



· 综述 ·

# 人工智能在类器官研究中的应用进展与挑战

吴洪基<sup>1</sup>, 王海霞<sup>2, 3, 4</sup>, 汪玲<sup>2,3,4</sup>, 罗小刚<sup>1</sup>, 邹冬玲<sup>2, 3, 4</sup>

1. 重庆大学生物工程学院, 重庆 400030;
2. 重庆大学附属肿瘤医院妇科肿瘤中心, 重庆 400030;
3. 重庆市卵巢癌专病医学研究中心, 重庆 400030;
4. 重庆大学附属肿瘤医院, 肿瘤转移与个体化诊治转化研究重庆市重点实验室, 类器官转化研究实验室, 重庆 400030

[摘要] 类器官是一种优异的肿瘤和干细胞研究模型, 对其生长或药筛等过程的各种类型数据进行分析, 有助于提升对类器官本身以及所代表疾病的了解。但人工观察和筛选类器官以及使用传统统计学方法在处理类器官数据时, 存在分析准确度与效率低、难度系数大、人工成本高以及带有一定主观性等问题。而人工智能在很多生物学和医学研究领域已被证明会产生卓越效果。将人工智能引入类器官研究, 有助于提升研究的客观性、准确性和速度, 从而使类器官能更好地实现疾病建模、药物筛选、个性化医疗等。首先, 类器官图像数据的人工智能分析取得了显著进展。结合深度学习的图像分析能够更精准地捕捉类器官的微观结构和变化, 提高对类器官形态和生长的自动识别能力, 达到较高的准确度, 节约研究时间与成本。其次, 对于类器官的组学数据, 人工智能技术的引入同样取得了重要突破: 可提高数据的处理效率以及发现潜在的基因表达模式, 为细胞发育和疾病机制的解析提供新的工具。再次, 类器官其他类型的数据如电信号和光谱等通过人工智能技术可实现对类器官类型和状态客观的分类, 为类器官的全面表征进行了有益的尝试。而在类器官重要应用领域—药物筛选方面, 人工智能可为过程监测和结果预测提供强有力的支持。通过高内涵显微镜图像和深度学习模型, 研究者们能够实时监测类器官对药物的响应, 实现了对药物作用的非侵入性检测, 使药物筛选更加精准和高效。然而, 尽管人工智能在类器官研究中取得了显著成果, 仍然存在一系列挑战。数据获取的难度、样本质量和样本量的不足、模型解释性的问题等制约了其广泛应用。为了克服这些问题, 未来的研究需要致力于提高数据的一致性, 增强模型解释性, 并探索多模态数据融合的方法, 以更全面、可靠地应用人工智能于类器官研究中。因此本文认为人工智能技术的引入为类器官研究带来了前所未有的机遇, 也取得了明显的研究进展。然而, 我们仍然需要跨学科的研究与合作, 共同应对挑战, 推动人工智能在类器官研究中的更深层次应用。未来, 人工智能有望在类器官研究中发挥更大的作用, 加速其向临床转化和精准治疗的应用进程。

[关键词] 类器官; 人工智能; 机器学习; 精准医疗

中图分类号: TP18 文献标志码: A DOI: 10.19401/j.cnki.1007-3639.2024.02.009

**Application progress and challenges of artificial intelligence in organoid research** WU Hongji<sup>1</sup>, WANG Haixia<sup>2, 3, 4</sup>, WANG Ling<sup>2,3,4</sup>, LUO Xiaogang<sup>1</sup>, ZOU Dongling<sup>2, 3, 4</sup> (1. Bioengineering College, Chongqing University, Chongqing 400030, China; 2. Department of Gynecologic Oncology, Chongqing University Cancer Hospital & Chongqing Cancer Institute & Chongqing Cancer Hospital, Chongqing 400030, China; 3. Chongqing Specialized Medical Research Center of Ovarian Cancer, Chongqing 400030, China; 4. Organoid Transformational Research Center, Chongqing Key Laboratory of Translational Research for Cancer Metastasis and Individualized Treatment, Chongqing University Cancer Hospital, Chongqing 400030, China)

Correspondence to: ZOU Dongling, E-mail: 13570049@qq.com

基金项目: 国家自然科学基金(82073129); 重庆市科技局项目(2023ZDXM029、2023MSXM043、cstc2022jxj1120039); 宁夏回族自治区重点研发计划项目(2023BEG02037); 重庆市教委科学技术研究项目(KJQN20230013); 重庆大学附属肿瘤医院科研能力提升专项(2023nlts005)。

作者简介: 吴洪基(ORCID: 0009-0007-4405-7411), 硕士研究生在读。

通信作者: 邹冬玲(ORCID: 0000-0001-9158-8783), 主任医师, 重庆大学附属肿瘤医院妇科肿瘤中心主任、生物样本库类器官专库主任、类器官转化研究实验室负责人, E-mail: 13570049@qq.com。

