Uncertainty-Aware Hypergraph Consistency Learning for Semi-supervised Medical Image Segmentation

Kexin Xie[1,2], Baoyao Yang[1⋆], Kejing Yin[2⋆], William K. Cheung[2], and Guibo Luo[3]

[1] Guangdong University of Technology, Guangzhou
ybaoyao@gdut.edu.cn
[2] Hong Kong Baptist University, Hongkong
[3] Guangdong Provincial Key Laboratory of Ultra High Definition Immersive Media Technology, Shenzhen Graduate School, Peking University

Abstract. Semi-supervised medical image segmentation (SSMIS) reduces reliance on costly manual annotations while providing pixel-level anatomical localization to support downstream clinical tasks. Consistency learning has emerged as a prominent paradigm in SSMIS, harnessing unlabeled data to improve segmentation fidelity. However, existing consistency-based methods predominantly enforce constraints at the pixel or feature level, overlooking the anatomical structural consistency essential for accurate delineation. This limitation is exacerbated by prediction uncertainty, particularly near lesion boundaries and in low-contrast regions, creating a dilemma: naïvely enforcing consistency misleads training, while discarding such regions risks losing diagnostically valuable information. To address these issues, we propose an Uncertainty-Aware Hypergraph Consistency Learning (UAHC) framework. UAHC elevates consistency learning to a high-order hypergraph level, capturing complex anatomical dependencies to enforce structural coherence beyond pairwise modeling. It introduces a dual-branch design that separately handles certain and uncertain regions via complementary constraints. The Certain-Region Hypergraph Consistency (CHC) module constructs hypergraphs to align teacher-student representations on stable areas, while the Uncertain-Region Hypergraph Consistency (UHC) module enables effective structural learning on ambiguous regions without being overwhelmed by noise. A novel hyperedge-level uncertainty weighting mechanism further adaptively emphasizes diagnostically significant yet uncertain anatomical patterns. Extensive experiments on four public benchmarks show that UAHC consistently outperforms state-of-the-art methods. Code available at: https://github.com/BaoyaoGroup/UAHC.

Keywords: Semi-supervised learning · Medical image segmentation · Hypergraph · Consistency Learning.

⋆ Co-corresponding authors.

1 Introduction

Medical image segmentation (MIS) is essential for computer vision and clinical diagnosis, supporting tasks such as lesion detection and organ delineation [3]. However, accurate segmentation typically relies on large amounts of expert-annotated data, which are costly and time-consuming to acquire [2]. Semi-supervised medical image segmentation (SSMIS) has emerged as a promising solution [13], leveraging abundant unlabeled data alongside limited labeled samples to reduce annotation burden while improving segmentation performance. Consistency learning has become a dominant paradigm in SSMIS [24], enabling robust representation learning by enforcing prediction consistency under perturbations. However, existing methods primarily enforce pixel- or feature-level constraints [20, 18, 26], overlooking intrinsic anatomical dependencies such as organ morphology and tissue continuity. Although graph-based extensions exist, they typically restrict structural modeling to pairwise interactions, failing to capture the complex high-order relationships inherent in medical images [22]. Another critical challenge is prediction uncertainty [27, 4], which severely undermines consistency learning. This issue is particularly acute in low-contrast regions and around lesion boundaries, where anatomical ambiguity makes stable predictions difficult [16, 23]. Uncertain predictions introduce a fundamental dilemma: naïvely enforcing consistency on these unreliable regions propagates noise and degrades model performance, whereas simply discarding them risks losing diagnostically valuable information embedded in these challenging areas. These challenges collectively highlight a crucial yet largely overlooked question in SSMIS: *how can we effectively model high-order anatomical structures while fully exploiting the valuable cues hidden within uncertain regions?*

