

# 基于深度学习的肝脏局灶性病变检测方法综述

董文益<sup>1</sup>, 杨伟东<sup>1</sup>, 唐冰慧<sup>1</sup>, 王琦<sup>2</sup>, 肖宏宇<sup>3</sup>

(1. 河北工业大学机械工程学院, 天津 300103;

2. 河北医科大学第四医院, 河北 石家庄 050011;

3. 中国人民解放军陆军航空兵学院, 北京 101123)

**摘 要:** 肝脏局灶性病变(FLLs)检测对疾病诊断和治疗至关重要。传统检测方法面临诸多挑战, 深度学习技术的应用为其带来新契机。鉴于此, 系统综述了基于深度学习的FLLs检测方法, 通过深入分析相关技术的优势与不足, 为FLLs检测技术的发展提供了具体的研究方向。首先对肝脏放射影像的公开数据集进行了整理归纳, 阐述数据预处理对提升模型性能的关键作用。其次, 对比分析了基于卷积神经网络、Transformer以及知识蒸馏等技术的2D与3D检测算法, 揭示了从局部特征建模到全局时空关联的技术演进路径。此外, 深入探讨了针对多期相影像的时序特征融合方法, 为动态病变表征提供了新思路。研究表明, 现有方法在检测精度与效率上取得突破, 但仍面临小病灶敏感性不足、跨设备泛化性弱及临床验证缺乏等挑战。未来研究需通过多中心数据协同、轻量化算法设计及可解释性增强等途径, 加速深度学习在肝脏病变辅助诊断中的临床转化与应用。

**关 键 词:** 深度学习; 肝脏局灶性病变; 目标检测; 计算机断层扫描; 多期相

中图分类号: R 575; TP 391.41; TP 18 DOI: 10.11996/JGj.2095-302X.2026010001

文献标识码: A 文章编号: 2095-302X(2026)01-0001-16

## Review of deep learning based methods for detecting focal liver lesions

DONG Wenyi<sup>1</sup>, YANG Weidong<sup>1</sup>, TANG Binghui<sup>1</sup>, WANG Qi<sup>2</sup>, XIAO Hongyu<sup>3</sup>

(1. School of Mechanical Engineering, Hebei University of Technology, Tianjin 300103, China;

2. Fourth Hospital of Hebei Medical University, Shijiazhuang Hebei 050011, China;

3. Chinese People's Liberation Army Aviation School, Beijing 101123, China)

**Abstract:** The detection of Focal Liver Lesions (FLLs) is crucial for disease diagnosis and treatment. Traditional detection methods face many challenges, and the application of deep-learning technology brings new opportunities. In view of this, this paper systematically reviewed the deep-learning-based FLLs detection methods, and provided specific research directions for the development of FLLs detection technology by analyzing the advantages and disadvantages of related technologies. First, the public datasets of liver radiological images were organized and summarized, and the key role of data preprocessing in improving model performance was expounded. Secondly, the 2D and 3D detection algorithms based on convolutional neural networks, Transformer, knowledge distillation, and other technologies were compared and analyzed, revealing the technical evolution path from local feature modeling to global spatio-temporal correlation. In addition, the temporal feature fusion methods for multi-phase images were examined in depth, providing new ideas for dynamic lesion characterization. The review showed that existing methods had achieved breakthroughs in detection accuracy and efficiency, but still faced challenges such as insufficient sensitivity to small lesions, weak cross-device generalization, and lack of clinical verification. Future research was recommended to accelerate the clinical transformation and application of deep learning in auxiliary diagnosis of liver

收稿日期: 2025-03-19; 定稿日期: 2025-06-18; 通信作者: 杨伟东, E-mail: yangweidong@hebut.edu.cn

Received: 19 March, 2025; Finalized: 18 June, 2025; Corresponding author: YANG Weidong, E-mail: yangweidong@hebut.edu.cn

基金项目: 2022年政府资助临床医学优秀人才培养项目(冀财预复[2022]180号)

Foundation items: Government funded Clinical Medicine Excellent Talent Training Project in 2022 (Ji Cai Yu Fu [2022]180)

lesions through multi-center data collaboration, lightweight algorithm design, and enhanced interpretability.

**Keywords:** deep learning; focal liver lesions; object detection; computerized tomography scan; multi-phase

肝脏局灶性病变(Focal Liver Lesions, FLLs)是一类常见的肝脏异常,涵盖了从良性病变如肝囊肿、肝血管瘤到恶性病变如原发性肝癌、转移性肝癌等多种类型,其临床表现多样,且部分病变在早期可能无明显症状,为患者的诊断和治疗带来了挑战。我国癌症的致死率相比于发达国家偏高,其主要原因之一是早诊率低<sup>[1-2]</sup>。

传统的 FLLs 检测方法包括血液检查、肝穿刺活检和医学成像等<sup>[3]</sup>。尽管活检目前仍被认为是大多数肝病诊断的金标准,但医学成像因其方便、无创、成本低等优点,成为 FLLs 筛查的有效检查手段之一。常用的影像学检查手段包括计算机断层扫描(Computed Tomography, CT)、磁共振成像(Magnetic Resonance Imaging, MRI)和超声(Ultrasound, US)。这些技术能够清晰地显示肝脏的内部结构,帮助医生识别病变的存在。面对复杂多变的 FLLs,传统的人工判读存在主观性,且与医生的经验息息相关。随着医疗技术的不断进步和医学影像数据的急剧增加,传统的检测方法面临着严峻的挑战。

随着影像组学概念的发展,医学图像处理领域将计算机视觉技术应用于计算机辅助诊断(Computer-aided Diagnosis, CAD)。深度学习作为人工智能的一个分支,具有强大的特征学习和模式识别能力,将其应用于肝脏医学图像的 FLLs 检测中,对提高检测准确率,以及减轻医生工作负担和减少误诊率都有着积极的作用。同时,加快推进了 CAD 在 FLLs 辅助诊断中的作用。

尽管已有综述探讨了深度学习在医学影像中的应用<sup>[4-5]</sup>,但现有研究多聚焦于单一模态或 2D 检测方法,缺乏对 3D 检测技术演进、多期相时序特征融合及轻量化模型优化的系统性分析。本文从算法动态发展的视角,全面对比了 2D 与 3D 检测技术的互补性与局限性,并深入探讨了多期相动态建模、无锚框检测及知识蒸馏等前沿技术对临床转化的推动作用。

## 1 数据准备

在医学影像分析领域,影像组学作为一门结合医学影像学、机器学习和统计学的新兴学科,通过分析大量的医学影像数据识别疾病特征,辅助诊断与治疗。其技术适用于多种医学影像模态的数据分

析,涵盖 CT, MRI 和 US 等。深度学习在影像组学的应用需大量数据集训练模型,数据预处理对确保模型准确性和泛化能力至关重要。现有公开数据集为肝脏病变检测研究提供资源,但直接用于 FLLs 目标检测的研究资源稀缺。本文将梳理肝脏影像公开数据集,并探讨数据预处理方法,同时引入后续将阐述的评价标准体系,为后续模型训练与算法评估奠定基础。

### 1.1 肝脏公开影像数据集

目前可供使用的公共数据集包含了来自临床影像学领域的大量放射影像数据,可以用于训练和测试深度学习模型。同时这些数据集中的标注信息提供了肿瘤位置和分类信息,使得深度学习模型可以自动地学习和理解不同类型的放射影像数据。通过发掘放射影像数据中的信息,提高肿瘤诊断和治疗的准确性和效率。

SLIVER07(CT Liver Tumor Segmentation Challenge 2007)<sup>[6]</sup>是 2007 年的公共医学影像数据集,可用于肝脏分割和肿瘤检测,包含 30 例患者 CT 扫描数据,其中训练集 20 例、测试集 10 例,有手动标注,对早期影像分割技术有重要意义。3D-IRCADB (3D Image Reconstruction for Comparison of Algorithm Database)<sup>[7]</sup>是法国消化系统癌症治疗研究所管理的一个知名数据集,包含 20 例患者 CT 扫描,男女各 10 例,有手动标注的肝脏和肿瘤区域,标注质量高,虽数据集规模较小但应用广泛,常与 LiTS (Liver Tumor Segmentation Challenge)<sup>[8]</sup>挑战赛数据集结合使用。LiTS 包含高分辨率 CT 扫描和肝脏肿瘤分割标签,共 131 例训练样本和 70 例测试样本,该数据集为肝脏图像领域最常用的数据集之一。CHAOS (Combined CT-MR Healthy Abdominal Organ Segmentation)<sup>[9]</sup>挑战赛数据集包含腹部 CT 和 MRI 图像,各包含 20 个训练集和 20 个测试集,提供肝脏、肾脏和脾脏的真实掩码。

MSD (Medical Segmentation Decathlon)<sup>[10]</sup>挑战赛在 2018 年 MICCAI (Medical Image Computing and Computer Assisted Intervention)期间举行。MSD 数据集由 10 个数据集组成,其中包括 2 个肝脏相关数据集。其一含 201 张多期相 CT 图像,其二含 443 张 CT 图像,但血管标注不准确。JFR (Journées

Francophones de Radiologie)<sup>[11]</sup>挑战赛的肝脏超声数据集第一版于 2018 年公开, 是为数不多的用于深度学习的超声数据集之一。含 367 例二维超声图像, 旨在解决公共健康问题, 并为人工智能社区搭建协作平台。

对于其他较少采用的数据集, 如 ATLAS<sup>[12]</sup>, TriALS 2024 Task1<sup>[13]</sup>, HCC-TACE-Seg<sup>[14]</sup> 和 LLD-MMRI2023<sup>[15]</sup>等, 其分别采用了不同的成像手

段(如 CE-MRI 和 CT 等), 涵盖了不同的病变类型或人群, 并各有特色。此外, 还有一些综合性数据集涵盖了多种类型的医学图像, 除了肝脏还包括其他器官, 如 Radiopaedia 数据集<sup>[16]</sup>、UK BioBank 数据集<sup>[17]</sup>等, 这种综合性数据集往往具有完善的社区生态。尽管已有上述数据集可用, 但肝脏相关的公开影像资源仍较为有限。表 1 汇总了当前肝脏影像领域常用的公开数据集及其关键信息。

