

## Classification of pulmonary nodules on PET/CT image based on deep belief network

MA Yuan<sup>1</sup>, WANG Feng<sup>2</sup>, HAN Yong<sup>1</sup>, ZHANG Feng<sup>1</sup>, LIANG Zhigang<sup>3</sup>,  
HUANG Jian<sup>4</sup>, YANG Zhi<sup>2\*</sup>, GUO Xiuhua<sup>1\*</sup>

(1. Department of Epidemiology and Health Statistics, School of Public Health, Capital Medical University, Beijing Key Laboratory of Clinical Epidemiology, Beijing 100069, China; 2. Department of Nuclear Medicine, Peking University Cancer Hospital, Beijing Institute for Cancer Research, Ministry of Education Key Laboratory of Carcinogenesis and Translational Research, Beijing 100036, China; 3. Department of Nuclear Medicine, Xuanwu Hospital Capital Medical University, Beijing 100053, China; 4. School of Mathematical Sciences, University College Cork, Cork T12 K8AF, Ireland)

**[Abstract]** **Objective** To observe classification effect of pulmonary nodules on PET/CT images with deep belief network (DBN). **Methods** PET/CT images of 216 patients with pulmonary nodules were collected, among them 339 pulmonary nodules were detected, including 190 benign and 149 malignant ones. Totally 2 055 ROI images were captured, including 1 069 of benign ones and 986 of malignant ones. Gray scale and size normalization were performed on ROI images, and then the lesions were detected with DBN. The network structure and training parameters were determined by experimental methods, and the results were evaluated by confusion matrix, overall accuracy, *Kappa* coefficient and other indicators. A support vector machine model (SVM) was also built with wavelet texture features based on nonsubsampling dual-tree complex contourlet transform (NSDTCT), using the same data as DBN. The results detected with DBN and SVM were compared. **Results** The results of DBN and SVM methods were 0.94 and 0.72 for overall accuracy, 0.96 and 0.66 for sensitivity, 0.92 and 0.96 for specificity, and 0.87 and 0.42 for *Kappa* coefficient, respectively. **Conclusion** The accuracy of DBN in identifying benign and malignant pulmonary nodules is better than that of SVM.

**[Keywords]** lung neoplasms; diagnosis; artificial intelligence; positron-emission tomography

**DOI:**10.13929/j.issn.1003-3289.2020.01.021

## 基于深度信念网络检测 PET/CT 图像肺结节良恶性

马 圆<sup>1</sup>, 王 风<sup>2</sup>, 韩 勇<sup>1</sup>, 张 凤<sup>1</sup>, 梁志刚<sup>3</sup>, 黄 健<sup>4</sup>, 杨 志<sup>2\*</sup>, 郭秀花<sup>1\*</sup>

(1. 首都医科大学公共卫生学院流行病学与卫生统计学系 北京市临床流行病学重点实验室, 北京 100069;  
2. 北京大学肿瘤医院暨北京市肿瘤防治研究所核医学科 恶性肿瘤发病机制及转化研究教育部  
重点实验室, 北京 100036; 3. 首都医科大学宣武医院核医学科, 北京 100053;  
4. 爱尔兰科克大学数学学院, 爱尔兰 科克 T12 K8AF)

