Collateral Circulation guided Multi-modality Fusion Network for Postoperative Infarct Prediction

Yichen Guo[1], Xinyi Zhao[1], Lisong Dai[2], Heming Dong[1], Lai Jiang[1,3⋆], Mai Xu[1], and Shengxi Li[1,3]

[1] School of Electronic and Information Engineering, Beihang University, Beijing, China
[2] Department of Radiology, Renmin Hospital of Wuhan University, Wuhan, China
[3] State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, China

Abstract. Acute ischemic stroke is one of the major causes of mortality and disability worldwide. Although thrombectomy is an effective intervention, it carries a lot of risks such as hemorrhage and vascular injury. Thus, it is crucial to accurately predicting postoperative infarct before intervention, providing the guidance for treatment. The existing perfusion imaging techniques relying on fixed thresholding approaches mostly fail to account for individual differences in collateral circulation recruitment, which has been proven to effectively reflect infarct severity. In this work, we take the first step toward integrating collateral circulation status into deep neural network, enabling the model to learn and capture hemodynamic cues for infarct prediction. Specifically, we establish the first brain computed tomography perfusion (CTP) dataset including collateral circulation status and further conduct a thorough analysis of its effectiveness in predicting infarcts. Based on the findings, we propose a novel multi-modal fusion module[4] that integrates spatiotemporal features of multiple modalities. Specifically, a bi-directional Mamba structure is developed to extract the sequential information, which is then fused with collateral priors via a mixture-of-experts mechanism. In addition, a two-stage infarct prediction module is developed to successively localize and segment the infarct region under the guidance of collateral circulation status. Finally, both infarct localization and segmentation performance of our method are validated to outperform 14 state-of-the-art methods.

Keywords: Postoperative infarct prediction · Collateral circulation status · CTP · Multi-modality fusion.

1 Introduction

Acute ischemic stroke, characterized by its high mortality rate and rapid progression, often leads to severe disability even with timely treatment [24]. Thrombectomy is effective on acute ischemic stroke, especially in patients with large vessel

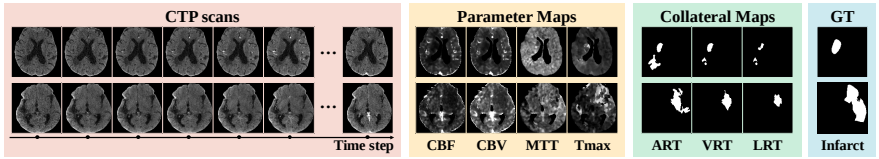
⋆ Corresponding author: Lai Jiang (jianglai.china@buaa.edu.cn)
[4] Codes are available at https://github.com/Frankenstein2026/CCGM

occlusion and minimal irreversible tissue damage on neuroimaging [5]. However, patients still experience poor outcomes due to futile recanalization and face risks such as hemorrhage and vascular injury [33]. Thus, accurately predicting post-operative infarct before intervention is crucial for guiding treatment decisions. Widely utilized in clinical practice, perfusion imaging assesses the preoperative ischemic condition of brain tissue by conducting multi-phase imaging. Nonetheless, existing techniques mostly rely on fixed thresholding approaches, which fail to account for individual differences in collateral circulation recruitment, leading to less accurate prediction. Therefore, it is crucial to develop a precise and automated methodology for infarct prediction leveraging collateral circulation status, which can effectively reflect cerebral perfusion and infarct severity [20].

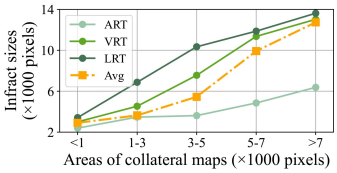
With the rapid development of artificial intelligence, deep neural networks (DNNs) have shown significant potential in medical image tasks, such as disease diagnosis [2, 26] and lesion segmentation [3, 32]. Among them, several DNN-based works are developed for postoperative infarct prediction upon multiple modalities, e.g., CT [22], MRI [21] CT perfusion (CTP) [27], and non-contrast CT (NCCT) [16]. For instance, Liu *et al.* [21] developed a residual attention network for infarct segmentation on MRI. Similarly, Marcus *et al.* [22] introduced a Transformer method to model global feature of CT scans. However, there remains a gap in developing collateral circulation based methods for infarct prediction, both in terms of data availability and DNN architecture. In this paper, we take the first step toward integrating collateral circulation status into DNN, enabling the model to learn the hemodynamic cues for infarct prediction.

