



DOI:10.11817/j.issn.1672-7347.2019.03.002

<http://xbyxb.csu.edu.cn/xbwk/fileup/PDF/201903233.pdf>

基于知识图谱的影像组学研究可视化分析

罗爱静^{1,2}, 姚山虎^{1,2,3}, 冯智超¹, 容鹏飞¹, 覃岳香¹, 王维^{1,2}

(中南大学 1. 湘雅三医院, 长沙 410013; 2. 医学信息研究湖南省普通高等学校重点实验室, 长沙 410013; 3. 信息安全与大数据研究院, 长沙 410083)

[摘要] **目的:** 使用知识图谱的方法对影像组学领域文献分布、研究力量分布和研究热点进行分析, 为相关研究提供参考。**方法:** 下载Web of Science数据库中影像组学文献题录。分别使用BICOM 2.0.1和SATI 3.2对发表年份、期刊、作者、关键词、科研机构数据进行清理和频次统计; 使用CiteSpace V4.4.R1构建国家(地区)合作网络知识图谱; 使用Ucinet 6构建核心著者、核心科研机构的合作网络知识图谱; 使用gCLUTO 1.0构建高频关键词双聚类图。**结果:** 共检索到700篇文献, 自2012年开始论文发表数量逐年快速增长。在该领域领先的主要有美国、中国、荷兰等国家, 形成5个重要科研机构合作群和10个重要著者合作群。根据高频关键词双聚类分析发现8个研究热点方向。**结论:** 影像组学作为一个新兴的研究方向处于快速发展阶段, 越来越多的国家、科研机构及多学科背景科研人员投入到相关研究中, 并不断有新的名词、新的方法出现在该领域。

[关键词] 影像组学; 知识图谱; 合作网络; 研究热点

Visualization analysis for radiomics research based on knowledge mapping

LUO Aijing^{1,2}, YAO Shanhu^{1,2,3}, FENG Zhichao¹, RONG Pengfei¹, QIN Yuexiang¹, WANG Wei^{1,2}

(1. Third Xiangya Hospital, Central South University, Changsha 410013;

2. Key Laboratory of Medical Information Research (Central South University), College of Hunan Province, Changsha 410013;

3. Research Institute of Information Security and Big Data, Central South University, Changsha 410083, China)

ABSTRACT

Objective: To illustrate the literature distribution, research power distribution, and research hotspots in the radiomics research by using knowledge mapping analysis, and to provide reference for relevant researchers.

Methods: Bibliographies from literature regarding radiomics in Web of Science database were

收稿日期(Date of reception): 2018-08-07

第一作者(First author): 罗爱静, Email: luoaj@mail.csu.edu.cn, ORCID: 0000-0001-5890-3470

通信作者(Corresponding author): 姚山虎, Email: yaoshanhu09@163.com, ORCID: 0000-0002-9782-4692

基金项目(Foundation item): 国家社科基金重点项目(17AZD037); 国家自然科学基金(81471715); 湖南省自然科学基金(2017JJ2369); 湖南省哲学社会科学项目(16YBA366)。This work was supported by Key Projects of the National Social Science Foundation (17AZD037); the National Natural Science Foundation (81471715); Natural Science Foundation of Hunan Province (2017JJ2369); Hunan Provincial Philosophy and Social Science Foundation (16YBA366), China.

downloaded. BICOM 2.0.1 and SATI 3.2 were used to clean and caculate the frequency of publication year, journal, author, key word, and research institution. CiteSpace V4.4.R1 was used to build the knowledge map of scientific research collaboration network between countries/regions. Ucinet 6 was used to build the knowledge map of scientific research collaboration network between core authors and institutions. gCLUTO 1.0 was applied to construct high-frequency keywords bi-clustering map.

Results: A total of 700 literature was screened. Since 2012 the number of publications has been growing rapidly year by year. The United States, China, and Netherlands were leaders in this field. There were 5 major scientific research institution cooperative groups and 10 major author cooperative groups. Eight research hotspots were clustered by using high-frequency key word bi-clustering analysis.

Conclusion: Radiomics is a new field and develops very fast. More and more countries, research institutions, and researchers with multidisciplinary background are going to participate in this filed. New terminology and new methods are going to appear in the field.

KEY WORDS

radiomics; knowledge mapping; collaboration network; hotspot

医学影像可以全面显示肿瘤整体形态、内部结构及功能信息,在肿瘤疾病诊疗方面发挥重要作用。常规医学影像评价多为肉眼主观分析形式,一定程度上限制了其临床应用价值。随着高通量影像特征提取技术的发展,多模态医学图像转变成了可挖掘的高维定量数据,被用于辅助临床决策,即影像组学^[1];与基因组学、蛋白质组学等数据相比,影像数据具有无创、全面、易获取(临床常规使用)、经济和多次数据(随访监测可形成时序数据)等优点^[2]。影像组学在肿瘤个体化诊疗、精准医学方面具有重要临床应用价值和前景^[3-4],发展极为迅速。在此笔者使用知识图谱方法对影像组学相关文献进行分析,构建合作关系网络知识图谱、关键词双聚类图,理清影像组学发展历程及现状,为今后相关研究提供参考。

1 资料与方法

1.1 资料来源

本文以Web of Science核心库为文献来源数据库。检索式:主题=“Radiomic” OR “Radiomics”。文献类型设定为Article, Review和Proceeding paper,时间不限,检索日期为2018年10月9日,以txt形式导出文献题录信息。

1.2 数据分析方法

采用CiteSpace V4.4.R1进行国家分布、学科分布情况分析,时间切片为每年1次、阈值设置为前50,

算法为Pathfinder。

使用BICOM 2.01进行著者、关键词、期刊数据的清理和频次统计,使用SATI 3.2进行科研机构数据的清理和频次统计;频次不小于5的作为核心著者、核心科研机构分别构建共现矩阵(根据普赖斯定律,以本领域最高产作者发文量的平方根的0.749倍为界,高于此发文量的作者即为该领域的核心著者,本研究最高产著者发文38篇,核心著者最低发文量为4.62篇,取整数为5篇;在此以核心著者最低发文量作为核心科研机构划分标准)。使用Ucinet 6构建著者共现网络、科研机构共现网络,进行核心-边缘社会网络分析^[5]以获得关键著者、科研机构;频次不小于10的作为高频关键词构建词篇矩阵。使用gCLUTO 1.0生成高频关键词双聚类树图^[6],设定聚类方法为重复二分法,相似系数选用Correlation Coefficient,聚类标准函数设定为 I_2 。

2 结果

2.1 文献分布情况

2.1.1 发表年代分布

共检索到700篇文献,其中2012, 2013, 2014, 2015, 2016, 2017, 2018年发表篇数分别为2(0.29%), 4(0.57%), 9(1.29%), 45(6.43%), 109(15.57%), 223(31.86%)和308(44.00%)。2012年发表首篇论文后其研究热度逐年上升,2015年出现大幅升高。

