

❖ 医学物理与工程学

Benign or malignant lung nodules classification model based on modified DenseNet

DAI Yaojun, YAN Shiju*, SONG Chengli

(School of Medical Instrument and Food Engineering, University of Shanghai for
Science and Technology, Shanghai 200093, China)

[Abstract] **Objective** To investigate the accuracy of classification of benign and malignant lung nodules with computer-aided diagnosis (CAD) systems. **Methods** DenseNet was used as the basic model, and 3D information of lung nodules was incorporated into training phase of the convolutional neural network via performing median density projection. Then the network structure was adapted to the classification of benign and malignant lung nodules. Traditional Cross Entropy Loss was replaced with Focal Loss to enable the network to focus on learning difficult-to-resolve lung nodules. **Results** The accuracy of the modified network model for classification of benign and malignant lung nodules was 89.93%, and the area under the curve was 0.947. **Conclusion** The adapted convolutional neural network model has high accuracy in the classification of benign and malignant lung nodules.

[Key words] Deep learning; Lung nodules; Benign or malignant classification; Lung image database consortium

DOI:10.13929/j.1003-3289.201710013

基于密集网络改进的肺结节良恶性分类模型

戴焱均, 闫士举*, 宋成利

(上海理工大学医疗器械与食品学院, 上海 200093)

[摘要] **目的** 探讨改进后的卷积神经网络模型对肺结节进行良恶性分类的准确率。**方法** 以分类模型密集网络(DenseNet)为基础模型,采用中间密度投影方法将肺结节的三维信息输入卷积神经网络进行训练,并针对肺结节良恶性分类问题适应性改进神经网络结构,将传统损失函数 Cross Entropy Loss 替换为 Focal Loss,使网络能着重学习难以分辨的肺结节。**结果** 改进后神经网络模型对良恶性肺结节分类的准确率为 89.93%,曲线下面积为 0.947。**结论** 适应性改进后的卷积神经网络模型判断良恶性肺结节准确率较高。

[关键词] 深度学习;肺结节;良恶性分类;肺部图像数据库联盟

[中图分类号] TP391.4; R563 **[文献标识码]** A **[文章编号]** 1003-3289(2018)07-1104-06

肺癌是发病率和死亡率居高不下的恶性肿瘤之一,早期检出至关重要。低剂量 CT 是检测早期无症状肺癌的有效方法,可发现 90% 的病灶^[1]。肺部病变常表现为肺结节,准确鉴别其良恶性非常重要。计算

机辅助诊断良恶性肺结节模型的鲁棒性好,且准确率高,其对良恶性肺结节的诊断分为两类。一类是基于从图像中提取量化特征(直方图特征^[2]、尺度不变特征转换特征^[3]、局部二进制模式特征^[4]及梯度方向直方图特征^[5]等),之后采用传统机器学习算法(支持向量机及随机森林等)对肺结节进行良恶性分类;但上述特征均由人工设计,难以对图像进行全面、深层次解析。另一类则直接基于图像本身或图像的某种变换图谱,采用层次更深的神经网络对肺结节进行良恶性分类,如 Song 等^[6-7]采用 LeNet^[8](4 层)卷积神经网络结

[第一作者] 戴焱均(1993—),男,江苏江阴人,在读硕士。研究方向:医学图像处理。E-mail: daiyaojun1993@163.com

[通信作者] 闫士举,上海理工大学医疗器械与食品学院,200093。

E-mail: yanshiju@usst.edu.cn

[收稿日期] 2017-10-09 **[修回日期]** 2018-04-05

构、Hussein 等^[9]采用基于 AlexNet^[10] (8 层) 卷积神经网络结构、Nibali 等^[11]基于 ResNet^[12] (18 层) 卷积神经网络结构等,但此类方法并未针对肺结节良恶性分类这一特定应用对现有模型进行有效改进,不能充分挖掘图像的深层次有用信息。本课题组提出基于 DenseNet^[13] (161 层) 改进的卷积神经网络,结合 DenseNet 具备残差结构、网络较深、能够融合浅层与深层特征进行分类的优势,针对肺结节良恶性分类这一特定问题对 DenseNet 的网络结构进行改进,将原有损失函数 Cross Entropy Loss 替换为 Focal Loss^[14],使得网络能够着重学习难以分辨的肺结节,以期提高准确率。

