Prototype-based Physiological Transfer Enables NCCT-only Hyperacute Stroke Tissue-Window Segmentation Under Missing Perfusion

Ying Dan*[1], Zhe Xu*[2,3], Longxi Zhou[4], Lu Zhang[5], Xiangyuan Ma[6], Yading Yuan(✉)[2,3], and Raymond Kai-yu Tong(✉)[1]

[1] Department of Biomedical Engineering, The Chinese University of Hong Kong, Hong Kong SAR, China
[2] Department of Radiation Oncology, Columbia University Irving Medical Center, New York, USA
[3] Data Science Institute, Columbia University, New York, USA
[4] Department of Biomedical Engineering, Southern University of Science and Technology, Shenzhen, China
[5] Zhejiang University School of Medicine, Hangzhou, China
[6] Department of Biomedical Engineering, Shantou University, Shantou, China
{zx2578, yy3332}@columbia.edu; kytong@cuhk.edu.hk

Abstract. Accurate delineation of ischemic core and salvageable penumbra is critical for guiding reperfusion therapy in hyperacute ischemic stroke. While CT perfusion (CTP) provides quantitative hemodynamic parameters for tissue-window assessment, its availability remains limited in many emergency settings. In contrast, non-contrast CT (NCCT) is nearly universally accessible but lacks explicit perfusion information, making core and penumbra segmentation highly challenging. We study stroke tissue-window segmentation under missing perfusion, where only NCCT is available at inference. In this setting, specialist NCCT models struggle due to subtle contrast and limited supervision, while reconstructing CTP from NCCT constitutes an ill-posed inverse problem prone to hemodynamic hallucination. To address this, we propose **ProPhyT**, a prototype-based physiological transfer framework that injects CTP-derived hemodynamic semantics into NCCT-based segmentation without reconstructing perfusion maps. ProPhyT consists of three processes: (i) Hemodynamic Prototype Support Bank Construction, where a self-supervised CTP encoder learns a hemodynamic embedding space and CTP features are clustered to form representative physiological prototypes; (ii) Cross-Modal Prototype Retrieval Learning, where an NCCT encoder is aligned to the CTP embedding space to enable effective prototype retrieval from NCCT alone; and (iii) Physiological Prototype-Conditioned Segmentation, where retrieved prototypes are converted into similarity-aware dense prompts to parameter-efficiently finetune a medical foundation model for core and penumbra segmentation. Extensive experiments demonstrate significant improvements over top-performing methods, supporting more reliable reperfusion decision-making in CTP-limited clinical settings.

* Equal Contribution.

Keywords: Hyperacute Ischemic Stroke · NCCT · Foundation Model.

1 Introduction

Acute ischemic stroke (AIS) is a leading cause of disability and mortality worldwide [6]. Reperfusion therapy remains the primary intervention for salvaging ischemic brain tissue and improving functional outcomes [21]. However, treatment decisions involve a critical trade-off: aggressive intervention may increase the risk of hemorrhagic transformation, whereas conservative management may forfeit salvageable tissue [5]. Accurate delineation of the ischemic core and penumbra is therefore essential for patient selection.

Traditionally, eligibility for reperfusion therapy has been determined by fixed time windows (e.g., 4.5-6 hours from onset) [21]. However, infarct progression varies substantially across patients due to differences in collateral circulation and vascular status [18, 30]. As a result, clinical practice has shifted toward a tissue-window paradigm, where imaging-based quantification of ischemic core and penumbra guides treatment decisions [1, 18]. CT perfusion (CTP) is the gold standard for tissue-window assessment (Fig. 1(b)), providing quantitative hemodynamic parameters. Clinically, the ischemic core is commonly defined by cerebral blood flow (CBF) $< 30\%$ of normal tissue, while critically hypoperfused (potentially salvageable) penumbra is identified by time-to-maximum (Tmax) of the deconvolved residue function > 6s, reflecting significant hemodynamic delay. When perfusion imaging is available, tissue-window segmentation is physiologically grounded and relatively straightforward. However, CTP requires specialized infrastructure, contrast injection, and longer acquisition time [23], limiting its availability in many emergency settings. In contrast, non-contrast CT (NCCT) is nearly universally accessible and serves as the first-line imaging modality. Unlike CTP, NCCT provides only static attenuation information and lacks explicit perfusion dynamics. Hyperacute ischemic lesions often exhibit subtle intensity changes, and core-penumbra differentiation cannot be directly inferred from intensity values alone. This leads to a fundamental challenge: *How can we perform tissue-window segmentation when perfusion information is missing?*

