

组学大数据和医学人工智能

王昕玥^{1,2,3}, 渠鸿竹^{1,2,3}, 方向东^{1,2,3}

1. 中国科学院北京基因组研究所(国家生物信息中心), 中国科学院基因组科学与信息重点实验室, 北京 100101
2. 中国科学院大学, 北京 100049
3. 中国科学院北京基因组研究所(国家生物信息中心), 基因组与精准医学检测技术北京市重点实验室, 北京 100101

摘要: 随着高通量测序技术和计算机科学的飞速发展, 组学数据量指数倍增长, 多组学分析优势逐渐显现, 人工智能应用也愈加广泛。本文介绍了近年来多组学数据分析和人工智能各自在医学领域的应用进展, 同时也介绍了两者相结合应用的案例以及优势, 最后简单阐述多组学分析和人工智能在现阶段面临的挑战, 旨在为医学行业提供新的研究思路, 助推精准医学发展应用。

关键词: 多组学; 人工智能; 医学; 精准医学

Omics big data and medical artificial intelligence

Xinyue Wang^{1,2,3}, Hongzhu Qu^{1,2,3}, Xiangdong Fang^{1,2,3}

1. CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences/China National Center for Bioinformation, Beijing 100101, China
2. University of Chinese Academy of Sciences, Beijing 100049, China
3. Beijing Key Laboratory of Genome and Precision Medicine Technologies, Beijing Institute of Genomics, Chinese Academy of Sciences/China National Center for Bioinformation, Beijing 100101, China

Abstract: With the rapid development of high-throughput sequencing technology and computer science, the amount of large omics data has increased exponentially, the advantages of multi-omics analysis have gradually emerged, and the application of artificial intelligence has become more and more extensive. In this review, we introduce the application progress of multi-omics data analysis and artificial intelligence in the medical field in recent years, and also show the cases and advantages of their combined application. Finally, we briefly explain the current challenges of multi-omics analysis and artificial intelligence in order to provide new research ideas for the medical industry and to promote the development and application of precision medicine.

Keywords: multi-omics; artificial intelligence; medicine; precision medicine

收稿日期: 2021-06-18; 修回日期: 2021-08-31

基金项目: 科技部重点研发计划(编号: 2020YFC2003405, 2016YFC0901700, 2018YFC0910700)资助[Supported by the National Key Research and Development Project of the Ministry of Science and Technology of the People's Republic of China (Nos. 2020YFC2003405, 2016YFC0901700, 2018YFC0910700)]

作者简介: 王昕玥, 在读博士研究生, 研究方向: 基因组学。E-mail: wangxinyue2019d@big.ac.cn

通讯作者: 渠鸿竹, 博士, 副研究员, 研究方向: 基因组学、精准医学大数据。E-mail: quhongzhu@big.ac.cn

方向东, 博士, 研究员, 研究方向: 医学遗传学、精准医学大数据。E-mail: fangxd@big.ac.cn

DOI: 10.16288/j.ycz.21-215

网络出版时间: 2021/10/11 09:01:14

URI: <https://kns.cnki.net/kcms/detail/11.1913.r.20211009.2351.001.html>

随着 5G 技术的成熟,大数据、物联网、云计算等也在快速发展,社会逐渐从信息时代进入智能时代,各行各业也都在争相推进智能化的脚步。在医疗行业中,医学影像、病患信息等数据海量且复杂,人工整理或解读效率低下,而人工智能可以快速精确地处理大数据,并挖掘其背后的潜在信息。与此同时,随着高通量测序技术的发展,单组学分析技术日益成熟与完善,而多组学大数据的整合分析,已成为研究者们探索生命机制的新方向,在此大背景下,结合组学数据的医疗行业的智能化发展建设已成必然趋势。本文介绍了近年来多组学整合分析和人工智能在医学领域中各自的应用以及结合应用的最新研究成果。

1 组学和人工智能的概念

组学(omics)主要包括基因组学(genomics)、转录组学(transcriptomics)、表观组学(epigenomics)、蛋白组学(proteomics)、代谢组学(metabolomics)、脂类组学(lipidomics)、免疫组学(immunomics)、糖组学(glycomics)等,是各类组学的统称。每一类组学为此类特征的系统集合,比如第一个被定义的基因组学,是对一个生物体所有基因进行集体表征和量化,并研究它们之间的相互关系及对生物体的影响。目前,单一组学的研究比较成熟,但挖掘到的数据信息受限,只能在某一层面解释特征与目标疾病的相关性,无法探究因果关系。多组学的整合分析,可使人们更宏观地了解到目标疾病的全貌,理解目标疾病的发生发展机制,为攻克疑难杂症提供更全面的信息^[1]。

