Semantic Difference Guidance for the Uncertain Boundary Segmentation of CT Left Atrial Appendage

Xin You[1], Ming Ding[2], Minghui Zhang[1], Yangqian Wu[1], Yi Yu[2(✉)], Yun Gu[1], and Jie Yang[1(✉)]

[1] Institute of Medical Robotics, Shanghai Jiao Tong University, Shanghai, China
{sjtu_youxin,jieyang}@sjtu.edu.cn
[2] Shanghai Xinhua Hospital, Shanghai Jiaotong University, Shanghai, China

Abstract. Atrial fibrillation (AF) is one of the most common types of cardiac arrhythmia, which is closely relevant to anatomical structures including the left atrium (LA) and the left atrial appendage (LAA). Thus, a thorough understanding of the LA and LAA is essential for the AF treatment. In this paper, we have modeled relative relations between the LA and LAA via deep segmentation networks for the first time, and introduce a new LA & LAA CT dataset. To deal with uncertain boundaries between the LA and LAA, we propose the semantic difference module (SDM) based on diffusion theory to refine features with enhanced boundary information. Besides, disconnections between the LA and LAA are frequently observed in the segmentation results due to uncertain boundaries of the LAA region and CT imaging noise. To address this issue, we devise another connectivity-refined network with the connectivity loss. The loss function exerts a distance regularization on coarse predictions from the first-stage network. Experiments demonstrate that our proposed model can achieve state-of-the-art segmentation performance compared with classic convolutional-neural-networks (CNNs) and recent Transformer-based models on this new dataset. Specifically, SDM can also outperform existing methods on refining uncertain boundaries. Codes are available at https://github.com/AlexYouXin/LA-LAA-segmentation.

Keywords: Left atrial appendage · Difference operator · Uncertain boundary · Image segmentation

1 Introduction

Atrial fibrillation (AF) has been one of the most common types of cardiovascular diseases and is closely related to the left atrium (LA) [18]. Beside this chamber structure, there is a finer anatomy termed with the left atrial appendage

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43990-2_12.

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14226, pp. 121–131, 2023.
https://doi.org/10.1007/978-3-031-43990-2_12

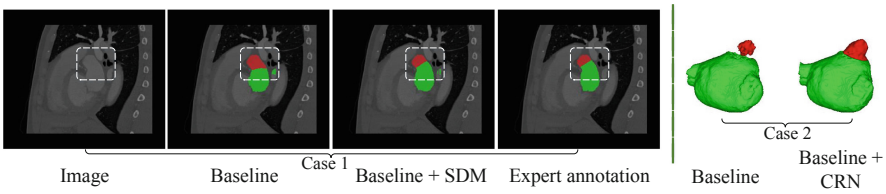

Fig. 1. Improved segmentation results of the baseline model after adding SDM on case 1 and after adding CRN on case 2. (Green Area: LA. Red Area: LAA). (Color figure online)

(LAA). The majority of strokes due to AF result from clots existing in the LAA [13]. A common measure for treatment is anticoagulant therapy. However, many patients have contraindications to this type of therapy. A more effective and feasible stroke prevention procedure is the left atrial appendage closure, which can avoid most of the drawbacks by anticoagulant therapy [12]. And the size of occlusion devices designed for patients is strongly associated with the anatomical interface between the LA and LAA [13]. Thus, enhancing the understanding for the structure of the LA and LAA is beneficial to carry out treatments for strokes due to AF. A normal pre-surgery imaging is Cardiac Computed Tomography (Cardiac CT), which is a popular physical inspection for diagnoses [18]. Thus, automatic and accurate segmentation of the LA and LAA from Cardiac CT images is essential to provide support for the diagnosis and treatment of various cardiovascular diseases.

