

Separating portal and hepatic veins on CT with a TotalSegmentator-compatible model trained only on public data: external validation and release

Meidai Kasai

meidai1223@gmail.com

Independent Researcher <https://orcid.org/0000-0002-4899-5663>

Research Article

Keywords: Portal vein, Hepatic vein, Vessel segmentation, TotalSegmentator, nnU-Net, Open model

Posted Date: October 27th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11500369/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

Purpose

Liver surgery planning requires separating the portal and hepatic veins on CT, yet, to the author's knowledge, no deep learning model for this has been released under a permissive license. Such a model was developed from public data, validated externally, and released.

Methods

An nnU-Net model, OpenPVHV, was trained on 224 cases of the public HVM dataset (Hepatic Vessel Map) with the configuration of the TotalSegmentator liver-vessel task. Fifty-six HVM cases formed the internal and the 40 VEELA cases the external test set; predictions for the 20 cases with hidden labels were scored by the challenge server. Metrics were the official centerline Dice (cIDice), intersection over union (IoU), and normalized surface distance (NSD). Comparators were the general-purpose model VoxTell, the TotalSegmentator liver-vessel task, the challenge participants, and a model trained with 62 added CRLM (colorectal liver metastases) cases.

Results

Mean cIDice of OpenPVHV (portal/hepatic) was 0.750/0.786 internally, 0.723/0.662 on the public external cases, and 0.734/0.732 on the hidden cases, where the portal vein ranked first on all three metrics above every team at the deadline and the organizers' baseline. OpenPVHV exceeded VoxTell on all six comparisons on the internal and public external cases, and the TotalSegmentator liver-vessel task on all vessels combined (cIDice 0.784 vs 0.528 internally, 0.706 vs 0.501 externally). Adding CRLM lowered all six values on the public external cases.

Conclusion

OpenPVHV separates the portal and hepatic veins on external data and may replace the vessel output of the TotalSegmentator liver-vessel task. Code and weights are released under Apache-2.0.

Introduction

In preoperative planning of hepatectomy, the portal territory containing the tumor and the course of the hepatic veins must be identified to determine the extent of resection while preserving as much liver parenchyma as possible [1]. The portal and hepatic veins and the liver segments are usually extracted with commercial software [2], whose licensing makes adoption costly.

Deep learning segmentation of the portal and hepatic veins has been reported [3–6], but each model was trained on non-public in-house data, and none of the reports states how the trained model can be obtained.

The public whole-body segmentation model TotalSegmentator [7] offers a liver-vessel task, but its output is a single intrahepatic-vessel class that does not separate the portal and hepatic veins [8]. In 2025, the VEELA challenge on portal and hepatic vein segmentation was held on public data [9], and its evaluation server has continued to accept submissions since the challenge closed [10]. Xie et al. released HVM, a two-center CT dataset in which radiologists annotated the portal and hepatic veins, under a CC BY license and described it in a 2026 paper [11]. Although the paper states that a trained baseline model is provided [11], no trained portal and hepatic vein model could be found at the linked repository as of October 2026 [12]. To the author's knowledge, only two trained models dedicated to separating the portal and hepatic veins had obtainable weights at that time, and neither provided a license for its weights [13, 14]. The general-purpose model VoxTell [15, 16] separates the two veins and reports Dice on 20 VEELA cases, but its weights are distributed under CC BY-NC-SA 4.0 and therefore cannot be bundled and redistributed on the same terms as a tool released under Apache-2.0, such as TotalSegmentator.

This study therefore developed OpenPVHV, a model that segments the portal and hepatic veins separately and runs alongside TotalSegmentator, trained on HVM, whose annotations are available under a CC BY license. Its performance was evaluated on internal and external test sets and compared with VoxTell, and the code and weights were released.

Methods

This is a retrospective study of public data.

Data: HVM [11] consists of 282 portal venous phase abdominal contrast-enhanced CT scans acquired at two centers in China; the annotations were produced by radiologists, either by correcting AI-generated initial masks or manually, and reviewed by a senior radiologist [11]. Because the file names of the released images and annotations did not correspond, the two were matched by image header and by the CT values inside the annotated vessels, and the 280 cases with a unique match were used. Most of the distributed images had clipped CT values: all 106 cases from one center were clipped to – 40 to 160 HU and 127 of the 174 cases from the other to – 120 to 180 HU, values outside the range having been replaced by the respective limit. All images were used as distributed. A random 20% from each center, 56 cases in total, formed the internal test set, and the remaining 224 cases were used for training.

