

Automatic Epicardial Adipose Tissue Segmentation in Cardiac CT with Position Regularization

Qinghe Yuan¹, Qiong Su², Zhenteng Li¹, Qi Wen¹, Zhikang Lin¹, Dajun Chai², and Sheng Lian¹(✉)

¹ Fuzhou University, Fuzhou, China

² Cardiovascular Department, Fujian Institute of Hypertension
The First Affiliated Hospital, Fujian Medical University, Fuzhou, China

Abstract. Cardiovascular diseases (CVDs) are a major global health concern. Epicardial adipose tissue (EAT) has been identified as playing a significant role in the pathogenesis and progression of cardiovascular diseases. While deep learning-based methods have shown promising results in EAT segmentation, they primarily treat EAT as a whole and do not consider the urgent clinical need for fine-grained segmentation at different locations. In this work, we propose a position-aware fine-grained EAT segmentation method that extends existing single-class coarse EAT segmentation to multi-class fine-grained segmentation of RV-, LV-, and PA-EAT. Our method utilizes a two-branch architecture, where one branch specializes in segmentation and the other focuses on precisely positioning centroids of various EATs, thereby enhancing model performance for EAT localization and boosting segmentation accuracy. By leveraging prior knowledge of spatial distributions of different tissues, our method demonstrates favorable performance on a challenging self-collected dataset and a public dataset. The proposed method has the potential to aid in the automatic fine-grained segmentation of EAT, enabling more detailed clinical diagnostic needs.

Keywords: Epicardial adipose tissue (EAT), segmentation, multi-class, position regularization.

1 Introduction

Cardiovascular diseases (CVDs) are a foremost global health concern, ranking the leading causes of mortality worldwide [1]. In China, CVD is estimated to affect approximately 330 million individuals, accounting for 46.74% of all deaths in rural areas and 44.26% in urban areas [2]. Such a factor not only imposes substantial economic burdens, but also places immense pressure on the nation's healthcare infrastructure. Epicardial adipose tissue (EAT) (depicted in Fig. 1(a)) is a unique fat depot situated between the myocardium and the visceral layer of the pericardium, which holds multiple implications in cardiology research and clinical practice [3]. Firstly, EAT is believed to play a significant role in the pathogenesis and progression of cardiovascular diseases such as coronary artery disease, atrial fibrillation, and heart failure [4]. Secondly, the volume, thickness and density of EAT have been correlated independently with several

✉: Corresponding author: Sheng Lian, shenglian@fzu.edu.cn

cardiovascular events [4-6]. Hence, quantitative analysis of EAT holds crucial clinical value. Notably, EAT is not randomly distributed around the cardiac region [4]. Studies including [7, 8] suggest that the thickness of EAT around the atrium is an independent correlate of atrial fibrillation burden, and increased left atrial EAT tissue may be associated with the onset and maintenance of atrial fibrillation [7]. Aliya et al. [8] further conducted a quantitative analysis of right ventricular EAT (**RV-EAT**), left ventricular EAT (**LV-EAT**), and peri-atrial EAT (**PA-EAT**) using CT imaging. They confirmed higher EAT content in both the RV- and LV-EAT among patients with arrhythmogenic right ventricular dysplasia (ARVD)/cardiomyopathy, indicating a correlation between increased EAT content and disease severity, particularly noticeable in LV-EAT. Therefore, quantitative analysis of EAT in different cardiac regions aids in assessing the risk of related cardiovascular diseases.

