e-ISSN 2083–4640
ISSN 1507–1367

RESEARCH PAPER

Reports of Practical Oncology and Radiotherapy
2022, Volume 27, Number 5, pages: 848–855
DOI: 10.5603/RPOR.a2022.0093
Submitted: 05.07.2022
Accepted: 05.08.2022

Image synthesis of effective atomic number images using a deep convolutional neural network-based generative adversarial network

Daisuke Kawahara[1], Shuichi Ozawa[1, 2], Akito Saito[1], Yasushi Nagata[1, 2]

[1]Department of Radiation Oncology, Institute of Biomedical and Health Sciences, Hiroshima University, Hiroshima, Japan
[2]Hiroshima High-Precision Radiotherapy Cancer Center, Hiroshima, Japan

ABSTRACT

Background: The effective atomic numbers obtained from dual-energy computed tomography (DECT) can aid in characterization of materials. In this study, an effective atomic number image reconstructed from a DECT image was synthesized using an equivalent single-energy CT image with a deep convolutional neural network (CNN)-based generative adversarial network (GAN).

Materials and methods: The image synthesis framework to obtain the effective atomic number images from a single-energy CT image at 120 kVp using a CNN-based GAN was developed. The evaluation metrics were the mean absolute error (MAE), relative root mean square error (RMSE), relative mean square error (MSE), structural similarity index (SSIM), peak signal-to-noise ratio (PSNR), and mutual information (MI).

Results: The difference between the reference and synthetic effective atomic numbers was within 9.7% in all regions of interest. The averages of MAE, RMSE, MSE, SSIM, PSNR, and MI of the reference and synthesized images in the test data were 0.09, 0.045, 0.0, 0.89, 54.97, and 1.03, respectively.

Conclusions: In this study, an image synthesis framework using single-energy CT images was constructed to obtain atomic number images scanned by DECT. This image synthesis framework can aid in material decomposition without extra scans in DECT.

Key words: deep learning; generative adversarial network; effective atomic number

Rep Pract Oncol Radiother 2022;27(5):848–855

Introduction

In a conventional single-energy computed tomography (SECT) image, the pixel value represents the photon attenuation of the tissue. Materials with similar absorbance have the same CT numbers and are difficult to distinguish [1].

Dual-energy CT (DECT) uses two different energy levels, which can determine the ratio of the photoelectric effect components and Compton scattering [2]. It has been used to distinguish between tissues and characterize materials. DECT can obtain a variety of data, including an effective atomic number (Z_{eff}) and iodine- and calcium-enhanced maps [3]. Revolution CT (GE Healthcare, Milwaukee, WI, USA) reconstructs 120 kVp equivalent images and Z_{eff} using the Gemstone Spectral Imaging (GSI) technique [4]. Z_{eff} decomposition

Address for correspondence: Daisuke Kawahara, Ph.D., Department of Radiation Oncology, Institute of Biomedical and Health Sciences, 1-2-3 Kasumi, Minami-ku, Hiroshima-shi, Hiroshima, Japan, tel: (+81) 82-257-1545, fax: (+81) 82-257-1546; e-mail: daika99@hiroshima-u.ac.jp

analysis can aid in the characterization of materials. Mileto et al. used Z_{eff} data to distinguish between non-enhancing renal cysts and enhancing masses [5]. Determining the electron density and effective atomic number is important to better understand the interaction of radiation and to accurately estimate the absorbed dose. For proton and carbon treatment planning, CT values are commonly converted into stopping power ratio (SPR_w) values using a conversion table for dose calculation [6]. However, this approach is restricted to specific human tissue compositions. Z_{eff} is useful for estimating the SPR_w for human tissues in complex anatomy [7]. However, it increases the radiation dose, scan time, and cost.