Existing approaches fall into two categories. First, specialist models trained directly on NCCT attempt to delineate core and penumbra from limited supervision [29, 11, 4]. However, due to low contrast and data scarcity, such models struggle to generalize. In our preliminary experiments, specialist models trained from scratch achieved Dice scores below 0.25 for core and 0.5 for penumbra, highlighting the intrinsic difficulty of NCCT-only tissue-window segmentation. Second, cross-modality synthesis methods attempt to generate CTP parameter maps from NCCT prior to segmentation. Yet this constitutes an ill-posed inverse problem: NCCT does not contain temporal hemodynamic dynamics, making faithful perfusion reconstruction theoretically underdetermined and prone to hemodynamic hallucination [9]. These limitations indicate that directly reconstructing missing perfusion signals may not be the optimal strategy. Recent medical foundation models [3, 17] pretrained on large-scale datasets provide strong

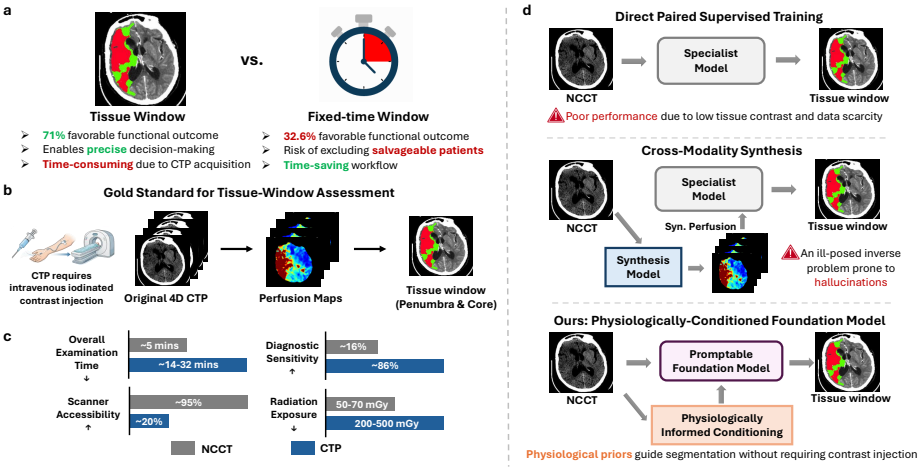

Fig. 1. (a-c) CTP-based tissue-window assessment is the gold standard but requires contrast, longer scans, and higher radiation. (d) Direct training suffers from low contrast, while perfusion synthesis is ill-posed. We propose a physiologically conditioned foundation model for contrast-free, tissue-aware NCCT-based segmentation.

generic visual priors and demonstrate superior generalization compared to conventional specialists. Yet, such models rely primarily on appearance cues and lack explicit physiological understanding of stroke hemodynamics. Visual fine-tuning alone is empirically insufficient under missing perfusion. Thus, we argue that physiological knowledge should be transferred at a higher semantic level and integrated into the segmentation foundation models.

To this end, we propose **ProPhyT**, a prototype-based physiological transfer framework for stroke tissue-window segmentation under missing perfusion. ProPhyT first conducts self-supervised pretraining [26] on available CTP and NCCT images to obtain modality-specific encoders. The pretrained CTP encoder establishes a hemodynamic-aware embedding space for constructing representative physiological prototypes, while the NCCT encoder is subsequently aligned to this space for cross-modal retrieval and conditioning. Specifically, ProPhyT consists of three steps: (i) Hemodynamic Prototype Bank Construction, where CTP representations are clustered in the hemodynamic embedding space to form representative physiological prototypes; (ii) Cross-Modal Prototype Retrieval Learning, where the NCCT encoder is aligned to the CTP embedding space to retrieve physiologically relevant prototypes; (iii) Physiological Prototype-Conditioned Segmentation, where similarity-weighted retrieved prototypes are aggregated into a patient-specific hemodynamic prior and projected into a spatially varying dense prompt to parameter-efficiently condition a medical foundation segmentation model for physiologically informed core and penumbra delineation. Extensive experiments demonstrate significant improvements in NCCT-only core and penumbra segmentation over recent top-performing ap-

proaches. By transferring high-level physiological semantics rather than reconstructing dense perfusion signals, ProPhyT provides an effective solution to segmentation under missing perfusion.

