MoE-based Feature Adapter for Prompt-free Binary Coronary Artery Segmentation in X-ray Angiography

Lin Xi[1,2], **Yingliang Ma**[2,3,*]
[1]University College London, United Kingdom
[2]University of East Anglia, United Kingdom
[3]King's College London, United Kingdom
xilin.chibchin@outlook.com, yingliang.ma@uea.ac.uk

Abstract

Accurate segmentation of coronary arteries in X-ray angiography videos is essential for quantitative coronary analysis and image-guided interventions. However, accurate segmentation remains challenging because coronary vessels are thin and exhibit low contrast, while the presence of catheters, guidewires, and complex anatomical background structures can further interfere with vessel delineation. Existing U-Net- and Transformer-based models provide strong baselines, but their shared feature-adaptation pathways may be insufficient for heterogeneous angiographic appearances. In this paper, we propose a prompt-free mixture-of-experts (MoE) feature adapter for binary coronary artery segmentation. Built upon parameter-efficient Vision Transformer adapters, the proposed method uses multiple lightweight experts with input-dependent top-k routing to adaptively refine vessel-related features while limiting active computational cost. Experiments on MOSXAV and external evaluation on XACV show that the proposed method outperforms representative baselines and improves cross-dataset generalisation. These results suggest that MoE-based adapter learning is effective for robust coronary artery segmentation in X-ray angiography videos.

***K*eywords** X-ray angiography · Coronary artery segmentation · Mixture-of-experts · Adapter learning

1 Introduction

X-ray coronary angiography is widely used for the diagnosis and treatment of coronary artery disease. It provides real-time visualisation of coronary arteries during clinical procedures and remains one of the most important imaging modalities in interventional cardiology. Therefore, automatic coronary artery segmentation from angiography videos is an important step for quantitative vessel analysis, treatment planning, image-guided interventions, and downstream computer-assisted diagnosis.

Despite recent advances in medical image segmentation, robust coronary artery segmentation in X-ray angiography remains challenging. Coronary arteries are thin, tortuous, and curvilinear, and they often appear partially discontinuous due to uneven contrast-agent filling and weak distal vessel visibility. These characteristics are consistent with the broader challenge of curvilinear structure segmentation, where thin tubular structures must be delineated under low contrast, uneven appearance, and complex background interference [1]. Such difficulties are further compounded by a low signal-to-noise ratio and confounding anatomical or device-related structures, both of which can substantially impair vessel delineation [2, 3]. In particular, catheters and guidewires may exhibit grey-level intensities similar to those of coronary vessels, while motion artefacts can blur vessel boundaries and introduce structural ambiguity. Together, these factors produce strong foreground–background ambiguity. Consequently, robust coronary artery segmentation requires effective modelling of fine curvilinear structures under low-contrast, discontinuous, and cluttered imaging conditions.

*Corresponding author

Encoder-decoder networks, such as U-Net [4] and Attention U-Net [5], have been widely used for medical image segmentation because they combine hierarchical feature extraction with spatial detail recovery. nnU-Net [6] further provides a strong self-configuring baseline by automatically adapting preprocessing, network design, and training strategies to a given dataset. More recently, Transformer-based segmentation models have shown strong potential for modelling long-range dependencies in medical images [7, 8]. For X-ray angiography, recent studies have explored semi-supervised segmentation with noisy pseudo-labels and introduced MOSXAV as a benchmark for angiography video segmentation [9], while few-shot video object segmentation has been investigated using local matching and spatio-temporal consistency [10]. However, robust binary coronary artery segmentation remains difficult because vessel and background regions exhibit highly heterogeneous visual patterns. Thin distal branches require the preservation of fine local details, weakly opacified vessels demand sensitivity to subtle intensity variations, and regions overlapping with catheters or guidewires require effective suppression of non-vessel responses. A single shared feature transformation may therefore lead to missed small branches, over-smoothed boundaries, or false positives. This motivates a more flexible feature adaptation mechanism that can refine vessel-related representations according to local appearance variations while maintaining robust global semantic understanding.