人工智能(artificial intelligence, AI)是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学^[2]。1950 年,“人工智能之父”图灵提出计算机是否拥有类智能这一问题。1956 年,美国达特茅斯会议首次提出“人工智能”这一术语,标志着人工智能这一新兴学科的正式诞生。其本质是通过大量的样本训练来模拟人脑的思维方式,或用以执行人类的行为活动。其研究涵盖领域十分广泛,主要分为自然语言处理、机器学习(machine learning, ML)、机器人、计算机视

觉、语言图像识别和知识图谱六大方向。目前,专家系统、护理机器人、手术机器人、计算机辅助诊断等产品都在医疗行业中起着重要作用^[3]。

机器学习是一种可用于实现人工智能的方法,其传统的算法包括决策树、逻辑回归、聚类、随机森林、支持向量机、深度学习等^[4]。从方法上来分,机器学习算法可以分为监督学习、无监督学习、半监督学习、深度学习和强化学习。其中,深度学习(deep learning, DL)是建立于模拟人脑进行分析学习的神经网络,典型的深度学习模型有循环神经网络、递归神经网络、卷积神经网络、深度信任网络、生成式对抗网络和自编码器等。人工智能、机器学习和深度学习的关系如图 1 所示。

2 组学大数据与人工智能在医学中的应用

2.1 多组学在医学中的应用

随着测序技术的发展及组学新技术的不断涌现,不同种类的组学数据指数级增长,对多组学大数据的整合分析,已成为科学家探索生命机制和疾病演变的新方向。

通过 DNA 甲基化和基因表达数据的整合分析,可以鉴定疾病发生发展相关的分子标志物。Mens 等^[5]



图 1 人工智能、机器学习和深度学习的关系示意图
Fig. 1 Schematic diagram of the relationship between artificial intelligence, machine learning, and deep learning

使用了来自大规模全基因组关联研究(genome-wide association study, GWAS)的公开数据来搜索与各种心脏代谢特征相关的 miRNA 相关序列的遗传变异,包括脂质和肥胖相关特征、血糖指数、血压、2 型糖尿病(type 2 diabetes, T2D)和冠心病(coronary heart disease, CHD)患病率,发现了 67 个已鉴定 miRNA 中的 180 个 SNP 与心脏代谢特征相关。然后,文章使用鹿特丹研究(The Rotterdam Study)参与者的 DNA 甲基化和 miRNA 表达数据进一步研究相关 miRNA 与心脏代谢特征之间的联系,其中 38 个 miRNA 的 CpG 位点的 DNA 甲基化水平与研究性状相关。此外作者从 67 个已鉴定 miRNA 中进一步发现了与研究性状相关的 8 个 miRNA。整合不同组学结果表明 miR-10b-5p、miR-148a-3p、miR-125b-5p 和 miR-100-5p 与心脏代谢的脂质性状密切相关,这些可以被视为潜在的 2 型糖尿病和冠心病早期诊断或进展的生物标志物。

Yuan 等^[6]使用多个项目的转录组和/或 DNA 甲基化数据证明了 45 种结直肠癌(colorectal cancer, CRC)亚型中 29 种的易感性可能是由基因调节的顺式作用介导的,并且确定了 66 个推测的易感基因。*DIP2B* 和 *SFMBT1* 通过破坏细胞行为(包括迁移、侵袭和上皮间质转化)在 CRC 的发生中起着至关重要的作用。

以序列为中心的蛋白质组、基因组和转录组数据的整合分析,可以为基因表达调控、信号网络、疾病亚型和临床预测提供新的见解。Cohen 等^[7]结合游离 DNA 突变和循环蛋白质生物标志物开发了一种新的基于血液的预测方法 CancerSEEK,不仅可以实现癌症早诊,还可以定位这些癌症的起源器官。研究将 CancerSEEK 应用于 1 005 名患有卵巢癌、肝癌、胃癌、胰腺癌、食道癌、结直肠癌、肺癌或乳腺癌的患者上,能够定位癌症的起源器官并鉴定出 5 种肿瘤类型(卵巢癌、肝癌、胃癌、胰腺癌和食道癌)的早期存在,其灵敏度介于 69%~98%,特异性为 99%。