2 Method

As illustrated in Fig. 2, our ProPhyT features three key processes: (i) Hemodynamic Prototype Support Bank Construction, where a self-supervised CTP encoder learns a hemodynamic embedding space and CTP features are clustered to form a prototype bank representing typical perfusion patterns; (ii) Cross-Modal Prototype Retrieval Learning, where paired NCCT-CTP data are used to align an NCCT encoder to the CTP embedding space, enabling retrieval of physiologically consistent prototypes from NCCT alone; and (iii) Physiological Prototype-Conditioned Segmentation, where retrieved prototypes are transformed into similarity-aware dense prompts to parameter-efficiently finetune the medical foundation model for core and penumbra prediction.

2.1 Hemodynamic Prototype Support Bank Construction

We first construct a hemodynamic prototype support bank from CTP to provide structured physiological semantics. Using CTP parametric maps from UniToBrain [20] and ISLES24 [22], a CTP encoder f_{CTP} is pretrained via Sparse and Hierarchical Masked Modeling (SparK) [26]. For each input, hierarchical features $\{F_l^{\text{CTP}}\}_{l=1}^4$ are extracted and fused by a Feature Pyramid Network (FPN) [15]-based aggregator $g(\cdot)$, yielding a compact embedding: $\mathbf{z}^{\text{CTP}} = g\big(\{F_l^{\text{CTP}}\}_{l=1}^4\big) \in \mathbb{R}^D$, where $D = 256$. Collecting embeddings from all CTP samples, we apply K-means clustering to discretize the embedding space into M (128 by default) representative physiological prototypes: $\mathcal{P}^{\text{CTP}} = \{\mathbf{p}_m^{\text{CTP}}\}_{m=1}^M = \texttt{KMeans}\big(\{\mathbf{z}^{\text{CTP}}\}\big)$, where each captures a distinct and recurrent physiological perfusion pattern (e.g., large infarct core, target mismatch, poor collaterals). The set $\mathcal{P}^{\text{CTP}}$ forms a compact hemodynamic prototype bank that summarizes dominant perfusion patterns and serves as physiological support for subsequent cross-modal retrieval.

2.2 Cross-Modal Prototype Retrieval Learning

The prototype bank constructed in Sec. 2.1 defines a CTP-specific hemodynamic embedding space. To enable NCCT-only inference, we learn an encoder that projects NCCT images into this space, enabling reliable retrieval of physiologically consistent prototypes. Similarly, we first pretrain the encoder via SparK [26] on NCCT images to obtain modality-aware representations. Subsequently, using paired NCCT-CTP data from ISLES24 [22], we align the two modalities at three complementary levels: structure, distribution, and instance discrimination.

Structure-Level Alignment. Hemodynamic semantics manifest at multiple spatial scales. To preserve structural correspondence, we align intermediate features using cosine similarity $\texttt{sim}(\cdot,\cdot)$: $\mathcal{L}_{\text{struct}} = \frac{1}{L}\sum_{l=1}^{L}\big(1 - \texttt{sim}\big(F_l^{\text{NCCT}}, F_l^{\text{CTP}}\big)\big)$.

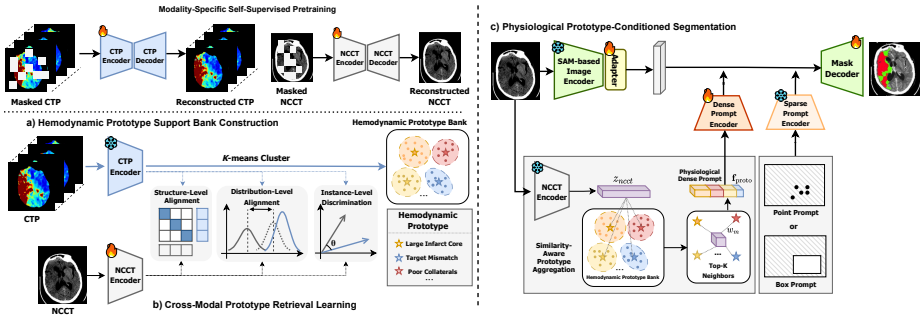

Fig. 2. Overview of ProPhyT, featuring three stages: (a) Hemodynamic Prototype Support Bank Construction learns a CTP embedding space and clusters physiological prototypes; (b) Cross-Modal Prototype Retrieval Learning aligns NCCT to the CTP space for prototype retrieval; (c) Physiological Prototype-Conditioned Segmentation converts retrieved prototypes into dense prompts to finetune the foundation model.

