EasyChair Preprint

№ 11596

Asymmetrical Dual-Cycle Adversarial Network for Material Decomposition and Synthesis of Dual-Energy CT Images

Xinrui Zhang, Ailong Cai, Shaoyu Wang, Ningning Liang, Yizhong Wang, Junru Ren, Lei Li and Bin Yan

December 19, 2023

Asymmetrical Dual-Cycle Adversarial Network for Material Decomposition and Synthesis of Dual-energy CT Images

Xinrui Zhang[1], AiLong Cai[1], Shaoyu Wang[1], Ningning Liang[1], Yizhong Wang[1], Junru Ren[1], Lei Li[1] and Bin Yan[1]

[1] Department of Henan Key Laboratory of Imaging and Intelligent Processing, PLA Strategic Support Force Information Engineering University, Zhengzhou, China

Abstract Dual-energy computed tomography (DECT) can identify the material properties with its excellent material quantitative analysis ability. However, the application of DECT is restricted by the problems of inaccuracy of energy spectrum estimation, non-linearity and inconsistency of imaging geometry, which will lead to the degradation of material distribution images. Hence, deep learning (DL)-based methods have become the state-of-the-art technique in DECT rely on its excellent feature recognition performance in the case of few spectrum prior. In this work, we propose an asymmetrical Dual-Cycle adversarial network (ADCNet) for both material decomposition and synthesis of dual-energy CT images, which has certain advantages in spectral CT multi-task parallel, improvement of image quality and radiation dose reduction. The experimental results show that the cycle framework achieves the adversarial learning of dual networks, and promotes the quality of generated images by introducing multiple mechanisms. Compared with the traditional DL-based methods, the proposed method has outstanding qualitative and quantitative indicators.

1 Introduction

Dual-energy computed tomography (DECT) utilizes the potential information in energy spectrum to achieve quantitative analysis of substances, which is highly promising for clinical applications. Although DECT has certain preponderance over conventional CT, a tiny disturbance in spectrum-imaging would bring out an inestimable impact on the material decomposition. Meanwhile, the radiation accumulation of DECT scanning is another question worthy of attention in clinical application. Exploring methods of reducing the radiation dose of DECT is also a key issue in the research field.

In order to effectively extract the intrinsic feature of the spectral CT images and improve the quality of decomposed images, deep learning (DL)-based methods have become the state-of-the-art technique in DECT. In 2019, Zhang *et al* [1] exploited the characteristics of DECT to optimize the traditional U-Net architecture and build a dual U-Net with butterfly structure. It shows that the dual U-Net architecture with information interaction has presented great potential in DECT material decomposition. To further improve the network performance, Shi *et al* [2] adopted the General Adversary Network (GAN) [3] to the dual U-Net structure and compared different GAN variants，creating a network called interactive Wasserstein GAN (DIWGAN). Based on this method, the effect of material decomposition has been further promoted. In 2022, Zhou *et al* [4] analysed the requirements of tradeoffs between the level of radiation and the quality of spectral CT images, and proposed a cycle adversarial network with multi-strategy to synthesize high-energy images from low- energy images. The bidirectional loop structure based on CycleGAN [5] have achieved promising results in the synthesis task of spectral CT images.

In this paper, we combine the two tasks: base material decomposition and synthesis of dual-energy CT images with an asymmetrical dual-cycle adversarial network (denoted ADCNet). In practical application, our method can use conventional CT image to synthetize dual-energy images to further reduce the radiation dose of CT scanning, and achieve accurate material decomposition at the same time, which is conducive to shorten the time of clinical diagnosis and promotes the practical application of DECT.

2 Materials and Methods

2.1 Dual-Cycle Adversarial Framework

Here, we first describe the composition of the dual-cycle generation adversarial network framework. To realize one-time conversion of multi-task in a integral framework, we design a double-entry and double-out architecture based on CycleGAN. The material decomposition module (Module I) and the image synthesis module (Module II) have been contained in the circle framework, as illustrated in Fig. 1.

Different from the traditional CycleGAN, the proposed ADCNet has been improved and innovated in network architecture, loss function and training methods.