Till now, many researches have focused on the automatic segmentation of the LA [24]. Compared with that, the LAA has not been sufficiently researched, particularly the relative relations between the LA and LAA [27]. In our work, we aim to design an automatic method for the correlation modeling between the LA and LAA, then give a quantitative and qualitative evaluation of segmentation performance. The LA and LAA both have large anatomical variations [9,27]. Besides, there are uncertain boundaries for the structure of the LAA, especially the interface between the LA and LAA. In contrast, cardiac tissues like the right atrium (RA), right ventricle (RV) and left ventricle (LV) have explicit boundaries, which can be easily and finely segmented [28]. As shown in Fig. 1, some cases show poor segmentation results on the uncertain boundary between the LA and LAA. There are many related works on the refinement for boundary segmentation, which can be grouped into three categories. The first strategy attempts to exert a strong loss constraint on boundaries via the multi-task learning paradigm [4]. Then some researchers apply a complex post-process to the segmentation of coarse boundaries, such as [26]. All the methods above mainly emphasize on refining predicted masks for high-quality images with clear boundaries, are not applicable for ambiguous or unclear boundaries [23] in the LA and LAA segmentation. Instead, the third strategy truly works, with the mechanism of enhancing deep features representing uncertain boundaries. Lee et al. [11] proposed a novel boundary-preserving block (BPB) with the ground-truth

structure information indicated by experts. Xie et al. [23] used the confidence map to evaluate the uncertainty of each pixel to enhance the segmentation of ambiguous boundaries. Furthermore, some loss functions [10,25] are specifically designed for the segmentation of uncertain boundaries.

Diffusion is a physical model aimed at minimizing the spatial concentration difference [15] and is widely used in computer vision [21,22]. In our work, we detailedly explain the process of refining uncertain boundaries based on diffusion theory. Then we propose a semantic guidance module based on differential operators to refine features from ambiguous boundaries, which is called semantic difference module (SDM). Here we introduce semantic information from deeper layers to guide the diffusion process. As a result of fuzzy boundaries of the LAA region and CT imaging noise, there exists a disconnection between the LA and LAA as shown in Fig. 1. Thus, we design another connectivity-refined network (CRN) combined with the connectivity loss, to deal with the connectivity of two regions. The contributions of our work are listed as follows:

(1) We introduce a new LA & LAA CT dataset. And as far as we are concerned, this is the first work based on deep neural networks, to model relative relations between the LA and LAA.
(2) We propose a novel semantic difference module based on diffusion theory to deal with the segmentation of uncertain boundaries.
(3) We apply a connectivity-refined network with the connectivity loss to refine coarse masks, then achieve the connectivity between the LA and LAA.
(4) Our proposed network achieves state-of-the-art segmentation performance on the LA and LAA. Specifically, SDM outperforms other methods related to refining the segmentation of uncertain boundaries.

2 Methodology

2.1 Preliminaries

Diffusion is a physical phenomenon, in which molecules spread from regions with higher concentrations toward regions with lower concentrations [15,16]. Then the whole system tends to be balanced. For a feature vector F to be smoothed, the diffusion process can be modeled as the following partial differential equation:

$$\frac{\partial F}{\partial t} = D \cdot \nabla^2 F \tag{1}$$

where D is the diffusivity function determining the diffusion speed along each direction, ∇ is the gradient operator. In our application, the stable state of $F(t)$ will show a more accurate localization for uncertain boundaries of the LAA.

Linear isotropic diffusion (D is equal to a constant) cannot be applied to complex scenes, because the diffusion velocity is the same in all directions. For a spatial-dependent function $D = D(x, y, z)$, the process is linear anisotropic. However, if we aim to extract refined boundary features, adopting linear diffusion processes will smooth both backgrounds and the edges. A more feasible

solution is to devise complex diffusion functions $D = D(F)$ with nonlinear characteristics [21,22]. As a result, the diffusion process exerts more smoothing to regions parallel to boundaries compared to regions vertical to these edges.

Detailedly, given a feature F where regions of uncertain boundaries are not highlighted, it is updated by the diffusion process in infinite time. The diffusion adjacent to ambiguous boundaries should be restrained, while the diffusion far away from boundaries is promoted. And the final state of the diffused feature will accurately localize uncertain boundaries of the LAA.

