

文章编号:1007-2780(2026)09-1339-13

基于多视角注意力融合的肺结核影像分类

闫小丽¹, 张蔚², 刘庭亭¹, 马超¹, 徐延忠³, 何志琴^{1*}

(1. 贵州大学 电气工程学院, 贵州 贵阳 550025;

2. 贵阳市公共卫生救治中心, 贵州 贵阳 550000;

3. 贵阳市花溪区人民医院, 贵州 贵阳 550025)

摘要:肺结核计算机断层扫描(Computed Tomography, CT)影像存在噪声干扰、病灶特征弱及三维空间信息利用不足等问题,影响分类性能。针对上述问题,本文提出一种融合图像增强与多维特征学习的肺结核CT三分类方法。该方法采用多尺度八方向各向异性扩散增强与渐进式N2F(Noise2Fast)去噪联合预处理策略,提高病灶纹理可见性;随后构建轻量化PRISM-Net(Parallel Representation with Iterative Spatial Multi-view Fusion Network),通过迭代式多视角融合策略(IMVF)挖掘不同空间方向的互补信息,并引入多维注意力模块(MDAM)增强三维特征表达能力。在150例独立测试集上进行验证,所提方法精确率达到94.68%,加权F1值达到94.63%,正常、活动性及非活动性肺结核类别AUC分别达到1.000 0、0.992 6和0.992 3,性能优于MambaOut-Kobe、MobileNetV3、MedVit及ResNet50等模型。

关键词:肺结核分类;图像增强;多视角融合;多维注意力机制;三维卷积神经网络

中图分类号:TP391.41;R521 文献标识码:A doi:10.3724/CJLCD.2026-0148 CSTR:32172.14.CJLCD.2026-0148

Classification of pulmonary tuberculosis images based on multi-view attention fusion

YAN Xiaoli¹, ZHANG Wei², LIU Tingting¹, MA Chao¹, XU Yanzhong³, HE Zhiqin^{1*}

(1. School of Electrical Engineering, Guizhou University, Guiyang 550025, China;

2. Guiyang Public Health Treatment Center, Guiyang 550000, China;

3. Guiyang Huaxi District People's Hospital, Guiyang 550025, China)

Abstract: Computed Tomography(CT) images of pulmonary tuberculosis suffer from issues including noise interference, weak lesion features, and insufficient utilization of three-dimensional spatial information, which impair classification performance. To address these problems, this paper proposes a three-class classification method for pulmonary tuberculosis CT images that integrates image enhancement and multi-dimensional feature learning. This method adopts a joint preprocessing strategy combining multi-scale eight-direction anisotropic diffusion enhancement and progressive N2F (Noise2Fast) denoising, which improves the visibility of lesion textures. Subsequently, a lightweight Parallel Representation with PRISM-Net (Iterative Spatial Multi-view Fusion Network) is constructed. The network utilizes an Iterative Multi-view

收稿日期:2026-08-02;修订日期:2026-08-19.

基金项目:贵州省科学技术基金(黔科合支撑[2023]一般096)

Supported by Guizhou Science and Technology Foundation (Qian [2023] 096)

*通信联系人, E-mail: zqhe@gzu.edu.cn

Fusion(IMVF) strategy to extract complementary information from different spatial orientations, and introduces a MDAM (Multi-Dimensional Attention Module) to enhance the expression capability of three-dimensional features. Validated on an independent test set comprising 150 cases, the proposed method achieves a precision of 94.68% and a weighted F1-score of 94.63%. The AUC (Area Under the Curve) for normal samples, active pulmonary tuberculosis, and inactive pulmonary tuberculosis reaches 1.000 0, 0.992 6, and 0.992 3 respectively, outperforming benchmark models including MambaOut-Kobe, MobileNetV3, MedVit and ResNet50.

Key words: pulmonary tuberculosis classification; image enhancement; multi-view fusion; multi-dimensional attention mechanism; three-dimensional convolutional neural network

1 引言

胸部影像学检查是肺结核诊断的重要依据,其中计算机断层扫描(Computed Tomography, CT)能够提供更加丰富的肺部结构信息。与传统X线胸片相比,CT具有更高的空间分辨率和组织分辨能力,可以显示树芽症、磨玻璃影、纤维化以及结节样改变等多种病理表现^[1-3]。然而,肺结核CT影像具有明显异质性,不同阶段病灶可能呈现不同形态,活动性与非活动性肺结核之间部分影像特征存在重叠,仅依靠人工阅片进行准确分类仍具有较大挑战。

深度学习的发展为医学影像自动分析提供了新的技术路径。近年来,机器学习与先进成像技术的深度融合已成为智能影像分析的重要发展方向^[4]。针对医学图像质量下降和弱结构信息恢复问题,基于图像增强和质量优化的方法逐渐受到关注,可通过改善局部对比度和增强有效区域响应提升视觉表达能力^[5-7]。卷积神经网络(CNN)能够从大量医学影像数据中自动学习层次化特征表示,在疾病分类、病灶检测和器官分割等任务中取得了较好的应用效果^[8-9]。针对CT影像分析,二维卷积网络通常以单张切片作为输入,但难以充分利用切片之间连续空间关系;相比之下,三维卷积神经网络(3D CNN)能够直接处理体积数据,在空间维度上建立更加完整的特征表示,因此逐渐应用于肺部疾病分析和CT辅助诊断任务^[10-11]。

尽管基于深度学习的方法取得了一定进展,肺结核CT分类仍面临多个方面的挑战。CT图像质量会影响模型对病灶特征的学习。实际临床CT数据通常受到扫描剂量、设备差异以及重

建算法影响,存在不同程度的噪声和灰度变化^[12]。传统医学影像增强方法通常依赖图像灰度分布和人工参数调整,通过对比度增强、亮度校正等方式改善视觉质量。例如,基于自适应伽马校正的方法能够有效提升胸部影像的对比度和病变区域可见性,为后续医学影像分析提供更清晰的结构信息^[13]。而深度学习去噪方法虽然能够通过数据驱动方式学习噪声与结构信息之间的映射关系,但通常依赖训练数据质量,在真实临床场景中的泛化能力仍需进一步提升^[14-15]。

此外,医学影像中的有效判别信息通常分布于多个特征维度。注意力机制通过动态调整特征权重,使模型能够增强关键区域响应并抑制无关信息^[16]。然而,现有注意力方法主要集中于通道关系或二维空间区域,对于三维CT数据中沿深度方向变化的切片信息关注不足。近年来,基于三维注意力机制的方法逐渐应用于医学影像分析,通过增强三维特征表示能力改善模型性能^[13]。肺结核病灶通常具有跨切片连续变化特征,因此需要进一步利用多维空间信息提升病灶表征能力。

