Enhance Regional Wall Segmentation by Style Transfer for Regional Wall Motion Assessment

Kaikai Liu[1]

Qing Lu[2], Boyang Li[2], Xiaoshan Li[3], Yiyu Shi[2] Jian Zhuang[3], Meiping Huang[3], Hongwen Fei[3], Erlei Zhang[1]

Xiaowei Xu[3]
https://xiaoweixu.github.io/

[1] School of Information and Technology, Northwest A&F University, Yanglin, China.

[2] Department of Computer Science and Engineering, University of Notre Dame, IN, USA

[3] Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China

Abstract

Regional wall motion assessment is critical in the diagnosis of coronary artery diseases and is usually performed using echocardiography in clinical practice. As manual assessment of regional wall motion is time-consuming and requires strong expertise, various automated methods have been proposed in which regional wall segmentation is the key step for detailed diagnosis and treatment. Generally, the existing methods suffer from low segmentation performance. In addition, there are no publicly available datasets which also introduces difficulties to related research in the community. In this paper, we propose to enhance regional wall segmentation by style transfer for regional wall motion assessment. **The motivation is that echocardiography images in 2D mode are inexpensive but the contrast is low while echocardiography images in myocardial contrast echocardiography (MCE) mode and left ventricle opacification (LVO) mode are invasive and expensive but with high contrast, so style transfer may help enhance the contrast of echocardiography images in 2D mode.** Specifically, we first transfer echocardiography images in 2D mode to echocardiography images in LVO or MCE mode using style transfer. Then, nnU-Net is adopted for regional wall segmentation. Note that the two networks are trained in an end-to-end manner. For evaluation, we collected a dataset of 198 patients each with three views (including A2C, A3C, and A4C) in three modes (including 2D mode, MCE mode, and LVO mode). Experimental results show that our framework can improve the Dice score by 7.53% compared with existing works. However, the Dice score is still below 60%, leaving much room for further improvement. The dataset and code in our study are released to the public [1].

Erlei Zhang and Xiaowei Xu are the corresponding authors.

1 Introduction

Coronary artery disease (CAD) is the third leading cause of death worldwide and kills 17.8 million people each year [2]. The assessment of left ventricular systolic function and particularly regional wall motion abnormality have become increasingly important in determining the severity and prognosis of CAD. Currently, echocardiography has been widely used for evaluating regional wall motion abnormalities [6]. However, manual assessment requires extensive experience, hence is subjective, time-consuming, and hard to reproduce [8, 17, 18, 19].

Currently, two automated approaches including rough evaluation and detailed evaluation have been proposed to solve the problem. In rough evaluation, only the abnormality is evaluated without detailed abnormal vessel positions. In detailed evaluation, detailed abnormal vessel positions are also provided. Recently, several works [9, 14] have adopted deep learning to perform a rough evaluation of regional wall motion abnormalities from echocardiography. Kusunose et al. [9] used deep neural networks to detect whether the regional wall motion is normal or abnormal, which was evaluated on a dataset consisting of 400 patients each with short-axis views in 2D mode. A recent work by Sanjeevi et al. [14] extracted temporal information from raw echocardiography images in 2D mode to enhance the detection performance. In the detailed evaluation approach, most of the works focus on regional wall segmentation [5, 22], which is the key step for further detailed diagnosis. Vasily et al. [22] adopted U-net [13] for regional wall segmentation with a dataset of 94 patients with apical 4-chamber view (A4C) in 2D mode. Huang et al. [5] also adopted U-net for regional wall segmentation and evaluation with a large dataset covering 5,772 patients each with 4 views including long-axis view, short-axis view at the papillary muscle level, A4C, and apical 2-chamber view (A2C) in 2D mode.

In clinical practice, detailed evaluation is preferred [6]. With the regional wall segmentation, it can be easily further divided into 18 segments and each of the three views (A2C, A3C, and A4C) can be divided into six segments [6]. Then, accurate abnormal vessel positions can be provided to help clinicians and surgeons perform precise diagnosis and treatment. However, there are mainly two challenges in regional wall segmentation. First, the segmentation performance is marginally acceptable. For example, the Dice score of regional wall segmentation is only 75.6% in [5], which is too low for further analysis and diagnosis. Second, existing works [5, 9, 14, 22] almost all adopted in-house datasets for evaluation, which not only introduce difficulties for fair comparison but also limit further investigation and development from the community.

