

影像组学介绍

冯源(译) 兰晓莉(审校)

华中科技大学同济医学院附属协和医院核医学科、分子影像湖北省重点实验室、生物靶向治疗教育部重点实验室,武汉 430022

本文首次发表在 *The Journal of Nuclear Medicine*, 2020, 61(4): 488-495

Introduction to radiomics

Marius E. Mayerhoefer^{1,2}, Andrzej Materka³, Georg Langs², Ida Häggström⁴, Piotr Szczypiński³, Peter Gibbs¹, Gary Cook^{5,6}

¹Department of Radiology, Memorial Sloan Kettering Cancer Center, New York, New York; ²Department of Biomedical Imaging and Image-Guided Therapy, Medical University of Vienna, Vienna, Austria; ³Institute of Electronics, Lodz University of Technology, Lodz, Poland; ⁴Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, New York; ⁵Cancer Imaging Department, School of Biomedical Engineering and Imaging Sciences, King's College London, London, United Kingdom; ⁶King's College London and Guy's and St. Thomas' PET Centre, St. Thomas' Hospital, London, United Kingdom

Corresponding author: Marius Mayerhoefer, Email: mayerhom@mskcc.org

原文 DOI:10.2967/jnumed.118.222893

【摘要】 影像组学是 1 个发展迅速的研究领域,涉及医学影像中的定量指标,即影像组学特征的提取。影像组学的特点是捕获组织和病变的特征,如异质性、形状等,并可单独或与人口统计学、组织学、基因组学或蛋白质组学等数据相结合用于解决临床问题。该文旨在提供该领域的介绍,内容涵盖基本的影像组学工作流程,包括特征的计算与选择、降维及数据处理;并讨论核医学的潜在临床应用,包括对基于 PET 影像组学的治疗反应和生存预测;还介绍当前影像组学的局限性,如对于所采集参数变化的灵敏度,以及其他常见问题。

【关键词】 影像组学;人工智能;机器学习;PET;单光子发射断层扫描

DOI:10.3760/cma.j.cn321828-20211130-00427

影像组学在核医学和一般医学成像领域中被广泛讨论。尽管这个术语并没有严格的定义,但影像组学通常是指从诊断图像中提取定量的、理想化可再现的信息,包括人眼难以识别或量化的复杂模式^[1-2]。从代谢显像技术(如 PET 和 SPECT)中提取出的影像组学特征在多种情况下可能会有用。

第一,影像组学可以用来捕获组织和病变的性质,如形状和异质性,以及它们在治疗或监测期间连续成像的变化。在肿瘤学中,对组织异质性的评估尤为重要:基因组学分析表明,肿瘤异质性的程度是影响生存预后的决定性因素,也是癌症得到控制的阻碍之一^[3-6]。研究表明,影像组学特征和细胞水平的异质性指数密切相关^[7-8]。活组织检查(简称活检)只能捕获到肿瘤的一小部分,且通常只是单个解剖部位的异质性,而影像组学可捕获到整个肿瘤体积上的异质性。影像组学特征也因此与肿瘤的侵袭性相关^[9]。影像组学还被认为可预测生存期、治疗反应等临床终点,并与基因组学、转录组学或蛋白质组学特征相关^[1,2,9]。尽管单个影像组学特征可能与基因组学或临床结果相关,但将影像组学提供的大量信息(通常有数百种特征,其中一部分为特定疾病的影像组学特征)通过机器学习技术进行处理,其影响会更大^[10-11]。

第二,影像组学的数据是可挖掘的,这意味着在足够大的数据集中,它们可以用来发现从前未知的疾病进程、进展和治疗反应的标志物和模式。这种所谓的人群成像方法^[12]可能使用来自不同成像方式(如 PET、CT 和 MRI)的非结构

化数据,这些数据是以特定的但可能不相关的临床目的从广义群体中获得的;或者也可以使用一些公布的数据,如德国国家 MRI 队列研究,这是 1 项多中心纵向观察性的大型队列研究中的单一成像测试^[13]。这样的影像组学数据可以通过使用无监督机器学习的方式与临床数据、实验室数据、组织学数据、基因组学数据或其他数据相结合。

阻碍影像组学研究进展的 1 个主要原因是医师与计算机科学家之间缺乏沟通,特别是缺乏共通语言,因此本文旨在向医师介绍影像组学的术语。笔者没有选择数学方程而是使用组合物,并辅以插图标注重点概念。和其他任何技术一样,影像组学也有其技术限制,包括图像采集和重建参数的灵敏度,这些后面也将讨论。最后,本文将讨论影像组学在核医学中的应用实例。

一、影像组学特征类型

了解不同类型的影像组学特征对于阅读影像组学研究论文或是进行影像组学研究不是必需的。但是,熟悉其核心原则可能有助于解释结果以及为特定用途预选特征。有关影像组学特征和基本方程更完整的列表,推荐参考《影像生物标志物标准化倡议白皮书》^[14]。

