

基于多维度融合的肺结节分类算法

堵红群¹, 李岳阳², 崔方正², 罗海驰³, 顾中轩²

1. 江南大学附属医院影像科, 江苏 无锡 214122; 2. 江南大学人工智能与计算机学院, 江苏 无锡 214122; 3. 江南大学物联网工程学院, 江苏 无锡 214122

【摘要】采用多维度模型融合的方法,提出一种肺结节分类算法。在肺结节假阳性减少算法基础上进行优化,在多尺度特征融合模块得到特征之后引入高层特征增强软激活映射模块,以增强模型的分类能力;针对实际分类过程中各类结节数据不平衡的问题,引入平衡均方差损失来改进模型的训练效果;采用三维和二维模型融合方式进一步提升模型分类性能。在Private Lung数据集上进行的实验证明本研究提出的模型分类准确度达到93.8%,优于现有方法。

【关键词】肺结节; 计算辅助诊断; 多尺度特征融合; 软激活映射; 平衡均方差损失

【中图分类号】R318; TP391.4

【文献标志码】A

【文章编号】1005-202X(2024)11-1428-09

Lung nodule classification algorithm based on multi-dimensional fusion

DU Hongqun¹, LI Yueyang², CUI Fangzheng², LUO Haichi³, GU Zhongxuan²

1. Department of Medical Imaging, Affiliated Hospital of Jiangnan University, Wuxi 214122, China; 2. School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, China; 3. School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China

Abstract: A novel algorithm based on multi-dimensional fusion is proposed for classifying lung nodules. Based on the algorithm for reducing false positives of pulmonary nodules, the optimization is carried out by introducing a high-level feature enhancement soft activation mapping module after obtaining features by the multi-scale feature fusion module to improve the classification ability. To address the imbalance of different nodule data in the actual classification, a balanced mean square error loss is adopted to improve the training effect of the model. A three-dimensional and two-dimensional model fusion method is used to further improve the classification performance. The experiment conducted on a Private Lung dataset proves that the proposed model has a classification accuracy of 93.8%, outperforming the existing methods.

Keywords: lung nodule; computer-aided diagnosis; multi-scale feature fusion; soft activation mapping; balanced mean square error loss

前言

目前,计算机辅助诊断(Computer-Aided Diagnosis, CAD)系统^[1]通常只能初步确定肺结节位置和大小,后续还需要影像科医生结合结节的大小、形态、CT影像特征、组织学特性以及代谢活性等多个

因素进一步对结节进行分类,耗费大量的时间和精力。因此研究结节分类算法对于完善肺结节CAD系统功能具有重大意义^[2]。目前基于CT图像的肺结节分类方法已有大量研究成果,主要分为机器学习方法和深度学习方法。Kawata等^[3]设计混合K-均值聚类算法,利用线性分类器区分良性和恶性结节。Wei等^[4]提出一种基于半监督的局部核回归算法,通过与正则项模型来区分良恶性肺结节,在临床数据中提取746个肺结节,验证了模型性能。Li等^[5]改进随机森林算法,通过结合灰度共生矩阵和Gabor滤波的方法提取肺结节特征,并进一步改进射频分类器用于肺结节良恶性分类。有研究表明对于肺结节分类,基于深度学习的卷积神经网络(Convolutional Neural

【收稿日期】2024-06-13

【基金项目】国家自然科学基金(U1836218)

【作者简介】堵红群,硕士,主任医师,研究方向:医学图像处理、人工智能在医疗影像中的应用, E-mail: 15301516381@163.com

【通信作者】李岳阳,博士,副教授,研究方向:医学图像处理、人工智能在医疗影像中的应用, E-mail: lyueyang@jiangnan.edu.cn

Network, CNN)的分类性能要比传统机器学习方法优秀。Shen等^[6]使用一种多裁剪CNN分类器,对LIDC-IDRI数据集中的880个低疑似恶性结节和495个高疑似恶性结节进行分类。Xie等^[7]提出一种基于多尺度结节输入的网络模型,利用结节边缘图像、肺实质图像和包含结节的图像作为输入,取得93.4%的精度。Lei等^[8]开发一种软激活映射(Soft Activation Mapping, SAM)模块,该方法可以关注到肺结节的细粒度特征,并对肺结节的形状和边缘特征进行分析,然后将高维度特征和SAM结合,进一步提出一种高层特征增强SAM(High-level Feature Enhancement Soft Activation Mapping, HESAM)模块来区分形状和边缘特征。Liu等^[9]探究肺结节分类和对应得分之间的相关性,并提出一种多任务深度学习框架,取得显著的效果。目前针对肺结节分类问题仍然存在着许多挑战,如肺结节分类的精度有待提高,如何有效解决数据不平衡问题,如何将多维度信息融合来提高分类性能等。

以上针对肺结节分类的研究通常采用LIDC-IDRI数据集进行训练验证,该数据集是由美国国家癌症研究所创建,一共收集1018个研究实例,每一个实例都由4位经验丰富的胸部放射科医生进行标注,该数据集仅针对肺结节良恶性进行判断,但是影像科医生在实际的诊断过程中,根据结节的形状、纹理和结节大小,还需要进行肺结节具体类别的判定,如实性结节、磨玻璃结节和混合磨玻璃结节等。