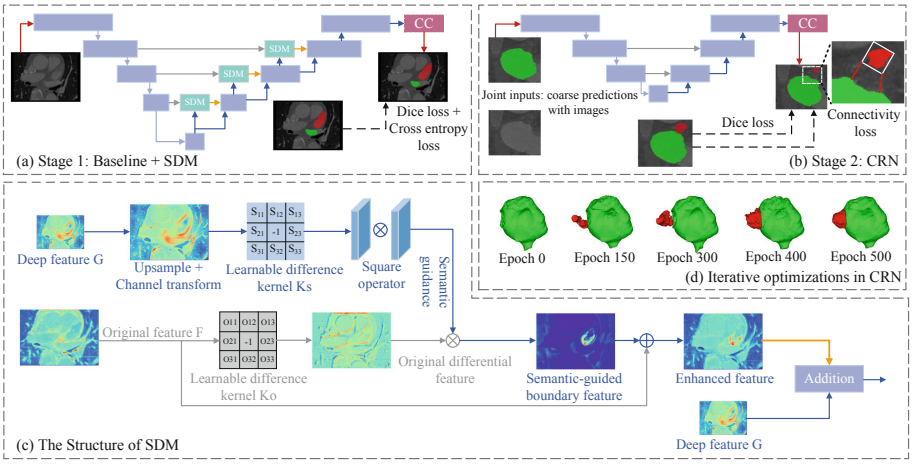

Fig. 2. Our proposed two-stage network. (a) Baseline model with the semantic difference module (CC: channel calibration). (b) Connectivity-refined network. (c) Boundary feature enhancement. (d) Iterative optimizations for the predicted mask of a specific testing case.

2.2 Semantic Difference Module

To localize fuzzy boundaries of the LAA, especially the interface between the LA and LAA, we propose the semantic difference module (SDM) to refine boundary features from these regions. Motivated by the diffusion process, we formulate the process of enhancing boundary features as solving a second-order partial differential equation. Due to the fact that semantic information is required to guide the localization of uncertain boundaries, we introduce the deep feature G from the precedent decoder layer to the diffusion process in each SDM.

Here we adopt ∇G as the semantic guidance map. And the square term $h(|\nabla G|^2)$ is deployed as function D to model nonlinear characteristics of the diffusion process, where h is a projection function. In terms of [15], Eq. 1 can be approximately solved via iterative updates as depicted by the following equations:

$$\hat{F}_{\boldsymbol{p}}^{t+1} = \sum_{\widetilde{\boldsymbol{p}} \in \delta_p} h(|G_{\widetilde{\boldsymbol{p}}} - G_{\boldsymbol{p}}|^2) \cdot (F_{\widetilde{\boldsymbol{p}}}^t - F_{\boldsymbol{p}}^t) \tag{2}$$

$$F_{\boldsymbol{p}}^{t+1} = \lambda \cdot F_{\boldsymbol{p}}^t + \nu \cdot \hat{F}_{\boldsymbol{p}}^{t+1} \tag{3}$$

where $\boldsymbol{p}$ is the index of feature maps, $\delta_{\boldsymbol{p}}$ is the local neighborhood centered at $\boldsymbol{p}$, λ and ν are weighting coefficients. Indeed, $F^t_{\widetilde{\boldsymbol{p}}} - F^t_{\boldsymbol{p}}$ is the differential information of original feature F^t at point $\boldsymbol{p}$, representing abundant boundary information, which contains complicated boundary features of anatomies as shown in Fig. 2.c, including the RA RV, etc. However, predicted boundaries between the LA and LAA are not accurate enough only with the diffusion process. Thus, the semantic difference guidance $|G_{\widetilde{\boldsymbol{p}}} - G_{\boldsymbol{p}}|^2$ is introduced to generate refined boundary feature $\hat{F}^{t+1}$. $\hat{F}^{t+1}$ will diffuse into the stable state as t increases, which can highlight boundaries between the LA and LAA, and suppress the activation on other boundary regions. And refined feature F^{t+1} is attained by fusing the original feature F^t with enhanced boundary feature $\hat{F}^{t+1}$.

