

# 基于颈部重构与双重感知的CT影像肺结节检测

张文倩<sup>a</sup>, 姚为<sup>a</sup>, 陈笑笑<sup>a</sup>, 刘皇池<sup>a</sup>, 徐胜舟<sup>ab\*</sup>

(中南民族大学 a. 计算机学院; b. 湖北省制造企业智能管理工程技术研究中心, 武汉 430074)

**摘要** 肺结节作为肺癌早期的重要影像学表现, 因尺寸微小且边界模糊, 检测过程中难以被有效捕获, 现有方法对微小结构与边界细节的表征不足, 容易导致漏检和误检. 为此, 提出了一种基于颈部重构与双重感知的肺结节检测模型(RDNet). 具体而言, 针对原颈部结构信息传递不足的问题, 重构颈部以补充浅层空间细节并减少微小结节的信息损失; 同时, 为减少检测头间冗余与计算开销, 设计了参数共享检测头(PSDH), 实现多尺度检测输出的一致性表达; 此外, 为增强模型对肺结节特征的刻画能力, 提出双重感知融合模块(DPFM), 集成多尺度语义信息与结构细节, 有效提升了特征表征能力. 在LUNA16数据集上, RDNet将AFROC提升至1.824, Recall提升8.33%, 在多个评价指标上均优于现有检测方法, 展现出良好的检测性能与实用价值. 此外, 天池数据集上的实验结果显示, RDNet具有良好的跨数据集泛化能力.

**关键词** 肺结节; 目标检测; 颈部重构; 双重感知

中图分类号 TP391.4 文献标志码 A 文章编号 1672-4321(2026)03-0373-10

doi:10.20056/j.cnki.ZNMDZK.20250864

## Lung nodule detection in computed tomography imaging based on neck redesign and dual perception

ZHANG Wenqian<sup>a</sup>, YAO Wei<sup>a</sup>, CHEN Xiaoxiao<sup>a</sup>, LIU Huangchi<sup>a</sup>, XU Shengzhou<sup>ab\*</sup>

(South-Central Minzu University, a. College of Computer Science; b. Hubei Provincial Engineering Research Center for Intelligent Management of Manufacturing Enterprises, Wuhan 430074, China)

**Abstract** Lung nodules, as important imaging manifestations in the early stage of lung cancer, are difficult to be effectively detected during the detection process due to their small size and blurred boundaries. Existing methods have insufficient representation of fine structures and boundary details, which tends to result in missed detections and false positives. To address limitations in small-object perception and feature representation, a novel detection model based on neck redesign and dual perception is proposed, termed RDNet. Specifically, to address the insufficient information transfer in the original neck structure, the neck is redesigned to supplement shallow spatial details and reduce information loss for small nodules. Parameter-Sharing Detection Head (PSDH) is introduced to reduce redundancy and computational cost while ensuring consistent representation of multi-scale detection outputs. Additionally, a Dual-Perception Fusion Module (DPFM) integrates multi-scale semantic features with structural details is presented to enhance feature representation ability of the model. Experimental results on the LUNA16 show that RDNet achieves an AFROC of 1.824 and increases Recall by 8.33%, outperforming existing methods across multiple metrics. These results confirm the effectiveness and value of the proposed method. Furthermore, experimental results on the Tianchi demonstrate that RDNet has strong cross-dataset generalization capability.

**Keywords** lung nodule; object detection; neck redesign; dual perception

肺癌是全球发病率和死亡率最高的恶性肿瘤之一. 根据2022年全球癌症统计数据<sup>[1]</sup>, 肺癌在所

收稿日期 2025-07-23

\* 通信作者 徐胜舟(1982-), 男, 副教授, 研究方向: 医学影像处理与分析, E-mail: xushengzhou@scuec.edu.cn

基金项目 湖北省自然科学基金资助项目(2025AFB688); 中南民族大学中央高校基本科研业务费专项资金资助项目(CZZ25005)

有癌症类型中同时位居发病率和死亡率首位,严重威胁人类生命健康.肺结节作为肺癌早期的重要影像学表现,其高效、准确的自动化检测对于提升早期筛查效率和改善患者预后具有重要的临床价值<sup>[2]</sup>.

随着低剂量计算机断层扫描(Computed Tomography, CT)技术的广泛应用,肺结节的早期筛查变得更加可行且有效,极大地推动了肺癌的早期诊断和干预<sup>[3]</sup>.尽管如此,肺结节在CT图像中呈现出高度的异质性和复杂性,给自动化检测带来了诸多挑战.首先,肺结节的尺寸跨度大,从仅数毫米的微小病灶到数厘米的实体肿块均可能出现,其中微小结节由于体积小、纹理模糊、边界不清晰,极易被漏检<sup>[4]</sup>;其次,肺结节的形态多变,可呈球形、椭圆形、扁平状,部分还伴随钙化、空洞、毛刺等影像特征<sup>[5]</sup>,对模型的结构感知能力提出更高的要求;第三,肺结节所处的组织环境复杂,往往与血管、胸壁等结构相邻,导致目标边界模糊、对比度低,进一步增加了检测难度.此外,CT图像还存在图像质量不一、层厚差异、伪影干扰等问题,也对检测算法的稳定性和泛化能力提出挑战<sup>[6]</sup>.