[ **Abstract** ] Organoids, recognized as invaluable models in tumor and stem cell research, assume a pivotal role in the meticulous analysis of diverse datasets pertaining to their growth dynamics, drug screening processes and related phenomena. However, the manual scrutiny and conventional statistical methodologies employed in handling organoid data often grapple with challenges such as diminished precision and efficiency, heightened complexity, escalated human resource requirements, and a degree of subjectivity. Acknowledging the remarkable efficacy of artificial intelligence (AI) in the realms of biology and medicine, the incorporation of AI into organoid research stands poised to enhance the objectivity, precision and expediency of analyses. This integration empowers organoids to more effectively fulfill objectives such as disease modeling, drug screening and precision medicine. Notably, significant strides have been made in AI-driven analyses of organoid image data. The amalgamation of deep learning into image analysis facilitates a more meticulous delineation of the microstructural intricacies and nuanced changes within organoids, achieving a level of accuracy akin to that of experts. This not only elevates the precision of organoid morphology and growth recognition, but also contributes to substantial time and cost savings in research endeavors. Furthermore, the infusion of AI technology has yielded breakthroughs in the processing of organoid omics data, resulting in heightened efficiency in data processing and the identification of latent gene expression patterns. This furnishes novel tools for comprehending cellular development and unraveling the intricate mechanisms underlying various diseases. In addition to image data, AI techniques applied to diverse organoid datasets, encompassing electrical signals and spectra, have realized an unbiased classification of organoid types and states, embarking on a comprehensive journey towards characterizing organoids holistically. In the pivotal domain of drug screening for organoids, AI emerges as a stalwart companion, providing robust support for real-time process monitoring and result prediction. Leveraging high-content microscopy images and sophisticated deep learning models, researchers can dynamically monitor organoid responses to drugs, effecting non-invasive detection of drug impacts and amplifying the precision and efficiency of drug screening processes. Despite the significant strides made by AI in organoid research, challenges persist, encompassing hurdles in data acquisition, constraints in sample quality and quantity, and quandaries associated with model interpretability. Overcoming these challenges necessitates dedicated future research efforts aimed at enhancing data consistency, fortifying model interpretability, and exploring methodologies for the seamless fusion of multimodal data. Such endeavors are poised to usher in a more comprehensive and dependable application of AI in organoid research. In summation, the integration of AI technology introduces unparalleled opportunities to organoid research, resulting in noteworthy advancements. Nevertheless, interdisciplinary research and collaborative efforts remain imperative to navigate challenges and propel the more profound integration of AI into organoid research. The future holds promise for AI to assume an even more prominent role in advancing organoid research toward clinical translation and precision medicine.

[ **Key words** ] Organoids; Artificial intelligence; Machine learning; Precision medicine

2009年, Sato等<sup>[1]</sup>报道了从成人组织中首次成功衍生出类器官, 因类器官与对应的人类器官拥有高度相似的组织学特征, 并能重现该器官的生理功能, 出现在疾病建模的最前沿, 成为研究人类疾病的一种有价值、相对不昂贵且便捷的方法。但目前类器官形态学判读主要依赖于专家的经验, 且类器官研究所产生的高通量数据增加了研究的复杂性, 对于类器官各种形态信息、组学信息等的有效整合与分析需要客观和高效的方法来辅助实现。

人工智能技术, 尤其是作为机器学习一支的深度学习, 近年来发展迅速。医疗数据大量产生和积累契合该数据驱动型方法, 推动了机器学习在医疗领域的全面应用, 提升了诊疗的客观性、准确性和效率。人工智能结合类器官大数据将帮助类器官研究专家减少基础性、重复性工作, 并

进行更深度的信息挖掘, 进而提高类器官自动化和标准化<sup>[2]</sup>, 促进类器官临床转化和精准治疗的实现。

本文旨在综述人工智能在类器官研究中的最新进展, 探讨人工智能在类器官研究应用中的潜能。

## 1 类器官研究进展

类器官来源于多种类型干细胞, 通过自组织方式形成三维结构, 因其能够模拟原生器官的结构和功能, 因此有助于人类疾病的建模、药物筛选以及体外替代组织或器官的再生。相对于患者来源的动物移植瘤模型, 类器官具有成本低、成功率高、培养周期短、操作简便等优势<sup>[3-4]</sup>。相比于传统的2D细胞系培养, 类器官直接来源于患者, 能够保留患者的基因组信息以及表观遗传学信息, 并且更符合体内条件下的三维结构<sup>[4]</sup>,

可以更准确地概括来源肿瘤的结构、特定功能、分子特征和基因组改变<sup>[5]</sup>。前列腺癌<sup>[6]</sup>、胰腺癌<sup>[7]</sup>、结直肠癌<sup>[3]</sup>、卵巢癌<sup>[8]</sup>等多种实体肿瘤的一类器官体外培养体系已成功建立。类器官在预测治疗反应、研究耐药相关机制、优化治疗策略和精准治疗等方面拥有巨大潜力<sup>[9-10]</sup>, 成为近十年内生物学领域的重要突破之一<sup>[11]</sup>。

类器官具有众多优越的特性, 但同时也给观察和分析带来了巨大的挑战, 作为三维培养物的一类器官拥有复杂的空间结构和表型, 且其在培养过程中强烈地变化和运动, 导致研究者们需要用合适的办法来定量、准确、快速和便捷地观察。其次, 类器官研究一般会产生大量的高通量数据, 数据的巨大规模、高纬度性、噪声广泛存在以及复杂的生物学过程使得人工筛选分析或者传统的统计学方法难以处理。如何对这些数据进行分析和挖掘, 以获得对类器官培养过程变化以及其代表的疾病发生、发展的理解等方面需要客观和准确的方法。

## 2 人工智能相关技术在医学领域的研究进展

近年来, 伴随着人工智能尤其是深度学习的迅速发展, 其方法和模型越来越多地被研究者结合患者的临床、病理学、影像学、基因组学、蛋白质组学、转录组学等数据, 用于诊疗全过程的决策, 以帮助医师快速、客观和准确地完成患者的筛查、诊断、治疗和预后, 制订个性化治疗方案。

机器学习是人工智能的一个技术分支, 能从数据中自动学习规律和模式来决策。经典的模型有随机森林与支持向量机。随机森林是由多个决策树组成的集成模型, 每个决策树基于随机子集得到, 其鲁棒性好, 不易受噪声和异常因素影响并且不需要大量调参操作。有研究成功运用其进行心血管疾病<sup>[12]</sup>和糖尿病<sup>[13]</sup>风险预测。支持向量机是另一种流行的机器学习算法, 其基本思想为寻找一个最佳的超平面, 以最大程度地分隔不同类别的数据点。在小样本高维数据上表现较好, 其对核函数等参数的敏感性很高, 并且大规模数据集计算时间和空间的效能都较高。有研究<sup>[14]</sup>尝试结合语音数据检测帕金森病。

