UniLiver: Gradient-Conditioned Unified Model for Multi-target Hepatic Segmentation

Yue Qiu, Shi Qiu (✉), Pheng-Ann Heng, and Chi-Wing Fu

Department of Computer Science and Engineering
Institute of Medical Intelligence and XR
The Chinese University of Hong Kong
{yqiu23, shiqiu, pheng, cwfu}@cse.cuhk.edu.hk

Abstract. Segmentations of hepatic vessels, Couinaud segments, and tumors are crucial for precision liver surgery. Existing methods mostly handle them independently, without exploiting their correlation in the segmentations. We propose UniLiver, a unified model for multi-target hepatic segmentation. Specifically, we introduce Multi-Target Gradient-Conditioned Adapter (MTGCA) to resolve the optimization conflicts. By this new means, we address the morphological variations in different hepatic targets and allow the shared encoder to serve each target with minimized mutual degradation. In addition, to handle vast scale and shape disparities among liver anatomical data, we design Target-Aware Spectral-Spatial Encoding (TA-SSE) to modulate spatial positional cues as per target geometric demands. Further, we present VCT Sequential Denoising (VCT-SD) to progressively refine predictions following the anatomical hierarchy from vessels to segments and tumors. Evaluated on the composite LiTS and MSD8 datasets, UniLiver achieves consistent improvements over state-of-the-art methods for specific hepatic targets, confirming its compelling performance of jointly modeling the anatomical dependencies for precision liver surgery. Code is available at https://github.com/YueQiu0911/UniLiver.

Keywords: Unified model · Hepatic vessel segmentation · Couinaud segmentation · Liver tumor segmentation

1 Introduction

Automated analysis of hepatic anatomy is crucial for surgical planning, requiring simultaneous delineation of hepatic vessels, Couinaud segments, and tumors [20,12,13,22]. Anatomically, these structures are interlinked: the vascular topology physically defines the functional liver segments (i.e., Couinaud segments), which in turn serve as the spatial containers for potential lesions [18,5,15,23]. However, existing medical AI models predominantly delineate these liver anatomical structures as independent segmentation targets, neglecting the structural synergies inherent in anatomical hierarchy. Usually, vessel segmentation [17,29,14,28,7] and tumor segmentation methods [11,27] overlook the vascular-segmental priors that could narrow the tumor search space. Although

recent studies attempt to incorporate vascular constraints for Couinaud segmentation, they typically rely on static ground-truth masks or separate preprocessing steps, preventing the two targets from mutually refining each other [30,31,6].

In this work, we propose UniLiver, a comprehensive unified model that fundamentally rethinks multi-structure hepatic segmentation not as a disparate collection of independent objectives but as a deeply interconnected physiological system. To achieve this, we leverage a joint representation learning strategy. By employing a shared backbone to extract common morphological features, UniLiver minimizes the redundant computation inherent in deploying multiple target-specific encoders. However, naively aggregating these heterogeneous objectives often leads to task interference and optimization instability: for example, the high-frequency features required for fine-grained vessel segmentation may conflict with the low-frequency, region-level context of volumetric segment delineation, causing gradient inconsistencies during backpropagation. Unlike typical joint learning architectures that employ static target weighting or shared-private encoder splits [24,21], we introduce Multi-Target Gradient-Conditioned Adapter (MTGCA) to resolve target interference by conditioning shared feature adaptation on inter-target gradient coherence. Particularly, our MTGCA maintains a per-target gradient bank and applies orthogonal decomposition to isolate the conflicting gradient component for each target at every training step, using this signal to adaptively correct the feature modulation parameters so that cooperative gradients are maintained while conflicting ones are suppressed.

To handle the vast morphological disparities among different liver anatomical targets, we design Target-Aware Spectral-Spatial Encoding (TA-SSE) that learns spectral filtering with multi-scale spatial encoding. This design is uniquely target-adaptive: a semantic embedding dynamically regulates positional cues, enhancing coordinate awareness for anatomically consistent structures (e.g., vessels) while suppressing it for location-invariant targets like tumors. Furthermore, we propose VCT Sequential Denoising (VCT-SD) to progressively refine predictions following the anatomical hierarchy. Specifically, our VCT-SD performs multiple refinement steps with directed cross-attention that propagates structural cues along the anatomical dependency chain, i.e., from vessels to segments and from both to tumors, allowing each segmentation target to capture increasingly refined geometric priors from their anatomical predecessors. To comprehensively evaluate the effectiveness of our proposed method, we conduct qualitative and quantitative evaluations on two most widely-used liver benchmarks, LiTS [2] and MSD8 [26], and compare against strong baselines and state-of-the-art methods.