Convolutional neural networks (CNN) have been successfully applied to image processing and synthesis. Previous studies have developed a deep learning approach using a CNN to perform DECT imaging using standard SECT data. These studies have focused on noise reduction from scanned and synthesized DECT images [8, 9]. Generative adversarial networks (GANs) have two different networks: a generator network that synthesizes images and a discriminator network that distinguishes between the reference and synthesized images [10]. Kida et al. adapted CycleGAN to synthesize PlanCT-like images from CBCT images to improve the quality of CBCT images [11]. Charyyev et al. proposed image synthesis of DECT from SECT and reconstructed the SPR map [12]. The corresponding SPR maps synthesized from DECT reduced the artifacts and noise levels compared with those from the original DECT. In our previous study, we proposed an image synthesis framework that uses single-energy CT images at 120 kVp to obtain fat-water and bone-water images [13]. These studies demonstrated that an image synthesis network with GAN could synthesize DECT images from SECT images.

Herein, we propose an image synthesis approach to obtain effective atomic number images reconstructed from DECT based on GAN architectures.

Materials and methods

Data acquisition

A total of 18,862 images from 29 patients approved by the institutional review board were used for the analysis. The DECT images for each patient were acquired using a Revolution DECT scanner (GE Healthcare, Princeton, NJ, USA). DECT was performed at tube voltages of 80 and 140 kV and exposure of 560 mA. The scanning parameters were rotation time of 1.0 s, slice thickness (ST) of 5 mm, and field of view of 360 mm. The Z_{eff} and equivalent SECT images were reconstructed using the GSI technique.

Deep learning model

An overview of the comparison between the synthesized and reference Z_{eff} images is shown in Figure 1. The Z_{eff} image was synthesized using a GAN. The GAN framework is illustrated in Figure 2. The 16-bit DICOM image was converted into an 8-bit RGB portable network graphics (PNG) image. The output 8-bit RGB PNG images synthesized from the two-dimensional (2D) CNN model were also converted into 16-bit DICOM images [14]. The range of pixel numbers in the effective atomic number images was 0–255. Thus, the unused pixel values (255–65,356) were eliminated in the 16-bit (0–65,536) images and converted into 8-bit images. The SECT and DECT images were rescaled using RescaleIntercept and RescaleSlope from the DICOM header as follows:

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

The proposed 2D CNN model with GAN includes a generator to estimate the Z_{eff} image and a discriminator to distinguish between the reference and synthesized Z_{eff} images. These generator and discriminator networks were trained simultaneously by evaluating . The generator consists of an encoder and a decoder. The encoding maps used eight convolutional layers each, followed by batch normalization and leaky-ReLU activation functions. The number of convolutional and deconvolution filters is shown in Figure 2. The stride was 2 and the kernel size was 4 × 4. The discriminator used seven convolution layers to extract features from the input image and produce the output image. The input images (x) to generator G were SECT images, and the target images (y) were the corresponding Z_{eff} images. Discriminator D was trained to return the loss to which the given image was synthesized. The loss was calculated as follows:

