

✧胸部影像学

Machine learning in differentiating pulmonary invasive adenocarcinoma from non-invasive adenocarcinoma manifested as pure ground-glass nodule

CAI Junhui^{1,2}, DUAN Shaofeng³, YUAN Hu^{1,2}, LYU Yan²,
WU Yawei², XU Qing², YE Jing^{2*}

(1. Graduate School of Dalian Medical University, Dalian 116044, China; 2. Department of Radiology, Northern Jiangsu People's Hospital, Yangzhou 225000, China;
3. GE Healthcare, Shanghai 210000, China)

[Abstract] **Objective** To investigate the value of machine learning model based on radiomic features in differentiating pulmonary invasive adenocarcinoma from non-invasive adenocarcinoma manifested as pure ground-glass nodule (pGGN). **Methods** A total of 87 lung adenocarcinomas CT presented as pGGN were analyzed retrospectively, including 32 cases with invasive adenocarcinoma (IAC) and 55 cases with non-IAC (17 adenocarcinomas in-situ [AIS] and 38 minimally invasive adenocarcinomas [MIA]). The software ITK-SNAP was used to draw ROI, and the radiomic features were extracted using A. K. analysis software. After screening the significant characteristic parameters, the feature dimensionality reduction was conducted with *Spearman* analysis and *Lasso* regression. The final feature parameters were selected to construct three machine learning models, including support vector machine (SVM), random forest (RF) and logistics regression (LR). Then 10-fold cross validation was used to get the optimal model, and ROC curve was drawn to evaluate the performance of 3 models. **Results** A total of 396 radiomic features were extracted, and 19 features were finally obtained after feature screening. Machine learning models SVM, RF and LR could effectively distinguish IAC from non-IAC, with the accuracies of 93.30%, 86.70% and 83.30%, and AUC of 0.94, 0.92 and 0.83 respectively. **Conclusion** Machine learning models based on radiomic features have good classified performances, which can effectively distinguish IAC from non-IAC manifested as pGGN preoperation.

[Keywords] lung neoplasms; adenocarcinoma; tomography, X-ray computed; machine learning

DOI:10.13929/j.issn.1003-3289.2020.03.021

机器学习鉴别表现为肺纯磨玻璃结节的浸润性腺癌与非浸润性腺癌

蔡俊辉^{1,2}, 段绍峰³, 袁 虎^{1,2}, 吕 燕², 吴雅蔚², 许 晴², 叶 靖^{2*}

(1. 大连医科大学研究生院, 辽宁 大连 116044; 2. 苏北人民医院放射科, 江苏 扬州 225000;
3. GE 医疗, 上海 210000)

[摘要] **目的** 探讨基于影像组学特征构建的机器学习模型鉴别表现为肺纯磨玻璃结节的浸润性腺癌与非浸润性腺癌的可行性。**方法** 回顾性分析经手术病理证实的 87 例 CT 表现为纯磨玻璃结节的肺腺癌患者, 其中浸润性腺癌 32 例, 非浸润性腺癌 55 例(原位癌 17 例, 微浸润性腺癌 38 例)。应用 ITK-SNAP 软件勾画 ROI, A. K. 软件提取影像组学特征。筛选有

[第一作者] 蔡俊辉(1993—), 男, 湖北十堰人, 在读硕士。研究方向: 胸部影像诊断。E-mail: 364896362@163.com

[通信作者] 叶靖, 江苏省苏北人民医院放射科, 225000。E-mail: yzhyejing@163.com

[收稿日期] 2019-08-09

[修回日期] 2019-12-05

意义的特征参数,以 *Spearman* 相关性分析和 *Lasso* 回归分析进行特征降维。选取降维后的特征参数,分别构建支持向量机(SVM)、随机森林(RF)及逻辑回归(LR)3 种机器学习模型,采用十折交叉验证法得到最优模型,绘制 ROC 曲线,评价 3 种模型的性能。**结果** 共提取 396 个影像组学特征,通过特征筛选后最终得到 19 个影像组学特征。SVM、RF、LR 3 种机器学习模型可有效鉴别浸润性腺癌与非浸润性腺癌,准确率分别为 93.30%、86.70%和 83.30%,AUC 分别为 0.94、0.92 和 0.83。**结论** 基于影像组学特征构建的机器学习模型有较好的分类性能,可于术前有效鉴别肺浸润性腺癌与非浸润性腺癌。

