

# 影像组学技术方法的研究及挑战

史张, 刘崎

**【摘要】** 随着医疗大数据时代的到来, 影像组学成为不可或缺的大数据分析及人工智能发展技术手段, 它通过高通量特征提取算法, 对影像图像进行定量分析, 充分深入挖掘和分析隐含在图像中的额外信息, 最高效地利用影像学检查结果, 从而提供有价值的诊断、预后或预测信息, 以支持个性化的临床决策和改善个体化的治疗选择。

**【关键词】** 影像组学; 人工智能; 研究技术

**【中图分类号】** R05; R33; R445.2; R814.42; R445.3 **【文献标识码】** A

**【文章编号】** 1000-0313(2018)06-0633-04

DOI: 10.13609/j.cnki.1000-0313.2018.06.018

开放科学(资料服务)标识码(OSID)



2016年,《国务院办公厅关于促进和规范健康医疗大数据应用发展的指导意见》发布,将健康医疗大数据纳入国家大数据战略布局。实现健康医疗大数据和人工智能发展,影像医学是不可或缺的重要手段。而影像组学(radiomics)突破了基于形态学及半定量分析的传统影像医学模式,融合了数字影像信息、统计学、机器学习等方法,采用了高通量特征提取算法,对影像图像进行定量分析,充分深入挖掘和分析隐含在图像中的额外信息,最高效地利用影像学检查结果。本文拟综述影像组学技术方法的研究现状及挑战。

## 影像组学的概念

从影像图像中提取的全部特征就是“影像组学”,而通过特征选择后所挑选出的那些具有预测价值的特征集合通常被称为“影像组学标签(radiomic signature)”<sup>[1]</sup>。影像组学源于计算机辅助检测或诊断(computer-aided detection or diagnosis, CAD),是将影像定量分析与机器学习方法结合起来<sup>[2]</sup>。

目前,影像组学的基本作用是通过大量的影像组学特征对肿瘤感兴趣区进行定量分析,从而可以提供有价值的诊断、预后或预测信息。影像组学的目的是探索和利用这些信息资源来开发诊断、预测或预后的影像组学模型,以支持个性化的临床决策和改善个体化的治疗选择<sup>[3]</sup>。

## 影像组学的工作流程

影像组学的工作流程分为图像采集、图像分割、图

像特征提取和量化、特征选择和建立模型5个阶段。

### 1. 图像采集

目前,主要通过CT、MRI和PET-CT等影像扫描方式来进行图像的采集。CT是影像组学研究中使用的最广泛的成像模式,其具有空间分辨率高的影像特点,可评估肿瘤和淋巴结的组织密度、形状、质地及纹理特征<sup>[4,5]</sup>。PET-CT可同时获得组织的密度及代谢信息,常用于肿瘤的检测和分期,是影像组学将功能成像与潜在肿瘤生物学行为直接进行关联的一种研究手段。MRI在软组织成像中表现突出,可提供高对比度的结构信息和功能信息,其中弥散加权成像(DWI)和动态对比增强磁共振成像(DCE-MRI)可以反应组织细胞结构及微血管生成情况,通过对这些图像的采集,可提取更有效的影像组学特征。

### 2. 图像分割

图像分割是指对感兴趣区部位的分割,也就是在影像图像上勾画出感兴趣区域,从而针对这一特定区域计算出影像组学特征。目前,图像分割的方法有人工分割法、半自动分割法及自动分割法3种。

手动分割法被应用于大多数影像学研究中,其优势在于准确度高并且对不规则的肿瘤边界勾画精细,但其受主观因素影响较大,可重复性低,且耗时久、效率低。

相反,自动或半自动分割法则表现出较高的可重复性和时效性,其中半自动分割法为目前影像组学图像分割的主要方法。例如,Mattonen等<sup>[6,7]</sup>对肺癌的研究中指出先由研究者根据实体肿瘤反应评估标准(the Response Evaluation Criteria in Solid Tumors, RECIST)测量肿瘤最长轴径,然后再用自动算法去勾勒实体肿瘤的区域,从而完成图像分割。

