2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

A Cascaded Deep Learning Framework for Detecting Aortic Dissection Using Non-contrast Enhanced Computed Tomography

Xiangyu Xiong[1], Xiuhong Guan[1], Chuanqi Sun[1], Tianjing Zhang[2], Hao Chen[1], Yan Ding[3], Zhangbo Cheng[4], Lei Zhao[3], Xiaohai Ma[3] and Guoxi Xie[1]

***Abstract*— Aortic dissection (AD) is a rare but potentially fatal disease with high mortality. The aim of this study is to synthesize contrast enhanced computed tomography (CE-CT) images from non-contrast CT (NCE-CT) images for detecting aortic dissection. In this paper, a cascaded deep learning framework containing a 3D segmentation network and a synthetic network was proposed and evaluated. A 3D segmentation network was firstly used to segment aorta from NCE-CT images and CE-CT images. A conditional generative adversarial network (CGAN) was subsequently employed to map the NCE-CT images to the CE-CT images non-linearly for the region of aorta. The results of the experiment suggest that the cascaded deep learning framework can be used for detecting the AD and outperforms CGAN alone.**

***Clinical relevance*— This work provides a novel framework to synthesize CE-CT images from NCE-CT images and concludes a criterion to detect aortic dissection using the synthesized images.**

I. INTRODUCTION

AD is a rare but life-threatening disease that has many complications [1]. It is associated with a high mortality of 1% per hour in untreated patients. Misdiagnosis of AD is devastating to the patients [2].

Accurate diagnosis of AD is an essential step for making correct treatment decisions to increase the patient's chance of survival. Contrast enhanced computed tomography (CE-CT) is the first line technique widely used in detecting AD. As the X-ray decay coefficient of the contrast medium is much higher than that of the vessel wall, the aorta after injecting contrast medium becomes much brighter than vessel wall on CE-CT image. This makes the dissection symptom that the dark intima within the bright aorta was easily identified on CE-CT images. Thus, CE-CT has high sensitivity (>95%) and specificity (>95%) for the diagnosis of AD. However, CE-CT needs to inject contrast medium, which cannot be used in the patients with acute renal failure or allergic reactions [1], [3]. In addition, some AD patients have very similar symptoms to other cardiovascular diseases, these patients may receive non-contrast enhanced CT (NCE-CT) but CE-CT for disease examinations initially, leading to the misdiagnosis and missing the treatment time window for AD patients. If NCE-CT could provide hints for the diagnosis of AD patients, the misdiagnosis would be reduced. However, the dissection symptom on NCE-CT images is not obvious because the aorta without contrast medium has similar contrast with the intima.

In recent years, several models have been developed to synthesize a CE-CT image from a NCE-CT or low dose CE-CT image. These models are almost based on convolutional neutral network including generative adversarial network (GAN) [4]. In 2018, an encoder-decoder deep convolutional network was proposed to generate cardiac contrast enhanced CT images from contrast-free CT thoracic scans, with the purpose of volumetric assessment of left heart chambers [5]. Similarly, an encoder-decoder convolutional neural network was proposed to synthesize the full dose brain contrast enhanced MR images from the zero dose and low dose MR images to reduce gadolinium dose [6]. A steerable GAN method was proposed to generates absent MRA from existing MR multi-contrast images [7]. It is noteworthy that several studies have applied GANs in image synthesis of cross-modality for radiotherapy planning [8], [9], [10], [11], [12]. Inspired by the previous work, we proposed a novel framework to synthesize the CE-CT images from NCE-CT images to detect AD.

The key contributions of this work can be summarized as: 1. A novel deep learning framework combines a segmentation network with a conditional generative adversarial network (CGAN) to synthesize CE-CT images of high quality to detect AD. 2. Experiments are conducted to show that our proposed framework can synthesize better results compared with CGAN alone 3. A reasonable criterion is concluded to demonstrate the improvements of detecting AD using the synthetic CE-CT images.