基于上述问题,本文提出了一种面向肺结核CT三分类任务的两阶段深度学习框架。该方法首先结合多尺度八方向各向异性扩散增强与渐进式N2F(Noise2Fast)去噪策略优化CT图像质量,提升微小病灶纹理表达能力;随后构建轻量化3D PRISM-Net(Parallel Representation with Iterative Spatial Multi-view fusion Network),通过迭代式多视角融合(Iterative Multi-View Fusion, IMVF)策略增强不同空间方向特征交互能力,并引入多维注意力模块(Multi-Dimensional Attention Module, MDAM),联合建模通道、空间及Z轴切片信息,提高模型对肺结核关键特征的感知能

力,实现活动性肺结核、非活动性肺结核与正常肺部的较准确地分类。

2 相关工作

2.1 CT图像增强与去噪方法

CT图像增强与去噪是医学影像分析的重要预处理步骤,主要用于降低噪声干扰并提升病灶结构可见性。传统方法主要包括滤波、直方图均衡化以及基于偏微分方程(Partial Differential Equation, PDE)的扩散方法,其中Perona-Malik各向异性扩散算法通过利用梯度信息调节扩散强度,在抑制噪声的同时保持边缘结构^[17]。然而,该类方法依赖人工参数设置,难以适应复杂噪声环境。随着深度学习的发展,基于卷积神经网络的CT增强方法逐渐应用于低剂量CT恢复任务,例如RED-CNN通过残差学习提高低剂量CT图像质量^[18],后续研究进一步结合动态卷积、

多尺度特征融合以及先验知识约束机制增强噪声抑制能力^[19]。此外,基于自适应增强和特征融合策略的方法能够有效改善图像对比度和细节表达,为医学影像增强提供新的思路^[5,14]。然而,监督式方法依赖高质量配对数据,而真实临床CT中难以获取无噪声参考图像。为降低数据依赖,Noise2Noise、Noise2Void以及Noise2Fast等自监督方法利用噪声图像自身统计规律完成恢复,其中Noise2Fast(N2F)通过快速采样策略降低训练复杂度,提高单幅图像去噪效率^[20]。但现有方法仍难以同时兼顾噪声抑制与微小病灶纹理保持。

2.2 三维CT特征提取与多视角融合方法

计算机断层扫描(CT)能够较为清晰地显示树芽征、磨玻璃影、纤维化以及结节样改变等多种肺结核相关病理表现,如图1所示。这些病灶通常具有复杂的三维空间分布和跨切片连续性,因此对CT体数据进行有效的空间特征建模成为肺结核智能诊断的重要研究方向。

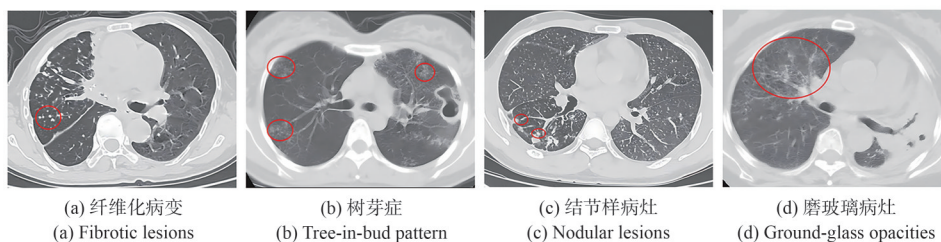


图1 肺结核CT影像中的典型病理表现

Fig. 1 Typical pathological features in CT imaging of tuberculosis

深度学习的发展推动了CNN在医学影像分类中的应用。早期二维CNN主要基于单张CT切片提取纹理和空间信息,但忽略了切片之间的连续结构关系。相比之下,三维卷积神经网络(3D CNN)能够同时建模深度、高度和宽度方向信息,在肺结节检测、肺癌筛查以及肺部疾病分类任务中表现出更强的空间特征提取能力^[12,21]。随后,ResNet、DenseNet以及轻量化3D网络结构被引入医学影像分析,通过残差连接和参数优化提升模型训练稳定性与计算效率^[22]。然而,单一视角3D CNN难以充分利用CT体数据中不同方向的互补信息,因此多视角学习方法逐渐受到关注。该类方法通常从轴向、冠状和矢状方向提取特征,并通过特征拼接、跨模态融合或注意力融合实现空间信息整合^[23-24]。

2.3 注意力机制在医学影像分类中的应用

注意力机制通过动态调整特征权重,使网络能够关注具有判别意义的信息,在医学影像分类任务中得到广泛应用。典型方法包括挤压-激励(Squeeze-and-Excitation, SE)模块,通过建模通道间依赖关系增强关键特征响应;卷积注意力模块(CBAM)进一步融合通道注意力和空间注意力,提高模型对重要区域的感知能力^[25],同时,基于注意力残差结构的特征重建方法也被证明能够有效增强关键信息表达能力^[26]。近年来,跨域交互注意力与自限制注意力等机制进一步增强了多源特征之间的信息交互能力,为复杂医学图像分析提供了新的研究思路^[27-28]。注意力机制逐渐扩展至三维医学影像,通过增强空间特征表示提升分类性能^[17]。然而,现有方法主要关注通道和二维空间信息,对于

CT 体数据中特有的 Z 轴方向连续变化建模不足。

3 方 法

3.1 多尺度八方向亚像素各向异性扩散增强

各向异性扩散通过梯度幅值自适应调节平滑强度:在平坦区加速去噪,在边缘处抑制扩散以保护细节。本文将梯度方向从传统的 4 个扩展至 8 个,通过反射填充获取邻域亚像素值,计算各方向的二阶差分。传导系数采用指数衰减与 Lorentz 函数的乘积形式:

$$c(g) = \exp\left(-\frac{g^2}{\epsilon^2}\right) \cdot \frac{1}{1 + \left(\frac{g}{k_{\text{adj}}}\right)^2}, \quad (1)$$

式中: g 为单方向亚像素梯度, ϵ 控制扩散衰减速率, k_{adj} 为自适应阈值。

在此基础上构建多尺度增强框架,设置 n_{scales} 级扩散尺度,逐级增加迭代次数与传导阈值,得到由细到粗的多层平滑特征图 $F_1, F_2, \dots, F_{n_{\text{scales}}}$ 。相邻尺度特征相减得到对应尺度的细节残差 $D_i = F_i - F_{i+1}$,通过细节图的信息熵自适应计算各尺度权重,高信息熵对应纹理丰富的病灶区

域,分配更高权重:

$$w_i = \frac{\text{ent}(D_i)}{\sum_{k=1}^{n_{\text{scales}}} \text{ent}(D_k)}, \quad (2)$$

式中, $\text{ent}(\cdot)$ 为图像信息熵计算函数。

最终将加权融合的细节特征叠加至最大尺度平滑基底,得到增强后图像:

$$I_{\text{enhance}} = B + \phi \cdot \sum_{i=1}^{n_{\text{scales}}} w_i \cdot D_i, \quad (3)$$