We design the semantic difference module based on Eq. 3 as illustrated by Fig. 2. Here we make an improvement on the calculation of differential maps. Motivated by the fact that there exists an anisotropic distribution for our LA and LAA dataset in x, y and z dimensions, traditional edge operator cannot finely extract the differential map of feature F. Therefore, we propose a learnable boundary operator, which bears different values in each position of the kernel. As shown in Fig. 2, we fix the center value as -1 to maintain the difference attribute of the edge kernel. The revised description of enhanced boundary feature is calculated by Eq. 4.

$$\hat{F}^{t+1}_{\boldsymbol{p}} = \sum_{\widetilde{\boldsymbol{p}} \in \delta_p} \omega_{\widetilde{\boldsymbol{p}}} \cdot |\alpha_{\widetilde{\boldsymbol{p}}} G_{\widetilde{\boldsymbol{p}}} - G_{\boldsymbol{p}}|^2 \cdot (\beta_{\widetilde{\boldsymbol{p}}} F^t_{\widetilde{\boldsymbol{p}}} - F^t_{\boldsymbol{p}}) \tag{4}$$

where $\alpha_{\widetilde{\boldsymbol{p}}}$ and $\beta_{\widetilde{\boldsymbol{p}}}$ refer to the learnable edge operator for feature F and semantic feature G respectively. And $\omega_{\widetilde{\boldsymbol{p}}}$ means a vanilla $3 \times 3 \times 3$ convolution kernel.

2.3 Connectivity-Refined Network

Due to the CT imaging noise, some cases show the phenomenon that the LAA is separated from the LA. To deal with disconnections between the LA and LAA, we propose the second-stage network called connectivity-refined network (CRN). Inspired by the metric of 95% Hausdorff distance (HD_{95}) [7], we figure out that a poor connectivity between the LA and LAA will bring a large HD_{95} value for the LAA segmentation, which is not what we expected. Therefore, we adopt another distance constraint loss called connectivity loss $\mathcal{L}_c$, besides the per-pixel Dice loss $\mathcal{L}_d$ in the training process of CRN. Specifically, we choose coarse predictions from validation datasets from the first stage, concatenated with original images as the input of CRN. We firstly localize the predicted LAA region via its unique label. Then to make the region of LAA connected with the LA, we can only focus on the voxels most adjacent to the boundary interface. For the training efficiency, we locate the predicted LAA region with the approximately minimum external cube C (Please refer to **supplementary material** for more details about the algorithm). Four vertexes V_i (i=1, 2, 3, 4) of C neighbored with the LA are selected to calculate the connectivity loss, which is indeed an improved minimal

distance from vertexes to the surface of LA. Here we note the point set in the LA surface as P, and each point from P is noted as P_j.

$$\mathcal{L}_c = \sigma\Big\{\frac{\sum_{i=1}^{4} \min_{P_j \in P} D(V_i, P_j)}{S}\Big\} - 0.5 \tag{5}$$

$$\mathcal{L} = \mathcal{L}_d + \lambda \times \mathcal{L}_c \tag{6}$$

where D means the Euclidean distance, σ is the sigmoid function. S is a scaling coefficient, and we set it as 20 according to the ablation study on this hyper-parameter (Please refer to **supplementary material** for more quantitative results). Besides, λ is set as 1 if the training epoch reaches more 300 epochs, or it is 0. When there is no LAA predictions in a cropped patch, $\mathcal{L}_c$ is equal to 0.

Table 1. (a) Results compared with other CNNs and Transformer-based models (**Values on both sides of '|' represent evaluation metrics for each stage of our model**). (b) Segmentation performance including our SDM and other methods related to deal with uncertain boundaries (LA: Left Atrium, LAA: Left Atrial Appendage. Bold: the best, Underlined numbers: the second best. **'–' indicates that loss functions do not change Parameters and FLOPs of the baseline model**).