2 Methodology

As Fig. 1 shows, UniLiver integrates MTGCA, TA-SSE, and VCT-SD into a DynUNet backbone [9]. In general, MTGCA mitigates multi-target optimization conflicts, TA-SSE extracts target-adaptive features, and VCT-SD progressively refines predictions following anatomical hierarchy.

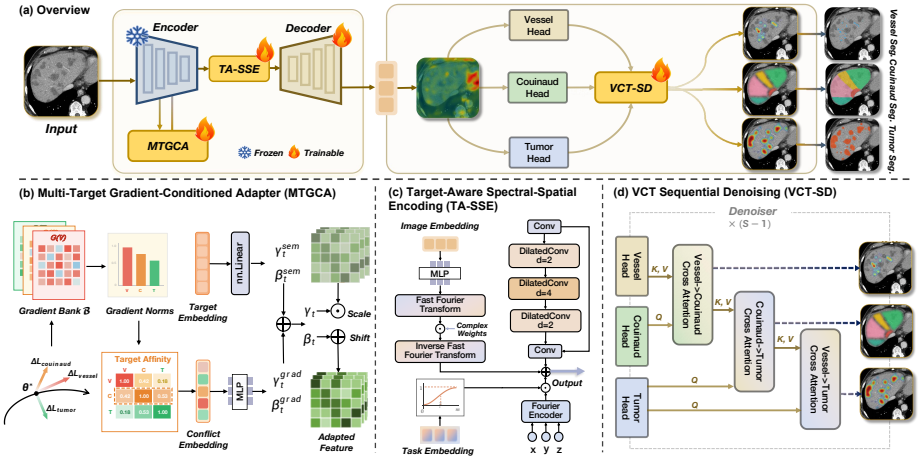

Fig. 1. (a) The overview of our UniLiver for multi-target hepatic segmentation (b) Multi-Target Gradient-Conditioned Adapter (MTGCA), (c) Target-Aware Spectral-Spatial Encoding (TA-SSE), and (d) VCT Sequential Denoising (VCT-SD). The framework employs an enhanced encoder-decoder backbone with MTGCA for adaptive multi-target feature modulation, TA-SSE for spatial-aware target embedding, and VCT-SD for iterative cross-target feature refinement via sequence cross-attention.

2.1 Multi-Target Gradient-Conditioned Adapter

For multi-target hepatic segmentation, the shared encoder is expected to implicitly represent the context of vascular boundaries, segmental regions, and tumor textures. These structurally distinct targets impose competing feature demands: for instance, capturing thin vessels requires the encoder to sharpen high-frequency edge responses, while delineating Couinaud segments requires it to maintain smooth regional consistency at the same layers, causing the shared parameters to oscillate between conflicting or compromised optimization directions. Conventional learning strategies rely on fixed or globally re-weighted target balancing, which cannot resolve such layer-wise varying conflicts. To address this, we propose the Multi-Target Gradient-Conditioned Adapter (MTGCA), which applies inter-target gradient coherence to condition feature modulation at each encoder stage, dynamically maintaining cooperative gradient directions while redirecting conflicting ones.

Formally, let $\nabla_\theta \mathcal{L}_t \in \mathbb{R}^p$ denote the gradient of target t with respect to shared parameters θ. We project these gradients onto a compact d-dimensional subspace using a random projection matrix $\mathbf{P} \in \mathbb{R}^{p \times d}$, and maintain a Gradient Bank $\mathcal{B} = \{\mathbf{g}_t\}_{t=1}^n$ storing the projected gradients for all n targets:

$$\mathbf{g}_t = \mathbf{P}^\top \nabla_\theta \mathcal{L}_t \in \mathbb{R}^d. \tag{1}$$

