

基于 3D FPN 的肺结节检测

张宇廷

(浙江理工大学 信息学院, 浙江 杭州 310000)

摘要:肺癌早期发现与诊断对提高肺癌患者生存率至关重要。肺癌主要是由恶性肺结节造成的,通过肺结节早期检测与诊断能够及时发现病情,显著提高肺癌存活率。随着深度学习网络在医学辅助诊断领域应用的迅速发展,现已有很多深度网络被应用于肺结节检测。通过统计检测目标肺结节半径分布情况,发现大部分肺结节半径较小。因此,结合 CT 影像数据的三维特性,提出使用 3D FPN 进行单阶段肺结节检测,能够解决肺小结节检测效果不佳的问题。在公用肺结节数据集 LUNA16 上验证了网络的有效性,CPM 值达到 0.893 2,相比其它肺结节检测网络,检测效果提高了 2%。

关键词:肺结节检测;单阶段检测;FPN;深度学习;LUNA16

DOI:10.11907/rjdk.201052

中图分类号:TP317.4

文献标识码:A

开放科学(资源服务)标识码(OSID):

文章编号:1672-7800(2020)009-0233-05



3D FPN-based Lung Nodule Detection

ZHANG Yu-ting

(School of Informatics and Electronics, Zhejiang Sci-Tech University, Hangzhou 310000, China)

Abstract: Early detection and diagnosis of lung cancer is essential to improve the survival rate of lung cancer patients. Lung cancer is mainly caused by malignant lung nodules. Through the detection and diagnosis of early lung nodules, the disease can be detected in time, and the survival rate of lung cancer is significantly improved. With the rapid development of the application of deep learning networks in the field of medical aided diagnosis, many deep networks have been proposed for the detection of lung nodules. By statistically detecting the radius distribution of the target lung nodules, it is found that most of the lung nodules have a small radius. Aiming at the problem of small detection target and the three-dimensional characteristics of CT image data, this paper proposes to use 3D FPN for single-stage pulmonary nodule detection, which can solve the problem of poor detection of small pulmonary nodules. The effectiveness of the network was verified on the common lung nodule data set LUNA16. The CPM value reached 0.8932. Compared with other lung nodule detection networks, the detection effect was improved by 2%.

Key Words: pulmonary nodule detection; single-stage detection; FPN; deep learning; LUNA16

0 引言

肺癌是发病率和死亡率最高的癌症之一,严重威胁着人类生命健康。与其它癌症一样,肺癌早期症状不太明显,只会出现咳嗽、胸闷等症状。然而一旦进入癌症晚期,存活率则大大降低,所以最好的方法是早诊断、早治疗。

现代医学都是使用胸腔 CT 影像作为成像工具,依据 CT 影像数据判断是否患有癌症^[1]。肺癌主要是由肺结节导致的,医生主要通过检查患者肺结节形状和大小判断其是否患有肺癌。然而,一个 CT 影像数据包含几百张 CT 切片,医生需要很长时间才能判断患者患有癌症的概率。对

于同一张 CT 切片,不同医生给出的诊断结果可能也不一样,因为医生容易疲劳,而且不同人经验也不同,容易出现误诊情况^[2]。

随着深度学习的兴起,计算机辅助诊断肺结节也开始采用深度学习技术,并取得了显著成效。将计算机辅助诊断肺结节系统应用于肺癌早期辅助诊断中,不仅可以挽救无数患者的生命,而且对于缓解医疗资源短缺问题及医患矛盾也有重大意义。随着肺结节检测网络的不断改进,部分辅助诊断系统甚至可以比一些医生做得更加出色,而且诊断所需时间更少。