因此,肺结节检测任务不仅要求模型具备多尺度建模能力以覆盖不同尺寸的病灶,还需具备结构感知能力以刻画边缘与空间关系,并兼顾轻量化以满足部署需求.YOLOv8<sup>[7]</sup>作为单阶段检测算法,在实时性与精度间取得了良好平衡,并已广泛应用于目标检测任务,因此本文将作为基线模型.然而,YOLOv8在肺结节检测中仍面临以下问题:颈部结构在特征融合过程中对浅层信息利用不足,难以充分捕捉微小结节特征;多尺度检测头相互独立,易产生冗余计算和跨尺度不一致;骨干网络与颈部结构中的卷积块对边缘与结构细节的建模能力有限.基于上述问题,本文在YOLOv8框架下提出一种基于颈部重构与双重感知的肺结节检测模型(Neck Redesign and Dual Perception Network, RDNet),其核心贡献包括:

(1)颈部重构:引入更高分辨率输入以增强浅层空间细节,显著提升对微小肺结节的定位能力;原结构的下采样会削弱小目标的有效表征,因此去除了下采样以降低信息损失;

(2)构建参数共享检测头(Parameter-Sharing Detection Head, PSDH):通过在多尺度分支间共享检测头参数,减少冗余结构与重复计算,显著压缩模型复杂度,并提升不同尺度肺结节的预测一致性与稳定性;

(3)设计双重感知融合模块(Dual-Perception Fusion Module, DPFM):由多尺度感知分支(Multi-Scale Perception Branch, MP)与结构感知分支(Structure Perception Branch, SP)组成,前者聚焦于提取丰富的语义上下文信息,后者通过引入空间边缘注意力机制(Spatial-Edge Attention, SEA),强化对形状与边缘细节的感知.两个分支互补协同,促进多尺度语义与结构特征的深度融合,从而提升对肺结节的判别能力与表达能力.

## 1 相关工作

当前肺结节检测研究主要聚焦于两种方法:基于卷积神经网络(Convolutional Neural Network, CNN)的方法和基于Transformer网络的方法<sup>[8]</sup>.CNN作为医学影像分析领域的传统技术,始终占据重要地位.例如,AGNES等<sup>[9]</sup>引入了一个结合增强型U-Net与卷积LSTM的两阶段框架,通过整合2D和3D特征来提升时空一致性.URREHMAN等<sup>[10]</sup>提出了一种双注意力CNN,利用通道和空间注意力机制增强特征表示.TANG等<sup>[11]</sup>提出了Circle-YOLO,这是一种无锚点检测器,采用圆形表征更好地适配肺结节.通过CSPA和改进的BCIoU损失函数,该模型在精度和速度上均得到提升.基于Transformer的模型中,MKINDU等<sup>[12]</sup>提出的诊断模型结合了贝叶斯优化与视觉Transformer,利用滑动窗口机制和超参数微调来提高胸部CT检测效果.CUI等<sup>[13]</sup>提出了SF2T框架,将3D Swin Transformer与特征金字塔结合用于候选结节生成,并采用创新的3D加权运动历史图像方法减少假阳性.TANG等<sup>[14]</sup>提出的LN-DETR是基于Transformer的肺结节检测模型,整合了PC-EMA实现多尺度特征提取、GS-CCFM进行跨尺度融合以及CTrans增强特征集成.

尽管近年来深度学习技术在肺结节检测任务中取得了显著进展,但在面向实际临床场景的肺结节检测中仍面临诸多挑战.一方面,现有模型在处理极小尺寸或边缘模糊的结节时,仍存在响应不足的问题,导致检测灵敏度下降;另一方面,主流多尺度融合机制在不同层级特征的整合过程中容易发生信息冗余或特征弱化,影响整体检测的一致性.此外,传统的卷积结构对肺结节复杂形态和结构信息的建模能力有限,难以精准捕捉其判别性细节.

## 2 方法

### 2.1 基于颈部重构与双重感知的肺结节检测模型

RDNet 整体架构如图 1 所示,由骨干网络、颈部结构和检测头组成.骨干网络通过 DPFM 和下采样提取关键特征,并使用 Conv 模块生成深层表征.颈

部结构结合 DPFM 与上采样实现多尺度特征融合,同时利用 SPPF(Spatial Pyramid Pooling-Fast)扩展感受野并增强全局信息.检测头通过 PSDH 对肺结节的位置与类别进行预测,同时降低计算复杂度并保证输出一致性.

