

基于多尺度感知与双流融合的尘肺病分期诊断^①



赵晨琦¹, 苏树智^{2,3}, 朱彦敏⁴, 戴勇³

¹(安徽理工大学 计算机科学与工程学院, 淮南 232001)

²(安徽理工大学 合肥综合性国家科学中心大健康研究院 职业医学与健康联合研究中心, 淮南 232001)

³(安徽理工大学 第一附属医院, 淮南 232007)

⁴(安徽理工大学 机电工程学院, 淮南 232001)

通信作者: 戴勇, E-mail: daiyong22@aliyuun.com

摘要: 针对尘肺病患者胸部 X 光片中病灶特征模糊与解剖关联弱可解释性的临床难题分期诊断困难问题, 本文设计了一种基于多尺度感知和 Transformer 的双流融合尘肺病分期诊断模型 (MDC-Net). MDC-Net 通过构建多尺度感知网络与细节保留型编码器的协同架构, 融合局部解剖特征与全局空间关联分析, 显著提升诊断准确性. 首先, 采用全局分组注意力卷积模块增强对微小结节特征的提取能力, 结合可逆残差编码器映射技术保留毛玻璃影等关键征象的纹理细节, 有效解决了纹理丢失问题. 同时, 设计了全局-局部双流融合机制, 通过层级化解剖 Transformer 建立肺门结构与胸膜线的空间关联模型堆叠栈, 打破空间关联建模的瓶颈. 实验结果表明, MDC-Net 在西南医科大学附属医院的 1760 例临床数据上, 对于 I-IV 期诊断的 AUC 值分别达到 0.9877、0.8449、0.7987、0.9912, 相较于 Vision Transformer 在整体准确率上提升了 5.42%, F1 指数提升了 4.92%, 尤其在难以区分的 I 期和 II 期诊断的 AUC 值分别提升了 0.0934 和 0.0426, 有效满足了临床尘肺病的特殊特征与分期诊断需求.

关键词: 尘肺病胸片; 纹理编码器; 可逆残差; 堆叠栈; Transformer

引用格式: 赵晨琦, 苏树智, 朱彦敏, 戴勇. 基于多尺度感知与双流融合的尘肺病分期诊断. 计算机系统应用, 2025, 34(12): 129-138. <http://www.c-s-a.org.cn/1003-3254/10021.html>

Staging Diagnosis of Pneumoconiosis Based on Multi-scale Perception and Dual-stream Fusion

ZHAO Chen-Qi¹, SU Shu-Zhi^{2,3}, ZHU Yan-Min⁴, DAI Yong³

¹(School of Computer Science and Engineering, Anhui University of Science & Technology, Huainan 232001, China)

²(Joint Research Center for Occupational Medicine and Health, Institute of Health and Medicine, Hefei Comprehensive National Science Center, Anhui University of Science & Technology, Huainan 232001, China)

³(The First Hospital of Anhui University of Science & Technology, Huainan 232007, China)

⁴(School of Mechatronics Engineering, Anhui University of Science & Technology, Huainan 232001, China)

Abstract: To address the clinical challenges in pneumoconiosis staging caused by ambiguous lesion features and weak anatomical interpretability in chest X-ray images, this study proposes a dual-stream fusion model for pneumoconiosis staging diagnosis, named MDC-Net, which integrates multi-scale perception and Transformer-based representation learning. By integrating a multi-scale perception network with a detail-preserving encoder, MDC-Net synergistically incorporates local anatomical features and global spatial correlation analysis to significantly improve diagnostic accuracy. A global grouped attention convolution module is introduced to enhance the extraction of micronodule features, while a reversible residual encoder is employed to preserve texture details of critical signs such as ground-glass opacities,

① 基金项目: 国家自然科学基金面上项目 (52374155); 安徽省高等学校自然科学研究项目 (重大项目) (2022AH040113); 安徽理工大学医学专项培育项目 (重大项目) (YZ2023H2A007); 安徽理工大学研究生创新基金 (2024cx2112); 淮南市指导性科技计划 (2023142); 合肥综合性国家科学中心大健康研究院职业医学与健康联合研究中心项目 (OMH-2023-05, OMH-2023-24)

收稿时间: 2025-05-13; 修改时间: 2025-06-03, 2025-06-17; 采用时间: 2025-06-30; csa 在线出版时间: 2025-11-04

CNKI 网络首发时间: 2025-11-05

effectively mitigating texture loss. Furthermore, a global-local dual-stream fusion mechanism is designed, in which a hierarchical anatomical Transformer builds spatial correlations between hilar structures and pleural lines, addressing limitations in spatial relationship modeling. Experimental results on 1 760 clinical cases from Southwest Medical University Hospital show that MDC-Net achieves AUC values of 0.9877, 0.8449, 0.7987, and 0.9912 for stages I to IV, respectively. Compared to Vision Transformer, the proposed model improves overall accuracy by 5.42% and the $F1$ -score by 4.92%. Particularly for the challenging differentiation between stages I and II, AUC values are increased by 0.0934 and 0.0426, respectively. These results effectively meet clinical demands for pneumoconiosis-specific feature characterization and precise staging.

Key words: pneumoconiosis chest radiograph; texture encoder; reversible residual; stacked stack; Transformer

1 引言

尘肺病^[1]作为常见的职业病之一,主要由长期吸入生产性粉尘引发的肺部弥漫性纤维化病变。在医学影像分析领域,基于传统机器学习的诊断方法主要遵循特征工程与模式识别两阶段范式:首先通过人工特征工程提取胸片的纹理特征和形态学指标,继而应用监督学习模型如随机森林(集成决策树框架)^[2]和支持向量机(核函数映射分类器)^[3],或借助非监督学习算法如K均值聚类(基于距离度量的数据分组技术)^[4]进行病理特征的自动化识别与分期。相较于深度学习端到端的特征学习机制,这类方法更依赖领域知识指导下的特征构造过程,虽有效,但未能区分病变严重程度,限制了临床应用。

