Automatic Segmentation of Hepatic and Portal Veins using SwinUNETR and Multi-Task Learning

Shanmugapriya Survarachakan[1]
shanmugapriya.survarachakan@ntnu.no[0000−0002−6576−8438], Michael Staff Larsen[1] michael.s.larsen@ntnu.no, Rahul Prasanna Kumar[2]
r.p.kumar@ous-research.no, and Frank Lindseth[1]
frankl@ntnu.no[0000−−0002−4979−9218]

[1] Norwegian University of Science and Technology, Department of Computer Science, Trondheim, Norway
[2] The Intervention Centre, Oslo University Hospital, Oslo, Norway

Abstract. Precise segmentation of the hepatic and portal veins plays a vital role in planning and guiding liver surgeries. This paper presents a novel approach using multi-task learning(MTL) within SwinUNETR architecture to segment both the hepatic and portal veins at the same time. The MTL framework is trained using Dice-Focal loss and designed with two decoder branches each for segmenting the hepatic and portal vein branches. The results from the clinical CT data have shown significant performance for both the hepatic and portal veins compared to the base model (SwinUNETR), especially at the early stages of training. Notably, the MTL model achieved statistically significant results for the portal vein segmentation compared to the base model after 100 epochs. Our proposed MTL model (SwinUNETR_MTL) achieved a dice similarity coefficient (DSC) of 0.8404 for the hepatic vein and a DSC of 0.8120 for the portal vein segmentation. Our findings suggest that the MTL model attains faster convergence and increased segmentation accuracy, making it a promising approach for segmenting complex structures in the clinical setup.

Keywords: Hepatic Veins · Portal Veins · Segmentation · Transformer · Multi-Task Learning.

1 Introduction

Primary liver cancer, which consists predominantly of hepatic cellular carcinoma (HCC), is the fifth most prevalent cancer worldwide, and colorectal cancer (CRC) is the third most common type of cancer[3]. Approximately 25% of CRC patients also have liver cancer at the time of initial diagnosis due to metastasis[8]. Precise segmentation of hepatic and portal veins and their relationship to a tumor is a critical task in treatment planning, making surgical decisions, and postoperative outcomes. Hepatic vessels aid in better visualization and serve as a landmark for the multi-modal registration and liver segment approximation[23]. Traditional segmentation approaches, including manual delineation and classical image processing techniques, are often time-consuming and prone to variability. Hepatic vessel segmentation becomes more tedious due to image-related challenges such as signal-to-noise ratio, limited contrast between the hepatic parenchyma and the vessels, low resolution, inhomogeneous background, etc.

Convolutional neural network (CNN) architectures, especially U-Net and its variants, have shown remarkable efficiency for segmentation tasks. However, these models often struggle with the complexity and variability of vascular structures [21], particularly in regions with minimal contrast. The U-Net architectures depend mainly on local context, limiting their ability to capture global context [21], an essential aspect for precisely segmentating vessel structures. The Swin-Transformer architecture has shown prominent success in overcoming some of the limitations of CNNs by leveraging attention mechanisms that capture both the local and the global contextual information [29]. SwinUNETR [11] integrates the SwinTransformer with the U-Net architecture and has significantly advanced medical image segmentation. Despite its potential, there is limited research exploring the application of SwinUNETR to hepatic and portal vein segmentation. This paper presents a novel approach to hepatic and portal vein segmentation using SwinUNETR in combination with a multi-task learning(MTL) framework. Our MTL framework is designed to simultaneously segment the hepatic and portal veins, using separate decoders for each task. The design enables the model to capture task-specific features by using the shared encoder representations, potentially enhancing the overall segmentation performance.

Our key contributions are as follows: 1. We introduce a SwinUNETR-based MTL framework for the hepatic and portal veins segmentation using separate decoders for each task to capture the unique features of each vascular structure. 2. We demonstrate the effectiveness of the MTL setup in improving the segmentation efficiency by allowing shared learning between the two closely related tasks. 3. We conducted an extensive evaluation of the proposed setup on inhouse and public datasets, which showed its superiority over the state-of-the-art single-task segmentation approaches.

The structure of the paper is as follows: Section 2 reviews related work on DL-based hepatic vessel segmentation and the use of a Transformer in combination with a MTL based framework for medical image segmentation. Section 3 details the clinical dataset and proposed methodology, including the MTL framework and separate decoder design. The experiments and results are briefed in Section 4, proceeding to discussion in Section 5. The conclusion and directions for future studies are presented in Section 6.