2.2 Loss Function

The input high- and low-energy images X_H^r , X_L^r are generated as truth data to provide dual-energy spectrum information for the material decomposition. On the other side, we have prepared bone and tissue images separated from the head data of patients, which are presented as X_B^r , X_S^r . To distinguish the real material images from the fake style images converted by generator G , and the real high- and low-energy images from the fake style images converted by the other generator F , we define the follwing loss functions:

$$\ell_{GAN}^1\left(G, D_{B,S}, X_{H,L}^r, X_{B,S}^r\right) = E_{x_{B,S}}\left[\log D_{B,S}(x_{B,S})\right] + E_{x_{H,L}}\left[\log\left(1 - D_{B,S}\left(G(x_{H,L})\right)\right)\right] \quad (1)$$

$$\ell_{GAN}^2\left(F, D_{H,L}, X_{B,S}^r, X_{H,L}^r\right) = E_{x_{H,L}}\left[\log D_{H,L}(x_{H,L})\right] + E_{x_{B,S}}\left[\log\left(1 - D_{H,L}\left(F(x_{B,S})\right)\right)\right] \quad (2)$$

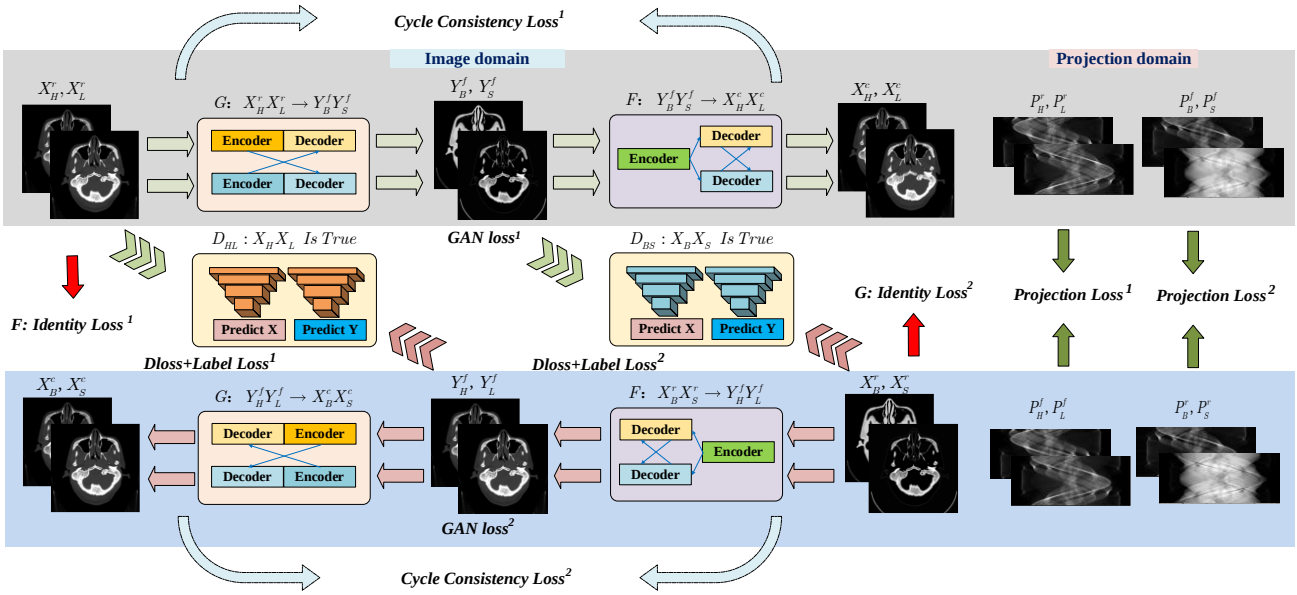

Figure 1: Dual-Cycle Adversarial Framework architecture. *1st line*: Module I of material decomposition. *2nd line*: Module II of synthesis of dual-energy CT images. From left to right are the loss functions in the image domain and projection domain.