Parameter-efficient adaptation provides a promising alternative for adapting large visual backbones to downstream medical segmentation tasks [11, 12, 13, 14, 15, 16]. AdaptFormer [17] introduces lightweight adapter modules for Vision Transformers and demonstrates that compact trainable branches can effectively adapt transformer representations with limited additional parameters. Nevertheless, the original adapter design relies on a single adapter branch, which applies the same feature transformation to all input representations. For coronary artery segmentation, such a uniform adaptation strategy may be insufficient to handle the heterogeneous angiographic appearances caused by thin vessels, low contrast, contrast-filled arterial segments, catheter overlap, and complex anatomical background structures [1, 2, 3]. This motivates an input-dependent feature adaptation strategy, where multiple expert branches can refine vessel-related representations according to local appearance variations, inspired by mixture-of-experts and conditional computation [18, 19, 20].

In this work, we propose a prompt-free mixture-of-experts feature-adapter framework for binary coronary artery segmentation in X-ray angiography videos. Instead of using a single AdaptFormer-style adapter, the proposed framework introduces multiple lightweight adapter experts and an input-dependent routing mechanism to dynamically combine the most relevant experts for each feature representation. Inspired by sparse mixture-of-experts learning [21], this design enables adaptive feature refinement for heterogeneous coronary angiography patterns while maintaining computational efficiency through sparse top-k routing. Since the method is prompt-free, it does not require manual prompts or additional prompt engineering during training or inference.

The main contributions of this paper are summarised as follows:

- We propose a MoE-based AdaptFormer-style lightweight adapters for input-dependent feature adaptation in coronary angiography segmentation.
- We introduce sparse top-k expert routing to improve adaptation flexibility while limiting active computational cost.
- We evaluate the proposed method on MOSXAV and XACV, demonstrating improved segmentation performance and cross-dataset generalisation over representative baselines.

2 Method

2.1 Problem Formulation

Given an X-ray angiography frame $\boldsymbol{I} \in \mathbb{R}^{H\times W}$, the goal is to predict a binary coronary artery segmentation mask $\boldsymbol{Y} \in \{0,1\}^{H\times W}$, where vessel pixels are labelled as foreground and all non-vessel pixels are labelled as background. The predicted probability map is denoted as $\hat{\boldsymbol{Y}} \in [0,1]^{H\times W}$.

The proposed model is formulated as:

$$\hat{\boldsymbol{Y}} = D(E(\boldsymbol{I};\theta_E,\theta_A);\theta_D), \tag{1}$$

where E is the transformer-based encoder, D is the segmentation decoder, θ_E and θ_D denote the backbone parameters, and θ_A denotes the parameters of the proposed MoE-based feature adapter modules.

2.2 Overview of MoE-based Feature Adapter

MoE-based Feature Adapter follows an encoder-decoder segmentation design. The encoder extracts hierarchical feature representations from the input angiography frame. At selected transformer blocks, we insert MoE adapter modules to

