GAPNET: GRANULARITY ATTENTION NETWORK WITH ANATOMY-PRIOR-CONSTRAINT FOR CAROTID ARTERY SEGMENTATION

LIN ZHANG*, CHENGGANG LU†, XIN-YANG SHI‡, CAIFENG SHAN §, JIONG ZHANG¶, DA CHEN‖, AND LAURENT D. COHEN**

Abstract. Atherosclerosis is a chronic, progressive disease that primarily affects the arterial walls. It is one of the major causes of cardiovascular disease. Magnetic Resonance (MR) black-blood vessel wall imaging (BB-VWI) offers crucial insights into vascular disease diagnosis by clearly visualizing vascular structures. However, the complex anatomy of the neck poses challenges in distinguishing the carotid artery (CA) from surrounding structures, especially with changes like atherosclerosis. In order to address these issues, we propose GAPNet, which is a consisting of a novel geometric prior deduced from an anatomical viewpoint. The use of anatomical prior allows the model to avoid segmentation contours whose topology violates the anatomical reality. Specifically, at the first stage, regional features are learned to identify the location of the target CA and to reduce the influence from the surrounding similar tissues. The second stage aims to improve the feature representation capability, by employing a delicately designed Feature Refinement Attention (FRA) module to capture boundary and detailed information alongside a new Multi-Scale Information Enhancement (MIE) module at the end of the decoder procedure. Experimental results demonstrate the superior performance of our approach on two carotid artery datasets, respectively achieving Dice scores of 0.76 and 0.83, proving the effectiveness of GAPNet in improving the accuracy of carotid artery segmentation in MR imaging.

Key words. Carotid artery segmentation, anatomical prior, topological prior, deep learning, isoperimetric theorem.

1. Introduction. Cardiovascular disease gets to be one of the leading causes of death globally [1]. Atherosclerosis is a chronic and progressive cardiovascular disease characterized by forming plaques in the arterial intima. These plaques can lead to arterial stenosis, hardening, and plaque rupture, which in turn cause serious complications such as thrombosis, myocardial infarction, and stroke. Therefore, regular examination of the carotid arteries and early detection of carotid atherosclerosis are essential. Magnetic Resonance (MR) black-blood vessel wall imaging (BB-VWI) can effectively display both normal and pathological arterial vessel walls and characterize atherosclerosis, providing important evidence for clinical diagnosis [2]. In clinical practice, manual carotid artery segmentation is time-consuming, subjective, and requires specialized training in vessel wall review [2]. In addition, the complex geometric structures of atherosclerotic lesions and carotid artery (as shown in Fig. 1.1) are also regarded as a crucial reason that yields difficulties for accurate segmentation.

---

*Shandong Artificial Intelligence Institute, Qilu University of Technology (Shandong Academy of Sciences), Jinan, China.

†School of Artificial Intelligence, Hebei University of Technology, Tianjin, China.

‡School of Computer, Electronics and Information, Guangxi University, Nanning, China.

§School of Intelligence Science and Technology, Nanjing University, Nanjing, China.

¶Institute of Biomedical Engineering, Ningbo lnstitute of Materials Technology and Engineering, Chinese Academy of Sciences

‖Shandong Artificial Intelligence Institute, Qilu University of Technology (Shandong Academy of Sciences), Jinan, China. (`dachen.cn@hotmail.com`)

**University Paris Dauphine, PSL Research University, CNRS, UMR 7534, CEREMADE, 75016 Paris, France. (`cohen@ceremade.dauphine.fr`)
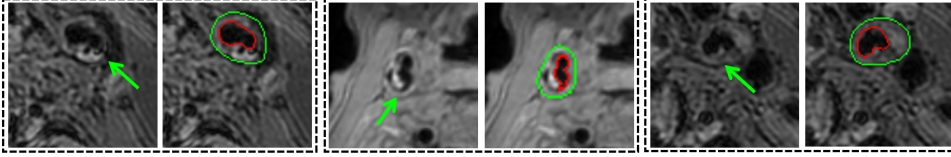
FIG. 1.1. *Typical challenges in CA segmentation task from MRI images. The green arrows point to the vessel walls undergoing complex deformations due to lesions.*