On the other hand, there are three modes including 2D mode, myocardial contrast echocardiography (MCE) mode, and left ventricle opacification (LVO) mode used in echocardiography images acquisition. 2D mode is the most widely used one as it is cheap and non-invasive, and existing works are also based on echocardiography images in 2D mode. However, this technique struggles to clearly display all the boundaries of the myocardium and valves. If the regional wall motion assessment using 2D mode is highly suspicious, LVO or MCE mode is then applied in which gas-filled microbubbles are injected to the left ventricle or the myocardium to enhance the contrast. However, the two modes are invasive and expensive. Here comes the question: **is it possible to enhance the contrast of echocardiography images in 2D mode to be the same as that in LVO or MCE mode by using deep learning?**

Motivated by the above analysis, we propose to enhance regional wall segmentation by style transfer for regional wall motion assessment. Particularly, we first transfer echocardiography images in 2D mode to high-contrast echocardiography images (e.g. echocardiography

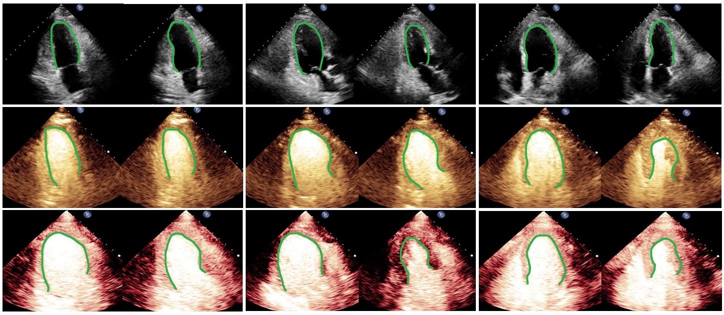

Figure 1: Examples in our dataset with three modes including (a) two-dimensional mode (2D), (b) left ventricle opacification (LVO) mode, and (c) myocardial contrast echocardiography (MCE) mode. Each mode contains three views including A2C, A3C, and A4C.

images in MCE mode and LVO mode) using style transfer. Then, nnU-Net [7] is adopted for regional wall segmentation. Note that the two networks are trained in an end-to-end manner. For evaluation, we collected a dataset of 198 patients each with three views (including A2C, A3C, and A4C) in three modes, including 2D mode, MCE mode, and LVO mode. Experimental results show that our framework can improve the Dice score by 7.53% compared with existing works. However, the Dice score is still below 60.0%, leaving much room for further improvement. The dataset and code in our study are released to the public [1]. We hope that the dataset can stimulate further research on this important and interesting topic.

2 Dataset

Our dataset consists of 198 patients where for each patient three views (A4C, A3C, and A2C) in three modes (2D mode, MCE mode, and LVO mode) are included as shown in Figure 2. Thus, there are totally 1,782 ultrasound videos, with a varying frame size of $640\times480\times(71-510)$. Note that in clinical practice, 2D mode, MCE mode, and LVO mode are all used for assessment of regional wall motion abnormality, while 2D mode is the most commonly used one. For each video, six frames are randomly selected, and the regional wall contour of the left ventricular is labeled as shown in Figure 1. There are a total of 594 wall labels, among which 45 walls are abnormal. Note that such low incidence rate is a reflection of the real-life statistics of clinical practice at our center. All labels were annotated by four experienced sonographers, and each taking about 5 minutes to finish.