影像组学特征大致可分为统计特征(包括基于直方图和纹理的特征)、基于模型的特征、基于变换的特征以及基于形状的特征^[15]。尽管影像组学特征既可以从二维 ROIs 中提取,也可以从三维感兴趣体积中提取,但为了提高可读性,

选择 ROI 作为两者的总称。此外,还可以计算未经修改的或离散化的灰度强度特征。灰度离散化(其中灰度的范围被减少到预定数值以提高稳定性和再现性)和特征值聚合(当以不同变体计算相同特征时,通过最简单的计算方式获得单个值)不在本文的范围内。

1.直方图特征。最简单的统计描述语是基于全局灰度直方图,包括灰度平均值、最大值、最小值、方差和百分位数^[14-15]。因为这些特征是基于单像素或单像素分析,所以被称为一阶特征。在 PET 中,常用的 SUV_{max} 、 SUV_{mean} 和 SUV 峰值(peak of SUV, SUV_{peak})都属于这一类。更复杂的特征包括偏度(skewness)和峰度(kurtosis),其描述了数据强度分布的形状:偏度反映数据分布曲线偏左(负偏,低于平均值)或偏右(正偏,高于平均值)的不对称性;峰度反映由于异常值导致的数据分布相对于高斯分布的拖尾性。其他特征包括直方图的熵(entropy)和均匀性(也称为能量)。值得注意的是,它们有别于同名的共生矩阵的对应项。

2.纹理特征。(1)绝对梯度(absolute gradient)。真正描述影像组学纹理特征的 1 个简单方法是分析绝对梯度,其反映了图像中灰度强度起伏的程度和急缓度。对于 2 个相邻的像素或体素来说,如果 1 个是黑色另 1 个是白色,则其梯度最高;而如果 2 个像素都是黑色(或都是白色),则该位置的梯度为零。灰度无论从黑色增加到白色(正梯度)或是白色减少到黑色(负梯度),均与梯度幅度无关。和直方图特征类似,梯度特征也包括梯度平均值、方差、偏度和峰度(图 1)^[14-15]。

(2)灰度共生矩阵(gray-level cooccurrence matrix, GLCM)。GLCM 是 1 个二阶灰度直方图^[16]。GLCM 在不同方向(二维分析的水平、垂直、对角线方向或三维分析的 13 个方向)上捕获具有预定义灰度强度的成对像素或体素的空间关系(图 2)。GLCM 特征包括熵(图 2),即对灰度不均匀性或随机性的度量;角二阶距(也被称为均匀性或能量),可以反映灰度的同质性和有序性;对比度,强调了在像素对或体素对中像素或体素之间的灰度差异^[14-16]。

(3)灰度游程长度矩阵(gray-level run-length matrix, GLRLM)。GLRLM 提供了在二维或三维中 1 个或多个方向上具有相同灰度的连续像素的空间分布信息^[17]。GLRLM 特征包括分数,其可评估在 ROI 内部分运行像素或体素的百分比,因此可反映颗粒性;长游程和短游程调强(逆)矩阵(图 1),其分别针对长游程和短游程的数据进行加权;灰度和游程不均匀性,其分别评估在不同灰度和游程长度上游程的分布^[14-15,17]。

(4)灰度区域大小矩阵(gray-level size zone matrix, GLSZM)和灰度区域距离矩阵(gray-level distance zone matrix, GLDZM)。GLSZM 基于同 GLRLM 相似的原理^[18],但本文中,具有相同灰度的互连的相邻像素或体素组(也称作“区域”)的数量统计构成了矩阵的基础(图 2)。纹理越均匀,矩阵就越宽越平。GLSZM 是针对相邻区域中不同像素或体素之间距离的计算。GLSZM 特征可以在二维(8 个相邻像素)或三维(26 个相邻体素)中计算,并遵循 GLRLM 定义,包括分数(目标区域内像素或体素的百分比),大、小区域调强以及其他^[14,18]。

作为 GLSZM 的变体,GLDZM 不仅评估具有相同灰度的互连的相邻像素或体素区域,还要求它们到 ROI 边缘的距离相同。因此 GLDZM 特征是纹理特征和形态学特征的混合,

