Original Article | Thoracic Imaging

eISSN 2005-8330
https://doi.org/10.3348/kjr.2020.0988
Korean J Radiol 2021;22(6):983-993

Korean Journal of Radiology
KJR

Cycle-Consistent Generative Adversarial Network: Effect on Radiation Dose Reduction and Image Quality Improvement in Ultralow-Dose CT for Evaluation of Pulmonary Tuberculosis

Chenggong Yan[1, 2], Jie Lin[1], Haixia Li[3], Jun Xu[4], Tianjing Zhang[3], Hao Chen[5], Henry C. Woodruff[2, 6], Guangyao Wu[2], Siqi Zhang[1], Yikai Xu[1], Philippe Lambin[2, 6]

[1]Department of Medical Imaging Center, Nanfang Hospital, Southern Medical University, Guangzhou, China; [2]The D-Lab, Department of Precision Medicine, GROW-School for Oncology and Developmental Biology, Maastricht University, Maastricht, The Netherlands; [3]Clinical and Technical Solution, Philips Healthcare, Guangzhou, China; [4]Department of Hematology, Nanfang Hospital, Southern Medical University, Guangzhou, China; [5]Jiangsu JITRI Sioux Technologies Co., Ltd., Suzhou, China; [6]Department of Radiology and Nuclear Imaging, GROW-School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, The Netherlands

Objective: To investigate the image quality of ultralow-dose CT (ULDCT) of the chest reconstructed using a cycle-consistent generative adversarial network (CycleGAN)-based deep learning method in the evaluation of pulmonary tuberculosis.

Materials and Methods: Between June 2019 and November 2019, 103 patients (mean age, 40.8 ± 13.6 years; 61 men and 42 women) with pulmonary tuberculosis were prospectively enrolled to undergo standard-dose CT (120 kVp with automated exposure control), followed immediately by ULDCT (80 kVp and 10 mAs). The images of the two successive scans were used to train the CycleGAN framework for image-to-image translation. The denoising efficacy of the CycleGAN algorithm was compared with that of hybrid and model-based iterative reconstruction. Repeated-measures analysis of variance and Wilcoxon signed-rank test were performed to compare the objective measurements and the subjective image quality scores, respectively.

Results: With the optimized CycleGAN denoising model, using the ULDCT images as input, the peak signal-to-noise ratio and structural similarity index improved by 2.0 dB and 0.21, respectively. The CycleGAN-generated denoised ULDCT images typically provided satisfactory image quality for optimal visibility of anatomic structures and pathological findings, with a lower level of image noise (mean ± standard deviation [SD], 19.5 ± 3.0 Hounsfield unit [HU]) than that of the hybrid (66.3 ± 10.5 HU, $p < 0.001$) and a similar noise level to model-based iterative reconstruction (19.6 ± 2.6 HU, $p > 0.908$). The CycleGAN-generated images showed the highest contrast-to-noise ratios for the pulmonary lesions, followed by the model-based and hybrid iterative reconstruction. The mean effective radiation dose of ULDCT was 0.12 mSv with a mean 93.9% reduction compared to standard-dose CT.

Conclusion: The optimized CycleGAN technique may allow the synthesis of diagnostically acceptable images from ULDCT of the chest for the evaluation of pulmonary tuberculosis.

Keywords: *Computed tomography; Radiation dose; Infection; Artificial intelligence; Deep learning*

Received: August 16, 2020 **Revised:** November 22, 2020 **Accepted:** December 21, 2020

This work was supported by the Natural Science Foundation of Guangdong Province, China (No. 2017A030310102, 2018030310343, and 2020B1515020008) and Medical Scientific Research Foundation of Guangdong Province (No. A2018014). Authors also acknowledge financial support from ERC advanced grant (ERC-ADG-2015 no 694812-Hypoximmuno), ERC-2020-PoC: 957565-AUTO.DISTINCT. Authors also acknowledge financial support from EUROSTARS (COMPACT-12053), the European Union's Horizon 2020 research and innovation programme under grant agreement: ImmunoSABR no 733008, MSCA-ITN-PREDICT no 766276, CHAIMELEON no 952172, EuCanImage no 952103, TRANSCAN Joint Transnational Call 2016 (JTC2016 CLEARLY no UM 2017-8295) and Interreg V-A Euregio Meuse-Rhine (EURADIOMICS no EMR4). This work was supported by the Dutch Cancer Society (KWF Kankerbestrijding), Project number 12085/2018-2.

Corresponding author: Yikai Xu, MD, PhD, Department of Medical Imaging Center, Nanfang Hospital, Southern Medical University, No. 1838 Guangzhou Avenue North, Guangzhou 510515, Guangdong, China.
• E-mail: yikaixu917@gmail.com

INTRODUCTION

Tuberculosis remains a widely distributed infection and a major global health problem [1,2]. Tuberculosis can be treated with a standard regimen for 6–9 months. The treatment may be continued for up to 2 years if M. tuberculosis bacteria are multidrug-resistant. Chest CT is an important diagnostic imaging modality for the management of pulmonary tuberculosis [3]. Health care personnel with tuberculosis should undergo a symptom evaluation and chest imaging to determine the risks and benefits of treatment every 3–6 months [4]. However, repeated follow-up CT poses concerns related to the potential risk of radiation-induced carcinogenesis. Recent studies have focused on dose reduction following the principle of "as low as reasonably achievable" while maintaining image quality and diagnostic accuracy [5].

Hybrid iterative reconstruction (IR) algorithms are routinely used in commercially available reconstruction techniques to decrease the radiation dose; they utilize statistical model-based denoising in image reconstruction [6]. The recent commercial implementation of iterative model reconstruction (IMR) is a more advanced and knowledge-based IR algorithm based on noise modeling techniques to further improve the spatial resolution and image quality [7]. Although IMR techniques enable substantial image noise reduction, several prior studies have shown their limitations related to over-smoothing or blotchy pixelated appearance [8,9].

Recently, several deep learning networks for image-to-image translation have been reported, and they have enabled remarkable radiation dose reduction by image denoising [10-13]. Specifically, generative adversarial networks (GANs) have achieved state-of-the-art performance in image generation [14]. These models often consist of a generator network focused on image synthesis and a discriminator network centered on discrimination. The two networks are trained simultaneously and challenge each other to achieve high-quality visual images from low-dose acquisitions. Kaplan and Zhu [15] proposed an end-to-end GAN-based framework that transformed low-dose PET images into estimates of the corresponding high-quality full-dose PET images. Recently, an unsupervised translation cycle-consistent GAN (CycleGAN) model was developed to convert cone-beam CT images to more accurately synthesize CT images for treatment planning [16]. To date, only a few clinical studies have evaluated the utility of GANs for noise reduction for low-dose CT [17,18]. Furthermore, published clinical studies on dose reduction that have compared deep learning techniques with commercially available IR methods are limited [10,12].

