肺癌 CT 影像分类与病灶分割方法研究

张钧浩，冯泓翔

大连东软信息学院健康医疗科技学院　辽宁大连

【摘要】以 IQ-OTH/NCCD 肺部 CT 数据集(1190 张)为对象,采用限制对比度自适应直方图均衡化(CLAHE)增强、高斯滤波、非锐化掩模（USM）锐化预处理提升病灶特征；利用 ImageNet 预训练视觉几何组 16 层网络（VGG16）两阶段迁移学习实现三分类；传统图像处理自动生成伪标签掩膜监督 U-Net 实现像素级病灶分割。实验表明：预处理 VGG16 准确率 98.74%，优于 ResNet50（97.48%）和 EfficientNet-B0（97.06%）；U-Net 分割像素准确率 99.29%、IoU 65.3%、Dice 78.6%。方法无需手工像素标注，实现分类与分割联合诊断。

【关键词】肺部 CT 影像；VGG16；迁移学习；U-Net；语义分割

【收稿日期】2026 年 8 月 5 日　　【出刊日期】2026 年 9 月 3 日　　【DOI】10.12208/j.ijcr.20260441

Research on lung cancer CT classification and lesion segmentation based on transfer learning

Junhao Zhang, Hongxiang Feng

Dalian Neusoft University of Information, School of Health and Medical Technology, Dalian, Liaoning

【**Abstract**】Taking the IQ-OTH/NCCD lung CT dataset (1190 images) as object, Contrast-Limited Adaptive Histogram Equalization (CLAHE) enhancement, Gaussian filtering and Unsharp Mask (USM) sharpening preprocessing are used to improve lesion features. Two-stage transfer learning based on ImageNet pretrained Visual Geometry Group 16-layer network (VGG16) realizes three-category classification. Pseudo-label masks generated by traditional image processing supervise U-Net for pixel-level lesion segmentation. Experiments show VGG16 accuracy reaches 98.74%, superior to ResNet50 (97.48%) and EfficientNet-B0 (97.06%). U-Net achieves pixel accuracy 99.29%, IoU 65.3% and Dice 78.6%. This method realizes joint diagnosis without manual pixel annotation.

【**Keywords**】Lung CT image; VGG16; Transfer learning; U-Net; Semantic segmentation

1 引言

肺癌是中国发病率和死亡率首位的恶性肿瘤，2022 年新发病例超 106 万例、死亡超 73 万例[1]。低剂量螺旋 CT 是唯一可降低肺癌死亡率的筛查手段,但医师阅片工作量大，微小病灶漏诊率高。现有 AI 方案多依赖人工像素标注，部署成本高，且多数仅实现单一分类。Shen 等[2]提出多裁剪 CNN 分类，Ardila 等[3]构建三维深度学习筛查系统；Ronneberger 等[4]提出的 U-Net 是医学分割标准架构。本文提出无手工像素标注的双模型联合方法：VGG16 迁移学习高精度三分类，伪标签监督 U-Net 分割。

2 相关理论与方法

（1）VGG16 网络与迁移学习

VGG16 采用 3×3 卷积核堆叠，共 13 层卷积+3 层全连接，5 个卷积块逐级提取特征[5]，结构如图 1 所示。

两阶段迁移学习：第一阶段冻结 13 层卷积，仅训练顶层分类器，损失为分类交叉熵：

$$L = -（1/N）\Sigma y_i \cdot \ln（p_i）\quad（1）$$

第二阶段解冻 block4 和 block5 共 6 层卷积，学习率 1×10^{-5} 细粒度调优，学习率衰减策略：

$$lr_t = lr_0 \times \beta^t\quad（2）$$

两阶段策略保留预训练通用特征，同时适配医学影像，缓解小样本过拟合。

（2）U-Net 语义分割网络

U-Net 呈 U 形对称：编码器 4 级下采样通道 64→1024，解码器 4 级上采样通道 512→64，跳跃连接融合底层细节与高层语义[6]，结构如图 2 所示。

输出层 Sigmoid 激活映射像素至[0,1]病灶概率，二元交叉熵与 Dice 损失：

作者简介：张钧浩（2003-）男，辽宁铁岭人，本科在读，研究方向：计算机视觉与医学图像处理；冯泓翔（2005-）男，河北石家庄人，本科在读，研究方向：深度学习与医学影像分析。

$$y = 1/（1 + e^{-x}）\tag{3}$$

$$L_BCE = -(1/N)\Sigma[y_i \cdot \ln(\hat{y}_i) + (1 - y_i) \cdot \ln(1 - \hat{y}_i)]\tag{4}$$

$$L_Dice = 1 - （2\Sigma y_i \cdot \hat{y}_i）/（\Sigma y_i + \Sigma \hat{y}_i）\tag{5}$$