表 1 肝脏常用放射数据集

Table 1 Commonly used liver radiation dataset

数据集名称	成像手段	维度	数据量/例	文件格式	任务类型	简介
SLIVER07 <sup>[6]</sup>	CT	3D	30	mhd	分割	早期经典数据集, 样本量较小, 常与其他数据集结合用于肝脏分割与肿瘤检测研究
3D-ircadb <sup>[7]</sup>	CT	3D	22	dicom,vk	分割	样本量较小但标注质量高, 含肝脏及肿瘤区域手动标注, 广泛应用于肝脏影像算法验证
LiTS <sup>[8]</sup>	CT	3D	201	.nii	分割	高分辨率 CT 数据集, 为肝脏肿瘤分割领域常用基准数据集
CHAOS <sup>[9]</sup>	CT, MRI	3D	40	dicom	分割	腹部多器官综合数据集, 提供肝脏、肾脏、脾脏真实掩码
MSD <sup>[10]</sup>	CT, MRI	3D	644	.nii.gz	分割	包含 10 个医学影像数据集, 其中肝脏相关数据集样本量较大, 但存在血管标注不准确问题
JFR <sup>[11]</sup>	US	2D	367	/	检测	首个公开的肝脏超声检测数据集
ATLAS <sup>[12]</sup>	CE-MRI	3D	90	.nii.gz	分割	采用对比增强 MRI 成像, 提供肝脏及肿瘤分割标注
TriALS 2024 Task1 <sup>[13]</sup>	CT	3D	60	.nii.gz	分割	针对非洲人群设计, 专注于门静脉期肝脏病变分割
HCC-TACE-Seg <sup>[14]</sup>	CT	3D	628	.dcm	分割	包含 105 例确诊的肝细胞癌(HCC)患者 CT 数据, 用于 HCC 相关检测与分割研究
LLD-MMRI2023 <sup>[15]</sup>	MRI	3D	394	.nii.gz	检测	多模态 MRI 数据集, 涵盖 7 种肝脏病变类型

上述所公开的肝脏数据集为肝脏病变检测研究提供了丰富的资源。然而现有的数据集大多服务于分割任务研究, 可直接用于 FLLs 目标检测的研究资源较为稀缺。但部分进行一定的处理后也可用于检测任务, 如采用将掩码标注转化为检测框的处理方法; 有些也可用于 FLLs 检测任务的前置任务, 如进行粗分割、配准等。

## 1.2 数据处理方法

对于肝脏医学图像的 FLLs 检测任务, 数据的质量和数量直接影响模型的检测能力和准确率。因此, 对数据进行预处理是深度学习应用于 FLLs 检测的关键步骤之一, 可减少模型的过拟合和欠拟合情况, 提高模型的泛化能力。数据预处理主要包括数据清洗、数据增强等步骤。数据清洗是对数据进行筛选和过滤, 去除无关数据, 提高数据的质量。数据增强是通过一系列变换和扩充, 扩充数据量并丰富其多样性, 提高模型的泛化能力。2 种方法相结合对图像进行处理, 以提高图像的质量和准确性, 从而更好地支持深度学习的应用。

### 1.2.1 数据清洗

肝脏影像通常包含大量的细节和噪声, 一般的数据清洗主要从灰度维度进行, 即去除图像的噪声、处理颜色亮度等。通常采用限定 HU 值的方法

进行处理。由于健康组织和恶性组织之间存在对比度, 一般使用对比度增强来突出特定组织, 并提高肝脏 CT 图像的质量<sup>[18]</sup>。研究者则常利用直方图均衡化来提高 CT 图像的对比度。

SHAH 等<sup>[19]</sup>考虑到 CT 图像中包含了许多不需要的组织, 因此限制了图像的 HU 值在[-100,200]的范围内, 另外为增加肝脏区域对比度, 进行了 CT 图像直方图均衡化。腹部 CT 切片图像预处理前后如图 1 所示。

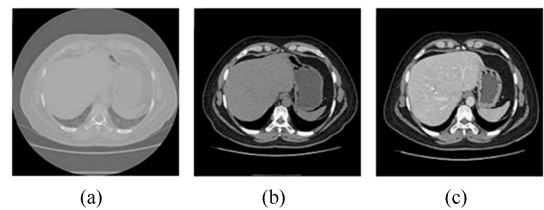


图 1 CT 图像的预处理步骤((a) 原始 CT 图像; (b) HU 阈值 CT 图像; (c) 对比度增强 CT 图像)<sup>[19]</sup>

Fig. 1 Preprocessing steps for CT images ((a) Original CT image; (b) HU threshold CT image; (c) Contrast enhanced CT images)<sup>[19]</sup>

采用多种方法对图像进行处理, 如 HE 等<sup>[20]</sup>为了消除肝脏周围其他器官或组织的影响, 在 HU 值为[-100,200]的范围内对 CT 图像进行剪切, 又将

图像重新采样到统一的分辨率,同时应用数据增强方法,以动态增加数据样本。KAGA 等<sup>[21]</sup>提出了 DLIR(Deep-Learning-based Image Reconstruction)算法用于 CT 图像降噪,与 40%的 ASiR-V 相比效果更好。即使过大的 DLIR 强度可能会丢失图像细节,但仍可作为传统方法的替代。

需要指出的是,上述方法都独立于 FLLs 检测算法,而通过对 FLLs 检测网络中的某些结构重新设计,可达到去噪的目的,并不属于图像预处理范畴,见第 2.1.1 节说明。

### 1.2.2 数据增强

在医学影像深度学习中,数据增强尤为重要,因为医学影像获取成本高,数据量有限,通过数据增强可生成更多训练数据,提高模型性能。

1) 传统形态学操作。ZHOU 等<sup>[22]</sup>采用多种数据增强方法,如随机翻转、随机旋转、亮度变换、高斯模糊和弹性变换等,对 CT 图像进行预处理。DADOUN 等<sup>[23]</sup>提出了一种基于谷歌大脑团队的边界框增强策略的数据增强方案,可以对 CT 图像进行旋转、剪切、锐化、平移、对比度和亮度调整等

操作。文献[18]通过在旋转范围参数中传递一个数字使得图像以  $0^{\circ}\sim 360^{\circ}$ 之间的任意角度随机旋转。SCHMAUCH 等<sup>[24]</sup>通过水平翻转对 CT 图像进行了增强,证明虽然肝脏不是对称器官,但该方法可略微提高性能。

2) 采用 GAN/RL 生成新的图像。近年来,越来越多的学者采用生成对抗网络(Generative Adversarial Network, GAN)等来生成新的肝脏图像以达到数据增强的目的,见表 2。FRID ADAR 等<sup>[25]</sup>使用 DCGAN(Deep Convolutional Generative Adversarial Networks)和 ACGAN(Auxiliary Conditional Generative Adversarial Networks)生成 CT 数据,该方法比传统的数据增强方法扩增的图像更有利于后续网络的训练。JIN 等<sup>[26]</sup>使用 FRGAN(Free-form Tumor Synthesis via Richer Generative Adversarial Network)生成肿瘤 CT 图像,提出了新的卷积特征增强扩张门控发生器,解决了合成过程中肿瘤与周围组织边界不确定的问题。这些方法均通过“已有病例”图像生成“新病例”的图像,扩充了网络的数据训练样本。

表 2 生成图像用于数据增强

Table 2 Generate images for data augmentation

方法	真实图像	生成图像	生成图像类型
DCGAN 和 ACGAN	CT	CT	新的病例
FRGAN	CT	CT	新的病例
cGAN	CT	PET	同一病例的不同模态
cGAN 和 cycleGAN	MRI	CT	同一病例的不同模态
Tripartite-GAN	MRI	CEMRI	注射了对比剂的图像
GRMM-GAN	MRI	CEMRI	注射了对比剂的图像
Pix-GRL	MRI	GDMRI	注射了对比剂的图像

除此之外, BEN-COHEN 和 GREENSPAN<sup>[27]</sup>使用 CGAN(Conditional Generative Adversarial Networks)通过 CT 图像生成 PET 图像; FU 等<sup>[28]</sup>使用 cGAN 和 cycleGAN(Cycle-consistent Generative Adversarial Networks)通过 MRI 图像生成 CT 图像; ZHAO 等<sup>[29]</sup>和 JIAO 等<sup>[30]</sup>使用其他技术融合 GAN 从对比剂注射前的 MR 合成对比增强的 MRI 图像。XU 等<sup>[31]</sup>提出了一种像素级图强化学习网络(Pix-GRL),可直接输入常规的非增强肝肿瘤 MRI 图像,输出增强的肝肿瘤图像,与传统的 GD 增强的 MRI 肝肿瘤图像相媲美。这些生成的图像可以与原有的标记数据关联,提高了后续网络的性能,同时避免了辐射带来的健康风险和检查费用

的增加<sup>[32]</sup>。

### 1.3 评价标准

肝脏局灶性病变目标检测算法的性能需通过以下核心指标进行量化评估:

1) 交并比与平均精度。交并比(Intersection over Union, IoU)衡量检测框与真实框的重合度,定义为

$$IoU = \frac{\text{检测框} \cap \text{真实框}}{\text{检测框} \cup \text{真实框}} \quad (1)$$

通常以阈值判定检测有效性。基于 IoU,平均精度(Mean Average Precision, mAP)综合评估模型的多类别检测能力<sup>[33]</sup>,即

$$AP = \int_0^1 P(R) dR \quad (2)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (3)$$

式中:  $P(R)$ 表示查准率-查全率曲线;  $N$  表示类别数。LiTS 挑战赛以  $mAP@0.5$  为基准推动算法横向对比<sup>[8]</sup>。

2) 敏感性(Recall)与特异性(Specificity)。敏感性反映模型检出真实病变的能力, 特异性衡量排除假阳性的效果, 其计算式为

$$Recall = \frac{TP}{TP + NP} \quad (4)$$

$$Specificity = \frac{TN}{TN + FP} \quad (5)$$

式中:  $TP$  表示真阳性;  $FN$  表示假阴性;  $TN$  表示真阴性;  $FP$  表示假阳性。二者平衡对临床误诊与漏诊风险控制至关重要。

3) F1 分数。查准率(Precision)表示模型预测为阳性的样本中实际为阳性的比例, 即

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

针对类别不平衡场景, F1 分数调和查准率与敏感性为

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

其值越接近 1, 表明模型鲁棒性越强。

4) 其他指标。临床场景中需结合假阳性率(FPs per case)等指标, 以平衡检测效率与误诊风险<sup>[34]</sup>。而在实时性场景(如术中导航)中, 推理速度(FPS)不可或缺, 需平衡精度与效率。

当前评价体系仍面临标准化不足的挑战, 例如不同研究中 IoU 阈值(如 0.50 或 0.75)的差异、数据集来源的多样性以及小病灶检测敏感度的定义不统一。未来需推动评价标准规范化, 强化多中心协作验证, 并结合动态指标, 以提升结果可比性并加速临床转化进程。

## 2 FLLs 检测中的深度学习技术

深度学习技术在 FLLs 检测中的应用已涵盖检测、分割及分类等任务, 其中本文聚焦于病灶检测, 即除分割和分类外的目标检测定位与标记任务。本章将以 CT 图像为核心分析对象, 并辅以 MRI 和 US 图像的检测方法, 系统梳理 2D 与 3D 的 FLLs 检测方法的技术演进, 为高精度、高效率的 FLLs 检测提供理论支撑。