$$\ell_{Cycle}(G,F) = E_{x_{H,L}}\left(\left\|F\left(G(x_{H,L})\right) - x_{H,L}\right\|_1\right) + E_{x_{B,S}}\left(\left\|G\left(F(x_{B,S})\right) - x_{B,S}\right\|_1\right) \quad (3)$$

$$\ell_{Idt}(G,F) = E_{x_{B,S}}\left[\left\|G(x_{B,S}) - x_{B,S}\right\|_1\right] + E_{x_{H,L}}\left[\left\|F(x_{H,L}) - x_{H,L}\right\|_1\right] \quad (4)$$

$$\ell_{edge}(G,F) = E_{x_{B,S}}\left[\left\|\nabla F(x_{B,S}) - \nabla x_{B,S}\right\|_2^2\right] + E_{x_{H,L}}\left[\left\|\nabla G(x_{H,L}) - \nabla x_{H,L}\right\|_2^2\right] \quad (5)$$

where ℓ^1_{GAN} and ℓ^2_{GAN} represent the loss of GAN at D_{BS} and D_{HL} respectively; $x_{H,L}$ and $x_{B,S}$ represent 2 spectral CT images and base materials. measuring the distance between the generated fake images and the real images and giving judgement. ℓ_{Cycle} aims to make the output images X^c_H and X^c_L (or X^c_B, X^c_S) consistent with X^r_H and X^r_L (or X^c_H, X^c_L) in style. The identity loss ℓ_{Idt} is also an indispensible link to maintain grey scale information of the materials. In addition, we introduce edge loss ℓ_{edge} to restore the texture and edge features of images. Y^f_H, Y^f_L, Y^f_B and Y^f_S represent the output of generators. Meanwhile, to further improve the quality of images, we try to excavate the projection information of the image from the projection domain as a loss function to constrain. We assume the projection of label images are P^r_H, P^r_L, P^r_B and P^r_S respectively. The generated images should be as close as possible to the label images. Therefore, the projection loss can be described by

$$\ell_{proj}(G,F) = E_{x_{B,S}}\left[\left\|P\left(F(x_{B,S})\right) - P(x_{B,S})\right\|_1\right] + E_{x_{H,L}}\left[\left\|P\left(G(x_{H,L})\right) - P(x_{H,L})\right\|_1\right] \quad (6)$$

The whole loss function is given by

$$\ell_{SUM} = \lambda_1\ell^1_{GAN} + \lambda_2\ell^2_{GAN} + \lambda_3\ell_{Cycle} + \lambda_4\ell_{Idt} + \lambda_5\ell_{edge} + \lambda_6\ell_{proj} \quad (7)$$

where $\{\lambda_i,\ i = 1,2,...,6\}$ represents the balance parameters of different loss functions.

2.3 ADCNet Architecture

In this paper, we proposed ADCNet which integrates multiple mechanisms to achieve the decomposition of bone and tissue from high- and low-energy CT images. The double-entry and double-out network with information interaction between the two paths can effectively acquire the internal characteristics of base materials from spectral CT images. On the contrast, to achieve the task of synthesis of DECT images, we design a single-entry and double-out network architecture to generate high- and low- energy images from fused images. Fig. 1 shows the structure of generator G, generator F and discriminator D. Moreover, we also introduce DANet module [6] which combines spatial attention and channel attention to enlarge the receptive field and restore the texture details of bones and tissues. In particular, we add butterfly architecture [1] to the deepest downsampling of ADCNet in generator G, in order to collect high-level abstract semantic information of high- and low-energy spectral images. In generator F, we also introduce multi-information interaction mechanism, and merge the two kinds of materials before entering the network to assist effective feature extraction.

2.4 Data Preparation & Parameter Setting

The data set comes from cranial cavity slice images of 7 patients with size 512 by 512. In experiments, we prepared 1505 actual data of bone and tissue and added 120 kVp and 80 kVp spectrum to the original slice images as our simulation dual-energy data. In the training, we selected 1400 pairs of high- and low-energy images as the training dataset to train the model, and 105 pairs of images as the test dataset to validate the network performance. As for parameters of ADCNet, each downsampling path of G and F includes 7 convolutional layers. the number of filters is

128, 256, 512, 512, 512 and 512, respectively. Two paths are connected by 2 cascaded residual blocks. In the process of training, the initial learning rate for adam was set to 0.0002 (momentum term: 0.5, β_1 : 0.5 , β_2 : 0.999). To ensure that the training process did not produce over fitting, we set the maximum epoch to 15. The training duration of the proposed model was about 20 hours.

