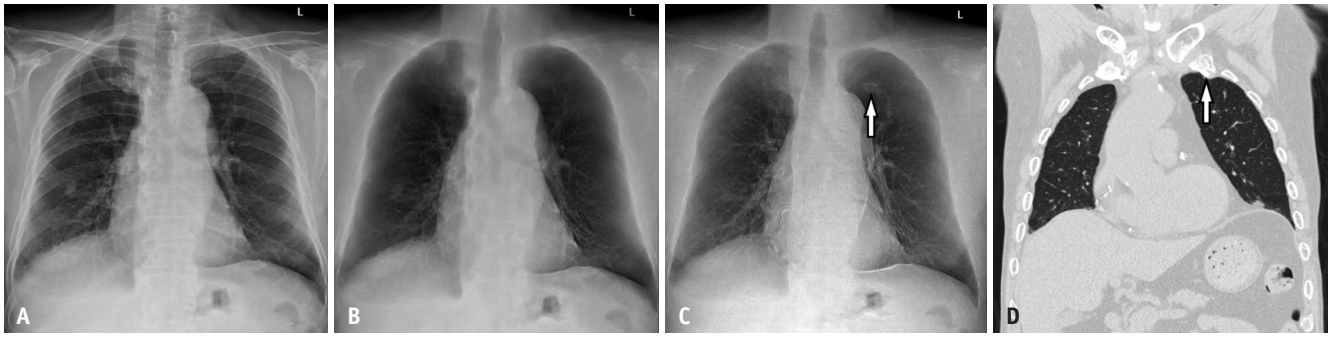

Fig. 7. A 77-year-old male with prostate cancer.
A. CXR shows a nodule opacity in the right lower lung zone due to an old fracture of the right 5th rib. **B.** BSp-DL showing no parenchymal lesion except the same opacity seen in CXR. **C.** BSt-DE showing additional nodular opacity in the left upper lobe (arrow). **D.** A coronal image of chest CT in lung window setting showing no parenchymal nodule in left upper lobe, but hypertrophied costochondral junction of the left 1st rib (arrow). BSp-DL= bone suppression imaging using deep learning, BSt-DE = bone subtraction imaging using dual energy technique, CXR = chest radiograph

found to be moderate for CXR (weighted kappa coefficient: 0.600) and substantial for both BSt-DE with CXR (weighted kappa coefficient: 0.707) and BSp-DL with CXR (weighted kappa coefficient: 0.697). Inter-observer agreements on person-wise diagnosis were substantial for both CXR (weighted kappa coefficient: 0.669) and BSp-DL with CXR (weighted kappa coefficient: 0.766) and almost perfect for BSt-DE with CXR (weighted kappa coefficient: 0.817).

False-Positive and False-Negative Findings

Causes of false-positive and false-negative findings were analyzed based on a retrospective review of chest CT and radiographic images in consensus by the two readers. While there were 23 (reader 1, n = 9 and reader 2, n = 14) false-negative nodules with the use of BSt-DE, there were 42 (reader 1, n = 18 and reader 2, n = 24) false-negative nodules when BSp-DL was used. We found that the majority of false-negative cases were small (≤ 10 mm) nodules, both with the use of BSt-DE (16/23, 69.6%) and the use of BSp-DL (35/42, 83.3%) (Fig. 6). Other false-negative nodules were due to misinterpretation of nodules as vascular structures or retrocardiac/infra-diaphragmatic locations of the nodules. There were 19 (reader 1, n = 8 and reader 2, n = 11) false positives detected using BSt-DE, and 13 (reader 1, n = 5 and reader 2, n = 8) false positives detected with the use of BSp-DL. False-positive nodules with the use of BSt-DE were due to vascular markings (n = 8), focal pleural thickening or parenchymal scar (n = 6), incompletely suppressed bones or costochondral junction (n = 4), or artifacts (n = 1) (Fig. 7). False-positive nodules with the use of BSp-DL were due to vascular markings (n = 6), parenchymal scars (n = 5), incompletely suppressed bones (n = 1), and nipple (n = 1).

DISCUSSION

Software-based bone suppression has been attempted through the delineation and segmentation of ribs and

https://doi.org/10.3348/kjr.2021.0146

Korean Journal of Radiology
KJR

clavicles on CXRs [18,24]. With recent advances in deep learning, neural-network-based bone suppression algorithms have been developed [13,25-28]. These algorithms have been programmed to eliminate bony structures by considering them as noise while preserving soft tissues [24]. However, several issues, such as uneven suppression, image blurring, and unnecessary suppression of soft tissue structures, remain to be solved for practical use [19,27,28]. In the present study, we used a bone suppression model based on the GAN framework to maintain the similarity between the target and generation domains [20]. The GAN can optimize image generators by adding discriminators as assisting components during training sessions [29]. To prevent image blurring, our model was designed to learn the frequency details of original images more effectively through the Harr 2D wavelet decomposition of the input data. The model outperformed other state-of-the-art methods in terms of quality metrics, such as peak signal-to-noise ratio and structural similarity index [20].