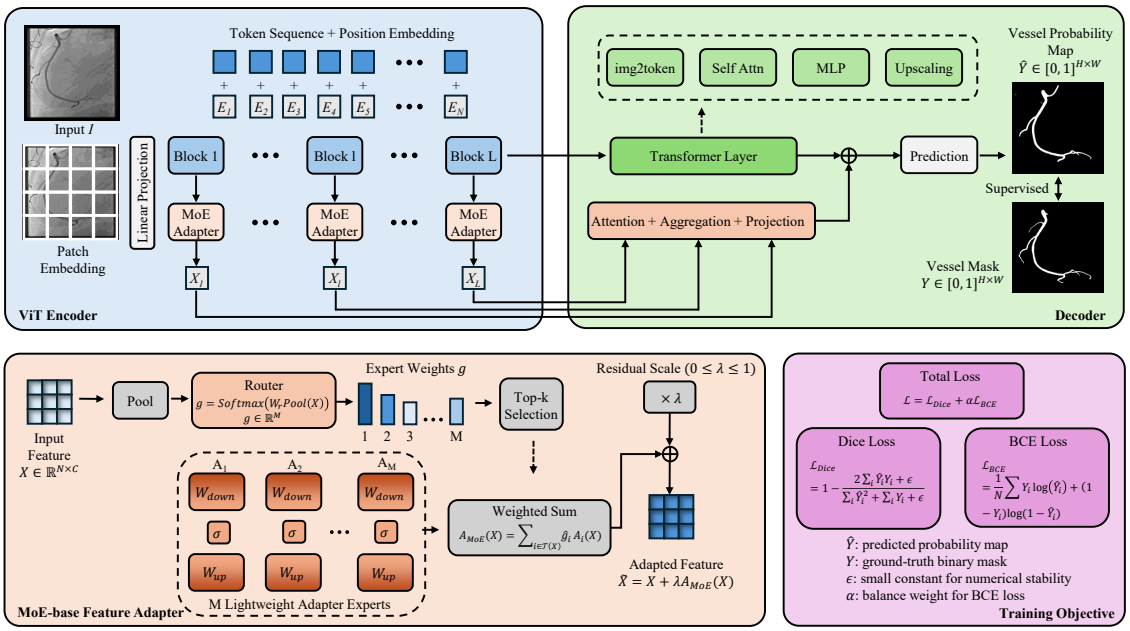

Figure 1: Overall architecture of the proposed network.

refine intermediate token features. The decoder progressively fuses multi-scale features and predicts the final binary vessel probability map.

The key component is the MoE-based feature adapter. Instead of using a single adapter branch, the proposed module contains multiple lightweight adapter experts and a router. For each input feature, the router predicts the importance of each expert. Only the top-k experts are activated, and their outputs are combined to generate an adaptive residual update.

2.3 AdaptFormer-style Lightweight Adapter

Given an intermediate token feature $\boldsymbol{X} \in \mathbb{R}^{N \times C}$, where N is the number of tokens and C is the channel dimension, a lightweight adapter performs bottleneck feature transformation:

$$A(\boldsymbol{X}) = \boldsymbol{W}_{up}\sigma(\boldsymbol{W}_{down}\boldsymbol{X}), \tag{2}$$

where $\boldsymbol{W}_{down} \in \mathbb{R}^{C \times d}$ projects the feature into a lower-dimensional bottleneck space, $\boldsymbol{W}_{up} \in \mathbb{R}^{d \times C}$ restores the original dimension, $d \ll C$, and σ is a non-linear activation function.

The adapted feature is obtained through residual refinement:

$$\tilde{\boldsymbol{X}} = \boldsymbol{X} + \lambda A(\boldsymbol{X}), \tag{3}$$

where λ controls the contribution of the adapter branch.

2.4 MoE-based Feature Adapter

A single adapter applies one fixed transformation to all angiographic patterns. To improve flexibility, we replace it with a MoE-based feature adapter containing M lightweight adapter experts:

$$\{A_1, A_2, ..., A_M\}. \tag{4}$$

The router predicts expert weights from the input feature:

$$\boldsymbol{g} = \mathrm{Softmax}(\boldsymbol{W}_r \mathrm{Pool}(\boldsymbol{X})), \tag{5}$$

where $\mathrm{Pool}(\cdot)$ aggregates token features and W_r is a learnable projection. The output $\boldsymbol{g} \in \mathbb{R}^M$ represents the importance of each expert.