本研究针对肺结节分类问题展开探讨,致力于提高模型对实性结节、磨玻璃结节和混合磨玻璃结节分类的精度,主要贡献如下:(1)在肺结节假阳性减少模型上进行优化,针对结节的大小多样性,在多尺度特征融合模块得到特征之后引入HESAM模块,以增强模型的分类能力。(2)针对在实际结节的分类过程中各类结节数据不均衡的情况,引入一种平衡均方误差损失函数(Balanced Mean Square Error Loss, BMSELoss),更适用于肺结节分类模型的训练。(3)借鉴影像科医生在实际诊断过程中根据多个连续的二维(Two-Dimensional, 2D)切片来进行肺结节分类的特点,针对2D和三维(Three-Dimensional, 3D)图像信息分别构建分类网络,然后将这2个维度的分类结果进行决策融合,进一步提高模型的分类性能。

1 肺结节分类框架

本研究方法是在CAD系统检测到肺结节的中心位置之后,针对检测到的肺结节进行分类,将结节分为实性结节、磨玻璃结节和混合磨玻璃结节。图1为

本文算法的整体架构,整个模型分为3D和2D两个分类网络,分别针对3D和2D图像信息进行处理,将两者得到的分类结果再进行决策融合,得到最终结果。

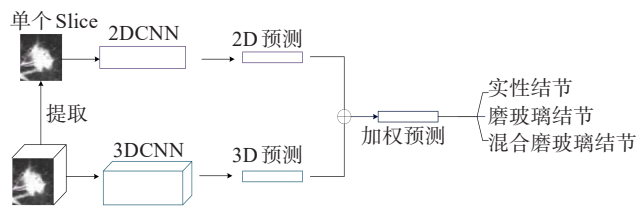


图1 肺结节分类算法框架

Figure 1 Framework of the lung nodule classification algorithm

在3D肺结节分类网络中,根据CAD系统检测到的肺结节中心位置,截取得到 $32 \times 32 \times 32$ 大小的3D结节数据块,输入3DCNN,得到分类结果,另外,根据3D结节数据块,得到其切片方向中心位置的2D切片,输入2DCNN,得到分类结果,将这两个维度的分类结果进行加权决策融合,得到最终分类结果。

2 3D肺结节分类网络

在实际诊断过程中可知,实性结节、磨玻璃结节和混合磨玻璃结节这3种肺结节具有相似的外观形态,特别是磨玻璃结节和混合磨玻璃结节非常相似,这给结节分类增加很大的难度。对于肺结节CAD系统,通常会包含两个步骤:候选结节检测和假阳性减少。对于假阳性减少算法,前期研究提出一种基于交叉注意力机制的肺结节假阳性减少网络^[10]。本研究在假阳性减少网络基础上进行优化,在多尺度特征融合模块得到特征之后引入HESAM模块,以增强网络的分类能力。

图2是3D肺结节分类网络示意图。网络的输入是 $32 \times 32 \times 32$ 大小的3D结节数据块,通过主干网络3DSENet50进行特征提取,获得4个层次的粗特征;然后,将粗特征传递到多尺度特征融合模块,得到64通道的融合特征;接着,通过残差块进行一次下采样操作对通道数进行调整,得到512维通道的聚合特征 $\mathbf{g}_k (k \in 1, 2, \dots, n)$ 。另外,由主干网络3DSENet50的最后一层特征,经过全局最大池化(Global Maximum Pooling, GMP)提取得到高层特征 $\mathbf{d}_k (k \in 1, 2, \dots, n)$,然后,将高层特征 \mathbf{d}_k 和聚合特征 \mathbf{g}_k 输入到HESAM模块中,其输出再输入到全连接层得到最终的3D输出向量,即模型分类结果。

2.1 多尺度特征融合模块

肺结节具有尺度可变、形状不规则、位置分布随

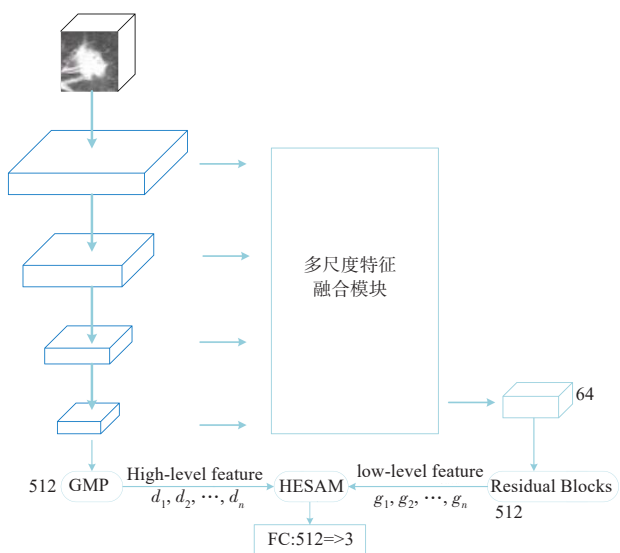


图2 3D肺结节分类网络

Figure 2 3D lung nodule classification network

机等特点,这大大增加了提取结节有效特征的难度。Zhang等^[11]提出一种具有尺度感知的特征融合模块。受此启发,本研究在假阳性减少网络^[10]中提出图2中所示的多尺度特征融合模块,其具体架构如图3所示。该模块可以将图2中的主干网络输出的4个层次(stage1~4)的粗特征 $x_i (i \in 1, 2, 3, 4)$ 经由特征细化模块(FR Block)进行细化处理,将细化后的特征 $s_i (i \in 1, 2, 3, 4)$ 通过一种类似特征金字塔的融合方式—交叉注意力融合机制得到64通道的融合特征。