式中: B 为最大扩散尺度对应的平滑特征图, ϕ 为全局增强系数。

3.2 N2F 的渐进式去噪模块

N2F 是一种基于单图像的无监督盲去噪模型,依靠棋盘式下采样构建微数据集,实现噪声解耦与快速高效降噪。但其采用单阶段固定去噪策略,数据集特征表征能力有限,无法适配 CT 图像多变的噪声强度,易出现强噪声细节丢失、弱噪声边缘过度平滑等问题,降噪鲁棒性不足。本文提出渐进式 N2F 去噪方法,整体流程如图 2 所示。该策略摒弃单次全局降噪模式,将复杂的 CT 图像去噪任务拆分为多级子任务,以由粗到细的迭代方式逐步优化图像特征。

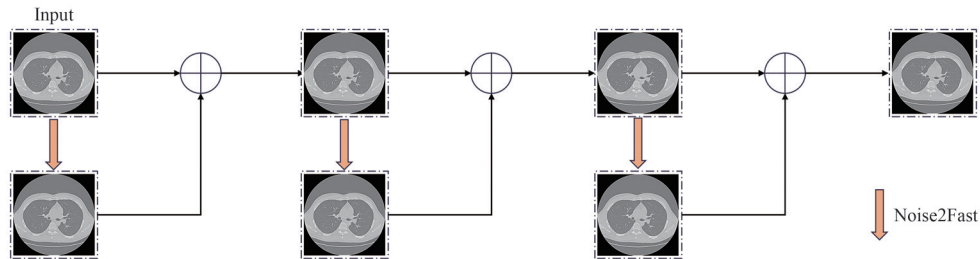


图 2 渐进式 N2F 流程

Fig. 2 Progressive N2F process

渐进式 N2F 以逐级迭代优化的图像为基准,仅对当前特征进行微调修正,有效规避单次映射导致的细节损毁,其迭代更新规则为 $I_t = I_{t-1} + \alpha_t \cdot \mathcal{D}_t(I_{t-1})$, $\mathcal{D}_t$ 为对应阶段去噪器, α_t 为渐进融合系数。

3.3 多维注意力模块

本文提出多维注意力模块(MDAM),同时从通道、空间与 Z 轴切片 3 个维度学习注意力权重,对骨干输出的三维特征进行多维度协同增强。模块内部结构如图 3 所示。

对于骨干网络输出的三维特征 $X \in$

$R^{C \times D \times H \times W}$, 其中 C 为特征通道数, D 为 Z 轴切片数, H, W 分别为特征图的高度与宽度, 3 条分支独立计算对应维度的注意力权重, 具体过程如下:

(1) 通道注意力分支

通道注意力分支采用压缩-激励机制, 通过全局三维平均池化对空间与深度维度进行全局聚合, 将原始三维特征压缩为长度为 C 的通道特征向量, 实现全局语义信息的嵌入:

$$z_c = \frac{1}{D \cdot H \cdot W} \sum_{d=1}^D \sum_{h=1}^H \sum_{w=1}^W X(:, d, h, w), \quad (4)$$

式中, $z_c \in R^C$ 为全局通道描述子。

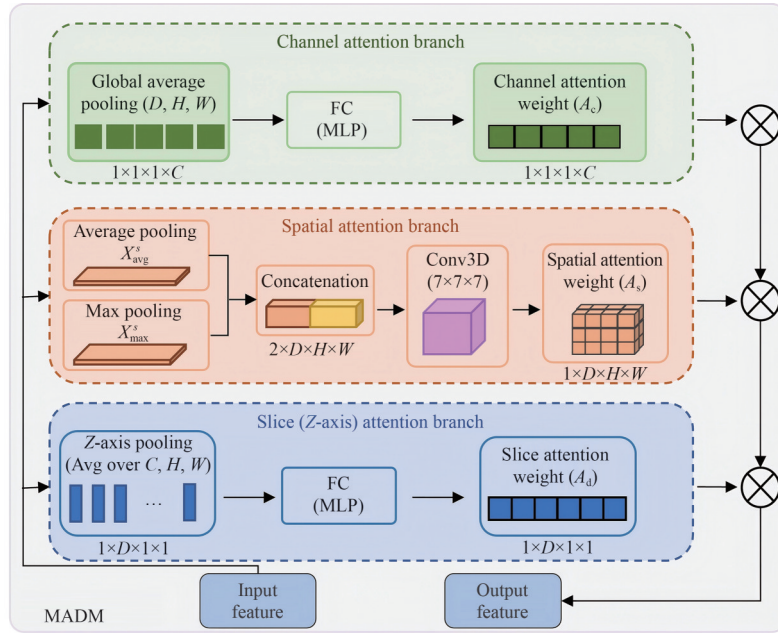


图3 多维注意力模块(MDAM)

Fig. 3 Multidimensional attention module (MDAM)

(2) 空间注意力分支

空间注意力分支沿通道维度对原始特征进行平均池化与最大池化,将通道维度压缩为2个特征映射,聚合通道维度的整体信息:

$$\begin{cases} X_{\text{avg}}^s = \frac{1}{C} \sum_{c=1}^C X(c, :, :, :) \\ X_{\text{max}}^s = \max_{c \in [1, C]} X(c, :, :, :) \end{cases}, \quad (5)$$

式中, $X_{\text{avg}}^s, X_{\text{max}}^s \in R^{1 \times D \times H \times W}$ 分别为通道维度的平均与最大池化结果。

(3) 切片注意力分支

切片注意力分支沿深度方向对特征进行全局聚合,先通过通道与空间维度的联合平均池化,得到长度为D的深度特征向量:

$$z_d = \frac{1}{C \cdot H \cdot W} \sum_{c=1}^C \sum_{h=1}^H \sum_{w=1}^W X(c, :, h, w), \quad (6)$$

式中, $z_d \in R^D$ 为深度描述子。

将3个分支输出的注意力权重通过广播机制与原始特征张量逐元素相乘,实现多维度协同增强:

$$X_{\text{out}} = X \otimes A_c \otimes A_s \otimes A_d, \quad (7)$$

式中: $\otimes$ 表示广播机制下的逐元素乘法运算, A_c, A_s, A_d 分别为通道注意力权重、空间注意力权重、全连接层权重。

3.4 迭代式多视角融合策略

本文基于同一CT体数据沿不同空间维度构

建正交视角特征。以原始轴位CT体数据作为主视角,同时采用各向异性最大池化沿3个空间维度分别进行信息压缩,生成深度、高度和宽度3个辅助视角。其中,深度视角通过沿Depth维聚合获得,高度视角通过沿Height维聚合获得,宽度视角通过沿Width维聚合获得。为保证后续网络处理的一致性,所有辅助视角均通过插值重采样恢复至与主视角相同的空间尺寸后输入共享骨干网络进行特征提取。

迭代式多视角融合(IMVF)模块以主视角特征为核心基准,首先通过骨干网络完成主视角CT图像的初始特征提取,构建基础特征表征。在此基础上,模型构建三轮递进式迭代校正链路,依次引入深度、高度、宽度3个正交辅助视角的特征信息,依托共享骨干网络完成各辅助视角的特征映射。每一轮迭代均通过SE通道注意力机制自适应挖掘辅助视角的有效病灶特征,生成专属通道校正权重,并通过公式(8)对主分支特征进行增量更新优化:

$$M_{\text{new}} = M_{\text{old}} + M_{\text{old}} \times (W_{\text{se}}(V^{(i)})). \quad (8)$$

