

结合深度卷积神经网络与影像学特征的 肺结节良恶性鉴别方法

高大川, 聂生东*

上海理工大学医疗器械与食品学院, 上海 200082

摘要 提出一种将卷积神经网络(CNN)学习特征与传统影像学特征结合的肺结节良恶性鉴别方法。首先,从电子计算机断层扫描(CT)图像中分割出肺结节区域,并使用传统机器学习方法提取结节区域的影像学特征;然后,使用截取的肺结节训练 3D-Inception-ResNet 模型,提取网络学习的 CNN 特征,组合两类特征,并利用随机森林(RF)模型进行特征选择;最后,采用支持向量机(SVM)、RF 等传统分类器对肺结节进行良恶性鉴别诊断。使用 LIDC-IDRI 数据库中的 1036 个肺结节进行实验验证,最终所提方法的分类准确率、敏感度、特异度及接受者操作特性曲线(ROC)下面积(AUC)分别达 94.98%、90.02%、97.03%及 97.43%。实验结果表明,所提方法能准确地判别肺结节的良恶性,并优于大部分主流方法。

关键词 图像处理; 肺癌早期诊断; CT 影像; 肺结节良恶性鉴别; LIDC-IDRI 数据库

中图分类号 TP391.41

文献标志码 A

doi: 10.3788/AOS202040.2410002

Method for Identifying Benign and Malignant Pulmonary Nodules Combing Deep Convolutional Neural Network and Hand-Crafted Features

Gao Dachuan, Nie Shengdong*

*School of Medical Instrument and Food Engineering, University of Shanghai for Science and Technology,
Shanghai 200082, China*

Abstract Here, we present a method for identifying benign and malignant pulmonary nodules that combines convolutional neural network(CNN)learning features and conventional hand-crafted features. First, the pulmonary nodules area is segmented from computed tomography (CT) images, and traditional machine learning methods are used to extract the image features of the nodule area. Then, the CNN features of network learning are extracted, using the intercepted pulmonary nodules to train the 3D-Inception-ResNet model, and the 2 kinds of features are combined, the random forest (RF) model is used for feature selection. Finally, support vector machine (SVM) and RF classifier are used to identify benign and malignant pulmonary nodules. The 1036 pulmonary nodules in the LIDC-IDRI database are used for experimental verification. Classification accuracy, sensitivity, specificity, and area under the receiver operating characteristic (ROC) curve (AUC) of the proposed method can reach 94.98%, 90.02%, 97.03%, and 97.43%, respectively. The proposed method can accurately distinguish benign and malignant lung nodules, more effectively than most existing mainstream methods, as shown by the experimental results.

Key words image processing; early diagnosis of lung cancer; CT image; classification of benign and malignant nodules; LIDC-IDRI database

OCIS codes 100.4996; 100.6950

1 引 言

肺癌已成为患者因恶性肿瘤死亡的首要原因,

发病率和死亡率呈逐年上升趋势^[1]。“早发现、早诊断、早治疗”是提高肺癌患者生存率的关键。与晚期肺癌相比,经早期诊断的肺癌患者存活率大大提高。

收稿日期: 2020-07-13; 修回日期: 2020-09-02; 录用日期: 2020-09-15

基金项目: 国家自然科学基金(81830052)、上海市分子影像学重点实验室项目(18DZ2260400)、上海市自然科学基金(20ZR1438300)

* E-mail: nsd4647@163.com

肺癌早期一般表现为肺结节,因此,对肺结节进行检测与诊断是降低肺癌患者死亡率的关键^[2]。

影像学检查是筛选肺癌的主要方法之一^[3]。在临床治疗中,医生会根据患者电子计算机断层扫描(CT)影像中结节的形态、位置及纹理等特征分析病变信息^[4]。但是由于我国医疗资源匮乏,放射科医生平均每天需要审阅百余位病人的影像,这无疑是一项耗时、费力、高强度的工作,并且不同放射医师对结节的诊断能力与自身资历有很大关系,具有很强的主观性^[5]。因此,计算机辅助诊断(CAD)系统应运而生,它能作为医生的“第二意见”辅助放射科医生进行临床诊断^[6]。

基于传统影像学方法的肺结节良恶性诊断总体流程主要包括肺结节分割、特征提取、特征优化、使用分类模型进行良恶性分类^[7]。通常,分类精度与3个因素有关:肺结节分割精度;肺结节提取特征的代表性;分类器的性能。因此常通过优化上述因素来提高肺结节良恶性分类精度。Dhara等^[8]提出了一种基于临床医生标注属性加权的肺结节分类模型,该模型建立结节属性与良恶性程度的权重映射,并计算172种结节特征,使用分类器预测结节属性,最后根据属性加权得到良恶性分类结果。作者使用LIDC-IDRI数据集对所提模型进行验证,最终在集成随机森林(RF)、集成支持向量机(SVM)及 K 最近邻域分类器中,模型的分类精度分别达84.89%、83.86%及79.59%。Kaya等^[9]使用半自动化技术分割肺结节,并计算肺结节形状、纹理及边缘等特征,最后利用SVM对结节进行分类。作者在三种样本配置情况下通过使用LIDC-IDRI数据集中的891个结节来评估方法的有效性(恶性评分为1和2的肺结节为良性,恶性评分为4和5的肺结节为恶性,三种配置分别认为评分3的结节为丢弃、良性及恶性),最终接受者操作特性(ROC)曲线下面积(AUC)分别为0.9505、0.8822及0.8488。

近年来,深度学习在医学图像分析领域取得显著进展,在肺结节良恶性分类领域,很多性能优良的网络模型被陆续提出。相较于传统的影像学特征分类模型,卷积神经网络(CNN)使用原始CT影像作为模型输入,自动学习特征并进行分类,省略了特征设计环节。一般情况下,性能优异的CNN模型都需要庞大的数据集进行训练学习。Nibali等^[10]采用课程学习、迁移学习、深度残差网络来提高分类性能,并提取了肺结节沿三个方向的切片(横截面、矢状面、冠状面),同时训练了三个ResNet-18网络,通