**[基金项目]** 国家自然科学基金项目(81773542)、国家青年科学基金项目(81703318)。

**[第一作者]** 马圆(1992—), 女, 江苏徐州人, 在读硕士。研究方向: 医学图像大数据方向。E-mail: 598941139@qq.com

**[通信作者]** 杨志, 北京大学肿瘤医院暨北京市肿瘤防治研究所核医学科, 恶性肿瘤发病机制及转化研究教育部重点实验室, 100036。  
E-mail: pekyz@163.com。

**[通信作者]** 郭秀花, 首都医科大学公共卫生学院流行病学与卫生统计学系, 100069。E-mail: statguo@ccmu.edu.cn

**[收稿日期]** 2019-02-19 **[修回日期]** 2019-12-24

**[摘要]** **目的** 观察深度信念网络(DBN)方法识别 PET/CT 图像良恶性肺结节的效果。**方法** 收集 216 例肺结节患者的 PET/CT 图像,共 339 个肺结节,其中良性 190 个、恶性 149 个;共截取 2 055 张 ROI 图像,良性 1 069 张,恶性 986 张。对 ROI 图像进行灰度、大小归一化处理后,采用 DBN 方法进行分类诊断。通过实验方法确定网络结构及训练参数,并以混淆矩阵、总体精度、*Kappa* 系数等指标评价分类结果。提取同一批图像数据非下采样双树复轮廓波变换(NSDTCT)的小波纹理参数,构建支持向量机分类模型(SVM),对比分析其与 DBN 的检测结果。**结果** DBN 和 SVM 方法测试集检测结果分别为总体精度 0.94 和 0.72、灵敏度 0.96 和 0.66、特异度 0.92 和 0.96 及 *Kappa* 系数 0.87 和 0.42。**结论** DBN 识别肺结节良恶性的准确性高于 SVM 方法。

**[关键词]** 肺肿瘤;诊断;人工智能;正电子发射断层显像术

**[中图分类号]** R734.2; TP181 **[文献标识码]** A **[文章编号]** 1003-3289(2020)01-0077-04

随着大数据时代到来,为解决数字化医学影像数量增长而增加的影像科医师阅片负担的问题,多种医学图像分析工具<sup>[1]</sup>被开发出来,其中影像组学备受关注<sup>[2]</sup>。影像组学通过提取图像中的大量特征量化肿瘤等疾病,定量肿瘤异质性<sup>[3]</sup>,其中图像特征往往需要进行人工设计,对研究者要求较高<sup>[4]</sup>;且线性方法难以提取高维数据信息,而各种浅层非线性学习算法对复杂函数的表示能力有限,对于数据特征提取的要求更高<sup>[5]</sup>。本研究探讨能自动提取图像特征的深度信念网络(deep belief network, DBN)方法的检测准确性,并与浅层非线性学习算法中分类效果较好的支持向量机(support vector machine, SVM)方法进行比较。

## 1 资料与方法

**1.1 一般资料** 收集 216 例肺结节患者的 PET/CT 图像,男 127 例,女 89 例,年龄 37~86 岁,平均(64.8±10.1)岁;其中 105 例为 2015 年 5 月—2018 年 4 月于首都医科大学宣武医院接受治疗的患者,63 例为 2018 年 1 月于北京市肿瘤防治研究所接受治疗的患者,48 例为 2012 年 10 月—2013 年 3 月于爱尔兰科克大学医院接受治疗的患者。本研究部分患者有 2 个及以上肺结节,共纳入 339 个肺结节,其中 190 个良性结节,149 个恶性结节。纳入标准:恶性结节经病理诊断或出院诊断肺癌并经随访证实;良性结节为病理诊断或经影像学随访确诊为良性病变。排除标准:①资料不完整;受化学治疗及放射治疗后;②融合图像上病变位置不清楚。

**1.2 勾画 ROI** 采用 UNITED IMAGING uMI S-96R 及 GE Discovery VCT PET/CT 机。在影像科医师指导下,采用 Image J 医学图像处理软件,逐层手动勾画 PET/CT 融合图像上每个肺结节病变区域,尽可能不包含病变以外肺组织。根据病变区域大小截取 2~10 张 ROI 图像,采用 Photoshop 等图像工具截取并保存为 jpg 格式(图 1)。最终获得 2 055 张 ROI 图像,其中良性 1 069 张,恶性 986 张。参照文献<sup>[6-7]</sup>方法对