From the Gradient Bank, we construct a target affinity matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ where each entry $a_{i,j} = \langle \hat{\mathbf{g}}_i, \hat{\mathbf{g}}_j \rangle$ represents the cosine similarity between normalized

gradients $\hat{\mathbf{g}}_i = \mathbf{g}_i / \|\mathbf{g}_i\|_2$. The conflict vector for target t is extracted as the t-th row of the affinity matrix: $\mathbf{c}_t = \mathbf{A}_{t,:} \in \mathbb{R}^n$, encapsulating the relational structure between target t and all other targets. We then integrate $\mathbf{c}_t$ for an extended feature-wise linear modulation. Given a feature tensor $\mathbf{F} \in \mathbb{R}^{B\times C\times D\times H\times W}$ and target embedding $\mathbf{e}_t \in \mathbb{R}^{d_e}$, the modulation is computed as:

$$\mathbf{F}' = (1 + \gamma_t(\mathbf{e}_t, \mathbf{c}_t)) \odot \mathbf{F} + \beta_t(\mathbf{e}_t, \mathbf{c}_t), \tag{2}$$

where the scale and shift parameters are decomposed as $\gamma_t = \gamma_t^{\text{sem}} + \gamma_t^{\text{grad}}$ and $\beta_t = \beta_t^{\text{sem}} + \beta_t^{\text{grad}}$, with $[\gamma_t^{\text{sem}}, \beta_t^{\text{sem}}] = \phi(\mathbf{e}_t)$ capturing semantic target identity, and $[\gamma_t^{\text{grad}}, \beta_t^{\text{grad}}] = \psi(\mathbf{c}_t)$ encoding gradient-derived relational signals. This decomposition enables the proposed adapter to enhance edge-sensitive features when vessel gradients dominate, smooth regional responses guided by segment gradients, and preserve texture-discriminative cues driven by tumor gradients. As a result, the shared encoder is exploited to serve all three hepatic segmentation targets without mutual compromise.

2.2 Target-Aware Spectral-Spatial Encoding

In volumetric medical imaging, hepatic structures such as vessels, Couinaud segments, and tumors exhibit distinct characteristics across both spatial and frequency domains: vessels appear as elongated tubular patterns with specific spatial continuity; Couinaud segments are defined by their anatomical positions relative to hepatic veins; and tumors present as localized lesions requiring precise boundary delineation. Existing encoder-decoder architectures usually operate in the spatial domain, overlooking the rich spectral information that can help distinguish spatially similar yet semantically distinct regions. To bridge this gap, we propose Target-Aware Spectral-Spatial Encoding (TA-SSE), a module that jointly models spectral characteristics and positional information at multiple scales of the network. Given an encoder output feature map $\mathbf{F} \in \mathbb{R}^{B\times C\times D\times H\times W}$, TA-SSE produces enhanced representations through the integration of frequency-domain filtering and positional awareness:

$$\hat{\mathbf{F}} = \mathbf{F} + \alpha \cdot \mathcal{G}_\omega(\mathbf{F}) + \beta \cdot \psi\left(\mathbf{F}, \mathbf{e}_t, \gamma(\mathbf{p})\right), \tag{3}$$

where $\mathcal{G}_\omega(\mathbf{F}) = \mathcal{F}^{-1}\left(\mathcal{F}(\mathbf{F}) \odot \sum_{k=1}^{K} r_k \mathbf{W}_c^{(k)}\right)$ denotes spectral filtering with learnable complex-valued filter banks $\mathbf{W}_c^{(k)} \in \mathbb{C}$, input-adaptive routing weights $\mathbf{r} = [r_1, \ldots, r_K]^\top \in \mathbb{R}^K$, and the 3D Fast Fourier Transform (FFT) $\mathcal{F}$. The spectral filtering term enables the network to selectively enhance or suppress frequency components based on semantic content. The positional term employs Fourier feature mapping $\gamma(\mathbf{p})$ to encode 3D coordinates $\mathbf{p} = (x, y, z) \in [-1, 1]^3$ into high-dimensional sinusoidal representations. Notably, the contribution of positional encoding is dynamically weighted by target embedding $\mathbf{e}_t$, allowing the network to learn target-specific reliance on spatial information: vessel segmentation benefits from stronger positional cues to maintain tubular continuity, while tumor detection may rely more on local appearance than absolute location. We

apply TA-SSE hierarchically at encoder skip connections with frequency-domain filtering, and at the bottleneck with additional multi-scale dilated convolutions, ensuring that both fine-grained local patterns and global structural context are effectively captured.