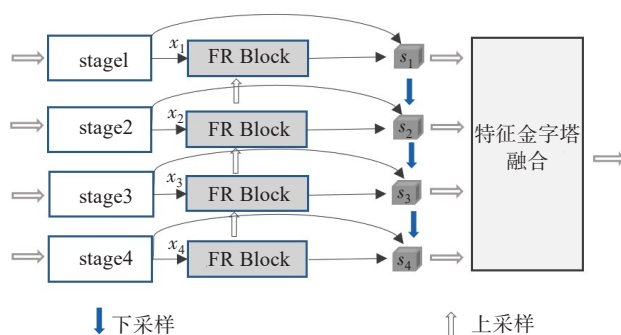


图3 多尺度特征融合模块

Figure 3 Multi-scale feature fusion module

图4是特征细化模块,其中 $x_i (i \in 1, 2, 3, 4)$ 是主干网络生成的粗特征, $s_i (i \in 1, 2, 3, 4)$ 是特征细化模块的输出特征。特征细化模块的特征融合方法可描述为:

$$x'_i = \begin{cases} fr_i(x_i), i = 4 \\ fr_i\left(\frac{\lambda_1 x_i + \lambda_2 \text{Up}(x'_{i+1})}{\lambda_1}\right), i \neq 4 \end{cases} \quad (1)$$

$$s_i = \begin{cases} \frac{\lambda'_1 x_i + \lambda'_2 x'_i}{\lambda'_1 + \lambda'_2}, i = 1 \\ \frac{\lambda'_1 x_i + \lambda'_2 x'_i + \lambda'_3 \text{Down}(s_{i-1})}{\lambda'_1 + \lambda'_2 + \lambda'_3}, i \neq 1 \end{cases} \quad (2)$$

其中, fr 代表特征细化模块,Up和Down分别表示上采样和下采样操作。

通过特征细化模块,对主干网络得到的粗特征

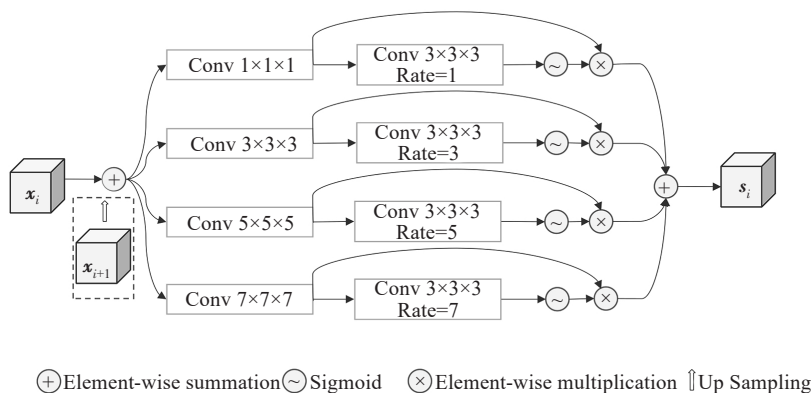


图4 特征细化模块

Figure 4 Feature refinement module

采用自上而下和自下而上的特征融合策略,将高分辨率的丰富上下文信息和低分辨率的高级区域语义信息进行有效融合,得到细化后特征。另外,使用4个分支进行特征细化处理,分别采用卷积核为1、3、5、7的3D卷积,对应膨胀系数分别为1、3、5、7,再采

用残差连接,避免梯度消失。

将细化后的特征 $[s_1, s_2, s_3, s_4]$ 通过特征金字塔的融合过程,输出通道数为64的特征向量 F_1 用于分类,如图5所示,其中CSFA表示交叉注意力融合模块。

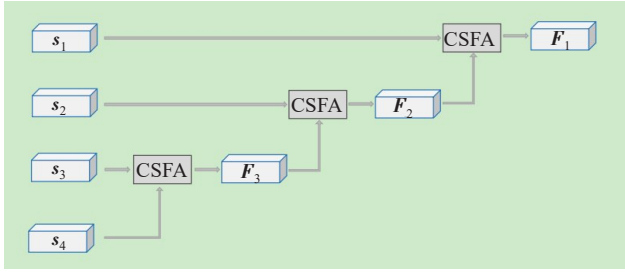


图5 特征金字塔的融合过程

Figure 5 Fusion process of the feature pyramid

2.2 HESAM 模块

在CAM^[12]中,验证了全局池化在对象定位中的作用,但是在局部或者像素稀疏区域的某些结节特征,全局池化的优势并不明显。因此,针对这一问题,Lei等^[8]提出一种SAM模块,并进一步优化,给出HESAM模块,用来进行肺结节良恶性分类,在LIDC-IDRI数据集上取得较好的效果。