*This research was supported by National Science Foundation of China (81971607)

[1]Xiangyu Xiong, Xiuhong Guan, Chuanqi Sun, Hao Chen, and Guoxi Xie are with the Department of Biomedical Engineering, Guangzhou Medical University, Guangzhou 511436, China (e-mail: xiangyu.xiong.adc@gmail.com; 2018217304@stu.gzhmu.edu.cn; chuanqisun077@gmail.com; hao.chen.gd@gmail.com; guoxixie@163.com).

[2]Tianjing Zhang is with the Philips Healthcare, Guangzhou 511436, China (e-mail: steve.zhang_1@philips.com).

[3]Yan Ding, Lei Zhao and Xiaohai Ma are with the Department of Interventional Diagnosis and Treatment, Beijing Anzhen Hospital, Capital Medical University, Beijing 100029, China (email: dingyan423761@163.com; s2006430@126.com; maxi8238@gmail.com).

[4]Zhangbo Cheng is with the Fujian Provincial Clinical College, Fujian Medical University, Fuzhou 350001, Fujian, China and the Department of Cardiovascular surgery, Fujian Provincial Hospital, 134 Dongjie Street, Fuzhou 350001, Fujian, China (email: cheng_eason@hotmail.com).

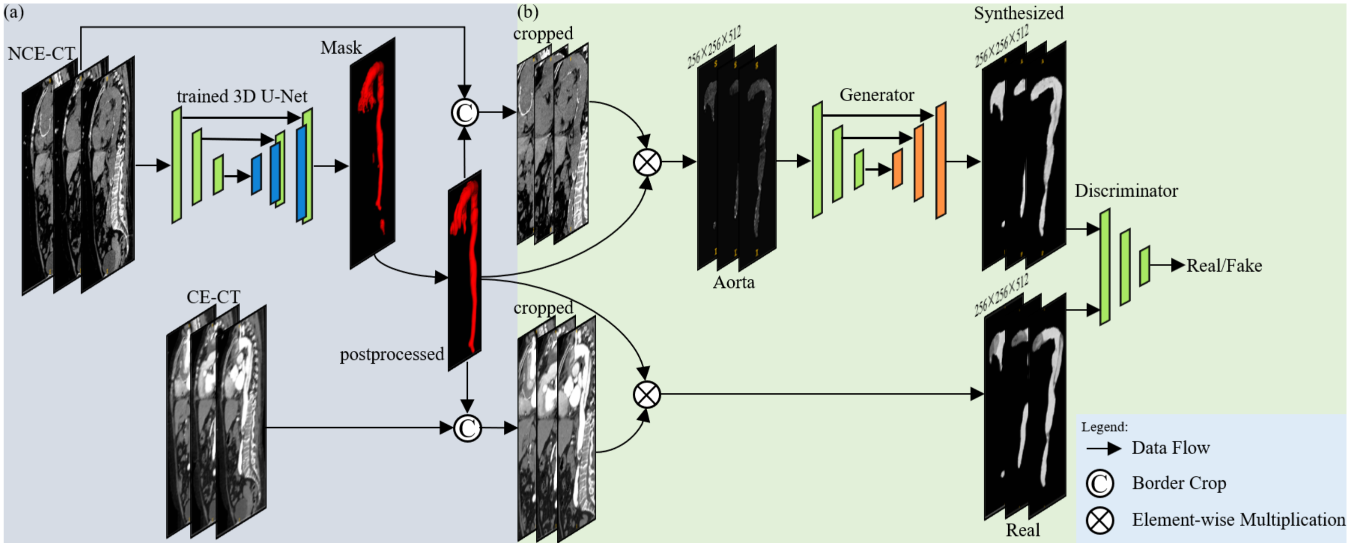

Fig. 1. An overview of proposed cascaded deep learning framework. (a) a 3D U-Net which is used to segment aorta from NCE-CT images. (b) a synthetic network is used to synthesize CE-CT images for aorta.