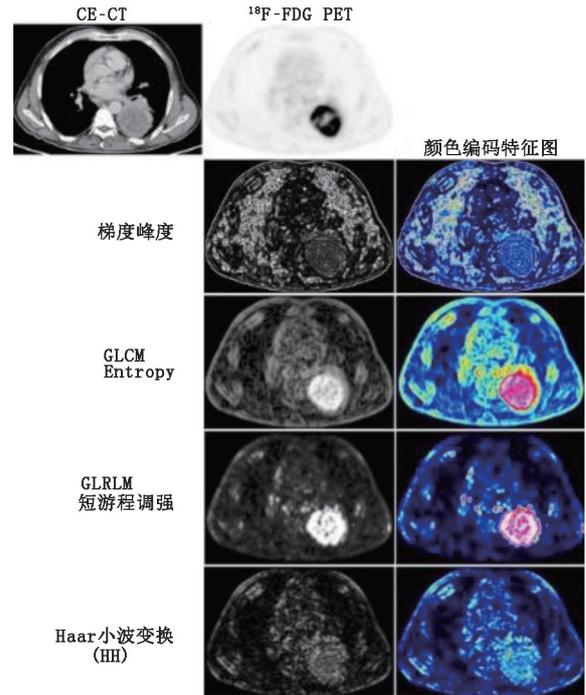


图 1 影像组学特征的直观表现:左肺下叶部分坏死肺癌的¹⁸F-FDG PET 和增强 CT 影像图。通过在 PET 图像上移动小矩形窗并计算每个位置的特征值生成的影像组学特征图,可反映整个肿瘤葡萄糖代谢异质性的不同方面。每个特征图描述 1 个影像组学特征,高值对应于灰度特征图上的高信号强度。颜色编码特征图可用于更好的可视化并用作 CT 的色彩叠加。CE 为对比增强,Entropy 为熵,GLCM 为灰度共生矩阵,GLRLM 为灰度游程长度矩阵,HH 为高通或双向高通滤波

这也反映在一些 GLDZM 特征名,如短距离高灰度调强^[18]。

(5)邻域灰度差矩阵(neighborhood gray-tone difference matrix, NGTDM)。NGTDM 量化了 1 个像素或体素的灰度与其相邻像素或体素在预定义距离内的平均灰度之间的差值总和^[19]。主要特征包括 NGTDM 的粗糙度、忙碌度和复杂度。粗糙度反映了中心像素或体素与其相邻像素或体素之间的灰度差异,从而捕获灰度强度的空间变化率,即由相对均匀灰度(空间强度变化率较低)较大区域组成的 ROI 具有较高的粗糙度值。忙碌度反映了中心像素或体素与其相邻像素或体素之间的快速灰度变化(强度变化的高空间频率),即灰度明显不同的小区组成的 ROI 具有更大的忙碌度。

(6)邻域灰度相关矩阵(neighborhood gray-level dependence matrix, NGLDM)。NGLDM 也是基于中心像素或体素与其相邻像素或体素之间的灰度关系^[20]。如果在定义的灰度差范围内满足相关性标准,则将预定义距离内的相邻像素或体素视为连接到中心像素或体素。然后分析 ROI 的中心像素或体素是否存在强度 i 和强度 j 相关的相邻像素或体素。和 GLRLM 相似,NGLDM 特征包括大、小相关调强,其反映异质性和同质性;以及灰度非均匀性和相关性均匀性,其反映整个 ROI 的灰度和灰度相关性的相似性^[14,20]。

3.基于模型的特征。基于模型的分析旨在解释空间灰度信息,以表征物体或形状。计算纹理生成的参数化模型并将其拟合到 ROI 中,其估计的参数被用作影像组学特征^[15]。

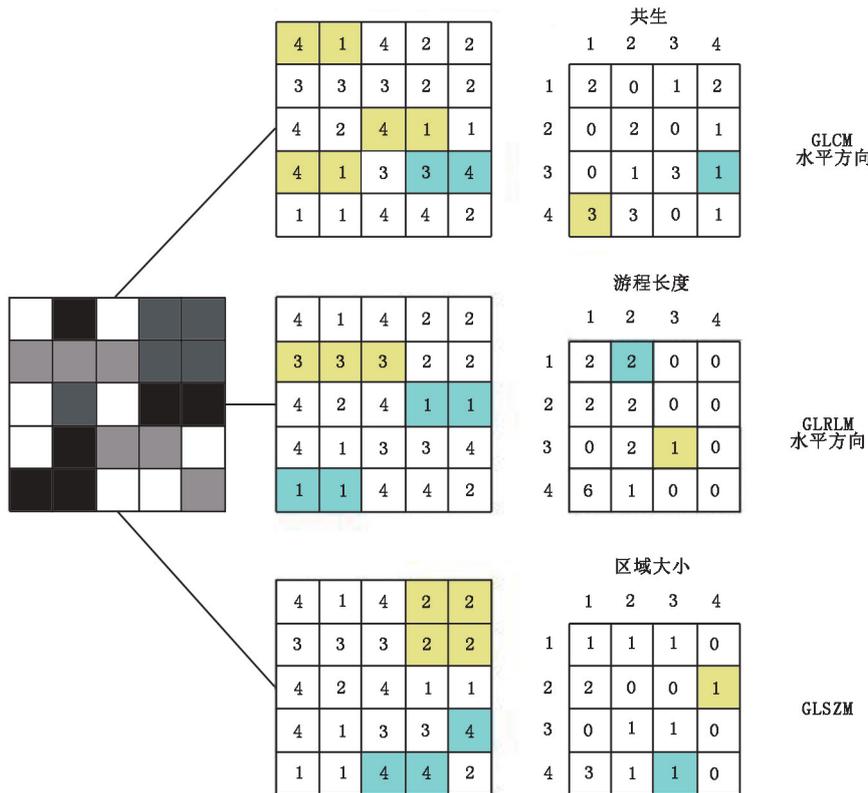


图 2 影像组学纹理特征的计算。灰度共生矩阵(GLCM)依赖于像素对(这里像素间距离=0),灰度游程长度矩阵(GLRLM)依赖于游程,灰度区域大小矩阵(GLSZM)依赖于有相同灰度的相邻像素区域