### 2.1 2DFLLs 检测算法

基于深度学习的 2D FLLs 检测技术通过多元

化方法应对小病变敏感性低与复杂背景干扰等核心挑战。有别于数据预处理, 有些学者在检测 FLLs 时, 使用了相关算法对肝脏影像进行了前置处理, 与之后的阶段共同组成 FLLs 检测系统。基础卷积神经网络(Convolutional Neural Network, CNN)通过优化特征提取网络提升局部细节捕捉能力, 而基于循环神经网络(Recurrent Neural Network, RNN)的序列建模方法可利用时序信息增强动态特征分析。值得注意的是, 基于 Transformer 的方法通过自注意力机制捕捉全局上下文依赖, 突破卷积局部性限制。随着检测效率需求的提升, 无锚框检测器与知识蒸馏技术逐步实现轻量化部署, 而强化学习方法通过动态决策优化检测边界。从局部到全局、从静态到动态, 2D 检测技术的演进为后续 3D 与多期相扩展提供了坚实基础。深度学习技术在 2D FLLs 检测中的应用如图 2 所示。

#### 2.1.1 前置处理操作

在肝癌检测任务中, 前置处理是提升模型性能的重要环节, 其技术路径主要包括基于图像块(Patch)的局部分析方法与机器学习 2 类技术。

基于图像块的经典方法在 CT 影像分析中应用广泛: 将原始图像划分为多个小块(即 Patch), 并对每个 Patch 分别进行 CNN 处理和分析, 合并结果以获取检测输出。JAVADI 等<sup>[35]</sup>针对样本数据的特点, 将 CT 切片划分为  $64 \times 64$  像素的 50%重叠 Patch, 经窗口调整和算法增强后再输入网络进行检测, 在 3Dircadb 数据集上实现 97%精度与 78%召回率, 有效提升小病变检测鲁棒性与相邻病变区分能力。ZHENG 等<sup>[36]</sup>使用了圆形 Patch, 更贴合肝脏病变通常为圆形的特点。针对小肿瘤特征微弱问题, 部分研究结合对比学习思想, 通过图像块的语义相似性分析增强局部注意力。如 LI 等<sup>[37]</sup>提出的 PCA (Patch-Contrastive Attention)模块, 将 CT 图像划分为  $64 \times 64$  的 Patch, 利用 Siamese 网络训练判断补丁是否含肿瘤, 生成注意力图后与原图拼接输入检测头。并通过显式对比“肿瘤块”与“背景块”的相似性, 有效突出了小肿瘤区域。

此外, 一些学者采用传统机器学习算法或深度神经网络, 通过粗略提取出肝脏区域或疑似病变区域的方法进行前置处理。WEI 等<sup>[38]</sup>提出 LiLNet 系统, 联合 3D 肝脏分割网络与目标检测模块提取肝内疑似病变区域, 有效过滤肝外背景干扰, 为后续分类提供聚焦病灶的 ROI(Region of Interest)输入, 提升了复杂背景下肝病灶的检测效率。

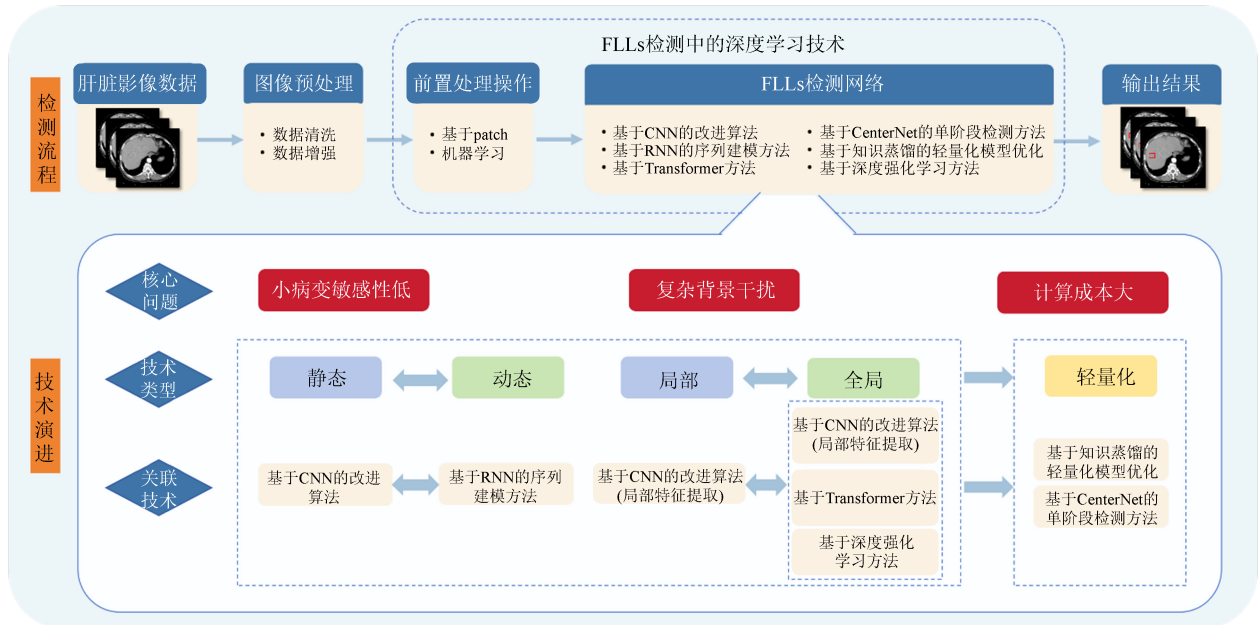


图 2 FLLs 检测中的深度学习技术

Fig. 2 Deep learning techniques in FLLs detection

NAVANEETHAKRISHNAN 等<sup>[39]</sup>使用黑洞熵模糊聚类(Black Hole Entropy Fuzzy C-Means Clustering, BHEFC)进行预处理和分割。这些方法通过排除非靶器官干扰,将检测范围约束至肝脏疑似区域,既降低了模型设计复杂度,也提升了训练效率与检测精度。

### 2.1.2 基于 CNN 的改进算法

CNN 作为广泛应用的深度学习算法,已经在目标检测等领域取得了显著的成果。CNN 凭借其强大的特征学习能力和对复杂模式的高度识别能力,在处理医学图像中的细微结构和病变方面尤为突出,为提高 FLLs 检测的准确率提供有力支撑<sup>[40]</sup>。然而,传统 CNN 受限于局部感受野与池化的空间信息损失,难以有效建模长程依赖及多尺度特征,尤其在噪声显著、病灶尺度多变的 CT 图像中面临挑战。

为更好地提取 FLLs 的特征,许多学者采用不同的方法优化了现有 CNN 特征提取过程。文献[24]通过使用 DenseNet 作为主干提取网络,并修改 Sigmoid 函数之前的局部预测应用因子,实现了对 CT 图像特征关注焦点的动态调整。然而,该方法仍受限于局部感受野,难以建模全局上下文。针对此, HUANG 等<sup>[41]</sup>引入了轻量级的非局部特征提取模块,以提取全局深度特征,提高了在 CT 图像中原发性肝癌的检测性能。此外该团队还提出了双线性核特征去噪模块,并将其集成到深度神经网络中,首次证明了局部深层特征去噪能提高网络对 CT 图像的肝癌表征性能。针对多尺度病灶检测

难题,肖宏宇等<sup>[42]</sup>设计了改进后的特征增强网络,采用多种方式的特征融合,包括自下而上、自上而下、横向直联和横向跳连等,以更有效地整合不同尺度的特征。吴德蓝<sup>[43]</sup>使用 ResNet101 和 FPN(Feature Pyramid Network)作为特征提取网络,以提取更丰富的特征信息,采用 NWD(Normalized Wasserstein Distance)作为 RPN(Region Proposal Network)中的度量指标,以更好地处理小病灶的标签分配问题。

在网络优化层面,研究者们通过结构改进提升了模型性能。KIM 等<sup>[44]</sup>提出微调结构(Fine-Tuned),通过启发式选择超参数组合,可训练参数量相比于 ResNet50, AlexNet, VGG-16 以及 Inception-ResNetV2 等传统模型大幅降低的情况下,其性能表现更优。另外,还有一些学者充分考虑 CT 图像特点,从数据输入的角度出发,提出了一些方法。文献[19]设计了一种新的 MDL-CNN 体系结构,通过将不同的 5 类图像(包含 4 种 FLLs 以及 1 种无 FLLs 图像)输入到不同层通道进行训练,以达到更好的定位和分类效果。

### 2.1.3 基于 RNN 的序列建模方法

RNN 及其衍生网络,则通过捕捉序列信息也在肝脏检测任务中得到应用,如长短期记忆网络(Long Short-Term Memory, LSTM)、分组卷积长短期记忆(Gated Convolutional Long Short-Term Memory, GCLSTM)等。LIANG 等<sup>[45]</sup>提出了一种多尺度不敏感的卷积和递归神经网络(Multi-Stream

Scale-insensitive Convolutional and Recurrent Neural Network, MSCRN)用于 CT 图像的肝肿瘤检测, 采用即插即用的 GCLSTM 模块来提取特征, 相较于先前方法表现更优。文献[39]提出了 AS-Jaya (ASO-Jaya)的优化方法对 RNN 进行了改进, 有效地更新了分类器的权重。进一步地, 将 AS-Jaya 应用于 Deep RNN 进行优化, 基于像素值更加有效地提取了特征, 提高了 CT 图像肝癌检测的效率。近期研究进一步探索多技术协同, 如 SUMELLIKA 和 PRASAD<sup>[46]</sup>提出集成模型(Combined Ensemble Model, CEM), 结合 ResNet50, RNN 与 XGBoost, 在肝癌检测中实现了 98.48%的准确率, 较传统方法提升显著。虽然 RNN 在肝脏检测中具有广泛的应用前景, 但也存在一些不足。RNN 网络的训练

过程需要更多的计算资源, 且易出现梯度消失和梯度爆炸等问题, 导致网络训练不稳定。

#### 2.1.4 基于 Transformer 方法

基于 Transformer 的方法方面, 文献[23]首次将 DETR 应用于 US 图像的 FLLs 检测, 将目标检测问题转化为一个端到端的序列到序列(Sequence-to-Sequence)问题, 并通过 Transformer 模型捕捉图像中目标的全局上下文信息, 如图 3 所示。相较于传统的目标检测算法, DETR 无需使用 RPN 或锚框(Anchor box), 直接预测图像中所有对象的类别和位置。SONG 等<sup>[47]</sup>则将 DETR 拓展至 CT 图像, 针对多期相 CT, 利用其全局建模能力直接预测病变框与类别, 检测精度较 Faster R-CNN 和 YOLO 系列显著提升。