It is well established that hypergraphs offer a powerful formalism for capturing high-order relationships among multiple entities [11, 17], making them well suited for modeling anatomical structures in medical images, where organs, lesions, and tissues exhibit complex, non-pairwise dependencies. To operationalize this insight, we construct hyperedges that connect sets of spatially or semantically related pixels, encoding group-wise contextual information reflective of intrinsic anatomical organization to enable high-order consistency. Specifically, we develop an Uncertainty-Aware Hypergraph Consistency Learning (UAHC) framework that jointly enforces hypergraph-level consistency and adaptive uncertainty modeling. UAHC reformulates consistency learning via hypergraphs, and dynamically partitions pixels into certain and uncertain regions within a dual-branch design featuring complementary modules tailored to each. The Certain-Region Hypergraph Consistency (CHC) module enforces robust structural alignment on stable areas, whereas the Uncertain-Region Hypergraph Consistency (UHC) module facilitates learning of complex patterns from ambiguous regions and suppresses noise via density-guided modulation. Additionally, UAHC propagates pixel-level uncertainty to quantify hyperedge-level confidence, adaptively reweighting hyperedges to emphasize uncertain yet clinically critical structures. By explicitly addressing both structural coherence and prediction ambiguity, UAHC establishes a new paradigm for SSMIS that holisti-

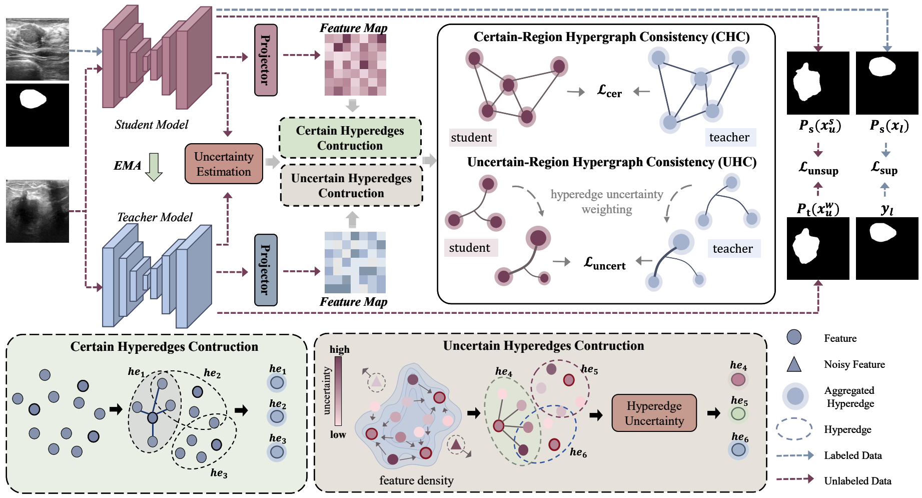

Fig. 1. Overview: UAHC models high-order anatomical structures via hypergraphs and introduces a dual-branch design: it aligns stable patterns in certain regions through the CHC module, while suppressing noise and exploring informative cues in uncertain regions via the UHC module.

cally models high-order anatomical dependencies while transforming uncertainty from a liability into a learnable signal.

Our main contributions include: 1) A novel hypergraph-based consistency learning framework that captures high-order anatomical dependencies via hyperedge representations, transcending limited pixel-level constraints. 2) Dual complementary modules (CHC & UHC) that align structures in certain regions while exploiting diagnostically valuable ambiguous areas. 3) Experiments on four benchmarks showing superior, coherent segmentation with limited supervision.

2 Method

As shown in Fig. 1, UAHC builds on the student–teacher MT architecture [21]. Given limited labeled data $\mathcal{D}_l = (x_l, y_l)$ and abundant unlabeled data $\mathcal{D}_u = x_u$, weakly augmented views x_u^w are fed to the teacher, and strongly augmented views x_u^s to the student. Both networks are equipped with a MLP projection head $\mathcal{P}(\cdot)$ that extracts pixel-wise features for subsequent hypergraph construction. UAHC first estimates pixel-wise uncertainty via prediction entropy and cross-model consensus, partitioning pixels into certain and uncertain regions. For each region, the teacher and student independently construct hypergraphs that encode high-order anatomical relationships via hyperedges. Certain-region hypergraphs are directly aligned in both feature and distribution spaces to enforce stable structural consistency. For uncertain regions, hyperedge construction incorporates uncertainty-weighted affinity to suppress noise, and consistency is adaptively reweighted by hyperedge uncertainty. This dual-branch hypergraph alignment enables robust learning from both reliable and challenging areas.