2 Related Work

2.1 DL-based hepatic vessel segmentation methods

Ibragimov et al.[14] extracted the intensity patterns of the portal vein using 3D CNNs to enhance the vein in the target image, followed by a Markov random field for segmentation refinement. In Kitrungrotsakul et al.[16], three CNNs were trained on sagittal, coronal, and axial slices, and the results were ensembled. In Survarachakan et al.[24], different vessel-enhanced images were used as input to the 3D U-Net, and the outputs were ensembled. Kazami et al.[15] followed a CNN-based MTL approach for vessel extraction, center voxel detection, and vessel-tree reconstruction, which resulted in better extraction of hepatic and portal veins. Yan et al.[31] used U-Net as the base model, replaced the encoder-decoder layers with multi-scale attention blocks, and introduced the attention-guided concatenation module between them. Yu et al.[33] proposed introducing residual blocks to all convolutional layers to propagate global and local information across the network. Hao et al.[10] proposed a hierarchical progressive multi-scale learning approach to learn semantic information about the vessels and a dual branch progressive 3D U-Net using downsampling and DS. In Gao et al. [7], the model utilized a laplacian salience filter to highlight the vessel-like regions coupled with pyramid DL architecture to capture different levels of features. Affane et al.[1] integrated a vessel enhancement filter into 3D U-Net variants to enhance the segmentation results. Kuang et al.[17] proposed a two-stage unsupervised domain adaptive framework to segment arterial and venous vessels from the liver CT. In Alirr et al.[2], a U-Net with a modified residual block to include skip connections was used. Wu et al.[27] extended a 2D Swin-Transformer to 3D with an inductive bias multi-head self-attention mechanism. Li et al.[18] proposed a multi-stage hierarchical framework with uncertainty-aware semi-supervised learning. Xu et al.[30] proposed a dual-stream encoder combining the convolutional and transformer blocks to extract the local features and spatial information. The majority of the works presented here focussed on segmenting the hepatic and portal veins as a single vessel class but segmenting them as two separate classes is of high clinical relevance.

2.2 Transformer and MTL frameworks for medical image segmentation

Tang et al.[26] proposed a transformer-based MTL network for the classification and segmentation of gastrointestinal tract lesions from endoscopic images. Yang et al.[32] emphasized the potential of a transformer-based MTL method for simultaneous segmentation and T-staging of nasopharyngeal carcinoma. For the concurrent survival prediction and semi-supervised segmentation of gliomas in brain MRI, Wu et al.[28] proposed a multi-modal fusion transformer with MTL. Li et al.[19] proposed a spacial dependence multi-task transformer for 3D knee segmentation and landmark localization from MRI. For simultaneous infiltrated brain area classification and segmentation of gliomas, a MTL transformer is proposed in Li et al.[20], where the model emphasized the shaped location and the boundary information to improve the tasks. Tagnamas et al.[25] proposed an efficient MTL framework by leveraging the strengths of both EffientNetV2 and adapted vision transformers for breast ultrasound (US) image segmentation and classification. Hao et al.[9] proposed MTL based on UNETR for complete organ segmentation as a primary task and partial organ segmentation as an auxiliary task. The proposed MTL architecture in Huang et al.[13] utilizes the combination of U-Net and transformers to finely segment retinal vessels and continuous prediction of diameter values. The use of transformer-based MTL frameworks for tubular structure segmentation has hardly been explored. Inspired by these works, the authors propose to use a SwinUNETR-based MTL framework, which has never been explored for the task of hepatic and portal vein segmentation.

3 Methodology

This section presents the datasets, the developed network architecture, and the evaluation metrics used in the experiments.

3.1 Dataset

This study uses data from the Oslo University Hospital's OSLO-COMET trial (Oslo Randomized Laparoscopic Versus Open Liver Resection for Colorectal Metastases Trial) (ClinicalTrials.gov: NCT01516710), with approval from relevant local and regional ethical committees. The dataset includes 57 contrast-enhanced CT images of patients diagnosed with colorectal liver metastasis referred for liver resection. The images were obtained from 4 CT machine manufacturers and 13 unique models. A medical doctor manually segmented these volumes to annotate liver parenchyma and visible vessels. The dataset was divided into three splits: 70% for training, 20% for validation, and 10% for testing.

3.2 Network Architecture

In this study we use SwinUNETR [11] as a baseline architecture which is among the most effective Vision Transformer (ViT) architectures used for medical image analysis. SwinUNETR integrates the key techniques, including the Hierarchical Shifting Windows method [22], which enables the network to focus on smaller objects while maintaining linear scalability. The Swin architecture is fused with the widely recognized 3D U-Net framework by incorporating Swin multi-headed attention blocks in the encoder and a CNN in the decoder. The SwinUNETR is extended to MTL [5] where the encoder is shared between two separate decoders (see Fig. 1), each dedicated to segmenting the specific veins (hepatic or portal). Each decoder follows a standard U-Net like design with upsampling layers and skip connections to reconstruct the high-resolution feature maps. Using separate decoders allows the model to learn specialized features for each vessel type, while the shared encoder facilitates knowledge transfer between the two segmentation tasks. The output of the decoder indicating the presence of the corresponding veins is concatenated channel-wise to provide the final segmentation with both vessels.