近年来,深度学习技术在尘肺病X光影像分析中取得显著进展^[5-7],如文献[8,9]通过实验分析了多种在ImageNet数据库上训练的卷积神经网络(CNN),包括VGG16/19、ResNet50、Inception-v3和DenseNet121,结合循环一致性生成对抗网络生成的合成图像,用于从小规模胸部X光数据集中诊断煤工尘肺病。研究表明,利用在ChestX-ray14和ImageNet数据库上训练的CheXNet显著提高了模型的诊断效率。此外,Öztürk等人^[10]指出传统的交叉熵损失函数无法充分描述多标签分期任务中病理共现关系的复杂性。因此,他们提出一种混合模式,结合CNN提取空间特征和Vision Transformer(ViT)捕获病理间的远程上下文信息,采用多分支策略以增强模型对单一类别和病理共现关系的敏感性。文献[11,12]则尝试混合CNN与ViT结构或引入辅助监督机制来增强模型的多尺度表达能力与可解释性。然而,这些方法仍面临着3方面显著局限:其一,现有模型在早期尘肺病灶识别方面精度不足,难以从微弱

纹理中区分病理与正常组织;其二,多数网络结构缺乏对解剖结构空间依赖的深层建模能力,尤其对肺门与胸膜区域间的空间关联缺乏有效捕捉;其三,特征融合方式多采用简单串联或加权平均,未能实现多尺度、跨语义信息的深度耦合。上述问题的根源主要在于两个方面:一是传统深度卷积神经网络架构受限于局部感受野,在处理高维空间依赖与连续分期变化时表现能力不足;二是当前方法对纹理细节和解剖结构多尺度表征的融合机制较为单一,限制了模型的分期判别能力。

为突破现有尘肺病影像分期方法在特征提取、空间建模和分期精度方面的技术瓶颈,本文提出了一种融合多尺度感知机制与双流交互结构的尘肺病智能分期诊断模型——MDC-Net。该模型系统性地引入3大关键技术模块,旨在充分挖掘X光影像中的细粒度病灶特征与解剖结构间的空间依赖关系,从而实现复杂病理特征的高效建模与精准分期。

本文主要工作为:(1)设计基于全局分组注意力机制的感知模块,提升网络对肺部微小结节、毛玻璃影、胸膜斑等细小病灶的敏感性与表达能力。通过在不同尺度下动态感知特征图间的信息交互,有效增强模型对局部细节与全局结构的综合感知能力;(2)引入可逆式残差神经网络结构,构建具备双向映射能力的细节提取模块,在保持特征表达完整性的同时,提升网络反向传播的稳定性与特征可逆性。该模块能够有效捕捉并保留尘肺病灶区域的精细纹理模式,提高早期病变的识别效率并提升模型训练的稳定性;(3)构建异构双流特征交互路径,其中局部通道聚焦于病灶关键区域的纹理细节建模,全局通道则通过层级式Transformer结构建模解剖结构之间的空间语义关联,建立

跨尺度、多区域之间的远程依赖关系. 两路信息在融合阶段实现深度耦合, 提升模型对复杂分期过渡状态的理解能力实现跨尺度特征的深度交互融合.

2 本文方法

本文提出模型总体框架如图1所示. 模型整体由3部分构成. 首先, 输入的X-ray图像被送入多尺度感知网络模块, 该模块基于全局分组注意力机制构建, 能够在不同分辨率特征之间实现信息交互与动态融合, 从而显著增强关键区域的感知能力并生成带权重的多尺度特征图. 并将融合感知到的特征图被并行传入双流结构: 一支为基于可逆神经网络 (invertible neural network, INN) 构建的可逆残差编码器, 专注于局部解剖细节的建模; 另一支为堆叠式Transformer架构, 侧

重捕获远距离区域间的全局语义依赖. 通过自适应融合机制, 对两支分支输出的特征进行动态加权, 综合其在空间分布与语义建模方面的互补性, 生成具有更高判别性的病灶表达. 最后, 该融合特征被输入至后端分期分类器中, 依次经过全局平均池化、展平和全连接层, 输出最终的多类别分期预测结果, 用于辅助尘肺病的临床诊断决策.

2.1 多尺度感知网络

为了提升肺部病灶特征表达能力, 将预处理后的X-ray图像输入到多尺度感知网络(MA)模块中, 其采用全局分组注意力机制以突破传统注意力机制中通道映射与并行卷积框架的局限, 增强局部和全局的跨通道信息交互, 以精准建模肺部X-ray图像中的关键病灶区域. 主体流程如图2所示.