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

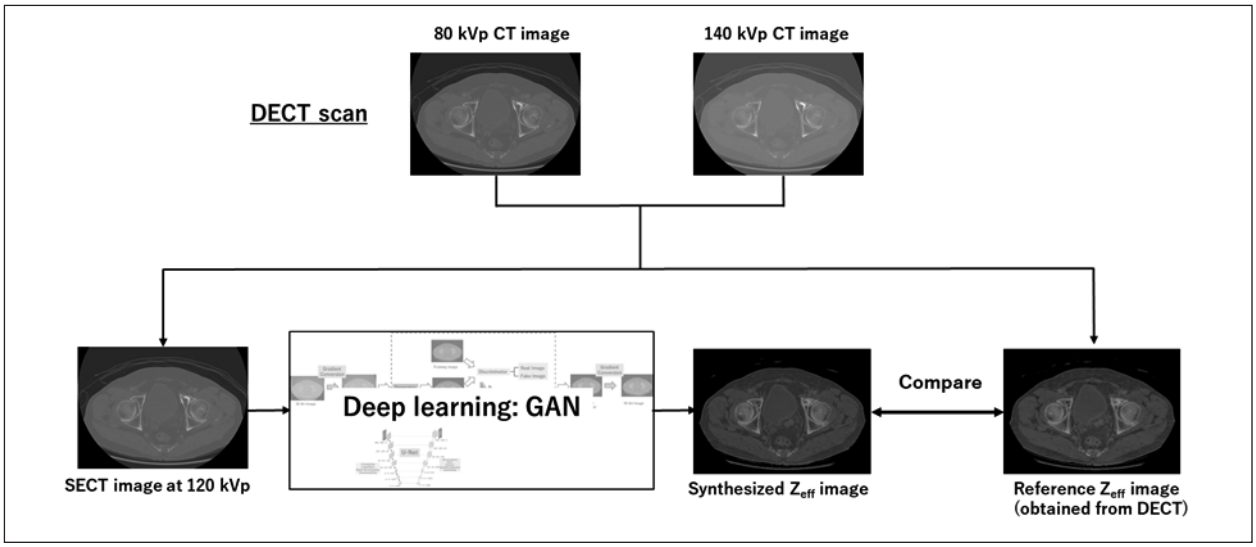

Figure 1. Comparison of synthesized and reference Z_{eff} images; the Z_{eff} image was synthesized from single-energy computed tomography (SECT) obtained from dual-energy computed tomography (DECT) with deep learning. The reference Z_{eff} image was obtained from the DECT; GAN — generative adversarial network

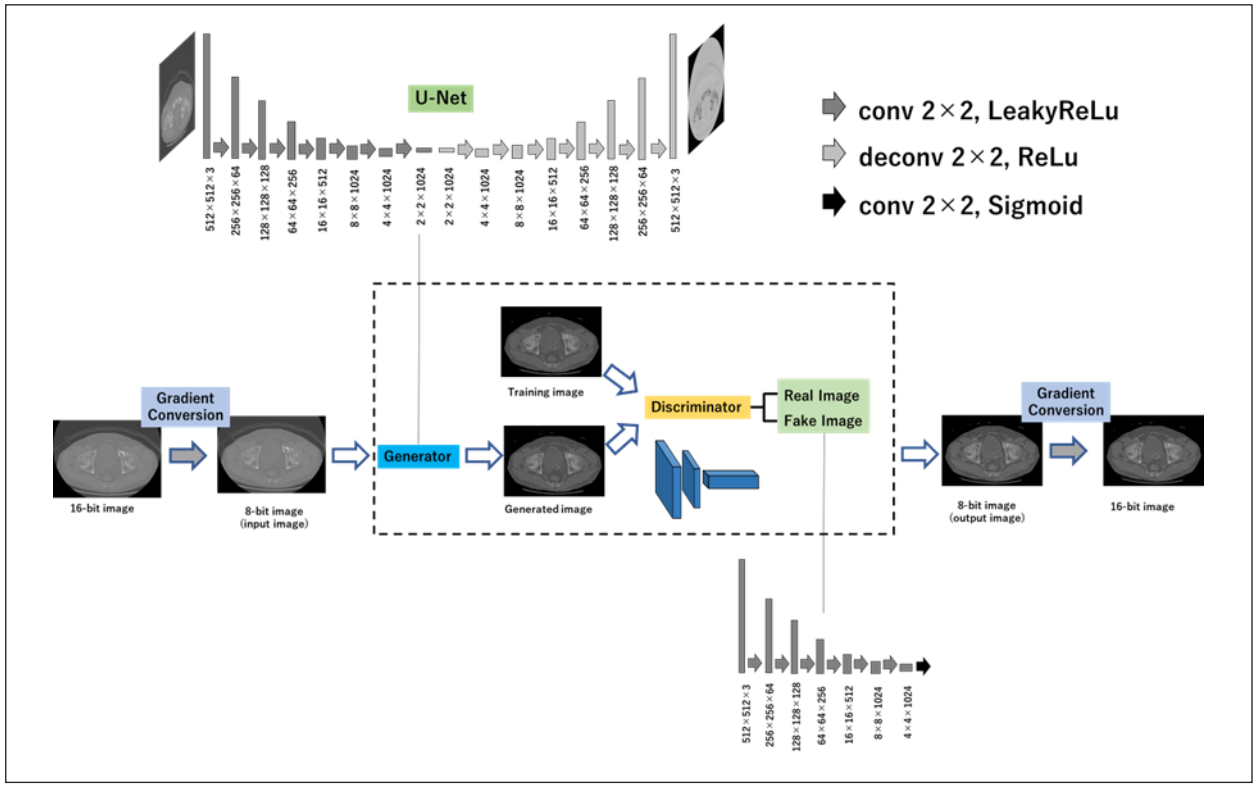

Figure 2. Generative adversarial network (GAN) architecture for the image synthesis of Z_{eff} images from single-energy computed tomography (SECT) images; for gradient conversion, 16-bit DICOM images were converted to 8-bit PNG images

