Title	Image synthesis with deep convolutional generative adversarial networks for material decomposition in dual-energy CT from a kilovoltage CT
Author(s)	Kawahara, Daisuke; Saito, Akito; Ozawa, Shuichi; Nagata, Yasushi
Citation	Computers in Biology and Medicine , 128 : 104111
Issue Date	2021-01
DOI	10.1016/j.compbiomed.2020.104111
Self DOI	
URL	https://ir.lib.hiroshima-u.ac.jp/00050458
Right	© 2021. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/ This is not the published version. Please cite only the published version. この論文は出版社版ではありません。引用の際には出版社版をご確認、ご利用ください。
Relation	

Title

Image synthesis with deep convolutional generative adversarial networks for material decomposition in dual-energy CT from a kilovoltage CT

Abstract

Generative Adversarial Networks (GANs) have been widely used and it is expected to use for the clinical examination and image. The objective of the current study was to synthesize material decomposition images of bone-water (bone(water)) and fat-water (fat(water)) reconstructed from dual-energy computed tomography (DECT) using an equivalent kilovoltage-CT (kV-CT) image and a deep conditional GAN. The effective atomic number images were reconstructed using DECT. We used 18,084 images of 28 patients divided into two datasets: the training data for the model included 16,146 images (20 patients) and the test data for evaluation included 1938 images (8 patients).

Image prediction frameworks of the equivalent single energy CT images at 120 kVp to the effective atomic number images were created. The image-synthesis framework was based on a CNN with a generator and discriminator. The mean absolute error (MAE), relative mean square error (MSE), relative root mean square error (RMSE), peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), and mutual information (MI) were evaluated. The Hounsfield unit (HU) difference between the synthesized and reference material decomposition images of bone(water) and fat(water) were within 5.3 HU and 20.3 HU, respectively. The average MAE, MSE, RMSE, SSIM, and MI of the synthesized and reference material decomposition of the bone(water) images were 0.8, 1.3, 0.9, 0.9, 55.3, and 0.8, respectively. The average MAE, MSE, RMSE, SSIM, and MI of the synthesized and reference material decomposition of the fat(water) images were 0.0, 0.0, 0.1, 0.9, 72.1, and 1.4, respectively. The proposed model can act as a suitable alternative to the existing methods for the reconstruction of material decomposition images of bone(water) and fat(water) reconstructed via DECT from kV-CT.

Keywords: Deep learning, Medical imaging, Artificial Intelligence, Dual-energy CT, Material Decomposition

Compliance with Ethical Standards:

Disclosure of potential conflicts of interest

Author Daisuke Kawahara declares that he has no conflict of interest.

Research involving human participants and/or animals

This article does not contain any studies with human participants or animals performed by any of the authors.

Ethnical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

Informed consent was obtained from all individual participants included in the study.

I. INTRODUCTION

Dual-energy computed tomography (DECT) uses two different energy spectra (low and high energy) that can be created by combining two datasets acquired over the same region [1]. DECT can obtain different information such as the effective atomic number, monochromatic energy CT (mCT) number, iodine-enhanced map, bone–water (bone(water)) density images, and fat–water (fat(water)) density images [2]. A (bone(water)) density image suppresses the water signal and enhances the calcium signal. It is created from the DECT data by estimating the amount of bone mineral, primarily composed of calcium, and subtracting this from the scanned original image [3, 4]. It is used for the diagnosis of bone marrow edema, which is a biomarker for arthritis, bone infarction, and hidden fractures. It is difficult to identify with traditional CT owing to the intrinsic low contrast of the involved tissues; it has been detected with magnetic resonance (MR) imaging [5]. MR represents the gold standard for soft tissue imaging and can provide quantitative fat-fraction measurements. A fat(water) image enhances the fat signal and suppresses other signals such as water and bone. Recently, Hyodo *et al*. reported that the DECT technique can estimate fat quantification in the liver [6].

A GE Revolution CT scanner (GE Healthcare, Milwaukee, WI) can reconstruct the effective atomic number, 120 kVp equivalent images, monochromatic energy CT, iodine contrast-enhanced, and calcium-enhanced images using a gemstone spectral imaging (GSI) technique [7]. The disadvantages of DECT are the increasing radiation dose, scan time, and cost.

Deep learning has been widely used for denoising applications [8, 9]. The denoising technique has been improved using a wavelet residual network, which synergistically combines the expressive power of deep learning and the performance guarantee from framelet-based algorithms [10]. The improvement of the image resolution has led to a reduction in the radiation dose. Convolutional neural networks (CNNs) have been successfully applied to image synthesis and image processing. Dong *et al*. performed super-resolution imaging using a CNN algorithm. Streak artifacts due to beam hardening and photon starvation have been potentially problematic. Zhang *et al*. suppressed the artifacts dramatically using the CNN-based metal artifact reduction framework, which fuses the information from the original and corrected images. For material decomposition, deep learning plays a significant role. Liao and Lyu simulated pseudo-high-energy images from low-energy CT images to improve the quality of the material decomposition with a simple U-Net architecture [11, 12]. Clark *et al*. used multi-energy CT with DECT and spectral CT for material decomposition with a U-Net-based CNN architecture [13]. Another approach to the crossover architecture that incorporates two material generation pathways for the bone(water) density image and water–bone (water(bone)) density images was introduced by Zhang *et al*. [14]. It used both kV-CT images at 80 and 140 kVp that used the DECT scan. These studies did not directly predict material decomposition images from the single energy CT (SECT) images.

In a recent study, an image-synthesis technique of cross modality with a generative adversarial network (GAN) was performed. GANs function by training two different networks: a generator network synthesizes an image, and a discriminator network distinguishes between the synthesized and reference images [15]. Florkow *et al*. proposed an image-synthesis framework of MR images to CT images with a two-dimensional (2D) CNN model [16]. For radiotherapy, the synthesis of PlanCT-like images from Cone beam computed tomography (CBCT) images with planning CT and CBCT datasets with GAN to improve the image quality of the CBCT was introduced in [17].