(a) Segmentation benchmark on LA & LAA								
Model	Params(M)	FLOPs(T)	Dice score (%) ↑			HD_{95} (mm) ↓		
			LAA	LA	Average	LAA	LA	Average
3D UNet [2]	16.47	0.514	76.64	95.75	86.19±13.0	10.25	6.94	8.59±5.0
ResUNet [3]	32.46	0.813	77.45	95.86	86.66±13.1	9.04	9.79	9.41±6.9
V-Net [14]	45.72	2.391	74.21	92.68	83.44±14.7	12.54	8.82	10.68±7.9
TransBTS [20]	35.61	0.613	73.65	95.05	84.35±12.5	10.85	15.53	13.19±11.1
UNETR [5]	93.54	0.458	75.97	95.26	85.62±12.9	11.41	15.32	13.37±10.1
TransUNet (3D) [1]	82.41	0.220	78.43	95.45	86.94±11.3	8.95	8.47	8.71±5.3
Swin UNETR [17]	62.19	0.975	76.32	95.94	86.13±12.9	11.45	7.71	9.58±5.7
UNeXt [19]	4.02	0.035	79.37	95.34	87.36±9.7	**8.38**	8.22	8.31±2.3
nnUNet [8]	30.79	0.835	81.16	**96.29**	88.72±9.6	8.67	7.17	7.92 ±2.8
Ours	23.47|1.92	0.666|0.130	**81.85**	95.96	**88.91** ±6.9	8.87	**6.63**	**7.75** ±2.4
(b) Comparison with other methods focused on uncertain boundaries								
Baseline	16.47	0.514	76.64	95.75	86.19±13.0	10.25	6.94	8.59±5.0
+ Boundary loss [10]	–	–	77.02	95.74	86.38 ±11.3	9.82	7.33	8.58 ±3.2
+ BU loss [25]	–	–	77.01	**96.16**	86.59 ±12.7	10.51	**6.78**	8.65 ±5.0
+ BPB [11]	27.97	0.717	77.49	96.02	86.76 ±12.8	9.96	6.97	8.46 ±4.7
+ CCM [23]	18.98	0.596	77.71	96.05	86.88 ±11.8	9.86	7.04	8.45 ±3.4
+ SDM (Ours)	23.47	0.666	**78.88**	95.91	**87.40** ±8.6	**9.76**	6.93	**8.35** ±3.5

On the ground that cases with a poor connectivity need to be refined in the training process of CRN, we increase sampling ratios of the whole LAA region from coarse masks of validation datasets. By cropping patches containing the boundary interface as network inputs, CRN will better learn the connectivity prior from the mapping between coarse predictions and expert annotations.

In the inference stage, final decoded features F_d are applied with the softmax operator to attain predicted masks. However, different channels of F_d bear different maximum values, which will affect segmentation performance. Thus, we propose the concept of channel calibration (CC). Before per-pixel selecting maximum values between channels, we uniform the maximum value of F_d channel by channel.

3 Experiment

3.1 Experimental Settings

Dataset. To evaluate the performance of our network, we conduct experiments on a new dataset, containing accurate annotations of LA and LAA provided by multiple experts. In detail, we collect 80 CT scans from 80 patients, which are split into 45/15/20 for training, validation and testing cases in Stage 1. In Stage 2, 50 predictions of the validation dataset in Stage 1 are generated by various models, in which there are 30 predicted masks with disconnections. Then we randomly split them as 35/15 for training and validation. Moreover, we choose the Dice score and HD_{95} as quantitative metrics.

Table 2. (a) Ablation study on the structure of the semantic difference module (SDM). DK refers to difference kernel) (b) Ablation study on the efficacy of our proposed key components, including SDM, connectivity-refined network (CRN), channel calibration (CC) and connectivity loss.