特征包括紧凑性和球形度,其描述了 ROI 的形状与圆形(二维分析)或球形(三维分析)的区别;以及密度,其依赖于构建包围 ROI 的最小定向边界框(或用于二维分析的矩形)^[14]。

二、采集参数和特征标准化

图像衍生指标,如 SUV 和影像组学特征对于图像采集设置、重建算法和图像处理非常敏感。Zwanenburg^[23]对 42 项 PET 影像组学研究进行了 Meta 分析,以评估特征的稳定性、重复性和标准化,其中 21 项符合量化评价标准。该研究分析了引起数据异质性的不同方面,包括采集参数的变化(如扫描持续时间、迭代次数和子集数目、重建类型和算法、空间分辨率)和图像处理的变化(分割方法和灰度离散化)。空间分辨率的影响最大(CV=3.63),其次是扫描持续时间(CV=2.93)、分割方法(CV=2.92)、重建方法(CV=2.30)、用以重建后平滑的高斯滤波宽度(CV=2.23)、迭代次数(CV=1.81)以及子集数目(CV=1.08);分割也有很大的影响(CV=2.92)^[23]。

自回归模型是基于模型方法的示例,其基于这样的思想:像素灰度是 4 个相邻像素灰度的加权和,包括左侧(θ-1)、左上侧(θ-2)、上侧(θ-3)和右上侧的像素(θ-4)。此外,σ 携带有关最小预测误差的方差信息,用于测量纹理的规律^[15]。

分形分析也可产生用于影像组学的特征,特别是分维,其反映了随着放大倍数、比例或分辨率的增加,结构细节增加的速率,因此可作为复杂性的度量。孔隙度分析是衡量旋转不变性或平移不变性缺失的 1 项特征,反映了不均匀性^[21]。

4. 基于变化的特征。基于变化的方法,包括 Fourier、Gabor 和 Haar 小波变换,用于分析不同空间的灰度模式。例如,离散的 Haar 小波变换可分析不同尺度下图像的频率内容^[15]。通过应用正交镜面滤波器、高通和低通滤波器,可对图像进行小波分解^[22]。虽然高通滤波器突出了灰度的变化并由此强调了图像细节,但低通滤波器在灰度方面对图像进行了平滑处理,去除了图像细节。经过信号分解之后,可以得到 1 组空间定向的频率通道,用于描述局部影像的可变性,并可将频率通道内的能量作为特征。在不同尺度上,双向高通滤波器(图 1)捕获对角线细节,高通滤波器后低通滤波器捕获垂直边缘,低通滤波器后高通滤波器捕获水平边缘,双向低通滤波器捕获最低频率^[15]。值得注意的是,小波变换不仅可用于生成影像组学特征,还可用于图像分割或纹理分析的预处理步骤。

5. 基于形状的特征。基于形状的特征描述了 ROI 的几何属性。许多基于形状的特征比其他影像组学特征的概念要更简单,如二维和三维、轴及其比率。基于网格(即三角形和四面体的小多边形)面积或体积的方法就更为复杂。具体

van Velden 等^[24]报道,大多数影像组学特征的重复性与简单 SUV 测量相当(60% 都有良好的重复性),且与重建方法相比,这些特征对勾画或分割更灵敏。其他研究还发现,影像组学的重复性类似于 SUV 的重复性^[25],且对图像的离散化非常灵敏^[26]。Lasnon 等^[27]发现点扩散函数和后过滤的有序子集最大期望值迭代法(ordered-subset expectation maximization, OSEM)产生的 SUV 和纹理信息与普通 OSEM 相当,但未过滤的点扩散函数图像显示出更高的异质性(在分层患者中可能更易区分)。Papp 等^[28]使用基于球形的体模,比较各个 PET 影像组学特征类别对空间分辨率和重建算法变化的灵敏度。为最大限度减小特征差异,他们推荐小体素、窄高斯后滤波,并像 Lasnon 等^[27]一样,使用点扩散函数建模。Yan 等^[29]发现箱体大小对影像组学特征影响很小,但迭代次数、后处理滤波器宽度和体素大小会影响其中多个影像组学特征(61 个特征中分别有 3、5 和 35 个特征 > 20% CV)。Galavis 等^[30]发现,由于不同的采集和重建参数(二维或三维 OSEM、迭代次数、处理后滤波器宽度和体素大小),80% 的研究纹理特征显示出很大的特异性(> 30%)。Pfahler 等^[31]指出,与使用标准 OSEM 或时间飞行算法重建的图像相比,使用点扩散函数重建的图像具有更高的重复性。

空间分辨率似乎对影像组学有显著影响,因此常对多中心成像数据进行重采样以改善扫描椅之间的分辨率差异影响。Whybra 等^[32]研究了 PET/CT 影像组学纹理特征和形状特征在三线性插值和样条插值的鲁棒性,以便对各向同性体素大小进行重采样。在接受测试的 141 个影像组学特征中,大约有 2/3 对 2 种重采样技术是鲁棒的,另有 21% 可能可以

