

# 基于深度学习的肺部肿瘤检测方法

陈强锐,谢世朋

(南京邮电大学 通信与信息工程学院,江苏 南京 210003)

**摘要:**随着现代计算机技术的发展与应用,计算机辅助诊断系统在医学影像分析领域的地位变得愈发重要。其技术的关键在于病灶的定位与分类。由于图像的特征提取十分复杂,若应用传统机器学习方法,则需对图像作大量的预处理。文中提出一种基于深度学习的肺部肿瘤检测方法,运用卷积神经网络对患者肺部肿瘤图像进行特征提取。结合区域建议网络预测肿瘤在图片中可能存在的位置,同时生成建议框。利用学习好的特征对目标区域进行分类并微调建议框的位置。该方法无需人工设计目标特征,通过卷积神经网络学习到的特征更加具有代表性,且能够较好地预测肿瘤的位置。在 NLST 以及 Kaggle 的数据集上对该方法进行了评估。实验结果表明,该方法具有较高的准确率和效率。

**关键词:**深度学习;肺部肿瘤检测;特征提取;卷积神经网络;区域建议网络

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2018)04-0201-04

doi:10.3969/j.issn.1673-629X.2018.04.043

## Lung Cancer Detection Method Based on Deep Learning

CHEN Qiang-ruì, XIE Shi-peng

(School of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

**Abstract:** With the development and application of modern computer technology, computer aided diagnosis system plays an important role in the field of medical image analysis. The key to this technology is the location and classification of lesions. As the feature extraction of the image is very complex, many image preprocessing methods must be used when the traditional machine learning method as a solution. We put forward a method for the detection of lung cancer based on deep learning. The images of lung tumors' features are extracted by convolutional neural network, and the locations in the picture where tumors may exist are predicted in combination with region proposal network, generating the proposal boxes simultaneously. The target area is classified by learned feature maps and the positions of the proposal boxes will be fine-tuned. The proposed method does not need to design the features manually, and the features learned by the convolutional neural network are more representative, which also could predict the location of the tumor. The experiments of evaluation on NLST and Kaggle data sets show that it has high accuracy and efficiency.

**Key words:** deep learning; lung tumor detection; feature extraction; convolutional neural network; region proposal network

## 0 引言

计算机辅助诊断(CAD)<sup>[1]</sup>是指运用计算机技术分析计算影像、病理等数据,辅助发现病灶,从而提高诊断的准确率。现代计算机技术的快速发展对其起到了极大的推进作用。目前CAD技术已经广泛应用于各种疾病的诊疗过程中,被形象地称之为医生的“第三只眼”。肺癌检测是最常见的CAD技术应用之一。近年来国内肺癌发病率呈逐年上升趋势,年平均增长率为1.63%。而肺癌的发病率及死亡率已居所有恶

性肿瘤之首。引入计算机辅助诊断技术对肺癌的早期发现与诊疗有着重要的积极作用,所以它在辅助医生诊疗肺癌领域有着巨大的发展前景。

对于肺癌检测的CAD技术关键在于肿瘤病灶的定位与分类。目前已经有许多传统机器学习的方法可以对肿瘤进行检测。文献[2]从灰度特征、纹理特征、形态特征等多个角度提取肺部CT图像的特征,并利用支持向量机<sup>[3]</sup>(SVM)对肺部CT图像进行分类分析。文献[4]针对肺结节的形态、位置、纹理、灰度等

收稿日期:2017-05-05

修回日期:2017-09-07

网络出版时间:2017-12-05

基金项目:国家自然科学基金(11547155);教育部-中国移动科研基金(MCM20150504);江苏省科技重点研发计划-产业前瞻与共性关键技术(BE2016001-4);南京邮电大学科研基金(NY214026, NY217035)

