

基于三维卷积神经网络的肺结节识别研究

冯雨, 易本顺*, 吴晨玥, 章云港

武汉大学电子信息学院, 湖北 武汉 430072

摘要 针对传统计算机辅助检测系统中肺结节检测存在大量假阳性的问题, 提出一种基于三维卷积神经网络的肺结节识别方法。首先, 将传统二维卷积神经网络扩展为三维卷积神经网络, 充分挖掘肺结节的三维特征, 增强特征的表达能力; 其次, 将密集连接网络与 SENet 相结合, 在加强特征传递和复用的同时, 通过特征重标定自适应学习特征权重; 另外, 引入 focal loss 作为网络的分类损失函数, 提高对难样本的学习。在 LUNA16 数据集上的实验结果表明: 与当前的主流深度学习算法相比, 所提网络模型在平均每组 CT 图像中假阳个数为 1 和 4 时的检出率达到了 0.911 和 0.934, CPM 得分为 0.891, 优于大部分主流算法。

关键词 图像处理; 计算机辅助检测; 肺结节; 三维卷积神经网络; 深度学习

中图分类号 TP391

文献标识码 A

doi: 10.3788/AOS201939.0615006

Pulmonary Nodule Recognition Based on Three-Dimensional Convolution Neural Network

Feng Yu, Yi Benshun*, Wu Chenyue, Zhang Yungang

Electronic Information School, Wuhan University, Wuhan, Hubei 430072, China

Abstract Herein, a method of pulmonary nodule recognition based on a three-dimensional (3D) convolution neural network (CNN) is proposed to overcome the problem of false positives in pulmonary nodule detection by traditional computer aided detection systems. First, a traditional two-dimensional CNN is extended to 3D CNN to fully extract the 3D features of pulmonary nodules and enhance the expressive ability of the features. Second, dense connection network and SENet are combined to enhance feature transfer and reuse, and feature weights are adaptively learned by feature recalibration. In addition, focal loss is introduced as the network classification loss to improve the learning of hard examples. The experimental results on the LUNA16 dataset demonstrate that the proposed network model achieves sensitivities of 0.911 and 0.934 at one and four false positives per scan, respectively, and the competition performance metric is up to 0.891, which is better than that of existing mainstream methods.

Key words image processing; computer aided detection; pulmonary nodule; three-dimensional convolution neural network; deep learning

OCIS codes 100.3008; 100.5010; 150.1135

1 引 言

肺癌是当今世界上发病率和死亡率最高的恶性肿瘤之一, 对人类健康危害极大, 如果能在早期发现就能有效提高患者的术后生存率^[1]。肺癌的早期形态表现为肺结节, 在肺部体现为类圆形的病灶, 其大小和形态各异, 且不易与肺部其他组织区分开来。一般情况下, 医生通过肉眼来识别结节, 具有很大难度, 大部分肺癌患者在确诊时都已到了晚期^[2]。因此, 肺结节的检测对早期肺癌的诊断治疗具有重大意义。

传统肺结节检测都是通过放射科医师观察患者肺部的电子计算机断层扫描(CT)图像来识别患者肺部是否存在结节, 而一次 CT 扫描往往会有数百张 CT 切片, 这大大增加了医师的工作量, 进而影响了诊断的准确性。近年来, 计算机辅助检测^[3-4](CAD)已逐步成为放射科医师诊断的“第三只眼睛”, 为医师提供诊断的“第二意见”, 从而提高了诊断的准确性。计算机辅助肺结节检测系统主要分为两部分: 疑似结节提取和假阳性去除。其中: 前者的主要目的是提高肺结节的检出率, 即尽可能地提高系统的敏感性, 但同时也会带来了大量的假阳性; 后

