Pathological Asymmetry-Guided Progressive Learning for Acute Ischemic Stroke Infarct Segmentation

Jiarui Sun, Qiuxuan Li, Yuhao Liu, Yichuan Liu, Gouenou Coatrieux, Jean-Louis Coatrieux, Yang Chen, Jie Lu

Pathological Asymmetry-Guided Progressive Learning for Acute Ischemic Stroke Infarct Segmentation

Jiarui Sun, Qiuxuan Li, Yuhao Liu, Yichuan Liu, Gouenou Coatrieux, *Senior Member, IEEE*, Jean-Louis Coatrieux, *Life Fellow, IEEE*, Yang Chen, *Senior Member, IEEE*, and Jie Lu

***Abstract*—Quantitative infarct estimation is crucial for diagnosis, treatment and prognosis in acute ischemic stroke (AIS) patients. As the early changes of ischemic tissue are subtle and easily confounded by normal brain tissue, it remains a very challenging task. However, existing methods often ignore or confuse the contribution of different types of anatomical asymmetry caused by intrinsic and pathological changes to segmentation. Further, inefficient domain knowledge utilization leads to mis-segmentation for AIS infarcts. Inspired by this idea, we propose a pathological asymmetry-guided progressive learning (PAPL) method for AIS infarct segmentation. PAPL mimics the step-by-step learning patterns observed in humans, including three progressive stages: knowledge preparation stage, formal learning stage, and examination improvement stage. First, knowledge preparation stage accumulates the preparatory domain knowledge of the infarct segmentation task, helping to learn domain-specific knowledge representations to enhance the discriminative ability for pathological asymmetries by constructed contrastive learning task. Then, formal learning stage efficiently performs end-to-end training guided by learned knowledge representations, in which the designed feature compensation module (FCM) can leverage the anatomy similarity between adjacent slices from the volumetric medical image to help aggregate rich anatomical context information. Finally, examination improvement stage encourages improving the infarct prediction from the previous stage, where the proposed perception refinement strategy (RPRS) further exploits the bilateral difference comparison to correct the mis-segmentation infarct regions by adaptively regional shrink and expansion. Extensive experiments on public and in-house NCCT datasets demonstrated the superiority of the proposed PAPL, which is promising to help better stroke evaluation and treatment.**

***Index Terms*—Acute Ischemic Stroke, Infarct Segmentation, Non-Contrast CT, Brain Bilateral Comparison, Progressive Learning**

I. Introduction

Stroke has become the second fatal disease, with approximately 70% of strokes being ischemic in etiology [1]. As we know, infarct extent is strongly associated with the functional outcome of AIS patients [2]. Computed tomography (CT) and magnetic resonance imaging (MRI) are the two main ways to estimate the infarct. MRI is relatively time-consuming, low available and expensive, which makes CT become the prime imaging when patients are on admission [3]. However, manually assessing the infarcts lacks efficiency and accuracy. Therefore, computer-aided and accurate quantitative estimation of AIS infarct on NCCT is an urgent need in the diagnosis and treatment process of stroke.

As demonstrated in [4, 5], the manifestations of ischemic tissues on NCCT are quite inconspicuous in the AIS early phase, which causes AIS infarct segmentation on NCCT to remain intractable. Especially, as described in Fig.1.(a), the pathological hypo-attenuation changes caused by ischemic tissues are subtle and be easily confounded by normal physiologic changes or healthy brain tissue [4]. To differentiate the infarct from normal brain tissue, radiologists usually utilize the symmetric comparison of the bilateral hemispheres in clinical practice [6]. As shown in the left of Fig.1.(b), they first find the ideal mid-sagittal line (MSL) and then compare bilateral hemisphere differences along the MSL to recognize the pathological asymmetry to further locate the infarct.

In recent years, due to the powerful feature representation capacity of the convolutional neural network (CNN), current deep learning (DL)-based methods have demonstrated great potential for AIS infarct segmentation. At first, these works [7–10] usually train a naive DL model to learn an end-to-end mapping from input data to label. Symmetry-based methods [4, 5, 11, 12] further propose to leverage and integrate the domain-specific knowledge of bilateral symmetric property of brain hemispheres to implement lesion information compensation, and achieve greater performance gains. **Limitation:** As depicted in the left of in Fig.1.(b), relevant regions in the bilateral hemisphere comparison include three relations: (1) pathological asymmetries caused by AIS infarcts (2)

This work was supported in part by the State Key Project of Research and Development Plan under Grants 2022YFC2401600 and 2022YFC2408500, in part by the National Natural Science Foundation of China under Grant T2225025, in part by the Key Research and Development Programs in Jiangsu Province of China under Grant BE2021703 and BE2022768, in part by the the National Key Research and Development Program of China under Grants 2022YFE0116700. (Co-first authors: Jiarui Sun and Qiuxuan Li, corresponding authors: Yang Chen and Jie Lu)

Jiarui Sun and Yichuan Liu is with the Laboratory of Image Science and Technology, Southeast University, Nanjing 210096, China (e-mail: 230198566@seu.edu.cn; 220225835@seu.edu.cn).

Qiuxuan Li and Jie Lu are with the department of Radiology and Nuclear Medicine, Xuanwu Hospital, Capital Medical University, and also with the Beijing Key Laboratory of Magnetic Resonance Imaging and Brain Informatics, Beijing 100053, China (e-mail: 18810637460@163.com; imaginglu@hotmail.com).

Yuhao Liu is with the department of Computer Science, City University of Hong Kong, Kowloon 999077, Hong Kong. (e-mail: yuhliu9-c@my.cityu.edu.hk)

Gouenou Coatrieux is with the IMT Atlantique, Inserm, LaTIM UMR1101, Brest 29000, France (e-mail: gouenou.coatrieux@telecom-bretagne.eu)

Jean-Louis Coatrieux is with the Laboratoire Traitement du Signal et de l'Image, Université de Rennes 1, F-35000 Rennes, France, and also with the National Institute for Health and Medical Research, F-35000 Rennes, France (e-mail: jean-louis.coatrieux@univ-rennes1.fr)

Yang Chen is with the Laboratory of Image Science and Technology, the School of Computer Science and Engineering, and the Key Laboratory of New Generation Artificial Intelligence Technology and Its Interdisciplinary Applications (Southeast University), Ministry of Education, Nanjing 210096, China (e-mail: chenyang.list@seu.edu.cn)

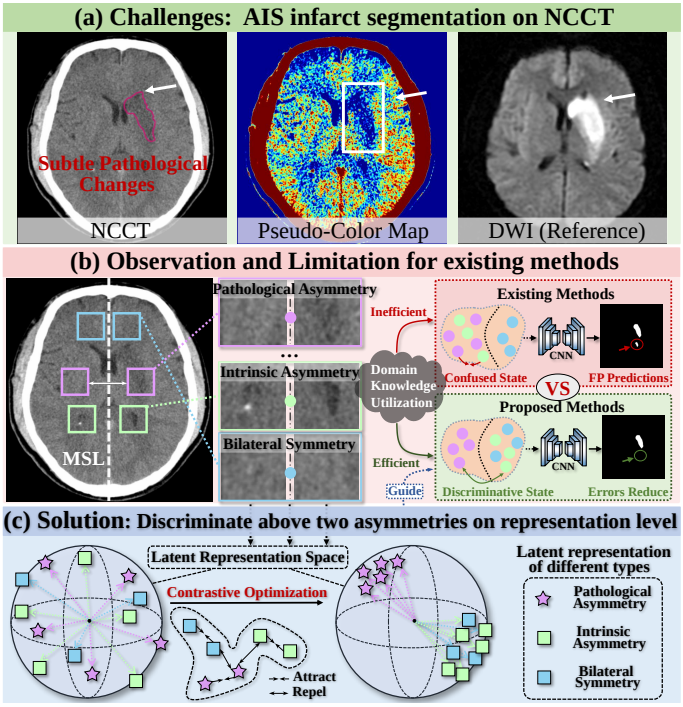


Fig. 1: (a) The AIS infarct segmentation on NCCT is a very challenging task due to subtle pathological changes. (b) The limitation of existing methods is that they commonly ignore or confuse the different contributions of intrinsic and pathological asymmetries in the bilateral hemisphere differences to segmentation. (c) Therefore, we propose to discriminate these two asymmetries on the representation level to fully leverage the domain knowledge for facilitating the infarct segmentation.

intrinsic asymmetries caused by normal physiologic changes (3) bilateral symmetries (no asymmetries). However, existing methods commonly ignore or confuse the different contributions of intrinsic and pathological asymmetries in the bilateral hemisphere differences to AIS infarct segmentation, resulting in extra error predictions for AIS infarct regions caused by inefficient domain knowledge utilization. **Solution:** To solve this problem, as depicted in Fig.1.(c), we propose to discriminate the two different asymmetries on the representation level to fully excavate the potential of this domain knowledge for the AIS infarct segmentation.

In this paper, we propose a pathological asymmetry-guided progressive learning (PAPL) method. PAPL mimics the step-by-step learning patterns observed in humans for progressively segmenting AIS infarct. Specifically, PAPL consists of three stages: knowledge preparation stage, formal learning stage, and examination improvement stage. Firstly, the knowledge preparation stage can accumulate the preparatory domain knowledge of the infarct segmentation task. Intuitively, it helps to learn domain-specific knowledge representations via the constructed contrastive learning task, enhancing the discriminative ability for pathological asymmetries. Then, the formal learning stage integrates domain knowledge by transferring the learned representations to help efficiently perform formal end-to-end training. In this stage, the designed feature compensation module (FCM) measures and learns the anatomy similarity between adjacent slices from the volumetric medical image, which aims to aggregate rich anatomical context information for supplementing infarct details. Finally, we proposed the regional perception refinement strategy (RPRS) in the examination improvement stage to gradually improve the prediction of the formal learning stage. Specifically, RPRS encourages the predicted infarct regions to adaptively shrink and expand by comparing signal intensity differences between the bilateral hemispheres, gradually correcting the mis-segmentation regions. In general, the contributions of our work include the followings:

- Inspired by human learning patterns, we propose a pathological asymmetry-guided progressive learning (PAPL) method including three progressive stages: (a) knowledge preparation stage, (b) formal learning stage, and (c) examination improvement stage, reasonably integrating domain-specific knowledge of brain hemispheres to boost the infarct segmentation performance on NCCT. Extensive experiments on public and in-house datasets demonstrated the superiority of the proposed PAPL. Besides, our method can be conveniently integrated into other segmentation backbones, making it highly applicable.
- In the knowledge preparation stage, we construct a domain-specific contrastive learning task to learn domain-specific knowledge representations and enhance the discriminative ability for pathological asymmetries.
- In the formal learning stage, we develop a feature compensation module (FCM), which measures and learns the anatomy similarity between adjacent slices from the volumetric medical image to aggregate rich anatomical context information for supplementing infarct details.
- We propose a regional perception refinement strategy (RPRS) for the examination improvement stage to gradually improve the mis-segmentation regions from the formal learning stage.

II. Related Work

In this section, we review AIS infarct segmentation, symmetric modeling-based stroke lesion segmentation and contrastive learning literature that are closely relevant to our work.

A. AIS Infarct Segmentation

Most traditional image segmentation methods usually can not handle the segmentation of AIS infarct well because of the insensitivity of NCCT imaging to infarct lesions of the acute phase [12]. Therefore, studies based on traditional segmentation methods are extremely limited. With the rapid development of DL in medical image analysis, DL-based segmentation methods [7, 13–17] show great potential for AIS infarct segmentation, achieving better performance. Kuang *et al.* [7] proposed a dense multi-path contextual method based on the generative adversarial network [18] to implement automated infarct segmentation by feeding rich context information to the generator and discriminator. Wu *et al.* [13] designed a cascaded two-stage CNN-based method to exploit the global and local information of images by an adaptive threshold strategy. Lin *et al.* [14] designed an R2U-Net-improved deep network with multiscale focal loss to mitigate the class imbalance problem caused by lesion targets being too small. Qi *et al.* [15] employs a depthwise separable convolution and feature similarity module to help capture long-range dependencies between the lesion

features for better segmentation. Kuang *et al.* [16] proposed a hybrid CNN and Transformer network with circular feature interaction for accurate AIS infarct segmentation. Zhu *et al.* [17] developed an encoder-decoder-based semantic model with multi-scale atrous convolution for better lesion feature extraction.