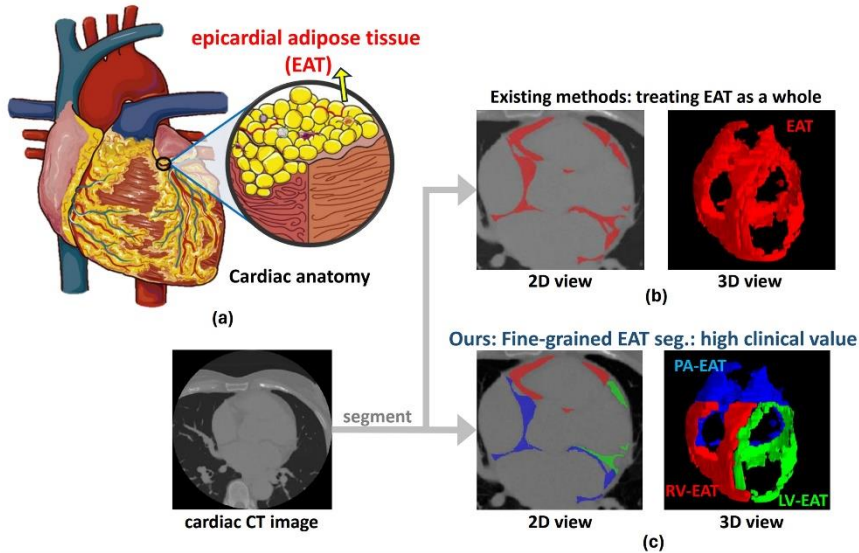


Fig. 1. (a) Anatomical diagram of Epicardial Adipose Tissue (EAT). (b) Existing EAT segmentation solution that treating EAT as a whole (in 2D and 3D view). (c) Our fine-grained EAT segmentation strategy that categories EAT into LV-, RV-, PA-EAT, which holds greater clinical significance. The original cardiac CT image for segmenting is located at the bottom left. (Best viewed in Color)

Automatic semantic segmentation is a crucial prerequisite for quantifying EAT. However, existing solutions treat EAT as a whole (Fig. 1(b)), failing to take into account the crucial clinical requirement for fine-grained segmentation and quantification of different EATs. According to its anatomical distribution and corresponding clinical significance, we categorize EAT into RV-, LV-, and PA-EAT, and attempt to perform fine-grained segmentation on EAT, as illustrated in Fig. 1(c). In contrast to other organ segmentation tasks, fine-grained segmentation of EAT poses greater challenges due to the following factors. (1) Class imbalance. Considering EAT is a very thin tissue, in cardiac CT images, the volume fraction of EAT is very low. For example, LV-EAT in

Fig. 1(c) accounts for only 0.29% of the overall image. Such extreme imbalance between foreground and background classes poses a greater challenge to segmentation models. (2) Complex tissue structure. As illustrated in Fig. 1(b, c), the distribution of EAT is irregular, with indistinct shape features and low differentiation from surrounding tissues, which reduces the accuracy and stability of segmentation.

For the task of EAT segmentation, there exist various methods including non-learning-based traditional methods, traditional machine learning-based methods, and deep learning-based methods. Non-learning-based traditional methods typically require hand-craft complex features and human intervention, while with suboptimal performance. For example, Coppini et al. [9] proposed a method that requires expert intervention to pinpoint the pericardium region of interest, then extract the approximate edge of EAT by combining thresholding and active contour model. Machine learning approaches, on the other hand, can extract the intrinsic features within images and boost their performance. For example, Rodrigues et al. [10] employed thresholding and map atlas initialization for image segmentation, then using features such as image centroid to train a random forest classifier for EAT segmentation. However, these methods suffer from limited model robustness and accuracy due to the need for handcrafted complex features. In contrast, deep learning-based methods have shown remarkable performance in EAT segmentation task, largely attributed to their powerful learning capacity. Commandeur et al. [11] proposed a multi-task two-stage segmentation model using CNNs to accomplish slice classification, pericardium segmentation, and EAT segmentation from chest CT images. They predicted EAT probability distribution in Cartesian and cylindrical coordinates, achieving automated EAT segmentation. Their multi-center validation achieved segmentation accuracy similar to human experts [12]. Other researchers attempted to modify U-Net model [13] to enhance its performance for EAT segmentation tasks [14-16]. However, current deep learning-based methods primarily treat EAT as a whole, failing to take into account the practical clinical necessities for fine-grained segmentation of EAT at different locations, e.g., RV-, LV-, PA-EAT.

In this paper, we integrate prior knowledge about the spatial distribution of different EATs into the segmentation model, proposing a centroid positioning branch. This branch can be easily embedded into existing segmentation frameworks, assisting the model in better localizing and perform fine-grained segmentation on different EATs. Evaluation on a challenging self-collected dataset and a public dataset demonstrated favorable performance. Our main contributions are three-fold: (1) We extend existing methods from single-class coarse EAT segmentation to multi-class fine-grained segmentation of RV-, LV-, and PA-EAT, catering to more detailed clinical diagnostic needs; (2) We propose a position-aware fine-grained EAT segmentation method that harnesses the prior knowledge of spatial distributions of different EAT as an additional regularization term. Such design enables the model to effectively learn and discern the relative positions of different EAT regions, thereby enhancing segmentation accuracy; (3) Our method, combined with the segmentation backbone network, achieved superior performance on a challenging self-collected EAT dataset and a public ACDC dataset.