本文通过分析检测目标的半径分布情况,发现很多肺小结节,因此提出使用 3D FPN 网络结构进行肺结节检

收稿日期:2020-01-12

作者简介:张宇廷(1994-),男,浙江理工大学信息学院硕士研究生,研究方向为计算机视觉。

©1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

测。该网络能够在 4 种不同大小的特征图上进行检测,解决了针对肺小结节识别率不高的问题。实验结果表明,本文提出的网络能够快速、有效地检测出肺结节。

1 相关研究现状

1.1 目标检测

从本质而言,肺结节检测任务也是属于目标检测算法。随着深度学习的发展,研究者们已提出很多优秀的目标检测方法,这些方法可大致分为两类:①两阶段目标检测;②单阶段目标检测。很多优秀的算法包含两个阶段(如 Faster-RCNN^[3]):第一阶段提出 proposal(是否包含对象)信息,第二阶段预测候选框位置,并进行物体类别识别。但目前的检测方法趋向于使用单阶段方法进行检测,在该阶段中,边界框和类概率同时被预测,如 YOLO^[4]、SSD^[5]和 FPN^[6]等。总体而言,单阶段方法速度较快,两阶段方法更加准确。将前沿的 2D 物体检测方法扩展到 3D 物体检测任务(如视频中的动作检测和体积检测)是有局限的,主要受到主流 GPU 内存限制,一般都不是直接将原图输入到网络中,而是随机将 3D 原图剪切成相同大小,再将这相同大小的 3D 图片输入到网络中。

1.2 肺结节检测

在传统肺结节检测方法中,主要依赖于手工设计特征提取及传统图像处理方法。Ginneken 等^[7]对 6 种传统算法进行全面回顾,并在包含 55 次扫描的 ANODE09 数据集上对其进行评估。然而,随着深度学习理论的快速发展,越来越多网络模型开始应用于肺结节检测中,相比传统候选结节检测算法,其摆脱了基于人工提取特征方法的局限性。目前,基于深度学习的肺结节检测效果已得到了大幅提高。

目前大多数肺结节检测都是采用两阶段方法进行的,具体过程包括:①候选框提取;②假阳性降低。最初使用 2D 卷积神经网络进行肺结节检测和假阳性降低,然后逐步拓展到 3D 卷积神经网络。如 Setio 等^[8]使用的检测网络由 3 个 2D 候选框检测器组成,分别针对固体、亚固体和大结节设计了肺结节检测器;Ding 等^[9]采用 2D 候选框提取网络(Faster RCNN^[10]中的 RPN 网络)对 CT 影像中的每个切片进行候选框提取,接着使用三维深度卷积神经网络减少假阳性样本数量;Dou 等^[10]第一次提出使用 3D 卷积神经网络进行肺结节检测,具体使用的网络结构是 3D FCN^[11],然后采用 3D 深度卷积神经网络进行假阳性降低,该方法在 LUNA16 数据集^[12]上取得的平均 FROC 分数为 83.9%。由于 3D 网络结构能够捕获更多 CT 影像数据特征,同时识别率更高,所以之后方法大多采用基于 3D 的网络结构;Zhu 等^[13]利用 3D DPN 网络进行肺结节检测,并使用含有 3D 双路径的 GBM 网络进行假阳性降低,该方法在 LUNA16 数据集上取得的平均 FROC 分数为 84.2%。从以上研究成果可以看出,越来越多学者开始将 2D 网络结构拓展成 3D 网络结构进行肺结节检测,并取得了较为理想的效果。

2 肺结节检测方法

本文提出的肺结节检测方法包括:肺实质提取和肺结节检测。其中,肺实质提取主要将肺部提取出来,同时去除一些无用的躯干信息,因为需要检测的肺结节全都在肺部区域。一个好的肺实质提取方法可以加快模型收敛速度、提高模型识别率,本文采用阈值法进行肺实质分割;肺结节检测主要任务是对目标区域进行回归与分类。本文提出的 3D FPN 网络结构也是单阶段肺结节检测,能够针对不同尺度的肺结节进行检测,在 LUNA16 肺结节数据集上表现优异。

