❖医学物理与工程学

Improved ConvNeXt model for identifying lung adenocarcinoma in CT images under small-sample conditions

YANG *Junjie*, CHEN *Zhaoxue**

(*Institute of Medical Imaging Technology, School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai* 200093, *China*)

[**Abstract**] **Objective** To observe the value of improved ConvNeXt model for identifying lung adenocarcinoma in CT images under small-sample conditions. **Methods** A total of 1 000 high-resolution CT images were obtained from publicly available LungCancer4Types dataset and divided into training set (n=613), test set (n=315) and validation set (n=72). Based on ConvNeXt model, hybrid attention Transformer (HAT) module and SCConv convolution module were embedded to construct an improved ConvNeXt model. The performance of the obtained model for identifying lung adenocarcinoma in CT images was evaluated using accuracy, precision, recall, F1 score and area under the curve (AUC) of receiver operating characteristic curve, which were all compared with those of GoogLeNet, EfficientNet, DenseNet and AlexNet models. The impact of introducing SCConv and HAT modules on model performance through ablation experiments was observed. **Results** The accuracy, precision, recall, F1 score and AUC of improved ConvNeXt model for identifying lung adenocarcinoma in test set was 95.91%, 95.78%, 96.93%, 0.964 and 0.975, respectively, all superior to the those of GoogLeNet, EfficientNet, DenseNet and AlexNet models. The baseline ConvNeXt model exhibited relatively lower accuracy, precision, recall and F1 score, the introduction of SCConv module significantly increasd model recall, and the introduction of HAT module significantly improved accuracy. The improved ConvNeXt model had the best overall performance. **Conclusion** Improved ConvNeXt model could effectively identify lung adenocarcinoma in CT images under small-sample conditions.

[**Keywords**] lung neoplasms; adenocarcinoma; tomography, X-ray computed; deep learning

DOI:10.13929/j.issn.1003-3289.2025.12.025

基于改进 ConvNeXt 模型在小样本条件下识别 CT 图像中的肺腺癌

杨俊杰,陈兆学*

(上海理工大学健康科学与工程学院医学影像技术研究所,上海 200093)

[摘　要] **目的** 观察基于改进 ConvNeXt 模型在小样本条件下识别 CT 图像中的肺腺癌的价值。**方法** 纳入来源于公开 LungCancer4Types 肺肿瘤高分辨率 CT 图像数据集的 1 000 幅 CT 图像,将其分为训练集(n=613)、测试集(n=315)与验证集(n=72)。在 ConvNeXt 模型基础上引入混合注意力 Transformer(HAT)模块和 SCConv 卷积模块构建改进 ConvNeXt 模型,利用准确率、精确度、召回率、F1 分数和受试者工作特征曲线的曲线下面积(AUC)评价模型识别肺腺癌效能,并与 GoogLeNet、EfficientNet、DenseNet 及 AlexNet 模型进行比较。以消融实验观察引入 SCConv 及 HAT 模块后对模型性能的影响。**结果** 改进 ConvNeXt 模型识别测试集肺腺癌的准确率、精确度、召回率、F1 分数及 AUC 分别为 95.91%、95.78%、96.93%、0.964 及 0.975,均高于 GoogLeNet、EfficientNet、DenseNet 及 AlexNet 模型。基础 ConvNeXt 模型识别肺腺癌的准确率、精确度、召回率及 F1 分数整体较低,引入 SCConv 模块后模型召回率显著提升,引入 HAT 模块后准确率显著提高;改进