2 Method

In this section, we first provide the method overview in Sec. 2.1, followed by the introduction of the proposed centroid positioning branch in Sec. 2.2, and finally present the training and inference processes in Sec. 2.3.

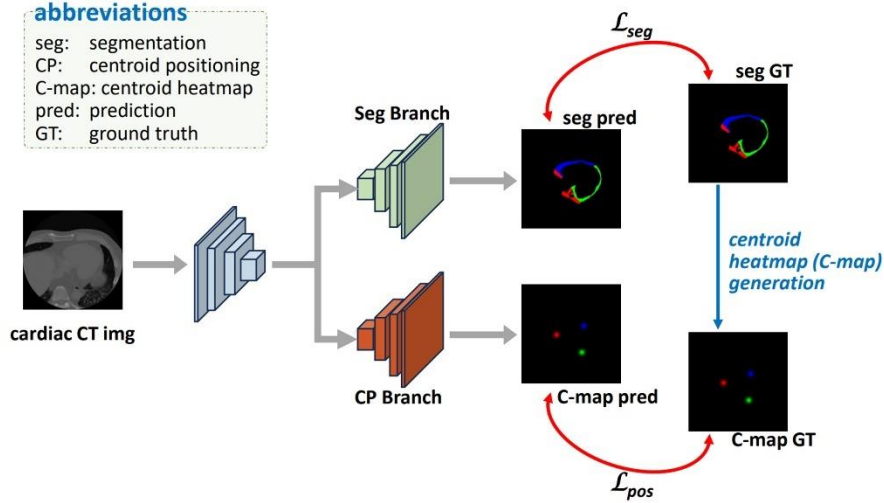


Fig. 2. The overall framework of our model, which adopts a two-branch architecture. The top branch is adopted for the segmentation task and the other branch is used for the centroid positioning task.

2.1 Method Overview

We illustrate the overall framework of our model in Fig. 2, which consists of two integral components: (1) The segmentation branch (on the upper half of Fig. 2) follows the conventional segmentation strategy and gives fine-grained segmentation of EAT into RV-, LV-, and PA-EAT. (2) The centroid positioning branch (on the lower half of Fig. 2) which predicts the centroid heatmap (C-map) of RV-, LV-, and PA-EAT. In this way, the centroid positioning branch serves as an additional regularization term, thereby enhancing the model’s capacity to localize different EAT regions accurately. We denote our dataset as $D = \{x_i, y_i\}_{i=1}^N$ where x_i and y_i denote the i -th image and corresponding segmentation mask. The centroid heatmap (C-map) for RV-, LV-, and PA-EAT is denoted as $H = \{\tilde{h}_i, h_i\}_{i=1}^N$, where \tilde{h}_i denotes the i -th heatmap prediction, and h_i is the corresponding ground truth. The segmentation branch adopts an encoder-decoder structure as the backbone (such as U-Net, ResU-Net, etc.), and outputs segmentation maps. The output feature of the encoder, denoted as f , is also fed into the decoder of the centroid positioning branch, ultimately producing C-map prediction \tilde{h}_i . Below, we elaborate on the details of the centroid positioning branch in Sec. 2.2.

2.2 Centroid Positioning Branch

To incorporate prior knowledge of the relative position among different EATs into the model, we introduce a centroid positioning branch. This branch shares the same encoder with the segmentation branch and concurrently predicts the centroids heatmap (C-map) for RV-, LV-, PA-EAT using a standalone decoder that has a similar structure to the segmentation branch. Specifically, the centroid coordinates for a certain class are derived from the segmentation branch by calculating the average of all pixel locations within the predicted class. After obtaining the centroid coordinates for RV-, LV-, and PA-EAT, we generate the Gaussian heatmap for each class, which goes as:

$$H(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\sigma^2}} \quad (1)$$

where (x_0, y_0) is the centroid coordinates for a certain class. We set $\sigma = 5$ to employ a Gaussian kernel with a standard deviation of 5 on the centroid coordinates. Then we formulate the C-map as the concatenation of all Gaussian heatmaps of different EATs, which is calculated as:

$$H_{C\text{-map}} = \text{concat}(H_{RV}, H_{LV}, H_{PA}) \quad (2)$$

where *concat* is the concatenate operation, while H_{RV}, H_{LV}, H_{PA} is the generated Gaussian heatmap for RV-, LV-, and PA-EAT's centroid, respectively.

