【电子与信息科学 / Electronics and Information Science】

基于多尺度特征融合的肺结节良恶性分类方法

顾军华^{1,2}, 孙哲然^{1,3}, 王 锋^{1,2}, 戚永军^{1,4}, 张亚娟^{1,2}

1) 河北工业大学河北省大数据计算重点实验室, 天津 300401; 2) 河北工业大学人工智能与数据科学学院, 天津 300401;
3) 河北工业大学电子信息工程学院, 天津 300401; 4) 北华航天工业学院信息技术中心, 河北廊坊 065000

摘 要: 为解决肺结节分类问题中肺电子计算机断层扫描(computed tomography, CT)图像特征提取不全面和随着卷积网络深度的加深导致的梯度消失问题, 提出一种基于多尺度特征融合网络(multi-scale feature fusion network, MSFFNet)的肺结节良恶性自动分类模型. 使用多尺度卷积操作对输入的肺结节 CT 图像分别进行不同范围的特征提取和特征的融合拼接, 解决特征提取不全面的问题; 引入 SE-ResNeXt 模块, 充分利用通道注意力机制, 有效解决特征信息丢失的问题; 输出肺结节良恶性的分类结果. 在大型公开可用的肺图像联合数据库(lung image database consortium, LIDC-IDRI)上进行实验, MSFFNet 模型的分类准确率达 97.2%, 特异性和敏感性分别为 96.14% 和 98.62%, 优于 SE-ResNeXt 等方法的分类效果.

关键词: 人工智能; 深度学习; 特征提取; 特征融合; 通道注意力; 肺结节分类; SE-ResNeXt; 电子计算机断层扫描图像

中图分类号: TP391

文献标志码: A

doi: 10.3724/SP.J.1249.2020.04417

Classification of benign and malignant pulmonary nodules based on multi-scale feature fusion

GU Junhua^{1,2}, SUN Zheran^{1,3}, WANG Feng^{1,2},
QI Yongjun^{1,4}, and ZHANG Yajuan^{1,2}

1) Hebei Province Key Laboratory of Big Data Computing, Hebei University of Technology, Tianjin 300401, P. R. China

2) School of Artificial Intelligence, Hebei University of Technology, Tianjin 300401, P. R. China

3) School of Electronics and Information Engineering, Hebei University of Technology, Tianjin 300401, P. R. China

4) Information Technology Center North China Institute of Aerospace Engineering, Langfang 065000, Hebei Province, P. R. China

Abstract: In order to solve the problem of incomplete feature extraction in lung computed tomography (CT) image and the gradient disappearance caused by network degradation with the deepening of convolution network, an automatic classification model of pulmonary nodule malignancy based on multi-scale feature fusion network (MSFFNet) is proposed. The multi-scale convolution operation is used to extract and fuse the features of different ranges of input CT images. The SE-ResNeXt module is introduced to make full use of the channel attention mechanism to effectively solve the problem of feature information loss. Finally, the classification results of benign and malignant pulmonary nodules are obtained. The accuracy of MSFFNet classification model achieves 97.2%, and the specificity and sensi-

Received: 2019-10-27; **Accepted:** 2020-03-21

Foundation: National Natural Science Foundation of China (61702157); Natural Science Foundation of Hebei Province (F2016202144)

Corresponding author: Professor GU Junhua. E-mail: jhgu_hebut@163.com

Citation: GU Junhua, SUN Zheran, WANG Feng, et al. Classification of benign and malignant pulmonary nodules based on multi-scale feature fusion [J]. Journal of Shenzhen University Science and Engineering, 2020, 37(4): 417-424. (in Chinese)

tivity achieve 96.14% and 98.62%, respectively, which are better than those of SE-ResNeXt.

Key words: artificial intelligence; deep learning; feature extraction; feature fusion; channel attention; pulmonary nodule classification; SE-ResNeXt; computed tomography (CT) image

肺癌是当前人类疾病中最常见的恶性肿瘤之一^[1], 早发现、早确诊是降低死亡率的希望所在. 医学上将胸部电子计算机断层扫描(computed tomography, CT)上的“肺部斑点”定义为肺部结节, 肺结节是肺癌最重要的初期表征之一, 如果能做到及早准确发现肺结节, 会大幅降低漏诊率和误诊率.

