

## 【影像组学】医学影像组学面临挑战和发展机遇

原创 2017-09-30 主作者: PKRC e学e用



赵周社 李昕 徐潇 杨俊 张陈

GE药业PKRC团队

医学影像组学 (Radiomics) 是指采用高通量和自动化计算方法将医学图像中病灶感兴趣区域 (region of interest, ROI) 的影像数据转化为具有特征空间的定量化数据, 并对特征空间区数据进行精准定量化分析, 以便从结构形态、组织细胞、分子和基因遗传等各种表型 (phenotype) 综合起来达到提高对疾病诊断效能、对疾病疗效进行精确评价和对疾病预后进行预测<sup>[1]</sup>。PET和SPECT图像是使用特异性放射性示踪剂(radioactive tracer), 获得组织细胞酶、受体和基因表达的定量化信息, 它具有对病灶组织和研究对象组织进行活体“组织和分子分型”的特点。CT、MRI和超声波图像是在解剖结构基础上, 提供功能、血流灌注和部分组织细胞代谢的信息。与PET和SPECT相比较, 影像组学技术更适合于对CT、MRI和超声波图像进行研究, 挖掘现在通过视觉阅片或常规定量化技术仍然没有发现的组织细胞特征。对于以解剖结构为基础的医学影像, 采用转化医学(translation medicine)理念、影像组学技术为工具, 实现精准医学(precision medicine)目标。影像组学技术已经被用于肿瘤、神经和心血管系统疾病的早期诊断、疗效评估研究中, 并且已经积累一定的

经验。但是，我们也需要看到，医学影像组学技术本身与分子影像（molecular imaging）技术、基因组学(genomics)、蛋白质组学(proteomics)和代谢组学(metabonomics)研究的方法存在本质的不同，从近期已经发表的资料中发现在采用影像组学技术对疾病研究过程中存在缺乏规范化、标准化流程和严格的质量控制体系，对于特征提取缺乏“金标准”支撑，建立预测模型理论和方法有待于进一步完善，对于影像组学研究的领域缺少行业专家的共识[2,3,4,5,6]。为此，本文对影像组学面临的挑战和发展机遇进行讨论。

## 1. 医学影像组学技术面临的挑战

Chalkidou A 等学者在“False Discovery Rates in PET and CT Studies with Texture Features: A Systematic Review”资料中系统性回顾性分析现在影像组学采用CT和PET图像研究中存在问题和技術缺陷<sup>[4]</sup>。目前采用的研究方法对于CT和PET图像纹理分析所致的“Type I error”是无法避免，导致过高估计预测效能。Larue RT资料中详细介绍影像组学研究过程从数据采集、病灶分割(segmentation)、特征提取(feature extraction)、建模(model building)和预测过程精准定量方法存在问题和挑战<sup>[5]</sup>。不同设备、同一设备不同扫描参数，使用或未使用对比剂等均对分析结果有严重的影响。Chen JH 和Asch SM发表的论文“Machine Learning and Prediction in Medicine — Beyond the Peak of Inflated Expectations”<sup>[6]</sup>，更是从不同角度提出在影像组学研究过程使用机器学习技术和方法应该高度重视的问题。为此，我们从以下几个方面介绍医学影像组学技术在实际研究应用中面临的挑战。

**(1) 影像组学中纹理特征基本生物学含义不清楚。**图像纹理具有其基本纹理单位，也被称为纹理基元（texton）<sup>[2]</sup>。人体组织是由组织细胞(cell)构成。然而，影像组学中最小的基本单元与人体组织结构之间到底存在什么关系？现在没有看到明确解释或给出可以接受的定义。在这种情况下，医学影像组学和建立在其基础上的人工智能(artificial intelligence, AI)研究急需将正常组织影像组学特征和正常组织解剖结构之间的关系搞清楚。此外，对比剂对纹理分析结果影响也没有得到充分的、高度认识。在基础理论不清楚、技术方法没有得到规范情况下，那么现在获得任何所谓“有意义”的研究结果都需要重新思考和认真的对待。也只有知道和掌握了影像组学中基本特征所表达的生物学含义，以及关键的影响因素后，才能推动影像组学的研究进