三轮迭代校正具备层级化的空间特征增益逻辑,各司其职、逐层优化特征表达。第一轮深度维度迭代聚焦轴向切片特征筛选,精准定位含病灶的关键切片区域,过滤无效背景特征;第二

轮与第三轮迭代依次补充冠状面、矢状面的空间结构信息,细化病灶的二维边界位置与三维空间分布关系。

3.5 PRISM-Net整体架构

PRISM-Net以经典残差网络 ResNet 为基础构建。考虑到 3D 卷积网络参数较大且数据规模有限,选择轻量化 ResNet10 作为骨干特征提取网络,以降低过拟合风险并提高计算效率。整体框架如图 4 所示,除原始 3D CT 体数据作为主视角(Main View)输入外,模型通过各向异性最大池化分别构建深度、高度和宽度 3 个辅助视角,以获取不同空间方向的互补特征。骨干网络由 Stem 层和 4 个残差阶段组成。Stem 层包含 $7 \times 7 \times 7$ 卷积、批归一化(BN)、ReLU 激活函数及

最大池化操作,用于完成初始特征提取与空间降采样。随后依次堆叠 4 个嵌入 MDAM 的基础模块(BasicBlock),输出通道数由 64 逐步扩展至 128、256 和 512,实现多层次特征提取。为控制模型复杂度并提升参数利用效率,主视角与 3 个辅助视角采用参数共享机制,共用同一套骨干网络权重。此外为缓解类别间特征重叠带来的过拟合问题,模型训练阶段采用标签平滑交叉熵损失函数作为优化目标,其表达式如式(9)所示:

$$L = - \sum_{i=1}^C y_i^{LS} \log p_i, \tag{9}$$

其中: $y_i^{LS} = (1 - \epsilon) y_i + \frac{\epsilon}{C}$, 式中 C 为类别数, p_i 为预测概率, ϵ 为平滑系数。

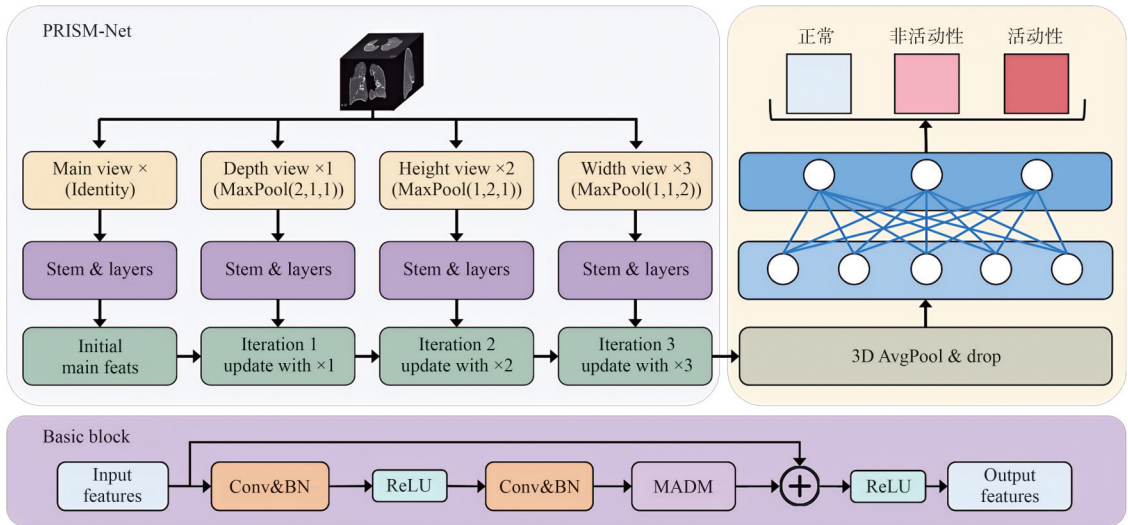


图 4 PRISM-Net分类诊断框架

Fig. 4 PRISM-Net classification and diagnosis framework

4 实验与分析

4.1 数据集

本文采用 3 组数据集作为实验验证的数据来源,包括公开标准数据集和本地数据集。公开标准数据集由 LIDC-IDRI (Lung16)^[29] 数据集和 COVID-19 数据集组成,COVID-19 数据集是 COVID-19_1^[30] 以及 COVID-19_2^[31] 进行合并的数据集。本地数据集是由项目合作医院提供的包含 5 mm 和 0.625 mm 两种厚度的数据集。为避免数据泄漏,所有数据均按照患者级别进行划

分,同一患者的 CT 数据仅被分配至训练集、验证集或测试集中的一个子集,不存在跨集合重复样本。为避免数据泄漏,归一化、数据增强以及渐进式 N2F 训练均仅在训练集上完成。

4.2 实验环境配置

实验的运行环境为 WSL-Ubuntu 22.04 操作系统,硬件环境是 12th Gen Intel(R) Core (TM) i9-12900K 处理器, NVIDIA GeForce RTX 4090-24G 显卡;软件配置为 PyTorch 深度学习框架以及 Python 3.9 编程语言。

为保证不同数据来源之间的数据分布一致

性,所有CT图像在训练前均采用Min-Max归一化。随后将所有输入3D CT样本的三维尺寸统一缩放至 $200 \times 256 \times 256$,中深度(Z轴)方向保留200层连续切片,轴位面(XY平面)分辨率统一为 256×256 ,以保留病灶的三维空间连续性特征。实验训练批处理大小为8,训练100个epoch,采用Adam优化器,设置初始学习率为0.001,同时设置权重衰减系数为0.001。当连续10轮验证损失未出现下降时,将学习率按0.5的比例进行衰减,动态调整优化步长以保障训练后期的收敛稳定性。训练阶段采用随机翻转和随机旋转进行数据增强。损失函数采用标签平滑交叉熵,平滑系数 ϵ 设置为0.1。

4.3 评价指标

实验选取精确率(Precision)、召回率(Recall)、F1-Score以及AUC(Area Under the Curve)作为评价指标全面评估PRISM-Net在多分类任务中的性能表现。此外,图像的客观指标定量分析中,选用峰值信噪比(Peak Signal to Noise Ratio, PSNR)、增强测量熵(Enhancement Measure Entropy, EME)和平均绝对误差(Mean Absolute Error, MAE)3种量化指标作为评价图像质量的标准。其中PSNR和MAE均以对应的原始CT图像作为参考图像进行计算,用于评估增强过程中图像结构与灰度信息的保持能力,EME用于评价图像局部对比度增强效果。各指标计算公式如式(10)~(18)所示:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (10)$$