The current study proposes an image-synthesis approach to material decomposition images of bone(water) and fat(water) reconstructed on DECT from the SECT of an equivalent kilovoltage CT (kV-CT) image at 120 kVp directly using a GAN-based CNN architecture.

II. MATERIALS AND METHODS

A) Data acquisition

The DECT image for each patient was acquired with a Revolution DECT scanner (GE Healthcare, Princeton, NJ, USA). The DECT scans were performed at tube voltages of 80 and 140 kVp and exposures of 560 mA. The other scanning parameters were a rotation time (RT) of 1.0 s, slice thickness of 0.5 mm, and field of view (FOV) of 360 mm. The material decomposition images of bone(water) and fat(water) and equivalent kV-CT images were reconstructed using the GSI technique. A total of 18,084 images from 28 patients were analyzed as part of an institutional review board-approved study.

B) Deep learning model

A 16-bit Digital Imaging and Communications in Medicine (DICOM) image was converted to an 8-bit red-green-blue (RGB) portable network graphics (PNG) image, and the output 8-bit RGB PNG image from the 2D CNN model was converted to 16-bit DICOM images. The pixel number in the CT image ranged from -1000 to 3079 Hounsfield units (HU). The unused pixel value was eliminated. Subsequently, the values of the pixels in the CT images were converted to 8-bit (0–255) images by dividing by 16, which is the value obtained by dividing the maximum pixel value, that is, 3079 HU, by 256. The process of radiomics analysis is presented in Fig. 1. The pixel values of the

DECT and kV-CT images were rescaled using the RescaleSlope and RescaleIntercept tags from the DICOM header as follows:

$$Image\ Data = (Image\ Data) \times RescaleSlope + RescaleIntercept + 1000. \quad (1)$$

Before calculating the radiomics features, we applied a medium smooth filter to the rescaled image data. An overview of the GAN network model is displayed in Fig. 1. This includes a generator (to estimate the material decomposition image) and discriminator (to distinguish the real material decomposition image from the generated image). The generator attempts to produce realistic images that confuse the discriminator. These CNN networks are trained simultaneously by evaluating $\theta_{G,D}$. The generator comprised an encoder and decoder. The encoder mapped the image from the input image of 512 × 512 resolution using a stack of eight convolutional layers, each followed by LeakyReLU activation functions and batch normalization. The number of convolutional filters was 64, 128, 256, 512, 1024, 1024, 1024, and 1024, respectively, with a kernel size of 4 × 4 and stride of 2. The decoder mirrored the encoding architecture, albeit utilizing fractionally strided convolution (deconvolution). The number of the deconvolution filters were 1024, 1024, 1024, 1024, 512, 256, 128, and 64, respectively, in the layers that utilized ReLU activation functions. The discriminator used seven convolution layers to extract the features from the image and generate the output image. The number of convolutional filters was 64, 128, 256, 512, 1024, 1024, and 1024, respectively, with a kernel size of 4 × 4 and stride of 2.

The deep learning model was a conditional GAN that required paired images from the kV-CT and DECT images that were co-registered with voxel-wise correspondence. The label was the kV-CT image before synthesis. The loss was evaluated using the generator and discriminator.

$$\mathcal{L}_{\mathrm{GAN}}(\mathrm{G,D}) = \mathbb{E}_{\mathrm{y}}[\log \mathrm{D}(\mathrm{y})] + \mathbb{E}_{\mathrm{x,z}}[\log(1 - \mathrm{D}(\mathrm{G}(\mathrm{x}))] \quad (2)$$

where *G* is the generator network, $\mathbb{E}$ is the expectation value dependent on both *x*, the set of kV-CT images, and *y*, the set of target images that are DECT images. Moreover, it included an additional loss based on the absolute difference between the synthesized DECT image and input kV-CT image (L1 norm loss). The L1 norm loss was calculated as

$$\mathcal{L}_{\mathrm{L1}}(\mathrm{G}) = \mathbb{E}_{\mathrm{x,y}}(|\mathrm{y} - \mathrm{G}(\mathrm{x})|_1). \quad (3)$$

The adversarial loss in Eq. (1) was calculated using the binary cross-entropy cost function. The final cost function used to optimize the network was a weighted summation of the losses in Eq. (2) and Eq. (3):

$$\theta_{\mathrm{G,D}} = \arg \underbrace{min}_{G} \underbrace{max}_{D} (\mathcal{L}_{\mathrm{L1}}(\mathrm{G}) + \mathcal{L}_{\mathrm{GAN}}(\mathrm{G,D})). \quad (4)$$

Here, hyperparameter lambda is the weighting factor for the L1 loss, which was set to 100. The proposed models were implemented using TensorFlow packages (V1.7.0, Python 2.7, CUDA 9.0) on a Ubuntu 16.04 LTS system. Three hundred epochs were used to operate the model on an 11-GB NVIDIA GeForce GTX 1080 GPU. All three models were trained with instance normalization and identical hyperparameters, except for the batch size. For each iteration, a mini-batch of 2D images was randomly selected from the training set.