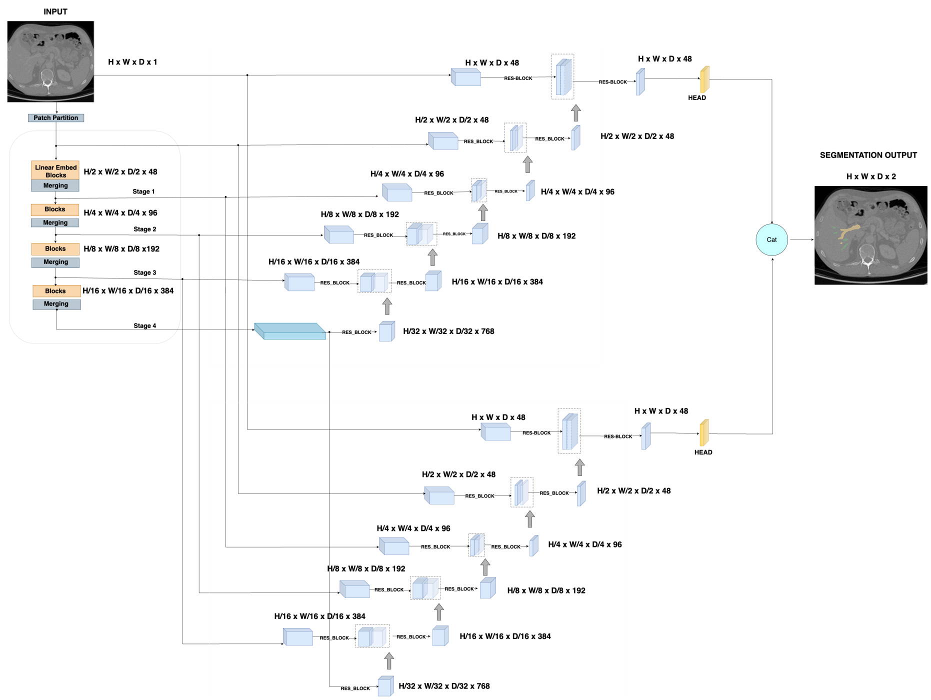

Fig. 1. Architecture of the SwinUNETR_MTL model.*Adapted from* [11]

3.3 Evaluation Metrics

The performance of the proposed framework was evaluated using a range of metrics to assess the quality of the segmentation. The Dice Similarity Coefficient (DSC) primarily computes the overlap between the predicted segmentations and ground truth, a key indicator of segmentation accuracy. It reflects how well the prediction aligns with the true structures. Sensitivity, specificity, and accuracy provide a balanced view of the model's performance in correctly segmenting the vessels vs non-vessel structures. These metrics were computed for both classes, providing a comprehensive evaluation of the models's accuracy, robustness, and reliability in clinical applications.

4 Experiments and Results

This section presents the experiments conducted and their corresponding results.

We compare the performance of the base SwinUNETR model with the MTL setup (SwinUNETR_MTL), where the MTL model is designed to leverage shared features across related tasks, enhancing learning efficiency. Several configurations of training parameters were evaluated, and the optimal settings for the dataset used are presented in Table 1. Specifically, increasing the spatial size from 96 to

Table 1. Training and data augmentation parameters

Training Parameters	**Value**	**Description**
maximum_epochs	600	Total training epochs
batch_size	4	Number of samples per batch
loss_fn	DiceFocalLoss	Loss function
lr	0.0001	Learning rate
lr_scheduler	cosine annealing	Adjusts learning rate over epochs
spacing	[1.0,1.0,1.0]	Voxel spacing for resampling
spatial_size	[128,128,128]	Target crop size
Augmentation Parameters	**Value**	**Description**
ScaleIntensityRanged	min=-80, max=250	Scales image intensity
RandFlipd_prob	0.5	Probability of random flip
RandRotate90d_prob	0.5	Probability of 90° rotation
RandGaussianSmoothd_prob	0.2	Probability of Gaussian smoothing
RandScaleIntensityd_prob	0.5	Probability of intensity scaling
RandShiftIntensityd_prob	0.5	Probability of intensity shifting
RandGaussianNoised_prob	0.2	Probability of adding noise
RandAffined_rotate_range	[0.15, 0.15, .15]	Rotation range for affine transform
RandAffined_scale_range	[0.2, 0.2, 0.2]	Scaling range for affine transform
RandAffined_prob	0.2	Probability of affine transform

128, combined with using the Dice-Focal loss function, resulted in better performance. All models were trained and evaluated using the MONAI[4] framework, following the same training parameters detailed in Table 1. To streamline the workflow and reduce manual dependencies - a key benefit in clinical settings, we avoided external preprocessing methods such as masking or cropping. Instead, the liver masks were used to determine the liver's bounding box, guiding the segmentation process within MONAI using a specific transform (transforms.CropForegroundd(keys = self.all_keys, source_key="liver")).