3 Method

3.1 Overview

As shown in Figure 2, the proposed framework includes two modules: style transfer, and regional wall segmentation. The style transfer module adopts cycle generative adversarial network (CycleGAN) [20] to transfer 2D echocardiography images to pseudo images in

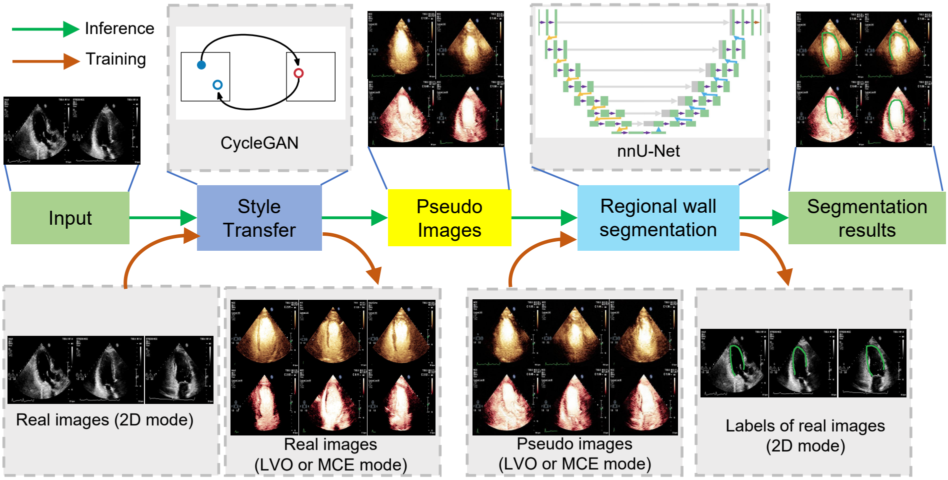

Figure 2: Overview of the proposed framework which includes two modules: style transfer, and regional wall segmentation. The style transfer module adopts cycle generative adversarial network (CycleGAN) [20] to transfer 2D echocardiography images to pseudo images in LVO or MCE mode. The regional wall segmentation module performs the segmentation using nnU-Net [7].

LVO or MCE mode. The regional wall segmentation module performs the segmentation using nnU-Net [7].

3.2 Style Transfer

Unpaired GAN specifically CycleGAN [20] is chosen in the style transfer module due to the fact that echocardiography images in 2D, LVO and MCE modes are unpaired data. First, the echocardiography acquisition is performed by sonographers thus is quite subjective with varying angles and quality of the captured views. Second, patients also have varying conditions including heart size, heartbeat frequency and breathing influence.

The adopted CycleGAN consists of two generators and two discriminators, utilizing generator networks based on ResNet [4] that contains 9 residual blocks. The generators transform input images from one modality domain to another, and the two generators are used to convert images between the two domains. The purpose of the two discriminators is to distinguish between real images and generated images within a given modality domain. One discriminator is used to differentiate between real and generated images in one domain, while the other is used to differentiate between real and generated images in the other domain. There are three loss functions used: $criterion_{GAN}$, which measures the difference between the generated images and real images, using mean squared error (MSE) loss; $criterion_{cycle}$, which measures the image restoration quality after cycling between the two modality domains, using $L1$ Loss; and $criterion_{identity}$, which measures the ability of the generator to preserve image features without changing the input image, also using $L1$ Loss in our configuration. More details of the configuration can be found at [20].

3.3 Regional Wall Segmentation

We adopt nnU-Net [7] for regional wall segmentation. The input is a 2D ultrasound image with a normalized size of 480×360. The number of resolution levels is 5, and the number of batch sizes is 50. We have made several improvements as follows. First, we adopt weighted Dice loss [10] and cross-entropy loss to enhance the performance. Particularly, the weights of the background, and the regional wall are 1 and 5, respectively. Second, we perform position correlation using several convolutional layers [12], and the final output is a point-wise multiplication of the U-net output and the output of the position correlation process. Third, we adopted instance normalization [15] after each convolutional layer to mitigate the ingredient missing problem.