深度学习是机器学习的热门领域, 其由多层神经网络叠加而成。相较于传统机器学习, 其能自动学习数据中的特征并组合转换为更高层、更抽象的特征, 可减少人工设计特征工作<sup>[15]</sup>, 但对数据的数量和质量都有较高要求。卷积神经网络是一种广泛被使用的网络结构, 通过卷积和池化操作来层次化地进行特征提取和学习, 用于处理和分析图像及视频数据。常见模型有Vgg、ResNet和U-NET等, 其应用已实现了分割脑肿瘤<sup>[16]</sup>, 对皮肤癌进行分类<sup>[17]</sup>以及使用视网膜图像检测糖尿病视网膜病变与糖尿病性黄斑水肿<sup>[17]</sup>, 能力水平可与专业医师媲美。另一种网络结构—循环神经网络拥有隐藏层对信息进行存储记忆, 每个时间步的输出和隐藏状态都受前一个时间步的影响, 常用于序列数据处理, 但当输入序列比较长时会出现长期依赖问题以及不能并行运算导致无法充分利用计算资源。有研究<sup>[18]</sup>报道将电子健康记录中事件之间的时间关系进行建模进行心力衰竭早期诊断。而Transformer使用自注意力机制, 会计算序列每个元素与其他元素的相似度, 实现全局依赖关系的捕捉且容易进行改进和扩展。Transformer最初用于自然语言中, 如Bert模型<sup>[19]</sup>, 随后在图像数据中也大放异彩<sup>[20]</sup>。研究者也尝试将其引入医学领域, 如使用影像学 and 临床特征在早期预测非小细胞肺癌患者的总生存期和无复发生存期<sup>[21]</sup>。值得一提的是, Transformer需要巨大的训练数据和计算资源。此外还有对抗生成网络, 其由生成器和判别器组成, 通过竞争训练生成器生成逼真数据, 用于生成高质量的图像、音频或其他数据。在医学领域, 研究者尝试利用其从患者MRI图像生成对应的CT图像以避免暴露于辐射<sup>[22]</sup>以及使用低质量图像进行超分辨率重建, 进而获得高质量超声图像<sup>[23]</sup>。但对抗生成网络也有缺点: 难以收敛, 不稳定, 因此需要更多训练技巧。

应用人工智能技术, 除了选择高效合适的模型外, 数据的可靠性直接影响到模型的鲁棒性。在医学领域, 数据稀缺<sup>[24-25]</sup>和类不平衡<sup>[22, 26]</sup>两大问题限制了人工智能在医学领域的进一步发展与应用。通过几何、色彩和噪声引入等直接对

原有数据进行变换,或是训练中在特征空间进行增强和利用对抗网络生成数据等技术,亦或微调预训练模型权重的迁移学习是解决数据稀缺的常用方法。针对类不平衡,可通过重采样和特征选择的数据预处理以及在模型层面纳入代价敏感策略来尝试解决。

### 3 人工智能在类器官研究中的应用

#### 3.1 人工智能在类器官图像分析中的应用

细胞图像蕴含海量信息,人工智能通过对细胞形态、功能及亚细胞形态、功能的定性和定量分析,获得大量的细胞特征信息,能够帮助研究者更好地把握细胞结构与功能的变化,从而理解细胞的生命过程及运作机制。

##### 3.1.1 提高类器官图像质量

类器官图像分析的前提条件是获取高质量的类器官图像,以便能充分和正确地提取图像特征,尤其在需要使用机器学习这一数据驱动型人工智能方法进行后续分析的时候。由于成像技术和类器官的自身条件,研究者往往无法得到高质量图像。例如类器官为避免光毒性而需要快速采集和降低光照,往往无法得到高质量图像,为此McAleer等<sup>[27]</sup>提出两种深度学习模型(基于U-NET和基于patch的回归模型),通过在双光子激发荧光下采集的类器官(人视网膜和皮肤类器官)低分辨率图像和对应的高分辨率图像组成的训练集上训练后,能够实现图像超分辨率重建,以达到使用减少光暴露来获取高分辨率类器官图像的目的,并在这两种类器官保留测试集上图像重建指标(均方误差和结构相似性)显示两种深度学习拥有良好的性能。

另外,成像技术的限制也会造就图像质量不佳,如定量相位成像技术会造成类器官相位包裹问题。Yang等<sup>[28]</sup>针对此问题提出了一种基于深度图像先验的类U-net深度学习模型,能通过对类器官相位包裹图像的拟合,从而达到无需训练集预训练模型来进行相位展开,进而修正类器官伪影还原真实图像的效果,与其他方法(未加权最小二乘算法、相位展开最大流量算法等)在小鼠肠道类器官相比较,效果良好。

##### 3.1.2 快速识别和分割类器官

大多数生物图像分析的第一步是目标识别和图像分割,帮助研究者定位和提取感兴趣区域以便进行更准确的分析和诊断。但类器官悬浮在三维空间中及其由数千个跨越数百微米的单个细胞组成,造就其庞大的规模。导致类器官成像常常出现遮挡、失焦、大小和形状上的巨大异质性以及高密度或高度稀疏的类器官分布,这给类器官的识别与分割造成巨大困难。有鉴于此,Kassis等<sup>[29]</sup>选择焦点中具有最多人肠道类器官的位置进行明场成像,再对该图像上的类器官进行位置、大小标注,在训练集上训练基于R-CNN的深度学习模型,训练完成后的模型能对入肠道类器官自动定位,标志质量(平均精度=80%)与人工相似,但速度更快。

二维成像虽然能一定程度地表征类器官,但三维成像更能展现它的复杂性和形态特征。Bao等<sup>[30]</sup>使用多尺度信息引导优化的神经网络EGO-Net对光学相干断层扫描技术得到的肝、胃和肠三类癌症类器官的三维图像进行监督分割,对比其他的单神经网络,其在精度、灵敏度、Dice和Jacard等指标上拥有最佳性能,尤其是对直径 $\geq 50 \mu\text{m}$ 的类器官。基于分割的后续分析表明,类器官的生长过程经历了体积增大、空腔形成和融合等形态学变化;此外,类器官的生长速度与初始尺度有关。Garrett等<sup>[31]</sup>通过扩展他们在组织学切片中描绘细胞核的深度学习方法扩展到三维并改进损失函数,在已标注的具有不同突变乳腺癌细胞系培养的类器官共聚焦图像上训练,随后在验证集分割效果与另一神经网络模型3D-Unet比较,F1为0.83,有所提高。