图2中,主干网络输出4个层次的粗特征,经过多尺度特征融合模块后,得到64通道的融合特征 F_1 ,通过残差块进行一次下采样操作对通道数进行调整,

得到512维通道的聚合特征 $g_k (k \in 1, 2, \dots, n)$ 。为增强网络的肺结节分类能力,在此引入HESAM模块,将聚合特征 g_k 与主干网络提取的最后一层特征经GMP得到的高层特征 d_k 进行融合。 d_k 与类别信息的直接关系比较小,仅仅与特定输入有关,将其与聚合特征 g_k 相结合,使其能捕捉到一些与结构相关的信息,如磨玻璃结节不规则的形状边缘,可以增强分类器的特征表达能力。

HESAM模块的架构如图6所示,可表示为:

$$A^k(x', y', z') = a(g_k(x, y, z)), k \in \{1, 2, \dots, n\} \quad (3)$$

$$H_k = h_k(A^k) \quad (4)$$

$$S_k = H_k + d_k \quad (5)$$

其中, g_k 表示第 k 个聚合特征图, a 表示自适应的平均池化(AVG Pool), $A^k(x', y', z')$ 表示自适应的平均池化在 g_k 的空间位置 (x, y, z) 生成的结果, h_k 表示在池化后的全连接操作,输出为 H_k 。通过这个全连接层,可有效融合各通道间的信息。 S_k 为 H_k 与高层特征 d_k 相加后的结果,作为最后一个分类器(Classifier)的输入,得到最终的3分类结果。分类器使用的是全连接层。

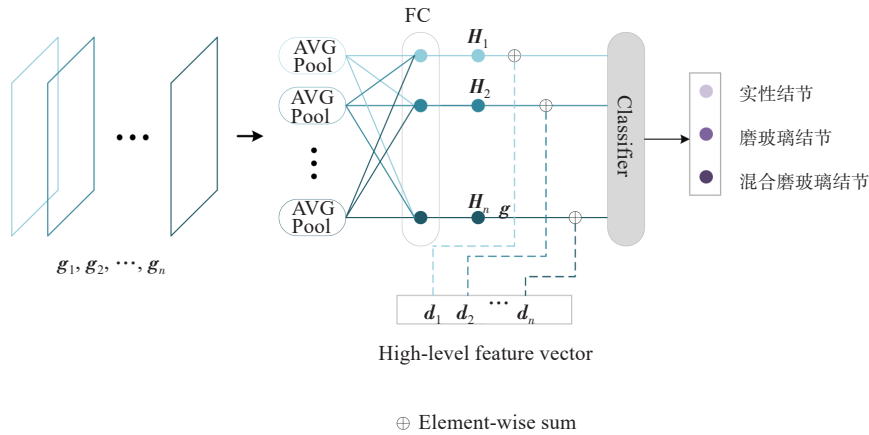


图6 高层特征增强软激活模块

Figure 6 High-level feature enhancement soft activation mapping (HESAM) module

3 BMSELoss

在实际结节的分类过程中,经常会出各类结节数据不均衡的情况。Ren等^[13]重新审视了常用的均方误差(Mean Square Error, MSE)损失函数,发现当训练数据不均衡时,MSE会受标签分布影响,从而训练结果会偏向分布比较常见的标签。因此,当测试集标签分布比较平均时,MSE的这一特点会导致模型的表现较差。Ren等^[13]从统计学的角度出发,使用概率的方法降低不平衡标签分布对MSE的影响。假

设训练集和测试集来自不同的分布 $p_{\text{train}}(x, y)$ 和 $p_{\text{val}}(x, y)$,模型输出的 $p_{\text{val}}(x, y)$ 通过贝叶斯公式和全概率公式可以实现两者的转换。贝叶斯定理可以表示为:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (6)$$

全概率公式可以表示为:

$$P(A) = \sum_{i=1}^n P(B_i)P(A|B_i) \quad (7)$$

$p_{\text{train}}(x, y)$ 可以表示为:

$$p_{\text{train}}(y|x) = \frac{p_{\text{train}}(y|x)}{\int_y p_{\text{train}}(y|x) dy'} \quad (8)$$

$$\begin{aligned} p_{\text{train}}(y|x) &= \frac{p_{\text{val}}(y|x) \cdot \frac{p_{\text{train}}(y)}{p_{\text{val}}(y)} \cdot \frac{p_{\text{val}}(x)}{p_{\text{train}}(x)}}{\int_y p_{\text{val}}(y|x) \cdot \frac{p_{\text{train}}(y')}{p_{\text{val}}(y')} \cdot \frac{p_{\text{val}}(x)}{p_{\text{train}}(x)} dy'} \\ &= \frac{p_{\text{val}}(y|x) \cdot \frac{p_{\text{train}}(y)}{p_{\text{val}}(y)}}{\int_y p_{\text{val}}(y|x) \cdot \frac{p_{\text{train}}(y')}{p_{\text{val}}(y')} dy'} \\ &= \frac{p_{\text{val}}(y|x) \cdot p_{\text{train}}(y)}{\int_y p_{\text{val}}(y|x) \cdot p_{\text{train}}(y') dy'} \end{aligned} \quad (9)$$

有了上述转换, BMSELoss 可以定义为:

$$\begin{aligned} L &= -\log p_{\text{train}}(y|x; \theta) = -\log \frac{p_{\text{val}}(y|x; \theta) \cdot p_{\text{train}}(y)}{\int_y p_{\text{val}}(y|x; \theta) \cdot p_{\text{train}}(y') dy'} \\ &\cong -\log N(y; y_{\text{pred}}, \sigma_{\text{noise}}^2 I) + \\ &\quad \log \int_y N(y'; y_{\text{pred}}, \sigma_{\text{noise}}^2 I) \cdot p_{\text{train}}(y') dy' \end{aligned} \quad (10)$$