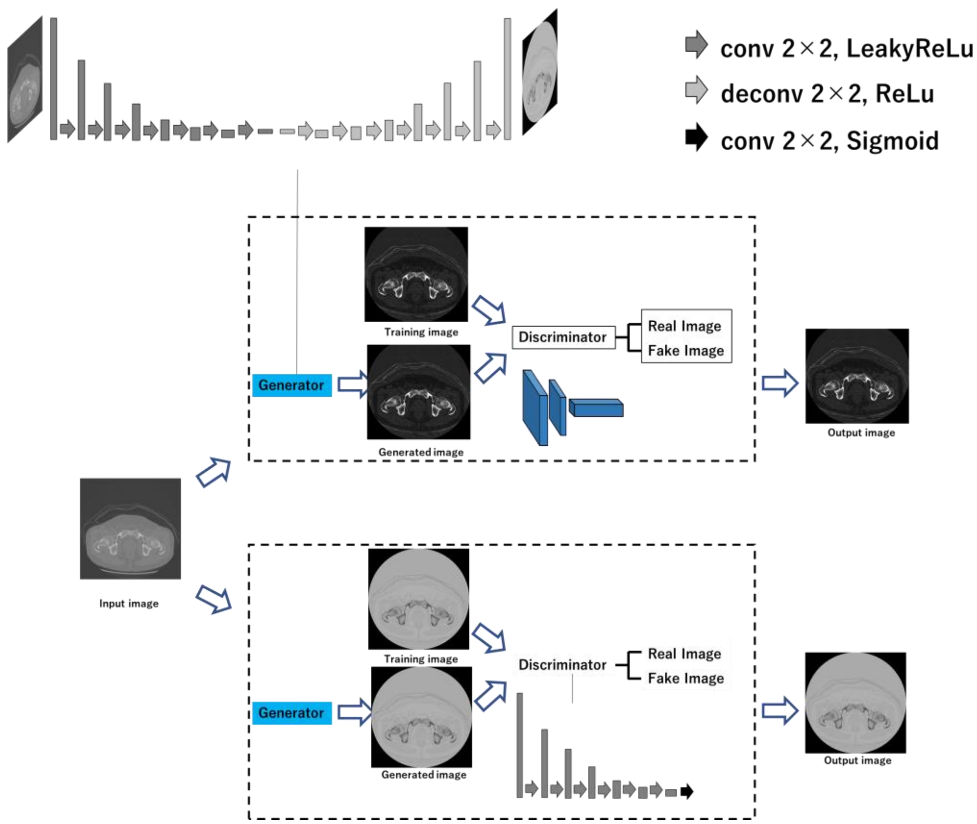

Fig. 1 GAN framework of the material decomposition images of bone(water) (upper) and fat(water) (lower). Given an input image in a source contrast, the

generator learns to generate an image of similar anatomy in a target contrast and the discriminator learns to discriminate between the synthesized and real pairs of the material decomposition images.

C) Evaluation

The prediction accuracy of the model for the synthesized and reference material decomposition images of bone(water) and fat(water) was evaluated using the following five metrics: relative mean absolute error (MAE), relative root mean square (RMSE), structural similarity index (SSIM), signal-to-noise ratio (PSNR), and mutual information (MI). These metrics are defined as follows:

$$MAE = \frac{1}{n_x n_y} \sum_{i,j}^{n_x n_y} \frac{|r(i,j) - t(i,j)|}{r(i,j)}. \quad (5)$$

Here, $r(i,j)$ is the value of pixel (i,j) in the planning CT image, $t(i,j)$ is the value of pixel (i,j) in the target image, and $n_x n_y$ is the total number of pixels. RMSE is defined as

$$RMSE = \sqrt{\frac{1}{n_x n_y} \sum_{i,j}^{n_x n_y} \left(\frac{r(i,j) - t(i,j)}{r(i,j)}\right)^2}. \quad (6)$$

The SSIM is computed based on consideration of the contrast, structure, and luminance to compute a similarity score between two images.

The SSIM between two images $\vec{x}$ and $\vec{y}$ can be computed as [18]

$$SSIM(\vec{x}, \vec{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{({\mu_x}^2 + {\mu_x}^2 + C_1)({\sigma_x}^2 + {\sigma_y}^2 + C_2)}, \tag{7}$$

$$C_1 = (k_1 Q)^2, \quad k_1 = 0.01, \tag{8}$$

$$C_2 = (k_2 Q)^2, \quad k_2 = 0.03, \tag{9}$$

where C_1 and C_2 are constants used to prevent a zero denominator and to maintain the stability of the formula. Q is the maximum CT value for the synthesized and reference images. The values of k_1 and k_2 are typically obtained from [19]. σ_x is an estimate in the discrete form

$$\sigma_x = \left(\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)^2\right)^{1/2}. \tag{10}$$

The correlation coefficient between $\vec{x}$ and $\vec{y}$ is defined as σ_{xy}. It is expressed as

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y), \tag{11}$$

where μ_x is the mean intensity and can be expressed as

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i. \tag{12}$$

The PSNR is calculated as

$$PSNR_{GL} = 10 \times log_{10}\left(\frac{(MAX)^2}{MSE}\right). \tag{13}$$

Here, MAE and MSE are the possible maximum signal intensity and mean square error (or difference) of the image, respectively. MI is used as a cross-modality similarity measure [20]. It is calculated as

$$I(r\colon t) = \sum_{m \in I_r} \sum_{n \in I_t} p(m,n) log\left(\frac{p(m,n)}{p(m)p(n)}\right), \tag{14}$$

where m and n are the intensities in the targeted monochromatic energy CT image I_r and predicted monochromatic energy CT image I_t, respectively. $p(m, n)$ is the joint probability density of I_r and I_t, whereas $p(m)$ and $p(n)$ are marginal densities. $p(m, n)$ can be calculated as follows:

$$p(m,n) = \frac{h(m,n)}{\sum_{m \in I_t} \sum_{n \in I_t} h(m,n)}, \tag{15}$$

where $h(m,n)$ is the histogram of the pixel values in the reference monoenergetic CT image I_r and synthesized monoenergetic CT image I_t. Furthermore, the difference in the synthesized and reference monoenergetic CT numbers in the region of interest (ROI) was evaluated for several slices, from the feet to chest, in a manually drawn ROI, as depicted in Fig. 2.