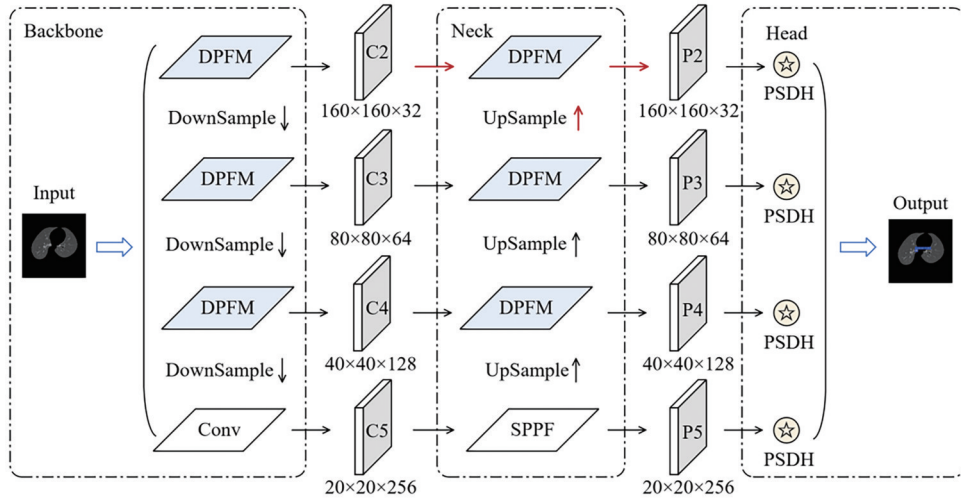


图 1 RDNet 结构图

Fig. 1 The structure of RDNet

#### 2.1.1 颈部重构

YOLOv8 的颈部结构以图 1 骨干网络生成的 C3、C4 和 C5 为输入,首先对最深层的 C5 进行 SPPF 处理,再逐级上采样,依次与 C4、C3 拼接融合,得到多尺度特征图 P5、P4 和 P3.随后对 P3 连续两次下采样,并分别与 P4、P5 融合,最终输出分辨率为  $80 \times 80$ 、 $40 \times 40$ 、 $20 \times 20$  的 3 个特征图.然而,受限于  $80 \times 80$  的最高输入分辨率,YOLOv8 的颈部结构在肺结节检测任务中难以有效保留微小结节的结构细节,极易造成漏检;而后续对 P3 的两次下采样还会进一步削弱浅层细节信息,降低模型对微小结构的敏感度.为此,本文对颈部结构进行了重构(见图 1),在原结构基础上引入更高分辨率的特征图 C2,将 C3 经 DPFM 处理后再上采样并与 C2 融合,生成空间细节更丰富的 P2,从而增强对小结节的定位能力,这一过程如图 1 中红色箭头所示.同时去除对 P3 的两次下采样操作,以减少额外的信息损耗.

#### 2.1.2 参数共享检测头(PSDH)

在多尺度目标检测任务中,传统检测头通常为每个尺度单独设计卷积分支,如图 2(a)所示,以适应不同层级特征的语义差异.这种结构虽具备一定的尺度适配能力,但也带来了大量参数冗余和跨尺度表达不一致的问题.尤其在肺结节检测中,肺结节

本身尺度就小、形态多变,往往只在特定尺度具有清晰表达,而传统检测头结构中各尺度间缺乏统一的表征机制,容易造成信息分散,进而引发小结节的误检与漏检,影响整体检测的稳定性和泛化能力.

为缓解上述问题,本文提出 PSDH(如图 2(b)所示),通过参数共享与通道对齐,在不同尺度间建立一致的建模路径.具体地,PSDH 接受来自颈部的 4 个不同尺度的特征图 P2、P3、P4、P5,其通道数分别为  $C_{in}^1$ 、 $C_{in}^2$ 、 $C_{in}^3$ 、 $C_{in}^4$ .首先,为确保多尺度特征的一致性,所有特征图首先经过集成组归一化(Group Normalization, GN)的  $1 \times 1$  卷积,统一至相同通道数  $C_{inp}$ ,GN 在小批量训练环境下表现稳定,有助于提升模型性能<sup>[15]</sup>;随后,统一后的特征图  $F'_2$ 、 $F'_3$ 、 $F'_4$ 、 $F'_5$  依次通过两层  $3 \times 3$  共享卷积模块,在所有尺度特征间实现显式权重共享,即各尺度分支共用同一组卷积参数,避免传统检测头每层独立建模带来的参数冗余,并促进跨层特征学习.在输出阶段,所有尺度特征图统一通过  $1 \times 1$  的共享卷积,分别完成分类与回归任务,以实现跨尺度的一致性参数学习.为增强尺度敏感性,在回归分支中引入可学习的尺度因子 Scale,自适应调整不同尺度特征的响应强度.

为量化检测头结构在参数效率上的差异,本文以实验中的具体配置为例进行对比分析.设  $k_1$  和  $k_3$

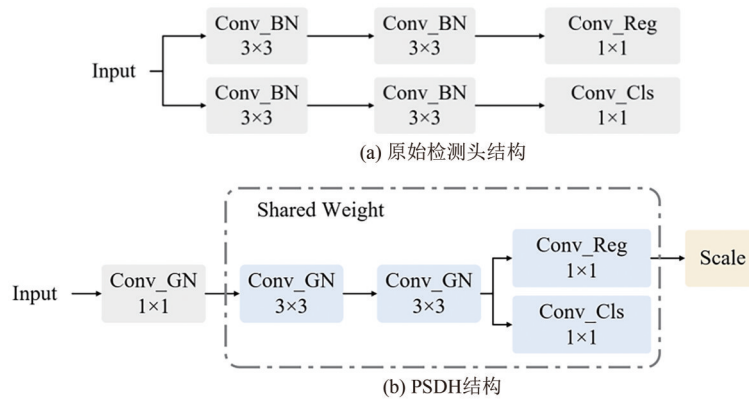


图2 原始检测头与PSDH示意图

Fig. 2 Diagram of original detection head and PSDH

分别表示卷积核大小为1和3,对于传统检测头,每个分支的中间通道  $C_{\text{dp}}=64$ ,输出通道  $C_{\text{out}}=256$ ,对于PSDH,中间通道和输出通道均为  $C_{\text{hid}}$ ,则传统检测头与PSDH的总参数量分别近似为:

$$P_{\text{baseline}} \propto \sum_{i=1}^4 [k_3^2 \times C_{\text{in}}^i \times (C_{\text{hid}} + C_{\text{out}}) + k_3^2 \times (C_{\text{hid}}^2 + C_{\text{out}}^2) + k_1^2 \times C_{\text{hid}}^2 + k_1^2 \times C_{\text{out}} \times 1], \quad (1)$$

$$P_{\text{PSDH}} \propto \sum_{i=1}^4 (k_1^2 \times C_{\text{in}}^i \times C_{\text{hid}} + 2 \times k_3^2 \times C_{\text{hid}}^2 + k_1^2 \times C_{\text{hid}}^2 + k_1^2 \times C_{\text{hid}} \times 1) \quad (2)$$

通过对比计算,PSDH相比传统检测头实现了91.2%的参数压缩率,同时加快推理速度,有效控制计算资源消耗.

### 2.1.3 双重感知融合模块(DPFM)

肺结节通常具有模糊边界、不规则形态,尤其在低剂量CT图像中,与周围组织的对比度较低.传统卷积网络由于感受野固定、结构建模能力有限,往往难以有效提取结节区域的特征信息,导致对肺结节的识别准确率较低.为缓解上述问题,本文设计了DPFM,用于联合建模多尺度语义信息与结构细节,提高模型对肺结节的判别能力与稳定性. DPFM的整体结构如图3所示,首先通过  $1 \times 1$  卷积对输入特征进行通道调整;随后构建两个并行分支:多尺度感知分支与结构感知分支,两者共同构成双重感知.最后通过残差连接进行融合,在强化结构表达的同时保持特征的完整性,促进梯度高效传播.