In carotid artery segmentation, traditional methods, such as variational models [3, 4, 5, 6, 7], usually require specialized domain knowledge, leading to poor generalization. With the development of fully convolutional networks, many UNet-based methods have emerged. Menchón-Lara et al. [8] employed a perceptron network segment ultrasound CA images. Shin et al. [9] utilized convolutional neural networks (CNNs) to segment ultrasound CA images. Alblas et al. [10] treated vessel wall segmentation as a multitask regression problem in polar coordinates, encouraging to find continuous and complete segmented structure of the vessel wall. Despite having improved the accuracy and efficiency of the solutions to the CA segmentation task, most of these methods rely on learning semantic features from images for segmentation. As an important shortcoming, the lack of geometric constraints, suffered by these methods, leads to unacceptable structural errors. Azzopardi et al. [11] proposed a geometrically constrained CNN and used amplitude and phase congruency data as input. It imposes shape constraints only considering convex shapes, while the CA may exhibit a certain degree of concavity and convexity simultaneously in reality, especially at bifurcations.

In this paper, we propose a novel granularity attention network and a penalty term based on geometric prior from an anatomical viewpoint, also referred to as anatomical prior, for CA segmentation. A core for the prior lies at the isoperimetric theorem which reveals the essential relation between the perimeter and area of a connected region. More specifically, we design a criterion in terms of the isoperimetric theorem to define the admissibility of a segmented structure. The granularity attention network consists of a two-stage segmentation network combined with the FRA module and the MIE module. This network first performs a coarse segmentation of the region, followed by a refinement process to enhance the network's representational capacity. It is designed to better distinguish the CA from other tissues within the complex anatomy of the neck. The penalty term relying on the introduced prior of the CA imposes constraint to comply with the anatomical structures.

The main contributions are as follows:

(a) We propose a granularity attention network optimized with anatomical prior constraint. The network and constraint ensure the completeness and accuracy of the segmentation by utilizing anatomical prior of the CA and performing feature extraction from coarse to fine granularity.

(b) A novel penalty term is proposed based on the anatomical prior to reduce structurally unacceptable segmentation errors. This prior that is taken as an efficient geometric constraint is able to encourage to detect the correct shapes of the CA.

(c) A granularity attention network with the FRA module and the MIE module is designed to enhance the segmentation accuracy. It captures refinement and multi-scale features through a two-stage network for coarse-to-fine-grained
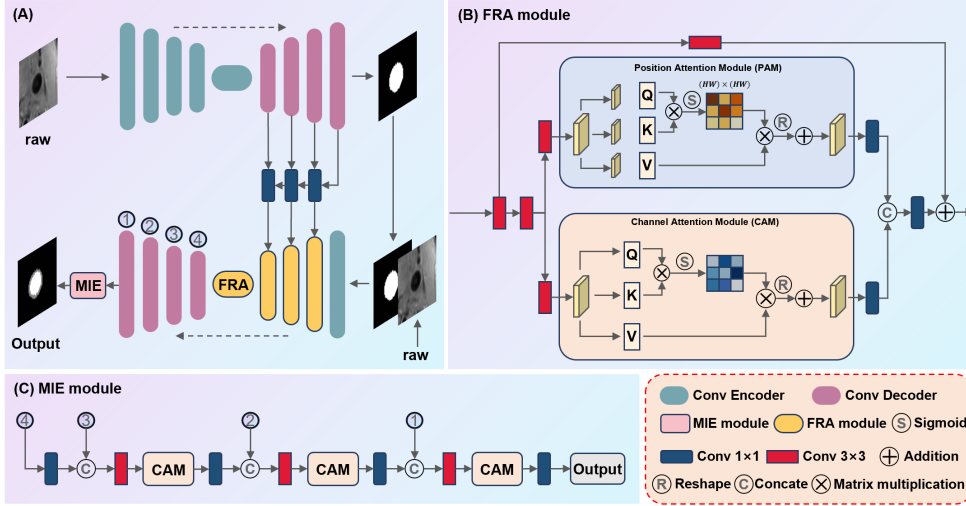
FIG. 2.1. *Diagram of the proposed GAPNet.The backbone of the network consists of two U-shaped networks embedded with the FRA module and the MIE module.*

segmentation.