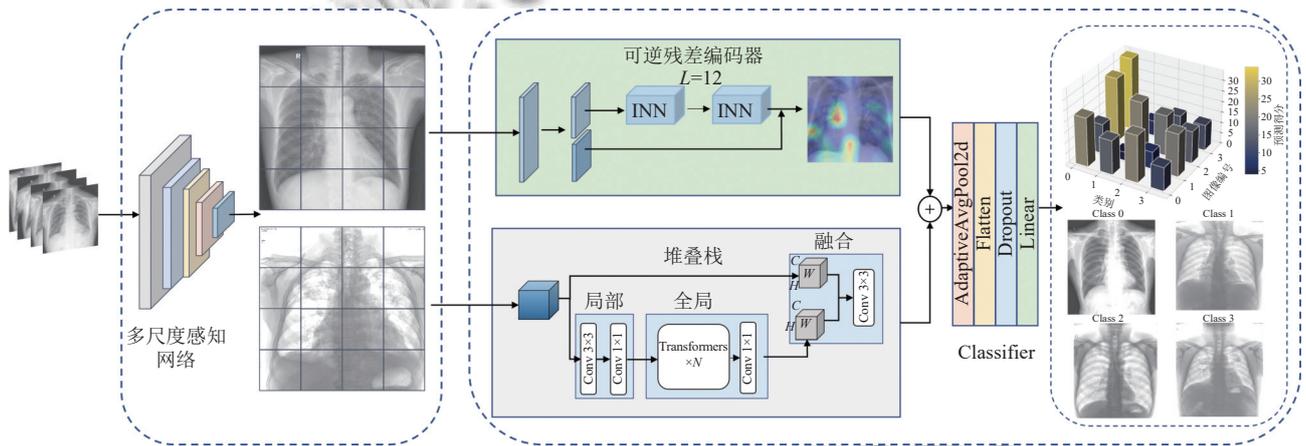


图1 MDC-Net网络结构图

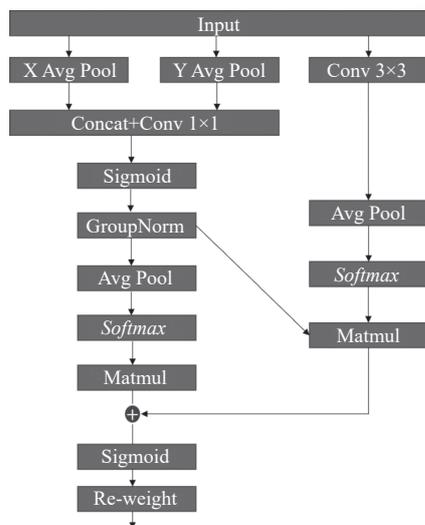


图2 MA模块结构图

在特征分组阶段, 针对输入的肺部病灶特征映射 $X \in R^{C \times H \times W}$, 利用全局平均池化^[13]提取全局的病灶特征 x_g .

$$x_g = \frac{1}{N} \sum_{i=1}^N x_i; N = H \times W \quad (1)$$

此后计算每个通道特征与 x_g 的相似度: $s_i = x_i^t x_g$ 并采用 Softmax 归一化: $S = \text{Softmax}(s)$, 并根据 s 进行通道划分, 将 X 细分为 G 组子特征, 使得各子特征均衡分配空间语义特征信息, 以强化特定病灶区域的表达. 在分组注意力计算阶段采用基于分组的注意力机制计算每组内的特征关联, 对于每个分组, 构造块状稀疏注意力掩码矩阵 $M = \{m_1, m_2, \dots, m_h\} \in R^h$, 确保每个注意力头仅关注所属组内的特征.

$$m_{ij} = \begin{cases} 1, & \text{如果 } x_i, x_j \text{ 属于同一组} \\ 2, & \text{否则} \end{cases} \quad (2)$$

随后利用分组掩码计算注意力:

$$head_i = \left(\text{Softmax} \left(\frac{Q_i K_i^T}{\sqrt{d}} \right) \cdot m_i \right) V_i \quad (3)$$

其中, Q 、 K 、 V 分别由输入特征 X 线性变换得到: $Q = XW_Q, K = XW_K, V = XW_V$.

在跨通道信息交互阶段, 为增强全局依赖关系, MA 在 G 组特征中引入通道间的全局注意力建模. 利用 1D 全局平均池化计算跨通道信息: $x_{\text{global},i} = \text{GAP}(x_i)$, $i \in \{1, \dots, G\}$, 通过全局空间压缩与通道交互学习动态生成通道注意力 $\alpha_i = \sigma(\text{Concat}(x_{\text{global},i}, S_i)W_\alpha + b_\alpha)$, 最终, 将所有通道的加权特征进行聚合.

$$X_{\text{final}} = \sum_{i=1}^G \alpha_i \cdot X_i \quad (4)$$

在多尺度特征融合阶段, MA 通过矩阵点积运算^[14]生成最终注意力权重图: $A = X_{\text{final}}W$, 并施加 S 形函数 (Sigmoid), 并通过权重再次校准输入 X , 以捕捉像素级成对关系: $A' = \sigma(A)$, 然后利用多头注意力进行多尺度特征融合.

$$X_{\text{output}} = \text{Concat}(head_1, head_2, \dots, head_h)W_h \quad (5)$$

将最终的 X_{output} 并行传入双流结构进行更进一步的病灶特征提取分析.

2.2 可逆残差编码器

为了充分捕获细节纹理特征信息, 在双流融合机制中第一分支采用基于可逆神经网络^[15]构建细节编码器进行精准捕捉, 如图 3 所示. 其利用残差式架构构建的如下编码器实现了输入与输出的双向残差映射, 确保特征提取的可逆性, 使每个输入对应唯一输出. 这种映射机制不仅赋予网络卓越的反向传播效率与梯度优化能力, 还能充分保留各模态特征的独特性, 如病灶区域毛玻璃影等. 将 MA 模块中生成的 X_{output} 作为设定输入特征表示为 $I \in R^{H \times W}$ 和 $V \in R^{H \times W \times 3}$, 则提取的细节特征可通过以下综合公式描述:

$$Y_I^D = E_D(Y_I^S), Y_V^D = E_D(Y_V^S) \quad (6)$$

其中, $E_D(\cdot)$ 表示细节编码器.

在后续特征交互过程中, 对病灶特征权重图施加逐元素相乘, 融合来自不同尺度路径的肺部病灶信息,

而逐元素相加进一步整合全局特征, 促进特征聚合. 在 Conv Block 结构中, 每个卷积模块包含多层次计算单元, 包括卷积层、ReLU 激活函数、批量归一化及最大池化. 通过逐层卷积处理, 使网络能够从 X-ray 图像的高频分量中提取更深层次的细节纹理特征.

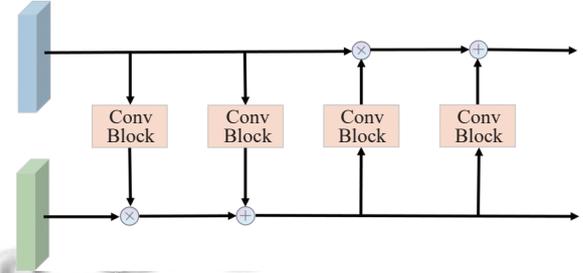


图 3 INN 结构示意图

该交互式归一化网络核心设计在于, 通过异尺度路径的特征提取与融合, 增强输入特征的表达能力, 并通过多级卷积块提取病灶的多尺度病灶细节纹理特征. 最终输出的特征是多级卷积块融合后的结果, 能够精准刻画尘肺图像数据的复杂模式与微观细节, 提升模型的识别能力与泛化表现.

2.3 堆叠栈

在双流融合机制的另一分支中, 本文构建了层级式 Transformer 架构, 通过参数高效配置实现输入张量的多尺度特征建模. 同样将 X_{output} 设为给定输入 $X \in R^{H \times W \times C}$, 首先采用 $n \times n$ 卷积核提取局部空间关联模式, 产生中间特征 $X_L \in R^{H \times W \times d}$, 其中 $d > C$, 该过程包含两个关键操作: (1) $n \times n$ 卷积执行邻域特征聚合; (2) 1×1 卷积实施通道维度线性变换, 进而提升维度强化病灶特征区分度.