The public CRLM dataset [17] was used in a comparison experiment on the effect of additional training data. CRLM is the Colorectal-Liver-Metastases collection of The Cancer Imaging Archive: 197 preoperative portal venous phase abdominal contrast-enhanced CT scans of patients with colorectal liver metastases from a single center in the United States, released under a CC BY 4.0 license with annotations of the liver, tumors, portal vein, and hepatic veins [17, 18]. Whereas HVM does not annotate the inferior vena cava, the hepatic vein annotation in some CRLM cases extended into the inferior vena

cava or beyond the liver; of the 80 cases with a slice thickness of 2.5 mm or less, therefore, 18 cases in which 30% or more of the hepatic vein voxels lay outside the liver annotation supplied with CRLM, or 20% or more overlapped the inferior vena cava obtained with the total task of TotalSegmentator 2.18, were excluded, and the remaining 62 cases were used only to train the comparison model (the HVM+CRLM model described below). CRLM was not used to train or evaluate the released OpenPVHV.

VEELA [9, 19] served as the external test set. VEELA consists of 40 portal venous phase abdominal contrast-enhanced CT scans of living liver donor candidates from a single center in Turkey, and its training and test data include images from the preceding CHAOS challenge [20], as described by the organizers [19]. The VEELA images were not clipped. The 20 cases with public reference labels were the challenge training data, and the rules permitted the use of external data and pretrained models [19]. In this study these 20 cases were used only for evaluation, not for training. Predictions for the 20 cases with hidden labels were submitted once to the evaluation server [10], which returned scores for Task 1, all vessels combined, and Task 2, portal and hepatic veins separately. HVM does not label the inferior vena cava, whereas VEELA labels its intrahepatic portion as hepatic vein [19].

Model: nnU-Net [21] derives its preprocessing and network configuration automatically from the training data, but this mechanism was not used; instead, the configuration of the TotalSegmentator liver-vessel model was adopted unchanged. The configuration is 3D full resolution with a voxel spacing of 1.5 mm through-plane and 0.80 mm in-plane, a patch size of $64 \times 192 \times 192$ voxels, a batch size of 2, and a six-stage U-Net. CT values were clipped to 8–268 HU and then standardized with a mean of 128.7 HU and a standard deviation of 54.6 HU. The output has three classes: background, portal vein, and hepatic vein. OpenPVHV was trained from random initialization on the 224 HVM cases for 1,000 epochs, and the weights of the final epoch were used. A model with the same settings was also trained on 286 cases, the 224 HVM cases plus the 62 CRLM cases, and is referred to below as the HVM+CRLM model. The HVM+CRLM model was used only for comparison and is not released. Training and inference used nnU-Net 2.7.0 and PyTorch 2.11.0 on Python 3.11. The weights of the TotalSegmentator liver-vessel model are distributed under Apache-2.0, but its training data, the hepatic vessel data of the Medical Segmentation Decathlon [8], are licensed CC BY-SA 4.0 [22], and whether that condition extends to the weights is unclear; these weights were therefore not used in the released model and were used only within the comparator, the TotalSegmentator liver-vessel task.

Inference

The liver was extracted with the 6-mm whole-body model of TotalSegmentator 2.18, and the vessel model was applied to a region of interest formed by expanding the bounding box of the liver by 20 mm in every direction. No test-time augmentation or postprocessing was applied, and the output was left empty when no liver was detected. This procedure is identical to the TotalSegmentator liver-vessel task, and the inference settings were fixed before evaluation without reference to the test results.

Comparators

Comparators for the separate portal and hepatic vein evaluation were restricted to trained models that output the portal and hepatic veins separately, have obtainable weights, and are accompanied by a peer-reviewed paper; VoxTell [15], a general-purpose 3D segmentation model whose target is specified by a free-text prompt, met these criteria. A search of GitHub, Hugging Face, Zenodo, and Grand Challenge in October 2026 with the five terms portal vein segmentation, hepatic vein segmentation, hepatic vessel segmentation, liver vessel segmentation, and portal hepatic vein found two other trained models with obtainable weights that separate the two veins [13, 14], but neither was described in a peer-reviewed paper or provided a license for its weights, so neither was included. For VoxTell, version v1.0, which the distributor designates for reproducing the results of the paper, was used [16]. The latest version, v1.1, is documented as including the paper's test sets in its training data [16] and therefore very likely contains the 20 public VEELA cases, so it was not used. The training data of v1.0 [15] include neither HVM nor VEELA. Because the TotalSegmentator liver-vessel task does not separate the portal and hepatic veins, it was compared on all vessels combined.