不同来源图像进行标准化处理,灰度化、归一化后,将图像大小统一为 32×32 像素。

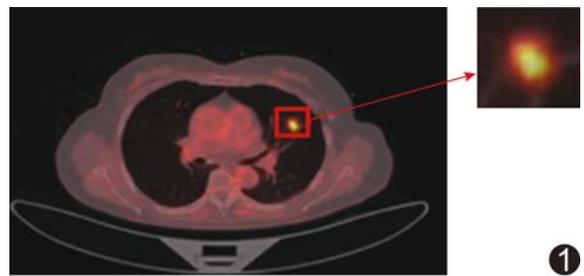


图 1 截取肺部 PET/CT 轴位图像 ROI

**1.3 确定 DBN 结构及训练参数** 将预处理后的肺结节 ROI 图像作为 DBN 输入图像(图 2),对网络进行构建、训练和评价。采用实验方法确定较优的网络结构,网络输入层有  $32 \times 32 = 1\,024$  个节点,中间隐藏层由若干受限玻尔兹曼机(restricted Boltzmann machine, RBM)构成。采用分类模型进行分类,标签为良性、恶性 2 类,因此最后网络输出层有 2 个节点。假设各隐含层的节点数目相同,且各隐含层均采用小数据集训练(mini batch)方法,各隐藏层具有相同学习率(learning rate)、隐含层训练次数(RBM epoch)及小数据集大小(minibatch size)。实验参数:DBN 隐藏层层数取值集合为{1,2,3,4,5,6},隐藏层节点数取值集合为{16,32,64,128,256,512},learning rate 取值集合为{0.8,0.1},RBM epoch 取值集合为{20,50},minibatch size 取值集合为{32,64}。将样本迭代次数统一为 1 000 次。

**1.4 构建 SVM 模型** 采用基于非下采样双树复轮廓波变换(non-subsampled dual-tree complex contourlet transform, NSDTCT)的小波纹理特征<sup>[8-9]</sup>构建 SVM 分类模型,对图像进行检测。首先对每幅 ROI 图像进行 NSDTCT 变换,得到 96 个子图像,再对每幅子图像提取纹理参数,包括 6 个灰度直方图特

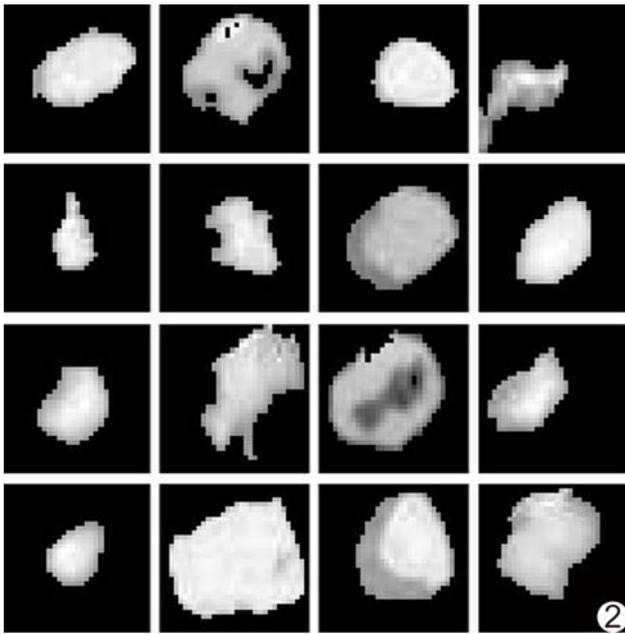


图 2 DBN 输入图像示例

征(均值、标准差、平滑度、三阶矩、一致性、熵), 14 个灰度共生矩阵特征(能量、灰度均值、逆差矩、熵、相关性、聚类趋势、对比度、同质度、方差、最大概率、和的均值、差的均值、和的熵、差的熵)和 3 个灰度邻域差分矩特征(粗糙度、对比度、频度), 并据此构建 SVM 模型。