随后为突破局部感受野限制, 系统将 X_L 沿空间维度划分为 N 个非重叠区域 $X_U \in R^{P \times N \times d}$. 每个区域经线性投影后输入多个 Transformer 编码器 ($N = 7$), 其核心计算单元为如式 (7) 所示的自注意力机制.

$$\text{Att}(Q^{(P)}, K^{(P)}, V^{(P)}) = \text{Softmax} \left(\frac{Q^{(P)}(K^{(P)})^T}{\sqrt{d_K}} \right) \cdot V^{(P)} \quad (7)$$

其中, 查询矩阵 Q 、键矩阵 K 、值矩阵 V 均由区域特征映射生成, d_K 为归一化因子. 注意力输出经层归一化与残差连接后, 通过前馈网络 (含 Swish 激活函数) 增强非线性.

$$\text{Swish}(x) = x \cdot (1 + e^{-\beta x})^{-1} \quad (8)$$

最终,各区域特征经反折叠操作恢复空间排布,与原始输入 X 进行通道拼接,再经 $n \times n$ 卷积实施跨尺度特征融合形成了一个通过层级化 Transformer 建立肺门结构与胸膜线的空间关联模型.该方法在保留卷积操作空间拓扑约束的前提下,通过 Transformer 堆叠建立跨区域依赖关系,其处理流程包含3个关键阶段:(1)局部特征展开;(2)全局交互建模;(3)多尺度特征重组,相较于传统卷积网络,在肺结节毛玻璃影、胸膜凹陷征等细微征象的跨尺度关联建模方面展现出显著优势.

2.4 双流融合机制

在如图4所示的双流融合机制中,模型并行引入了可逆残差编码器与堆叠栈两个结构,分别针对输入特征的细节表示与全局语义进行建模.二者在对于注意力权重图的特征提取阶段独立执行,以增强不同尺度与不同感受野下的信息表达能力.

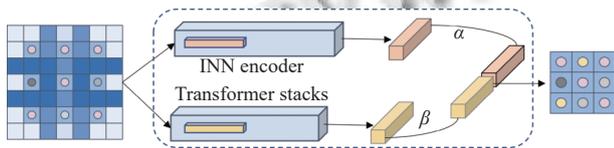


图4 双流融合机制流程图

为实现高效的信息整合,本研究引入了自适应加权融合策略,其中 α 和 β 分别是每个通道特征图通过卷积层获得的融合权重,从而对两个子模块输出的特征图进行加权融合,从而动态平衡局部与全局特征的重要性.融合后的特征被传入为分期诊断任务定制的后端分类器.该分类器首先通过全局平均池化提取全图级别的表示,再经展平操作转化为线性层输入格式.随后引入可选的 Dropout 正则化,以降低过拟合风险并提升模型泛化能力.最终,全连接层输出用于多类别分期预测的类别概率分布,为损失计算与模型优化提供基础.

3 数据集和实验

3.1 数据集

本研究构建的尘肺病影像学分期数据集基于西南医科大学附属医院前瞻性收集的数字化胸片影像资料.图5为部分示例图,数据采集周期为2017年1月–2020年6月,共纳入1760例采用数字放射摄影系统(DR)获取的标准后前位胸片影像,所有病例均经临床确诊存在尘肺相关病理改变.病例筛选遵循多维度质量控制标准是入选对象需满足:(1)明确职业性粉尘暴露史;(2)影像质量符合国际劳工组织(ILO)胸片摄片规范;

(3)阳性病例须经具备尘肺诊断资质的医疗机构复核确认.排除标准涵盖合并气胸、胸腔积液等可能干扰影像判读的胸膜病变,以及接受过肺切除术等改变胸廓解剖结构的病例.队列特征分析显示:入组样本年龄谱覆盖20–95岁(性别构成:男性1501例,女性259例),职业性粉尘暴露持续时间2–29年(均值 6.2 ± 5.6 年).经分层诊断验证,确诊阳性病例1248例(70.9%),阴性对照512例(29.1%).本研究经西南医科大学附属医院伦理审查委员会批准(KY 2020200),机构审查委员会基于回顾性研究性质豁免知情同意要求,所有影像数据采集和处理流程均严格遵循机构标准化协议.

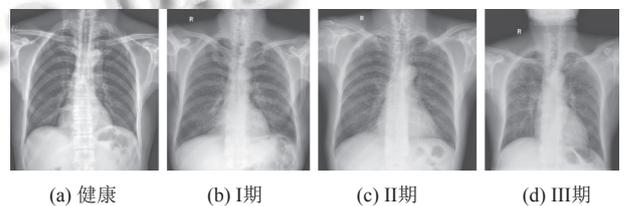


图5 尘肺病分期数据示例图(ILO指引)

本文采用国际劳工组织(ILO)标准化尘肺病影像诊断框架,组建由1名注册放射科医师与4名职业医学专家构成的多学科团队进行独立盲法阅片.诊断过程中系统采集3类关键影像表征参数:①微结节性病灶(小阴影)的分布密度、累及肺区范围、形态学特征(圆形/不规则)及直径分级;②纤维化团块(大阴影)的定位与结构特征;③胸膜异常的定量评估,包括胸膜斑厚度测量、肋膈角闭锁程度及弥漫性胸膜增厚范围.诊断决策树整合了职业暴露史中的石棉接触记录与胸膜复合病变的协同诊断价值,其中胸膜斑的钙化特征与膈肌黏连程度被确立为分期判定的核心指标.基于共识诊断机制,病例被归入4阶分期系统:健康标记为0(健康对照)、I期标记为1(早期纤维化)、II期标记为2(进展期结构性改变)、III期标记为3(晚期广泛纤维化),详见表1所示的图像分期汇总表.