计算机辅助诊断结果能够为医生提供第二意见. 尽管已经有肺癌辅助诊断系统被用于协助放射科医生阅读胸部 CT, 但一些容易发展成肺癌的结节常会因表现出类似良性病变而难以区分, 如肺脏错构瘤、肺结核球和霉菌球赫尔炎性假瘤等, 易出现假阴性导致漏诊; 也可能将非病变解读为病变, 或将良性病变误解为恶性, 导致假阳性结果. 因此, 判别肺结节恶性度的计算机辅助诊断技术仍需不断的发展和改进. 早期传统的医学图像分类方法依赖于手动提取特征和分类器分类两个步骤, 特征提取是肺结节良恶性分类的关键. 传统的肺结节良恶性特征提取方法主要从生长速度、纹理特性、形状特性和多特征融合来设计. ELBAZ 等^[2]将两步标记方法用于精确地监测相应结节之间的体积变化, 能够有效判断肺结节良恶性, 缺点是需长时间跟踪结节变化. WAY 等^[3]先分割结节, 再提取纹理特征来训练一个线性判别分类器. HAN 等^[4]对多个方向上的多个相邻体素之间的关系进行纹理特征分析. 常用良恶性分类器包括支持向量机和 k 近邻等. SCHILHAM 等^[5]提取的结节特征并输入到 k 近邻分类器中进行训练, 但该方法发生假阳性的概率很高.

基于深度学习的肺癌辅助诊断日渐成为该领域最为活跃的研究方向之一, 它可将特征表示和分类器分类结合到一个模型当中, 进行一体化的学习, 特别是深度卷积神经网络, 能自动学习到图像不同层次的特征信息, 实现端到端的分类. 卷积神经网络(convolutional neural networks, CNN)局部感知的特点决定了其在提取自然图像或者医学图像特征的过程是从局部到全局、从低级特征到高级语义特征, 许多方法通过改进卷积层的数量和组合方式, 来达到分类效果的提升. KUMAR 等^[6]应用自动编码器网络提取肺结节特征, 采用二叉决策树对结节良恶性进行分类, 并在肺图像联合数据库(lung

image database consortium, LIDC-IDRI)^[7]上获得了 75.01% 的准确率, 该方法在敏感性方面有较大提升, 但分类准确率相对较差. 多视角方法在肺结节诊断中被广泛应用. SETIO 等^[8]基于肺结节的实性结节、亚实性结节和实性大结节 3 种形态, 设计了对应的二维卷积检测模型, 对于每个候选对象, 从不同方向的平面提取一组二维切片, 输入网络中, 并使用特定融合方法将结果进行组合. XIE 等^[9]对每个三维结节提取了 9 种不同视角下的二维图像, 除了在每个二维平面切片以外, 在三维立方体的不同的对角方向也进行了切片, 从不同视角表征结节质异性, 并采用十折交叉进行了验证, 在 LIDC-IDRI 数据库上最终达到了平均 92.53% 的准确率. 以上的多视角结节良恶性诊断方法从不同角度提取结节特征, 但仍不可避免的存在因结节信息丢失, 难以对图像进行全面和深层次解析而导致结节诊断精度较低的问题. DEY 等^[10]设计了 4 个双通路 3 维卷积神经网络 MoDenseNet (multi-output dense connection) 用于肺结节的诊断, 选取 LIDC 数据集中的 686 个结节进行训练, 获得了 90.4% 的准确率, 分类准确率虽然获得了较大提升, 但因模型中采用的密集网络易造成特征爆炸, 不易于训练. ZUO 等^[11]利用卷积神经网络用作肺结节的边缘检测, 并将模型改进为一个适合于图像分类任务的多分辨率模型, 使用 LUNA16 数据集中 3 种不同大小的结节图片进行多分辨率测试, 准确率均超过 92.81%.