多尺度感知分支(如图3左侧所示)主要面向肺结节尺寸多变所带来的感受野适应性问题.为增强模型对不同尺度目标的响应能力,该分支引入并行的深度可分离卷积结构,分别采用  $3 \times 3$  与  $7 \times 7$  的可重参数化卷积核,以捕捉不同空间尺度下的语义特征.在训练阶段,卷积与批归一化(Batch

Normalization, BN)联合建模;推理阶段则采用重参数化策略<sup>[16]</sup>将其融合为等效的单一卷积结构,有效降低推理计算的复杂度;同时,特征融合过程中引入加权求和机制与SiLU激活函数,进一步提升非线性表达能力,强化对多尺度语义差异的建模效果.结构感知分支(如图3右侧所示)主要针对结节边界模糊与结构信息表达不足的问题,旨在提升模型对肺结节空间结构与局部细节的感知能力.该分支以所提出的SEA模块为核心,通过显式建模局部结构关系,增强对结构敏感区域的响应能力;紧接着引入由  $1 \times 1$  卷积、GELU激活以及再次  $1 \times 1$  卷积组成的轻量级特征变换模块,用于进一步优化特征表达.

如图4所示,SEA模块将输入特征沿通道维度划分为  $g$  个子组  $X \in \mathbb{R}^{B \times C/g \times H \times W}$ ,以提升局部建模能力与计算效率.在空间注意力(SA)子模块中,为捕捉结构的水平与垂直依赖关系,每个分组特征首先分别沿高度与宽度方向做平均池化,获得对应的结构特征图  $X_H \in \mathbb{R}^{B \times C/g \times 1 \times W}$  和  $X_W \in \mathbb{R}^{B \times C/g \times H \times 1}$ ,将二者拼接后输入  $1 \times 1$  卷积,随后通过Sigmoid激活函数生成双向空间注意力权重.该注意力权重用于对原始分组特征  $X$  进行加权调整,增强结节区域在水平与垂直方向上的响应.加权后的输出通过GN处理,得到结构感知表示  $X_1$ .在边缘注意力(EA)子模块中,首先通过拉普拉斯算子对  $X$  进行卷积,得到边缘响应图,并与原特征相加得到边缘增强特征  $X_2 = X + \text{LapConv}(X)$ .随后,分别对  $X_1$  和  $X_2$  执行自适应池化与通道维度平均,形成两个注意力向量.该过程中,  $X_1$  的全局响应用于加权  $X_2$  的空间特征,  $X_2$  的全局响应则反向作用于  $X_1$ ,具体表示为:

$$\begin{aligned} W_1 &= \text{Softmax}(\text{AvgPool}(X_1)), \\ W_2 &= \text{Softmax}(\text{AvgPool}(X_2)), \end{aligned} \quad (3)$$