2.1 Uncertainty Region Estimation

To enable region-adaptive supervision, we first estimate pixel-wise uncertainty based on two complementary indicators: prediction entropy and cross-model consensus. Entropy quantifies single-model confidence, with higher values indicating greater ambiguity in class prediction, while cross-model consensus reflects prediction disagreement, a key signal that a region may be ambiguous or the models lack confidence, even if each individual entropy is low. Their combination provides a robust and holistic uncertainty estimate for guiding region partitioning. Formally, entropy for teacher and student predictions at pixel i is computed as:

$$U_t(i) = -\sum_a \tilde{Y}_t^a(i) \log \tilde{Y}_t^a(i), \quad U_s(i) = -\sum_a \tilde{Y}_s^a(i) \log \tilde{Y}_s^a(i), \tag{1}$$

where $\tilde{Y}_s^a(i)$ and $\tilde{Y}_t^a(i)$ denote the softmax probabilities for class a output by the teacher and student models, respectively.

Consensus between the two models is captured by a binary mask $B(i)$ defined at the prediction level, as it reflects the final decision agreement: $B(i) = 1$ if $\arg\max(\tilde{Y}_t(i)) = \arg\max(\tilde{Y}_s(i))$; otherwise, $B(i) = 0$. Here, $\tilde{Y}_t(i)$ and $\tilde{Y}_s(i)$ represent the full softmax vectors at pixel i. Combining entropy and consensus, pixels are partitioned into certain (χ_c) and uncertain (χ_u) regions:

$$\begin{aligned} \chi_c &= \{z_i \mid \max(U_s(i), U_t(i)) < \tau \wedge B(i) = 1\}, \\ \chi_u &= \{z_i \mid \max(U_s(i), U_t(i)) > \tau \vee B(i) \neq 1\}, \end{aligned} \tag{2}$$

where z_i is the projected feature of pixel i. Following the dynamic thresholding strategy of [25], the threshold τ is linearly increased during training, progressively incorporating more reliable pixels into the certain region as the model becomes more reliable. This dynamic partition evolves over iterations, enabling supervision strategies that adapt to regional confidence levels.

2.2 Certain-Region Hypergraph Consistency

Pixels in certain regions exhibit high confidence and stability, providing reliable features for structural alignment. To capture high-order anatomical dependencies, we construct a certain-region hypergraph $\mathcal{G}_c = (\mathcal{V}_c, \mathcal{E}_c)$, where nodes $\mathcal{V}_c = \{z_i\}_{i=1}^{N_c}$ are feature from χ_c, and hyperedges $\mathcal{E}_c = \{e_j\}_{j=1}^{M_c}$ connects semantically related nodes to capture high-order contextual dependencies. Following [12], hyperedges are constructed via the k-nearest neighbors (KNN) based on feature similarity. For each node z_i, we select its top-k neighbors $\mathcal{N}_k(z_i)$ and form a hyperedge $e_i = z_i \cup \mathcal{N}_k(z_i)$, with edge weights $w_{ij} = \|z_i - z_j\|_2$. This captures local semantic correlations. The adjacency structure is shared between the teacher and student models to enable consistent alignment. Each hyperedge is represented by the mean aggregation of its nodes, *i.e.*, $h_j = \frac{1}{|e_j|} \sum_{z_i \in e_j} z_i$.

Hypergraph Alignment. To enforce structural consistency between teacher and student models, we align their hypergraph features at both local and global levels. Locally, we align each hyperedge representation via mean squared error:

$$\mathcal{L}_{\mathrm{he}}(j) = \|h_j^s - h_j^t\|_2^2, \tag{3}$$

where h_j^s and h_j^t are student and teacher hyperedge features.

Globally, we model inter-hyperedge relationships to capture long-range interactions and maintain holistic anatomical coherence. By computing pairwise similarity distributions among hyperedges, we minimize the Kullback-Leibler (KL) divergence between student and teacher:

$$\mathcal{L}_{\mathrm{dist}}(j) = \mathrm{KL}\left(\ \mathrm{softmax}(H_s H_s^\top/\tau)_j \mid \mathrm{softmax}(H_t H_t^\top/\tau)_j \ \right), \tag{4}$$

where H represents aggregated hyperedge features, τ is a temperature parameter, and the subscript j indexes the similarity distribution of hyperedge e_j against all others.