II. DATASET AND PREPROCESSING

A. Dataset collection

Total of 154 subjects, including 65 AD patients and 89 volunteers without AD but with other cardiovascular diseases, were recruited for collecting the paired datasets at two centers: Beijing Anzhen Hospital, Beijing, China, and Fujian Provincial Hospital, Fujian, China. All subjects volunteered to participate in this study which was approved by the Institutional Review Board. Each patient was sequentially conducted NCE-CT and then CE-CT with the same scan position, coverage, and parameters. The matrix size of acquired image is 512×512 with a resolution range from 0.625mm×0.625mm to 0.977mm×0.977mm. The range of slice thickness is from 0.625 mm to 1.250 mm. The CT images were acquired from GE MEDICAL SYSTEMS, TOSHIBA, and SIEMENS devices with KVP from 100 to 120. ECG was used during data collection to reduce motion artifacts. Furthermore, the data was collected at the end of respiration with breath-holding to reduce misregistration of the two paired datasets.

B. Dataset selection

24 paired datasets from patients without AD were randomly selected for segmentation of aorta. The rest 130 paired datasets, including 65 AD patients' datasets and 65 volunteers' datasets, were further selected for synthetic network. Five-fold cross validation protocol was used to evaluate the performance of proposed framework, with a experimental setting (80% as training & 20% as testing).

C. Dataset Preprocessing

To further reduce misregistration of the two paired datasets obtain from each subject, we use elastix tool of Slicer3D software [13] to make a rigid registration for the paired NCE-CT and CE-CT images. To prepare the mask of aorta for training segmentation network, the aorta was then manually contoured from CE-CT images which were obtained from the 24 volunteers without AD via the Slicer3D software. The segmentation network was used to segment the aorta from NCE-CT images. The NCE-CT images and CE-CT images were then multiplied by the obtained masks of aorta to prepare a paired dataset for synthetic network.

It notes that NCE-CT and CE-CT images are resampled to the average voxel spacing of $0.731 \times 0.731 \times 0.951\text{mm}^3$ using linear interpolation. The paired volumes for synthetic network are cropped and padded to a size of 256×256×512 for training acceleration. The volumes contain entire aorta that ranges from neck to pelvis. Due to abnormal CT value of aortic stent in the volume of postoperative patients, intensity values of NCE-CT volumes and CE-CT volumes is cut to [0, 90] and [0, 700], respectively.

III. METHODOLOGY

The cascaded deep learning framework was shown in Fig. 1, which contains a 3D segmentation network and a CGAN synthetic network. The 3D segmentation network was firstly used to segment aorta from the NCE-CT images and CE-CT images, and then CGAN was subsequently employed to map the NCE-CT images to the CE-CT images non-linearly for the entire aorta.

A. Segmentation of aorta

Encoder-decoder architectures based on deep convolutional neural networks, using skip connections to combine semantic information with spatial information, have been widely used in medical image segmentation, such as 2D U-Net [14] and 3D-UNet [15]. Compared with 2D U-Net, 3D U-Net has an advantage of using image information along with slice direction. In contrast to the networks which require expert knowledge and computing resources, nnU-Net [16] is recently developed as an out-of-the-box tool that can generate