where G is the generator network and is the expected value dependent on both the SECT images (x) and target images (y). Moreover, it includes an additional loss, based on the absolute difference between the input and synthesized images (L1 norm loss). The L1 norm loss is calculated as follows:

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

Adversarial loss is calculated using the binary cross-entropy cost function. The final cost function is calculated as follows:

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

The proposed image synthesis model was implemented using TensorFlow packages (V1.7.0, Python 2.7, CUDA 10.0) on an Ubuntu 16.04 LTS system. The number of epochs was 300. The dataset consisted of 18,826 DECT images scanned from the chest to the pelvis of 29 patients. The data were split into two sets: 16,726 images (21 patients) for training the models and 2100 images (8 patients) for testing.

Evaluation

The prediction accuracy of the model for the reference and synthesized Z_{eff} images was evaluated based on six metrics. The relative mean absolute error (MAE) and mean absolute percentage error (MAPE) were derived as follows:

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

$$MAPE = \frac{1}{n_x n_y} \sum_{i,j}^{n_x n_y} \frac{|r(i,j) - t(i,j)|}{r(i,j)} \quad (6)$$

where and are the values of pixels in the reference Z_{eff} and target Z_{eff} images, respectively, and is the total number of pixels. The relative root mean square error (RMSE) is defined as

$$RMSE = \sqrt{\frac{1}{n_x n_y} \sum_{i,j}^{n_x n_y} (r(i,j) - t(i,j))^2} \quad (7)$$

The structural similarity index (SSIM) considers luminance, structure, and contrast between two images. The SSIM between two images and can be computed as

$$SSIM(\vec{x}, \vec{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_x^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (8)$$

$$C_1 = (k_1 Q)^2, \; k_1 = 0.01 \quad (9)$$

$$C_2 = (k_2 Q)^2, \; k_2 = 0.03 \quad (10)$$

where and are constants used to prevent a zero denominator. is the maximum Z_{eff} value for the reference and synthesized Z_{eff} images. The values of and are typically obtained from [15], and is an estimate in discrete form as follows:

$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_x)^2 \right)^{1/2} \quad (11)$$

The correlation coefficient between and is defined as , and is given by

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_x)(y_i - \mu_y) \quad (12)$$

and is the mean intensity, which is given by

$$\mu_x = \frac{1}{N} \sum_{i=1}^{N} x_i \quad (13)$$

The peak signal-to-noise ratio (PSNR) is calculated as

$$PSNR_{GL} = 10 \times log_{10} \left(\frac{(MAX)^2}{MSE} \right) \quad (14)$$

The mutual information (MI) [16] is calculated as

$$I(r:t) = \sum_{m \in I_r} \sum_{n \in I_t} p(m,n) log \left(\frac{p(m,n)}{p(m)p(n)} \right) \quad (15)$$

where m and n are the intensities in the reference I_r and synthesized Z_{eff} images, and the predicted I_t and Z_{eff} image, respectively. $p(m)$ and $p(n)$ are the marginal densities, and $p(m, n)$ is the joint probability density of I_r and I_t. Moreover, the difference between the reference and synthesized Z_{eff} images in the region of interest (ROI) was evaluated for several slices in the images from the chest to pelvis, as shown in Figure 3.

Results

The losses of the generator, discriminator, and L1 norm are shown in Figure 4. The training time was approximately 154.8 ± 3.2 h. The times to synthesize the Z_{eff} images using the trained models were approximately 7.8–8.2 images/s.

Figures 5 and 6 show samples of the synthetic Z_{eff} images at the pelvic and chest levels. A difference between the reference and synthetic Z_{eff} images was found on the body surface and at the edge of the heart. Table 1 presents the numerical and percentage differences in the Z_{eff} values between the synthetic and reference Z_{eff} images. The numerical and percentage differences of the Z_{eff} value were within 0.86 and 9.5%, respectively, in all ROIs. Table 2 lists the average MAE, MSE, RMSE, PSNR, and MI computed over multiple slices from the pelvis to the chest slices.