The total certain-region consistency loss combines both objectives:

$$\mathcal{L}_{\mathrm{certain}} = \mathbb{E}_{e_j \in \mathcal{E}_c}\left[\mathcal{L}_{\mathrm{he}}(j) + \lambda \mathcal{L}_{\mathrm{dist}}(j)\right], \tag{5}$$

with balancing parameter λ. This joint alignment encourages feature-level consistency while preserving global anatomical structure in high-confidence regions.

2.3 Uncertain-Region Hypergraph Consistency

High-uncertainty regions are diagnostically critical yet inherently noise-prone, making direct structural alignment as described in Sec. 2.2 suboptimal and risk of noise propagation. To enable robust learning, we suppress noise during uncertain hyperedge construction via feature density modulation and reweight consistency objectives at the hyperedge level, allowing UAHC to effectively capture complex, diagnostically significant structures.

Denote uncertain-region hypergraph $\mathcal{G}_u = (\mathcal{V}_u, \mathcal{E}_u)$, where nodes $\mathcal{V}_u = \{z_i\}_{i=1}^{N_u}$ are feature from χ_u, and hyperedges $\mathcal{E}_u = \{e_j\}_{j=1}^{M_u}$ connects semantically related nodes. In addition to feature similarity, node uncertainty should be considered during hyperedge construction, as highly uncertain regions often contain substantial erroneous predictions that can distort structural modeling and misguide alignment in consistency learning. Therefore, we construct hyperedges among uncertain regions using a composite affinity measure that incorporates both feature similarity and modulated uncertainty:

$$w_{ij} = \exp\left(-\|z_i - z_j\|_2/\tau\right) \cdot \left(\ 1 + \gamma(\ U'(i) + U'(j) \ ) \ \right), \tag{6}$$

where γ controls the influence of uncertainty. Similar to the construction of certain-region hypergraphs, the top-k neighbors of each node z_i under this weighting scheme form a hyperedge $e_j \in \mathcal{E}_u$. Here, $U'(i) = \widetilde{U}(i) \cdot \delta_i$ is a noise-suppressed uncertainty estimate that preserves diagnostically relevant ambiguous structures

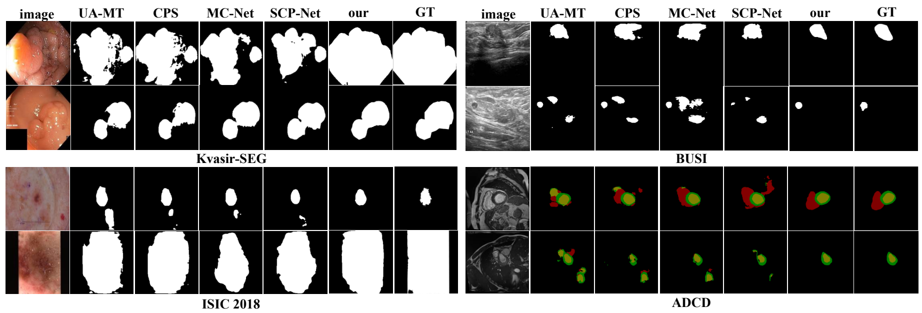

Fig. 2. Visual comparison of SSMIS methods using 5% labeled data on four datesets

while suppressing noise via density modulation. $\tilde{U}(i)$ denotes the fused uncertainty from teacher and student networks. The modulation factor δ_i is derived from feature density, as noisy pixels often exhibit both high uncertainty and feature-wise isolation [29]. For each node z_i, we compute:

$$\delta_i = \exp(-\frac{1}{\beta \cdot k'} \sum_{j=1}^{k'} \|z_i - z_j\|_2), \tag{7}$$

where k' is the number of nearest neighbors and β controls sensitivity. This formulation downweights isolated, noisy nodes while preserving uncertain but structurally coherent regions.

Hyperedge uncertainty and alignment. Hyperedge uncertainty captures the collective reliability of structurally related pixels, which offers a holistic view that complements pixel-wise ambiguity to guide adaptive weighting in consistency learning. Therefore, we aggregate node-wise uncertainties within each hyperedge $e_j \in \mathcal{E}_u$ to obtain its overall uncertainty measure: $U_e(j) = \frac{1}{|e_j|} \sum_{z_i \in e_j} U'(i)$. This collective uncertainty is used to reweight the local and global consistency learning objectives, formulated as:

$$\mathcal{L}_{\text{uncert}} = \mathbb{E}_{e_j \in \mathcal{E}_u} \left[(1 + \gamma U_e(j))(\mathcal{L}_{\text{he}}(j) + \lambda \mathcal{L}_{\text{dist}}(j)) \right]. \tag{8}$$

where $\mathcal{L}_{\text{he}}$ and $\mathcal{L}_{\text{dist}}$ follow the definitions in Eq. (5). The reweighting scheme adaptively emphasizes hyperedges that are uncertain yet diagnostically significant, while suppressing those dominated by noise, thereby enabling more robust structural learning in challenging regions.