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (11)$$

$$F1 = \frac{2TP}{2TP + FP + FN}, \quad (12)$$

$$\text{AUC} = \int_0^1 \text{TPR}(\text{FPR})d(\text{FPR}), \quad (13)$$

$$\text{TPR} = \frac{TP}{TP + FN}, \quad (14)$$

$$\text{FPR} = \frac{FP}{FP + TN}, \quad (15)$$

$$\text{PSNR} = 10 \lg \frac{\text{MAX}^2}{\text{MSE}}, \quad (16)$$

$$\text{EME} = \frac{1}{k_1 \cdot k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} 20 \cdot \ln \left(\frac{I_{\max}(l, k)}{I_{\min}(l, k)} \right), \quad (17)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (18)$$

其中:式(10)精确率衡量预测为正样本结果中真实正样本占比;式(11)召回率表征真实正样本被检出比例;式(12)F1为精确率与召回率的调和平均,综合二者性能;式(13)AUC为ROC曲线下积分面积,依托式(14)真阳性率TPR(等同于召回率)与式(15)假阳性率FPR计算,评估模型分类能力;式(16)PSNR由图像最大像素值与均方误差求得,反映图像失真程度;式(17)EME通过图像分块灰度极值比值评估局部对比度;式(18)MAE计算像素灰度绝对误差均值,衡量图像灰度偏差。

4.4 分类模型与主流方法性能对比

4.4.1 各类图像增强方法视觉效果对比分析

为充分验证本文所提图像增强算法的综合性能,从亮度调节、对比度优化、细节保持与噪声抑制多个维度开展横向对比实验。为保证实验公平性,所有增强方法均在相同二维CT切片数据上实现并采用一致的实验设置进行评估。其中CLAHE^[32]、AGCWD和文献[13]方法均为非学习型增强方法;本文N2F模块仅利用训练集数据进行自监督训练,未使用任何外部预训练模型。本实验在Lung16与COVID-19两类数据集上进行,各类增强方法的定量分析结果和可视化结果如表1、图5所示。

表1 在各数据集上的测试结果

Tab. 1 Test results on each dataset

数据集	方法	EME $\uparrow$	PSNR $\uparrow$	MAE $\downarrow$
Lung16数据集	CLAHE	25.403 8	23.080 1	13.759 3
	AGCWD	33.221 8	18.928 4	23.520 5
	文献[13]	32.548 7	21.300 5	18.861 0
	Ours	38.321 5	29.871 5	4.002 7
COVID-19数据集	CLAHE	15.917 8	21.430 5	15.797 8
	AGCWD	41.729 4	20.070 7	20.644 6
	文献[13]	42.006 3	22.279 6	16.116 5
	Ours	53.893 4	28.221 8	5.944 6

注:粗体数字表示最优值

结果显示,本文方法在Lung16和COVID-19数据集上均取得较优结果,定量指标整体优于对比方法,表明增强后图像具有更好的局部对比度表现。同时,较高的PSNR和较低的MAE说明

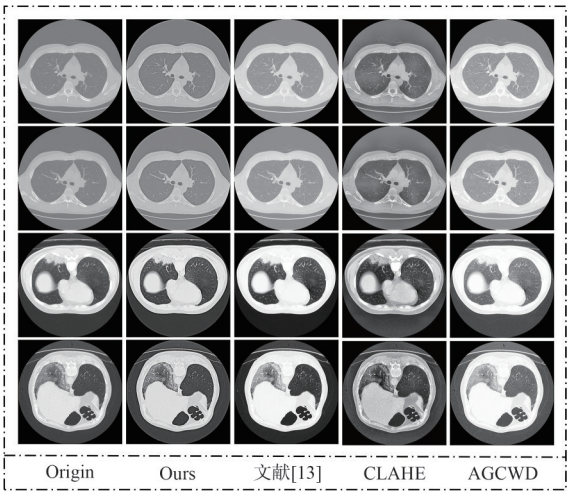


图 5 图像增强效果对比

Fig. 5 Comparison of image enhancement effects

本文方法在提升图像视觉效果的过程中,能够较好地保持原始 CT 图像的灰度分布与结构信息,避免过度增强带来的失真。相比于 CLAHE、AGCWD 及文献[13]方法,本文方法在增强肺实质纹理和血管边缘等关键结构细节方面表现更为稳定,并可有效抑制增强过程中噪声放大现象。

此外,为验证图像增强模块对下游任务的实际价值,本文进一步开展对比实验,结果如表 2 所示。实验结果表明,引入图像增强模块后,分类

表 2 图像增强对分类性能影响分析

Tab. 2 Analysis on the impact of image enhancement on classification performance

方法	Precision	Recall	F1 Score
增强前	0.925 7	0.923 4	0.922 7
增强后	0.946 9	0.946 7	0.946 3

模型的 Precision、Recall 及 F1 Score 均得到提升,说明增强后的 CT 图像能够为 PRISM-Net 提供更加清晰且具有判别性的病灶特征,从而提升肺结核分类性能。

4. 4. 2 主干网络对比

为验证本文所提模型的分类性能优势,将嵌入多维注意力模块的 ResNet10 模型与 ResNet18、VGG16 等经典三维卷积网络开展对比实验。所有对比实验保持数据集划分、损失函数、训练超参数完全一致,以加权精确率、加权召回率及模型参数量作为综合评价指标,在同一测试集上完成性能评估,实验结果如表 3、图 6 所示。

表 3 主干网络对比实验

Tab. 3 Comparative experiments on backbone networks

模型算法	评价指标		
	Precision	Recall	参数量/M
ResNet10 (3D)	0.903 2	0.892 3	14.357 1
ResNet18 (3D)	0.912 4	0.901 2	33.161 5
ResNet34 (3D)	0.916 1	0.897 1	63.471 2
Vgg16 (3D)	0.842 2	0.827 3	195.156

注:粗体数字表示最优值

结合表 2、图 5 结果能够发现,相较于 ResNet18、VGG16 等网络,ResNet10 在分类指标上具备较好性能,同时拥有更低的参数量与计算开销。受限于本数据集规模,深度更大的网络更容易出现过拟合现象,泛化性能受限。依托轻量化的网络结构优势,ResNet10 能够很好地适配本研究任务,在此基础上引入多维注意力模块(MDAM),可以进一步挖掘特征信息,提升肺结核 CT 图像分类效果。