作者简介:陈强锐(1994-),男,硕士研究生,研究方向为图像处理;谢世朋,副教授,研究方向为图像处理。

网络出版地址:cnki.net/kcms/detail/61.1450.TP.20171205.1434.114.html

17 个特征,利用主成分分析(PCA)方法,对特征集进行优化选择和降维处理,最后利用基于 SVM 的分类器对样本进行分类检测。文献[5]开发了一种提取肺结节特征信息的转换器,并结合随机森林算法利用提取到的特征信息对模型进行训练和测试。这些方法较好地解决了分类问题,但实验结果也并非十分完美,仍有提升的空间,尤其在复杂的特征提取方法和结节的位置预测方面。

深度学习<sup>[6]</sup>的概念于 2006 年由 Hinton 等提出,作为机器学习的一个分支,由于其善于发现高维数据的复杂结构,使用泛化目标的学习过程可以自动学习好的特征,所以近年来发展十分迅速。深度学习算法模仿人脑的机制来解释并处理数据,在语音识别、图像识别、自然语言理解等领域取得了重大突破。卷积神经网络(CNN)则是深度学习在图像领域的一个重要应用,已成为众多科学领域的研究热点之一。卷积神经网络对图像的特征提取十分方便,只需设置合理的网络结构与网络参数,每对图像进行一次卷积操作就可以生成一个特征图。通过加深网络结构可以学习到图像更深层次的特征。相较于人工设计特征的方法,卷积神经网络的适用范围更广,对图像的特征提取更加深刻全面。国内知名学者也曾表明深度学习将取代人工特征加机器学习的方法而逐渐成为主流图像识别方法<sup>[7]</sup>。文中采用深度学习的方法,将卷积神经网络应用到肺部肿瘤图像的特征提取中,结合区域生成网络预测肺癌结节的位置,并通过实验验证该方法的有效性。

## 1 肿瘤检测

肿瘤检测的流程如图 1 所示。首先让整幅图像经过卷积神经网络进行特征提取,然后针对获得的特征图像生成肿瘤位置建议框,最后对各个建议框包围的图像进行分类并微调建议框的位置。

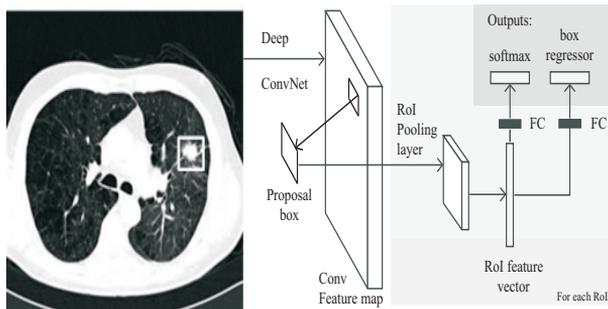


图 1 肺部肿瘤检测流程

### 1.1 特征提取

对于图像的特征提取方法有很多,它们大多依赖于人工的设计。例如算法 HOG<sup>[8]</sup>和 SIFT<sup>[9]</sup>。但是 SIFT 和 HOG 依赖于图像中梯度的方向直方图<sup>[10]</sup>,并

不具备普适性。

1998 年,LeCun 提出的卷积神经网络是第一个真正多层结构学习算法<sup>[11]</sup>,并成功地将其应用到手写数字识别中,标志着卷积神经网络对图像特征提取的第一次成功应用。由于当时的硬件计算条件不足以支撑更深层次的网络,只适合做小图片的识别,因此对于大规模数据识别效果不佳。直至 2012 年 ImageNet 图像分类竞赛中,AlexNet 凭借 Top-5 错误率低于上一年十个百分点的优秀成绩,使得 CNN 受到了研究者的重视<sup>[12]</sup>。

卷积神经网络的再次流行主要得益于非线性激活函数 Relu 和防止过拟合方法 Dropout 的提出,当然还有大数据训练以及 GPU 并行计算的发展。文中使用的特征提取网络是基于 Zeiler 和 Fergus 提出的 ZF<sup>[13]</sup>网络。