收稿日期: 2019-01-22; 修回日期: 2019-02-23; 录用日期: 2019-03-11

* E-mail: yibs@whu.edu.cn

者则是对候选结节进行识别,以剔除假阳性结节,保留真正的肺结节,以用于后续病理分析。因此,研究肺结节的假阳性抑制具有重要意义。

近年来,深度学习算法以其强大的特征提取能力在图像识别等方面取得了巨大进步,越来越多的学者将其用于计算机辅助医疗。在肺结节识别方面,文献[5]提出了一种多视图卷积神经网络(CNN),从肺结节的立体特性出发,融合肺结节的二维(2D)切面信息来剔除假阳性结节,但该方法本质上是对2D信息的融合,无法充分提取肺结节的三维(3D)空间信息。文献[6]提出了一种多尺度输入的三维卷积神经网络(3DCNN),在充分利用肺结节空间信息的同时提出多尺度输入,以适应肺结节自身的尺度变化,但多尺度输入所带来的多个平行支路大大增加了网络的计算量。文献[7]提出了一种深层三维残差网络,并将其用于肺结节特征的提取,在下采样过程中通过添加中心裁剪操作来提取结节的多尺度特征,进一步提升网络性能,但该深层三维残差网络拥有千万级的参数量,训练难度较大。文献[8]考虑到医学图像训练样本集分布不平衡以及难样本难以训练等问题,提出了一种在线样本选择机制,每次选择对loss影响较大的样本,使训练更加有效。

本文针对肺结节检测中假阳性过高的问题,提出了一种改进的基于3DCNN的肺结节识别算法。其特点如下:首先,针对传统二维卷积神经网络(2DCNN)对肺结节特征提取不充分、缺乏三维空间信息等问题,将3DCNN用于肺结节的特征提取;其次,在添加密集连接模块^[9]以强化特征的传播复用能力、缓解梯度消失的同时,引入SE(squeeze-and-excitation)模块^[10]对特征提取过程中的特征进行重标定,自适应学习特征权重,进一步提高识别的准确率;最后利用focal loss损失函数^[11]解决训练样本中正负样本严重不平衡以及难样本难以训练的问题。

2 所提算法

2.1 3D卷积

肺部CT图像属于断层图像,从观察角度来看属于三维立体视图,而传统的2D卷积只能提取单张切片的平面特征,从而导致肺结节的立体信息丢失,因此,本研究采用3D卷积,以保证特征信息的充分提取。由图1可以看出,与2D卷积不同的是,3D卷积输入多了一个深度维度,在CT图像中体现

为多张连续切片,因此其卷积核也会相应增加一个维度^[12-13]。在图1(a)中, H 、 W 分别为传统二维图像数据的高和宽, k 为卷积核的大小;图1(b)中的 H 、 W 、 D 分别表示三维数据的高、宽和深,其中卷积核的大小由二维中的 $k \times k$ 变为 $k \times k \times d$ 。值得注意的是,3D卷积虽然可以有效提取三维数据特征,但由于维度增加,其参数与计算量也会随之增加,因此在一定程度上会限制网络的深度^[14]。

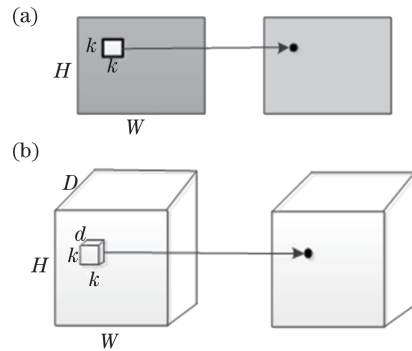


图1 二维与三维卷积。(a)二维卷积;(b)三维卷积
Fig. 1 2D and 3D convolutions. (a) 2D convolution;
(b) 3D convolution

2.2 SE-Dense 模块

DenseNet通过跨层密集连接操作互相连接模块中的所有层,使得模块中每一层的输入都与其前面所有层的输出特征相关联,增强了网络中各层信息之间的流通^[9]。其优点在于:通过这种密集连接,网络中任何两层都有直接连接,可在一定程度上减轻训练过程中的梯度消散;另外,密集连接加强了特征的传播复用,使得通过少量的卷积核即可以得到大量的特征。设网络第 l 层的输出为 \mathbf{x}_l ,则一个密集连接模块第 l 层的输出可表示为

$$\mathbf{x}_l = H_l[(\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{l-1})], \quad (1)$$

式中: $H_l[\cdot]$ 操作为第 l 层的非线性映射; $(\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{l-1})$ 为将前面的特征图进行合并操作。

文献[10]提出的SE模块结构如图2所示,主要包含三个部分:首先,利用全局平均池化(global pooling)操作将输入各特征通道压缩成一个具有全局响应的实数,输出维度与输入特征图的数目相同,这一步骤被称为squeeze操作;其次,利用全连接层(FC)与Sigmoid激活函数组成的门控机制来学习各个特征通道的权重,实现excitation操作;最后,将得到的特征权重值加权到输入的特征通道上,从而实现特征图的重标定操作,完成scale操作。与传统的从深度和宽度角度提升网络性能模块不同的是,SE模块通过自适应学习的方式来自动获取每个