Evaluation conditions

VoxTell inference used the distributed Python package vox tell 0.1.2 and the v1.0 weights with default settings, with the prompts “portal vessels” and “hepatic vessels”. On the 20 public VEELA cases, the mean Dice over the whole image with these settings was 0.460 for the portal vein and 0.412 for the hepatic veins, matching the values the VoxTell paper reports for the same 20 cases [15]. Because datasets and models differ in whether extrahepatic vessels are annotated or output, the reference and all model outputs on the internal test set and the 20 public cases were scored only within the liver mask from the 6-mm TotalSegmentator model dilated by 10 mm.

Metrics

For the internal test set and the 20 public cases, the three official VEELA metrics [19], centerline Dice (cIDice) [23], intersection over union (IoU), and normalized surface distance (NSD) [24], were computed for the portal and hepatic veins separately with the official evaluation code. The official code [25] assumes a voxel spacing of $3 \times 2 \times 1$ mm for every case and computes NSD at a tolerance of 1 mm. A class with an empty output, for which the official code leaves cIDice undefined, was scored 0 on all three metrics and included in the averages. The hidden test set was scored on the same three metrics by the evaluation server. OpenPVHV was submitted in October 2026, after the challenge deadline; for Task 1 the portal and hepatic vein outputs were merged. Ranks were determined per metric against the submissions on the leaderboard at the deadline of 22 May 2025, 14 from four teams for Task 1 and 13 from four teams for Task 2, both per submission and against each team's best value [10]. GLIMS, a hybrid CNN–Transformer 3D segmentation model previously published by the organizers, was the best of the three baselines they trained on the VEELA training data [19, 26]; its values are shown for comparison but were not included in the ranking.

Statistical analysis

Metric values are reported as means. Differences between OpenPVHV and VoxTell, the HVM+CRLM model, and the TotalSegmentator liver-vessel task were tested with the two-sided Wilcoxon signed-rank test on paired per-case values. Multiple comparisons were corrected by the Bonferroni method, treating all comparisons tested for a given pair of models and dataset as one family. The family comprised six comparisons, three metrics for each vein, for the comparisons with VoxTell and the HVM+CRLM model, and three comparisons, the metrics for all vessels combined, for the comparison with the TotalSegmentator liver-vessel task. Output vessel volumes are compared descriptively without testing.

Use of generative AI

Claude (Opus 5.5 and Fable 5.1; Anthropic) was used in October 2026 for writing scripts, collecting and checking references, drafting the manuscript, and translating it into English. All numbers were compiled from execution logs, and the author reviewed and revised the final text and takes full responsibility for it.

Results

Internal test

On the internal test set, mean portal vein cIDice was 0.750 for OpenPVHV and 0.386 for VoxTell v1.0, and mean hepatic vein cIDice 0.786 and 0.448; OpenPVHV was significantly higher on all six comparisons, including IoU and NSD (Bonferroni-corrected $p < 0.001$; Table 1). Against TotalSegmentator, mean cIDice for all vessels combined was 0.784 for OpenPVHV and 0.528 for the liver-vessel task, OpenPVHV being significantly higher on all three comparisons ($p < 0.001$; Table 1). Representative cases are shown in Fig. 1. In one case with severe hepatic steatosis and CT values clipped to – 40 to 160 HU, no liver was detected and the output was empty. Median processing time per case, including liver extraction and file input/output, was 42 s on an RTX 4070.

Table 1
Comparison of OpenPVHV with existing tools (means)

		HVM internal test (56 cases)			VEELA public (20 cases)		
Vessel	Metric	OpenPVHV	Comparator	p	OpenPVHV	Comparator	p
(a) VoxTell v1.0							
Portal vein	clDice	0.750	0.386	< 0.001	0.723	0.426	< 0.001
	IoU	0.667	0.423	< 0.001	0.554	0.400	< 0.001
	NSD	0.735	0.383	< 0.001	0.726	0.466	< 0.001
Hepatic vein	clDice	0.786	0.448	< 0.001	0.662	0.416	< 0.001
	IoU	0.625	0.311	< 0.001	0.387	0.269	< 0.001
	NSD	0.750	0.379	< 0.001	0.647	0.386	< 0.001
(b) TotalSegmentator liver-vessel task (portal and hepatic veins combined)							
All vessels	clDice	0.784	0.528	< 0.001	0.706	0.501	< 0.001
	IoU	0.659	0.282	< 0.001	0.456	0.292	< 0.001
	NSD	0.757	0.295	< 0.001	0.701	0.398	< 0.001
p values: two-sided Wilcoxon signed-rank test, Bonferroni-corrected over six comparisons in (a) and three in (b) per dataset. In (b), portal and hepatic vein labels were merged as vessel in prediction and reference, and the vessel label (excluding tumor) of the TotalSegmentator 2.18 liver-vessel task was used. HVM case c1_093, in which no liver was detected, was included with all metrics scored 0 for OpenPVHV and TotalSegmentator; its VoxTell output was scored on the whole image. HVM is the source of the OpenPVHV training data, so the HVM values favor OpenPVHV.							

