

Review

# Accuracy of Medical Image–Based Deep Learning for Detecting Microvascular Invasion in Hepatocellular Carcinoma: Systematic Review and Meta-Analysis

Wei Feng<sup>1\*</sup>, MM; Bo Qu<sup>2\*</sup>, MM; Shuo Han<sup>3</sup>, PhD

<sup>1</sup>Department of Ultrasound, the Fourth Affiliated Hospital, China Medical University, Shenyang, China

<sup>2</sup>Department of Urology, Jinqiu Hospital of Liaoning Province, Shenyang, China

<sup>3</sup>Department of Cardiology, the Fourth Affiliated Hospital, China Medical University, Shenyang, China

\*these authors contributed equally

## Corresponding Author:

Shuo Han, PhD

Department of Cardiology

the Fourth Affiliated Hospital, China Medical University

No. 4, Chongshan East Road, Huanggu District

Shenyang 110032

China

Phone: 86 18900912036

Email: [hans@cmu.edu.cn](mailto:hans@cmu.edu.cn)

## Abstract

**Background:** Hepatocellular carcinoma (HCC) is a leading cause of cancer-related mortality worldwide. Microvascular invasion (MVI) is a critical pathological indicator of postoperative recurrence and poor prognosis in patients with HCC. Some researchers have explored the diagnostic accuracy of deep learning (DL) based on various imaging modalities for MVI.

**Objective:** This meta-analysis aimed to systematically evaluate the preoperative diagnostic performance of DL models using medical images to predict MVI in HCC, and to investigate the impact of different imaging modalities and validation strategies on model performance and generalizability.

**Methods:** PubMed, Cochrane Library, Embase, and Web of Science were searched up to October 16, 2025. Studies investigating the detection of MVI in HCC using imaging-based DL techniques were eligible. Studies focusing solely on image segmentation were excluded. The Quality Assessment of Diagnostic Accuracy Studies-2 tool was used to assess risk of bias. A bivariate mixed-effects meta-analysis was performed to calculate the pooled sensitivity, specificity, and area under the summary receiver operating characteristic curve (SROC). Subgroup analyses were conducted by imaging modality and validation set generation method.

**Results:** This meta-analysis included 52 studies with 19,531 patients with HCC. The pooled analysis revealed that imaging-based DL models had an overall sensitivity of 0.80 (95% CI 0.78-0.83), a specificity of 0.82 (95% CI 0.80-0.85), and an SROC of 0.88 for MVI prediction. Subgroup analysis showed that models based on preoperative contrast-enhanced computed tomography performed excellently, with a sensitivity of 0.84 (95% CI 0.79-0.88), a specificity of 0.83 (95% CI 0.77-0.88), and an SROC of 0.90. These results suggest that contrast-enhanced computed tomography is the most promising noninvasive method for current clinical applications. Meanwhile, DL models using pathological sections achieved the highest diagnostic performance: a sensitivity of 0.91 (95% CI 0.87-0.94), a specificity of 0.90 (95% CI 0.68-0.97), and an SROC of 0.92. This establishes the ultimate benchmark for performance optimization for all noninvasive models. A key finding was that model performance was less consistent in independent external validation (SROC: 0.85) than in internal validation (SROC: 0.90). This discrepancy indicates that overreliance on internal validation may overestimate model efficacy and underscores the decisive role of rigorous external validation in assessing real-world generalizability.

**Conclusions:** This study is the first to systematically assess the use of imaging-based DL for diagnosing MVI in HCC. The results demonstrate a significant potential for these models in predicting MVI. However, their clinical applicability requires rigorous evaluation, given the scarcity of independent external validation cohorts, notable heterogeneity among them, and the observed decline in model performance. Therefore, prospective, multicenter studies following standardized reporting

guidelines are a critical future direction. These studies should also focus on developing integrated algorithms that translate histopathological insights into preoperative imaging data to establish robust clinical tools.

**Trial Registration:** PROSPERO CRD42024613733; <https://www.crd.york.ac.uk/PROSPERO/view/CRD42024613733>

*J Med Internet Res* 2026;28:e82000; doi: [10.2196/82000](https://doi.org/10.2196/82000)

**Keywords:** deep learning; hepatocellular carcinoma; medical imaging; microvascular invasion ; artificial intelligence

## Introduction

Hepatocellular carcinoma (HCC) is the most common pathological subtype of primary liver cancer, accounting for about 90% of cases [1]. Globally, HCC is the fourth leading cause of cancer-related deaths [2]. According to recent epidemiological data, the age-standardized incidence and mortality rates of HCC are highest in Africa and the Western Pacific region. Over 70% of global HCC cases occur in Asia [1]. Despite advancements in treatment modalities for HCC, including liver transplantation, surgical resection, transarterial chemoembolization, local ablation, targeted therapy, and immunotherapy, the 5-year relative survival rate remains below 20% [3]. Even after complete surgical tumor removal, around 50%-70% of patients with HCC experience tumor recurrence within 5 years postsurgery [4]. Consequently, HCC has become a significant oncological burden that threatens human life.

Microvascular invasion (MVI) is the pathological process by which tumor cells invade the microvascular structures of the liver tissue surrounding an HCC lesion. MVI occurs in approximately 30%-50% of cases. It is a significant factor in HCC recurrence after surgery and is associated with poor prognoses in patients with HCC [4,5]. Studies have shown that individuals with HCC and MVI-positive status have significantly lower 5-year disease-free survival and overall survival rates than those with MVI-negative status [6, 7]. Notably, MVI status directly influences treatment strategy selection. Wide-range hepatectomy (resection of  $\geq 3$  liver segments) is recommended for individuals at high preoperative risk for MVI. This method has a significantly lower 5-year cumulative recurrence rate than limited resection (26.6% vs 58.3%,  $P=.040$ ) [8]. However, other studies have found that adjuvant hepatic arterial infusion chemotherapy does not significantly improve the survival of high-risk MVI-positive individuals compared with the untreated group ( $P=.61$ ). Nevertheless, hepatic arterial infusion chemotherapy significantly improves the prognosis of low-risk patients with MVI ( $P<.001$ ) [9]. Furthermore, individuals with MVI undergoing radiofrequency ablation have a significantly higher recurrence risk than those undergoing radical surgery ( $P<.05$ ) [10]. Therefore, accurately identifying MVI preoperatively is significant for formulating individualized, comprehensive treatment regimens and improving patient prognosis [11].

Currently, a definitive diagnosis of MVI relies on a postoperative pathological examination. However, this examination is subject to biases resulting from the quality of slide preparation and interobserver heterogeneity. These factors may lead to diagnostic inaccuracies. Furthermore, the

absence of preoperative MVI information restricts its use in personalized treatment decisions. Therefore, developing efficient MVI auxiliary detection tools is crucial for optimizing clinical management of HCC.

While imaging examinations are crucial for evaluating MVI, predictions based on traditional imaging features rely heavily on radiologists' subjective interpretations. A systematic review and meta-analysis of 19 studies involving 1920 patients revealed that traditional contrast-enhanced features on magnetic resonance imaging (MRI) showed poor overall diagnostic performance in predicting MVI. Only peritumoral enhancement in the arterial phase exhibited moderate diagnostic accuracy. The combined efficacy of other features, such as peritumoral hypointensity in the hepatobiliary phase and irregular margins, was insufficient to meet the requirements for precise preoperative clinical assessment [12]. Recent progress in data mining techniques has accelerated the growth of radiomics. This technique assists in the analysis of imaging features (eg, shape, intensity, and texture) that are difficult for the human eye to perceive. It can overcome some of the limitations of subjectivity. However, radiomic features are mostly low- or mid-level and susceptible to noise interference. They may also not fully reflect tumor heterogeneity [13,14]. Zhang et al [15] noted that radiomics based on single-modality medical imaging is inherently limited. Due to constraints in imaging principles, such approaches can only reflect partial tumor information. Furthermore, when features are extracted using the entire tumor as the region of interest, information about intratumoral heterogeneity is inevitably lost. Additionally, these single-modality radiomic features are susceptible to image noise and variations in scanning parameters, which further compromise the model's ability to capture tumor heterogeneity comprehensively. In contrast, deep learning (DL) uses multi-layer neural networks and an end-to-end learning mode to directly extract multi-level abstract high-order features from original images. This improves the predictive performance, interpretability, and generalizability of models. DL is expected to provide a new paradigm for the preoperative, noninvasive assessment of MVI [16,17]. However, existing studies often focus on a single imaging modality or have small sample sizes. These studies lack a systematic comparison of DL model performance across different imaging modalities, which limits the interpretation of the advantages of DL in detecting MVI and poses challenges to the development or update of intelligent auxiliary diagnostic tools. Consequently, this meta-analysis was conducted to systematically evaluate the diagnostic efficacy of DL models based on medical images for MVI, as well as to explore the impact of different imaging modalities, validation strategies, and sources of heterogeneity on model

performance and generalizability, to provide evidence-based support for the development or update of future intelligent auxiliary diagnostic tools.

## Methods

### Study Registration

This meta-analysis was prospectively registered with the PROSPERO (CRD42024613733). This systematic review and meta-analysis of diagnostic test accuracy was reported in accordance with the PRISMA-DTA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Diagnostic Test Accuracy studies) [18] guidelines in [Checklist 1](#). Due to the absence of subject information collection and its lack of impact on clinical diagnosis and treatment, ethical approval and informed consent were waived.

### Eligibility Criteria

#### Inclusion Criteria

The inclusion criteria are as follows:

- Original research with full text published in English (including cohort, case-control, and cross-sectional studies).
- MVI status in HCC individuals had to be confirmed by histopathology or biopsy.
- Studies had to develop complete DL models based on medical images to detect MVI status in patients with HCC.
- English-language studies.

#### Exclusion Criteria

The exclusion criteria are as follows:

- Meta-analyses, reviews, guidelines, or expert opinions.
- Only differential factor analysis was implemented without a comprehensive DL model.
- Studies lacking outcome measures of predictive accuracy for the machine learning model, including sensitivity, C-index, accuracy, specificity, precision,  $F_1$ -score, and confusion matrix.
- Only image segmentation was performed.

### Search Strategies and Data Sources

We conducted a systematic literature search in accordance with the PRISMA-S (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for literature searches; completed checklist is available in [Checklist 2](#)) [19]. Relevant English-language publications were retrieved from PubMed, Web of Science, the Cochrane Library, and Embase, with the search covering all records up to October 16, 2025. The search strategy used both MeSH (Medical Subject Headings) and free-text keywords, including MeSH terms such as “Carcinoma, Hepatocellular,” “liver cell carcinoma,” and “deep learning.” Boolean operators were used to integrate MeSH terms and free-text terms, constructing tailored search queries for each database. Furthermore, the reference lists of identified review articles were

manually screened to locate any additional eligible studies. No prospective study registries were searched, and no attempts were made to obtain unpublished data or to contact study authors for further information. The complete search strings for each database are provided in [Table S1 in Multimedia Appendix 1](#).

### Study Selection

Retrieved articles were imported into EndNote. After removing duplicates, the remaining articles were reviewed by title and abstract to identify initially eligible articles. Then, the full texts were downloaded and screened to determine the final eligible articles. Two researchers (WF and BQ, with 6 and 4 years of meta-analysis experience, respectively) performed the review independently. Interresearcher agreement during the literature screening was assessed using the  $\kappa$  coefficient ( $\kappa=0.93$ ). Any disagreements were resolved in a consensus meeting with a third researcher (SH, with 10 years of experience in meta-analysis).

### Data Extraction

Before data extraction, a standardized spreadsheet was generated. The content to be extracted included the following: publication year, patient source, author, image source, manual segmentation, number of patients with MVI, total number of patients with HCC, number of patients with MVI in the training set, number of patients with HCC in the training set, validation set generation method, number of patients with MVI in the validation set, number of patients with HCC in the validation set, confusion matrix, sensitivity, specificity, precision, and accuracy.

Two researchers (WF and BQ, with 14 and 16 years of medical experience, respectively) carried out the data extraction independently. They then cross-checked their results. Any discrepancies were resolved through consultation with a third researcher (SH, with 13 years of medical experience).

### Risk of Bias in Studies

The Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool was used to evaluate the overall risk of bias (ROB) and applicability of the eligible studies. The QUADAS-2 instrument encompasses 4 domains: index test, patient selection, flow and timing, and reference standard. Each domain includes specific questions. Answers to these questions are categorized as “yes,” “no,” or “unclear.” These answers correspond to ROB ratings of “unclear,” “high,” or “low.” If all key questions within a domain received a “yes” answer, the domain was rated as having a low ROB. If any key question received a “no” answer, a potential ROB was indicated, and the researcher judged the ROB according to established guidelines. An “unclear” rating was assigned when the literature did not provide sufficient detail for the researcher to make a judgment.

Two researchers completed the QUADAS-2 assessment independently. Any discrepancies were resolved through discussion with a third researcher.

## Synthesis Methods

A meta-analysis was performed using a bivariate mixed-effects model based on diagnostic 2x2 tables [20]. For studies that did not report these tables directly, we derived them from the available specificity, sensitivity, positive predictive value, accuracy,  $F_1$ -score, and case numbers. Throughout the analysis, the explicitly defined independent validation cohort from each study served as the unit of analysis. Each data point corresponded to distinct and nonoverlapping patient samples, which ensured the independence of the pooled results and mitigated potential bias from data reuse. Using the bivariate mixed-effects model, we computed the pooled estimates for specificity, sensitivity, negative likelihood ratio (LR-), positive likelihood ratio (LR+), diagnostic odds ratio (DOR), and area under the summary receiver operating characteristic (SROC) curve, along with their 95% CIs [20]. The Spearman correlation coefficient was used to evaluate the threshold effect and its contribution to between-study heterogeneity. Small-study effects were assessed using Deeks funnel plot asymmetry test. For subgroups with fewer than 10 studies, a Doi plot was used to informally assess publication bias. The degree of bias was determined based on the absolute value of the Luis Furuya-Kanamori (LFK) index. A value less than

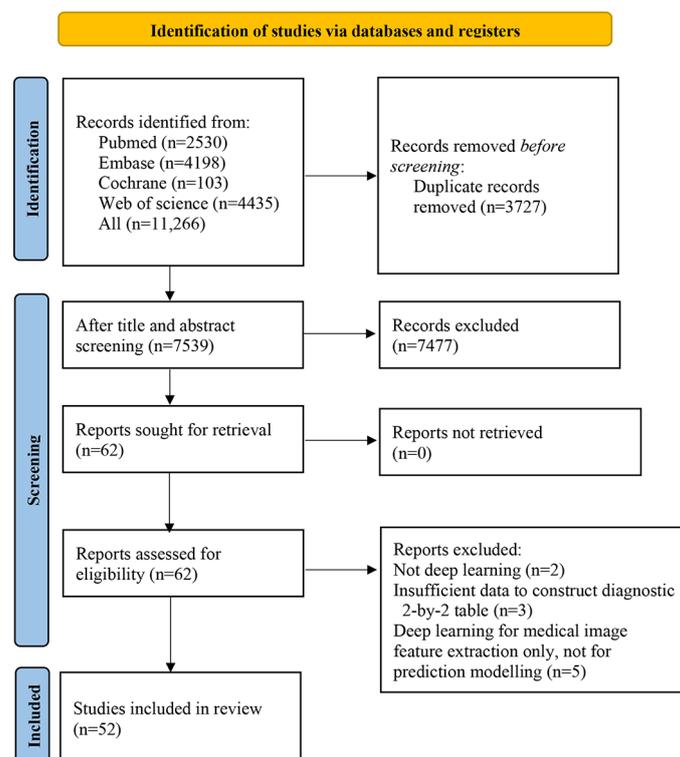
1 suggests minor publication bias, a value between 1 and 2 indicates moderate publication bias, and a value exceeding 2 suggests substantial publication bias. During the meta-analysis, the validation set was used, and subgroup analyses were performed according to the validation set generation method and image type. A  $P$  value less than .05 was considered statistically significant.