2.4 Overall Objective

The total training objective of UAHC combines four terms:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{ce}\big(P_{\text{s}}(x_l),\, y_l\big) + \mathcal{L}_{ce}\big(P_{\text{s}}(x_u^s),\, P_{\text{t}}(x_u^w)\big) + \theta_1 \cdot \mathcal{L}_{\text{certain}} + \theta_2 \cdot \mathcal{L}_{\text{uncert}}, \tag{9}$$

where $\mathcal{L}_{ce}$ denotes cross-entropy, P_{s} and P_{t} are the student and teacher predictions, which enforces supervised learning on labeled data while encouraging

Table 1. Comparison of semi-supervised segmentation on four datasets.

Method	Scans used		Kvasir-SEG		ISIC-2018		BUSI		ACDC		
	Lab.	**Unlab.**	**Dice**	**Jaccard**	**Dice**	**Jaccard**	**Dice**	**Jaccard**	**Dice**	**Jaccard**	**ASD**
MT [21]	**5%**	**95%**	80.75	68.01	87.82	78.69	60.38	45.71	48.85	37.48	22.49
UA-MT [27]			82.17	70.08	88.63	79.80	63.01	47.57	59.68	48.74	6.88
CPS [7]			83.35	71.78	86.31	76.64	67.85	52.13	65.48	55.29	2.71
MC-Net [24]			83.97	72.69	84.41	74.13	68.56	55.43	64.60	53.40	4.29
URPC [19]			86.58	76.58	88.37	79.52	69.80	54.07	48.01	37.65	4.88
SCP-Net [28]			79.46	66.28	86.16	76.12	69.23	53.79	68.11	57.59	4.88
ABD [8]			86.06	76.47	87.38	79.72	71.34	57.22	87.89	78.70	2.37
UARC [25]			86.87	76.91	89.20	80.71	72.45	57.97	77.21	79.05	3.95
UAHC			**87.62**	**78.38**	**90.04**	**82.21**	**73.10**	**58.41**	**89.39**	**81.87**	**1.12**
MT	**10%**	**90%**	84.46	73.69	87.99	79.08	62.96	47.23	79.15	67.88	2.42
UA-MT			85.58	75.10	88.84	80.32	65.85	50.74	84.96	74.98	2.10
CPS			87.15	78.08	86.31	76.64	69.85	54.13	85.85	76.21	1.87
MC-Net			87.13	77.72	85.12	74.82	71.30	56.43	85.51	75.90	1.68
URPC			86.59	76.59	89.40	81.08	71.71	56.82	82.62	71.84	1.68
SCP-Net			86.33	76.56	89.50	81.16	73.13	58.22	87.20	78.08	1.67
ABD			87.52	79.36	86.71	78.79	73.87	60.41	89.42	80.84	1.73
UARC			88.18	79.41	89.79	81.76	74.92	60.92	88.94	80.62	1.62
UAHC			**88.71**	**80.22**	**90.92**	**83.47**	**75.95**	**62.16**	**91.09**	**84.71**	**0.59**

Table 2. Ablation studies on Kvasir-SEG *(5% labeled data, NS: Noise Suppression).*

ResNet[14]+ASPP[6] *(Baseline)*	CHC	UHC		Dice	Jaccard	mIoU	GFLOPs
		w/o NS	with NS				
✓				86.16	76.02	86.13	62.77
✓	✓			87.10	77.47	87.05	64.76 (+1.99)
✓		✓		86.73	76.98	86.41	62.88 (+0.11)
✓	✓	✓		87.09	77.35	87.04	64.87 (+2.10)
✓	✓		✓	**87.62**	**78.38**	**87.26**	64.91 (+2.14)

teacher–student consistency on unlabeled data. The losses $\mathcal{L}_{\text{certain}}$ (Sec. 2.2) and $\mathcal{L}_{\text{uncert}}$ (Sec. 2.3) enforce hypergraph consistency in their respective regions. θ_1 and θ_2 are used to balance the contributions of the two components.