与半自动分割法相比,自动分割法可以实现排除人为因素,更好地达到自动化、可重复且效率高的效

作者单位:200433 中国人民解放军海军军医大学第一附属医院放射科

作者简介:史张(1987—),男,陕西渭南人,硕士研究生,住院医师,主要从事心脑血管疾病影像诊断与评估工作。

通讯作者:刘崎, E-mail:liuqimd@126.com

基金项目:国家自然科学基金(81670396; 31470910; 31600755; 81701689)资助

果。目前,应用于影像组学的完全自动分割法还没有统一的方案及标准,但是自动分割技术已经初有成效,其中,CAD自动分割乳腺肿瘤已应用十余年,其分割结果的准确性跟人工手动分割相比几无差别<sup>[8]</sup>。这些都表明实现感兴趣区的自动分割法将是未来影像组学图像分割的一个重要研究方向。

### 3. 图像特征提取和量化

影像组学特征可以分为形状特征、一阶直方图特征、二阶直方图或纹理特征。还有一些获取于特定图像的影像组学特征(如PET中的SUV度量),以及仅适用于多模式数据集的分形和融合特征。

**形状特征:**包括描述ROI大小的特征,例如体积、表面积、二维和三维的最大直径以及有效直径(与ROI具有相同体积的球体直径),以及描述ROI与球体的相似程度的特征,如表面体积比、致密度、偏心度、球形度等。

**一阶直方图特征:**描述与ROI内的体素强度分布有关的特征,不包含它们之间相互的空间作用,可通过直方图分析计算得到,包括均数、中位数、最小值、最大值、标准差、偏度和峰度。这些特征可反应所测体素的对称性、均匀性以及局部强度分布变化。

**二阶直方图或纹理特征:**是描述体素空间分布强度等级的特征。图像纹理是指在强度水平可感知或可测量的空间变化,它被视为一个灰度级,是一种视觉感知的图像局部特征的综合<sup>[9]</sup>。二阶特征包括:灰度共生矩阵(gray level co-occurrence matrix, GLCM)、灰度级长矩阵(gray level run-length matrix, GLRLM)、灰度级带矩阵(gray level size zone matrix, GLSZM)和邻域灰度差分矩阵(neighborhood gray-tone difference matrix, NGTDM)。

GLCM是一个其行列数表示灰度值、单元格包含灰度值处于一定关系(角度、距离)次数的矩阵,也称为二阶直方图。在GLCM上计算的特征包括熵(二阶熵,与异质性有关)、能量(也被定义为角二次矩,再次描述图像的均匀性)、对比度(其测量局部变化)、同质性(图像局部灰度均衡性的度量)、不相似性和相关性。

GLRLM是二维矩阵,其中每个元素 $(i, j)$ 描述了灰度级 $i$ 在指定的方向上连续出现的次数<sup>[10]</sup>,灰度运行是在图像中预设方向上具有相同强度的连续体素的长度。

GLSZM是在行和列处的元素存储具有灰度级和大小的区域(具有相同灰度级的连接体素)数量的矩阵。GLSZM包括描述小/大区和低/高灰度区分布的特征<sup>[11]</sup>。

NGTDM第 $i$ 项是所有具有灰色调 $i$ 的像素与其周围邻域像素平均值之差的总和。NGTDM的影像

组学特征包括粗糙度、对比度、冗繁度、复杂度、纹理强度等。

**融合和分形特征:**融合特征与多模态图像数据集相关,其可通过配准技术与几何图像对齐。分形特征是分析评估不同层面表面的自相似性和粗糙度,这些区域的复杂性由Hausdorff的分形维(fractal dimension, FD)量化呈现,该分形维是一个模式的自我重复纹理,并具有放大特征<sup>[12]</sup>。

### 4. 特征选择

如何将高维特征数据进行降维是影像组学工作流程中的重要一步,也称之为特征选择。最简单的特征选择方法是根据变量的稳定程度或相关性制定一个评分标准,以此标准对变量进行筛选。另外,在消除相关特征时,使用相关矩阵消除高度相关的特征,通过消除那些高度相关的特性,剩下“非冗余”的特征集,其中常用的方法有LASSO(least absolute shrinkage and selection operator)Cox回归模型、最大相关最小冗余(maximum relevance and minimum redundancy, mRMR)、评估特征RE相关性(relevance in estimating features, RELIEF)、主成分分析法(principal component analysis, PCA)等。