Sun等<sup>[32]</sup>则选择磁粒子成像,对移植到小鼠肾囊的经超顺磁性氧化铁纳米颗粒标记的胰岛类器官进行成像,随后使用机器学习算法k-means进行三维图像分割,完成对感兴趣区域的监测和其含铁量的估计,实现对体内移植胰岛类器官的量化监测。Gritti等<sup>[33]</sup>开发了MOrgAna软件使用机器学习解决不同显微镜设备、放大倍率和视野获得的图像问题;随后使用不同拍摄条件下获得的人脑类器官、小鼠胚胎类器官和肠类

器官与CellProfiler及OrganoSeg两种工具进行比较,在Jaccard距离、精度和准确性等标准指标上有所提高。但其要求每张图像只有一个对象,可能限制其应用。

### 3.1.3 帮助类器官进行形态学分类

类器官形态分类可以帮助我们更好地理解类器官结构、状态和功能之间的关系。例如,Abdul等<sup>[34]</sup>开发了一种基于深度学习的工具(D-CryptO)并对结直肠类器官进行分类,经过标注图片训练后,能以98.00%的准确率区分透明和不透明的类器官以及以90.87%的准确率区分球形和出芽类器官。基于此的类器官传代、短期暴露于外部刺激和药物治疗的形态学分析能捕捉结肠类器官对不同作用下反应方式的变化。Okamoto等<sup>[35]</sup>将结直肠癌类器官按形态分为6类后作为分类训练集,训练分类器支持向量机和DensNet201,结果显示,分类准确度与人工相当。整个模块训练后实现全自动、高准确率地对图像中结直肠癌类器官进行识别并分类,为其后续不同形态类器官基因差异表达分析提供支撑。

Kegeles等<sup>[36]</sup>使用单个视网膜类器官第5天明场图像和第9天荧光报告结果为分化标签制作训练集,对比训练4个神经网络(ResNet50v2、VGG19、Xception和DenseNet121)后选择ResNet50v2实现根据第5天明场图像实现非侵入式预测单个视网膜类器官分化。在测试集上结果显示,卷积神经网络在分化早期阶段预测视网膜类器官分化方面,准确度(0.84)高于人工(0.67±0.06),证实神经网络在报告基因表达开始之前可以成功地识别和预测类器官的视网膜分化。

Mergenthaler等<sup>[37]</sup>为解决特征选择对类器官带来的挑战,开发了数据驱动的基于体素的特征结合机器学习的方法对三维图像数据进行自动定量表型分析,并研究了不同癌基因表达(SNAIL、Bcl-XL和Bcl-2等)下对人乳腺腺泡类器官的形态影响。结果显示,该方法相比于使用二维形态和预定义的三维特征能实现更小的计算量和更好的聚类与分类效果,并显示与先前类器官恶性转化中癌基因作用的研究结果吻合。

### 3.1.4 实现对类器官的自动追踪

类器官的显著优点是可以进行实时观察,以研究空间和时间上的复杂细胞生命活动过程<sup>[38]</sup>。为了能对类器官实现跨时间追踪,Hradecká等<sup>[39]</sup>提出了一个工作流程,使用人工标注的真实类器官图像和由条件生成对抗网络生成的图像数据训练卷积神经网络U-Net,实现类器官分割。结合后续的形状相似约束和实例分割纠正的跟踪步骤来关联相应的类器官。在不同表型的小鼠乳腺上皮类器官序列上对其进行验证,在检测精度、分割精度和边界定位误差等指标上达到了与人工相当的性能。但该研究使用的数据是局部图像,每次需要重新对焦所需区域,可能会限制其应用。而Bian等<sup>[40]</sup>则使用拼接和景深合成处理高通量采集的小鼠肝脏类器官和肺泡类器官图片,得到全景明场图。经过标注后作为训练集,训练目标检测神经网络SSD模型,进而识别图像中单个类器官,再以单个类器官不同时刻图像和位置标注为一个训练样本的形式作为训练集,训练基于ResNet的跟踪模型,实现对单个类器官的跟踪。Kok等<sup>[41]</sup>则希望跟踪类器官中的细胞,以便在单细胞水平上研究细胞的生长和稳态。使用U-Net神经网络架构,通过注释细胞核来训练网络预测每个细胞核的中心位置,结合之后的链接算法和手动纠错完成细胞追踪,实现半自动类器官细胞追踪,与手动追踪相比速度提升。

### 3.1.5 人工智能在其他类器官图像处理中的应用

除了分割、检测和分类等计算机视觉常规任务,研究者尝试将生物学信息融入类器官图像研究中。Ballweg等<sup>[42]</sup>使用机器学习随机森林算法处理由胃类器官的延时共聚焦显微镜收集的时间过程数据分解得到的静态特征,分析了分子水平事件(肌动蛋白聚合和解聚)与细胞水平的行为(损伤修复、细胞剥落)的联系,揭示了肌动蛋白聚合时间在决定胃修复过程中涉及的细胞水平行为中的关键作用,为我们了解单个胃类器官如何调节修复提供了数据驱动的见解。Libby等<sup>[43]</sup>将机器学习、数学模式和基因工程优化相结合,创建了一种数据驱动的方法,通过敲除显示影响

干细胞集落组织的基因*CDH1*和*ROCK1*来控制多能干细胞自组织。使用扩展的Cellular Potts模型在计算机上对体外系统进行计算复制,使机器学习驱动的参数优化能够产生所需的模式,从而实现通过机器学习对多能干细胞行为的模型驱动探索,准确预测形态发生动力学,进而实现多细胞模式的空间控制及更好地设计类器官。