2.1 数据处理

本文实验采用肺结节数据集 LUNA16,该数据集包括 888 个 CT 扫描,每一个扫描都由 4 位经验丰富的放射科医生进行注释。在检测候选结节前,肺实质分割是十分必要的,因为肺外器官和组织(如胸骨)可能会对检测造成极其不利的影响。为了减少相关影响,本文采用阈值分割与形态学相结合的方法对原始数据进行肺实质分割,具体如图 1 所示。首先,利用阈值将原始 CT 切片图像(a)变为二值化图像(b);接着,清除二值化图像(b)的边界信息,得到图像(c);然后,在肺实质区域图像中保留最大的两块连通区域,得到图像(d);最后,得到左肺和右肺的掩膜图像(e),与原始 CT 切片结合,得到肺实质分割图像(f)。

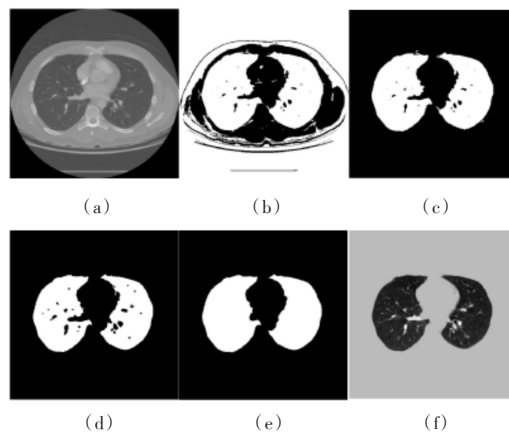


图 1 肺结节数据集处理过程

2.2 3D FPN 网络结构

一个网络的提出往往依赖于检测目标分布情况。如图 2 所示,本文统计了肺结节半径分布情况,发现检测目标中大部分是较小的肺结节,半径集中在 3~10mm 之间。如果使用常规方法,只是在低分辨率特征图上进行检测,会出现小肺结节信息过少的情况,从而导致漏检。本文提出的 3D FPN 肺结节检测网络基于 2D FPN 网络结构。2D FPN 在执行检测任务时具有两大优点:①能够检测出小目标;②能够提高网络检测的准确性。考虑到肺结节数据本身就是三维的,3D 网络结构可以更好地挖掘肺结节数据的空间信息。基于此,本文提出使用 3D FPN 网络进行肺结节检测。

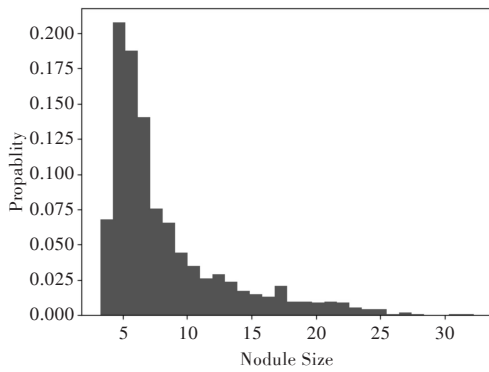


图2 LUNA16数据集中肺结节半径分布

本文提出的3D FPN网络结构如图3所示。网络输入为 $128 \times 128 \times 128 \times 1$ (深度 \times 高度 \times 宽度 \times 通道数),首先在下采样过程中,通过含有3D Residual Block的4个Blocks进行特征提取,在每个Blocks后通过最大池化降低特征图大小;然后进入上采样过程,通过含有3D Residual Block的3个Blocks进行特征提取,在每个Blocks前进行转置卷积