Yang 等^[8]为了确定早期结 CRC 的潜在靶点,对来自 II 期 CRC 患者进行了无标记蛋白质组学分析,共鉴定出 2 968 种蛋白质,再从癌症基因组图谱(The Cancer Genome Atlas, TCGA)结肠腺癌库中检索到相应的 RNA 测序数据,筛选出 111 种关键候选蛋白,

最终由病理图像数据确定了三个潜在靶点:蛋白质精氨酸脱亚胺酶 2 (PADI2)、IgG 结合蛋白的 Fc 片段(FCGBP)和磷酸丝氨酸转氨酶 1。

肉瘤代表一组高度组织学和分子异质性的罕见恶性肿瘤,预后较差。Zhang 等^[9]从癌症蛋白质组图谱(The Cancer Proteome Atlas, TCPA)上下载了肉瘤患者的蛋白质组表达谱以及临床信息,发现 55 种蛋白质与患者的总生存期(overall survival, OS)相关。基于蛋白质组学特征开发了肉瘤患者的预后模型,模型包括七种蛋白: AMPKALPHA、CHK1、S6、ARID1A、RBM15、ACETYLATUBULINLYS40 和 MSH6。再使用 TCGA 中肉瘤患者的转录组数据集验证预后模型的性能,证明模型可能是指导临床实践的有效工具。

这些研究表明,多组学大数据的集成为了解跨细胞组织多个层面的因果关系提供了机会,在医学领域应用中具有巨大潜力,可以得到比单组学分析更全面,更精确的结果。

2.2 人工智能在医学中的应用

2.2.1 人工智能在影像中的应用

医学影像主要包括医学影像计算机断层扫描(computer tomography, CT)、磁共振成像(magnetic resonance imaging, MRI)、正电子发射计算机断层显像(positron emission tomography-computer tomography, PET-CT)、X 射线、超声等技术^[10]。人工智能在医学影像上已经应用到图像分割、图像分类、图像配准和目标检测等场景中^[11],其可以实现图像质量的改善提升,且对于理解图像信息,辅助诊断分类都起着极大的作用,还可以迅速完成图像分割配准等更高级的操作。

在图像配准方面,De Silva 等^[12]使用了一种可以纵向对齐多模态视网膜图像的深度学习配准算法。在彩色眼底照相(color fundus photography, CFP)、眼底自发荧光(fundus autofluorescence, FAF)和红外反射(infrared reflectance, IR)三种图像模式实验中,单模态纵向配准实验获得 54~59 μm 的误差,相较于传统的配准方法(平均误差在 39~53 μm 范围内),展示了更高的准确性。对于多模态横断面配准实验,该方法平均误差在 66~69 μm 范围内,而传统方法则错

误颇多。在图像分类方面, Jiang 等^[13]对 111 例乳房动态对比材料增强(dynamic contrast material-enhanced, DCE)MRI 检查的图像分别通过传统的计算机辅助评估软件和 AI 进行分类测试, 发现使用 AI 方法图像分类准确性更高, 平均 AUC 从 0.71 提高到 0.76。在图像分割方面, Hoseini 等^[14]提出了一种大容量深度卷积神经网络(deep convolutional neural network, DCNN)的方法, 可以更准确, 更快速地分割脑部 MRI 图像, 在 BraTS2016 脑肿瘤数据集的完整区域、核心区域和增强区域的准确性分别为 0.90、0.85 和 0.84。

2.2.2 人工智能在癌症研究中的应用

人工智能在肿瘤学研究的各个方面的应用增长迅速, 应用范围包括癌症风险预测、癌症检测和分类分期、癌症药物发现和再利用、预后分析等, 有助于医生为每位患者制定高度个性化的癌症预防和治疗计划。