Initially, we compared the convergence of the base model against our proposed MTL model. Both models were trained for a maximum of 600 epochs. The validation DSC (y-axis) plotted against the number of epochs (x-axis) for the base model and the MTL model shows that the MTL model (brown) consistently outperformed the base model (red) during the initial stages of training (see Fig 2). The plot shows that the curve from the MTL rises quickly and attains stability at a higher DSC sooner than the base model. Also, the convergence of the MTL model appears more stable (i.e. less fluctuations between epochs), indicating faster and more efficient learning. During training, the checkpoints were saved at every 100 epochs, and the performance of the models was evaluated on the test set at every 100 epochs using the saved checkpoints. Table 2 further highlights that the MTL approach provides substantial benefits over the base model, particularly in the early stages of training.

Early Epochs (100-300): At 100 epochs, the MTL model achieved a DSC of 0.7895 for the hepatic vein and 0.7868 for the portal vein, marking a significant improvement over the base model's scores by 7% and 10.5%, respectively.

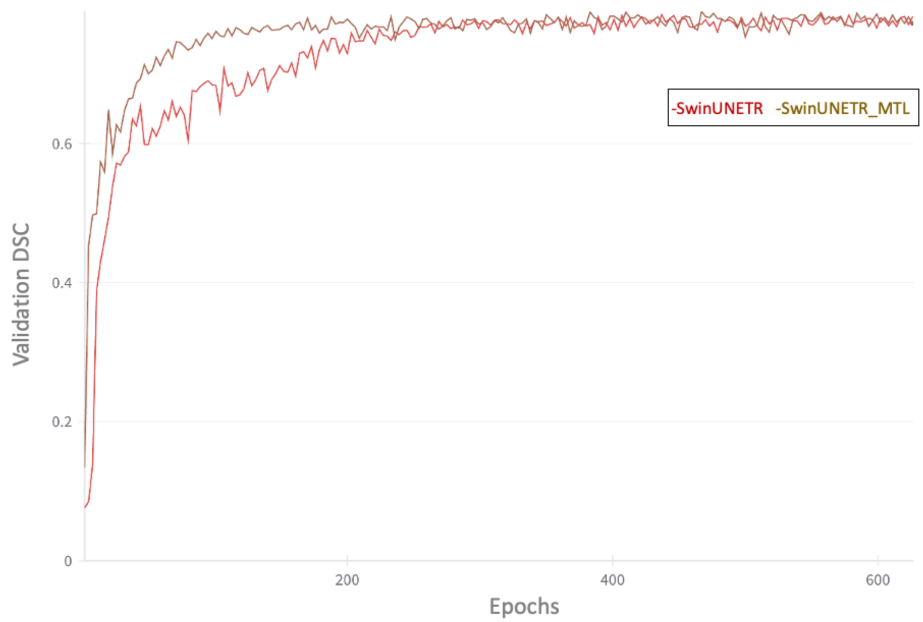

Fig. 2. Validation dice plot

To assess the statistical significance of these observations, paired t-tests were conducted. For the portal vein, the t-test at 100 epochs indicated a statistically significant difference in mean DSC ($p = 0.0438$), confirming the advantage of MTL during the early training phase. This indicates that MTL effectively captures the necessary features early in training. By 200 epochs, the gap narrows for the hepatic veins, but MTL still maintains a higher DSC of 0.8216 compared to the base model's DSC of 0.8048. For portal vein segmentation, a 4.2% improvement in DSC demonstrates the sustained advantages of MTL as training progresses. At 300 epochs, MTL reaches its peak performance for the hepatic vein with a DSC of 0.8404, 3.1% higher than the base model's 0.8151. For the portal vein, MTL achieves a DSC of 0.8120, slightly surpassing the base model's DSC of 0.8093.

Later Epochs (400-600): At 400 epochs, the MTL model continues to outperform the base model, showing a 2.4% and 3.6% improvement in DSC for the hepatic and portal veins, respectively. However, by 500 epochs, the base model slightly surpasses the MTL approach in both classes. At 600 epochs, both models yield nearly identical results, with the base model marginally outperforming MTL for the hepatic vein, while MTL slightly outperforms the base model for the portal vein.

These findings suggest that the MTL framework offers a pronounced advantage in the early stages of training, achieving optimal performance sooner and potentially reducing the need for prolonged training. Consequently, we chose

Table 2. Performance comparison between SwinUNETR and SwinUNETR_MTL based on Dice score on the test set at every 100 epochs

Epochs	Hepatic Vein		Portal Vein	
	SwinUNETR	SwinUNETR_MTL	SwinUNETR	SwinUNETR_MTL
100	0.7383	**0.7895**	0.7121	**0.7868**
200	0.8048	**0.8216**	0.7804	**0.8135**
300	0.8151	**0.8404**	0.8093	**0.8120**
400	0.8161	**0.8356**	0.7897	**0.8182**
500	**0.8303**	0.8255	**0.8213**	0.8208
600	**0.8303**	0.8221	0.8144	**0.8180**

the model trained for 300 epochs as the optimal model for further comparisons. Overall, the MTL approach effectively improves segmentation performance during the early and mid stages of training. Although the benefits of MTL may diminish with extended training, the early gains in DSC make it a valuable strategy, particularly in applications where reducing training time without compromising performance is crucial. The visualization of the segmentation results from both the models comparing the ground truth is shown in Fig. 3.