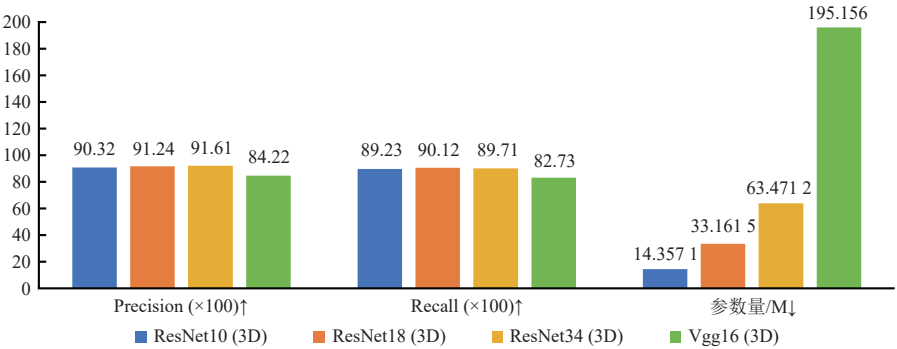


图 6 主干网络对比

Fig. 6 Comparison of backbone networks

4.4.3 PRISM-Net对比实验分析

为验证 PRISM-Net 在小样本肺结核诊断任务中的性能优势,本文将所提 PRISM-Net 与多种主流图像分类模型开展横向对比实验,对比模型包含 MambaOut-Kobe^[33]、MobileNetV3^[34]、MedVit^[35]以及 ResNet50^[36]。采用合作医院提供的 CT 数据按照 8:1:1 划分为训练集、测试集与验证集,测试集共计 150 例样本,其中正常样本 51 例、活动性肺结核样本 50 例、非活动性肺结核样本 49 例,各项对比结果如表 4。

表 4 不同方法在独立测试数据集上的对比(所有指标均采用加权平均)

Tab. 4 Comparison of different methods on the independent test dataset (all metrics are weighted averages)

模型算法	评价指标			
	Precision	Recall	F1 Score	AUC
MedVit	0.902 1	0.874 8	0.885 2	0.932 0
Resnet50	0.915 4	0.885 5	0.904 7	0.934 1
MambaOut-Kobe	0.934 1	0.897 9	0.916 8	0.963 2
MobileNetV3	0.924 3	0.903 1	0.911 2	0.958 4
PRISM-Net	0.946 9	0.946 7	0.946 3	0.994 3

注:粗体数字表示最优值

根据表 4 定量结果,PRISM-Net 取得较高的 Precision(94.69%)、Recall(94.67%)和 F1 Score(94.63%)。该结果充分说明,针对三维肺部病灶任务,模型无需依赖深层主干网络;在 ResNet10 浅层架构基础上引入多视角融合与注意力机制,可显著强化模型特征提取能力,较准确地识别肺结核,兼顾模型轻量化与识别精度。

PRISM-Net 引入 MDAM 模块以增强特征表征能力。在控制参数量、维持轻量化特性的条件下,模型各项指标均衡发展,并大幅超越对比方法,充分证明多维度注意力修正策略具备良好作用。模型采用参数共享多视角机制,可在主视角与 3 个辅助视角上复用同一套参数完成特征优化,提升模型对病灶特征的辨识深度,改善传统模型容易出现的漏诊、误诊失衡问题,显著提升召回率。

为全面评估 PRISM-Net 的分类性能,绘制三分类任务对应的 ROC 曲线与混淆矩阵,结果

如图 7、图 8 所示。类别定义:Class_1 为活动性肺结核,Class_2 为非活动性肺结核,Class_3 为正常肺部样本。

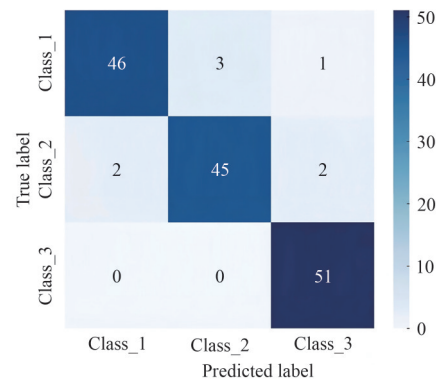


图 7 测试集混淆矩阵

Fig. 7 Test set confusion matrix

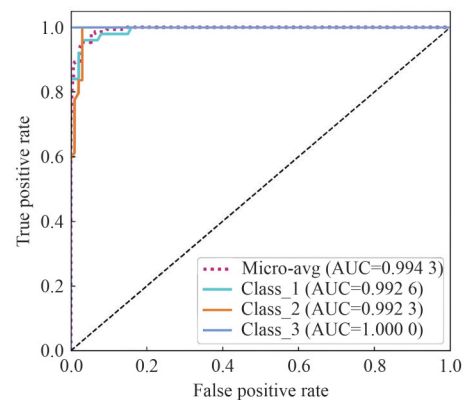


图 8 测试集 ROC 曲线

Fig. 8 ROC curves of the test set

图 7、图 8 结果表明,模型对正常肺部样本具有较强识别能力,召回率达到 100%;同时,各类别 AUC 均保持较高水平,正常类、活动性肺结核和非活动性肺结核 AUC 分别达到 1.000 0、0.992 6 和 0.992 3,平均 AUC 为 0.994 3。尽管部分活动性与非活动性肺结核因病灶特征重叠及微小病灶信息损失产生误判,但模型整体仍表现出较好的分类准确性、稳定性和鲁棒性。

4.5 消融实验分析

4.5.1 基于多尺度增强图像消融分析

为验证八方向滤波、亚像素特征增强与渐进式 N2F 去噪各模块的有效性,本文基于 Lung16 数据集开展模块消融实验。为便于定量分析,将八方向滤波策略、亚像素优化策略、渐进式 N2F

去噪策略分别定义为模块 A、B、C,各模块组合下的图像量化评价结果如表 5、图 9 所示。

表 5 消融实验结果

Tab. 5 Results of the ablation experiment

A	B	C	EME ↑	PSNR ↑	MAE ↓
×	×	×	33.346 2	25.606 6	7.207 6
✓	×	×	35.907 0	28.875 4	5.922 9
✓	✓	×	36.293 6	26.411 7	5.562 7
✓	✓	✓	38.321 5	29.871 5	4.002 7

注: 粗体数字表示最优值

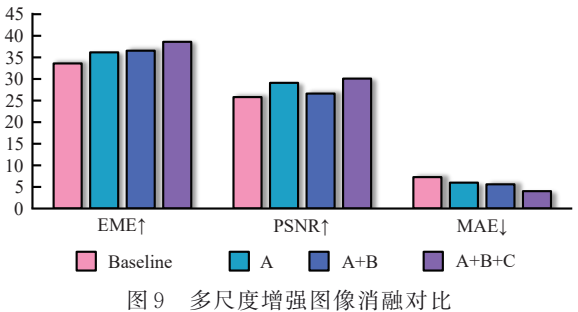


图 9 多尺度增强图像消融对比

Fig. 9 Ablation study on multi-scale image enhancement

表 6 消融实验 (指标数值均为 3 个类别指标的平均数)

Tab. 6 Ablation experiments (all reported metric values are the average of three class-specific metrics)

模型算法	MDAM	IMVF	评价指标			
			Precision ↑	Recall ↑	F1-Score ↑	参数量/M ↓
Baseline	×	×	0.903 2	0.892 3	0.901 2	14.357 1
Model A	✓	×	0.912 1	0.904 8	0.905 6	14.401 6
Model B	×	✓	0.938 4	0.940 5	0.939 5	14.465 2
PRISM-Net	✓	✓	0.946 8	0.946 7	0.946 3	14.549 7