### 3.2 人工智能在类器官大数据处理中的应用

#### 3.2.1 便捷评估类器官保真度

癌症模型的泛化性和功效源于它对所研究肿瘤的保真度,为了能客观公正而快速地评价肿瘤类器官对患者疾病忠实反映程度,Chen等<sup>[44]</sup>利用机器学习方法结合其培养结直肠癌类器官的单细胞转录组学数据对类器官亚型进行无偏聚类。结果显示,类器官中的异质性仍然保持良好,展示出高度的多样性。Peng等<sup>[45]</sup>采用机器学习算法,开发出计算工具,可以排除平台和物种的干扰,进而利用转录组学数据定量评估癌症模型与22种自然发生的肿瘤类型和36种亚型之间的相似性,并开发出计算工具。随后用该工具来测量肿瘤类器官、细胞系、源自患者的异种移植物和基因工程小鼠模型与自然产生人类肿瘤的相似性,结果显示,患者类器官和基因工程小鼠相比其他两种模型拥有更高的转录保真度。

#### 3.2.2 深度挖掘组学数据

使用人工智能方法对类器官组学数据进行分析,可帮助研究者识别不同细胞类型和细胞状态转换,揭示细胞发育和疾病发展的动态过程。Feng等<sup>[46]</sup>使用基于机器学习随机森林的标签转移方法结合原代人类胎儿心脏细胞的细胞注释(细胞类型、腔室和偏侧性),实现客观的基于RNA-seq对类器官系统分化的细胞进行表征。通过该方法对比和分析野生型细胞系衍生的类器官与携带三尖瓣下移畸形相关基因的类器官,确定腔室发育缺陷。He等<sup>[47]</sup>开发了一种机器学习框架,实现在细胞分辨率上比较大脑和类器官的基因表达分析,以识别保守和特异的发育轨迹,以及发育表达的基因和功能。Devall等<sup>[48]</sup>利用机器学习算法和RNA-seq来解释致癌物对正常人结肠上皮细胞的早期转录组学和细胞学效应,确定

了将结肠类器官暴露于单一剂量的致癌物混合物(包括MeIQx、PhIP、BaP和NDEA)24 h后基因表达的差异。该研究结果观察到在暴露于致癌物质的结肠类器官中存在强大的转录组反应,揭示了细胞组成的选择性变化。Kim等<sup>[49]</sup>使用基于图像的机器学习方法分析14 d胚状体类器官的单细胞转录组学数据来推断细胞状态和谱系轨迹,为后续转录和表观遗传学调控因子的验证奠定了基础。

#### 3.3 人工智能在类器官其他类型数据中的应用

类器官电信号和光谱等类型的数据逐渐被研究者采用,以实现更全面地表征并揭示类器官的特征和生长过程。Hasib等<sup>[50]</sup>使用局部场电位信号结合机器学习支持向量机和深度学习卷积神经网络两种方案,实现高性能区分6个月(AUC分别为0.993和0.996)和12个月(AUC分别为0.881和0.997)下*CHD2*突变类器官与对照类器官。该研究结果揭示仅少数局部场电位特征包含突变特异性信息,并显示了其随时间的变化,证明了使用局部场电位信号表征类器官的可行性。Tubbesing等<sup>[51]</sup>使用拉曼共聚焦法获取不同因子处理后被固定的小鼠唾液腺类器官细胞密集区域的拉曼光谱,将光谱与其对应处理因子标签作为训练集,训练机器学习RSVM模型,随后在活的小鼠唾液腺类器官进行相同操作,结果显示,EFG处理、FEF2处理、对照组相互区分准确率为76%~81%,因此实现对不同分化状态类器官进行降低基质胶信号干扰、无标记和无损的分类。Becker等<sup>[52]</sup>使用神经网络FeaSel-Net和机器学习方法主成分分析与k均值聚类结合不同药物处理下的膀胱癌类器官荧光和拉曼光谱数据来寻找生物标志物,随后使用挑选的拉曼光谱生物标志物进行分类,虽然准确率仅在73%~87%,但考虑到其仅使用了不到原光谱数据的1%,降低了数据的复杂性,改进了用于药物特异性反应的分类。

#### 3.4 人工智能在类器官药筛方面的应用

药物筛选是类器官应用的一个重要方面,人工智能的介入能更好地实时监测和快速采集与分析数据以反馈调整和优化筛选过程。此外,药物

筛选预测类器官对于药物和治疗方案的响应, 人工智能可提升筛选的效率、成功率, 减少实验的成本和时间以加速个体化治疗的进度。

#### 3.4.1 监测药物作用于类器官的过程

基于ATP的活力检测和活体染料是确定药物效果的传统方法, 但有着明显的缺点: 侵入式测定会在一定程度上影响类器官药筛过程, 细胞水平的测定需要跨细胞整理并掩盖类器官异质性。而使用人工智能结合类器官图像能实现非侵入、定量和实时快速地监测类器官对药物的响应。Spiller等<sup>[53]</sup>通过高内涵显微镜获取178个结直肠癌类器官不同时间点的明场图像, 统计单个类器官形态测量和纹理信息, 训练机器学习线性模型。结果显示, 在保留的测试集上, 与专家和DRAQ7染色的一致性分别为78%和61%。该研究随后还使用该模型非破坏性跟踪3例不同患者的类器官对伊立替康和奥沙利铂的7天动态响应, 展现出捕捉患者特异性药物反应的能力。Larsen等<sup>[54]</sup>建立基于条件生成对抗网络(RCA网络)的治疗分析模型, 该模型利用混合损失层来改进生成的荧光读数, 并包括一个额外的分支来预测明场图像对应的数据。该模型接受9种不同类型的癌症29 165配对明场和3通道荧光图像的训练, 最终能够使用明场图像预测对应荧光图像和活力读数, 即可以通过简单的光学显微镜预测荧光活体染料染色表型, 实现药物反应监测, 从而降低了高通量筛选的成本, 也节省了时间。与基于代谢的测定相比, 通过光学显微镜预测药物反应的RCA网络能以很高的重现性和较少的生物量输入来确定药物敏感性, 特别是使用纵向延时成像, 这为克服个性化体外药物测试中的时间障碍提供了可能性。

#### 3.4.2 预测类器官对药物的反应

药物筛选的周期长、费用高、失败率高, 类器官结合人工智能技术对药物作用效果实现了准确预测和分析, 可提高研发效率, 提高药物安全性, 优化个性化临床治疗方案。Kong等<sup>[55]</sup>使用机器学习方法(岭回归、线性回归和支持向量回归), 结合结直肠癌和膀胱癌类器官药物基因组筛查数据构建类器官药物反应预测模型,