对于癌症诊断, 主要是对癌症患者的图像和非癌性病变的影像学图像利用人工智能方法提取差异特征, 构建诊断模型。Jeyaraj 等^[15]开发了一种具有两个分层的新结构深度卷积神经网络(convolutional neural network, CNN), 分析样本多维高光谱图像辅助口腔癌诊断, 准确度为 94.5%。Luo 等^[16]利用 1 036 496 个内窥镜图像开发了一种胃肠道人工智能诊断系统(GRAIDS), 经外部验证和与专业内镜医师的诊断结果比较, GRAIDS 表现出良好的诊断性能, 其准确性为 0.915~0.977, 灵敏度为 0.942, 堪比专业内镜医师(0.945)。Li 等^[17]开发了一个基于 CNN 的新模型来分析窄带成像放大内窥镜(magnifying endoscopy with narrow band imaging, M-NBI)观察到的胃粘膜病变图像, 用于胃癌的早期诊断, 该模型准确率、敏感性和特异性分别为 90.91%、91.18%和 90.64%, 其中敏感性显著高于专家组判定, 准确率和特异性与专家组未有差异。同样, 基于 4204 张影像利用 CNN 方法建立的黑色素瘤诊断模型, 灵敏度、特异性都超过了皮肤科医生的检验结果^[18], 表明了人工智能在大数据时代的优越性。

在风险预测方面, 人工智能可辅助医生决策, 降低患者就医成本, 提高其生活质量。Kudo 等^[19]使用仅接受内镜切除术或 T1 CRC 手术切除术的人

选患者的临床病理记录, 构建了一个 T1 大肠癌淋巴结转移风险的人工智能预测模型(artificial neural network, ANN), 该模型(AUC = 0.84)在识别初次内镜手术切除后的转移淋巴结风险上效果明显优于美国指南(AUC=0.77), 为 T1 期大肠癌患者是否进行淋巴结清除手术, 提供了辅助预测手段。

肿瘤的分级用于评价恶性肿瘤侵袭转移的程度, 指导临床治疗和患者预后管理。Bulten 等^[20]使用患者穿刺活检切片和病理报告开发了一项可根据 Gleason 评分标准对前列腺活检样本分级评分的深度系统, 这个系统在诊断和分级方面性能卓越, 区分恶性肿瘤的 AUC 为 0.990, 判断 ≥ 2 级和 ≥ 3 级的 AUC 分别为 0.978 和 0.974, 该系统对 100 份活检样本的分级结果(kappa=0.854)优于专家判定(kappa= 0.819), 可降低人工判定带来的主观性错误并节约时间成本。

在预后分析方面, 人工智能也表现出良好的预测效果, Arya 等^[21]提出了一个两阶段的人类乳腺癌预后预测多模式模型, 即第一阶段使用卷积神经网络提取特征, 第二阶段输入提取的特征进而在基于堆栈的集成模型中得到分类结果, 该集成模型的结果比现有的多模式 CNN 方法更好(AUC=0.93, 准确度为 90.2%)。

2.2.3 人工智能在辅助医学上的应用

人工智能还能辅助医生做一些识别、监控管理疾病的工作, 将人工智能的数据整合、分析与判断能力与人类医生的诊疗经验相结合, 提供辅助医疗的处理逻辑, 分担医生压力、提高工作效率。

对于术后分析, Frit 等^[22]收集了来自医疗中心接受气管插管手术的患者数据构建了一个可以预测术后 30 天死亡率的多路径卷积神经网络深度学习模型, 在纳入的 95,907 名患者数据中, 最终有 941 名患者在手术后 1 个月内死亡。该模型预测术后 30 天死亡率结果 AUC 为 0.867, 95%CI 为 0.835~0.899。

对于监控管理慢性病, 以往用于测量视网膜血管口径的方法离不开人工操作, 图片质量好坏直接影响测定时间长短, 且人工测量差异较大。为了提高效率, Cheung 等^[23]开发了一种深度学习 CNN 模型(SIVA-DLS), 可以全自动地从视网膜照片中去测量视网膜血管的口径, 且结果与人工测量结果具有

高度的一致性, 相关系数在 0.82~0.95 之间。除此之外, 该研究还证明了 SIVA-DLS 测量的视网膜血管口径与心血管疾病有关。

人工智能在医学中的应用如表 1 所示。除此之外, 医学领域的康复护理机器人在很大程度上也应用了人工智能技术, 可以帮助行动不便的病人日常生活和术后康复^[3], 也可以辅助医生进行手术, 提高医疗效率, 节省医疗资源, 应用十分广泛。