2.1.2 来源期刊分布

涉及期刊188种、学术会议29场。根据布拉德福定律^[7], 核心区期刊有8种(193篇文献, 占比32.82%), 包括*Scientific Reports* 51篇、*Medical Physics* 27篇、*European Radiology* 23篇、*Journal of Medical Imaging* 22篇、*PLoS One* 20篇、*Physics in Medicine and Biology* 18篇、*Oncotarget* 17篇、*Radiology* 15篇。由国际光学工程学会(Society of Photo-Optical Instrumentation Engineers, SPIE)举办的医学影像学术论坛是该领域的重要会议(68篇文献, 占比60.71%), 2015, 2016, 2017和2018年载文量分别为1, 13, 20和34篇。

2.1.3 学科分布情况

涵盖65个学科, 发文量占比10%以上的有: 影像医学与核医学、肿瘤学、工程学、科学技术-其他主题、多学科科学及生物医学工程学, 其载文量分别为370篇(52.86%), 158篇(22.57%), 77篇(11.00%), 75

篇(10.71%), 75篇(10.71%)及70篇(10.00%)。

2.2 研究力量分布情况

2.2.1 国家(地区)发文量及合作情况

共有43个国家(地区), 发文量排名前10的有美国344篇(49.14%)、中国145篇(20.71%)、荷兰57篇(8.14%)、德国52篇(7.43%)、加拿大48篇(6.86%)、法国46篇(6.57%)、意大利38篇(5.43%)、英国29篇(4.14%)、韩国22篇(3.14%)、日本19篇(2.71%), 网络中心度分别为0.70, 0.00, 0.42, 0.15, 0.24, 0.10, 0.10, 0.12, 0.00, 0.03。美国、荷兰处于国家(地区)合作网络的中心位置, 于2012年合作发表了最早的两篇文章, 与加拿大开展了深入合作。中国于2015年才发表首篇论文, 但随后发文量增长迅速, 与美国开展了较深入合作。

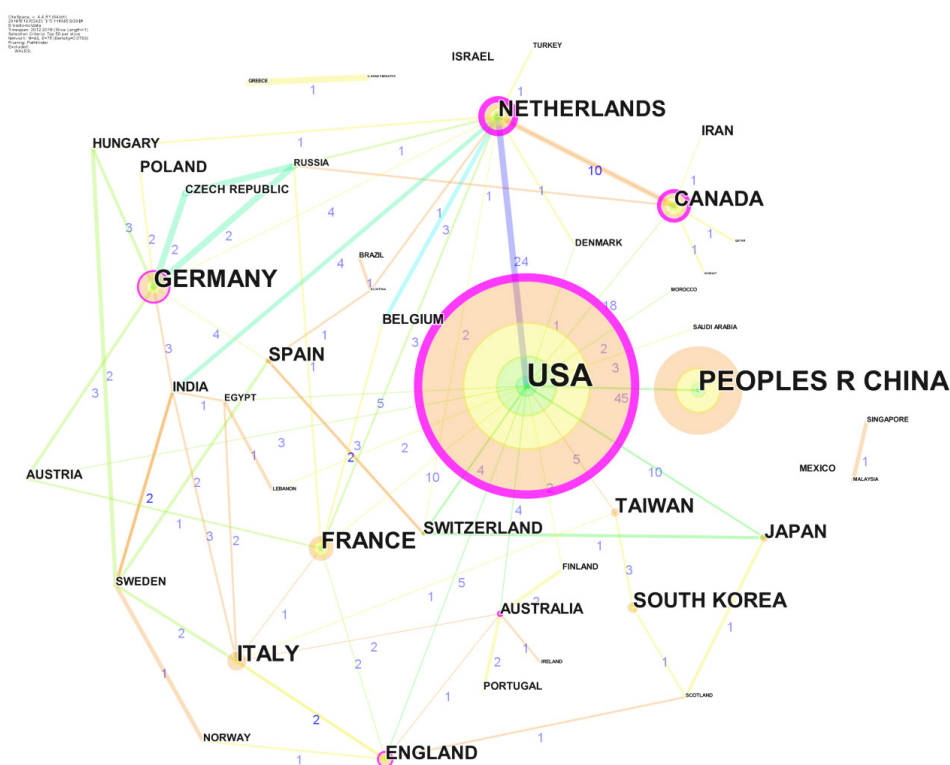


图1 国家(地区)合作网络知识图谱

Figure 1 Scientific collaboration network map between countries/regions

The size of nodes in the figure reflects how many papers the country/region published; the number on the connections between nodes indicates how many papers they co-published; the thickness of the connections reflects the cooperation density (the more cooperation, the thicker the line)

2.2.2 科研机构发文量及合作情况

共有789家科研机构, 其中核心科研机构有72家, 包括美国26家、中国19家、加拿大5家、法国5家、荷兰5家、意大利4家、德国3家、西班牙2家、韩

国1家、印度1家、瑞士1家。

基于图2A开展核心-边缘社会网络分析发现, 哈佛大学、马斯特里赫特大学、南佛罗里达大学处于该网络的中心位置(网络中心度0.823)。基于图2发现

如下重要科研机构合作群: 1)以南佛罗里达大学、哈佛大学、马斯特里赫特大学、阿姆斯特丹自由大学、多伦多大学为核心, 包括天津医科大学、奈梅亨大学、麻省总医院和滑铁卢大学的科研机构合作群; 2)德克萨斯大学MD Anderson肿瘤中心、德克萨斯大学休斯顿健康科学中心和贝勒医学院科研机构

合作群; 3)中国科学院、广东省医学科学院/广东省人民医院、南方医科大学、华南理工大学和首都医科大学科研合作群; 4)德国癌症研究中心和海德堡大学科研合作群; 5)凯斯西储大学和克利夫兰医学中心科研合作群。