现有方法中, 许多模型在效率和精度方面得到了有效提升, 但相对复杂的模型往往容易过拟合, 而简单模型则会出现提取的图像特征单一, 对图像的表达会比较片面, 难以充分表示医学图像信息的缺点. 特别是针对肺结节这类微小目标识别分类问题时, 更需要在多个条件下提取特征, 做到特征的全面表达.

本研究针对由于肺结节结构复杂, 其形态容易根据其位置、尺寸和类型的不同而动态变化, 导致肺结节 CT 图像特征提取不全面的问题, 提出一个新的多尺度特征融合网络 (multi-scale feature fusion network, MSFFNet), 用于肺结节良恶性自动分类, 在公开数据集 LIDC-IDRI 上做了大量的实验, 达到了 97.2% 的准确率.

1 特征提取与融合模型 MSFFNet

本研究提出的肺结节分类模型 MSFFNet 旨在提取更高层次的语义信息和更全面的结节特征. 首先, 对输入的肺结节 CT 图像在单层卷积层中用多个不同大小的卷积核来卷积, 在不同尺度提取不同

范围的肺结节特征信息; 之后, 聚合多种不同感受野(receptive field)的特征获得性能增益, 在通道维度中拼接这些特征, 融合特征被输入到 3 个连续的 SE-ResNeXt 模块中, 以提取更高层次的语义信息; 最后, 特征经过全局平均池化和全连接操作, 经过 softmax 函数^[12]输出良恶性分类结果. 图 1 为本研究提出的肺结节良恶性分类模型.

图 1 多尺度特征融合的肺结节良恶性分类模型

Fig. 1 Classification model of pulmonary nodules based on multi-scale feature fusion

1.1 多尺度特征提取与融合

卷积核尺寸决定了特征图感受野的大小. 若感受野太大, 可能会阻碍模型对微小目标的分类; 反之, 若感受野太小, 会使网络忽略输入数据的全局信息. 卷积操作通常是将卷积核作为窗口在输入图像上以一定的步长滑动, 并在相应的位置作加权和. 一张长宽都为 m 的肺 CT 图像, 若步长为 z , $k(k \leq m)$ 个边长为 n 的卷积核卷积, 会得到 k 个大小为 $(m-n)/z+1$ 的特征图. 考虑到肺结节的位置多是非固定状态, 且不同肺结节的形态和大小也存在很大差异, 因此, 相应的结节特征也需从不同范围中提取. 不同尺度的卷积核表示不同范围的感受野, 可获取更全面的结节特征信息. 图 1 中的(a)

为多尺度特征提取与融合过程. 其中, I 为网络的多尺度提取部分, conv 表示卷积操作, 具体方法是将输入的肺结节图像分别经过 4 个不同尺寸的卷积核 (1×1 、 3×3 、 5×5 和 7×7), 深度为 64 的卷积层, 步长为 2, 使用填充(padding)操作保持输出特征图大小与输入保持一致. 特征融合过程如图 1 中 II 所示. 卷积后获得的 4 种特征进行通道拼接(concatenate), 网络在执行这一步骤的同时完成特征融合过程. 多尺度特征融合过程即特征向量的拼接, 表示为

$$\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3, \mathbf{Z}_4] \quad (1)$$

其中, $\mathbf{Z}_1$ 、 $\mathbf{Z}_2$ 、 $\mathbf{Z}_3$ 和 $\mathbf{Z}_4$ 分别为卷积获得的 4 种特征向量, 对应的特征通道数分别为 m 、 n 、 p 和 q . $\mathbf{Z}$ 对

应的通道数为 C , $C = m + n + p + q$.

1.2 特征重调

特征重调是通过特征图进行两次维度的变换, 捕捉通道之间的特征关系, 将特征按照重要性程度排序, 获得结节分类的判别特征, 使特征具有高度的层次性. 此过程主要依靠引入 SE-ResNeXt 网络模块来实现.