2.3 VCT Sequential Denoising

The target-specific feature maps produced by the backbone modules inevitably involve prediction noise arising from the entanglement of multiple correlated anatomical structures within a shared representation space. To mitigate this, we introduce a VCT Sequential Denoising (VCT-SD) module that progressively enhances target-specific representations through structured inter-target interactions over multiple refinement steps.

Central to our formulation is the observation that the three target structures, i.e., vessel V, Couinaud segments C, and tumor T, are associated by a directed anatomical dependency: vessel topology defines Couinaud segment partitioning, while both vascular proximity and segmental identity are associated with tumor localization. We encode this hierarchy as a directed acyclic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{V, C, T\}$ denotes the target nodes and $\mathcal{E} = \{V \to C,\ C \to T,\ V \to T\}$ encodes directed conditioning edges. Each edge is instantiated as a cross-attention block in which the downstream target queries structural cues from the upstream target: i.e., Couinaud features leverage vessel features to perceive branching topology, and tumor features leverage both vessel and segment features to incorporate vascular proximity and segmental context. The three anatomical segmentation pathways are operated in topological order, ensuring that vessel representations are refined first, while subsequent segmentation targets progressively incorporate enriched structural priors.

Practically, the refinement unfolds over S iterative steps, each conditioned on a sinusoidal step embedding $\mathbf{e}_s$ that encodes the current stage, allowing the module to apply coarse corrections in early steps and fine-grained adjustments in later ones. Formally, the single-step update is:

$$\mathbf{F}_j^{(s+1)} = \mathbf{F}_j^{(s)} + \sum_{i:\,(i\to j)\,\in\,\mathcal{E}} \mathcal{A}_{i\to j}\Big(\mathbf{F}_j^{(s)},\ \mathbf{F}_i^{(s)},\ \mathbf{e}_s\Big); \quad s = 0, 1, \ldots, S-1; \tag{4}$$

where $\mathcal{A}_{i\to j}$ is a cross-attention block in which target j attends to target i. Upon completion of S steps, the enhanced features $\mathbf{F}_V^{(S)}, \mathbf{F}_C^{(S)}, \mathbf{F}_T^{(S)}$ are projected through target-specific output layers and forwarded to the segmentation heads.

3 Experiments

3.1 Experimental Setup

To validate the effectiveness of UniLiver, we utilize the LiTS [2] and MSD8 [26] datasets for hepatic vessel, tumor, and Couinaud segmentation targets. Both datasets are randomly split into training and testing sets with a ratio of 8:2. In

Table 1. Quantitative comparisons of three hepatic segmentation targets. UniLiver (Single) is trained on a single dataset, while UniLiver (Multi) is trained on the combined two datasets. Red denotes the best, and orange denotes the second best.

Target 1: Couinaud Segmentation

Method	LiTS Dataset			MSD8 Dataset		
	Dice↑	ACC↑	ASD↓	Dice↑	ACC↑	ASD↓
3D UNet [4]	74.91	63.06	1.9598	88.31	79.29	1.8299
Swin UNetR [8]	73.53	62.36	1.5375	85.33	74.79	0.9309
Dual-Atten [10]	79.43	68.05	1.6914	85.75	75.43	1.0271
UNetR++ [25]	79.27	67.90	1.0781	88.59	79.17	1.1180
APVN [30]	88.98	80.59	2.9741	89.56	81.33	1.7776
UniLiver (Single)	90.15	82.21	0.6285	89.83	81.85	**0.7310**
UniLiver (Multi)	**91.15**	**83.88**	**0.3470**	**90.07**	**82.21**	0.7469

Target 2: Vessel Segmentation

Method	LiTS Dataset				MSD8 Dataset			
	Dice↑	VOE↓	ASSD↓	HD95↓	Dice↑	VOE↓	ASSD↓	HD95↓
TransFusionNet[19]	54.17	60.12	4.6603	21.26	57.73	57.96	5.3986	26.10
VesselFM [28]	55.49	60.46	3.1734	19.04	56.08	60.32	3.6045	19.19
3D UNet[4]	60.67	54.52	4.8852	25.68	63.66	50.81	5.6512	33.39
MFSNet [1]	56.03	59.58	4.8485	26.17	64.90	51.13	4.6503	25.46
LSFNet [7]	64.62	51.18	4.5359	23.84	60.74	54.72	6.3835	34.24
UniLiver (Single)	**69.18**	**46.68**	1.2917	10.63	67.48	48.66	2.0121	16.33
UniLiver (Multi)	69.15	46.71	**1.2672**	**10.61**	**67.57**	**48.54**	**1.9111**	**15.22**