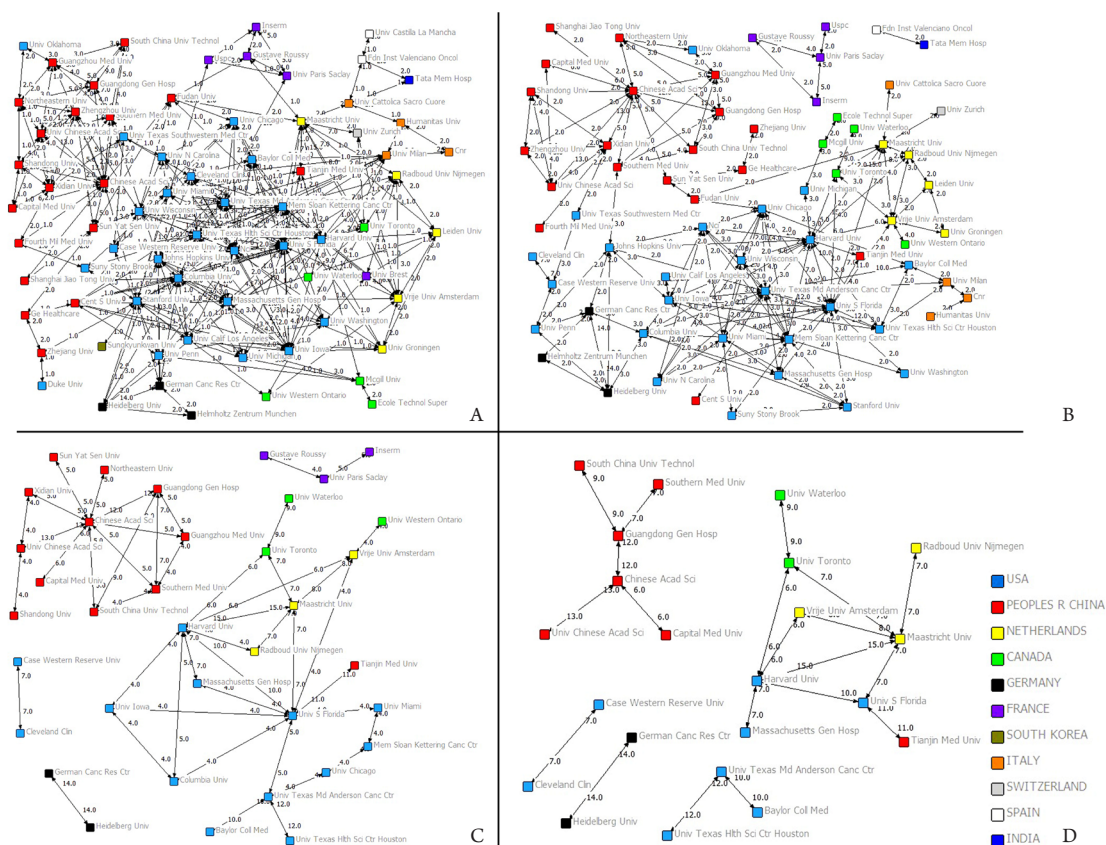


图2 核心科研机构合作网络知识图谱

Figure 2 Scientific collaboration network map between core institutions

A: Co-published papers no less than 1; B: Co-published papers no less than 2; C: Co-published papers no less than 4; D: Co-published papers no less than 6. The number on the connections between nodes indicates how many papers they co-published

2.2.3 著者发文量及合作情况

共有3 371位著者, 其中核心著者169位。基于图3A数据开展核心-边缘社会网络分析, 发现包括Gillies RJ, Aerts HJWL, Lambin P, Leijenaar RTH, Parmar C, Balagurunathan Y, Goldgof DB, Schabath MB, Velazquez ER, Hall LO, Liu Ying, Grossmann P, Stringfield O, Mak RH, Ye ZX, Li Qian, Gatenby RA, Van Elmpt W, Dekker A在内的19位学者处于网络中心位置(网络中心度为0.806)。基于图3发现存在如下重要著者合作群: 1)以Aerts HJWL, Lambin P为代表的哈佛大学、马斯特里赫特大学著者合作群; 2)以Gillies RJ, Goldgof DB为代表的南佛罗里达大学李莫

菲特癌症中心/研究所和计算机科学与工程学院著者合作群; 3)以Court LE为代表的德克萨斯大学MD Anderson肿瘤中心著者合作群; 4)以Tan J, Liu ZY和Liang CH为代表的中国科学院、广东省医学科学院/广东省人民医院著者合作群; 5)以Madabhushi A为代表的凯斯西储大学著者合作群; 6)以Giger ML为代表的芝加哥大学著者合作群; 7)以Khalvati F, Wong A等为代表的多伦多大学、滑铁卢大学著者合作群; 8)以Wang YY为代表的复旦大学著者合作群; 9)以Akbari H, Davatzikos C为代表的宾夕法尼亚大学著者合作群; 10)以Hadjiski L为代表的密歇根大学著者合作群。