## Results

### Study Selection

Database retrieval yielded 11,266 articles, of which 7539 remained after duplicate removal. Following title and abstract screening, 62 articles were selected for a full-text review. The full-text review subsequently excluded 10 records: 2 non-DL studies, 3 studies with insufficient data to construct diagnostic 2x2 tables, and 5 studies that used DL solely for medical image segmentation without establishing HCC MVI prediction models. Finally, 52 articles [14,16,21-70] met the eligibility criteria. The detailed process is illustrated in Figure 1.

**Figure 1.** PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the study screening and selection process for this meta-analysis.



### Study Characteristics

This meta-analysis included 52 studies published between 2019 and 2025. All of the studies used histopathological diagnosis as the gold standard for MVI. All 52 studies were case-control studies, encompassing 19,531 individuals with HCC, of whom 8161 were MVI cases. Regarding population source, 32 (61.5%) of the studies were single-center, 19

(36.5%) studies were multicenter, and 1 (1.9%) study was based on a registry database.

These studies primarily used single-modality imaging techniques for image modeling methods: contrast-enhanced computed tomography (CECT, n=16), contrast-enhanced magnetic resonance imaging (CEMRI, n=19), MRI (n=5), contrast-enhanced ultrasound (CEUS, n=5), and pathological

sections (n=2). Five additional studies integrated multimodal imaging (computed tomography + positron emission tomography: 1, CECT +CEMRI: 4) for modeling. Regarding image segmentation methods, 38 of the 52 (73.1%) studies used manual segmentation, while the remaining studies used fully automatic segmentation (n=5) or semiautomatic segmentation (n=9).

Concerning model validation strategies, 23 studies used randomly sampled internal validation, 15 used performance through cross-validation, and 14 used external validation with independent cohorts (Table 1).

**Table 1.** Basic characteristics of the 52 eligible studies evaluating medical image-based deep learning models for preoperatively detecting microvascular invasion (MVI) in patients with hepatocellular carcinoma.

Author (year of publication)	Basic characteristics					Training versus validation cohort characteristics					
	Source of the patients <sup>a</sup>	Imaging modality	Segmentation	Diagnostic criteria for MVI	Patients with MVI	Sample size	Patients with MVI in the training set	Sample size in the training set	Type of validation	Patients with MVI in the validation set	Sample size in the validation set
Zhang et al (2024) <sup>a</sup> [21]	Single center +registration database (TCGA) <sup>b</sup>	Pathological section	Semiautomatic segmentation	Pathological diagnosis	576	1111	328	530	Random sampling (7:3), external validation (registration database TCGA)	Validation set 1: 143; Validation set 2: 105	Validation set 1: 223; Validation set 2: 358
Lei et al (2024) [22]	Multicenter	CECT <sup>c</sup> +CEMRI <sup>d</sup>	Semiautomatic segmentation	Pathological diagnosis	104	345	89	301	Random sampling (6:4)	15	44
Liu et al (2024) [23]	Single center	CEMRI	Manual segmentation	Pathological diagnosis	123	265	97	211	Random sampling (8:2)	26	54
Zhou et al (2024) [24]	Multicenter	CECT	Semiautomatic segmentation	Pathological diagnosis	54	140	38	98	Random sampling (7:3)	16	42
He et al (2024) [25]	Multicenter	CECT	Semiautomatic segmentation	Pathological diagnosis	299	640	172	368	Internal validation	Validation set 1: 63; Validation set 2: 64	Validation set 1: 134; Validation set 2: 138
Zhong et al (2024) [26]	Single center	CEMRI	Manual segmentation	Pathological diagnosis	82	173	57	120	Random sampling (7:3)	25	53
Wang et al (2024) <sup>a</sup> [27]	Multicenter	CEMRI	Manual segmentation	Pathological diagnosis	274	725	109	234	Random sampling (4:1); External validation (multicenter)	Validation set 1: 20; Validation set 2.1: 82; Validation set 2.2: 37; Validation set 2.3: 26	Validation set 1: 58; Validation set 2.1: 212; Validation set 2.2: 111; Validation set 2.3: 110
Yu et al (2024) <sup>a</sup> [28]	Multicenter	CECT	Semiautomatic segmentation	Pathological diagnosis	78	205	39	119	Random sampling (7:3); External validation (multicenter)	Validation set 1: 15; Validation set 2: 24	Validation set 1: 44; Validation set 2: 42
Ma et al (2025) [29]	Single center	CEMRI	Manual segmentation	Pathological diagnosis	52	117	42	94	Random sampling (8:2); 10-fold cross-validation	10	23
Zhang et al (2024) <sup>a</sup> [30]	Multicenter	CEUS <sup>e</sup>	Manual segmentation	Pathological diagnosis	219	576	175	461	Single center random sampling (8:2); External validation (multicenter)	— <sup>f</sup>	—
Wang et al (2023) <sup>a</sup> [31]	Multicenter	CECT+CEMRI	Manual segmentation	Pathological diagnosis	150	397	119	297	External validation (multicenter)	31	100
You et al (2023) [32]	Single center	CEMRI	Semiautomatic segmentation	Pathological diagnosis	70	210	56	168	Random sampling (4:1); 5-fold cross-validation	14	42
Qin et al (2023) [33]	Multicenter	CEUS	Manual segmentation	Pathological diagnosis	92	252	71	198	Random sampling (8:2)	21	54

Author (year of publication)	Basic characteristics					Training versus validation cohort characteristics					
	Source of the patients <sup>a</sup>	Imaging modality	Segmentation	Diagnostic criteria for MVI	Patients with MVI	Sample size	Patients with MVI in the training set	Sample size in the training set	Type of validation	Patients with MVI in the validation set	Sample size in the validation set
Li et al (2023) [34]	Single center	CEMRI	Manual segmentation	Pathological diagnosis	146	283	117	226	Random sampling (4:1); 5-fold cross-validation	29	57
Cao et al (2023) [35]	Single center	CECT	Manual segmentation	Pathological diagnosis	149	559	120	448	Random sampling (4:1)	29	111
Wang et al (2023) [36]	Single center	CEMRI	Manual segmentation	Pathological diagnosis	109	233	76	163	Random sampling (7:3)	33	70
Ye et al (2023) [37]	Multicenter	CECT+PET <sup>g</sup>	Manual segmentation	Pathological diagnosis	41	100	29	70	Random sampling (7:3); 5-fold cross-validation	12	30
Xu et al (2023) [38]	Single center	CECT	Manual segmentation	Pathological diagnosis	99	305	79	244	Random sampling (8:2); 5-fold cross-validation	20	61
Deng et al (2022) [39]	Single center	CECT+CEMRI	Manual segmentation	Pathological diagnosis	44	103	35	82	Random sampling (4:1); 5-fold cross-validation	9	21
Li et al (2022) [40]	Multicenter	CECT	Manual segmentation	Pathological diagnosis	433	1116	346	892	Random sampling (4:1)	87	224
Chen et al (2022) <sup>a</sup> [41]	Multicenter	Pathological section	Automatic segmentation	Pathological diagnosis	224	470	137	270	Random sampling external validation (multicenter)	Validation set 1: 43; Validation set 2: 44	Validation set 1: 80; Validation set 2: 120
Zhang et al (2022) [42]	Single center	CEUS	Manual segmentation	Pathological diagnosis	150	436	103	301	Random sampling (3:1); Internal validation	Validation set 1: 35; Validation set 2: 12	Validation set 1: 02; Validation set 2: 33
Liu et al (2022) [43]	Single center	MRI <sup>h</sup>	Manual segmentation	Pathological diagnosis	43	114	28	74	Random sampling	15	40
Sun et al (2022) [44]	Single center	CECT	Semiautomatic segmentation	Pathological diagnosis	134	358	77	193	Random sampling	Validation set 1: 23; Validation set 2: 34	Validation set 1: 61; Validation set 2: 104
Wang et al (2022) [45]	Single center	CECT	Automatic segmentation	Pathological diagnosis	68	138	54	110	Random sampling (8:1:1); 5-fold cross-validation	Validation set 1: 7; Validation set 2: 7	Validation set 1: 14; Validation set 2: 14
Xiao et al (2022) <sup>a</sup> [46]	Multicenter	CECT	Automatic segmentation	Pathological diagnosis	1103	2096	458	876	Random sampling (3:1); External validation (multicenter)	Validation set 1: 152; Validation set 2: 327; Validation set 3: 166	Validation set 1: 292; Validation set 2: 578; Validation set 3: 350
Yang et al (2022) [47]	Single center	CECT	Manual segmentation	Pathological diagnosis	36	283	25	198	Random sampling (198:85)	11	85
Sun et al (2022) [48]	Single center	CEMRI	Manual segmentation	Pathological diagnosis	185	321	86	149	Internal validation	99	172
Dai et al (2022) [49]	Single center	CECT	Manual segmentation	Pathological diagnosis	215	400	172	320	Random sampling (80:10:10)	Validation set 1: 21; Validation set 2: 22	Validation set 1: 40; Validation set 2: 40
Zhang et al (2021) [50]	Single center	CEMRI	Manual segmentation	Pathological diagnosis	92	237	61	158	Random sampling	31	79
Liu et al (2021) <sup>a</sup> [51]	Multicenter	CECT	Manual segmentation	Pathological diagnosis	135	473	68	216	Random sampling (70:30); External	Validation set 1: 28;	Validation set 1: 93;

Author (year of publication)	Basic characteristics					Training versus validation cohort characteristics					
	Source of the patients <sup>a</sup>	Imaging modality	Segmentation	Diagnostic criteria for MVI	Patients with MVI	Sample size	Patients with MVI in the training set	Sample size in the training set	Type of validation	Patients with MVI in the validation set	Sample size in the validation set
Wei et al (2021) <sup>a</sup> [52]	Multicenter	CECT + CEMRI	Manual segmentation	Pathological diagnosis	270	750	216	635	validation (multicenter) External validation (prospective, multicenter)	validation set 2: 39	Validation set 2: 164
Zhou et al (2021) [53]	Single center	CEMRI	Manual segmentation	Pathological diagnosis	—	114	—	—	Random sampling	—	—
Wang et al (2021) [54]	Single center	MRI	Manual segmentation	Pathological diagnosis	43	100	24	60	Random sampling	19	40
Zeng et al (2021) [55]	Single center	MRI	Manual segmentation	Pathological diagnosis	38	98	25	64	Random sampling; 4-fold cross-validation	13	34
Gao et al (2021) [56]	Single center	MRI	Manual segmentation	Pathological diagnosis	74	225	56	168	Random sampling	18	57
Jiang et al (2020) [57]	Single center	CECT	Manual segmentation	Pathological diagnosis	220	405	176	324	Random sampling (8:2)	44	81
Song et al (2021) [58]	Single center	CEMRI	Manual segmentation	Pathological diagnosis	225	601	174	461	Random sampling	51	140
Men et al (2019) [59]	Single center	CEMRI	Manual segmentation	Pathological diagnosis	28	63	21	47	4-fold cross-validation	7	16
Zhou et al (2022) [60]	Single center	CECT	Manual segmentation	Pathological diagnosis	145	466	97	311	3-fold cross-validation	48	155
Chu et al (2022) [61]	Single center	CEMRI	Manual segmentation	Pathological diagnosis	51	133	35	93	Random sampling (7:3)	16	40
Huang et al (2022) [62]	Single center	MRI	Manual segmentation	Pathological diagnosis	43	114	32	86	4-fold cross-validation	11	28
Wang et al (2025) [63]	Single center	CEUS	Semiautomatic segmentation	Pathological diagnosis	142	318	99	222	Random sampling; 5-fold cross-validation	43	96
Cen et al (2025) [64]	Single center	CECT	Semiautomatic segmentation	Pathological diagnosis	68	192	47	134	Random sampling (7:3)	21	58
Huang et al (2025) [65]	Multicenter	CEMRI	Manual segmentation	Pathological diagnosis	124	300	87	210	Random sampling; 5-fold cross-validation	37	90
Miao et al (2025) <sup>a</sup> [66]	Multicenter	CECT	Manual segmentation	Pathological diagnosis	206	483	136	311	Random sampling (8:2); External validation (multicenter)	Validation set 1: 32; Validation set 2: 38	Validation set 1: 77; Validation set 2: 95
Zhu et al (2025) <sup>a</sup> [14]	Multicenter	CEMRI	Manual segmentation	Pathological diagnosis	120	304	90	216	External validation (multicenter, retrospective)	30	88
Dong et al (2025) <sup>a</sup> [67]	Single center	CEMRI	Manual segmentation	Pathological diagnosis	188	519	100	263	Random sampling (4:1); External validation (multicenter)	Validation set 1: 26; Validation set 2: 27; Validation set 3: 35	Validation set 1: 66; Validation set 2: 93; Validation set 3: 97
Zhang et al (2025) [68]	Single center	CEMRI	Automatic segmentation	Pathological diagnosis	142	270	114	216	5-fold cross-validation	28	54
Zheng et al (2025) <sup>a</sup> [16]	Multicenter	CEMRI	Automatic segmentation	Pathological diagnosis	292	589	154	317	Random sampling (7:3); External validation (multicenter)	Validation set 1: 52; Validation set 2: 86	Validation set 1: 106; Validation set 2: 166
Zhao et al (2025) <sup>a</sup> [69]	Multicenter	CEMRI	Manual segmentation	Pathological diagnosis	51	145	25	66	10-fold cross-validation; External	26	79

Author (year of publication)	Basic characteristics					Training versus validation cohort characteristics					
	Source of the patients <sup>a</sup>	Imaging modality	Segmentation	Diagnostic criteria for MVI	Patients with MVI	Sample size	Patients with MVI in the training set	Sample size in the training set	Type of validation	Patients with MVI in the validation set	Sample size in the validation set
Qin et al (2025) [70]	Single center	CEUS	Manual segmentation	Pathological diagnosis	65	164	44	114	validation (multicenter) Random sampling (7:3); 10-fold cross-validation	21	50

<sup>a</sup>Studies used external validation with an independent cohort.

<sup>b</sup>TCGA: The Cancer Genome Atlas.

<sup>c</sup>CECT: contrast-enhanced computed tomography.

<sup>d</sup>CEMRI: contrast-enhanced magnetic resonance imaging.

<sup>e</sup>CEUS: contrast-enhanced ultrasound.

<sup>f</sup>Not available.

<sup>g</sup>PET: positron emission tomography.

<sup>h</sup>MRI: magnetic resonance imaging.

## ROB in Studies

Regarding patient selection, all eligible studies included consecutive or random cases. According to the QUADAS-2 assessment criteria, this study design carries an inherent high ROB in the “Patient Selection” domain. Consequently, all included studies received a “high” rating for ROB in this domain.

Due to the adoption of supervised DL, model training was based on clear pathological outcomes. However, since DL models predict by extracting inherent image features, rather than directly relying on clinical covariates, their training process’s dependence on known outcomes did not result in diagnostic information leakage. This led to a low ROB. Manual segmentation was used in 38 studies, which could have introduced operator subjectivity and led to a high ROB.