External test, 20 cases with public labels

On the 20 external cases with public labels, mean portal vein clDice was 0.723 for OpenPVHV and 0.426 for VoxTell, and mean hepatic vein clDice 0.662 and 0.416; OpenPVHV was significantly higher on all six comparisons ($p < 0.001$; Table 1). Against TotalSegmentator, mean clDice for all vessels combined was 0.706 for OpenPVHV and 0.501 for TotalSegmentator, OpenPVHV being significantly higher on all three comparisons ($p < 0.001$; Table 1).

External test, 20 cases with hidden labels

In Task 1, mean cIDice of OpenPVHV was 0.740, ranking second on cIDice, fourth on IoU, and first on NSD against each team's best, and below the organizers' baseline GLIMS [19] on all three metrics (cIDice 0.769). In Task 2, cIDice of OpenPVHV was 0.734 for the portal vein and 0.732 for the hepatic veins; the portal vein ranked first on all three metrics, and the hepatic veins second on cIDice, third on IoU, and first on NSD. In Task 2, OpenPVHV exceeded GLIMS on five of the six values, all but hepatic vein IoU (Table 2).

Table 2
VEELA hidden 20 cases (means)

Vessel	Metric	OpenPVHV	Rank vs team best (of 5)	Rank per submission	Best other submission	Organizers' baseline GLIMS
Task 1: all vessels	cIDice	0.740	2	2 / 15	0.789	0.769
	IoU	0.495	4	7 / 15	0.534	0.588
	NSD	0.730	1	1 / 15	0.704	0.734
Task 2: portal vein	cIDice	0.734	1	1 / 14	0.704	0.634
	IoU	0.547	1	1 / 14	0.474	0.511
	NSD	0.728	1	1 / 14	0.625	0.643
Task 2: hepatic vein	cIDice	0.732	2	4 / 14	0.833	0.678
	IoU	0.460	3	10 / 14	0.532	0.565
	NSD	0.718	1	1 / 14	0.698	0.684
Ranks are among the deadline leaderboard submissions (saved 16 Oct 2026 [10]) and OpenPVHV. Team best: highest value per metric among a team's submissions; best other submission: highest value among submissions other than OpenPVHV. GLIMS values are from Tables 4 and 5 of [19] and are not ranked.						

Additional training data (additional comparison): On the internal test set, the HVM+CRLM model, trained with the 62 additional CRLM cases, was not higher than OpenPVHV on any of the six comparisons, with differences in means below 0.01, and only hepatic vein IoU and NSD reached statistical significance (Bonferroni-corrected $p = 0.007$ and 0.045 , respectively). On the 20 public external cases, the HVM+CRLM model was significantly lower on all six comparisons ($p < 0.001$; Table 3): mean portal vein cIDice was 0.723 for OpenPVHV and 0.506 for the HVM+CRLM model, and mean hepatic vein cIDice 0.662 and 0.483, and the output vessel volume of the HVM+CRLM model was 0.41 times (portal vein) and 0.49 times (hepatic veins) that of OpenPVHV as the median of the per-case ratios.

Table 3
OpenPVHV vs the HVM+CRLM model trained with 62 additional CRLM cases (means)

		HVM internal test (56 cases)			VEELA public (20 cases)		
Vessel	Metric	OpenPVHV	HVM+CRLM	p	OpenPVHV	HVM+CRLM	p
Portal vein	clDice	0.750	0.750	1.000	0.723	0.506	< 0.001
	IoU	0.667	0.666	0.136	0.554	0.401	< 0.001
	NSD	0.735	0.735	1.000	0.726	0.477	< 0.001
Hepatic vein	clDice	0.786	0.783	0.130	0.662	0.483	< 0.001
	IoU	0.625	0.621	0.007	0.387	0.271	< 0.001
	NSD	0.750	0.746	0.045	0.647	0.420	< 0.001
p values are two-sided Wilcoxon signed-rank p values with Bonferroni correction over six comparisons per dataset. Scoring as in Table 1, including case c1_093.							