B. Symmetric Modeling-Based Stroke Lesion Segmentation

Current studies [4, 5, 11, 19–23] based on brain symmetric modeling show that incorporating the domain knowledge of bilateral symmetric property of brain hemispheres into deep learning models is an effective way for further boosting stroke lesion segmentation. Liang *et al.* [4] proposed a symmetry-enhanced attention module to efficiently model anatomical symmetry by capturing context information from the opposite side of the brain. Ni *et al.* [5] designed an asymmetry disentanglement network with a novel tissue-type aware regularization constraint, which can compensate the extracted asymmetric information onto images for better infarct segmentation. Kuang *et al.* [11] employed a triplet CNN network with comparison disparity block abstract, aiming to implicitly code the asymmetry prior into the network. Wang *et al.* [19] developed a novel lesion symmetry network to reasonably utilize the brain quasi-symmetry, and demonstrates that proper utilization of the quasi-symmetry property is very useful. Bao *et al.* [20] developed a differential feature augmentation module to highlight the semantic asymmetries of deep features from the encoder and applied mirror position-based difference augmentation module to further magnify the discrepancy. Gruber *et al.* [21] presented a novel input enhancement strategy combined with a suitable non-stationary loss to fully exploit the quasi-symmetry property. Yu *et al.* [22] designed symmetry-inspired online data augmentation, which can help the model locate the lesions much easier. Yang *et al.* [23] mimic the priors on symmetry property by introducing self-attention to explore the interaction between each entity and the complete spatial domains including symmetrical regions.

Symmetric modeling-based methods further facilitate segmentation when reasonably leveraging the domain knowledge of the bilateral asymmetry. However, these methods generally confuse the different contributions of intrinsic and pathological asymmetries in the bilateral comparison to segmentation, leading to extra error predictions for the normal brain tissues.

C. Contrastive Learning

Contrastive learning (CL) is a powerful technique for representation learning [24]. The main purpose of the CL is to construct suitable contrastive sample pairs for the proxy task to reasonably leverage and integrate domain- or task-specific knowledge, which helps learn discriminative representations to further boost the model performance. These CL methods including MOCO [25], SimCLR [26] and BYOL [27] were first proposed for natural image tasks. They construct massive contrastive pairs by generating augmented views to facilitate various downstream tasks. Unlike natural images, medical images usually contain much more diverse domain-specific knowledge (ie., organ anatomy structure or vessel topology characteristics). Current CL works [28–35] have demonstrated the great potential of facilitating medical image segmentation by constructing suitable contrastive sample pairs to make full use of diverse domain knowledge. Zeng *et al.* [28] generate contrastive sample pairs based on the position of a slice in volumetric medical images, which can leverage the domain-specific knowledge of medical images as adjacent slices typically contain similar anatomical structures. Chaitanya and Hu *et al.* [29, 30] define contrastive sample pairs based on the fact that volumetric images of the same anatomical regions for different subjects contain similar organ and tissue structure content when capturing the same field-of-view (FOV). You *et al.* [31–34] further propose two new perspectives on addressing existing limitations of works [29, 30] and demonstrating significant advantages when effectively leveraging anatomy knowledge: (1) work [31] utilizes a relaxed softly labeling way to construct contrastive sample pairs, helping to alleviate the issues of false negatives caused by the standard binary supervision. (2) works [32–34] design a range of effective anatomy-based methods including the loss function, feature decomposition and sampling strategy, etc for better head and tail class separations in task scenarios, optimizing learned multi-level anatomical representations when facing the challenge of long-tailed class distribution. Yang *et al.* [35] is enlightened by the domain knowledge in histopathology: different tissues are composed by different cells and extracellular matrices.

These existing CL methods demonstrate the importance of reasonably leveraging domain-specific knowledge in constructing the contrastive sample pairs. However, they cannot be naively applied to our task due to significant differences in domain knowledge. Therefore, we are the first to introduce the CL to fully develop the domain-specific knowledge of the bilateral symmetric property of brain hemispheres for enhancing the feature representation for the pathological asymmetries.

III. METHODOLOGY

This section introduces the proposed PAPL for AIS infarct segmentation. Firstly, Fig.2 describes the overall architecture. Then, three progressive stages of the PAPL in Fig.2 are specified sequentially in subsections III.C-E.

A. Overall architecture

Before PAPL starts, to easily utilize bilateral symmetry property, original NCCT scans are firstly aligned to be in standard symmetric space, keeping the brain region in the center of the image and horizontally symmetric. Then, as described in Fig.2, the overall pipeline of the proposed PAPL consists of three progressive stages: (1) the knowledge preparation stage, (2) the formal learning stage, and (3) the examination improvement stage. Firstly, the knowledge preparation stage learns discriminative representation for pathological asymmetry via the constructed CL proxy task based on the inter- and intra-sample contrastive training scheme, which helps accumulate the preparatory domain-specific knowledge of the infarct segmentation task. Then, the formal learning stage efficiently performs the formal segmentation learning by transferring learned knowledge representation. Meanwhile, the designed feature compensation module (FCM) can leverage the anatomy similarity between adjacent slices, which helps aggregate rich anatomical context information for enhancing the feature representation of the slice participating in end-to-end training.

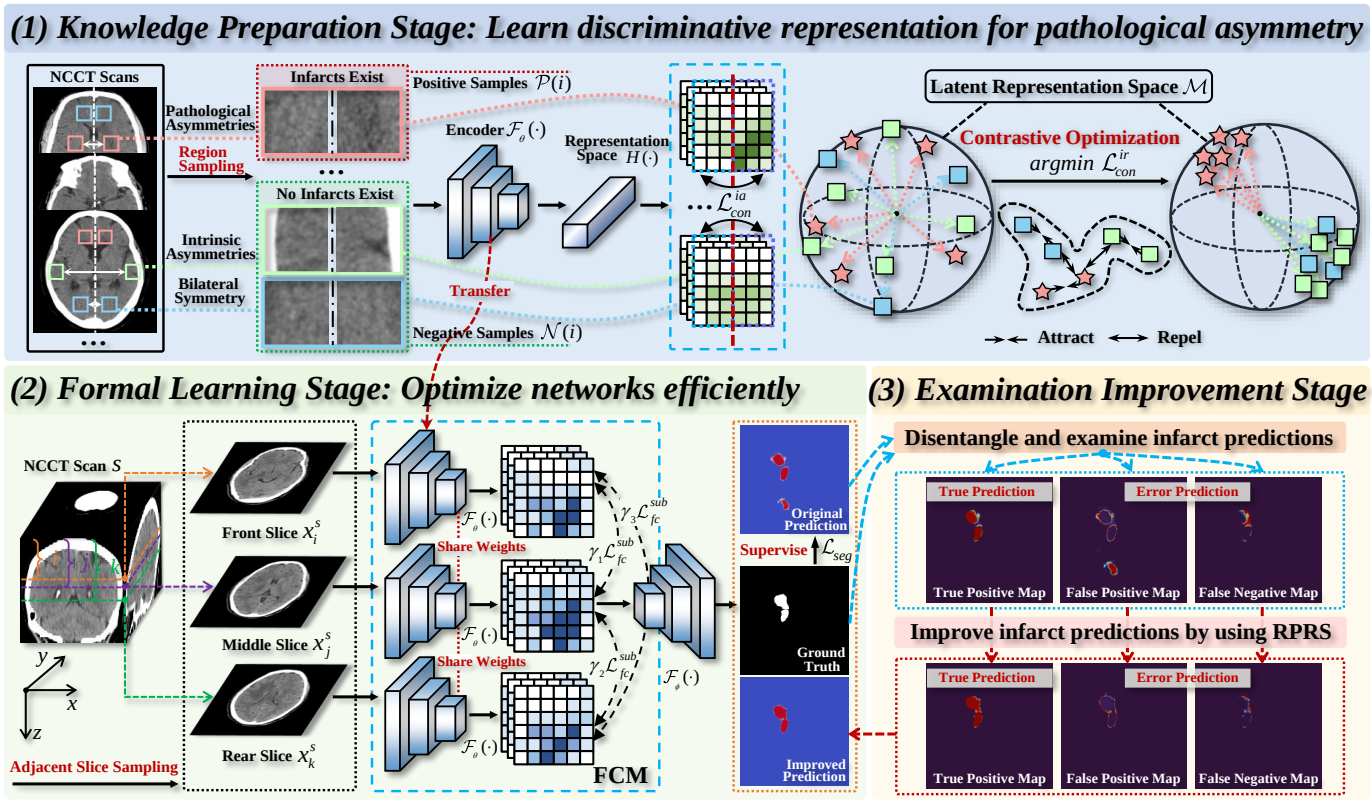


Fig. 2: The overall illustration of the proposed pathological asymmetry-guided progressive learning (PAPL), including three progressive stages: (1) knowledge preparation stage, (2) formal learning stage, and (3) examination improvement stage.

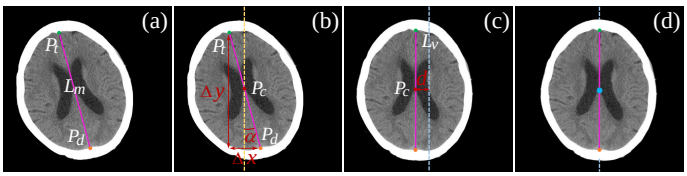


Fig. 3: The illustration of the symmetry space transformation of the NCCT Scans. By the affine transformation including rotation (b) and translation (c), the given slice can be transformed from the unaligned state (a) to the aligned state (d).

Finally, the examination improvement stage improves the infarct predictions from the formal learning stage, in which the proposed regional perception refinement strategy (RPRS) can encourage suitable regional shrinkage and expansion by performing bilateral intensity difference comparison to mitigate mis-segmentation.

B. Symmetry Space Transformation

1) ***Problem Statement****:* Since the brain poses of patients are distinct during NCCT imaging, it is difficult to perform directly symmetry-based calculations on the original images. As shown in Fig.3, to support PAPL to utilize easily bilateral symmetry property, the unaligned NCCT images will be transformed into the aligned state to stay in the standard symmetric space, keeping the brain region in the center of the image and horizontally symmetric. The steps of symmetry space transformation are based on the fact that the brain region stays bilateral symmetric along MSL when MSL is aligned to be in the vertical center of the slices.

2) ***Preliminary Setting****:* The aligning process is conducted on a slice-by-slice basis. For a given slice (Fig.3 (a)), the MSL L_m can be obtained by the pre-trained midline detection network [36]. In L_m, the top point $P_t(x_t, y_t)$ and the down point $P_d(x_d, y_d)$ can be accessed directly. The middle point $P_c(x_c, y_c)$ (Fig.3 (b)) of L_m can is formulated by P_t and P_d:

$$P_c(x_c, y_c) = P_c(\frac{x_t + x_d}{2}, \frac{y_t + y_d}{2}) \,, \tag{1}$$

Meanwhile, the point P_c is also the center of the entire brain region.