Target 3: Tumor Segmentation

Method	LiTS Dataset			MSD8 Dataset		
	G-Dice↑	C-Dice↑	VOE↓	G-Dice↑	C-Dice↑	VOE↓
TransUNet [3]	74.07	44.88	64.86	76.04	47.46	49.75
nnUNet [9]	74.65	54.59	62.04	72.56	44.87	57.92
RollingUNet [16]	72.67	42.50	67.86	77.11	50.42	54.31
TexLiverNet [11]	80.20	64.23	52.98	76.21	64.45	51.45
SBCNet [27]	82.13	65.44	51.97	77.72	67.12	49.49
UniLiver (Single)	78.54	69.54	48.60	80.23	69.35	48.71
UniLiver (Multi)	**83.86**	**69.68**	**47.46**	**80.56**	**71.82**	**45.07**

addition to the standard evaluation on each benchmark, we also report results under a joint-training setting, where we train a unified model on the combined training splits of both datasets and test on each dataset respectively. All experiments are conducted on a single NVIDIA RTX 4090 GPU using PyTorch. For fair comparisons, we train and test all comparison methods with the same experimental settings and configurations. We adopt a DynUNet backbone [9] initialized with pre-trained VesselFM weights [28]. We apply the AdamW optimizer with an initial learning rate of 2×10^{-5}, coupled with a step decay strategy that reduces the learning rate by a factor of 0.2 every 5 epochs. The training batch size and epochs are set to 2 and 150 respectively. We apply the combination of

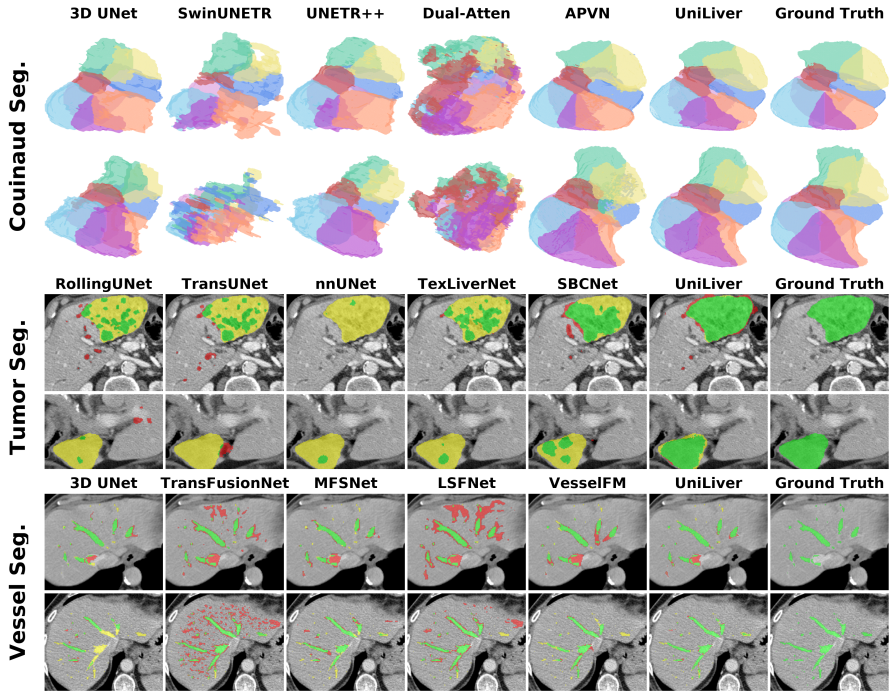

Fig. 2. Qualitative comparison of three hepatic segmentation targets. For tumor and vessel segmentation, green indicates *correct* predictions, red denotes *incorrect* predictions, and yellow marks *missed* regions.

Dice and Cross-Entropy losses for the three segmentation targets. Empirically, we set the target loss weights to $\lambda_{vessel} = 1.0$, $\lambda_{tumor} = 3.0$, and $\lambda_{couinaud} = 1.0$ to balance the optimization.