其中, N 表示 NLL 损失函数。Ren 等^[13]通过 Batch-based Monte-Carlo (BMC) 的方法解出损失函数。不需要对 $p_{\text{train}}(y')$ 建模, 而是将所有的标签样本都看成是 $p_{\text{train}}(y')$ 的随机样本。那么 BMSELoss 可以化简为:

$$L = -\log \frac{\exp(-\|y_{\text{pred}} - y\|_2^2 / \tau)}{\sum_{y' \in B_y} \exp(-\|y_{\text{pred}} - y'\|_2^2 / \tau)} \quad (11)$$

其中, $\tau = 2\sigma_{\text{noise}}^2$, σ_{noise} 为噪声超参数。

本研究采用的数据集中, 磨玻璃结节相对实性结节和混合磨玻璃结节数量较多, 但后面两种结节的总体数量还是比较多的, 没有出现数据分布极度不均衡的问题。因此, 采用 BMSELoss 作为损失函数可以进一步提高模型对于不同样本数量的结节的感知能力。

4 2D 切片分类结果融合

在实际工作中, 影像科医生一般通过多个连续的 2D CT 切片来判断结节的类型。相较于 2D 图像信息, 3D 图像信息更多关注的是结节的整体 3D 形状特征, 但混合磨玻璃结节和磨玻璃结节在空间形状上具有高度相似的形态特征, 仅仅使用 3D 分类网络对结节进行分类容易造成误判。本研究采用 3D 数据块切片方向中心位置的一张 2D CT 切片作为 2D 分类网络 (ResNet50) 的输入, 综合考虑 2D 和 3D 网络的输

出, 将这两个维度的分类结果进行加权决策融合。网络的最终输出表示为:

$$\text{pred}_{\text{out}} = w_1 \times \text{soft max}(\text{pred}_{2D}) + w_2 \times \text{soft max}(\text{pred}_{3D}) \quad (12)$$

其中, soft max 表示多分类使用的激活函数, w_1 和 w_2 分别代表 2D 和 3D 分类结果的权重。

5 实验结果与分析

5.1 实验环境及参数设置

本研究所用实验平台是 Linux, 操作系统是 Ubuntu20.04, 显卡为 4 块 NVIDIA GeForce3090, 开发语言是 Python, 深度学习框架是 Pytorch1.11。采用分类模型对肺结节进行分类, 网络训练的具体参数设置如表 1 所示。

表 1 网络训练参数

Table 1 Network training parameters

算法参数	参数设置
批大小 (Batchsize)	64
学习率 (Learning Rate)	0.001
超参数 σ (Parameter σ)	1.0
优化器 (Optimizer)	AdamW
动量因子 (Momentum)	0.9
权重衰减 (Weight decay)	0.000 000 8
最大迭代次数 (Max epoch)	300

5.2 数据集介绍

本实验的数据集是某医院提供的私有临床数据集 (仅供学习研究), 称为 Private Lung。该数据集收集了 2021~2022 年该医院影像科收录的部分脱敏数据 (不包含患者的隐私信息)。本研究选取层厚为 1 mm 的 CT 图像构成数据集。Private Lung 包含两部分: CT 数据和标注文件。CT 数据采用 DICOM 的格式进行保存, 标注信息包括 ID、结节中心位置坐标、结节类型。数据集一共包括 490 个结节, 其中实性结节 144 个、磨玻璃结节 223 个、混合磨玻璃结节 123 个。数据集按照 ID 随机划分, 训练集和测试集的比例为 8:2。

5.3 数据集预处理

本实验的 Private Lung 数据集中, 肺部 CT 的存储方式是 DICOM 格式。采用 SimpleITK 读取目标结节位置的 32 张相邻切片, 使用阈值分割的方法进行肺实质分割, 然后以目标结节为中心点截取 $32 \times 32 \times 32$ 大小的数据块。

对于2D数据,提取目标位置切片,也就是数据块的中心切片。目标位置相邻的切片也包含类似的2D形态信息,所以又选取4张相邻切片,作为一种数据增强方法,这也符合影像科医生通常会根据连续多张2D切片来判断结节分类的实际情况。

对于3D数据,本研究采取以下数据增强方法:旋转、镜像和翻转,并添加偏移中心点的方法,截取部分结节。

5.4 评价指标

本研究是一个三分类问题,评价分类模型性能最常用的方法是混淆矩阵。混淆矩阵通常由以下4类组成:(1)TP,真实值是 Positive,模型认为是 Positive 的数量;(2)FN,真实值是 Positive,模型认为是 Negative 的数量;(3)FP,真实值是 Negative,模型认为是 Positive 的数量;(4)TN,真实值是 Negative,模型认为是 Negative 的数量。

本研究采用准确率、精确率、召回率和特异性作

为评价指标,计算方式如下:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

(13)

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

(14)

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

(15)

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

(16)