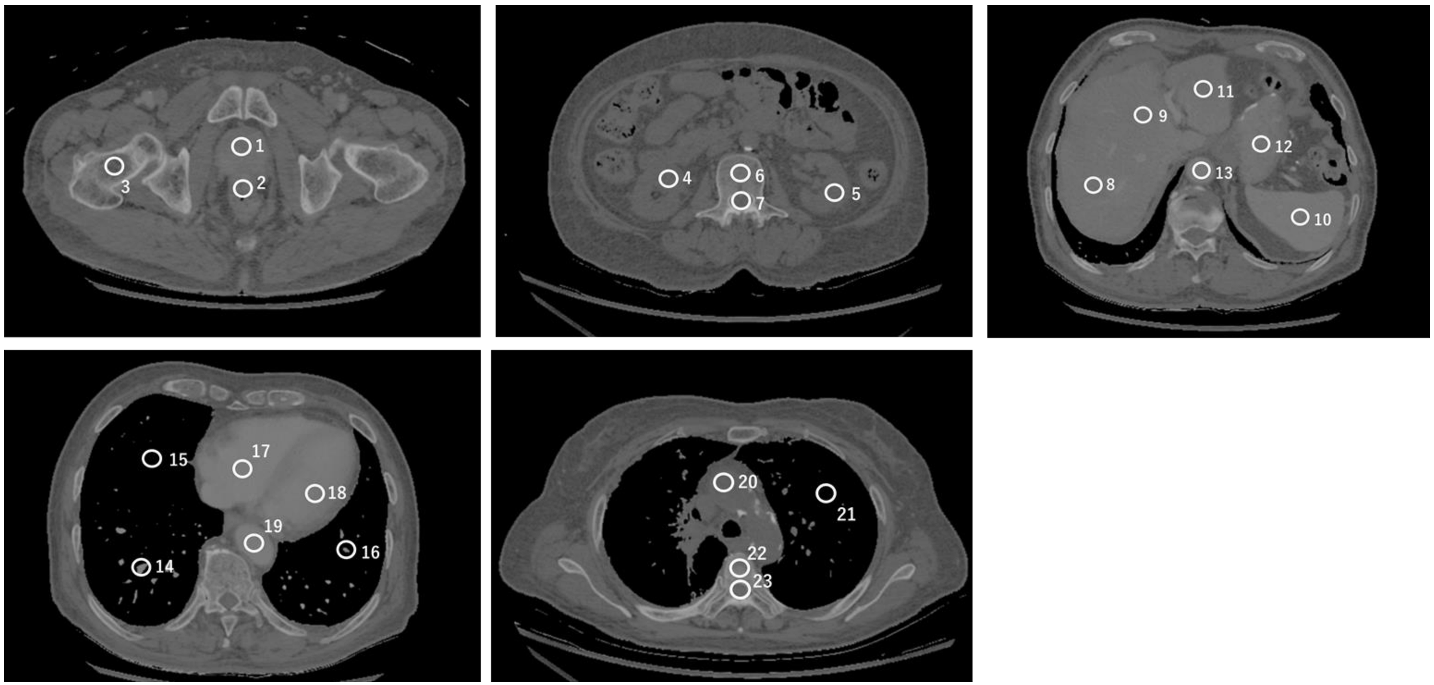

Fig. 2 Method of measurement in the evaluation of the HU in the material decomposition images from feet to chest slice. The average and SD values of the HU were measured by creating a circular ROI, 2 cm in diameter.

III. RESULTS

The time required to create the image synthesized model was approximately 142.2 ± 3.1 h for conversion of the kV-CT to the material decomposition images of the bone(water) and fat(water). The rate to create the synthesized monochromatic energy CT images using all the trained models was approximately 7.2–8.1 images/s. The generator loss, discriminator loss, and L1 norm loss in each prediction model are displayed in Figs. 3 and 4.

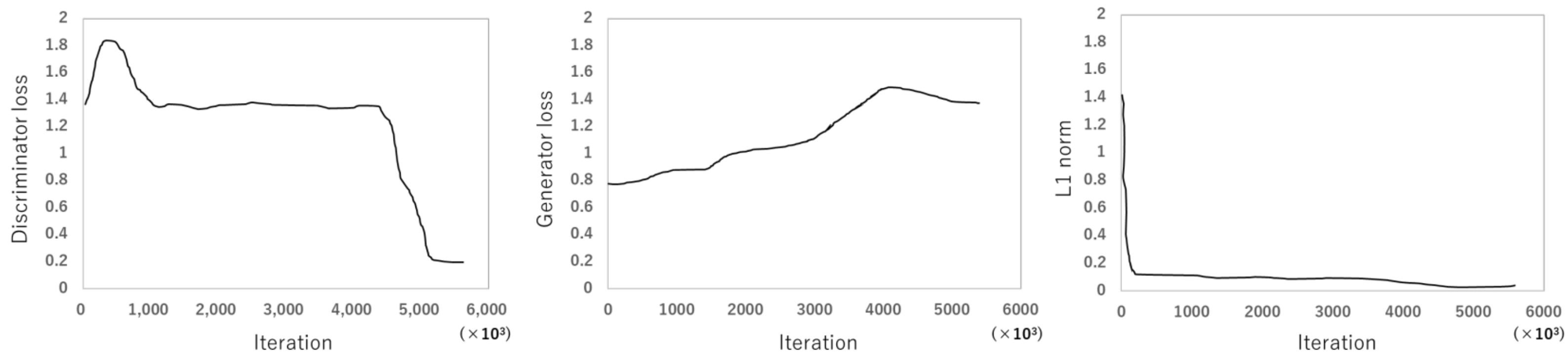

Fig. 3 Average training losses in the generator and discriminator in the CT-based prediction model for conversion of kV-CT to the material decomposition images of bone(water).

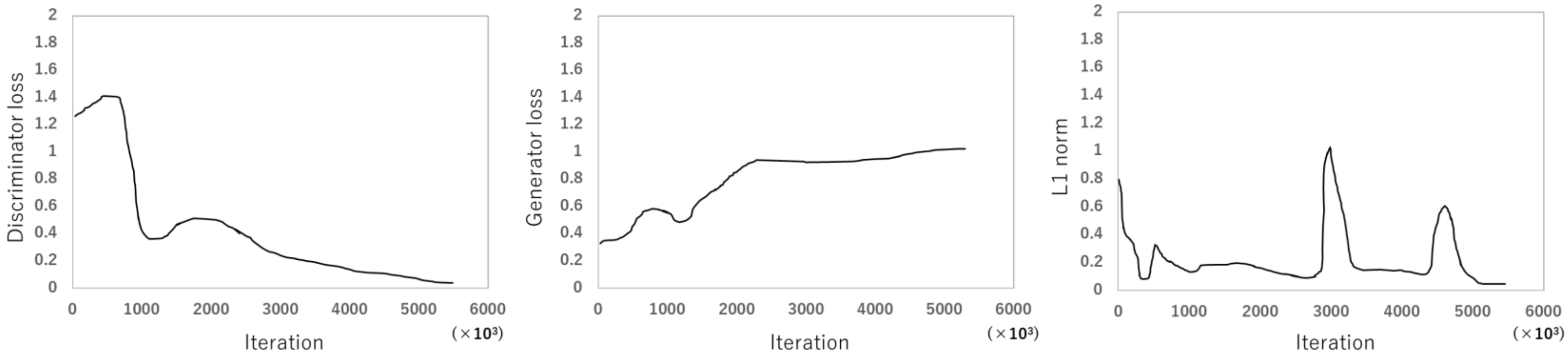

Fig. 4 Average training losses in the generator and discriminator in the CT-based prediction model for conversion of kV-CT to the material decomposition images of fat(water).