[第一作者] 杨俊杰(2001—),男,江苏连云港人,在读硕士。研究方向:医学影像技术。E-mail: 2250828541@qq.com

[通信作者] 陈兆学,上海理工大学健康科学与工程学院医学影像技术研究所,200093。E-mail: chenzhaoxue@163.com

[收稿日期] 2025-07-26　[修回日期] 2025-11-13

ConvNeXt 模型整体性能最优。**结论**　改进 ConvNeXt 模型能在小样本条件下有效识别 CT 图像中的肺腺癌。

［**关键词**］　肺肿瘤；腺癌；体层摄影术，X 线计算机；深度学习

［**中图分类号**］　R734.2；R814.42　［**文献标识码**］　A　［**文章编号**］　1003-3289(2025)12-2056-05

肺癌是全球发病率和死亡率最高的恶性肿瘤，其中肺腺癌约占 40%[1-2]。病理学是评估肺癌分期及类型的金标准；胸部 CT 对早期筛查与术前评估具有关键作用，提升其检出与鉴别能力对于精准诊断肺腺癌和个体化治疗尤为重要。深度学习（deep learning，DL）模型已在评估肺肿瘤病理分型方面展现出一定潜力，但受限于小样本数据和模型可解释性不足，其临床应用仍面临挑战。本研究提出一种基于 CT 图像的改进 ConvNeXt 模型，观察其在小样本条件下识别肺腺癌的价值。

1　资料与方法

1.1　研究对象　纳入来源于公开 LungCancer4Types 肺肿瘤高分辨率 CT 图像数据集中的 1 000 幅 CT 图，均对应术后病理完成精确标注[3]，涵盖腺癌、鳞状细胞癌、大细胞癌及小细胞癌共 4 种肺癌常见病理类型；将标注为腺癌的图像定义为正类（label＝1）、非腺癌为负类（label＝0）；除标签重映射外，均依照原数据集设置，将图像分为训练集（n＝613）、测试集（n＝315）与验证集（n＝72）。

1.2　预处理　采用最小-最大归一化将原始 CT 图像灰度值线性映射至［0，1］区间，以弱化噪声和量纲差异对于训练网络的影响；利用旋转（90°、180°、270°）、水平/垂直轴向翻转等几何变换方法将训练集数据规模扩充 5 倍。

1.3　混合注意力 Transformer（hybrid attention Transformer，HAT）模块　HAT 模块包含浅层特征提取、深层特征提取和图像重建 3 部分。浅层特征提取通过一个卷积层提取输入图像的初始特征；深层特征提取由多个残差混合注意力组（residual hybrid attention group，RHAG）构成，每个 RHAG 包含若干混合注意力块（hybrid attention block，HAB）、1 个重叠交叉注意力块（overlapping cross attention block，OCAB）及 1 个带残差连接的 3×3 卷积层，其中的 HAB 在 Swin-Transformer 窗口自注意力基础上增加通道注意力分支、引入全局统计信息，OCAB 通过滑动窗口方式建立跨窗口的 Query-Key/Value 交互、扩大有效感受野并强化局部与全局信息融合；图像重建利用像素重组上采样融合特征并输出重构结果[4-5]。见图 1。

1.4　SCConv 卷积模块　SCConv 主要由空间重构单元（spatial reconstruction unit，SRU）和通道重构单元（channel reconstruction unit，CRU）串联而成[6]。SRU 利用归一化层缩放因子（γ）评估特征重要性，并基于门控机制进一步区分重要（高 γ）与非重要（低 γ）特征，之后通过交叉重构增强信息流动、抑制冗余。CRU 按一定比例对通道进行分组与压缩，结合分组卷积、1×1 卷积和注意力加权，以自适应融合不同通道分支的特征。见图 2。