过融合三个平面的分类结果获得最终分类结果。作者使用LIDC-IDRI数据集对该网络进行验证,准确率为89.90%。Shen等^[11]提出了一种多尺度CNN模型来区分良性结节与恶性结节。首先,在结节区域提取出多个尺度的结节块,包括 $32 \times 32 \times 32$ 像素块、 $64 \times 64 \times 64$ 像素块、 $96 \times 96 \times 96$ 像素块;然后,通过训练三个CNN模型来学习深度特征;最后使用RF和SVM进行分类。作者使用LIDC-IDRI数据集中的880个良性和495个恶性结节对模型进行评估,准确率达86.84%。

目前,基于CT影像的肺结节良恶性分类研究主要都是基于上述两类方法而单独进行的,很少有文献结合两类方法展开研究。本文结合两种不同方法,提出一种结合深度CNN与影像学特征的肺结节良恶性鉴别方法。该方法使用传统机器学习方法提取出肺结节影像学特征,并结合3D-Inception-ResNet模型的高级特征,利用RF模型进行特征选择,最后采用SVM、RF等传统分类器进行分类鉴别。实验结果表明,两类不同特征的组合使用大大丰富了结节特征的代表性,这既克服了传统影像学特征不能充分反映病灶区域信息的缺点,又弥补了在数据量不足的情况下3D CNN分类效果不好的劣势。所提方法对肺结节良恶性分类的准确率有明显提高。

2 基本原理

使用LIDC-IDRI数据集开展实验,首先使用4名放射科专家勾画的肺结节边界作为结节分割的结果;然后对每个肺结节样本提取灰度、形状、纹理、小波域等多种传统影像学特征,并训练改进3D-Inception-ResNet模型,得到肺结节CNN特征;组合两类特征,使用RF模型进行特征选择;最后采用SVM、RF等传统分类器进行分类鉴别。图1为所提方法的流程图,主要包括三部分:传统肺结节影像学特征提取部分;CNN提取特征部分;分类部分。

2.1 传统肺结节影像学特征提取部分

使用LIDC-IDRI数据集^[12]中4名放射科专家勾画的肺结节边界作为结节分割的结果。该数据库提供1010个病例的1018套胸腔CT图像数据,每个病例中包含病人的DICOM图像、4位放射科专家标注的结节位置和临床征象信息。首先,排除切片厚度大于3 mm的CT影像中的结节,研究表明,厚层CT影像不利于肺结节良恶性鉴别^[13];然后,挑选出至少被3名专家同时勾画的结节,并通过P-MAP融合方法^[14]确定结节分割结果;最终,得到

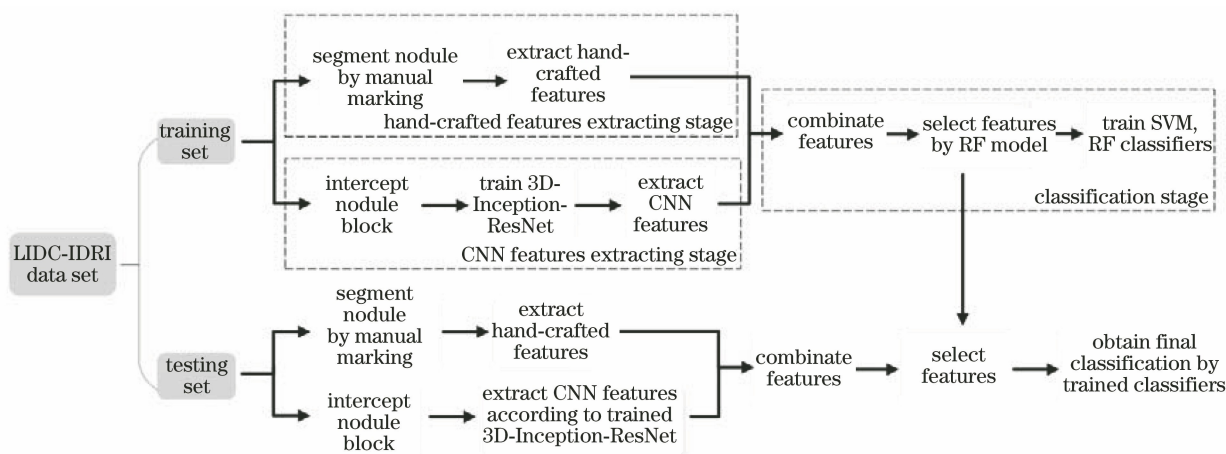


图 1 所提方法的流程图

Fig. 1 Flow chart of proposed method

1036 个 $64 \times 64 \times 64$ 肺结节像素块,其中良性结节 380 个,恶性结节 300 个,良恶性不确定结节 356

个。图 2 为结节融合方法示意图,图 2(e)为至少 3 名专家所勾画区域的公共部分。

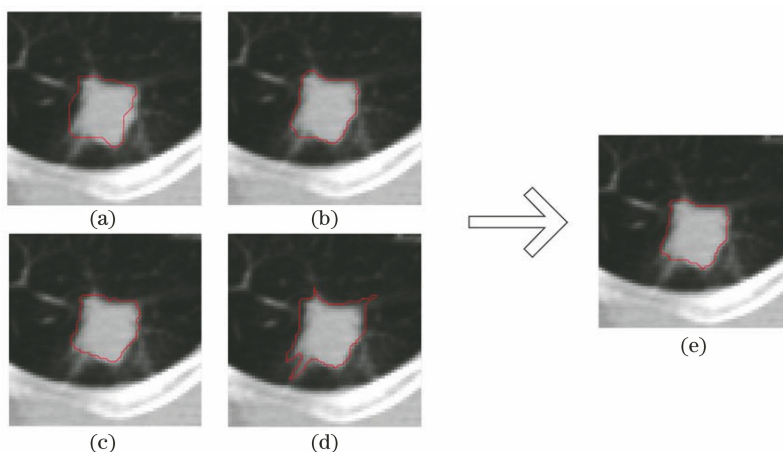


图 2 结节融合方法示意图。(a)~(d) 4 名放射医师手动分割的肺结节区域;(e)结节融合方法分割的肺结节区域

Fig. 2 Schematic of nodule fusion method. (a)~(d) Pulmonary nodule areas are manually segmented by four radiologists; (e) pulmonary nodule area segmented by nodule fusion method