Figs. 5–10 display samples obtained by cross-modality generation for the synthesized and reference material decomposition images of bone(water) and fat(water). Table 1 presents the difference in HU values between the synthesized and reference material decomposition images of bone(water). The difference between the synthesized and reference material decomposition images is within 5.3 HU. The difference of the monochromatic energy CT number is within the appropriate range of the SD values in all ROIs. Table 2 indicates the HU difference between the synthesized and reference material decomposition images of fat(water). The difference of the synthesized and reference material decomposition images is within 20.3 HU for the fat(water) images. The difference of the monochromatic energy CT number is within the appropriate range of the SD values in all ROIs. The RMSE in all ROIs was 0.6 for the material decomposition image of bone(water) and 1.2 for the material decomposition image of fat(water).

Tables 3 and 4 display the average MAE, MSE, RMSE, PSNR, and MI computed from feet to chest slices for the material decomposition images of bone(water) and fat(water). The MAE, MSE, and RMSE were less for the material decomposition image of fat(water). The PSNR and MI were greater for the material decomposition image of fat(water). There was no difference in the SSIM.

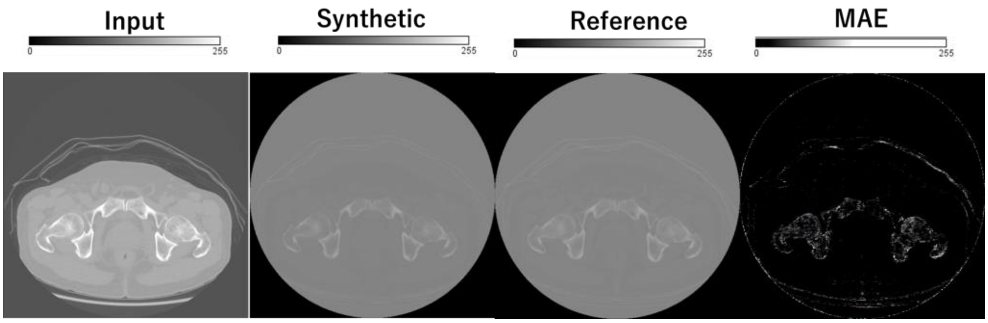

Fig. 5 Samples obtained from a material decomposition image of bone(water) at pelvic level: input image is the equivalent kV-CT image at 120 kVp, synthesized and reference images are the material decomposition images of bone(water), and MAE is the difference between the synthesized and reference bone(water) images.

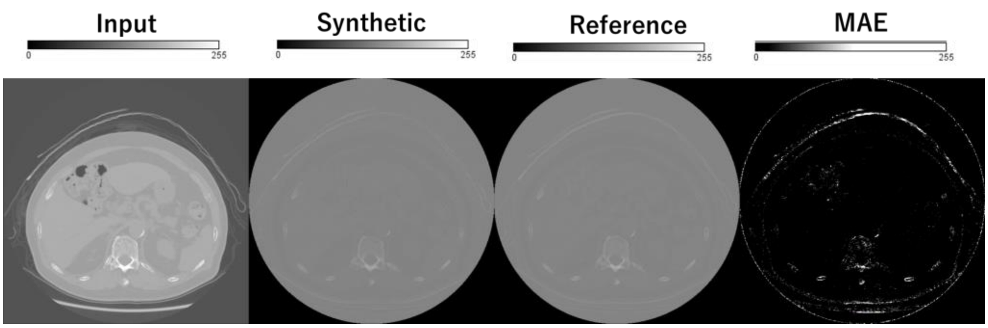

Fig. 6 Samples obtained from a material decomposition image of bone(water) at abdominal level: input image is the equivalent kV-CT image at 120 kVp, synthesized and reference images are the material decomposition images of bone(water), and MAE is the difference between the synthesized and reference bone(water) images.

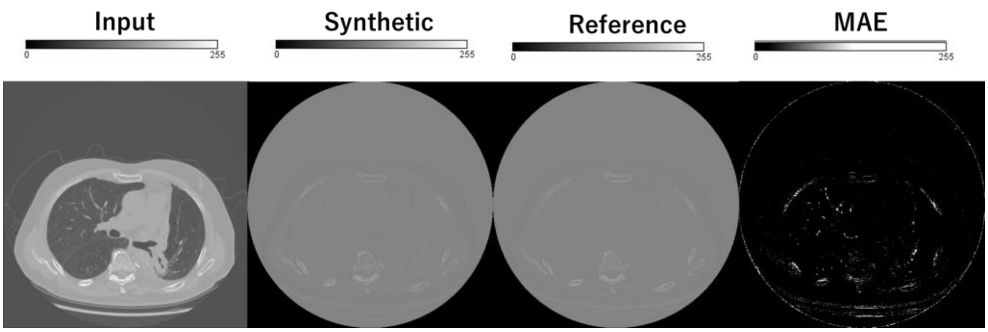

Fig. 7 Samples obtained from a material decomposition image of bone(water) at chest level: input image is the equivalent kV-CT image at 120 kVp, synthesized and reference images are the material decomposition images of bone(water), and MAE is the difference between the synthesized and reference bone(water) images.