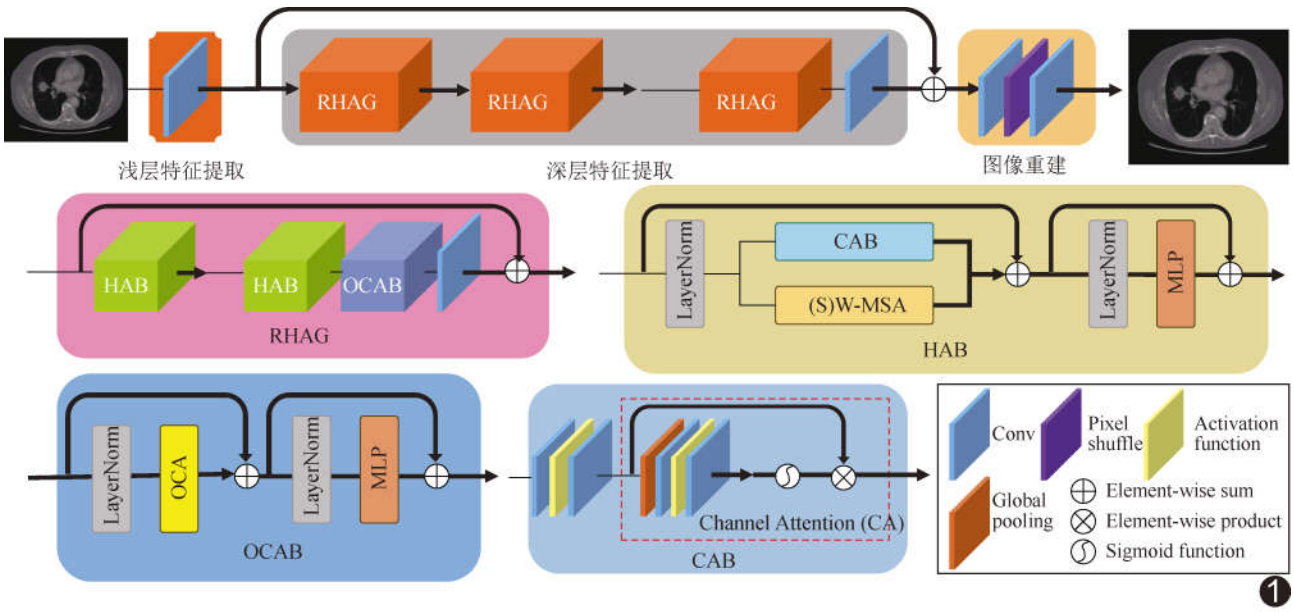

图 1　HAT 模块架构图

1.5　改进 ConvNeXt 模型　基础 ConvNeXt 模型如图 3 所示，对其加以如下改进：①以 ConvNeXt-T 为基准结构，以 SCConv 模块替代各阶段卷积块中原有深度可分离卷积，以减少冗余特征、提升参数利用率；②在SCConv 后插入 HAT 模块，遵循“先压缩冗余，再增强注意”顺序，使网络在精简特征基础上进一步强化对于关键信息的关注；③保持原有 Stage 1～4 分层结构及 Patch Embedding 配置不变，对 SCConv 和 HAT 模块的通道分割比例、卷积核大小等关键超参数进行适配性调整，以兼顾性能与计算开销。见图 4。