3.4 Training Details

Note that the two networks including CycleGAN and nnU-Net are trained in an end-to-end manner. The specific process is as follows. First, echocardiography images in 2D mode and LVO or MCE mode are used for the training of CycleGAN. Then, the trained style transfer model is used to perform style transfer on echocardiography images in 2D mode, generating pseudo echocardiography images in LVO or MCE mode. Next, nnU-Net is trained on the pseudo echocardiography images for a number of epochs while the weights in CycleGAN is frozen. Finally, the two networks are jointly trained for several epochs to enhance their performance. As we also labeled the regional walls in echocardiography images in LVO and MCE modes, we evaluate the performance when nnU-Net is trained with both pseudo echocardiography images, and echocardiography images in LVO and MCE modes. More details can be referred to Table 1.

For the individual training of CycleGAN, we use the Adam optimizer and employ a learning rate scheduler to gradually reduce the learning rates during training. The number of epochs is set to 200, the batch size to 16, the learning rate of the Adam optimizer to 0.0002, and β_1 and β_2 in the Adam optimizer to 0.5 and 0.999, respectively. Learning rate decay starts from the $100th$ epoch. For the individual training of nnU-Net, The number of epochs is set to 1000, with other training configurations such as data augmentation, data normalization, and learning rate decay strategies adhering to its original implementation [7].

4 Experiments

4.1 Experiments Setup

All experiments were conducted on an Nvidia A40 GPU with 48GB of memory. The CycleGAN and nnU-Net were implemented using the PyTorch framework [11]. To ensure a robust evaluation, we employed k-fold cross-validation. Given the limited number of abnormal segments, we utilized k values of 2, 3, 4, and 5 to obtain an average performance that accurately reflects the efficacy of our approach. In each cross-validation setting, 9,881 2D images were split by the ratios of 1-1/k-0.1, 0.1, and 1/k for training, validation, and testing, respectively. Note that the partitioning was conducted to ensure the ratio of cases with abnormal wall motion in each set to be roughly equal. For the evaluation of our experiment, Dice score and Hausdorff distance were used as the metrics.

2*Method	Training and test dataset configuration			2*Dice score	2*Hausdorff distance
	CycleGAN	nnU-Net	Test		
Existing works [5, 22]	NA	2D	2D	43.52%	23.67
2*Proposed method	2D, LVO	2D	2D	49.26%	19.68
	2D, MCE	2D	2D	52.83%	18.61
Refer-Imple A	NA	LVO	LVO	63.91%	18.60
Refer-Imple B	NA	MCE	MCE	66.37%	16.44
Refer-Imple C	NA	2D, LVO, MCE	2D	48.86%	18.42
1*Refer-Imple D	2D, LVO	2D, LVO, MCE	2D	53.61%	19.67
Refer-Imple E	2D, MCE	2D, LVO, MCE	2D	57.06%	12.86

Table 1: Segmentation performance in Dice score (in %) and Hausdorff distance (pixels) of existing works [5, 22] and our proposed method. Note that labels of regional wall segmentation are not required for the training of CycleGAN, and are required for the training of nnU-Net. Refer-Imple is short for reference implementation.

4.2 Quantitative Results

Quantitative results are shown in Table 1. Our proposed method can improve the segmentation performance by 5.74% and 9.31% compared with existing works [5, 22] on echocardiography images in LVO and MCE modes, respectively. The Hausdorff distance is also improved by 3.99 and 5.06 on echocardiography images in LVO and MCE modes, respectively. The possible reason is that style transfer works as some sort of feature extraction which can enhance the contrast in echocardiography images in 2D mode.

Performance of five reference implementations (Refer-Imple) are also reported for discussion. First, we can easily notice that nnU-Net with only echocardiography images in LVO (Refer-Imple A) or MCE (Refer-Imple B) mode can achieve the optimal segmentation performance, which is expected. In clinical practice, the two modes can obtain high contrast of the myocardium boundary as discussed in Section 1. Second, nnU-Net trained with echocardiography images in all the three modes (Refer-Imple C) can obtain a higher segmentation performance than existing works. A possible reason is that echocardiography images in LVO and MCE mode can help detect more features in echocardiography images in 2D mode. This phenomenon also confirms the reason of improvement in our proposed method, in which the style transfer module trained by echocardiography images in 2D, and LVO or MCE mode can help enhance the contrast. Third, compared with our proposed methods, nnU-Net trained with the addition of echocardiography images in LVO (Refer-Imple D) and MCE (Refer-Imple E) modes improves the segmentation performance by 4.23% and 4.35%, respectively, compared with existing methods. The same reason discussed above also applies here. We can further notice that though with the help of echocardiography images in LVO and MCE modes, Refer-Imple D and Refer-Imple E achieve high improvements compared with existing works but still low than Refer-Imple A and Refer-Imple B. The main reason is that echocardiography images in LVO and MCE modes are enhanced with more contrast features which does not exist in echocardiography images in 2D mode.