The purpose of this study was to train a CycleGAN-based denoising network to generate high-quality images from ultralow-dose CT (ULDCT) at chest radiography dose levels and evaluate and compare the quality of CycleGAN-reconstructed images with those of the hybrid and knowledge-based IMR algorithms in patients with pulmonary tuberculosis.

MATERIALS AND METHODS

This single-center, Health Insurance Portability and Accountability Act-compliant study was approved by the Institutional Review Board. Written informed consent was obtained from all patients.

Patient Selection

Between June 2019 and November 2019, 129 consecutive patients with clinical indications of tuberculosis on chest CT were enrolled (Fig. 1). The inclusion criteria were as follows: 1) age of ≥ 18 years; 2) characteristic symptoms (fever, cough, and night sweats) and microbiological confirmation with positive sputum microbiology results; 3) initial diagnosis or follow-up under tuberculostatic medication. The exclusion criteria were as follows: no typical morphological features indicative of tuberculosis (n = 15) and obesity with body mass index of ≥ 25 kg/m^2 (n = 11). In this study, 103 patients with pulmonary

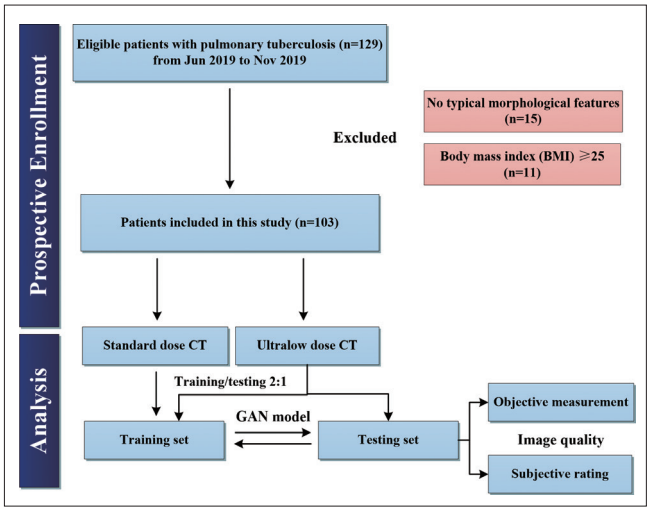

Fig. 1. Flowchart showing participant inclusion and exclusion criteria. GAN = generative adversarial network

https://doi.org/10.3348/kjr.2020.0988

Korean Journal of Radiology
KJR

tuberculosis (61 men, 42 women; mean age ± standard deviation [SD], 40.8 ± 13.6 years; age range, 18–68 years; mean body mass index ± SD, 20.5 ± 2.4 kg/m^2) were finally included.

CT Technique and Image Reconstruction

All examinations were performed using a 256-slice CT scanner (Brilliance iCT; Philips Healthcare). The standard-dose CT (SDCT) examinations were performed at 120 kVp with automatic exposure control at a dose right index of 13, which is routinely used in clinical practice. An additional set of ULDCT research images were acquired with the tube voltage lowered to 80 kVp, and the fixed-tube current-time product was reduced to 10 mAs. All remaining scanning parameters were kept constant for the two unenhanced image series with a pitch of 0.984:1 and detector configuration of 128 x 0.625 mm. The raw data were reconstructed using the hybrid IR algorithm (iDose) and knowledge-based IMR (Philips Healthcare), with a section thickness of 1 mm and a resolution of 512 x 512 pixels.

Generative Adversarial Network for Noise Reduction

A total of 103 cases were randomly divided into training and testing sets using a ratio of approximately 2:1 (training [68 patients]/testing [35 patients]). According to our preliminary results (Supplementary Table 1, Supplementary Fig. 1) and a previously published report [16], the deep learning algorithm for denoising was optimized based on CycleGAN rather than WGAN-VGG or Pix2Pix. We implemented the network architectures in Pytorch based on the GitHub project (https://github.com/UserWendy/GAN-denoising). The original CycleGAN pipeline contains two networks: a generator (G) and a discriminator (D), as illustrated in Figures 2 and 3. G contains G1 and G2, which simultaneously learn the mappings from ULDCT → SDCT and SDCT → ULDCT, respectively. D consists of D1 and D2:D1 discriminates the real SDCT images from the source domain and synthetic SDCT images from G1, whereas D2 discriminates the real images from the target ULDCT domain and synthetic ULDCT images from G2. In this adversarial learning process, networks are simultaneously optimized to generate high-quality images. The computation time of the optimized CycleGAN denoising model was 0.05–0.07 seconds per image. The details of the optimized CycleGAN network are provided in the Supplementary Materials. To validate the performance of our proposed methods, the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) were used for image quality evaluation [17].

Objective Image Analysis

Regions of interest (ROIs) were manually drawn within the descending aorta at the level of the carina and around