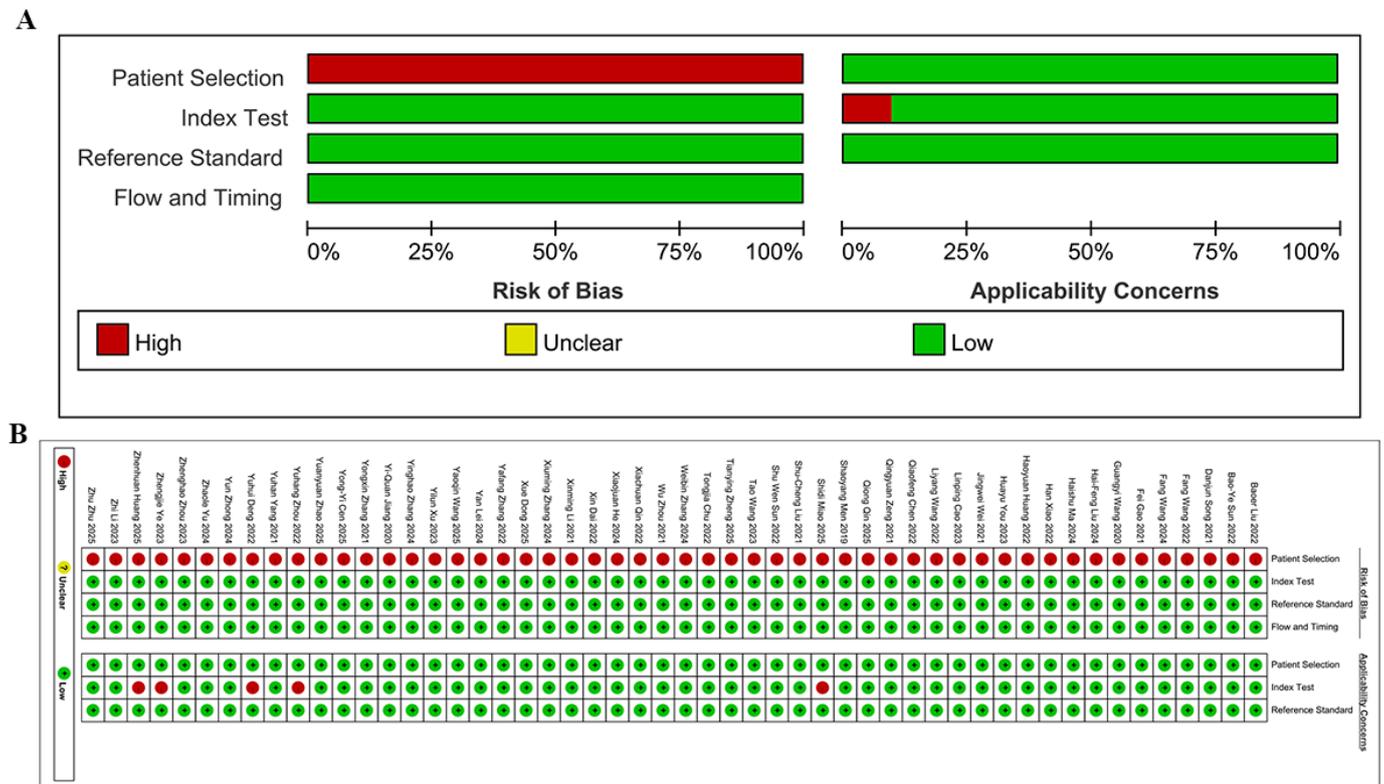
Regarding the implementation of the gold standard, all studies used histopathological diagnosis as the gold standard for MVI, ensuring the objectivity and consistency of disease

classification. This indicated a low ROB in the implementation of the gold standard.

Regarding the item “the match between the conduct and interpretation of the index test and the review question” in the QUADAS-2 scale, 5 (9.6%) out of the 52 studies did not directly report specificity values. Since specificity is a key indicator for verifying the match between a DL model and a clinical question, the absence of such data may lead to an incomplete assessment of a model’s diagnostic efficacy and weaken the reliability of a study’s conclusions. Therefore, these 5 studies were determined to have a high ROB. The remaining 47 (90.4%) studies fully reported diagnostic performance indicators, ensuring the transparency in test conduct and interpretation, and indicating a low ROB.

There was a reasonable and appropriate time interval between imaging examinations and pathological diagnoses in all eligible studies. Therefore, there did not appear to be a significant impact on the cases’ process (Figure 2).

**Figure 2.** Methodological quality assessment of the included studies based on the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) scale. (A) Summary, (B) Individual studies [14,16,21-70].



**Meta-Analysis**

**Overall**

The model’s accuracy was validated using 68 diagnostic fourfold tables. The Spearman correlation coefficient was 0.02, indicating a minimal threshold effect. This effect accounted for none of the observed between-study

heterogeneity. The pooled analysis revealed the following results: sensitivity 0.80 (95% CI 0.78-0.83,  $I^2=65.52%$ ), specificity 0.82 (95% CI 0.80-0.85,  $I^2=79.13%$ ), LR+4.6 (95% CI 3.9-5.3), LR- 0.24 (95% CI 0.21-0.27), DOR 19 (95% CI 15-25), and SROC 0.88 (95% CI 0.56-0.98; Table 2 and Figures 3 and 4).

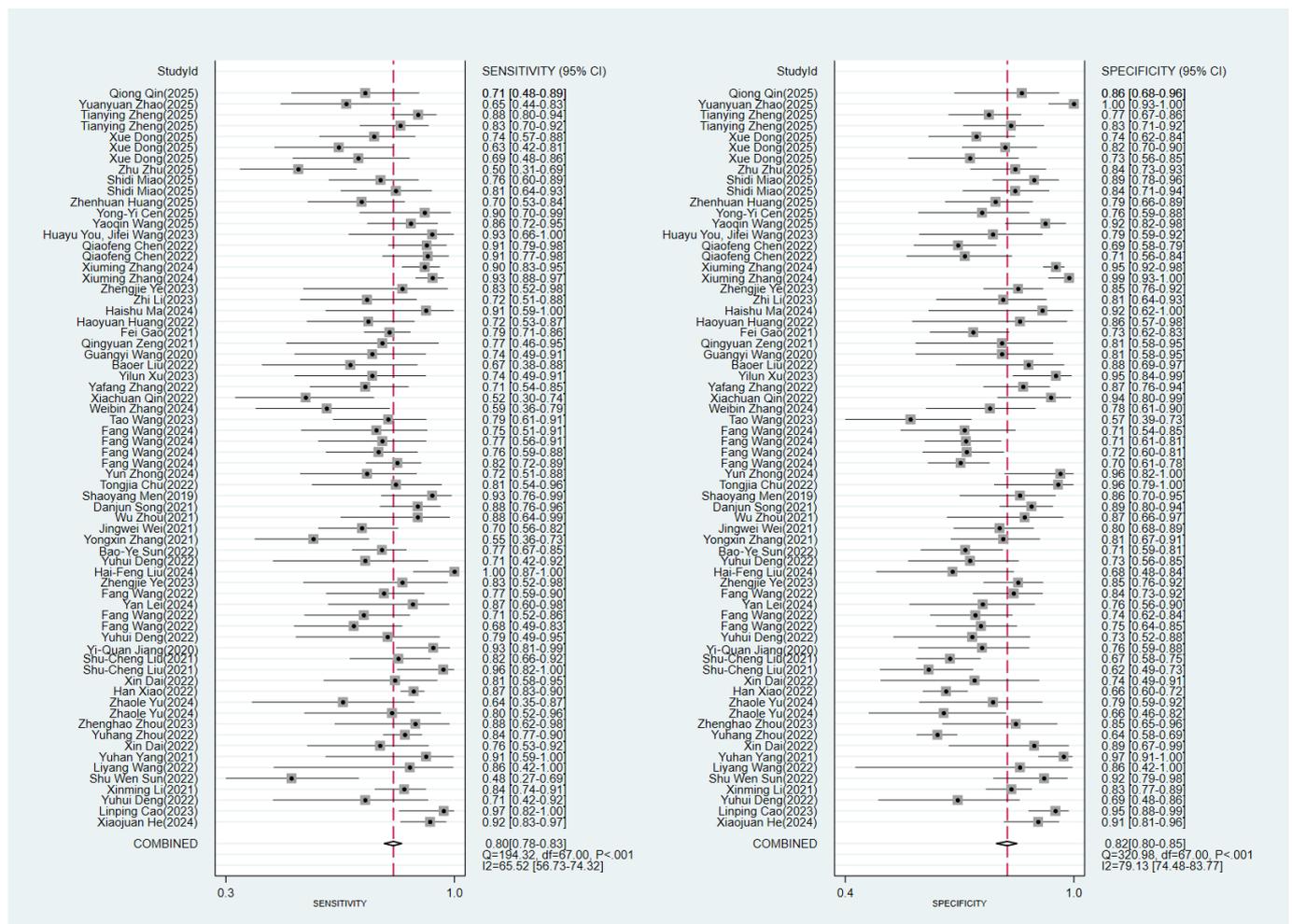
**Table 2.** Meta-analysis results of deep learning for microvascular invasion diagnosis under different image sources and validation set generation methods.

Subgroup	N	SENS <sup>a</sup> (95% CI)	SPEC <sup>b</sup> (95% CI)	PLR <sup>c</sup> (95% CI)	NLR <sup>d</sup> (95% CI)	DOR <sup>e</sup> (95% CI)	SROC <sup>f</sup> (95% CI)	Deeks	Doi
Overall	68	0.80 (0.78-0.83)	0.82 (0.80-0.85)	4.6 (3.9-5.3)	0.24 (0.21-0.27)	19 (15-25)	0.88 (0.56-0.98)	0.77	— <sup>g</sup>
Validation set generation methods									
Internal validation	49	0.82 (0.79-0.85)	0.83 (0.80-0.86)	4.9 (4.1-6.0)	0.22 (0.18-0.26)	23 (17-32)	0.90 (1.00-0.00)	0.85	—
External validation	19	0.77 (0.72-0.82)	0.80 (0.74-0.85)	3.9 (3.0-5.0)	0.29 (0.23-0.36)	13 (9-20)	0.85 (1.00-0.00)	0.22	—
The number of cases in the validation set									
<100	38	0.79 (0.75-0.83)	0.83 (0.79-0.87)	4.7 (3.8-5.8)	0.25 (0.20-0.30)	19 (14-26)	0.88 (0.63-0.97)	0.15	—
≥100	30	0.82 (0.78-0.85)	0.81 (0.77-0.85)	4.3 (3.4-5.5)	0.23 (0.18-0.28)	19 (13-29)	0.88 (0.66-0.97)	0.5	—
Image sources									
CECT <sup>h</sup>	20	0.84 (0.79-0.88)	0.83 (0.77-0.88)	5.0 (3.6-6.9)	0.19 (0.15-0.25)	26 (16-42)	0.90 (1.00-0.00)	0.09	—
CECT internal validation	16	0.86 (0.80-0.90)	0.85 (0.78-0.90)	5.7 (3.8-8.4)	0.17 (0.12-0.24)	33 (19-59)	0.92 (1.00-0.00)	0.19	—
CECT external validation	4	0.82 (0.73-0.88)	0.76 (0.64-0.85)	3.4 (2.3-5.0)	0.24 (0.17-0.34)	14 (9-23)	0.86 (1.00-0.00)	—	-1.14
CEUS <sup>i</sup>	5	0.70 (0.58-0.80)	0.88 (0.82-0.92)	5.6 (3.5-9.0)	0.34 (0.23-0.51)	17 (7-37)	0.89 (1.00-0.00)	—	0.46
CEMRI <sup>j</sup>	24	0.78 (0.73-0.83)	0.81 (0.76-0.85)	4.0 (3.2-5.0)	0.27 (0.22-0.33)	15 (10-22)	0.86 (1.00-0.00)	0.15	—

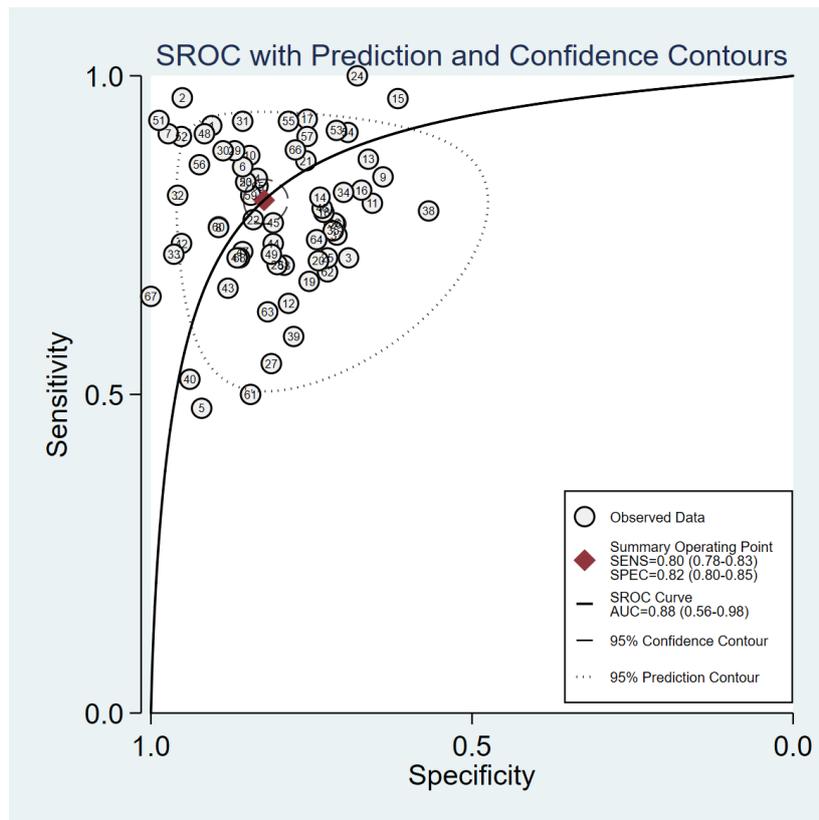
Subgroup	N	SENS <sup>a</sup> (95% CI)	SPEC <sup>b</sup> (95% CI)	PLR <sup>c</sup> (95% CI)	NLR <sup>d</sup> (95% CI)	DOR <sup>e</sup> (95% CI)	SROC <sup>f</sup> (95% CI)	Deeks	Doi
CEMRI internal validation	16	0.81 (0.75-0.86)	0.81 (0.75-0.85)	4.3 (3.2-5.6)	0.24 (0.17-0.32)	18 (11-30)	0.88 (1.00-0.00)	0.17	—
CEMRI external validation	8	0.74 (0.65-0.82)	0.80 (0.71-0.87)	3.7 (2.7-5.2)	0.32 (0.24-0.43)	12 (7-18)	0.84 (1.00-0.00)	—	1.74
MRI <sup>k</sup>	5	0.76 (0.68-0.82)	0.80 (0.70-0.87)	3.7 (2.5-5.5)	0.30 (0.23-0.40)	12 (7-21)	0.84 (1.00-0.00)	—	3.11
Multimodal	10	0.74 (0.68-0.80)	0.79 (0.75-0.83)	3.6 (2.9-4.4)	0.33 (0.26-0.41)	11 (7-16)	0.82 (1.00-0.00)	0.44	—
Multimodal internal validation	6	0.79 (0.69-0.87)	0.79 (0.73-0.85)	3.8 (2.8-5.3)	0.26 (0.17-0.41)	15 (7-29)	0.86 (1.00-0.00)	—	1.6
Multimodal external validation	4	0.71 (0.64-0.78)	0.78 (0.73-0.83)	3.3 (2.6-4.2)	0.36 (0.28-0.47)	9 (6-14)	0.82 (1.00-0.00)	—	1.93
Pathological Sections	4	0.91 (0.87-0.94)	0.90 (0.68-0.97)	9.2 (2.5-33.6)	0.09 (0.06-0.15)	97 (20-465)	0.92 (1.00-0.00)	—	3.08

- <sup>a</sup>SENS: sensitivity.
- <sup>b</sup>SPEC: specificity.
- <sup>c</sup>PLR: positive likelihood ratio.
- <sup>d</sup>NLR: negative likelihood ratio.
- <sup>e</sup>DOR: diagnostic odds ratio.
- <sup>f</sup>SROC: summary receiver operating characteristic.
- <sup>g</sup>Not applicable.
- <sup>h</sup>CECT: contrast-enhanced computed tomography.
- <sup>i</sup>CEUS: contrast-enhanced ultrasound.
- <sup>j</sup>CEMRI: contrast-enhanced magnetic resonance imaging.
- <sup>k</sup>MRI: magnetic resonance imaging.