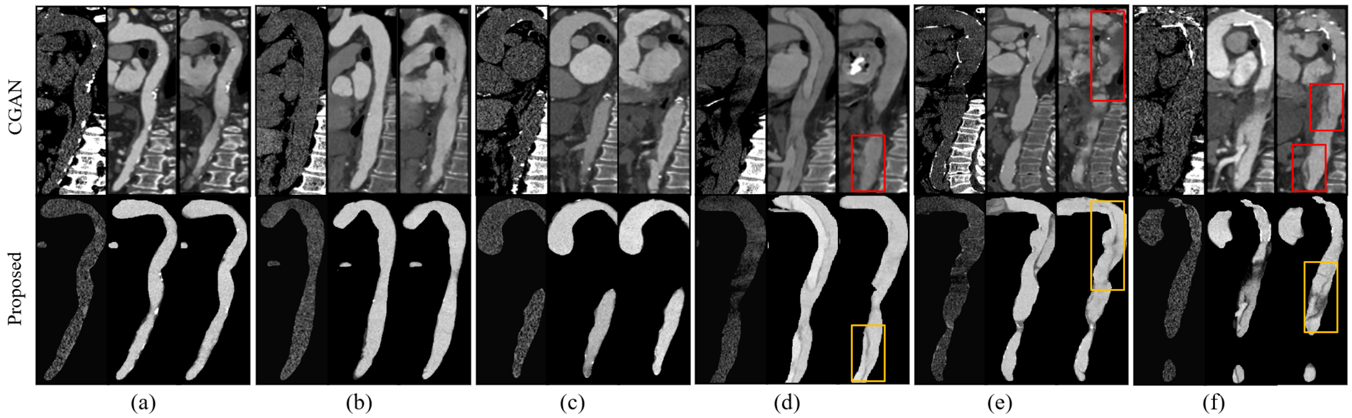

Fig. 2. Synthesis of CE-CT images from NCE-CT images via CGAN (first row) and proposed framework (second row). The figure contains 3 subjects without AD: (a), (b), (c) and 3 subjects with AD: (d), (e) and (f). Each subject is displayed by three images: NCE-CT image (left), real CECT image (middle) and synthesized CE-CT image (right). The figure demonstrates that our proposed framework synthesizes more realistic CE-CT images for aorta compared with CGAN alone.

state-of-the-art segmentation of various target structures and configure itself automatically. Thus, 3D nnU-Net was used for the segmentation of the aorta.

As shown in Fig. 1(a), the trained 3D U-Net, namely 3D nnU-Net, is employed to obtain the mask of aorta from NCE-CT images. The border of the mask is then smoothed in postprocessing and multiplied to the paired NCE-CT and CE-CT images for obtaining the region of aorta.

The segmentation experiment was conducted by a NVIDIA TITAN RTX GPU with 24GB memory. The patch size of 3D U-Net is defined as the size of $256\times256\times512$.

B. Synthesis of CE-CT images

The synthetic network is CGAN, inspired by the GAN-based method proposed by Isola et al. [17]. As shown in Fig. 1(b), the generator of the CGAN is a "U-Net" like architecture that shares low semantic information between NCE-CT images and synthesized CE-CT images. The discriminator of the CGAN is based on a 70×70 patchGAN [17]. It detects the synthesized CE-CT image is real or fake and guided the generator to synthesize images with sharp details.

To increase train samples for the CGAN, paired images are flipped by mirroring. We train the CGAN with an initial learning rate of 2×10^{-4} via Adam optimizer, and the batch size is set as 1. The CGAN is implemented with Pytorch framework and we trained it on a NVIDIA TITAN GPU with 24GB memory for 200 epochs.

IV. EXPERIMENTS AND RESULTS

A. Qualitive Evaluation

Our proposed framework is compared with CGAN with same train and test datasets. The CGAN synthesizes CE-CT images directly, whereas the proposed framework synthesizes CE-CT images for aorta only. As shown in Fig. 2, the proposed framework can synthesize clearer images than using CGAN alone. Specifically, in Fig. 2(a), (b), and (c), the proposed framework synthesizes clearer and smoother images for aorta without AD. In Fig. 2(d), the proposed framework synthesizes an intimal flap only in abdominal aorta (see the yellow box), whereas the CGAN synthesizes nothing in aorta (see the red box). In Fig. 2(e), the proposed framework synthesizes an intimal flap in thoracic aorta (see the yellow box), whereas the CGAN synthesizes a blurry and inhomogeneous aorta (see the red box). In Fig. 2(f), for a postoperative subject, the proposed framework synthesizes a winding intimal flap in abdominal aorta (see the yellow box), whereas the CGAN synthesizes two short and blurry intimal flaps near the aortic wall (see the red box).