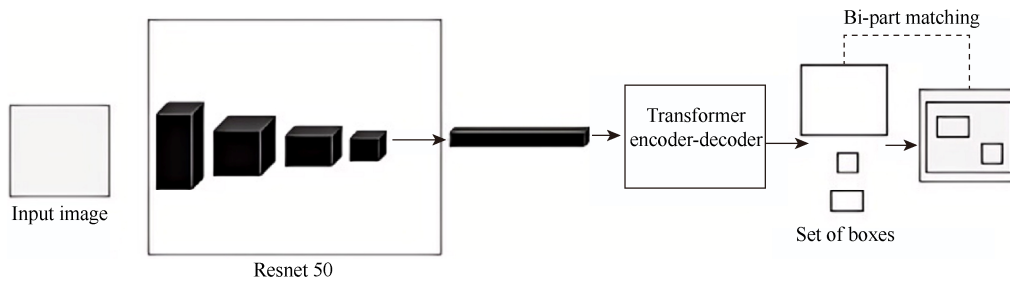


图 3 DETR 结构示意图<sup>[23]</sup>

Fig. 3 DETR structure diagram<sup>[23]</sup>

丁熠玮<sup>[48]</sup>结合 Transformer 和卷积操作, 进一步提出了 CSPTR (Cross Stage Partial Transformer) 模块, 增强了特征的表达能力, 在 Transformer 层中, 使用了多头自注意力结构, 该结构支持模型在多个子空间中并行地学习不同尺度的特征表示, 提高了模型对全局特征的捕捉能力, 同时保持了计算效率和实时性, 使得模型更适合部署在移动智能设备上。

Transformer 能捕捉全局信息, 不受卷积操作的局限性, 处理不同尺度的目标时表现出色<sup>[34]</sup>, 然而由于计算复杂性较高, 研究人员正在探索各种优化策略和改进方法, 以提高检测 Transformer 的效率和性能。

#### 2.1.5 基于 CenterNet 的单阶段检测方法

目标检测领域包括双阶段和单阶段检测 2 个重要方法。双阶段方法如 Faster R-CNN, 通过生成候选框再进行分类和精修, 检测精度高但速度慢; 单阶段方法如 YOLO 系列则直接在图像上定位和分类目标, 检测速度更快, 但其依赖密集锚框设计与后处理导致计算冗余与微小目标漏检问题。相比

YOLO 系列的单阶段检测器, CenterNet<sup>[49]</sup>提出了一种基于关键点的无锚框范式, 通过直接预测物体中心点及其尺寸, 以简化检测流程并实现端到端训练, 在保持实时性的同时显著提升了多尺度目标的适应性。

基于此优势, PENG 和 YANG<sup>[50]</sup>提出了一种基于 CenterNet 改进的 CT 肝脏肿瘤检测方法, 将 CT 图像转换为多通道数据来捕获连续性信息, 并在 DLA(Deep Layer Aggregation)主干网络中嵌入 ESE(Effective Squeeze-Excitation)通道注意力机制, 避免通道维度减少带来的信息损失。特别地, 设计了边界盒注意力机制, 通过计算预测中心点和真实框的相似度确定注意力权重, 从而减少冗余边界框, 解决了肿瘤中心点预测不准确的问题。进一步地, MA 等<sup>[51]</sup>提出的 TDCenterNet 通过密集残差注意力网络与 Transformer 模块的协同设计, 显著增强了 CenterNet 的全局特征捕捉能力, 并结合多尺度空洞卷积优化小病灶检测性能, 为单阶段无锚框检测算法在医学图像中的应用提供了创新范式。这 2 项研究凸显了 CenterNet 框架通过结构轻量化与

注意力机制融合,在平衡检测精度与效率方面的独特价值。

### 2.1.6 基于知识蒸馏的轻量化模型优化

知识蒸馏(Knowledge distillation)作为模型压缩领域的重要技术,最早由 HINTON 等<sup>[52]</sup>提出,其核心思想是通过知识迁移机制,将高精度但复杂的教师模型(Teacher model)中蕴含的“暗知识”(Dark knowledge)传递给轻量级学生模型(Student model)。可使学生网络在参数更少的情况下达到与教师网络相似甚至更好的准确性,在医学影像分析领域展现出显著优势。

WANG 等<sup>[53]</sup>基于知识蒸馏提出了用于癌症检测的 SSD-KD(Self-Supervised Diverse Knowledge

Distillation),融合多种知识辅助轻量级模型学习,并结合自监督训练策略,显著提升了轻量级学生模型的特征表达能力,其对癌症诊断至关重要。文献[51]进一步创新提出基于主动注意力图的知识蒸馏方法,如图 4 所示,采用 YOLOx 作为教师网络,TDCenterNet 作为学生网络,通过注意力转移机制实现知识蒸馏。考虑到肝脏病变在 CT 图像中常表现为病变大小悬殊、边界模糊且与周围组织对比度低的特点,该方法在教师网络的低层、中层和高层分别建立特征蒸馏路径,通过分析不同网络层激活图的特征差异,强制学生网络学习教师网络对肝脏病变关键区域的注意力分布,以有效解决传统方法在低对比度下漏检小病灶的问题。

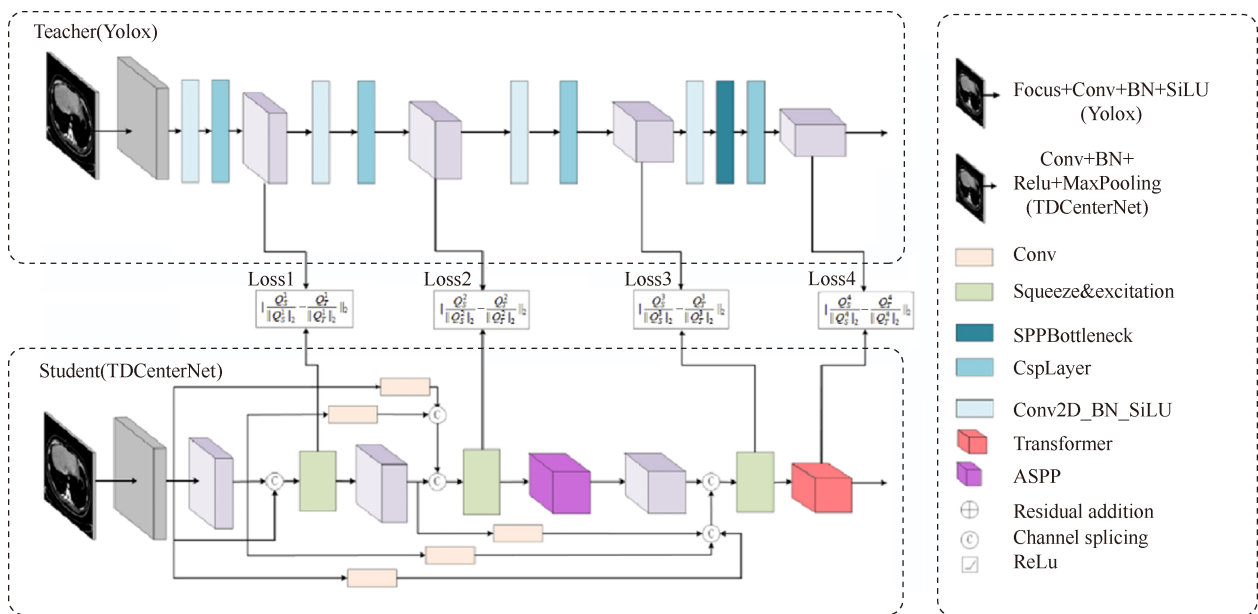


图 4 TDCenterNet 知识蒸馏示意图<sup>[51]</sup>

Fig. 4 TDCenterNet knowledge distillation diagram<sup>[51]</sup>

### 2.1.7 基于深度强化学习方法

强化学习(Reinforcement Learning, RL)是一种机器学习方法,通过智能体与环境的交互学习最优的决策策略。RL 可以与深度学习结合,通过深度神经网络来近似 RL 中的策略或价值函数,称为深度强化学习。深度强化学习与深度学习结合可提高决策制定能力,以此更好地解决复杂任务。其中 GRL (Graph Reinforcement Learning Network)是深度强化学习的一种,通过图结构转换状态空间,捕捉依赖关系和复杂非线性关系,在医学图像目标检测等任务中,可有效地处理图像中的噪声、遮挡等复杂情况,提高了图像处理的准确性和鲁棒性。

文献[31]提出了一种基于像素级图结构的强化

学习网络(Pixel-level Graph Reinforcement Learning Network, Pix-GRL),如图 5 所示,其可用于 MRI 图像分析。该网络包含 2 项改进,一是将图驱动的上下文感知代理(Graph-driven context-aware agent)模块,用于特征探索和动作获取,减少算法的复杂性;二是将双层互补奖励(Dual-level complementary reward)方法,用于代理优化,以防止代理在训练过程中陷入局部最优解。GRL 技术应用前景广阔,通过 RL 和图神经网络相结合,可以处理具有复杂关系的数据。但 GRL 仍面临如何处理具有高度变化和复杂结构的医学图像、如何解决长期依赖等问题的挑战,以及其在医学领域的应用仍需要进一步研究和探索。

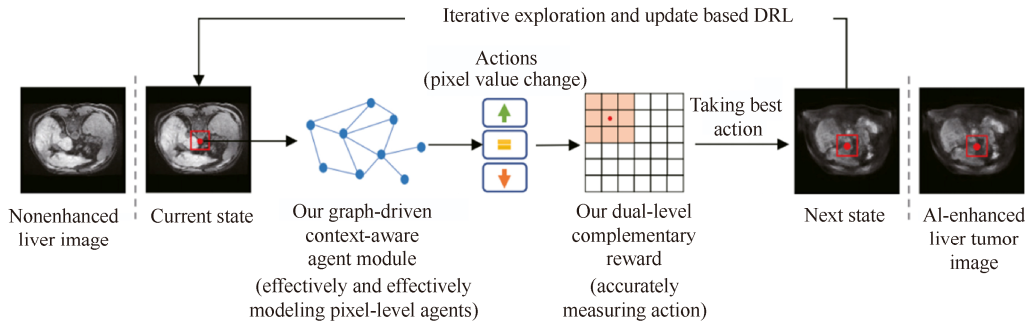


图 5 Pix-GRL 结构示意图<sup>[31]</sup>  
Fig. 5 Pix GRL structure diagram<sup>[31]</sup>

## 2.2 3D FLLs 检测算法

尽管 2D 检测通过注意力机制与轻量化设计提升了效率, 但其空间信息缺失限制了复杂病变识别。3D 方法通过体积数据分析与混合维度策略弥补此缺陷, 下文将重点讨论 3D 技术如何整合时空上下文。3D 影像技术的发展为肝脏病变检测提供了更丰富的信息维度, 已成为精准识别病变的关键所在。基于 CNN 的 3D 影像肝脏病变检测方法, 其卷积运算和多层架构有强大的特征提取与模式识别能力, 为肝脏病变检测带来了新的机遇。