3 Results

To evaluate the performance of different modules in ADCNet, we design ablation experiments on 20 typical test data. The experiment is divided into 2 independent parts. Part I verifies the performance of information interaction, residual block, butterfly structure and DANet attention module on the original Dual-CycleGAN. As is shown in Table 1, the PSNR of decomposed Bone and Tissue increased by 5.06dB and 6.67dB respectively after adding all of the modules. We can see that ADCNet with four hybrid modules has better quantitative indicators inTo evaluate the performance of different modules in the feature extraction of bone and tissue texture. Part II verifies the improvement effect of generator *F* using fused inputs. As is shown in Fig. 2, the symmetrical dual-cycle network with 2 same generators (SDCNet) is compared with the ADCNet of 2 fused inputs in spatial dimension (ADCNet-SF) and in channel dimension (ADCNet-CF). Obviously, these two forms of ADCNet have shorter training time than SDCNet, due to the fused architecture sharing the parameters in downsampling path (the number of parameters of SDCNet and ADCNet is 387M and 329M, respectively). Especially, the quantitative indicators of PSNR and RMSE also illustrate that the ADCNet-CF has better performance far beyond than the other two networks in the synthesis of dual-energy CT images, due to the high similarity of fused material images with spectral CT images.

In addition, we select a group of head data to evaluate the qualitative performance of different networks. Fig. 3 demonstrates the comparison of four state-of-the-art networks on material decomposition including FCN, Pix2Pix [7], DIWGAN and the proposed ADCNet. The ROI (Region of Interest) of bone shows the structure of the cochlea and frontal lobe, and the ROI of tissue shows the structure of the lateral ventricle and cerebellar. Fig.4 shows the residual images of the results with the label images. Compared with the conventional FCN, the SSIM and RMSE of bone images are greatly improved by 0.3246 and 0.0188, and tissue images by 0.072 and 0.0508, which means the proposed ADCNet has marvel performance in material decomposition. Although we find that there still exists salt and pepper noise in the material images, the later operation of median filtering can also help us to further improve the quality of images. At the same time, ADCNet also achieves the synthesis of dual-energy CT images owing to its special dual-cycle architecture, even better beyond the SDCNet and ADCNet-SF. Fig.5 shows the pixel values of a certain section of 120kVp and 80kVp synthetic spectral CT images. The expected result is closer to the ground truth from the structure of the dual-energy images. In Fig.3, we also discussed the accuracy of synthetic dual-energy images from a quantitative perspective, which reached 45.85dB and 45.93dB in PSNR, and over 0.99 in SSIM. To sum up, ADCNet has achieved good results in both tasks of base material decomposition and synthesis of dual-energy CT images.

Tabel 1: Ablation experiments of Part I. From left to right are information interaction, residual block, butterfly structure and DANet attention module.

ID	Interac	Res	Butterfly	DANet	PSNR-Bone	PSNR-Tissue
1	√				37.36	29.17
2	√	√			39.58	31.58
3	√	√	√		41.10	32.92
4	√	√		√	41.43	34.30
5	√	√	√	√	42.42	35.84

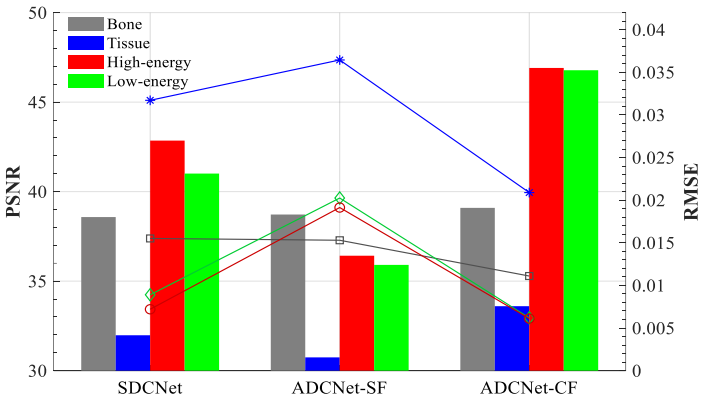

Figure 2: Ablation experiments of Part II. (1)SDCNet: the symmetrical dual-cycle network with 2 same generators G (2)ADCNet-SF: the asymmetrical dual-cycle network of 2 fused inputs in spatial dimension. (3)ADCNet-CF: the asymmetrical dual-cycle network of 2 fused inputs in channel dimension. *Note:* Bar chart represents PSNR, and line chart represents RMSE.