LASSO Cox回归模型,是最常用的特性选择方法之一<sup>[13]</sup>,它是一种收缩和变量选择方法的回归模型,它使罚对数似然函数最大化,并适用于高维数据的回归。

mRMR通过计算一组特性和结果变量之间的相互信息(mutual information, MI),对输入的特征进行排名,最大化MI、最小化更高排名的MI平均值,从而达到降维的效果。

RELIEF是最近比较热门的特征选择方法之一,其方法是根据属性之间的区分进行排名。与其他方法相比,RELIEF可以有效地评估强相互依赖的特征,并显示出更高的预测准确性。RELIEF算法能够检测特征之间的上下关联信息,从而更准确地处理存在依赖关系的情况。

PCA是将多个变量通过线性变换以选择出少量重要变量的一种多元统计分析方法,即能将相关性强的影像组学特征合并为主成分,各主成分间相互独立从而实现将高维空间简化为二维或三维空间。

### 5. 建立模型

目前,有许多机器学习的方法可被用于建立基于影像组学特征的预测和分类模型,其中很多都是之前服务于CAD的。在影像组学建模中,logistic回归模型因其简单易行,成为最受欢迎且常用的监督分类器;另外,常用的机器学习模型还有随机森林(random forest)、支持向量机(SVM)、人工神经网络(Artificial neural networks, ANNs)、聚类分析(clustering analy-

sis)、留-法交叉验证(leave-one out cross validation, LOOCV)、自举法(bootstrapping)。

**随机森林:**是基于机器学习中一个较普遍的决策树概念,将假设表示为连续的“如果-然后”,这一点很类似于人类推理。在该方法中,训练一组决策树,并且该算法引入两个级别的随机化。所有的决策树被训练后,对所有单株树进行预测,并将选出最高频的一类作为最终结果<sup>[12]</sup>。

**SVM:**是一种先前用于CAD的有辨识监督机器学习的技术,它利用边界将数据点分为两类(如反应者或无反应者),该方法在CAD微钙化检测中有较高的判别力,并经常用于影像组学。

**ANNs:**是经典的机器学习方法之一,也被用于基于超分类的基因组学研究。卷积神经网络(Convolutional Neural Network, CNN)是一种特殊的前馈神经网络,用于学习直接对原始图像进行操作的图像数据的分层表现,通过尝试自动提取高度表现的成像特征, CNN 消除了对图像预处理和特征选择的依赖性。该方法在MRI对三阴乳腺癌的判别中要优于其他预测模型<sup>[14]</sup>。

**聚类分析:**属于无监督分类(unsupervised classifiers),其可将抽象对象集合分组,在分组的过程中对类似的对象组成的多个类进行分析,目标是在相似的基础上收集数据来进行分类。共识聚类(consensus clusterin)是常用于降维的聚类分析。有报道称共识聚类可以在440个特征的集合空间中识别出13个非冗余特征聚类<sup>[15]</sup>。

**LOOCV:**是用于内部验证的最常用的技术,也称为“jackknife”。该算法除了一个数据点被用于测试外其余所有数据均用于训练或拟合,这个过程在每次LOOCV迭代中重复,使得每个数据点只剩下一次。

**自举法(bootstrapping):**由原始数据生成大量自举样本,引导数据集是一系列数据(特征,结果),每个数据来自患者队列中随机选择的患者,每个自举样本重复建模。该方法提供了模型参数和模型评估指数AUC或CI值的分布,从而可以评估特征的不确定性。

## 影像组学的质量评估

随着精准医疗的提出和发展,近年来影像组学的研究逐年增多,因其提取特征及建模的方法和形式具有多样性,而有研究表明目前影像组学所建立的预测模型质量不佳<sup>[16]</sup>,因此,迫切需要制定统一的评估标准和指南,使影像组学成为一个科学、严谨、可评判的研究领域。2017年10月,在Nature Reviews Clinical Oncology上发表的文章中提到影像组学质量评分(radiomics quality score, RQS)及其标准<sup>[17]</sup>,从而帮

助我们科学评估之前做过的以及未来将要进行的影像组学研究。RQS共有16个项目,每个项目下方均有2~3个选项,测试者只需要勾选即可,满分36分。RQS标准对预测模型的所有方面都需要进行全面和清晰的评估和打分,以尽量减少偏差,从而提高预测模型的实用性。文章中指出,目前RQS可以在线进行打分,并建议进行影像组学研究时应通过RQS的评估。