Discussion

OpenPVHV, trained on public data alone and without any VEELA data, exceeded every team at the deadline and the organizers' baseline on the portal vein of the hidden test set, and exceeded VoxTell [15] and the TotalSegmentator liver-vessel task on all three official metrics on the internal test set and the 20 public cases. OpenPVHV is therefore an option for vessel segmentation in place of the TotalSegmentator liver-vessel task, which also outputs tumors. The external hepatic vein evaluation was unfavorable to OpenPVHV because of the definition of the inferior vena cava described below; OpenPVHV nevertheless exceeded VoxTell. In the hidden test, OpenPVHV fell short of the organizers' baseline GLIMS only on all vessels combined (Task 1), mainly in IoU, and on hepatic vein IoU in Task 2. Both are measures of volume overlap and may reflect the intrahepatic inferior vena cava, which GLIMS learned from the VEELA training data and OpenPVHV does not output.

On the external test, hepatic vein values were lower than portal vein values, which partly reflects how the inferior vena cava is defined. Because HVM does not annotate the inferior vena cava and OpenPVHV does not output it, hepatic vein scores against the VEELA reference, which includes the intrahepatic inferior vena cava in the hepatic veins [19], are lowered by this difference. The extent annotated as portal and hepatic veins thus differs even among public datasets. In CRLM, counting automatically with the TotalSegmentator heart and vessel masks over the 80 cases with a slice thickness of 2.5 mm or less, the hepatic vein annotation extended beyond the extrahepatic inferior vena cava into the heart in 4

cases, and the portal vein annotation reached 30 mm or more to the left of the midline, that is, into the splenic vein, in 5 cases. The 4 cases reaching the heart had been excluded from training by the criteria described in the Methods, whereas 3 of the 5 cases including the splenic vein were among the 62 cases used to train the HVM+CRLM model. The whole-body task of TotalSegmentator likewise outputs the portal and splenic veins as a single label [8]. The splenic vein is not needed for hepatectomy planning, whereas pancreatic surgery requires the anatomy from the splenic vein to the main portal trunk; the extent that should be annotated depends on the intended operation. Future public datasets should define the annotated extent according to the intended operation and state it explicitly.

Two further factors may lower the external values. First, the input distributions differed. Most HVM images were clipped at 160 or 180 HU, so only 1–3% of vessel voxels in the training data exceeded 180 HU, whereas VEELA vessels were more strongly enhanced, with about 30% of their voxels between 180 and 268 HU. Second, in a study that re-annotated the hepatic vessel data of the Medical Segmentation Decathlon, two readers assigned different portal or hepatic labels to part of the vessels in 19 of 61 cases [27]; external values therefore also reflect differences in how reference labels were produced.

Adding the 62 CRLM cases did not improve external performance and lowered it on VEELA. Since the output vessel volume was about half that of OpenPVHV, more vessels were probably missed. The cause was not determined, but as noted above, some CRLM annotations include extrahepatic structures such as the splenic vein and inferior vena cava, and the mismatch with the HVM annotation extent may have contributed. In addition, when the HVM (280 cases) and CRLM (80 cases) annotations were compared after resampling to 1 mm isotropic voxels, the median centerline length of the portal vein was 1,442 mm in HVM and 726 mm in CRLM, with 84 and 58 branch end points, so the CRLM portal vein annotations contain fewer thin branches, which is consistent with more vessels being missed (hepatic veins 1,386 and 1,168 mm). Moreover, although 197 CRLM cases are public, many have a slice thickness above 2.5 mm, and only 62 could be used for training. More training data does not guarantee better performance: training on a small set of accurately annotated images has been reported to perform as well as training on ten times as many noisily annotated images [28]. At least at the scale of a few hundred cases, consistency of annotation extent and image conditions may matter more than the number of cases.

A limitation is that the training data come from two centers in one country and the external test set, VEELA, consists of CT scans of living liver donor candidates, so no external validation was performed in livers with tumors or cirrhosis. Moreover, only portal venous phase CT was evaluated, the liver was not detected in one internal test case and the output was empty, and usefulness for surgical planning was not evaluated.

Conclusion

OpenPVHV, a portal and hepatic vein segmentation model trained on public data, was developed, and its code and weights were released under the Apache-2.0 license. Without any VEELA data in training, OpenPVHV separated the portal and hepatic veins on the external hidden test set at a level comparable

to the challenge participants. It exceeded the general-purpose model VoxTell on all three metrics for both veins and may serve as an alternative to the TotalSegmentator liver-vessel task for vessel segmentation.

Declarations

Funding: This research received no external funding.

Competing interests: The author has no competing interests to declare.

Research involving human participants: This article does not contain any studies with human participants performed by the author. Only anonymized public data were used; the ethics approvals for HVM and CRLM are described in the source publications [11, 17]. VEELA is public data whose images include those of the preceding CHAOS challenge, for which ethics approval and anonymization are described in the CHAOS paper [20].