1.2.1 SE-ResNeXt 基本理论

在 MSFFNet 模型中引入的 SE-ResNeXt 模块是将缩聚-激发网络 (squeeze-and-excitation networks, SENet^[13]) 模块插入到 ResNeXt^[14] 中得到的, 它同时具备了两者的优点, 即在网络层数递增的同时, 减小网络退化, 又能建立特征通道之间的关联, 提升结节分类效果.

SENet 可作为单独模块插入现有的多种分类网络中. 其中, 最关键的两个步骤是缩聚 (squeeze) 操作和激发 (excitation) 操作. 前者利用全局平均池化在空间维度压缩特征, 把每个通道的二维特征变成一个实数, 该实数代表了特征通道上响应全局分布; 后者通过学习参数 W 为每个特征通道生成权重, 此权重被用来显式地对特征通道之间的相关性进行建模.

SENet 图像分类算法流程为:

1) 用 3×3 大小的卷积核, 对 C' 个通道, 大小为 $H' \times W'$ 的输入图像 X 进行卷积 (加权和计算), 卷积后得到 C 个大小为 $H \times W$ 的特征图 U , 其第 c 个特征图 u_c 的输出^[13] 为

$$u_c = v_c * X = \sum_{l=1}^{C'} v_c^l * x^l \quad (2)$$

其中, 该卷积核的权重组合为 V , $V = [v_1, v_2, \dots, v_C]$, v_c 是第 c 个卷积核; v_c^l 为第 c 个卷积核作用于第 l 个输入; $x^l = (x_{ij})^l$ 表示第 l 个输入的特征图, $l = 1, 2, \dots, C'$.

2) 进行 squeeze 操作. 将 C 个 $H \times W$ 大小的特征图 U 作为输入, 经过全局平均池化, 转换成 C 个大小为 1×1 的特征输出 m . 因此, 第 c 个特征图 u_c 的全局平均池化输出^[13] 为

$$m_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_{ijc} \quad (3)$$

3) 进行 excitation 操作. 如式 (4), 将 m 先后和网络学习到的参数 W_1 和 W_2 相乘, 即两个全连接操作. 经过这两个全连接操作, 实现了通道维度的变换, 得到输出权重为

$$s = \sigma(g(m, W)) = \sigma(W_2 \delta(W_1 m)) \quad (4)$$

其中, $\delta(\cdot)$ 为 ReLU 激活层; $\sigma(\cdot)$ 为 sigmoid 激活函数.

ResNeXt 网络模型是由残差网络为基础改进得到的, 高度模块化且易于搭建, 通过聚合相同的拓扑结构, 增加了分支卷积策略, path 为分支数, 则可在不增加参数复杂度的前提下提高模型准确率. 图 2 为 path = 32 时的 ResNeXt 结构示意图.

图 2 path = 32 的 ResNeXt 结构^[14]

Fig. 2 ResNeXt structure of path = 32^[14]

1.2.2 特征重调过程

在 MSFFNet 模型中, 肺 CT 图像经图 1 (a) 过程后, 将得到的特征图输入到 SE-ResNeXt 图像分类模块, 经过一个 3×3 的卷积操作, 再将 3 个连续的 SE-ResNeXt 模块嵌入网络, 对应的输出通道数分别为 64、128 和 256 个, 在充分增加网络深度的前提下提高了模型准确率, 缓解了网络退化, 同时能够捕捉通道间特征依赖关系. 对融合后的特征进行扁平化处理, 再把多维特征一维化, 作为卷积层到全连接层的过渡, 经过全连接层输出 256 维特征, 利用 softmax 函数分别输出良性结节和恶性结节的分类概率.