3.2 Comparison with State-of-the-art Methods

Couinaud Segmentation. As shown in Tab. 1, UniLiver achieves a Dice of 91.15% on LiTS, surpassing APVN [30] by 2.17%, along with the lowest ASD of 0.3470 mm. As illustrated in the first row of Fig. 2, UniLiver produces segment boundaries that most closely align with the ground truth, whereas competing methods such as SwinUNETR [8] and Dual-Atten [10] exhibit noticeable segment misclassification and fragmented inter-segment boundaries.

Vessel Segmentation. UniLiver consistently outperforms all compared methods across both datasets. Compared to the LSFNet [7], our framework improves Dice by 4.53% on LiTS and reduces HD95 from 23.84 mm to 10.61 mm, suggesting improved recovery of distal vessel branches that are typically difficult to segment, as illustrated in the third row of Fig. 2, where UniLiver faithfully recov-

Table 2. Ablation Study of key modules. The best results are highlighted in bold. M, T, and S denote the MTGCA, TA-SSE, and VCT-SD modules, respectively.

	Mod.			Couinaud			Vessel				Tumor		
	M	T	S	Dice↑	ACC↑	ASD↓	Dice↑	VOE↓	ASSD↓	HD95↓	G-Dice↑	C-Dice↑	VOE↓
LiTS				87.95	78.82	0.567	66.52	49.48	1.487	11.89	75.81	54.12	58.42
	✓			88.89	80.33	0.539	68.01	48.00	1.506	12.76	78.18	63.99	53.77
	✓	✓		89.32	80.95	0.480	68.37	47.57	1.414	11.83	79.29	66.76	51.52
	✓	✓	✓	**91.15**	**83.88**	**0.347**	**69.15**	**46.71**	**1.267**	**10.61**	**83.86**	**69.68**	**47.46**
MSD8				87.05	77.55	1.049	67.00	49.23	1.961	**14.37**	68.98	61.09	49.23
	✓			87.88	78.73	1.010	67.39	48.80	1.916	14.56	69.86	57.18	60.78
	✓	✓		88.13	79.25	0.918	67.56	48.74	1.968	15.38	72.99	65.28	52.64
	✓	✓	✓	**90.07**	**82.21**	**0.747**	**67.57**	**48.54**	**1.911**	15.22	**80.56**	**71.82**	**45.07**

Table 3. Ablation of Spatial and Spectral components in TA-SSE on LiTS dataset.

TA-SSE		Couinaud			Vessel				Tumor		
T_{spe}	T_{spa}	Dice↑	ACC↑	ASD↓	Dice↑	VOE↓	ASSD↓	HD95↓	G-Dice↑	C-Dice↑	VOE↓
✓		90.27	82.39	0.5855	68.94	47.51	1.2765	10.6591	75.92	68.16	49.85
	✓	89.65	81.68	0.4041	69.09	47.39	1.2773	10.6628	79.66	62.55	55.10
✓	✓	**91.15**	**83.88**	**0.3470**	**69.15**	**46.71**	**1.2672**	**10.6111**	**83.86**	**69.68**	**47.46**

ers fine peripheral branches while other methods either miss them entirely (e.g., MFSNet [1]) or introduce substantial false positives (e.g., TransFusionNet [19]).
Tumor Segmentation. UniLiver achieves the best performance in both Global Dice (G-Dice) and Per-Case Dice (C-Dice). On the MSD8 dataset, our model reaches a G-Dice of 80.56% and a C-Dice of 71.82%, outperforming the recent SBCNet [27] by 2.84% and 4.70%, respectively. The notable gain in C-Dice indicates more consistent detection across individual patients, including challenging cases with small or multiple lesions. As shown in Fig. 2, UniLiver correctly identifies nearly all tumor regions, whereas methods such as TransUNet [3] exhibit large missed regions, and RollingUNet [16] produces scattered false positives.

3.3 Ablation Study

To evaluate our modules, we conduct ablation studies on LiTS and MSD8 (Tab. 2) against a vanilla multi-target DynUNet baseline. Introducing MTGCA brings Couinaud Dice, G-Dice, and C-Dice improvements of 0.94%, 2.37%, and 9.87% on LiTS. Further integrating TA-SSE reaches 89.32% Couinaud Dice and reduces vessel ASSD to 1.414 mm. Finally, incorporating VCT-SD achieves the best overall metrics, notably yielding substantial tumor C-Dice improvements of 15.56% and 10.73% over the baseline on LiTS and MSD8.