To address the above issue, we propose a novel collateral circulation guided multi-modality (CCGM) fusion network for postoperative infarct prediction. Firstly, we establish the first CTP dataset incorporating the collateral circulation status, including the preoperative CTP scans, 4 perfusion parameter maps, 3 collateral maps and the postoperative infarct maps. Through the analysis of our dataset, we find the strong correlation of the collateral maps and brain infarct. Inspired by the findings, our CCGM method is proposed with a multi-modality fusion (MMF) module to fuse the CTP scans, the parameter maps and the collateral maps. Specifically, a Mamba-based structure is developed to extract the spatiotemporal features of CTP scans, and then a mixture-of-experts mechanism is introduced to fuse the multi-modal information, according to the specialized knowledge of each parameter map. Besides, an infarct prediction module is developed to successively localize and segment the infarct, under the condition of collateral maps. Finally, both infarct localization and segmentation performance of our CCGM method are validated to outperform 14 state-of-the-art methods. In summary, the main contributions of this paper are three-fold.

- We establish the first brain CTP dataset with collateral priors for postoperative infarct prediction, and analyze the effectiveness of collateral circulation.
- We propose a novel multi-modal fusion module to fuse the spatiotemporal features from input modalities, based on the developed bi-directional Mamba structure and mixture-of-experts mechanism.
- We introduce a two-stage conditional network, which learns to localize and segment the brain infarct guided by the collateral supply status.



(a) Examples of CTPPC dataset.



(b) Correlation of collateral maps and infarct sizes.

CTP scans		✓	✓	✓	✓	✓	✓	✓
Parameter maps			✓					✓
ART				✓			✓	✓
VRT					✓		✓	✓
LRT						✓	✓	✓
Segmentation Methods	[4]	29.45	34.86	42.24	34.17	40.91	**43.89**	**48.00**
	[6]	28.32	40.24	45.23	42.17	43.86	**48.03**	**50.83**

(c) Dice scores (%) of infarct prediction.

Fig. 1. (a)-Randomly selected examples of our CTPPC dataset with CTP scans, derived images and infarct maps. (b)-The correlation between the areas of collateral maps and brain infarcts. (c)-Segmentation results utilizing different derived modalities.

2 Dataset Establishment and Analysis

2.1 Dataset Establishment

We establish the first brain CTP dataset with collateral maps for postoperative infarct prediction, namely CTPPC, which includes the preoperative imaging data from 161 patients, containing 3,439 brain slices. Specifically, the dataset comprises the whole-brain volumetric CTP imaging, along with 7 derived medical modalities of 4 parameter maps and 3 collateral maps, and the labeled groud-truth infart maps. As shown in Fig. 1-(a), the CTP scans and the derived images are all registered and reshaped into the same size. All the data in CTPPC were collected from the topmost hospitals in China, which was conducted according to the Declaration of Helsinki.

- **Preoperative CTP scans.** Before intervention, the CTP scans of the patient is obtained via the intravenous administration of a contrast agent, and then followed by continuous multi-phase imaging of the brain, consisting of 15 sequential scans in a brief period. This imaging technique monitors the distribution of the contrast agent within the brain, allowing for the assessment of changes in tissue density throughout critical biological phases, i.e., arterial inflow, tissue perfusion, and venous outflow.
- **Preoperative perfusion parameter maps.** The parameter maps are computed based on CTP scans using commercial analysis toolkit (Shukun, Beijing, China), including cerebral blood flow (CBF), cerebral blood volume (CBV), mean transit time (MTT), and Tmax [17]. They deliver detailed information on brain tissue perfusion, facilitating the detection of ischemic areas and the assessment of the extent and severity of cerebral ischemia.
- **Preoperative collateral status maps.** In the CTP scans, we identify 3 critical phases of pathophysiological significance, i.e., peak arterial phase, peak venous phase, and late venous phase [10, 25, 29]. Specifically, we calculate the global voxel grayscale value differences between each voxel in these

phases and the first phase (no contrast agent inflow). Voxels lacking effective collateral supply are identified if the grayscale values remain unchanged, which are designated as arterial-phase risk tissue (ART), venous-phase risk tissue (VRT), and late-venous-phase risk tissue (LRT), respectively.
– **Postoperative infarct areas.** As for the ground-truth of the infarct prediction, the final infarct areas are manually outlined on postoperative diffusion-weighted imaging (DWI) or NCCT images. Two experienced radiologists, blinded to clinical information, independently conducted all manual segmentations, with any differences resolved by consensus.