Note that existing methods [5] obtains a Dice score of 75.6% in regional wall segmentation. However, ours is even lower which is mainly caused by the dataset difference. The

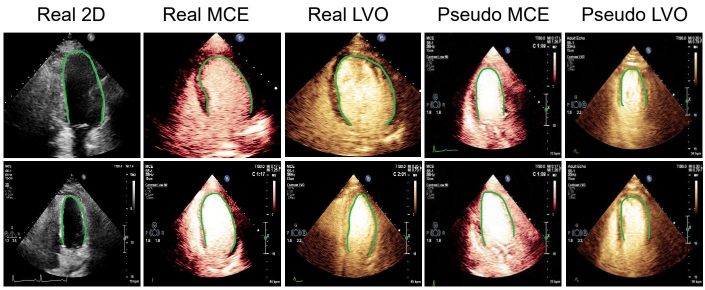

Figure 3: Examples of good segmentation results of the proposed framework with corresponding echocardiography images. Segmentation results of echocardiography images in LVO and MCE modes are also provided for comparison.

echocardiography machines, images, and labeling strategies are quite different between our dataset and theirs. In addition, their implementations are not public available, which also introduces some difficulties for fair comparison. Note that we tried our best to implement their framework with the same configuration of nnU-Net in our implementation. This also shows the urgent need of publicly available datasets and codes to facilitate related research.

4.3 Qualitative Results

Qualitative results including good and bad segmentation results are shown in Figure 3 and Figure 4, respectively. Our framework can clearly segment the regional walls without breaks if the image quality is good and there are no complex context as shown in Figure 3. Otherwise bad segmentation appears as shown in Figure 4. For example, in Figure 4(a), the apical segment is not detected because of the blurred border in the echocardiography image. In Figure 4(b) and (c), there are segmentation breaks in the basal segments and mid segments. This type of error leads to an inaccurate differentiation between the end-systolic and end-diastolic volumes and thus the misclassification of regional wall motion abnormalities.

There are mainly three causes that may lead to segmentation errors including echo loss [3], lung air [21], and papillary muscles [16]. Figure 4 shows several examples of wall missing due to echo loss. The apical part suffers from a blurring boundary which can not be easily detected. Note that echo loss is caused by a variety of reasons including tissue density, acoustic frequency and wrong operation. In clinical practice, sonographer experts may obtain the wall contour based on their experience. We may add an image quality assessment module to filter out low-quality images like the ones with echo loss to mitigate this problem. Figure 4(b) shows some examples with interrupted wall contours at the mid-segment regions of the heart due to lung air. Because most echocardiography scans can last between 15 and 45 minutes, respiration of the patients is inevitable and some air can reach near the heart during the process. As the density of the lung air and the background is very close (especially in the 2D mode), the boundaries of the heart will disappear from the echocardiography images (especially in the 2D mode), making the segmentation much more difficult. Figure 4(c) shows examples of segmentation errors at the basal segments due to papillary muscles. Since papillary muscles move during acquisition, the shape and location of papillary muscles can vary. They may appear in some frames and disappear in some others. The shape of wall