特征提取是将图像信息数字化的重要转换手段。特征分布的差异性是肺结节良恶性鉴别的重要判据,如恶性结节具有分叶征、毛刺征、血管相融、囊性气隙等临床征象,而良性结节通常位于边缘位置且表面光滑^[15-16]。因此,对每个肺结节样本提取纹理、灰度、形状等 103 种传统影像学特征^[17]。

2.2 卷积神经网络提取特征部分

搭建了 3D-Inception-ResNet 模型,相较于传统二维网络模型,三维模型能够学习更丰富的结节空间信息,使 CNN 提取的高级特征更具代表性^[18]。Inception-ResNet 模型由 Szegedy 等^[19]提出,在 Inception 模块中引入 ResNet 残差结构后,该模型具备更优异的分类性能。

图 3 为搭建的 3D-Inception-ResNet 分类模型结构。首先根据 LIDC-IDRI 数据库中放射科专家的标注信息,从 $512 \times 512 \times S \times 1$ 原始 CT 影像中截取出以肺结节为中心的 $64 \times 64 \times 64 \times 1$ 像素块,其中 S 为 CT 影像切片数($S > 64$);然后,在三个维度(即 x 轴、 y 轴和 z 轴)对每个肺结节样本进行平移、翻转、旋转等操作来扩增实验数据,并将增强后的数据中心采样至 $48 \times 48 \times 48 \times 1$ 像素块,以消除数据增强操作对结节块边界值的影响;最后,使用扩增数据训练 3D-Inception-ResNet 模型。模型共包含 3 个卷积层、5 个池化层、3 个 Inception-ResNet 模块、1 个输出层。本文提取了 3D-Inception-ResNet 模型分类输出前一层的特征,总共 512 种 CNN 特征。

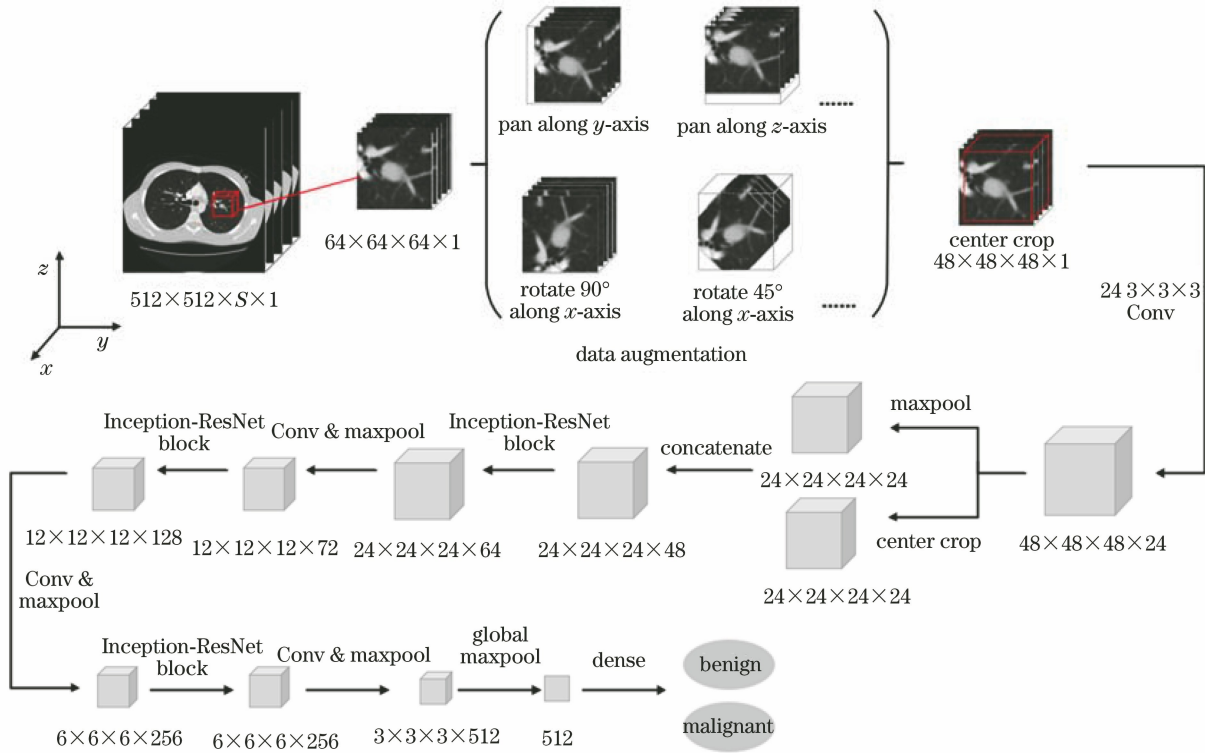


图 3 3D-Inception-ResNet 模型架构

Fig. 3 Architecture of 3D-Inception-ResNet model

图 4 为 Inception-ResNet 模块。与原始 Inception 模块相比, Inception-ResNet 模块增加了 shortcut 结构, 基于输入数据共有 4 条卷积分支, 分别模拟不同大小的感受野; 叠加使用 $3 \times 3 \times 3$ 卷积核, 可以在替代大卷积核的同时减少模型参数量; 最后对各卷积分支所得特征图进行级联, 使用 $1 \times 1 \times 1$ 卷积操作对齐维度, 将结果与输入数据相加, 得到 Inception-ResNet 模块输出。