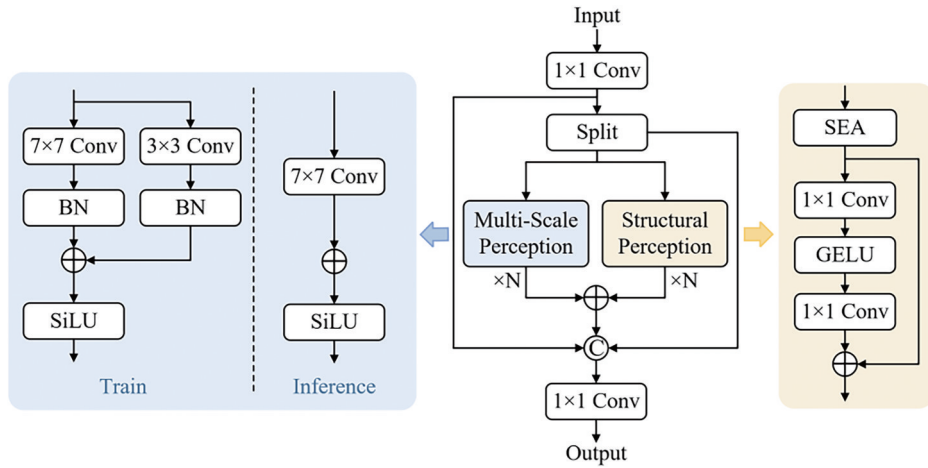


图 3 DPFM 示意图

Fig. 3 Diagram of DPFM

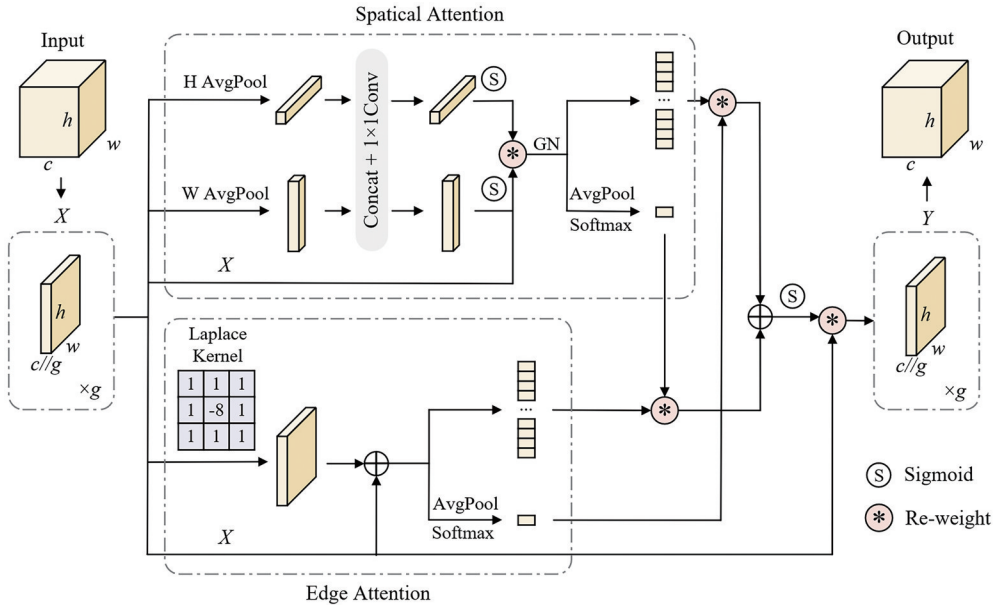


图 4 SEA 示意图

Fig. 4 Diagram of SEA

$$R_1 = W_1 \cdot \text{Flatten}(X_2), R_2 = W_2 \cdot \text{Flatten}(X_1). \quad (4)$$

两个权重矩阵分别与对应特征图做点乘操作后相加融合, 得到最终的融合注意力权重  $A_f$ :

$$A_f = \sigma(R_1 + R_2), \quad (5)$$

其中  $\sigma(\cdot)$  表示激活.

最终输出特征通过将该融合注意力权重与原始输入逐像素相乘, 并重新 reshape 为输入维度:

$$Y = A_f \odot X. \quad (6)$$

该注意力机制通过 SA 与 EA 的交互建模, 有效整合了结构信息与边缘信息, 强化了模糊区域的特征表达.

### 3 实验与结果

#### 3.1 数据集及实验环境

##### 3.1.1 实验数据与预处理

本文采用的医学影像数据来自 LUNA16 数据集<sup>[17]</sup>与天池数据集, 分别包含 888 例与 1000 例低剂量胸部 CT 扫描影像及其对应的结节标注信息. 两个数据集均按 6:2:2 划分为训练集、验证集与测试集, 为确保评估的客观性, 同一病例的 CT 图像始终划分至同一子集. LUNA16 作为主要实验数据集, 用于模型训练、参数优化、性能评估以及消融分析; 天池数据集用于评估模型在不同数据分布下的泛化能力.

考虑到低剂量 CT 图像普遍存在对比度低、噪声

干扰强以及解剖结构复杂等问题,本文对原始图像进行了肺实质区域提取的预处理.具体而言,LUNA16数据集采用官方提供的肺部掩膜文件对图像进行像素级筛选;而天池数据集由于未提供掩膜文件,本文基于灰度阈值分割与形态学处理自动提取肺实质区域.通过该预处理步骤,有效剔除了无关背景与非肺部结构,突出病灶区域的显著性与可辨识性.预处理前后的图像效果对比如图5所示.

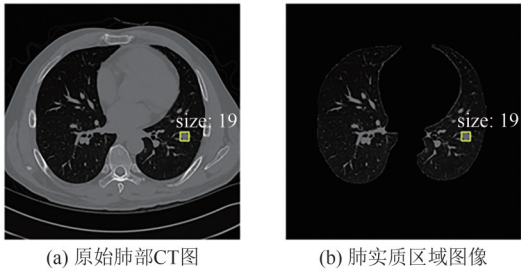


图5 肺部CT图像预处理

Fig. 5 Lung CT image preprocessing

### 3.1.2 实验环境与训练参数设置

本文基于PyTorch深度学习框架,采用Python编程语言,在搭载NVIDIA RTX 4080显卡的Windows 10操作系统环境下完成模型训练.优化器选用随机梯度下降法(SGD),以实现高效的参数更新与收敛控制.在超参数设置方面,综合实验稳定性与性能表现,本文将Batch Size设为32.根据损失函数收敛曲线的变化趋势,将训练轮数Epoch设置为300.初始学习率设定为0.01,权重衰减系数为0.0005,以在训练初期实现快速收敛,同时保持后期优化的稳定性与泛化能力.

### 3.2 评价指标

本文采用的评价指标包括精确率(Precision)、召回率(Recall)、平均精度(Average Precision, AP)、自由响应受试者工作特征曲线下面积(AFROC)、参数量(Parameters)、浮点运算总次数(GFLOPs).其中,Precision与Recall依赖于TP(真阳性)、FP(假阳性)和FN(假阴性)进行统计,其计算公式如下:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (7)$$

$$\text{Recall} = \frac{TP}{TP + FN}. \quad (8)$$

AP定义为精确率-召回率曲线下的面积,用于评价模型在不同阈值下的综合性能,公式如下:

$$\text{AP} = \int_0^1 P(R) dR. \quad (9)$$

$\text{AP}_{50}$ 表示交并比(Intersection over Union, IoU)

阈值为0.5时的平均精度,其中IoU用于衡量预测框与真实框的重叠程度.

AFROC是在自由响应曲线基础上计算得到的面积值,用于综合评估模型在不同假阳性水平下的整体检测效果.该曲线以每图像假阳性数(False Positives Per Image, FPPI)为横轴,假阳性率(True Positive Rate, TPR)为纵轴.为确保评价结果具有临床参考意义,本文将FPPI范围限制在 $[0, 2]$ ,从而减少过高假阳性对指标的干扰.