Distribution-Level Alignment. Due to intrinsic modality gaps, NCCT and CTP features exhibit systematic distribution shifts. We reduce global bias [24] by aligning feature statistics within each batch. Let $\mu_{\text{CTP}} = \mathbb{E}[F^{\text{CTP}}]$ and $\mu_{\text{NCCT}} = \mathbb{E}[F^{\text{NCCT}}]$, and define the mean shift $\delta = \mu_{\text{CTP}} - \mu_{\text{NCCT}}$. We minimize $\mathcal{L}_{\text{dist}} = \left\|(F^{\text{NCCT}} + \delta) - F^{\text{CTP}}\right\|_2^2$, compensating systematic modality bias.

Instance-Level Discrimination. Reliable prototype retrieval requires discriminative alignment at the instance level. We adopt an InfoNCE-style contrastive loss [19], pulling paired embeddings together while separating unrelated samples: $\mathcal{L}_{\text{inst}} = -\log \frac{\exp(\texttt{sim}(\mathbf{z}^{\text{NCCT}}, \mathbf{z}^{\text{CTP}})/\tau)}{\sum_j \exp(\texttt{sim}(\mathbf{z}^{\text{NCCT}}, \mathbf{z}_j^{\text{CTP}})/\tau)}$.

The **overall alignment objective** is $\mathcal{L}_{\text{align}} = \mathcal{L}_{\text{struct}} + \mathcal{L}_{\text{dist}} + \mathcal{L}_{\text{inst}}$. After training, f_{NCCT} maps NCCT images into the CTP hemodynamic embedding space, enabling reliable retrieval from the prototype support bank.

2.3 Physiological Prototype-Conditioned Segmentation

Similarity-Aware Prototype Aggregation. After alignment, the NCCT encoder (frozen in this stage) produces a query embedding $\mathbf{z}^{\text{NCCT}}$ in the hemodynamic space. We view the prototype bank $\mathcal{P}^{\text{CTP}} = \{\mathbf{p}_m\}_{m=1}^{M}$ as a discrete approximation of the hemodynamic manifold; prototype retrieval then amounts to manifold integration by aggregating nearby physiological states rather than committing to a single (potentially noisy) match. Cosine similarities between the query and all prototypes are computed as $s_m = \texttt{sim}(\mathbf{z}^{\text{NCCT}}, \mathbf{p}_m)$ and temperature-scaled attention over the Top-K neighbors (denoted as $\mathcal{N}_K$) yields:

$$w_m = \frac{\exp(s_m/\tau_r)}{\sum_{j\in\mathcal{N}_K} \exp(s_j/\tau_r)}, \quad \mathbf{f}_{\text{proto}} = \sum_{m\in\mathcal{N}_K} w_m\, \mathbf{p}_m. \tag{1}$$

The resulting descriptor $\mathbf{f}_{\text{proto}}$ represents a weighted integration of nearby physiological states, robustly representing a patient-specific hemodynamic prior.

Table 1. Summary of datasets utilized in this study.

Dataset	Samples	Available Modality		Annotation		Role
		NCCT	CTP Maps	Core	Penumbra	
UniToBrain [20]	115	-	✓	-	-	Prototype Bank Construction
ISLES 2024 [22]	140	✓	✓	✓†	-	Cross-Modal Retrieval Learning
CPAISD [27]	112	✓	-	✓	✓	Target Benchmark

† ISLES 2024 provides the final infarct volume observed on follow-up MRI, which is used here as a proxy for the core region.

Physiologically Conditioned Finetuning. The aggregated descriptor $\mathbf{f}_{\text{proto}}$ is fed into a trainable two-layer MLP with ReLU activation, termed the dense prompt projection layer, which transforms the global hemodynamic representation into a spatially varying guidance map aligned with the dimensions of SAM's dense prompt embedding [10]. This projection injects physiological priors into the decoder, encouraging attention toward regions consistent with the retrieved perfusion patterns. Following SAM-style interactive training [10, 3], sparse prompts are simulated by randomly sampling either a bounding box or a single positive point with equal probability. During finetuning, the image encoder is frozen except for its Adapter layers. The mask decoder and the dense prompt projection layers are trainable. The segmentation loss combines Dice loss $\mathcal{L}_{\text{dice}}$, which optimizes region overlap, and Focal loss $\mathcal{L}_{\text{focal}}$ [16], which mitigates class imbalance: $\mathcal{L}_{\text{seg}} = \mathcal{L}_{\text{dice}} + 20\,\mathcal{L}_{\text{focal}}$. Additionally, we supervise the mask quality prediction head with a mask IoU regression loss $\mathcal{L}_{\text{iou}}$ [10], defined as the mean squared error between the predicted IoU score and the actual IoU computed from the predicted and ground-truth masks.