网络结构如图 2 所示。

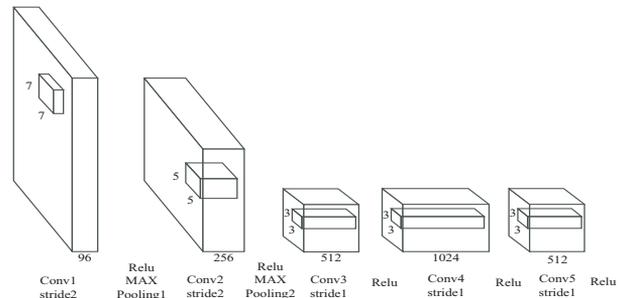


图 2 特征提取网络结构

网络总共含有 5 层可共享特征的卷积层,每一卷积层都应用了一个 Relu 激活函数。卷积层第一层与第二层的卷积核大小分别为  $7 * 7$  和  $5 * 5$ ,卷积步长为 2。卷积层后三层的卷积核大小为  $3 * 3$ ,卷积步长为 1。卷积层的第一层与第二层之后分别设置了一个最大值池化层。这种网络结构保证了神经网络既能充分学习到图像特征,又能够防止过拟合。

### 1.2 区域建议网络

传统生成区域建议框的方法如 selective search<sup>[14]</sup>,是将输入图像分割成许多相邻的小区域,计算相邻区域的相似度且合并相似度最高的相邻区域然后重复迭代,并融合各种多样化策略。虽然这种方法相较之前 exhaustive search<sup>[15]</sup>方法提升很大,但是其不能通过 GPU 加速导致耗时很大。或者是像 YOLO<sup>[16]</sup>中将原始图片分成  $S * S$  个网格,将生成建议框看成单一的回归问题,虽然解决了耗时问题,但是这种方法对于处在网格边缘或者尺寸较小的肿瘤,生成的建议框效果较差。

上述两种方法的输入都是原始图像,并未充分利用 CNN 输出的特征提取后的图像。文中使用的区域建议网络是 RPN,结构如图 3 所示。其输入是 CNN 网

络输出的特征图像。该方法在特征图像上设计了一个滑动窗口,在每一个窗口的中心生成 3 种不同大小、3 种不同宽纵比,总计 9 个固定的建议框。同时将每一个滑动窗口映射成为一个 256 维的向量,该向量输出给两个同级的全连接层 cls layer 与 reg layer,分别用作建议框的回归和分类。

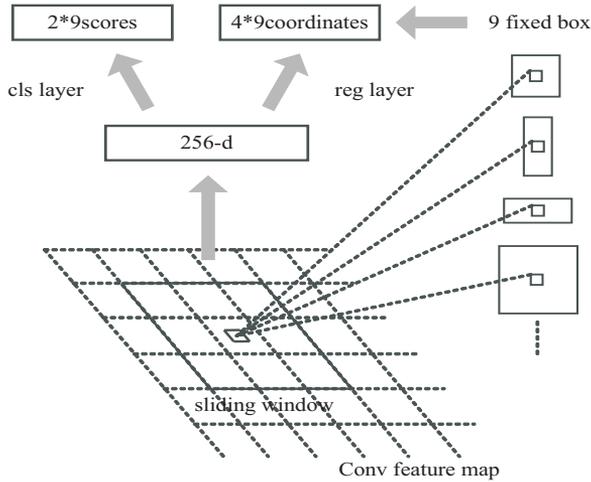


图 3 区域建议网络

文中方法利用目标区域与 ground-truth 面积的交并比 (IOU) 对建议框进行筛选, IOU 定义为:

$$IOU = \frac{A \cap B}{A \cup B} \quad (1)$$

其中,  $A$  与  $B$  分别表示不同区域的面积。

选用 IOU 大于 0.7 的建议框作为目标建议框, 将 IOU 小于 0.3 的建议框记为背景样本, 弃用其余以及跨越图像边界的建议框。每张图片最后至多生成 300 个建议框, 并利用其对应的向量对建议框内容进行分类且微调其位置。

