Supporting information for " Preoperative diagnosis of hepatocellular carcinoma patients with bile duct tumor thrombus using deep learning method "

Table S1 Overall survival rates of HCC patients with and without BDTT.

	HCC with BDTT			HCC without BDTT		
Reference	1 year	3 year	5 year	1 year	3 year	5 year
Lu et al. [1]	77.0%	42.0%	23.0%	80.0%	60.0%	48.0%
Wang et al. [2]	81.8%	50.0%	37.5%	90.9%	66.9%	55.9%
Shao et al. [3]	70.3%	25.9%	7.4%	90.6%	54.0%	37.7%
Rammohan et al. [4]	82.0%	48.0%	10.0%	90.0%	55.0%	38.0%
Shiomi et al. [5]	-	46.6%	28.0%	-	63.2%	47.5%

Table S2 D-TPRs of 16 patients in the case group (average D-TPR: 0.92).

ID	tDBD	dDBD	D-TPR	ID	tDBD	dDBD	D-TPR
1	5	5	1.00	9	18	15	0.83
2	10	8	0.80	10	11	11	1.00
3	11	11	1.00	11	4	4	1.00
4	11	9	0.82	12	18	17	0.94
5	10	9	0.90	13	9	8	0.89
6	5	5	1.00	14	8	7	0.88
7	7	7	1.00	15	17	14	0.82
8	7	7	1.00	16	16	13	0.81

tDBD: the number of true DBDs, dDBD: the number of successfully detected DBDs, D-TRP = dDBD/tDBD.

Appendix A Model configurations

Main configurations of the model include the initial sizes and aspect ratios of the anchors, the intersection over union (IoU) threshold and the non-maximum suppression (NMS) threshold, and the parameters for image standardization. To determine adequate size and aspect ratio of the anchors, statistical analysis was carried out based on the annotations, including plotting histogram and calculating quantiles. From Fig. S1 and Table S3, the aspect ratios of the bounding boxes concentrate between 0.50 and 3.50. Considering computational efficiency, the initial aspect ratios of the anchors were set to five values: 0.33, 0.50, 1.00, 2.00, 3.00. From Table S4, the minimum, 25% quantile, median, 75% quantile, and maximum of the bounding boxes' sizes are approximately 16^2, 32^2, 48^2, 68^2 and 168^2 pixels, respectively. Therefore, the initial anchor sizes were set according to these quantiles with each size corresponding to one feature map output by FPN.

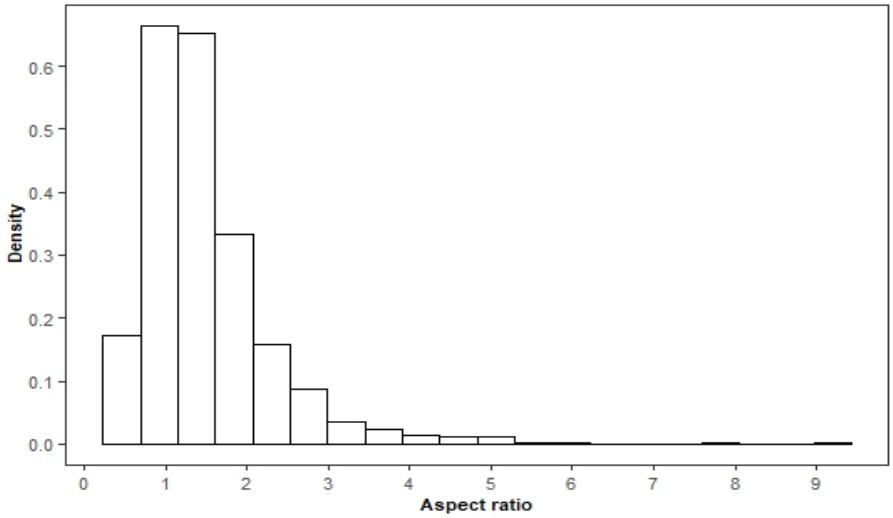

Fig. S1 Histogram of aspect ratios of bounding boxes.

Table S3 Quantiles of aspect ratios of the bounding box.