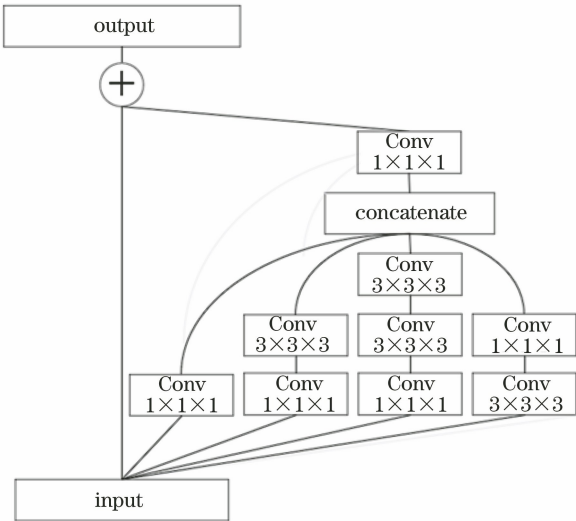


图 4 Inception-ResNet 模块

Fig. 4 Inception-ResNet module

对于二分类网络, Softmax 函数被定义为

$$p_i = \frac{\exp y_i}{\exp y_0 + \exp y_1}, \quad i = 0, 1, \quad (1)$$

式中: y_0, y_1 均为 CNN 的输出; p_i 为肺结节样本预测为良性或恶性的概率。通过最小化损失函数来训练网络模型, 损失函数被定义为

$$L = -\frac{1}{N} \sum [q \log p_1 + (1 - q) \log p_0], \quad (2)$$

式中: N 为批处理大小, 设置为 16; q 为肺结节样本的真实标签, $q = 0$ 或 1, 即良性或恶性。使用一阶 Adam 算法优化目标函数^[20], 初始学习率为 0.0015, 每训练 8 轮次后, 学习率将缩小 1/10, 直至训练结束, 总共学习 30 轮次。

2.3 分类部分

为了更大程度地丰富肺结节特征的代表性, 组合两类特征, 图 5 为特征集可视化图。行表示结节, 共包括 LIDC-IDRI 数据集中的 680 个肺结节样本; 列表示特征, 包含传统影像学特征 103 个。3D-Inception-ResNet 模型提取特征 512 个, 共 615 个特征。由于影像学特征与 CNN 特征产生方式不同, 在两类特征的过渡点存在明显的灰度差异, 特征被归一化至 $[0, 1]$, 以便可视化。接下来使用组合特征集完成分类鉴别。

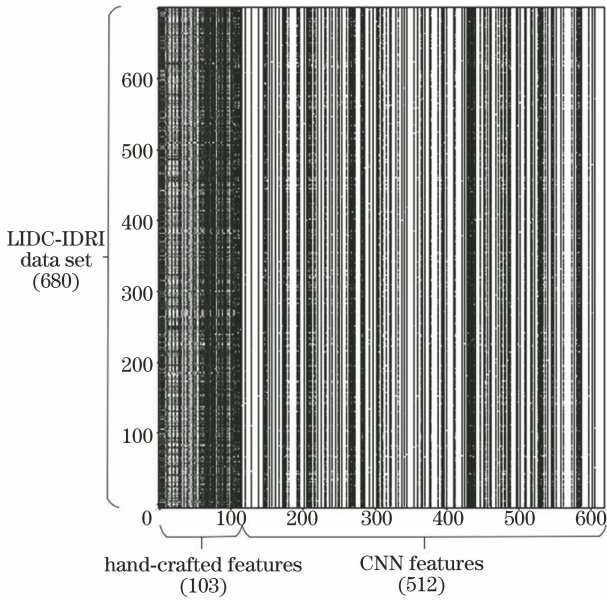


图 5 特征集可视化图

Fig. 5 Visualization map of features

准确、有效的特征向量可以提高肺结节的分类精度,但当特征冗余达到一定程度时,不仅会增加整个模型的运算量,而且还会降低肺结节分类的精度。

因此,特征优化有助于提高分类模型的性能。采用 RF 模型进行特征选择,挑选出权重高于 0.01 的图像特征构成特征子集^[21]。最后,将优化特征子集输入 SVM、RF 等传统分类器中,实现肺结节良恶性分类鉴别。

3 分析与讨论

3.1 实验设置

实验设置如表 1 所示,包含 3 种实验:1)分别基于传统影像学特征、3D-Inception-ResNet 模型、结合 CNN 特征与影像学特征的分类实验;2)在三种不同样本配置方案下的分类实验,样本配置方案如表 2 所示,三种样本配置方案分别认为恶性度为 3 的结节为丢弃、良性、恶性;3)与 3D-DenseNet、3D-ResNet 等经典网络架构的对比实验。为确保实验的可靠性,训练集与验证集采用十折交叉验证方法确定,所有分类器都被设置为默认参数,CNN 模型是基于 TensorFlow 框架实现的,在配备有 Intel Xeon E5-2673 CPU 2.4GHz 处理器,NVIDIA GeForce GTX 1080 显卡的工作站上完成实验。

表 1 实验设置说明

Table 1 Description of experiment setting

Experiment	Sub experiment	Description
1	a	Classification based on traditional hand-crafted features
	b	Classification based on 3D-Inception-ResNet model
	c	Classification combining CNN features and hand-crafted features
	d	Classification on the Shanghai Chest Hospital dataset
2		Classification under three different sample configuration schemes
3		Contrast of different architectures (including DenseNet and ResNet)

表 2 三种肺结节样本配置方案

Table 2 Configuration scheme of three pulmonary nodule samples

Description	Configuration	Number of benign nodules	Number of malignant nodules
'1', '2' as benign and '4', '5' as malignant	1	380	300
'1', '2', '3' as benign and '4', '5' as malignant	2	736	300
'1', '2' as benign and '3', '4', '5' as malignant	3	380	656