5.5 实验结果及分析

5.5.1 消融实验 选取肺结节假阳性分类模型^[10]作为基线网络(Baseline),该模型已在LUNA16数据集中被证明为一种有效的分类方法,但是仅仅针对于肺结节假阳性减少任务。针对本研究要解决的3D肺结节分类问题,将该模型进行优化改进,然后在 Private Lung 数据集中进行肺结节定性分类实验。3D肺结节分类网络各个模块的消融实验如表2所示。

表2 3D肺结节分类网络构成模块消融实验
Table 2 Ablation study on the component modules of the 3D lung nodule classification network

组合	准确率	实性结节			磨玻璃结节			混合磨玻璃结节		
		精确率	召回率	特异性	精确率	召回率	特异性	精确率	召回率	特异性
Baseline	0.865	0.837	0.936	0.928	0.841	0.938	0.865	0.867	0.781	0.934
Baseline+BMSELoss	0.867	0.846	0.943	0.905	0.841	1.000	0.885	0.867	0.797	0.955
SAM	0.876	1.000	0.857	1.000	0.818	0.973	0.869	0.833	0.769	0.944
HESAM	0.887	0.933	0.8	0.968	0.923	0.973	0.951	0.793	0.885	0.917
HESAM+BMSELoss	0.898	1.000	0.829	1.000	0.837	0.973	0.885	0.885	0.885	0.958

基线网络是前期研究^[10]提出的肺结节假阳性分类模型(3DSENet50+FRBlock+CFSA+DisAlignLinear+GRWLoss),在 Private Lung 数据集上取得0.865的准确率。本研究在基线网络的基础上添加SAM模块,即对CFSA模块输出的聚合特征做软特征图映射操作,加入SAM模块后,模型的准确率提升1.1%。对3DSENet50主干网络提取的最后一层特征做GMP处理后,和SAM模块的输出相加输入到分类器中,即采用HESAM模块后,分类网络的准确率又上升1.1%。

为验证BMSELoss更适合 Private Lung 数据集的样本分布,将Baseline的DisAlignLinear^[14]替换成普通的线性层,与DisAlignLinear对应的GRWLoss替换成BMSELoss,经过对比发现准确率上升0.2%,说明BMSELoss可能更适合作为训练 Private Lung 数据集的损失函数。在采用HESAM模块基础上,使用BMSELoss后准确率上升1.1%。

表3所示的是2D与3D模型多维度融合的实验结果。2D分类网络采用Resnet50^[15]训练,准确率达0.854,仍然可以有效提取到2D切片中包含的语义信息;但是因为模型参数量小,训练时间约为3D模型的3.3%。本研究提取3D数据块中最中间的切片输入训练好的2D网络中,将两个模型的预测结果加权融合,得到最终的分类结果。通过这种多维度的模型融合方法,与单独使用3D分类模型相比,准确率上升4%。

为有效地融合2D和3D信息,进行消融实验来调整式(15)中的决策融合参数 w_1 和 w_2 ,如表4所示。结果显示,在参数设置为[0.5, 0.5]时,模型准确率达到最优值0.938,这明确表明2D信息在分类任务中仍具有关键作用;而当取[0.9, 0.1]或[0.1, 0.9]时,即主要侧重于单一的2D或3D信息时,模型准确率均有降低。实验结果表明将2D和3D信息的分类结果进行有效融合,是非常有效的。

表3 2D与3D模型多维度融合
Table 3 Multi-dimensional fusion of 2D and 3D models

维度	准确率	实性结节			磨玻璃结节			混合磨玻璃结节		
		精确率	召回率	特异性	精确率	召回率	特异性	精确率	召回率	特异性
2D	0.854	0.966	0.800	0.984	0.854	0.946	0.902	0.786	0.846	0.917
3D	0.898	1.000	0.829	1.000	0.837	0.973	0.885	0.885	0.885	0.958
2D+3D	0.938	0.971	0.943	0.984	0.900	0.973	0.934	0.958	0.885	0.986

表4 决策融合参数消融实验
Table 4 Ablation study on decision fusion parameters

$[w_1, w_2]$	准确率	实性结节			磨玻璃结节			混合磨玻璃结节		
		精确率	召回率	特异性	精确率	召回率	特异性	精确率	召回率	特异性
[0.9, 0.1]	0.888	0.938	0.857	0.968	0.854	0.946	0.902	0.88	0.846	0.958
[0.6, 0.4]	0.898	0.938	0.857	0.968	0.875	0.946	0.918	0.885	0.885	0.958
[0.5, 0.5]	0.938	0.971	0.943	0.984	0.900	0.973	0.934	0.958	0.885	0.986
[0.4, 0.6]	0.908	1.000	0.857	1.000	0.857	0.973	0.902	0.885	0.885	0.958
[0.1, 0.9]	0.897	1.000	0.829	1.000	0.837	0.937	0.885	0.885	0.885	0.958

5.5.2 与现有算法比较 为验证本文算法的有效性,在 Private Lung 数据集上进行实验,并选取部分现有算法^[6, 8, 16-17]进行对比,如表5所示,可见本文方法的分类性能显著超越其他算法,实现最优0.938的准确率。