**Figure 3.** Meta-analysis forest plots: specificity and sensitivity of image-based deep learning in microvascular invasion diagnosis [14,16,21-70].



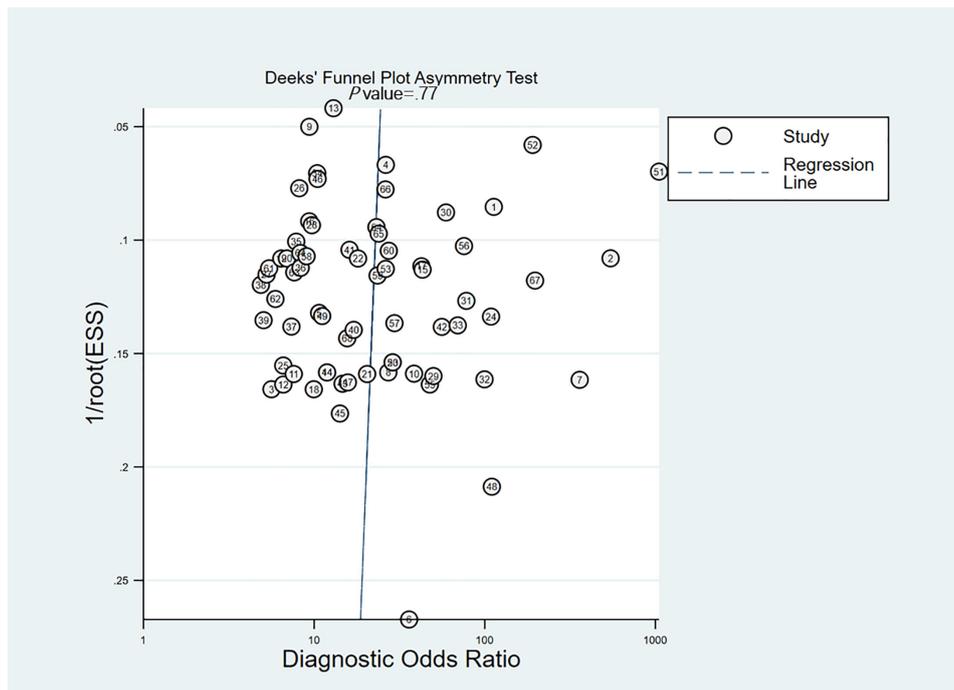
**Figure 4.** Meta-analysis SROC: specificity and sensitivity of image-based deep learning in microvascular invasion diagnosis. SENS: sensitivity; SPEC: specificity; SROC: summary receiver operating characteristic.



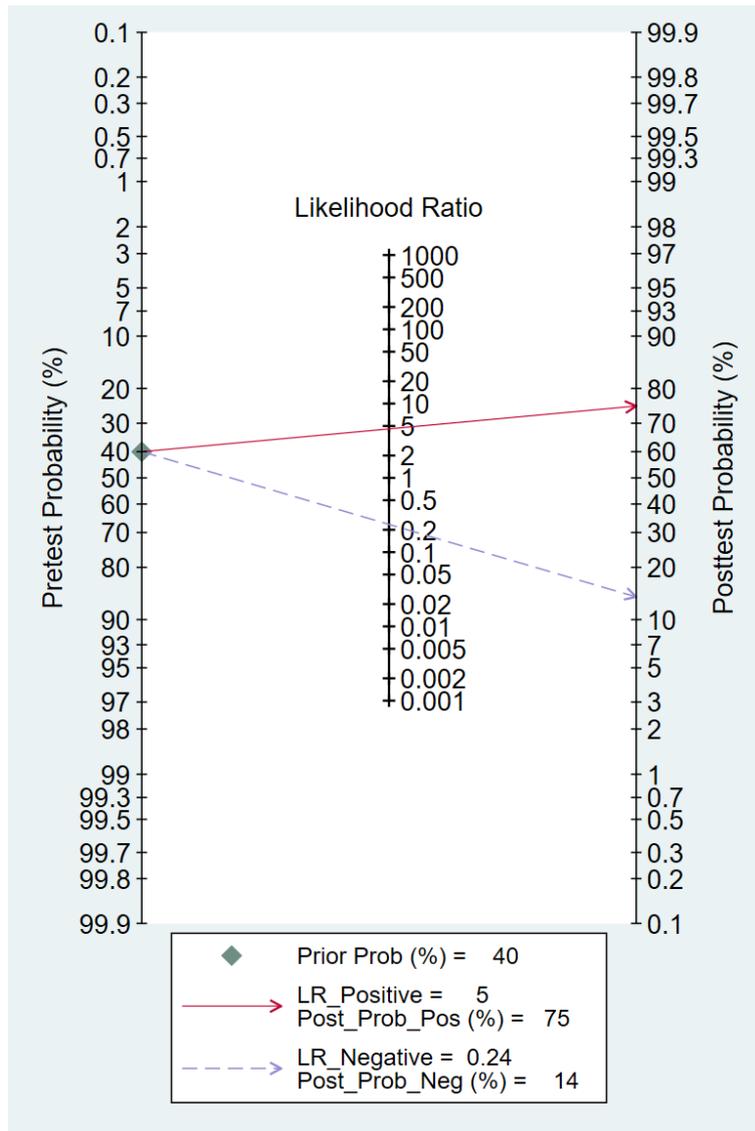
Deeks' funnel plot indicated no significant small-study effects ( $P=.77$ , Table 2 and Figure 5). When the prior probability of MVI was set to 40%, the posterior probabilities corresponding to positive and negative DL model detection results were 75% and 14%, respectively. Fagan's nomogram analysis

showed that a positive detection result increased the posterior probability by 35% compared with the prior probability, whereas a negative detection result decreased it by 26% (Figure 6).

**Figure 5.** Deeks' funnel plot from meta-analysis of the specificity and sensitivity of image-based deep learning in microvascular invasion diagnosis.



**Figure 6.** Fagan’s nomogram from meta-analysis of the specificity and sensitivity of image-based deep learning in microvascular invasion diagnosis.



### Subgroup Analysis by Validation Set Generation Method

#### Internal Validation

The model’s accuracy was validated using 49 diagnostic fourfold tables in internal validation. The Spearman correlation coefficient was 0.04, suggesting a minimal threshold effect. This effect accounted for none of the observed between-study heterogeneity. The pooled analysis yielded the following results: sensitivity 0.82 (95% CI 0.79-0.85,  $I^2=61.27\%$ ), specificity 0.83 (95% CI 0.80-0.86,  $I^2=72.22\%$ ), LR+4.9 (95% CI 4.1-6.0), LR-0.22 (95% CI 0.18-0.26), DOR 23 (95% CI 17-32), and SROC 0.90 (95% CI 1.00-0.00; [Table 2](#) and [Figures S1A,B in Multimedia Appendix 2](#)).

Deeks’ funnel plot illustrated no significant small-study effects ( $P=.85$ , [Table 2](#) and [Figure S1C in Multimedia Appendix 2](#)). When the prior probability of MVI was set to 40%, the corresponding posterior probabilities for positive and negative DL model detection results were 77% and 13%, respectively. Fagan’s nomogram analysis showed that

a positive detection result increased the posterior probability by 37%, compared with the prior probability; a negative detection result decreased the posterior probability by 27% ([Figure S1D in Multimedia Appendix 2](#)).

#### External Validation

The model’s accuracy was validated using 19 diagnostic fourfold tables in external validation. The Spearman correlation coefficient was -0.15, indicating a minimal threshold effect. This effect accounted for 2% of the observed between-study heterogeneity. The pooled analysis yielded the following results: sensitivity 0.77 (95% CI 0.72-0.82,  $I^2=73.40\%$ ), specificity 0.80 (95% CI 0.74-0.85,  $I^2=83.19\%$ ), LR+3.9 (95% CI 3.0-5.0), LR-0.29 (95% CI 0.23-0.36), DOR 13 (95% CI 9-20), and SROC 0.85 (95% CI 1.00-0.00; [Table 2](#) and [Figures S2A,B in Multimedia Appendix 3](#)).

Deeks’ funnel plot revealed no significant small-study effects ( $P=.22$ , [Table 2](#) and [Figure S2C in Multimedia Appendix 3](#)). When the prior probability of MVI was set to 39%, the corresponding posterior probabilities for positive and negative DL model detection results were 71% and

15%, respectively. Fagan's nomogram analysis showed that a positive detection result increased the posterior probability by 32% compared with the prior probability, whereas a negative detection result decreased it by 24% (Figure S2D in [Multimedia Appendix 3](#)).

### Validation Set Size < 100 Cases

Among studies with a validation set size <100 cases, the model's accuracy was validated using 38 diagnostic fourfold tables. The Spearman correlation coefficient was  $-0.44$ , suggesting a minimal threshold effect. This effect accounted for 19% of the observed between-study heterogeneity. The pooled analysis yielded the following results: sensitivity 0.79 (95% CI 0.75-0.83,  $I^2=56.49\%$ ), specificity 0.83 (95% CI 0.79-0.87,  $I^2=67.03\%$ ), LR+4.7 (95% CI 3.8-5.8), LR-0.25 (95% CI 0.20-0.30), DOR 19 (95% CI 14-26), and SROC 0.88 (95% CI 0.63-0.97; [Table 2](#) and [Figure S3A,B in Multimedia Appendix 4](#)).

Deeks' funnel plot revealed no significant small-study effects ( $P=.15$ , [Table 2](#) and [Figure S3C in Multimedia Appendix 4](#)). When the prior probability of MVI was set to 41%, the corresponding posterior probabilities for positive and negative DL model detection results were 77% and 15%, respectively. Fagan's nomogram analysis displayed that a positive detection result increased the posterior probability by 36% compared with the prior probability, whereas a negative detection result decreased it by 26% (Figure S3D in [Multimedia Appendix 4](#)).

### Validation Set Size $\geq$ 100 Cases

Among studies with a validation set size  $\geq$ 100 cases, the model's accuracy was validated using 30 diagnostic fourfold tables. The Spearman correlation coefficient was 0.47, indicating a minimal threshold effect. This effect accounted for 22% of the observed between-study heterogeneity. The pooled analysis yielded the following results: sensitivity 0.82 (95% CI 0.78-0.85,  $I^2=69.35\%$ ), specificity 0.81 (95% CI 0.77-0.85,  $I^2=85.58\%$ ), LR+4.3 (95% CI 3.4-5.5), LR-0.23 (95% CI 0.18-0.28), DOR 19 (95% CI 13-29), and SROC 0.88 (95% CI 0.66-0.97; [Table 2](#) and [Figure S4A,B in Multimedia Appendix 5](#)).

Deeks' funnel plot revealed no significant small-study effects ( $P=.50$ ; [Table 2](#) and [Figure S4C in Multimedia Appendix 5](#)). When the prior probability of MVI was set to 39%, the corresponding posterior probabilities for positive and negative DL model detection results were 73% and 13%, respectively. Fagan's nomogram analysis showed that a positive detection result increased the posterior probability by 34% compared with the prior probability, whereas a negative detection result decreased it by 26% (Figure S4D in [Multimedia Appendix 5](#)).

## Subgroup Analysis by Image Source

### DL Based on CECT

In CECT, the model's accuracy was validated using 20 diagnostic fourfold tables. The Spearman correlation

coefficient was  $-0.28$ , suggesting a minimal threshold effect. This effect accounted for 8% of the observed between-study heterogeneity. The pooled analysis revealed the following results: sensitivity 0.84 (95% CI 0.79-0.88,  $I^2=71.67\%$ ), specificity 0.83 (95% CI 0.77-0.88,  $I^2=87.84\%$ ), LR+5.0 (95% CI 3.6-6.9), LR- 0.19 (95% CI 0.15-0.25), DOR 26 (95% CI 16-42), and SROC 0.90 (95% CI 1.00-0.00; [Table 2](#) and [Figure S5 in Multimedia Appendix 6](#)). Deeks funnel plot revealed no significant small-study effects ( $P=.09$ , [Table 2](#)).

Among CECT-based models with internal validation, the accuracy was validated using 16 diagnostic fourfold tables. The Spearman correlation coefficient was  $-0.26$ , indicating a minimal threshold effect. This effect accounted for 7% of the observed between-study heterogeneity. The pooled analysis yielded the following results: sensitivity 0.86 (95% CI 0.80-0.90,  $I^2=71.85\%$ ), specificity 0.85 (95% CI 0.78-0.90,  $I^2=88.74\%$ ), LR+5.7 (95% CI 3.8-8.4), LR-0.17 (95% CI 0.12-0.24), DOR 33 (95% CI 19-59), and SROC 0.92 (95% CI 1.00-0.00; [Table 2](#) and [Multimedia Appendix 7](#)). Deeks' funnel plot revealed no significant small-study effects ( $P=.19$ , [Table 2](#)).

Among CECT-based models with external validation, the accuracy was validated using four diagnostic fourfold tables. The pooled analysis yielded the following results: sensitivity 0.82 (95% CI 0.73-0.88,  $I^2=62.47\%$ ), specificity 0.76 (95% CI 0.64-0.85,  $I^2=77.88\%$ ), LR+3.4 (95% CI 2.3-5.0), LR- 0.24 (95% CI 0.17-0.34), DOR 14 (95% CI 9-23), and SROC 0.86 (95% CI 1.00-0.00; [Table 2](#) and [Multimedia Appendix 7](#)). Subsequent Doi plot analysis showed moderate publication bias among the included studies (LFK index= $-1.14$ , [Table 2](#)).

### DL Based on CEUS

In CEUS, the model's accuracy was validated using 5 diagnostic fourfold tables. The pooled analysis yielded the following results: sensitivity 0.70 (95% CI 0.58-0.80,  $I^2=58.75\%$ ), specificity 0.88 (95% CI 0.82-0.92,  $I^2=30.18\%$ ), LR+5.6 (95% CI 3.5-9.0), LR-0.34 (95% CI 0.23-0.51), DOR 17 (95% CI 7-37), and SROC 0.89 (95% CI 1.00-0.00; [Table 2](#) and [Multimedia Appendix 8](#)). Analysis using the Doi plot revealed minimal publication bias among the included studies (LFK index=0.46, [Table 2](#)).

### DL Based on CEMRI

In CEMRI, the model's accuracy was validated using 24 diagnostic fourfold tables. The Spearman correlation coefficient was  $-0.22$ , suggesting a minimal threshold effect. This effect accounted for 5% of the observed between-study heterogeneity. The pooled analysis yielded the following results: sensitivity 0.78 (95% CI 0.73-0.83,  $I^2=61.70\%$ ), specificity 0.81 (95% CI 0.76-0.85,  $I^2=65.14$ ), LR+4.0 (95% CI 3.2-5.0), LR- 0.27 (95% CI 0.22-0.33), DOR 15 (95% CI 10-22), and SROC 0.86 (95% CI 1.00-0.00; [Table 2](#) and [Multimedia Appendix 9](#)). Deeks funnel plot revealed no significant small-study effects ( $P=0.15$ , [Table 2](#)).

Among CEMRI models with internal validation, the accuracy was validated using 16 diagnostic fourfold tables.

The Spearman correlation coefficient was 0.17, indicating a minimal threshold effect. This effect accounted for 3% of the observed between-study heterogeneity. The pooled analysis yielded the following results: sensitivity 0.81 (95% CI 0.75-0.86,  $I^2=57.28\%$ ), specificity 0.81 (95% CI 0.75-0.85,  $I^2=61.32\%$ ), LR+4.3 (95% CI 3.2-5.6), LR- 0.24 (95% CI 0.17-0.32), DOR 18 (95% CI 11-30), and SROC 0.88 (95% CI 1.00-0.00; [Table 2](#) and [Figure S9 in Multimedia Appendix 10](#)). Deeks funnel plot revealed no significant small-study effects ( $P=.17$ , [Table 2](#)).