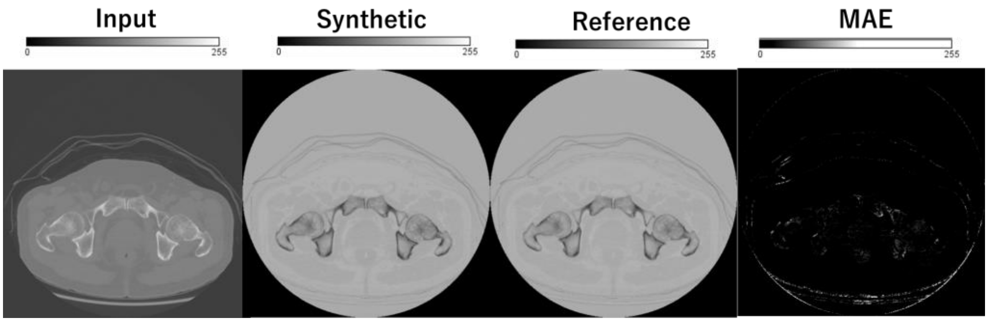

Fig. 8 Samples obtained from a material decomposition image of fat(water) at pelvic level: input image is the equivalent kV-CT image at 120 kVp, synthesized and reference images are the material decomposition images of fat(water), and MAE is the difference between the synthesized and reference fat(water) images.

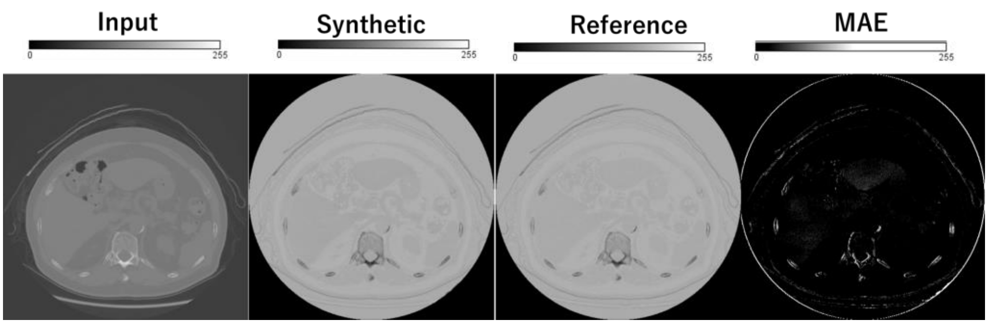

Fig. 9 Samples obtained from a material decomposition image of fat(water) at abdominal level: input image is the equivalent kV-CT image at 120 kVp, synthesized and reference images are the material decomposition images of fat(water), and MAE is the difference between the synthesized and reference fat(water) images.

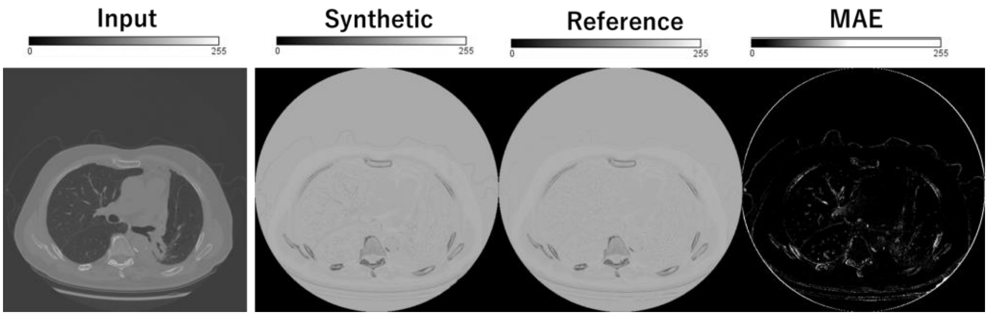

Fig. 10 Samples obtained from a material decomposition image of bone(water) at pelvic level: input image is the equivalent kV-CT image at 120 kVp, synthesized and reference images are the material decomposition images of fat(water), and MAE is the difference between the synthesized and reference fat(water) images.

Table 1. Difference (Δ) and SD values of synthesized and reference material decomposition of bone(water).

	Bone(water)	
	Δ (HU)	SD (HU)
①	1.7	5.0
②	0.8	11.5
③	3.0	26.1
④	-1.8	5.2
⑤	-2.2	4.1

⑥	1.2	11.7
⑦	-1.1	5.84
⑧	1.4	3.9
⑨	2.7	4.5
⑩	3.0	4.8
⑪	-0.9	5.1
⑫	2.9	4.4
⑬	4.6	4.8
⑭	-0.1	4.5
⑮	3.2	5.7
⑯	5.3	6.2
⑰	-0.3	4.1
⑱	-3.2	4.3
⑲	-0.1	4.4
⑳	1.3	3.6
㉑	-4.6	14.3
㉒	3.4	15.5
㉓	-1.1	5.1

Table 2. Difference (Δ) and SD values between synthesized and reference material decomposition of fat(water).

	Fat(water)	
	Δ (HU)	SD (HU)
①	-5.3	9.3
②	-5.0	8.2
③	-7.4	7.8
④	-3.6	31.5

⑤	-6.7	12.9
⑥	-8.9	14.8
⑦	-2.9	9.8
⑧	-5.9	9.7
⑨	4.0	17.6
⑩	-0.7	18.1
⑪	0.0	58.5
⑫	-3.6	9.1
⑬	-6.9	8.9
⑭	-8.1	9.2
⑮	-5.9	20.8
⑯	-3.1	8.6
⑰	-6.4	10.6
⑱	-1.7	12.3
⑲	5.3	12.5
⑳	-3.0	42.9
㉑	7.5	182.0
㉒	-7.0	51.9
㉓	-8.8	93.1

Table 3. Average MAE, MSE, RMSE, PSNR, and SSIM computed from feet to chest slices for the material decomposition images of bone(water).

MAE		MSE		RMSE		PSNR		SSIM		MI	
Average	SD	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD
0.8	0.7	1.3	1.2	0.9	0.7	55.3	15.0	0.9	0.1	0.8	0.1

Table 4. Average MAE, MSE, RMSE, PSNR, and SSIM computed from feet to chest slices for the material decomposition images of fat(water).