Previous studies have confirmed the positive effects of software-based bone suppression on lung nodule detection compared to the use of CXR alone [15-19,28,30]. Nonetheless, only a few studies have compared the performance of software-based suppression and DES techniques [21,22]. Szucs-Farkas et al. [31] showed that software-based bone suppression imaging can provide similar detection rates for pulmonary nodules with the DES technique. However, the improvement in detection did not reach statistical significance compared to CXR detection. Li et al. [32] have shown improved accuracy for detecting small lung cancers by radiologists when software-based bone suppression and DES images were additionally used with CXR. However, the accuracy of software-based bone suppression imaging is inferior to that of the DES technique. A recent study by Hong et al. [22] showed that commercial bone suppression software can produce improved bone-eliminated chest radiographic images compared to the DES technique in terms of structural similarity index and visual quality. However, the authors did not assess diagnostic performance in their study.

In the present study, the use of BSp-DL with CXR showed better diagnostic performance than CXR alone. Its performance was comparable to that of BSt-DE in both nodule-wise and person-wise analyses. The use of BSp-DL or BSt-DE increased the detectability of nodules overlapping with bones on CXR. The use of BSp-DL and BSt-DE improved the detectability of nodules in the upper/mid lung zone. However, there were no significant differences in the detection of lesions in the lower lung zone. Since the ribs and clavicles overlap in the upper lung zone, bone elimination techniques could be more helpful in detecting lesions in the upper lung than in the lower lung zone [33]. While there was no significant difference in the detection of nodules in the central lung zone or nodules > 10 mm, between the two techniques, BSp-DL was inferior to BSt-DE in detecting the sub-centimeter and peripheral nodules. Due to the variability and complexity of overlying cardiovascular structures, detection of central lesions might be less affected by the use of bone elimination techniques [3]. The inferior detectability of small and peripheral nodules with BSp-DL to BSt-DE might reflect the inherent weakness of bone suppression algorithms that use a de-noising approach. Ribs are more closely overlapped in the peripheral area on CXR, forming a lateral arc. Although our program outperformed existing state-of-the-art methods in preserving the frequency details of original images, small peripheral nodules could fade out while de-noising the overlying bones.

When image quality was compared, the score for visualization of pulmonary vessels in lung fields was higher in BSt-DE than in BSp-DL. However, the score for visualization of vessels in the cardiac and diaphragmatic areas was higher in BSp-DL than in BSt-DE. Similar to small peripheral nodules, pulmonary vessel area might have been suppressed in BSp-DL, particularly in peripheral lung fields where small caliber vessels are overlaid by assembled bony structures. The reason for the inferior visualization of vessels in the cardiac and diaphragmatic areas in BSt-DE is unclear. This may be partly attributed to misregistration of vessels in these areas caused by the inevitable motion of the heart and diaphragm [32].

False-negative lesions were more frequent with the use of BSp-DL, while false positive lesions were more recurrent with BSt-DE. False negative findings were mostly due to small lesion sizes in both techniques, as shown in a previous study [9]. False-positive lesions were thought to be caused by increased prominence of normal structures such as vessels, unsuppressed bone/costal cartilage, and parenchymal scars with diminishing bones in both techniques [17,34].

The limitations of this study include the small number of cases in each subgroup, which could have affected the results. Although a recent study [33] using a pre-commercial bone suppression algorithm has shown results

 https://doi.org/10.3348/kjr.2021.0146

similar to ours in nodule detectability relevance to anatomic location, further studies with larger numbers of cases are needed for validation. Second, we did not determine the optimal radiation dose for DES imaging through a phantom study. Instead, we used protocols suggested by the manufacturer. This might have affected the image quality as well as detectability of nodule with BSt-DE. Finally, our bone suppression model was an in-house program which has not been widely tested. Because our model being a new attempt, performance tests comparing with other types of deep learning methods are also needed to determine its weakness and strengths for clinical applications.

In conclusion, BSp-DL (a bone suppression imaging based on the GAN framework) can improve radiologists' detection of nodules on CXRs. It showed comparable performance to that of the DES technique. Nevertheless, further technical improvements are needed for better delineation of small and peripherally located nodules on CXRs.

Supplement

The Supplement is available with this article at https://doi.org/10.3348/kjr.2021.0146.