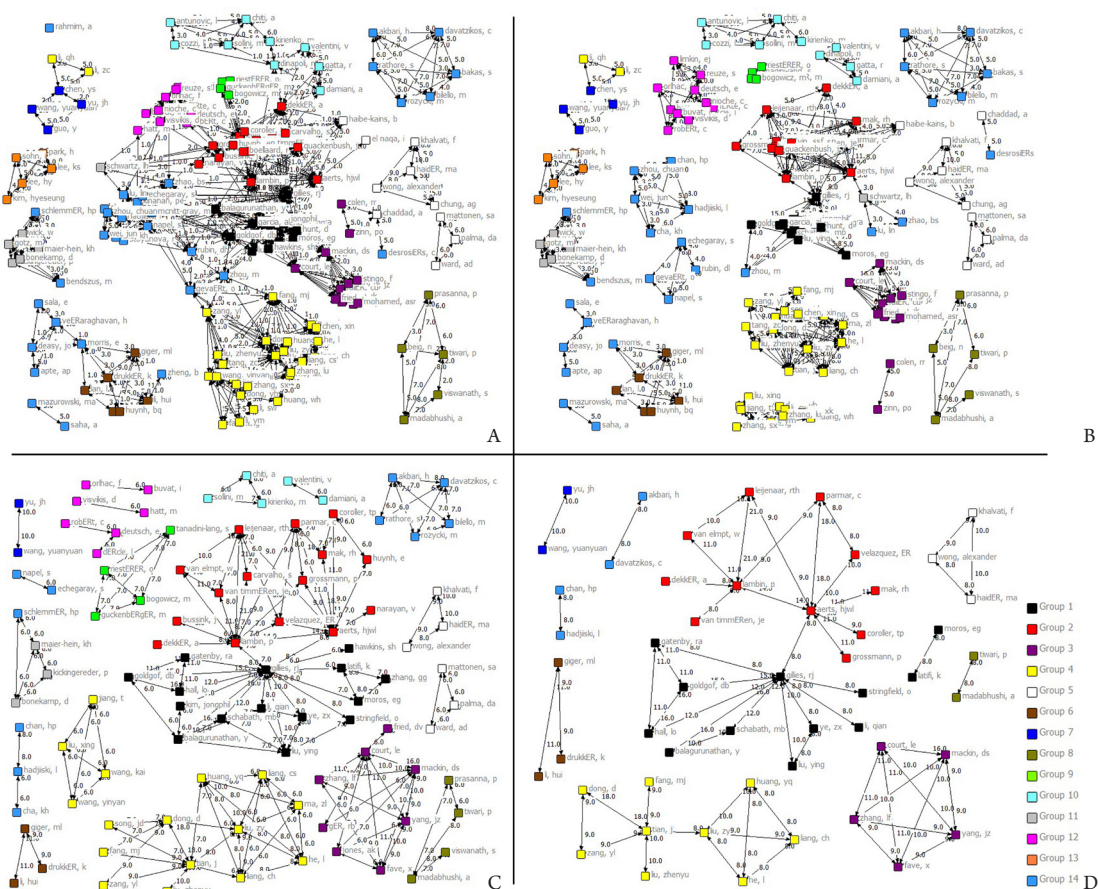


图3 核心著者合作网络知识图谱

Figure 3 Scientific collaboration network map between core authors

A: Co-published papers no less than 1; B: Co-published papers no less than 3; C: Co-published papers no less than 6; D: Co-published papers no less than 8. Authors in Group 1 belong to the University of South Florida (Ye ZX, Liu Ying, and Li Q also belong to Tianjin Medical University); authors in Group 2 belong to Harvard University or Maastricht University; authors in Group 3 belong to the University of Texas MD Anderson Cancer Center, Baylor College of Medicine, or University of Texas Health Science Center at Houston; authors in Group 4 belong to Chinese Academy of Sciences, Guangdong Academy of Medical Sciences and General Hospital, Southern Medical University, Capital Medical University, Northeastern University, Shandong University, etc, which located in China; authors in Group 5 belong to University of Toronto, McGill University, University of Waterloo, University of Western Ontario, École de Technologie Supérieure, etc, which located in Canada; authors in Group 6 belong to University of Chicago; authors in Group 7 belong to Fudan University, Sun Yat-sen University, etc, which located in China; authors in Group 8 belong to Case Western Reserve University; authors in Group 9 belong to University of Zurich; authors in Group 10 belong to Humanitas University, Catholic University of Milan; authors in Group 11 belong to German Cancer Research Center, Heidelberg University, etc, which located in Germany; authors in Group 12 belong to Institut National de la Sante et de la Recherche Medicale (abbreviated as INSERM), University of Paris-Saclay, Gustave Roussy, Université de Bretagne Occidentale, etc, which located in France; authors in Group 13 belong to Sungkyunkwan University; authors in Group 14 belong to others

2.3 研究热点分析

疾病关键词出现频次如下：胶质母细胞瘤 (glioblastoma multiforme, GBM) 55次、肺癌52次、乳腺癌34次、非小细胞肺癌29次、头颈部肿瘤18次、脑肿瘤16次、前列腺癌15次、肺腺癌13次、肝细胞癌12次、子宫颈癌8次、食管癌8次、直肠癌8次、卵巢癌7次、结肠直肠癌6次、阿尔兹海默病5次、肾细胞癌4

次、脂肪肝3次、胃癌3次、肝肿瘤3次、口咽癌3次、胰腺癌3次、原发性中枢神经系统淋巴瘤3次。频次快速增长的方法关键词有机器学习、生存分析、计算机辅助诊断、分类、分割、深度学习、进展、卷积神经网络、计算机辅助检测、特征选择、支持向量机等(表1)。

基于高频关键词构建双聚类图(图4)，有如下研

究热点：1)GBM生存分析及预测、影像基因组学研究；2)肺癌(包括肺腺癌)转移预测；3)基于PET/CT图像纹理特征的肿瘤基因异质性研究；4)基于卷积神经网络和深度学习的乳腺肿瘤影像组学研究、计算机辅助诊断；5)头颈肿瘤放射治疗影像组学研究；6)前列腺癌计算机辅助检测；7)肿瘤(包括肝细胞癌)影像图像分割；8)基于定量影像的影像特征可重复性研究。

表1 高频关键词情况表

Table 1 High-frequency key words on radiomics research					
序号	关键字段	出现频次	序号	关键字段	出现频次
1	radiomics	340	22	segmentation	20
2	MRI	124	23	metastasis	20
3	CT	100	24	deep learning	20
4	texture features	96	25	heterogeneity	20
5	glioblastoma	55	26	prognosis	20
6	lung cancer	52	27	cancer	19
7	machine learning	48	28	head and neck cancer	18
8	PET	46	29	brain tumor	16
9	quantitative imaging	45	30	convolutional neural networks	16
10	biomarkers	41	31	prostate cancer	15
11	PET/CT	35	32	lung adenocarcinoma	13
12	breast cancer	34	33	image processing	12
13	radiogenomics	30	34	hepatocellular carcinoma	12
14	non-small cell lung cancer	29	35	computer-aided detection	12
15	image features	28	36	precision medicine	11
16	prediction	24	37	feature selection	11
17	survival analysis	24	38	support vector machine	10
18	image analysis	23	39	treatment response	10
19	classification	22	40	radiotherapy	10
20	medical imaging	22	41	reproducibility	10
21	computer-aided diagnosis	22			

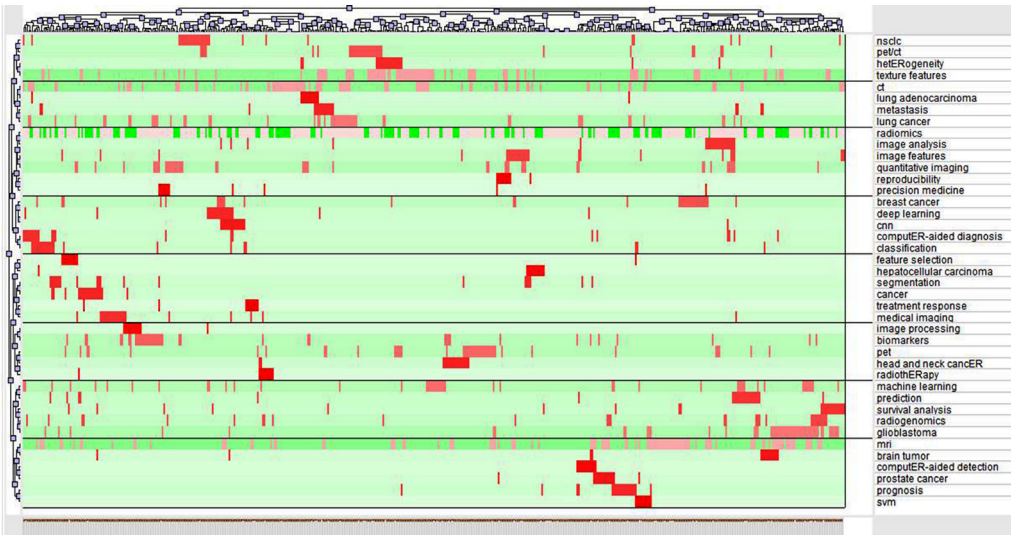


图4 高频关键词双聚类图

Figure 4 Visualized matrix of bi-clustering of high-frequency key words and articles on radiomics research

In the visualization matrix, the rows clustering represent the high-frequency key words clustering, and key words are listed on the right of the matrix. The column clustering represents articles' clustering, and articles are listed at the bottom of the matrix. The colors represent the values in the original datum matrix. It is represented in white when close to zero, whereas the progressively deeper red indicates a larger value, and the progressively deeper green indicates a larger negative value. Different clusterings are separated by black lines