## 影像组学面临的挑战

影像组学尚处于起步阶段,虽然在临床疾病的诊断、疗效监测及预后评估中都取得了一些成果<sup>[18]</sup>,但是影像组学依然有一些局限和不足,工作流程中也有可能改进的地方。

### 1. 可重复性

可重复性是在相同或几乎相同的条件和采集参数下的精度测量,并通过“测试-重新测试”分析进行评估,比较对同一患者采集图像的结果<sup>[19]</sup>。有研究表明,在相同成像参数设置和半自动分割下获得的影像组学特征可重复性较高(一致性指数 $>0.9$ )。相反,当测量系统或参数设置不同时,测量的重复性或稳定性欠佳。

对于CT,通过比较从17台不同类型CT获取的水模图像中提取影像组学特征间的变异度, Mackin等<sup>[20]</sup>发现扫描仪间的变异度与同一台扫描仪的相比是有差异的。有研究认为影像组学特征的变异主要是由重建方法所引起的,重建算法的变异度显著高于测量者之间的变异度<sup>[21]</sup>。在PET-CT/MR中,纹理特征受不同采集模式、重建算法及所用参数设置的影响较大<sup>[22]</sup>,例如迭代次数、后滤波级、输入数据的噪声、矩阵大小以及离散块大小等。

分割代表了在基因组工作流程中最关键的步骤之一,因为许多提取的特征可能取决于分割的区域,而肿瘤可能有不清楚或复杂的边界,这可能会导致结果的不一致性和低可重复性。研究表明尽管手动勾画分割出现观察者间的差异度较高,且耗时久,但它仍是金标准。也有研究结果显示,基于3D-Slicer软件平台,采用半自动分割法对CT扫描肺肿瘤的分割比手动勾画的区域可重复性更好<sup>[23]</sup>。

### 2. 样本量

影像组学是大数据时代的产物,充足的特征数据及构建数据库是影像组学研究前提。目前,许多影像组学研究大多是小样本量的研究<sup>[24]</sup>,而样本量不足所产生的小数据集会降低模型预测准确率并增加过度拟合的风险。

### 3. 统计与数据库

大多数影像组学研究并没有在独立队列研究中得

到充分验证,从而患者人群的普适性不高。目前,只有一个小组对其影像组学模型进行了外部验证<sup>[24]</sup>。多中心图像数据的共享可以成为构建大数据库的解决方案,并可作为外部验证的高质量数据库。而多学科合作建立影像组学数据库将是后续努力的方向,而建成的数据库可用于影像组学和基因表型数据的存储、检索和分析。

#### 4. 标准化

通常,大多数影像组学研究使用的图像是从多个研究机构的各种扫描方案或来自不同供应商的扫描仪。我们认为,需要通过扫描方案及重建算法的标准化来降低输入数据的变异度,尤其是多中心的研究更应如此。建议从影像组学研究开始就设定好一个方案,并按照 RQS 标准严格把控,以实现标准化,从而提高研究质量。