## 2 实验

### 2.1 损失函数及训练

实验采用的数据集来自于 NLST 以及 Kaggle。从不同肺癌患者的 CT 中挑选出 451 张切片, 并将 DICOM 文件格式转换成  $512 * 512$  像素的 JPG 图像。随机挑选出 361 幅图像并对这些图像中肺癌特征区域进行标注, 作为训练集, 剩余的 90 幅图像则作为测试集。文中采用的检测框架为 Faster RCNN<sup>[17]</sup>, 对每幅图像的损失函数定义为:

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (2)$$

其中,  $i$  为建议框的索引号;  $p_i$  为建议框  $i$  中含有肿瘤的概率;  $p_i^*$  为 ground-truth 的标签, 如果含有肿瘤即为 1, 不含数据为 0;  $t_i$  为代表建议框坐标的四维向

量;  $t_i^*$  为代表 ground-truth 坐标的四维向量。

分类损失函数  $L_{cls}$  定义为:

$$L_{cls}(p_i, p_i^*) = -\log[p_i^* p_i + (1 - p_i^*)(1 - p_i)] \quad (3)$$

建议框回归损失函数  $L_{reg}$  定义为:

$$L_{reg} = R(t_i - t_i^*) \quad (4)$$

其中,  $R$  是鲁棒的损失函数  $smooth_{L1}$ <sup>[18]</sup> 定义为:

$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases} \quad (5)$$

网络训练时损失函数超参数  $\lambda$  设置为 10, 卷积神经网络的基础学习速率设置为 0.001, 权重衰减设置为 0.0005, 在配置为 Intel Core i7-2600 3.4 GHz 处理器, 4 G 内存, 8 G 显存, 显卡为 GTX1070 的服务器上训练总时长约 4 个小时。

### 2.2 实验结果分析

经过实验, 测试集共 90 幅 CT 图像成功检测出 86 幅, 其中典型的检测效果如图 4 所示。

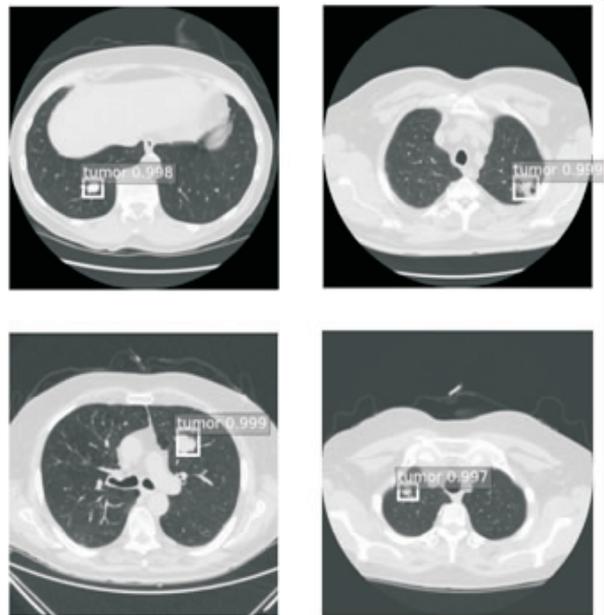


图 4 检测结果

由图 4 可见, 肺部肿瘤可以被文中方法准确定位, 并且具有较好的检测效果。

文中准确率定义为:

$$Accuracy = \frac{\sum_j I_{det} P_r(j)}{N} \quad (6)$$

其中,  $j$  为测试图像的索引号;  $I_{det}$  为检测结果标记, 检测成功为 1, 检测失败为 0;  $P_r(j)$  为预测概率;  $N=90$  为参与检测的图像总数。

利用相同的数据集在 YOLO\_V2 框架上进行了重复实验, 并将两次的实验结果进行了对比, 结果如表 1 所示。

表 1 不同框架检测结果对比

检测框架	准确率/%	训练时间/h	检测耗时/(s/幅)
Faster RCNN	95.1	4	0.194
YOLO_V2	80.7	20	0.136

由表 1 可以看出,文中方法的检测效果更佳,准确率高达 95.1% 且高于 YOLO\_V2 约 15 个百分点,并且训练时间也大大缩短。虽然每幅图像的检测用时略高于 YOLO\_V2,但是综合考虑各方面因素,文中方法对于肺部肿瘤的检测有着更好的性能。