3) ***Alignment Transformation****:* Based on the above settings, MSL L_m can be aligned to be in the vertical axis of symmetry (VTI) of the slice by the affine transformation including rotation and translation. To rotate the brain region to keep the horizontally symmetric state (Fig.3 (c)), the rotation angle α of L_m is calculated as:

$$\Delta x = x_d - x_t \,, \tag{2}$$

$$\Delta y = y_d - y_t \,, \tag{3}$$

$$\alpha = arctan(\frac{\Delta x}{\Delta y}) \,, \tag{4}$$

where $arctan(\cdot)$ is the arctangent function. To translate the center P_c of the horizontally symmetric brain region to the VTI L_v (Fig.3 (c)), the horizontal translation distance d is measured as:

$$d = x_i - x_c \,. \tag{5}$$

where x_i is the horizontal distance from L_v to the image boundary. Finally, the NCCT images can be aligned to be in the standard symmetric space (Fig.3 (d)) after the above affine transformation.

C. Knowledge Preparation Stage

1) ***Purpose and Step****:* The knowledge preparation stage encourages the model to learn the discriminative knowledge representations for pathological asymmetries via a learnable encoder $\mathcal{F}_\theta(\cdot)$. For this purpose, a domain-specific CL proxy task is constructed for this stage. In this process, the contrastive sample pairs are first constructed reasonably, and then the inter- and intra-sample contrastive training scheme can facilitate gaining the robust discriminative representation for pathological asymmetries.

2) ***Contrastive Sample Pair Construction****:* Inspired by the domain-specific knowledge in Fig.1(b), the proxy task is designed to address the problem that discriminates whether the pathological asymmetries exist in given a paired patch regions (the region and its mirrored region) along the MSL, which helps integrate the discriminative ability for pathological asymmetry into the $\mathcal{F}_\theta(\cdot)$. As described in Fig.1(b), the included three relations in relevant paired patch regions along MSL can be further divided into two states in Fig.2(1): (1) existing pathological asymmetries caused by AIS infarcts and (2) no existing pathological asymmetries (intrinsic asymmetries and bilateral symmetries). Since the objective aims to discriminate pathological asymmetries, the no-pathological asymmetries including intrinsic asymmetries and bilateral symmetries do not need to be further subdivided. Furthermore, following the designed region mask sampling strategy, the contrastive sample set $\mathcal{D}:=\{(x_i, y_i)\}_{i=1}^{N}$ including positive and negative samples can be constructed to support the learning of the domain-specific knowledge representation, where N means the total number of generated samples x_i, and y_i indicates the positivity or negativity of contrastive sample x_i. As depicted in Fig.4, our strategy is straightforward: first, it samples randomly (ie., variable sampling windows sizes) the paired patch regions along MSL from the aligned slices and masks (i.e., removes) the remaining regions, generating x_i. Especially, sampling is performed multiple times on each slice, and one x_i is generated after only one sampling. Then, the positivity/negativity (ie., y_i) of x_i is determined according to the proportion of infarct area in the sampled paired regions. Finally, by sampling slice by slice, a large number of contrastive samples x_i can be obtained to support gaining effective knowledge representation.

3) ***Inter- and Intra-Sample Contrastive Training****:* As depicted in stage Fig.2(1), the constructed proxy task is conducted by the inter- and intra-sample contrastive training scheme. It consists of the parallel inter- and intra-sample contrastive losses. The former can help to contrast and discriminate the different latent representations with pathological or no-pathological asymmetries on the sample-level, while the latter can measure the similarity of the feature distributions of the intra-sample bilateral hemispheres to provide clear evidence and constraint for the former.

The deep image features of the sample x_i are firstly extracted by the learnable encoder $\mathcal{F}_\theta(\cdot)$ and are projected to latent

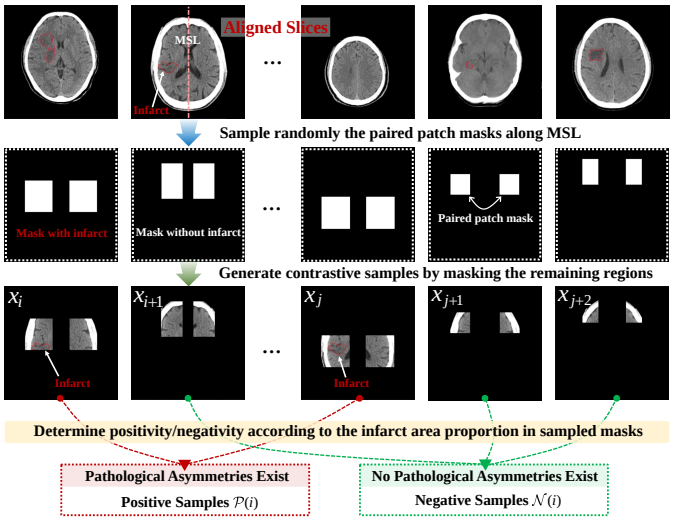


Fig. 4: The construction of positive and negative samples by the designed region mask sampling strategy.

representation space $\mathcal{M}$ by the projection head $H(\cdot)$ as a L2-normalized d-dimension embedding v_i, which can be formulated as:

$$v_i = H(\mathcal{F}_\theta(x_i)) \,, \tag{6}$$

To measure the distance between the given two embeddings v_i and v_j in $\mathcal{M}$, cosine similarity metric $\mathcal{S}(\cdot)$ is formulated as the following:

$$\mathcal{S}(v_i, v_j) = \frac{(v_i)^{\mathrm{T}} \cdot v_j}{||v_i||_2 ||v_j||_2} \,, \tag{7}$$

where $||\cdot||_2$ represents the $L2$ norm. Then, the encoder $\mathcal{F}_\theta(\cdot)$ can be trained in supervised contrastive learning way [37] based on the constructed contrastive sample set $\mathcal{D}$. Given a randomly sampled batch, $\mathcal{I}$ denotes the sample indexes. Setting x_i as the anchor sample, $\mathcal{P}(i):=\{j \in \mathcal{I} \mid y_j = y_i,\ j \neq i\}$ is the index set of all augmented positive samples that are with the same supervisory signal y_i of the anchor sample x_i. Naturally, the augmented negative samples are these with different y_i values for x_i. The index set of these negative samples is defined as $\mathcal{N}(i):=\{j \in \mathcal{I} \mid y_j \neq y_i,\ j \neq i\}$. To optimize the $\mathcal{F}_\theta(\cdot)$ to minimize the distance between positive samples and maximize the distance between positive and negative samples in $\mathcal{M}$, the inter-sample contrastive loss $\mathcal{L}_{con}^{inter}$ is defined as:

$$\mathcal{L}_{con}^{inter}(v_i, v_j, v_k) = \frac{-1}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} \mathcal{L}_{con}^{inter,i}(v_i, v_j, v_k), \tag{8}$$

$$\mathcal{L}_{con}^{inter,i}(v_i, v_j, v_k) = \frac{-1}{|\mathcal{P}(i)|} \sum_{j \in \mathcal{P}(i)} \log \frac{\exp(\mathcal{S}(v_i, v_j)/\tau)}{\sum_{k \in \mathcal{I} \setminus i} \exp(\mathcal{S}(v_i, v_k)/\tau)}. \tag{9}$$

where $|\cdot|$ denotes the element number of the given set, $\tau \in \mathbb{R}^+$ is a temperature scaling parameter. To constrain that positive and negative samples are discriminated strictly according to the pathological asymmetries, the intra-sample contrastive loss $\mathcal{L}_{con}^{intra}$ is incorporated to provide extra constraints when optimizing the encoder $\mathcal{F}_\theta(\cdot)$. $\mathcal{L}_{con}^{intra}$ is motivated based on the fact that the feature distributions of the paired regions in bilateral hemispheres of a given sample x_i is dissimilar when the pathological asymmetry exists. $\mathcal{L}_{con}^{intra}$ is defined as:

$$\mathcal{L}_{con}^{intra}(z_i, z_j) = \frac{1}{|\mathcal{I}|-1}\Big(\sum_{i\in\mathcal{P}(i)} D_{KL}(z_i^l||z_i^r) - \sum_{j\in\mathcal{N}(i)} D_{KL}(z_j^l||z_j^r)\Big), \tag{10}$$

$$D_{KL}(z_i^l||z_i^r) = \sum_{k=1}^{N} z_i^l(k) log\frac{z_i^l(k)}{z_i^r(k)}, \tag{11}$$

$$[z_i^l|z_i^r] = [z_i] = \sigma(\mathcal{F}_\theta(x_j)), \tag{12}$$

$$[z_i^l|z_i^r] = \left[\begin{array}{ccc|ccc} z_{i,1,1}\cdots z_{i,1,p} & & & z_{i,1,p+1}\cdots z_{i,1,2p} & & \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ z_{i,q,1}\cdots z_{i,q,p} & & & z_{i,q,p+1}\cdots z_{i,q,2p} & & \end{array}\right], \tag{13}$$

where $D_{KL}(\cdot||\cdot)$ computes the Kullback–Leibler divergence [38] of given two feature distributions, and $\sigma(\cdot)$ is sigmoid activation function. The total loss including inter-sample and intra-sample contrastive training for optimizing $\mathcal{F}_\theta(\cdot)$ is defined as:

$$\mathcal{L}_{KP} = \lambda_1\mathcal{L}_{con}^{inter} + \lambda_2\mathcal{L}_{con}^{intra}. \tag{14}$$

where λ_1 and λ_2 are weight factors of inter- and intra-sample contrastive losses respectively. Finally, the encoder $\mathcal{F}_\theta(\cdot)$ can gain the discriminative representation for pathological asymmetry by optizing $\mathcal{L}_{KP}$.

D. Formal Learning Stage

As depicted in Fig.2(2), the pre-trained encoder $\mathcal{F}_\theta(\cdot)$ with learned knowledge representations is transferred into the formal learning stage to enable the segmentation network $\mathcal{F}$ including the encoder $\mathcal{F}_\theta(\cdot)$ and decoder $\mathcal{F}_\phi(\cdot)$ possess discriminative ability for pathological asymmetries caused by AIS infarcts. Guided by gained discriminative ability, $\mathcal{F}$ efficiently performs formal end-to-end training. In this stage, the hybrid segmentation loss $\mathcal{L}_{seg}$ including Dice Loss [39] and Cross-Entropy Loss is employed to train $\mathcal{F}$ based on slice level. However, the conventional slice-based training approaches [20, 40, 41] usually fail to fully excavate the anatomical context information inside the volumetric medical images. Inspired by this disadvantage, the feature compensation module (FCM) is designed based on the fact that within a given NCCT scan, slices that are closer in the spatial position (ie., the smaller distance intervals along the sagittal plane direction) should contain more similar anatomical structures. Furthermore, the extracted features from these adjacent slices also should be more similar when employing the same CNN. Therefore, FCM measures and then learns the similarity between adjacent slices, which can compensate the deep image features of adjacent slices to the same slice for aggregating rich anatomical context information. As described in Fig.2(2), the three encoders $\mathcal{F}_\theta(\cdot)$ with shared weights take as input three adjacent slices x_i^s, x_j^s and x_k^s from the same NCCT scan s and predict the high-dimensional deep features output by the encoder $\mathcal{F}_\theta$, where $i<j<k$. In each training batch, each three adjacent slices including front slice x_i^s, middle slice x_j^s and rear slice x_k^s is obtained using three randomly slice-level sampling from different positions along the sagittal plane direction of the same NCCT scan. Then the feature compensation loss $\mathcal{L}_{fc}$ is defined to enforce the anatomical information similarity of these central features from adjacent slices:

$$\mathcal{L}_{fc}(x_i^s, x_j^s, x_k^s) = \frac{1}{N}\sum_{i=1}^{N}\Big(\gamma_1\mathcal{L}_{fc}^{sub}(x_i^s, x_j^s) + \gamma_2\mathcal{L}_{fc}^{sub}(x_j^s, x_k^s) + \gamma_3\mathcal{L}_{fc}^{sub}(x_i^s, x_k^s)\Big), \tag{15}$$

$$\mathcal{L}_{fc}^{sub}(x_i^s, x_j^s) = ||\mathcal{F}_\theta(x_i^s) - \mathcal{F}_\theta(x_j^s)||_1 + ||\mathcal{F}_\theta(x_i^s) - \mathcal{F}_\theta(x_j^s)||_2 \tag{16}$$

$$\gamma_1 = \frac{1}{j-i} - \frac{1}{N}, \gamma_2 = \frac{1}{k-j} - \frac{1}{N}, \gamma_3 = \frac{1}{k-i} - \frac{1}{N}, \tag{17}$$

where $||\cdot||_1$ represents the $L1$ norm, and N denotes the number of the slices of the given NCCT scan s. Here γ_1, γ_2 and γ_3 controls adaptively the weight of the similarity according to the corresponding the distance interval between each two slices among three sampled slices x_i^s, x_j^s and x_k^s. When the distance interval of two slices is smaller, γ is larger. This indicates that the anatomical context information provided by the adjacent slices is more meaningful and should have a stronger influence on the feature compensation. Conversely, when the distance interval is larger, γ is smaller. In this case, the FCM does not impose constraints on the far-away adjacent slices. This is because the anatomical context information they provide tends to be discontinuous and may not contribute significantly to the feature compensation. In this way, FCM compensates the deep image features of adjacent front and rear slices x_i^s,x_k^s to the middle slice x_j^s to aggregate rich anatomical context information. As shown in Fig.2(2), only the extracted features from the middle slice x_j^s, which contain the richer anatomical context information, are forwarded to the decoder $\mathcal{F}_\phi(\cdot)$ for participating in the calculation of the segmentation loss and the next stage.