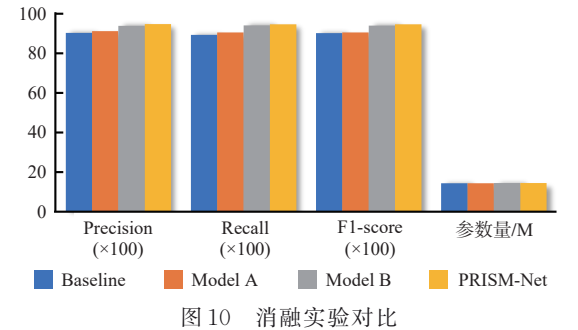


图 10 消融实验对比

Fig. 10 Comparison of ablation experiments

基于表 6 的消融实验量化结果,本实验通过指标提升幅度定量分析 MDAM 与 IMVF 模块的

实验结果表明,所提出的图像增强策略能够有效提升 CT 图像质量。模块 A 通过八方向滤波增强多维结构信息,使 EME 提升 7.67%、PSNR 提升 12.7%,MAE 降低 17.8%;进一步加入模块 B 后,增强了高频细节表达能力,但同时引入了部分噪声;结合渐进式 N2F 去噪模块 C 后,EME 提升 5.58%、PSNR 回升 13.09%、MAE 降低 28.04%,有效抑制增强过程中的噪声干扰,实现了病灶细节增强与噪声控制之间的平衡。

4.5.2 MDAM 与 IMVF 模块消融实验结果分析

为验证多维注意力模块(MDAM)与迭代式多视角融合策略(IMVF)的作用,设计递进消融实验。基线模型(Baseline)为原始 3D ResNet10,仅接收单路主视角输入,不含注意力模块。Model A 在 Baseline 的所有 BasicBlock 中嵌入 MDAM,以探究微观特征增强带来的性能提升;Model B 基于 Baseline 引入 IMVF 多视角迭代融合方案,基础模块不增设 MDAM,用于验证多视角融合策略的有效性。完整模型 PRISM-Net 集成 MDAM 与 IMVF 两大模块。消融实验结果如表 6、图 10 所示。

性能增益。以原始 3D ResNet10 为基线模型,单独嵌入 MDAM 模块后,模型 Precision、Recall 与 F1-Score 分别提升 0.98%、1.40% 与 0.49%,且仅产生微量的参数量增量。相比 MDAM 模块的小幅增益,多视角迭代融合(IMVF)策略对模型性能的提升效果更为显著。仅引入 IMVF 模块时,模型 Precision、Recall、F1-Score 相较基线模型分别提升 3.90%、5.40%、4.25%,分类性能实现大幅优化,同时仅引入极低的参数开销。

当 MDAM 与 IMVF 模块协同作用构成完整 PRISM-Net 时,两类模块形成高效互补的增益

效应,模型性能突破单模块上限,取得最优分类效果。相较于基线模型,完整模型的 Precision、Recall、F1-Score 分别提升 4.83%、6.10%、5.00%,增益幅度显著优于单一模块配置。其中,IMVF 模块从宏观层面拓宽模型特征感知维度,弥补单视角全局特征缺失问题;MDAM 模块则对融合特征进行微观精细化校正与增强,抑制无效特征干扰。

5 结 论

本文围绕肺结核 CT 影像中病灶特征弱、空间信息利用不足以及图像质量差异较大的问题,

构建了融合图像增强与多维特征学习的三分类诊断框架。研究表明,通过联合图像质量优化、多视角空间信息融合以及多维注意力建模,能够有效提升模型对肺结核关键病灶特征的表征能力,增强活动性肺结核、非活动性肺结核与正常肺部之间的区分能力。实验验证了图像增强、多视角融合和注意力机制在肺结核 CT 分类任务中的协同作用,说明充分利用三维空间结构信息有助于提高模型诊断性能和稳定性。受数据规模和样本来源限制,模型的泛化能力仍有进一步提升空间,未来将结合多中心数据和更丰富的临床信息开展研究,以提高模型在复杂临床场景中的适用性和应用价值。

参 考 文 献:

- [1] World Health Organization. Global tuberculosis report 2023 [R]. Geneva: World Health Organization, 2023.
- [2] NIJIATI M, GUO L, ABULIZI A, *et al.* Deep learning and radiomics of longitudinal CT scans for early prediction of tuberculosis treatment outcomes [J]. *European Journal of Radiology*, 2023, 169: 111180.
- [3] 马依迪丽·尼加提,米日古丽·达毛拉,张斌,等. 基于CT深度学习的活动性/非活动性肺结核分类模型构建及验证[J]. 暨南大学学报(自然科学与医学版),2023,44(1):69-77,86.
NIJIATI·MAYIDILI, DAMAOLA·MIRIGULI, ZHANG B, *et al.* Establishment and verification of active/non-active tuberculosis classification model based on CT deep learning [J]. *Journal of Jinan University (Natural Science & Medicine Edition)*, 2023, 44(1): 69-77, 86. (in Chinese)
- [4] WU N F, SUN Y X, HU J T, *et al.* Intelligent nanophotonics: when machine learning sheds light [J]. *eLight*, 2025, 5(1): 5.
- [5] 万丙辰,张选德. 基于元学习和失真感知的图像质量评价[J]. 液晶与显示,2024,39(11):1519-1531.
WAN B C, ZHANG X D. Image quality assessment based on meta-learning and distortion perception [J]. *Chinese Journal of Liquid Crystals and Displays*, 2024, 39(11): 1519-1531. (in Chinese)
- [6] CHENG S Y, CHANG S B, LI Y Z, *et al.* Enhanced multiscale human brain imaging by semi-supervised digital staining and serial sectioning optical coherence tomography [J]. *Light: Science & Applications*, 2025, 14(1): 57.
- [7] ZHANG Z L, KONG L Y, ZHANG L H Y, *et al.* Structured light meets machine intelligence [J]. *eLight*, 2025, 5(1): 26.
- [8] 梁礼明,朱晨锟,阳渊,等. 跨级特征自适应融合的暗光图像增强算法[J]. 液晶与显示,2024,39(6):856-866.
LIANG L M, ZHU C K, YANG Y, *et al.* Cross-level feature adaptive fusion network for low-light image enhancement [J]. *Chinese Journal of Liquid Crystals and Displays*, 2024, 39(6): 856-866. (in Chinese)
- [9] KAMATH V, BHAT V G, RAJU G, *et al.* Application of fluorescence lifetime imaging-integrated deep learning analysis for cancer research [J]. *Light: Advanced Manufacturing*, 2025, 6(3): 522-545.
- [10] ESTEVA A, CHOU K, YEUNG S, *et al.* Deep learning-enabled medical computer vision [J]. *npj Digital Medicine*, 2021, 4(1): 5.
- [11] ARDILA D, KIRALY A P, BHARADWAJ S, *et al.* End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography [J]. *Nature Medicine*, 2019, 25(6): 954-961.
- [12] SIDDIQUI E A, CHAURASIA V, SHANDILYA M. Classification of lung cancer computed tomography images using a 3-dimensional deep convolutional neural network with multi-layer filter [J]. *Journal of Cancer Research and Clinical Oncology*, 2023, 149(13): 11279-11294.