图 1 (b) 展示了本研究使用的 SE-ResNeXt 结构. 其中, C 为输入或输出特征的通道数; C/r 表示将特征维度降低到输入的 $1/r$, 以此来降低计算量, 本研究设 $r = 4$, 经过一个全连接层并采用 ReLU 激活函数激活后, 降维的特征升回到原来的维度, 大幅减少了计算复杂度. 经过 sigmoid 函数激活, 获得归一化的权重, 它将输出特征映射到 (0, 1) 的区间, 在用于特征相差不大时展现出较好的分类问题效果, 且单调递增又容易被优化. 经过 sigmoid 函数激活后, 归一化的权重被加权到每个通道的特征上, 从而可获得通道特征的重要性排序. 特征重排序操作过程为

$$\mathbf{x}_c = \mathbf{F}_{\text{scale}}(\mathbf{u}_c, s_c) = s_c \cdot \mathbf{u}_c \quad (7)$$

其中, $\mathbf{X} = (\mathbf{x}_c)$, $c = 1, 2, \dots, C$; $\mathbf{u}_c$ 为第 c 个通道特征的二维矩阵; s_c 为对应的权重. 当前特征通道二维矩阵的每个值都乘以该权重, 这些权重代表了每个特征通道的重要程度. 将它们加权到先前 C 个通道的特征上, 得到 $\mathbf{X}$, 完成对原特征的重新标定, 增强重要的特征, 削弱不重要的特征, 进一步提升了模型的分类效果.

2 MSFFNet 模型配置与训练

2.1 网络配置

MSFFNet 网络配置如表 1. 其中, multi_conv 为多尺度卷积; conv 为卷积操作; stride 为步长, 卷积步骤的分支数 path = 8; fc 为全连接, 其参数为 C/r 和 C , C 为特征通道数, 当 $r = 4$ 时, C/r 为 excitation 操作中将特征维度降低到输入的 $1/4$.

表 1 网络配置

Table 1 The configuration of the module

网络层	输出维度	卷积操作
multi_conv	64	$\begin{cases} \text{conv}, 1 \times 1 \\ \text{conv}, 3 \times 3 \\ \text{conv}, 5 \times 5 \\ \text{conv}, 7 \times 7 \end{cases}$
concatenate	256	—
conv	64	conv, 3×3 , 64, stride = 1
SE-ResNeXt_1	64	$\begin{cases} \text{conv}, 1 \times 1, 64 \\ \text{conv}, 3 \times 3, 64 \text{ path} = 8 \\ \text{fc}[16, 64] \end{cases} \times 3$
SE-ResNeXt_2	128	$\begin{cases} \text{conv}, 1 \times 1, 128 \\ \text{conv}, 3 \times 3, 128 \text{ path} = 8 \\ \text{fc}[32, 128] \end{cases} \times 3$
SE-ResNeXt_3	256	$\begin{cases} \text{conv}, 1 \times 1, 256 \\ \text{conv}, 3 \times 3, 256 \text{ path} = 8 \\ \text{fc}[64, 256] \end{cases} \times 3$
avg_pooling	256	conv, 2×2 , 256, stride = 1
fc + softmax	2	—

2.2 损失函数

MSFFNet 使用如式(8)的交叉熵损失函数 $H(p, q)$ 来衡量真实概率分布与预测概率分布之间的差

异, 函数值越小, 模型预测效果越好. 交叉熵便于梯度下降反向传播, 利于优化, 常用于识别图像类别问题.

$$H(p, q) = - \sum_x p(x) \ln q(x) \quad (8)$$

其中, p 为类别标签; q 为模型输出的实际值. 交叉熵用来判定实际值和预测值的接近程度, 因此, 通常最后一个输出层的节点数量与最终分类的目标数量相等, 在本研究的 MSFFNet 模型中, 表示输出一个二维数组, 其每个维度对应一个类别.

2.3 训练参数

设学习率为 0.01, 对 MSFFNet 肺结节良恶性分类模型进行训练. 设衰减率为 5×10^{-4} , 加入 L_2 正则化和 dropout^[15] 防止过拟合, 同时加入批量归一化^[16], 以解决梯度消失问题, 并设输入数据的批量大小 batch size 为 16, 模型总迭代次数为 1 744 次, 模型在 160 次遍历后收敛到最优值.