1 原理与方法

1.1 数据 本文所用数据集来源于公开的肺部图像数据库联盟 (Lung Image Database Consortium, LIDC),其内包含 1 018 组 CT 序列,每组序列对应一个 XML 注释文件,由 4 名高年资放射医师完成注释,包括肺结节(最大径 > 3 mm)的坐标和恶性程度评分 (1~5 分)。

选取层厚 ≤ 3 mm^[15-16] 的 CT 序列作为研究对象,共计 888 组、1 186 个肺结节,由 3 名及以上医师同时标注,取其均值作为总评分, > 3 分为恶性结节, < 3 分为良性结节, = 3 分为良恶性无法确定,予以舍弃。最终共获得 1 030 个肺结节的图像数据。

1.2 原理

1.2.1 中间密度投影 (median intensity projection, MIP) 卷积神经网络输入的是二维图像,但肺结节为三维结构,故采用 MIP^[9]对三维肺结节的外接立方体 I 分别沿 x 、 y 、 z 轴进行投影(图 1),每个轴投影获得

的 MIP 图像 φ 为:

$$\begin{aligned}\varphi(x, y) &= \underset{z}{\text{med}}[I(x, y, z)] \\ \varphi(y, z) &= \underset{x}{\text{med}}[I(x, y, z)] \\ \varphi(x, z) &= \underset{y}{\text{med}}[I(x, y, z)]\end{aligned}\quad (1)$$

其中 med 代表取中位数运算。将计算获得的 3 张二维 MIP 图像拼接为 1 个三维张量 Φ , $\Phi = [\varphi(x, y), \varphi(y, z), \varphi(x, z)]$,可视 Φ 为一张三通道的红绿蓝 (red green blue, RGB) 彩色图像,并能被输至卷积神经网络中进行训练。

1.2.2 数据集划分及扩充 对三维肺结节数据进行 MIP 处理后,获得 1 030 张 RGB 三通道图像(图 2)。采用十折交叉验证方法,将 1 030 个肺结节图像随机分为 10 组,对其中 9 组加椒盐噪声、高斯噪声、均值滤波、中值滤波、高斯滤波及双边滤波,随机旋转后将图像扩充 50 倍作为训练集;以剩余的 1 组作为测试集。重复上述步骤 10 次,获得交叉验证结果。

1.2.3 损失函数计算 为在神经网络迭代训练中不断调整参数、达到最优模型,需定义 1 个损失函数作为优化目标。

假设共有 n 张图像,令 $D_i = \{x_i, y_i\}$, $i \in (1, \dots, n)$,其中 D_i 代表数据集中第 i 个样本, x_i 代表第 i 张图像, y_i 代表 x_i 的良恶性。以其中 1 个训练样本图像 x 为例,神经网络进行正向传播,计算获得神经网络最后一层全连接层 (fully connected layer, FC) 的输出向量 v , $v = \{v_1, \dots, v_k\}$, $k \in (1, 2)$,向量 v_1 与 v_2 分别对应肺结节的良性与恶性特征。采用 SoftMax 为神经网络分类器,预测概率计算公式为:

$$p(k|x) = \frac{e^{v_k}}{\sum_{i=1}^2 e^{v_i}} \quad (2)$$

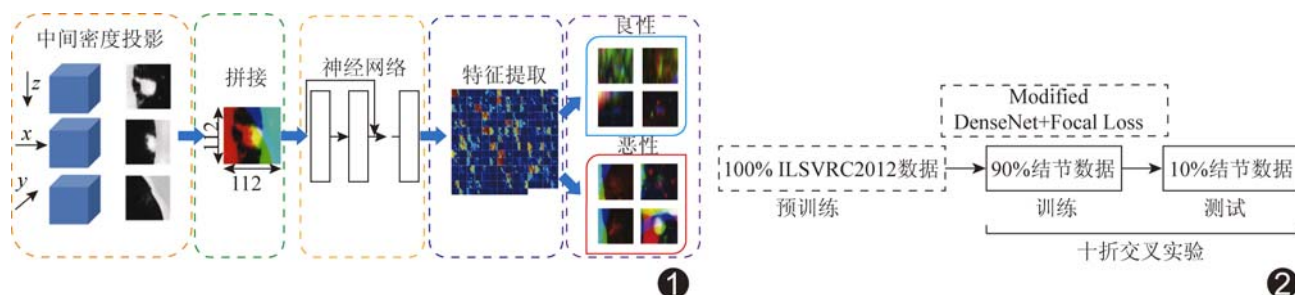


图 1 基于密集网络改进的肺结节良恶性分类流程图 图 2 训练与测试流程图 (ILSVRC2012:2012 大规模视觉识别挑战赛^[17]数据; Modified DenseNet+Focal Loss: 基于密集网络改进的肺结节良恶性分类模型)

神经网络常采用 Cross Entropy Loss 损失函数, n 个样本的损失函数 L 计算公式为:

$$L = - \sum_{i=1}^n \sum_{k=1}^2 \log(p(k|x)) \quad (3)$$

Focal Loss 损失函数计算公式为:

$$L = - \alpha (1 - p(k|x))^{\gamma} \log p(k) \quad (4)$$

式(4)中,超参数 α 由训练样本中的正负样本比例决定,由于本课题组在图像扩充时保持正负样本数量一致,故此处 α 取 1。

当神经网络对正确类别预测概率值较大(如当概率值 > 0.6)时,损失函数获得的损失值较小,权重在反向传播时获得的偏导数也较小,故权重调整幅度不大;而当神经网络对正确类别预测概率较小时,损失函数获得的损失值较大,权重在反向传播时获得的偏导数也较大,故权重调整幅度较大(图 3)。因此,神经网络即可着重学习“难以分辨”的肺结节。

1.2.4 残差结构 当网络层数较深时,训练会产生梯度消失^[12]。本课题组采用残差结构^[12]构建网络,见图 4,其中 x 代表前置层输入,不采用残差结构时,串联的网络相当于拟合函数 $y = F(x)$;采取残差结构后,神经元拟合函数优化为 $y = h(x) = F(x) + x$,代表残差结构网络拟合的函数功能为串联网络与输入 x 相加。进行误差反向传播时,由于 $F(x)$ 对 x 的偏导数不会恒定为数值 -1,保证了输出 y 对输入 x 偏导数不会恒等于 0,可减轻由层数递增导致的梯度为零的问题(图 4)。

残差结构还能将浅层与深层图像特征拼接,重复拼接可使网络最后一层 FC 层与肺结节图像的各层图像特征间接相连,可有效融合图像的浅层与深层特征,以对肺结节进行良恶性分类,充分挖掘图像蕴含的有

用信息,有望提高准确率^[13]。

1.2.5 模型微调 卷积神经网络需大量图像数据进行训练,而筛选后的肺结节图像数量较少,难以充分优化网络权重,本课题组采用模型微调的方法予以解决。首先采用大规模开源数据[2012 大规模视觉识别挑战赛^[17] (Large Scale Visual Recognition Challenge 2012, ILSVRC2012)]对模型进行预训练,然后对预训练后的模型进行参数微调。此法不仅能有效提升神经网络模型分类的准确性,还可大幅度加快损失函数的收敛速度。本研究后续实验均采用预训练模型微调参数的方法。