(a) Ablation study on SDM						
Model	Dice score (%) ↑			HD_{95} (mm) ↓		
	LAA	LA	Average	LAA	LA	Average
+ SDM	**78.88**	**95.91**	**87.40**	9.76	**6.93**	**8.35**
w/o learnable DK	76.60 (↓ 2.28)	95.77 (↓ 0.14)	86.18 (↓ 1.22)	10.53 (↑ 0.77)	7.30 (↑ 0.37)	8.92 (↑ 0.57)
w/o original feature F	77.15 (↓ 1.73)	95.71 (↓ 0.20)	86.43 (↓ 0.97)	**9.71** (↓ 0.05)	7.16 (↑ 0.23)	8.44 (↑ 0.09)
w/o semantic guidance	77.82 (↓ 1.06)	95.49 (↓ 0.42)	86.66 (↓ 0.74)	9.82 (↑ 0.06)	7.38 (↑ 0.45)	8.60 (↑ 0.25)
(b) Ablation study on key components						
Baseline	76.64	95.75	86.19	10.25	6.94	8.59
+ SDM	78.88	95.91	87.40	9.76	6.93	8.35
+ CC	77.44	**96.06**	86.75	9.70	6.84	8.27
+ SDM + CC	79.72	96.03	87.88	9.51	6.90	8.21
+ CRN	79.51	95.62	87.57	9.42	6.89	8.16
+ CRN w/o connectivity loss	78.24	95.65	86.95	9.93	6.95	8.44
+ SDM + CRN	81.32	95.64	88.48	9.08	6.78	7.93
+ SDM + CRN + CC	**81.85** (↑ 5.21)	95.96 (↑ 0.21)	**88.91** (↑ 2.72)	**8.87** (↓ 1.38)	**6.63** (↓ 0.31)	**7.75** (↓ 0.84)

Implementation Details. In the first stage, we choose the vanilla 3D UNet as our baseline model, trained for 2000 epochs. And we utilize a combination of cross entropy loss and Dice loss followed by [8]. For the second stage, CRN is trained for 500 epochs, which is a smaller 3D UNet with only $1.92M$ parameters.

And we utilize Dice loss and connectivity loss as illustrated by Eq. 6. We train all models using AdamW optimizer. With the linear warm-up strategy, the initial learning rate is set as $5e{-}4$ with a cosine learning rate decay scheduler, and weight decay is set as $1e{-}5$. The size of cropped patches is $160 \times 160 \times 192$. All models are implemented based on Pytorch and trained on 2 NVIDIA Tesla V100 GPUs.

3.2 Experimental Results

Table 1 consists of two sub-figures (a) and (b). According to Table 1.a, our proposed network outperforms classic CNNs and recent Transformer-based models. Specifically, our model shows superior to powerful nnUNet [8] on four metrics, and there exist 0.69% increase on the Dice score of LAA, 0.54 mm decrease on the HD_{95} of LA. Compared with nnUNet, our two-stage model requires less computational cost and is free from complicated multi-model ensembles. However, the Dice score of our model on LA is inferior to nnUNet. We argue that our model is aimed at improving the segmentation mask of LAA, which bears a different structure from LA. And the iterative optimization process of CRN in Fig. 2 shows the refinement for LAA segmentation. More results about the generalization of CRN can be found in **supplementary material**. Besides, Swin UNETR [17] takes the lead in Transformer-based models, which is not as good as our model because Transformer-based models are data-hungry [6]. Another phenomenon worth to mention is that our model shows an ordinary performance on the HD_{95} of LAA, which is owing to the appearance of outliers. And not only our model but other CNNs and Transformer-based models suffer from the existence of outliers in predictions (More visualization results can be found in **supplementary material**). We will address this issue in the future research. In Table 1.b, we choose 3D UNet as the baseline model. SDM shows better segmentation performance compared with other methods on improving the segmentation of uncertain boundaries. And Fig. 2 illustrated that SDM can enhance uncertain boundary features, especially the boundary interface between LA and LAA.