2.3 Training and Inference.

Training procedure. For the loss function of the segmentation branch (\mathcal{L}_{seg}), we adopt DiceCE loss (\mathcal{L}_{DCE}), which goes as:

$$\mathcal{L}_{seg} = \mathcal{L}_{DCE} = \mathcal{L}_{Dice} + \mathcal{L}_{CE} \quad (3)$$

Specifically, $\mathcal{L}_{Dice} = -\sum_{i=1}^n \left(1 - 2 \cdot \frac{t_i p_i}{t_i + p_i}\right)$ and $\mathcal{L}_{CE} = -\sum_{i=1}^n t_i \log(p_i)$, where t_i is the ground truth label, and p_i is the predicted probability for the i -th class. For the loss function of the centroid positioning branch (\mathcal{L}_{pos}), we adopt BCE loss (\mathcal{L}_{BCE}) which serves to measure the difference between predicted heatmaps \tilde{h}_i and ground truth h_i . The formulation is given as follows:

$$\mathcal{L}_{pos} = \mathcal{L}_{BCE}(\tilde{h}_i, h_i) = -(\tilde{h}_i \log_{h_i} h_i + (1 - \tilde{h}_i) \log_{(1-h_i)} h_i) \quad (4)$$

During training, the total loss (\mathcal{L}_{ttl}) of our network is formulated as:

$$\mathcal{L}_{total} = \lambda_1 \mathcal{L}_{seg} + \lambda_2 \mathcal{L}_{pos} \quad (5)$$

We empirically set weights $\lambda_1 = \lambda_2 = 0.5$.

Inference procedure. Given a test CT image x_i , we directly obtain the fine-grained EAT segmentation prediction \hat{y}_i through the segmentation branch.

Table 1. Comparison of experimental results for various segmentation models with/without the proposed position-aware fine-grained EAT segmentation method on the EAT dataset.

Model	DSC(↑)	Jaccard(↑)	95HD(↓)	ASD(↓)	BIoU(↑)
U-Net[13]	62.23	25.49	20.72	3.11	12.49
+ Ours	63.34	26.02	15.74	3.50	13.06
SegNet [17]	67.67	29.76	13.66	1.72	14.13
+ Ours	68.23	31.11	12.22	1.72	14.70
ResUNet[18]	65.21	27.85	17.68	2.72	13.61
+ Ours	68.25	30.82	12.38	1.97	15.50
PSPNet [19]	70.01	34.01	9.55	1.86	15.53
+ Ours	71.63	34.78	9.10	1.66	16.57
Swin-UNet [25]	67.40	37.19	7.92	1.74	17.74
+Ours	69.25	37.56	7.92	1.60	17.94
MedT [24]	65.46	35.41	10.08	1.71	16.67
+Ours	66.72	34.28	10.45	1.70	16.38
Li et al. [20]	71.28	35.65	9.33	1.73	17.27
+ Ours	72.74	35.68	8.72	1.53	17.54
Commandeur et al. [12]	69.50	32.10	10.46	1.71	14.46
+Ours	70.88	32.49	10.08	1.64	14.82
Hoori et al. [21]	64.07	27.05	16.15	3.41	13.52
+ Ours	64.86	27.91	14.45	3.23	14.02
Zhao et al. [22]	63.26	25.35	16.78	3.16	12.74
+Ours	64.86	27.06	15.68	3.10	13.82

3 Experiments and Results

3.1 Materials and Evaluation Metrics

Dataset description. In this study, a self-collected **EAT dataset** and a public **ACDC dataset** [26] were adopted for model learning and evaluation. The EAT dataset contains 97 cardiac CT cases with RV-, LV-, and PA-EAT manually annotated and validated by three professionals who are either experienced physicians or medical Ph.D. students with over three years of tenure. All these cases were obtained from *The First Affiliated Hospital of Fujian Medical University*. The acquired data is fully anonymized and has been approved by the corresponding ethics committee (No: MRCTA, ECFAH of FMU [2021]072). All the CT images were scanned by a GE 256-row Revolution CT machine. The data is gathered and annotated in the axial view, with an average of 67 slices per case, totaling 6,575 slices. The spacing between slices is either 2.5mm or 3mm, and the resolution is 512*512 pixels. The ACDC dataset is a short-axis MR-cine dataset containing MR images of 100 patients. All the short-axis slices are with a slice thickness of 5 to 8 mm, and the resolution goes from 0.83 to 1.75 mm^2 /pixel. Each image is

labelled with the left ventricle(LV), myocardium(Myo) and right ventricle(RV) of that patient.

Table 2. Comparison of experimental results for various segmentation models with/without the proposed centroid positioning branch on the ACDC dataset.