表5 与现有算法比较
Table 5 Comparison with the existing algorithms

方法	准确率	实性结节			磨玻璃结节			混合磨玻璃结节		
		精确率	召回率	特异性	精确率	召回率	特异性	精确率	召回率	特异性
Shen等 ^[6]	0.816	0.875	0.8	0.937	0.775	0.838	0.852	0.808	0.808	0.931
Zhu等 ^[16]	0.887	0.967	0.829	0.984	0.857	0.973	0.902	0.846	0.846	0.944
Lei等 ^[8]	0.877	0.931	0.771	0.968	0.860	1.000	0.902	0.846	0.846	0.944
Al-Shabi等 ^[17]	0.897	0.938	0.857	0.968	0.860	1.000	0.902	0.913	0.808	0.972
本文方法	0.938	0.971	0.943	0.984	0.900	0.973	0.934	0.958	0.885	0.986

5.5.3 模型预测可视化展示 为进一步分析多维度融合分类模型的性能,在此给出了分类之后的混淆矩阵,如图7所示,其中标签0、1、2分别代表实性结节、磨玻璃结节、混合磨玻璃结节。本研究还对误判结节做了切片可视化处理,如图8所示。

由于磨玻璃结节的样本数量相对实性结节和混合磨玻璃结节较多,3D分类模型的训练效果相对较好,但是融合2D分类模型的结果后,磨玻璃分类的召回率下降了,有1个结节分类错误,即图8a,分析其原因,3D分类模型可能包含更多的空间信息,更容易区

分磨玻璃结节,而从2D切片角度很难将两者区分,所以造成了误判。实性结节有2个误判,如图8b和图8f所示,其原因可能是实性结节因为预测中心点位置的偏差,所以中心切片密度会偏小,从而造成误判。总体来说,在加入2D切片融合后,实性结节的召回率上升了。混合磨玻璃结节共有3个误判,其中有1个被认为是实性结节,如图8c~e所示,判断分析其原因,应该是混合磨玻璃结节具有实性结节和磨玻璃结节的部分形态信息有关。

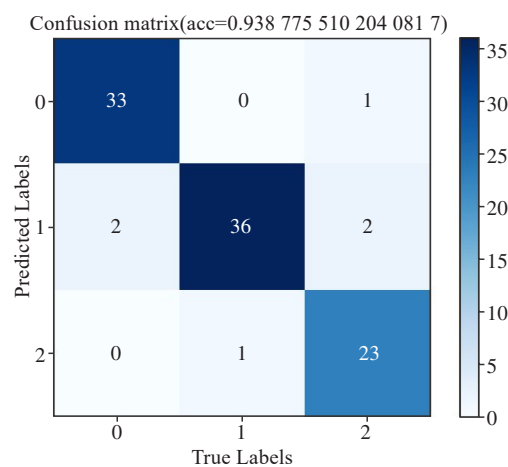


图7 混淆矩阵
Figure 7 Confusion matrix

6 总结

本研究基于多维度模型融合策略,提出一种肺结节分类的算法。在多尺度融合模块后引入HESAM模块以增强模型精度,缓解数据不均衡问题,并结合2DCNN和3DCNN实现高效决策融合。在Private Lung数据集上验证了本文算法的有效性,可以较好地地区分实性结节、磨玻璃结节和混合磨玻璃结节。在未来的研究和实际临床应用中,可以基于更多数据对提出的算法进行验证、优化和部署。此外,可以考虑与其他肺结节检测算法,如nnDetection模型^[18]、3-D OSAF-YOLOv3模型^[19]、Multi-level 3D DenseNet^[20]、NoduleNet^[21]、Fast

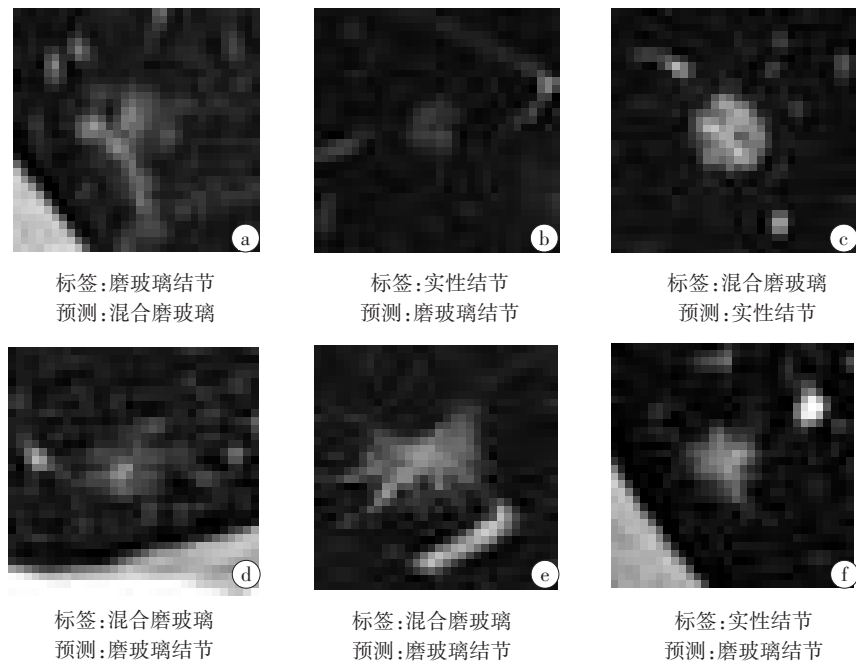


图8 误判结节切片
Figure 8 Misclassified nodule slices

Marching Method^[22]、基于YOLOv3的肺结节模型^[23]、CPM-net^[24]等进行联合使用,以确保算法在真实医疗场景中的可靠性和适用性。这样的联合使用可能有助于提高检测的准确性,并为医生提供更全面和可靠的辅助,进一步推动肺癌早期筛查和诊断的进展。

【参考文献】

[1] Amer HM, Abou-Chadi FE, Kishk SS, et al. A CAD system for the early detection of lung nodules using computed tomography scan images[J]. Int J Onl Biomed Eng, 2019, 15(4): 40-52.