### 2.2.1 基础 3D 网络架构

3D CNN 作为一种经典解决方案, 在肺结节检测中已经取得了很好的效果<sup>[54-55]</sup>。与 2D CNN 框架相比, 3D CNN 算法能够编码更为丰富的空间信息, 且借助分层结构在深度方向可提取更多具有判别性的表示。

TRIVIZAKIS 等<sup>[56]</sup>将 2D CNN 扩展到 3D, 以提高深度学习在癌症分类任务中的性能。2D 和 3D 架构均分为训练阶段和特征提取与分类两阶段, 将 Patch 和原始切片输入到训练阶段, 并分别在

2D 或 3D 架构中进行训练与分类; 另一阶段则是使用训练后的网络丢弃全连通层和输入层作为特征提取网络, 使用支持向量机进行特征提取和分类。实验结果表明, 3D 深度学习架构在肝脏肿瘤分类方面表现出色, 具有更高的准确性和 F1 分数。此外, 在面对不同来源、不同质量的图像数据时, 3D 检测模型通常能够表现出更好的适应性和稳定性。

为了实现对三维 CT 扫描数据的高效处理和肝脏病变的准确检测, CAI 等<sup>[57]</sup>提出了 VULD (Volumetric Universal Lesion Detection)模型架构, 如图 6 所示。该架构的亮点在于结合了伪 3D 卷积 (Pseudo-3D convolutions)和其他先进的深度学习技术。伪 3D 卷积通过将标准的二维卷积滤波器扩展为可处理 3 个维度(即轴向、冠状和矢状)的滤波器, 从而实现了三维数据的处理。由于伪 3D 有效地结合了二维卷积的效率和三维数据处理的必要性, 为模型提供了强大的特征提取能力, VULD 的三维处理能力使其能够更准确地捕捉和识别肝脏中的病变。

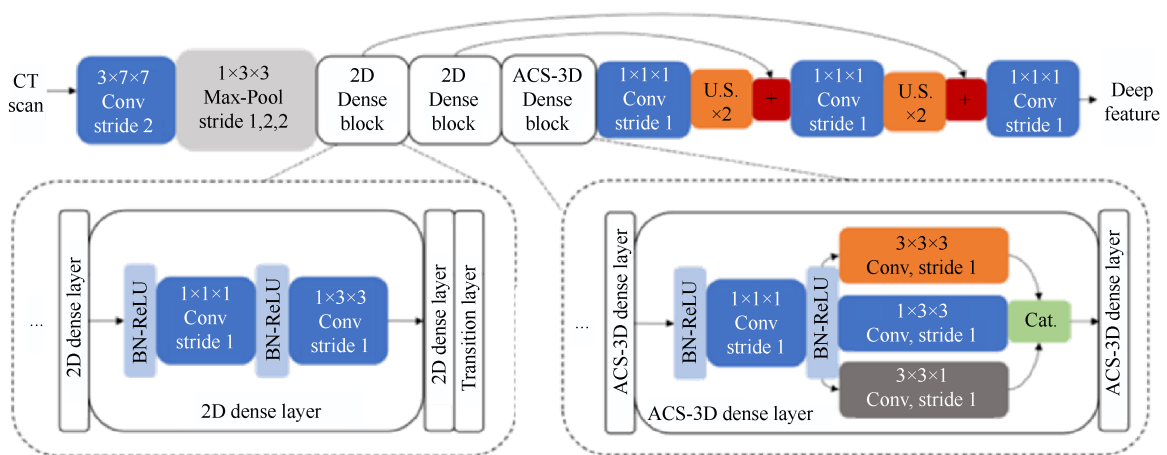


图 6 伪 3D 卷积(P3DC)主干<sup>[57]</sup>  
Fig. 6 Pseudo 3D convolution (P3DC) backbone<sup>[57]</sup>

YING 等<sup>[58]</sup>提出的多中心临床 AI 系统中的病灶检测部分由现有的 2D CSwin Transformer 扩展到 3D 版本。CSwin Transformer (Cross-Shaped Window Transformer) 是一种改进的 Transformer 变体, 通过引入特定的窗口形状来捕获图像中的局部特征。在 2D CSwin Transformer 中, 窗口形状被设计为能够覆盖图像中局部区域, 并通过自注意力机制处理区域间的关系。该团队将 3D CSwin Transformer 作为 Faster R-CNN 框架的特征提取后端, 以处理 CT 图像中的丰富 3D 上下文信息。

### 2.2.2 混合维度方法

VATS 等<sup>[59]</sup>提出一种融合 CNN 与 Transformer 的混合模型, 用于对比增强超声图像中 FLLs 的检测与分类。该模型通过 R(2+1)D 卷积模块提取 3D 时空局部特征, 将传统 3D 卷积分解为 2D 空间与 1D 时间卷积, 提升时空特征处理效率。实验中, R(2+1)D 模块联合 G1Transformer 全局注意力机制, 实现局部细节与全局结构的互补融合, 解决了 CEUS 图像中病变对比度低、边界模糊的挑战。该 3D 架构在多中心 CEUS 数据集上实现 AUC 0.823 7 和准确率 0.78, 较纯 2D CNN 模型分别提升 9.8% 和

11.4%, 且推理时间仅 52 ms, 兼顾精度与实时性。

基于 CNN 的方法在 3D 影像肝脏病变检测中得到了广泛应用, 有些学者结合 2.5D 和 3D 来实现三维的病变检测。文献[22]使用了一个包含 3 个 2.5D Faster R-CNN w/FPN 和一个 3D ResNet-18 的多模态与多尺度卷积神经网络结构, 其结合了 2.5D 和 3D 的方法, 利用 CT 影像的三期相信息进行 FLLs 的自动检测与分类。尽管该模型在检测直径较小或密度接近肝实质与血管周围病变时存在困难, 但其多模态与多尺度的结构仍为肝脏病变检测提供了新的思路。

类似地, CHEN 等<sup>[60]</sup>提出 2D 和 3D CNN 框架用于多期相增强 CT 图像的肝脏病变检测, 如图 7 所示。在 2D 检测框架中选择了 2.5D 图像输入和 Retina U-Net 算法, 在 3D 检测框架中选择了 2D CenterNet 网络的 3D 扩展, 使用焦点损失函数检测病变中心, 并使用平滑 L1 损失函数回归病变半径。实验表明与 2D CNN 相比, 3D CNN 在检测小病变 ( $< 20 \text{ mm}$ ) 时表现出更高的敏感性, 特别是在 4 个假阳性(FPs)下, 3D 网络的敏感性为 86.8%, 而 2D 网络为 75.9% ( $p < 0.001$ )。

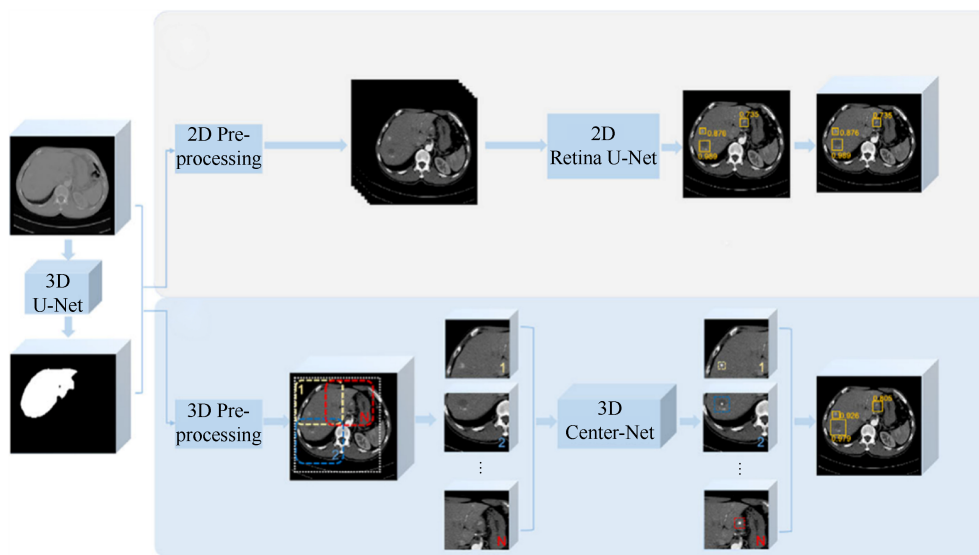


图 7 2D 和 3D 检测网络框架<sup>[60]</sup>

Fig. 7 2D and 3D detection network frameworks<sup>[60]</sup>

### 2.2.3 其他创新方法

ZHOU 等<sup>[61]</sup>将 3D 目标检测技术引入肝癌病变识别, 实现肝病变的自动化识别与纵向追踪。该研究采用自适应框架 nnDetection, 分为检测模型构建与弱监督优化两阶段: 前者基于 Retina-Unet 网络, 整合区域提议网络(RPN)与特征金字塔网络(FPN), 通过 3D 边界框回归实现病变定位; 后者创新采用

直径生成伪掩码替代全监督分割掩码, 仅需标注病变最大直径所在 2D 切片的外接圆区域, 大幅降低标注成本。实验表明, 该模型在验证集和外部测试集上分别达到 0.806 和 0.726 的均值平均精度, 并在 FROC 指标上显著优于 MedYOLO 等对比模型。该方法不仅提升了腹部 CT 中复杂病变检测的鲁棒性, 还为临床 RECIST 疗效评估提供了高一致性的

直径测量 ( $p < 1.6 \times 10^{-4}$ ), 凸显了自适应框架与弱监督学习在 3D 肝脏病变检测中的潜力。

XU 等<sup>[62]</sup>提出了一种基于 3D RPN (3D Region Proposal Network)(可以删除英文全称)的高效 CT 图像多器官定位方法, 如图 8 所示。3D RPN 结构用于生成候选的三维边界框(B-boxes), 并能够定位腹部器官中的肝脏, 有望未来改进用于定位 3D 空间中的肝脏病变检测。

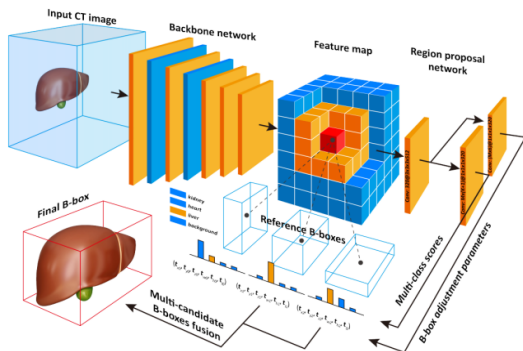


图 8 多器官定位方法示意图<sup>[62]</sup>

Fig. 8 Schematic diagram of multi organ localization method <sup>[62]</sup>

综上所述, 当前 3D 影像肝脏病变检测领域已取得了一系列研究成果, 在提高检测准确性和效率方面均取得了显著进展。然而, 面对复杂多变的肝脏病变, 现有方法仍存在一些局限性, 如对小病变的检测能力不足和对复杂病变结构的识别不够准