Model	DSC(\uparrow)	Jaccard(\uparrow)	95HD(\downarrow)	ASD(\downarrow)	BIoU(\uparrow)
U-Net[13]	90.24	82.97	2.24	0.68	76.73
+ Ours	91.67	84.94	1.67	0.29	79.05
SegNet [17]	90.31	82.75	1.21	0.40	76.12
+ Ours	91.05	83.95	1.18	0.36	76.75
Res-UNet[18]	91.48	84.68	1.50	0.33	79.19
+ Ours	91.87	85.33	1.30	0.46	79.89
PSPNet [19]	88.43	79.81	1.30	0.38	73.06
+ Ours	88.54	79.85	1.97	0.51	73.45
Swin-UNet [25]	88.20	79.45	2.61	0.83	71.64
+Ours	88.93	80.50	1.48	0.46	72.09
MedT [24]	88.57	80.18	1.94	0.58	72.55
+Ours	89.61	80.50	1.30	0.43	73.23

Evaluation metrics. To demonstrate the effectiveness of our model, we conduct the five-fold cross-validation on EAT dataset, where 77 cases were utilized for training purposes, leaving the remaining 20 cases for testing. For the ACDC dataset, we use 70 cases for training, 10 cases for validation and 20 cases for testing. We adopt two types of evaluation metrics. One is from the perspective of *region overlap*, including Dice Similarity Coefficient (Dice) and Jaccard Index (Jaccard). The other is from the perspective of *boundary similarity*, including 95% Hausdorff Distance (95HD), Average Surface Distance (ASD), and Boundary Intersection-over-Union (BIoU).

3.2 Implementation Details

We implement our model using PyTorch [23] 1.13 on a machine with two NVIDIA RTX A4000, each with 16GB GPU memory. The images of both datasets were uniformly resized to $256 * 256$, and then normalized to zero mean and unit variance for better convergence. Random flipping with ratio of 0.5 is performed for data augmentation. We utilized the Adam optimizer with a learning rate of $1e - 4$ and a weight decay factor of $1e - 5$. For the EAT dataset, we set the training epoch and batch size to 40 and 8, respectively. For the ACDC dataset, training employed 40 epochs with a batch size of 12.