来挑选潜在药物的生物标志物。随后使用挑选的生物标志物准确预测了114例5-FU治疗的结直肠癌患者和77例顺铂(cisplatin, DDP)治疗的膀胱癌患者的效果, 接受治疗的预测应答者比预测无应答者有更好的生存结果( $P_{5-FU}=0.014$ ;  $P_{DDP}=0.01$ )。这种将网络分析结合到机器学习框架中的方法实现了使用类器官模型准确识别稳健的药物反应生物标志物以改善治疗结果, 拉近了类器官与临床的距离。Park等<sup>[56]</sup>在确定患者来源的直肠癌类器官的照射反应与实际放疗结果之间的相关性后, 使用类器官生存分数建立基于机器学习方法随机森林的放射治疗结果预测模型, 用于预测良好应答和低应答, 模型准确度值分别为81.5%和92.1%。Esmail等<sup>[57]</sup>因脑类器官培养昂贵、耗时, 着眼于脑类器官的计算机模拟。Esmail等<sup>[57]</sup>使用递归神经网络、支持向量机和进化系统等方法, 联合开发了包含4 516个基因/蛋白质或表型概念和41 493非零关系的系统, 结果显示, 与公布的湿实验数据结果相比, 无偏二项检验显示其成功预测了全脑类器官常见的9种神经细胞类型、7种腹侧-背侧区域、6种大脑皮层、酸碱状态、细胞应激等的表达, 这种低成本和易于定制类器官模拟可用于许多神经系统疾病模拟并带来重要的治疗见解。有研究<sup>[58]</sup>对异染性脑白质营养不良的全脑类器官进行模拟, 并模拟861种单药和双药组合对该类器官的作用, 提出几种双药组合作为潜在的治疗方案, 将为罕见病的发病机制、疾病进展和治疗的评估提供新的手段。

## 4 未来与挑战

尽管人工智能在类器官研究中逐渐应用并展现出良好的性能, 但也面临如下困难。首先, 某些类器官数据难以获取, 数据集样本量和完备性不佳使得机器学习这一数据驱动型方法建立的模型可能在单一来源的样本上虽表现突出但泛化性能差。同时, 不同研究中数据质量、来源和格式等的差异使得很难实现合并以充分利用来改善这种情况。其次, 很多研究报道仅采用准确性或受试者工作特征(receiver operating characteristic, ROC)曲线等性能指标, 但这些指标可能并不能

充分客观地评价模型。只有对输入到输出的因果关系进行说明,打破人工智能模型的黑盒性质,提高可解释性才能更好地提高信任度。再次,多数研究使用单一类型数据或针对单一任务进行建模,但单一模态数据可能存在信息缺失以及许多研究任务之间拥有相关性,存在特征共享。进行多任务和多模态建模可以提高数据的利用效率与模型泛化性,并提供多角度、更全面分析。此外,现有的应用大部分集中于高通量成像和分析技术的集成,更多的研究范式需要被挖掘,例如使用微流体和其他芯片实验室技术将类器官与人工智能连接起来。这些系统可以精确控制类器官周围的环境,使研究人员能够研究复杂的生物学过程,并对单个细胞或细胞群的生物学行为进行预测<sup>[59]</sup>。

人工智能技术促进了类器官研究的准确度、客观性和复杂数据的处理能力。尽管仍存在不少障碍,但我们相信随着越来越多生物学、医学和人工智能领域的科研人员参与到类器官的研究中,人工智能会在类器官研究中发挥重要作用,加速类器官向临床转化和精准治疗。

**利益冲突声明:** 所有作者均声明不存在利益冲突。

#### 作者贡献声明:

吴洪基: 调查研究、撰写初稿、审读和修订。

王海霞: 撰写初稿、审读和修订。

汪玲: 写作、审读和修订。

罗小刚: 写作、审读和修订。

邹冬玲: 总体构思、写作指导、资金支持。

#### [参 考 文 献]

- [1] SATO T, VRIES R G, SNIPPERT H J, et al. Single Lgr5 stem cells build crypt-villus structures *in vitro* without a mesenchymal niche [J]. *Nature*, 2009, 459(7244): 262–265.
- [2] 中国抗癌协会肿瘤多学科诊疗专业委员会, 中国抗癌协会肿瘤内分泌专业委员会. 肿瘤类器官诊治平台的质量控制标准中国专家共识(2022年版) [J]. *中国癌症杂志*, 2022, 32(7): 657–668.  
The Society of Cancer Multidisciplinary Diagnosis and Treatment, China Anti-Cancer Association, the Society of Cancer Endocrinology, China Anti-Cancer Association. *Chinese experts consensus on quality control standards for tumor organoids diagnosis and treatment platform (2022 version)* [J]. *China Oncol*, 2022, 32(7): 657–668.
- [3] VAN DE WETERING M, FRANCIES H E, FRANCIS J M, et al. Prospective derivation of a living organoid biobank of colorectal cancer patients [J]. *Cell*, 2015, 161(4): 933–945.
- [4] LI X D, FRANCIES H E, SECRIER M, et al. Organoid cultures recapitulate esophageal adenocarcinoma heterogeneity providing a model for clonality studies and precision therapeutics [J]. *Nat Commun*, 2018, 9(1): 2983.
- [5] 王若彤, 王欣, 沈波. 类器官在肿瘤转化医学中的应用和进展 [J]. *中国癌症杂志*, 2022, 32(11): 1105–1114.  
WANG R T, WANG X, SHEN B. Application and progress of organoids in tumor translational medicine [J]. *China Oncol*, 2022, 32(11): 1105–1114.
- [6] GAO D, VELA I, SBONER A, et al. Organoid cultures derived from patients with advanced prostate cancer [J]. *Cell*, 2014, 159(1): 176–187.
- [7] BOJ S F, HWANG C I, BAKER L A, et al. Organoid models of human and mouse ductal pancreatic cancer [J]. *Cell*, 2015, 160(1/2): 324–338.
- [8] HILL S J, DECKER B, ROBERTS E A, et al. Prediction of DNA repair inhibitor response in short-term patient-derived ovarian cancer organoids [J]. *Cancer Discov*, 2018, 8(11): 1404–1421.
- [9] SMITH R C, TABAR V. Constructing and deconstructing cancers using human pluripotent stem cells and organoids [J]. *Cell Stem Cell*, 2019, 24(1): 12–24.
- [10] 芦瑜, 席雨梦, 何晓明, 等. 共培养策略在类器官研究中的应用进展 [J]. *中国癌症杂志*, 2023, 33(3): 293–302.  
LU Y, XI Y M, HE X M, et al. Advances in the application of co-culture strategies in organoids [J]. *China Oncol*, 2023, 33(3): 293–302.
- [11] 马可鑫, 陈晓芳, 李玥, 等. 肿瘤类器官的应用及展望 [J]. *中国肿瘤*, 2022, 31(4): 284–291.  
MA K X, CHEN X F, LI Y, et al. Application and prospect of tumor organoids [J]. *China Cancer*, 2022, 31(4): 284–291.
- [12] TOSHNIWAL D, GOEL B, SHARMA H. Multistage classification for cardiovascular disease risk prediction [C] // KUMAR N, BHATNAGAR V. *International Conference on Big Data Analytics*. Cham: Springer, 2015: 258–266.
- [13] MANI S, CHEN Y K, ELASY T, et al. Type 2 diabetes risk forecasting from EMR data using machine learning [J]. *AMIA Annu Symp Proc*, 2012, 2012: 606–615.
- [14] BEHROOZI M, SAMI A. A multiple-classifier framework for Parkinson's disease detection based on various vocal tests [J]. *Int J Telemed Appl*, 2016, 2016: 6837498.
- [15] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. *Nature*, 2015, 521(7553): 436–444.
- [16] ZHAO X M, WU Y H, SONG G D, et al. A deep learning model integrating FCNNs and CRFs for brain tumor segmentation [J]. *Med Image Anal*, 2018, 43: 98–111.