[关键词] 肺肿瘤; 腺癌; 体层摄影术, X 线计算机; 机器学习

[中图分类号] R737.9; R814.42 **[文献标识码]** A **[文章编号]** 1003-3289(2020)03-0405-06

随着高分辨率 CT 的广泛应用和胸部低剂量 CT 筛查的普及,肺磨玻璃结节(ground-glass nodule, GGN)的检出率日益提高^[1]。根据其是否存在实性成分,GGN 又可分为纯磨玻璃结节(pure ground-glass nodule, pGGN)和部分实性 GGN。持续存在的 pGGN 大多是浸润前病变[包括不典型腺瘤样增生和原位腺癌(adenocarcinoma in situ, AIS)]或微浸润性腺癌(minimally invasive adenocarcinoma, MIA)^[2],但部分浸润性腺癌(invasive adenocarcinoma, IAC)CT 亦可表现为 pGGN,肉眼很难与 MIA、AIS 区分。文献^[3]报道 I A 期 IAC 患者 5 年无病生存率为 74.6%,而 AIS 和 MIA 手术完全切除后患者 5 年无病生存率为 100%。术前区别浸润性腺癌与非浸润性腺癌不仅可减轻患者心理负担,倾向 IAC 时,还可尽早干预并改善预后。影像组学用于辅助诊断疾病、评估疗效及预后,可帮助临床决策^[4]。本研究观察采用机器学习方法鉴别表现为 pGGN 的 IAC 与非 IAC 的可行性。

1 资料与方法

1.1 一般资料 收集 2016 年 9 月—2018 年 12 月苏北人民医院经手术病理证实为肺腺癌的患者。纳入标

准为:①术前接受 CT 平扫检查,肺窗(窗宽 1 500 HU、窗位 -600 HU)发现 pGGN,纵隔窗(窗宽 400 HU、窗位 30 HU)其内无钙化及实性成分;②图像质量好,无运动及呼吸伪影,与纵隔或胸膜无粘连;③拥有完整薄层重建图像,5 mm < GGN 直径 < 30 mm;④无其他肿瘤病史,检查前未经治疗。共 82 例患者(87 个 pGGN)纳入研究,男 25 例,女 57 例,年龄 23~70 岁,平均(52.5±10.6)岁;IAC 32 例,非 IAC 55 例(包括 17 例 AIS 和 38 例 MIA)。

1.2 仪器与方法 采用 GE Light Speed VCT 及 GE Discovery CT 750HD 进行图像采集。嘱患者仰卧,从肺尖扫描至后肋膈角。扫描参数:管电压 120 kV,自适应管电流,螺距 0.984:1,旋转时间 0.6 s,FOV 400 mm×400 mm,层厚及层间距均为 5 mm,重建层厚 1.25 mm。由 1 名具有 5 年胸部影像学诊断经验的医师在不知晓病理结果的情况下使用 ITK-SNAP 3.8.0 软件沿病灶边缘尽量避开周围大血管与支气管手动逐层勾画 ROI;再由 1 名从事胸部诊断 20 年以上的主任医师确认分割结果(图 1~3)。最终软件自动生成整个病灶的三维容积感兴趣区(volume of interest, VOD),以 NiFTI 格式将其导出并保存。