1.5 图像及统计学分析 采用 MATLAB R2018a 软件和 R 软件 (<https://www.r-project.org/>), 以 MATLAB 软件 DeepLearnToolbox (<https://github.com/rasmusbergpalm/DeepLearnToolbox>) 分析 DBN 的构建、训练和评价, 以 80% 数据作为训练集, 20% 作为测试集。采用混淆矩阵、总体精度和  $Kappa$  系数评价分类结果。  $0 < Kappa \leq 0.20$  为极低一致性,  $0.21 < Kappa \leq 0.40$  为一般一致性,  $0.41 < Kappa \leq 0.60$  为中等一致性,  $0.61 < Kappa \leq 0.80$  为高度一致性,  $0.81 < Kappa \leq 1$  为几乎完全一致。基于 NSDTCT 的小波纹理特征建立 SVM 模型对良恶性肺结节进行分类, 随机抽取 80% 数据作为训练集, 以 R 软件的“e1071”工具包建立 SVM 模型, 以剩余 20% 数据作为测试集对分类诊断模型进行测试。分类器默认选择 C-classification, 采用 5 折交叉验证方法评估模型的准确率, 并在测试集上进行验证。

## 2 结果

2.1 DBN 网络结构、训练参数实验结果 隐层网络层从 1 到 6, 准确率、 $Kappa$  系数取得最大值时的分类结果。隐层数为 4 时, DBN 分类效果最好, 见表 1。从

{16, 32, 64, 128, 256, 512} 中选取隐层网络节点数, 准确率、 $Kappa$  系数取得最大值时的分类结果, 当隐层节点数为 512 时 DBN 分类效果最好, 见表 2。

2.2 DBN 建模及分类结果分析 根据实验结果, 采用的 DBN 网络模型结构为: 隐层网络层数为 4, 每个隐层的网络节点数为 512, 训练次数为 1 000 次。RBM 训练参数为: learning rate 为 0.1, 迭代次数为 20, minibatch size 为 32。测试集分类结果与患者资料几乎一致 ( $Kappa=0.87$ ), 见表 3。

2.3 SVM 分类结果分析 每张 ROI 图像共提取 2 208 个纹理特征建立 SVM 模型。采用 5 折交叉验证方法评估模型准确率的平均值为 0.69, 测试集分类结果与标签数据一致性中等 ( $Kappa=0.42$ , 表 3)。

表 1 各 DBN 隐层网络层的分类结果

隐层网络层数	准确率	$Kappa$
1	0.937	0.874
2	0.934	0.868
3	0.932	0.864
4	0.949	0.898
5	0.944	0.888
6	0.944	0.888

表 2 各 DBN 隐层节点数的分类结果

隐层节点数	准确率	$Kappa$
16	0.881	0.763
32	0.888	0.776
64	0.925	0.849
128	0.925	0.849
256	0.937	0.874
512	0.949	0.898

表 3 DBN 与 SVM 对测试集的检测结果

方法	总体精度	$Kappa$	灵敏度	特异度
DBN	0.94	0.87	0.96	0.92
SVM	0.72	0.42	0.67	0.76

## 3 讨论

深度学习方法可模拟大脑学习过程, 基于海量训练数据构建深层次模型学习数据中的隐含特征<sup>[10]</sup>。深度学习由人工神经网络模型发展而来, 通过多个非线性处理层对数据进行逐层抽象, 即将当前网络层的输出作为下一网络层的输入, 经过逐层堆叠和归纳, 得到更高级的表示, 从而获得不同层面的抽象特征。DBN 具有多层非线性结构, 对复杂数据有极强的非线性映射能力, 通过有监督和无监督学习过程的交替, 能较好地完

成特征学习与分类任务<sup>[11]</sup>。而传统影像组学方法往往涉及到分割图像、提取与选择特征和选择分类器各步骤,其相比深度学习需要更多的先验知识,且耗费时间与精力较多。本研究采用深度学习对肺结节图像进行分类,并与传统影像组学方法进行比较。

本研究将图像直接作为 DBN 的输入,通过实验方法确定最佳的网络结构和训练参数,采用混淆矩阵、总体精度、Kappa 系数作为分类评价指标。其中混淆矩阵反映分类结果与真实标签信息,是总体精度和 Kappa 系数指标计算的基础。混淆矩阵对角线的元素值越大,分类精度越高。总体精度是总体评价分类的准确率(accuracy)。传统影像组学方法采用的是在多尺度分析工具中表现较为优越的 NSDTCT 小波纹理特征及在高维、非线性数据分类中表现突出的 SVM 模型<sup>[12]</sup>。DBN 由若干层 RBM 堆叠而成,通过无监督学习的方法训练各层 RBM 为神经网络的权值提供一个合适的初始值,再通过监督学习的方式自动发掘图像中更深层次特征,从而提高分类精度<sup>[13]</sup>。比较本研究中 2 种方法图片分类结果,发现 DBN 分类效果优于 SVM,正确率提高 20%,提示 DBN 无论方法方便程度还是检测准确程度均优于纹理特征结合 SVM 的传统分析方法。