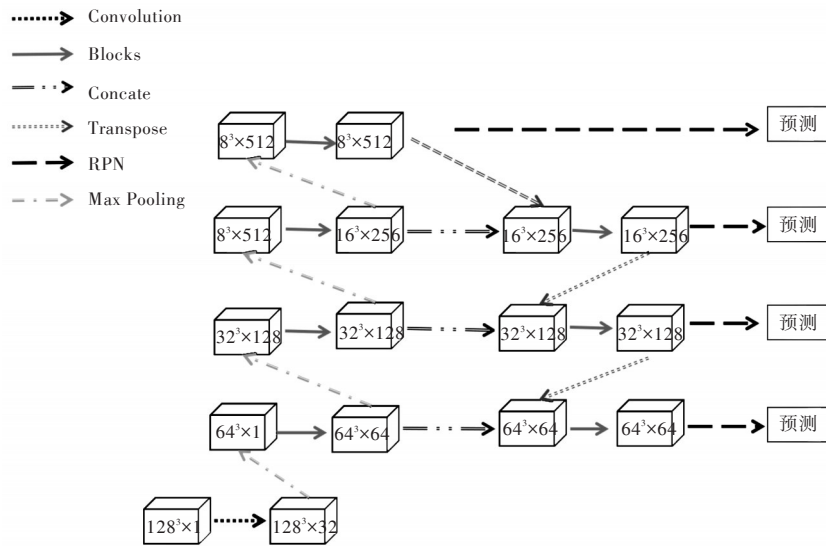


图3 3D FPN网络结构

2.3 3D Residual Block

基于ResNet^[15]中的基础结构Bottleneck block,提出3D Residual block,具体网络结构如图4所示。首先,特征图 $a^3 \times c$ 通过大小为 $1 \times 1 \times 1$ 的卷积核、BN(Batch Normalization)和激活函数Leaky ReLU^[16]改变特征通道数量,得到特征图I;然后,特征图I通过大小为 $3 \times 3 \times 3$ 的卷积核、BN和激活函数Leaky ReLU进行特征提取,得到特征图II;接下来,特征图II通过大小为 $1 \times 1 \times 1$ 的卷积核和BN,得到特征图III($a^3 \times c1$);最后,将特征图III与输入的特征图在通道上进行相加操作,再通过激活函数Leaky ReLU得到最后结果。需要注意的是,在进行相加操作前,需要判断输入和输出通道大小是否相同,如图4中的判断框所示,如果不相同,需要通过一个卷积操作和BN将通道大小变为一致。

2.4 损失函数设计

在基于锚框的检测任务中,锚框损失函数包括两部

(Tranpose)操作增加特征图大小,在每个Blocks后将下采样对应大小的特征信息融合进来;最后,利用3D RPN思想,分别在4个维度上($8^3 \times 512, 16^3 \times 256, 32^3 \times 128, 64^3 \times 64$)进行肺结节检测,具体需要在4个维度中每个像素上设置3个锚框,网络输出维度是 $8^3 \times 3 \times 5, 16^3 \times 3 \times 5, 32^3 \times 3 \times 5, 64^3 \times 3 \times 5$,其中3代表设置anchor数量,5代表每个anchor的信息。

每个锚框需要5个值(a_x, a_y, a_z, a_d, c)表示,坐标值(a_x, a_y, a_z)表示一个锚框的中心坐标, a_d 表示锚框半径, c 表示是肺结节的可能性。锚框半径是根据经验设置的,具有一定的人为认知局限性。本文采用K-means在数据集中拟合更加合理的半径,将计算得到的半径值取整为 $\{5, 7, 10, 12, 15, 20, 22, 25\}$ 。在大小为 8^3 的特征图中设置锚框半径为 $\{20, 22, 25\}$,在大小为 16^3 的特征图中设置锚框半径为 $\{12, 15, 20\}$,在大小为 32^3 的特征图中设置锚框半径为 $\{7, 10, 12\}$,在大小为 64^3 的特征图中设置锚框半径为 $\{5, 7, 10\}$ 。