Availability of Data and Material

The datasets generated or analyzed during the study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

Author Contributions

Conceptualization: Kyung Nyeo Jeon, Il Dong Yun. Data curation: Kyungsoo Bae, Kyung Nyeo Jeon. Formal analysis: Kyungsoo Bae, Dong Yul Oh, Kyung Nyeo Jeon. Funding acquisition: Kyung Nyeo Jeon. Investigation: Dong Yul Oh, Kyung Nyeo Jeon. Methodology: Kyungsoo Bae, Dong Yul Oh. Project administration: Kyungsoo Bae, Kyung Nyeo Jeon. Resources: Dong Yul Oh, Il Dong Yun, Kyung Nyeo Jeon. Software: Dong Yul Oh, Il Dong Yun. Supervision: Kyung Nyeo Jeon. Validation: Kyungsoo Bae, Kyung Nyeo Jeon. Visualization: Kyung Nyeo Jeon. Writing—original draft: Kyungsoo Bae, Dong Yul Oh, Kyung Nyeo Jeon. Writing—review & editing: all authors.

ORCID iDs

Kyungsoo Bae
https://orcid.org/0000-0002-2545-2682
Dong Yul Oh
https://orcid.org/0000-0002-8393-7960
Il Dong Yun
https://orcid.org/0000-0001-9893-8897
Kyung Nyeo Jeon
https://orcid.org/0000-0003-2267-0366

Funding Statement

This work was supported by a fund of Research Promotion Program, Gyeongsang National University Hospital, 2019.

REFERENCES

1. Geitung JT, Skjaerstad LM, Göthlin JH. Clinical utility of chest roentgenograms. *Eur Radiol* 1999;9:721-723
2. Toomes H, Delphendahl A, Manke HG, Vogt-Moykopf I. The coin lesion of the lung. A review of 955 resected coin lesions. *Cancer* 1983;51:534-537
3. Del Ciello A, Franchi P, Contegiacomo A, Cicchetti G, Bonomo L, Larici AR. Missed lung cancer: when, where, and why? *Diagn Interv Radiol* 2017;23:118-126
4. Austin JH, Romney BM, Goldsmith LS. Missed bronchogenic carcinoma: radiographic findings in 27 patients with a potentially resectable lesion evident in retrospect. *Radiology* 1992;182:115-122
5. de Groot PM, Carter BW, Abbott GF, Wu CC. Pitfalls in chest radiographic interpretation: blind spots. *Semin Roentgenol* 2015;50:197-209
6. Shah PK, Austin JH, White CS, Patel P, Haramati LB, Pearson GD, et al. Missed non-small cell lung cancer: radiographic findings of potentially resectable lesions evident only in retrospect. *Radiology* 2003;226:235-241
7. Oda S, Awai K, Funama Y, Utsunomiya D, Yanaga Y, Kawanaka K, et al. Detection of small pulmonary nodules on chest radiographs: efficacy of dual-energy subtraction technique using flat-panel detector chest radiography. *Clin Radiol* 2010;65:609-615
8. Li F, Engelmann R, Doi K, MacMahon H. Improved detection of small lung cancers with dual-energy subtraction chest radiography. *AJR Am J Roentgenol* 2008;190:886-891
9. Szucs-Farkas Z, Patak MA, Yuksel-Hatz S, Ruder T, Vock P. Single-exposure dual-energy subtraction chest radiography: detection of pulmonary nodules and masses in clinical practice. *Eur Radiol* 2008;18:24-31
10. Kelcz F, Zink FE, Peppler WW, Kruger DG, Ergun DL, Mistretta CA. Conventional chest radiography vs dual-energy computed radiography in the detection and characterization of pulmonary nodules. *AJR Am J Roentgenol* 1994;162:271-278
11. Kashani H, Varon CA, Paul NS, Gang GJ, Van Metter R,