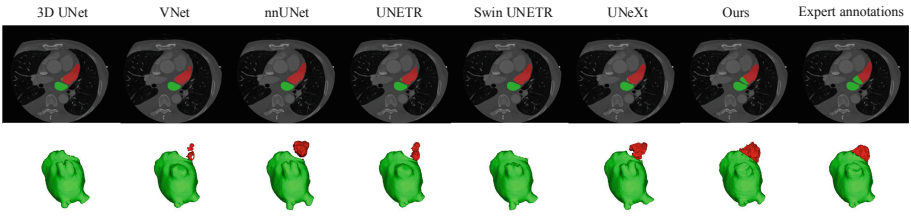

Fig. 3. Qualitative segmentation results of other CNNs and Transformer-based models. (Green Area: LA. Red Area: LAA). (Color figure online)

3.3 Ablation Studies

Table 2 shows ablation studies on key components and on the detailed structure of SDM. From Table 2.a, SDM, CRN and CC can all boost segmentation performance of the baseline model. Besides, we visualize qualitative results in Fig. 3. For the first row, our model give a more accurate localization for the boundary interface, which proves the effectiveness of SDM. The second row is a strong proof that CRN can effectively improve the connectivity between LA and LAA. Then we probe into the efficacy of connectivity loss $\mathcal{L}_c$ by removing it from CRN, which results in a $0.51mm$ increase on the HD_{95} of LAA, which is reasonable because $\mathcal{L}_c$ is indeed a distance regularization on LAA boundaries.

In Table 2.b, we investigate the significance of each component in SDM. (1) Without learnable difference kernels for differential maps of feature F and semantic feature G, the Dice score for LAA declines sharply, which reveals we need to focus on the anisotropic distribution of this dataset. (2) In SDM, we adopt a residual block by fusing the original feature F and the boundary feature. With F removed, there is some details and texture information missing, resulting in a performance drop. (3) Finally, semantic guidance from deeper features is deployed to guide the extraction for uncertain boundaries between LA and LAA. Besides, some false boundaries are restrained, which is illustrated in Fig. 2.

4 Conclusion

In this paper, we introduce a new CT dataset on LA and LAA, then carry out the segmentation task. Detailedly, we explain the refined process for the segmentation of uncertain boundaries via diffusion theory. Based on this, we apply SDM to successfully improve the segmentation for uncertain boundaries between LA and LAA. Then CRN with the connectivity loss can deal with the poor connectivity between two structures. Detailed quantitative and qualitative results have demonstrated the efficacy of two proposed elements.

Acknowledgement. This work is supported in part by National Key R&D Program of China (2019YFB1311503), the Shanghai Sailing Program (20YF1420800), the Shanghai Health and Family Planning Commission (202240110) and Xinhua Hospital affiliated with the School of Medicine (XHKC2021-07).

References

1. Chen, J., et al.: Transunet: transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306 (2021)
2. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-net: learning dense volumetric segmentation from sparse annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 424–432. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_49