1.2.6 网络输入结构改进 由于肺结节图像本身较小,放大操作图像可影响神经网络对微小肺结节良恶性分类的学习,故尝试缩小神经网络输入尺寸。参数微调需各层数据张量尺寸相同,故本研究共尝试 4 种改进方法,如图 5 中 $1 \times 3 \times 224 \times 224$ 代表批处理数量为 1,输入特征图像通道数为 3,输入特征图像长度为 224 像素、宽度为 224 像素;均可保证后续网络输入仍为 $1 \times 96 \times 112 \times 112$,与后置层张量尺寸保持一致,可在后置层上进行参数微调。

2 结果

2.1 超参数 γ 的选择 采用 Ubuntu14.04、NVIDIA GeForce GTX 1080 显卡、Intel Xeon E5-1620 v4 3.5 GHz 处理器。当超参数 γ 取不同数值时,Focal Loss 中损失函数对分类概率的函数曲线会发生变化(图 3)。 $\gamma=1, 2, 3, 4, 5$ 时,模型对肺结节良恶性进行分类的准确率分别为 89.41%、89.93%、88.64%、89.57%、89.33%,其中 $\gamma=2$ 时模型对肺结节良恶性分类的准确率最高。

2.2 选择网络结构 采用 DenseNet 原始结构及 4 种

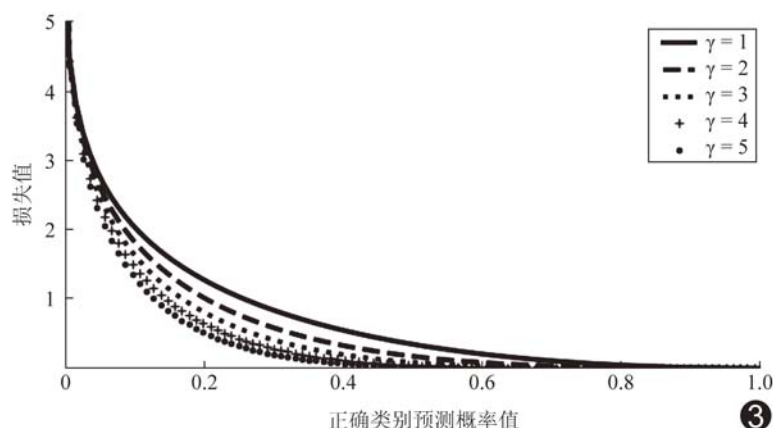


图 3 Focal Loss 损失函数图

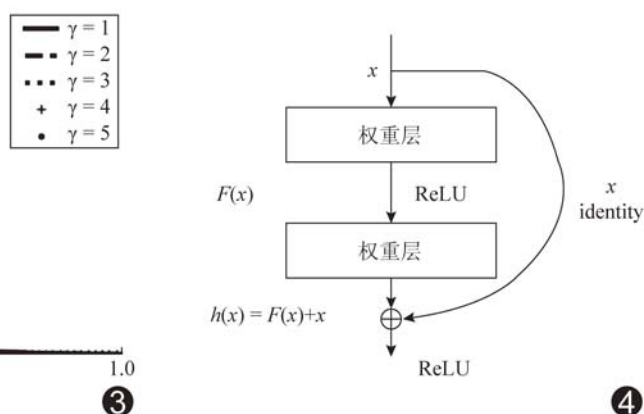


图 4 残差结构示意图 (ReLU:线性整流函数;identity:恒等变换)

改进网络结构以相同数据进行训练,未改进版本、改进版本 1、2、3、4 对应模型分类肺结节良恶性的准确率分别为 87.76%、88.52%、89.34%、88.94%、89.93%,其中第 4 种改进版本的准确率最高(图 5)。