Yorkston J, et al. Diagnostic performance of a prototype dual-energy chest imaging system ROC analysis. *Acad Radiol* 2010;17:298-308
12. Kuhlman JE, Collins J, Brooks GN, Yandow DR, Broderick LS. Dual-energy subtraction chest radiography: what to look for beyond calcified nodules. *Radiographics* 2006;26:79-92
13. Gasarev M, Kuleev R, Khan A, Rivera AR, Khattak AM. Deep learning models for bone suppression in chest radiographs. Proceedings of the 2017 IEEE Conference on Computational Interlligence in Bioinformatics and Computational Biology (CIBCB); 2017 Aug 23-25; Manchester, UK: IEEE; 2017; p. 1-7
14. Yang W, Chen Y, Liu Y, Zhong L, Qin G, Lu Z, et al. Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain. *Med Image Anal* 2017;35:421-433
15. Freedman MT, Lo SC, Seibel JC, Bromley CM. Lung nodules: improved detection with software that suppresses the rib and clavicle on chest radiographs. *Radiology* 2011;260:265-273
16. Schalekamp S, van Ginneken B, Meiss L, Peters-Bax L, Quekel LG, Snoeren MM, et al. Bone suppressed images improve radiologists' detection performance for pulmonary nodules in chest radiographs. *Eur J Radiol* 2013;82:2399-2405
17. Hong GS, Do KH, Lee CW. Added value of bone suppression image in the detection of subtle lung lesions on chest radiographs with regard to reader's expertise. *J Korean Med Sci* 2019;34:e250
18. Oğul H, Oğul BB, Ağıldere AM, Bayrak T, Sümer E. Eliminating rib shadows in chest radiographic images providing diagnostic assistance. *Comput Methods Programs Biomed* 2016;127:174-184
19. Schalekamp S, van Ginneken B, Koedam E, Snoeren MM, Tiehuis AM, Wittenberg R, et al. Computer-aided detection improves detection of pulmonary nodules in chest radiographs beyond the support by bone-suppressed images. *Radiology* 2014;272:252-261
20. Oh DY, Yun ID. Learning bone suppression from dual energy chest X-rays using adversarial networks. Arxiv.org Web site. https://arxiv.org/abs/1811.02628. Published November 5, 2018. Accessed November 19, 2020
21. Kang E, Min J, Ye JC. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. *Med Phys* 2017;44:e360-e375
22. Hong GS, Do KH, Son AY, Jo KW, Kim KP, Yun J, et al. Value of bone suppression software in chest radiographs for improving image quality and reducing radiation dose. *Eur Radiol* 2021;31:5160-5171
23. Chakraborty DP, Zhai X. Analysis of data acquired using ROC paradigm and its extensions. Mran.microsoft Web site. https://mran.microsoft.com/snapshot/2015-05-08/web/packages/RJafroc/vignettes/RJafroc.pdf. Accessed May 13, 2021
24. van Ginneken B, Stegmann MB, Loog M. Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. *Med Image Anal* 2006;10:19-40
25. Suzuki K, Abe H, MacMahon H, Doi K. Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN). *IEEE Trans Med Imaging* 2006;25:406-416
26. Li F, Engelmann R, Pesce L, Armato SG 3rd, Macmahon H. Improved detection of focal pneumonia by chest radiography with bone suppression imaging. *Eur Radiol* 2012;22:2729-2735
27. Zarshenas A, Liu J, Forti P, Suzuki K. Separation of bones from soft tissue in chest radiographs: anatomy-specific orientation-frequency-specific deep neural network convolution. *Med Phys* 2019;46:2232-2242
28. Oda S, Awai K, Suzuki K, Yanaga Y, Funama Y, MacMahon H, et al. Performance of radiologists in detection of small pulmonary nodules on chest radiographs: effect of rib suppression with a massive-training artificial neural network. *AJR Am J Roentgenol* 2009;193:W397-W402
29. Zhou Z, Zhou L, Shen K. Dilated conditional GAN for bone suppression in chest radiographs with enforced semantic features. *Med Phys* 2020;47:6207-6215
30. Endo K, Kaneko A, Horiuchi Y, Kasuga N, Ishizaki U, Sakai S. Detectability of pulmonary nodules on chest radiographs: bone suppression versus standard technique with single versus dual monitors for visualization. *Jpn J Radiol* 2020;38:676-682
31. Szucs-Farkas Z, Schick A, Cullmann JL, Ebner L, Megyeri B, Vock P, et al. Comparison of dual-energy subtraction and electronic bone suppression combined with computer-aided detection on chest radiographs: effect on human observers' performance in nodule detection. *AJR Am J Roentgenol* 2013;200:1006-1013
32. Li F, Engelmann R, Pesce LL, Doi K, Metz CE, Macmahon H. Small lung cancers: improved detection by use of bone suppression imaging--comparison with dual-energy subtraction chest radiography. *Radiology* 2011;261:937-949
33. Miyoshi T, Yoshida J, Aramaki N, Matsumura Y, Aokage K, Hishida T, et al. Effectiveness of bone suppression imaging in the detection of lung nodules on chest radiographs: relevance to anatomic location and observer's experience. *J Thorac Imaging* 2017;32:398-405
34. Li F, Hara T, Shiraishi J, Engelmann R, MacMahon H, Doi K. Improved detection of subtle lung nodules by use of chest radiographs with bone suppression imaging: receiver operating characteristic analysis with and without localization. *AJR Am J Roentgenol* 2011;196:W535-W541

 https://doi.org/10.3348/kjr.2021.0146