MAE		MSE		RMSE		PSNR		SSIM		MI	
Average	SD	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD
0.0	0.0	0.0	0.0	0.1	0.0	72.1	3.4	0.9	0.0	1.4	0.1

IV. DISCUSSION

DECT enables the separation of several additional materials including calcium, fat, and uric acid from a single kV-CT. It provides anatomic knowledge with functional information [21]. DECT requires post-processing to obtain material decomposition images. It requires 5 to 10 min of additional interpretation time after the scan [22]. Conversely, the proposed image-synthesized system can reconstruct a DECT image within 1 min automatically. GE scanners use dual X-ray sources; Siemens Healthiness scanners use dual X-ray sources and two data acquisition systems [23, 24]. Thus, the dual-source scanner cost is considerably greater than the standard SECT scanner. The Philips Healthcare scanner acquires DECT projection data using a layered detector. The high-energy data and low-energy data are collected by the posterior and anterior detector layers, respectively [25]. All DECT data acquisition techniques pose a significant burden on the CT system hardware. Zhao *et al*. reported that, due to this fact, DECT scanners are not widely used in less-developed regions [26]. In addition to the increased cost and complexity of the imaging system, DECT can also increase the radiation dose owing to the additional CT scan. The image-synthesized approach with deep learning is expected to reduce the scanning radiation dose and imaging cost by synthesizing the DECT from the single kV-CT image.

Li *et al*. proposed a new image-domain method for DECT decomposition that combines conventional penalized weighted-least squares estimation with regularization based on a mixed union of learned transforms model [27]. The RMSE of the density image was 0.04–0.05. To compare the RMSE on the same scale, the material decomposition image was converted to a density image with a theoretical value. Here, the theoretical density values were 1 g/cm^3 for the water and 1.92 g/cm^3 for the bone used by Li. The RMSE in all ROIs was 0.55 for the material decomposition image of bone(water) and 1.2 for the material decomposition image of fat(water). The RMSE of the density value in all ROIs was 0.01 for the material decomposition image of bone(water) and 0.02 for the material decomposition image of fat(water), which was significantly less than that of Li *et al*. Although other studies have proposed the image synthesis of a DECT, the accuracy of the image synthesis could not be directly compared because there were no studies evaluating the HU value. Moreover, previous studies used other reconstruction images such as low- and high-energy kV-CT, multi-energy images, and virtual non-contrast images [11–13]. These models required multiple images or additional reconstruction. The current study proposed a prediction model of the material decomposition images of bone(water) and fat(water) from a single kV-CT image with GAN architectures. The HU difference between the reference and synthesized material decomposition images of bone(water) and fat(water) were less than 5.3 HU and 20.3 HU, respectively. The material decomposition images of the bone(water) had a smaller MAE, MSE, and RMSE, and a greater PSNR and MI than the fat(water). The bone was highlighted and the other organ with a similar density to water in the material decomposition images of the bone(water), were less prominent. For the U-Net-based CNN employed by the previous study, the label was image-wise. Conversely, the model proposed in the current study used the label that was pixel-wise. Although the previous study could extract local imaging features, it was required to register the input and output images for model training. The current study used the kV-CT and material decomposition images reconstructed from the DECT image. No differences were observed in the alignment of the images.

There is a possibility that certain patients could be affected by the beam-hardening artifact. This could cause errors in the model training in the correlation of the training of the kV-CT and material decomposition images of the bone(water). However, these differences in each ROI were within the SD range. The proposed model could produce highly accurate DECT images within the noise estimations from the kV-CT images.

There are limitations to the current study. First, the current study used 120 kV-CT images reconstructed from DECT. The difference in the image quality of 120 kV-CT images scanned via SECT and the equivalent 120 kV-CT images reconstructed from DECT for clinical patients were evaluated by Tawfil *et al*. The subjective image quality scores between the DECT and SECT groups did not indicate a significant difference. Thus, 120 kV-CT images from the DECT images can be used as equivalent to the SECT images [28]. Moreover, the patients used for the current study were randomly selected, regardless of the presence or absence of disease. Therefore, evaluation of the lesion detectability could not be performed. In addition to evaluating the image similarities for complete images, we confirmed that the HU difference was within the noise in the local region from the pelvic to the chest level. Further studies will be performed to examine the quality of the synthesized images compared to the original images in terms of diagnostic performance.

V. CONCLUSION

The current study proposed an image-synthesis framework using a GAN-based CNN architecture for kV-CT to material decomposition images of bone(water) and fat(water) scanned by DECT. The proposed image synthesis model showed a highly image quality and the difference of the monochromatic energy CT number is within the appropriate range of the SD values in the local region. Synthesized medical image generation can be a cost-effective approach for developing automated diagnostic technologies.

References

1. Alvarez RE, Macovski A (1976) Energy-selective reconstructions in X-ray computerized tomography. Phys Med Biol 21:733–744.
2. Johnson, T. R. C. *et al*. Material differentiation by dual energy CT: initial experience. European Radiology 2006; 17: 1510–1517.
3. McCollough, C. H., Leng, S., Yu, L. & Fletcher, J. G. Dual- and Multi-Energy CT: Principles, Technical Approaches, and Clinical Applications. Radiology 2015; 276: 637–653.
4. Guggenberger R, Gnannt R, Hodler J, *et al*. (2012) Diagnostic performance of dual-energy CT for the detection of traumatic bone marrow lesions in the ankle: comparison with mr imaging. Radiology 264:164–173.
5. Coriasco M, Rampado O, Bradac GB. Elementi di risonanza magnetica, dal protone alle sequenze per le principali applicazioni diagnostiche. Springer (2014).
6. Hyodo T, Yada N, Hori M1, *et al*. Multimaterial Decomposition Algorithm for the Quantification of Liver Fat Content by Using Fast-Kilovolt-Peak Switching Dual-Energy CT: Clinical Evaluation. Radiology. 2017 Apr;283(1):108-118.

7. Slavic *et al*, Technology White Paper, GSI Xtream on Revolution (TM) CT, 2017.

8. Chen Y, Yu W, Pock T. On learning optimized reaction diffusion processes for effective image restoration. In: Proceeding-IEEE Conference on Computer Vision and Pattern Recognition. 2015;1: 5261-5269.