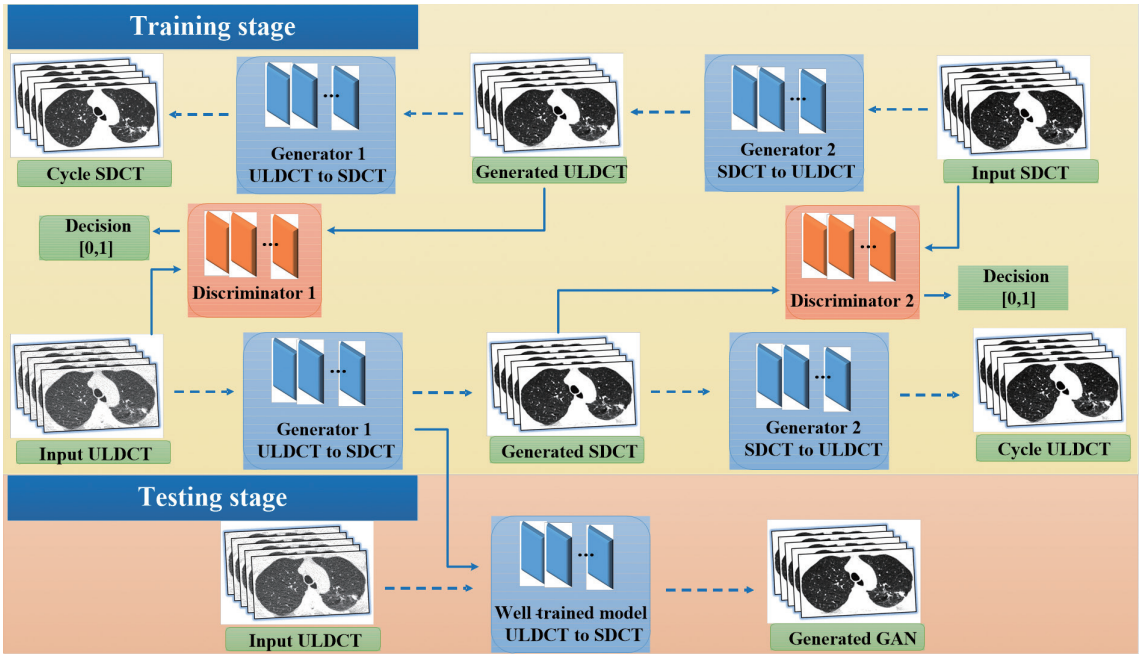

Fig. 2. A schematic view of the proposed cycle-consistent GAN frameworks for image-to-image translation during the training and testing stage. The generator and discriminator networks were simultaneously optimized to generate high-quality images from low-dose acquisitions. GAN = generative adversarial network, SDCT = standard dose CT, ULDCT = ultralow dose CT

https://doi.org/10.3348/kjr.2020.0988

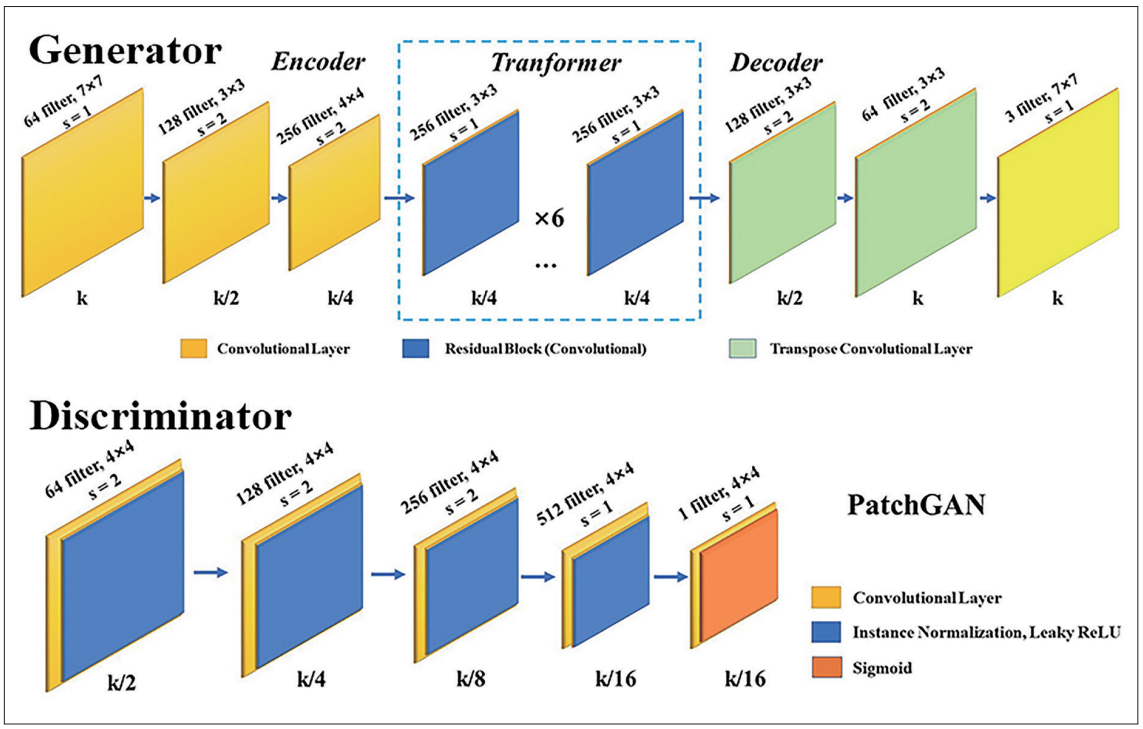

Fig. 3. The structure of the generator and discriminator network in the proposed cycle-consistent GAN. GAN = generative adversarial network, ReLU = rectified linear units

the most significant pulmonary lesions (longest diameter > 1 cm) by an experienced thoracic radiologist. Large vessels and bronchi were avoided when the ROIs were placed. The signal-to-noise ratio (SNR) in the aorta and contrast-to-noise ratio (CNR) for the lesion and the aorta were calculated using the following formulas: SNR = $|ROI_A/N|$ and CNR = $|ROI_L - ROI_A| /N$, where the image noise, N, was defined as the SD of CT attenuation in the aorta, and ROI_L and ROI_A were the mean CT values within the ROIs of the lesion and the aorta, respectively. The image noise was normalized based on the effective dose (ED) for the 80- and 120-kVp protocols using a figure of merit (FOM) defined as $FOM_{noise} = 1/(N^2 \cdot ED)$ [9]. These FOM values are mainly used for objective noise assessment that is independent of the radiation output.

Subjective Image Analysis

The subjective visual assessment was independently performed by two experienced thoracic radiologists (with 5 and 9 years of experience, respectively). All images were anonymized, blinded, and analyzed in random order. Standard-dose iDose images were used as reference for diagnostic interpretation. A consensus reading was used to resolve disagreements between the two observers. The overall image quality, noise, visibility of anatomical structures (proximal bronchi and vessels, peripheral bronchi and vessels, fissures, lymph nodes, and pericardium), and pathological findings (centrilobular nodules, consolidations, tree-in-bud, cavitation, calcification, and fibrosis) were evaluated on a 5-point scale (5 indicating best through to 1 indicating worst, Table 1) [19].

Assessment of Radiation Dose

The volume CT dose index (mGy) and dose-length product (mGy·cm) were recorded from the dose information page. The estimated ED was calculated by multiplying the dose-length product by a conversion factor (k = 0.014 mSv mGy^{-1} cm^{-1}).

Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics for Windows, Version 21.0 (IBM Corp.). For the baseline patient characteristics, the Student t test for continuous variables and the χ^2 test for categorical variables were used to compare the training and testing data. Continuous variables including PSNR, SSIM, CT number, noise, SNR, CNR, and FOM_{noise} were analyzed using repeated-measures analysis of variance. The Wilcoxon signed-rank test was used to assess the subjective image quality scores. Interobserver reliability was assessed by calculating the weighted κ coefficient. The paired student's *t* test was

https://doi.org/10.3348/kjr.2020.0988

Korean Journal of Radiology
KJR

Table 1. Grading Scale for Subjective Image Analysis

Grading Scale	Overall Image Quality	Noise	Visibility of Anatomic Structures and Pathological Findings
1	Excellent image quality without artifacts	Visually no image noise	Excellent visibility of details with sharp demarcation of structures
2	Good image quality with minor artifacts	No significant noise	Good visibility of the structures, with unrestricted image evaluation
3	Fair image quality with moderate artifacts	Acceptable noise	Moderate visibility of the interface structures, with slight restricted assessment
4	Poor image quality with substantial artifacts	More than acceptable noise	Poorly defined structures, with uncertainty about the evaluation

performed to compare the radiation dose parameters of the two successive scans. Bonferroni adjustment was applied to the four comparisons of the multiple groups, with $p < 0.008$ indicating a statistically significant difference.

RESULTS

The baseline patient characteristics showed no statistically significant differences in the training and testing sets (all, $p > 0.05$), as shown in Table 2. In the testing set, the main CT findings were centrilobular nodules (27/35, 77.1%), followed by fibrosis (60.0%), consolidation (34.3%), tree-in-bud (20.0%), and cavitation (17.1%). Mediastinal calcification, in addition to pulmonary abnormalities, was observed in 18 (51.4%) patients. Representative cases of pulmonary tuberculosis are presented in Figures 4 and 5.

Table 2. Clinical Characteristics of the Full Cohort

	Training Set (n = 68)	Testing Set (n = 35)	*P*
Age, years*	40.1 ± 13.9	42.1 ± 13.1	> 0.05
Sex			> 0.05
Men	39 (57.4)	22 (62.9)	
Women	29 (42.6)	13 (37.1)	
BMI, kg/m^2*	20.5 ± 2.2	20.6 ± 2.7	> 0.05
Imaging findings			> 0.05
Centrilobular nodules	58 (85.3)	27 (77.1)	
Tree-in-bud	19 (27.9)	7 (20.0)	
Fibrosis	39 (57.4)	21 (60.0)	
Calcification	22 (32.4)	18 (51.4)	
Cavitation	15 (22.1)	6 (17.1)	
Consolidation	26 (38.2)	12 (34.3)	

Unless otherwise specified, data are the number of patients, with percentages in parentheses. *Data are means ± standard deviation. *P* values for differences were calculated with the student *t* test or χ^2 test. BMI = body mass index

Objective Image Analysis

The objective measurements of image quality are summarized in Table 3. For the ULDCT images, the proposed GAN method demonstrated PSNR and SSIM that were significantly higher than IMR and iDose (both $p < 0.001$). The mean CT attenuation values of the aorta significantly differed with the various acquisition and reconstruction techniques ($p < 0.001$). For the significant pulmonary lesions, the CT attenuation values of the SDCT, ULDCT with GAN, IMR, and iDose images were not significantly different ($p = 0.908$). The mean image noise ± SD (19.5 ± 3.0 Hounsfield unit [HU]) of the GAN-generated images was significantly lower than that of iDose (66.3 ± 10.5 HU, $p < 0.001$) and similar to that of IMR (19.6 ± 2.6 HU, $p = 0.908$). Among the ultralow-dose CT, the GAN-generated images showed the highest SNRs and CNRs for the pulmonary lesions, followed by IMR and iDose (GAN vs. IMR, $p > 0.05$; GAN vs. iDose, $p < 0.001$). The image noise, SNRs, and CNRs for the ultralow-dose GAN and standard-dose iDose images were similar, with no significant differences ($p = 0.042$–0.471). Moreover, our FOM results showed that when the radiation dose was kept constant, the ULDCT GAN and IMR images yielded significantly lower noise than SDCT (both $p < 0.001$).

Subjective Image Analysis

The results of the subjective image quality assessment are summarized in Tables 4 and 5. For the overall diagnostic quality and noise ratings, the ultralow-dose GAN images were typically graded as "good image quality with minor artifacts" (28/35, 80.0%) and "no significant image noise" (26/35, 74.3%). For the visibility of normal anatomic structures and pathological findings on ultralow-dose images, the hybrid IR iDose images were substantially improved with GAN. However, the network generated artifacts in the images of 3 of 35 patients (8.6%), although they did not affect the diagnostic interpretation. A minor

 https://doi.org/10.3348/kjr.2020.0988

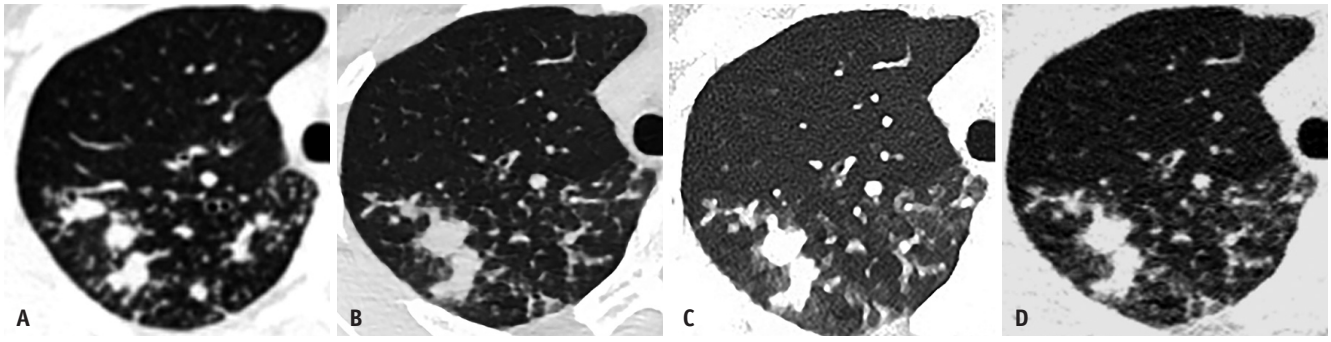

Fig. 4. SDCT iDose images (A) and ULDCT images of multiple centrilobular nodules with slight ground-glass opacity indicative of pulmonary tuberculosis in the lung window reconstructed with GAN, IMR, and iDose (B-D).
The visibility of pathological findings in ULDCT GAN images (**B,** output) is substantially better than that of iDose (**D,** input) and visually similar to that of SDCT iDose **(A)**. Note the minor blotchy appearance in the IMR image **(C)**; the accompanying ground-glass opacity is slightly exaggerated. GAN = generative adversarial network, IMR = iterative model reconstruction, SDCT = standard dose CT, ULDCT = ultralow dose CT