2.3 MIP 实验 观察 5 种 MIP 方案对肺结节进行良恶性分类准确性的影响,包括 A 方案(不采用 MIP)、B 方案(采用 MIP,直接将 3 个面放缩至 112×112 大小进行拼接)、C 方案(采用 MIP,用图像均值作为填充值,填充为正方形后放缩至 112×112,之后进行拼接)、D 方案(采用 MIP,用 0 像素值作为填充值,填充后放缩至 112×112,之后进行拼接)、E 方案(采用 MIP,用 4096 像素值作为填充值,填充后放缩至 112×112,之后进行拼接),见图 6。由于层厚及扫描层数的关系,肺结节在不同方向的投影尺度相异。为后续输入神经网络的便利性,从 CT 中截取肺结节立方体时,保持肺结节的横断面为正方形;而为进行三通道拼接,需对肺结节进行矢状位与冠状位缩放。方案 A、B、C、D、E 对肺结节良恶性进行分类的准确率分别为 89.19%、89.93%、89.54%、88.63%、89.32%,即采用直接将 3 个面放缩至 112×112 大小进行拼接的策略时,准确率最高。

2.4 模型诊断肺结节良恶性的效能

2.4.1 同一方法不同变量对肺结节良恶性进行分类的准确率 采用 DenseNet 方法,以不同变量分别进行 8 组测试实验,观察模型对肺结节良恶性进行分类的准确率,每组实验均训练 50 轮。其中采用 MIP、改变输入尺寸及 Focal Loss 训练网络的准确率较高(表 1、图 7)。

表 1 不同变量下肺结节良恶性分类结果

组别	中间密度 投影	输入尺寸 (像素)	是否使用 Focal Loss 损失函数	准确率 (%)	训练时间 (h)
1	否	224×224	否	85.87	21
2	否	224×224	是	86.31	21
3	否	112×112	否	88.43	19
4	否	112×112	是	89.19	19
5	是	224×224	否	87.09	21
6	是	224×224	是	87.76	21
7	是	112×112	否	88.79	19
8	是	112×112	是	89.93	19

2.4.2 不同方法对肺结节进行良恶性分类的准确率

本课题组提出的网络结构具有残差结构,可构建较深的网络,且不发生梯度消失,较好地融合了浅层与深层的图像特征,其对肺结节良恶性分类的准确率、敏感度、特异度及曲线下面积(area under roc curve, AUC)均较高(表 2)。对改进前、改进后 DenseNet 分别进行 10 组测试实验,对获得的 2 组测试集分类准确率进行假设检验分析,改进后 DenseNet 对肺结节良恶性分类的准确率高出改进前 DenseNet(P=0.017,图 7)。

表 2 不同方法诊断恶性肺结节的效能

方法	网络 层数	输入尺寸 (像素)	准确率 (%)	敏感度 (%)	特异度 (%)	AUC
Song Q 等 ^[6] 的方法	4	28×28	84.15	83.96	84.32	0.916
Nibali A 等 ^[11] 的方法	18	64×64	89.90	91.07	88.64	0.946
DenseNet	161	224×224	85.87	85.03	86.90	0.868
改进后 DenseNet	161	112×112	89.93	91.14	89.07	0.947