校正。尽管如此,这 2 种插值技术的绝对差异在一定程度上是相当大的,因此,应该一贯使用同一种插值技术。

最近提出了用于 PET 重建的深度学习,可用于传统重建图像的后处理^[33]、迭代重建框架^[34]、将 PET 数据直接映射到图像上^[35]。这些方法能恢复或重建比传统 OSEM 更高质量的 PET 图像,并在不降低分辨率的情况下减少图像噪声。由于噪声(由计数统计和迭代次数等因素控制)可能严重影响图像的计算特征,深度学习可能会使影像组学特征更具鲁棒性。

三、特征的调和、选择和缩减

1.特征调和。调和是 1 个数学后处理技术,用于图像采集、重建和分析之后消除对影像组学特征的所谓批次效应(即采集参数变化的中心相关效应)。因此,调和不是用于图像而是直接应用于数字影像组学特征值。当前最流行的技术是 ComBat 调和,最初是用来描述基因组学数据,现已被验证可用于消除影像组学特征中心效应,并保留病理生理学信息^[36]。已经有几项研究将该技术应用于 PET 影像组学^[37-38]。

2.特征选择与降维。影像组学图像分析完成后,必须确定统计模型中解决临床问题(如区分良恶性病变)的相关特征。

尽管从理论上说,提取出的数百个候选影像组学特征(图 3)可用于预测模型的输入,但所需模型参数的数量将会呈指数增长,因此必须去除或转换大量的候选特征,这个过程称作降维。影像组学特征常表现出高度的相关性,表明数据冗余,意味着某些特征可被丢弃,而另一些特征可通过主成分分析或线性判别分析,被分组并被替换成具有代表性的特征。在这些代表性特征中,显示最高自然生物学范围(如患者之间的变异)的信息特征将被优先选择^[39]。Parmar 等^[40]对 14 种影像组学特征选择方法进行了比较,包括基于互信息的方法和 12 个机器学习分类器;Leger 等^[41]也使用了类似的方法,他们

关注的是时间-事件存活率。Lian 等^[42]开发了用于从 PET 图像中获取影像组学特征以预测肿瘤治疗结果的系统。

在不利用目标变量(如良性或恶性)的情况下减少冗余的降维技术(如主成分分析)很流行,但通常是混合变量,并使初始影像组学特征集中的预测因子的后续跟踪复杂化。一旦在特征选择中考虑了预测目标,就必须注意避免所谓的过度拟合,以免导致对预测精读的过度乐观估计。如果特征的数量足够多,即使是在随机的数据中也能检测到相关性。

四、模型构建与分类

选择影像组学特征后,其被用来预测当前的目标变量,如疾病的有无或肿瘤的类型;或是未来的变量,如治疗反应或复发时间。

目标变量可以是回归模型预测的标量(如以月为单位的存活期),也可以是分类模型(图 3)预测的分类类别(如反应状态或受体阳性)^[43]。机器学习作为人工智能的分支领域,在过去 20 年中得到了快速发展。机器学习模型基于训练示例学习高维输入-影像组学特征和目标变量之间的关系。一个关键的能力是综合利用预测因子或特征组,即所谓的多变量模式,而不是单变量或质量单变量回归。支持向量机是首批非常成功的模型之一,唯一不足是需要仔细选择特征^[44]。Bagging 方法和 boosting 方法,如随机森林,通过在训练过程中集成特征选择和有效采样,引入了鲁棒分类器或回归器的训练^[45]。这一能力是当前许多基于预测贡献度来选择特征的影像组学方法的核心。这些方法通过将重点从专家引导的特征选择转向直接让分类器评估候选特征,并识别那些具有预测值的特征,这些方法已经在汽车(行人识别)到遗传学(识别关联)等多个领域取得了非常成功的应用^[46]。最近,神经网络重新崛起,成为强大的分类和特征重建模型。由于