B. Quantitative Evaluation

Three metrics, i.e., peak signal-to-noise ratio (PSNR), structural similarity (SSIM), mean absolute error (MAE), were used to evaluate the performance of CGAN and the proposed framework on synthesizing CE-CT images from NCE-CE images. For a fair comparison, all metric scores are calculated on pixels of aorta. Therefore, the region of aorta is extracted from the real CE-CT images and synthesized CE-CT images by multiplying the mask. As listed in Table I, the proposed framework significantly improves the PSNR scores and MAE scores and performs with a slight increase in SSIM scores compared to the CGAN.

C. Clinical diagnostic performance

Clinical performance of the proposed network was evaluated by an experienced radiologist. As the synthetic CE-CT images are essentially different from the real CE-CT images, the radiologist was trained to learn a criterion for the diagnosis of AD using a fold of the synthesized and corresponding real CE-CT first. The other four folds are then used to test clinical diagnostic performance. The criteria of determining AD on the synthetic CE-CT images includes a wider diameter of aorta, inhomogeneous CT density, a long intimal flap and consecutive short intimal flaps that are parallel or tilted to edge of aortic wall. Otherwise, the synthetic CE-CT images are defined without AD. Table II shows the diagnostic results using the synthetic CE-CT

TABLE I
COMPARISION BETWEEN CGAN AND PROPOSED FRAMEWORK USING THREE METRICS WITH 5-FOLD CROSS VALIDATION. (MEAN±STD)

Experiments	CGAN		
	PSNR	SSIM	MAE
Fold1	31.154±4.049	0.986±0.004	2.367±1.470
Fold2	30.548±3.654	0.986±0.006	2.499±1.329
Fold3	32.077±4.098	0.987±0.004	2.105±1.156
Fold4	31.602±3.451	0.985±0.006	2.143±0.957
Fold5	31.102±3.335	0.985±0.006	2.373±1.287
Average	31.297±3.763	0.985±0.005	2.297±1.261
Experiments	**Proposed framework**		
	PSNR	SSIM	MAE
Fold1	33.776±2.847	0.990±0.003	1.247±0.561
Fold2	32.639±3.691	0.989±0.005	1.654±1.278
Fold3	34.590±3.088	0.991±0.003	1.196±0.787
Fold4	33.806±3.446	0.990±0.004	1.375±0.917
Fold5	33.447±2.915	0.989±0.004	1.377±0.731
Average	33.652±3.271	0.990±0.004	1.369±0.899

images obtained by the proposed framework and CGAN. It demonstrates that the proposed framework has significantly higher accuracy and sensitivity compared to CGAN alone. The proposed framework has an approximate specificity with CGAN alone.

TABLE II
COMPARISION BETWEEN CGAN AND PROPOSED FRAMEWORK BY PERCEPTUAL STUDY ON FOUR FOLDS.

Experiments	CGAN		
	Accuracy	Sensitivity	Specificity
Fold1	80.77%	90.91%	71.43%
Fold2	65.38%	45.45%	78.57%
Fold3	69.23%	69.23%	69.23%
Fold4	57.69%	69.23%	46.15%
Average	68.27%	68.71%	66.35%
Experiments	**Proposed framework**		
	Accuracy	Sensitivity	Specificity
Fold1	76.92%	91.67%	64.29%
Fold2	80.77%	100%	64.29%
Fold3	80.77%	100%	61.54%
Fold4	84.61%	92.31%	76.92%
Average	80.77%	96.00%	66.76%

V. DISCUSSION AND CONCLUSION

The aortic wall and blood have slightly different X-ray decay coefficients on NCE-CT images [18], making the proposed framework for synthesizing CE-CT images from NCE-CT images to detect AD become possible. It notes that the sensitivity of the proposed framework is high with an average of 96%. This indicates the proposed framework can reduce the misdiagnosis and prompt the patient to make a further examination. ECG simplifies the rigid registration and can lead to negligible registration errors. For future work, we will combine the synthesized network with an auxiliary convolutional neural network to automatically classify AD/Non-AD subjects using the synthesized CE-CT images.

In summary, our experimental results demonstrate that the proposed framework can be used for detecting AD and significantly outperformed CGAN alone.