Table 3. Comparison of different methods for the hepatic and portal Vein segmentation based on DSC, Accuracy, Sensitivity and Specificity.

Method	Hepatic Vein				Portal Vein			
	DSC	Acc (%)	Sen (%)	Sp (%)	DSC	Acc (%)	Sen (%)	Sp (%)
U-Net	0.7658	99.45	75.11	99.78	0.4237	98.71	**87.84**	98.77
UNETR	0.8109	99.59	75.89	**99.89**	0.7678	99.77	73.89	99.91
SwinUNETR	0.8151	99.56	81.79	99.79	0.8093	99.82	74.93	**99.95**
SwinUNETR_MTL	**0.8404**	**99.63**	**82.24**	99.85	**0.8120**	**99.82**	76.82	99.94

We also compared the performance of the MTL model with the most popular architectures used in medical image segmentation applications, such as U-Net[6] and UNETR[12]. In addition to DSC, we used other metrics mentioned in Section 3.3 to evaluate performance. All models were evaluated on test data after 300 epochs of training. Table 3 shows that the MTL model achieves the highest DSC and accuracy for both hepatic and portal veins, indicating superior overall segmentation quality. It also performs best in terms of sensitivity for the hepatic vein. However, U-Net has the highest sensitivity for the portal vein, suggesting that U-Net detects more true positives but at the cost of over-segmentation (lower DSC). Relative to other models, MTL balances sensitivity and specificity well, leading to more precise and reliable segmentations. All the models perform well in specificity (99.7%), with MTL being best in achieving a balance between avoiding false positives and maintaining segmentation precision. To summarize,

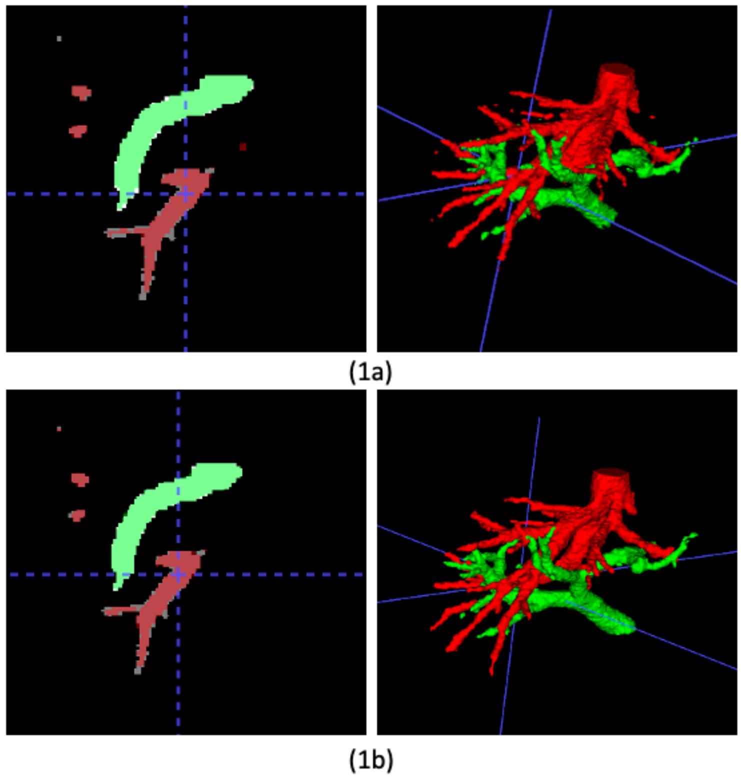

Fig. 3. Axial and 3D views of the hepatic and portal vein predictions overlayed on ground truth. (1a) From the base model (SwinUNETR), (1b) From the MTL model (SwinUNETR) viewed in ITK-Snap viewer[34]

MTL outperforms the other models by offering more accurate and balanced segmentation across both veins.