分:分类任务(L_{cls})和回归任务(L_{reg})。对于第 i 个anchor的LOSS函数定义如下:

$$L(p_i, p_i^*) = L_{cls}(p_i, p_i^*) + p_i^* L_{reg}(t_i, t_i^*) \quad (1)$$

其中, p_i 代表锚框 i 预测为肺结节的概率, p_i^* 代表锚框 i 的真实值。如果锚框 i 与真实值的边界框重叠,且交集(IoU)大于0.5,则将其视为正锚框(P_i^*);如果锚框与真实值边界框的IoU均小于0.02,则将其视为负锚框(P_i^*)。

t_i 代表锚框 i 预测的相对坐标信息, t_i^* 代表真实结节的相对坐标信息,具体使用($t_x^*, t_y^*, t_z^*, t_d^*$)表示 t_i^* ,定义如公式(2)所示。

$$t_i^* = (t_x^*, t_y^*, t_z^*, t_d^*) = \left(\frac{b_x - a_x}{a_d}, \frac{b_y - a_y}{a_d}, \frac{b_z - a_z}{a_d}, \log\left(\frac{b_d}{a_d}\right) \right) \quad (2)$$

其中, (b_x, b_y, b_z, b_d)是原始空间中真实肺结节的坐标和直径, (a_x, a_y, a_z, a_d)表示锚框 i 的坐标和半径。

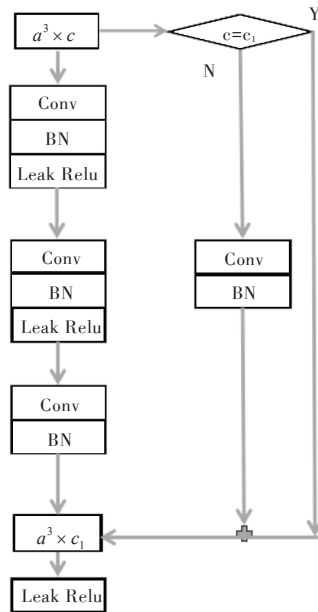


图 4 3D Residual block 网络结构

由于负锚框数量远大于正锚框,大量负样本将会主导梯度下降方向。考虑到正负样本的不平衡性,本文将分类损失函数由交叉熵损失函数变为 Focal LOSS 函数^[17],以更好地均衡正负样本对模型的重要程度,防止过度偏向负样本,造成网络分类效果变差。Focal LOSS 具体定义如公式(3)所示,其中 γ 为关注度参数(Focusing Parameter),当 $\gamma \geq 0$ 时,网络模型会关注那些容易错分类的群体,通过调节 γ 的取值,可适当度量易分类和难分类的界限,并分别计算它们在整体损失中所占比例。

$$FL(p_i) = -(1 - p_i)^\gamma \log(p_i) \quad (3)$$

使用 Smoth L1 LOSS 计算回归损失函数,定义为:

$$L_{reg}(t_i, t_i^*) = \text{Smoth L1}(t_i, t_i^*) = \begin{cases} 0.5(t_i - t_i^*)^2 & \text{if } |t_i - t_i^*| \leq 1 \\ |t_i - t_i^*| - 0.5 & \text{otherwise} \end{cases} \quad (4)$$

3 实验结果与分析

3.1 LUNA16 数据集

本文使用的肺结节数据集是 LUNA16,其是最大公用肺结节数据集 LIDC-IDRI^[18]的子集,LIDC-IDRI 包括 1 018 个低剂量的肺部影像。LIDC-IDRI 删除了切片厚度大于 3mm 和肺结节小于 3mm 的 CT 影像,剩下的即为 LUNA16 数据集。LUNA16 中有 888 幅 CT 影像,其中包含 1 186 个至少由 3 位放射学专家达成一致、提供位置框标注的肺结节。

3.2 实验细节

实验硬件环境如下:CPU 为 Inter Xeon Gold 6148,GPU 为 Tesla V100,操作系统为 Linux,编程语言为 Python,深度学习框架为 Pytorch。由于整个候选结节检测网络结构是三维的,需要较多的 GPU 资源,所以实验在 8 个 Tesla V100 上运行。因 GPU 内存的限制,Batch-Size 不能设置