4 Conclusion

In this work, we proposed a new dual-cycle network architecture (ADCNet), which can achieve dual tasks in material decomposition and synthesis of dual-energy CT images. Compared with the other four typicall networks in terms of quantitative and qualitative indicators, our method demonstrated outstanding ability in material identification and accurate generation of energy spectrum images under the none-spectral prior condition, which can assist multi-functional integrated spectral CT imaging more efficiently and accurately. Additionally, Our method can achieve lower dose as the DECT images can be totally synthetic. In the future, we will further carry out relevant experiments based on the actual problems in medical diagnosis.

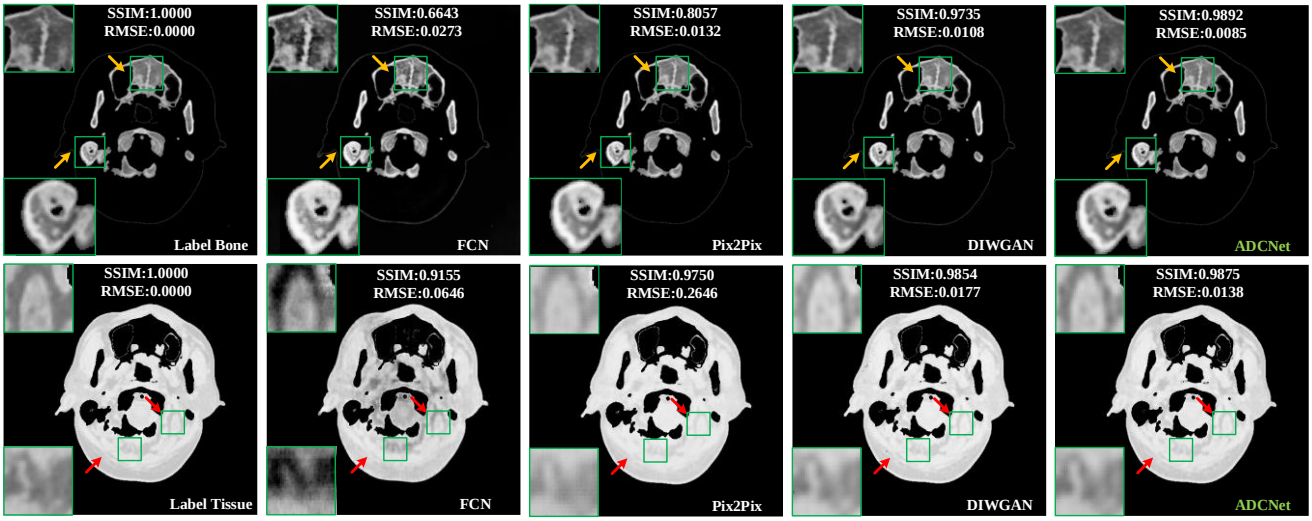

Figure 3: Networks on material decomposition. From left to right are FCN, Pix2Pix, DIWGAN and the proposed ADCNet. From top to bottom are bone and tissue. The 1st line represents the decomposed bone images, and the 2nd line represents the decomposed tissue images.The SSIM and RMSE values are written at the top.