E. Examination Improvement Stage

As depicted in Fig.2(3), in the examination improvement stage of the PAPL, the infarct prediction from the formal learning stage is examined and improved gradually. Intuitively, the optimization of the segmentation network $\mathcal{F}$ is constrained by the designed regional perception refinement strategy (RPRS) in each training batch. Given a slice-label pair (x_i^s,y_i^s), the infarct prediction M_{pr} for the infarcts (i.e., the predictive probability map) output by the decoder $\mathcal{F}_\phi(\cdot)$ can be formulated as:

$$M_{pr} = softmax(\omega\mathcal{F}_\phi(\mathcal{F}_\theta(x_i^s))), \tag{18}$$

where $\omega \in (0, +\infty)$ controls how much uncertainty is permitted. The uncertainty is very low when $\omega \to +\infty$ and the values are closer to either 0 or 1. To further refine the infarct prediction M_{pr}, RPRS is utilized to guide the regional shrink and expansion of M_{pr} while taking full consideration of the different saliences of the hypo-attenuation changes of AIS infarcts. Then RPRS can be integrated into the optimization of the segmentation network $\mathcal{F}$ in the form of a loss function $\mathcal{L}_{rprs}$:

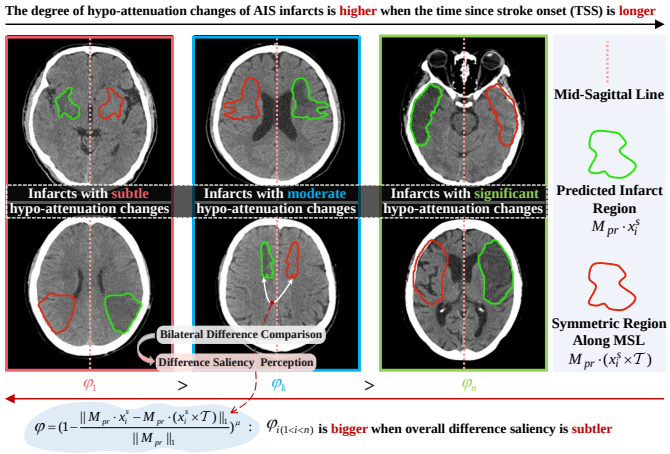


Fig. 5: Evaluating the difference saliency degree φ in infarct-related bilateral region intensity comparison.

$$\mathcal{L}_{rprs}(x_i^s, y_i^s, M_{pr}) = \frac{1}{M} \sum_{i=1}^{M}\Big(-\varphi\mathcal{L}_{rprs}^1(x_i^s, M_{pr}) + \mathcal{L}_{rprs}^2(y_i^s, M_{pr}) + \mathcal{L}_{rprs}^3(y_i^s, M_{pr})\Big), \tag{19}$$

where $\mathcal{L}_{rprs}$ is the overall optimizing objective of the RPRS, which consists of three sub-constraint terms $\mathcal{L}_{rprs}^1$, $\mathcal{L}_{rprs}^2$, $\mathcal{L}_{rprs}^3$. $\mathcal{L}_{rprs}^1$ encourages M_{pr} to expand adaptively to further discover unpredicted infarct regions by performing the infarct-related bilateral region difference comparison while retaining the truly predicted infarct regions in M_{pr}, which helps reduce the false negative prediction. As described in Fig.5, this comparison is conducted by calculating the overall signal intensity difference of infarct-related bilateral image regions including the predicted infarct region and its symmetric region along MSL, in which these regions can be extracted according to M_{pr}. The difference value calculated by $\mathcal{L}_{rprs}^1$ will be further increased when M_{pr} expands and includes extra new infarcts. On the contrary, when M_{pr} expands but does not include new infarcts, the value of $\mathcal{L}_{rprs}^1$ almost will be not affected when containing the intensity difference of new predicted infarct-related bilateral image regions. However, this situation may introduce additional error predictions. For this issue, $\mathcal{L}_{rprs}^2$ is used to limit the region expansion range of M_{pr} to the false negative region of current predicted results, which aims to avoid introducing these extra errors during region expansion. Besides, $\mathcal{L}_{rprs}^3$ is utilized to encourage M_{pr} shrinking by controlling the false positive prediction in M_{pr}. $\mathcal{L}_{rprs}^1$, $\mathcal{L}_{rprs}^2$ and $\mathcal{L}_{rprs}^3$ are formulated as respectively:

$$\varphi = (1 - \frac{||M_{pr} \cdot x_i^s - M_{pr} \cdot (x_i^s \times \mathcal{T})||_1}{||M_{pr}||_1})^{\mu}, \tag{20}$$

$$\mathcal{L}_{rprs}^1(x_i^s, M_{pr}) = ||M_{pr} \cdot x_i^s - M_{pr} \cdot (x_i^s \times \mathcal{T})||_1, \tag{21}$$

$$\mathcal{L}_{rprs}^2(y_i^s, M_{pr}) = ||M_{fn}||_1 = ||y_i^s - M_{pr} \cdot y_i^s||_1, \tag{22}$$

$$\mathcal{L}_{rprs}^3(y_i^s, M_{pr}) = ||M_{fp}||_1 = ||M_{pr} - M_{pr} \cdot y_i^s||_1, \tag{23}$$

where $\mathcal{T}$ is the mirror transformation matrix with the same size as M_{pr}. μ is adjustable factor. The false positive map M_{fp} and the false negative map M_{fn} can be calculated using the label y_i^s. Especially, φ is designed to perceive the overall difference saliency of infarct-related bilateral regions when performing the bilateral difference comparison. The motivation for setting is based on the fact that the longer time since stroke onset (TSS) is associated with a higher degree of hypo-attenuation changes of ischemic tissue, leading to more significant contrast differences of AIS infarcts in NCCT images (As shown in from-left-to-right images of Fig.5). Naturally, φ is employed to perceive the difference saliency of infarct-related bilateral regions and adaptively control the weights of different training samples according to the hypo-attenuation change degree of AIS infarcts during the region expansion, which facilitates mining these more difficult samples with subtle hypo-attenuation changes gradually. Therefore, the mis-segmentation in M_{pr} can be improved gradually by optimizing $\mathcal{L}_{rprs}$.

Finally, integrating FCM and RPRS, the total training loss of $\mathcal{F}$ is formulated, implementing joint training for the formal learning stage and examination improvement stage of the PAPL:

$$\mathcal{L}_{total} = \beta_1\mathcal{L}_{seg} + \beta_2\mathcal{L}_{fc} + \beta_3\mathcal{L}_{rprs}. \tag{24}$$

where β_1, β_2 and β_3 are the weight factors of the segmentation, the FCM and the RPRS terms respectively.

IV. Experiment Setup

A. Datasets

The effectiveness of the proposed PAPL is evaluated based on the following two NCCT datasets of acute ischemic stroke.

*1) **AISD**:* The public dataset AISD [4] consists of 397 NCCT scans of AIS patients, which contains 345 training-validation scans and 52 testing scans. The interval from stroke symptom onset to NCCT imaging is within 24 hours. After NCCT imaging, the corresponding diffusion-weighted MRI (DWI) of the stroke patients within 24 hours is the reference standard, and then the doctors manually annotate infarcts on NCCT.

*2) **XWHD**:* The in-house dataset XWHD consists of 171 NCCT scans of AIS patients, which contains 130 training-validation scans and 41 testing scans. XWHD was retrospectively collected from Xuanwu Hospital, Capital Medical University, which meets the following multiple inclusion criteria: (a) The interval from stroke symptom onset to NCCT imaging is within 24 hours and DWI is within 72 hours after taking the NCCT. (b) The AIS patients with large vessel occlusions in the anterior circulation. (c) The patients with other brain diseases (such as chronic infarct, white matter degeneration, brain tumor, and intracranial hemorrhage) were excluded. (d) Ineligible imaging due to head movement or artifacts was excluded. As shown in Table I, National Institutes of Health Stroke Scale (NIHSS) is a standardized assessment tool used to comprehensively evaluate the severity of stroke, which assesses the neurological deficits in AIS patients through 11 scoring items. The NIHSS scoring range is from 0 to 42 points, where a higher score indicates more severe neurological deficits. The time since stroke onset (TSS) represents the time interval from stroke onset to the completion of NCCT imaging. The visibility of the early pathological hypo-attenuation changes is significantly associated with TSS. Lesion location describes the specific location of AIS lesions in the bilateral brain hemisphere. Therefore, the characteristics

TABLE I
Statistical characteristics distribution of AIS patients on XWHD. The Age and TSS are reported in the form of the (median,quartile). The NIHSS is reported in the form of the mean±standard deviation.

Characteristics	Value
Demographic	
Patients	171
Male	121
Female	50
Age	65(54,73)
Clinical indicators	
TSS	6.02(4.07,9.22)
NIHSS	12.40±5.69
Lesion location	
left	84
right	87

distribution of enrolled AIS patients is diverse, which can be as close to the real clinical assessment scenario as possible. Finally, the follow-up DWI imaging after reperfusion therapy was used as a reference standard to manually help annotate infarcts on NCCT from three senior clinical experts.

B. Implementation Details

The proposed PAPL is implemented based on PyTorch 1.8.0 and works on NVIDIA GeForce RTX 3090 GPUs with CUDA 11.4. The Adam optimizer is employed for parameter optimization. In stage 1, the learning rate, the batch size, the training epoch, and the temperature scaling parameter are set to $1 \times e^{-4}$, 256, 200 and 0.1 respectively. Besides, the weight factors λ_1 and λ_2 in inter- and intra-sample contrastive losses are set to 0.6 and 0.3. For the region mask sampling strategy, the sampling step follows a uniform distribution, and for each given slice: (1) the width and height of each patch region are set to $(10, 70)$ and $(10, 80)$, respectively. (2) the sampling range (the start and end in the x/y direction) is determined by the mask range of the whole brain regions. (3) the sampling strategy is performed 8 times and generates 8 contrastive samples. Among them, one specific sampling that can cover the whole infarct region is included to ensure the generation of positive samples. In stages 2 and 3, the learning rate, the batch size, and the training epoch are set to $3 \times e^{-4}$, 32, 180, respectively. Besides, the weight factors β_1, β_2 and β_3 in $\mathcal{L}_{total}$ are initialized to 0.8, 0.8 and 0.2. Especially, considering the fact that the segmentation quality is not good enough at the beginning of training, the weight factor β_3 gradually increases and the increase coefficient is set to 1.2 per 20 epochs with the improvement of the segmentation quality.