9. Wang Z, Bovik AC, Sheikh HR, et al. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process. 2004;13:1–14.

10. Kang E, Chang W, Yoo J, *et al*. Deep Convolutional Framelet Denosing for Low-Dose CT via Wavelet Residual Network. IEEE Trans Med Imaging, 37(6):1358-1369.

11. Liao Y, Wang Y, Li S, *et al*. Pseudo dual energy CT imaging using deep learning-based framework: basic material estimation. Medical Imaging 2018: Physics of Medical Imaging 10573.

12. Lyu T, Wu Z, Zhang Y, *et al*. Dual-energy CT imaging from single-energy CT data with material decomposition convolutional neural network. arXiv:2006.00149v1.

13. Clark DP, Holbrook M, Badea CT. Multi-energy CT decomposition using convolutional neural networks. Proceedings of the SPIE. 2018;10573.

14. Zhang W, Zhan H, Wang L, *et al*. Image domain dual material decomposition for dual-energy CT using butterfly network. Med Phys. 2019 May;46(5):2037-2051.

15. Goodfellow I, Pouget-Abadie J, Mirza M, *et al*. Generative adversarial nets. In: Advances in neural information processing systems; 2014:2672–2680.

16. Florkow MC, Zijlstra F, Willemsen K, *et al*. Deep learning-based MR-to-CT synthesis: The influence of varying gradient echo-based MR images as input channels. Magn Reson Med. 2020 Apr;83(4):1429-1441.

17. Kida S, Kaji S, Nawa K, *et al*. Cone-beam CT to Planning CT synthesis using generative adversarial networks. arXiv:1901.05773v1.

18. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process. 2004;13:1–14.

19. Qi M, Li Y, Wu A, et al. Multi-sequence MR image-based synthetic CT generation using a generative adversarial network for head and neck MRI-only radiotherapy. Med Phys. 2020 Apr;47(4):1880-1894.

20. Zhao W, Lv T, Gao P, *et al*. Dual-energy CT imaging using a single-energy CT data is feasible via deep learning. ArXiv. 2019; 1906.04874.

21. Patino M, Prochowski A, Agrawal MD, Material Separation Using Dual-Energy CT: Current and Emerging Applications. Radiographics. 2016 Jul-Aug;36(4):1087-105.

22. Ohana M, Jeung MY, Labani A, *et al*. Thoracic dual energy CT: acquisition protocols, current applications and future developments. Diagn Interv Imaging. 2014 Nov;95(11):1017-26.

23. Matsumoto K, Jinzaki M, Tanami Y, Ueno A, Yamada M and Kuribayashi S, Virtual monochromatic spectral imaging with fast kilovoltage switching: improved image quality as compared with that obtained with conventional 120-kVp CT, Radiology 259, 257 (2011).

24. Johnson TR, Krauss B, Sedlmair M, Grasruck M, Bruder H, Morhard D, Fink C, Weckbach S, Lenhard M, Schmidt B *et al*., Material differentiation by dual energy CT: initial experience, European Radiology 17, 1510 (2007).

25. Hao J, Kang K, Zhang L and Chen Z, A novel image optimization method for dual-energy computed tomography, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 722, 34 (2013).

26. Zhao W, Lv T, Lee R, *et al*. Obtaining dual-energy computed tomography (CT) information from a single-energy CT image for quantitative imaging analysis of living subjects by using deep learning. Pac Symp Biocomput. 2020;25:139-148.

27. Li Z, Ravishankar S, Long Y, et al. DECT-MULTRA: Dual-Energy CT Image Decomposition With Learned Mixed Material Models and Efficient Clustering. arXiv:1901.00106v2.

28. Tawfik AM, Kerl JM, Razek AA, *et al*. Image quality and radiation dose of dual-energy CT of the head and neck compared with a standard 120-kVp acquisition. AJNR Am J Neuroradiol. 2011 Dec;32(11):1994-9.

Figure captions

Fig. 1. GAN framework of the material decomposition images of bone(water) (upper) and fat(water) (lower). Given an input image in a source contrast, Generator learns to generate an image of similar anatomy in a target contrast and Discriminator learns to discriminate between synthesized and real pairs of the material decomposition images.

Fig. 2 Method of measurement in the evaluation of the HU in the material decomposition images from feet to chest slice. The average and SD values of the HU were measured by creating a circular ROI, 2 cm in diameter.

Fig. 3 Average training losses in the generator and discriminator in the CT-based prediction model for conversion of kV-CT to the material decomposition images of bone(water).

Fig. 4 Average training losses in the generator and discriminator in the CT-based prediction model for conversion of kV-CT to the material decomposition images of fat(water).

Fig. 5 Samples obtained from a material decomposition image of bone(water) at pelvic level: input image is the equivalent kV-CT image at 120 kVp, synthesized and reference images are the material decomposition images of bone(water), and MAE is the difference between the synthesized and reference bone(water) images.

Fig. 6 Samples obtained from a material decomposition image of bone(water) at abdominal level: input image is the equivalent kV-CT image at 120 kVp, synthesized and reference images are the material decomposition images of bone(water), and MAE is the difference between the synthesized and reference bone(water) images.

Fig. 7 Samples obtained from a material decomposition image of bone(water) at chest level: input image is the equivalent kV-CT image at 120 kVp, synthesized and reference images are the material decomposition images of bone(water), and MAE is the difference between the synthesized and reference bone(water) images.

Fig. 8 Samples obtained from a material decomposition image of bone(water) at pelvic level: input image is the equivalent kV-CT image at 120 kVp, synthesized and reference images are the material decomposition images of bone(water), and MAE is the difference between the synthesized and reference bone(water) images.

Fig. 9 Samples obtained from a material decomposition image of bone(water) at abdominal level: input image is the equivalent kV-CT image at 120 kVp, synthesized and reference images are the material decomposition images of bone(water), and MAE is the difference between the synthesized and reference bone(water) images.

Fig. 10 Samples obtained from a material decomposition image of bone(water) at pelvic level: input image is the equivalent kV-CT image at 120 kVp, synthesized and reference images are the material decomposition images of bone(water), and MAE is the difference between the synthesized and reference bone(water) images.