3 Experiments

Datasets. We evaluate our method on the **CPAISD** dataset [27], which contains 112 hyperacute stroke cases (onset $<$ 24h) and, to our knowledge, is currently the only public dataset providing both ischemic core and penumbra annotations derived with reference to CTP while releasing NCCT images only, making it well suited for the missing-perfusion setting studied in this work. Following the official split, 92/10/10 cases are used for training/validation/testing. To compensate for the absence of perfusion imaging and enable physiological knowledge transfer, we additionally leverage two auxiliary datasets with perfusion modalities: **ISLES 2024** [22] (140 cases with paired NCCT and 4D CTP, with follow-up infarct labels used as a proxy for the core) and **UniToBrain** [20] (115 cases with 4D CTP and perfusion parameter maps). These datasets are used for cross-modal representation alignment and physiological prototype bank construction, while CPAISD serves exclusively as the target benchmark. Dataset details are summarized in Table 1.

Implementation. The framework is implemented in PyTorch on an NVIDIA A100 GPU. All slices are resized to 256×256 pixels. The single-channel NCCT images are min-max normalized to $[0, 1]$, and three-channel inputs follow ImageNet standardization. Random flipping, rotation, and affine transformation are applied for data augmentation. We adopt SAM-Med2D [3] as the backbone

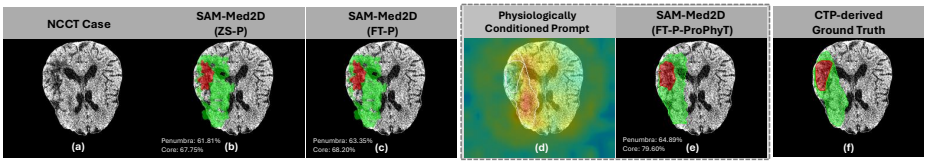

Fig. 3. Exemplar results with SAM-Med2D across different settings (ZS-P, FT-P, FT-P-ProPhyT (ours)). Core is shown in red; penumbra in green.

due to its large-scale medical adaptation (∼19.7M masks) and flexible prompting modes, supporting point, box, and dense mask inputs for parameter-efficient finetuning. The NCCT and CTP encoders both use ResNet-50 architectures. Top-5 CTP hemodynamic prototypes are retrieved with temperature $\tau_r = 0.1$ in a 256-d embedding space. The model is trained for 15 epochs using the Adam optimizer with a learning rate of 1×10^{-4} and a batch size of 2. Performance is evaluated using Dice score and 95% Hausdorff distance (HD95). Code will be available at `https://github.com/CUHK-BMEAI/ProPhyT`.

Comparison Study. Table 2 reports the quantitative comparison on CPAISD. AIS-specific models [13, 11, 31, 25], conventional architectures [7, 12, 2, 14], and perfusion synthesis-based approaches [8] demonstrate partial feasibility; however, their performance remains limited, highlighting the intrinsic difficulty of inferring physiologically defined tissue states from low-contrast NCCT alone. AISCT-SAM, which models bilateral hemisphere differences for infarct segmentation, also yields modest performance, likely because asymmetry-based modeling does not explicitly encode perfusion dynamics necessary for distinguishing core and penumbra. Foundation models exhibit stronger generalization capacity. Under zero-shot box prompting (ZS-P), SAM-Med2D achieves balanced performance on both core and penumbra, with an average Dice of 67.34%, substantially outperforming MedSAM (44.77%). Accordingly, we adopt SAM-Med2D as our backbone. However, naive prompt-based finetuning (FT-P) utilized in [3] fails to yield consistent gains and slightly reduces the average Dice, suggesting that structural adaptation alone may amplify intensity-driven biases that are weakly aligned with perfusion-defined labels. In contrast, our physiologically conditioned version, SAM-Med2D (FT-P-ProPhyT), achieves 74.28% Dice for core and 81.03% for penumbra, with consistent improvements across both tissue types. This supports the hypothesis that transferring CTP-derived physiological representations introduces complementary and effective hemodynamic priors beyond structural finetuning. Exemplar results are shown in Fig. 3. For scenarios where clinicians cannot provide reliable spatial cues from structural intensity alone, we further provide a fully automatic mode (FT-Auto-ProPhyT) that couples physiological dense priors with detector-derived coarse spatial cues [28], achieving an average Dice of 58.57%, with a modest performance trade-off relative to the interactive setting.