For the data pre-processing on AISD and XWHD datasets, the raw intensity values first are cropped to the range [30, 100] Hounsfield units (HU). Then, these raw NCCT scans are normalized in the z-score way, and finally resampled to $256 \times 256 \times N$, where N is the original number of the slices.

C. Comparison Settings

1) ***Backbone Settings****:* In the experiments, the U-Net [42] networks consisting of the encoder $\mathcal{F}_\theta(\cdot)$ and the decoder $\mathcal{F}_\phi(\cdot)$ are respectively employed as the segmentation backbone to evaluate the effectiveness of the proposed PAPL.

2) ***Comparison Methods****:* The proposed PAPL is compared with the 12 state-of-the-art (SOTA) segmentation methods, including U-Net [42], CENet [44], Attention U-Net [43], U-Net++ [45], TransU-Net [46], ISP-Net [17], X-Net [15], CLCI-Net [47], SEAN [4], ADN [5], SAN-Net [22] and IS-Net [23]. The first four methods achieved satisfactory results on medical image segmentation tasks in the past few years. TransU-Net [46] is a Transformer-based segmentation method that captures rich global information from the images. ISP-Net [17], X-Net [15] and CLCI-Net [47] are designed for stroke lesion segmentation. SEAN [4], ADN [5], SAN-Net [22] and IS-Net [23] are brain symmetric modeling-based SOTA methods that are designed for AIS infarct segmentation. For the aforementioned methods, we follow the source code released by the author. Besides, all methods follow the same pre-processing steps for a fair comparison.

3) ***Evaluation Metrics****:* To quantitatively evaluate the segmentation performance of different methods and demonstrate the advantages of the proposed PAPL, the relevant experiments are performed via multiple volume-based evaluation metrics [48] including Dice similarity coefficient (DSC), Jaccard Coefficient(JC), Average Symmetric Surface Distance (ASSD), Precision (PRE), and Sensitivity (SEN).

V. Experiment Results

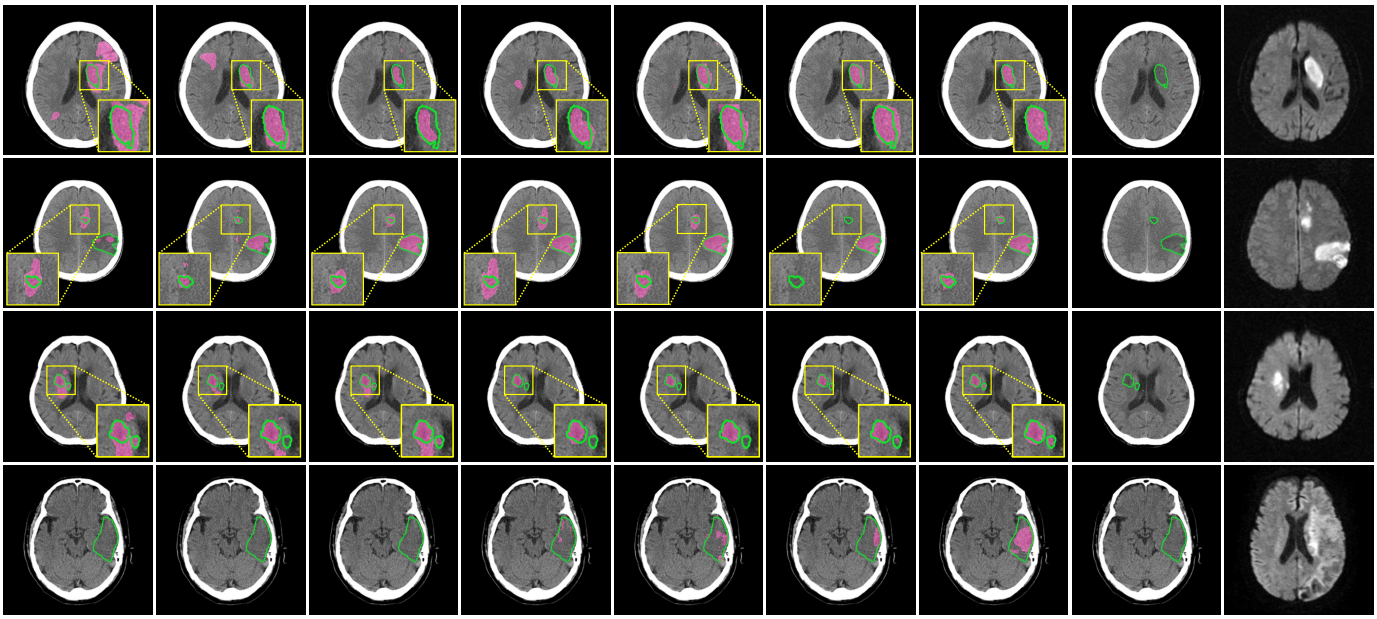
A. Segmentation Comparison With the State-of-the-art Methods

Table II gives the quantitative comparison of the proposed PAPL and different SOTA segmentation methods on the AISD and XWHD datasets. The best results are highlighted in boldface. General medical image segmentation methods including U-Net [42], CENet [44], Attention U-Net [43] and U-Net++ [45] are difficult to capture subtle hypo-attenuation changes caused by the AIS infarct, which only achieve poor segmentation performance. TransU-Net [46] can provide long-distance dependent information to help perceive the bilateral difference of the image. Therefore, the TransU-Net [46] can achieve better performance compared to the aforementioned methods. The ablation studies for the aligning process show that these segmentation methods including U-Net [42], CENet [44] and U-Net++ [45] and TransU-Net [46] can gain further performance improvement when the NCCT scans are aligned to standard symmetric space. One reason is that the aligning process reduces the differences in the spatial distribution between training images and test images and makes the spatial distributions between them closer. Besides, these stroke lesion-specific segmentation methods including ISP-Net [17], X-Net [15] and CLCI-Net [47] can more effectively extract the hypo-attenuation features of AIS infarcts and achieve better segmentation performance. Especially, these brain symmetric modeling-based methods including SEAN [4], ADN [5], SAN-Net [22] and IS-Net [23] can gain further improvement by incorporating the domain knowledge of the brain hemispheres into the deep networks. For our PAPL, it can suppress undesirable effects for the infarct segmentation from the intrinsic asymmetries caused by normal physiological changes. Therefore, the proposed PAPL achieved the best segmentation performance of 0.5468 and 0.4899 DSCs on the AISD and XWHD datasets respectively.

TABLE II

Performance comparison between other methods and our PAPL on the AISD and XWHD datasets. † means stroke lesion-specific segmentation methods and ⋆ means brain symmetric modeling-based methods. The best and second performances are in bold and underlined respectively.

Method	Aligned	AISD					XWHD				
		DSC ↑	JC ↑	ASSD ↓	PRE ↑	SEN ↑	DSC ↑	JC ↑	ASSD ↓	PRE ↑	SEN ↑
Comparative BaseLines											
U-Net [42]	×	0.4578	0.3307	8.3910	0.5181	0.4799	0.4240	0.2927	9.5496	0.5461	0.3744
Attention U-Net [43]	×	0.4710	0.3389	6.8993	0.5126	0.5088	0.4440	0.3151	8.6390	0.5336	0.4260
CENet [44]	×	0.4745	0.3446	6.2368	0.6109	0.4655	0.4355	0.2966	7.3242	0.4511	**0.5297**
U-Net++ [45]	×	0.4631	0.3274	7.6546	0.4710	0.5152	0.4542	0.3222	5.6319	0.5196	0.4652
TransU-Net [46]	×	0.4792	0.3508	6.7035	0.5719	0.4824	0.4399	0.3033	5.6246	0.4922	0.4851
ISP-Net† [17]	✓	0.4844	0.3453	6.2874	0.5767	0.4662	0.4737	0.3296	5.3786	0.5050	0.5100
X-Net† [15]	✓	0.4947	0.3595	7.0283	0.5740	0.4990	0.4603	0.3209	5.5470	0.4961	0.4825
CLCI-Net† [47]	✓	0.5068	0.3696	4.9309	0.6554	0.4606	0.4754	0.3361	6.1644	0.5234	0.5011
SEAN†⋆ [4]	✓	0.5047	-	-	-	-	-	-	-	-	-
ADN†⋆ [5]	✓	0.5245	-	-	-	-	-	-	-	-	-
SAN-Net†⋆ [22]	✓	0.5089	0.3708	4.7585	0.6099	0.4992	0.4788	0.3396	5.0202	0.5564	0.5262
IS-Net†⋆ [23]	✓	0.5124	0.3767	5.4794	0.6260	0.4935	0.4685	0.3325	5.2903	0.6076	0.4441
U-Net [42]	✓	0.4866	0.3538	6.7984	0.6103	0.4670	0.4354	0.3037	8.0282	0.4830	0.4828
CENet [44]	✓	0.4975	0.3648	6.5559	0.6238	0.4828	0.4421	0.3090	5.8224	0.5366	0.4284
U-Net++ [45]	✓	0.4923	0.3605	6.3707	0.5969	0.4921	0.4709	0.3302	6.5311	0.6135	0.4222
TransU-Net [46]	✓	0.5045	0.3740	5.3703	0.6966	0.4675	0.4587	0.3192	5.5921	0.5475	0.4604
PAPL†⋆($ours$)	✓	**0.5468**	**0.4080**	**4.7404**	**0.6985**	**0.5154**	**0.4899**	**0.3541**	**4.6173**	**0.6204**	0.4923



(a) U-Net (b) U-Net++ (c) TransU-Net (d) ISP-Net (e) X-Net (f) CLCI-Net (g) PAPL (h) Ground Truth (i) DWI

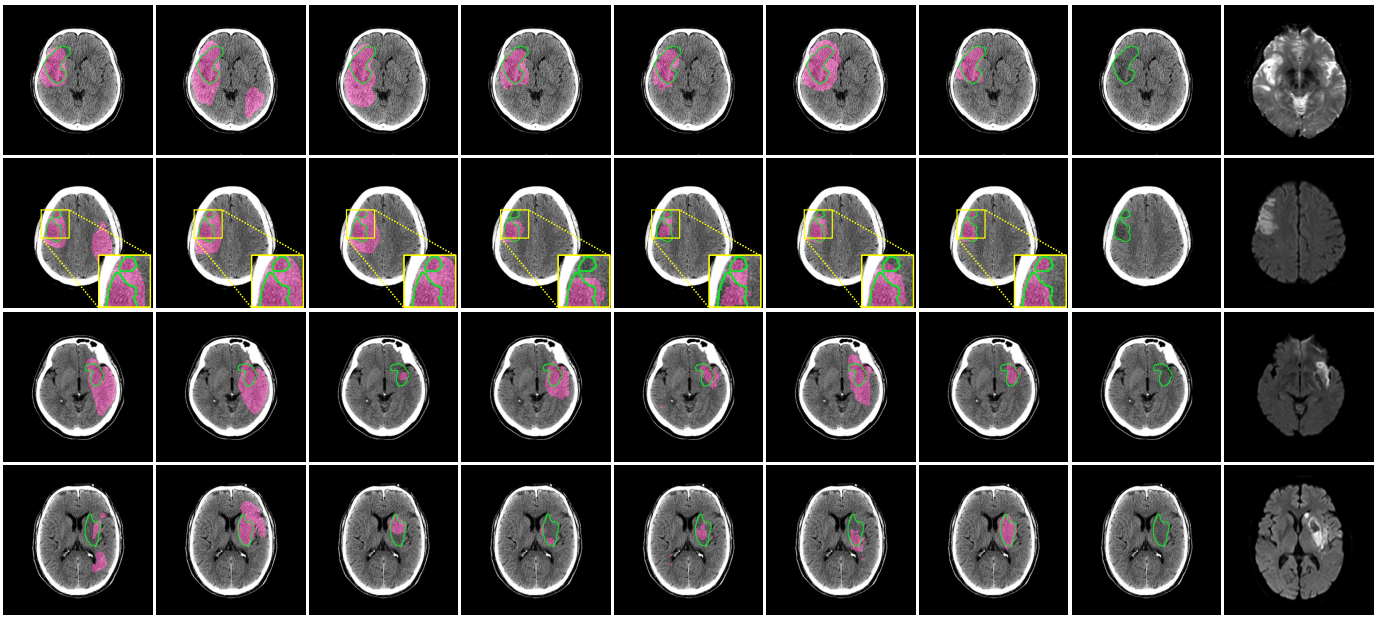
Fig. 6: Visualizing segmentation comparison on AISD dataset.