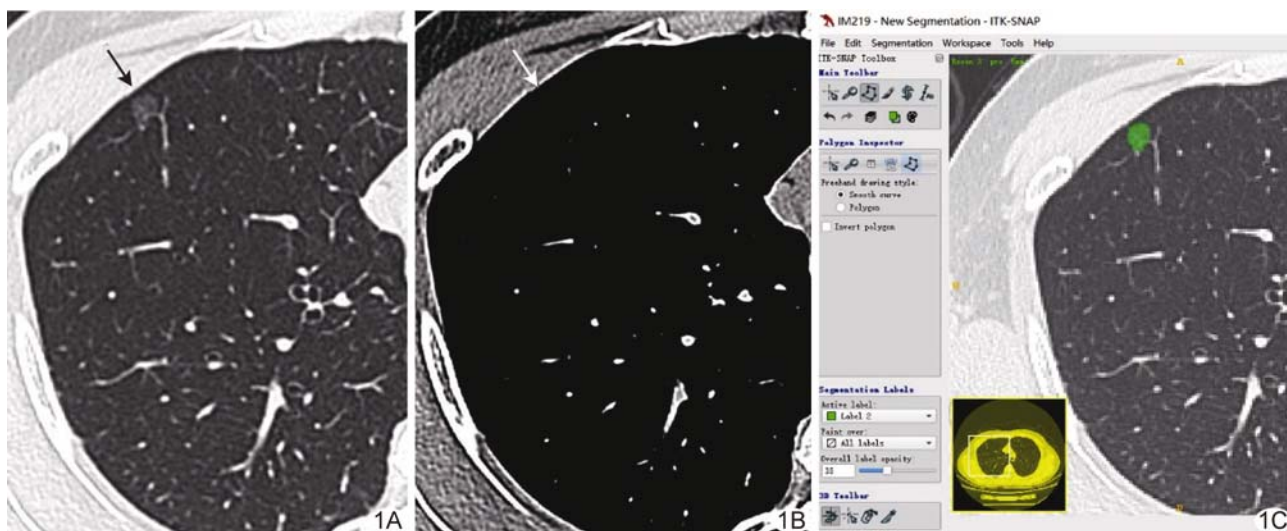


图 1 患者女,35 岁,右肺上叶 AIS A. 薄层肺窗; B. 薄层纵隔窗; C. 手动勾画的 ROI(绿色区域) (箭示要勾画区域)

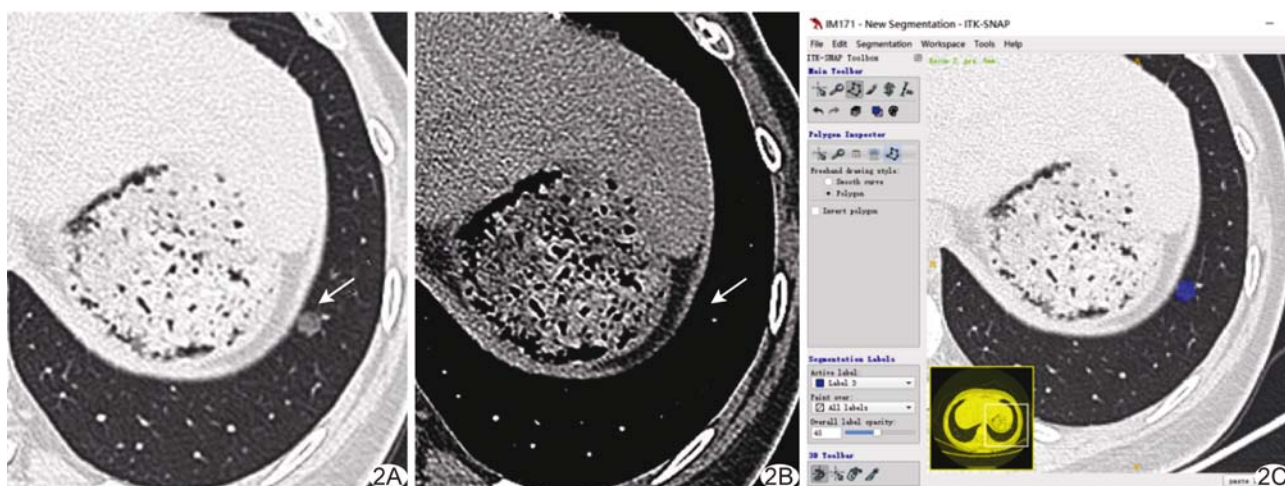


图2 患者女,42岁,左肺下叶 MIA A. 薄层肺窗; B. 薄层纵隔窗; C. 手动勾画的 ROI(蓝色区域) (箭头要勾画区域)