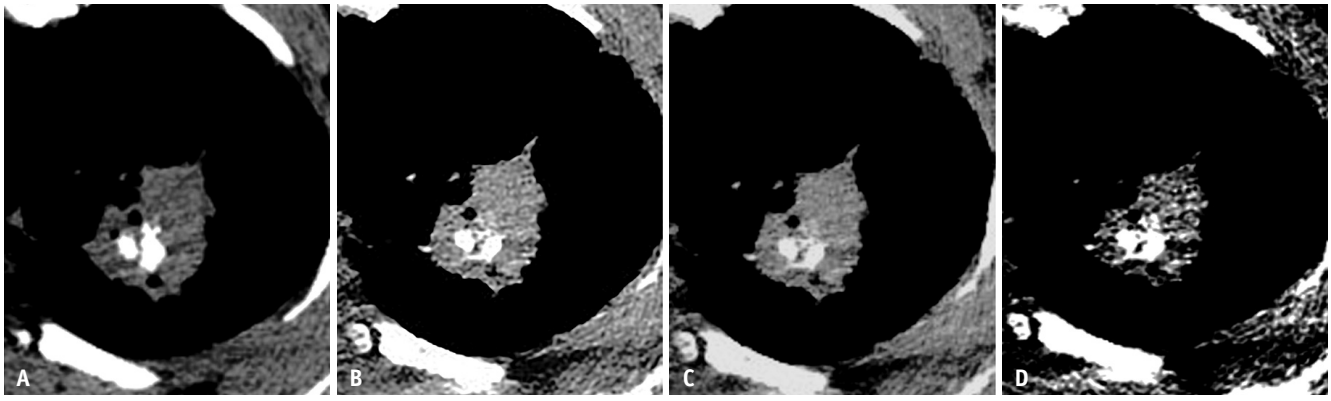

Fig. 5. CT images of multiple calcifications in lobular consolidation shown in the mediastinal window of the standard dose CT iDose (A) and ULDCT reconstructed with GAN, iterative model reconstruction, and iDose (B-D).
Note the lower image noise and improved image quality for the ULDCT GAN (**B,** output) compared with iDose (**D,** input). GAN = generative adversarial network, ULDCT = ultralow dose CT

blotchy image appearance was observed in ULDCT IMR images (Fig. 4), which mildly limited the visualization of the CT features. Additionally, the interobserver agreement was substantial (κ = 0.61–0.74) for the subjective image quality scores.

Radiation Dose

The mean value ± SD of volume CT dose index and the dose-length product for SDCT were 163.0 ± 141.7 mGy and 3.6 ± 0.6 mGy·cm, whereas those for ULDCT were 8.3 ± 0.6 mGy and 0.2 ± 0.0 mGy·cm, respectively. Compared with the standard-dose scan, there was a 93.9% reduction in the average ED using the ultralow-dose scan (2.3 ± 2.0 mSv vs. 0.12 ± 0.01 mSv; $p < 0.001$).

DISCUSSION

Health care personnel with a positive tuberculosis test should be monitored with follow-up chest imaging if they begin treatment with a recommended regimen [4]. Multidrug-resistant tuberculosis, for which first-line drugs are ineffective, requires extensive chemotherapy for up to 2 years of treatment. CT is a critical imaging modality for the assessment of chronic pulmonary diseases. However, cumulative radiation exposure from repeated CT has raised significant public safety concerns. Several prior studies have shown that various drastic dose reduction techniques are limited by the inherent trade-off between greater image noise and diminished image quality [20,21]. In this study, we proposed a CycleGAN-based deep neural network framework with feature matching for denoising ultralow-dose chest CT images for follow-up and visualization of pulmonary tuberculosis. Our results showed that the CycleGAN model enabled a substantial reduction in the radiation dose (93.9%, ED, 0.12/2.3 mSv) delivered for chest CT and generated diagnostically acceptable images

https://doi.org/10.3348/kjr.2020.0988

Korean Journal of Radiology
KJR

Table 3. Objective Measurements of Image Quality Parameters

Parameter	SDCT	ULDCT			*P*		
	iDose	GAN	IMR	iDose	SDCT vs. GAN	GAN vs. IMR	GAN vs. iDose
PSNR	NA	32.4 ± 3.3	31.0 ± 2.9	30.4 ± 2.2	NA	< 0.001	< 0.001
SSIM	NA	0.83 ± 0.03	0.77 ± 0.04	0.62 ± 0.05	NA	< 0.001	< 0.001
CT $value_{aorta}$ (HU)	44.3 ± 5.5	54.1 ± 17.6	51.7 ± 9.0	54.9 ± 10.5	< 0.001	0.379	0.776
CT $value_{lesion}$ (HU)	48.7 ± 111.5	26.6 ± 81.1	42.5 ± 169.9	31.3 ± 167.4	1.000	1.000	1.000
Noise (HU)	16.6 ± 2.7	19.5 ± 3.0	19.6 ± 2.6	66.3 ± 10.5	0.042	0.908	< 0.001
SNR	2.8 ± 6.4	2.9 ± 1.1	2.7 ± 0.6	0.4 ± 2.2	0.471	0.286	< 0.001
CNR	3.9 ± 5.0	3.5 ± 2.5	3.4 ± 3.3	1.5 ± 1.6	0.058	0.902	< 0.001
FOM_{noise} 10^{-3}	2.0 ± 0.9	25.0 ± 13.2	23.4 ± 7.1	2.1 ± 0.7	< 0.001	0.362	< 0.001

Data are means ± standard deviation. *P* values for differences were calculated with the repeated-measures analysis of variance with Bonferroni test for post-hoc comparisons. CNR = contrast-to-noise ratio, FOM = figure of merit, GAN = generative adversarial network, HU = Hounsfield unit, IMR = iterative model reconstruction, PSNR = peak signal-to-noise ratio, SDCT = standard dose CT, SNR = signal-to-noise ratio, SSIM = structural similarity index, ULDCT = ultralow dose CT

Table 4. Qualitative Assessment of the Overall Image Quality and Noise

Parameter	SDCT	ULDCT			*P*		
	iDose	GAN	IMR	iDose	SDCT vs. GAN	GAN vs. IMR	GAN vs. iDose
Overall image quality					0.005	0.083	< 0.001
5, excellent	16 (45.7)	6 (17.1)	4 (11.4)	0			
4, good	19 (54.3)	28 (80.0)	29 (82.9)	1 (2.9)			
3, fair	0	1 (2.9)	2 (5.7)	22 (62.9)			
2, poor	0	0	0	12 (34.3)			
1, nondiagnostic	0	0	0	0			
Noise					0.796	0.008	< 0.001
5, no	10 (28.6)	9 (25.7)	3 (8.6)	0			
4, little	25 (71.4)	26 (74.3)	32 (91.4)	0			
3, acceptable	0	0	0	4 (11.4)			
2, more than acceptable	0	0	0	31 (88.6)			
1, unacceptable	0	0	0	0			

Data are the number of patients, with percentages in parentheses. *P* values for differences were calculated with the Wilcoxon signed-rank test. GAN = generative adversarial network, IMR = iterative model reconstruction, SDCT = standard dose CT, ULDCT = ultralow dose CT

with sufficient image quality for the evaluation of lung diseases.