Fig.6 and Fig.7 visualize the segmentation results of the different methods on the AISD and the XWHD. It can be easily seen that the predicted results from our PAPL are more consistent with infarct lesion boundaries of ground truth than other methods. Especially, as shown in the 2nd row of Fig.6 and 3rd row of Fig.7, the PAPL achieves superior segmentation for small infarct lesions. Besides, as shown in the 4th row of Fig.6 and Fig.7, the PAPL exhibits higher sensitivity for the AIS infarcts with extremely subtle hypo-attenuation changes.

B. Ablation Studies of Proposed PAPL

(1) Alter Different Backbones. The proposed PAPL can be conveniently integrated into other encoder-decoder-based segmentation backbones, making it highly applicable. Table III gives the related quantitative results on the AISD and the XWHD datasets, when different backbones including U-Net [42], CENet [44], U-Net++ [45] are equipped with the PAPL. Although the some metrics of some backbones decreased slightly, overall performance is improved. These results support the generality that modeling reasonably pathological symmetries of brain hemispheres can facilitate infarct segmentation.

(2) Combine With Different Submodules. To quantitatively analyze the effectiveness of each submodule in three stages of the PAPL, different submodules are combined for ablation analysis. Table IV lists the DSC scores on the AISD dataset when combining different submodules. It can be seen that it



(a) U-Net (b) U-Net++ (c) TransU-Net (d) ISP-Net (e) X-Net (f) CLCI-Net (g) PAPL (h) Ground Truth (i) DWI

Fig. 7: Visualizing segmentation comparison on XWHD dataset.

TABLE III
Segmentation performance on the AISD and XWHD datasets when altering backbones.

Method	AISD					XWHD				
	DSC ↑	JC ↑	ASSD ↓	PRE ↑	SEN ↑	DSC ↑	JC ↑	ASSD ↓	PRE ↑	SEN ↑
U-Net [42] (w/o PAPL)	0.4866	0.3538	6.7984	0.6103	0.4670	0.4354	0.3037	8.0282	0.4830	0.4828
U-Net [42] (w/ PAPL)	0.5468	0.4080	4.7404	**0.6985**	0.5159	0.4899	0.3541	4.6173	**0.6204**	0.4923
CENet [44] (w/o PAPL)	0.4975	0.3648	6.5559	0.6238	0.4828	0.4421	0.3090	5.8224	0.5366	0.4284
CENet [44] (w/ PAPL)	**0.5488**	**0.4086**	**4.7126**	0.6116	**0.5597**	0.4720	0.3335	6.5567	0.5282	0.4901
U-Net++ [45] (w/o PAPL)	0.4923	0.3605	6.3707	0.5969	0.4921	0.4709	0.3302	6.5311	0.6135	0.4222
U-Net++ [45] (w/ PAPL)	0.5474	0.4085	4.8079	0.6978	0.5053	**0.4970**	**0.3549**	**4.4108**	0.5145	**0.5420**

TABLE IV
Ablation analysis of different submodules on the AISD dataset.

Stage 1		Stage 2	Stage 3	DSC ↑
Inter-Contrastive	Intra-Contrastive	FCM	RPRS	
				0.4866
✓				0.4960
✓	✓			0.5188
		✓		0.4975
			✓	0.5015
		✓	✓	0.5102
✓	✓	✓		0.5285
✓	✓		✓	0.5400
✓	✓	✓	✓	0.5468

can gain higher DSC scores than the baseline when combining anyone in three submodules and the combination of all submodules achieves the best results. Especially, the segmentation improvement is subtle when only considering the inter-sample contrastive training. The improvement increases significantly when considering both the inter- and intra-sample contrastive training. One reason is that the intra-sample contrastive training can provide clear evidence and constraint of learning the relevant discriminative knowledge for pathological asymmetry, which makes the learned representations more robust and gains stable segmentation improvement.

C. Effectiveness Analysis of Different Stage in Proposed PAPL

(1) Effectiveness Analysis of Knowledge Preparation Stage. To analyze the effectiveness of the knowledge preparation stage, the t-Distributed Stochastic Neighbor Embedding (t-SNE) method [49] is employed to visualize the learned knowledge representations. As visualized in Fig.8, the representations of the positive and negative samples generated by the randomly initialized network are mixed in the latent space. As the training progresses, these samples with similar representations are attracted together and the samples with dissimilar representations are repelled away. Finally, with inter- and intra-sample contrastive training, the representations for the same class are tightly clustered and those for different classes are also clearly separated, while some hard samples are indistinguishable only using inter-sample contrastive training. This indicates that the knowledge preparation stage can help the network gain the discriminative representation for the pathological asymmetry. This aligns with the results in Table IV where integrating stage 1 achieves the higher DSC scores (1st and 3rd rows).

(2) Effectiveness Analysis of Formal Learning Stage. To analyze the effectiveness of the formal learning stage, the segmentation results for the adjacent slices are visualized. As described in Fig.9, the continuity and consistency of segmentation results (1st column) for the adjacent successive

TABLE V
Performance comparison between different methods for small (≤70 ml) and big volume (>70 ml) infarcts on the AISD dataset.

Method	Small-volume infarct (≤70ml)				Big-volume infarct (>70ml)				Average			
	DSC ↑	JC ↑	ASSD ↓	PRE ↑	DSC ↑	JC ↑	ASSD ↓	PRE ↑	DSC ↑	JC ↑	ASSD ↓	PRE ↑
U-Net [42]	0.4402	0.3097	7.7730	0.5808	0.7548	0.6093	1.1783	0.7725	0.4886	0.3558	6.7574	0.6103
CENet [44]	0.4498	0.3190	7.5875	0.5892	0.7598	0.6168	0.8788	**0.8144**	0.4975	0.3649	6.5544	0.6239
U-Net++ [45]	0.4443	0.3149	7.3636	0.5641	0.7563	0.6115	0.9097	0.7771	0.4923	0.3606	6.3697	0.5969
TransU-Net [46]	0.4430	0.3127	7.6525	0.5863	**0.7726**	**0.6322**	**0.6494**	0.7772	0.4938	0.3619	6.5740	0.6157
PAPL (ours)	**0.5081**	**0.3700**	**5.4625**	**0.6819**	0.7606	0.6171	0.7691	0.7895	**0.5470**	**0.4080**	**4.7397**	**0.6985**

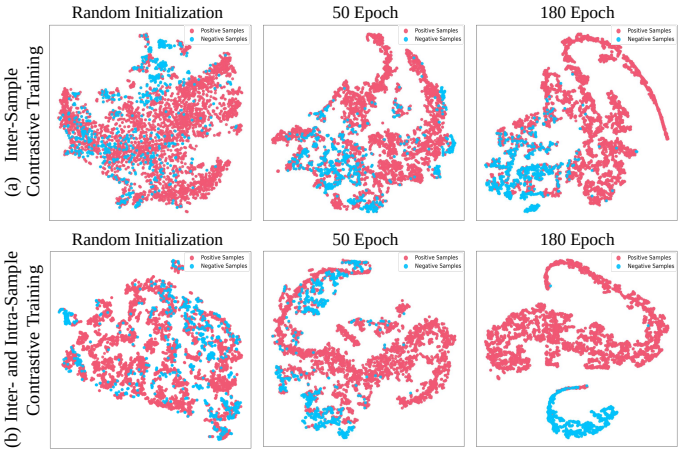


Fig. 8: Visualizing learned knowledge representations in knowledge preparation stage. Red and blue points are embeddings of the positive and negative samples respectively after applying t-SNE

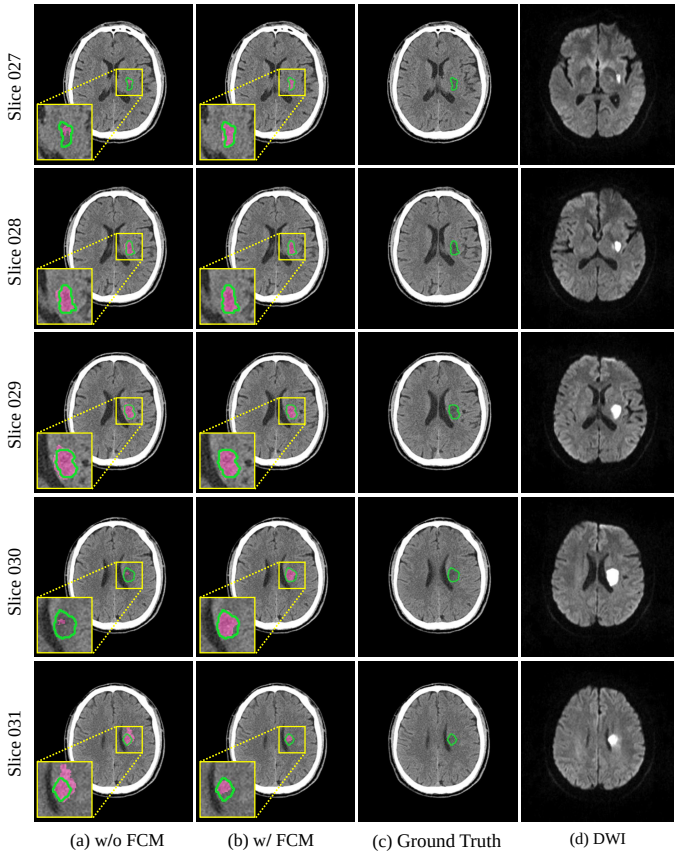


Fig. 9: Visualizing segmentation for the adjacent slices.

slices are poor when not employing the FCM in the formal learning stage. Then, the continuity and consistency improve significantly when employing the FCM (2nd column). This indicates that FCM can leverage the anatomy similarity between adjacent slices to aggregate rich anatomical context information, helping get better segmentation continuity and consistency. One reason is that FCM increases the explorable anatomical space, which is akin to expanding the computational receptive field of the feature extraction process along the sagittal plane direction.

(3) Effectiveness Analysis of Examination Improvement Stage. To analyze the effectiveness of the examination improvement stage, the feature maps output by different positions of the decoder and related prediction results including prediction probability map, true positive map, false positive map, and false negative map are given. As described in Fig.10, integrating RPRS can help gradually refine extracted deep features when epochs increase, which effectively mitigates mis-segmentation (yellow arrow in 7th and 8th columns) including the false infarct and the normal tissue predictions. By contrast, the mis-segmentations (5th and 6th columns) are not further improved when not employing the RPRS. This indicates that the RPRS can help further refine the infarct predictions.

D. Performance Evaluation on Different Volume Infarcts

(1) Quantitative Evaluation on Different Volume Infarcts. Accurate estimation of the infarct volume is crucial for the effective AIS treatment. In clinical practice, the 70 ml cutoff is one of the important basis for selecting the appropriate treatment decision including tissue plasminogen activator (t-PA) thrombolysis and mechanical thrombectomy, etc [50–52]. Therefore, setting 70 ml as the cutoff in the infarct volume estimation to evaluate the performance of small- (≤70 ml and big-volume (>70 ml) infarcts is important clinical significance. Table V gives the quantitative comparison between different methods for the small- and big-volume infarcts on the AISD dataset. For small-volume infarcts, the PAPL outperforms other methods (around 6% DSC and 10% PRE improvement). For big-volume infarcts, the PAPL is closer to the best performance (around 1% DSC and 2% PRE gap). Further, the proposed PAPL obtains the superior average performance for both large- and small-volume infarcts. These results indicate that the PAPL can achieve better quantitative estimation for both the small- and big-volume infarcts.