将文中方法的检测结果与利用传统机器学习方法的检测结果进行了对比,结果如表 2 所示。

表 2 不同方法检测结果对比

检测方法	准确率/%
文献[2]中方法	89.374
文献[4]中方法	87.2
文中方法	95.1

由表 2 可以看出,文中方法的检测准确率可达 95.1%,相较于其他检测方法,具有更高的准确率,其准确率提升约 6 个百分点。由于文中方法可以对肿瘤图像自动地进行特征提取,并且深层次的网络结构可以提取到更好的特征,所以检测效果明显有所提高。并且该方法还可以准确定位出肿瘤的具体位置,相较于其他方法具有一定的优势。

### 3 结束语

提出了基于深度学习的肺部肿瘤的检测方法,通过卷积神经网络对 CT 图像进行特征提取,最终定位出图像中肿瘤的位置。该方法在 NLST 以及 Kaggle 数据集上取得了较好的实验效果,对于肿瘤检测领域具有一定的积极意义。然而,该方法只考虑到肿瘤图像的二维特征,若能够对肿瘤进行三维特征提取,其检测效果则会更好。接下来的工作将向此方面发展。

#### 参考文献:

[1] 舒荣宝,王成林.电子计算机辅助诊断(CAD)的原理及临床应用[J].中国CT和MRI杂志,2004,2(2):55-56.  
 [2] 马杨林.基于SVM的肺部CT图像特征提取及分类研究[D].成都:西华大学,2012.  
 [3] SCHÖLKOPF B, SMOLA A. Learning with kernels: support vector machines, regularization, optimization, and beyond [J]. IEEE Transactions on Neural Networks, 2005, 16(3): 781.

[4] 张婧.基于SVM的肺结节自动识别方法研究[D].广州:华南理工大学,2011.  
 [5] KOUZANI A Z, LEE S L A, HU E J. Lung nodules detection by ensemble classification [C]//IEEE international conference on systems, man and cybernetics. [s.l.]: IEEE, 2008: 324-329.  
 [6] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521(7553): 436-444.  
 [7] 余凯,贾磊,陈雨强,等.深度学习的昨天、今天和明天 [J]. 计算机研究与发展, 2013, 50(9): 1799-1804.  
 [8] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.  
 [9] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]//IEEE computer society conference on computer vision & pattern recognition. Washington DC, USA: IEEE Computer Society, 2005: 886-893.  
 [10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//IEEE conference on computer vision and pattern recognition. Washington DC, USA: IEEE Computer Society, 2014: 580-587.  
 [11] 孙志军,薛磊,许阳明,等.深度学习研究综述[J].计算机应用研究,2012,29(8):2806-2810.  
 [12] 李彦冬,郝宗波,雷航.卷积神经网络研究综述[J].计算机应用,2016,36(9):2508-2515.  
 [13] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks [J]. Lecture Notes in Computer Science, 2013, 8689: 818-833.  
 [14] UIJLINGS J R, SANDE K E, GEVERS T, et al. Selective search for object recognition [J]. International Journal of Computer Vision, 2013, 104(2): 154-171.  
 [15] HARZALLAH H, JURIE F, SCHMID C. Combining efficient object localization and image classification [C]//International conference on computer vision. [s.l.]: [s.n.], 2010: 237-244.  
 [16] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//IEEE conference on computer vision and pattern recognition. Washington DC, USA: IEEE Computer Society, 2016: 779-788.  
 [17] 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, 2017, 39(6): 1137-1149.  
 [18] GIRSHICK R. Fast R-CNN [C]//International conference on computer vision. [s.l.]: [s.n.], 2015: 1440-1448.