确等。

### 3 针对多期相图像的检测方法

多期相 CT 是一种医学成像技术, 通过多次扫描同一部位获取不同时间点的图像, 提供更全面的信息。该技术通常借助造影剂增强对比度, 主要包括平扫期(Non-Contrast Phase, NC)、动脉期(Arterial Phase, AP)、门静脉期(Portal Venous Phase, PVP)及延迟期(Delayed Phase, DP) 4 个时间段的影像采集。STOLLMAYER 等<sup>[63]</sup>研究表明, 多期相图像比单期相图像能提供更丰富的信息, 有助于提升神经网络模型解决分类问题的性能, 从而更准确地检测病变, 为临床诊断提供可靠依据。然而多期相图像的 FLLs 检测也面临挑战, 如在不同期相图像中呈现不同的形状和大小, 其位置也会有所不同, 网络需要能处理这种变化, 以准确检测病变。因此对于多期相图像的 FLLs 检测, 需要进一步研究新的方法和技术, 以提高检测的准确性和稳定性。

针对多期相医学图像的 FLLs 检测方法的改进需要从多方面入手, 本节主要从基于图像输入方式和模型架构改进来进行介绍。

#### 3.1 基于图像输入方式

基于多期相图像的输入方式主要分为串联、并联和混联 3 种, 如图 9 所示。其不直接依赖特定的 FLLs 检测算法, 而是通过在网络框架中不同的输入方式改善多期相图像的 FLLs 检测。

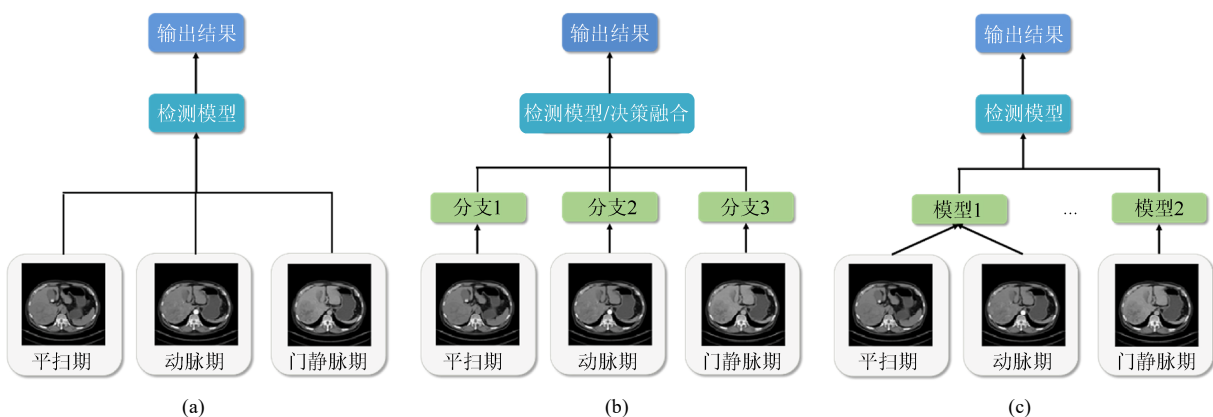


图 9 输入方式(以三期相 CT 图像为例)((a) 串联; (b) 并联; (c) 混联)

Fig. 9 Input method (taking three-phase CT images as an example) ((a) Serial connection; (b) Parallel connection; (c) Mixed union)

1) 串联方式。多期相图像按时间顺序或某种特定顺序串联, 形成一个序列, 或以多通道一起输入进网。这以此学习不同期相之间的时间依赖关系, 从而更好地捕捉病变在不同阶段的特征变化。

LEE 等<sup>[64]</sup>未针对多期相 CT 图像改进后的检测网络, 而是直接将动脉、门静脉和延迟 3 个期相图像作为 3 个通道输入, 类似于 RGB 彩色图像。CHENG 等<sup>[65]</sup>针对 VULD(Volumetric Universal Lesion

Detection)算法只能接受一个期相输入的局限性进行了改进,设计了一种灵活的深度学习 FLLs 检测算法,称为异相体积检测(Hetero Phase Volumetric Detection, HPVD)。该网络可以接受一个或任意某几个期相的三维 CT 图像作为输入。

2) 并联方式。多期相图像被分别输入到不同的分支或子网络中进行处理。每个分支或子网络独立学习每个时间点的特征表示,并通过特征融合或组合来实现综合的决策。且可充分利用每个时间点的信息,各个时间点的处理过程相互独立,相对减少了对时序关系的考虑。文献[29]提出了一种基于 CNN 的策略,用于在多期相 CT 图像中自动检测和分类肝脏病变。采用了多模态和多尺度结构,首先将 3 个期相的图像分别经过 2.5D Faster R-CNN 进行特征提取,然后将提取的特征输入到一个 3D ResNet-18 中进行病变的检测和分类。此外, KITRUNGROTSAKUL 等<sup>[66]</sup>提出了一种新的多尺度相位注意力深度层聚合网络(Multi-Scale Phase Attention Deep Layer Aggregation, MSPA-DLA++)。该方法在网络的深层阶段结合了来自不同相位的特征,而不是在网络的早期阶段直接进行合并。通过这种方式, MSPA-DLA++能够更有效地利用每个相位的特定信息,并增强对病变检测的准确性。

3) 混合融合方式。是一种结合了前两者优点的更灵活的方式。在混联输入方式中,先对部分期相图像采用串联输入方式进行初步特征提取,然后与其他期相图像采用并联输入方式送入后续的神经网络模型进行综合分析。GUO 等<sup>[67]</sup>提出了一种基于对比增强超声(Contrast-Enhanced Ultrasound,

CEUS)的两阶段学习框架,从 CEUS 视频中采用了从动脉期、门静脉期和延迟期选择的 3 个典型图像,分别组合成 A-P, PV-P 和 D-P 共 3 对图像。在第一阶段,分别对 3 对图像进行深度典型相关分析(Deep Canonical Correlation Analysis, DCCA),然后生成 6 个视图特征。而在第二阶段,将多视图特征馈送到基于多核学习(Multiple Kernel Learning, MKL)的分类器,以提供肝肿瘤的预测诊断结果。该方法可以提供更准确和高效的肝肿瘤诊断,但将其用于 CT 图像有待研究。

### 3.2 基于模型架构的改进方法

除了从多期相图像输入的角度改进算法,还可以针对模型架构本身进行改进。如采用注意力机制增强对多期相图像的处理,注意力机制可以自动学习不同时间点的重要性权重,从而使网络更加关注对任务有贡献的时间点。WANG 等<sup>[68]</sup>提出了一种基于深度学习的预测模型对多期相 CT 图像进行处理,其中包含期相内注意力(Intra-phase attention)和期相间注意力(Inter-phase attention)。期相内注意力主要关注同一期相中的特征,而期相间注意力则集中于不同期相之间的重要信息。

另有学者通过结合 CNN 与 RNN 来同时挖掘多期相图像在时间上和空间上的特征。GAO 等<sup>[69]</sup>提出了一个名为 STIC (SpatialExtractor-TemporalEncoder-Integration-Classifier)的深度学习模型, STIC 模型包含 4 个不同的模块,如图 10 所示。其中 SpatialExtractor 模块是一个深度 CNN,使用卷积层提取 CECT 图像的详细空间特征, TemporalEncoder 模块使用门控 RNN 来挖掘不同 CECT 相位之间的

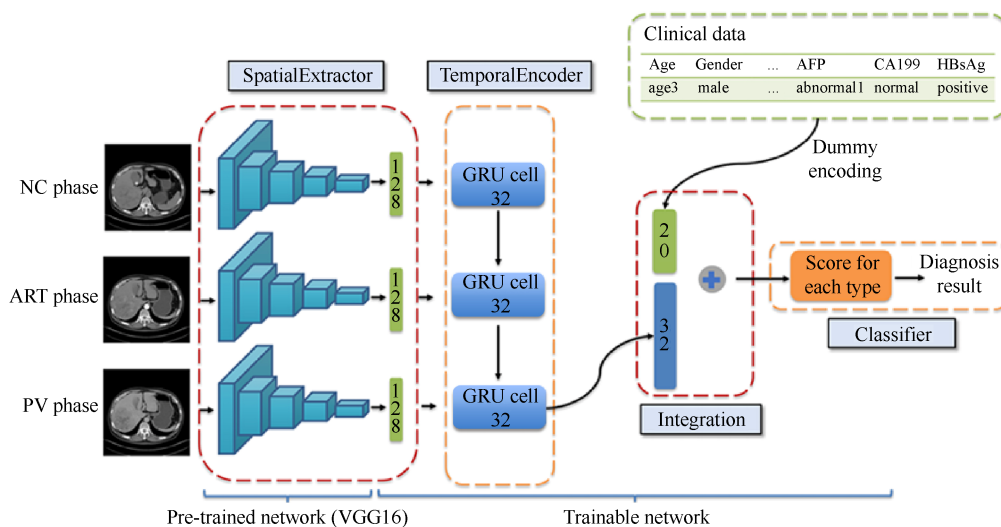


图 10 STIC 模块<sup>[69]</sup>

Fig. 10 STIC module<sup>[69]</sup>

变化模式。

此外, 迁移学习也可用于多期相图像的处理中。迁移学习通过在一个或多个期相上进行预训练, 然后将学到的知识迁移到其他期相中, 从而减少对大量标注数据的需求。CHO 等<sup>[70]</sup>评估了将单序列训练的模型应用于多序列肝脏 MRI 的可行性。同时训练了一个基于 HBP (Hepatic Arterial Phase) 的 HCC 检测网络, 适用于 T1WI (T1-Weighted Imaging) 和 PVP (Portal Venous Phase), 但因 T2WI (T2-Weighted Imaging) 上 HCC 的信号与其他期相不同, 并不适用该网络。这表明, 基于单序列 MRI 训练的计算机辅助诊断系统可以应用于其他序列图像, 特别是那些使用相似图像参数的图像。但不同期相 CT 图像 HCC 的信号均有差异, 该方法能否适用于多期相 CT 仍需进一步验证。

综上所述, 多期相图像可以通过串联或并联以及更灵活的混联方式进入网络, 同时可以应用注意力机制、结合 CNN 和 RNN 以及迁移学习等其他机制来增强处理效果, 以提高多期相图像的特征表示和分析能力, 进而为 FLLs 检测提供更准确的结果。并且多期相技术在 2D 与 3D 检测中呈现差异化应用趋势, 2D 方法通过多分支特征融合或通道串联实现期相信息整合, 而 3D 方法则利用体积数据动态分析来提升复杂病变识别能力。当前研究正逐步从 2D 单期相分析向 3D 多期相融合演进, 以充分挖掘时空维度的诊断价值。