(2) Consistency Evaluation For Infarct Volume Estimation. The infarct volume of the AIS patients can be estimated by combining the segmentation results and real imaging spacing of the NCCT. Based on this, to further evaluate the estimation

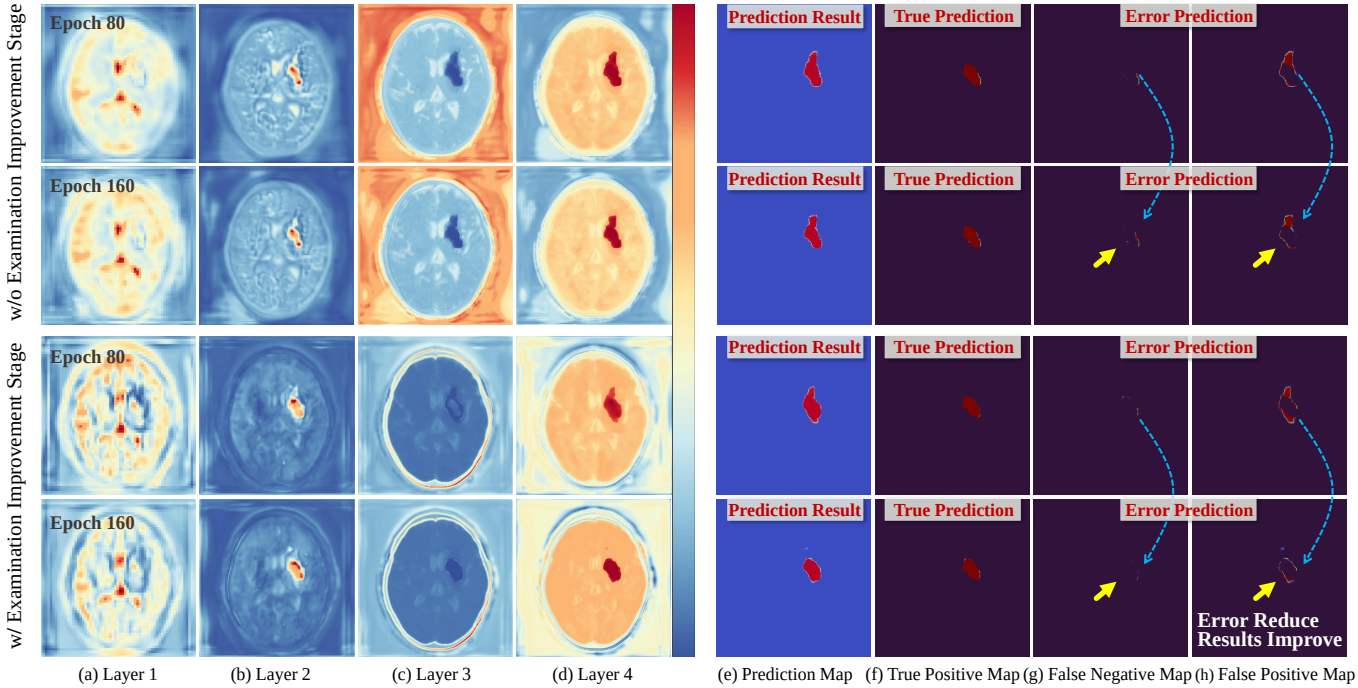


Fig. 10: Visualizing feature maps and predicted results from different training epochs.

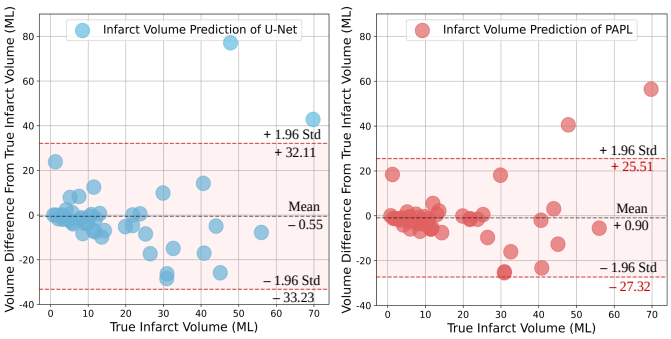


Fig. 11: The Bland-Altman plot of backbone and proposed methods for small-volume infarcts.

performance on small-volume infarcts, the Bland-Altman plot is calculated to analyze the difference between the real and estimated infarct volume. As described in Fig.11, the calculated difference of our PAPL for most cases is in 95% confidence interval (CI), and the range of the CI is smaller compared to the backbone method. Therefore, for volume estimation of small-volume infarcts, our PAPL has better predictive stability and higher consistency between the estimated and real infarcts.

E. Performance Comparison with Other Self-Supervised CL Methods

To validate the advantages of the proposed CL method of the Knowledge Preparation Stage on learning domain-specific representation, a performance comparison is conducted between the proposed method and existing advanced self-supervised CL methods [28, 29] on the AISD dataset. Specifically, by applying different self-supervised CL methods in the Knowledge Preparation Stage and then disabling the other last two stages of PAPL, the performance of these methods alone on the AIS infarct segmentation task is compared. As shown in Table VI, compared with other methods, the proposed method can better facilitate AIS infarct segmentation. This indicates the learned domain-specific knowledge representation using the proposed CL method can enhance the discriminative ability for pathological asymmetries and efficiently improve AIS infarct segmentation while methods [28, 29] may not have the ability.

TABLE VI
Segmentation performance when altering different self-supervised CL methods in Knowledge Preparation Stage.

Knowledge Preparation Stage			DSC ↑
PCL [28]	GCL [29]	Ours	
			0.4866
✓			0.5012
	✓		0.4943
		✓	0.5188

F. Effects of weight parameter settings in proposed PAPL

To explore the influence of the weight parameters in these hybrid loss functions of Eq.(14) and Eq.(24) on segmentation performance and further ascertain appropriate parameter values, different coefficient combinations of these weight parameters respectively are set for ablation analysis.

The effect of the λ_1 and λ_2. The experiments in TABLE VII investigate the performance under multiple weight coefficient combinations in Eq.(14) including setting λ_1 to 1, 0.8, 0.6, 0.5 and 0.4, and setting λ_2 to 0.1, 0.2, 0.3, 0.5 and 0.8. The results show that it can gain the highest DSC scores when λ_1 is set to 0.6, and λ_2 is set to 0.3. Therefore, λ_1 and λ_2 are empirically set to 0.6 and 0.3 respectively in the work.

The effect of the β_1, β_2 and β_3. The experiments in TABLE VIII investigate the performance under multiple weight

TABLE VII
Ablation analysis of weight parameter λ_1 and λ_2 settings.

λ_1	λ_2	DSC ↑
1	0.1	0.4983
0.8	0.2	0.5094
0.6	0.3	**0.5188**
0.5	0.5	0.5143
0.4	0.8	0.4935

TABLE VIII
Ablation analysis of weight parameter β_1, β_2 and β_3 settings.

β_1	β_2	β_3	DSC ↑
1	0.5	0	0.5213
0.9	0.6	0	0.5156
0.8	0.8	0	0.5285
0.6	0.9	0	0.5227
0.5	1	0	0.5179
0.8	0.8	0.4	0.5201
0.8	0.8	0.2	**0.5468**
0.8	0.8	0.1	0.5410

coefficient combinations in Eq.(24) including setting β_1 to 1, 0.9, 0.8, 0.6 and 0.5, setting β_2 to 0.5, 0.6, 0.8, 0.9 and 1 and setting β_3 to 0, 0.1, 0.2 and 0.4. In the process, the two-stage screening way is employed. Specifically, the first step involves determining the weights β_1 and β_2, while the second step focuses on the weight β_3. The results show that it can gain the highest DSC scores when β_1 is set to 0.8, β_2 is set to 0.8 and β_3 is set to 0.2. Therefore, β_1, β_2 and β_3 are empirically set to 0.8, 0.8, and 0.2 respectively in the work.

VI. Discussion and Conclusion

The changes of ischemic brain tissue on NCCT are subtle and difficult to identify in the early stage of AIS. Therefore, it remains a very challenging task for early AIS infarct segmentation on NCCT. Current methods generally ignore or confuse the different contributions of intrinsic and pathological asymmetries in the bilateral hemisphere differences to AIS infarct segmentation when integrating this domain knowledge into DL methods. Inspired by this, we propose a pathological asymmetry-guided progressive learning (PAPL) method for facilitating AIS infarct segmentation. PAPL mimics the step-by-step learning patterns observed in humans, including three progressive stages: knowledge preparation stage, formal learning stage, and examination improvement stage. Based on this preview-learn-correct progressive strategy, we can achieve further better AIS infarct segmentation. Extensive experiments on the public AISD and in-house XWHD datasets demonstrated the superiority of the proposed PAPL, which is promising to help better evaluate brain tissue status and predict prognosis. Besides, our method can be easily integrated into other segmentation backbones, making it highly applicable.

Limitation. (1) The proposed MGCL can achieve better segmentation performance for the AIS infarcts. Even so, for these hyperacute infarcts (4th row in Fig.6 and Fig.7), the hypo-attenuation changes are more subtle and our method also does not perform well enough. (2) The proposed method includes three stages and involves a certain number of hyperparameters. When applying to new scenarios or datasets, it may take some effort to adjust and balance these parameters to ensure better performance. (3) Though our study involved multiple NCCT datasets, they were acquired respectively in a single stroke center. Multi-center or cross-center validation can generalize the derived model, hoping to be applied to clinical settings and help better stroke evaluation.

Future work. (1) The proposed PAPL can achieve the best volume estimation performance. However, compared to large-volume infarct segmentation, there is still intractable for small-volume infarct segmentation. In future work, we will explore how to further improve the small-volume infarct segmentation. (2) Compared to single-modal sequence, multi-modal CT-MRI sequence can provide complementary biological information for achieving better disease evaluation and treatment planning [53, 54]. In clinical practice, limited by the urgency of AIS onset and medical conditions, collecting large-scale paired multi-modal sequences is very difficult. In comparison, the single-modality NCCT sequence of AIS patients is easier to obtain due to its quick imaging and cost-effectiveness [55]. Naturally, extracting privileged knowledge from multi-modal sequences and utilizing them to improve further the unimodal segmentation performance when only single-modal NCCT sequence can be accessed, can match current clinical needs well. In future work, we will also conduct the study on the cross-modal knowledge transfer method to further facilitate the segmentation performance on NCCT.

References

[1] C. W. Tsao, A. W. Aday, Z. I. Almarzooq, A. Alonso, A. Z. Beaton, M. S. Bittencourt, A. K. Boehme, A. E. Buxton, A. P. Carson, Y. Commodore-Mensah *et al.*, "Heart disease and stroke statistics—2022 update: a report from the american heart association," *Circulation*, vol. 145, no. 8, pp. e153–e639, 2022.

[2] G. W. Albers, M. P. Marks, S. Kemp, S. Christensen, J. P. Tsai, S. Ortega-Gutierrez, R. A. McTaggart, M. T. Torbey, M. Kim-Tenser, T. Leslie-Mazwi *et al.*, "Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging," *New England Journal of Medicine*, vol. 378, no. 8, pp. 708–718, 2018.

[3] T. Vos, S. S. Lim, C. Abbafati, K. M. Abbas, M. Abbasi, M. Abbasifard, M. Abbasi-Kangevari, H. Abbastabar, F. Abd-Allah, A. Abdelalim *et al.*, "Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019," *The Lancet*, vol. 396, no. 10258, pp. 1204–1222, 2020.

[4] K. Liang, K. Han, X. Li, X. Cheng, Y. Li, Y. Wang, and Y. Yu, "Symmetry-enhanced attention network for acute ischemic infarct segmentation with non-contrast ct images," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VII 24*. Springer, 2021, pp. 432–441.