得过大,本文将其设置为 32。

对每一个模型训练 300 个 epochs,训练使用的优化方法是 Adam^[19]。初始学习率为 0.01,每隔 100 epochs,学习率下调为原来的 0.1 倍,最后 epochs 将学习率下调到 0.000 1。

3.3 评价指标

肺结节检测结果可看作一个二分类问题,即肺结节和非肺结节分类情况。因此,肺结节实际上是分类问题中的正例,非肺结节是负例。对正例与负例进行预测存在 4 种情况:①正例预测为正例(记为真阳性);②正例预测为负例(记为假阴性);③负例预测为正例(记为假阳性);④负例预测为负例(记为真阴性)。预测结果与实际结果混淆矩阵如表 1 所示。

表 1 预测结果与实际结果混淆矩阵

真实结果	预测结果	
	肺结节	非肺结节
肺结节	真阳性	假阴性
非肺结节	假阳性	真阴性

肺结节检测中常用的性能评估指标有:敏感度、FROC 曲线和 CPM 值。

敏感度(Sensitivity)即灵敏度,是指实际为正样本的数据被检测为正样本,检测到的正样本占全部正样本的比例。敏感度越大,说明对正样本的检测效果越好。敏感度公式如下:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100\% \quad (5)$$

FROC(Free-response Receiver Operating Characteristic)曲线的横坐标表示平均每个样本包含的不同 FP 值,如:1/8、4、8 等,纵坐标为敏感度。其中横坐标在 1/8、1/4、1/2、1、2、4 和 8 上对应敏感度的平均值称为 CPM(Competition Performance Metric)。一般使用 CPM 值作为评估肺结节检测网络的指标,CPM 值越高,说明肺结节检测效果越好。

3.4 实验结果

本文在公用数据集 LUNA16 上测试模型性能,为了更全面地评估模型表现情况,采用 10 倍交叉验证方法。将 LUNA16 数据集分成 10 份,第一次使用 9 份数据集进行训练,剩下 1 份数据集进行测试。重复该过程 10 次,并保证 10 次测试的数据集不同,将所有测试结果汇总,根据评价指标 CPM 值进行模型评估。

表 2 比较了很多优秀的肺结节检测网络,可以发现本文提出的 3D FPN 在肺结节检测中取得了最好成绩,CPM 值达到 0.893 2,同时在假阳性例数为 8 时,网络敏感度达到 0.969,验证了网络的有效性。

3.5 实验结果分析

为了检验模型的有效性,以及 3D FPN 肺结节检测模型对小结节有足够的敏感度,本文对 3D FPN 模型进行修改,并在最后一个尺度上进行检测,标记为 3D FPN*。通过统计假阳性例数为 8 时的未检测肺结节数量,对比实验结果如表 3 所示。

表 2 各模型在 LUNA16 数据集上表现

Model	1/8	1/4	1/2	1	2	4	8	CPM
DIAG_ConvNet	0.692	0.771	0.809	0.863	0.895	0.914	0.923	0.838
Dou et al	0.659	0.754	0.819	0.865	0.906	0.933	0.946	0.893
DeepLung								0.842
Ding et al	0.784	0.853	0.887	0.922	0.938	0.944	0.946	0.8911
Ours (3D FPN)	0.752	0.846	0.876	0.918	0.944	0.948	0.969	0.8932

注:表格空白的是文献没有给出

表 3 对比实验结果

Model	小于 10mm	大于 10mm	小于 10mm 占比
3D FPN	24	13	0.648
3D FPN*	37	13	0.74

对比实验结果证明了 3D FPN 网络的有效性,与其它方法相比,本文提出的网络结构对小肺结节也有很好的检测效果。

图 5 展示了 3D FPN 网络检测部分结果,所有图片全部来自于测试数据集。首先,对测试数据集进行数据预处理,即肺实质分割;然后,将其切割成一个个大小为 $128^3 \times 1$ 的部分;最后,将 $128^3 \times 1$ 的数据直接输入到 3D FPN 网络中,得到肺结节检测结果。从图中可以发现,网络能够检测出不同大小的肺结节,说明网络对于不同尺寸的肺结节具有高度敏感性。