特征通道的重要程度,然后依照这个重要程度去提升有用的特征,并抑制对当前任务用处不大的特征,可以在不显著增加参数数量的同时有效提升模型的性能。

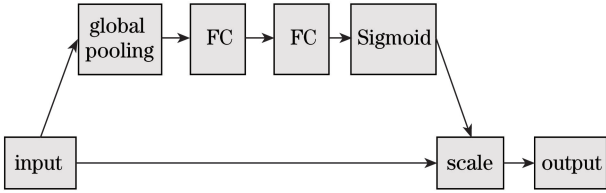


图 2 SE 模块结构

Fig. 2 Structure of SE block

将传统二维密集连接网络扩展到三维时,由于

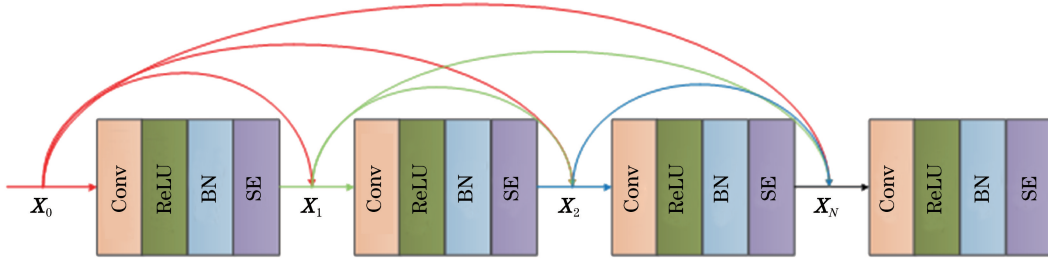


图 3 SE-Dense 模块示意图

Fig. 3 Schematic of SE-Dense block

2.3 focal loss 函数

当样本类别分布严重不均衡时,常常会导致模型无法学习少例样本。在假阳性去除研究中,所使用的数据集中用于训练候选结节的正例样本(真结节)与负例样本(假阳性结节)的数量比高达 1:420。在训练过程中,正例样本往往被淹没在负例样本里,导致网络得不到有效训练。此外,由于肺结节自身形态、尺度等的差异以及肺实质内部结构复杂,往往会产生许多难样本,而这些难样本往往是模型需要侧重学习的,因此,为了有效解决这些问题,本研究将文献[11]提出的 focal loss 函数作为网络训练时的损失函数,即

$$L_F(p_t) = -\alpha_t(1 - p_t)^\gamma \log_2(p_t), \quad (2)$$

$$p_t = \begin{cases} p, & R_{lbl} = 1 \\ 1 - p, & R_{lbl} = 0 \end{cases}, \quad (3)$$

$$\alpha_t = \begin{cases} \alpha, & R_{lbl} = 1 \\ 1 - \alpha, & R_{lbl} = 0 \end{cases}, \quad (4)$$

式中: p 为样本的预测概率; R_{lbl} 为样本标签; α 和 γ 为两个超参数。 α 为用以解决样本类别不平衡的一个参数。 γ 用来提高难样本的 loss 权重,从而使网络倾向于难样本的学习,提高网络的判别能力。通过比较后,在后续实验中将 α 设置为 0.6,将 γ 设置为 2,用以解决样本类别不均以及难样本难以学习

参数和计算量增加,往往会限制模型的深度,因此,本研究将密集连接模块与 SE 模块相结合,即 SE-Dense 模块,从特征通道的角度出发来提升网络性能。其结构如图 3 所示,其中, $X_n (n=0,1,\dots,N)$ 表示密集连接中每一层的输入,Conv 表示 3D 卷积层,ReLU 表示修正线性单元,BN 表示批量归一化,之后将 Conv-ReLU-BN-SE 视为一层。从图 3 中可以看出,SE-Dense 模块既可以通过密集连接实现原始输入信息的无损传输,又可以通过特征重标定来自动学习全局信息,从而获得每个特征通道的重要程度,然后根据重要程度来增强有益特征并抑制无用特征,最终实现特征通道的自适应校准。