To reduce computation and encourage expert specialisation, we use top-k routing. Let $\mathcal{T}_k(\boldsymbol{g})$ denote the indices of the top-k expert weights. The normalised routing weight is:

$$\hat{g}_i = \frac{g_i}{\sum_{j \in \mathcal{T}_k(\boldsymbol{g})} g_j}, \quad i \in \mathcal{T}_k(\boldsymbol{g}). \tag{6}$$

The MoE adapter output is:

$$A_{\mathrm{MoE}}(\boldsymbol{X}) = \sum_{i \in \mathcal{T}_k(\boldsymbol{g})} \hat{g}_i A_i(\boldsymbol{X}). \tag{7}$$

The final adapted feature is:

$$\tilde{\boldsymbol{X}} = \boldsymbol{X} + \lambda A_{\mathrm{MoE}}(\boldsymbol{X}). \tag{8}$$

This design enables complementary expert refinements for heterogeneous vessel and background appearances.

2.5 Multi-level Adapter Insertion

Coronary arteries appear at different spatial scales, from thick proximal segments to thin distal branches. We therefore insert MoE adapters into multiple transformer stages. Let $\{\boldsymbol{X}_1, \boldsymbol{X}_2, ..., \boldsymbol{X}_L\}$ denote encoder features from different stages. The adapted feature at stage l is:

$$\tilde{\boldsymbol{X}}_l = \begin{cases} \boldsymbol{X}_l + \lambda_l A^l_{\mathrm{MoE}}(\boldsymbol{X}_l), & l \in \mathcal{S}, \\ \boldsymbol{X}_l, & l \notin \mathcal{S}, \end{cases} \tag{9}$$

where $\mathcal{S}$ denotes the selected adapter insertion stages. The adapted multi-level features are then passed to the decoder for binary mask prediction.

2.6 Training Objective

The model is trained using a combination of binary Dice loss and binary cross-entropy loss:

$$\mathcal{L} = \mathcal{L}_{Dice} + \alpha \mathcal{L}_{BCE}, \tag{10}$$

where α balances the contribution of the BCE term, dice loss encourages foreground overlap, while BCE provides pixel-wise supervision for vessel and background classification.

3 Experiments

3.1 Datasets

3.1.1 MOSXAV.

We use MOSXAV [9, 10] as the main training and testing dataset. MOSXAV is a benchmark dataset for X-ray angiography video segmentation. In this work, we formulate the task as binary coronary artery segmentation. Specifically, the vessel annotation is used as foreground, and all remaining pixels are treated as background. All models are trained on the MOSXAV training set, selected using the validation set, and evaluated on the MOSXAV test set.

3.1.2 XACV.

To evaluate external generalisation, we test the trained models on the XACV [22] dataset. XACV is an X-ray coronary angiography video dataset with manually annotated vessel segmentation ground truth. For XACV, we directly use the available binary vessel segmentation ground truth for external evaluation. In other words, the models trained on MOSXAV are directly evaluated on XACV to assess cross-dataset robustness.

Table 1: Binary coronary artery segmentation results on the MOSXAV validation and test sets.

Method	Validation Set				Test Set			
	Dice ↑	IoU ↑	Precision ↑	Recall ↑	Dice ↑	IoU ↑	Precision ↑	Recall ↑
U-Net [4]	81.15	69.02	84.10	79.98	35.55	23.22	26.55	71.87
Attention U-Net [5]	81.76	69.86	86.50	79.11	38.68	25.85	29.43	70.27
nnU-Net [6]	83.14	72.03	92.01	73.58	45.76	28.59	40.33	73.04
Pseudo Diffusion [9]	80.82	66.89	82.36	72.63	34.78	22.97	24.26	68.11
AdapterSeg [17]	83.83	72.82	85.37	83.53	50.57	36.13	40.60	**81.96**
MaskVSC [24]	78.70	65.77	81.40	78.15	35.56	23.19	24.80	72.40
nnWNet [23]	84.07	72.99	**93.03**	74.62	46.18	29.17	41.26	74.29
Ours	**84.75**	**74.05**	85.22	**85.97**	**53.47**	**39.09**	**46.29**	77.74

3.2 Implementation Details

All input frames are resized to 1024×1024. Pixel intensities are normalised to [-1,1]. The models are trained using AdamW with an initial learning rate of 3e-4 and a weight decay of 0.01. Training is performed for 40 epochs. The best model is selected according to the validation Dice score.