2.2 Dataset Analysis

Based on our CTPPC dataset, we perform data analysis and obtain the following findings about the correlation between collateral maps and brain infarcts.
Finding 1*: The areas of collateral map and infarct region are highly correlated.*
Analysis: Here, we investigate the correlation between the areas of collateral maps and the postoperative infarct region. Specifically, we calculate the areas of the collateral map and the infarct sizes by counting the corresponding pixels. Then, the collateral maps are uniformly splitted into five groups according to the areas, i.e., <1000, 1000-3000, ..., >7000 pixels. For each group, the averaged infarct sizes are depicted in Fig. 1-(b) and are shown to increase with the areas of the collateral maps. The above results show that the areas of collateral maps and infarct sizes are highly correlated. This completes the analysis of *Finding 1*.
Finding 2: *Collateral maps can obviously improve the infarct prediction accuracy of existing DNN models.*
Analysis: Here, we further explore the impact of collateral maps for infarct prediction. Specifically, 2 widely-used segmentation models (Swin-UNet [4] and TransUNet [6]) are leveraged to predict the postoperative infarcts given different derived images concatenated with CTP scans. The results are evaluated by the segmentation accuracy in terms of Dice score [9]. As shown in Fig.1-(c), together with the CTP scans, the collateral maps tend to bring significant performance gain, which is higher than that of the parameter maps. The above results indicate that the collateral maps can obviously contribute to infarct prediction, which complete the analysis of *Finding 2*.

3 Method

In this section, we introduce our collateral circulation guided multi-modality (CCGM) fusion network for brain infarct prediction. As shown in Fig. 2, our CCGM method is developed with a multi-modality fusion (MMF) module to fuse the modalities of CTP scans, parameter maps and collateral maps, as well as an infarct prediction module to successively localize and segment the infarct regions, under the guidance of collateral circulation status. Specifically, in MMF module, a Mamba structure is developed to extract the spatiotemporal feature of the T-length CTP scans $\{\mathbf{x}_t \in \mathbb{R}^{H\times W\times 1}\}_{t=1}^{T}$ with a size of $H \times W$, which is

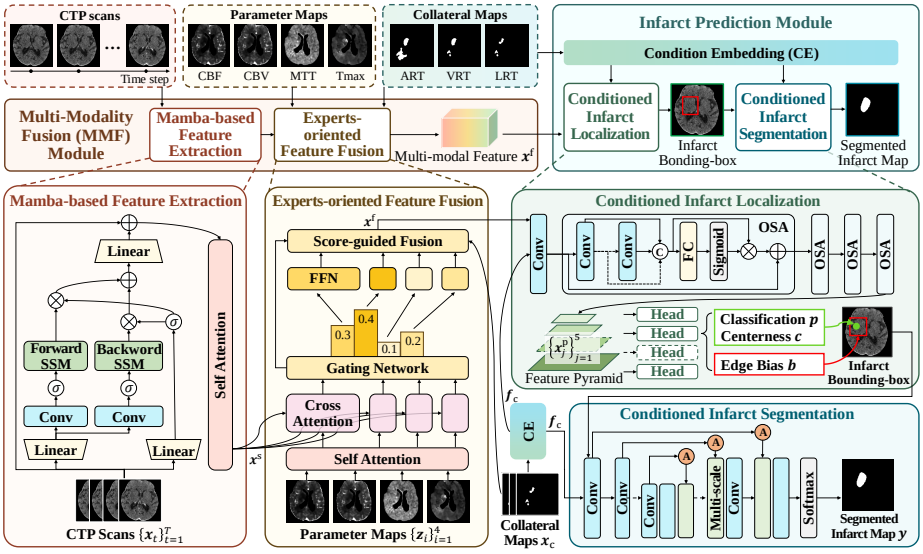


Fig. 2. Architecture of the proposed CCGM network, in which SSM and FFN denote state space model and feed forward network, respectively.

then fused with the derived parameter maps $\{\mathbf{z}_i \in \mathbb{R}^{H\times W\times 1}\}_{i=1}^4$ and collateral maps $\mathbf{x}^{\mathrm{c}} \in \mathbb{R}^{H\times W\times 3}$ via a mixture-of-experts mechanism, to obtain the multi-modal feature $\mathbf{x}^{\mathrm{f}}$ for the following module. In infarct prediction module, inspired by our findings, the collateral maps are initially embedded as collateral prior $\mathbf{f}^{\mathrm{c}}$ for infarct prediction. Finally, conditioned by $\mathbf{f}^{\mathrm{c}}$, the multi-modal feature $\mathbf{x}^{\mathrm{f}}$ is leveraged to localize and segment the infarct region in a two-stage manner.