表1 数据汇总

类别	尘肺病	健康
0	—	512
1	352	—
2	478	—
3	418	—

3.2 实验设置

本文整合华中区域某胸科专科医院构建的尘肺病影像生物标记物数据库,该机构尘肺病诊疗团队由

5名具有10年以上尘肺病影像诊断经验的副主任医师及3名参与国家级尘肺病科研项目的临床医师组成。数据集采用前瞻性-回顾性混合构建模式,所有数字化胸片均通过联影uDR780i数字化放射成像系统采集,原始数据经三重校验流程确保影像质量符合ILO B级标准。数据架构实施多中心验证策略,核心诊断组由2名具有尘肺病诊断资质认证的影像学专家独立完成初始判读,争议病例提交多学科会诊(MDT)平台进行仲裁分析。质量控制体系特别设立职业暴露因素校正模块,通过粉尘接触时间指数(DETI)与胸膜钙化评分(PCS)的量化分析,构建诊断效能验证队列。

在数据预处理阶段,实施多模态影像预处理框架,采用3阶优化策略提升数据质量:首先执行基于Z-score标准化算法的全域像素值归一化,通过动态范围调整建立 $N(0, 1)$ 标准分布,消除设备间异质性;继而应用ROI智能定位技术进行自适应裁剪,采用双三次插值算法统一规范化为 224×224 像素矩阵;最后部署非均匀亮度校正模块,结合自适应直方图均衡化与Gamma校正($\gamma = 1.2$)实现跨样本光度一致性控制。针对尘肺病影像小样本学习场景,设计形态学约束增强策略,构建包含空间变换域($\pm 15^\circ$ 随机旋转、垂直镜像)、噪声扰动域($\sigma = 0.02$ 高斯噪声注入)及病理模拟域($5 \times 5 - 15 \times 15$ 像素随机区域擦除)的多维增广体系^[16]。该方案通过参数化控制确保解剖结构保真度,训练集有效扩增至原始数据的5倍规模,同时生成对抗性样本提升模型鲁棒性。

在上述基础上,本研究提出了基于4分期拓扑优化的尘肺病分期诊断体系,采用3阶段训练架构:①混沌模拟退火优化算法初始化网络参数,采用AdamW优化器进行优化,且 $\text{weight_decay} = 1E-3$;②基础学习率设置为 $1E-4$,结合余弦退火调度机制^[17],其中退火周期 T_{\max} 设为10,最小学习率 η_{\min} 为0,并在每轮训练后进行更新;③使用焦点损失,定义如下:

$$FL(p_t) = -\alpha_t \cdot (1 - p_t)^\gamma \cdot \log(p_t) \quad (9)$$

其中, $FL(p_t)$ 表示对一个样本计算得到的焦点损失值, p_t 是模型对正/负类标签的预测概率, α_t 是用于处理类别不均衡的平衡因子,设置为0.25,当 $y = 1$ 时取 α ,当 $y = 0$ 时取 $1 - \alpha$, γ 是聚焦因子,设置为2。实验配置NVIDIA RTX 4060/CUDA 11.7平台,PyTorch端到端训练(batch=64, epoch=50), Dropout=0.2,禁用迁移学习

以验证特征独立性,通过冻结预训练权重对比实验和3次随机种子均值化确保结果可靠性。

3.3 评价指标

为了全面评价模型的性能,本文采用了多个常用的分期指标来评估分期结果,包括ROC曲线、准确率(accuracy)、召回率(recall)、精确率(precision)、F1指数(F1-score)以及混淆矩阵等可以对模型进行评价的图表。由于本研究涉及多类别分期任务,因此计算并报告每个类别的平均值以更全面地评价模型表现,以下为各项指标的计算公式。

$$\text{accuracy} = \sum_{s=0}^{s-1} \frac{TP_s}{N} \quad (10)$$

$$\text{recall} = \frac{TP_s}{TP_s + FN_s} \quad (11)$$

$$\text{precision} = \frac{TP_s}{TP_s + FP_s} \quad (12)$$

$$F1\text{-score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (13)$$

其中, N 为总样本数, TP_s 为表示 s 类的真阳数, FP_s 为表示 s 类的假阳数, FN_s 为表示 s 类的假阴数。

3.4 对比实验

如表2所示,本文系统性评估了多种代表性深度学习架构在尘肺分期任务中的性能边界,包括残差学习的ResNet50^[18]、运用稠密连接思想的DenseNet-121^[13]和使用Transformer的ViT^[19]、轻量级神经网络MobileNet^[20]、CNN与ViT之间的桥梁网络ConvNeXt^[21],以及经典的卷积神经网络VGG16^[22],以及最新图像分类算法MedMamba^[23]。各自代表了不同算法架构特点为模型的分期诊断性能提供多方面的对比和评估。横向对比表明,相较于ViT在整体准确率上提升了5.42%,F1指数提升了4.92%,说明了MDC-Net的多尺度感知网络融合框架实现了较优的准确率-敏感性平衡,验证了跨架构优势集成的临床必要性。

表2 不同模型在尘肺分期数据集上的性能比较(%)

模型	准确率	精确率	召回率	F1-score
ResNet50	72.85	73.63	72.85	72.82
DenseNet121	75.42	76.26	75.42	75.47
VGG16	75.71	75.94	75.71	75.60
ViT	72.28	71.92	72.82	72.28
MobileNet	76.85	77.58	76.57	76.85
ConvNeXt	77.08	77.16	77.21	77.00
MedMamba	70.39	68.76	67.98	69.08
MDC-Net	77.70	77.56	77.69	77.20

为系统评估模型在训练过程中的优化稳定性与收敛性能,可视化了如图6所示的训练集与测试集损失函数变化曲线。从曲线走势可以观察到,模型在前35轮训练过程中仍处于震荡调整阶段,损失值存在一定波动;而在第35轮后,损失收敛速度显著提升,整体趋于平稳,表明模型已进入稳定学习区间。在50轮训练后,两条曲线基本保持收敛状态,未出现显著的过拟合现象,最终训练轮数设置为50轮。进一步从表3中的性能评估结果可见,MDC-Net在尘肺病分期诊断任务中取得了显著优于主流方法的性能表现,尤其在边界模糊、分期判别困难的I期与II期上,模型依然保持出色的分类能力,AUC分别达到0.8449与0.7987。相较于ResNet50模型,本方法在上述两个阶段的诊断准确率分别提升了11.35%与8.48%;并且其AUC值相较于传统ViT架构,则分别提升了9.34%与4.26%。上述结果不仅验证了该模型在建模复杂病灶特征上的有效性,也反映出其在实际分期任务中具备更强的判别能力与泛化表现。