For the MoE-based feature adapter, we use $M = 4$ experts and activate top-$k = 2$ experts for each feature. The adapter bottleneck ratio is set to 1. MoE adapters are inserted into all ViT stages. The loss weight α is set to 0.7.

We compare the proposed method with several representative baselines, including U-Net [4], Attention U-Net [5], nnU-Net [6], nnWNet [23], AdapterSeg [17], MaskVSC [24], and a recently validated pseudo-label-based diffusion segmentation model for binary vessel segmentation [9].

3.3 Evaluation Metrics

We evaluate binary segmentation performance using Dice score, Intersection over Union (IoU), precision, and recall. Dice and IoU measure mask overlap, while precision and recall assess false-positive suppression and vessel recovery, respectively. Since F1 is equivalent to Dice for binary foreground segmentation when computed from pixel-level precision and recall, we omit F1 to avoid redundancy.

4 Results

4.1 Results on MOSXAV

Table 1 reports the quantitative results on MOSXAV. On the validation set, the proposed method achieves the best Dice, IoU, and recall, indicating improved vessel recovery and overall overlap performance. Although nnU-Net and nnWNet obtain higher precision, their lower recall suggests a more conservative prediction tendency. In contrast, our method provides a better precision-recall balance.

On the test set, our method achieves the best Dice and IoU, outperforming the strongest baseline, AdapterSeg. It also obtains the highest precision, suggesting improved suppression of vessel-like background responses. While AdapterSeg achieves the highest recall, its lower precision indicates more false-positive predictions. Overall, these results demonstrate that MoE-based feature adaptation improves segmentation robustness under challenging angiographic appearances.

4.2 Results on XACV

Table 2 reports external testing results on XACV, where all models are trained on MOSXAV and evaluated without fine-tuning. The proposed method achieves the best Dice, IoU, and recall, showing stronger cross-dataset generalisation than the compared baselines. Although nnWNet obtains slightly higher precision, its lower recall indicates a more conservative prediction behaviour. In contrast, our method maintains competitive precision while recovering more vessel pixels, leading to the best overall overlap performance.

These results suggest that the proposed MoE-based adapter does not merely overfit to MOSXAV but improves robustness under unseen imaging distributions, where vessel contrast, background structures, and acquisition characteristics may differ from the training data.

Table 2: External binary vessel segmentation results on XACV.

Method	Dice ↑	IoU ↑	Precision ↑	Recall ↑
U-Net [4]	65.58	49.62	69.32	62.99
Attention U-Net [5]	64.30	48.25	71.28	59.44
nnU-Net [6]	67.02	51.23	77.61	59.66
Pseudo Diffusion [9]	62.37	45.81	69.58	59.33
AdapterSeg [17]	68.41	52.76	76.20	62.95
MaskVSC [24]	60.56	44.37	69.20	54.65
nnWNet [23]	68.22	52.03	**77.91**	60.94
Ours	**70.25**	**54.95**	76.84	**65.37**

Table 3: Ablation study on MOSXAV.

Variant	Dice ↑	IoU ↑	Precision ↑	Recall ↑	Params
Baseline backbone	48.37	32.77	62.19	60.72	308M
+ Single adapter	83.83	72.82	85.37	83.53	325M
+ MoE adapter, all experts	84.25	73.91	85.08	84.88	356M
+ MoE adapter, top-1	83.76	73.19	85.13	83.47	356M
+ MoE adapter, top-k	**84.75**	**74.05**	**85.22**	**85.97**	356M

4.3 Ablation Study

Table 3 presents the ablation study on MOSXAV. The baseline backbone performs poorly without adapter-based adaptation, while adding a single adapter substantially improves Dice and IoU, confirming the importance of parameter-efficient feature adaptation. Replacing the single adapter with the proposed MoE adapter further improves performance, indicating that multiple expert pathways provide complementary refinements for heterogeneous angiographic appearances.