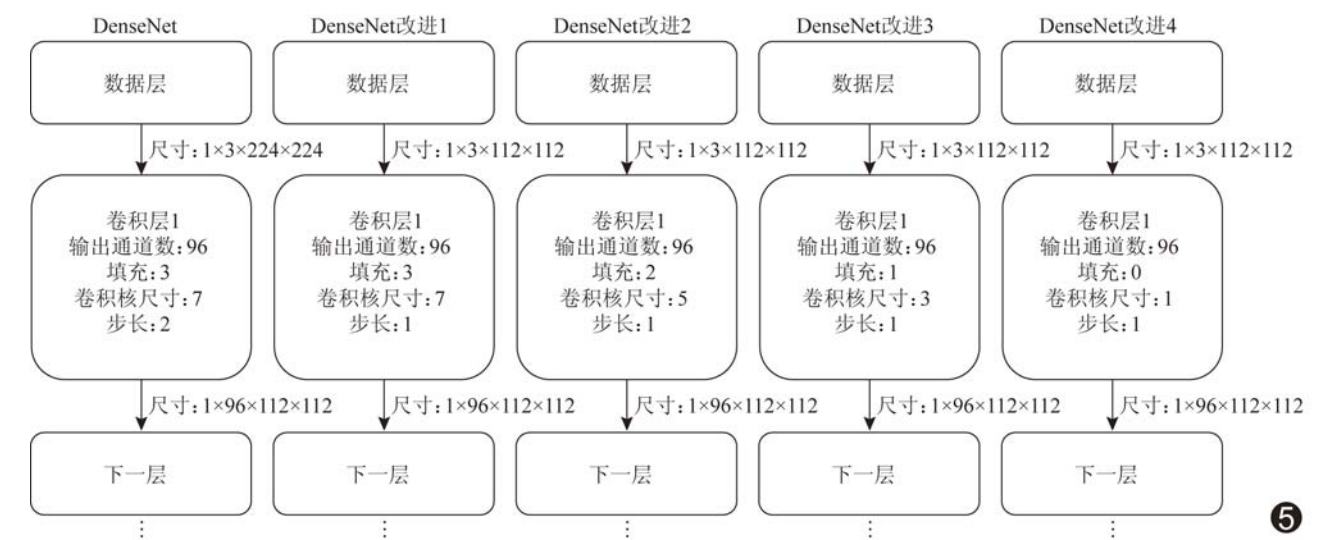


图 5 DenseNet 及其 4 种改进版本示意图

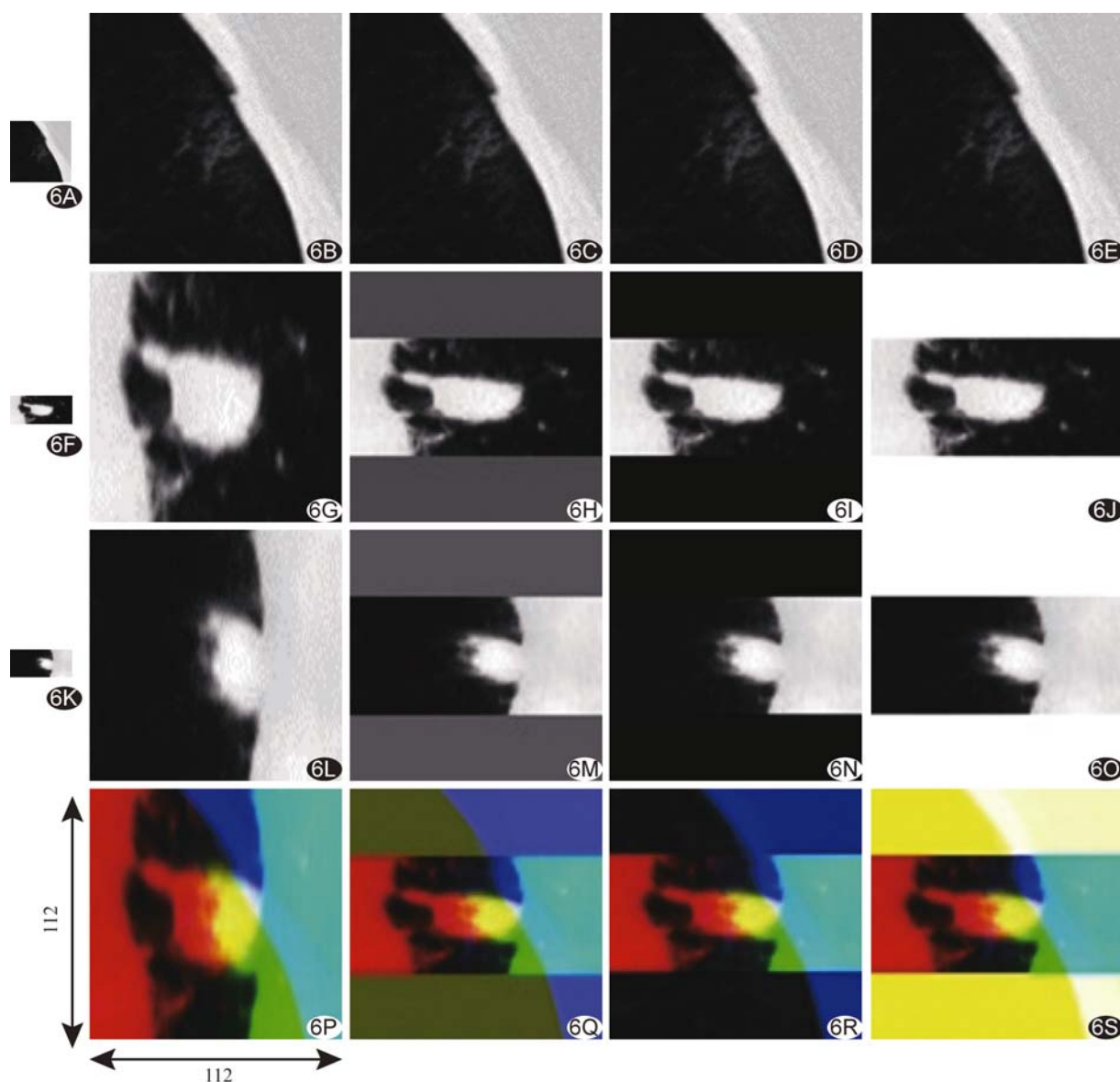


图 6 不同方案 MIP 示意图 A~E. 分别为采用 A、B、C、D、E 方案的轴位图；F~J. 分别为采用 A、B、C、D、E 方案的矢状位图；K~O. 分别为采用 A、B、C、D、E 方案的冠状位图；P~S. 采用 B、C、D、E 方案的拼接图

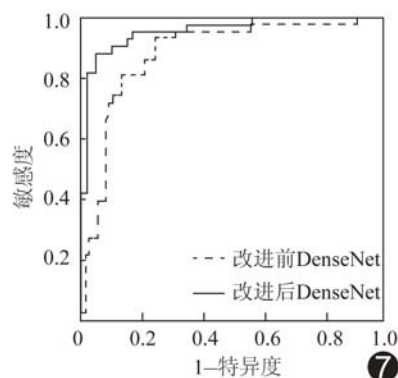


图 7 改进前与改进后的 DenseNet 分类 ROC 曲线对比

3 讨论

现有的深度神经网络方法虽能自动提取肺结节图像特征,但多采用浅层网络(小于 50 层)或既有网络对肺结节进行良恶性分类,未观察深层网络肺结节的良恶性,也未对肺结节良恶性分类问题进行适应性改进。

本课题组以 MIP 进行图像处理,可更全面地观察肺结节的三维信息;并采用旋转、滤波等图像处理方法对肺结节图像进行适当的数据扩充,使神经网络能获得更好的训练。本课题组还对密集网络的结构进行了适应性改进,使模型具备深层网络(161 层),融合浅层与深层特征学习肺结节分类,输入尺寸更适应肺结节图像,并将图像检测领域的 Focal Loss 函数添加至肺

结节良恶性分类中,使网络可针对性地学习难以分辨的肺结节,提升了模型对肺结节进行良恶性分类的准确率。今后可构建三维卷积神经网络对肺结节进行良恶性分类,而对于其他影像学设备的图像,如 PET 图像,也可考虑将之输入网络进行辅助学习。

[参考文献]