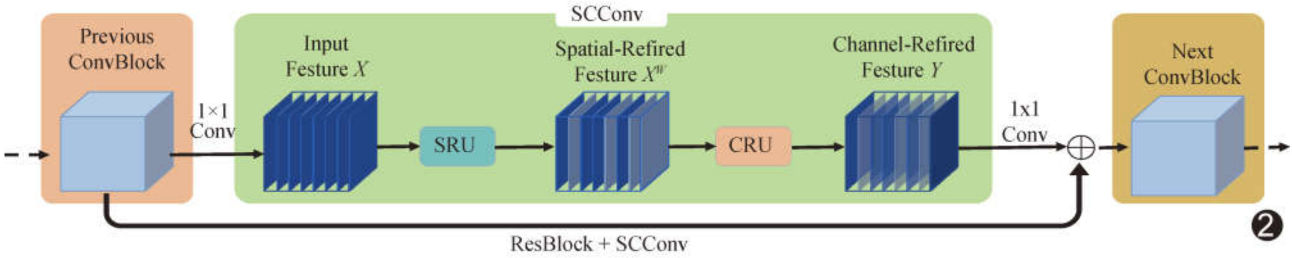

图 2　SCConv 模块示意图

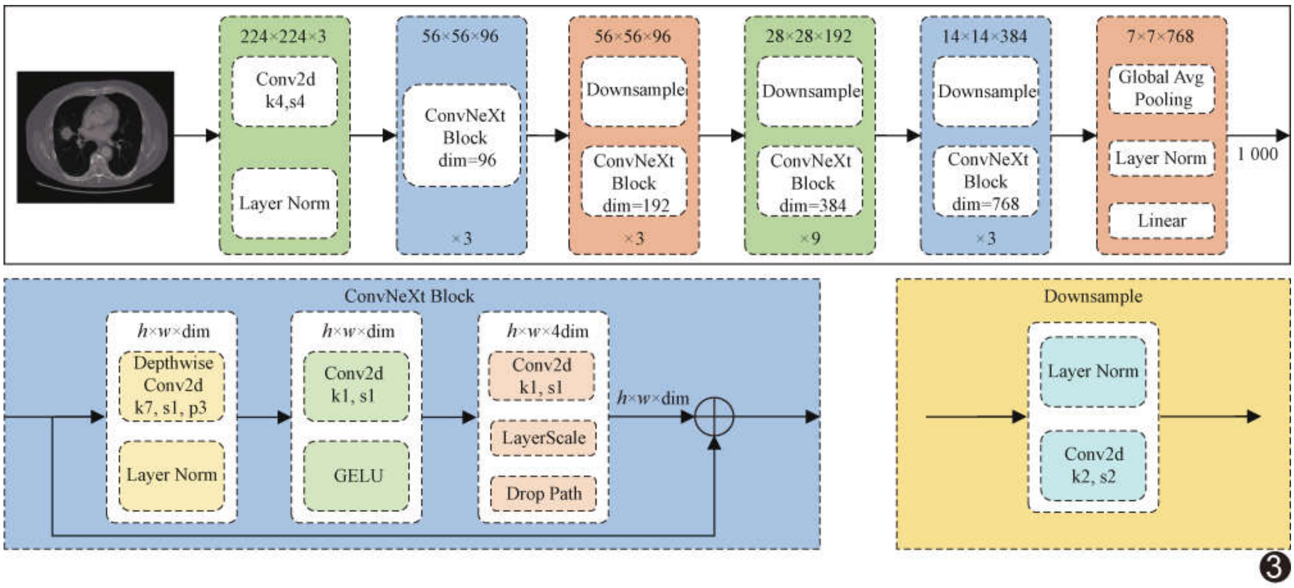

图 3　基础 ConvNeXt 模型示意图

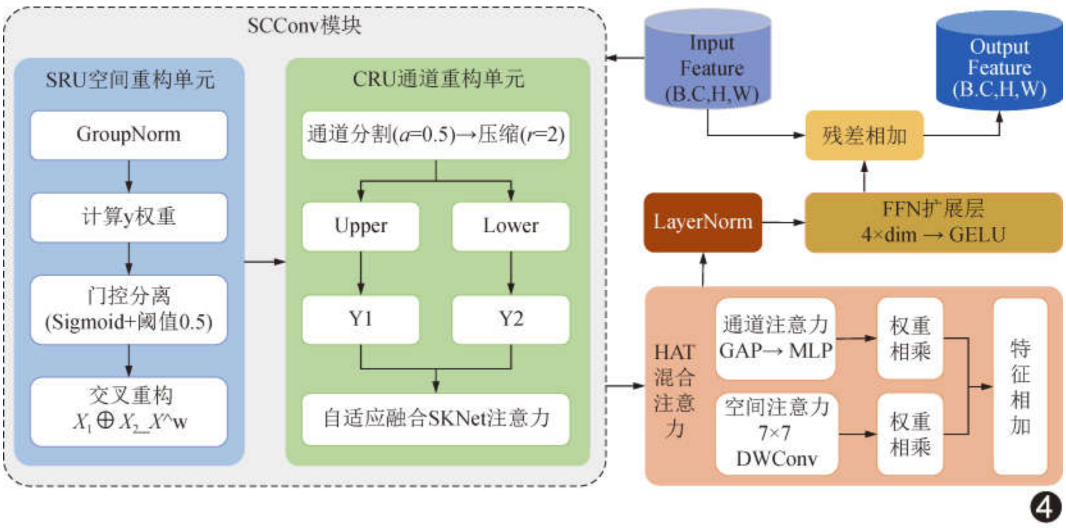

图 4　改进 ConvNeXt 模型重点模块示意图

1.6 实验环境及评价指标 在 Microsoft Windows 11 操作系统下训练并测试模型,CPU 型号 AMD Ryzen 7 5800H,GPU 型号 GeForce RTX 3060、可用显存 12GB;基于 CUDA 11 加速框架进行训练,使用 Python 3.7 编程语言及 Pytorch DL 编程框架。

将所获模型与通过 ImageNet 等公开数据集验证的 GoogLeNet、EfficientNet、DenseNet 及 AlexNet 模型进行对比。对各模型训练时的超参数进行标准化,训练过程中均采用 Adam 梯度优化算法,设初始学习率为 0.001,模型 BatchSize 设为 32,训练轮次 100,选择交叉熵作为损失函数。

以训练集数据训练模型,每轮训练结束后基于验证集数据计算模型各项性能指标,综合每轮训练数据观察模型收敛速度等信息;最后以测试集数据对模型性能进行最终测试。采用准确率、精确度、召回率、F1 分数及受试者工作特征曲线的曲线下面积(area under the curve, AUC)评价模型在小样本条件下识别 CT 图像中肺腺癌的效能。

1.7 消融实验 于基础 ConvNeXt 模型中依次加入 SCConv 模块、HAT 模块及 SCConv 模块+HAT 模块(即改进 ConvNeXt),观察上述因素对模型性能的影响,并以梯度加权类激活映射(gradient-weighted class activation mapping, Grad-CAM)叠加图可视化。

2 结果

改进 ConvNeXt 模型识别测试集肺腺癌的准确率、精确度、召回率、F1 分数及 AUC 分别为 95.91%、95.78%、96.93%、0.964 及 0.975,均高于 GoogLeNet、EfficientNet、DenseNet 及 AlexNet 模型(表 1)。