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

图 1　VGG16 网络层级结构

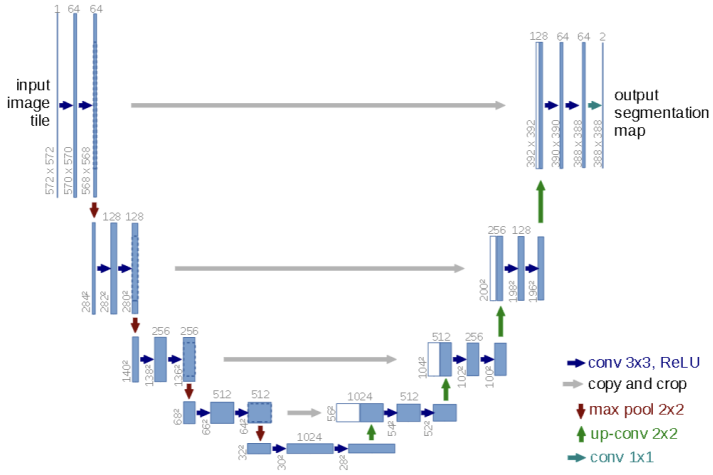


图 2　U-Net 网络架构

3 数据集与图像预处理

采用 IQ-OTH/NCCD 公开数据集，1190 张 CT 切片：Normal 415 张、Benign 481 张、Malignant 294 张。预处理流水线[7]：①CLAHE 均衡化（clipLimit=2.0，8×8 子区域）；②5×5 高斯滤波；③USM 锐化（α=1.5）；④归一化[0,1]，统一缩放 224×224。核心公式：

$$h'（k）= \min（h（k），T）+ \text{redistribute}（k）\tag{6}$$

$$G（x，y）=（1/2\pi\sigma^2）\cdot \exp[-（x^2 + y^2）/2\sigma^2]\tag{7}$$

$$I_sharp = I + \alpha \cdot （I - I_blur）\tag{8}$$

数据集按 7∶1∶2 分层随机抽样为训练集（833 张）、验证集（119 张）、测试集（238 张），训练阶段随机翻转、±15°旋转增强。

4 双模型构建

VGG16 分类：冻结阶段（学习率 10^{-4}，Adam，batch=32，EarlyStopping patience=8）训练顶层分类器；微调阶段解冻 block4~block5（学习率 10^{-5}）。分别训练 VGG16-Orig（原始图）和 VGG16-Enh（预处理图）。

U-Net 分割：伪标签自动生成——CLAHE 增强→高斯滤波→Otsu 二值化→形态学开闭运算→连通域面

积筛选→伪标签掩膜。U-Net 训练采用二元交叉熵损失，Adam 优化器，学习率 10^{-4}，训练 30 轮。

整体流程：CT 图像经预处理后并行送入双分支，VGG16 输出分类置信度，U-Net 输出病灶掩膜，共同构成诊断报告，流程如图 3 所示。



图 3　系统整体算法流程图

5 实验与结果分析

实验环境：Intel i7-12700H，16GB 内存，RTX 3060（6GB），Python 3.9，TensorFlow 2.15。分类评价：准确率、精确率、召回率、F1；分割评价：像素准确率、IoU、Dice，核心公式：

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \tag{9}$$

$$IoU = TP/(TP + FP + FN) \quad Dice = 2TP/(2TP + FP + FN) \tag{10}$$

（1）VGG16 分类与对比实验

VGG16 双模型性能见表 1，VGG16-Enh 准确率 98.74%，比 VGG16-Orig 提升 2.02 个百分点，预处理显著增强特征辨识度。同条件下与 ResNet50、EfficientNet-B0 对比（表 2），VGG16-Enh 以 98.74% 准确率领先，其 3×3 卷积堆叠在小样本医学影像上迁移效果更优，参数量适中不易过拟合。

（2）U-Net 分割与消融实验

U-Net 分割指标与消融实验如表 3、表 4。分割像素准确率 99.29%、IoU 65.3%、Dice 78.6%，模型能定位大部分病灶区域，边界精度受伪标签噪声和类别不平衡（病灶像素<1%）限制。消融实验表明：去除 CLAHE 分类下降 2.02 个百分点，去除 ROI 约束 IoU 下降 6.6 个百分点，去除 Dropout+早停测试准确率降至 94.50%，各模块均有显著贡献。