2. Method.

2.1. Granularity Attention Network. The complex anatomy of the neck region, where various tissues such as blood vessels, nerves, and soft tissues are densely packed and closely positioned. To precisely segment the CA from complex structures, we propose a granularity attention network based on a two-stage architecture, as shown in Fig. 2.1A. The backbone of the network consists of two U-shaped networks. The first stage is employed to extract coarse-grained target regions. It mainly focuses on identifying the approximate contours of the target, performing a rough segmentation of the target area in the image. The decoder output feature maps of the first-stage network are passed through 1×1 convolutions for multi-layer feature fusion before being transmitted to the second-stage encoder. These feature maps contain rich high-level information, and this feature-sharing approach helps the network better understand contextual information, thereby improving segmentation accuracy. The second stage is employed to refine the CA wall and lumen segmentation. The second stage segmentation builds on the first stage segmentation, focusing more on refining the details within the target region. The second stage uses the target area information provided by the first stage to concentrate on more precise segmentation, distinguishing the CA wall and lumen within the target region. In the second stage, we embed the feature refinement attention (FRA) module and the multi-scale information enhancement (MIE) module. The FRA module is employed to refine the higher-level features. The MIE module is used to aggregate information across multiple scales.

The FRA module (as shown in Fig. 2.1B) first performs initial feature extraction and aggregation using two 3×3 convolutions on different layers of features from the first stage decoder. Then, we employ position attention and channel attention modules [12] to weight the features with attention, enabling the network to focus on the channels and spatial positions of interest adaptively. Next, we integrate features from channel attention and position attention, leveraging the strengths of both atten-
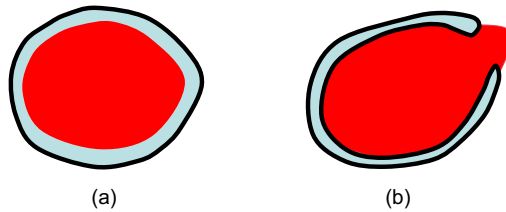
Fig. 2.2. *(a) and (b) respectively illustrate the normal and abnormal segmentation results CA. The regions indicated by cyan color are the segmented CA walls whose external boundary contours are indicated by black solid lines.*

tion mechanisms to enhance feature representation and discrimination capabilities. Finally, we use 1×1 convolutions to adjust the number of parameters, making the model more lightweight and improving computational efficiency and speed. Additionally, through residual connections [13], the original input information is preserved, enhancing gradient flow and improving network training effectiveness. The FRA module can effectively promote cross-layer and cross-stage information interaction. The MIE module (as shown in Fig. 2.1C) integrates a channel attention module and aggregates multi-scale features from the decoder. It is employed to enhance the reconstruction of information. Specifically, we use 3×3 convolutions to progressively fuse hierarchical features from different decoder layers, fully utilizing multi-scale information to improve the accuracy of segmentation details. Then, channel attention is applied to enhance important feature information, increasing the model's ability to capture critical information. Additionally, feature dimensions are adjusted through 1×1 convolutions during multi-scale feature aggregation, reducing model complexity and computational cost.

2.2. Penalty Terms from the Anatomical Prior. In medical image segmentation, the anatomical prior penalty term refers to a technique that penalizes the segmentation results based on the prior knowledge of known anatomical structures. This penalty term is typically used to guide the segmentation algorithm to follow the known anatomical structures or biological rules when generating segmentation results, thereby improving the accuracy and interpretability of the segmentation. Standard loss functions, such as CrossEntropy loss or Mean Squared Error loss, typically compare the output to the ground truth and quantify their differences. These loss functions usually focus on overall matching between predictions of the model and ground truth. In contrast, anatomical prior constraints provide richer information, aiding the model in better understanding the image content. In this study, we incorporate anatomical prior knowledge of the CA as constraints by adding it as an additional term to the loss function. This additional term penalizes inconsistencies between the output and the anatomical prior, thereby enhancing the robustness and accuracy of the model. The definition of the additional term is as follows.
Topology Prior: The CA vessel walls in MRI images usually appear as a *narrow closed band-shape*. However, most of the existing CA segmentation approaches often suffer from the anatomically incorrect leaking problem, where the segmented walls are broken and non-closed. To address this issue on topology changes, we introduce a novel geometric penalization term based on the *isoperimetric theorem*, encouraging a closed narrow band wall structure.