3.1 Multi-modality Fusion (MMF) Module

As shown in Fig. 2, in MMF module, the CTP scans $\{\mathbf{x}_t\}_{t=1}^T$ are first input to a bi-directional Mamba to extract the spatiotemporal feature. Subsequently, the extracted feature is integrated with the collateral maps $\mathbf{x}^{\mathrm{c}}$ using the expert knowledge of each parameter map $\mathbf{z}_i$, within a transformer-based architecture. The details are introduced below.

Mamba-based Feature Extraction. Inspired by [19], a bi-directional Mamba structure is developed to capture the spatiotemporal feature of the sequential CTP scans. Specifically, the input CTP scans $\{\mathbf{x}_t\}_{t=1}^T$ are first projected into L non-overlapping 16×16 patches $\mathbf{x}^{\mathrm{p}} \in \mathbb{R}^{L\times 256}$. Given the projected patches, the Mamba block $\mathrm{M}(\cdot)$ merges the spatial and temporal features via both forward and backward state space models, followed by a self attention block $\mathrm{Self}(\cdot)$ to extract the attented feature $\mathbf{x}^{\mathrm{s}}$:

$$\mathbf{x}^{\mathrm{s}} = \mathrm{Self}(\mathrm{M}(\mathbf{x}^{\mathrm{p}} + \mathbf{p}^{\mathrm{s}} + \mathbf{p}^{\mathrm{t}})), \tag{1}$$

where $\mathbf{p}^{\mathrm{s}}$ and $\mathbf{p}^{\mathrm{t}}$ denote the spatial and temporal embedding, respectively.

Experts-oriented Feature Fusion.

Since the parameter maps can reflect different blood flow properties, a mixture-of-experts mechanism [35] is developed to further represent the spatiotemporal feature $\mathbf{x}^{\mathrm{s}}$ of CTP, by taking each parameter map as a specialized expert. Specifically, the 4 parameter maps $\{\mathbf{z}_i\}_{i=1}^4$ are first concatenated and encoded via a self attention block to merge the inter-channel information into $\{\mathbf{z}_i^{\mathrm{s}}\}_{i=1}^4$. Then, the CTP feature $\mathbf{x}^{\mathrm{s}}$ is fused with each parameter map $\mathbf{z}_i^{\mathrm{s}}$ by the cross attention blocks $\{\mathrm{Cross}_i(\cdot)\}_{i=1}^4$ into $\{\mathbf{h}_i\}_{i=1}^4$:

$$\mathbf{h}_i = \mathrm{Cross}_i(\mathbf{x}^{\mathrm{s}}, \mathbf{z}_i^{\mathrm{s}}), \quad \mathbf{z}_i^{\mathrm{s}} = \mathrm{Self}([\mathbf{z}_1, \cdots, \mathbf{z}_4])_i. \tag{2}$$

Next, following [35], a gating network $\mathrm{G}(\cdot)$ is leveraged to score the expertises of the cross attention features $\{\mathbf{h}_i\}_{i=1}^4$ corresponding to different parameter maps. Given the scores $\{g_i\}_{i=1}^4$, the outputs of the following feed forward networks $\mathrm{FFN}(\cdot)$ are fused by a score-guided fusion block, comprising of a weighted addition and a Dense block [13] $\mathrm{F}^{\mathrm{DB}}(\cdot)$. Consequently, as illustrated in Fig. 2, the experts-oriented feature fusion can be formulated as,

$$\mathbf{x}^{\mathrm{f}} = \mathrm{F}^{\mathrm{DB}}\left(\left[\sum\nolimits_{i=1}^{4} g_i \mathrm{FFN}_i(\mathbf{h}_i), \quad \mathbf{x}^{\mathrm{c}}\right]\right), \; g_i = \mathrm{Softmax}(\mathrm{G}(\mathbf{h}))_i. \tag{3}$$

Finally, the fused multi-modal feature $\mathbf{x}^{\mathrm{f}}$ merged with diverse expert knowledge is leveraged to the following infarct prediction module.