3 Expriments

We evaluate UAHC on four public medical segmentation datasets: Kvasir-SEG [15] (1,000 polyp images), ISIC 2018 [9] (2,594 skin lesions), BUSI [1] (780 breast ultrasound), and ACDC [5] (1,600 cardiac MRI slices). Experiments are conducted on a single NVIDIA RTX 3090 GPU. We use SGD with learning rate 0.005, weight decay 0.0001, batch size 8, and resize inputs to 224×224. Training runs for 100 epochs, with the first 75 using only labeled data and the last 25 incorporating our UAHC framework on unlabeled data to prevent overfitting. RandAugment [10] provides strong/weak augmentations. Each dataset is split 80%/10%/10% for training/validation/testing. All comparison methods follow identical settings and augmentations. Key hyperparameters are empirically set: uncertainty weight $\gamma = 4$, noise suppression $\beta = 0.5$, hyperedge size $K = 16$, distribution weight $\lambda = 0.3$, loss weights $\theta_1 = 0.5$, $\theta_2 = 1$, and feature density neighbors $k' = 5$. Further analysis of these hyperparameters is provided later.

Result Comparisons with SOTAs. Table 1 presents quantitative results under 5% and 10% labeled data. UAHC consistently outperforms both general and uncertainty-aware SSMIS methods across all datasets and metrics. For low-contrast image datasets like BUSI, UAHC remains robust and stable. On

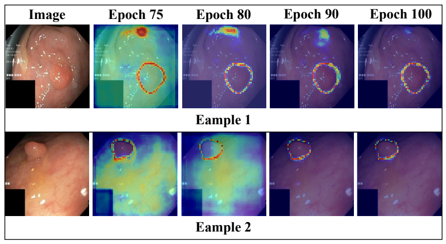

Fig. 3. Uncertainty reduction over training epochs after applying UAHC

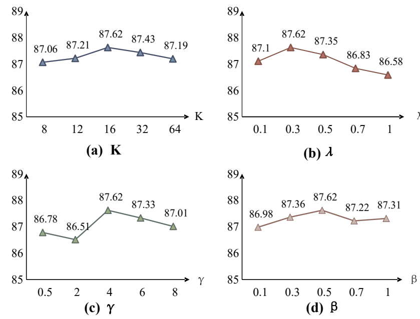

Fig. 4. Analysis of Hyperparameters

the multi-class ACDC dataset, boundary metrics (ASD) confirm that UAHC produces accurate and sharp boundaries around lesions. Overall, these results demonstrate that UAHC is capable of handling complex anatomies, low-contrast regions, and multi-class segmentation tasks across diverse datasets. Fig. 2 shows that UAHC generates more accurate, complete masks with smooth boundaries, while baseline methods tend to under-segment lesions (Kvasir-SEG, ISIC 2018) or over-segment structures (BUSI, ACDC). Notably, on ACDC, UAHC achieves precise segmentation with accurate boundaries, highlighting its strength in modeling high-order anatomical relations and effectively handling uncertain regions.

Uncertainty Reduction. As illustrated in Fig. 3, prediction uncertainty progressively declines throughout training, accompanied by sharper boundaries and reduced background noise. This trend confirms that UAHC effectively disentangles valuable uncertainty from noise, enabling the model to refine complex lesion structures while improving overall confidence and stability.

Ablation Studies. Table 2 shows that incorporating the CHC or UHC module alone improves performance: CHC enforces structural consistency in stable regions, while UHC facilitates learning from ambiguous yet informative areas. Without NS, combining CHC and UHC yields marginal gains, as the absence of suppression allows noise to propagate and disrupt consistency learning. With NS, UHC effectively filters out noise, allowing CHC and UHC to complement each other synergistically. The full UAHC achieves the highest Dice score of 87.62%, demonstrating the complementary nature of the two modules.

Efficiency. As shown in Table 2, the additional computational overhead of UAHC mainly stems from the CHC module, which constructs and aligns hypergraphs in certain regions. Despite this, the full framework adds only 2.14 GFLOPs compared to the baseline (62.77 GFLOPs), demonstrating that the performance gains are achieved with minimal additional cost.