图3 患者女,37岁,右肺下叶 IAC A. 薄层肺窗; B. 薄层纵隔窗; C. 手动勾画的 ROI(黄色区域) (箭头要勾画区域)

1.3 数据预处理与特征提取 先将原始图像(DICOM 格式)及 VOI(NiFTI 格式)分别导入 GE A. K. (Artificial Intelligent Kit)软件进行图像预处理,包含重采样(resample)、降噪(denoising)和亮度标准化(intensity standardization);随后将预处理后图像重新导入 A. K. 软件,软件自动获取包括直方图特征(histogram features)、形态学特征(form factor features)、灰度共生矩阵特征(GLCM features)及基于灰度共生矩阵的 Haralick 特征、游程矩阵特征(RLM features)和灰度连通区域矩阵特征(GLSAZM features),共 396 个特征参数。

1.4 特征筛选与统计学分析 由 A. K. 软件内置 R 语言算法自动完成特征筛选。数据处理步骤:①以单因素方差分析检验单一变量或指标在不同分类及不同集合之间差异是否有统计学意义,对连续变量使用 Mann-Whitney U 检验或 t 检验,筛选后剩余 275 个特

征;②以 Spearman 相关性分析进行特征去冗余,剔除自相关性系数大于 0.9 ($|r| > 0.9$)者,剩余 89 个特征;③以 Lasso 回归模型通过交叉验证法进行特征降维,选出最具鉴别意义的 19 个特征。根据样本和特征之间的层级关联性绘制特征热力图,图中小方块深浅颜色对应每个特征下每个样本的 z 分数(z-score),其定义为:

$$z \text{ 分数} = \frac{x_{I,J} - x_J}{\sigma_J}$$

其中, $x_{I,J}$ 是第 I 个样本的第 J 个特征。 x_J 和 σ_J 是第 J 个特征关于所有样本的均值和标准差。

以此 19 个特征参数构建随机森林(random forest, RF)、支持向量机(support vector machine, SVM)和逻辑回归(logistic regression, LR)3 种机器学习模型。采用十折交叉验证法选择最优模型参数,确定最优模型,测试组内数据,得到各模型的准确率

(accuracy); 绘制 ROC 曲线, 评价 3 种模型鉴别 IAC 与非 IAC 的效能。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 特征参数 经过特征提取和特征筛选, 最终得到

19 个特征参数(表 1), 其权重见图 4, 灰度共生矩阵特征中的惯性矩(inertia)所占权重最大。特征热力图(图 5)显示相同特征对于 IAC 与非 IAC 有较好的区分度。