[2] 谢新林, 肖毅, 续欣莹. 基于神经网络架构搜索的肺结节分类算法[J]. 计算机应用, 2022, 42(5): 1424-1430.

Xie XL, Xiao Y, Xu XY. Lung nodule classification algorithm based on neural network architecture search [J]. Journal of Computer Applications, 2022, 42(5): 1424-1430.

[3] Kawata Y, Niki N, Ohmatsu H, et al. Hybrid classification approach of malignant and benign pulmonary nodules based on topological and histogram features[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2000. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000: 297-306.

[4] Wei GH, Ma H, Qian W, et al. Lung nodule classification using local kernel regression models with out-of-sample extension[J]. Biomed Signal Process Control, 2018, 40: 1-9.

[5] Li XX, Li B, Tian LF, et al. Automatic benign and malignant classification of pulmonary nodules in thoracic computed tomography based on RF algorithm[J]. IET Image Process, 2018, 12(7): 1253-1264.

[6] Shen W, Zhou M, Yang F, et al. Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification[J]. Pattern Recognit, 2017, 61: 663-673.

[7] Xie HT, Yang DB, Sun NN, et al. Automated pulmonary nodule detection in CT images using deep convolutional neural networks[J]. Pattern Recognit, 2019, 85: 109-119.

- [8] Lei YM, Tian YK, Shan HM, et al. Shape and margin-aware lung nodule classification in low-dose CT images *via* soft activation mapping[J]. Med Image Anal, 2020, 60: 101628.
- [9] Liu LH, Dou Q, Chen H, et al. MTMR-net: multi-task deep learning with margin ranking loss for lung nodule analysis[C]//Proceedings of the Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer International Publishing, 2018: 74-82.
- [10] Gu ZX, Li YY, Luo HC, et al. Cross attention guided multi-scale feature fusion for false-positive reduction in pulmonary nodule detection[J]. Comput Biol Med, 2022, 151(Pt A): 106302.
- [11] Zhang YT, Li HM, Du J, et al. 3D multi-attention guided multi-task learning network for automatic gastric tumor segmentation and lymph node classification[J]. IEEE Trans Med Imaging, 2021, 40(6): 1618-1631.
- [12] Zhou BL, Khosla A, Lapedriza A, et al. Learning deep features for discriminative localization[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 2921-2929.
- [13] Ren JW, Zhang MY, Yu CJ, et al. Balanced MSE for imbalanced visual regression[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2022: 7916-7925.
- [14] Zhang SY, Li ZM, Yan SP, et al. Distribution alignment: a unified framework for long-tail visual recognition [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2021: 2361-2370.
- [15] He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 770-778.
- [16] Zhu WT, Liu CC, Fan W, et al. DeepLung: deep 3D dual path nets for automated pulmonary nodule detection and classification[C]//2018 IEEE Winter Conference on Applications of Computer Vision (WACV). Piscataway, NJ, USA: IEEE, 2018: 673-681.
- [17] Al-Shabi M, Shak K, Tan M. 3D axial-attention for lung nodule classification[J]. Int J Comput Assist Radiol Surg, 2021, 16(8): 1319-1324.
- [18] Baumgartner M, Jäger PF, Isensee F, et al. nnDetection: a self-configuring method for medical object detection[C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2021. Cham: Springer International Publishing, 2021: 530-539.
- [19] Huang YS, Chou PR, Chen HM, et al. One-stage pulmonary nodule detection using 3-D DCNN with feature fusion and attention mechanism in CT image[J]. Comput Methods Programs Biomed, 2022, 220: 106786.
- [20] Lu XQ, Gu Y, Yang LD, et al. Multi-level 3D densenets for false-positive reduction in lung nodule detection based on chest computed tomography[J]. Curr Med Imaging, 2020, 16(8): 1004-1021.
- [21] Tang H, Zhang CP, Xie XH. NoduleNet: decoupled false positive reduction for pulmonary nodule detection and segmentation[C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2019. Cham: Springer International Publishing, 2019: 266-274.
- [22] Paing MP, Choomchuay S. 3D pulmonary nodules detection using fast marching segmentation[J]. J Fundam Appl Sci, 2017, 9(4S): 319-339.
- [23] Ramachandran SS, George J, Skaria S, et al. Using YOLO based deep learning network for real time detection and localization of lung nodules from low dose CT scans [C]//Medical Imaging 2018: Computer-Aided Diagnosis. Bellingham, WA, USA: SPIE, 2018: 105751I.
- [24] Song T, Chen JN, Luo XD, et al. CPM-net: a 3D center-points matching network for pulmonary nodule detection in CT scans[C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2020. Cham: Springer International Publishing, 2020: 550-559.

(编辑:谭斯允)