3 实验结果与分析

3.1 数据集与预处理

3.1.1 LIDC-IDRI 数据集

LIDC-IDRI (<https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>) 数据集包括 1 018 个肺结节 CT 扫描实例, 结节直径范围为 3 ~ 30 mm, 由 4 个放射科医生提供 2 个阶段的诊断标注, 每个图像包含至少 1 个肺结节. 所有结节的恶性程度, 由放射科医生在等级 1 ~ 5 之间进行评判, 评判等级越高表示恶性程度越高, 患癌风险越大. 图 3 展示了良性和恶性结节的形态.

3.1.2 数据增强

由于肺结节相对于整张 CT 图像来说, 面积占比往往非常小, 即使是恶性结节, 也只占很小一部分, 肺实质等构造对模型的训练会产生一定程度的影响; 此外, 良性结节的图像数量往往多于恶性结节的, 这种类别不平衡也会影响分类性能. 因此, 本研究采用以下预处理方法: 对原始图片进行像素提取, 以去除原始结节中潜在的冗余信息, 每张图片的大小最终为 64×64 像素. 随机选取部分图像并对每张图像都进行 -10° 、 -5° 、 5° 和 10° 的旋转操作, 再采用增加高斯噪声和椒盐噪声的方法对图像进行变换, 以增强模型的鲁棒性. 预处理后每类结节图像数量相等, 共包含 16 385 张肺结节图像, 由于评分为 3 的未知等级的结节存在不确定性, 影响

图 3 五种不同恶性度的肺结节

Fig. 3 Five lung nodules with different malignancies

模型的训练效果, 因此, 本研究仅从这些样本中取评分为 1 和 2 的结节将其划分为良性, 评分 4 和 5 的结节划分为恶性, 最终每个等级样本分布情况如表 2.

表 2 各等级样本数据量分布

Table 2 The data distribution of each level

样本等级	训练集	测试集
1	2 950	327
2	2 950	327
4	2 950	327
5	2 950	327
合计	11 800	1 308

3.2 评价指标

在肺结节分类问题中, 准确率 (accuracy) 是判定模型分类性能的重要指标. 本研究设准确率为分类正确的样本数与总体样本数之比, 即

$$\text{accuracy} = T / (T + N) \quad (9)$$

其中, T 为正确分类的样本数; N 为被错误分类的样

本数; $T + N$ 为总样本数.

此外, 低误诊率和低错诊率也是肺结节分类研究所追求的重要目标. 本研究选用受试者工作特征 (receiver operation characteristic, ROC) 曲线、敏感性 (sensitivity) 和特异性 (specificity) 作为评定指标. 敏感性指真正有病的患者中被确诊为患者的数量, 也称真阳性率, 如式 (10). 特异性表示真正无病的患者中被确诊为无病的数量, 如式 (11).

$$\text{sensitivity} = TP / (TP + FN) \quad (10)$$

$$\text{specificity} = TN / (FP + TN) \quad (11)$$

其中, TP 为真阳性, 即真实数据和预测数据都是阳性; FP 为假阳性, 表示真实数据为阴性, 预测结果为阳性; TN 为真阴性, 表示真实数据为阴性, 预测结果为阴性; FN 为假阴性, 表示真实数据为阳性, 预测结果为阴性.

4.3 实验结果对比

为找到可以使模型的分类表现达到最优的卷积核尺寸组合, 本研究分别验证了单一尺寸、2 种、3 种、4 种以及 5 种不同尺寸的卷积核对肺结节分类表现的影响, 实验结果如图 4.

图 4 多尺度卷积核的种类对分类性能的影响

Fig. 4 (Color online) The influence of the types of multi-scale convolution kernels on classification performance

由图 4 可见, 单一尺寸的卷积核分类效果最差; 采用 4 种不同尺寸的卷积操作时, 所得准确率最优; 采用 5 种不同尺寸的卷积时, 所得准确率和敏感性都出现了一定程度的下降, 而特异性指标升高, 这意味着模型的误诊率相对较低, 但准确率不高. 因此, 综合准确率、特异性和敏感性指标后, 本研究采用 4 种不同尺寸的卷积核进行多尺度卷积, 以期使模型的分类效果达到最佳.