We denote by $\xi : \mathbb{M} \to [0, 1]$ the segmentation prediction of the introduced model,

where $\mathbb{M} \subset \mathbb{R}^2$ is the open bounded image domain. The segmentation region $\mathfrak{S} \subset \mathbb{M}$ can be recovered from the prediction ξ and a thresholding value $\lambda \in (0, 1)$ such that $\xi(x) > \lambda$ means that the point x is inside the segmentation region $\mathfrak{S}$, i.e. $x \in \mathfrak{S}$, and outside $\mathfrak{S}$, otherwise. Let χ_ξ be a binary function associated with the thresholding value λ, which reads as

$$\chi_\xi(x) = \begin{cases} 1, & \text{if } \xi(x) > \lambda \\ 0, & \text{otherwise.} \end{cases}$$

In this case, the length of the boundary $\partial \mathfrak{S}$ can be denoted by

$$\mathcal{L}(\partial \mathfrak{S}) = \int_{\mathbb{M}} \|\nabla \chi_\xi(x)\| dx.$$

Moreover, the area of the region $\mathfrak{S}$ reads as

$$\mathcal{A}(\mathfrak{S}) = \int_{\mathbb{M}} \chi_\xi(x) dx.$$

The isoperimetric theorem states that the length $\mathcal{L}(\partial \mathfrak{S})$ and the area $\mathcal{A}(\mathfrak{S})$ obey $\mathcal{L}(\partial \mathfrak{S})^2 \geq 4\pi \mathcal{A}(\mathfrak{S})$. It measures the relationship between the perimeter of a closed curve and the area it encloses. Therefore, in terms of the isoperimetric theorem, we define a ratio for measuring the shape of the segmentation region as follows:

$$\mu(\mathfrak{S}) := \frac{4\pi \mathcal{A}(\mathfrak{S})}{\mathcal{L}(\partial \mathfrak{S})^2}. \tag{2.1}$$

One can see that the ratio $\mu(\mathfrak{S}) \in [0, 1]$. In particular, if the region $\mathfrak{S}$ is close to a disk, the ratio $\mu(\mathfrak{S}) \approx 1$, while when the segmentation region $\mathfrak{S}$ appears as a *non-circular narrow closed band* shape, $\mu(\mathfrak{S})$ gets to be small. Specific to the CA vessel wall segmentation task, the leaking problem usually yields a narrow band shape of *strong concavity*, as shown in Fig. 2.2(b), which leads to a very low value of $\mu(\mathfrak{S})$. For the normal case, the segmented CA vessel walls can be approximately delineated via either a disk-like shape, or an elliptical-like shape, or a union of multiple disk-like shapes, leading to $\mu(\mathfrak{S}) \approx 1$.

Providing that the boundary $\partial \mathfrak{S}$ represents the *external* contour of the segmented CA vessel wall structure. We define that a segmented CA vessel wall is admissible if the ratio $\mu(\mathfrak{S}) > \tau$. Specifically, we consider the following penalization term

$$\mathcal{P}_{\text{topology}} := \max \{0, -(\mu(\mathfrak{S}) - \tau)\}, \tag{2.2}$$