Among CEMRI models with external validation, the accuracy was validated using 8 diagnostic fourfold tables. The Spearman correlation coefficient was  $-0.72$ , suggesting a significant threshold effect. This effect accounted for 52% of the observed between-study heterogeneity. The pooled analysis yielded the following results: sensitivity 0.74 (95% CI 0.65-0.82,  $I^2=70.90\%$ ), specificity 0.80 (95% CI 0.71-0.87,  $I^2=72.55\%$ ), LR+3.7 (95% CI 2.7-5.2), LR-0.32 (95% CI 0.24-0.43), DOR 12 (95% CI 7-18), and SROC 0.84 (95% CI 1.00-0.00; [Table 2](#) and [Figure S9 in Multimedia Appendix 10](#)). Further Doi plot analysis revealed moderate publication bias among the included studies (LFK index=1.74, [Table 2](#)).

### DL Based on MRI

In MRI, the model's accuracy was validated using 5 diagnostic fourfold tables. The pooled analysis revealed the following results: sensitivity 0.76 (95% CI 0.68-0.82,  $I^2=0.00\%$ ), specificity 0.80 (95% CI 0.70-0.87,  $I^2=0.00\%$ ), LR+3.7 (95% CI 2.5-5.5), LR-0.30 (95% CI 0.23-0.40), DOR 12 (95% CI 7-21), and SROC, 0.84 (95% CI 1.00-0.00; [Table 2](#) and [Multimedia Appendix 11](#)). Analysis using the Doi plot revealed substantial publication bias among the included studies (LFK index=3.11, [Table 2](#)).

### DL Based on Multimodal Imaging

In multimodal medical images, the model's accuracy was validated using 10 diagnostic fourfold tables. The pooled analysis yielded the following results: sensitivity 0.74 (95% CI 0.68-0.80,  $I^2=0.00\%$ ), specificity 0.79 (95% CI 0.75-0.83,  $I^2=11.32\%$ ), LR+3.6 (95% CI 2.9-4.4), LR- 0.33 (95% CI 0.26-0.41), DOR 11 (95% CI 7-16), and SROC 0.82 (95% CI 1.00-0.00; [Table 2](#) and [Multimedia Appendix 12](#)). Deeks' funnel plot revealed no significant small-study effects ( $P=.44$ , [Table 2](#)).

In multimodal internal validation, the model's accuracy was validated using 6 diagnostic fourfold tables. The pooled analysis revealed the following results: sensitivity 0.79 (95% CI 0.69-0.87,  $I^2=0.00\%$ ), specificity 0.79 (95% CI 0.73-0.85,  $I^2=31.69\%$ ), LR+3.8 (95% CI 2.8-5.3), LR- 0.26 (95% CI 0.17-0.41), DOR 15 (95% CI 7-29), and SROC 0.86 (95% CI 1.00-0.00; [Table 2](#) and [Multimedia Appendix 13](#)). Analysis using the Doi plot revealed moderate publication bias among the included studies (LFK index=1.6; [Table 2](#)).

In multimodal external validation, the model's accuracy was validated using 4 diagnostic fourfold tables. The pooled analysis showed the following results: sensitivity 0.71 (95% CI 0.64-0.78,  $I^2=0.00\%$ ), specificity 0.78 (95% CI 0.73-0.83,

$I^2=0.00\%$ ), LR+3.3 (95% CI 2.6-4.2), LR- 0.36 (95% CI 0.28-0.47), DOR 9 (95% CI 6-14), and SROC 0.82 (95% CI 1.00-0.00; [Table 2](#) and [Multimedia Appendix 13](#)). Subsequent Doi plot analysis indicated moderate publication bias (LFK index=1.93; [Table 2](#)).

### DL Based on Pathological Sections

In pathological sections, the model's accuracy was validated using 4 diagnostic fourfold tables. The pooled analysis revealed the following results: sensitivity 0.91 (95% CI 0.87-0.94,  $I^2=0.00\%$ ), specificity 0.90 (95% CI 0.68-0.97,  $I^2=94.83\%$ ), LR+9.2 (95% CI 2.5-33.6), LR- 0.09 (95% CI 0.06-0.15), DOR 97 (95% CI 20-465), and SROC 0.92 (95% CI 1.00-0.00; [Table 2](#) and [Multimedia Appendix 14](#)). Further Doi plot analysis revealed substantial publication bias (LFK index=3.08; [Table 2](#)).

## Discussion

### Summary of the Main Findings

The current meta-analysis revealed that the modeling methods used for DL-based prediction of HCC MVI primarily used CECT, CEUS, CEMRI, MRI, multimodal imaging, and pathological image techniques. DL models based on medical imaging showed favorable overall diagnostic performance in predicting HCC MVI, with a pooled sensitivity of 0.80 (95% CI 0.78-0.83) and specificity of 0.82 (95% CI 0.80-0.85). Further analysis across imaging modalities revealed that CECT-based models achieved the highest diagnostic efficacy, showing a sensitivity of 0.84 (95% CI 0.79-0.88) and specificity of 0.83 (95% CI 0.77-0.88). Models based on CEUS exhibited particularly high specificity (0.88, 95% CI 0.82-0.92). Furthermore, models using pathological slides, considered the diagnostic reference standard, attained the highest overall performance, with a sensitivity of 0.91 (95% CI 0.87-0.94) and specificity of 0.90 (95% CI 0.68-0.97). Therefore, these models appear promising as a diagnostic approach for MVI in HCC.

### Comparison With Other Previous Reviews

This study also noted that some researchers have discussed the use of machine learning for MVI in HCC. Xiao et al [71] and Liang et al [72], who focused on MRI and ultrasound radiomics, respectively, validated the predictive potential of single modalities, with pooled area under the curves (AUCs) of 0.87 and 0.81, respectively. However, their analyses included a limited number of studies and focused only on a single imaging modality, limiting the generalizability of their conclusions. Li et al [73] integrated multiple imaging modalities across 22 studies (involving 4129 participants), reporting a pooled AUC of 0.90 for radiomic models. However, their analysis primarily incorporated traditional machine learning models and did not stratify performance based on the validation set generation method. This limitation may lead to an overly optimistic assessment of model generalizability.

Unlike previous systematic reviews that focused on traditional machine learning models, our meta-analysis focuses on the value of DL algorithms for diagnosing HCC MVI. Methodologically, this study provides a deeper exploration of the sources of performance heterogeneity through subgroup analyses based on the validation set generation method and image source. A particular methodological strength is the inclusion of pathological sections, the diagnostic gold standard, as a benchmark for performance optimization. This meta-analysis synthesized data from 52 studies involving 19,531 patients and provided more robust and reliable conclusions than those from analyses with smaller sample sizes.

### ***Influence of Imaging Modalities on DL***

The imaging modalities used to construct DL models can be categorized into 2 main types: noninvasive and invasive. This study showed that the properties of different modalities and their ability to extract biological features directly affected the diagnostic efficacy of corresponding models.

Among noninvasive imaging modalities, CECT can effectively capture tumor heterogeneity, enhancement patterns, and peritumoral microenvironment changes due to its high spatial resolution and multi-phase dynamic imaging capabilities. These features support the superior diagnostic performance of DL models (AUC=0.90) [74,75]. Models developed from CEUS achieved commendable performance (AUC=0.89) and exhibited a high pooled specificity (0.88). This finding suggests that CEUS's real-time hemodynamic properties may be valuable in ruling out MVI-negative cases. However, its comparatively lower sensitivity (0.70) concurrently suggests ongoing challenges in consistently identifying MVI-positive features. Integrating Sonazoid-based functional imaging with conventional ultrasound characteristics and serum markers in the future may be a promising way to improve performance [76]. CEMRI has unique advantages in depicting tumor boundaries and detecting subtle changes in the peritumoral liver parenchyma due to its excellent soft tissue contrast [77]. The model's performance (sensitivity: 0.78, specificity: 0.81, AUC=0.86) surpasses the assessment that relies on traditional visual features, as reported by Wu et al [78] (sensitivity: 0.55, specificity: 0.87, AUC=0.80). However, its widespread adoption is limited by long examination times, high costs, and dependence on equipment. Nonenhanced MRI models also demonstrate predictive potential (sensitivity: 0.76, specificity: 0.80, AUC=0.84). Notably, multimodal fusion models, which are designed to integrate complementary information, have not demonstrated significant advantages in the studies included (AUC=0.82). This may be due to the simplified fusion strategies or data heterogeneity.

Among invasive imaging modalities, DL models based on pathological sections demonstrated the highest diagnostic efficacy in this study (AUC=0.92). This highlights the gold-standard status of pathology in evaluating MVI. However, the inherent invasiveness of pathological examination precludes its use in preoperative decision-making. Therefore, a central challenge for future research is to

effectively transfer and integrate the gold-standard-level diagnostic insights from pathological sections into preoperative, noninvasive imaging. This objective requires leveraging advanced methodologies, such as transfer learning, cross-modal fusion, and generative models. Augmenting existing advantageous modalities, such as CECT, with these advancements could ultimately pave the way for a clinically viable preoperative "virtual biopsy."

### ***Image Segmentation***

Accurate image segmentation is essential for building reliable DL models. However, the literature incorporated in this meta-analysis primarily relies on manual segmentation. This method can introduce subjective bias (38/52, 73.1%). While advanced network architectures have achieved expert-level precision in segmenting normal liver tissue, as demonstrated by Dice coefficients ranging from 0.968 to 0.982 [79, 80], comparative data from the Liver Tumor Segmentation challenge reveal significant discrepancies. Specifically, the accuracy of liver tumor segmentation (Dice: 0.739) is considerably lower than that of liver parenchyma segmentation (Dice: 0.963) [81]. This discrepancy primarily stems from the heterogeneity of HCC lesions, suboptimal image contrast, and a lack of high-quality annotated data. Concurrently, selecting the appropriate segmentation strategy is critical. While 3D segmentation comprehensively captures spatial heterogeneity, the more clinically feasible 2D approach sacrifices substantial volumetric information [82]. Given the demand for submillimeter-level precision in MVI prediction, there are 2 key technological pathways for enhancing model stability. First, exploring segmentation paradigms that require fewer annotations, such as weakly or self-supervised learning, can reduce dependency on annotations. Second, developing novel network architectures designed specifically to address HCC heterogeneity and boundary ambiguity is equally crucial.

### ***Validation Set Generation Method***

The rigor of validation strategies is paramount for evaluating the real-world generalizability of DL models in predicting MVI of HCC [83,84]. The present analysis reveals that, despite exemplary performance during internal validation, consistent performance declines emerge in independent external validation cohorts. This finding clearly shows that relying too much on internal validation can lead to overestimating a model's true efficacy [85]. Concern regarding validation strategies is not unique to DL research. A recent meta-analysis [86] that focused on MRI-based radiomics for predicting HCC recurrence and MVI similarly concluded that the current predominance of internal validation results in an overestimation of model generalizability as well. Consequently, any proclaimed superior performance may substantially diminish when confronting real-world heterogeneity, if the evaluation framework remains confined to internal validation, irrespective of the underlying algorithm—be it DL or radiomics. Therefore, promoting rigorous external validation and establishing standardized, cross-institutional imaging protocols are essential steps toward reliable clinical translation in this field.

## Heterogeneity Analysis

There is substantial heterogeneity among the included studies. Notably, high levels of heterogeneity persist within subgroups, even after stratifying analyses by imaging modality and validation strategy. This observation objectively reflects the inherent complexity of artificial intelligence (AI)-based medical imaging research and is a common challenge for meta-analyses in this field. The heterogeneity primarily stems from 3 levels. Technically, variations in critical parameters, including imaging equipment, magnetic field strength, and slice thickness, directly influence image texture and quality. These variations are a significant technical source of variability in model performance. Methodologically, diversity in study design, validation strategies, segmentation techniques, and DL network architectures introduces additional variation in model construction and performance interpretation. Clinically, differences in patient populations regarding geographic distribution, underlying liver disease etiology, and disease stage may also affect model performance and generalizability. While this heterogeneity limits the direct interpretability of the pooled results to some extent, it accurately reflects the diversity in methodology and clinical practice within the field. Future investigations should adhere to the Findable, Accessible, Interoperable, and Reusable principles, providing detailed reporting of imaging acquisition parameters, model architectures, and training specifics. Such comprehensive reporting will facilitate the in-depth exploration of heterogeneity sources via methods such as meta-regression. This will promote the identification of key influencing factors and the standardization of methodologies.

## Methods for ROB Assessment

This systematic review used a composite strategy to assess the ROB. First, all included studies were rigorously evaluated according to the QUADAS-2 guidelines. The results showed that all primary studies were rated as having a high ROB in the "Patient Selection" domain of QUADAS-2 due to the widespread use of the retrospective case-control design. While this outcome aligns with the QUADAS-2 assessment principles, it also reveals a limitation of the tool when evaluating machine learning-based diagnostic studies that use retrospective data. The tool struggles to differentiate nuances in data construction quality among studies.

To conduct a more granular assessment of data-level bias risk, a supplemental analysis was performed using items from the Quality Assessment of Diagnostic Accuracy Studies for Artificial Intelligence (QUADAS-AI) tool targeting "Study Participant Selection." The QUADAS-AI tool provides more detailed criteria for this dimension, including an explicit description of data source, size, and quality characteristics, use of open-source datasets, a clear rationale for splitting data into training, validation, and test sets, performance of image preprocessing, and provision of scanner model information. The analysis using QUADAS-AI items showed that all 52 studies appropriately described the source, size, and quality of the input data and clearly defined the patient inclusion criteria. Among these studies, only one used an open-source dataset. All studies provided a rationale for the data split.

Image preprocessing was performed in all studies. However, 11 studies did not report the scanner model used for image acquisition (Table S2 in [Multimedia Appendix 15](#) [14,16,21-70]).

## Advantages and Limitations

A primary strength of this research is that it is the first large-scale, systematic meta-analysis to evaluate medical imaging-based DL models for predicting MVI in HCC. The analysis included 52 studies with 19,531 patients, providing the field with comprehensive evidence. The analytical process strictly adhered to the PRISMA guidelines, and bias risk was evaluated using QUADAS-2. These measures ensured methodological rigor and transparency. However, several limitations warrant consideration. First, the training cohorts in most primary studies were small. Only 3 studies had a sample size greater than 1000 patients. The robustness of DL models depends heavily on large volumes of high-quality data. Therefore, restricted training sample sizes are a potential methodological limitation. This likely contributes to overfitting in some models, which is probably an internal reason for the observed performance degradation during external validation. This finding underscores the fundamental importance of acquiring large-scale, high-quality datasets to enhance model generalizability in the development of current DL models. Second, the 95% CIs for the pooled AUCs and for most subgroup analyses were exceptionally wide. This constrained the interpretation of the result precision to some extent. This primarily stems from the significant heterogeneity among the included studies. The limited number of studies in subgroup analyses exacerbated data sparsity. Third, Doi plot analyses for some subgroups indicated moderate to substantial publication bias. These analyses showed that the pooled results for these subgroups may be influenced by unpublished negative studies, which poses a risk of overestimating diagnostic performance. Fourth, the rigor of the validation strategies needs to be improved. Most studies relied on internal validation. Only a few conducted stringent external validation. This reliance may introduce an optimistic bias into the overall assessment of the models' real-world generalizability. While the overall performance on external validation sets was discussed, the limited quantity of external validation data prevented a more in-depth subgroup analysis of validation strategies by different imaging modalities. Fifth, the vast majority of primary studies inadequately reported model calibration metrics or details about network complexity. This omission hinders a quantitative, systematic evaluation of predictive reliability and overfitting risk at the review level and impacts the comprehensive assessment of clinical applicability. This reflects a general deficiency in the transparency of methodological reporting within the current research landscape.