REFERENCES

[1] Joanna Gawinecka, Felix Schönrath, and Arnold von Eckardstein. Acute aortic dissection: pathogenesis, risk factors and diagnosis. *Swiss medical weekly*, 147:w14489, 2017.

[2] Pawan D Patel and Rohit R Arora. Pathophysiology, diagnosis, and management of aortic dissection. *Therapeutic advances in cardiovascular disease*, 2(6):439–468, 2008.

[3] A Karthikesalingam, PJE Holt, RJ Hinchliffe, MM Thompson, and IM Loftus. The diagnosis and management of aortic dissection. *Vascular and endovascular surgery*, 44(3):165–169, 2010.

[4] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014.

[5] Gianmarco Santini, Lorena M Zumbo, Nicola Martini, Gabriele Valvano, Andrea Leo, Andrea Ripoli, Francesco Avogliero, Dante Chiappino, and Daniele Della Latta. Synthetic contrast enhancement in cardiac ct with deep learning. *arXiv preprint arXiv:1807.01779*, 2018.

[6] Enhao Gong, John M Pauly, Max Wintermark, and Greg Zaharchuk. Deep learning enables reduced gadolinium dose for contrast-enhanced brain mri. *Journal of magnetic resonance imaging*, 48(2):330–340, 2018.

[7] Sahin Olut, Yusuf H Sahin, Ugur Demir, and Gozde Unal. Generative adversarial training for mra image synthesis using multi-contrast mri. In *International workshop on predictive intelligence in medicine*, pages 147–154. Springer, 2018.

[8] Dong Nie, Roger Trullo, Jun Lian, Caroline Petitjean, Su Ruan, Qian Wang, and Dinggang Shen. Medical image synthesis with context-aware generative adversarial networks. In *International conference on medical image computing and computer-assisted intervention*, pages 417–425. Springer, 2017.

[9] Jelmer M Wolterink, Anna M Dinkla, Mark HF Savenije, Peter R Seevinck, Cornelis AT van den Berg, and Ivana Išgum. Deep mr to ct synthesis using unpaired data. In *International workshop on simulation and synthesis in medical imaging*, pages 14–23. Springer, 2017.

[10] Heran Yang, Jian Sun, Aaron Carass, Can Zhao, Junghoon Lee, Zongben Xu, and Jerry Prince. Unpaired brain mr-to-ct synthesis using a structure-constrained cyclegan. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 174–182. Springer, 2018.

[11] Yafen Li, Wen Li, Pin He, Jing Xiong, Jun Xia, and Yaoqin Xie. Ct synthesis from mri images based on deep learning methods for mri-only radiotherapy. In *2019 International Conference on Medical Imaging Physics and Engineering (ICMIPE)*, pages 1–6. IEEE, 2019.

[12] Cheng-Bin Jin, Hakil Kim, Mingjie Liu, Wonmo Jung, Seongsu Joo, Eunsik Park, Young Saem Ahn, In Ho Han, Jae Il Lee, and Xuenan Cui. Deep ct to mr synthesis using paired and unpaired data. *Sensors*, 19(10):2361, 2019.

[13] Ron Kikinis, Steve D Pieper, and Kirby G Vosburgh. 3d slicer: a platform for subject-specific image analysis, visualization, and clinical support. In *Intraoperative imaging and image-guided therapy*, pages 277–289. Springer, 2014.

[14] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[15] Adeleh Bitarafan, Mahdi Nikdan, and Mahdieh Soleymanibaghshah. 3d image segmentation with sparse annotation by self-training and internal registration. *IEEE Journal of Biomedical and Health Informatics*, 2020.

[16] Fabian Isensee, Paul F Jaeger, Simon AA Kohl, Jens Petersen, and Klaus H Maier-Hein. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2):203–211, 2021.

[17] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.

[18] Michael E Phelps, Edward J Hoffman, and Michel M Ter-Pogossian. Attenuation coefficients of various body tissues, fluids, and lesions at photon energies of 18 to 136 kev. *Radiology*, 117(3):573–583, 1975.