- [13] YADAV V K, SINGHAI J. Adaptive gamma correction for automatic contrast enhancement of chest-X-ray images affected by various lung diseases [J]. *Multimedia Tools and Applications*, 2024, 83(29): 73457-73475.
- [14] YANG S H, PU Q, LEI C T, *et al.* Low-dose CT denoising with a high-level feature refinement and dynamic convolution network [J]. *Medical Physics*, 2023, 50(6): 3597-3611.
- [15] FU M H, DUAN Y H, CHENG Z P, *et al.* Total-body low-dose CT image denoising using a prior knowledge transfer technique with a contrastive regularization mechanism [J]. *Medical Physics*, 2023, 50(5): 2971-2984.
- [16] LI Z Y, LIU Y, CHEN Y, *et al.* Dual-domain fusion deep convolutional neural network for low-dose CT denoising [J]. *Journal of X-Ray Science and Technology*, 2023, 31(4): 757-775.
- [17] LI H, NAN Y, DEL SER J, *et al.* Large-kernel attention for 3D medical image segmentation [J]. *Cognitive Computation*, 2024, 16(4): 2063-2077.
- [18] PERONA P, MALIK J. Scale-space and edge detection using anisotropic diffusion [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, 12(7): 629-639.
- [19] CHEN H, ZHANG Y, KALRA M K, *et al.* Low-dose CT with a residual encoder-decoder convolutional neural network [J]. *IEEE Transactions on Medical Imaging*, 2017, 36(12): 2524-2535.
- [20] 张建强, 何秋生. 上下文感知的低光图像增强算法[J]. 液晶与显示, 2025, 40(5): 751-760.
ZHANG J Q, HE Q S. Context aware low-light image enhancement algorithm [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(5): 751-760. (in Chinese)
- [21] LEQUYER J, PHILIP R, SHARMA A, *et al.* A fast blind zero-shot denoiser [J]. *Nature Machine Intelligence*, 2022, 4(11): 953-963.
- [22] HUANG Z, LI Y, WANG Y, *et al.* Lightweight three-dimensional convolutional neural network for medical image classification [J]. *IEEE Journal of Biomedical and Health Informatics*, 2023, 27(8): 3955-3966.
- [23] 周涛, 程倩茹, 张祥祥, 等. 基于 DCIF-GAN 的肺部肿瘤 PET/CT 跨模态医学图像融合[J]. 光学精密工程, 2024, 32(2): 221-236.
ZHOU T, CHENG Q R, ZHANG X X, *et al.* PET/CT cross-modal medical image fusion of lung tumors based on DCIF-GAN [J]. *Optics and Precision Engineering*, 2024, 32(2): 221-236. (in Chinese)
- [24] 邸敬, 郭文庆, 任莉, 等. 基于结构功能交叉神经网络的多模态医学图像融合[J]. 光学精密工程, 2024, 32(2): 252-267.
DI J, GUO W Q, REN L, *et al.* Multimodal medical image fusion method based on structural functional cross neural network [J]. *Optics and Precision Engineering*, 2024, 32(2): 252-267. (in Chinese)
- [25] WOO S, PARK J, LEE J Y, *et al.* CBAM: convolutional block attention module [C]//*Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018: 3-19.
- [26] 闫纲琦, 梁宗林, 宋延嵩, 等. 基于注意力残差网络的快照式多光谱相机图像重构[J]. 中国光学(中英文), 2024, 17(6): 1316-1328.
YAN G Q, LIANG Z L, SONG Y S, *et al.* Image reconstruction of snapshot multispectral camera based on an attention residual network [J]. *Chinese Optics*, 2024, 17(6): 1316-1328. (in Chinese)
- [27] 邸敬, 梁婵, 刘冀钊, 等. 基于跨域交互注意力和对比学习引导的红外与可见光图像融合[J]. 中国光学(中英文), 2025, 18(2): 317-332.
DI J, LIANG C, LIU J Z, *et al.* Infrared and visible image fusion guided by cross-domain interactive attention and contrastive learning [J]. *Chinese Optics*, 2025, 18(2): 317-332. (in Chinese)
- [28] 朱新军, 赵浩森, 王红一, 等. 基于轻型自限制注意力的结构光相位及深度估计混合网络[J]. 中国光学(中英文), 2024, 17(1): 118-127.
ZHU X J, ZHAO H M, WANG H Y, *et al.* A hybrid network based on light self-limited attention for structured light phase and depth estimation [J]. *Chinese Optics*, 2024, 17(1): 118-127. (in Chinese)
- [29] HENDRIX W, HENDRIX N, SCHOLTEN E T, *et al.* Deep learning for the detection of benign and malignant pulmonary nodules in non-screening chest CT scans [J]. *Communications Medicine*, 2023, 3(1): 156.
- [30] MAFTOUNI M, LAW A C C, SHEN B. A robust ensemble-deep learning model for COVID-19 diagnosis based on an integrated CT scan images database [C]//*Proceedings of the IISE Annual Conference and Expo 2021*. Online:

- IISE, 2021: 632-637.
- [31] JENSSEN H B, SAKINIS T. MedSeg COVID dataset 1 [DB/OL]. (2021-01-05). <https://doi.org/10.6084/m9.figshare.13521488>.
- [32] 上官泓廷,刘玉红. 医学图像增强方法综述[J]. 中国医学物理学杂志,2023,40(4):410-415.
SHANGGUAN H T, LIU Y H. A review of medical image enhancement methods [J]. *Chinese Journal of Medical Physics*, 2023, 40(4): 410-415. (in Chinese)
- [33] LIUFU Y L, SU R H, WEN Y H, *et al.* A CT-based deep learning approach to differentiate multiple primary lung cancers, metastases, and benign nodules [J]. *BMC Cancer*, 2026, 26(1): 177.
- [34] HOWARD A, SANDLER M, CHEN B, *et al.* Searching for MobileNetV3 [C]//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision*. Seoul: IEEE, 2019: 1314-1324.
- [35] MANZARI O N, AHMADABADI H, KASHIANI H, *et al.* MedViT: a robust vision transformer for generalized medical image classification [J]. *Computers in Biology and Medicine*, 2023, 157: 106791.
- [36] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 770-778.

作者简介:



闫小丽,女,硕士研究生,2025年于河南工程学院获得学士学位,主要从事医学影像处理方面的研究。E-mail: gs.xlyan25@gzu.edu.cn



何志琴,女,硕士,教授,2005年于贵州大学获得硕士学位,主要从事人工智能、机器人控制以及医学影像诊断方面的研究。E-mail: zqhe@gzu.edu.cn