基础 ConvNeXt 模型识别肺腺癌的准确率、精确度、召回率及 F1 分数均最低;引入 SCConv 模块后模型召回率显著提升,引入 HAT 模块后模型准确率显著提高,同时引入 HAT 及 SCConv 模块后的改进 ConvNeXt 模型整体性能最优(表 2)。Grad-CAM 叠加图显示,改进 ConvNeXt 模型对于肺实质内病灶区域产生显著高响应,而对非瘤肺组织和纵隔区域响应较低,与实际病灶分布较为一致(图 5)。

表 1 各模型识别肺腺癌效能

模型	集别	准确率(%)	精确度(%)	召回率(%)	F1 分数	AUC
GoogLeNet 模型	测试集	78.11	80.25	82.52	0.780	0.884
	验证集	91.16	94.18	91.14	0.926	0.946
EfficientNet 模型	测试集	86.35	86.86	89.81	0.874	0.885
	验证集	77.13	82.48	73.42	0.776	0.814
DenseNet 模型	测试集	90.79	91.58	92.45	0.915	0.921
	验证集	85.96	89.98	81.71	0.857	0.887
AlexNet 模型	测试集	66.03	78.58	70.36	0.742	0.799
	验证集	88.72	89.86	86.65	0.882	0.901
改进 ConvNeXt 模型	测试集	95.91	95.78	96.93	0.964	0.975
	验证集	93.51	95.32	92.84	0.910	0.958

表 2 消融实验中各模型识别肺腺癌效能

模型	准确率(%)	精确度(%)	召回率(%)	F1 分数	可训练参数量(M)
基础 ConvNeXt 模型	88.21	84.32	85.22	0.847	28
基础 ConvNeXt 模型+SCConv 模块	91.62	82.23	90.33	0.861	33
基础 ConvNeXt 模型+HAT 模块	93.83	92.32	84.61	0.883	42
改进 ConvNeXt 模型	96.62	95.66	93.92	0.948	47