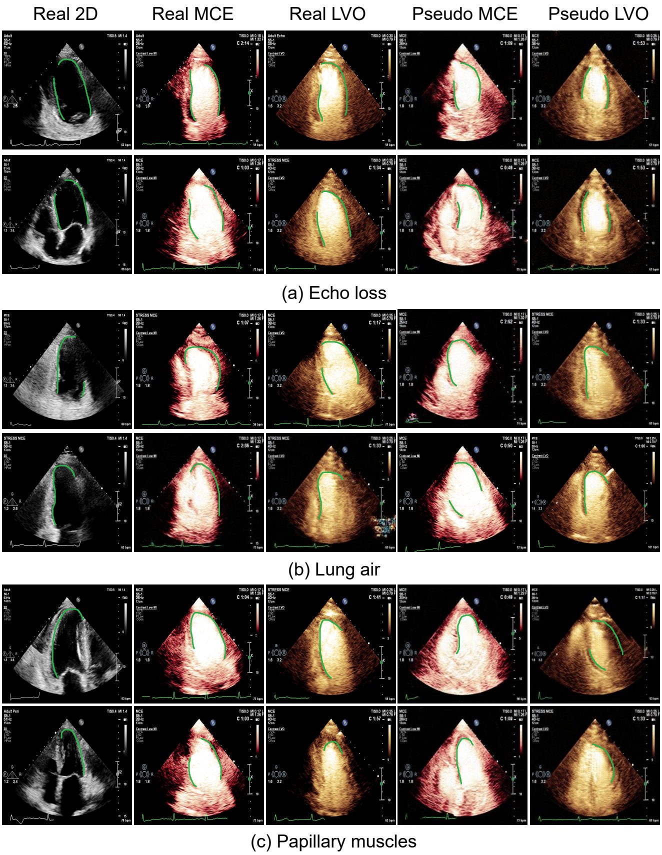

Figure 4: Examples of bad segmentation results of the proposed framework which are due to (a) echo loss, (b) lung air, and (c) papillary muscles. Segmentation results of echocardiography images in LVO and MCE modes are also provided for comparison.

contours with papillary muscles is quite different from those without as shown in Figure 4(c).

4.4 Results of Style Transfer

Qualitative results are shown in Figure 5. Overall, the pseudo echocardiography images in LVO and MCE modes obtain a quite similar style compared with the real ones, respectively. Compared with the original echocardiography images in 2D mode, the corresponding pseudo ones in LVO and MCE modes seems to have enhanced boundaries of the myocardium as indicated by green circles in Figure 5. In addition, the contrast of the mitral valve seems to be improved as noted by dashed blue circles in Figure 5. The contrast enhancement of the myocardium boundary and the mitral valve well explains the segmentation performance improvement of our proposed method. However, we can notice that real echocardiography images still have higher contrast (clear myocardium boundaries) than pseudo ones.