## Clinical Implications and Future Perspectives

This study indicates that medical imaging-based DL models, particularly those using preoperative CECT, demonstrate promising diagnostic performance in predicting MVI of HCC. These models have the potential to assist in personalized

surgical planning. However, translating them into clinical practice faces multiple challenges. One primary issue is establishing clinical decision thresholds. While this analysis quantified predictive probabilities via Fagan nomograms, there is currently a lack of evidence-based guidelines defining “at what predicted probability of MVI the surgical margin should be adjusted.” Future work must integrate clinical outcome data and empirically explore the net benefit of different thresholds using methods such as decision curve analysis. Second, these models’ generalizability and deployment feasibility need urgent enhancement. Performance degradation during external validation suggests susceptibility to variations in imaging protocols, equipment, and patient populations. Furthermore, model interpretability is crucial for gaining clinical trust. This necessitates developing transparent methods for presenting decision rationale.

To bridge the gap between “high performance” and “high utility,” future efforts must focus on 3 interconnected levels. At the research level, the focus must shift from model construction to rigorous, prospective, multicenter validation to unequivocally assess generalizability. At the algorithmic level, it is essential to explore cross-modal information fusion, especially using transfer learning, to bring “gold standard”-level diagnostic insights from histopathological sections to preoperative, noninvasive imaging. At the clinical level, establishing standardized, cross-institutional

imaging protocols and developing decision support systems that integrate seamlessly into clinical workflows is imperative. This integrated approach is vital for reliably translating technology from innovation to tangible patient benefit.

## Conclusions

This systematic review and meta-analysis demonstrate that medical imaging-based DL models, especially those leveraging preoperative CECT, hold significant promise for the noninvasive preoperative prediction of MVI in HCC. Unlike previous reviews that focused on radiomics or single imaging modalities, this study conducted a comprehensive comparison across multiple modalities. The study also emphasizes the critical role of external validation in the real-world generalizability of a model. However, substantial heterogeneity across studies and the performance degradation observed during independent external validation suggest that their generalizability to the real world must be confirmed through more rigorous study designs. Consequently, future research should prioritize establishing model robustness via prospective, multicenter external validation, coupled with efforts to standardize methodologies and improve reporting transparency. A critical step toward reliable clinical translation and achieving the ultimate goal of a “virtual biopsy” is developing algorithms that can translate pathology-grade diagnostic insights into preoperative, noninvasive imaging.

## Funding

This study was funded by the General Project of the Liaoning Provincial Department of Education (grant number JYTMS20230109) and the Liaoning Provincial Science and Technology Program Joint Plan (Natural Science Foundation General Program) (grant number 2025-MSLH-742). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Authors' Contributions

All authors contributed to the study conception and design. WF and BQ prepared the original draft of the manuscript and contributed to the methodology, formal analysis, and investigation. SH was responsible for conceptualization and funding acquisition and provided supervision. SH also reviewed and edited the manuscript. All authors commented on previous versions of the manuscript and also read and approved the final manuscript.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Flow diagram illustrating the search strategy.

[\[DOCX File \(Microsoft Word File\), 23 KB-Multimedia Appendix 1\]](#)

## Multimedia Appendix 2

(A) Meta-analysis forest plots: specificity and sensitivity of image-based deep learning (DL) in microvascular invasion (MVI) diagnosis in internal validation; (B) Meta-analysis summary receiver operating characteristic: specificity and sensitivity of image-based DL in MVI diagnosis in internal validation; (C) Deeks funnel plot from meta-analysis of specificity and sensitivity of image-based DL in MVI diagnosis in internal validation; (D) Fagan nomogram from meta-analysis of specificity and sensitivity of image-based DL in MVI diagnosis in internal validation.

[\[PNG File \(Portable Network Graphics File\), 607 KB-Multimedia Appendix 2\]](#)

## Multimedia Appendix 3

(A) Meta-analysis forest plots: specificity and sensitivity of image-based deep learning (DL) in microvascular invasion (MVI) diagnosis in external validation; (B) Meta-analysis summary receiver operating characteristic: specificity and sensitivity

of image-based DL in MVI diagnosis in external validation; (C) Deeks funnel plot from meta-analysis of specificity and sensitivity of image-based DL in MVI diagnosis in external validation; (D) Fagan nomogram from meta-analysis of specificity and sensitivity of image-based DL in MVI diagnosis in external validation.

[\[PNG File \(Portable Network Graphics File\), 403 KB-Multimedia Appendix 3\]](#)

---

#### **Multimedia Appendix 4**

(A) Meta-analysis forest plots: specificity and sensitivity of image-based deep learning (DL) in microvascular invasion (MVI) diagnosis in Validation set size of <100 cases; (B) Meta-analysis summary receiver operating characteristic: specificity and sensitivity of image-based DL in MVI diagnosis in Validation set size < 100 cases; (C) Deeks funnel plot from meta-analysis of specificity and sensitivity of image-based DL in MVI diagnosis in the validation set size of <100 cases; (D) Fagan nomogram from meta-analysis of specificity and sensitivity of image-based DL in MVI diagnosis in the validation set size of <100 cases.

[\[PNG File \(Portable Network Graphics File\), 519 KB-Multimedia Appendix 4\]](#)

---

#### **Multimedia Appendix 5**

(A) Meta-analysis forest plots: specificity and sensitivity of image-based deep learning (DL) in microvascular invasion (MVI) diagnosis in Validation set size of  $\geq 100$  cases; (B) Meta-analysis summary receiver operating characteristic: specificity and sensitivity of image-based DL in MVI diagnosis in Validation set size of  $\geq 100$  cases; (C) Deeks funnel plot from meta-analysis of specificity and sensitivity of image-based DL in MVI diagnosis in the validation set size  $\geq 100$  cases; (D) Fagan nomogram from meta-analysis of specificity and sensitivity of image-based DL in MVI diagnosis in Validation set size of  $\geq 100$  cases.

[\[PNG File \(Portable Network Graphics File\), 529 KB-Multimedia Appendix 5\]](#)

---

#### **Multimedia Appendix 6**

(A) Meta-analysis forest plots: specificity and sensitivity of contrast-enhanced computed tomography (CECT)-based deep learning (DL) in microvascular invasion (MVI) diagnosis; (B) Meta-analysis SROC: specificity and sensitivity of CECT-based DL in MVI diagnosis.

[\[PNG File \(Portable Network Graphics File\), 320 KB-Multimedia Appendix 6\]](#)

---

#### **Multimedia Appendix 7**

(A) Meta-analysis forest plots: specificity and sensitivity of contrast-enhanced computed tomography (CECT)-based deep learning (DL) in microvascular invasion (MVI) diagnosis in internal validation; (B) Meta-analysis SROC: specificity and sensitivity of CECT-based DL in MVI diagnosis in internal validation; (C) Meta-analysis forest plots: specificity and sensitivity of CECT-based DL in MVI diagnosis in external validation; (D) Meta-analysis summary receiver operating characteristic: specificity and sensitivity of CECT-based DL in MVI diagnosis in external validation.

[\[PNG File \(Portable Network Graphics File\), 395 KB-Multimedia Appendix 7\]](#)

---

#### **Multimedia Appendix 8**

(A) Meta-analysis forest plots: specificity and sensitivity of contrast-enhanced ultrasound (CEUS)-based deep learning (DL) in microvascular invasion (MVI) diagnosis; (B) Meta-analysis summary receiver operating characteristic: specificity and sensitivity of CEUS-based DL in MVI diagnosis.

[\[PNG File \(Portable Network Graphics File\), 173 KB-Multimedia Appendix 8\]](#)

---

#### **Multimedia Appendix 9**

(A) Meta-analysis forest plots: specificity and sensitivity of contrast-enhanced magnetic resonance imaging (CEMRI)-based deep learning (DL) in microvascular invasion (MVI) diagnosis; (B) Meta-analysis summary receiver operating characteristic: specificity and sensitivity of CEMRI-based DL in MVI diagnosis.

[\[PNG File \(Portable Network Graphics File\), 374 KB-Multimedia Appendix 9\]](#)

---

#### **Multimedia Appendix 10**

(A) Meta-analysis forest plots: specificity and sensitivity of contrast-enhanced magnetic resonance imaging (CEMRI)-based deep learning (DL) in microvascular invasion (MVI) diagnosis in internal validation; (B) Meta-analysis summary receiver operating characteristic: specificity and sensitivity of CEMRI-based DL in MVI diagnosis in internal validation; (C) Meta-analysis forest plots: specificity and sensitivity of CEMRI-based DL in MVI diagnosis in external validation; (D) Meta-analysis SROC: specificity and sensitivity of CEMRI-based DL in MVI diagnosis in external validation.

[\[PNG File \(Portable Network Graphics File\), 424 KB-Multimedia Appendix 10\]](#)

---

#### **Multimedia Appendix 11**

(A) Meta-analysis forest plots: specificity and sensitivity of magnetic resonance imaging (MRI)-based deep learning (DL) in microvascular imaging (MVI) diagnosis; (B) Meta-analysis summary receiver operating characteristic: specificity and sensitivity of MRI-based DL in MVI diagnosis.

[\[PNG File \(Portable Network Graphics File\), 171 KB-Multimedia Appendix 11\]](#)

### Multimedia Appendix 12

(A) Meta-analysis forest plots: specificity and sensitivity of multimodal imaging-based deep learning (DL) in microvascular invasion (MVI) diagnosis; (B) Meta-analysis summary receiver operating characteristic: specificity and sensitivity of multimodal imaging-based DL in MVI diagnosis.

[\[PNG File \(Portable Network Graphics File\), 222 KB-Multimedia Appendix 12\]](#)

### Multimedia Appendix 13

(A) Meta-analysis forest plots: specificity and sensitivity of multimodal imaging-based deep learning (DL) in microvascular invasion (MVI) diagnosis in internal validation; (B) Meta-analysis summary receiver operating characteristic (SROC): specificity and sensitivity of multimodal imaging-based DL in MVI diagnosis in internal validation; (C) Meta-analysis forest plots: specificity and sensitivity of multimodal imaging-based DL in MVI diagnosis in external validation; (D) Meta-analysis SROC: specificity and sensitivity of multimodal imaging-based DL in MVI diagnosis in external validation.

[\[PNG File \(Portable Network Graphics File\), 313 KB-Multimedia Appendix 13\]](#)

### Multimedia Appendix 14

(A) Meta-analysis forest plots: specificity and sensitivity of pathological sections-based deep learning (DL) in microvascular invasion (MVI) diagnosis; (B) Meta-analysis summary receiver operating characteristic: specificity and sensitivity of pathological sections-based DL in MVI diagnosis.

[\[PNG File \(Portable Network Graphics File\), 164 KB-Multimedia Appendix 14\]](#)

### Multimedia Appendix 15

Summary of bias risk for each study included in the paper according to the QUADAS-AI (artificial intelligence-specific quality assessment of diagnostic accuracy studies) domains.

[\[PDF File \(Adobe File\), 203 KB-Multimedia Appendix 15\]](#)

### Checklist 1

PRISMA-DTA checklist.

[\[DOCX File \(Microsoft Word File\), 32 KB-Checklist 1\]](#)

### Checklist 2

PRISMA-S checklist.

[\[DOCX File \(Microsoft Word File\), 19 KB-Checklist 2\]](#)

### References

1. Hwang SY, Danpanichkul P, Agopian V, et al. Hepatocellular carcinoma: updates on epidemiology, surveillance, diagnosis and treatment. *Clin Mol Hepatol*. Feb 2025;31(Suppl):S228-S254. [doi: [10.3350/cmh.2024.0824](https://doi.org/10.3350/cmh.2024.0824)] [Medline: [39722614](https://pubmed.ncbi.nlm.nih.gov/39722614/)]
2. Brown ZJ, Tsilimigras DI, Ruff SM, et al. Management of hepatocellular carcinoma: a review. *JAMA Surg*. Apr 1, 2023;158(4):410-420. [doi: [10.1001/jamasurg.2022.7989](https://doi.org/10.1001/jamasurg.2022.7989)] [Medline: [36790767](https://pubmed.ncbi.nlm.nih.gov/36790767/)]
3. Vogel A, Meyer T, Sapisochin G, Salem R, Saborowski A. Hepatocellular carcinoma. *Lancet*. Oct 15, 2022;400(10360):1345-1362. [doi: [10.1016/S0140-6736\(22\)01200-4](https://doi.org/10.1016/S0140-6736(22)01200-4)] [Medline: [36084663](https://pubmed.ncbi.nlm.nih.gov/36084663/)]
4. Zhang ZH, Jiang C, Qiang ZY, et al. Role of microvascular invasion in early recurrence of hepatocellular carcinoma after liver resection: a literature review. *Asian J Surg*. May 2024;47(5):2138-2143. [doi: [10.1016/j.asjsur.2024.02.115](https://doi.org/10.1016/j.asjsur.2024.02.115)] [Medline: [38443255](https://pubmed.ncbi.nlm.nih.gov/38443255/)]
5. Fuster-Anglada C, Mauro E, Ferrer-Fàbrega J, et al. Histological predictors of aggressive recurrence of hepatocellular carcinoma after liver resection. *J Hepatol*. Dec 2024;81(6):995-1004. [doi: [10.1016/j.jhep.2024.06.018](https://doi.org/10.1016/j.jhep.2024.06.018)] [Medline: [38925272](https://pubmed.ncbi.nlm.nih.gov/38925272/)]
6. Li SH, Mei J, Cheng Y, et al. Postoperative adjuvant hepatic arterial infusion chemotherapy with FOLFOX in hepatocellular carcinoma with microvascular invasion: a multicenter, phase III, randomized study. *JCO*. Apr 1, 2023;41(10):1898-1908. [doi: [10.1200/JCO.22.01142](https://doi.org/10.1200/JCO.22.01142)]
7. Kang I, Jang M, Lee JG, et al. Subclassification of microscopic vascular invasion in hepatocellular carcinoma. *Ann Surg*. 2021;274(6):e1170-e1178. [doi: [10.1097/SLA.0000000000003781](https://doi.org/10.1097/SLA.0000000000003781)]