为验证 SE-ResNeXt 模块数量对肺结节诊断效果的影响, 分别对比了不同数量 SE-ResNeXt 模块的诊断结果, 如图 5.

图5 SE-ResNeXt 模块数量对分类准确率的影响

Fig. 5 The effect of the number of SE-ResNeXt blocks on the classification accuracy

由图5可见,前期肺结节良恶性分类准确率随着SE-ResNeXt模块数量的增加而逐渐提高,但后期随着模块数量的增加,参数增多,模型复杂度亦增大,当SE-ResNeXt模块数量增至4个时,分类准确率开始下降.因此可认为,SE-ResNeXt模块数量为3时,MSFFNet模型拥有最优的结节良恶性分类效果,此时模型的分类准确率为97.2%,敏感性为98.62%,特异性为96.14%.

分别将MSFFNet与卷积神经网络(convolution neural networks, CNNs)^[17]和SE-ResNeXt^[13]的实验结果进行比较,得到ROC曲线如图6.由图6可见,CNNs模型分类效果相对较差,MSFFNet模型肺结节良恶性的分类效果最优.

图6 不同肺结节良恶性分类方法的ROC曲线

Fig. 6 ROC curve of different classification methods of pulmonary nodules

图7是采用MSFFNet模型和3种传统经典肺结节分类模型,包括CNNs^[17]、堆栈自编码器(stacked autoencoder, SAE)^[17]和VGG^[18],在准确率、特异性和敏感性方面的分类结果.由图7可见,MSFFNet模型可将良恶性分类准确率提升到97.2%,性能最佳.

图7 MSFFNet与经典分类模型的性能比较

Fig. 7 Comparison between MSFFNet and classical classification models

为验证MSFFNet分类模型性能,本研究将改进前的SE-ResNeXt分类模型以及近两年提出的其他几种肺结节良恶性分类方法在LIDC-IDRI预处理数据集上进行比较,结果如表3.由表3可见,MSFFNet模型不论是在准确率还是特异性和敏感性方面都具备更强的表现力,在模型的多卷积设计以及特征融合策略方面对比单独的SE-ResNeXt分类网络,MSFFNet模型的肺结节的良恶性分类结果优于其他分类模型.

表3 LIDC数据集上不同肺结节分类方法比较¹⁾

Table 3 Performance comparisons of different pulmonary nodule classification methods on the LIDC dataset %

方 法	准确率	敏感性	特异性
AL-SHABI等 ^[19]	88.46	88.66	—
XIE等 ^[9]	91.60	86.52	93.38
MoDenseNet ^[10]	90.40	90.47	90.33
SE-ResNeXt ^[13]	95.31	93.27	96.82
MSFFNet	97.20	98.62	96.14

¹⁾ 灰底数值为该指标上取得的最优结果

结 语

本研究提出一种基于卷积神经网络的多尺度特征融合的肺结节恶性度分类方法,首次在肺结节分类问题中采用SE-ResNeXt模块,在加深网络深度,提取更高级别语义信息的同时,减少了网络退化问题,表明多卷积的特征融合方法可改善网络模型的性能.对比实验结果证明,模型在肺结节良恶性分类方面能够表现出有效的分类结果.下一步我们将从超参数设计和损失函数等方面进行优化,结合现有计算机辅助诊断系统或网络端,在临床应用中进一步验证模型的可行性.

基金项目: 国家自然科学基金资助项目 (61702157); 河北省自然科学基金资助项目 (F2016202144)

作者简介: 顾军华 (1966—), 河北工业大学教授、博士生导师。研究方向: 智能信息处理和计算机视觉。
E-mail: jhgu_hebut@163.com

引文: 顾军华, 孙哲然, 王 锋, 等. 基于多尺度特征融合的肺结节良恶性分类方法[J]. 深圳大学学报理工版, 2020, 37(4): 417-424.