2.3 组学大数据与人工智能的联合应用

组学大数据和人工智能各自在医学领域都发挥着巨大的潜能与优势, 两者结合应用, 即高维数据集的可用性加上高性能计算机以及创新的机器学习架构^[24], 不仅可以提高数据利用率, 更能优化单组学或非人工智能研究的结果。

卵巢癌是女性生殖器官常见的恶性肿瘤之一, 初始症状不明显, 很难及早发现, 复发率较高, 治愈率较差, 死亡率超过宫颈癌及子宫内膜癌之和, 高居妇科癌症首位。Hira 等^[25]开发了基于可变自动编码器(variational autoencoder, VAE)改进的最大平

均差异 VAE (maximum mean discrepancy-variational autoencoder, MMD-VAE)技术, 而后开发了 VAE 和 MMD-VAE 的深度学习框架。VAE 是一种训练被正则化以避免过度拟合, 并确保潜在空间具有良好的特性来支持生成过程的自编码器。MMD-VAE 则是将 VAE 的损失函数替换成使用最大平均差异。通过单一组学(基因组、转录组、表观组)、综合二组学(基因组+转录组、表观组+转录组、基因组+表观组)和三组学(基因组+转录组+表观组)的数据分析, 研究卵巢癌的癌症样本识别、分子亚型聚类 and 分类以及生存分析。结果显示, VAE 和 MMD-VAE 优于现有的降维(PCA、t-SNE)技术, 且在亚型聚类和分类以及生存分析中, 综合二组学和三组学的结果比基于单组学更好, 且在多组学数据集分析中 MMD-VAE 性能比 VAE 更出色。

Zhao 等^[26]构建了一个可扩展且可解释的深度学习框架 DeepOmix 用来集成多组学数据和生存预测, 该团队使用突变、拷贝数变化、基因表达和 DNA 甲基化四种组学数据, 将 DeepOmix 应用在八个不同的癌症(膀胱尿路上皮癌、乳腺浸润癌、头颈部鳞

表 1 人工智能在医学中的应用

Table 1 The application of artificial intelligence in medicine

应用	研究内容	方法	结果	参考文献
影像	黄斑变性	DL _{OVLP}	单模态纵向配准实验平均误差为 54~59 μm , 多模态横断面配准实验平均误差为 66~69 μm	[12]
	乳腺 DCE MRI	AI 系统	AI 分析平均 AUC 从 0.71 提高到 0.76	[13]
	分割脑部 MRI 图像	大容量深度卷积神经网络	完整区域, 核心区域和增强区域的准确性分别为 0.90、0.85 和 0.84	[14]
癌症研究	口腔癌诊断	分层深度卷积神经网络	准确度为 94.5%	[15]
	上消化道癌诊断	GRAIDS	内部验证集准确性为 0.955, 外部验证集准确性结果为 0.915~0.977, 且灵敏度堪比专业内镜医师	[16]
	胃粘膜病变	卷积神经网络	敏感性显著高于专家组结果, 准确率和特异性与专家组未有差异	[17]
	黑色素瘤图像	卷积神经网络	CNN 的灵敏度和特异性均高于专家	[18]
	结直肠癌转移至淋巴结风险	人工神经网络	AUC = 0.84	[19]
	前列腺活检评分	深度学习	良恶性 AUC = 0.990 (0.982~0.996); 观察者数据集中, kappa 比专家组高	[20]
	乳腺癌预后预测	基于堆叠集成的卷积神经网络	AUC=0.93, 准确度为 90.2%	[21]
辅助医学	预测术后 30 天死亡率	多路径卷积神经网络	AUC=0.867 (0.835~0.899)	[22]
	测量视网膜血管口径	SIVA-DLS	与人工测量结果相关系数在 0.82~0.95 之间	[23]

状细胞癌、低级别胶质瘤、肾透明细胞癌、肺腺癌、卵巢浆液性囊腺癌和胃腺癌)数据集的预后分析,经与其他五种最新方法(BLockForest、DeepHit、DeepSurv、glmBoost、IPF_LASSO)比较,在其中的六个数据集中,DeepOmix 的预后预测结果表现均为最佳。文章中还提到 DeepOmix 除了可以使用四种类型的组学数据之外,还可以集成更复杂的蛋白质数据。