Strategies aimed at reducing radiation dose have become particularly important owing to increasing concerns about high radiation burdens and the risk of lifetime cancer associated with CT examinations. Several recent studies have focused on the development of efficient IR techniques to provide substantial dose reduction with acceptable image quality [22-24]. Messerli et al. [25] proved that ultralow-dose CT obtained at 100 kVp and 70 mAs and reconstructed with model-based IR allowed for lung volumetry and the quantification of emphysema. We performed the ULDCT at 80 kVp and 10 mAs with a mean ED of 0.12 mSv during this study. The feasibility of GAN in dose-reduced CT scans at the dose level of plain radiography may benefit patients who require multiple follow-up studies, especially young patients. Another study [24] reported that ultralow-dose CT with IMR resulted in approximately 60% dose reduction (ED, 0.3/0.7 mSv) with an improved depiction of abnormal findings of pulmonary invasive fungal infection, compared with low-dose CT. However, the sharpness of ground-glass opacities in model-based IR was rated as blurred for some cases. Likewise, the minor blotchy and pixelated appearances were also associated with ultralow-dose IMR images in this study. The subjective image rating of centrilobular nodules on ULDCT IMR images was lower than that of GAN-generated images. This may be due to the aggressive denoising by the advanced IR algorithm, which affected the sharp visualization and subjective analysis scores [26].

Compared with model-based IR approaches, the main advantage of the GAN method is that the network learns image statistics in an entirely data-driven manner instead of

 https://doi.org/10.3348/kjr.2020.0988

Table 5. Subjective Image Quality Scores of Visibility of Normal Anatomic Structure and Pathological Findings

Parameter	SDCT iDose	ULDCT			*P*		
		GAN	IMR	iDose	SDCT vs. GAN	GAN vs. IMR	GAN vs. iDose
Normal structures							
Peripheral bronchi and vessels	4.43 ± 0.50	4.20 ± 0.58	4.06 ± 0.59	2.71 ± 0.52	0.059	0.059	< 0.001
Proximal brochi and vessels	4.80 ± 0.41	4.60 ± 0.50	4.49 ± 0.51	3.43 ± 0.61	0.020	0.157	< 0.001
Fissures	4.37 ± 0.55	3.63 ± 0.55	3.31 ± 0.53	2.49 ± 0.56	< 0.001	0.002	< 0.001
Lymph nodes (n = 25)	4.38 ± 0.58	2.96 ± 0.36	3.46 ± 0.51	2.04 ± 0.20	< 0.001	0.021	< 0.001
Pericardium	4.46 ± 0.51	3.29 ± 0.46	3.26 ± 0.44	2.11 ± 0.32	< 0.001	0.739	< 0.001
Pathological findings							
Centrilobular nodules (n = 25)	4.64 ± 0.49	4.28 ± 0.46	4.00 ± 0.00	3.12 ± 0.33	0.013	0.008	< 0.001
Fibrosis (n = 21)	4.43 ± 0.51	4.14 ± 0.57	4.00 ± 0.00	2.71 ± 0.56	0.034	0.257	< 0.001
Calcification (n = 18)	4.67 ± 0.49	3.94 ± 0.64	4.28 ± 0.46	3.11 ± 0.58	0.002	0.014	0.036
Consolidation (n = 12)	4.92 ± 0.29	4.50 ± 0.52	4.08 ± 0.29	3.08 ± 0.52	0.025	0.025	0.002
Tree-in-bud (n = 7)	4.29 ± 0.49	4.00 ± 0.00	3.71 ± 0.49	2.86 ± 0.38	0.157	0.157	0.011
Cavitation (n = 6)	4.83 ± 0.41	4.17 ± 0.41	4.17 ± 0.41	3.00 ± 0.00	0.046	0.317	0.014

Data are mean ± standard deviation of the subjective scores. Grading for subjective image analysis of normal and pathological findings was performed evaluated on a 5-point scale (5 indicating best through to 1 indicating worst). *P* values for differences were calculated with the Wilcoxon signed-rank test. $P < 0.008$ indicates a statistically significant difference. GAN = generative adversarial network, IMR = iterative model reconstruction, SDCT = standard dose CT, ULDCT = ultralow dose CT

depending on hand-tuned regularizations. In recent years, GANs have gained much attention in medical imaging in many traditional and novel applications, such as synthesis, segmentation, reconstruction, classification, registration, and others [10]. GANs are special models in which a generator network generates realistic images to confuse the discriminator, while the discriminator network recognizes whether the generated images are real or fake, which work in competition with each other. For example, Wolterink et al. [17] reported that training with an adversarial network improved coronary artery calcification quantification over standard filtered back-projection reconstruction for low-dose cardiac CT. The quantitative PSNR results are consistent with ours. With a similar architecture, Ouyang et al. [27] proposed a conditional GAN-based model to generate a high-quality full-dose PET image from the corresponding low-dose PET image for amyloid status assessment. Consistent with recently published studies [12,17,28], our CycleGAN model converted ultralow-dose CT images into standard-dose visually similar images without any blurring or blotchy appearance. However, GAN for CT denoising may suffer from artifacts present in the synthesized images (8.6%). This can be attributed to the potential mode-collapsing behavior of the GAN when the generator network generates limited outputs regardless of the input [29]. Fortunately, the likelihood of the problem was significantly reduced after the implementation of an inverse path in the network architectures. Furthermore, the drift of CT density, a frequent problem of dose-reduced exams, was still not well resolved for the GAN technique in our study. The minimal alteration of attenuation values is directly related to the different tube voltages and reconstruction techniques [20].