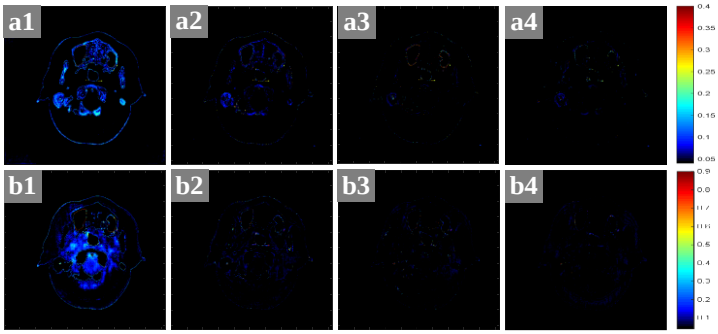

Figure 4: The residual images of the results of FCN, Pix2Pix, DIWGAN and ADCNet with label images. We use absolute error $\Re = \|y - y^*\|$ to express the residual images. (a1)-(a4): The residual images of bone. (b1)-(b4): The residual images of Tissue. *Note*: The range of colorbox is [0.05, 0.4] g/mm^3 and [0.1, 0.9] g/mm^3, respectively.

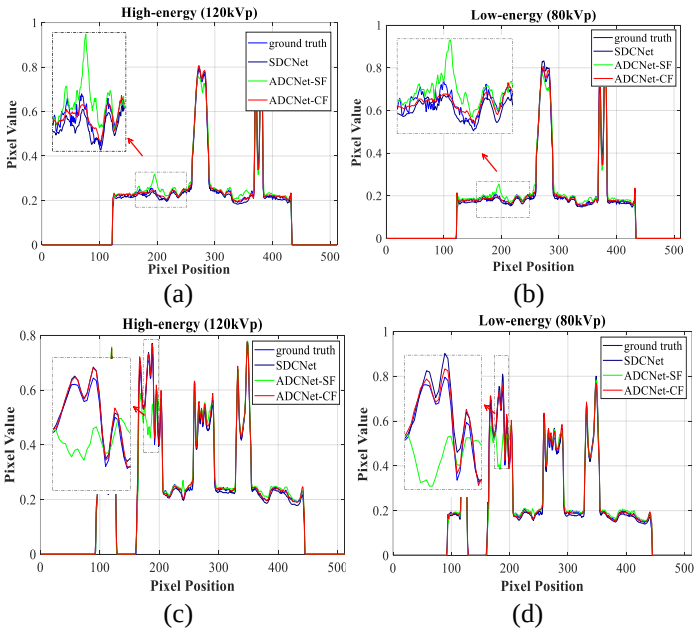

Figure 5: Profile plots of synthetic dual-energy CT images with different networks including SDCNet, ADCNet-SF and ADCNet-CF. (a), (b): Horizontal profile plots of 120kVp and 80kVp synthetic images. (c), (d): Vertical profile plots of 120kVp and 80kVp synthetic images.

5 Acknowledgments

This work was supported by the National Key Research and Development Project of China (Grant No. 2020YFC1522002), the National Natural Science Foundation of China (Grant No. 62101596), the National Natural Science Foundation of China (Grant No. 62201616) and the China Postdoctoral Science Foundation (Grant No. 2019M663996).

References

[1] Zhang W, Zhang H, Cai A, Wang L, Li L and Yan B, et al. "Image domain dual material decomposition for dual-energy CT using butterfly network". *Medical Physics*. 46.5 (2019), pp. 2037-2051. DOI: 10.1002/mp.13489.

[2] Shi Z, Li H, Cao Q, Wang Z, Cheng M. "A material decomposition method for dual-energy CT via dual interactive Wasserstein generative adversarial networks". *Medical Physics*. 48.6 (2021), pp. 2891-2905. DOI: 10.1002/mp.14828.

[3] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, et al. "Generative Adversarial Networks" 2014. arXiv:1406.2661.

[4] Zhou H, Liu X, Wang H, Chen Q, et al. "The synthesis of high-energy CT images from low-energy CT images using an improved cycle generative adversarial network". *Quantitative Imaging in Medicine and Surgery* 12.1 (2022), pp. 28-42. DOI: 10.21037/qims-21-182.

[5] Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks". 2020. arXiv:1703.10593.

[6] Li Y, Chen X, Zhu Z, Xie L, Huang G et al. "Attention-guided Unified Network for Panoptic Segmentation" 2019. arXiv:1812.03904.

[7] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros. "Image-to-Image Translation with Conditional Adversarial Networks". 2018. arXiv:1611.07004.