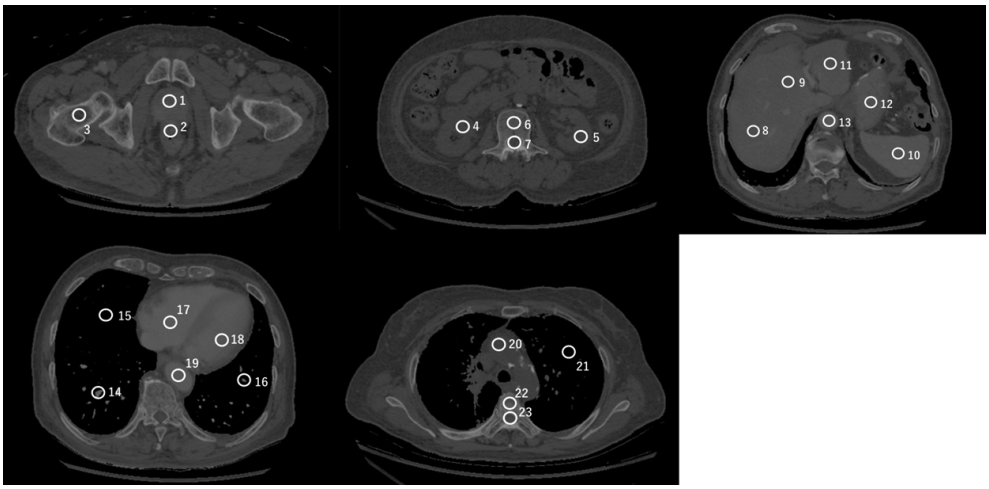

Figure 3. Measurement region in the evaluation of the Z_{eff} value from pelvis to chest slices; the average and standard deviation values of Z_{eff} were measured by creating a circular region of interest (ROI) of 2 cm in diameter

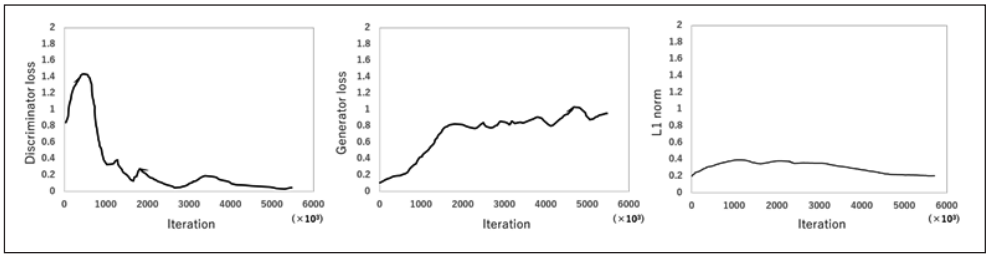

Figure 4. Average training losses in the discriminator, generator, and L1 norm for the training model

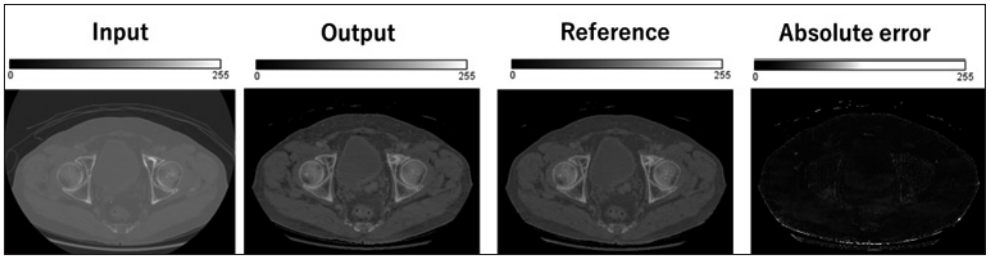

Figure 5. Samples of cross-modality Z_{eff} image generation results at pelvic level: input, output, and reference are the equivalent SECT image at 120 kVp, synthetic Z_{eff} images, and reference Z_{eff} images, respectively. The absolute error was calculated using the synthetic and reference Z_{eff} images