- [1] Schneider W, Bortfeld T, Schlegel W. Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions. *Phys Med Biol*, 2000, 45(2):459-478.
- [2] Uchiyama Y, Katsuragawa S, Abe H, et al. Quantitative computerized analysis of diffuse lung disease in high-resolution computed tomography. *Med Phys*, 2003, 30(9):2440-2454.
- [3] Farag A, Ali A, Graham J, et al. Evaluation of geometric feature descriptors for detection and classification of lung nodules in low dose CT scans of the chest. *Conf Proc IEEE Eng Med Biol Soc*, 2011, 48(1):169-172.
- [4] Sørensen L, Shaker SB, De Bruijne M. Quantitative analysis of pulmonary emphysema using local binary patterns. *IEEE Trans Med Imaging*, 2010, 29(2):559-569.
- [5] Song Y, Cai WD, Zhou Y, et al. Feature-based image patch approximation for lung tissue classification. *IEEE Trans Med Imaging*, 2013, 32(4):797-808.
- [6] Song Q, Zhao L, Luo X, et al. Using deep learning for classification of lung nodules on computed tomography images. *J Healthc Eng*, 2017(1):8314740.
- [7] Sun WQ, Zheng B, Qian W. Computer aided lung cancer diagnosis with deep learning algorithms//SPIE. *Medical Imaging: Computer-aided diagnosis*. San Diego: SPIE, 2016:9785
- [8] LeCun Y, Boser B, Denker JS, et al. Handwritten digit recognition with a back-propagation network// Touretzky DS. *Advances in neural information processing systems 2*. San Francisco: Morgan Kaufmann Publishers Inc, 1990:396-404.
- [9] Hussein S, Gillies R, Cao KL, et al. TumorNet: Lung nodule characterization using multi-view Convolutional Neural Network With Gaussian Process. *Conf Proc IEEE Eng Med Biol Soc*, 2017:1007-1010.
- [10] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM*, 2017, 60(6):84-90.
- [11] Nibali A, He Z, Wollersheim D. Pulmonary nodule classification with deep residual networks. *Int J Comput Assist Radiol Surg*, 2017, 12(10):1799-1808
- [12] He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition//IEEE. *The IEEE Conference on Computer Vision and Pattern Recognition (CPVR)*. Las Vegas: IEEE, 2016:770-778.
- [13] Huang G, Liu Z, Weinberger KQ, et al. Densely connected convolutional networks//IEEE. *The IEEE Conference on Computer Vision and Pattern Recognition (CPVR)*. Las Vegas: IEEE, 2017:4700-4708.
- [14] Lin T, Goyal P, Girshick R, et al. Focal loss for dense object detection//IEEE. *The IEEE Conference on Computer Vision and Pattern Recognition (CPVR)*. Las Vegas: IEEE, 2017: 2980-2988.
- [15] Naidich DP, Bankier AA, Macmahon H, et al. Recommendations for the management of subsolid pulmonary nodules detected at CT: A statement from the Fleischner Society. *Radiology*, 2013, 266(1):304-317.
- [16] Manos D, Seely JM, Taylor J, et al. The lung reporting and data system (LU-RADS): A proposal for computed tomography screening. *Can Assoc Radiol J*, 2014, 65(2):121-134.
- [17] Deng J, Dong W, Socher R, et al. ImageNet: A Large-Scale hierarchical image database// IEEE. *The IEEE Conference on Computer Vision and Pattern Recognition, VOLS 1-4*. Las Vegas: IEEE, 2009:248-255.

关键词

关键词又称主题词,是位于摘要之后,在论文中起关键作用的、最能说明问题的、代表论文特征的名词或词组。它通常来自于题目,也可以从论文中挑选。一般每篇论文要求 2~5 个关键词。每个关键词都可以作为检索论文的信息,若选择不当,会影响他人的检索效果。医学上现在主要使用美国《医学索引》(Index Medicus)的医学主题词表(Medical Subject Headings, MeSH)最新版作为规范,亦可参考中国医学科学院情报研究所翻译地英汉对照《医学主题词注释字顺表》。非主题词表的关键词为自由词,只有必要时,才可排列于最后。有些新词也可选用几个直接相关的主题词进行搭配。