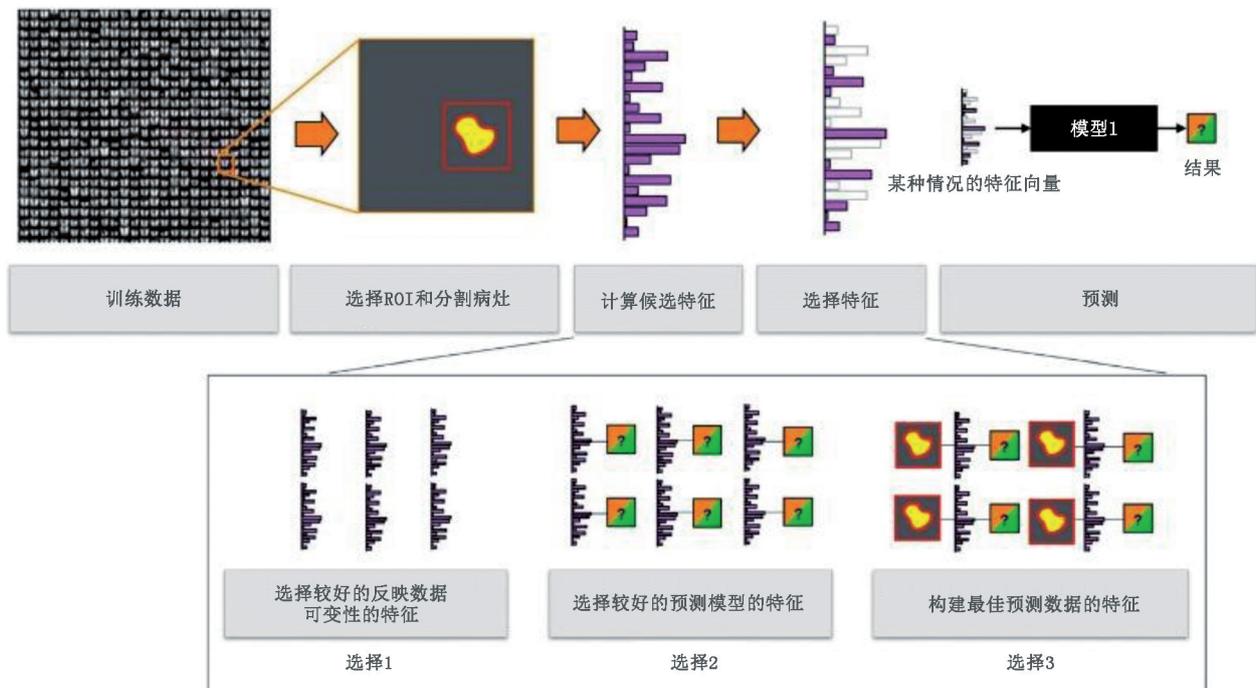


图 3 影像组学工作流程。首先,定义 ROI 或分割病灶。对于 ROI 或病灶,通常提取大量候选特征。选择有效反映数据可变性的特征,或能够更好服务于特定预测模型的特征。深度学习方法不是从预定义特征集中选择,而是直接将特征重建和建模联系起来以进一步提高预测准确性和可靠性

卷积神经网络 (convolutional neural network, CNN) 不是从预定义中有限的候选特征集中选择,而是通过图像数据本身构建最佳特征,所以在给定足量训练数据的情况下,CNN 的性能优于特征选择方案^[47]。CNN 的局限性在于需要大量的训练数据,最近迁移学习等方法利用问题域中视觉特征的相似性,成功解决了这一限制。

测试集的拟合优度是近似函数预测目标函数输出的有效度量。对于分类变量,假阳性、假阴性或衍生变量反映了相关信息。由于许多方法都有参数,所以可用 ROC 曲线来表示参数范围内的预测精度。在交叉验证方案中,训练集和测试集在可用数据集中轮换,可减少有限的数据集大小,但必须谨慎使用。一旦算法设计及其参数通过中间交叉验证获得大量信息,提供独立测试结果的能力就会丧失。最好的做法是在最终验证期间保留 1 组仅在算法完成后才可触及的数据。关于机器学习的更多信息可以在 Uribe 等^[11]的文章中找到。

五、隐患与质量控制

在进行影像组学的研究时有几个注意事项,其中一些已经在前面简要提及,具体包括需要高质量、无伪影的图像,最好选用通过同质图像采集协议和重建技术获得,或在无法达到图像质量要求时可在影像组学分析之前或之后使用适当的校正(如在多中心回顾性研究中);足够的数据集大小和数据集完整性;以及单独的训练数据集和验证数据集。除了这些有意义的影像组学研究的前提之外,还有 2 个经常影响影像组学研究的隐患:类别失衡和过度拟合。

1. 类别失衡。在随机临床试验之外,类别失衡很常见。特别是在使用常规临床数据的回顾性研究中,感兴趣疾病有和无的发生率相同的情况很少。例如,在弥漫性大 B 细胞淋巴瘤患者中,大约有 16% 的患者骨髓受累。在评估¹⁸F-FDG PET 影像组学检测骨髓受累的效果时,必须考虑骨髓受累患者(16%)和不受累患者(84%)之间比例的失衡。将样本中所有的病例都归入无骨髓受累组的分类器有着 84% 看似不错的准确性,但这在临床上没有用,因为其在 PET 图像上无法区分受累和未受累的骨髓(补充表 1;补充材料可在 <https://jnm.snmjournals.org> 中获得)。因此,不仅应报告总体的准确性,还要报告各类别的准确性、灵敏度、特异性。

2. 过度拟合和欠拟合。如果模型在函数近似方面没有充分平衡,则可能会遇到过度拟合或在较小程度上遇到欠拟合的情况。当具有大量输入参数或太多自由度的模型可记忆数据时,就会发生过度拟合,因此模型中不仅包括相关的特定疾病的特征,还包括反映图像噪声和随机波动的特征(补充图 1)。在训练期间,这样的模型在提供给它的数据点上给出正确的分类结果,但其对训练数据集以外的点的响应是错误的——该模型无法概括信息。为避免过度拟合,需要运用正则化来平滑模型函数,或需要减少输入特征的数量,从而减少所需模型参数的数量。使用单独的数据集进行验证有助于检测过度拟合:如果训练数据集中的错误减少,但验证数据集中的错误开始增加,则需要停止训练。另一方面,模型无法在训练和验证数据集中正确分类数据时会发生欠拟合,如模型过于简单时。如果发生欠拟合,可能需要额外的数据或切换到不同的模型。