3 讨论

基于发表年代分布可将影像组学发展历程分为两个阶段:1)早期探索及验证期(2012—2014年),发文量缓慢增长,涉及影像组学概念及处理框架提出、方法学研究、肿瘤疾病探索及验证等;2)2015年开始发文量出现逐年大幅增长,涉及多种肿瘤疾病,进入快速增长期。2014年Lambin等^[3]验证了影像组学在肺癌和头颈癌临床预后评估中的效果,认为影像组学具有巨大临床应用前景;同年Doroshov等^[8]指出影像组学是转化医学未来10年的重点发展方向之一。北美放射学会召开了主题为“影像组学:从临床图像到组学”的专题峰会。这些研究和会议促使更多学者关注并开展影像组学研究,发文量逐年攀升。

基于研究力量分布可发现,美国和荷兰在影像组学发展历程中发挥着引领作用,处于研究核心位置;中国、加拿大、德国和法国等为影像组学发展做出了重要贡献。我国2015年发表了首篇影像组学研究论文,随后论文数增长迅速,拥有了一批核心科研机构和著者,但国家(地区)合作网络中心度偏低,提示我国影像组学研究影响力和论文质量有待提升。

基于学科分布和著者所属单位可发现,影像组学研究是医工多学科交叉,涉及影像医学、肿瘤学、计算机科学和生物医学工程等学科领域学者,其研究发展是由国际间多学科合作推动的结果。

恶性实体肿瘤表现为多层次的时空异质性,与其侵袭性生物学行为及个体差异相关,是影响治疗结果、临床结局的重要因素。以CT, PET, PET/CT等为核心的医学影像可以清晰显示病灶组织结构、功能代谢差异;纹理分析通过病灶图像灰度/体素空间分布及关联模式量化中存在的差异进而评估肿瘤异质性,是影像组学研究的一类重要特征。当前影像组学用于多种肿瘤良恶性鉴别、组织病理分型、治疗计划及治疗反应、生存、转移和复发预测;此外有学者开展了基于影像组学路径的乳腺肿瘤及前列腺癌计算机辅助诊断/检测、前列腺癌辅助筛查^[9-10]。在此以肺癌和GBM为代表进行研究热点的进展分析,并简要介绍机器学习及深度学习在影像组学中的应用。

肺癌是影像组学研究最多的疾病,涉及肺癌的病灶分割、组织病理类型、转移、治疗反应、生存期的预测,以及肺结节良恶性鉴别等。Aerts等^[4]发现基于非小细胞肺癌(non-small cell lung cancer, NSCLC)患者CT图像特征的聚类组间有明显临床分期、组织表型和生存期差异,影像组

学标签可用于生存期预测,并与基因表达模式相关,证明了其在肺癌研究中的重要价值。在病灶分割方面,Aerts课题组^[11-12]比较了3D-slicer平台上NSCLC患者CT图像病灶基于区域生长算法的半自动分割与手动分割效果,包括病灶一致性、体积变化、不确定区域大小、病理相关性和特征再现性,发现半自动分割效果优于手动分割。在组织病理类型鉴别方面,肺癌患者CT图像影像组学特征可用于组织学亚型预测^[13-14];有研究者^[15-17]建立了区分肺原位/微侵袭性腺癌和侵袭性腺癌的CT影像组学标签;CT图像影像组学特征也可用于肺癌患者EGFR, KRAS突变以及ALK, ROS1和RET融合突变预测^[18-21]。在转移预测方面,肺癌患者CT图像影像组学特征及标签可用于远处转移及淋巴结转移预测^[14, 22];Zhong等^[23]建立了基于CT图像的肺腺癌隐匿性纵隔淋巴结转移影像组学预测模型及标签。在治疗反应预测方面,早期NSCLC患者在吉非替尼治疗前的CT图像特征可用于预测治疗反应^[24];影像组学特征可用于II~III期NSCLC患者新辅助放射和化学药物治疗后病理完全缓解和肉眼残余病灶预测^[25];Luo等^[26]基于NSCLC患者放射治疗前SNPs, mRNAs, 细胞因子, PET影像组学特征,以及肺和肿瘤广义等效均匀剂量数据构建了能够预测放射治疗后肿瘤局部控制或放射性肺炎的多目标贝叶斯网络模型。在生存预测方面,NSCLC患者PET/CT的CT图像纹理异质性特征可用于患者生存预测^[27];早期NSCLC患者立体定向放射治疗前CT图像影像组学特征与总体生存期相关^[28];一些影像组学标签可用于早期(I~II期)NSCLC患者无病生存期的预测^[29]。在肺结节良恶性鉴别方面,用于肺癌筛查的低剂量CT图像特征被发现也可用于预测肺结节癌变^[30];CT图像特征和支持向量机(support vector machine, SVM)算法可用于肺结节良恶性的鉴别^[31];学者^[32]使用卷积神经网络提取肺结节CT图像深度特征,发现其与经典影像组学特征融合后具有更好的结节良恶性分类效能。

GBM具有肿瘤异质性显著、分子分型多^[33]、中位生存期短等特点,影像特征-生物信息、生存分析成为GBM影像组学研究的重点。研究^[34-36]发现GBM患者术前MRI伦勃朗视觉感受图像[Visually Accesable REMBRANDT(Repository for Molecular Brain Neoplasia Data) images, VASARI]特征、增强肿瘤坏死和瘤周水肿区影像特征、体积特征与生存期、肿瘤基因表达模式和生物学行为密切相关,可用于肿瘤分子亚型预测。在影像特征-生物信息研究方面, MRI影像组学特征可用于表皮生长因子受体(epidermal growth factor receptor, EGFR)^[37-38], 异柠檬酸脱氢酶(isocitrate dehydrogenase, IDH)和1p/19q^[39-40], MGMT(O-6-

methylguanine-DNA methyltransferase)^[41-42], p53突变^[43]的预测。在生存分析方面, VASARI特征^[44-45]、体积相关特征^[46]、病灶空间多样性特征^[38]、纹理特征^[47-50]可用于GBM患者生存期预测; 对于新诊断未标准治疗的GBM患者^[51-52]和贝伐单抗治疗前的复发性GBM患者^[53]建立了预测无进展生存期和总生存期的影像组学标签和模型; Lao等^[54]使用迁移学习方法建立了可用于GBM患者总生存期预测的卷积神经网络深度特征影像组学标签和模型。

机器学习、深度学习(尤其是卷积神经网络、迁移学习等)在特征筛选、发现特征关系方面具有独特优势, 尤其深度学习能够从图像中自动提取影像特征(或许比人工定义特征更有效), 不需要太多背景知识和复杂的数据处理流程, 从而能有效提升影像组学模型/标签效能。包括支持向量机、卷积神经网络、人工神经网络、贝叶斯网络、决策树、K近邻、logistics回归、随机森林、boosting等机器学习方法在影像组学研究中得到广泛应用。Parmar^[55-56], Wu^[13], Ypsilantis^[57]等学者分别在头颈肿瘤生存期预测、肺癌组织分类、新辅助化学药物治疗反应等影像组学分析中对多种机器学习算法效果进行了对比研究。深度学习方法在影像组学相关的分类、感兴趣区确定、图像分割、特征提取、图像配准、图像重建等方面得到了应用, 涉及深度卷积神经网络(包括迁移学习方法)、生成式对抗网络(generative adversarial network, GAN)等深度学习网络框架。