Hyperparameters. Fig. 4 analyzes four key hyperparameters: (a) Nodes $K = 16$ optimally capture local context; smaller values miss it, larger ones add noise. (b) Distribution weight $\lambda = 0.3$ best balances global and local alignment. (c) Uncertainty coefficient $\gamma = 4$ emphasizes uncertain regions, improving boundaries. (d) Noise suppression $\beta = 0.5$ optimally balances uncertainty exploitation and noise. Notably, UAHC performance is relatively insensitive to θ_1 and θ_2.

4 Conclusion

We propose UAHC, a semi-supervised framework that leverages hypergraph modeling to capture high-order anatomical relationships. By harnessing uncertainty as an informative guide, our CHC and UHC modules collaboratively enforce structure-aware consistency, achieving state-of-the-art results across multiple medical segmentation benchmarks. Future work includes extending UAHC to 3D volumetric data and incorporating task-specific anatomical priors.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grants 62472105 and 62302413, in part by the Natural Science Foundation of Guangdong Province under Grants 2025A1515011385 and 2024A1515010186, in part by the Youth Scientific Talent Cultivation Program of the Guangdong Provincial Association for Science and Technology under Grant SKXRC2026513, in part by the Research Foundation of Guangdong-Hong Kong-Macao Applied Mathematics Center under Grant 2025A1515060016, in part by the Health and Medical Research Fund under Grants 23220312 and 09230217, as well as the Guangdong Provincial Key Laboratory of Ultra High Definition Immersive Media Technology under Grant 2024B1212010006.

Disclosure of Interests. The authors declare that they have no known competing interests that could have appeared to influence the work reported in this paper.

References

1. Al-Dhabyani, W., Gomaa, M., Khaled, H., Fahmy, A.: Dataset of breast ultrasound images. Data in Brief p. 104863 (Feb 2020). https://doi.org/10.1016/j.dib.2019.104863
2. Aljuaid, A., Anwar, M.: Survey of supervised learning for medical image processing. SN Computer Science **3**(4), 292 (2022)
3. Ashtari, P., Sima, D.M., De Lathauwer, L., Sappey-Marinier, D., Maes, F., Van Huffel, S.: Factorizer: A scalable interpretable approach to context modeling for medical image segmentation. Medical image analysis **84**, 102706 (2023)
4. Basak, H., Yin, Z.: Pseudo-label guided contrastive learning for semi-supervised medical image segmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 19786–19797 (2023)
5. Bernard, O., Lalande, A., Zotti, C., Cervenansky, F., Yang, X., Heng, P.A., Cetin, I., Lekadir, K., Camara, O., Ballester, M.A.G., et al.: Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: is the problem solved? IEEE transactions on medical imaging **37**(11), 2514–2525 (2018)
6. Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European conference on computer vision (ECCV). pp. 801–818 (2018)
7. Chen, X., Yuan, Y., Zeng, G., Wang, J.: Semi-supervised semantic segmentation with cross pseudo supervision. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 2613–2622 (2021)
8. Chi, H., Pang, J., Zhang, B., Liu, W.: Adaptive bidirectional displacement for semi-supervised medical image segmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 4070–4080 (2024)
9. Codella, N., Rotemberg, V., Tschandl, P., Çelebi, M., Dusza, S., Gutman, D., Helba, B., Kalloo, A., Liopyris, K., Marchetti, M., Kittler, H., Halpern, A.: Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). Cornell University - arXiv,Cornell University - arXiv (Feb 2019)
10. Cubuk, E.D., Zoph, B., Shlens, J., Le, Q.V.: Randaugment: Practical automated data augmentation with a reduced search space. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (Jun 2020). https://doi.org/10.1109/cvprw50498.2020.00359, http://dx.doi.org/10.1109/cvprw50498.2020.00359
11. Dong, P., Guo, Y., Gao, Y., Liang, P., Shi, Y., Wu, G.: Multi-atlas segmentation of anatomical brain structures using hierarchical hypergraph learning. IEEE Transactions on Neural Networks and Learning Systems **31**(8), 3061–3072 (2020). https://doi.org/10.1109/TNNLS.2019.2935184
12. Feng, Y., You, H., Zhang, Z., Ji, R., Gao, Y.: Hypergraph neural networks. In: Proceedings of the AAAI conference on artificial intelligence. vol. 33, pp. 3558–3565 (2019)
13. Han, K., Sheng, V.S., Song, Y., Liu, Y., Qiu, C., Ma, S., Liu, Z.: Deep semi-supervised learning for medical image segmentation: A review. Expert Systems with Applications **245**, 123052 (2024)
14. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)