Informed consent: HVM and CRLM are anonymized retrospective data, and the requirement for consent was waived in the source studies [11, 17].

Data availability: HVM is available from Zenodo at <https://doi.org/10.5281/zenodo.19885789> under CC BY 4.0, CRLM from The Cancer Imaging Archive at <https://doi.org/10.7937/QXK2-QG03> under CC BY 4.0 [18], and VEELA from Synapse as syn65471967 [10]. The case correspondence table, the training/test split, and the list of CRLM cases used to train the HVM+CRLM model with the reasons for exclusion are released with the code.

Code and weight availability: The code is released under the Apache-2.0 license at <https://github.com/meimeimei1223/OpenPVHV>, and the weights on Zenodo (<https://doi.org/10.5281/zenodo.23619753>). OpenPVHV is provided as a standalone script, not as an official TotalSegmentator task. The evaluation/ directory of the same repository contains the scripts that reproduce the inference used for evaluation, the HVM split by random seed and case name, and the HVM image–annotation correspondence table. The HVM and CRLM images and annotations themselves are not included. VEELA was used only for evaluation, and its images, reference labels, and predictions are not included in the repository. Neither the VoxTell weights and outputs nor the HVM+CRLM model weights are included.

Author contributions: M.K. conceived and designed the study, performed all experiments and analyses, and wrote the manuscript.

References

1. Miyata A, Arita J, Kawaguchi Y, Hasegawa K, Kokudo N (2020) Simulation and navigation liver surgery: an update after 2,000 virtual hepatectomies. *Glob Health Med* 2:298–305.

<https://doi.org/10.35772/ghm.2020.01045>

2. Kazami Y, Kaneko J, Keshwani D, Kitamura Y, Takahashi R, Mihara Y et al (2023) Two-step artificial intelligence algorithm for liver segmentation automates anatomic virtual hepatectomy. *J Hepatobiliary Pancreat Sci* 30:1205–1217. <https://doi.org/10.1002/jhbp.1357>
3. Keshwani D, Kitamura Y, Ihara S, Iizuka S, Simo-Serra E (2020) TopNet: topology preserving metric learning for vessel tree reconstruction and labelling. *MICCAI 2020. Lecture Notes in Computer Science*, vol 12266. Springer, Cham, pp 14–23. https://doi.org/10.1007/978-3-030-59725-2_2
4. Li S, Li XG, Zhou F et al (2024) Automated segmentation of liver and hepatic vessels on portal venous phase computed tomography images using a deep learning algorithm. *J Appl Clin Med Phys* 25:e14397. <https://doi.org/10.1002/acm2.14397>
5. Xie T, Zhou J, Zhang X et al (2024) Fully automated assessment of the future liver remnant in a blood-free setting via CT before major hepatectomy via deep learning. *Insights Imaging* 15:164. <https://doi.org/10.1186/s13244-024-01724-6>
6. Chierici A, Guzzi L, Goffart S et al (2025) Fully automatic artificial intelligence liver anatomy segmentation in the management of colorectal liver metastases. *Cureus* 17:e86072. <https://doi.org/10.7759/cureus.86072>
7. Wasserthal J, Breit HC, Meyer MT et al (2023) TotalSegmentator: robust segmentation of 104 anatomic structures in CT images. *Radiol Artif Intell* 5:e230024. <https://doi.org/10.1148/ryai.230024>
8. Wasserthal J (2026) TotalSegmentator (version 2.18.0), README. <https://github.com/wasserth/TotalSegmentator> (commit a026545). Accessed 14 Oct 2026
9. Toprak T, Yazıcı ZA, Öksüz İ, Koska İÖ, Bulucu P, Gezer NS et al (2025) VEELA challenge - vessel extraction and extrication for liver analysis. In: 2025 IEEE 35th International Workshop on Machine Learning for Signal Processing (MLSP), pp 1–6. <https://doi.org/10.1109/MLSP62443.2025.11204261>
10. VEELA 2025 challenge (2025) Vessel Extraction and Extrication for Liver Analysis. Synapse. <https://www.synapse.org/Synapse:syn65471967>. Accessed 16 Oct 2026
11. Xie T, Li X et al (2026) Hepatic Vessel Map (HVM): an expert-annotated CT dataset for clinically applicable AI in liver vascular segmentation and surgical planning. *Sci Data* 13:1259. <https://doi.org/10.1038/s41597-026-07550-3>
12. tao-yi (2026) Missing pretrained models for Portal_veins_segment and Hepatic_veins_segment. zhengyidong135/nnunet, GitHub issue #1. <https://github.com/zhengyidong135/nnunet/issues/1>. Accessed 14 Oct 2026
13. WCM-HPB (2026) Hepatic-Vessel-Segmentation, release hepatic-vessels-20260921; RELEASE.md as first committed (4704c42, 22 Sep 2026) and current README (39bba02, 1 Oct 2026). <https://github.com/WCM-HPB/Hepatic-Vessel-Segmentation>. Accessed 14 Oct 2026. Archived at Software Heritage, swb:1:snp:2e16f750de271f29a7ebb13bcea491d8ffcf3d9a
14. Onsakon S (2026) Hepatic-Vessel-Segmentation: improving hepatic vessel segmentation with VSNet on a reannotated Task08 dataset (independent study report, January 2026). GitHub.