3. 影像组学评分。基于 16 项具有不同权重的标准,Lambin 等^[48]制定了与模态无关的影像组学质量评分方法,

最多可达 36 分。尽管根据已公布的建议使用标准化的图像采集协议是这些标准的 1 项,但其影响很小;可能是出版日期的原因,没有明确提及基于软件的多中心数据校正或协调技术;相反,支持使用体模来评估影像组学特征值。对前瞻性设计和试验注册给予特定的权重,两者各占 7 分;验证数据集的使用也很重要,占 5 分。使用特征减少技术来降低过度拟合的风险也是 1 个相关标准,占 3 分。值得注意的是,与当前的“金标准”(影像组学和基于影像的 TNM 分期)相比,影像组学附加值的评估以及临床相关性和实用性,都是重要的因素,各占 2 分。该评分还建议将影像组学数据和临床数据、分子数据和基因组学数据相结合^[48]。

六、临床应用

和大量使用传统的 PET 指标进行结果预测或组织表征的研究相比,对纹理、形状或直方图特征的应用在文献中还很少见。这一部分将回顾精选的文章,以突出有前景的临床应用,并讨论其局限性。

1. 影像基因组学。影像基因组学将影像数据和生物学联系起来。影像基因组学在非小细胞肺癌(non-small cell lung cancer, NSCLC)中的应用引起了人们特别的兴趣。Nair 等^[49]研究了在 NSCLC 患者中¹⁸F-FDG PET/CT 的肿瘤代谢体积(metabolic tumor volume, MTV)和直方图特征与基因组数据的相关性。在 25 例肿瘤切除前接受 PET/CT 检查患者的训练数据集中,14 个影像组学特征和 3 个主要成分均与单基因或共表达基因簇的表达相关。4 个基因(LY6E、RNF149、MCM6 和 FAP)与影像组学特征和患者生存相关。与影像组学研究不同的是,证实这些相关性的测试和验证队列比训练队列大得多(63 和 84 例患者)。Yip 等^[50]研究了 348 例 NSCLC 患者的¹⁸F-FDG PET/CT 影像组学特征(直方图、GLCM、GLRLM、GRSZM、NGTDM 和形状)与表皮生长因子或 Kirsten 大鼠肉瘤型三磷酸鸟苷酶家族(Kirsten rat sarcoma GTPase family, KRAS)基因突变的关系。虽然有 8 个纹理特征(以及 SUV 和 MTV)和表皮生长因子突变状态相关,1 个 GLCM 特征甚至可以预测表皮生长因子的阳性突变状态,但没有和 KRAS 突变相关的特征。由于缺乏验证队列,数据来自 8 台不同 PET/CT,该项研究具有一定的局限性,尽管对体素强度进行了重新采样,但体素大小却没有重新采样。

Pyka 等^[51]使用 *O*-(2-[¹⁸F]氟乙基)-*L*-酪氨酸} *O*-(2-[¹⁸F]-fluoroethyl)-*L*-tyrosine, ¹⁸F-FET} PET/CT 影像组学方法来区分 113 例高级别胶质瘤患者的肿瘤级别,采用同一 PET/CT 设备,避免了后续重新采样或调和的需要。4 个 NGTDM 特征能够区分 III 级和 IV 级肿瘤;通过与 MTV 相结合,取得了进一步的优化。值得注意的是,研究没有使用验证集。为预测术前不常获得的 1 种诊断标志物——异柠檬酸脱氢酶的突变,Lohmann 等^[52]回顾性分析了¹⁸F-FET PET/CT 影像组学。84 例患者中有 56 例接受了 PET 检查,余 28 例接受了 PET/MR 检查——也就是说,几乎所有的采集参数(包括分辨率和重建算法)在 2 个亚组之间都不同。从中提取了 33 个特征(直方图、GLCM、GLRLM、GLSZM 和形状),84 例患者中有 26 例显示了异柠檬酸脱氢酶突变。为避免过度拟合,相关特征的数量减少到 2 个,并采用 5 倍和 10 倍交叉验证。最终准确性高达 81%,但灵敏度相当低,可能是由于类别失衡和

采集参数异质性共同导致。

Rajkumar 等^[53]检测了从¹²⁵I-A5B7 抗癌胚抗原抗体的小动物 SPECT 显像中提取出的 GLCM 特征是否可用于鉴别转移性结直肠癌的表型。在 14 只结直肠癌肝转移的小鼠中,他们发现未分化的转移灶明显比分化良好的转移灶更不均匀,这从 3 个 SPECT 纹理特征中也得到反映,这些特征还可反映抗血管治疗效果。