Comparing the performance of the MTL model with recently published methods on hepatic vessel segmentation is challenging. As noted earlier, most of the methods referenced in Section 2 were evaluated on complete vessel segmentation rather than distinguishing between hepatic and portal veins. Methods that segment hepatic vessels into two classes either use in-house clinical datasets or public datasets adapted to their workflow. For example, Xu et al.[30] used the Medical Segmentation Decathlon (MSD) dataset, adapting the whole liver vessel segmentation into two-class hepatic and portal vein segmentation, where the adapted ground truth is not publicly available. We found that Li et al.[18] used the 3DIRCAD dataset to evaluate their method. Therefore, the proposed MTL model trained on the in-house dataset was further finetuned on the 3DIRCAD dataset and compared with Li et al. Instead of training the model from scratch, the new model was trained using the checkpoint trained on the in-house dataset. This transfer learning approach reduces the time and resources required to train

Table 4. Performance comparison on the 3DIRCAD dataset

Method	Hepatic Vein			Portal Vein		
	Acc(%)	Sen(%)	Sp(%)	Acc(%)	Sen(%)	Sp(%)
Li et al.	98.9	61.5	98.3	**99.7**	62.4	99.4
SwinUNETR_MTL	**99.6**	**67.9**	**99.9**	**99.7**	**65.9**	**99.9**

the model on a new dataset smaller than the original dataset. Table 4 shows that our proposed method, using the MTL strategy, surpasses Li et al.'s method in hepatic vein accuracy (99.6% vs. 98.9%) and maintains similar accuracy for the portal vein (99.7%). MTL also outperforms Li et al. in sensitivity for both veins, with an increase in hepatic vein sensitivity from 61.5% to 67.9% and portal vein sensitivity from 62.4% to 65.9%. This indicates that the MTL model detects more true positives and performs better at identifying the veins. MTL also shows higher specificity for both hepatic and portal veins, which demonstrates better avoidance of false positives, suggesting MTL delivers more precise segmentation results. To summarize, the MTL model outperforms Li et al.'s method in all key metrics and shows a significant boost in sensitivity for both veins.

5 Discussion

This study highlights the impact of using MTL, especially with SwinUNETR architecture for the hepatic and portal vein segmentation. The results highlight the critical insights of using MTL for complex medical image segmentation tasks, particularly concerning training efficiency, segmentation accuracy, and potential clinical relevance.

Impact on Training Efficiency and Early Convergence

The most significant advantage of using the MTL framework is the potential of the MTL model to enhance training efficiency, especially by promoting early convergence. The analysis of the results suggests that the MTL model consistently outperformed the base SwinUNETR model, particularly in the early training phase, as shown by the results at 100 epochs. For instance, the MTL achieved an increased DSC of 0.7868 compared to 0.7383 from the base model for the hepatic vein. A more pronounced improvement in DSC of 0.7868 was achieved by the MTL compared to 0.7121 from the base model for portal vein. These early gains are critical in scenarios where rapid model development is essential. The MTL's ability to utilize shared representations across tasks accelerates the learning process, requiring fewer epochs to achieve higher accuracy. This is a significant benefit, especially in clinical settings where the time and computational resources are limited.

Enhancing Segmentation Accuracy

In addition to accelerated training, MTL consistently improves segmentation accuracy in complex tasks like portal vein segmentation. The DSC across multiple epochs reveals that MTL persistently boosts accuracy over the base model, especially in the early stages of training. Even at 300 epochs, MTL achieved a DSC of 0.8404 for the hepatic vein, significantly higher than the base model's DSC of 0.8151. For the portal vein segmentation, MTL maintained a superior performance, showing its potential to capture complex anatomical features. The consistent performance advantage of MTL over the base model highlights its role in enhancing reliability and precision in medical image segmentation tasks, which is very crucial in clinical applications. Compared with other segmentation architectures like U-Net and UNETR, the MTL approach showed superior performance on most metrics, as shown in Table 3. Furthermore, when finetuned on 3DIRCAD datasets, MTL surpassed the method proposed in Li et al. [18]. This implies MTL generalizes well across different datasets and offers a competitive edge in accuracy.

Clinical Relevance and Implications

The potential clinical relevance of MTL is profound, especially in its ability to accelerate model training and deliver accurate segmentation results. In clinical procedures such as post-operative planning and intra-operative navigation, the need for precise and timely segmentation is paramount. MTL's ability to achieve faster convergence and higher accuracy means it can produce high-quality segmentations much more quickly than traditional methods. This efficiency can streamline clinical workflows, and potentially enhance surgical outcomes. Additionally, improved segmentation of complex structures such as portal veins can lead to more accurate diagnosis and better treatment decisions.

Challenges with Centerline Extraction

While this research highlighted the advantages of applying MTL for the vessel segmentation tasks, specific challenges emerged while attempting centreline extraction (CLE) and using Hausdorff Distance (HD) as a metric to assess the quality of the segmentation accuracy. CLE was performed on the ground truth and the segmented results, and HD was calculated. However, the ground truth labels were annotated on a slice-by-slice basis and lacked 3D continuity, resulting in fragmented centrelines in the ground truth. The fragmentation led to increased HD values that did not reflect the actual segmentation accuracy. In this context, using HD gives no insight into segmentation performance, so this part of the analysis was omitted in the final evaluation. This highlights the importance of ensuring that the metrics used for 3D segmentation assessments are appropriate for the data and the labeling process. In future work, the labeling process has to be improved to obtain more accurate segmentation and extensive evaluation.