Ablation Analysis. We perform an ablation study to evaluate each component (Table 3). Replacing CTP-derived prototypes with learnable tokens (Abla-1)

Table 2. Quantitative comparison (mean ± std) of different methods. † denotes AIS-specific method. The best results are **bolded**.

Method	Core		Penumbra		Avg.	
	Dice (%) ↑	HD95 (mm) ↓	Dice (%) ↑	HD95 (mm) ↓	Dice (%) ↑	HD95 (mm) ↓
CL-SegNet† [13]	19.75±20.98	74.69±43.05	30.30±26.84	85.54±22.42	25.03	80.12
I^2PC-Net† [11]	15.54±19.56	64.01±35.73	39.96±31.03	63.22±38.94	27.75	63.62
CLCI-Net† [31]	12.15±21.25	76.81±38.42	44.49±33.41	56.82±39.16	28.32	66.82
U-KAN [14]	10.31±20.20	64.76±38.32	36.67±30.49	55.81±34.68	23.49	60.29
BEA-Net [12]	22.75±23.60	55.21±42.52	38.88±33.92	56.92±39.19	30.82	56.07
nnUNet [7]	11.00±21.25	65.66±37.84	33.36±31.33	67.11±41.69	22.18	66.39
nnUNet (Syn.)† [7, 8]	13.14±22.92	58.18±34.88	43.78±33.30	45.95±36.22	28.46	52.07
Swin-UNet [2]	12.54±18.95	76.96±31.75	45.73±29.31	67.18±40.59	29.14	72.07
Swin-UNet (Syn.)† [2, 8]	16.59±20.96	54.33±30.85	47.92±35.13	40.80±34.70	32.26	47.57
AISCT-SAM† [25]	13.45±16.13	97.11±14.71	36.23±26.18	82.13±35.33	24.84	89.62
MedSAM (ZS-P) [17]	10.76±19.53	20.63±10.88	78.77±11.33	9.51±5.83	44.77	15.07
MedSAM (FT-P) [17]	14.92±21.92	22.69±17.72	**86.36±9.28**	**5.11±4.50**	50.64	13.90
SAM-Med2D (ZS-P) [3]	57.76±14.54	7.53±3.74	76.91±9.00	9.84±6.20	67.34	8.69
SAM-Med2D (FT-P) [3]	60.80±9.92	7.31±2.50	68.95±11.54	9.09±3.90	64.88	8.20
SAM-Med2D (FT-P-**ProPhyT**)	**74.28±10.26**	**5.17±3.01**	81.03±8.45	6.94±4.20	**77.66**	**6.06**
SAM-Med2D (FT-Auto-**ProPhyT**)	56.13±23.15	16.84±25.29	61.01±18.73	13.27±18.41	58.57	15.06

Table 3. Ablation analysis.

Method	Core		Penumbra	
	Dice (%) ↑	HD95 (mm) ↓	Dice (%) ↑	HD95 (mm) ↓
Full Model: SAM-Med2D (FT-P-**ProPhyT**)	**74.28±10.26**	**5.17±3.01**	**81.03±8.45**	**6.94±4.20**
Hemodynamic Prototype Bank				
Abla-1: learnable tokens (not CTP-derived)	60.31±14.06	10.33±3.88	76.20±10.21	10.27±4.71
Cross-model Alignment				
Abla-2: w/o alignment	68.57±11.44	6.08±2.88	74.57±8.60	8.86±3.42
Similarity-Aware Prototype Aggregation & Conditioning				
Abla-3: w/o physiological transfer (baseline)	60.80±9.92	7.31±2.50	68.95±11.54	9.09±3.90
Abla-4: w/o retrieval (random)	64.48±12.40	7.58±2.80	73.17±9.07	8.50±3.54
Abla-5: w/o similarity weighting (uniform)	70.34±14.34	6.05±3.32	75.13±8.59	8.10±4.71
Abla-6: w/o Top-K retrieval (Top-1)	65.82±12.19	6.85±3.22	74.24±12.96	10.01±5.73
Abla-7: w/o Top-K sparsity (all)	52.85±19.64	18.43±29.42	72.82±9.47	10.26±5.96
Abla-8: w/o dense prompt projection (broadcast)	71.58±10.14	6.08±2.92	80.78±8.74	7.16±4.14