入一个新的台阶。简单的讲，医学影像组学的最基本单元现在还没有搞清楚，急需将最基本的理论问题和基本概念搞清楚。

**(2) 缺乏规范化、标准化流程和严格的质量控制体系。**影像组学采用定量化技术提取病灶的特征，对于任何影响定量化研究的技术因素都应该具有一套严格的、规范化、标准化的流程和质量管理体系。比如，图像采集模式（2D/3D）、图像滤波函数、图像重建方法（FBP/迭代重建）、是否使用对比剂等， 这些因素都直接影响特征提取和分析的结果。表1是对CT、MRI和PET图像进行定量化研究前首选需要考虑的部分影响因素。当选择不同的扫描方式、图像重建算法和病灶提取方法时，获得的结果之间存在明显的差别，甚至于出现相反的结果。比如，在平扫描、动脉期或静脉期勾画病灶的ROI分割出的病灶体积可以相差一倍以上。临床经验证实，采用增强扫描的图像能够获得精确的病灶边界，明显提高病灶组织分割的精度。由于影像组学技术在医学领域研究应用的时间相对较短，迄今国际上仍然没有建立起影像组学研究的规范化流程，缺乏严格的质量控制体系。目前，报道的研究结果多数是回顾性研究，这样在数据收集、病灶的提取（分割）、特征提取均没有严格的流程和可以参照的标准。导致不同作者、不同研究中心、即使同一个研究中心采用不同的成像设备报道的数据之间差别很大。在影像组学研究中急需对流程、操作进行规范，并建立共识的“参考标准”，以便进行数据的对比研究。

表1 影响三种医学影像图像定量化研究结果部分因素

影响因素		CT	MRI	PET
扫描参数	扫描模式	螺距、mA、Kv	不同序列，T1/T2	动态、静态、门控
	扫描选择矩阵	512/256	512/340/192	256/192
	对比剂/放射性药物	是否增强扫描，动脉、静脉	是否增强扫描，动脉、静脉	不同PET示踪剂
	特殊扫描模式	轴扫	ZTE/Magic	TOF，PSF
介入试验		有/无	有/无	有/无
图像重建	重建方式	迭代类型	K空间填充方式	迭代重建参数选择
	滤波函数	滤波函数	滤波函数	滤波函数+点扩展函数
	矩阵	512/256	512/340/192	256/192
病灶分割	病灶分割方法	自动、半自动、手动	自动、半自动、手动	自动、半自动、手动
	病灶选择	2D/3D	2D/3D	2D/3D

备注：TOF（time of flight）飞行时间技术；ZTE（zero Echo time）零回波时间；Magic 一次扫描获得多幅图像。

**（3）误认为获得图像特征越多越好。**影像组学技术研究中多数选择纹理特征（texture feature）统计分析方法。纹理的统计分析可以在空域或频域进行。在图像空域，基于统计的方法包括：矩、自相关函数、灰度共生矩阵(gray level co-occurrence matrix, GLCM)、和灰度游程矩阵（gray level run-length matrix, GLRLM）等；在频域有频谱分析方法，如傅里叶域变换（Fourier transform），高斯变换（gaussian-transform），小波变换（wavelet-transform）等。这些不同的方法都是从不同角度来提取病灶的特征，提取出的特征之间存在明显重复性，不仅仅是增加工作量，更严重的是要面对去冗余(redundancy)问题。目前，去冗余仍然没有统一的参考标准，存在可怕的随意性。这样有可能将真正有价值的组织特征给删除掉。在不同类型的纹理提取方法之间，在去冗余过程中如何对数据进行标准化处理，也是一个很大问题。由此可见，在没有获得公认的“标准”前提下，认为纹理特征越多越好的思路显然是一种概念和技术的误导。尤其是有些研究其图像特征数量远远大于研究的样本数，尽管获得结果看似很好，但是是否具有临床价值值得深入研究。

**（4）组织提取特征缺乏客观的分子基础“金标准”。**表2是不同影像技术、组学技术研究采用的参考标准。可以看出，无论是传统的肿瘤组织学分型，还是精准医学要求的分子分型，以及基因组学研究都有严格的“金标准”或“参考标准”。但是，迄今我们从图像特征中提取有价值的特征的标准是什么？这些有价值或有效地特征所代表的生物学意义是什么？从报道的文献来看，基本是肿瘤良性（“0”）或恶性（“1”），转移（“1”）或无转移（“0”），基因表达（“1”）或无表达（“0”），或生存时间等。影像组学研究中将众多的组织纹理特征与分类结果“0”或“1”进行回归建模，但是我们却并不知道每一个组织纹理特征所代表的真正生物学意义。目前采用预测模型实质上也是一个分类器（classifier），并采用这个分类器对病灶进行良性或恶性分类。然而，这种分类与病变组织的分子分型之间并无实质性关系，也就是讲尽管采用先进的计算机图像处理、纹理分析和AI技术，但是其参考标准仍然是组织病理，基于组织病理推测的生物标志物变化缺乏理论依据。

尽管可以采用反向传播算法（Backpropagation）、深度置信网络（Deep Belief Nets, DBNs）到现在卷积神经网络（Convolutional Neural Networks, CNNs），循环神经网络（Recurrent Neural network, RNN）等技术来提高分类的准确性，但是对于缺乏“金标准”情况