where τ is a thresholding value and is set as $\tau = 0.6$ through this paper. When $\mu(\mathfrak{S}) > \tau$, the value of $\mu(\mathfrak{S}) - \tau$ is positive, indicating that the topological structure of the segmentation region is acceptable. In this case, taking the negative of $\mu(\mathfrak{S}) - \tau$ results in a value less than 0, ensuring that no penalty is applied in this case. For segmentation results that are below the thresholding value τ, the value of $\mu(\mathfrak{S}) - \tau$ is less than 0. After taking the negative of $\mu(\mathfrak{S}) - \tau$, the result is greater than zero, applying a penalty to these abnormal segmentation results, thus encouraging the model to generate satisfactory segmentation predication.
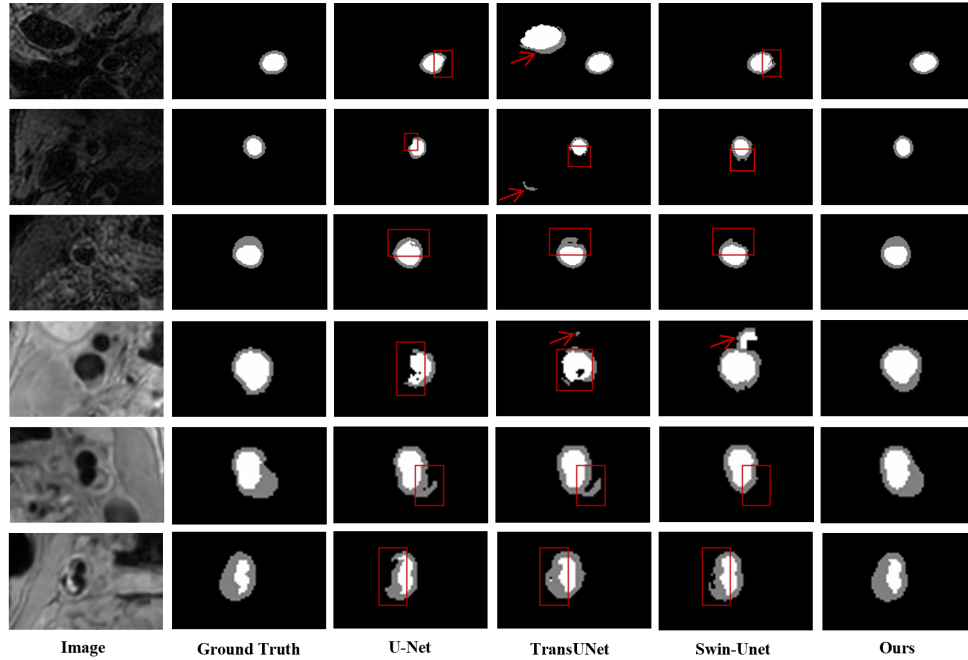
| Image | Ground Truth | U-Net | TransUNet | Swin-Unet | Ours |

FIG. 3.1. *Comparison of different methods. Arrows indicate excessive segmentation and boxes denote incomplete segmentation structures or insufficient details.*

3. Experimental Results.

3.1. Dataset. We used data from the Carotid Vessel Wall Segmentation and Atherosclerosis Diagnosis Challenge (COSMOS 2022) [2] and the Carotid Artery Vessel Wall Segmentation Challenge 2021(CAVWSC 2021) [14] to validate our method. The COSMOS 2022 dataset consists of 50 3D MR scans. We randomly split 40 cases (80%) for training and validation, reserving 10 cases (20%) for testing. Among the 40 cases used for training and validation, we obtained 934 annotated slices, with 80% randomly assigned for training and 20% for validation. The testing data comprised 268 annotated slices. For the CAVWSC dataset, we obtained 1737 annotated slices for training and validation, with 80% randomly assigned for training and 20% for validation. The testing set contained 1754 annotated slices. Each image was resized to a resolution of 224×224 for experimentation.

3.2. Implementation Details. Our method uses the PyTorch framework, and experiments are conducted on NVIDIA Tesla V100. We employ the AdamW optimizer with a base learning rate of 0.0001. In the first stage, the network uses Dice loss and cross-entropy loss to supervise the region segmentation; in the second stage, the network uses a combination of Dice loss, cross-entropy loss, and penalty term to supervise the target segmentation. For the first 5 epochs, we utilize the Warm-up learning rate strategy, followed by the Polynomial Decay strategy for learning rate decay after 5 epochs. The batch size is set to 8, and we train for 300 epochs. Data augmentation is performed by adding brightness and Gaussian noise and applying random scaling, rotation, shifting, and cropping. We adopt the Dice coefficient (Dice) and Hausdorff distance (HD) as evaluation metrics. A 5-fold cross-validation is used to evaluate the model.