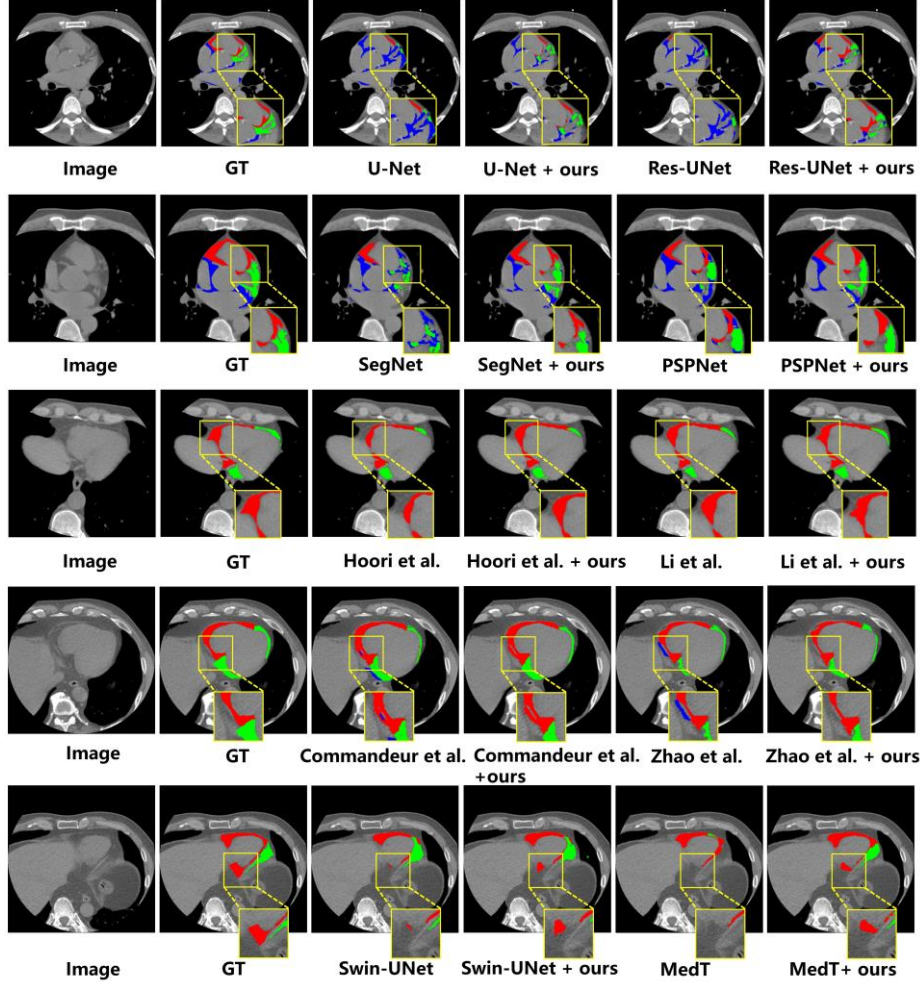


Fig. 3. Qualitative results of different backbone segmentation models with/without the proposed centroid positioning branch on the EAT dataset.

3.3 Comparative Experiments

We evaluate the proposed model on the EAT dataset, comparing it with both classical medical image segmentation models including: (1) U-Net [13], (2) SegNet [17], (3) Res-UNet [18], (4) PSPNet [19], (5) Swin-UNet [25], (6) MedT [24]; And the existing promising EAT segmentation models including: (5) Li et al. [20]: A new lightweight U-shaped segmentation model utilizing residual multi-scale dilated convolution blocks. (6) Hoori et al. [21]: A novel bisect technique, dividing the heart into upper and lower halves, adjusts the image input sequence to enhance the model’s comprehension. (7) Commandeur et al. [12]: A multi-task model that includes slice classification and

segmentation tasks, enhancing the performance of the model. (8) Zhao et al. [22]: A modified 2D Dense U-Net specifically designed for EAT segmentation.

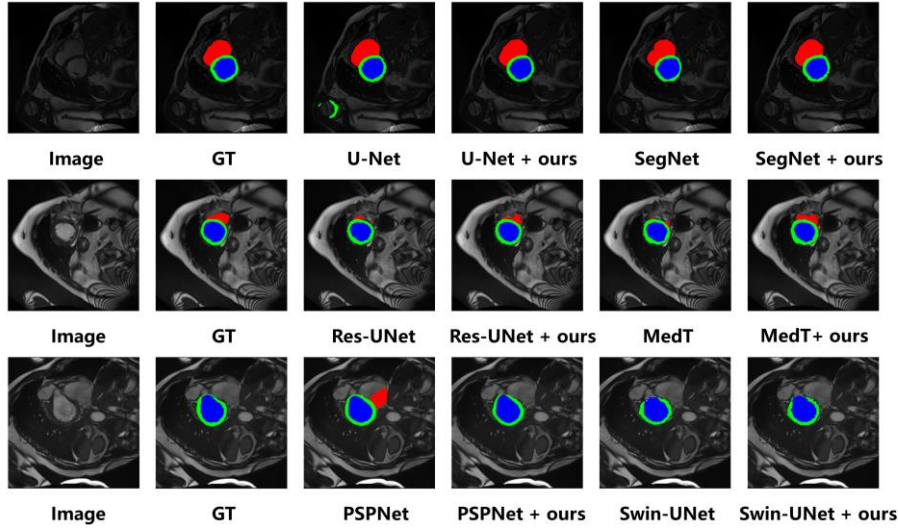


Fig. 4. Qualitative results of classic medical segmentation model with/without the proposed centroid positioning branch on the ACDC dataset.

Quantitative results. The quantitative comparing results are summarized in Table 1 and Tabel 2. From these tables, we have the following observations. (1) *Our model can help classical segmentation models achieve promising improvement across different metrics.* As illustrated by the second to the fifth rows of Table 1, after integrating our model, all comparing classical segmentation models achieve improvements across most metrics. For example, after integrating our model, the DSC (\uparrow), Jaccard (\uparrow), and BIoU (\uparrow) of Res-UNet have each improved by 3.04%, 2.97% and 1.89%, respectively. While the 95HD (\downarrow) and ASD (\downarrow) have decreased by 5.3 and 0.75 voxels, respectively. Table 2 demonstrates the effectiveness of our model on the ACDC dataset. Our model on the UNet improved the segmentation performance by 1.43%, 1.97% and 2.32% in terms of DSC (\uparrow), Jaccard (\uparrow) and BIoU (\uparrow); (2) *The existing EAT segmentation methods, after being integrated with our model, demonstrate consistent enhancements across different metrics.* Considering the existing EAT segmentation models, our model can also aid in segmenting different EATs. For example, compared with the original model proposed by Zhao et al. [22], after introducing our model, the DSC (\uparrow), Jaccard (\uparrow), and BIoU (\uparrow) have each improved by 1.6%, 1.71% and 1.08%, respectively. While the 95HD (\downarrow) have decreased by 1.1 voxel.