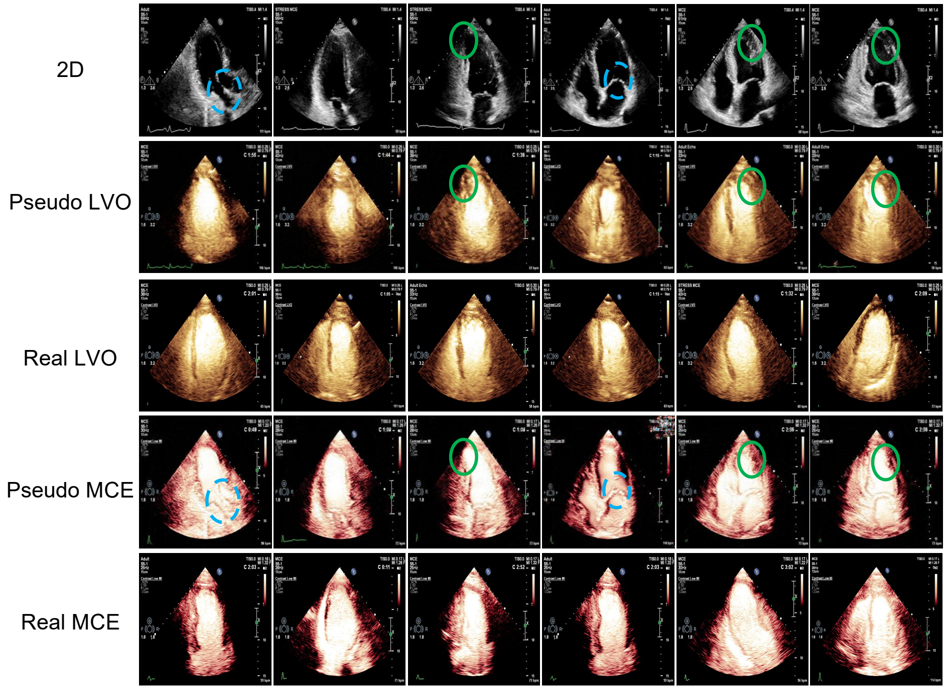

Figure 5: Qualitative results of the style transfer module (CycleGAN). Pseudo echocardiography images seems to obtains enhanced boundaries of the myocardium and the mitral valve as indicated by the green and dashed blue circles, respectively.

4.5 Discussion

In this paper, we transformed the echocardiography images in 2D mode into pseudo ones in LVO or MCE mode for segmentation. We collected a dataset for segmentation and the results shows our method can successfully improve the segmentation performance by 7.5% on average. Thus, with no effort in labeling of echocardiography images in LVO and MCE modes, our method can enhance the segmentation effectively with the help of style transfer. This shows great potential in clinical practice. Cheap but less accurate diagnosis methods can be enhanced by expensive but accurate ones, which can help improve efficiency and patient experience, and reduce costs in clinical practice.

However, our research has certain limitations. Firstly, we need to expand the scale of the dataset to further improve the segmentation performance as the current data size is still limited for training and evaluation. Secondly, we need to validate our framework in more real-world clinical scenarios (e.g. more echocardiography machines) to ensure its stability and reliability under various conditions. Third, the segmentation performance of our method is quite limited for clinical practice, and future research can attempt to adopt other advanced deep learning techniques to optimize the performance.

5 Conclusion

In this paper, we proposed to enhance regional wall segmentation by style transfer for regional wall motion assessment. Particularly, we first transfer 2D echocardiography images to high-contrast echocardiography images (e.g. echocardiography images in LVO and MCE modes) using style transfer. Then, nnU-Net is adopted for regional wall segmentation. For evaluation, we collected a dataset of 198 patients each with three views (including A2C, A3C, and A4C) in three modes, including 2D mode, MCE mode, and LVO mode. Experimental results show that our framework can improve the Dice score by 7.5% on average compared with existing works. Thus, with echocardiography images in LVO and MCE modes, our method can enhance the segmentation effectively with the help of style transfer, which shows promising potential in clinical practice. Cheap but less accurate diagnosis methods can be enhanced by expensive but accurate ones, which can help improve efficiency and patients' experience and reduce costs in clinical practice. The dataset and code in our study are released to the public [1] to stimulate further research in this important and interesting topic.

6 Acknowledgment

This work was approved by the Research Ethics Committee of Guangdong General Hospital, Guangdong Academy of Medical Science under Protocol No. KY-N-2022-048-01. This work was supported by the Science and Technology Planning Project of Guangdong Province, China (No. 2019B020230003), Guangdong Peak Project (No. DFJH201802), Guangzhou Science and Technology Planning Project (No. 202206010049), the National Natural Science Foundation of China (No. 62006050, No. 62276071), Guangdong Special Support Program-Science and Technology Innovation Talent Project (No. 0620220211), Guangdong Basic and Applied Basic Research Foundation (No. 2022A1515010157, No. 2022A1515011650), Guangzhou Science and Technology Planning Project (No. 202102 080188), and the QinChuangYuan High-Level Innovation and Entrepreneurship Talent Program of Shaanxi (No. 2021QCYRC4-50).

References

[1] Our dataset. *https://github.com/XiaoweiXu/Segment-level-Assessment-of-Regional-Wall-Motion-Abnormality-from-Echocardiography-Images*.

[2] Jonathan C Brown, Thomas E Gerhardt, and Edward Kwon. Risk factors for coronary artery disease. 2020.

[3] Sergey V Dushin and Sergey S Shavrin. Efficient echo cancellation in single carrier duplex satellite systems. In *2019 International Conference on Engineering and Telecommunication (EnT)*, pages 1–5. IEEE, 2019.