GAN is a promising method for image-to-image translation with several variants, including WGAN [17], CycleGAN [16,29], and pix2pix conditional GAN [30]. In this study, breath movement between two successive CT scans resulted in a mismatch in pixel-to-pixel correspondence. Theoretically, CycleGAN is a favorable model for inconsistent input data in medical imaging generation [29]. This model, as part of the denoising task, could enable a good style transfer with or without matching paired data [18], which outperforms other paired approaches such as WGAN and pix2pix (Supplementary Materials). The smart design of reverse mapping and cycle consistent loss will facilitate a more accurate translation from inconsistent ultralow-dose to standard-dose visually similar CT images. In our study, the innovative inclusion of an extra L1 loss in the original CycleGAN model further strengthens the confidence in denoising performance. In this optimized model, the generator is designed not only to fool the discriminator but also to smoothen generated images using a Gaussian filter in an L1 sense. Previous approaches have found it useful to use the GAN model with a traditional loss function [31,32]. In addition, the CycleGAN approaches make real-time

https://doi.org/10.3348/kjr.2020.0988

Korean Journal of Radiology
KJR

reconstruction possible (0.05–0.07 second per image) once the adversarial network is trained.

Our study had several limitations. First, the sample size was relatively small. It is anticipated that an increase in training data will improve the performance of the deep learning framework. Second, this population included patients with a narrow body size range. Extremely large or obese patients need to be investigated in separate studies. Third, because of the inherent imaging characteristics of the different reconstruction methods, it was infeasible to completely blind the radiologists during subjective image analysis. Additionally, the higher attenuation of the pulmonary lesion (and chest wall) was noted in the GAN images. It may have affected the accuracy of the CT diagnosis in individual patients and should be investigated in future studies. Fourth, the SD of the CT values in the descending aorta was relatively high in the GAN-generated image, which may have been caused by large dosimetric errors around the paraspinal regions. CT density is critical for diagnosis, but the CT values of the pulmonary lesions measured in GAN images seem lower than those in other iteratively reconstructed images. This potentially limits the performance and generalizability of the current deep learning model in daily clinical practice. Further network design and optimization of this system are required. Lastly, the visibility of anatomic structures, including fissures and the pericardium, in the GAN-generated images was still not as good as that of standard-dose images. Future studies will focus on further improving the local structure of the GAN model to address this problem.

In conclusion, we developed and evaluated a CycleGAN-based deep learning technique for generating denoised ultralow-dose CT images with preserved image quality. The optimized CycleGAN technique may allow the synthesis of diagnostically acceptable images from ULDCT of the chest for the evaluation of pulmonary tuberculosis.

Supplement

The Data Supplement is available with this article at https://doi.org/10.3348/kjr.2020.0988.

Conflicts of Interest

Dr Philippe Lambin reports, within and outside the submitted work, grants/sponsored research agreements from Varian medical, Oncoradiomics, ptTheragnostic, Health Innovation Ventures and DualTpharma. He received an advisor/presenter fee and/or reimbursement of travel costs/external grant writing fee and/or in-kind manpower contribution from Oncoradiomics, BHV, Merck and Convert pharmaceuticals. Dr Lambin has shares in the company Oncoradiomics SA and Convert pharmaceuticals SA and is co-inventor of two issued patents with royalties on radiomics (PCT/NL2014/050248, PCT/NL2014/050728) licensed to Oncoradiomics and one issue patent on mtDNA (PCT/EP2014/059089) licensed to ptTheragnostic/DNAmito, three non-patentable invention (software) licensed to ptTheragnostic/ DNAmito, Oncoradiomics and Health Innovation Ventures. Dr Woodruff has (minority) shares in the company Oncoradiomics. Haixia Li and Tianjing Zhang serve in the Philips Healthcare. Hao Chen is an employee to the Jiangsu JITRI Sioux Technologies Co., Ltd. They all provided technical support on deep learning programming. None of the authors had control of the data and information that might present a conflict of interest for the other employee or consultant authors.

Acknowledgments

The authors thank Xiaochen Huai, MM, Philips Healthcare, China, as well as Jin Qi, PhD, School of Information and Communication Engineering, University of electronic science and technology of China, for their support in this study.

Author Contributions

Conceptualization: Yikai Xu, Philippe Lambin. Data curation: Chenggong Yan, Jie Lin. Formal analysis: Chenggong Yan, Haixia Li. Funding acquisition: Chenggong Yan, Jun Xu, Yikai Xu, Philippe Lambin. Investigation: Jun Xu, Guangyao Wu, Siqi Zhang. Methodology: Jun Xu, Hao Chen, Guangyao Wu. Project administration: Tianjing Zhang, Henry C. Woodruff. Resources: Jie Lin, Jun Xu, Siqi Zhang. Software: Haixia Li, Hao Chen. Supervision: Yikai Xu, Henry C. Woodruff, Philippe Lambin. Validation: Tianjing Zhang, Henry C. Woodruff. Visualization: Guangyao Wu. Writing—original draft: Chenggong Yan, Jie Lin, Haixia Li. Writing—review & editing: Jun Xu, Tianjing Zhang, Henry C. Woodruff, Guangyao Wu, Hao Chen, Siqi Zhang, Yikai Xu, Philippe Lambin.

ORCID iDs

Chenggong Yan
https://orcid.org/0000-0002-2968-4595
Jie Lin
https://orcid.org/0000-0003-1089-4143

Haixia Li
https://orcid.org/0000-0002-4818-4107
Jun Xu
https://orcid.org/0000-0001-7621-3617
Tianjing Zhang
https://orcid.org/0000-0002-7809-6845
Hao Chen
https://orcid.org/0000-0001-5887-039X
Henry C. Woodruff
https://orcid.org/0000-0001-7911-5123
Guangyao Wu
https://orcid.org/0000-0003-0658-2956
Siqi Zhang
https://orcid.org/0000-0002-1382-0708
Yikai Xu
https://orcid.org/0000-0003-1607-5077
Philippe Lambin
https://orcid.org/0000-0001-7961-0191