TABLE 3.1
Performance comparisons for vessel wall and lumen segmentation.

Methods	CAVWSC 2021				COSMOS 2022			
	Vessel Wall		Lumen		Vessel Wall		Lumen	
	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓
U-Net [15]	0.7295	10.5739	0.9154	9.7920	0.8054	5.6000	0.9152	4.8744
UNet++ [16]	0.7432	10.3248	0.9130	9.3279	0.8226	4.7868	0.9213	4.3572
Attention U-Net [17]	0.7415	9.5014	0.9184	8.7724	0.8214	4.8154	0.9246	5.8245
DualAttentionU-Net [18]	0.7356	7.5142	0.9203	7.9179	0.8232	4.0071	0.9297	3.5967
Res-UNet [19]	0.7473	8.3080	0.9190	7.9907	0.8212	3.8654	0.9266	3.6677
TransUNet [20]	0.7528	7.4204	0.9223	7.1371	0.8209	4.4920	0.9295	5.5275
Swin-Unet [21]	0.7485	6.3345	0.9160	4.3895	0.8050	3.6810	0.9200	3.1113
Proposed	**0.7666**	**4.6540**	**0.9328**	**3.5642**	**0.8397**	**2.7375**	**0.9345**	**2.6059**

3.3. Comparison with State-Of-The-Art Models. We compared the proposed method with seven advanced medical image segmentation methods, namely U-Net [15], UNet++ [16], Attention U-Net [17], Dual Attention U-Net [18], Res-UNet [19], TransUNet [20], and Swin-Unet [21]. We conducted the same training and testing on both the COSMOS and CAVWSC datasets using these methods, and the results are shown in Table 3.1. Our method performs optimally in Dice and HD evaluation metrics for segmenting the vessel wall and lumen. Specifically, for the challenging task of vessel wall segmentation, our method achieves a Dice of 0.7666 and an HD of 4.6540 on the CAVWSC dataset and a Dice of 0.8397 and an HD of 2.7375 on the COSMOS dataset.

Fig. 3.1 exhibits the qualitative comparison results that depict the CA in MRI images. The boxes highlight erroneous segmentations caused by complex diseased vessel wall structures or noise interference in the images, producing discontinuities in the results that do not align with the real anatomical structure. On the other hand, the red arrows indicate that TransUNet and Swin-Unet mistakenly identify tissues with similar features as the CA. In contrast, our method excels in finely and accurately segmenting vessel wall structures, yielding smoother segmentation contours.

The heatmap in Fig. 3.2 further confirms that our network can focus on more complete vessel wall areas, particularly in complex diseased vessel wall structures. In summary, our method excels in accurately locating targets in complex images, overcoming interference from nearby similar tissues, and producing more detailed and precise segmentation results, thereby significantly improving result accuracy.

3.4. Ablation Study. To validate the effectiveness of our proposed methods M1: anatomical prior constraint, M2: FRA module, and M3: MIE module in the CA segmentation task, we utilized a two-stage granularity network as the backbone and gradually incorporated our methods for ablation experiments. As shown in Table 3.2, the experimental results indicate that adding our proposed methods individually to the backbone, or combining them in pairs in different ways, all had a positive impact on the experimental results. Finally, the comprehensive GAPNet integrating M1, M2, and M3 achieved the best overall performance on both datasets. On the other hand, to verify the effectiveness of the first stage in extracting coarse-grained target regions for
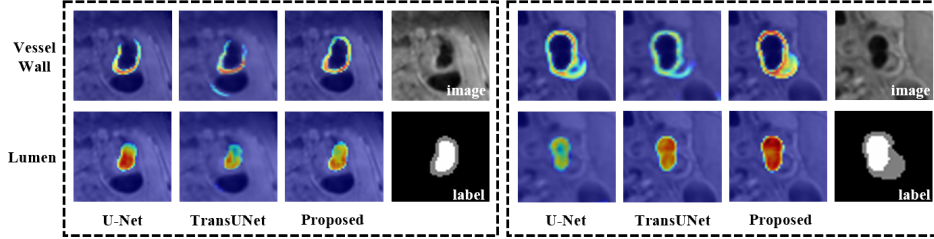
FIG. 3.2. *Visualization of the heat map from the final layer of the decoder.*