Among different routing strategies, top-k routing achieves the best Dice, IoU, precision, and recall. Compared with top-1 routing or activating all experts, top-k routing provides a better balance between expert specialisation and feature diversity.

5 Conclusion

We presented a MoE-based feature adapter framework for prompt-free binary coronary artery segmentation in X-ray angiography videos. The proposed method extends AdaptFormer-style lightweight adapters by introducing multiple expert branches and input-dependent top-k routing, enabling adaptive feature refinement for heterogeneous angiographic appearances.

Experiments on MOSXAV and external testing on XACV demonstrate improved segmentation performance and stronger cross-dataset generalisation compared with representative baselines. These results indicate that MoE-based adapter learning is a promising strategy for robust coronary artery segmentation. Future work will explore temporal modelling across angiography videos and evaluate its impact on downstream clinical measurements such as vessel diameter estimation and stenosis quantification.

Acknowledgements

This work was supported by EPSRC UK (EP/X023826/1).

Disclosure of Interests

The authors have no competing interests to declare that are relevant to the content of this article.

References

[1] Lei Mou, Yitian Zhao, Huazhu Fu, Yonghuai Liu, Jun Cheng, Yalin Zheng, Pan Su, Jianlong Yang, Li Chen, Alejandro F. Frangi, Masahiro Akiba, and Jiang Liu. Cs2-net: Deep learning segmentation of curvilinear structures in medical imaging. *Medical Image Analysis*, 67:101874, 2021. 1, 2

[2] Zijun Gao, Lu Wang, Reza Soroushmehr, Alexander Wood, Jonathan Gryak, Brahmajee Nallamothu, and Kayvan Najarian. Vessel segmentation for x-ray coronary angiography using ensemble methods with deep learning and filter-based features. *BMC Medical Imaging*, 22(1):10, 2022. 1, 2

[3] Lin Xi, Yingliang Ma, Ethan Koland, Sandra Howell, Aldo Rinaldi, and Kawal S. Rhode. Catheter detection and segmentation in x-ray images via multi-task learning. *International Journal of Computer Assisted Radiology and Surgery*, 21:163–173, 2025. 1, 2

[4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI*, pages 234–241, 2015. 2, 5, 6

[5] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y. Hammerla, Bernhard Kainz, Ben Glocker, and Daniel Rueckert. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, 2018. 2, 5, 6

[6] Fabian Isensee, Paul F. Jaeger, Simon A. A. Kohl, Jens Petersen, and Klaus H. Maier-Hein. nnu-net: A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18:203–211, 2021. 2, 5, 6

[7] Jieneng Chen, Yongyi Lu, Qihang Yu, Xiangde Luo, Ehsan Adeli, Yan Wang, Le Lu, Alan L. Yuille, and Yuyin Zhou. Transunet: Transformers make strong encoders for medical image segmentation. *arXiv preprint arXiv:2102.04306*, 2021. 2

[8] Hu Cao, Yueyue Wang, Joy Chen, Dongsheng Jiang, Xiaopeng Zhang, Qi Tian, and Manning Wang. Swin-unet: Unet-like pure transformer for medical image segmentation. In *European Conference on Computer Vision Workshops*, pages 205–218, 2022. 2

[9] Lin Xi, Yingliang Ma, Cheng Wang, Sandra Howell, Aldo Rinaldi, and Kawal S. Rhode. Robust noisy pseudo-label learning for semi-supervised medical image segmentation using diffusion model. In *Deep Generative Models: 5th MICCAI Workshop, DGM4MICCAI 2025*, volume 16128 of *Lecture Notes in Computer Science*, pages 12–23, 2026. 2, 4, 5, 6