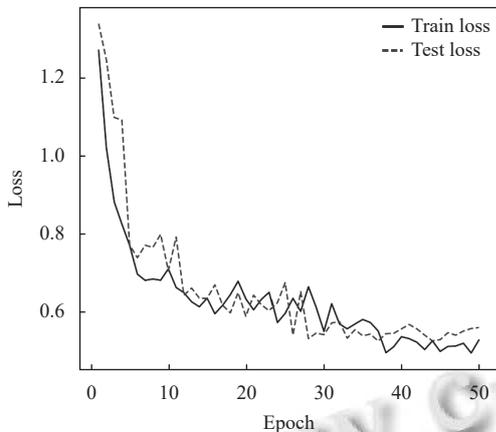


图6 模型损失变化曲线

表3 不同深度学习模型在尘肺病分期数据集上的AUC值

模型	健康	I期	II期	III期
ResNet50	0.9151	0.7314	0.7139	0.9488
DenseNet121	0.9663	0.7394	0.7969	0.9607
VGG16	0.9860	0.8293	0.7047	0.9702
ViT	0.9454	0.7515	0.7561	0.9425
MobileNet	0.9847	0.8348	0.7872	0.9769
ConvNeXt	0.9807	0.8421	0.7748	0.9813
MedMamba	0.9246	0.7405	0.7826	0.9428
MDC-Net	0.9877	0.8449	0.7987	0.9912

进一步结合图7所示的ROC曲线结果,本文模型在健康(Class 0)与III期(Class 3)患者的识别中表现

出显著的分类能力,分别对应极高的AUC值,体现出模型对疾病早期排查和晚期预警的双重临床价值:既可有效减少对健康个体的误判,避免过度随访;又可精准识别重度患者,助力及时干预和治疗。

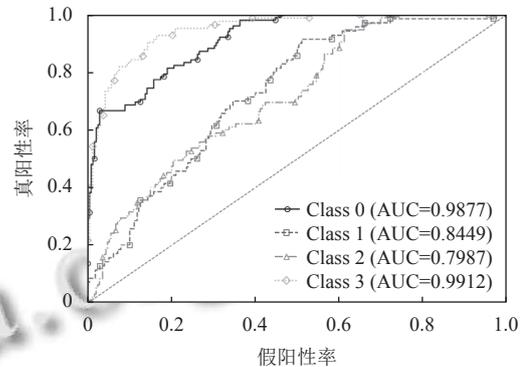


图7 模型ROC曲线

尤其值得关注的是模型在II期(Class 2)与III期(Class 3)之间的区分能力——尽管这两个阶段在影像表现上差异较小、临床症状常有交叉,且传统人工阅片易误判,但模型依然在II期中达到AUC=0.7987,反映出其对这一诊断难点的良好适应性。II期作为尘肺病由轻转重的关键过渡期,具有结构纹理复杂、边界不清等特征,模型通过深度提取纤维化相关的细粒度特征,成功识别出这一“亚临界阶段”的潜在变化,显示出其在辅助分期诊断中的实际应用潜力。另外从图8的多分类混淆矩阵分析可观察到,本模型在各分期之间的分类上具有明显优势,尤其在健康与I期、II期与III期等边界模糊阶段,模型表现出更清晰的类间判别能力,明显优于传统方法。这一性能提升主要归功于所提出的双流融合机制,其在特征提取阶段能够动态捕捉尘肺病肺纤维化区域中的细微纹理与局部结构变化,有效强化了对病灶渐进性演变特征的建模能力。该机制的引入不仅优化了分期决策边界的清晰度,还为临床鉴别尘肺病不同进展阶段提供了更加稳定且可解释的辅助诊断依据,为复杂的尘肺病分期诊断提供了新思路。

本研究提出的MDC-Net方法在尘肺病分期诊断任务中展现出显著优势。针对尘肺病相邻分期(II/III期)的尘肺病胸片影像特征连续渐变特性,MDC-Net实现了较优分期鉴别性能。

如图9所示,在非连续分期(0/I期与III期)的分期任务中,其准确率达到较优水平,显著优于传统分期

评估方法对弥漫性纤维化与胸膜斑形成的量化分析能力.值得注意的是,该方法在各项评价指标中均保持性能均衡性,特别是在不同病理特征的敏感性及特异性平衡方面表现突出,验证了模型在复杂多分期场景中的临床适用价值.

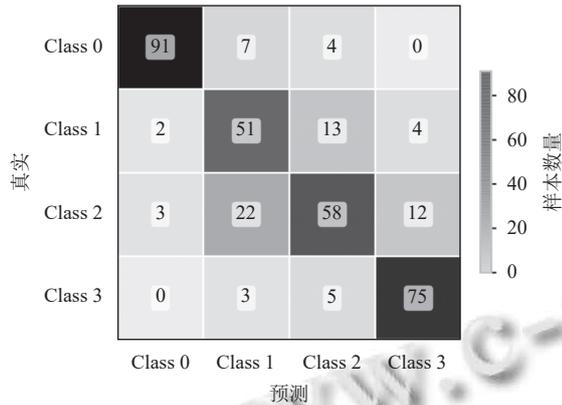


图8 模型混淆矩阵

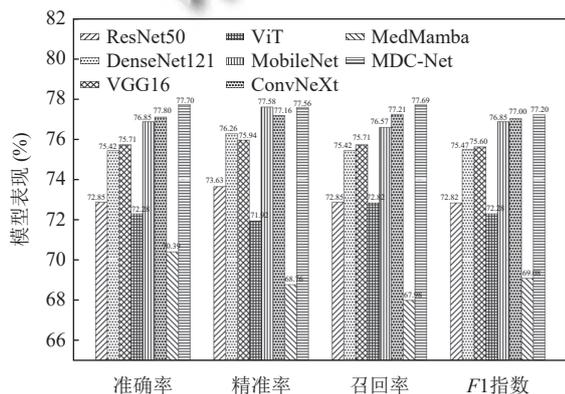


图9 不同深度模型在尘肺病分期诊断性能指标对比

3.5 消融实验

在本文提出的 MDC-Net 模型中,利用全局分组注意力机制提升肺部病灶特征表达能力,增强局部和全局的跨通道信息交互,以精准建模肺部 X-ray 图像中的关键病灶区域,并结合纹理编码器和通过 Transformer 堆叠栈建立肺门结构与胸膜线的全局-局部双流融合机制空间关联模型.为了验证这些模块的有效性,在本研究所使用的尘肺病数据集上进行了消融研究,实验结果如表 4 所示,其中 MA 为全局分组注意力机制基线模型, TB 为基于残差式可逆神经网络构建的细节纹理编码器, SS 为构建的层级式 Transformer 架构,即对原始尘肺病数据进行样本处理后进行训练,从而从消融研究中查验到每个组成模块都在发挥互补的作用.