本研究 DBN 分类结果测试集的准确率达 0.94,灵敏度和特异度分别为 0.96 和 0.92,其分类效果优于 SVM 方法。一项 Meta 分析研究<sup>[14]</sup>显示,通过合并分析<sup>18</sup>F-FDG PET/CT 诊断肺癌的灵敏度、特异度、诊断优势比并进行异质性检验,<sup>18</sup>F-FDG PET/CT 诊断肺癌的合并灵敏度为 0.88、特异度为 0.56、诊断优势比为 9.10,绘制并计算 SROC 曲线下面积为 0.81、统计量  $Q^*$  为 0.75,提示基于 DBN 的计算机辅助诊断技术识别肺结节的效果优于医生基于 PET/CT 图像诊断。

综上所述,深度学习方法对图像的分类结果较佳,但尚无法给出明确的临床解释<sup>[15]</sup>。将深度学习方法与临床指标或影像学指标相结合,具有重要临床意义。

## [参考文献]

- [1] TAHA A A, HANBURY A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool[J]. BMC Med Imaging, 2015, 15(1):29.
- [2] GILLIES R J, KINAHAN P E, HRICAK H. Radiomics: Images are more than pictures, they are data[J]. Radiology, 2016, 278(2):563-577.
- [3] LAMBIN P, LEIJENAR R T H, DEIST T M, et al. Radiomics: The bridge between medical imaging and personalized medicine[J]. Nat Rev Clin Oncol, 2017, 14(12):749-762.
- [4] YIP S S, AERTS H J. Applications and limitations of radiomics[J]. Phys Med Biol, 2016, 61(13):R150-166.
- [5] AKTER S, WAMBA S F. Big data and disaster management: A systematic review and agenda for future research[J]. Ann Oper Res, 2019, 283(1-2):939-959.
- [6] KATHER J N, KRISAM J, CHAROENTONG P, et al. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study[J]. Plos Med, 2019, 16(1):e1002730.
- [7] HOSNY A, PARMAR C, COROLLER T P, et al. Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study[J]. PloS Med, 2018, 15(11):e1002711.
- [8] 马圆, 韩鸿毅, 孙燕北, 等. 基于非下采样双树复轮廓波变换的小波纹理特征识别肺良恶性结节 CT 图像[J]. 中国医学影像技术, 2019, 35(2):272-276.
- [9] 赵杰, 李易瑾, 刘帅奇. 结合模糊逻辑和 SCM 的 NSDTCT 域红外和可见光图像融合[J]. 小型微型计算机系统, 2018, 39(2):352-356.
- [10] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786):504-507.
- [11] HINTON G. Deep belief networks[J]. Scholarpedia, 2009, 4(5):5947.
- [12] 马圆, 陈斯鹏, 田思佳, 等. 基于 PET/CT 图像纹理参数的肺结节诊断模型[J]. 北京生物医学工程, 2017, 36(3):257-261.
- [13] 吕启, 窦勇, 牛新, 等. 基于 DBN 模型的遥感图像分类[J]. 计算机研究与发展, 2014, 51(9):1911-1918.
- [14] 王龙龙, 王警建, 韩冰, 等. <sup>18</sup>F-FLT 与 <sup>18</sup>F-FDG PET/CT 诊断肺癌的 Meta 分析[J]. 中国医学影像学杂志, 2015(1):50-55.
- [15] STEAD W W. Clinical implications and challenges of artificial intelligence and deep learning [J]. JAMA, 2018, 320(11):1107-1108.