substantially degrades performance, confirming that gains arise from physiologically grounded representations rather than additional parameters. Removing cross-modal alignment (Abla-2) also leads to notable degradation, showing that mapping NCCT features into a hemodynamic-consistent space is essential for meaningful retrieval. Disabling physiological transfer entirely (Abla-3) markedly degrades performance, confirming that typical foundation-model finetuning is insufficient under missing perfusion. Random retrieval (Abla-4) causes a significant decline, highlighting the necessity of similarity-aware matching. Uniform weighting (Abla-5) and Top-1 retrieval (Abla-6) both impair performance, demonstrating that similarity-weighted Top-K aggregation stabilizes retrieval. Aggregating over all prototypes (Abla-7) significantly degrades performance, suggesting that locality prevents interference from distant physiological states. Replacing spatial dense prompt projection with prototype broadcasting (Abla-8) also degrades performance, validating the importance of structured conditioning. Overall, these findings support the progressive and synergistic design of ProPhyT.

4 Conclusion

In this paper, we introduced ProPhyT, a prototype-based physiological transfer framework for NCCT-only stroke tissue-window segmentation under missing perfusion. ProPhyT learns a hemodynamic embedding space from perfusion data, constructs representative physiological prototypes, and enables cross-modal retrieval to inject physiological dense prompts into the medical foundation model finetuning. Extensive experiments showed that ProPhyT consistently outperforms top-performing baselines for ischemic core and penumbra delineation. By enabling physiologically informed segmentation directly from widely accessible NCCT, ProPhyT improves early tissue-window assessment and supports timely reperfusion decisions in resource-limited settings.

Acknowledgments. This research has been partially funded through the generous support of Herbert and Florence Irving/the Irving Trust. This work is also supported by ITC Mainland-Hong Kong Joint Funding Scheme (MHKJFS, 2022) (MHP/097/22)

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

References

1. Albers, G.W., Marks, M.P., Kemp, S., et al.: Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. New England Journal of Medicine **378**(8), 708–718 (2018)
2. 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)
3. Cheng, J., Ye, J., Deng, Z., Chen, J., Li, T., Wang, H., Su, Y., Huang, Z., Chen, J., Jiang, L., et al.: Sam-med2d. arXiv preprint arXiv:2308.16184 (2023)
4. Christensen, S., Demeestere, J., Verhaaren, B.F., Heit, J.J., Von Stein, E.L., Madill, E.S., Kennedy Loube, D., Dugue, R., Rengarajan, S., Mlynash, M., et al.: Semi-automated detection of early infarct signs on noncontrast ct improves interrater agreement. Stroke **54**(12), 3090–3096 (2023)
5. Emberson, J., Lees, K.R., Lyden, P., et al.: Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: a meta-analysis of individual patient data from randomised trials. The Lancet **384**(9958), 1929–1935 (2014)
6. Feigin, V.L., Stark, B.A., Johnson, C.O., et al.: Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the global burden of disease study 2019. The Lancet Neurology **20**(10), 795–820 (2021)
7. 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)
8. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1125–1134 (2017)