#### 参考文献:

- [1] Kotrotsou A, Zinn PO, Colen RR. Radiomics in brain tumors: an emerging technique for characterization of tumor environment[J]. *Magnetic Resonance Imaging Clinics of North America*, 2016, 24(4):719-729.
- [2] Parekh V, Jacobs MA. Radiomics: a new application from established techniques[J]. *Expert review of Precision Medicine and Drug Development*, 2016, 1(2):207-226.
- [3] Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data[J]. *Radiology*, 2016, 278(2):563-577.
- [4] Giesel FL, Schneider F, Kratochwil C, et al. Correlation between SUVmax and CT radiomic analysis using lymph node density in PET/CT-based lymph node staging[J]. *J Nuclear Medicine: Official Publication, Society of Nuclear Medicine*, 2017, 58(2):282-287.
- [5] Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach[J]. *Nature Communications*, 2014, 5:4006. DOI: 10.1038/ncomms5006. PMID:24892406
- [6] Mattonen SA, Palma DA, Johnson C, et al. Detection of local cancer recurrence after stereotactic ablative radiation therapy for lung cancer: physician performance versus radiomic assessment[J]. *International J Radiation Oncology, Biology, Physics*, 2016, 94(5):1121-1128.
- [7] Mattonen SA, Tetar S, Palma DA, et al. Imaging texture analysis for automated prediction of lung cancer recurrence after stereotactic radiotherapy [J]. *J Medical Imaging (Bellingham, Wash)*, 2015, 2(4):041010.
- [8] Li H, Zhu Y, Burnside ES, et al. MR imaging radiomics signatures for predicting the risk of breast cancer recurrence as given by research versions of mammaPrint, oncotype DX, and PAM50 gene assays[J]. *Radiology*, 2016, 281(2):382-391.
- [9] Gnep K, Fargeas A, Gutierrez-Carvajal RE, et al. Haralick textural features on T<sub>2</sub>-weighted MRI are associated with biochemical recurrence following radiotherapy for peripheral zone prostate cancer[J]. *J Magnetic Resonance Imaging: JMIR*, 2017, 45(1):103-117.
- [10] Wu W, Parmar C, Grossmann P, et al. Exploratory study to identify radiomics classifiers for lung cancer histology[J]. *Frontiers in Oncology*, 2016, 6(71):1-11.
- [11] Lian C, Ruan S, Denoex T, et al. Selecting radiomic features from FDG-PET images for cancer treatment outcome prediction [J]. *Medical Image Analysis*, 2016, 32:257-268. DOI:10.1016/j.media.2016.05.007. PMID:27236221
- [12] Yang D, Rao G, Martinez J, et al. Evaluation of tumor-derived MRI-texture features for discrimination of molecular subtypes and prediction of 12-month survival status in glioblastoma[J]. *Medical Physics*, 2015, 42(11):6725-6735.
- [13] Huang YQ, Liang CH, He L, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer[J]. *J Clinical oncology: Official J the American Society of Clinical Oncology*, 2016, 34(18):2157-2164.
- [14] Ypsolantis PP, Siddique M, Sohn HM, et al. Predicting response to neoadjuvant chemotherapy with PET imaging using convolutional neural networks[J]. *PLoS One*, 2015, 10(9):e0137036.
- [15] Paimar C, Leijenaar RT, Grossmann P, et al. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer[J]. *Scientific Reports*, 2015, 5(11044). DOI: 10.1038/srep11044. PMID:PMC4937496
- [16] Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement[J]. *Br J Surgery*, 2015, 102(3):148-158.
- [17] Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine[J]. *Nature Reviews Clinical Oncology*, 2017, 14(12):749-762.
- [18] 李振辉, 丁莹莹. 积极推动影像组学的临床应用研究[J]. *放射学实践*, 2017, 32(12):1233-1214.
- [19] Hatt M, Tixier F, Pierce L, et al. Characterization of PET/CT images using texture analysis: the past, the present any future? [J]. *Eur J Nuclear Medicine and Molecular Imaging*, 2017, 44(1):151-165.
- [20] Mackin D, Fave X, Zhang L, et al. Measuring computed tomography scanner variability of radiomics features[J]. *Investigative Radiology*, 2015, 50(11):757-765.
- [21] Kim H, Park CM, Lee M, et al. Impact of reconstruction algorithms on CT radiomic features of pulmonary tumors: analysis of intra- and inter-reader variability and inter-reconstruction algorithm variability[J]. *PLoS One*, 2016, 11(10):e0164924.
- [22] Desseroit MC, Tixier F, Weber WA, et al. Reliability of PET/CT shape and heterogeneity features in functional and morphologic components of non-small cell lung cancer tumors: a repeatability analysis in a prospective multicenter cohort[J]. *J Nuclear Medicine*, 2017, 58(3):406-411.
- [23] Rios Velazquez E, Aerts HJ, Gu Y, et al. A semiautomatic CT-based ensemble segmentation of lung tumors: comparison with oncologists' delineations and with the surgical specimen[J]. *Radiation Therapy and Oncology*, 2012, 105(2):167-173.
- [24] Napel S, Giger M. Special section guest editorial: radiomics and imaging genomics: quantitative imaging for precision medicine [J]. *J Medical Imaging (Bellingham, Wash)*, 2015, 2(4):041001.