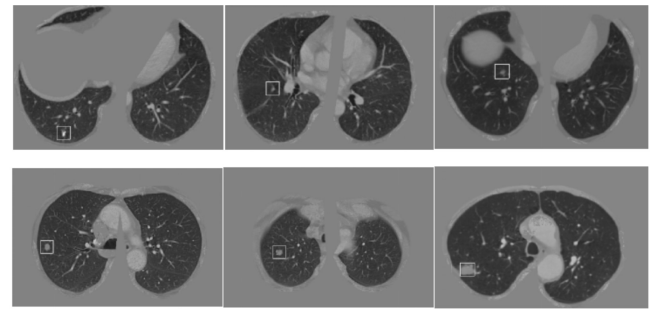


图 5 检测效果展示

4 结语

通过统计检测目标肺结节的半径,发现检测目标分布不均衡,大部分都是肺小结节。因此,本文提出基于 3D FPN 的肺结节检测网络,该网络结构能够结合深层语义信息,同时引入多尺度预测,大大提升了对肺小结节的检测效果。在公用肺结节数据集 LUNA16 上验证了网络的有效性。与现有目标检测算法相比,本文算法能够检测出更多肺结节,对于肺小结节也有更高的敏感性。

虽然本文提出的 3D FPN 网络已取得很好的效果,但是基础网络模块采用 ResNet 结构,无法很好地挖掘肺结节深度信息。接下来将通过更改基础网络模块,使整个网络能够更高效地挖掘深度信息,以获得更好的检测效果。

参考文献:

[1] INFANTE M, CAVUTO S, LUTMAN F R. A randomized study of lung cancer screening with spiral computed tomography[J]. American Journal of Respiratory and Critical Care Medicine, 2009, 180(5): 445-453.

[2] LIAO F, LIANG M, LI Z, et al. Evaluate the malignancy of pulmonary nodules using the 3D deep leaky noisy-or network[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(11): 3484-3495.

[3] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39(6): 1137-1149.

[4] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.

[5] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: single shot multibox detector[C]. European Conference on Computer Vision, 2016: 21-37.

[6] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.

[7] GINNEKEN B V, ARMATO S G, HOOP B D, et al. Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: the anode09 study[J]. Medical Image Analysis, 2010, 14(6): 707-722.

[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] DING J, LI A, HU Z, et al. Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks[C]. International Conference on Medical Image Computing and Computer-Assisted Intervention, 2017: 559-567.

[10] DOU Q, CHEN H, JIN Y, et al. Automated pulmonary nodule detection via 3D ConvNets with online sample filtering and hybrid-loss residual learning[C]. 20th International Conference on Medical Image Computing and Computer Assisted Intervention, 2017: 630-638.

[11] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 39(4): 640-651.

[12] SETIO AAA, TRAVERSO A, DE B 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.

[13] ZHU W, LIU C, 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). IEEE, 2018: 673-681.

[14] CHEN Y, LI J, XIAO H, et al. Dual path networks[C]. Advances in Neural Information Processing Systems, 2017: 4467-4475.

[15] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. Computer Vision and Pattern Recognition. IEEE, 2016: 770-778.

[16] MAAS A L, HANNUN A Y, NG A Y. Rectifier nonlinearities improve neural network acoustic models [C]. ICML Workshop on Deep Learning for Audio, Speech and Language Processing, 2013: 1-6.

[17] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.

[18] MCNITT-GRAY M F, ARMATO III S G, MEYER C R. The lung image database consortium (LIDC) data collection process for nodule detection and annotation[J]. Academic Radiology, 2007, 14(12): 1464-1474.

[19] KINGMA D P, BA J. Adam: a method for stochastic optimization [C]. International Conference on Learning Representations, 2015: 1-5.

(责任编辑:黄健)