表 1　VGG16 双模型分类性能对比

评价指标	VGG16-Orig	VGG16-Enh
准确率/%	96.72	98.74
精确率/%	95.81	98.52
召回率/%	96.03	98.31
F1 分数/%	95.92	98.41

表 2　不同骨干网络分类性能对比

模型	准确率/%	F1 分数/%	参数量/M
VGG16-Enh（本文）	98.74	98.41	138
ResNet50	97.48	96.92	25.6
EfficientNet-B0	97.06	96.35	5.3

表 3　U-Net 分割性能指标

评价指标	数值
像素准确率 PA/%	99.29
交并比 IoU/%	65.30
Dice 系数/%	78.60

表 4　消融实验结果

编号	去除模块	分类准确率/%	分割 IoU/%
消融 1	CLAHE 预处理	96.72	62.1
消融 2	ROI 胸腔约束	98.74	58.7
消融 3	Dropout+早停	94.50	60.3

（3）训练曲线分析

双模型训练曲线如图 4。VGG16 两阶段训练准确率快速收敛，U-Net 验证准确率 25 轮后基本稳定，早停策略抑制过拟合。

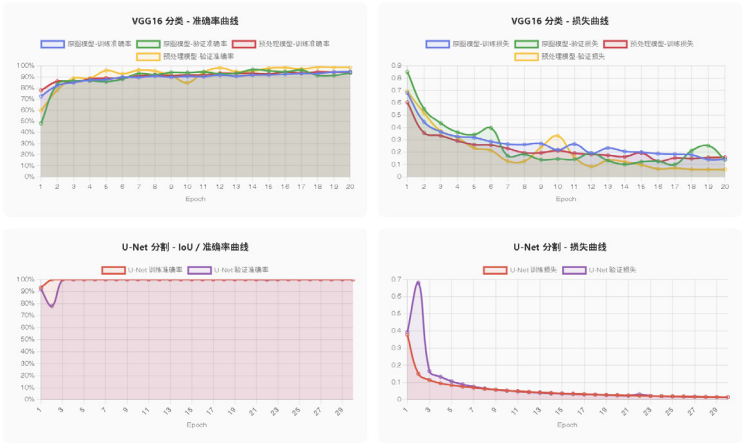


图 4　双模型训练准确率与损失函数曲线

6 结论与展望

本文提出基于迁移学习的肺癌 CT 分类与病灶分割方法：VGG16 两阶段迁移学习分类准确率 98.74%，伪标签监督 U-Net 分割 IoU 65.3%、Dice 78.6%，消融实验验证各模块有效性。方法无需手工像素标注，部署门槛低，适配基层医疗机构。

未来引入 Transformer 骨干和 3D CT 序列分割[8]，增强 Grad-CAM 可解释性；引入多中心数据集和人工标注真值掩膜；对接 PACS 系统开展多中心验证。

参考文献

[1] 国家癌症中心. 2024 年全国癌症统计报告[R]. 北京: 国家癌症中心, 2024.

[2] SHEN W, ZHOU M, YANG W, et al. Multi-crop convolutional neural networks for lung nodule classification[J]. Pattern Recognition, 2016, 61: 630-639.

[3] ARDILA D, KIRALY A P, BHARADWAJ S, et al. End-to-end lung cancer screening with three-dimensional deep learning[J]. Nature Medicine, 2019, 25(6): 954-961.

[4] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation[M]. Cham: Springer, 2015.

[5] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]. San Diego: ICLR, 2015.

[6] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. Las Vegas: IEEE, 2016.

[7] Dominyka Stragyte, Gvidas Mikalauskas, Katrina Gaidulevic, Renata Paukstaitiene, Kestutis Stasaitis,Vidas Raudonis,Skaidra Valiukeviciene.A Structured Review of Deep Learning Approaches and Image-Preprocessing Techniques for Automated Contact Allergy Patch Test Interpretation.[J].Medical sciences (Basel, Switzerland), 2026, 14 (2):322-322.

[8] TIAN Kaiyuan, QIAO Linbo, LIU Baihui, JIANG Gongqingjian, LI Shanshan, LI Dongsheng. A survey on memory-efficient transformer-based model training in AI for science[J].Frontiers of Computer Science,2026,20 (11):

版权声明：©2026 作者与开放获取期刊研究中心（OAJRC）所有。本文章按照知识共享署名许可条款发表。
https://creativecommons.org/licenses/by/4.0/

OPEN ACCESS