Xu 等^[27]提出了一种新的分层集成深度灵活神经森林框架(HI-DFNForest),集成多组学数据用于肿瘤亚型分类。文章针对乳腺浸润性癌、多形性胶质母细胞瘤和卵巢癌三种肿瘤,分别使用 DNA 甲基化数据、miRNA 表达数据、基因表达数据或者将三类数据整合来进行亚型分类。结果显示无论在哪种肿瘤数据集中,多组学整合数据分类的准确性(分别为 0.846、0.885、0.840)都高于单组学(DNA 甲基化: 0.731、0.596、0.640; miRNA 表达: 0.769、0.539、0.640; 基因表达: 0.808、0.865、0.760)。

多组学分析产生的数据较之单一组学分析更多更复杂,针对具体的生物学科学问题,多组学也可以凭借多维度多角度的优势更为全面地解释该问题。所以多组学相较单组学更为需要人工智能的辅助。而将人工智能技术应用于多组学大数据上,将会是实现精准医疗和个性化医疗的重要步骤。现有研究表明多组学数据与人工智能结合应用,结果会更优于单组学分析。

3 医学应用的挑战

大数据产业的发展和人工智能的兴起,促进了数据量的增长。国家政策的扶持,为组学大数据和人工智能的发展带来了前所未有的机遇,同时也面临诸多挑战。

大规模各种组学数据的产生,为疾病的发病机制的研究,提供了丰富的数据资源。但是组学数据的分析由于缺乏统一的标准,根据不同流程,产生不同的结果,导致数据整合过程艰难复杂;不同种类组学数据质量参差不齐,由于某些组学技术还处于初步发展阶段,检测方法尚不成熟,导致数据质量差;组学数据种类多,亟需开发相互整合的分析

技术;组学数据由于在机构、医院之间分散保存,缺乏有效的集成技术,阻碍了利用大批量的数据进行人工智能的训练测试。

人工智能应用于医学研究同样面临多方面的挑战。现下研究多为监督学习,但是无监督学习更贴近人工智能,大量的数据集还需要熟练的医生来进行标注,此过程需要耗费大量的人力物力精力^[28]。人工智能计算过程不透明,“黑箱”的可解释性不强^[29],在医院层面存在责任划分归属问题。并且各家医院之间信息共享率低,开发出来的人工智能算法不一定具有普遍适应性。此外,随着越来越多人工智能产品投入临床使用,数据安全形势逐渐严峻,病人的隐私保护问题凸显,要开发相应技术并且具备相应的法律法规和伦理规范去约束管制^[30]。

4 结语与展望

组学大数据和人工智能相辅相成,组学大数据的增长与整合使得临床对人工智能的需求日益增大,而人工智能的发展又可以很好地利用组学大数据推动医疗行业的发展。虽存在诸多问题,但前景一片广阔,加之机遇良多,因此,更需要从国家政府到企业研究机构再到医护人员与病患的一致合作,为未来精准医学和个性化医学的实现打下坚实基础。

参考文献(References):

- [1] 安绍维. 多组学大数据整合分析推动人类未来的健康发展. 张江科技评论, 2019, (6): 12–14. [DOI]
- [2] Yu YZ, Shi DJ, Ma JC, Zhou Z. Advances in application of artificial intelligence in medical image analysis. *Chin J Med Imag Technol*, 2019, 35(12): 1808–1812.
俞益洲, 石德君, 马杰超, 周振. 人工智能在医学影像分析中的应用进展. 中国医学影像技术, 2019, 35(12): 1808–1812. [DOI]
- [3] 李贺. 人工智能在医学领域中的应用研究. 科技风, 2020, (17): 7. [DOI]
- [4] Zhao XT, Yang YD, Qu HZ, Fang XD. Applications of machine learning in clinical decision support in the omic era. *Hereditas(Beijing)*, 2018, 40(9): 693–703.
赵学彤, 杨亚东, 渠鸿竹, 方向东. 组学时代下机器学习方法在临床决策支持中的应用. 遗传, 2018, 40(9): 693–703. [DOI]