采用以下评价指标衡量分类模型性能:1)准确度(A),正确分类的样本占所有样本的比例;2)敏感度(S_{EN}),正确分类的恶性结节占所有真实恶性结节的比例;3)特异度(S_{PE}),正确分类的良性结节占所有真实良性结节的比例;4)AUC、ROC 曲线是一种显示分类器真阳性率与假阳性率之间折中的图像化方法^[22]。AUC 通常用于评估分类模型的性能。 N_{TP} 和 N_{TN} 分别表示正确分类的恶性结节、良性结节样本数; N_{FP} 和 N_{FN} 分别表示错误预测为恶性结节、良性结节样本数。

$$A = \frac{(N_{TP} + N_{TN})}{(N_{TP} + N_{TN} + N_{FP} + N_{FN})}, \quad (3)$$

$$S_{EN} = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (4)$$

$$S_{PE} = \frac{N_{TN}}{N_{TN} + N_{FP}}. \quad (5)$$

3.2 结果与分析

3.2.1 结合 CNN 特征与影像学特征的肺结节良恶性分类

表 3 为实验 1 分类结果,其中 NB 为朴素贝叶

斯分类器, KNN 为 K 最近邻分类器, LDA 为线性判别分类器。组合特征在 RF 分类器与 SVM 中表现优异, 分别取得 92.75% 与 94.98% 的分类精度。与传统影像学特征分类方法相比, 融合 CNN 特征

的分类方法的各评价指标显著提高, 证明 3D-Inception-ResNet 模型提取了能有效反映病灶区域信息的 CNN 特征, 从而进一步提高了肺结节良恶性分类精度。

表 3 实验 1 分类结果比较

Table 3 Comparison of classification results in experiment 1

Method	$A / \%$	$S_{EN} / \%$	$S_{PE} / \%$	AUC / %
Hand-crafted features+Gaussian-NB	89.69	85.18	91.43	95.19
Hand-crafted features+KNN	90.72	83.33	94.03	95.54
Hand-crafted features+RF	91.64	89.42	93.06	96.78
Hand-crafted features+LDA	91.75	86.21	94.12	96.18
Hand-crafted features+SVM	91.81	88.66	94.59	96.53
3D-Inception-ResNet	91.44	92.87	91.09	96.27
Combined features+Gaussian-NB	91.75	86.21	94.11	95.99
Combined features+KNN	91.81	88.66	94.59	96.53
Combined features+RF	92.75	92.12	93.34	97.11
Combined features+LDA	90.72	88.46	91.55	96.46
Combined features+SVM	94.98	90.02	97.03	97.43

从 LIDC-IDRI 数据库中选择 4 个典型结节证明所提方法的有效性, 如图 6 所示。图 6(a) 结节为良性结节, 图 6(b) 为恶性结节, 基于传统影像学特征方法对这两种结节的预测都是正确的, 而 3D-Inception-ResNet 模型却预测错误; 相反地, 图 6(c)

为良性结节, 图 6(d) 为恶性结节, 3D-Inception-ResNet 模型对这两种结节的预测都是正确的, 而传统影像学特征方法却预测错误。所提方法能正确分类鉴别这 4 种结节, 充分说明, 组合特征能更大程度丰富肺结节特征, 提高结节分类精度。

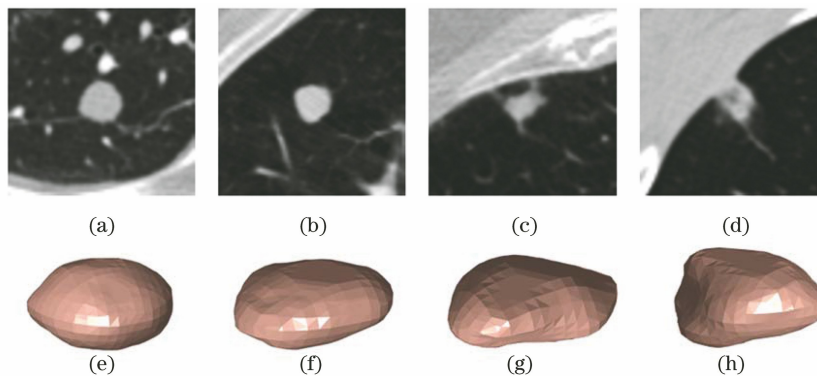


图 6 LIDC-IDRI 数据库中 4 个典型结节。(a)~(d) 肺结节二维切片; (e)~(h) 对应结节的三维显示图

Fig. 6 Four typical nodules in LIDC-IDRI database. (a)~(d) 2D slices of nodules; (e)~(h) 3D displays of corresponding nodules

除此之外, 在保证所有数据匿名的情况下, 使用上海市胸科医院 (SCH) 的 36 名肺癌诊断患者数据进行额外验证, 每例数据均包括 CT 影像和病理检验报告, 根据医生标注结果与病理活检结果, 最终得到 66 个肺结节样本, 其中良性结节有 24

个, 恶性结节有 42 个。该数据仅作为测试数据来验证所提方法的临床可实施性。表 4 为实验结果, SCH 测试集数据仍然能达到较高的分类精度, 充分说明所提方法具有鲁棒性, 是一种有效的计算机辅助诊断方法。

表 4 SCH 数据集分类结果

Table 4 Classification results of SCH dataset

Dataset	$A / \%$	$S_{EN} / \%$	$S_{PE} / \%$	AUC / %
LIDC-IDRI	94.98	90.02	97.03	97.43
SCH	90.91	88.10	95.83	95.58

3.2.2 三种不同样本配置方案下的肺结节良恶性分类

表 5 为三种不同样本配置方案的分类结果比较。图 7 为三种配置方案的 ROC 曲线对比。配置 1 的总体分类性能优于配置 2 和配置 3, 证明去除专家标注的恶性度为‘3’的不确定结节后, 结节恶

性度特征更明确, 分类结果更好。配置 2 的分类性能优于配置 3, 证明专家标注的恶性度为‘3’的不确定结节本质上更趋向于良性结节, 这与 Dhara 等^[8]实验结论一致。实验结果表明, 所提方法对不确定性结节也具有较弱的分辨能力。

表 5 三种不同样本配置方案的分类结果比较

Table 5 Comparison of classification results in three different sample configuration schemes