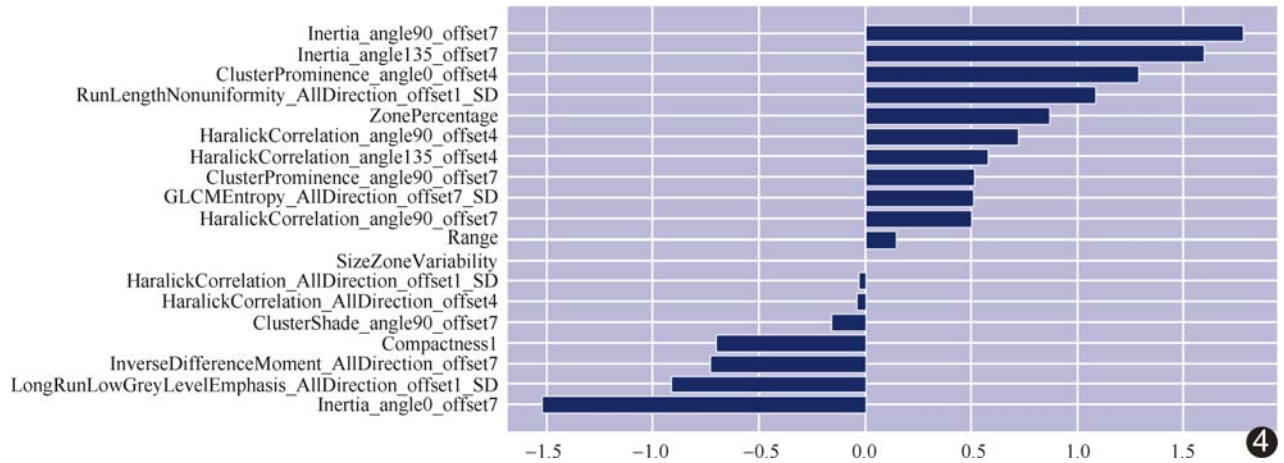


图 4 不同组学特征在机器学习模型中所占权重示意图 横轴柱体的长度代表特征所占权重, 纵轴对应不同组学特征

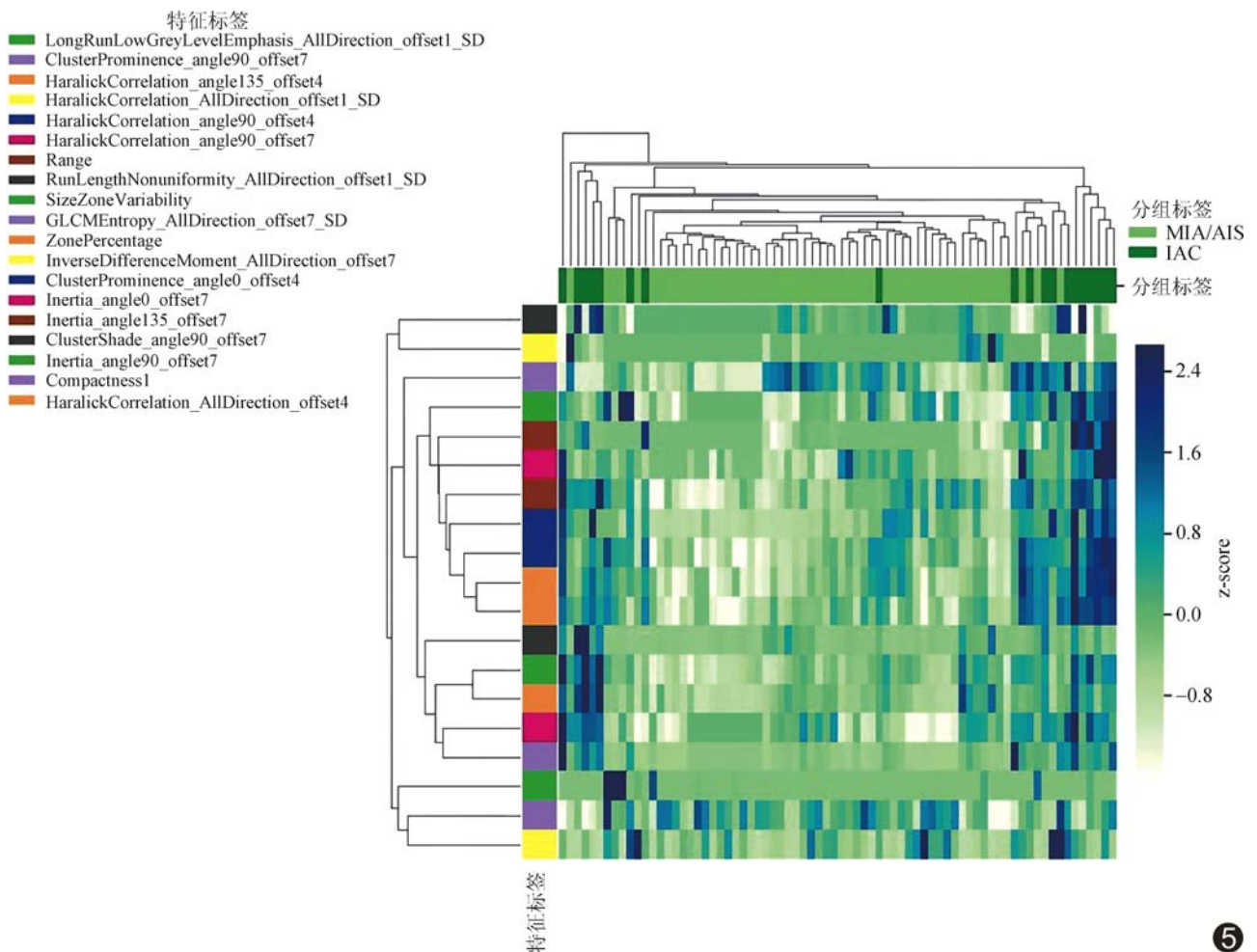


图 5 特征热力图 横轴代表不同样本, 浅绿色和深绿色分别表示 IAC 与非 IAC, 纵轴不同颜色分别代表 19 个组学特征, 图中每个小方块代表不同样本对应的不同特征, 颜色越深表示 z 分数越高

表 1 用于构建机器学习模型的 19 个影像组学特征

特征类别	特征名称
直方图特征	Range
形态学特征	Compactness1
灰度连通矩阵特征	Zone Variability
	Zone Percentage
游程矩阵特征	Long Run Emphasis_AllDirection_offset1_SD
	Run Length Nonuniformity_AllDirection_offset1_SD
纹理特征	Cluster Prominence_angle0_offset4
	Cluster Prominence_angle90_offset7
	Cluster Shade_angle90_offset7
灰度共生矩阵特征	GLCMEntropy_AllDirection_offset7_SD
	Inertia_angle0_offset7
	Inertia_angle135_offset7
	Inertia_angle90_offset7
	HaralickCorrelation_AllDirection_offset1_SD
	HaralickCorrelation_AllDirection_offset4
	Inverse Difference Moment_AllDirection_offset7
	HaralickCorrelation_angle135_offset4
	HaralickCorrelation_angle90_offset4
	HaralickCorrelation_angle90_offset7

2.2 诊断效能 3 种机器学习模型组内测试所得的准确率、AUC、敏感度及特异度见表 2, ROC 曲线见图 6。各模型之间 AUC 差异具有统计学意义(P 均 <0.05)。

表 2 3 种机器学习模型的性能比较

模型	准确率(%)	AUC	敏感度(%)	特异度(%)
SVM	93.30	0.94	87.50	95.50
RF	86.70	0.92	66.70	100.00
LR	83.30	0.83	63.60	94.70