REFERENCES

1. Zumla A, George A, Sharma V, Herbert RH, Oxley A, Oliver M. The WHO 2014 global tuberculosis report—further to go. *Lancet Glob Health* 2015;3:e10-e12
2. Furin J, Cox H, Pai M. Tuberculosis. *Lancet* 2019;393:1642-1656
3. Jeong YJ, Lee KS. Pulmonary tuberculosis: up-to-date imaging and management. *AJR Am J Roentgenol* 2008;191:834-844
4. Sosa LE, Njie GJ, Lobato MN, Morris SB, Buchta W, Casey ML, et al. Tuberculosis screening, testing, and treatment of U.S. health care personnel: recommendations from the national tuberculosis controllers association and CDC, 2019. *MMWR Morb Mortal Wkly Rep* 2019;68:439-443
5. Kalra MK, Sodickson AD, Mayo-Smith WW. CT radiation: key concepts for gentle and wise use. *Radiographics* 2015;35:1706-1721
6. Stiller W. Basics of iterative reconstruction methods in computed tomography: a vendor-independent overview. *Eur J Radiol* 2018;109:147-154
7. Moloney F, James K, Twomey M, Ryan D, Grey TM, Downes A, et al. Low-dose CT imaging of the acute abdomen using model-based iterative reconstruction: a prospective study. *Emerg Radiol* 2019;26:169-177
8. Yamada Y, Jinzaki M, Tanami Y, Shiomi E, Sugiura H, Abe T, et al. Model-based iterative reconstruction technique for ultralow-dose computed tomography of the lung: a pilot study. *Invest Radiol* 2012;47:482-489
9. Yan C, Liang C, Xu J, Wu Y, Xiong W, Zheng H, et al. Ultralow-dose CT with knowledge-based iterative model reconstruction (IMR) in evaluation of pulmonary tuberculosis: comparison of radiation dose and image quality. *Eur Radiol* 2019;29:5358-5366
10. Shin YJ, Chang W, Ye JC, Kang E, Oh DY, Lee YJ, et al. Low-dose abdominal CT using a deep learning-based denoising algorithm: a comparison with CT reconstructed with filtered back projection or iterative reconstruction algorithm. *Korean J Radiol* 2020;21:356-364
11. Chen KT, Gong E, de Carvalho Macruz FB, Xu J, Boumis A, Khalighi M, et al. Ultra-low-dose 18F-florbetaben amyloid PET imaging using deep learning with multi-contrast MRI inputs. *Radiology* 2019;290:649-656
12. Zhao T, McNitt-Gray M, Ruan D. A convolutional neural network for ultra-low-dose CT denoising and emphysema screening. *Med Phys* 2019;46:3941-3950
13. Yin X, Zhao Q, Liu J, Yang W, Yang J, Quan G, et al. Domain progressive 3D residual convolution network to improve low-dose CT imaging. *IEEE Trans Med Imaging* 2019;38:2903-2913
14. Finck T, Li H, Grundl L, Eichinger P, Bussas M, Mühlau M, et al. Deep-learning generated synthetic double inversion recovery images improve multiple sclerosis lesion detection. *Invest Radiol* 2020;55:318-323
15. Kaplan S, Zhu YM. Full-dose PET image estimation from low-dose PET image using deep learning: a pilot study. *J Digit Imaging* 2019;32:773-778
16. Liang X, Chen L, Nguyen D, Zhou Z, Gu X, Yang M, et al. Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy. *Phys Med Biol* 2019;64:125002
17. Wolterink JM, Leiner T, Viergever MA, Isgum I. Generative adversarial networks for noise reduction in low-dose CT. *IEEE Trans Med Imaging* 2017;36:2536-2545
18. Yang Q, Yan P, Zhang Y, Yu H, Shi Y, Mou X, et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Trans Med Imaging* 2018;37:1348-1357
19. Pontana F, Billard AS, Duhamel A, Schmidt B, Faivre JB, Hachulla E, et al. Effect of iterative reconstruction on the detection of systemic sclerosis-related interstitial lung disease: clinical experience in 55 patients. *Radiology* 2016;279:297-305
20. Padole A, Digumarthy S, Flores E, Madan R, Mishra S, Sharma A, et al. Assessment of chest CT at CTDIvol less than 1 mGy with iterative reconstruction techniques. *Br J Radiol* 2017;90:20160625
21. Ye K, Zhu Q, Li M, Lu Y, Yuan H. A feasibility study of pulmonary nodule detection by ultralow-dose CT with adaptive statistical iterative reconstruction-V technique. *Eur J Radiol* 2019;119:108652
22. Tang H, Liu Z, Hu Z, He T, Li D, Yu N, et al. Clinical value of a new generation adaptive statistical iterative reconstruction (ASIR-V) in the diagnosis of pulmonary nodule in low-dose chest CT. *Br J Radiol* 2019;92:20180909
23. Hwang EJ, Park CM. Clinical implementation of deep learning in thoracic radiology: potential applications and challenges. *Korean J Radiol* 2020;21:511-525

24. Yan C, Xu J, Liang C, Wei Q, Wu Y, Xiong W, et al. Radiation dose reduction by using CT with iterative model reconstruction in patients with pulmonary invasive fungal infection. *Radiology* 2018;288:285-292
25. Messerli M, Ottilinger T, Warschkow R, Leschka S, Alkadhi H, Wildermuth S, et al. Emphysema quantification and lung volumetry in chest X-ray equivalent ultralow dose CT - Intra-individual comparison with standard dose CT. *Eur J Radiol* 2017;91:1-9
26. Padole A, Singh S, Ackman JB, Wu C, Do S, Pourjabbar S, et al. Submillisievert chest CT with filtered back projection and iterative reconstruction techniques. *AJR Am J Roentgenol* 2014;203:772-781
27. Ouyang J, Chen KT, Gong E, Pauly J, Zaharchuk G. Ultra-low-dose PET reconstruction using generative adversarial network with feature matching and task-specific perceptual loss. *Med Phys* 2019;46:3555-3564
28. Yi X, Babyn P. Sharpness-aware low-dose CT denoising using conditional generative adversarial network. *J Digit Imaging* 2018;31:655-669
29. Kang E, Koo HJ, Yang DH, Seo JB, Ye JC. Cycle-consistent adversarial denoising network for multiphase coronary CT angiography. *Med Phys* 2019;46:550-562
30. Koike Y, Akino Y, Sumida I, Shiomi H, Mizuno H, Yagi M, et al. Feasibility of synthetic computed tomography generated with an adversarial network for multi-sequence magnetic resonance-based brain radiotherapy. *J Radiat Res* 2020;6:92-103
31. Tie X, Lam SK, Zhang Y, Lee KH, Au KH, Cai J. Pseudo-CT generation from multi-parametric MRI using a novel multi-channel multi-path conditional generative adversarial network for nasopharyngeal carcinoma patients. *Med Phys* 2020;47:1750-1762
32. Kida S, Kaji S, Nawa K, Imae T, Nakamoto T, Ozaki S, et al. Visual enhancement of cone-beam CT by use of CycleGAN. *Med Phys* 2020;47:998-1010

 https://doi.org/10.3348/kjr.2020.0988