TABLE 3.2
Ablation results for vessel wall and lumen segmentation.

| Methods | CAVWSC 2021 | | | | COSMOS 2022 | | | |
| | Vessel Wall | | Lumen | | Vessel Wall | | Lumen | |
	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓
w/o Stage1+M1	0.7515	7.0755	0.9222	6.5112	0.8168	4.0999	0.9258	3.6526
w/o Stage1	0.7554	6.3529	0.9202	5.7233	0.8194	3.9221	0.9273	3.7459
Backbone	0.7569	5.8109	0.9154	5.4130	0.8243	3.7496	0.9253	3.1835
Backbone+M1	0.7622	5.9202	0.9269	5.3959	0.8276	3.2927	0.9259	2.8477
Backbone+M2	0.7636	5.2730	0.9290	4.7746	0.8351	3.2658	0.9321	2.8156
Backbone+M3	0.7605	5.6511	0.9268	4.7736	0.8257	3.2222	0.9232	2.7375
Backbone+M1+M2	0.7628	5.3616	0.9288	4.6485	0.8328	3.0883	0.9300	2.5990
Backbone+M1+M3	0.7633	5.3028	0.9273	4.5919	0.8287	3.4332	0.9223	2.9572
Backbone+M2+M3	0.7632	5.1912	0.9258	4.4615	0.8358	3.0950	0.9343	**2.4486**
Backbone+M1+M2+M3	**0.7666**	**4.6540**	**0.9328**	**3.5642**	**0.8397**	**2.7375**	**0.9345**	2.6059

this task, we conducted experiments by removing the first stage network separately. The experimental results showed that, compared to GAPNet, removing the first stage coarse-grained extraction network had a negative impact on the experimental results. On the CAVWSC dataset, the Dice coefficients for segmenting the vessel wall and lumen decreased by 1.12% and 1.26%, respectively; on the COSMOS dataset, they decreased by 2.03% and 0.72%, respectively. All metrics on both datasets showed a significant decline, indicating that the first stage coarse-grained extraction network helps the network accurately locate target regions, achieving more precise segmentation. In experiments removing the first stage coarse-grained extraction network, we further validated the effectiveness of M1 through ablation. As shown in Table 3.2, the results indicate that introducing M1 positively impacts the model's performance, especially in the vessel wall region. This demonstrates the effectiveness of anatomical priors for carotid vessel segmentation.

4. Conclusion and Future Work. In this work, we introduce an effective method for fully automated and precise segmentation of CA in MRI. Our approach introduces a novel granularity attention network, enabling segmentation from coarse to fine-grained levels and enhancing the ability to capture boundary and detail information through the FRA module and the MIE module. Additionally, anatomical prior constraints are introduced to adjust the segmentation results, thereby improv-

ing segmentation completeness and accuracy. Comprehensive experimental results demonstrate that our method achieves excellent performance on two publicly available datasets, further demonstrating the accuracy and robustness of the model segmentation.

We note that the proposed segmentation model indeed does not take into account geometric regularization such as curvature-based length, and more types of shape priors, for instance the star convexity shape prior which is an important cue for defining the expected segmentation contours in the CA segmentation task. Future work will be devoting to solving these limitations.

REFERENCES

[1] Tsao, C.W., Aday, A.W., Almarzooq, Z.I., Alonso, A., Beaton, A.Z., Bittencourt, M.S., Boehme, A.K., Buxton, A.E., Carson, A.P., Commodore-Mensah, Y., et al.: Heart disease and stroke statistics—2022 update: a report from the american heart association. Circulation **145**(8) (2022) e153–e639

[2] Chen, H., Zhao, X., Dou, J., Du, C., Yang, R., Sun, H., Yu, S., Zhao, H., Yuan, C., Balu, N.: Carotid vessel wall segmentation and atherosclerosis diagnosis challenge(2022). https://vessel-wall-segmentation-2022.grand-challenge.org/