- [5] Mens MMJ, Maas SCE, Klap J, Weverling GJ, Klatser P, Brakenhoff JPI, van Meurs JBJ, Uitterlinden AG, Ikram MA, Kavousi M, Ghanbari M. Multi-omics analysis reveals microRNAs associated with cardiometabolic traits. *Front Genet*, 2020, 11: 110. [DOI]
- [6] Yuan Y, Bao JD, Chen ZS, Villanueva AD, Wen WQ, Wang FQ, Zhao DJ, Fu XH, Cai QY, Long JR, Shu XO, Zheng DY, Moreno V, Zheng W, Lin WQ, Guo XY. Multi-omics analysis to identify susceptibility genes for colorectal cancer. *Hum Mol Genet*, 2021, 30(5): 321–330. [DOI]
- [7] Cohen JD, Li L, Wang YX, Thoburn C, Afsari B, Danilova L, Douville C, Javed AA, Wong F, Mattox A, Hruban RH, Wolfgang CL, Goggins MG, Molin MD, Wang TL, Roden R, Klein AP, Ptak J, Dobbyn L, Schaefer J, Silliman N, Popoli M, Vogelstein JT, Browne JD, Schoen RE, Brand RE, Tie J, Gibbs P, Wong HL, Mansfield AS, Jen J, Hanash SM, Falconi M, Allen PJ, Zhou SB, Bettgowda C, Diaz LA, Tomasetti C, Kinzler KW, Vogelstein B, Lennon AM, Papadopoulos N. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science*, 2018, 359(6378): 926–930. [DOI]
- [8] Yang W, Shi J, Zhou Y, Liu TJ, Zhan FL, Zhang K, Liu N. Integrating proteomics and transcriptomics for the identification of potential targets in early colorectal cancer. *Int J Oncol*, 2019, 55(2): 439–450. [DOI]
- [9] Zhang BY, Yang L, Wang X, Fu DG. Identification of a survival-related signature for sarcoma patients through integrated transcriptomic and proteomic profiling analyses. *Gene*, 2021, 764: 145105. [DOI]
- [10] Han D, Li QH, Cai W, Xia YW, Ning J, Huang F. Research and application of artificial intelligence in medical imaging. *Big Data*, 2019, 5(1): 39–67.
韩冬, 李其花, 蔡巍, 夏雨薇, 宁佳, 黄峰. 人工智能在医学影像中的研究与应用. 大数据, 2019, 5(1): 39–67. [DOI]
- [11] Wang Y, Li CF. The new research progress of artificial intelligent methods in medical image processing. *Chin J Med Phys*, 2013, 30(3): 4138–4143.
王弈, 李传富. 人工智能方法在医学图像处理中的研究新进展. 中国医学物理学杂志, 2013, 30(03): 4138–4143. [DOI]
- [12] De Silva T, Chew EY, Hotaling N, Cukras CA. Deep-learning based multi-modal retinal image registration for the longitudinal analysis of patients with age-related macular degeneration. *Biomed Opt Express*, 2020, 12(1): 619–636. [DOI]
- [13] Jiang YL, Edwards AV, Newstead GM. Artificial intelligence applied to breast MRI for improved diagnosis. *Radiology*, 2021, 298(1): 38–46. [DOI]
- [14] Hoseini F, Shahbahrami A, Bayat P. An efficient implementation of deep convolutional neural networks for MRI segmentation. *J Digit Imaging*, 2018, 31(5): 738–747. [DOI]
- [15] Jeyaraj PR, Nadar ERS. Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm. *J Cancer Res Clin Oncol*, 2019, 145(4): 829–837. [DOI]
- [16] Luo HY, Xu GL, Li CF, He LJ, Luo LN, Wang ZX, Jing BZ, Deng YS, Jin Y, Li Y, Li B, Tan WC, He CS, Seeruttun SR, Wu QB, Huang J, Huang DW, Chen B, Lin SB, Chen QM, Yuan CM, Chen HX, Pu HY, Zhou F, He Y, Xu RH. Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study. *Lancet Oncol*, 2019, 20(12): 1645–1654. [DOI]
- [17] Li L, Chen YS, Shen Z, Zhang XQ, Sang JZ, Ding Y, Yang XY, Li J, Chen M, Jin CH, Chen CL, Yu CH. Convolutional neural network for the diagnosis of early gastric cancer based on magnifying narrow band imaging. *Gastric Cancer*, 2020, 23(1): 126–132. [DOI]
- [18] Brinker TJ, Hekler A, Enk AH, Berking C, Haferkamp S, Hauschild A, Weichenthal M, Klode J, Schadendorf D, Holland-Letz T, von Kalle C, Fröhling S, Schilling B, Utikal JS. Deep neural networks are superior to dermatologists in melanoma image classification. *Eur J Cancer*, 2019, 119: 11–17. [DOI]
- [19] Kudo SE, Ichimasa K, Villard B, Mori Y, Misawa M, Saito S, Hotta K, Saito Y, Matsuda T, Yamada K, Mitani T, Ohtsuka K, Chino A, Ide D, Imai K, Kishida Y, Nakamura K, Saiki Y, Tanaka M, Hoteya S, Yamashita S, Kinugasa Y, Fukuda M, Kudo T, Miyachi H, Ishida F, Itoh H, Oda M, Mori K. Artificial intelligence system to determine risk of T1 colorectal cancer metastasis to lymph node. *Gastroenterology*, 2021, 160(4): 1075–1084.e2. [DOI]
- [20] Bulten W, Pinckaers H, van Boven H, Vink R, de Bel T, van Ginneken B, van der Laak J, Hulsbergen-van de Kaa C, Litjens G. Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study. *Lancet Oncol*, 2020, 21(2): 233–241. [DOI]
- [21] Arya N, Saha S. Multi-modal classification for human breast cancer prognosis prediction: Proposal of deep-learning based stacked ensemble model. *IEEE/ACM Trans Comput Biol Bioinform*, 2020, doi: 10.1109/TCBB.2020.3018467. [DOI]