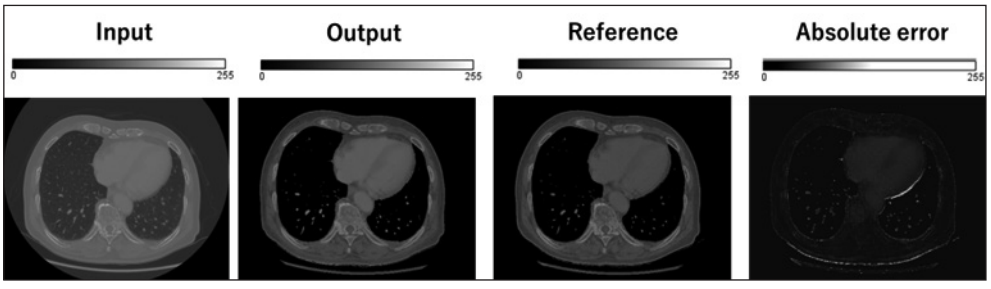

Figure 6. Samples of cross-modality Z_{eff} image generation results at chest level: input, output, and reference are the equivalent single-energy computed tomography (SECT) image at 120 kVp, synthetic Z_{eff} images, and reference Z_{eff} images, respectively. The absolute error was the calculated using the synthetic and reference Z_{eff} images

Table 1. Numerical (Δ) and percentage differences of the Z_{eff} value between synthetic and reference Z_{eff} images. The numerical and percentage differences of the Z_{eff} value were within 0.86 and 9.5% in all ROIs from the chest to pelvis

Measurement region	Δ	%
1	0.69	8.77
2	0.72	9.19
3	0.41	5.98
4	0.21	2.74
5	0.19	2.47
6	0.53	6.41
7	-0.11	-1.47
8	0.73	8.89
9	0.77	9.58
10	0.79	9.42
11	0.79	9.75
12	0.76	9.46
13	-0.06	–0.73
14	0.29	8.04
15	0.12	7.90
16	0.76	8.13
17	0.82	9.54
18	0.84	9.95
19	0.86	9.84
20	0.02	–0.26
21	–0.01	8.12
22	0.06	0.70
23	–0.08	–0.80

The standard deviation (SD) from the pelvis to the chest slices was significantly smaller for all evaluation items.

Discussion

In this study, an image synthesis model for Z_{eff} images from SECT images using a deep learning approach was proposed. The numerical difference between the Z_{eff} values of synthesized and reference Z_{eff} images was within 9.95% in some regions, from the pelvis to chest slice. Mitchell et al. evaluated the Z_{eff} values obtained from DECT by comparing them with theoretical Z_{eff} values. The Catphan phantom (The Phantom Laboratory, Salem, NY, USA) had a Z_{eff} value accuracy of 15% when no lung inserts were used [17]. This suggests that the synthesized Z_{eff} image was in good agreement with the reference image within the uncertainty of the Z_{eff} image obtained from DECT.

The SD of the Z_{eff} values in the lung region of the Z_{eff} images was larger than that of other regions. This is because the lungs have a non-uniform structure. A previous study also showed that the measured Z_{eff} values of the inhaled lung insert in the CIRS 062M phantom were significantly different from the theoretical Z_{eff} values [18]. Thus, an accurate Z_{eff} image reconstructed from DECT is an essential input for deep learning. Further studies are needed to synthesize Z_{eff} values in the lung region using high-quality DECT images.