- [ 17 ] ESTEVA A, KUPREL B, NOVOA R A, et al. Dermatologist-level classification of skin cancer with deep neural networks [ J ] . *Nature*, 2017, 542: 115–118.
- [ 18 ] CHOI E, SCHUETZ A, STEWART W F, et al. Using recurrent neural network models for early detection of heart failure onset [ J ] . *J Am Med Inform Assoc*, 2017, 24(2): 361–370.
- [ 19 ] DEVLIN J, CHANG M W, LEE K, et al. Pre-training of deep bidirectional transformers for language understanding [ J ] . 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NaacL Hlt 2019), Vol 1, 2019: 4171–86.
- [ 20 ] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted Windows [ C ] // 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada. IEEE, 2021: 9992–10002.
- [ 21 ] LIAN J, DENG J J, HUI E S, et al. Early stage NSCLS patients' prognostic prediction with multi-information using transformer and graph neural network model [ J ] . *Elife*, 2022, 11: e80547.
- [ 22 ] NIE D, TRULLO R, LIAN J, et al. Medical image synthesis with context-aware generative adversarial networks [ J ] . *Med Image Comput Comput Assist Interv*, 2017, 10435: 417–425.
- [ 23 ] ZHU Y P, ZHOU Z C, LIAO G J, et al. CsrGAN: medical image super-resolution using a generative adversarial network [ C ] // 2020 IEEE 17th International Symposium on Biomedical Imaging Workshops (ISBI Workshops). Iowa City, IA, USA. IEEE, 2020: 1–4.
- [ 24 ] BANSAL M A, SHARMA D R, KATHURIA D M. A systematic review on data scarcity problem in deep learning: solution and applications [ J ] . *ACM Comput Surv*, 54(10s): 208.
- [ 25 ] GE Y, GUO Y T, DAS S, et al. Few-shot learning for medical text: a review of advances, trends, and opportunities [ J ] . *J Biomed Inform*, 2023, 144: 104458.
- [ 26 ] GUO H X, LI Y J, SHANG J, et al. Learning from class-imbalanced data: review of methods and applications [ J ] . *Expert Syst Appl*, 2016, 73(2/3).
- [ 27 ] MCALEER S, FAST A, XUE Y T, et al. Deep learning-assisted multiphoton microscopy to reduce light exposure and expedite imaging in tissues with high and low light sensitivity [ J ] . *Transl Vis Sci Technol*, 2021, 10(12): 30.
- [ 28 ] YANG F S, PHAM T A, BRANDENBERG N, et al. Robust phase unwrapping via deep image prior for quantitative phase imaging [ J ] . *IEEE Trans Image Process*, 2021, 30: 7025–7037.
- [ 29 ] KASSIS T, HERNANDEZ-GORDILLO V, LANGER R, et al. OrgaQuant: human intestinal organoid localization and quantification using deep convolutional neural networks [ J ] . *Sci Rep*, 2019, 9(1): 12479.
- [ 30 ] BAO D, WANG L, ZHOU X F, et al. Automated detection and growth tracking of 3D bio-printed organoid clusters using optical coherence tomography with deep convolutional neural networks [ J ] . *Front Bioeng Biotechnol*, 2023, 11: 1133090.
- [ 31 ] WINKELMAIER G, PARVIN B. An enhanced loss function simplifies the deep learning model for characterizing the 3D organoid models [ J ] . *Bioinformatics*, 2021, 37(18): 3084–3085.
- [ 32 ] SUN A X, HAYAT H, LIU S H, et al. 3D *in vivo* magnetic particle imaging of human stem cell-derived islet organoid transplantation using a machine learning algorithm [ J ] . *Front Cell Dev Biol*, 2021, 9: 704483.
- [ 33 ] GRITTI N, LIM J L, ANLAŞ K, et al. MOrgAna: accessible quantitative analysis of organoids with machine learning [ J ] . *Development*, 2021, 148(18): dev199611.
- [ 34 ] ABDUL L, XU J, SOTRA A, et al. D-CryptO: deep learning-based analysis of colon organoid morphology from brightfield images [ J ] . *Lab Chip*, 2022, 22(21): 4118–4128.
- [ 35 ] OKAMOTO T, NATSUME Y, DOI M, et al. Integration of human inspection and artificial intelligence-based morphological typing of patient-derived organoids reveals interpatient heterogeneity of colorectal cancer [ J ] . *Cancer Sci*, 2022, 113(8): 2693–2703.
- [ 36 ] KEGELES E, NAUMOV A, KARPULEVICH E A, et al. Convolutional neural networks can predict retinal differentiation in retinal organoids [ J ] . *Front Cell Neurosci*, 2020, 14: 171.
- [ 37 ] MERGENTHALER P, HARIHARAN S, PEMBERTON J M, et al. Rapid 3D phenotypic analysis of neurons and organoids using data-driven cell segmentation-free machine learning [ J ] . *PLoS Comput Biol*, 2021, 17(2): e1008630.
- [ 38 ] RIOS A C, CLEVERS H. Imaging organoids: a bright future ahead [ J ] . *Nat Methods*, 2018, 15(1): 24–26.
- [ 39 ] HRADECKÁ L, WIESNER D, SUMBAL J, et al. Segmentation and tracking of mammary epithelial organoids in brightfield microscopy [ J ] . *IEEE Trans Med Imaging*, 2023, 42(1): 281–290.
- [ 40 ] BIAN X S, LI G, WANG C, et al. A deep learning model for detection and tracking in high-throughput images of organoid [ J ] . *Comput Biol Med*, 2021, 134: 104490.
- [ 41 ] KOK R N U, HEBERT L, HUELSZ-PRINCE G, et al. OrganoidTracker: efficient cell tracking using machine learning and manual error correction [ J ] . *PLoS One*, 2020, 15(10): e0240802.
- [ 42 ] BALLWEG R, ENGEVIK K A, MONTROSE M H, et al. Extracting insights from temporal data by integrating dynamic modeling and machine learning [ J ] . *Front Physiol*, 2020, 11: 1012.
- [ 43 ] LIBBY A R G, BRIERS D, HAGHIGHI I, et al. Automated design of pluripotent stem cell self-organization [ J ] . *Cell Syst*, 2019, 9(5): 483–495.e10.
- [ 44 ] CHEN K Y, SRINIVASAN T, LIN C, et al. Single-cell transcriptomics reveals heterogeneity and drug response of human colorectal cancer organoids [ C ] // 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Honolulu, HI, USA. IEEE, 2018: 2378–2381.
- [ 45 ] PENG D, GLEYZER R, TAI W H, et al. Evaluating the