表 4 消融实验在不同通道模块的尘肺病诊断分期的性能 (%)

模块	准确率	精准率	召回率	F1-score
MA+TB	68.85	70.58	68.85	66.47
MA+SS	74.57	73.44	74.57	73.63
MA+TB+SS	77.70	77.56	77.69	77.20

在使用 MA 模块处理原始 X 光胸片的基础上,引入 TB 和 SS 模块后,尘肺分期诊断性能实现系统性突破:整体准确率值分别提升 8.85% 和 3.13%,且 F1 指数提升了 10.73% 与 3.57%,这一突破性进展源于双流异构架构的协同创新,不仅得益于 TB 中通过可逆残差结构实现参数效率优化,精准定位局灶性纤维化区域,尤其增强对胸膜斑块与肺纹理扭曲的检测鲁棒性和 SS 中层级式 Transformer 堆叠栈的多头自注意力机制建立跨尺度病灶关联,有效建模尘肺结节与肺门血管、肋骨伪影的解剖空间约束关系,更得益于二者构建的双流并行双通道交互式的模型架构,它为模型带来了対病灶特征的远程依赖关系的学习捕获能力,在尘肺病 X 光胸片的解析过程中,这意味着模型现在可以更有效地捕获和整合病灶区域及其周边的关键信息,从而构建出更为丰富和细致的病灶特征表示.

为了进一步验证 MDC-Net 性能,在尘肺病分期诊断数据集中随机选择 4 张胸片,并在图 10 可视化了相应学习特征图.

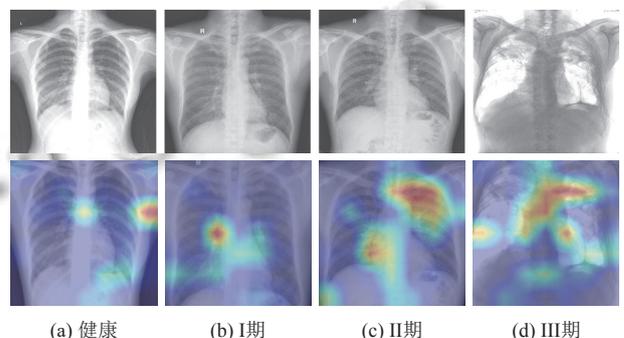


图 10 可视化学习特征图

图 10 中第 1 列为原始胸片,第 2 列为对应的可视化特征图,结果表明,该模型能够精准捕捉尘肺病患者 X 光胸片的不透明区域,并学习尘肺病相关的特征模式且能够在重度尘肺病影像中表现出更高的识别准确度.然而,正常胸片与早期尘肺病影像(图 10(a)和(b))的相似性较高,尤其是在早期病变阶段,影像特征与正常组织界限模糊,增加了识别难度.特征可视化分析验证了模型的决策逻辑,提升了模型的可解释性,印证了微小病灶(小于 2 mm 纤维化结节)与正常肺纹理的视

觉相似性对诊断造成的干扰,显著增强了诊断结果的可解释性。

4 结论与展望

本文模型通过引入全局分组注意力机制以增强感受野和病灶特征提取能力,设计可逆残差纹理编码器以精准保留细节纹理,并构建基于层级式 Transformer 堆叠栈的双流架构,实现局部与全局特征的深度交互融合。多通道信息在并行路径中协同处理,有效提升了尘肺病多阶段影像的识别精度与诊断鲁棒性。在实验验证中,用来自西南医科大学附属医院 1760 例尘肺病组建的数据集实验结果表明, MDC-Net 在不仅能够有效区分健康和 III 期,而且在难以区分的 I 期和 II 期相比其他模型也有较好提升,以证明 MDC-Net 能够适应临床尘肺病的特殊特征与分期任务要求,进一步凸显了医学人工智能领域的潜力。在保证高精度的基础上,如何实现实时性全面区分不同患病程度是今后的研究方向。