- <https://github.com/SteveOat/Hepatic-Vessel-Segmentation>. Accessed 19 Oct 2026
15. Rokuss M, Langenberg M, Kirchhoff Y et al (2026) VoxTell: free-text promptable universal 3D medical image segmentation. In: Proc IEEE/CVF CVPR, pp 37538–37557. arXiv:2511.11450
 16. Rokuss M (2026) VoxTell: model card and weights, versions v1.0 and v1.1 (CC BY-NC-SA 4.0). Hugging Face. <https://huggingface.co/mrokuss/VoxTell>. Accessed 20 Oct 2026
 17. Simpson AL, Peoples J, Creasy JM, Fichtinger G, Gangai N, Keshavamurthy KN et al (2024) Preoperative CT and survival data for patients undergoing resection of colorectal liver metastases. Sci Data 11:172. <https://doi.org/10.1038/s41597-024-02981-2>
 18. Simpson AL, Peoples J, Creasy JM, Fichtinger G, Gangai N, Lasso A et al (2023) Preoperative CT and survival data for patients undergoing resection of colorectal liver metastases (Colorectal-Liver-Metastases), version 2 [Data set]. The Cancer Imaging Archive. <https://doi.org/10.7937/QXK2-QG03>
 19. Yazıcı ZA, Gezer NS, Öksüz İ, Koska İÖ, Toprak T, Bulucu P et al (2026) VEELA: a clinically-constrained benchmark for liver vessel segmentation in computed tomography angiography. <https://doi.org/10.48550/arXiv.2605.22357>. arXiv:2605.22357v1
 20. Kavur AE, Gezer NS, Barış M, Aslan S, Conze PH, Groza V et al (2021) CHAOS Challenge - combined (CT-MR) healthy abdominal organ segmentation. Med Image Anal 69:101950. <https://doi.org/10.1016/j.media.2020.101950>
 21. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH (2021) nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods 18:203–211. <https://doi.org/10.1038/s41592-020-01008-z>
 22. Antonelli M, Reinke A, Bakas S, Farahani K, Kopp-Schneider A, Landman BA et al (2022) The Medical Segmentation Decathlon. Nat Commun 13:4128. <https://doi.org/10.1038/s41467-022-30695-9>
 23. Shit S, Paetzold JC, Sekuboyina A, Ezhov I, Unger A, Zhylyka A et al (2021) cIDice: a novel topology-preserving loss function for tubular structure segmentation. Proc IEEE/CVF CVPR 16555–16564. <https://doi.org/10.1109/CVPR46437.2021.01629>
 24. Nikolov S, Blackwell S, Zverovitch A, Mendes R, Livne M, De Fauw J et al (2021) Clinically applicable segmentation of head and neck anatomy for radiotherapy: deep learning algorithm development and validation study. J Med Internet Res 23:e26151. <https://doi.org/10.2196/26151>
 25. Yazıcı ZA (2026) VEELA2025_Evaluation: evaluation scripts of the VEELA 2025 challenge (commit 65fd000, 8 Mar 2026). GitHub. https://github.com/yaziciz/VEELA2025_Evaluation. Accessed 24 Oct 2026
 26. Yazıcı ZA, Öksüz İ, Ekenel HK (2024) GLIMS: attention-guided lightweight multi-scale hybrid network for volumetric semantic segmentation. Image Vis Comput 146:105055. <https://doi.org/10.1016/j.imavis.2024.105055>
 27. Salvi M, Gassino R, Ranzini MBM, Ferillo G, Vacca L, Isu G et al (2026) Refined annotations of intrahepatic portal and hepatic veins from thin-slice CT images. Sci Data. <https://doi.org/10.1038/s41597-026-08346-1>

28. Karimi D, Dou H, Warfield SK, Gholipour A (2020) Deep learning with noisy labels: exploring techniques and remedies in medical image analysis. Med Image Anal 65:101759. <https://doi.org/10.1016/j.media.2020.101759>