Schaeffer evaluated the accuracy of the Z_{eff} between the theoretical and measurement Z_{eff} from DECT. The MAPE was 6.3% for the body phantom and 3.2% for the head phantom [19]. The current study showed that the MAPE of the Z_{eff} was 1.16% ± 0.14 % with the GAN method. Moreover, Garcia et al. proposed a method of the extraction the Z_{eff} for the DECT image based on an Karhunen-Loeve expansion of the atomic cross section per electron [20]. The MAPE between the theoretical and calculated value was 4.1% ± 0.3%. Schaeffer et al. evaluated the accuracy of Z_{eff} from DECT. It suggests that the synthesized Zeff image showed a good agreement within the uncertainty of the Z_{eff} image obtained from DECT and the accuracy of the estimation for the Zeff was superior to the conventional method. Although the other evaluation metrics were used in the image synthesis study, it has never been used for the Z_{eff} image synthesis. These results of the evaluation metrics

Table 2. Evaluation metrics of Z_{eff} image synthesis from pelvis to chest slice

MAE		MAPE		MSE		RMSE		PSNR		SSIM		MI	
Avg	SD	Avg	SD	Avg	SD	Avg	SD	Avg	SD	Avg	SD	Avg	SD
0.09	0.01	1.16	0.14	0.21	4.2E–03	0.45	4.6E–03	54.97	0.09	0.89	0.01	1.03	0.12

MAE — mean absolute error; MAPE — mean absolute percentage error; MSE — relative mean square error; RMSE — relative root mean square error; PSNR — peak signal-to-noise ratio; SSIM — structural similarity index; MI — mutual information; Avg — average; SD — standard deviation

would be important in the image synthesis or conversion to the Z_{eff} from the DECT image in further studies.

An equivalent SECT image was used in this study. Kamiya et al. compared equivalent and conventional SECT images [21]. Although the radiation dose was reduced for the equivalent SECT image, the image quality was equivalent in both the quantitative and qualitative evaluations. Thus, the proposed model can be applied to conventional SECT images.

Zhao et al. proposed an image synthesis method to map low-energy to high-energy images using a two-stage CNN [16]. Zhao et al. evaluated virtual non-contrast imaging using DECT from SECT [16]. This might contribute to the prediction of perfusion imaging, urinary stone characterization, cardiac imaging, and angiography from SECT images. Our model extends the possibility of predicting the DECT image from the SECT image and contributes to the material decomposition with the predicted DECT image. Thus, the proposed image synthesis model can significantly simplify the DECT system design and reduce scanning and imaging costs. For radiation diagnosis, the Z_{eff} image should assist in lesion detection. The current study showed the possibility of efficient image synthesis of Z_{eff} images for material decomposition from a simple analysis. Further studies will be performed to evaluate the detectability of the lesions.

Conclusion

In this study, an image synthesis framework using single-energy CT images to generate atomic number images scanned by DECT was proposed. This image synthesis framework can aid in determining material decomposition without extra scans in DECT.

Conflict of interest

None declared.

Funding

None declared.

Ethnical approval

The current study does not involve any experimentation on human participants or animals.

Informed consent

The current study does not involve any experimentation on human participants or animals.

Acknowledgements

None declared.