3.2 Infarct Prediction Module

As shown in Fig. 2, the infarct prediction module includes a collateral prior guided infarct localization network to predict the infarct bounding box in an anchor-free manner, as well as a collateral prior guided infarct segmentation network to obtain the infarct map based on a multi-scale UNet. Besides, inspired by our findings, a condition embedding block composed of a Dense block is leveraged to encode the feature of collateral maps into the conditional guidance $\mathbf{f}^{\mathrm{c}}$. The details are introduced as follows.

Conditioned Infarct Localization. The infarct localization is developed on the top of FCOS [8], with feature pyramid network and multi-scale prediction heads. Specifically, the input multi-modality feature $\mathbf{x}^{\mathrm{f}}$ with the collateral condition $\mathbf{f}^{\mathrm{c}}$ are first encoded by a VoVNetV2 backbone with one-shot aggregation (OSA) [18] blocks $\mathrm{F}_j^{\mathrm{OSA}}(\cdot)$ into the feature pyramid $\{\mathbf{x}_j^{\mathrm{p}}\}_{j=1}^5$:

$$\mathbf{x}_j^{\mathrm{p}} = \begin{cases} \mathrm{Conv}([\mathbf{x}^{\mathrm{f}}, \mathbf{f}^{\mathrm{c}}]), & j = 1, \\ \mathrm{F}_j^{\mathrm{OSA}}(\mathbf{x}_{j-1}^{\mathrm{p}}), & j = 2, \cdots 5, \end{cases} \tag{4}$$

where j denotes the pyramid level. Then, each single-scale head individually predicts the bounding box from corresponding $\mathbf{x}_j^{\mathrm{p}}$. As illustrated in Fig. 2, each prediction head includes three outputs: the classification score $\mathbf{c} \in \mathbb{R}^{H \times W \times 2}$, the centerness $\mathbf{p} \in \mathbb{R}^{H \times W \times 1}$, and the bias $\mathbf{b} \in \mathbb{R}^{H \times W \times 4}$ from the center (u, v) to the four edges, i.e., $\mathbf{b}(u, v) = (l, r, t, d)$. Finally, the pixel of head $\tilde{j}$ with the

Table 1. Results for infarct prediction.

Methods	Dice(%) ↑	HD95(mm) ↓	ASSD(mm) ↓	Acc.(%) ↑	Prec.(%) ↑	Recall(%) ↑
SAN-Net	30.35±24.22	61.05±32.91	21.28±18.57	91.35±5.19	32.17±29.31	43.74±31.85
CA-Net	31.03±28.68	73.24±44.74	36.00±30.40	93.70±4.75	39.53±37.63	33.39±31.30
SiNGR	31.17±33.70	59.29±41.74	30.72±35.09	93.39±5.17	44.72±37.39	35.76±38.98
Bi-JROS	34.30±29.50	38.96±20.47	17.54±14.78	93.65±4.86	40.73±35.21	40.27±34.80
Factorizer	34.56±28.17	39.02±20.42	17.75±14.23	92.20±4.50	38.24±34.42	47.42±35.74
FAT-Net	35.64±29.43	68.58±36.79	30.77±25.73	92.33±5.31	39.63±32.90	50.69±38.48
UNet	43.01±32.17	45.42±36.73	18.05±20.99	93.58±3.87	50.62±37.16	47.47±36.67
nnUNet	43.76±29.71	62.76±42.71	27.00±26.10	92.80±4.60	42.12±33.31	58.01±33.75
Swin-unet	48.00±27.27	50.38±38.80	20.26±23.05	93.80±3.93	50.81±31.82	60.60±31.79
Transunet	50.83±27.34	48.31±37.74	18.11±18.95	93.19±3.69	54.56±32.48	60.70±31.13
Attn UNet	52.57±28.07	42.69±35.59	18.63±19.75	90.34±3.84	53.35±32.52	60.63±31.51
CCGM	56.34±28.14	38.06±37.81	17.13±23.69	94.07±6.45	59.63±34.45	65.60±32.50

Table 2. Results of IoU for infarct localization.