影像组学模型常优于传统临床模型, 影像特征可作为临床、基因组学数据的补充用于肿瘤的临床决策支持及精准诊疗, 但其稳定性、可重复性和成本效益^[58]需通过多阶段、多方面验证和探索^[59]。目前影像组学研究和临床应用尚存在如下问题^[60-61]:

1) 处理过程缺乏标准和规范^[62], 结果可能不易重复, 泛化能力不强。影像特征受处理过程多种因素影响, 包括成像设备、扫描参数、器官(运动)状态、图像预处理方法、图像分割结果等; 特征提取方法体系尚未标准化, 存在多种特征定义、计算公式及提取算法; 影像组学软件种类多, 尽管有研究^[63]表明不同软件聚类结果存在一致性, 但图像处理、特征提取及筛选、模型构建存在差异。发表研究结果时附录详细处理过程、算法及参数, 共享原始数据, 及筛选稳定的组学特征是解决上述问题的重要措施。其他学者可据此验证、使用和提升研究结果, 促进处理过程优化、规范化和标准化。Lambin等^[3]提出了影像组学质量评分体系(the radiomics quality score, RQS)来评价、规范和提升影像组学研究,

RQS列出了影像组学各个步骤的质量要求及评分细则, 但操作相对复杂, 评价具有主观性, 部分评价内容只适用于特定主题研究。有学者^[64-65]在相关系统综述研究中使用RQS进行影像组学研究质量评价, 发现评价员间的RQS评分存在一定差异, 大多数影像组学研究的多项评分较低, 涉及病灶分割、结果验证、评判金标准、潜在临床应用、费用分析和数据开放共享等方面。

2) 目前有50~5 000个定量特征, 但患者样本量仍较小, 且多为单中心回顾性研究, 导致特征选择偏向、模型过拟合或欠拟合、研究结果假阳性率高。原始图像数据共享、多中心研究是解决这个问题的重要途径。多个国际学术组织正在进行推动原始图像共享(形成共享库)及标准化工作; 处理流程的规范可保证多中心研究图像、数据和结果的一致性、有效性。国家/国际图像共享网络(关联病理、临床和组学数据)、参考病例数据库将是影像组学研究的重要基础。

3) 部分影像特征与肿瘤异质性、基因分型或表达模式等生物学特征相关, 但其潜在生物学机制仍不清楚, 需通过体外实验或动物实验等进一步探讨和验证。

此外深度学习应用目前亦存在一些问题^[66-68]: 黑箱问题, 决策过程不透明, 很难从临床角度进行解释; 模型效能严重依赖于训练样本数量及质量, 模型复杂、参数多, 训练过程中需要关注过拟合问题; 可复用性差, 训练好的模型被应用到新的环境中时需要迁移学习以保证其性能。但这些问题不影响深度学习在影像组学研究中的探索应用, 一些学者^[69-70]亦就上述问题进行研究并提出了部分解决方案。

病灶分割多采用手动或半自动分割, 其准确性受操作人员经验、环境和工作量影响, 可重复性较差、耗时长, 精确、高效和可重复的全自动、半自动分割算法将是影像组学研究的重要内容。

影像组学作为一个新兴的研究方向处于快速发展阶段, 越来越多的国家、科研机构及学者参与其中, 但相关研究仍存在一些问题。随着处理流程规范化和标准化、影像共享网络建设、参考病例数据集建设、自动/半自动图像分割方法开发、高效分析预测方法应用, 以及基于影像生物标志物路线图^[59]的验证及确认, 影像组学将在真正意义上成为临床辅助决策工具, 为肿瘤个体化/精准诊疗提供帮助和指导。