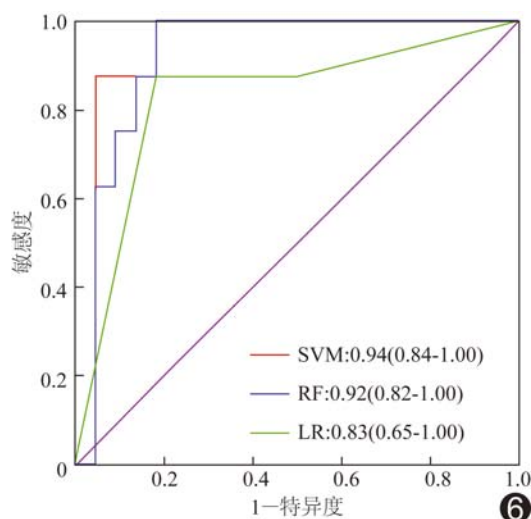


图 6 3 种机器学习模型组内测试 ROC 曲线图

3 讨论

不同于部分实性结节, pGGN 被认为是惰性生

长, 需要进行长周期随访。受 CT 扫描层厚及部分容积效应的影响, IAC 即使存在 >5 mm 间质浸润, 薄层 CT 仍可表现为 pGGN^[2], 难以与非 IAC 区分。既往研究^[5-6]认为影像组学特征对表现为 pGGN 的肺腺癌的侵袭性有较高诊断价值, 但国内外有关影像组学研究中, 鲜有基于影像组学特征构建机器学习模型鉴别表现为肺 pGGN 的 IAC 与非 IAC 的报道。本研究将 MIA 与 AIS 归为非 IAC, 且未纳入不典型腺瘤样增生, 分组更具临床意义^[2-3]。范丽等^[7-8]采用相同分组方法, 但其研究对象中包含部分实性 GGN。2017 版 Fleischner 结节处理指南^[9]指出, 偶发部分实性 GGN 较 pGGN 更倾向于 IAC, 需要密切随访; 临床工作中, 考虑患者心理压力, 对此类结节多在短期随访无变化后便进行手术治疗。对于偶发 pGGN 患者, 医师往往建议长期随访观察, 患者心理也可耐受, 故 pGGN 中尽早发现 IAC 更具临床价值。

自 2012 年 LAMBIN 等^[10]提出影像组学概念以来, 作为精准医疗中的一种辅助决策工具, 影像组学已被证实可用于诊断肿瘤、评估预后、评价疗效及监测复发^[4]。影像组学研究的最终目的是利用影像组学标签构建模型, 对病灶分类和预测。基于影像组学特征构建的机器学习模型简单易行, 是有效实用的分类工具^[11], 其用于肿瘤分级与分型的可行性已获证实。李小虎等^[12]对 34 例肾透明细胞癌患者进行 CT 容积纹理分析, 构建 RF 鉴别低级别和高级别肾透明细胞癌, AUC 为 0.88, 敏感度和特异度分别为 0.79 和 0.82, 证实基于 CT 纹理特征的机器学习模型可用于评估肾透明细胞癌的恶性度。林斌等^[13]发现 SVM 构建的预测模型可鉴别肺黏膜相关淋巴组织淋巴瘤与肺 IAC。李智慧等^[14]分析高分辨率 T2WI, 构建影像组学模型, 验证了 RF、决策树、K 近邻及 LR 4 种机器学习模型对直肠良恶性病变具有一定鉴别诊断价值, 准确率分别为 84.61%、69.23%、84.61% 和 82.05%。

本研究基于影像组学特征构建 SVM、RF、LR 3 种机器学习模型, 对于鉴别 IAC 与非 IAC 均表现出很好的分类性能, 其中 SVM 模型最优, 主要得益于 SVM 对于小样本数据具有更优秀的分类性能^[15]。在用于构建机器学习模型的 19 个影像组学特征中, 灰度共生矩阵特征占主要权重(10/19), 其中惯性矩(inertia)特征是鉴别 IAC 与非 IAC 的最重要的特征。惯性矩反映图像的清晰度和纹理沟槽的深度, 对比度与纹理凹槽深度呈正比, 凹槽越深, 产生的清晰度越高, 反之则对比度越小、图像越模糊。IAC 和非 IAC 病理机制不同, 前者存在较

多浸润成分,使部分肺泡塌陷、肺泡间隔增厚^[2],与肿瘤中未受累肺泡组织形成不同深度的纹理沟槽;而非 IAC 肿瘤细胞多是贴壁式生长^[1],密度较淡且均匀,对比度可能不高。影像组学可发挥其独特的优势,在高通量数据中发现肉眼无法分辨的规律。

本研究以 2 台不同 CT 机采集图像,并对原始图像和 ROI 进行重采样、降噪和亮度标准化 3 步预处理后提取特征,减小了个体化差异及采集误差,保证了数据的可靠性。本研究的主要不足之处在于样本量较少,采用十折交叉验证获得最优模型后,只能重复使用组内数据测试模型的性能。后续将通过多中心研究增加病例数,对模型进行组外验证,以增加结果的可信度。另外,本研究中勾画 ROI 难以避免主观性和随机性,计划引入计算机半自动勾画联合人工校准,以确保实验的可重复性,减少不必要的人力消耗。