- transcriptional fidelity of cancer models [ J ] . *Genome Med*, 2021, 13(1): 73.
- [ 46 ] FENG W, SCHRIEVER H, JIANG S, et al. Computational profiling of hiPSC-derived heart organoids reveals chamber defects associated with NKX2-5 deficiency [ J ] . *Commun Biol*, 2022, 5(1): 399.
- [ 47 ] HE C F, KALAFUT N C, SANDOVAL S O, et al. BOMA, a machine-learning framework for comparative gene expression analysis across brains and organoids [ J ] . *Cell Rep Methods*, 2023, 3(2): 100409.
- [ 48 ] DEVAL M, DAMPIER C H, EATON S, et al. Novel insights into the molecular mechanisms underlying risk of colorectal cancer from smoking and red/processed meat carcinogens by modeling exposure in normal colon organoids [ J ] . *Oncotarget*, 2021, 12(19): 1863-1877.
- [ 49 ] KIM I S, WU J Y, RAHME G J, et al. Parallel single-cell RNA-seq and genetic recording reveals lineage decisions in developing embryoid bodies [ J ] . *Cell Rep*, 2020, 33(1): 108222.
- [ 50 ] HASIB M, LYBRAND Z, ESTEVEZ V N, et al. Characterizing hESCs organoids from electrical signals with machine learning [ C ] // 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI). Chicago, IL, USA. IEEE, 2019: 1-4.
- [ 51 ] TUBBESING K, MOSKWA N, KHOO T C, et al. Raman microspectroscopy fingerprinting of organoid differentiation state [ J ] . *Cell Mol Biol Lett*, 2022, 27(1): 53.
- [ 52 ] BECKER L, FISCHER F, FLECK J L, et al. Data-driven identification of biomarkers for *in situ* monitoring of drug treatment in bladder cancer organoids [ J ] . *Int J Mol Sci*, 2022, 23(13): 6956.
- [ 53 ] SPILLER E R, UNG N, KIM S, et al. Imaging-based machine learning analysis of patient-derived tumor organoid drug response [ J ] . *Front Oncol*, 2021, 11: 771173.
- [ 54 ] LARSEN B M, KANNAN M, LANGER L F, et al. A pan-cancer organoid platform for precision medicine [ J ] . *Cell Rep*, 2021, 36(4): 109429.
- [ 55 ] KONG J, LEE H, KIM D, et al. Network-based machine learning in colorectal and bladder organoid models predicts anti-cancer drug efficacy in patients [ J ] . *Nat Commun*, 2020, 11(1): 5485.
- [ 56 ] PARK M, KWON J, KONG J, et al. A patient-derived organoid-based radiosensitivity model for the prediction of radiation responses in patients with rectal cancer [ J ] . *Cancers*, 2021, 13(15): 3760.
- [ 57 ] ESMAIL S, DANTER W R. NEUBOrg: Artificially induced pluripotent stem cell-derived brain organoid to model and study genetics of Alzheimer's disease progression [ J ] . *Front Aging Neurosci*, 2021, 13: 643889.
- [ 58 ] ESMAIL S, DANTER W R. Artificially induced pluripotent stem cell-derived whole-brain organoid for modelling the pathophysiology of metachromatic leukodystrophy and drug repurposing [ J ] . *Biomedicines*, 2021, 9(4): 440.
- [ 59 ] BAI L, WU Y, LI G F, et al. AI-enabled organoids: construction, analysis, and application [ J ] . *Bioact Mater*, 2024, 31: 525-548.

(收稿日期: 2023-07-24 修回日期: 2023-10-3)