利益冲突声明: 作者声称无任何利益冲突。

参考文献

- [1] Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data[J]. *Radiology*, 2016, 278(2): 563-577.
- [2] 苏会芳, 周国锋, 谢传森, 等. 放射组学的兴起和研究进展[J]. *中华医学杂志*, 2015, 95(7): 553-556.
- SU Huifang, ZHOU Guofeng, XIE Chuanmiao, et al. The rise and research progress of radiomics[J]. *National Medical Journal of China*, 2015, 95(7): 553-556.
- [3] Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine[J]. *Nat Rev Clin Oncol*, 2017, 14(12): 749-762.
- [4] Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach[J]. *Nat Commun*, 2014, 5: 4006.
- [5] 曹霞, 崔雷. 基于SNA的国外医学信息学领域合著网络研究[J]. *现代情报*, 2016, 36(3): 129-134.
- CAO Xia, CUI Lei. Co-authorship network research of foreign medical informatics based on social network analysis[J]. *Modern Information*, 2016, 36(3): 129-134.
- [6] 方丽, 崔雷. 利用双聚类 and 突发检测算法探测学科前沿及知识基础的比较分析——以h指数研究领域为例[J]. *情报杂志*, 2015, 34(2): 79-83.
- FANG Li, CUI Lei. Comparison of Biclustering analysis and burst detection algorithm in detecting research fronts and intellectual base: A case study of h-index[J]. *Journal of Intelligence*, 2015, 34(2): 79-83.
- [7] 王敏, 宋彬, 黄子星, 等. 基于文献计量学的近5年放射组学研究热点分析[J]. *中华放射学杂志*, 2016, 50(12): 974-977.
- WANG Min, SONG Bin, HUANG Zixing, et al. Bibliometric and research hotspots analysis of radiomics in recent five years[J]. *Chinese Journal of Radiology*, 2016, 50(12): 974-977.
- [8] Doroshow JH, Kummar S. Translational research in oncology—10 years of progress and future prospects[J]. *Nat Rev Clin Oncol*, 2014, 11(11): 649-662.
- [9] Wang J, Wu CJ, Bao ML, et al. Machine learning-based analysis of MR radiomics can help to improve the diagnostic performance of PI-RADS v2 in clinically relevant prostate cancer[J]. *Eur Radiol*, 2017, 27(10): 4082-4090.
- [10] Algohary A, Viswanath S, Shiradkar R, et al. Radiomic features on MRI enable risk categorization of prostate cancer patients on active surveillance: Preliminary findings[J]. *J Magn Reson Imaging*, 2018, 48(3): 818-828.
- [11] Velazquez ER, Parmar C, Jermoumi M, et al. Volumetric CT-based segmentation of NSCLC using 3D-Slicer[J]. *Sci Rep*, 2013, 3: 3529.
- [12] Parmar C, Rios Velazquez E, Leijenaar R, et al. Robust Radiomics feature quantification using semiautomatic volumetric segmentation[J]. *PLoS One*, 2014, 9(7): e102107.
- [13] Wu W, Parmar C, Grossmann P, et al. Exploratory study to identify radiomics classifiers for lung cancer histology[J]. *Front Oncol*, 2016, 6: 71.
- [14] Ferreira Junior JR, Koenigkam-Santos M, Cipriano FEG, et al. Radiomics-based features for pattern recognition of lung cancer histopathology and metastases[J]. *Comput Methods Programs Biomed*, 2018, 159: 23-30.
- [15] She Y, Zhang L, Zhu H, et al. The predictive value of CT-based radiomics in differentiating indolent from invasive lung adenocarcinoma in patients with pulmonary nodules[J]. *Eur Radiol*, 2018, 28(12): 5121-5128.
- [16] Fan L, Fang M, Li Z, et al. Radiomics signature: a biomarker for the preoperative discrimination of lung invasive adenocarcinoma manifesting as a ground-glass nodule[J]. *Eur Radiol*, 2018, 29(2): 889-897.
- [17] Yuan M, Zhang YD, Pu XH, et al. Comparison of a radiomic biomarker with volumetric analysis for decoding tumour phenotypes of lung adenocarcinoma with different disease-specific survival[J]. *Eur Radiol*, 2017, 27(11): 4857-4865.
- [18] Zhang L, Chen B, Liu X, et al. Quantitative biomarkers for prediction of epidermal growth factor receptor mutation in non-small cell lung cancer[J]. *Transl Oncol*, 2018, 11(1): 94-101.
- [19] Gevaert O, Echegaray S, Khuong A, et al. Predictive radiogenomics modeling of EGFR mutation status in lung cancer[J]. *Sci Rep*, 2017, 7: 41674.
- [20] Liu Y, Kim J, Balagurunathan Y, et al. Radiomic features are associated with EGFR mutation status in lung adenocarcinomas[J]. *Clin Lung Cancer*, 2016, 17(5): 441-448.e6.
- [21] Yoon HJ, Sohn I, Cho JH, et al. Decoding tumor phenotypes for ALK, ROS1, and RET fusions in lung adenocarcinoma using a radiomics approach[J]. *Medicine (Baltimore)*, 2015, 94(41): e1753.
- [22] Coroller TP, Grossmann P, Hou Y, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma[J]. *Radiother Oncol*, 2015, 114(3): 345-350.
- [23] Zhong Y, Yuan M, Zhang T, et al. Radiomics approach to prediction of occult mediastinal lymph node metastasis of lung adenocarcinoma[J]. *AJR Am J Roentgenol*, 2018, 211(1): 109-113.
- [24] Aerts HJ, Grossmann P, Tan Y, et al. Defining a radiomic response phenotype: A pilot study using targeted therapy in NSCLC[J]. *Sci Rep*, 2016, 6: 33860.
- [25] Coroller TP, Agrawal V, Narayan V, et al. Radiomic phenotype features predict pathological response in non-small cell lung cancer[J]. *Radiother Oncol*, 2016, 119(3): 480-486.
- [26] Luo Y, McShan DL, Matuszak MM, et al. A multiobjective Bayesian networks approach for joint prediction of tumor local control and radiation pneumonitis in nonsmall-cell lung cancer (NSCLC) for

- response-adapted radiotherapy[J]. *Med Phys*, 2018, 45(8): 3980-3995.
- [27] Win T, Miles KA, Janes SM, et al. Tumor heterogeneity and permeability as measured on the CT component of PET/CT predict survival in patients with non-small cell lung cancer[J]. *Clin Cancer Res*, 2013, 19(13): 3591-3599.
- [28] Huynh E, Coroller TP, Narayan V, CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer[J]. *Radiother Oncol*, 2016, 120(2): 258-266.
- [29] Huang Y, Liu Z, He L, et al. Radiomics signature: A potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer[J]. *Radiology*, 2016, 281(3): 947-957.
- [30] Hawkins S, Wang H, Liu Y, et al. Predicting malignant nodules from screening CT scans[J]. *J Thorac Oncol*, 2016, 11(12): 2120-2128.
- [31] Jun Wang, Xia Liu, Di Dong, et al. Prediction of malignant and benign of lung tumor using a quantitative radiomic method[J]. *Conf Proc IEEE Eng Med Biol Soc*, 2016, 2016: 1272-1275.
- [32] Paul R, Hawkins SH, Schabath MB, et al. Predicting malignant nodules by fusing deep features with classical radiomics features[J]. *J Med Imaging (Bellingham)*, 2018, 5(1): 011021.
- [33] 冷月爽, 王小宜, 廖伟华, 等. 影像组学在胶质瘤临床研究中的应用进展[J]. *中南大学学报(医学版)*, 2018, 43(4): 354-359.
- LENG Yueshuang, WANG Xiaoyi, LIAO Weihua, et al. Radiomics in gliomas: A promising assistance for glioma clinical research[J]. *Journal of Central South University. Medical Science*, 2018, 43(4): 354-359.
- [34] Gutman DA, Cooper LA, Hwang SN, et al. MR imaging predictors of molecular profile and survival: multi-institutional study of the TCGA glioblastoma data set[J]. *Radiology*, 2013, 267(2): 560-569.
- [35] Gevaert O, Mitchell LA, Achrol AS, et al. Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features[J]. *Radiology*, 2014, 273(1): 168-174.
- [36] Grossmann P, Gutman DA, Dunn WD Jr, et al. Imaging-genomics reveals driving pathways of MRI derived volumetric tumor phenotype features in Glioblastoma[J]. *BMC Cancer*, 2016, 16: 611.
- [37] Akbari H, Bakas S, Pisapia JM, et al. In vivo evaluation of EGFRvIII mutation in primary glioblastoma patients via complex multiparametric MRI signature[J]. *Neuro Oncol*, 2018, 20(8): 1068-1079.
- [38] Lee J, Narang S, Martinez J, et al. Spatial habitat features derived from multiparametric magnetic resonance imaging data are associated with molecular subtype and 12-month survival status in glioblastoma multiforme[J]. *PLoS One*, 2015, 10(9): e0136557.
- [39] Lu CF, Hsu FT, Hsieh KL, et al. Machine learning-based radiomics for molecular subtyping of gliomas[J]. *Clin Cancer Res*, 2018, 24(18): 4429-4436.
- [40] Shofty B, Artzi M, Ben Bashat D, et al. MRI radiomics analysis of molecular alterations in low-grade gliomas[J]. *Int J Comput Assist Radiol Surg*, 2018, 13(4): 563-571.
- [41] Li ZC, Bai H, Sun Q, et al. Multiregional radiomics features from multiparametric MRI for prediction of MGMT methylation status in glioblastoma multiforme: A multicentre study[J]. *Eur Radiol*, 2018, 28(9): 3640-3650.
- [42] Xi YB, Guo F, Xu ZL, et al. Radiomics signature: A potential biomarker for the prediction of MGMT promoter methylation in glioblastoma[J]. *J Magn Reson Imaging*, 2018, 47(5): 1380-1387.
- [43] Li Y, Qian Z, Xu K, et al. MRI features predict p53 status in lower-grade gliomas via a machine-learning approach[J]. *Neuroimage Clin*, 2017, 17: 306-311.
- [44] Mazurowski MA, Desjardins A, Malof JM. Imaging descriptors improve the predictive power of survival models for glioblastoma patients[J]. *Neuro Oncol*, 2013, 15(10): 1389-1394.
- [45] Peeken JC, Hesse J, Haller B, et al. Semantic imaging features predict disease progression and survival in glioblastoma multiforme patients[J]. *Strahlenther Onkol*, 2018, 194(6): 580-590.
- [46] Mazurowski MA, Zhang J, Peters KB, et al. Computer-extracted MR imaging features are associated with survival in glioblastoma patients[J]. *J Neurooncol*, 2014, 120(3): 483-488.
- [47] Beig N, Patel J, Prasanna P, et al. Radiogenomic analysis of hypoxia pathway is predictive of overall survival in Glioblastoma[J]. *Sci Rep*, 2018, 8: 7.
- [48] Chaddad A, Sabri S, Niazi T, et al. Prediction of survival with multi-scale radiomic analysis in glioblastoma patients[J]. *Med Biol Eng Comput*, 2018, 56(12): 2287-2300.
- [49] Upadhyaya T, Morvan Y, Stindel E, et al. A framework for multimodal imaging-based prognostic model building: Preliminary study on multimodal MRI in Glioblastoma Multiforme[J]. *IRBM*, 2015, 36(6): 345-350.
- [50] Prasanna P, Patel J, Partovi S, et al. Radiomic features from the peritumoral brain parenchyma on treatment-naïve multi-parametric MR imaging predict long versus short-term survival in glioblastoma multiforme: Preliminary findings[J]. *Eur Radiol*, 2017, 27(10): 4188-4197.
- [51] Kickingereder P, Neuberger U, Bonekamp D, et al. Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma[J]. *Neuro Oncol*, 2018, 20(6): 848-857.
- [52] Kickingereder P, Burth S, Wick A, et al. Radiomic profiling of glioblastoma: identifying an imaging predictor of patient survival with improved performance over established clinical and radiologic risk models[J]. *Radiology*, 2016, 280(3): 880-889.
- [53] Kickingereder P, Götz M, Muschelli J, et al. Large-scale radiomic profiling of recurrent glioblastoma identifies an imaging predictor for