[3] Seabra, J.C., Pedro, L.M., e Fernandes, J.F., Sanches, J.M.: A 3-d ultrasound-based framework to characterize the echo morphology of carotid plaques. IEEE Transactions on Biomedical Engineering **56**(5) (2009) 1442–1453

[4] Yang, X., Jin, J., He, W., Yuchi, M., Ding, M.: Segmentation of the common carotid artery with active shape models from 3d ultrasound images. In: Medical Imaging 2012: Computer-Aided Diagnosis. Volume 8315., SPIE (2012) 718–725

[5] Ukwatta, E., Awad, J., Ward, A., Buchanan, D., Samarabandu, J., Parraga, G., Fenster, A.: Three-dimensional ultrasound of carotid atherosclerosis: semiautomated segmentation using a level set-based method. Medical Physics **38**(5) (2011) 2479–2493

[6] Hossain, M.M., AlMuhanna, K., Zhao, L., Lal, B.K., Sikdar, S.: Semiautomatic segmentation of atherosclerotic carotid artery wall volume using 3d ultrasound imaging. Medical Physics **42**(4) (2015) 2029–2043

[7] Chen, D., Zhang, J., Cohen, L.D.: Minimal paths for tubular structure segmentation with coherence penalty and adaptive anisotropy. IEEE Transactions on Image Processing **28**(3) (2018) 1271–1284

[8] Menchón-Lara, R.M., Sancho-Gómez, J.L., Bueno-Crespo, A.: Early-stage atherosclerosis detection using deep learning over carotid ultrasound images. Applied Soft Computing **49** (2016) 616–628

[9] Shin, J., Tajbakhsh, N., Hurst, R.T., Kendall, C.B., Liang, J.: Automating carotid intima-media thickness video interpretation with convolutional neural networks. In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. (2016) 2526–2535

[10] Alblas, D., Brune, C., Wolterink, J.M.: Deep-learning-based carotid artery vessel wall segmentation in black-blood mri using anatomical priors. In: Medical Imaging 2022: Image Processing. Volume 12032., SPIE (2022) 237–244

[11] Azzopardi, C., Camilleri, K.P., Hicks, Y.A.: Bimodal automated carotid ultrasound segmentation using geometrically constrained deep neural networks. IEEE Journal of Biomedical and Health Informatics **24**(4) (2020) 1004–1015

[12] Fu, J., Liu, J., Tian, H., Li, Y., Bao, Y., Fang, Z., Lu, H.: Dual attention network for scene

segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. (2019) 3146–3154

[13] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. (2016) 770–778

[14] Zhao, X., Li, R., Hippe, D.S., Hatsukami, T.S., Yuan, C., Investigators, C.I., et al.: Chinese atherosclerosis risk evaluation (care ii) study: a novel cross-sectional, multicentre study of the prevalence of high-risk atherosclerotic carotid plaque in chinese patients with ischaemic cerebrovascular events—design and rationale. Stroke and vascular neurology **2**(1) (2017)

[15] Ronneberger, O., Fischer, P., Brox, T.: 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) 234–241

[16] Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J.: Unet++: A nested u-net architecture for medical image segmentation. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4, Springer (2018) 3–11

[17] Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., et al.: Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999 (2018)

[18] Yu, H., Zha, S., Huangfu, Y., Chen, C., Ding, M., Li, J.: Dual attention u-net for multi-sequence cardiac mr images segmentation. In: Myocardial Pathology Segmentation Combining Multi-Sequence Cardiac Magnetic Resonance Images: First Challenge, MyoPS 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Proceedings 1, Springer (2020) 118–127

[19] Xiao, X., Lian, S., Luo, Z., Li, S.: Weighted res-unet for high-quality retina vessel segmentation. In: 2018 9th international conference on information technology in medicine and education (ITME), IEEE (2018) 327–331

[20] Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., Lu, L., Yuille, A.L., Zhou, Y.: Transunet: Transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306 (2021)

[21] Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., Wang, M.: Swin-unet: Unet-like pure transformer for medical image segmentation. In: European Conference on Computer Vision, Springer (2022) 205–218