2. 临床结果预测。临床医师对于早期评估治疗反应和预测生存率非常感兴趣,因为这可能有助于治疗选择和患者分层,并证明治疗转换的合理性。在 358 例 I ~ III 期 NSCLC 患者中,Arshad 等^[54]使用来自 7 个机构的治疗前¹⁸F-FDG 数据来评估¹⁸F-FDG PET 影像组学对放疗或化疗后总体生存率的预测。除了传统的 PET 指标外,还计算了直方图、形状和纹理特征(从原始图像和小波变换图像中提取的 GLCM, GLRLM 和 NGTDM),并通过最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)回归与加权线性特征组合的方式进行降维,但没有对采集参数的变化进行校正。共有 133 个数据集用于训练,其中 204 例患者进行了内部验证,21 例患者进行了外部验证。组合的影像组学特征正确地预测了验证队列中 14 个月的生存差异和测试队列中没有生存差异的情况。

为预测 III ~ IV a 期鼻咽癌患者的无病生存期,Peng 等^[55]使用通过同一扫描仪和固定采集方案获得的图像,评估了基于预处理¹⁸F-FDG PET/CT 影像组学信号和列线图。训练数据集有 470 例患者,验证数据集有 237 例患者。影像组学特征(包括直方图、形状、GLCM 和 GLRLM 特征)均为手动选择,通过深度学习 CNN 自动选择,LASSO Cox 回归分析用于特征降维。影像组学的列线图被证明优于基于临床数据和血浆 Epstein-Barr 病毒 DNA(公认的鼻咽癌预后生物标志

物)的列线图。影像组学的列线图能将患者分为 2 个 5 年无病生存率不同的风险组;除了标准的放化疗之外,只有影像组学高危组显示出诱导化疗的益处。有趣的是,研究没有评估影像组学特征与其他数据(如 DNA)的组合。然而,Lv 等^[56]在 1 项类似的涉及 128 例鼻咽癌患者的研究中评估了这项策略;在该研究中,影像组学和临床数据的结合略微提高了对无进展生存期的预测。

Jiang 等^[57]研究了 214 例胃癌患者的¹⁸F-FDG PET 影像组学特征(直方图、形状、GLCM、GLRLM、GLSZM 和 NGTDM)在无病生存期和总生存期预测中的作用。在该研究中,训练队列(132 例患者)使用同一扫描仪进行检查,验证队列(82 例患者)使用来自不同供应商的不同扫描仪进行检查,提供了真正的外部验证。尽管体素大小和其他采集参数在 2 组队列中不同,但影像组学评分是建立在以 LASSO 回归选择的特征基础上的,且和 TNM 分期或肿瘤标志物糖类抗原 19-9 相比可更好预测总生存期和无病生存期。同样,影像组学预测模型中没有包括 CT 影像组学特征或临床、实验室数据。

七、结论

影像组学是 1 种复杂的影像分析技术,有可能在精准医学领域中确立自己的地位。影像组学特征不仅与基因组学相关,还可以提供整个肿瘤体积上肿瘤异质性的补充信息,以改善生存预测,并因此可能有助于患者分层。对核医学来说,凭借其可提供定量生物数据的传统优势,影像组学可能为核医学的进展提供新的方向,不仅可以作为临床决策工具,还可作为发现疾病新分子通路的研究工具。也要注意,制定并严格遵守标准化的图像采集和重建协议至关重要。本文参考文献具体见 <http://jnm.snmjournals.org>。

(收稿日期:2021-11-30)

· 读者 · 作者 · 编者 ·

2023 年本刊可直接用缩写的常用词汇

ATP(adenosine-triphosphate),三磷酸腺苷

AUC(area under curve),曲线下面积

CI(confidence interval),可信区间

CT(computed tomography),计算机断层摄影术

CV(coefficient of variation),变异系数

DNA(deoxyribonucleic acid),脱氧核糖核酸

FDG(fluorodeoxyglucose),脱氧葡萄糖

HAV(hepatitis A virus),甲型肝炎病毒

Hb(hemoglobin),血红蛋白

HBsAg(hepatitis B surface antigen),乙型肝炎表面抗原

HBV(hepatitis B virus),乙型肝炎病毒

HCV(hepatitis C virus),丙型肝炎病毒

MRI(magnetic resonance imaging),磁共振成像

PBS(phosphate buffered solution),磷酸盐缓冲液

PCR(polymerase chain reaction),聚合酶链反应

PET(positron emission tomography),正电子发射断层摄影术

PLT(platelet count),血小板计数

RBC(red blood cells),红细胞

RNA(ribonucleic acid),核糖核酸

ROC(receiver operating characteristic),受试者工作特征

ROI(region of interest),感兴趣区

SPECT(single photon emission computed tomography),单光子发射计算机断层摄影术

SUV(standardized uptake value),标准摄取值

SUV_{max}(maximum standardized uptake value),最大标准摄取值

SUV_{mean}(mean standardized uptake value),平均标准摄取值

WBC(white blood cells),白细胞

WHO(World Health Organization),世界卫生组织

本刊编辑部