8. Kim NR, Bae H, Hwang HS, et al. Preoperative prediction of microvascular invasion with gadoxetic acid-enhanced magnetic resonance imaging in patients with single hepatocellular carcinoma: the implication of surgical decision on the extent of liver resection. *Liver Cancer*. Apr 24, 2024;13(2):181-192. [doi: [10.1159/000531786](https://doi.org/10.1159/000531786)]
9. Ma L, Zhang C, Wen Y, et al. Imaging-based surrogate classification for risk stratification of hepatocellular carcinoma with microvascular invasion to adjuvant hepatic arterial infusion chemotherapy: a multicenter retrospective study. *Int J Surg*. Jan 1, 2025;111(1):872-883. [doi: [10.1097/JS9.0000000000001903](https://doi.org/10.1097/JS9.0000000000001903)]
10. Lee S, Kang TW, Song KD, et al. Effect of microvascular invasion risk on early recurrence of hepatocellular carcinoma after surgery and radiofrequency ablation. *Ann Surg*. 2021;273(3):564-571. [doi: [10.1097/SLA.0000000000003268](https://doi.org/10.1097/SLA.0000000000003268)]
11. Xu XF, Diao YK, Zeng YY, et al. Association of severity in the grading of microvascular invasion with long-term oncological prognosis after liver resection for early-stage hepatocellular carcinoma: a multicenter retrospective cohort study from a hepatitis B virus-endemic area. *Int J Surg*. Apr 1, 2023;109(4):841-849. [doi: [10.1097/JS9.000000000000325](https://doi.org/10.1097/JS9.000000000000325)] [Medline: [36974673](https://pubmed.ncbi.nlm.nih.gov/36974673/)]
12. Wang L, Jia M, Wen X, Shen J, Yang H. Diagnostic value of magnetic resonance imaging features of microvascular invasion in hepatocellular carcinoma: a meta-analysis. *Diagn Interv Radiol*. Sep 2022;28(5):428-440. [doi: [10.5152/dir.2022.21108](https://doi.org/10.5152/dir.2022.21108)] [Medline: [36218148](https://pubmed.ncbi.nlm.nih.gov/36218148/)]
13. Xie XY, Chen R. Research progress of MRI-based radiomics in hepatocellular carcinoma. *Front Oncol*. 2025;15:1420599. [doi: [10.3389/fonc.2025.1420599](https://doi.org/10.3389/fonc.2025.1420599)] [Medline: [39980543](https://pubmed.ncbi.nlm.nih.gov/39980543/)]
14. Zhu Z, Wu K, Lu J, et al. Gd-EOB-DTPA-enhanced MRI radiomics and deep learning models to predict microvascular invasion in hepatocellular carcinoma: a multicenter study. *BMC Med Imaging*. Mar 31, 2025;25(1):40165094. [doi: [10.1186/s12880-025-01646-9](https://doi.org/10.1186/s12880-025-01646-9)]
15. Zhang L, Wang Y, Peng Z, et al. The progress of multimodal imaging combination and subregion based radiomics research of cancers. *Int J Biol Sci*. 2022;18(8):3458-3469. [doi: [10.7150/ijbs.71046](https://doi.org/10.7150/ijbs.71046)]
16. Zheng T, Zhu Y, Jiang H, et al. MRI-based topology deep learning model for noninvasive prediction of microvascular invasion and assisting prognostic stratification in HCC. *Liver Int*. Mar 2025;45(3):e16205. [doi: [10.1111/liv.16205](https://doi.org/10.1111/liv.16205)] [Medline: [39992060](https://pubmed.ncbi.nlm.nih.gov/39992060/)]
17. Zhu Y, Wang P, Wang B, et al. Accelerated multi-b-value DWI using deep learning reconstruction: image quality improvement and microvascular invasion prediction in BCLC stage A hepatocellular carcinoma. *Acad Radiol*. Jul 2025;32(7):3924-3937. [doi: [10.1016/j.acra.2025.01.043](https://doi.org/10.1016/j.acra.2025.01.043)]
18. McInnes MDF, Moher D, Thombs BD, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies. *JAMA*. Jan 23, 2018;319(4):388. [doi: [10.1001/jama.2017.19163](https://doi.org/10.1001/jama.2017.19163)]
19. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *J Med Libr Assoc*. Apr 1, 2021;109(2):174-200. [doi: [10.5195/jmla.2021.962](https://doi.org/10.5195/jmla.2021.962)] [Medline: [34285662](https://pubmed.ncbi.nlm.nih.gov/34285662/)]
20. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol*. Oct 2005;58(10):982-990. [doi: [10.1016/j.jclinepi.2005.02.022](https://doi.org/10.1016/j.jclinepi.2005.02.022)]
21. Zhang X, Yu X, Liang W, et al. Deep learning-based accurate diagnosis and quantitative evaluation of microvascular invasion in hepatocellular carcinoma on whole-slide histopathology images. *Cancer Med*. Mar 2024;13(5):e7104. [doi: [10.1002/cam4.7104](https://doi.org/10.1002/cam4.7104)] [Medline: [38488408](https://pubmed.ncbi.nlm.nih.gov/38488408/)]
22. Lei Y, Feng B, Wan M, et al. Predicting microvascular invasion in hepatocellular carcinoma with a CT- and MRI-based multimodal deep learning model. *Abdom Radiol*. May 2024;49(5):1397-1410. [doi: [10.1007/s00261-024-04202-1](https://doi.org/10.1007/s00261-024-04202-1)]
23. Liu HF, Wang M, Lu YJ, et al. CEMRI-based quantification of intratumoral heterogeneity for predicting aggressive characteristics of hepatocellular carcinoma using habitat analysis: comparison and combination of deep learning. *Acad Radiol*. Jun 2024;31(6):2346-2355. [doi: [10.1016/j.acra.2023.11.024](https://doi.org/10.1016/j.acra.2023.11.024)] [Medline: [38057182](https://pubmed.ncbi.nlm.nih.gov/38057182/)]
24. Zhou Z, Xia T, Zhang T, et al. Prediction of preoperative microvascular invasion by dynamic radiomic analysis based on contrast-enhanced computed tomography. *Abdom Radiol*. Feb 2024;49(2):611-624. [doi: [10.1007/s00261-023-04102-w](https://doi.org/10.1007/s00261-023-04102-w)]
25. He X, Xu Y, Zhou C, et al. Prediction of microvascular invasion and pathological differentiation of hepatocellular carcinoma based on a deep learning model. *Eur J Radiol*. Mar 2024;172:111348. [doi: [10.1016/j.ejrad.2024.111348](https://doi.org/10.1016/j.ejrad.2024.111348)] [Medline: [111348](https://pubmed.ncbi.nlm.nih.gov/4111348/)]
26. Zhong Y, Chen L, Ding F, Ou W, Zhang X, Weng S. Assessing microvascular invasion in HBV-related hepatocellular carcinoma: an online interactive nomogram integrating inflammatory markers, radiomics, and convolutional neural networks. *Front Oncol*. 2024;14:1401095. [doi: [10.3389/fonc.2024.1401095](https://doi.org/10.3389/fonc.2024.1401095)] [Medline: [39351352](https://pubmed.ncbi.nlm.nih.gov/39351352/)]
27. Wang F, Zhan G, Chen QQ, et al. Multitask deep learning for prediction of microvascular invasion and recurrence-free survival in hepatocellular carcinoma based on MRI images. *Liver Int*. Jun 2024;44(6):1351-1362. [doi: [10.1111/liv.15870](https://doi.org/10.1111/liv.15870)] [Medline: [38436551](https://pubmed.ncbi.nlm.nih.gov/38436551/)]

28. Yu Z, Liu Y, Dai X, Cui E, Cui J, Ma C. Enhancing preoperative diagnosis of microvascular invasion in hepatocellular carcinoma: domain-adaptation fusion of multi-phase CT images. *Front Oncol.* 2024;14:1332188. [doi: [10.3389/fonc.2024.1332188](https://doi.org/10.3389/fonc.2024.1332188)] [Medline: [38333689](https://pubmed.ncbi.nlm.nih.gov/38333689/)]
29. Ma H, Wang L, Sun L, et al. Preoperative prediction of microvascular invasion in hepatocellular carcinoma from multi-sequence magnetic resonance imaging based on deep fusion representation learning. *IEEE J Biomed Health Inform.* May 2025;29(5):3259-3271. [doi: [10.1109/JBHI.2024.3451331](https://doi.org/10.1109/JBHI.2024.3451331)] [Medline: [39196745](https://pubmed.ncbi.nlm.nih.gov/39196745/)]
30. Zhang W, Guo Q, Zhu Y, et al. Cross-institutional evaluation of deep learning and radiomics models in predicting microvascular invasion in hepatocellular carcinoma: validity, robustness, and ultrasound modality efficacy comparison. *Cancer Imaging.* Oct 22, 2024;24(1):142. [doi: [10.1186/s40644-024-00790-9](https://doi.org/10.1186/s40644-024-00790-9)] [Medline: [39438929](https://pubmed.ncbi.nlm.nih.gov/39438929/)]
31. Wang F, Chen Q, Chen Y, et al. A novel multimodal deep learning model for preoperative prediction of microvascular invasion and outcome in hepatocellular carcinoma. *Eur J Surg Oncol.* Jan 2023;49(1):156-164. [doi: [10.1016/j.ejso.2022.08.036](https://doi.org/10.1016/j.ejso.2022.08.036)] [Medline: [36333180](https://pubmed.ncbi.nlm.nih.gov/36333180/)]
32. You H, Wang J, Ma R, et al. Clinical interpretability of deep learning for predicting microvascular invasion in hepatocellular carcinoma by using attention mechanism. *Bioengineering (Basel).* Aug 9, 2023;10(8):948. [doi: [10.3390/bioengineering10080948](https://doi.org/10.3390/bioengineering10080948)] [Medline: [37627833](https://pubmed.ncbi.nlm.nih.gov/37627833/)]
33. Qin X, Zhu J, Tu Z, Ma Q, Tang J, Zhang C. Contrast-enhanced ultrasound with deep learning with attention mechanisms for predicting microvascular invasion in single hepatocellular carcinoma. *Acad Radiol.* Sep 2023;30 Suppl 1:S73-S80. [doi: [10.1016/j.acra.2022.12.005](https://doi.org/10.1016/j.acra.2022.12.005)] [Medline: [36567144](https://pubmed.ncbi.nlm.nih.gov/36567144/)]
34. Li Z, Wang Y, Zhu Y, et al. Modality-based attention and dual-stream multiple instance convolutional neural network for predicting microvascular invasion of hepatocellular carcinoma. *Front Oncol.* 2023;13:1195110. [doi: [10.3389/fonc.2023.1195110](https://doi.org/10.3389/fonc.2023.1195110)] [Medline: [37434971](https://pubmed.ncbi.nlm.nih.gov/37434971/)]
35. Cao L, Wang Q, Hong J, et al. MVI-TR: a transformer-based deep learning model with contrast-enhanced CT for preoperative prediction of microvascular invasion in hepatocellular carcinoma. *Cancers (Basel).* Feb 28, 2023;15(5):1538. [doi: [10.3390/cancers15051538](https://doi.org/10.3390/cancers15051538)] [Medline: [36900327](https://pubmed.ncbi.nlm.nih.gov/36900327/)]
36. Wang T, Li Z, Yu H, et al. Prediction of microvascular invasion in hepatocellular carcinoma based on preoperative Gd-EOB-DTPA-enhanced MRI: comparison of predictive performance among 2D, 2D-expansion and 3D deep learning models. *Front Oncol.* 2023;13:987781. [doi: [10.3389/fonc.2023.987781](https://doi.org/10.3389/fonc.2023.987781)] [Medline: [36816963](https://pubmed.ncbi.nlm.nih.gov/36816963/)]
37. Ye Z, Zhang J, Wu N, et al. PET-guided attention for prediction of microvascular invasion in preoperative hepatocellular carcinoma on PET/CT. *Ann Nucl Med.* Apr 2023;37(4):238-245. [doi: [10.1007/s12149-023-01822-3](https://doi.org/10.1007/s12149-023-01822-3)] [Medline: [36723705](https://pubmed.ncbi.nlm.nih.gov/36723705/)]
38. Xu Y, Chen Y, Wu J, Pan J, Liao C, Su H. The utility of vision transformer in preoperatively predicting microvascular invasion status of hepatocellular carcinoma. *HPB (Oxford).* May 2023;25(5):533-542. [doi: [10.1016/j.hpb.2023.01.015](https://doi.org/10.1016/j.hpb.2023.01.015)]
39. Deng Y, Jia X, Yu G, et al. Can a proposed double branch multimodality-contribution-aware TripNet improve the prediction performance of the microvascular invasion of hepatocellular carcinoma based on small samples? *Front Oncol.* 2022;12:1035775. [doi: [10.3389/fonc.2022.1035775](https://doi.org/10.3389/fonc.2022.1035775)] [Medline: [36387069](https://pubmed.ncbi.nlm.nih.gov/36387069/)]
40. Li X, Qi Z, Du H, et al. Deep convolutional neural network for preoperative prediction of microvascular invasion and clinical outcomes in patients with HCCs. *Eur Radiol.* Feb 2022;32(2):771-782. [doi: [10.1007/s00330-021-08198-w](https://doi.org/10.1007/s00330-021-08198-w)]
41. Chen Q, Xiao H, Gu Y, et al. Deep learning for evaluation of microvascular invasion in hepatocellular carcinoma from tumor areas of histology images. *Hepatol Int.* Jun 2022;16(3):590-602. [doi: [10.1007/s12072-022-10323-w](https://doi.org/10.1007/s12072-022-10323-w)]
42. Zhang Y, Wei Q, Huang Y, et al. Deep learning of liver contrast-enhanced ultrasound to predict microvascular invasion and prognosis in hepatocellular carcinoma. *Front Oncol.* 2022;12:878061. [doi: [10.3389/fonc.2022.878061](https://doi.org/10.3389/fonc.2022.878061)] [Medline: [35875110](https://pubmed.ncbi.nlm.nih.gov/35875110/)]
43. Liu B, Zeng Q, Huang J, et al. IVIM using convolutional neural networks predicts microvascular invasion in HCC. *Eur Radiol.* Oct 2022;32(10):7185-7195. [doi: [10.1007/s00330-022-08927-9](https://doi.org/10.1007/s00330-022-08927-9)]
44. Sun SW, Xu X, Liu QP, et al. LiSNet: an artificial intelligence -based tool for liver imaging staging of hepatocellular carcinoma aggressiveness. *Med Phys Mex Symp Med Phys.* Nov 2022;49(11):6903-6913. [doi: [10.1002/mp.15972](https://doi.org/10.1002/mp.15972)]
45. Wang L, Wu M, Li R, Xu X, Zhu C, Feng X. MVI-Mind: a novel deep-learning strategy using computed tomography (CT)-based radiomics for end-to-end high efficiency prediction of microvascular invasion in hepatocellular carcinoma. *Cancers (Basel).* 2022;14(12):2956. [doi: [10.3390/cancers14122956](https://doi.org/10.3390/cancers14122956)]
46. Xiao H, Guo Y, Zhou Q, et al. Prediction of microvascular invasion in hepatocellular carcinoma with expert-inspiration and skeleton sharing deep learning. *Liver Int.* Jun 2022;42(6):1423-1431. [doi: [10.1111/liv.15254](https://doi.org/10.1111/liv.15254)] [Medline: [35319151](https://pubmed.ncbi.nlm.nih.gov/35319151/)]
47. Yang Y, Zhou Y, Zhou C, Ma X. Deep learning radiomics based on contrast enhanced computed tomography predicts microvascular invasion and survival outcome in early stage hepatocellular carcinoma. *Eur J Surg Oncol.* May 2022;48(5):1068-1077. [doi: [10.1016/j.ejso.2021.11.120](https://doi.org/10.1016/j.ejso.2021.11.120)]