Configuration	Method	A / %	S _{EN} / %	S _{PE} / %	AUC / %
Configuration 1	Combined features+RF	92.75	92.12	93.34	97.11
	Combined features+SVM	94.89	90.02	97.03	97.43
Configuration 2	Combined features+RF	86.04	70.87	90.20	91.70
	Combined features + SVM	85.86	70.20	90.55	91.12
Configuration 3	Combined features+RF	80.38	85.11	72.37	87.26
	Combined features+SVM	79.73	82.35	74.91	85.31

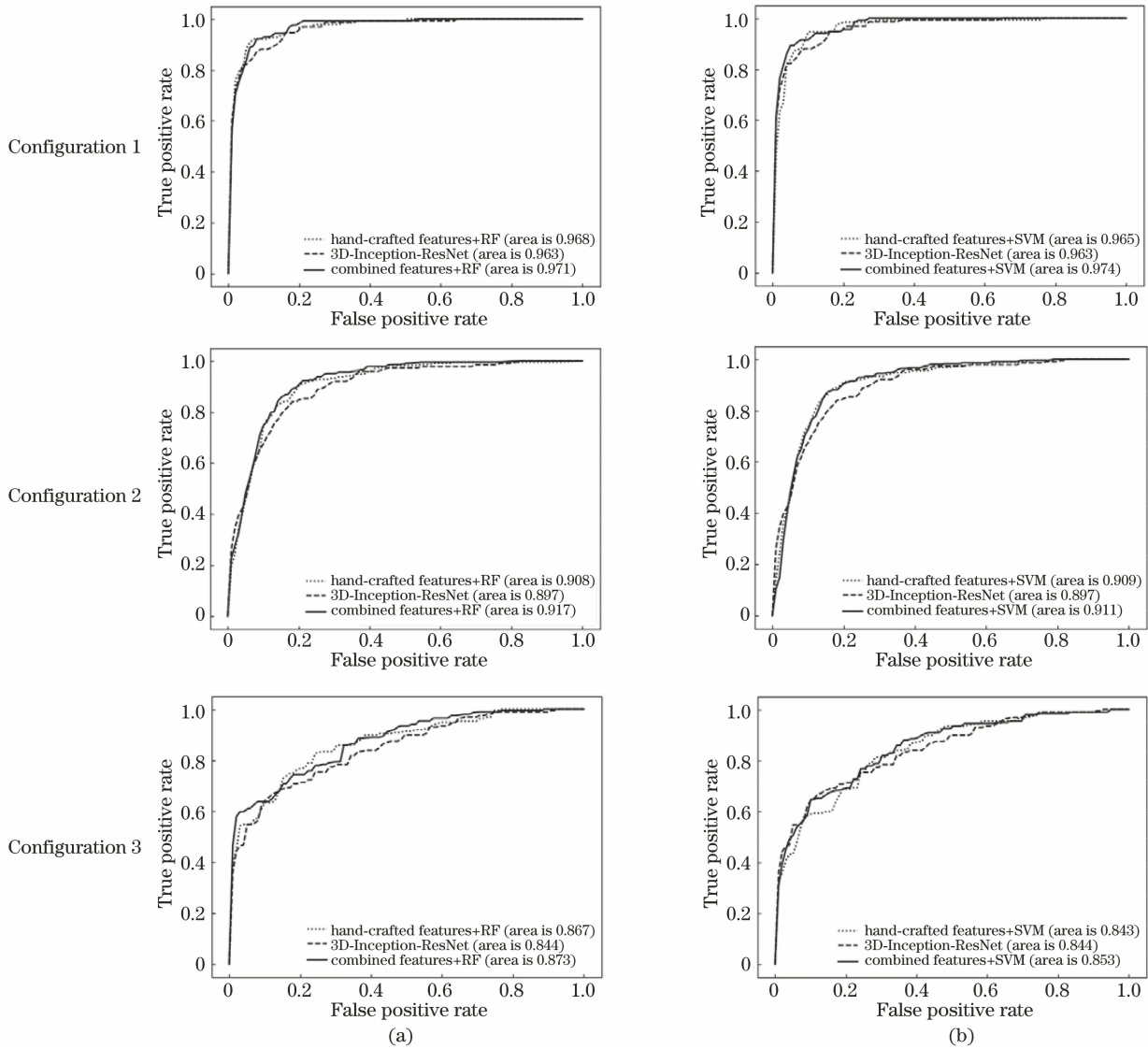


图 7 在不同分类器中的 ROC 曲线。(a) RF; (b) SVM

Fig. 7 ROC curves in different classifiers. (a) RF; (b) SVM

3.2.3 经典卷积神经网络架构对比

为了进一步评估所提方法的有效性,在相同数据集划分情况下,使用 3D-DenseNet^[23]、3D-

ResNet^[24-25]等经典 CNN 架构提取 CNN 特征,并结合影像学特征进行分类实验。图 8 为搭建的 3D-DenseNet 与 3D-ResNet 模型图。