3. Diakogiannis, F.I., Waldner, F., Caccetta, P., Wu, C.: Resunet-A: a deep learning framework for semantic segmentation of remotely sensed data. ISPRS J. Photogrammetry Remote Sens. **162**, 94–114 (2020)
4. Hatamizadeh, A., Terzopoulos, D., Myronenko, A.: End-to-end boundary aware networks for medical image segmentation. In: Suk, H.-I., Liu, M., Yan, P., Lian, C. (eds.) MLMI 2019. LNCS, vol. 11861, pp. 187–194. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32692-0_22
5. Hatamizadeh, A., et al.: UnetR: transformers for 3D medical image segmentation. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 574–584 (2022)
6. He, K., et al.: Transformers in medical image analysis: a review. Intell. Med. (2022)
7. Huttenlocher, D.P., et al.: Comparing images using the hausdorff distance. TPAMI **15**(9), 850–863 (1993)
8. Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.H.: NNU-net: a self-configuring method for deep learning-based biomedical image segmentation. Nat. Methods **18**(2), 203–211 (2021)
9. Jin, C., et al.: Left atrial appendage segmentation using fully convolutional neural networks and modified three-dimensional conditional random fields. IEEE J. Biomed. Health Inform. **22**(6), 1906–1916 (2018)
10. Kervadec, H., et al.: Boundary loss for highly unbalanced segmentation. Med. Image Anal. **67**, 101851 (2021)
11. Lee, H.J., Kim, J.U., Lee, S., Kim, H.G., Ro, Y.M.: Structure boundary preserving segmentation for medical image with ambiguous boundary. In: CVPR, pp. 4817–4826 (2020)
12. Leventić, H., Benčević, M., Babin, D., Habijan, M., Galić, I.: A survey of left atrial appendage segmentation and analysis in 3D and 4d medical images. arXiv preprint arXiv:2205.06486 (2022)
13. Leventić, H., et al.: Left atrial appendage segmentation from 3D CCTA images for occluder placement procedure. Comput. Biol. Med. **104**, 163–174 (2019)
14. Milletari, F., Navab, N., Ahmadi, S.-A.: V-net: fully convolutional neural networks for volumetric medical image segmentation. In: 2016 Fourth International Conference on 3D Vision (3DV), pp. 565–571. IEEE (2016)
15. Sapiro, G.: Geometric Partial Differential Equations and Image Analysis. Cambridge University Press, Cambridge (2006)
16. Tan, H., Sitong, W., Pi, J.: Semantic diffusion network for semantic segmentation. NeurIPS **35**, 8702–8716 (2022)
17. Tang, Y., et al.: Self-supervised pre-training of swin transformers for 3D medical image analysis. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 20730–20740 (2022)
18. Tobon-Gomez, C., et al.: Benchmark for algorithms segmenting the left atrium from 3D CT and MRI datasets. IEEE Trans. Med. Imaging **34**(7), 1460–1473 (2015)
19. Valanarasu, J.M.J., Patel, V.M.: UNeXt: MLP-based rapid medical image segmentation network. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) MICCAI 2022. LNCS, vol. 13435, pp. 23–33. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-16443-9_3
20. Wang, W., Chen, C., Ding, M., Yu, H., Zha, S., Li, J.: TransBTS: multimodal brain tumor segmentation using transformer. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12901, pp. 109–119. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87193-2_11

21. Weickert, J.: Coherence-enhancing diffusion filtering. Int. J. Comput. Vision **31**(2–3), 111 (1999)
22. Weickert, J., Ter Haar Romeny, B.M., Viergever, M.A.: Efficient and reliable schemes for nonlinear diffusion filtering. IEEE Trans. Image Process. **7**(3), 398–410 (1998)
23. Xie, Y., Liao, H., Zhang, D., Chen, F.: Uncertainty-aware cascade network for ultrasound image segmentation with ambiguous boundary. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) MICCAI 2022. LNCS, vol. 13434, pp. 268–278. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-16440-8_26
24. Xiong, Z., et al.: A global benchmark of algorithms for segmenting the left atrium from late gadolinium-enhanced cardiac magnetic resonance imaging. Med. Image Anal. **67**, 101832 (2021)
25. Yeung, M., Yang, G., Sala, E., Schönlieb, C.-B., Rundo, L.: Incorporating boundary uncertainty into loss functions for biomedical image segmentation. arXiv preprint arXiv:2111.00533 (2021)
26. Yuan, Y., Xie, J., Chen, X., Wang, J.: SegFix: model-agnostic boundary refinement for segmentation. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, J.-M. (eds.) ECCV 2020. LNCS, vol. 12357, pp. 489–506. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58610-2_29
27. Zheng, Y., Yang, D., John, M., Comaniciu, D.: Multi-part modeling and segmentation of left atrium in c-arm CT for image-guided ablation of atrial fibrillation. IEEE Trans. Med. Imaging **33**(2), 318–331 (2013)
28. Zhuang, X., Shen, J.: Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI. Med. Image Anal. **31**, 77–87 (2016)