## 4 结论和展望

深度学习在肝脏局灶性病变检测中取得显著进展。本文梳理了基于深度学习的 FLLs 检测方法, 从数据准备、算法演进到多期相图像处理进行了全面分析。在数据层面, 公开影像资源的整合与预处理方法的创新为模型训练奠定基础, 但直接适用于目标检测任务的资源仍较为稀缺。算法层面, 2D 检测通过 CNN 和 Transformer 与知识蒸馏等技术在精度与效率间取得平衡, 而 3D 检测通过空间上下文建模显著提升了复杂病变的识别能力。多期相图像处理则通过时序特征融合与动态建模进一步挖掘了医学影像的时空诊断价值。这些成果为临床辅助诊断提供了重要技术支撑。

然而, 当前深度学习在肝脏局灶性病变检测领域仍面临诸多问题。其一, 标注数据稀缺以及不同设备存在差异, 且多个研究使用单一的公开数据集缺乏外部测试, 导致数据质量与多样性不足以及模

型泛化能力受限。其二, 医学图像本身复杂多变, 算法性能存在瓶颈: 2D 检测算法缺乏对动态病理演变的建模; 3D 检测在小病变和复杂病变结构检测中存在局限, 且训练成本较高。其三, 模型缺乏可解释性且缺乏验证, 较少应用于临床实践并被患者和医学界认可。针对上述问题, 未来研究可从以下方向突破:

1) 数据驱动的技术创新。未来研究可通过开发基于联邦学习的多中心协作框架, 在严格保护患者隐私的基础上实现跨机构数据共享与联合建模, 有效缓解医疗影像数据稀缺性问题。同时融合物理驱动与扩散模型技术, 开展病理特征可控的合成数据生成, 在增强数据多样性的同时显著降低人工标注成本。此外, 可构建动态标注平台整合主动学习机制与医生实时反馈, 通过迭代优化标注质量推动医疗影像数据的标准化进程。

2) 算法设计的临床导向优化可探索 2D/3D 混合架构与神经架构搜索 (Neural Architecture Search, NAS) 等新技术结合应用, 兼顾计算效率与三维空间特征提取能力。算法研发应注重引入因果推理和元学习机制, 增强模型对多期相影像的时序关联分析能力以及对罕见病变的识别适应性。同时通过硬件-算法协同设计路径, 开发 FPGA 加速的轻量化边缘计算模型, 满足临床实时检测与便捷部署的实践需求。

3) 可解释性与临床验证体系构建。集成注意力热图、特征反演等可解释性 AI 技术, 实现模型关注区域与病理特征的直观可视化关联, 有效提升临床医生对 AI 检测结果的信任度。通过多中心前瞻性临床试验建立包含病变检出率、假阳性临床影响权重等指标的标准化评估体系, 量化验证模型对诊疗效率的实际提升效果。开发可编辑修正功能的医工协同决策系统, 构建交互反馈闭环, 促进 AI 与临床工作的融合。

4) 跨学科深度融合创新。建立“临床问题定义-算法开发-临床验证”闭环研究范式, 通过深度医工交叉合作确保技术研发与临床需求的精准对接。探索 AI 技术与光子计数 CT、超高频超声等新型成像设备的协同创新, 从数据采集源头提升病变表征能力, 形成成像技术创新与 AI 算法进步的良性互动。

## 参考文献 (References)

- [1] 施俊, 汪琳琳, 王珊珊, 等. 深度学习在医学影像中的应用综述[J]. 中国图象图形学报, 2020, 25(10): 1953-1981.

- SHI J, WANG L L, WANG S S, et al. Applications of deep learning in medical imaging: a survey[J]. *Journal of Image and Graphics*, 2020, 25(10): 1953-1981 (in Chinese).
- [2] 左艳, 黄钢, 聂生东. 深度学习在医学影像智能处理中的应用与挑战[J]. *中国图象图形学报*, 2021, 26(2): 305-315.  
ZUO Y, HUANG G, NIE S D. Application and challenges of deep learning in the intelligent processing of medical images[J]. *Journal of Image and Graphics*, 2021, 26(2): 305-315 (in Chinese).
- [3] 陈孝平. 中国肝癌诊疗发展历程[M]. 北京: 人民卫生出版社, 2021:215.  
CHEN X P. The development of diagnosis and treatment of liver cancer in China[M]. Beijing: People's Medical Publishing House, 2021:215 (in Chinese).
- [4] 韩冰, 顾劲扬. 深度学习神经网络在肝癌诊疗中的研究及应用前景[J]. *中华肝脏外科手术学电子杂志*, 2023, 12(5): 480-485.  
HAN B, GU J Y. Research and application prospect of deep learning neural network in diagnosis and treatments for liver cancer[J]. *Chinese Journal of Hepatic Surgery (Electronic Edition)*, 2023, 12(5): 480-485 (in Chinese).
- [5] LAKSHMIPRIYA B, POTTAKKAT B, RAMKUMAR G. Deep learning techniques in liver tumour diagnosis using CT and MR imaging—a systematic review[J]. *Artificial Intelligence in Medicine*, 2023, 141: 102557.
- [6] HEIMANN T, VAN GINNEKEN B, STYNER M A, et al. Comparison and evaluation of methods for liver segmentation from CT datasets[J]. *IEEE Transactions on Medical Imaging*, 2009, 28(8): 1251-1265.
- [7] SOLER L, HOSTETTLER A, AGNUS V, et al. 3D image reconstruction for comparison of algorithm database[EB/OL]. [2024-11-26]. <https://www.kiphub.com/paper/666c985cd183372eaf854a82>.
- [8] BILIC P, CHRIST P, LI H B, et al. The liver tumor segmentation benchmark (LiTS)[J]. *Medical Image Analysis*, 2023, 84: 102680.
- [9] KAVUR A E, GEZER N S, BARIŞ M, et al. CHAOS challenge-combined (CT-MR) healthy abdominal organ segmentation[J]. *Medical Image Analysis*, 2021, 69: 101950.
- [10] ANTONELLI M, REINKE A, BAKAS S, et al. The medical segmentation decathlon[J]. *Nature Communications*, 2022, 13(1): 4128.
- [11] LASSAU N, ESTIENNE T, DE VOMECOURT P, et al. Five simultaneous artificial intelligence data challenges on ultrasound, CT, and MRI[J]. *Diagnostic and Interventional Imaging*, 2019, 100(4): 199-209.
- [12] QUINTON F, POPOFF R, PRESLES B, et al. A tumour and liver automatic segmentation (ATLAS) dataset on contrast-enhanced magnetic resonance imaging for hepatocellular carcinoma[J]. *Data*, 2023, 8(5): 79.
- [13] MICCAI. TriALS 2024: triphasic-aided liver lesion segmentation challenge[EB/OL]. (2024-01-18)[2024-11-26]. <https://www.synapse.org/Synapse:syn53285416/wiki/>.
- [14] MOAWAD A W, FUENTES D, MORSHID A, et al. Multimodality annotated HCC cases with and without advanced imaging segmentation[EB/OL]. [2024-11-26]. <https://doi.org/10.7937/TCIA.5FNA-0924>.
- [15] LOU M, LIU X Q, ZHANG Y Q, et al. Liver lesion diagnosis challenge on multi-phase MRI[EB/OL]. (2023-04-18) [2024-11-26]. <https://zenodo.org/records/7852363>.
- [16] WU C Y, ZHANG X M, ZHANG Y, et al. Towards generalist foundation model for radiology by leveraging web-scale 2D&3D medical data[EB/OL]. [2024-11-25]. <https://arxiv.org/abs/2308.02463>.
- [17] WILMAN H R, KELLY M, GARRATT S, et al. Characterisation of liver fat in the UK Biobank cohort[J]. *PLoS One*, 2017, 12(2): e0172921.
- [18] HAMEED U, UR REHMAN M, REHMAN A, et al. A deep learning approach for liver cancer detection in CT scans[J]. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2024, 11(7): 2280558.
- [19] SHAH S, MISHRA R, SZCZUROWSKA A, et al. Non-invasive multi-channel deep learning convolutional neural networks for localization and classification of common hepatic lesions[J]. *Polish Journal of Radiology*, 2021, 86: 440-448.
- [20] HE K, LIU X M, SHAHZAD R, et al. Advanced deep learning approach to automatically segment malignant tumors and ablation zone in the liver with contrast-enhanced CT[J]. *Frontiers in Oncology*, 2021, 11: 669437.
- [21] KAGA T, NODA Y, FUJIMOTO K, et al. Deep-learning-based image reconstruction in dynamic contrast-enhanced abdominal CT: image quality and lesion detection among reconstruction strength levels[J]. *Clinical Radiology*, 2021, 76(9): 710.e15-710.e24.
- [22] ZHOU J R, WANG W Z, LEI B W, et al. Automatic detection and classification of focal liver lesions based on deep convolutional neural networks: a preliminary study[J]. *Frontiers in Oncology*, 2021, 10: 581210.
- [23] DADOUN H, ROUSSEAU A L, DE KERVILER E, et al. Deep learning for the detection, localization, and characterization of focal liver lesions on abdominal US images[J]. *Radiology: Artificial Intelligence*, 2022, 4(3): e210110.
- [24] SCHMAUCH B, HERENT P, JEHANNO P, et al. Diagnosis of focal liver lesions from ultrasound using deep learning[J]. *Diagnostic and Interventional Imaging*, 2019, 100(4): 227-233.
- [25] FRID ADAR M, DIAMANT I, KLANG E, et al. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification[J]. *Neurocomputing*, 2018, 321: 321-331.
- [26] JIN Q G, CUI H, SUN C M, et al. Free-form tumor synthesis in computed tomography images via richer generative adversarial network[J]. *Knowledge-Based Systems*, 2021, 218: 106753.
- [27] BEN-COHEN A, GREENSPAN H. Liver lesion detection in CT using deep learning techniques[M]//ZHOU S K, RUECKERT D, FICHTINGER G. *Handbook of Medical Image Computing and Computer Assisted Intervention*. New York: Academic Press, 2020: 65-90.
- [28] FU J, SINGHRAO K, CAO M S, et al. Generation of abdominal synthetic CTs from 0.35T MR images using generative adversarial networks for MR-only liver radiotherapy[J]. *Biomedical Physics & Engineering Express*, 2020, 6(1): 015033.
- [29] ZHAO J F, LI D W, KASSAM Z, et al. Tripartite-GAN: synthesizing liver contrast-enhanced MRI to improve tumor detection[J]. *Medical Image Analysis*, 2020, 63: 101667.
- [30] JIAO C Z, LING D N, BIAN S, et al. Contrast-enhanced liver magnetic resonance image synthesis using gradient regularized multi-modal multi-discrimination sparse attention fusion GAN[J]. *Cancers*, 2023, 15(14): 3544.
- [31] XU C C, ZHANG D, CHONG J, et al. Synthesis of gadolinium-enhanced liver tumors on nonenhanced liver MR images using pixel-level graph reinforcement learning[J]. *Medical Image Analysis*, 2021, 69: 101976.
- [32] 张颖麟, 胡衍, 东田理沙, 等. 生成对抗式网络及其医学影像应用研究综述[J]. *中国图象图形学报*, 2022, 27(3): 687-703.  
ZHANG Y L, HU Y, TOSHISAWA R, et al. A review of generative adversarial networks and the application in medical image[J]. *Journal of Image and Graphics*, 2022, 27(3): 687-703