的问题。

2.4 肺结节识别网络模型

本研究网络结构如图 4 所示,对于输入三维数据,其后有 4 个 SE-Dense 结构进行特征的提取和重标定,其层数分别为 4、6、8、10,每个结构中输出的特征图(即增长率)分别为 12、12、12、24,各层卷积核大小为 $3 \times 3 \times 3$,在每个 SE-Dense 结构后都添加一个核大小为 $2 \times 2 \times 2$ 、步长为 2 的最大池化层(max pooling),用来将特征图的三维长度缩小至原来的 1/2。另外,为提高计算效率,在 SE-Dense1、SE-Dense2、SE-Dense3 中都添加 Bottleneck 层,即

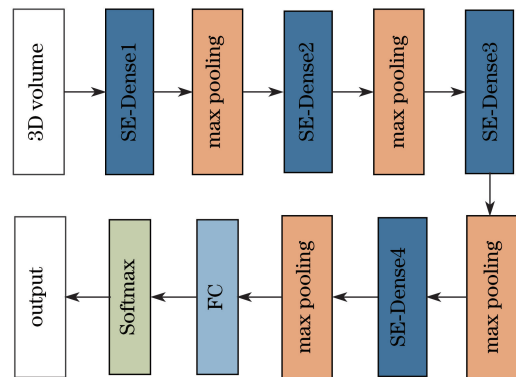


图 4 网络结构模型

Fig. 4 Model of network structure

在其输出后添加 $1 \times 1 \times 1$ 的卷积层对输出的特征数量进行压缩,压缩后的通道数都设置为 12。最后,通过 Softmax 分类器输出分类结果。

3 实验结果分析

所提算法的实验环境基于 Windows 10 操作系统,硬件环境为 Intel Core i7-6850K 与 Nvidia GeForce GTX 1080Ti (11 GB 内存),编程开发环境为 CUDA-Toolkit8.0,编程语言为 Python 3.6,深度学习框架为 keras。

3.1 实验数据集与预处理

实验数据来源于公开竞赛数据集 LUNA16^[15],该数据集来自于更大的公开数据库 LIDC-IDRI^[16]。该竞赛数据集包含了 888 张低剂量 CT 图像,共 1186 个肺结节,同时给出了肺结节中心位置和直径的标注。将该竞赛数据集用于候选结节检测和疑似结节分类,在疑似结节分类中,LUNA16 给出了候选结节的病人序列号、中心坐标以及其对应的标签信息,共计 551065 个候选结节,其中真结节只有 1120 个。将 LUNA16 数据集随机分成 5 个子集进行交叉验证。评价标准采用 FROC (free-response receiver operating characteristic) 准则,该准则通过计算 CPM (competition performance metric)^[15] 来衡量算法的性能。CPM 指平均每组 CT 图像中假阳个数 (FPs per scan) 为 1/8、1/4、1/2、1、2、4、8 时的平均检出率。

在数据预处理方面,由于 CT 扫描图像可能来自于不同的设备,切片的厚度也不完全相同,所以首先将数据重采样至 $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$,保证数据的一致性。然后将所有的 CT 影像数据根据 HU (Hounsfield unit) 值进行过滤,只保留在 $(-1000 \sim$

$400 \text{ HU})$ 区间的值,以减小肺部无效组织对实验的影响。

3.2 训练

本研究以 LUNA16 提供的候选结节质心为中心裁取尺寸为 $24 \times 36 \times 36$ 的数据块模型训练数据。由于实验数据中正负样本严重失衡,直接训练很容易导致模型无法学习正样本信息,因此需要对正样本进行数据扩增,主要包括在 XYZ 各方向进行 $[-2, 2]$ 的随机平移、旋转、翻转等操作。在进行平移操作时,先对候选结节的质心做平移再裁切出数据块,而非对裁切出来的数据块做平移。另外,由于内存限制,同时为了提高训练效率,在训练阶段随机选取 50% 的负样本参与训练。

本研究模型的超参数配置如下:权重初始化采用文献[17]提出的正态分布方法,通过 Adam 算法来优化网络,训练的 batch size 取 128,epoch 设置为 40。

3.3 实验结果

为验证密集连接模块、SE 模块、focal loss 对识别性能的影响,对网络结构进行以下调整:1)使用传统 3DCNN 作为分类网络,即删除 SE-Dense 模块,去掉 focal loss 损失函数,该网络记为 Model_1;2)删去 SE-Dense 结构中 SE 模块,即去掉特征重标定操作,转为普通的三维密集连接模块结构,并采用传统的交叉熵损失函数,该网络记为 Model_2;3)只去掉 focal loss 损失函数,该网络记为 Model_3。