Qualitative results. Fig. 3 illustrates the qualitative results of different backbone segmentation models on the EAT dataset with/without the proposed centroid positioning branch. The results demonstrate that segmentation models with the centroid positioning branch yield enhanced segmentation results, especially in terms of improved differentiation between various adjacent types of EAT. As shown in the first two rows of Fig. 3, the segmentation results using only backbone segmentation show large category prediction errors, and the model distinguishes categories more accurately with the inclusion of the centroid positioning branch. Fig. 4 shows the segmentation results of our model on ACDC dataset. The segmentation results likewise confirm the effectiveness of our model. The myocardium appears in the lower left corner of the U-Net plot in the first row of Fig. 4 with a large positional offset. With the addition of centroid positioning branch network, the positional accuracy of all types of location is improved while the segmentation accuracy is enhanced.

4 Conclusion

In this paper, we propose a position-aware fine-grained EAT segmentation method that harnesses the prior knowledge of the spatial distribution of different EATs as an additional regularization. Our method extends existing solutions from single-class coarse EAT segmentation to multi-class fine-grained segmentation of RV-, LV-, PA-EAT, catering to more detailed clinical diagnostic needs. Our model can be easily integrated into existing segmentation backbones, and achieve favorable performance on a challenging self-collected EAT dataset and a public ACDC dataset. We believe that our work can contribute to the development of more accurate and efficient EAT segmentation methods, ultimately improving the diagnosis and treatment of cardiovascular diseases.

Acknowledgments. This work was supported by the Natural Science Foundation of Fujian Province of China, Youth Project (No.2023J05117); the University-Industry Research Joint Innovation Project of Science and Technology, Fujian Province (No. 2020Y4016); Education and Scientific Research Project for Middle-aged and Young Teachers in Fujian Province (No. JAT220017).

References

1. G. A. Roth, G. A. Mensah, C. O. Johnson, G. Addolorato, E. Ammirati, L. M. Baddour, N. C. Barengo, A. Z. Beaton, E. J. Benjamin, C. P. Benziger et al., “Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the gbd 2019 study,” *Journal of the American College of Cardiology*, vol. 76, no. 25, pp. 2982–3021, 2020.
2. H. Sheng-Shou et al., “Report on cardiovascular health and diseases in china 2021: an updated summary,” *Journal of Geriatric Cardiology*, vol. 20, pp. 1–32, 2023.
3. Iacobellis, M. C. Ribaldo, F. Assael, E. Vecci, C. Tiberti, A. Zappaterreno, U. Di Mario, and F. Leonetti, “Echocardiographic epicardial adipose tissue is related to anthropometric and clinical parameters of metabolic syndrome: a new indicator of cardiovascular risk,” *The Journal of Clinical Endocrinology & Metabolism*, vol. 88, no. 11, pp. 5163–5168, 2003