### 3.3 实验结果分析

#### 3.3.1 不同检测算法对比实验

为全面评估RDNet模型的检测性能,本文与SSD<sup>[18]</sup>、RetinaNet<sup>[19]</sup>、Faster-RCNN<sup>[20]</sup>、YOLOv8、GELAN<sup>[21]</sup>、YOLOv10<sup>[22]</sup>、RT-DETR<sup>[23]</sup>、YOLOv11<sup>[24]</sup>以及YOLOv12<sup>[25]</sup>进行对比,结果如表1所示.其中,与基线模型相比, $\text{AP}_{50}$ 提升了1.87%,Precision提高了2.74%,Recall实现了显著的8.33%的增幅.作为衡量检出能力的核心指标,Recall对于肺结节检测至关重要,其提升能够显著减少漏检,保障临床筛查的敏感性.此外,RDNet的AFROC指标优于其他模型,表明其在低假阳性情形下具有优异性能,在召回率与假阳性率之间取得了良好平衡.从参数优化来看,通过结构优化将参数量从3.01 M减少至1.38 M,大幅降低了存储需求并提升了部署效率.尽管GFLOPs较基线模型略有增加,但考虑到整体性能的显著提升,这种计算成本的增加是可以接受的.

图6展示了两个肺结节CT切片的检测示例,用于评估各模型在复杂场景下的表现.其中,金标准以黄色方框表示,蓝色标注表示RDNet的检测结果及置信度,红色标注表示其他模型的检测结果.对于病例1,结节位于低对比度背景区域,边界模糊.与大多数基准模型相比,RDNet能够准确定位结节的位置,生成的边界框更为紧凑,且置信度达到0.84,高于其他模型,表现出良好的检测稳定性.病例2中的结节位于肺实质边缘,位置贴近肺壁,结构上属于孤立型小结节.这类结节因靠近肺部边缘,易被误判为肺组织,增加了漏检与误检风险.从图中可见,SSD漏检了其中一个结节,Faster R-CNN则产生了明显误报.而RDNet依然能够稳定识别两个真实结节,分别给出0.75与0.73的高置信度预测结果,有效降低了误检与漏检的发生率,进一步验证了其在复杂边缘区域的鲁棒性与泛化能力.

#### 3.3.2 改进模型的消融实验

为验证各模块在肺结节检测任务中的独立贡

表 1 不同方法的肺结节检测结果  
Tab. 1 Detection results of lung nodules by different methods

Model	Year	AP <sub>50</sub> /%	Precision/%	Recall/%	AFROC	Parameters/M	GFLOPs
SSD	2016	62.56	72.82	56.80	1.725	13.06	15.03
RetinaNet	2017	76.67	77.84	52.00	1.771	32.20	127.21
Faster-RCNN	2016	63.11	64.49	71.20	1.573	41.35	121.28
YOLOv8	2023	83.70	85.27	74.07	1.783	3.01	4.10
GELAN	2024	81.44	80.08	77.20	1.805	1.88	3.54
YOLOv10	2024	77.64	83.01	64.50	1.760	2.71	4.20
RT-DETR	2024	72.96	73.71	67.60	1.789	27.42	52.58
YOLOv11	2024	82.53	84.01	75.68	1.794	2.59	3.22
YOLOv12	2025	81.56	77.53	77.28	1.705	2.52	2.99
RDNet	2025	85.57	88.01	82.40	1.824	1.38	4.50

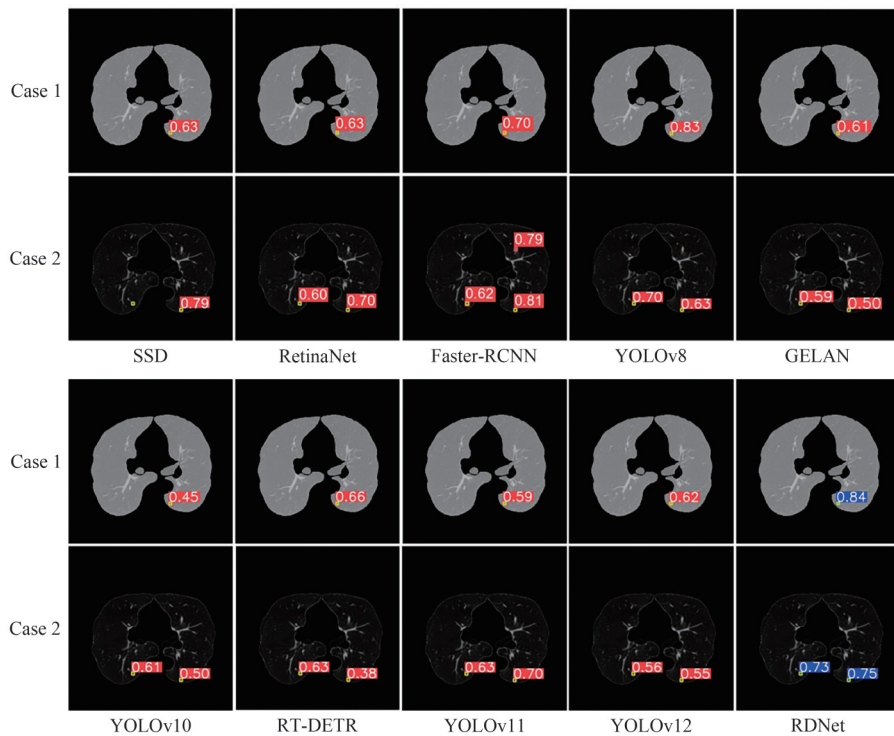


图 6 不同方法的肺结节检测结果图

Fig. 6 Figures of lung nodules detection results by different methods

献, 本文设计了系统性的消融实验 (如表 2 所示). 引入颈部重构后模型的精度与召回率均得到显著提升. 为提升跨尺度表达一致性, 本文进一步引入 PSDH, 在降低参数量与计算量的同时, 提升 AFROC

至 1.809, 验证其在复杂区域中的检测稳定性. 最终加入 DPFM 模块, 显著提升 Recall 至 82.40%、AFROC 升至 1.824, 同时保持较低的参数量与计算负担, 展现出优异的检测性能与效率平衡.