综上所述,基于影像组学特征构建的机器学习模型可有效鉴别表现为肺 pGGN 的 IAC 与非 IAC,在不明显增多假阳性的前提下检出 IAC,术前辅助医师进行临床决策。

[参考文献]

- [1] CHANG B, HWANG J H, CHOI Y H, et al. Natural history of pure ground-glass opacity lung nodules detected by low-dose CT scan[J]. *Chest*, 2013, 143(1):172-178.
- [2] LEE H Y, CHOI Y L, LEE K S, et al. Pure ground-glass opacity neoplastic lung nodules: Histopathology, imaging, and management[J]. *AJR Am J Roentgenol*, 2014, 202(3): W224-233.
- [3] ZHANG J, WU J, TAN Q, et al. Why do pathological stage IA lung adenocarcinomas vary from prognosis: A clinicopathologic study of 176 patients with pathological stage IA lung adenocarcinoma based on the IASLC/ATS/ERS classification[J]. *J Thorac Oncol*, 2013, 8(9):1196-1202.
- [4] GILLIES R J, KINAHAN P E, HIRCAK H. Radiomics: Images are more than pictures, they are Data[J]. *Radiology*, 2016, 278(2):563-577.
- [5] 崔效楠,刘颖,叶兆祥,等.影像组学特征对肺纯磨玻璃结节侵袭性腺癌与非侵袭性腺癌的鉴别价值[J]. *国际医学放射学杂志*, 2018, 41(4):375-378.
- [6] 王亚丽,朱慧媛,毛海霞,等.影像组学特征对肺部 10 mm 以下纯磨玻璃结节侵袭性的诊断价值[J]. *中国医学计算机成像杂志*, 2017, 23(6):513-516.
- [7] 范丽,方梦捷,董迪,等.影像组学对磨玻璃结节型肺腺癌病理亚型的预测效能[J]. *中华放射学杂志*, 2017, 51(12):912-917.
- [8] 罗婷,张峥,李昕,等.CT 图像纹理分析鉴别诊断磨玻璃密度肺腺癌的浸润性[J]. *中国医学影像技术*, 2017, 33(12):1788-1791.
- [9] MACMAHON H, NAIDICH D P, GOO J M, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: From the Fleischner society 2017[J]. *Radiology*, 2017, 284(1):228-243.
- [10] LAMBIN P, RIOS-VELAZQUEZ E, LEIJENAR R, et al. Radiomics: Extracting more information from medical images using advanced feature analysis[J]. *Eur J Cancer*, 2012, 48(4): 441-446.
- [11] LANGS G, ROHRICH S, HOFMANNINGER J, et al. Machine learning: From radiomics to discovery and routine[J]. *Radiologe*, 2018, 58(Suppl 1):1-6.
- [12] 李小虎,CAI WENLI,裴子璐,等.肾脏 CT 容积纹理分析及机器学习相结合的影像组学评价肾透明细胞癌病理分级的价值初探[J]. *中华放射学杂志*, 2018, 52(5):344-348.
- [13] 林斌,王韬,沈可人,等.肺黏膜相关淋巴组织淋巴瘤与肺腺癌鉴别的影像组学研究[J]. *中华放射学杂志*, 2018, 52(10): 766-769.
- [14] 李智慧,陆建平,沈浮,等.基于磁共振 T2WI 高分辨率图像的影像组学模型对直肠良恶性病变的鉴别诊断价值[J]. *中国医学计算机成像杂志*, 2018, 24(5):388-392.
- [15] JIANG H, CHING W K, CHEUNG W S, et al. Hadamard Kernel SVM with applications for breast cancer outcome predictions[J]. *BMC Syst Biol*, 2017, 11(Suppl 7):138.