表 1 对比了三种网络与本研究所提网络在 LUNA16 数据集上的性能差异,可以看出:1)密集连接模块可有效提高网络的识别性能;2)通过特征重标定操作可有效提高特征的有效性,提高模型的性能;3)与传统的交叉熵损失函数相比,focal loss 损失函数的效果更好。

表 1 不同结构在 LUNA16 数据集上的肺结节识别性能对比

Table 1 Comparison of pulmonary nodule recognition performance by different network structures on LUNA16 dataset

Method	False positives per scan							CPM
	1/8	1/4	1/2	1	2	4	8	
Model_1	0.629	0.735	0.807	0.865	0.901	0.917	0.928	0.826
Model_2	0.734	0.812	0.869	0.901	0.918	0.927	0.929	0.870
Model_3	0.754	0.821	0.888	0.913	0.930	0.933	0.937	0.883
Proposed	0.807	0.843	0.877	0.911	0.925	0.934	0.939	0.891

为了验证所提算法的优越性,在 LUNA16 数据集上,将所提算法与文献[5]、文献[6]、文献[18]、文献[7]提出的主流深度学习算法对疑似结节的平均检出率进行对比,对比结果如表 2 所示。由

表 2 可见,所提算法的 CPM 得分值 (0.891) 与文献[7]的结果 (0.892) 相近,且远优于其他算法,当每组 CT 图像中假阳个数为 1 和 4 时,检出率分别达到了 0.911 和 0.934。

表 2 不同算法在 LUNA16 数据集上的肺结节识别性能
Table 2 Pulmonary nodule recognition performance by different algorithms on LUNA16

Algorithm	False positives per scan							CPM
	0.125	0.25	0.5	1	2	4	8	
Ref. [6]	0.678	0.738	0.816	0.848	0.879	0.907	0.922	0.827
Ref. [5]	0.692	0.710	0.809	0.863	0.895	0.914	0.923	0.838
Ref. [18]	0.760	0.794	0.833	0.860	0.876	0.893	0.906	0.846
Ref. [7]	0.802	0.847	0.886	0.909	0.925	0.936	0.941	0.892
Proposed	0.807	0.843	0.877	0.911	0.925	0.934	0.939	0.891

图 5 为所提算法在 LUNA16 数据集上对部分疑似结节的识别效果,图 5(a)为真结节的预测概率,图 5(b)为假阳性结节的预测概率, P 表示预测

概率。可以看出,所提算法对不同尺度、不同形态的肺结节均具有良好的识别效果,同时也可以准确识别出各种假阳性结节。

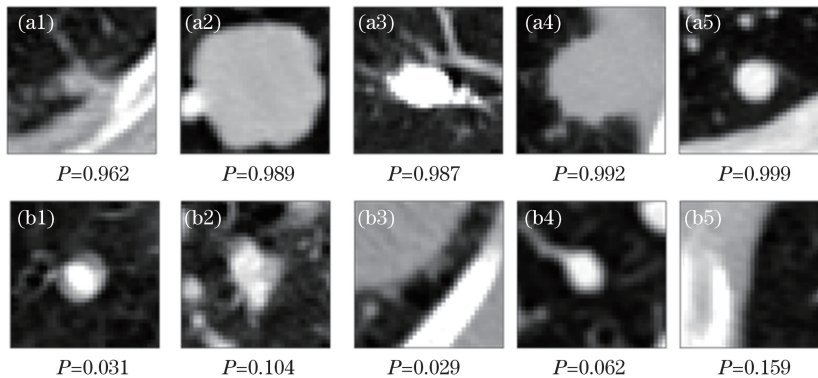


图 5 候选结节的识别结果。(a)真结节的预测概率;(b)假阳性结节的预测概率

Fig. 5 Recognition results of candidate nodules. (a) Prediction probability of true nodule;
(b) prediction probability of pseudopositive nodule

4 结 论

针对肺结节检测中存在大量假阳性、传统网络识别分类性能不足等问题,提出了一种基于 3DCNN 的肺结节识别方法。该方法首先将传统 2DCNN 扩展到 3DCNN,充分挖掘肺结节的三维特征;其次,在不显著增加模型参数和计算量的情况下,将密集连接模块与 SE 模块相结合,以提高特征的传播复用能力,增强有效特征,抑制无用特征,从而提高网络性能;最后,引入 focal loss 函数作为训练的损失函数,在减小数据集中样本分布不均带来的影响的同时,加强对难样本的学习,从而进一步提高网络的识别效果。实验结果表明,所提算法的平均检出率达到了 0.891,优于其他深度学习算法。