参考文献 / References:

- [1] BRAY F, FERLAY J, SOERJOMATARAM I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. CA: A Cancer Journal for Clinicians, 2018, 68: 394-424.
- [2] EI-BAZ A, SETHU P, GIMEL'FARB G, et al. Elastic phantoms generated by microfluidics technology: validation of an imaged-based approach for accurate measurement of the growth rate of lung nodules [J]. Biotechnology Journal, 2011, 6(2): 195-203.
- [3] WAY T W, HADJIISKI L M, SAHINER B, et al. Computer-aided diagnosis of pulmonary nodules on CT scans: segmentation and classification using 3D active contours [J]. Medical Physics, 2006, 33(7): 2323-2337.
- [4] HAN Fangfang, ZHANG Guopeng, WANG Huafeng, et al. A texture feature analysis for diagnosis of pulmonary nodules using LIDC-IDRI database [C]// IEEE International Conference on Medical Imaging Physics and Engineering. Shenyang, China, IEEE, 2013: 14-18.
- [5] SCHILHAM A M R, VAN GINNEKEN B, LOOG M. A computer-aided diagnosis system for detection of lung nodules in chest radiographs with an evaluation on a public database [J]. Medical Image Analysis, 2006, 10(2): 247-258.
- [6] KUMAR D, WONG A, CLAUSI D A. Lung nodule classification using deep features in CT images [C]// Proceedings of the 12th Conference on Computer and Robot Vision. Halifax, Canada: IEEE Computer Society, 2015: 133-138.
- [7] MCNITT-GRAY M F, ARMATO III S G, MEYER C R, et al. The lung image database consortium (LIDC) data collection process for nodule detection and annotation [J]. Academic Radiology, 2007, 14(12): 1464-1474.
- [8] SETIO A A A, CIOMPI F, LITJENS G, et al. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks [J]. IEEE Transactions on Medical Imaging, 2016, 35(5): 1160-1169.
- [9] XIE Yutong, XIA Yong, ZHANG Jianpeng, et al. Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT [J]. IEEE Transactions on Medical Imaging, 2018, 38(4): 991-1004.
- [10] DEY R, LU Zhongjie, HONG Yi. Diagnostic classification of lung nodules using 3D neural networks [C] // The 15th International Symposium on Biomedical Imaging. Washington D C: IEEE, 2018: 774-778.
- [11] ZUO Wangxia, ZHOU Fuqiang, LI Zuoxin, et al. Multi-resolution CNN and knowledge transfer for candidate classification in lung nodule detection [J]. IEEE Access, 2019, 7: 32510-32521.
- [12] LIU Weiyang, WEN Yandong, YU Zhiding, et al. Large-margin softmax loss for convolutional neural networks [C]// Proceedings of the 33rd International Conference on Machine Learning. New York, USA: International Machine Learning Society (IMLS), 2016: 507-516.
- [13] HU Jie, SHEN Li, SUN Gang, et al. Squeeze-and-excitation networks [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 7132-7141.
- [14] XIE Saining, GIRSHICK R, DOLLÁR P, et al. Aggregated Residual transformations for deep neural networks [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017: 1492-1500.
- [15] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from overfitting [J]. Journal of Machine Learning Research, 2014, 15(56): 1929-1958.
- [16] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [EB/OL] (2015-02-11) [2015-03-02]. <https://arxiv.org/pdf/1502.03167.pdf>.
- [17] SONG Qingzeng, ZHAO Lei, LUO Xingke, et al. Using deep learning for classification of lung nodules on computed tomography images [J]. Journal of Healthcare Engineering, 2017, 2017(1): 1-7.
- [18] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL] // (2014-09-04) [2015-04-10]. <https://arxiv.org/pdf/1409.1556.pdf>.
- [19] AL-SHABI M, LAN B L, CHAN W Y, et al. Lung nodule classification using deep local; global networks [J]. International Journal of Computer Assisted Radiology and Surgery, 2019, 14(10): 1815-1819.

【中文责编: 英 子; 英文责编: 木 柯】