[10] Lin Xi, Yingliang Ma, and Xiahai Zhuang. Few-shot video object segmentation in x-ray angiography using local matching and spatio-temporal consistency loss. *Neural Networks*, 200:108808, 2026. 2, 4

[11] Neil Houlsby, Andrei Giurgiu, Stanislaw Jastrzebski, Bruna Morrone, Quentin de Laroussilhe, Andrea Gesmundo, Mona Attariyan, and Sylvain Gelly. Parameter-efficient transfer learning for nlp. In *International Conference on Machine Learning*, pages 2790–2799, 2019. 2

[12] Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022. 2

[13] Menglin Jia, Luming Tang, Bor-Chun Chen, Claire Cardie, Serge Belongie, Bharath Hariharan, and Ser-Nam Lim. Visual prompt tuning. In *European Conference on Computer Vision*, pages 709–727, 2022. 2

[14] Zhe Chen, Yuchen Duan, Wenhai Wang, Junjun He, Tong Lu, Jifeng Dai, and Yu Qiao. Vision transformer adapter for dense predictions. In *International Conference on Learning Representations*, 2023. 2

[15] Junde Wu, Wei Ji, Yuanpei Liu, Huazhu Fu, Min Xu, Yanwu Xu, and Yueming Jin. Medical sam adapter: Adapting segment anything model for medical image segmentation. *arXiv preprint arXiv:2304.12620*, 2023. 2

[16] Kaidong Zhang and Dong Liu. Customized segment anything model for medical image segmentation. *arXiv preprint arXiv:2304.13785*, 2023. 2

[17] Shoufa Chen, Chongjian Ge, Zhan Tong, Jiangliu Wang, Yibing Song, Jue Wang, and Ping Luo. Adaptformer: Adapting vision transformers for scalable visual recognition. In *Advances in Neural Information Processing Systems*, pages 16664–16678, 2022. 2, 5, 6

[18] Robert A. Jacobs, Michael I. Jordan, Steven J. Nowlan, and Geoffrey E. Hinton. Adaptive mixtures of local experts. *Neural Computation*, 3(1):79–87, 1991. 2

[19] Noam Shazeer, Azalia Mirhoseini, Krzysztof Maziarz, Andy Davis, Quoc V. Le, Geoffrey E. Hinton, and Jeff Dean. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. In *International Conference on Learning Representations*, 2017. 2

[20] William Fedus, Barret Zoph, and Noam Shazeer. Switch transformers: Scaling to trillion parameter models with simple and efficient sparsity. *Journal of Machine Learning Research*, 23(120):1–39, 2022. 2

[21] Noam Shazeer, Azalia Mirhoseini, Krzysztof Maziarz, Andy Davis, Quoc V. Le, Geoffrey E. Hinton, and Jeff Dean. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. In *International Conference on Learning Representations*, 2017. 2

[22] Chia-Hung Wu, Sheng-Hung Chen, Chih-Yao Hu, Hao-Yu Wu, Kai-Hung Chen, Yen-Yu Chen, Chih-Hsien Su, Chia-Kai Lee, and Yen-Liang Liu. DeNVeR: Deformable neural vessel representations for unsupervised video vessel segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2025. 4

[23] Yanfeng Zhou, Lingrui Li, Le Lu, and Minfeng Xu. nnwnet: Rethinking the use of transformers in biomedical image segmentation and calling for a unified evaluation benchmark. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 20852–20862, 2025. 5, 6

[24] Yi Zhou, Thiara Sana Ahmed, Meng Wang, Eric A. Newman, Leopold Schmetterer, Huazhu Fu, Jun Cheng, and Bingyao Tan. Masked vascular structure segmentation and completion in retinal images. *IEEE Transactions on Medical Imaging*, 44(6):2492–2503, 2025. 5, 6