Figures

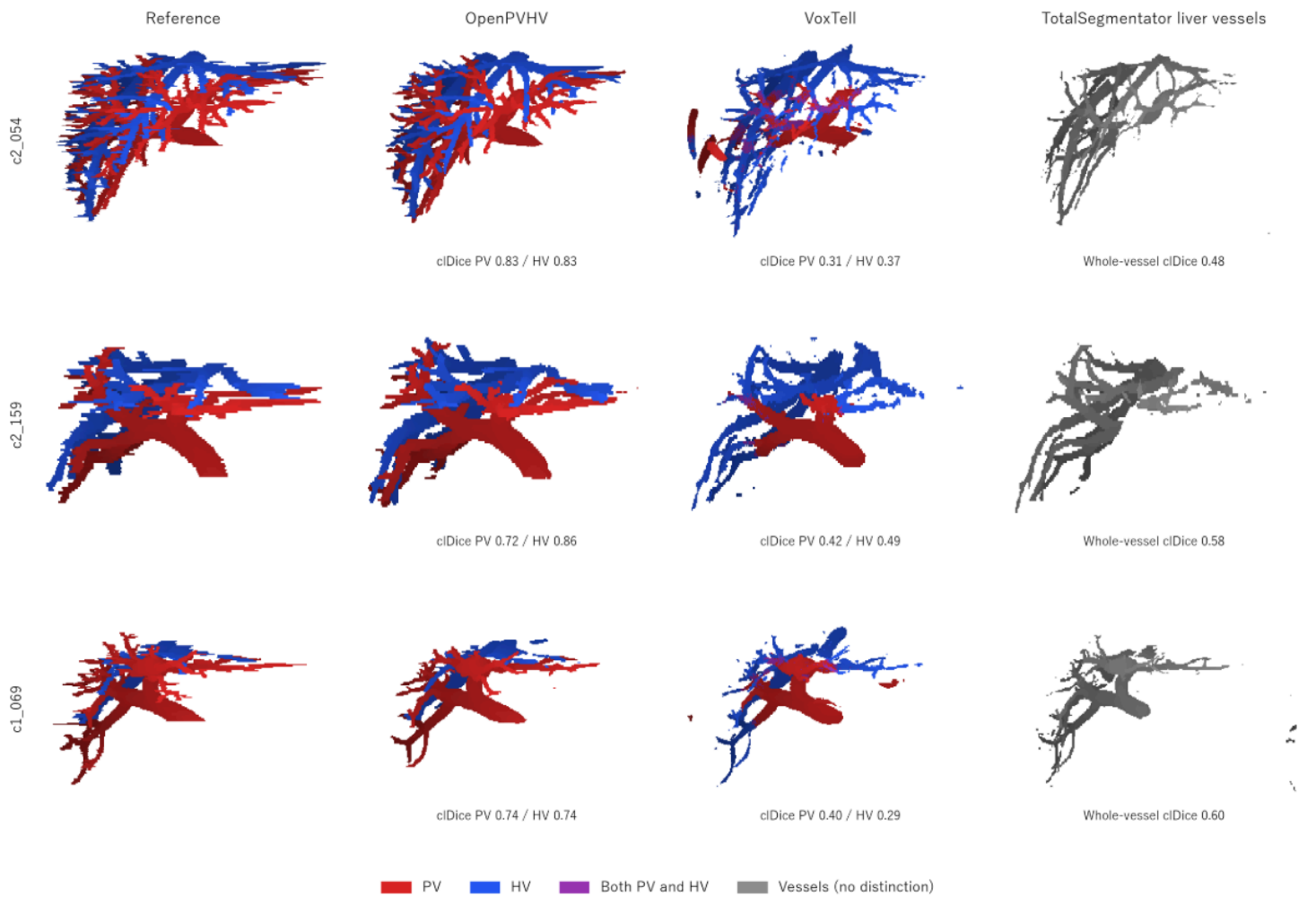


Figure 1

Three HVM internal-test cases at the 75th, 50th, and 25th percentiles of OpenPVHV cIDice (mean of portal and hepatic veins) among the 55 cases with non-empty output. Vessel surfaces are viewed from the front; brighter is closer. Red, portal vein (PV); blue, hepatic vein (HV); purple, VoxTell output where the two prompts overlap; grey, TotalSegmentator liver-vessel task, which does not separate the two veins. Values are cIDice within the liver mask dilated by 10 mm; the VoxTell output is shown within the same region. Reference: HVM [11] (CC BY 4.0)