Methods	FCOS	CenterNet	CenterMask	**CCGM (ours)**
IoU (%)	32.93±21.97	34.42±23.36	40.99±28.86	49.32±30.48

highest score of classification and centerness among the pixels of all the heads determines the infarct center $(\tilde{u}, \tilde{v})$ and the edge biases $\mathbf{b}_{\tilde{j}}(\tilde{u}, \tilde{v})$:

$$\tilde{j}, \tilde{u}, \tilde{v} = \max_{j,u,v} \mathbf{c}_j(u,v)\mathbf{p}_j(u,v), \quad \mathbf{b}_{\tilde{j}}(\tilde{u}, \tilde{v}) = (\tilde{l}, \tilde{r}, \tilde{t}, \tilde{d}). \tag{5}$$

Finally, following [18], the centerness $\mathbf{c}$, classification score $\mathbf{p}$ and the bounding box $(\tilde{u}, \tilde{v}, \tilde{l}, \tilde{r}, \tilde{t}, \tilde{d})$ are supervised by the cross-entropy loss, focal loss and intersection over union (IoU) loss, respectively.

Conditioned Infarct Segmentation. For segmentation, the multi-modality feature $\mathbf{x}^{\mathrm{f}}$ is first combined with collateral condition $\mathbf{f}^{\mathrm{c}}$ and cropped into $\mathbf{x}_0^{\mathrm{bb}} = \mathrm{Crop}([\mathbf{x}^{\mathrm{f}}, \mathbf{f}^{\mathrm{c}}])$. On the top of [23], the segmentation model consists of a UNet encoder, decoder and attention gates as $\{\mathrm{E}_k(\cdot), \mathrm{D}_k(\cdot), \mathrm{A}_k(\cdot)\}_{k=1}^{K}$, where k denotes the layer level of UNet. Let $\mathbf{x}_k^{\mathrm{bb}}$ denote the output of each UNet encoder, and the infarct map within the predicted bounding box $\mathbf{y}^{\mathrm{bb}}$ can be calculated as,

$$\mathbf{f}_k^{\mathrm{bb}} = \mathrm{A}_k(\mathrm{E}_k(\mathbf{x}_k^{\mathrm{bb}}) + \mathrm{D}_k(\mathrm{E}_k(\mathbf{x}_k^{\mathrm{bb}}))) \cdot \mathrm{E}_k(\mathbf{x}_k^{\mathrm{bb}}), \quad \mathbf{y}^{\mathrm{bb}} = \mathrm{Softmax}(\mathbf{f}_K^{\mathrm{bb}}), \tag{6}$$

where $\mathrm{D}_k(\cdot) = \mathrm{Conv}_k(\mathrm{MConv}_{\mathrm{k}}(\cdot))$. Here, $\mathrm{MConv}(\cdot)$ denotes the multi-scale convolution [15]. Then, given the predicted bounding box, the cropped map $\mathbf{y}^{\mathrm{bb}}$ can be projected back to the segmented infarct map $\mathbf{y}$. Finally, the segmentation is supervised by the Dice loss and focal loss.

4 Experiments

4.1 Experimental settings

We conduct extensive experiments to evaluate the performance on infarct prediction. Specifically, training and test slices are randomly divided as a ratio of 4/1, with a size of 256×256 pixels. For training, the parameters are updated by SGD

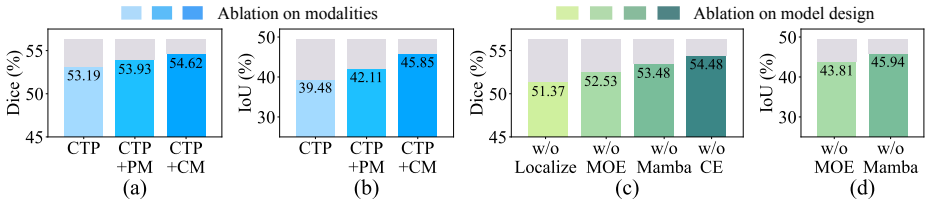


Fig. 3. (a)/(b)-Segmentation/localization results of ablation on modalities. (c)/(d)-Segmentation/localization results of ablation on model design.

optimizer with an initial learning rate of 1×10^{-4} and weight decay of 3×10^{-5}. Note that the segmentation and localization networks are trained separately. All the experiments are conducted with an AMD EPYC 7542 32-core CPU and an NVIDIA RTX 4090 GPU, with a training batch size of 8.