参考文献

- 1 Wang JR, Song MY, Fan DP, *et al.* Radiologist-inspired symmetric local-global multi-supervised learning for early diagnosis of pneumoconiosis. *Expert Systems with Applications*, 2025, 276: 127173. [doi: [10.1016/j.eswa.2025.127173](https://doi.org/10.1016/j.eswa.2025.127173)]
- 2 Yu X, Xiong T, Yu L, *et al.* Gut microbiome and metabolome profiling in coal workers' pneumoconiosis: Potential links to pulmonary function. *Microbiology Spectrum*, 2024, 12(11): e0004924. [doi: [10.1128/spectrum.00049-24](https://doi.org/10.1128/spectrum.00049-24)]
- 3 Chen ZJ, Shi JQ, Zhang Y, *et al.* Screening of serum biomarkers of coal workers' pneumoconiosis by metabolomics combined with machine learning strategy. *International Journal of Environmental Research and Public Health*, 2022, 19(12): 7051. [doi: [10.3390/ijerph19127051](https://doi.org/10.3390/ijerph19127051)]
- 4 梁军,洪泽泓,余松森.基于改进粒子群优化算法和遗传变异的图像分割模型. *计算机应用*, 2023, 43(6): 1743-1749.
- 5 Wang ZC, Hu MN, Zeng M, *et al.* Intelligent image diagnosis of pneumoconiosis based on wavelet transform-derived texture features. *Computational and Mathematical Methods in Medicine*, 2022, 2022(1): 2037019. [doi: [10.1155/2022/2037019](https://doi.org/10.1155/2022/2037019)]
- 6 Liu YF, Wu J, Zhou JW, *et al.* Identification of high-risk population of pneumoconiosis using deep learning segmentation of lung 3D images and radiomics texture analysis. *Computer Methods and Programs in Biomedicine*, 2024, 244: 108006. [doi: [10.1016/j.cmpb.2024.108006](https://doi.org/10.1016/j.cmpb.2024.108006)]
- 7 Sun WJ, Wu DS, Luo Y, *et al.* A fully deep learning paradigm for pneumoconiosis staging on chest radiographs. *IEEE Journal of Biomedical and Health Informatics*, 2022, 26(10): 5154-5164. [doi: [10.1109/JBHI.2022.3190923](https://doi.org/10.1109/JBHI.2022.3190923)]
- 8 Devnath L, Luo SH, Summons P, *et al.* An accurate black lung detection using transfer learning based on deep neural networks. *Proceedings of the 2019 International Conference on Image and Vision Computing New Zealand (IVCNZ)*. Dunedin: IEEE, 2019. 1-6. [doi: [10.1109/IVCNZ48456.2019.8960961](https://doi.org/10.1109/IVCNZ48456.2019.8960961)]
- 9 Alam MS, Wang DD, Sowmya A. AMFP-net: Adaptive multi-scale feature pyramid network for diagnosis of pneumoconiosis from chest X-ray images. *Artificial Intelligence in Medicine*, 2024, 154: 102917. [doi: [10.1016/j.artmed.2024.102917](https://doi.org/10.1016/j.artmed.2024.102917)]
- 10 Öztürk S, Turalı MY, Çukur T. HydraViT: Adaptive multi-branch transformer for multi-label disease classification from chest X-ray images. *Biomedical Signal Processing and Control*, 2025, 100: 106959. [doi: [10.1016/j.bspc.2024.106959](https://doi.org/10.1016/j.bspc.2024.106959)]
- 11 Ashraf SMN, Mamun MA, Abdullah HM, *et al.* SynthEnsemble: A fusion of CNN, vision Transformer, and hybrid models for multi-label chest X-ray classification. *Proceedings of the 26th International Conference on Computer and Information Technology (ICCIT)*. Cox's Bazar: IEEE, 2023. 1-6.
- 12 Marikkar U, Atito S, Awais M, *et al.* LT-ViT: A vision transformer for multi-label chest X-ray classification. *Proceedings of the 2023 IEEE International Conference on Image Processing (ICIP)*. Kuala Lumpur: IEEE, 2023. 2565-2569. [doi: [10.1109/ICIP49359.2023.10222175](https://doi.org/10.1109/ICIP49359.2023.10222175)]
- 13 Kumar RL, Kakarla J, Isunuri BV, *et al.* Multi-class brain tumor classification using residual network and global average pooling. *Multimedia Tools and Applications*, 2021, 80(9): 13429-13438. [doi: [10.1007/s11042-020-10335-4](https://doi.org/10.1007/s11042-020-10335-4)]
- 14 Zhang CM, He J, Shang L. An X-ray image classification method with fine-grained features for explainable diagnosis of pneumoconiosis. *Personal and Ubiquitous Computing*, 2024, 28(2): 403-415. [doi: [10.1007/s00779-023-01730-3](https://doi.org/10.1007/s00779-023-01730-3)]
- 15 Wang JM, Lu T, Huang X, *et al.* Pan-sharpening via conditional invertible neural network. *Information Fusion*, 2024, 101: 101980. [doi: [10.1016/j.inffus.2023.101980](https://doi.org/10.1016/j.inffus.2023.101980)]
- 16 Ren XT, Chu SR, Ji GH, *et al.* OMSF2: Optimizing multi-

- scale feature fusion learning for pneumoconiosis staging diagnosis through data specificity augmentation. *Complex & Intelligent Systems*, 2025, 11(1): 109. [doi: [10.1007/s40747-024-01729-0](https://doi.org/10.1007/s40747-024-01729-0)]
- 17 Magdy A, Ismail KN, Mohamed MH, *et al.* PolyRes-Net: A polyhierarchical residual network for decoding anatomical complexity in medical image segmentation. *IEEE Access*, 2025,13:15312–15323.[doi:[10.1109/ACCESS.2024.3475877](https://doi.org/10.1109/ACCESS.2024.3475877)]
- 18 Hassan E, Hossain MS, Saber A, *et al.* A quantum convolutional network and ResNet(50)-based classification architecture for the MNIST medical dataset. *Biomedical Signal Processing and Control*, 2024, 87: 105560. [doi: [10.1016/j.bspc.2023.105560](https://doi.org/10.1016/j.bspc.2023.105560)]
- 19 Li B, Wang JH, Zhao YH, *et al.* Vitamin D3 reverses immune tolerance and enhances the cytotoxicity of effector T cells in coal pneumoconiosis. *Ecotoxicology and Environmental Safety*, 2024, 271: 115972. [doi: [10.1016/j.ecoenv.2024.115972](https://doi.org/10.1016/j.ecoenv.2024.115972)]
- 20 Yang Y, Zhang L, Ren L. Semi-supervised medical image segmentation via cross teaching between MobileNet and MobileViT. *Image and Vision Computing*, 2024, 150: 105196. [doi: [10.1016/j.imavis.2024.105196](https://doi.org/10.1016/j.imavis.2024.105196)]
- 21 Benchallal F, Hafiane A, Ragot N, *et al.* ConvNeXt based semi-supervised approach with consistency regularization for weeds classification. *Expert Systems with Applications*, 2024, 239: 122222. [doi: [10.1016/j.eswa.2023.122222](https://doi.org/10.1016/j.eswa.2023.122222)]
- 22 S YK, Jeya JJ, R MT, *et al.* Explainable lung cancer classification with ensemble transfer learning of VGG16, ResNet50 and InceptionV3 using grad-cam. *BMC Medical Imaging*, 2024, 24(1): 176. [doi: [10.1186/s12880-024-01345-x](https://doi.org/10.1186/s12880-024-01345-x)]
- 23 Yue YB, Li ZZ. MedMamba: Vision mamba for medical image classification. arXiv:2403.03849, 2024.

(校对责编: 张重毅)