Furthermore, Tab. 3 explicitly ablates the spectral (T_{spe}) and spatial (T_{spa}) branches of TA-SSE on the LiTS dataset. Their combination synergistically aggregates frequency-domain contexts and fine-grained spatial details, exhibiting complementarity over individual branches. This unified design boosts Couinaud Dice to 91.15%, vessel Dice to 69.15%, and tumor G-Dice to 83.86%.

4 Conclusion

In this work, we propose a unified hepatic segmentation framework with three core components: MTGCA dynamically modulates features based on gradient coherence, TA-SSE adaptively fuses spectral-spatial representations to handle morphological disparities, and VCT-SD progressively enforces anatomical hierarchy through iterative refinement. Extensive experiments on LiTS and MSD8 datasets demonstrate state-of-the-art performance across all three targets. In future work, we aim to scale UniLiver with larger and more diverse liver imaging datasets to develop a foundational model for comprehensive hepatic anatomical segmentation.

Acknowledgments. The work described in this paper was supported by the Research Grants Council of the Hong Kong Special Administrative Region, China, under Project T45-401/22-N.

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

References

1. Basak, H., Kundu, R., Sarkar, R.: Mfsnet: A multi focus segmentation network for skin lesion segmentation. Pattern Recognition **128**, 108673 (2022)
2. Bilic, P., Christ, P., Li, H.B., Vorontsov, E., Ben-Cohen, A., Kaissis, G., Szeskin, A., Jacobs, C., Mamani, G.E.H., Chartrand, G., et al.: The liver tumor segmentation benchmark (lits). Medical image analysis **84**, 102680 (2023)
3. Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., Lu, L., Yuille, A.L., Zhou, Y.: Transunet: Transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306 (2021)
4. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3d u-net: learning dense volumetric segmentation from sparse annotation. In: International conference on medical image computing and computer-assisted intervention. pp. 424–432. Springer (2016)
5. d'Albenzio, G., Palomar, R., et al.: Patient-specific functional liver segments based on centerline classification of the hepatic and portal veins. Computer Assisted Surgery (2025)
6. Fan, W., Fang, H., Li, R., Lin, Y., An, C., Luo, X.: Anatomy-aware frequency-attention transformer networks for liver couinaud ct/mr segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 55–65. Springer (2025)
7. Gao, Z., Zong, Q., Wang, Y., Yan, Y., Wang, Y., Zhu, N., Zhang, J., Wang, Y., Zhao, L.: Laplacian salience-gated feature pyramid network for accurate liver vessel segmentation. IEEE Transactions on Medical Imaging **42**(10), 3059–3068 (2023)
8. Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H., Xu, D.: Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. arXiv preprint arXiv:2201.01266 (2022)
9. 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)