4.2 Evaluation on infarct prediction

We compare our performance with the state-of-the-art methods for brain infarct prediction. For infarct segmentation, we finetune 11 medical image segmentation methods on our CTPPC dataset, inluding SiNGR [7], Bi-JROS [11], SAN-Net [31], Factorizer [1], Swin-unet [4], FAT-Net [30], Transunet [6], nnUNet [14], CA-Net [12], Attention UNet [23] and UNet [28]. Note that all the modalities are leveraged as the input of the compared algorithms. We use Dice score, 95% Hausdorff distance (HD95), average symmetric surface distance (ASSD), accuracy, precision and recall rate to evaluate the segmentation performance. As shown in Tab. 1, our CCGM outperforms all the compared methods in terms of all 6 metrics, which achieves at least 3.77% improvement of Dice score, 0.9mm reduction of HD95, and 0.41mm reduction of ASSD. Besides, for infarct localization, Tab. 2 also shows that our method outperforms 3 bounding box prediction methods: CenterNet [34], FCOS [8] and CenterMask [18], in terms of IoU. The above results validate the effectiveness of our method for infarct prediction.

4.3 Ablation study

Here, we conduct ablation experiments to validate the effectiveness of modality fusion and model design. Specifically, we first investigate the performance of infarct segmentation and localization under different modal fusion: 1) CTP, 2) CTP and parameter maps (PMs) and 3) CTP and collateral maps (CMs). As shown in Fig. 3-(a), the Dice scores degrade from 56.34% to 53.19%, 53.93% and 54.62%, respectively. Similar results of Localization can be found in Fig. 3-(b). This validates the effectiveness of our motivation for multi-modality fusion. Besides, to validate the effectiveness of model design, we also ablate the model structure as: 1) w/o localization: directly segment the infarct using the whole multi-modal feature $\mathbf{x}^{\mathrm{f}}$, 2) w/o MOE: use only one set of cross attention and FFN, 3) w/o Mamba: discard the mamba-based feature extraction and 4) w/o CE: directly use the collateral maps as condition. As can be seen in Fig. 3-(c)/(d), all the ablation settings lead to performance degradation. The above results validate the effectiveness of the model designs for our CCGM method.

5 Conclusion

In this paper, we have proposed a collateral circulation guided multi-modality fusion network for postoperative infarct prediction. Firstly, we established the first brain CTP dataset with collateral circulation status for postoperative infarct prediction. Based on our dataset, we obtained several findings about the strong correlation of collateral maps and brain infarcts. Motivated by the findings, we proposed a multi-modality fusion module to fuse the CTP scans and the derived images based on the developed Mamba structure and mixture-of-experts mechanism, as well as a collateral prior guided infarct prediction module for localization and segmentation. Finally, the extensive experiments compared with 14 state-of-the-art methods validate the effectiveness our CCGM method.

Disclosure of Interests. This work was supported by NSFC under Grants 62401027, 62206011 and 62450131, and Beijing Natural Science Foundation under Grant L223021, and the Fundamental Research Funds for the Central Universities.

References

1. Ashtari, P., Sima, D.M., De Lathauwer, L., Sappey-Marinier, D., Maes, F., Van Huffel, S.: Factorizer: A scalable interpretable approach to context modeling for medical image segmentation. Medical image analysis **84**, 102706 (2023)
2. Bie, Y., Luo, L., Chen, Z., Chen, H.: Xcoop: Explainable prompt learning for computer-aided diagnosis via concept-guided context optimization. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 773–783. Springer (2024)
3. Bui, P.N., Le, D.T., Choo, H.: Visual-textual matching attention for lesion segmentation in chest images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 702–711. Springer (2024)
4. 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. pp. 205–218. Springer (2022)
5. Chen, H., Lee, J.S., Michel, P., Yan, B., et al.: Endovascular stroke thrombectomy for patients with large ischemic core: a review. JAMA neurology (2024)
6. Chen, J., Lu, Y., Yu, Q., Luo, X., et al.: Transunet: Transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306 (2021)
7. Dang, T., Nguyen, H.H., et al.: Singr: Brain tumor segmentation via signed normalized geodesic transform regression. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 593–603. Springer (2024)
8. Detector, A.F.O.: Fcos: a simple and strong anchor-free object detector. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(4) (2022)
9. Dice, L.R.: Measures of the amount of ecologic association between species. Ecology **26**(3), 297–302 (1945)
10. Dundamadappa, S., Iyer, K., Agrawal, A., Choi, D.J.: Multiphase ct angiography: a useful technique in acute stroke imaging—collaterals and beyond. American Journal of Neuroradiology **42**(2), 221–227 (2021)
11. Fan, X., et al.: Bi-level learning of task-specific decoders for joint registration and one-shot medical image segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11726–11735 (2024)