[4] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[5] Mu-Shiang Huang, Chi-Shiang Wang, Jung-Hsien Chiang, Ping-Yen Liu, and Wei-Chuan Tsai. Automated recognition of regional wall motion abnormalities through

deep neural network interpretation of transthoracic echocardiography. *Circulation*, 142 (16):1510–1520, 2020.

[6] Borja Ibanez, Stefan James, Stefan Agewall, Manuel J Antunes, Chiara Bucciarelli-Ducci, Héctor Bueno, Alida LP Caforio, Filippo Crea, John A Goudevenos, Sigrun Halvorsen, et al. 2017 esc guidelines for the management of acute myocardial infarction in patients presenting with st-segment elevation: The task force for the management of acute myocardial infarction in patients presenting with st-segment elevation of the european society of cardiology (esc). *European heart journal*, 39(2):119–177, 2018.

[7] 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.

[8] Zhenge Jia, Jianxu Chen, Xiaowei Xu, John Kheir, Jingtong Hu, Han Xiao, Sui Peng, Xiaobo Sharon Hu, Danny Chen, and Yiyu Shi. The importance of resource awareness in artificial intelligence for healthcare. *Nature Machine Intelligence*, pages 1–12, 2023.

[9] Kenya Kusunose, Takashi Abe, Akihiro Haga, Daiju Fukuda, Hirotsugu Yamada, Masafumi Harada, and Masataka Sata. A deep learning approach for assessment of regional wall motion abnormality from echocardiographic images. *Cardiovascular Imaging*, 13(2_Part_1):374–381, 2020.

[10] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, pages 565–571. IEEE, 2016.

[11] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.

[12] Christian Payer, Darko Štern, Horst Bischof, and Martin Urschler. Multi-label whole heart segmentation using cnns and anatomical label configurations. In *International Workshop on Statistical Atlases and Computational Models of the Heart*, pages 190–198. Springer, 2017.

[13] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015.

[14] G Sanjeevi, Uma Gopalakrishnan, Rahul Krishnan Pathinarupothi, and Thushara Madathil. Automatic diagnostic tool for detection of regional wall motion abnormality from echocardiogram. *Journal of Medical Systems*, 47(1):13, 2023.

[15] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*, 2016.

[16] Benson R Wilcox, Andrew C Cook, and Robert H Anderson. *Surgical anatomy of the heart*. Cambridge university press, 2005.

[17] Xiaowei Xu, Tianchen Wang, Yiyu Shi, Haiyun Yuan, Qianjun Jia, Meiping Huang, and Jian Zhuang. Whole heart and great vessel segmentation in congenital heart disease using deep neural networks and graph matching. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II 22*, pages 477–485. Springer, 2019.

[18] Xiaowei Xu, Tianchen Wang, Jian Zhuang, Haiyun Yuan, Meiping Huang, Jianzheng Cen, Qianjun Jia, Yuhao Dong, and Yiyu Shi. Imagechd: A 3d computed tomography image dataset for classification of congenital heart disease. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part IV 23*, pages 77–87. Springer, 2020.

[19] An Zeng, Chunbiao Wu, Wen Xie, Jin Hong, Meiping Huang, Jian Zhuang, Shanshan Bi, Dan Pan, Najeeb Ullah, Kaleem Nawaz Khan, et al. Imagecas: A large-scale dataset and benchmark for coronary artery segmentation based on computed tomography angiography images. *Computerized Medical Imaging and Graphics*, page 102287, 2023.

[20] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.

[21] Yun Zhu, Xenophon Papademetris, Albert Sinusas, and James Duncan. Automated segmentation of real-time 3d echocardiography using an incompressibility constraint. *Echocardiography—New Techniques*, 2012.

[22] Vasily Zyuzin, Porshnev Sergey, Andrey Mukhtarov, Tatyana Chumarnaya, Olga Solovyova, Anastasia Bobkova, and Vladislav Myasnikov. Identification of the left ventricle endocardial border on two-dimensional ultrasound images using the convolutional neural network unet. In *2018 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)*, pages 76–78. IEEE, 2018.