6 Conclusion

In this study, we studied the effect of using MTL with SwinUNETR architecture to segment the hepatic and portal veins. The key findings suggest, that using MTL significantly improves the segmentation of both the vessel branches, particularly in the early stages of training. This is evident by statistically significant improvement in DSC at 100 epochs. In addition to the increased segmentation accuracy, training with the MTL strategy leads to faster model convergence than the base model. In clinical settings, the MTL approach could reduce the time and computational resources required for model training, ultimately leading to more timely and precise surgical interventions. Future research should focus on applying MTL to other multi-structure segmentation tasks. Expanding the validation to larger and more diverse datasets could ensure the generalizability of the MTL model. To conclude, using MTL with SwinUNETR has shown promising results in complex medical segmentation tasks like hepatic and portal vein segmentation. MTL offers promising results in terms of both time and accuracy that could enhance clinical outcomes, particularly in liver surgery.

Acknowledgement

This work is supported by H2020-MSCA-ITN Marie Sklodowska-Curie Actions, Innovative Training Networks (ITN)-H2020 MSCA ITN 2016 GA EU project number 722068 High Performance Soft Tissue Navigation (HiPerNav).

Disclosure of Interests

The authors have no competing interests.

References

1. Affane, A., Lamy, J., Lebre, M.A., Vacavant, A.: Robust deep 3-d architectures based on vascular patterns for liver vessel segmentation. Informatics in Medicine Unlocked **34**, 101111 (2022)
2. Alirr, O.I., Rahni, A.A.A.: Hepatic vessels segmentation using deep learning and preprocessing enhancement. Journal of applied clinical medical physics **24**(5), e13966 (2023)
3. on Cancer, I.A.f.R., Organization, W.H.: Globocan 2020. Globocan 2020 **419**, 1–2 (2020)
4. Cardoso, M.J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., Murrey, B., Myronenko, A., Zhao, C., Yang, D., et al.: Monai: An open-source framework for deep learning in healthcare. arXiv preprint arXiv:2211.02701 (2022)
5. Caruana, R.: Multitask learning. Machine learning **28**, 41–75 (1997)
6. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3d u-net: learning dense volumetric segmentation from sparse annotation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19. pp. 424–432. Springer (2016)

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. van der Geest, L.G., Lam-Boer, J., Koopman, M., Verhoef, C., Elferink, M.A., de Wilt, J.H.: Nationwide trends in incidence, treatment and survival of colorectal cancer patients with synchronous metastases. Clinical and Experimental Metastasis **32**(5), 457–465 (2015). https://doi.org/10.1007/s10585-015-9719-0
9. Hao, Q., Tian, S., Yu, L., Wang, J.: Mfunetr: A transformer-based multi-task learning network for multi-organ segmentation from partially labeled datasets. Biomedical Signal Processing and Control **85**, 105081 (2023). https://doi.org/https://doi.org/10.1016/j.bspc.2023.105081
10. Hao, W., Zhang, J., Su, J., Song, Y., Liu, Z., Liu, Y., Qiu, C., Han, K.: Hpm-net: Hierarchical progressive multiscale network for liver vessel segmentation in ct images. Computer Methods and Programs in Biomedicine **224**, 107003 (2022)
11. Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H.R., Xu, D.: Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In: International MICCAI brainlesion workshop. pp. 272–284. Springer (2021)
12. Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., Roth, H.R., Xu, D.: Unetr: Transformers for 3d medical image segmentation. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision. pp. 574–584 (2022)
13. Huang, C., Wang, Z., Yuan, G., Xiong, Z., Hu, J., Tong, Y.: Pksea-net: A prior knowledge supervised edge-aware multi-task network for retinal arteriolar morphometry. Computers in Biology and Medicine **172**, 108255 (2024). https://doi.org/https://doi.org/10.1016/j.compbiomed.2024.108255
14. Ibragimov, B., Toesca, D., Chang, D., Koong, A., Xing, L.: Combining deep learning with anatomical analysis for segmentation of the portal vein for liver sbrt planning. Physics in Medicine & Biology **62**(23), 8943 (2017)
15. Kazami, Y., Kaneko, J., Keshwani, D., Takahashi, R., Kawaguchi, Y., Ichida, A., Ishizawa, T., Akamatsu, N., Arita, J., Hasegawa, K.: Artificial intelligence enhances the accuracy of portal and hepatic vein extraction in computed tomography for virtual hepatectomy. Journal of Hepato-Biliary-Pancreatic Sciences (11 2021). https://doi.org/10.1002/jhbp.1080
16. Kitrungrotsakul, T., Han, X.H., Iwamoto, Y., Lin, L., Foruzan, A., Xiong, W., Chen, Y.W.: Vessel net: A deep convolutional neural network with multi pathways for robust hepatic vessel segmentation. Computerized Medical Imaging and Graphics **75** (06 2019). https://doi.org/10.1016/j.compmedimag.2019.05.002
17. Kuang, H., Yang, Z., Zhang, X., Tan, J., Wang, X., Zhang, L.: Hepatic vein and arterial vessel segmentation in liver tumor patients. Computational Intelligence and Neuroscience **2022**(1), 2303733 (2022)
18. Li, C., Ma, W., Sun, L., Ding, X., Huang, Y., Wang, G., Yu, Y.: Hierarchical deep network with uncertainty-aware semi-supervised learning for vessel segmentation. Neural Computing and Applications pp. 1–14 (2022)
19. Li, X., Lv, S., Li, M., Zhang, J., Jiang, Y., Qin, Y., Luo, H., Yin, S.: Sdmt: spatial dependence multi-task transformer network for 3d knee mri segmentation and landmark localization. IEEE transactions on medical imaging **42**(8), 2274–2285 (2023)
20. Li, Y., Zheng, K., Li, S., Yi, Y., Li, M., Ren, Y., Guo, C., Zhong, L., Yang, W., Li, X., et al.: A transformer-based multi-task deep learning model for simultaneous infiltrated brain area identification and segmentation of gliomas. Cancer Imaging **23**(1), 105 (2023)