下的任何再好的分类技术也是无济于事的。现在的医学影像组学技术和医学领域的AI被一些所谓“专家”误导其发展的方向。

表2 影像技术和组学技术标准

研究项目	核心技术	参考标准
医学影像临床诊断效能	CT、MRI、超声波、PET、SPECT 等	组织病理结果
分子影像技术	探针探针	免疫组化
基因组学	分子探针	基因序列
蛋白质学	分子探针	质谱、免疫荧光技术
影像组学	纹理分析	？

**(5) 缺乏标准的建模方式，获得远离实际的预测结果。**现在采用的监督学习（supervised learning）建模方法是将标记或标注的数据分成训练组和验证组数据，采用训练组数据建立模型，然后对验证组数据进行验证。采用这种方式获得的结果确实明显提高了诊断效能或预测结果的准确性，但是我们需要考虑的是这种研究模式是否存在“用自己的左手来证明自己的右手解剖结构和形状”？大数据和影像组学技术为我们提供挖掘疾病机理、精确预测疾病预后的平台，但是我們也需要清楚意識到疾病診斷已經進入到精準醫學的分子分型階段，不能簡單地把理工領域使用的模型簡單的搬運到精準醫學研究領域。我們更需要考慮的是採用所謂的影像組學和人工智能技術對肺結節病灶進行良性或惡性預測，其真正的臨床價值在哪裡？這種“良性”或“惡性”診斷與肺癌精準治療需要知道的EGFR、Kras、C-Met和VEGFR等表達之間有什麼關係？所以，我們需要慎重考慮如何在分子分型基礎進行建模和進行人工智能化分析研究。



2. 医学影像组学技术发展机遇

尽管医学影像组学技术在实际应用中存在很多急需解决的问题，但是作为一种能够自动提取人视觉不能发现的影像特征的技术，在以解剖结构为基础上（CT、MRI和超声波）图像研究

中具有重要的价值。只要我们正确认识、系统掌握影像组学的技术和方法，那么必将对基于解剖结构基础上的图像研究起到推动作用。

**(1) 推动从常规的解剖、功能和代谢影像到分子分型发展。**尽管PET和SPECT图像能够显示特异性信息。但是，现在常规的解剖、功能和代谢图像（CT、MR和超声波）无法提供特异生物标志物信息，这就需要我们采用影像组学技术从这些医学影像中提取组织特征信息，并提供这些影像组学特征与生物标志物之间的独特关系。这样就能为临床提供更有价值的信息，推动影像技术从诊断到指导个体化治疗方向发展，推动精准医学朝着普及化方向发展。

**(2) 加速从影像组学到影像基因组学(radiogenomics)的发展。**为了更好适应精准医学发展，就需要我们加速推动从影像组学到影像基因组学发展。疾病发生与其致病基因之间存在内在的、必然的联系。在影像组学研究阶段可以采用免疫组化作为对照研究，但是研究的目标是组织影像特征与表型基因之间的关系。所以，我们需要加速从影像组学到影像基因组学的发展。影像基因组学是影像组学研究的方向和目标。

综上所述，影像组学在提取影像组织特征具有重要的价值。在进行影像组学研究时需要高度注意研究过程的规范方法、选择客观的标准进行对照，建立科学的模型实现对疾病智能化诊断和治疗效果的预后的预测。加速从影像组学到影像基因组学发展。

## 参考文献

- 1.Sala E, Mema E, Himoto Y,et al. Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. Clinical Radiology 2017;72:3-10
2. Depeursinge A, Foncubierta-Rodriguez A and Van De Ville D, et al. Three-dimensional solid texture analysis in biomedical imaging: Review and opportunities. Medical image analysis. 2014;18: 176-196.
3. 董永生 著 图像纹理分类理论与方法 科学出版社 北京 2017
4. Chalkidou A, O'Doherty MJ and Marsden PK. False Discovery Rates in PET and CT Studies with Texture Features: A Systematic Review. PLoS ONE 10(5): e0124165. doi:10.1371
5. Larue RT, Defraene G and De Ruysscher D,et al. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. Br J Radiol 2017; 90: 20160665
- 6.Chen JH and Asch SM. Machine Learning and Prediction in Medicine — Beyond

the Peak of Inflated Expectations. n engl j med 2017; 376;26

【影像组学】精彩回顾：

【影像组学】影像组学研究进展介绍

欢迎点击左下角的“阅读原文”，进入e学e用对比剂学院，发现更多精彩~

同时，我们期待您的留言，任何您感兴趣的技术话题或其他宝贵建议，都请发送给我们吧（一脸期待）~



全线学院，百花齐放  
专业课程，海纳百川  
直播活动，精彩纷呈  
名院专栏，高手论道  
论坛切磋，百家争鸣

e  
学  
e  
用

公众号：e学e用  
网站：<http://edu.gehealthcare.cn>  
首页|学院|活动|医院|论坛



阅读原文