10. Jia, X., Qian, C., Yang, Z., Xu, H., Han, X., Ren, H., Wu, X., Ma, B., Yang, D., Min, H.: Boundary-aware dual attention guided liver segment segmentation model. KSII Transactions on Internet and Information Systems (TIIS) **16**(1), 16–37 (2022)
11. Jiang, X., Zhou, Z., Wang, H., Wang, G., Fang, Z.: Texlivernet: Leveraging medical knowledge and spatial-frequency perception for enhanced liver tumor segmentation. In: 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI). pp. 1–5. IEEE (2025)
12. Lèbre, M.A., Magnin, B., et al.: Automatic segmentation methods for liver and hepatic vessels from ct and mri volumes, applied to the couinaud scheme. Computers in Biology and Medicine (2019)
13. Li, C., Tong, Y., Chen, K., Yang, Z., Li, R., Qiu, S., Chan, J.Y.K., Heng, P.A., Dou, Q.: Clipgs: Clippable gaussian splatting for interactive cinematic visualization of volumetric medical data. In: MICCAI. pp. 127–137. Springer (2025)
14. Li, R., Wang, L., et al.: 3d graph-connectivity constrained network for hepatic vessel segmentation. IEEE Journal of Biomedical and Health Informatics (2021)
15. Liu, Q., Qiu, S., Wang, Y., Wu, X., Chok, K.S.H., Fu, C.W., Heng, P.A.: Coordinated 2d-3d visualization of volumetric medical data in xr with multimodal interactions. In: IEEE Visualization and Visual Analytics (VIS). IEEE (2025)
16. Liu, Y., Zhu, H., Liu, M., Yu, H., Chen, Z., Gao, J.: Rolling-unet: Revitalizing mlp's ability to efficiently extract long-distance dependencies for medical image segmentation. In: Proceedings of the AAAI conference on artificial intelligence. vol. 38, pp. 3819–3827 (2024)
17. Machta, Y., Vignon-Clementel, I., et al.: Enhancing the automatic segmentation and analysis of 3d liver vasculature models. arXiv preprint arXiv: (2024)
18. Majno, P., Mentha, G., Toso, C., Morel, P., Peitgen, H.O., Fasel, J.H.: Anatomy of the liver: an outline with three levels of complexity–a further step towards tailored territorial liver resections. Journal of hepatology **60**(3), 654–662 (2014)
19. Meng, X., Zhang, X., Wang, G., Zhang, Y., Shi, X., Dai, H., Wang, Z., Wang, X.: Exploiting full resolution feature context for liver tumor and vessel segmentation via integrate framework: Application to liver tumor and vessel 3d reconstruction under embedded microprocessor. arXiv preprint arXiv:2111.13299 (2021)
20. Oliveira, D.A.B., Correia, M.M., et al.: Segmentation of liver, its vessels and lesions from ct images for surgical planning. BioMedical Engineering OnLine (2011)
21. Qin, X., Wang, X., Yan, J.: Towards consistent multi-task learning: Unlocking the potential of task-specific parameters. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 10067–10076 (2025)
22. Qiu, S., Li, R., Liu, Q., Tong, Y., Qiu, Y., Wang, Y., Li, Y., Fu, C.W., Heng, P.A.: A collaborative extended reality prototype for 3d surgical planning and visualization. arXiv preprint arXiv:2601.19303 (2026)
23. Qiu, Y., Tong, Y., Zhang, Y., Liu, Q., Pei, J., Qiu, S., Heng, P.A., Fu, C.W.: Cvhslicer 2.0: Immersive and interactive visualization of chinese visible human data in xr environments. arXiv preprint arXiv:2503.15507 (2025)
24. Royer, A., Blankevoort, T., Ehteshami Bejnordi, B.: Scalarization for multi-task and multi-domain learning at scale. Advances in Neural Information Processing Systems **36**, 16917–16941 (2023)
25. Shaker, A., Maaz, M., Rasheed, H., Khan, S., Yang, M.H., Khan, F.S.: Unetr++: delving into efficient and accurate 3d medical image segmentation. IEEE Transactions on Medical Imaging **43**(9), 3377–3390 (2024)
26. Simpson, A.L., Antonelli, M., Bakas, S., Bilello, M., Farahani, K., Van Ginneken, B., Kopp-Schneider, A., Landman, B.A., Litjens, G., Menze, B., et al.: A large an-

notated medical image dataset for the development and evaluation of segmentation algorithms. arXiv preprint arXiv:1902.09063 (2019)
27. Wang, K.N., Li, S.X., Bu, Z., Zhao, F.X., Zhou, G.Q., Zhou, S.J., Chen, Y.: Sbc-net: Scale and boundary context attention dual-branch network for liver tumor segmentation. IEEE Journal of Biomedical and Health Informatics **28**(5), 2854–2865 (2024)
28. Wittmann, B., Wattenberg, Y., Amiranashvili, T., Shit, S., Menze, B.: vesselfm: A foundation model for universal 3d blood vessel segmentation. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 20874–20884 (2025)
29. Zhang, D., Comaniciu, D., et al.: Graph attention network based pruning for reconstructing 3d liver vessel morphology from contrasted ct images. arXiv preprint arXiv:2006. (2020)
30. Zhang, X., Ali, S., Liu, T., Zhao, X., Cui, Z., Han, M., Ma, S., Zhu, J., Kang, Y., Wang, L., et al.: Robust and smooth couinaud segmentation via anatomical structure-guided point-voxel network. Computers in Biology and Medicine **182**, 109202 (2024)
31. Zhang, X., Liu, Y., Ali, S., Zhao, X., Sun, M., Han, M., Liu, T., Zhai, P., Cui, Z., Zhang, P., et al.: Anatomical-aware point-voxel network for couinaud segmentation in liver ct. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 465–474. Springer (2023)