Prob	0%	5%	25%	50%	75%	95%	100%
Quantile	0.35	0.59	0.96	1.26	1.80	2.92	9.09

The IoU value is defined as the area of the overlapping part of the bounding box b_D generated by the model and the ground truth box b_G labeled by the doctor divided by the total area of the two boxes:

$$IoU(b_G, b_D) = \frac{\text{area}(b_G \cap b_D)}{\text{area}(b_G \cup b_D)}, \tag{A.1}$$

where area$(\cdot)$ is the function calculating the area of a region. IoU threshold is used to determine if an anchor is foreground or background. The non-maximum suppression (NMS) threshold is used to filter out anchors with high degree of overlap, i.e., the redundant bounding boxes with higher overlap than NMS threshold will be removed when outputting detection results. The IoU threshold of the foregroundwas set to 0.70 for RPN so

Table S4 Quantiles of sizes of the bounding boxes.

Prob	0%	25%	50%	75%	100%
Quantile	8.00^2	16.31^2	24.37^2	34.15^2	83.79^2

that the criteria for determining whether an anchor is positive is strict, and the NMS thresholds were set to 0.70 and 0.50 for RPN and R-CNN separately to reduce the number of anchors and improve training efficiency.

For Faster R-CNN, the input image should be standardized and then scaled to a fixed size before extracting features.The images were normalized so that the sample mean and the standard deviation were 0 and 1, respectively. Then, the image were magnified four times its original size (twice the x-axis and twice the y-axis) so that the DBD could be enlarged without excessived deformation.

Appendix B Training details

Due to the limited data, data augmentation was conducted in the training process. There are many approaches to implementing data augmentation. Due to the tiny DBD size and the grayscale characteristic of CT images, some methods (e.g., changing brightness and adding noise) are not suitable for the current study. During our training process, four techniques for data augmentation were used by turns: random horizontal and vertical translation, random cropping, random clockwise or counter-clockwise rotation up to 10 degrees, and random horizontal or vertical flips with a probability equal to 0.5, 0.5, 0.5, and 0.6.

In the model training, we take the strategy of transfer learning due to the limited data. Specifically, the Faster R-CNN model was pretrained on MS COCO[6] data set using framework MMdetection[7], then we continued to train the model for another 24 epochs (schedule 2x) with stochastic gradient descent (SGD) algorithm as the optimizer. The initial learning rate was set to 0.0025, and was decayed by 0.1 at epoch 16 and

epoch 22. The L2 penalty coefficient (also known as weight decay) was set to 0.0001. The model was trained on one graphic card (GTX 2080Ti; NVIDIA, Santa Clara, Calif) using MMdetection.

References

[1] Lu W, Tang H, Yang Z, et al, A proposed modification for the Barcelona clinic liver cancer staging system: Adding bile duct tumor thrombus status in patients with hepatocellular carcinoma[J]. The American Journal of Surgery, 2020, 220(4): 965-971.

[2] Wang D D, Wu L Q, Wang Z S, Prognosis of hepatocellular carcinoma with bile duct tumor thrombus after R0 resection: A matched study[J]. Hepatobiliary & Pancreatic Diseases International, 2016, 15(6): 626-632.

[3] Shao W, Sui C, Liu Z, et al, Surgical outcome of hepatocellular carcinoma patients with biliary tumor thrombi[J]. World Journal of Surgical Oncology, 2011, 9-2.

[4] Rammohan A, Sathyanesan J, Rajendran K, et al. Bile duct thrombi in hepatocellular carcinoma: Is aggressive surgery worthwhile?[J]. HPB, 2015, 17(6): 508-513.

[5] Shiomi M, Kamiya J, Nagino M, et al. Hepatocellular carcinoma with biliary tumor thrombi: Aggressive operative approach after appropriate preoperative management[J]. Surgery, 2001, 129(6): 692-698.

[6] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: Common objects in context. In: Fleet D, Pajdla T, Schiele B, editors. Computer Vision-ECCV 2014. 2014: 740-755.

[7] Chen K, Wang J, Pang J, et al. MMDetection: Open MMLab Detection Toolbox and Benchmark. 2019. https://arxiv.org/abs/1906.07155. Accessed December 29, 2022.