- stratifying anti-angiogenic treatment response[J]. Clin Cancer Res, 2016, 22(23): 5765-5771.
- [54] Lao J, Chen Y, Li ZC, et al. A deep learning-based radiomics model for prediction of survival in glioblastoma multiforme[J]. Sci Rep, 2017, 7: 10353.
- [55] Parmar C, Grossmann P, Bussink J, et al. Machine learning methods for quantitative radiomic biomarkers[J]. Sci Rep, 2015, 5: 13087.
- [56] Parmar C, Grossmann P, Rietveld D, et al. Radiomic machine-learning classifiers for prognostic biomarkers of head and neck cancer[J]. Front Oncol, 2015, 5: 272.
- [57] Ypsilantis 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.
- [58] Keek SA, Leijenaar RT, Jochims A, et al. A review on radiomics and the future of theragnostics for patient selection in precision medicine[J]. Br J Radiol, 2018, 91(1091): 20170926.
- [59] O'Connor JP, Aboagye EO, Adams JE, et al. Imaging biomarker roadmap for cancer studies[J]. Nat Rev Clin Oncol, 2017, 14(3): 169-186.
- [60] Avanzo M, Stancanello J, El Naqa I. Beyond imaging: The promise of radiomics[J]. Phys Med, 2017, 38: 122-139.
- [61] Limkin EJ, Sun R, Dercle L, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology[J]. Ann Oncol, 2017, 28(6): 1191-1206.
- [62] 仇清涛, 段敬豪, 巩贯忠, 等. 影像组学可重复性问题研究进展[J]. 中华放射肿瘤学杂志, 2018, 27(3): 327-330.
- QIU Qingtao, DUAN Jinghao, GONG Guanzhong, et al. Research progress on radiomics reproducibility[J]. Chinese Journal of Radiation Oncology, 2018, 27(3): 327-330.
- [63] Gan J, Wang J, Zhong H, et al. MO-DE-207B-09: A Consistent test for radiomics softwares[J]. Medical Physics, 2016, 43(6): 3706.
- [64] Sanduleanu S, Woodruff HC, de Jong EEC, et al. Tracking tumor biology with radiomics: A systematic review utilizing a radiomics quality score[J]. Radiother Oncol, 2018, 127(3): 349-360.
- [65] Valdora F, Houssami N, Rossi F, et al. Rapid review: radiomics and breast cancer[J]. Breast Cancer Res Treat, 2018, 169(2): 217-229.
- [66] Savadjiev P, Chong J, Dohan A, et al. Demystification of AI-driven medical image interpretation: past, present and future[J]. Eur Radiol, 2019, 29(3): 1616-1624.
- [67] 李赞铎, 宫恩浩, 李睿, 等. 深度学习技术与医学影像——现状及未来[J]. 中华放射学杂志, 2018, 52(5): 321-326.
- LI Yunduo, GONG Enhao, LI Rui, et al. Deep learning technology and medical imaging—present and future[J]. Chinese Journal of Radiology, 2018, 52(5): 321-326.
- [68] Thrall JH, Li X, Li Q, et al. Artificial intelligence and machine learning in radiology: Opportunities, challenges, pitfalls, and criteria for success[J]. J Am Coll Radiol, 2018, 15(3 Pt B): 504-508.
- [69] Paul R, Schabath M, Balagurunathan Y, et al. Explaining deep features using radiologist-defined semantic features and traditional quantitative features[J]. Tomography, 2019, 5(1): 192-200.
- [70] Vial A, Stirling D, Field M, et al. The role of deep learning and radiomic feature extraction in cancer-specific predictive modelling: a review[J]. Transl Cancer Res, 2018, 7(3): 803-816.
- (本文编辑 陈丽文)

本文引用: 罗爱静, 姚山虎, 冯智超, 容鹏飞, 覃岳香, 王维. 基于知识图谱的影像组学研究可视化分析[J]. 中南大学学报(医学版), 2019, 44(3): 233-243. DOI:10.11817/j.issn.1672-7347.2019.03.002

Cite this article as: LUO Aijing, YAO Shanhu, FENG Zhichao, RONG Pengfei, QIN Yuexiang, WANG Wei. Visualization analysis for radiomics research based on knowledge mapping[J]. Journal of Central South University. Medical Science, 2019, 44(3): 233-243. DOI:10.11817/j.issn.1672-7347.2019.03.002