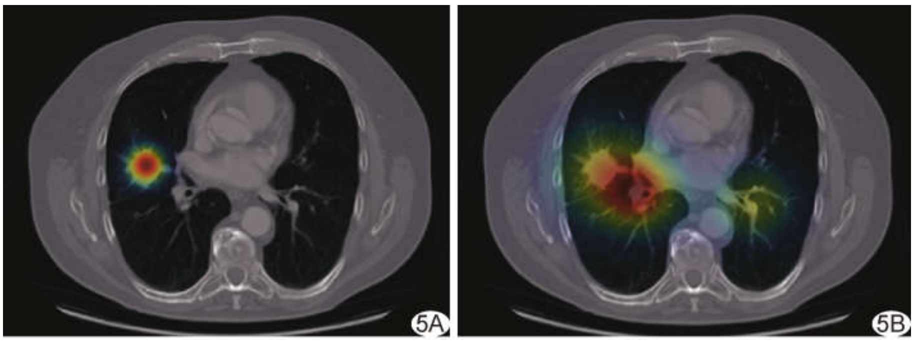

图 5 改进 ConvNeXt 模型(A)与基础 ConvNeXt 模型(B)识别肺腺癌的 Grad-CAM 叠加图

3 讨论

多项研究[7-11]证实,DL CT 图像分析技术用于检测肺结节、评估恶性风险及预测其组织学亚型等具有显著优势,但通常依赖大规模训练数据,模型结构较为复杂且可解释性有限,限制了其临床推广应用[12-14]。本研究重点关注 DL 模型在小样本条件下的鲁棒性与可解释性,基于 LungCancer4Types 单一公开数据集训练、测试模型并对其验证,以更好地控制变量,避免多数据集直接融合引入噪声及各公开数据库病理标注体系不一致等问题。

ConvNeXt 是一类面向视觉任务的新型卷积网络,在保持纯卷积主干结构的同时引入大卷积核、层归一化、高斯误差线性单元(Gaussian error linear units, GELU)激活和分层设计等视觉 Transformer 中的关键元素,在

ImageNet 等基准数据集中表现出不亚于主流 Transformer 模型的性能[15]。近年 Transformer 网络在图像复原和超分辨等任务中的表现已优于传统卷积神经网络(convolutional neural network, CNN),但纯粹自注意力结构使其在局部细节建模和跨窗口信息交互方面存在不足。CHEN 等[16]提出的 HAT 模块结合通道注意力与基于窗口的自注意力,通过引入重叠窗口交互机制有效缓解了窗口划分带来的块状伪影及信息隔离问题。本研究借鉴 HAT 的混合注意力与重叠交叉注意力的设计思路,将其嵌入 ConvNeXt 主干的特征提取阶段,以增强模型对于肺腺癌病灶区域显著特征的聚焦能力,使其能在保持较高计算效率的同时具备更强的全局建模能力。SCConv 是一种即插即用的卷积模块,通过降低特征的空间冗余和通道冗余而在有限计算量下提升网络的表征能力;相比仅在单一维度(空间或通道)进行优化的传统方法,SCConv 能在不改变整体网络拓扑结构的前提下直接替换标准卷积层,已为多种主干网络所采用。

GoogLeNet 引入 Inception 多通路结构,可在单层中并行获取不同尺度信息;EfficientNet 利用复合尺度扩展策略而在算法效率与模型性能之间达到较好平衡;DenseNet 通过特征层间的密集连接促进梯度传播与深层特征共享;AlexNet 则体现了基础卷积结构在特征提取方面的能力。相比上述各类型 DL 架构,本研究所获模型识别测试集肺腺癌的准确率、精确度、召回率、F1 分数和 AUC 均更高,分别为 95.91%、95.78%、96.93%、0.964 及 0.975,证实改进 ConvNeXt 模型在小样本条件下识别肺腺癌效能优异。此外,改进 ConvNeXt 模型在精确度与召回率间达到更好的平衡,训练收敛速度更快,表明该架构在样本量有限条件下仍具有较强的稳定性与泛化能力。本研究消融实验进一步证明 SCConv 和 HAT 模块均对 ConvNeXt 性能提升起关键作用,联合应用二者更有助于在小样本场景下降低 CT 图像冗余特征、提升有效特征密度。Grad-CAM 结果显示,改进 ConvNeXt 模型关注区域主要集中于肺实质病灶而非非瘤肺组织,与临床诊断逻辑一致,提升了模型决策过程的可解释性。