[5] H. Ni, Y. Xue, K. Wong, J. Volpi, S. T. Wong, J. Z. Wang, and X. Huang, "Asymmetry disentanglement network for interpretable acute ischemic stroke infarct segmentation in non-contrast ct scans," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VIII*. Springer, 2022, pp. 416–426.

[6] P. Vilela, "Acute stroke differential diagnosis: Stroke mimics," *European journal of radiology*, vol. 96, pp. 133–144, 2017.

[7] H. Kuang, B. K. Menon, and W. Qiu, "Automated infarct segmentation from follow-up non-contrast ct scans in patients with acute ischemic stroke using dense multi-path contextual generative adversarial network," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part III 22*. Springer, 2019, pp. 856–863.

[8] A. Marcus, P. Bentley, and D. Rueckert, "Concurrent ischemic lesion age estimation and segmentation of ct brain using a transformer-based network," in *Machine Learning in Clinical Neuroimaging: 5th International Workshop, MLCN 2022, Held in Conjunction with MICCAI 2022, Singapore, September 18, 2022, Proceedings*. Springer, 2022, pp. 52–62.

[9] H. El-Hariri, L. A. S. M. Neto, P. Cimflova, F. Bala, R. Golan, A. Sojoudi, C. Duszynski, I. Elebute, S. H. Mousavi, W. Qiu *et al.*, "Evaluating nnu-net for early ischemic change segmentation on non-contrast computed

tomography in patients with acute ischemic stroke," *Computers in biology and medicine*, vol. 141, p. 105033, 2022.

[10] J. Lu, Y. Zhou, W. Lv, H. Zhu, T. Tian, S. Yan, Y. Xie, D. Wu, Y. Li, Y. Liu *et al.*, "Identification of early invisible acute ischemic stroke in non-contrast computed tomography using two-stage deep-learning model," *Theranostics*, vol. 12, no. 12, p. 5564, 2022.

[11] H. Kuang, B. K. Menon, S. I. Sohn, and W. Qiu, "Eis-net: Segmenting early infarct and scoring aspects simultaneously on non-contrast ct of patients with acute ischemic stroke," *Medical Image Analysis*, vol. 70, p. 101984, 2021.

[12] W. Qiu, H. Kuang, E. Teleg, J. M. Ospel, S. I. Sohn, M. Almekhlafi, M. Goyal, M. D. Hill, A. M. Demchuk, and B. K. Menon, "Machine learning for detecting early infarction in acute stroke with non–contrast-enhanced ct," *Radiology*, vol. 294, no. 3, pp. 638–644, 2020.

[13] G. Wu, X. Chen, J. Lin, Y. Wang, and J. Yu, "Identification of invisible ischemic stroke in noncontrast ct based on novel two-stage convolutional neural network model," *Medical Physics*, vol. 48, no. 3, pp. 1262–1275, 2021.

[14] S.-Y. Lin, P.-L. Chiang, P.-W. Chen, L.-H. Cheng, M.-H. Chen, P.-C. Chang, W.-C. Lin, and Y.-S. Chen, "Toward automated segmentation for acute ischemic stroke using non-contrast computed tomography," *International journal of computer assisted radiology and surgery*, vol. 17, no. 4, pp. 661–671, 2022.

[15] K. Qi, H. Yang, C. Li, Z. Liu, M. Wang, Q. Liu, and S. Wang, "X-net: Brain stroke lesion segmentation based on depthwise separable convolution and long-range dependencies," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part III 22*. Springer, 2019, pp. 247–255.

[16] H. Kuang, Y. Wang, J. Liu, J. Wang, Q. Cao, B. Hu, W. Qiu, and J. Wang, "Hybrid cnn-transformer network with circular feature interaction for acute ischemic stroke lesion segmentation on non-contrast ct scans," *IEEE Transactions on Medical Imaging*, 2024.

[17] H. Zhu, Y. Chen, T. Tang, G. Ma, J. Zhou, J. Zhang, S. Lu, F. Wu, L. Luo, S. Liu *et al.*, "Isp-net: Fusing features to predict ischemic stroke infarct core on ct perfusion maps," *Computer Methods and Programs in Biomedicine*, vol. 215, p. 106630, 2022.

[18] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[19] Y. Wang, A. K. Katsaggelos, X. Wang, and T. B. Parrish, "A deep symmetry convnet for stroke lesion segmentation," in *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2016, pp. 111–115.

[20] Q. Bao, S. Mi, B. Gang, W. Yang, J. Chen, and Q. Liao, "Mdan: mirror difference aware network for brain stroke lesion segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 4, pp. 1628–1639, 2021.

[21] N. Gruber, M. Haltmeier, A. ter Telgte, J. Schwab, E. Gizewski, and M. Galijasevic, "Non-stationary deep lifting with application to acute brain infarct segmentation," in *Medical Imaging with Deep Learning*, 2022.

[22] W. Yu, Z. Huang, J. Zhang, and H. Shan, "San-net: Learning generalization to unseen sites for stroke lesion segmentation with self-adaptive normalization," *Computers in Biology and Medicine*, vol. 156, p. 106717, 2023.

[23] H. Yang, C. Huang, X. Nie, L. Wang, X. Liu, X. Luo, and L. Liu, "Is-net: Automatic ischemic stroke lesion segmentation on ct images," *IEEE Transactions on Radiation and Plasma Medical Sciences*, 2023.

[24] A. v. d. Oord, Y. Li, and O. Vinyals, "Representation learning with contrastive predictive coding," *arXiv preprint arXiv:1807.03748*, 2018.

[25] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 9729–9738.

[26] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *International conference on machine learning*. PMLR, 2020, pp. 1597–1607.

[27] J.-B. Grill, F. Strub, F. Altché, C. Tallec, P. Richemond, E. Buchatskaya, C. Doersch, B. Avila Pires, Z. Guo, M. Gheshlaghi Azar *et al.*, "Bootstrap your own latent-a new approach to self-supervised learning," *Advances in neural information processing systems*, vol. 33, pp. 21 271–21 284, 2020.

[28] D. Zeng, Y. Wu, X. Hu, X. Xu, H. Yuan, M. Huang, J. Zhuang, J. Hu, and Y. Shi, "Positional contrastive learning for volumetric medical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2021, pp. 221–230.

[29] K. Chaitanya, E. Erdil, N. Karani, and E. Konukoglu, "Contrastive learning of global and local features for medical image segmentation with limited annotations," *Advances in Neural Information Processing Systems*, vol. 33, pp. 12 546–12 558, 2020.

[30] X. Hu, D. Zeng, X. Xu, and Y. Shi, "Semi-supervised contrastive learning for label-efficient medical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2021, pp. 481–490.

[31] C. You, W. Dai, Y. Min, L. Staib, and J. S. Duncan, "Bootstrapping semi-supervised medical image segmentation with anatomical-aware contrastive distillation," in *International conference on information processing in medical imaging*. Springer, 2023, pp. 641–653.

[32] C. You, W. Dai, Y. Min, L. Staib, J. Sekhon, and J. S. Duncan, "Action++: improving semi-supervised medical image segmentation with adaptive anatomical contrast," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2023, pp. 194–205.

[33] C. You, W. Dai, F. Liu, Y. Min, H. Su, X. Zhang, X. Li, D. A. Clifton, L. Staib, and J. S. Duncan, "Mine your own anatomy: Revisiting medical image segmentation with extremely limited labels," *arXiv preprint arXiv:2209.13476*, 2022.

[34] C. You, W. Dai, Y. Min, F. Liu, D. Clifton, S. K. Zhou, L. Staib, and J. Duncan, "Rethinking semi-supervised medical image segmentation: A variance-reduction perspective," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[35] P. Yang, X. Yin, H. Lu, Z. Hu, X. Zhang, R. Jiang, and H. Lv, "Cs-co: A hybrid self-supervised visual representation learning method for h&e-stained histopathological images," *Medical Image Analysis*, vol. 81, p. 102539, 2022.

[36] S. Wang, K. Liang, C. Pan, C. Ye, X. Li, F. Liu, Y. Yu, and Y. Wang, "Segmentation-based method combined with dynamic programming for brain midline delineation," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2020, pp. 772–776.

[37] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, and D. Krishnan, "Supervised contrastive learning," *Advances in neural information processing systems*, vol. 33, pp. 18 661–18 673, 2020.

[38] S. Kullback and R. A. Leibler, "On information and sufficiency," *The annals of mathematical statistics*, vol. 22, no. 1, pp. 79–86, 1951.

[39] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 fourth international conference on 3D vision (3DV)*. Ieee, 2016, pp. 565–571.

[40] F. Bougourzi, C. Distante, F. Dornaika, and A. Taleb-Ahmed, "Pdatt-unet: Pyramid dual-decoder attention unet for covid-19 infection segmentation from ct-scans," *Medical Image Analysis*, vol. 86, p. 102797, 2023.

[41] X. Wang, Y. Yuan, D. Guo, X. Huang, Y. Cui, M. Xia, Z. Wang, C. Bai, and S. Chen, "Ssa-net: Spatial self-attention network for covid-19 pneumonia infection segmentation with semi-supervised few-shot learning," *Medical Image Analysis*, vol. 79, p. 102459, 2022.

[42] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.

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

[44] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, and J. Liu, "Ce-net: Context encoder network for 2d medical image segmentation," *IEEE transactions on medical imaging*, vol. 38, no. 10, pp. 2281–2292, 2019.

[45] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "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, pp. 3–11.

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

[47] H. Yang, W. Huang, K. Qi, C. Li, X. Liu, M. Wang, H. Zheng, and S. Wang, "Clci-net: Cross-level fusion and context inference networks for lesion segmentation of chronic stroke," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part III 22*. Springer, 2019, pp. 266–274.

[48] A. A. Taha and A. Hanbury, "Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool," *BMC medical imaging*, vol. 15, no. 1, pp. 1–28, 2015.

[49] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne." *Journal of machine learning research*, vol. 9, no. 11, 2008.

[50] E. C. Jauch, J. L. Saver, H. P. Adams Jr, A. Bruno, J. Connors, B. M. Demaerschalk, P. Khatri, P. W. McMullan Jr, A. I. Qureshi, K. Rosenfield *et al.*, "Guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the american heart association/american stroke association," *Stroke*, vol. 44, no. 3, pp. 870–947, 2013.

[51] R. G. Nogueira, A. P. Jadhav, D. C. Haussen, A. Bonafe, R. F. Budzik, P. Bhuva, D. R. Yavagal, M. Ribo, C. Cognard, R. A. Hanel *et al.*, "Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct," *New England Journal of Medicine*, vol. 378, no. 1, pp. 11–21, 2018.

[52] G. Turc, P. Bhogal, U. Fischer, P. Khatri, K. Lobotesis, M. Mazighi, P. D. Schellinger, D. Toni, J. De Vries, P. White *et al.*, "European stroke organisation (eso)–european society for minimally invasive neurological therapy (esmint) guidelines on mechanical thrombectomy in acute ischaemic strokeendorsed by stroke alliance for europe (safe)," *European stroke journal*, vol. 4, no. 1, pp. 6–12, 2019.

[53] C. Chen, Q. Dou, Y. Jin, Q. Liu, and P. A. Heng, "Learning with privileged multimodal knowledge for unimodal segmentation," *IEEE transactions on medical imaging*, vol. 41, no. 3, pp. 621–632, 2021.

[54] Y. Chen, Y. Pan, Y. Xia, and Y. Yuan, "Disentangle first, then distill: A unified framework for missing modality imputation and alzheimer's disease diagnosis," *IEEE Transactions on Medical Imaging*, 2023.

[55] B. C. Campbell, "Optimal imaging at the primary stroke center," *Stroke*, vol. 51, no. 7, pp. 1932–1940, 2020.