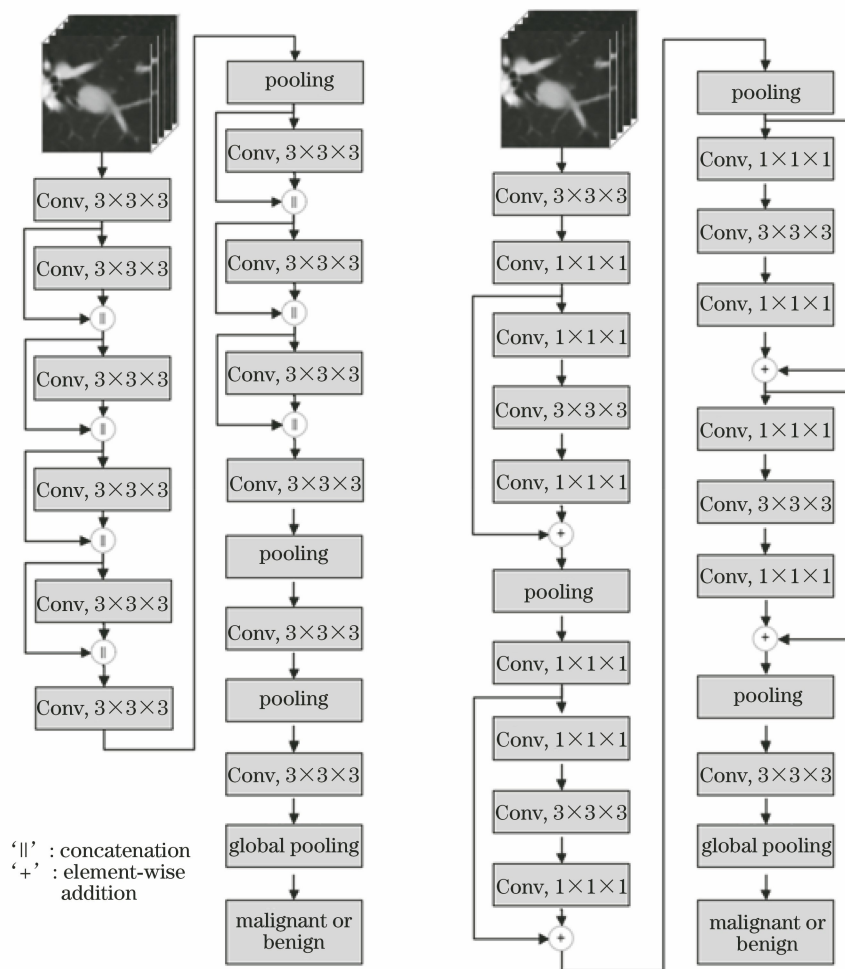


图 8 经典 CNN 架构。(a) 3D-DenseNet 模型;(b) 3D-ResNet 模型

Fig. 8 Classic CNN architecture. (a)3D-DenseNet model; (b) 3D-ResNet model

表 6 为不同 3D CNN 的分类结果对比,实验结果表明:在 3D CNN 分类模型中,3D-Inception-ResNet 的分类性能优于 3D-DenseNet 与 3D-

ResNet;相较于 CNN 分类模型,结合传统影像学 with CNN 特征的分类模型具有更优异的分类性能。

表 6 不同 CNN 架构的分类结果对比

Table 6 Comparison of classification results of different CNN architectures

Architecture	A /%	S _{EN} /%	S _{PE} /%	AUC /%
3D-DenseNet	89.69	87.10	94.55	92.64
3D-DenseNet combining hand-crafted features	90.72	90.00	89.47	94.20
3D-ResNet	84.69	76.74	90.91	88.13
3D-ResNet combining hand-crafted features	90.82	90.48	91.07	95.38
3D-Inception-ResNet	91.44	92.87	91.09	96.27
3D-Inception-ResNet combining hand-crafted features	94.98	90.02	97.03	97.43

3.2.4 最新研究成果对比

为了更客观地评估所提方法,对所提方法和近些年使用的其他相似方法进行对比。表 7 为所提方

法与其他方法的结果比较。Causey 等^[26]使用 NoduleX 分类模型进行肺结节良恶性分类,通过构建 2D-CNN21 与 2D-CNN47 网络提取高级特征,并

将其与传统影像学特征组合后,使用 RF 分类器进行分类鉴别,实验表明,此方法可取得更高的分类精度。Li 等^[27]提取 29 种影像学特征,并构建 3D-AlexNet、3D-VGG16Net 及 3D-Multi-cropNet 模型

提取输出层分类结果,组合两类特征后,使用顺序前向特征选择方法选出最优特征子集,并使用 SVM 分类鉴别。实验结果表明,所提方法的整体分类性能优于这两种方法。

表 7 不同方法的结果比较

Table 7 Comparison of the results of different methods

Method	Number of nodules	A /%	S _{EN} /%	S _{PE} /%	AUC /%
Method in Ref. [26]	664	93.20	87.90	98.50	97.10
Method in Ref. [27]	1226	85.62±2.37	81.21±6.20	89.56±1.17	90.45±2.58
Proposed method	1036	94.98	90.02	97.03	97.43

4 结 论

提出了一种结合深度学习与影像学特征的肺结节良恶性分类模型,并使用 LIDC-IDRI 数据集进行实验验证。与其他现有的良恶性分类模型相比,所提方法具有以下优势。1)基于医生勾画的金标准分割肺结节,省略了设计算法分割肺结节的步骤,具有更高的可靠性。2)提取了 103 种传统影像学特征与 512 种 CNN 特征,组合两类特征并使用基于随机森林的特征选择,减少了大量不相关特征和冗余特征。所提方法充分运用了传统影像学特征与 CNN 高级抽象特征的优势,既克服了传统影像学特征不能充分反映病灶区域信息的问题,又弥补了在数据量不足的情况下 3D CNN 分类效果不好的劣势。3)尝试了众多传统分类器,在 SVM 中取得最佳分类效果,实验 1 的分类准确率、敏感度、特异度及 AUC 分别达 94.98%、90.02%、97.03% 及 97.43%,实验结果表明,所提方法能准确地判别肺结节的良恶性。4)设置了 3 种不同的样本配置方案,3 种方案分别认为专家标注的恶性度为‘3’的结节为丢弃、良性和恶性,最终 AUC 分别为 97.43%、91.70% 和 87.26%,结果表明,所提方法对不确定性结节也具有较弱的分辨能力。5)对 3D-DenseNet、3D-ResNet、3D-Inception-ResNet 模型的性能进行对比,结果表明,相较于单一 3D CNN 分类模型,结合传统影像学特征与 CNN 特征的分类模型具有更优异的分类性能,其中 3D-Inception-ResNet 模型的实验效果最佳。