References

1. Yoon W, Seo JJ, Kim JK, et al. Contrast enhancement and contrast extravasation on computed tomography after intra-arterial thrombolysis in patients with acute ischemic stroke. Stroke. 2004; 35: 876–881, doi: 10.1161/01.STR.0000120726.69501.74, indexed in Pubmed: 14988575.
2. McCollough CH, Leng S, Yu L, et al. Dual- and Multi-Energy CT: Principles, Technical Approaches, and Clinical Applications. Radiology. 2015; 276: 637–653, doi: 10.1148/radiol.2015142631, indexed in Pubmed: 26302388.
3. Johnson TRC, Krauss B, Sedlmair M, et al. Material differentiation by dual energy CT: initial experience. Eur Radiol. 2007; 17(6): 1510–1517, doi: 10.1007/s00330-006-0517-6, indexed in Pubmed: 17151859.
4. Slavic S. Madhav P., Profio M. et al. Technology White Paper, GSI Xtream on RevolutionTM CT. https://www.gehealthcare.com/-/media/069734962cbf45c1a5a01d-1cdde9a4cd.pdf.
5. Mileto A, Allen BC, Pietryga JA, et al. Characterization of Incidental Renal Mass With Dual-Energy CT: Diagnostic Accuracy of Effective Atomic Number Maps for Discriminating Nonenhancing Cysts From Enhancing Masses. AJR Am J Roentgenol. 2017; 209(4): W221–W230, doi: 10.2214/AJR.16.17325, indexed in Pubmed: 28705069.
6. Wohlfahrt P, Möhler C, Hietschold V, et al. Clinical Implementation of Dual-energy CT for Proton Treatment Planning on Pseudo-monoenergetic CT scans. Int J Radiat Oncol Biol Phys. 2017; 97(2): 427–434, doi: 10.1016/j.ijrobp.2016.10.022, indexed in Pubmed: 28068248.
7. Wohlfahrt P, Möhler C, Richter C, et al. Evaluation of Stopping-Power Prediction by Dual- and Single-Energy Computed Tomography in an Anthropomorphic Ground-Truth Phantom. Int J Radiat Oncol Biol Phys. 2018; 100(1): 244–253, doi: 10.1016/j.ijrobp.2017.09.025, indexed in Pubmed: 29079119.
8. 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, indexed in Pubmed: 31797593.
9. Lyu T, Zhao W, Zhu Y, et al. Estimating dual-energy CT imaging from single-energy CT data with material decomposition convolutional neural network. Med Image Anal. 2021; 70: 102001, doi: 10.1016/j.media.2021.102001, indexed in Pubmed: 33640721.
10. Goodfellow I, Pouget-Abadie J, Mirza M. Generative adversarial nets. In: Welling M. ed. Advances in neural information processing systems. Neural Information Processing Systems Foundation, Inc., Montreal 2015: 2672–2680.
11. Kida S, Kaji S, Nawa K. Cone-beam CT to Planning CT synthesis using generative adversarial networks. arXiv: 1901.05773v1.

12. Charyyev S, Wang T, Lei Y, et al. Learning-based synthetic dual energy CT imaging from single energy CT for stopping power ratio calculation in proton radiation therapy. Br J Radiol. 2022; 95(1129): 20210644, doi: 10.1259/bjr.20210644, indexed in Pubmed: 34709948.
13. Kawahara D, Saito A, Ozawa S, et al. Image synthesis with deep convolutional generative adversarial networks for material decomposition in dual-energy CT from a kilovoltage CT. Comput Biol Med. 2021; 128: 104111, doi: 10.1016/j.compbiomed.2020.104111, indexed in Pubmed: 33279790.
14. Zhou X, Takayama R, Wang S, et al. Deep learning of the sectional appearances of 3D CT images for anatomical structure segmentation based on an FCN voting method. Med Phys. 2017; 44: 5221–5233, doi: 10.1002/mp.12480, indexed in Pubmed: 28730602.
15. Wang Z, Bovik AC, Sheikh HR. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process. 2004; 13(4): 600–612.
16. 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.
17. Mitchell MM, Christodoulou EG, Larson SL. Accuracies of the synthesized monochromatic CT numbers and effective atomic numbers obtained with a rapid kVp switching dual energy CT scanner. Med Phys. 2011; 38(4): 2222–32, doi: 10.1118/1.3567509, indexed in Pubmed: 21626956.
18. Kawahara D, Ozawa S, Yokomachi K, et al. Synthesized effective atomic numbers for commercially available dual-energy CT. Rep Pract Oncol Radiother. 2020; 25(4): 692–697, doi: 10.1016/j.rpor.2020.02.007, indexed in Pubmed: 32684854.
19. Schaeffer CJ, Leon SM, Olguin CA, et al. Accuracy and reproducibility of effective atomic number and electron density measurements from sequential dual energy CT. Med Phys. 2021; 48(7): 3525–3539, doi: 10.1002/mp.14916, indexed in Pubmed: 33932301.
20. Garcia LI, Azorin JF, Almansa JF. A new method to measure electron density and effective atomic number using dual-energy CT images. Phys Med Biol. 2016; 61(1): 265–279, doi: 10.1088/0031-9155/61/1/265, indexed in Pubmed: 26649484.
21. Kamiya K, Kunimatsu A, Mori H, et al. Preliminary report on virtual monochromatic spectral imaging with fast kVp switching dual energy head CT: comparable image quality to that of 120-kVp CT without increasing the radiation dose. Jpn J Radiol. 2013; 31(4): 293–298, doi: 10.1007/s11604-013-0185-9, indexed in Pubmed: 23408047.