综上所述,基于改进 ConvNeXt 模型能在小样本条件下有效识别 CT 图像中的肺腺癌。本研究的主要局限性:数据均来源于单一公开数据集,且未进一步划分非腺癌亚型,限制了模型对于更细粒度分型任务的适用性;缺乏来自不同研究中心、不同扫描设备或成像参数的外部验证数据,对于模型用于跨机构数据的鲁棒性与迁移性能仍需进一步评估。

利益冲突:全体作者声明无利益冲突。

作者贡献:杨俊杰研究设计、数据分析、撰写文章;陈兆学指导、研究实施、修改文章。

[参考文献]

[1] 蔡林,朱称心,张兴龙,等.全球肺癌流行数据解读[J].中华流行病学杂志,2024,45(4):585-590.

[2] 郑雅琳,邸吉廷,熊焰,等.Ⅰ～Ⅱ期肺腺癌 PD-L1 表达的临床病理意义[J].中华病理学杂志,2021,50(7):796-798.

[3] UDDIN M B, SULTANA J. LungCancer4Types: ImageDataset[EB/OL]. [2025-07-16]. https://data. mendeley. com/datasets/bncfhxbzrv/1.

[4] LIANG J, CAO J, SUN G, et al. Swinir: Image restoration using swin transformer[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Montreal: IEEE,2021:1833-1844.

[5] ZHANG Y, LI K, LI K, et al. Image super-resolution using very deep residual channel attention networks[C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich: Springer, 2018:286-301.

[6] LI J, WEN Y, HE L. Scconv: Spatial and channel reconstruction convolution for feature redundancy[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Paris: IEEE,2023:6153-6162.

[7] ZHU W, JIN Y, MA G, et al. Classification of lung cancer subtypes on CT images with synthetic pathological priors[J]. Med Image Anal, 2024,95:103199.

[8] XU X, WANG C, GUO J, et al. DeepLN: A framework for automatic lung nodule detection using multi-resolution CT screening images[J]. Knowledge-Based Systems, 2020,189:105128.

[9] LEE T, LEE K H, LEE J H, et al. Prognostication of lung adenocarcinomas using CT-based deep learning of morphological and histopathological features: A retrospective dual-institutional study[J]. Eur Radiol, 2024,34(5):3431-3443.

[10] MASSION P P, ANTIC S, ATHER S, et al. Assessing the accuracy of a deep learning method to risk stratify indeterminate pulmonary nodules[J]. Am J Respir Crit Care Med, 2020,202(2):241-249.

[11] BALDWIN D R, GUSTAFSON J, PICKUP L, et al. External validation of a convolutional neural network artificial intelligence tool to predict malignancy in pulmonary nodules[J]. Thorax, 2020,75(4):306-312.

[12] ZHANG H, CHEN L, GU X, et al. Faithful learning with sure data for lung nodule diagnosis[J]. arXiv preprint arXiv, 2202:12515.

[13] 高杭宇,杨阳,李建英.人工智能技术在肺结节影像学诊断中的应用:进展、挑战与展望[J].临床个性化医学,2024,3(4):1896-1902.

[14] WANG C, SHAO J, LV J, et al. Deep learning for predicting subtype classification and survival of lung adenocarcinoma on computed tomography[J]. Transl Oncol, 2021,14(8):101141.

[15] LIU Z, MAO H, WU C Y, et al. A ConvNet for the 2020s[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE,2022:11976-11986.

[16] CHEN X, WANG X, ZHOU J, et al. Activating more pixels in image super-resolution transformer[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Paris: IEEE,2023:22367-22377.