4. G. Iacobellis, "Epicardial adipose tissue in contemporary cardiology," *Nature reviews cardiology*, vol. 19, no. 9, pp. 593–606, 2022.
5. A. A. Mahabadi, M. H. Berg, N. Lehmann, H. Kalsch, M. Bauer, K. Kara, N. Dragano, S. Moebus, K.-H. Jockel, R. Erbel *et al.*, "Association of epicardial fat with cardiovascular risk factors and incident myocardial infarction in the general population: the heinz nixdorf recall study," *Journal of the American College of Cardiology*, vol. 61, no. 13, pp. 1388–1395, 2013.
6. R. M. Abazid, O. A. Smettei, M. O. Kattea, S. Sayed, H. Saqqah, A. M. Widyan, and M. P. Opolski, "Relation between epicardial fat and subclinical atherosclerosis in asymptomatic individuals," *Journal of thoracic imaging*, vol. 32, no. 6, pp. 378–382, 2017.
7. C. Beyer, L. Tokarska, M. Stuhlinger, G. Feuchtner, F. Hintringer, S. Honold, L. Fiedler, M.-S. Schonbauer, R. Schönbauer, and F. Plank, "Structural cardiac remodeling in atrial fibrillation," *Cardiovascular Imaging*, vol. 14, no. 11, pp. 2199–2208, 2021.
8. M. Aliyari Ghasabeh, A. S. Te Riele, C. A. James, H. V. Chen, C. Tichnell, B. Murray, J. Eng, B. G. Kral, H. Tandri, H. Calkins *et al.*, "Epicardial fat distribution assessed with cardiac ct in arrhythmogenic right ventricular dysplasia/cardiomyopathy," *Radiology*, vol. 289, no. 3, pp. 641–648, 2018.
9. G. Coppini, R. Favilla, P. Marraccini, D. Moroni, and G. Pieri, "Quantification of epicardial fat by cardiac ct imaging," *The open medical informatics journal*, vol. 4, p. 126, 2010.
10. E. O. Rodrigues, F. F. C. de Moraes, and A. Conci, "On the automated segmentation of epicardial and mediastinal cardiac adipose tissues using classification algorithms," *arXiv preprint arXiv:2208.14352*, 2022.
11. F. Commandeur, M. Goeller, J. Betancur, S. Cadet, M. Doris, X. Chen, D. S. Berman, P. J. Slomka, B. K. Tamarappoo, and D. Dey, "Deep learning for quantification of epicardial and thoracic adipose tissue from non-contrast ct," *IEEE transactions on medical imaging*, vol. 37, no. 8, pp. 1835–1846, 2018.
12. F. Commandeur, M. Goeller, A. Razipour, S. Cadet, M. M. Hell, J. Kwiecinski, X. Chen, H.-J. Chang, M. Marwan, S. Achenbach *et al.*, "Fully automated ct quantification of epicardial adipose tissue by deep learning: a multicenter study," *Radiology: Artificial Intelligence*, vol. 1, no. 6, p. e190045, 2019.
13. O. Ronneberger, P. Fischer, and T. 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*. Springer, 2015, pp. 234–241.
14. Q. Zhang, J. Zhou, B. Zhang, W. Jia, and E. Wu, "Automatic epicardial fat segmentation and quantification of ct scans using dual u-nets with a morphological processing layer," *IEEE Access*, vol. 8, pp. 128 032–128 041, 2020.
15. Z. Li, L. Zou, and R. Yang, "A neural network-based method for automatic pericardium segmentation," in *Proceedings of the 2nd International Conference on Computer Science and Software Engineering*, 2019, pp. 45–49.
16. D. Molnar, O. Enqvist, J. Ulen, M. Larsson, J. Brandberg, A. A. Johnsson, E. Bjornson, G. Bergström, and O. Hjelmgren, "Artificial intelligence based automatic quantification of epicardial adipose tissue suitable for large scale population studies," *Scientific reports*, vol. 11, no. 1, p. 23905, 2021.
17. V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

18. S. Lian, L. Li, G. Lian, X. Xiao, Z. Luo, and S. Li, "A global and local enhanced residual unet for accurate retinal vessel segmentation." *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 3, pp. 852–862, 2021.
19. H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2881–2890.
20. Y. Li, S. Song, Y. Sun, N. Bao, B. Yang, and L. Xu, "Segmentation and volume quantification of epicardial adipose tissue in computed tomography images," *Medical Physics*, vol. 49, no. 10, pp. 6477–6490, 2022.
21. A. Hoori, T. Hu, J. Lee, S. Al-Kindi, S. Rajagopalan, and D. L. Wilson, "Deep learning segmentation and quantification method for assessing epicardial adipose tissue in ct calcium score scans," *Scientific Reports*, vol. 12, no. 1, p. 2276, 2022.
22. Z. Zhao, Y. Iwamoto, Y. Tezuka, H. Okada, K. Maeda, A. Wada, A. Kashiwagi, and Y.-W. Chen, "Automatic segmentation of visible epicardium using deep learning in ct image," in *Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery: Volume 1*. Springer, 2020, pp. 577–584.
23. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga et al., "Pytorch: An imperative style, high-performance deep learning library," *Advances in neural information processing systems*, vol. 32, 2019.
24. Valanarasu J M J, Oza P, Hacıhaliloğlu I, et al. Medical transformer: Gated axial-attention for medical image segmentation[C]//Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24. Springer International Publishing, 2021: 36-46.
25. Cao H, Wang Y, Chen J, et al. Swin-unet: Unet-like pure transformer for medical image segmentation[C]//European conference on computer vision. Cham: Springer Nature Switzerland, 2022: 205-218.
26. Bernard O, Lalande A, Zotti C, et al. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved?[J]. *IEEE transactions on medical imaging*, 2018, 37(11): 2514-2525.