- [22] Fritz BA, Cui ZC, Zhang MH, He YJ, Chen YX, Kronzer A, Abdallah AB, King CR, Avidan MS. Deep-learning model for predicting 30-day postoperative mortality. *Br J Anaesth*, 2019, 123(5): 688–695. [DOI]
- [23] Cheung CY, Xu DJ, Cheng CY, Sabanayagam C, Tham YC, Yu M, Rim TH, Chai CY, Gopinath B, Mitchell P, Poulton R, Moffitt TE, Caspi A, Yam JC, Tham CC, Jonas JB, Wang YX, Song SJ, Burrell LM, Farouque O, Li LJ, Tan G, Ting DSW, Hsu W, Lee ML, Wong TY. A deep-learning system for the assessment of cardiovascular disease risk via the measurement of retinal-vessel calibre. *Nat Biomed Eng*, 2021, 5(6): 498–508. [DOI]
- [24] Bhinder B, Gilvary C, Madhukar NS, Elemento O. Artificial intelligence in cancer research and precision medicine. *Cancer Discov*, 2021, 11(4): 900–915. [DOI]
- [25] Hira MT, Razzaque MA, Angione C, Scrivens J, Sawan S, Sarker M. Integrated multi-omics analysis of ovarian cancer using variational autoencoders. *Sci Rep*, 2021, 11(1): 6265. [DOI]
- [26] Zhao LH, Dong QY, Luo CL, Wu Y, Bu DC, Qi XN, Luo YF, Zhao Y. DeepOmix: A scalable and interpretable multi-omics deep learning framework and application in cancer survival analysis. *Comput Struct Biotechnol J*, 2021, 19: 2719–2725. [DOI]
- [27] Xu J, Wu P, Chen YH, Meng QF, Dawood H, Dawood H. A hierarchical integration deep flexible neural forest framework for cancer subtype classification by integrating multi-omics data. *BMC Bioinformatics*, 2019, 20(1): 527. [DOI]
- [28] Liu FW, Li HJ, Zhang YH, Li RS, Wang ZS, Tang XY. Application of artificial intelligence in medical imaging diagnosis. *Beijing Biomed Eng*, 2019, 38(2): 206–211. 刘丰伟, 李汉军, 张逸鹤, 李若松, 王尊升, 唐晓英. 人工智能在医学影像诊断中的应用. *北京生物医学工程*, 2019, 38(2): 206–211. [DOI]
- [29] 许家睿. 人工智能在辅助医疗领域现状与未来发展趋势概述. *中国新通信*, 2021, 23(1): 232–234. [DOI]
- [30] Zhou WL. The classification and ethical reflection of artificial intelligence medical application. *Chin Med Ethic*, 2020, 33(7): 826–830. 周琬琳. 人工智能医学应用的分类与伦理问题反思. *中国医学伦理学*, 2020, 33(7): 826–830. [DOI]

(责任编辑: 朱波峰)