表 2 不同模块的消融实验

Tab. 2 Ablation experiments with different modules

Model	AP <sub>50</sub> /%	Precision/%	Recall/%	AFROC	Parameters/M	GFLOPs
Baseline	83.70	85.27	74.07	1.783	3.01	4.10
+ 颈部重构	85.57	85.43	78.40	1.736	5.38	28.31
+ 颈部重构+ PSDH	84.47	85.27	78.00	1.809	1.58	5.09
+ 颈部重构+ PSDH + DPFM	85.57	88.01	82.40	1.824	1.38	4.50

### 3.3.3 颈部重构可视化对比分析

为深入分析颈部结构对检测性能的影响, 本

文对其输出特征图进行热力图可视化, 以直观展示不同区域对结节的响应敏感性与激活分布. 如

图 7 所示,原始颈部结构在浅层特征图 P3' 中虽能捕捉部分高响应区域,但随着层级加深,P4' 与 P5' 对病灶区域的激活显著减弱,几乎无法形成有效响应,反映出特征在多次上下采样过程中存在信息丢失与扩散的问题.相比之下,本文构建的重构

颈部结构在各尺度特征图中均能保持对目标区域的良好响应,激活分布更加集中,展现出更强的空间聚焦能力与层间特征互补性.其中,新增的高分辨率浅层分支 P2 进一步增强了对小目标区域的感知能力.

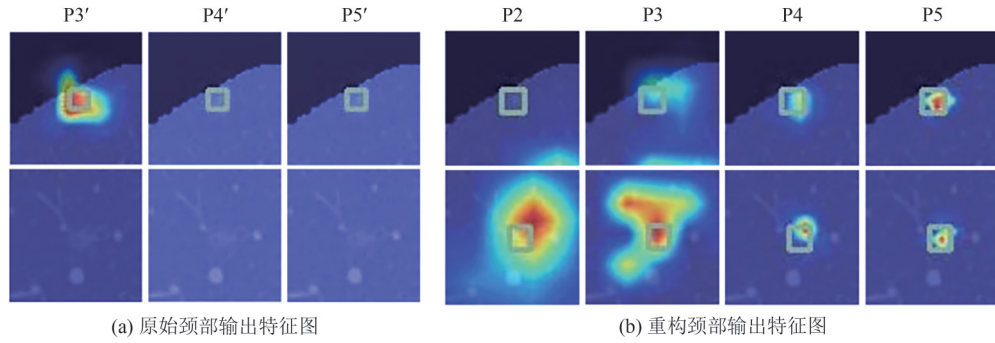


图 7 颈部结构的特征响应可视化

Fig. 7 Feature response visualization of neck structures

### 3.3.4 双重感知融合模块的消融实验

为评估 DPFM 模块中各子模块的独立作用与协同效应,本文设计了针对其内部 MP、SA 和 EA 的消融实验(如表 3 所示).其中,MP 用于提取多感受野上下文信息,SA 引导特征聚焦关键区域,EA 则强化结构边缘以辅助结节定位.实验结果表明,单独引入 MP

即可将 Precision 提升至 87.35%;在此基础上进一步加入 EA 或 SA,均带来稳健性能提升,体现了边界建模与空间引导机制的有效性.最终,三者协同构建完整的 DPFM 模块后, Precision、Recall 与 AFROC 实现同步提升,显著优于任一子模块或其简单组合,验证了各子模块之间的功能互补性与协同增益效果.

表 3 DPFM 的内部消融实验

Tab. 3 The internal ablation experiment of DPFM

MP	SA	EA	AP <sub>50</sub> /%	Precision/%	Recall/%	AFROC	Parameters/M	GFLOPs
√			80.51	87.35	74.00	1.822	1.31	4.25
√	√		80.66	82.37	74.76	1.798	1.37	4.48
√		√	83.00	87.97	73.20	1.806	1.37	4.49
√	√	√	<b>85.57</b>	<b>88.01</b>	<b>82.40</b>	<b>1.824</b>	1.38	4.50

图 8 展示了 DPFM 模块的消融实验热力图,系统分析了各子模块对特征响应能力的提升作用.原始卷积结构中,响应较为分散,难以聚焦于结节区域,且伴随显著的背景干扰.引入 MP 后,响应区域明显集中于目标区域,表明多尺度路径有助于增强模型对小目标的感知能力.在此基础上加入 SA,进一步强化了模型对结节内部结构的感知,使目标形态更完整.而在 MP 基础上引入 EA,有效增强了边缘轮廓的表达能力,显著改善结节边界的辨识度.最终在 RDNet 中融合 MP、SA 与 EA 后,模型实现了对结构细节与边缘信息的全面感知,响应区域更加集中、准确,呈现出最优的特征表征效果.

### 3.3.5 跨数据集检测性能验证

为验证 RDNet 在未知数据上的泛化能力,本文选用天池数据集进行跨数据集评估.为兼顾计算效率与迁移适应性,所有模型在训练过程中均采用冻

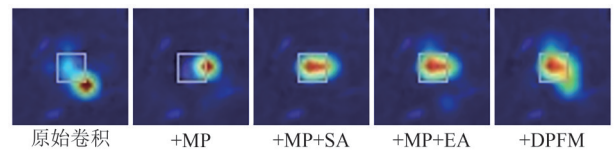


图 8 DPFM 消融实验热力图

Fig. 8 Ablation study heatmaps of the DPFM

结骨干网络的策略,训练轮数基于初步实验中的收敛趋势统一设为 50 轮,结果如表 4 所示,尽管目标数据在分布上存在差异,RDNet 在相同的训练策略与轮次条件下,仍取得了优于其他模型的检测性能,充分体现了其结构设计的有效性与跨数据集的泛化能力.