- (in Chinese).
- [33] EVERINGHAM M, VAN GOOL L, WILLIAMS C K I, et al. The PASCAL visual object classes (VOC) challenge[J]. *International Journal of Computer Vision*, 2010, 88(2): 303-338.
- [34] KIM D W, LEE G, KIM S Y, et al. Deep learning-based algorithm to detect primary hepatic malignancy in multiphase CT of patients at high risk for HCC[J]. *European Radiology*, 2021, 31(9): 7047-7057.
- [35] JAVADI A, FORUZAN A H, CHEN Y W. Accurate delineation of multiple CT liver lesions by deep processing of image patches[C]//The 7th International Conference on Machine Learning and Machine Intelligence. New York: ACM, 2024: 235-239.
- [36] ZHENG R C, WANG L N, WANG C Y, et al. Feasibility of automatic detection of small hepatocellular carcinoma ( $\leq 2$  cm) in cirrhotic liver based on pattern matching and deep learning[J]. *Physics in Medicine & Biology*, 2021, 66(8): 085014.
- [37] LI X Y, XIAO H, WENG Z P, et al. PCA-YOLO: a small liver tumor detection model with patch-contrastive attention [EB/OL]. [2024-11-26]. <https://openreview.net/pdf?id=iEBZjNZ63T>. *Medical Imaging with Deep Learning*.
- [38] WEI Y, YANG M Y, ZHANG M, et al. Focal liver lesion diagnosis with deep learning and multistage CT imaging[J]. *Nature Communications*, 2024, 15(1): 7040.
- [39] NAVANEETHAKRISHNAN M, VAIRAMUTHU S, PARTHASARATHY G, et al. Atom search-jaya-based deep recurrent neural network for liver cancer detection[J]. *IET Image Processing*, 2021, 15(2): 337-349.
- [40] 张雪怡, 孙博, 陈癸霖, 等. 计算机视觉在医学图像中的应用及其对罕见肌肉骨骼疾病的影响[J]. *罕见病研究*, 2023, 2(4): 589-595.  
ZHANG X Y, SUN B, CHEN G L, et al. Computer vision in medical imaging and its impact on the rare musculoskeletal diseases[J]. *Journal of Rare Diseases*, 2023, 2(4): 589-595 (in Chinese).
- [41] HUANG H Y, XIE Y Y, WANG G Y, et al. DLNLF-net: denoised local and non-local deep features fusion network for malignancy characterization of hepatocellular carcinoma[J]. *Computer Methods and Programs in Biomedicine*, 2022, 227: 107201.
- [42] 肖宏宇, 杨伟东, 王琦. 基于深度学习的双期相 CT 肝癌检测算法[J]. *光电子·激光*, 2025, 36(6): 664-672.  
XIAO H Y, YANG W D, WANG Q. Dual-phase CT liver cancer detection algorithm based on deep learning[J]. *Journal of Optoelectronics·Laser*, 2025, 36(6): 664-672 (in Chinese).
- [43] 吴德蓝. 基于改进 Faster R-CNN 的肝脏 CT 图像小病灶检测 [D]. 南宁: 广西大学, 2022.  
WU D L. Detection of small lesions in liver CT images based on improved Faster R-CNN[D]. Nanning: Guangxi University, 2022 (in Chinese).
- [44] KIM J, MIN J H, KIM S K, et al. Detection of hepatocellular carcinoma in contrast-enhanced magnetic resonance imaging using deep learning classifier: a multi-center retrospective study[J]. *Scientific Reports*, 2020, 10(1): 9458.
- [45] LIANG D, LIN L F, CHEN X, et al. Multi-stream scale-insensitive convolutional and recurrent neural networks for liver tumor detection in dynamic CT images[C]//2019 IEEE International Conference on Image Processing. New York: IEEE Press, 2019: 794-798.
- [46] SUMALLIKA T, PRASAD R S. A combined ensemble model (CEM) for a liver cancer detection system[J]. *International Journal of Advanced Computer Science and Applications*, 2024, 15(2): 163-172.
- [47] SONG J, HU Y C, ZHANG J H, et al. Detection of focal liver lesions in CT images using a transformer-based end-to-end detection model[C]//2024 IEEE 13th Global Conference on Consumer Electronics. New York: IEEE Press, 2024: 1061-1064.
- [48] 丁熠玮. 基于深度学习的 CT 图像肝脏肿瘤分析[D]. 南京: 东南大学, 2021.  
DING Y W. DEEP Learning based liver lesion analysis on CT images[D]. Nanjing: Southeast University, 2021 (in Chinese).
- [49] ZHOU X Y, WANG D Q, KRÄHENBÜHL P. Objects as points[EB/OL]. [2024-11-26]. <https://doi.org/10.48550/arXiv.1904.07850>.
- [50] PENG X F, YANG X W. Liver tumor detection based on objects as points[J]. *Physics in Medicine & Biology*, 2021, 66(23): 235009.
- [51] MA J L, OUYANG K, MA Z P, et al. Transformer dense center network for liver tumor detection[J]. *Biomedical Signal Processing and Control*, 2024, 91: 106066.
- [52] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network[EB/OL]. [2024-11-26]. <https://doi.org/10.48550/arXiv.1503.02531>.
- [53] WANG Y W, WANG Y H, CAI J Y, et al. SSD-KD: a self-supervised diverse knowledge distillation method for lightweight skin lesion classification using dermoscopic images[J]. *Medical Image Analysis*, 2023, 84: 102693.
- [54] GONG L, JIANG S, YANG Z Y, et al. Automated pulmonary nodule detection in CT images using 3D deep squeeze-and-excitation networks[J]. *International Journal of Computer Assisted Radiology and Surgery*, 2019, 14(11): 1969-1979.
- [55] ZHU W T, LIU C C, FAN W, et al. DeepLung: deep 3D dual path nets for automated pulmonary nodule detection and classification[C]//2018 IEEE Winter Conference on Applications of Computer Vision. New York: IEEE Press, 2018: 673-681.
- [56] TRIVIZAKIS E, MANIKIS G C, NIKIFORAKI K, et al. Extending 2-D convolutional neural networks to 3-D for advancing deep learning cancer classification with application to MRI liver tumor differentiation[J]. *IEEE Journal of Biomedical and Health Informatics*, 2019, 23(3): 923-930.
- [57] CAI J Z, YAN K, CHENG C T, et al. Deep volumetric universal lesion detection using light-weight pseudo 3D convolution and surface point regression[C]//The 23rd International Conference on Medical Image Computing and Computer Assisted Intervention. Cham: Springer, 2020: 3-13.
- [58] YING H N, LIU X Q, ZHANG M, et al. A multicenter clinical AI system study for detection and diagnosis of focal liver lesions[J]. *Nature Communications*, 2024, 15(1): 1131.
- [59] VATS S, SINGH C R, MEHTA S. Next-generation diagnostic tools: the role of hybrid CNN-transformer models in liver cancer detection[C]//The 15th International Conference on Computing Communication and Networking Technologies. New York: IEEE Press, 2024: 1-5.
- [60] CHEN Q Q, ZHU Y J, CHEN Y N, et al. Applicability of multidimensional convolutional neural networks on automated detection of diverse focal liver lesions in multiphase CT images[J]. *Medical Physics*, 2023, 50(5): 2872-2883.
- [61] ZHOU J, XIA Y J, XUN X L, et al. Deep learning-based detect-then-track pipeline for treatment outcome assessments in immunotherapy-treated liver cancer[J]. *Journal of Imaging Informatics in Medicine*, 2025, 38(1): 380-393.
- [62] XU X A, ZHOU F G, LIU B, et al. Efficient multiple organ localization in CT image using 3D region proposal network[J].

- IEEE Transactions on Medical Imaging, 2019, 38(8): 1885-1898.
- [63] STOLLMAYER R, BUDAI B K, TÓTH A, et al. Diagnosis of focal liver lesions with deep learning-based multi-channel analysis of hepatocyte-specific contrast-enhanced magnetic resonance imaging[J]. World Journal of Gastroenterology, 2021, 27(35): 5978-5988.
- [64] LEE G, KIM J, LEE J G, et al. Automatic hepatocellular carcinoma lesion detection with dynamic enhancement characteristic from multi-phase CT images[C]//International Forum on Medical Imaging in Asia 2019. Singapore: SPIE, 2019: 203-208.
- [65] CHENG C T, CAI J Z, TENG W, et al. A flexible three-dimensional heterophase computed tomography hepatocellular carcinoma (HCC)detection algorithm for generalizable and practical HCCscreening[J]. Hepatology Communications, 2022, 6(10): 2901-2913.
- [66] KITRUNGROTSAKUL T, XU Y Y, CHEN Q Q, et al. MSPA-DLA++: a multi-scale phase attention deep layer aggregation for lesion detection in multi-phase CT images[EB/OL]. [2025-01-29]. <https://openreview.net/forum?id=8m8XbVvgOm>.
- [67] GUO L H, WANG D, QIAN Y Y, et al. A two-stage multi-view learning framework based computer-aided diagnosis of liver tumors with contrast enhanced ultrasound images[J]. Clinical Hemorheology and Microcirculation, 2018, 69(3): 343-354.
- [68] WANG W B, WANG F, CHEN Q Q, et al. Phase attention model for prediction of early recurrence of hepatocellular carcinoma with multi-phase CT images and clinical data[J]. Frontiers in Radiology, 2022, 2: 856460.
- [69] GAO R T, ZHAO S, AISHANJIANG K, et al. Deep learning for differential diagnosis of malignant hepatic tumors based on multi-phase contrast-enhanced CT and clinical data[J]. Journal of Hematology & Oncology, 2021, 14(1): 154.
- [70] CHO Y, HAN Y E, KIM M J, et al. Computer-aided hepatocellular carcinoma detection on the hepatobiliary phase of gadoteric acid-enhanced magnetic resonance imaging using a convolutional neural network: feasibility evaluation with multi-sequence data[J]. Computer Methods and Programs in Biomedicine, 2022, 225: 107032.