参 考 文 献

- [1] Sluimer I, Schilham A, Prokop M, *et al.* Computer analysis of computed tomography scans of the lung: a survey[J]. IEEE Transactions on Medical Imaging, 2006, 25(4): 385-405.
- [2] Wang C M. Key techniques for lung nodule detection and classification based on chest imaging[D]. Shenzhen: University of Chinese Academy of Sciences, 2018: 1-3.
王昌森. 基于胸部影像的肺结节检测与分类关键技术研究[D]. 深圳: 中国科学院大学, 2018: 1-3.
- [3] Messay T, Hardie R C, Rogers S K. A new computationally efficient CAD system for pulmonary nodule detection in CT imagery[J]. Medical Image Analysis, 2010, 14(3): 390-406.
- [4] Setio A A A, Jacobs C, Gelderblom J, *et al.* Automatic detection of large pulmonary solid nodules in thoracic CT images[J]. Medical Physics, 2015, 42(10): 5642-5653.
- [5] 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.
- [6] Dou Q, Chen H, Yu L Q, *et al.* Multilevel contextual

- 3-D CNNs for false positive reduction in pulmonary nodule detection[J]. *IEEE Transactions on Biomedical Engineering*, 2017, 64(7): 1558-1567.
- [7] Jin H S, Li Z Y, Tong R F, *et al.* A deep 3D residual CNN for false-positive reduction in pulmonary nodule detection[J]. *Medical Physics*, 2018, 45(5): 2097-2107.
- [8] Dou Q, Chen H, Jin Y M, *et al.* Automated pulmonary nodule detection via 3D ConvNets with online sample filtering and hybrid-loss residual learning[M]//Descoteaux M, Maier-Hein L, Franz A, *et al.* *Springer Proceedings in Physics*. Singapore: Springer, 2017: 630-638.
- [9] Huang G, Liu Z, Maaten L V D, *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, 2017: 2261-2269.
- [10] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 7132-7141.
- [11] Lin T Y, Goyal P, Girshick R, *et al.* Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 2999-3007.
- [12] Li Z Q, Zhu R F, Gao F, *et al.* Hyperspectral remote sensing image classification based on three-dimensional convolution neural network combined with conditional random field optimization[J]. *Acta Optica Sinica*, 2018, 38(8): 0828001.
李竺强, 朱瑞飞, 高放, 等. 三维卷积神经网络模型联合条件随机场优化的高光谱遥感影像分类[J]. *光学学报*, 2018, 38(8): 0828001.
- [13] Lü X Q, Wu L, Gu Y, *et al.* Detection of low dose CT pulmonary nodules based on 3D convolution neural network[J]. *Optics and Precision Engineering*, 2018, 26(5): 1211-1218.
吕晓琪, 吴凉, 谷宇, 等. 基于三维卷积神经网络的低剂量CT肺结节检测[J]. *光学精密工程*, 2018, 26(5): 1211-1218.
- [14] Miao G, Li C F. Detection of pulmonary nodules CT images combined with two-dimensional and three-dimensional convolution neural networks[J]. *Laser & Optoelectronics Progress*, 2018, 55(5): 051006.
苗光, 李朝锋. 二维和三维卷积神经网络相结合的CT图像肺结节检测方法[J]. *激光与光电子学进展*, 2018, 55(5): 051006.
- [15] Setio A A A, Traverso A, de Bel T, *et al.* Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge[J]. *Medical Image Analysis*, 2017, 42: 1-13.
- [16] Messay T, Hardie R C, Tuinstra T R. Segmentation of pulmonary nodules in computed tomography using a regression neural network approach and its application to the lung image database consortium and image database resource initiative dataset[J]. *Medical Image Analysis*, 2015, 22(1): 48-62.
- [17] He K M, Zhang X Y, Ren S Q, *et al.* Delving deep into rectifiers: surpassing human-level performance on ImageNet classification[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 1026-1034.
- [18] Sakamoto M, Nakano H, Zhao K, *et al.* Multi-stage neural networks with single-sided classifiers for false positive reduction and its evaluation using lung X-Ray CT images[M]//Battiato S, Gallo G, Schettini R, *et al.* *Image Analysis and Processing-ICIAP 2017*. Cham: Springer, 2017: 370-379.