所提方法的不足与改进之处:1)使用临床医生勾画的金标准作为肺结节分割结果,尚未实现肺结节的自动分割;2)使用的分类器参数大多采用软件默认值,需进一步评估默认参数对实验结果的影响;3)仅使用 LIDC-IDRI 数据库验证所提方法的有效性,今后还需要更多数据库和临床数据进行验证。

参 考 文 献

- [1] Torre L A, Siegel R L, Ward E M, et al. Global cancer incidence and mortality rates and trends: an update [J]. *Cancer Epidemiology, Biomarkers & Prevention*, 2016, 25(1): 16-27.
- [2] Chen W Q, Zheng R S, Baade P D, et al. Cancer statistics in China, 2015 [J]. *CA: A Cancer Journal for Clinicians*, 2016, 66(2): 115-132.
- [3] Yu H, Liu S Y, Li H M, et al. Application of imaging examination technology in the diagnosis of lung cancer [J]. *Journal of Diagnostics Concepts & Practice*, 2010, 9(2): 124-128.
于红, 刘士远, 李惠民, 等. 影像学检查技术在肺癌诊断中的应用 [J]. *诊断学理论与实践*, 2010, 9(2): 124-128.
- [4] Feng Y, Yi B S, Wu C Y, et al. Pulmonary nodule recognition based on three-dimensional convolution neural network [J]. *Acta Optica Sinica*, 2019, 39(6): 0615006.
冯雨, 易本顺, 吴晨玥, 等. 基于三维卷积神经网络的肺结节识别研究 [J]. *光学学报*, 2019, 39(6): 0615006.
- [5] Renfrew D L, Franken E A, Berbaum K S, et al. Error in radiology: classification and lessons in 182 cases presented at a problem case conference [J]. *Radiology*, 1992, 183(1): 145-150.
- [6] Zia ur Rehman M, Javaid M, Shah S I A, et al. An appraisal of nodules detection techniques for lung cancer in CT images [J]. *Biomedical Signal Processing and Control*, 2018, 41: 140-151.
- [7] Firmino M, Morais A H, Mendonça R M, et al. Computer-aided detection system for lung cancer in computed tomography scans: review and future prospects [J]. *BioMedical Engineering OnLine*, 2014, 13(1): 41.
- [8] Dhara A K, Mukhopadhyay S, Dutta A, et al. A combination of shape and texture features for classification of pulmonary nodules in lung CT images [J]. *Journal of Digital Imaging*, 2016, 29(4): 466-

- 475.
- [9] Kaya A, Can A B. A weighted rule based method for predicting malignancy of pulmonary nodules by nodule characteristics [J]. *Journal of Biomedical Informatics*, 2015, 56: 69-79.
- [10] Nibali A, He Z, Wollersheim D. Pulmonary nodule classification with deep residual networks [J]. *International Journal of Computer Assisted Radiology and Surgery*, 2017, 12(10): 1799-1808.
- [11] Shen W, Zhou M, Yang F, et al. Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification [J]. *Pattern Recognition*, 2017, 61: 663-673.
- [12] Armato S G, McLennan G, Bidaut L, et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans [J]. *Medical Physics*, 2011, 38(2): 915-931.
- [13] Manos D, Seely J M, Taylor J, et al. The lung reporting and data system (LU-RADS): a proposal for computed tomography screening [J]. *Canadian Association of Radiologists Journal*, 2014, 65(2): 121-134.
- [14] Ochs R, Kim H J, Angel E, et al. Forming a reference standard from LIDC data: impact of reader agreement on reported CAD performance [J]. *Proceedings of SPIE*, 2007, 6514: 65142A.
- [15] McWilliams A, Tammemagi M C, Mayo J R, et al. Probability of cancer in pulmonary nodules detected on first screening CT [J]. *The New England Journal of Medicine*, 2013, 369(10): 910-919.
- [16] Snoeckx A, Reyntiens P, Desbuquoit D, et al. Evaluation of the solitary pulmonary nodule: size matters, but do not ignore the power of morphology [J]. *Insights into Imaging*, 2018, 9(1): 73-86.
- [17] Zhou Z G, Folkert M, Iyengar P, et al. Multi-objective radiomics model for predicting distant failure in lung SBRT [J]. *Physics in Medicine and Biology*, 2017, 62(11): 4460-4478.
- [18] Yang J, Wang S, Zhou P. Recognition and classification for three-dimensional model based on deep voxel convolution neural network [J]. *Acta Optica Sinica*, 2019, 39(4): 0415007.
- 杨军, 王顺, 周鹏. 基于深度体素卷积神经网络的三维模型识别分类 [J]. *光学学报*, 2019, 39(4): 0415007.
- [19] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, Inception-ResNet and the impact of residual connections on learning [EB/OL]. (2016-08-23) [2020-07-12]. <https://arxiv.org/abs/1602.07261>.
- [20] Kingma D P, Ba J. Adam: a method for stochastic optimization [EB/OL]. (2017-01-30) [2020-07-12]. <https://arxiv.org/abs/1412.6980>.
- [21] Díaz-Uriarte R, de Andrés S A. Gene selection and classification of microarray data using random forest [J]. *BMC Bioinformatics*, 2006, 7: 3.
- [22] Fawcett T. An introduction to ROC analysis [J]. *Pattern Recognition Letters*, 2006, 27(8): 861-874.
- [23] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2261-2269.
- [24] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [25] Huang S, Li F F, Chen Q. Computed tomography image classification algorithm based on improved deep residual network [J]. *Acta Optica Sinica*, 2020, 40(3): 0310002.
- 黄盛, 李菲菲, 陈虬. 基于改进深度残差网络的计算断层扫描图像分类算法 [J]. *光学学报*, 2020, 40(3): 0310002.
- [26] Causey J L, Zhang J Y, Ma S, et al. Highly accurate model for prediction of lung nodule malignancy with CT scans [J]. *Scientific Reports*, 8(1): 9286.
- [27] Li S L, Xu P P, Li B, et al. Predicting lung nodule malignancies by combining deep convolutional neural network and handcrafted features [J]. *Physics in Medicine and Biology*, 2019, 64(17): 175012.