## 4 总结

本文提出的肺结节检测模型 RDNet,在颈部重

表4 Tianchi数据集的肺结节检测结果

Tab. 4 The lung nodule detection results of the Tianchi dataset

Model	Year	AP <sub>50</sub> /%	Precision/%	Recall/%	AFROC	Parameters/M	GFLOPs
YOLOv8	2023	73.87	83.21	67.07	1.849	3.01	4.10
GELAN	2024	69.02	81.62	60.15	1.759	1.88	3.54
YOLOv10	2024	59.47	67.54	53.98	1.646	2.71	4.20
RT-DETR	2024	63.37	74.74	62.78	1.798	27.42	52.58
YOLOv11	2024	73.11	<b>84.94</b>	65.41	1.852	2.59	3.22
YOLOv12	2025	74.27	83.13	68.42	1.825	2.52	<b>2.99</b>
RDNet	2025	<b>76.50</b>	78.59	<b>74.44</b>	<b>1.854</b>	<b>1.38</b>	4.50

构、PSDH模块与DPFM设计方面进行了有效探索,展现出良好的结构潜力与性能表现.通过优化颈部结构,RDNet有效增强了浅层与深层特征的融合,提升了对小病灶的敏感度.PSDH模块通过共享参数策略,减少了多尺度预测中的冗余计算,同时提升了检测结果的一致性.DPFM结合深度可分离卷积与注意力机制,提升了模型对小结节细节的捕捉能力.这些创新有效平衡了模型的计算效率与检测性能,显著提升了对小目标的识别效果.

#### 参 考 文 献

- [1] BRAY F, LAVERSANNE M, SUNG H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. CA: A Cancer Journal for Clinicians, 2024, 74(3): 229-263.
- [2] WEN M, ZHENG Q, JI X, et al. Precise diagnosis and prognosis assessment of malignant lung nodules: A narrative review [J]. Journal of Thoracic Disease, 2024, 16(11): 7999-8013.
- [3] KUMAR S, KUMAR H, KUMAR G, et al. A methodical exploration of imaging modalities from dataset to detection through machine learning paradigms in prominent lung disease diagnosis: A review [J]. BMC Medical Imaging, 2024, 24(1): 30.
- [4] KALITA A J, BORUAH A, DAS T, et al. Biomedical imaging: Advances in artificial intelligence and machine learning [M]. Singapore: Springer, 2024.
- [5] LU P X, LU H Z, YI Y X. Diagnostic Imaging of Drug Resistant Pulmonary Tuberculosis [M]. Singapore: Springer Nature Singapore, 2023.
- [6] CRASTA L J, NEEMA R, PAIS A R. A novel Deep Learning architecture for lung cancer detection and diagnosis from Computed Tomography image analysis [J]. Healthcare Analytics, 2024, 5: 100316.
- [7] VARGHESE R, M S. YOLOv8: A novel object detection algorithm with enhanced performance and robustness [C]//2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS). Chennai: IEEE, 2024: 1-6.
- [8] ARKIN E, YADIKAR N, XU X, et al. A survey: Object detection methods from CNN to transformer [J]. Multimedia Tools and Applications, 2023, 82(14): 21353-21383.
- [9] AKILA AGNES S, ANITHA J, ARUN SOLOMON A. Two-stage lung nodule detection framework using enhanced UNet and convolutional LSTM networks in CT images [J]. Computers in Biology and Medicine, 2022, 149: 106059.
- [10] URREHMAN Z, QIANG Y, WANG L, et al. Effective lung nodule detection using deep CNN with dual attention mechanisms [J]. Scientific Reports, 2024, 14: 3934.
- [11] TANG C, ZHOU F, SUN J, et al. Circle-YOLO: An anchor-free lung nodule detection algorithm using bounding circle representation [J]. Pattern Recognition, 2025, 161: 111294.
- [12] MKINDU H, WU L, ZHAO Y. Lung nodule detection in chest CT images based on vision transformer network with Bayesian optimization [J]. Biomedical Signal Processing and Control, 2023, 85: 104866.
- [13] CUI F, LI Y, LUO H, et al. SF<sub>2</sub>T: Leveraging Swin Transformer and Two-stream networks for lung nodule detection [J]. Biomedical Signal Processing and Control, 2024, 95: 106389.
- [14] TANG J, CHEN X, FAN L, et al. LN-DETR: An efficient Transformer architecture for lung nodule detection with multi-scale feature fusion [J]. Neurocomputing, 2025, 633: 129827.
- [15] TIAN Z, SHEN C, CHEN H, et al. FCOS: A simple and strong anchor-free object detector [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(4): 1922-1933.
- [16] DING X, ZHANG X, MA N, et al. RepVGG: Making VGG-style ConvNets great again [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021: 13728-13737.

- [17] SETIO A A A, TRAVERSO A, DE BEL T, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge [J]. *Medical Image Analysis*, 2017, 42: 1-13.
- [18] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector [C]//*Computer Vision-ECCV 2016*. Cham: Springer, 2016: 21-37.
- [19] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017: 2999-3007.
- [20] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [21] WANG C Y, YE H I, MARK LIAO H Y. YOLOv9: Learning what you want to Learn using programmable gradient information [C]//*Computer Vision-ECCV 2024*. Cham: Springer, 2025: 1-21.
- [22] WANG A, CHEN H, LIU L, et al. Yolov10: Real-time end-to-end object detection [J]. *Advances in Neural Information Processing Systems*, 2024, 37: 107984-108011.
- [23] ZHAO Y, LV W, XU S, et al. DETRs beat YOLOs on real-time object detection [C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2024: 16965-16974.
- [24] KHANAM R, HUSSAIN M. Yolov11: An overview of the key architectural enhancements [J]. *arXiv preprint arXiv*: 2024, 2410.17725.
- [25] TIAN Y, YE Q, DOERMANN D. Yolov12: Attention-centric real-time object detectors [J]. *arXiv preprint arXiv*: 2025, 2502.12524.

(责编 曹东,校对 雷建云)