15. Jha, D., Smedsrud, P., Riegler, M., Halvorsen, P., Lange, T., Johansen, D., Johansen, H.: Kvasir-seg: A segmented polyp dataset. arXiv: Image and Video Processing,arXiv: Image and Video Processing (Nov 2019)
16. Kurz, A., Hauser, K., Mehrtens, H.A., Krieghoff-Henning, E., Hekler, A., Kather, J.N., Fröhling, S., Von Kalle, C., Brinker, T.J.: Uncertainty estimation in medical image classification: systematic review. JMIR Medical Informatics **10**(8), e36427 (2022)
17. Liu, X., Tang, Z., Wu, Y., Zhai, R., Dong, X., Du, Z., Cao, S.: Mfhs: Mutual consistency learning-based foundation model integrates hypergraph for semi-supervised medical image segmentation. Pattern Recognition p. 112721 (2025)
18. Lu, Y., Li, W., Cui, Z., Zhang, Y.: Beyond low-dimensional features: Enhancing semi-supervised medical image semantic segmentation with advanced consistency learning techniques. Expert Systems with Applications **261**, 125456 (2025)
19. Luo, X., Liao, W., Chen, J., Song, T., Chen, Y., Zhang, S., Chen, N., Wang, G., Zhang, S.: Efficient semi-supervised gross target volume of nasopharyngeal carcinoma segmentation via uncertainty rectified pyramid consistency. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part II 24. pp. 318–329. Springer (2021)
20. Shu, Y., Li, H., Xiao, B., Bi, X., Li, W.: Cross-mix monitoring for medical image segmentation with limited supervision. IEEE Transactions on Multimedia **25**, 1700–1712 (2022)
21. Tarvainen, A., Valpola, H.: Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. Advances in neural information processing systems **30** (2017)
22. Wang, M., Li, J., Su, H., Yin, N., Yang, L., Li, S.: Graphcl: Graph-based clustering for semi-supervised medical image segmentation. arXiv preprint arXiv:2411.13147 (2024)
23. Wang, Y., Xiao, B., Bi, X., Li, W., Gao, X.: Mcf: Mutual correction framework for semi-supervised medical image segmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 15651–15660 (2023)
24. Wu, Y., Ge, Z., Zhang, D., Xu, M., Zhang, L., Xia, Y., Cai, J.: Mutual consistency learning for semi-supervised medical image segmentation. Medical Image Analysis **81**, 102530 (2022)
25. Wu, Y., Li, X., Zhou, Y.: Uncertainty-aware representation calibration for semi-supervised medical imaging segmentation. Neurocomputing **595**, 127912 (2024)
26. Yang, Y., Sun, G., Zhang, T., Wang, R., Su, J.: Semi-supervised medical image segmentation via weak-to-strong perturbation consistency and edge-aware contrastive representation. Medical Image Analysis **101**, 103450 (2025)
27. Yu, L., Wang, S., Li, X., Fu, C.W., Heng, P.A.: Uncertainty-aware self-ensembling model for semi-supervised 3d left atrium segmentation. In: Medical image computing and computer assisted intervention–MICCAI 2019: 22nd international conference, Shenzhen, China, October 13–17, 2019, proceedings, part II 22. pp. 605–613. Springer (2019)
28. Zhang, Z., Ran, R., Tian, C., Zhou, H., Li, X., Yang, F., Jiao, Z.: Self-aware and cross-sample prototypical learning for semi-supervised medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 192–201. Springer (2023)
29. Zhao, X., Qi, Z., Wang, S., Wang, Q., Wu, X., Mao, Y., Zhang, L.: Rcps: Rectified contrastive pseudo supervision for semi-supervised medical image segmentation. IEEE Journal of Biomedical and Health Informatics **28**(1), 251–261 (2023)