48. Sun BY, Gu PY, Guan RY, et al. Deep-learning-based analysis of preoperative MRI predicts microvascular invasion and outcome in hepatocellular carcinoma. *World J Surg Oncol*. Jun 8, 2022;20(1):189. [doi: [10.1186/s12957-022-02645-8](https://doi.org/10.1186/s12957-022-02645-8)] [Medline: [35676669](https://pubmed.ncbi.nlm.nih.gov/35676669/)]
49. Dai X, Cao S, Guo Y, et al. Preoperative identification of microvascular invasion in hepatocellular carcinoma based on multi-modal and multi response convolutional neural network. Presented at: Fourteenth International Conference on Graphics and Image Processing (ICGIP 2022); Oct 21-23, 2022; Nanjing, China. [doi: [10.1117/12.2680693](https://doi.org/10.1117/12.2680693)]
50. Zhang Y, Lv X, Qiu J, et al. Deep learning with 3D convolutional neural network for noninvasive prediction of microvascular invasion in hepatocellular carcinoma. *Magn Reson Imaging*. Jul 2021;54(1):134-143. [doi: [10.1002/jmri.27538](https://doi.org/10.1002/jmri.27538)]
51. Liu SC, Lai J, Huang JY, et al. Predicting microvascular invasion in hepatocellular carcinoma: a deep learning model validated across hospitals. *Cancer Imaging*. Oct 9, 2021;21(1):56. [doi: [10.1186/s40644-021-00425-3](https://doi.org/10.1186/s40644-021-00425-3)] [Medline: [34627393](https://pubmed.ncbi.nlm.nih.gov/34627393/)]
52. Wei J, Jiang H, Zeng M, et al. Prediction of microvascular invasion in hepatocellular carcinoma via deep learning: a multi-center and prospective validation study. *Cancers (Basel)*. May 14, 2021;13(10):2368. [doi: [10.3390/cancers13102368](https://doi.org/10.3390/cancers13102368)]
53. Zhou W, Jian W, Cen X, et al. Prediction of microvascular invasion of hepatocellular carcinoma based on contrast-enhanced MR and 3D convolutional neural networks. *Front Oncol*. 2021;11:588010. [doi: [10.3389/fonc.2021.588010](https://doi.org/10.3389/fonc.2021.588010)] [Medline: [33854959](https://pubmed.ncbi.nlm.nih.gov/33854959/)]
54. Wang G, Jian W, Cen X, et al. Prediction of microvascular invasion of hepatocellular carcinoma based on preoperative diffusion-weighted MR using deep learning. *Acad Radiol*. Nov 2021;28:S118-S127. [doi: [10.1016/j.acra.2020.11.014](https://doi.org/10.1016/j.acra.2020.11.014)]
55. Zeng Q, Liu B, Xu Y, Zhou W. An attention-based deep learning model for predicting microvascular invasion of hepatocellular carcinoma using an intra-voxel incoherent motion model of diffusion-weighted magnetic resonance imaging. *Phys Med Biol*. Sep 21, 2021;66(18):185019. [doi: [10.1088/1361-6560/ac22db](https://doi.org/10.1088/1361-6560/ac22db)]
56. Gao F, Qiao K, Yan B, et al. Hybrid network with difference degree and attention mechanism combined with radiomics (H-DARnet) for MVI prediction in HCC. *Magn Reson Imaging*. Nov 2021;83:27-40. [doi: [10.1016/j.mri.2021.06.018](https://doi.org/10.1016/j.mri.2021.06.018)] [Medline: [34147593](https://pubmed.ncbi.nlm.nih.gov/34147593/)]
57. Jiang YQ, Cao SE, Cao S, et al. Preoperative identification of microvascular invasion in hepatocellular carcinoma by XGBoost and deep learning. *J Cancer Res Clin Oncol*. Mar 2021;147(3):821-833. [doi: [10.1007/s00432-020-03366-9](https://doi.org/10.1007/s00432-020-03366-9)] [Medline: [32852634](https://pubmed.ncbi.nlm.nih.gov/32852634/)]
58. Song D, Wang Y, Wang W, et al. Using deep learning to predict microvascular invasion in hepatocellular carcinoma based on dynamic contrast-enhanced MRI combined with clinical parameters. *J Cancer Res Clin Oncol*. Dec 2021;147(12):3757-3767. [doi: [10.1007/s00432-021-03617-3](https://doi.org/10.1007/s00432-021-03617-3)] [Medline: [33839938](https://pubmed.ncbi.nlm.nih.gov/33839938/)]
59. Men S, Ju H, Zhang L, Zhou W. Prediction of microvascular invasion of hepatocellular carcinoma with contrast-enhanced MR using 3D CNN and LSTM. Presented at: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI); Oct 8-11, 2019; Venice, Italy. [doi: [10.1109/ISBI.2019.8759582](https://doi.org/10.1109/ISBI.2019.8759582)]
60. Zhou Y, Sun SW, Liu QP, Xu X, Zhang Y, Zhang YD. TED: two-stage expert-guided interpretable diagnosis framework for microvascular invasion in hepatocellular carcinoma. *Med Image Anal*. Nov 2022;82:102575. [doi: [10.1016/j.media.2022.102575](https://doi.org/10.1016/j.media.2022.102575)] [Medline: [36063747](https://pubmed.ncbi.nlm.nih.gov/36063747/)]
61. Chu T, Zhao C, Zhang J, et al. Application of a convolutional neural network for multitask learning to simultaneously predict microvascular invasion and vessels that encapsulate tumor clusters in hepatocellular carcinoma. *Ann Surg Oncol*. Oct 2022;29(11):6774-6783. [doi: [10.1245/s10434-022-12000-6](https://doi.org/10.1245/s10434-022-12000-6)] [Medline: [35754067](https://pubmed.ncbi.nlm.nih.gov/35754067/)]
62. Huang H, Liu B, Zhang L, Xu Y, Zhou W, editors. Transformer Based Multi-Task Deep Learning with Intravoxel Incoherent Motion Model Fitting for Microvascular Invasion Prediction of Hepatocellular Carcinoma. Springer Nature Switzerland; 2022. [doi: [10.1007/978-3-031-16449-1\\_26](https://doi.org/10.1007/978-3-031-16449-1_26)]
63. Wang Y, Xie W, Li C, et al. Automated microvascular invasion prediction of hepatocellular carcinoma via deep relation reasoning from dynamic contrast-enhanced ultrasound. *Comput Med Imaging Graph*. Sep 2025;124:102606. [doi: [10.1016/j.compmedimag.2025.102606](https://doi.org/10.1016/j.compmedimag.2025.102606)] [Medline: [40680384](https://pubmed.ncbi.nlm.nih.gov/40680384/)]
64. Cen YY, Nong HY, Huang XX, et al. Computed tomography-based deep learning and multi-instance learning for predicting microvascular invasion and prognosis in hepatocellular carcinoma. *World J Gastroenterol*. Aug 14, 2025;31(30):109186. [doi: [10.3748/wjg.v31.i30.109186](https://doi.org/10.3748/wjg.v31.i30.109186)] [Medline: [40933208](https://pubmed.ncbi.nlm.nih.gov/40933208/)]
65. Huang Z, Huang W, Jiang L, et al. Decision fusion model for predicting microvascular invasion in hepatocellular carcinoma based on multi-MR habitat imaging and machine-learning classifiers. *Acad Radiol*. Apr 2025;32(4):1971-1980. [doi: [10.1016/j.acra.2024.10.007](https://doi.org/10.1016/j.acra.2024.10.007)] [Medline: [39472207](https://pubmed.ncbi.nlm.nih.gov/39472207/)]
66. Miao S, Sun M, Li X, et al. Deep learning-based prediction of microvascular invasion and survival outcomes in hepatocellular carcinoma using dual-phase CT imaging of tumors and lesser omental adipose: a multicenter study. *Acad Radiol*. Oct 2025;32(10):5789-5801. [doi: [10.1016/j.acra.2025.07.015](https://doi.org/10.1016/j.acra.2025.07.015)]

67. Dong X, Jia X, Zhang W, et al. Interpretable and generalizable deep learning model for preoperative assessment of microvascular invasion and outcome in hepatocellular carcinoma based on MRI: a multicenter study. *Insights Imaging*. Jul 3, 2025;16(1):151. [doi: [10.1186/s13244-025-02035-0](https://doi.org/10.1186/s13244-025-02035-0)] [Medline: [40610844](https://pubmed.ncbi.nlm.nih.gov/40610844/)]
68. Zhang Y, Liu H, Zhu L, et al. Modality-aware distillation network for microvascular invasion prediction of hepatocellular carcinoma from MRI images. *IEEE Trans Biomed Eng*. Jun 2025;72(6):1825-1836. [doi: [10.1109/TBME.2024.3523921](https://doi.org/10.1109/TBME.2024.3523921)]
69. Zhao Y, Huang X, Sun M, et al. Predicting microvascular invasion plus cytokeratin 19 expression positivity in hepatocellular carcinoma based on EOB-MRI using multitask deep learning. *Hepatoma Res*. 2025;10:12. [doi: [10.20517/2394-5079.2024.143](https://doi.org/10.20517/2394-5079.2024.143)]
70. Qin Q, Pang J, Li J, et al. Transformer model based on Sonazoid contrast-enhanced ultrasound for microvascular invasion prediction in hepatocellular carcinoma. *Med Phys*. Jul 2025;52(7):e17895. [doi: [10.1002/mp.17895](https://doi.org/10.1002/mp.17895)] [Medline: [40384312](https://pubmed.ncbi.nlm.nih.gov/40384312/)]
71. Xiao Q, Zhu W, Tang H, Zhou L. Ultrasound radiomics in the prediction of microvascular invasion in hepatocellular carcinoma: a systematic review and meta-analysis. *Heliyon*. Jun 2023;9(6):e16997. [doi: [10.1016/j.heliyon.2023.e16997](https://doi.org/10.1016/j.heliyon.2023.e16997)]
72. Liang G, Yu W, Liu S, et al. The diagnostic performance of radiomics-based MRI in predicting microvascular invasion in hepatocellular carcinoma: a meta-analysis. *Front Oncol*. 2022;12:960944. [doi: [10.3389/fonc.2022.960944](https://doi.org/10.3389/fonc.2022.960944)] [Medline: [36798691](https://pubmed.ncbi.nlm.nih.gov/36798691/)]
73. Li L, Wu C, Huang Y, Chen J, Ye D, Su Z. Radiomics for the preoperative evaluation of microvascular invasion in hepatocellular carcinoma: a meta-analysis. *Front Oncol*. 2022;12:831996. [doi: [10.3389/fonc.2022.831996](https://doi.org/10.3389/fonc.2022.831996)] [Medline: [35463303](https://pubmed.ncbi.nlm.nih.gov/35463303/)]
74. He K, Liu X, Shahzad R, et al. Advanced deep learning approach to automatically segment malignant tumors and ablation zone in the liver with contrast-enhanced CT. *Front Oncol*. 2021;11:669437. [doi: [10.3389/fonc.2021.669437](https://doi.org/10.3389/fonc.2021.669437)] [Medline: [34336661](https://pubmed.ncbi.nlm.nih.gov/34336661/)]
75. Kinoshita M, Ueda D, Matsumoto T, et al. Deep learning model based on contrast-enhanced computed tomography imaging to predict postoperative early recurrence after the curative resection of a solitary hepatocellular carcinoma. *Cancers (Basel)*. Apr 4, 2023;15(7):2140. [doi: [10.3390/cancers15072140](https://doi.org/10.3390/cancers15072140)] [Medline: [37046801](https://pubmed.ncbi.nlm.nih.gov/37046801/)]
76. Lu D, Wang LF, Han H, et al. Prediction of microvascular invasion in hepatocellular carcinoma with conventional ultrasound, Sonazoid-enhanced ultrasound, and biochemical indicator: a multicenter study. *Insights Imaging*. Oct 28, 2024;15(1):261. [doi: [10.1186/s13244-024-01743-3](https://doi.org/10.1186/s13244-024-01743-3)] [Medline: [39466459](https://pubmed.ncbi.nlm.nih.gov/39466459/)]
77. Yan M, Zhang X, Zhang B, et al. Deep learning nomogram based on Gd-EOB-DTPA MRI for predicting early recurrence in hepatocellular carcinoma after hepatectomy. *Eur Radiol*. Jul 2023;33(7):4949-4961. [doi: [10.1007/s00330-023-09419-0](https://doi.org/10.1007/s00330-023-09419-0)] [Medline: [36786905](https://pubmed.ncbi.nlm.nih.gov/36786905/)]
78. Wu Y, Zhu M, Liu Y, Cao X, Zhang G, Yin L. Peritumoral imaging manifestations on Gd-EOB-DTPA-enhanced MRI for preoperative prediction of microvascular invasion in hepatocellular carcinoma: a systematic review and meta-analysis. *Front Oncol*. 2022;12:907076. [doi: [10.3389/fonc.2022.907076](https://doi.org/10.3389/fonc.2022.907076)] [Medline: [35814461](https://pubmed.ncbi.nlm.nih.gov/35814461/)]
79. Li C, Tan Y, Chen W, et al. Attention unet++: a nested attention-aware u-net for liver CT image segmentation. Presented at: 2020 IEEE International Conference on Image Processing (ICIP); Oct 25-28, 2020:345-349; Abu Dhabi, United Arab Emirates. 2020. [doi: [10.1109/ICIP40778.2020.9190761](https://doi.org/10.1109/ICIP40778.2020.9190761)]
80. Huang H, Lin L, Tong R, et al. UNet 3+: a full-scale connected unet for medical image segmentation. Presented at: ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); May 4-8, 2020:1055-1059; Barcelona, Spain. [doi: [10.1109/ICASSP40776.2020.9053405](https://doi.org/10.1109/ICASSP40776.2020.9053405)]
81. Bilic P, Christ P, Li HB, et al. The liver tumor segmentation benchmark (LiTS). *Med Image Anal*. Feb 2023;84:102680. [doi: [10.1016/j.media.2022.102680](https://doi.org/10.1016/j.media.2022.102680)] [Medline: [36481607](https://pubmed.ncbi.nlm.nih.gov/36481607/)]
82. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. Feb 2016;278(2):563-577. [doi: [10.1148/radiol.2015151169](https://doi.org/10.1148/radiol.2015151169)]
83. Riley RD, Archer L, Snell KIE, et al. Evaluation of clinical prediction models (part 2): how to undertake an external validation study. *BMJ*. Jan 15, 2024;384:e074820. [doi: [10.1136/bmj-2023-074820](https://doi.org/10.1136/bmj-2023-074820)]
84. Collins GS, Dhiman P, Ma J, et al. Evaluation of clinical prediction models (part 1): from development to external validation. *BMJ*. Jan 8, 2024;384:e074819. [doi: [10.1136/bmj-2023-074819](https://doi.org/10.1136/bmj-2023-074819)]
85. Strandberg R, Jepsen P, Hagström H. Developing and validating clinical prediction models in hepatology - an overview for clinicians. *J Hepatol*. Jul 2024;81(1):149-162. [doi: [10.1016/j.jhep.2024.03.030](https://doi.org/10.1016/j.jhep.2024.03.030)] [Medline: [38531493](https://pubmed.ncbi.nlm.nih.gov/38531493/)]
86. Maung ST, Tangkijvanich P, Chaiteerakij R. Magnetic resonance imaging radiomics for predicting hepatocellular carcinoma recurrence following resection or ablation: a systematic review and meta-analysis. *Abdom Radiol (NY)*. Sep 22, 2025. [doi: [10.1007/s00261-025-05184-4](https://doi.org/10.1007/s00261-025-05184-4)] [Medline: [40981987](https://pubmed.ncbi.nlm.nih.gov/40981987/)]

## Abbreviations

**AI:** artificial intelligence

**AUC:** area under the curve  
**CECT:** contrast-enhanced computed tomography  
**CEMRI:** contrast-enhanced magnetic resonance imaging  
**CEUS:** contrast-enhanced ultrasound  
**DL:** deep learning  
**DOR:** diagnostic odds ratio  
**HCC:** hepatocellular carcinoma  
**LFK:** Luis Furuya-Kanamori  
**LR:** likelihood ratio  
**MeSH:** Medical Subject Headings  
**MRI:** magnetic resonance imaging  
**MVI:** microvascular invasion  
**PRISMA-DTA:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Diagnostic Test Accuracy  
**PRISMA-S:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for literature searches  
**QUADAS-2:** Quality Assessment of Diagnostic Accuracy Studies 2  
**QUADAS-AI:** Quality Assessment of Diagnostic Accuracy Studies for Artificial Intelligence  
**ROB:** risk of bias  
**SROC:** summary receiver operating characteristic

*Edited by Stefano Brini; peer-reviewed by Long Zhou, Soe Thiha Maung; submitted 07.Aug.2025; accepted 08.Jan.2026; published 02.Mar.2026*

*Please cite as:*

*Feng W, Qu B, Han S*

*Accuracy of Medical Image-Based Deep Learning for Detecting Microvascular Invasion in Hepatocellular Carcinoma: Systematic Review and Meta-Analysis*

*J Med Internet Res 2026;28:e82000*

*URL: <https://www.jmir.org/2026/1/e82000>*

*doi: [10.2196/82000](https://doi.org/10.2196/82000)*

© Wei Feng, Bo Qu, Shuo Han. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 02.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.