21. Liu, J., Zhao, D., Shen, J., Geng, P., Zhang, Y., Yang, J., Zhang, Z.: Hrd-net: High resolution segmentation network with adaptive learning ability of retinal vessel features. Computers in Biology and Medicine **173**, 108295 (2024)
22. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B.: Swin transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 10012–10022 (2021)
23. Selle, D., Preim, B., Schenk, A., Peitgen, H.O.: Analysis of vasculature for liver surgical planning. IEEE transactions on medical imaging **21**, 1344–57 (12 2002). https://doi.org/10.1109/TMI.2002.801166
24. Survarachakan, S., Pelanis, E., Khan, Z.A., Kumar, R.P., Edwin, B., Lindseth, F.: Effects of enhancement on deep learning based hepatic vessel segmentation. Electronics **10**(10) (2021). https://doi.org/10.3390/electronics10101165
25. Tagnamas, J., Ramadan, H., Yahyaouy, A., Tairi, H.: Multi-task approach based on combined cnn-transformer for efficient segmentation and classification of breast tumors in ultrasound images. Visual Computing for Industry, Biomedicine, and Art **7**(1), 2 (2024)
26. Tang, S., Yu, X., Cheang, C.F., Liang, Y., Zhao, P., Yu, H.H., Choi, I.C.: Transformer-based multi-task learning for classification and segmentation of gastrointestinal tract endoscopic images. Computers in Biology and Medicine **157**, 106723 (2023). https://doi.org/https://doi.org/10.1016/j.compbiomed.2023.106723
27. Wu, M., Qian, Y., Liao, X., Wang, Q., Heng, P.A.: Hepatic vessel segmentation based on 3d swin-transformer with inductive biased multi-head self-attention. BMC Medical Imaging **23**(1), 91 (2023)
28. Wu, W., Yan, J., Zhao, Y., Sun, Q., Zhang, H., Cheng, J., Liang, D., Chen, Y., Zhang, Z., Li, Z.C.: Multi-task learning for concurrent survival prediction and semi-supervised segmentation of gliomas in brain mri. Displays **78**, 102402 (2023)
29. Xiao, H., Li, L., Liu, Q., Zhu, X., Zhang, Q.: Transformers in medical image segmentation: A review. Biomedical Signal Processing and Control **84**, 104791 (2023). https://doi.org/https://doi.org/10.1016/j.bspc.2023.104791, `https://www.sciencedirect.com/science/article/pii/S1746809423002240`
30. Xu, J., Jiang, W., Wu, J., Zhang, W., Zhu, Z., Xin, J., Zheng, N., Wang, B.: Hepatic and portal vein segmentation with dual-stream deep neural network. Medical Physics (2024)
31. Yan, Q., Wang, B., Zhang, W., Luo, C., Xu, W., Xu, Z., Zhang, Y., Shi, Q., Zhang, L., You, Z.: Attention-guided deep neural network with multi-scale feature fusion for liver vessel segmentation. IEEE Journal of Biomedical and Health Informatics **25**, 2629–2642 (2021)
32. Yang, K., Dong, X., Tang, F., Ye, F., Chen, B., Liang, S., Zhang, Y., Xu, Y.: A transformer-based multi-task deep learning model for simultaneous t-stage identification and segmentation of nasopharyngeal carcinoma. Frontiers in Oncology **14** (2024)
33. Yu, W., Fang, B., Liu, Y., Gao, M., Zheng, S., Wang, Y.: Liver vessels segmentation based on 3d residual u-net. In: 2019 IEEE International Conference on Image Processing (ICIP). pp. 250–254. IEEE (2019)
34. Yushkevich, P.A., Gao, Y., Gerig, G.: Itk-snap: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 3342–3345 (2016). https://doi.org/10.1109/EMBC.2016.7591443