9. Khan, W., See, K.B., Kato, S., Huang, Z., Lazarte, A., Douglas, K., Lou, X., Peng, T.J., Rajderkar, D., Rees, J., et al.: Physiology-Informed Generative Multi-Task Network for Contrast-Free CT Perfusion. arXiv preprint arXiv:2505.22673 (2025)
10. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., et al.: Segment anything. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 4015–4026 (2023)
11. Kuang, H., Tan, X., Wang, J., Qu, Z., Cai, Y., Chen, Q., Kim, B.J., Qiu, W.: Segmenting ischemic penumbra and infarct core simultaneously on non-contrast ct of patients with acute ischemic stroke using novel convolutional neural network. Biomedicines **12**(3), 580 (2024)
12. Kuang, H., Wang, Y., Liang, Y., Liu, J., Wang, J.: BEA-Net: Body and edge aware network with multi-scale short-term concatenation for medical image segmentation. IEEE Journal of Biomedical and Health Informatics **27**(10), 4828–4839 (2023)
13. 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)
14. Li, C., Liu, X., Li, W., Wang, C., Liu, H., Liu, Y., Chen, Z., Yuan, Y.: U-kan makes strong backbone for medical image segmentation and generation. In: Proceedings of the AAAI conference on artificial intelligence. vol. 39, pp. 4652–4660 (2025)
15. Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S.: Feature pyramid networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2117–2125 (2017)
16. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision. pp. 2980–2988 (2017)
17. Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B.: Segment anything in medical images. Nature communications **15**(1), 654 (2024)
18. Nogueira, R.G., Jadhav, A.P., Haussen, D.C., Bonafe, A., Budzik, R.F., Bhuva, P., Yavagal, D.R., Ribo, M., Cognard, C., Hanel, R.A., et al.: Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. New England Journal of Medicine **378**(1), 11–21 (2018)
19. Parulekar, A., Collins, L., Shanmugam, K., Mokhtari, A., Shakkottai, S.: Infonce loss provably learns cluster-preserving representations. In: The thirty sixth annual conference on learning theory. pp. 1914–1961. PMLR (2023)
20. Perlo, D., Tartaglione, E., Gava, U., D'Agata, F., Benninck, E., Bergui, M.: Unitobrain dataset: a brain perfusion dataset. In: international conference on image analysis and processing. pp. 498–509. Springer (2022)
21. Powers, W.J., Rabinstein, A.A., Ackerson, T., et al.: Guidelines for the early management of patients with acute ischemic stroke: 2019 update to the 2018 guidelines for the early management of acute ischemic stroke: a guideline for healthcare professionals from the american heart association/american stroke association. Stroke **50**(12), e344–e418 (2019)
22. Riedel, E.O., de la Rosa, E., Baran, T.A., Petzsche, M.H., Baazaoui, H., Yang, K., Musio, F.A., Huang, H., Robben, D., Seia, J.O., et al.: Isles'24–a real-world longitudinal multimodal stroke dataset. arXiv preprint arXiv:2408.11142 (2024)
23. Shen, J., Li, X., Li, Y., Wu, B.: Comparative accuracy of CT perfusion in diagnosing acute ischemic stroke: A systematic review of 27 trials. PloS One (2017)
24. Sun, B., Saenko, K.: Deep coral: Correlation alignment for deep domain adaptation. In: European Conference on Computer Vision. pp. 443–450. Springer (2016)

25. Tan, X., Qu, Z., Wang, J., Kuang, H.: A clinical knowledge-driven fine-tuning strategy for applying foundation model to fully automatic acute ischemic stroke lesion segmentation on non-contrast ct scans. In: 2025 IEEE International Conference on Acoustics, Speech and Signal Processing. pp. 1–5. IEEE (2025)
26. Tian, K., Jiang, Y., Diao, Q., Lin, C., Wang, L., Yuan, Z.: Designing bert for convolutional networks: Sparse and hierarchical masked modeling. arXiv preprint arXiv:2301.03580 (2023)
27. Umerenkov, D., Kudin, S., Peksheva, M., Pavlov, D.: Core-penumbra hyperacute ischemic stroke dataset. Scientific Data **12**(1), 707 (2025)
28. Varghese, R., Sambath, M.: Yolov8: A novel object detection algorithm with enhanced performance and robustness. In: 2024 International conference on advances in data engineering and intelligent computing systems (ADICS). pp. 1–6. IEEE (2024)
29. Wang, X., Meng, Y., Dong, Z., Cao, Z., He, Y., Sun, T., Zhou, Q., Niu, G., Ding, Z., Shi, F., et al.: Segmentation of infarct lesions and prognosis prediction for acute ischemic stroke using non-contrast CT scans. Computer Methods and Programs in Biomedicine **258**, 108488 (2025)
30. Xiong, Y., Li, S., Wang, C., Sun, D., Li, Z., Gu, H., Jin, A., Dong, Q., Liu, L., Miao, Z., et al.: Chinese stroke association guidelines on reperfusion therapy for acute ischaemic stroke 2024. Stroke and Vascular Neurology (2025)
31. Yang, H., Huang, W., Qi, K., Li, C., Liu, X., Wang, M., Zheng, H., Wang, S.: Clci-net: Cross-level fusion and context inference networks for lesion segmentation of chronic stroke. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 266–274. Springer (2019)