12. Gu, R., Wang, G., Song, T., Huang, R., Aertsen, M., Deprest, J., et al.: Ca-net: Comprehensive attention convolutional neural networks for explainable medical image segmentation. IEEE transactions on medical imaging **40**(2), 699–711 (2020)
13. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4700–4708 (2017)
14. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021)
15. Khened, M., Kollerathu, V.A., Krishnamurthi, G.: Fully convolutional multi-scale residual densenets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. Medical image analysis **51**, 21–45 (2019)
16. Kuang, H., Wang, Y., Liu, J., et al.: Hybrid cnn-transformer network with circular feature interaction for acute ischemic stroke lesion segmentation on non-contrast ct scans. IEEE Transactions on Medical Imaging **43**(6), 2303–2316 (2024)
17. Laughlin, B., Chan, A., Tai, W.A., Moftakhar, P.: Rapid automated ct perfusion in clinical practice. Pract Neurol **2019**, 41–55 (2019)
18. Lee, Y., Park, J.: Centermask: Real-time anchor-free instance segmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 13906–13915 (2020)
19. Li, K., et al.: Videomamba: State space model for efficient video understanding. In: European Conference on Computer Vision. pp. 237–255. Springer (2024)
20. Liebeskind, D.S.: Collateral circulation. Stroke **34**(9), 2279–2284 (2003)
21. Liu, L., Kurgan, L., Wu, F.X., Wang, J.: Attention convolutional neural network for accurate segmentation and quantification of lesions in ischemic stroke disease. Medical Image Analysis **65**, 101791 (2020)
22. Marcus, A., Bentley, P., Rueckert, D.: Concurrent ischemic lesion age estimation and segmentation of ct brain using a transformer-based network. IEEE Transactions on Medical Imaging **42**(12), 3464–3473 (2023)
23. Oktay, O., Schlemper, J., Le Folgoc, 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. In: Medical Imaging with Deep Learning (2018)
24. Organization, W.H.: Global status report on noncommunicable diseases 2014 (2014)
25. Potter, C.A., Vagal, A.S., Goyal, M., Nunez, D.B., Leslie-Mazwi, T.M., Lev, M.H.: Ct for treatment selection in acute ischemic stroke: a code stroke primer. Radiographics **39**(6), 1717–1738 (2019)
26. Qiu, X., Wang, F., Sun, Y., Lian, C., Ma, J.: Towards graph neural networks with domain-generalizable explainability for fmri-based brain disorder diagnosis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 454–464. Springer (2024)
27. Robben, D., Boers, A.M., Marquering, H.A., Langezaal, L.L., Roos, Y.B., van Oostenbrugge, R.J., van Zwam, W.H., Dippel, D.W., Majoie, C.B., van der Lugt, A., et al.: Prediction of final infarct volume from native ct perfusion and treatment parameters using deep learning. Medical image analysis **59**, 101589 (2020)
28. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 234–241. Springer (2015)
29. Wang, Z., Xie, J., Tang, T.Y., Zeng, C.H., Zhang, Y., Zhao, Z., Zhao, D.L., Geng, L.Y., Deng, G., Zhang, Z.J., et al.: Collateral status at single-phase and multiphase

ct angiography versus ct perfusion for outcome prediction in anterior circulation acute ischemic stroke. Radiology **296**(2), 393–400 (2020)
30. Wu, H., Chen, S., Chen, G., et al.: Fat-net: Feature adaptive transformers for automated skin lesion segmentation. Medical image analysis **76**, 102327 (2022)
31. Yu, W., Huang, Z., Zhang, J., Shan, H.: San-net: Learning generalization to unseen sites for stroke lesion segmentation with self-adaptive normalization. Computers in Biology and Medicine **156**, 106717 (2023)
32. Zhong, Y., Tang, C., Yang, Y., Qi, R., Zhou, K., Gong, Y., Heng, P.A., Hsiao, J.H., Dou, Q.: Weakly-supervised medical image segmentation with gaze annotations. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 530–540. Springer (2024)
33. Zhou, T., Yi, T., Li, T., Zhu, L., Li, Y., Li, Z., Wang, M., Li, Q., et al.: Predictors of futile recanalization in patients undergoing endovascular treatment in the direct-mt trial. Journal of NeuroInterventional Surgery **14**(8), 752–755 (2022)
34. Zhou, X., Wang, D., Krähenbühl, P.: Objects as points. In: arXiv preprint arXiv:1904.07850 (2019)
35. Zhou, Y., Lei, T., Liu, H., Du, N., et al.: Mixture-of-experts with expert choice routing. Advances in Neural Information Processing Systems **35**, 7103–7114 (2022)