人工智能在甲状腺细胞病理学诊断中的研究进展

林　宇[1,3]，罗渝昆[2]，李　杰[3]，王宏伟[4]
[1]解放军医学院，北京　100853；[2]解放军总医院第一医学中心超声诊断科，北京　100853；[3]解放军总医院第一医学中心病理科，北京　100853；[4]解放军总医院第四医学中心病理科，北京　100048

摘要：甲状腺癌的发病率呈逐年上升趋势，超声引导下细针穿刺是相对简便、微创、准确的诊断方法。近年来，人工智能在众多医学领域广泛应用，显著提高了疾病诊断的效率和准确性。本文结合当前人工智能在甲状腺细胞病理学诊断中的最新研究文献，对人工智能在甲状腺细胞病理学诊断中的应用现状进行综述，分析人工智能应用于甲状腺细胞病理学所面临的困难和挑战，并对其未来的应用前景进行展望。

关键词：人工智能；甲状腺；病理；细胞学；数字图像
中图分类号：R581　　**文献标志码：**A　　**文章编号：**2095-5227(2023)11-1297-05
DOI：10.12435/j.issn.2095-5227.2023.053
引用本文：林宇，罗渝昆，李杰，等. 人工智能在甲状腺细胞病理学诊断中的研究进展［J］. 解放军医学院学报，2023，44（11）：1297-1300, 封三.

Research advances in artificial intelligence in cytopathologic diagnosis of thyroid

LIN Yu[1,3], LUO Yukun[2], LI Jie[3], WANG Hongwei[4]
[1]Chinese PLA Medical School, Beijing 100853, China; [2]Department of Ultrasound, the First Medical Center, Chinese PLA General Hospital, Beijing 100853, China; [3]Department of Pathology, the First Medical Center, Chinese PLA General Hospital, Beijing 100853, China; [4]Department of Pathology, the Fourth Medical Center, Chinese PLA General Hospital, Beijing 100048, China
Corresponding author: WANG Hongwei. Email: whw6626@sina.com

Abstract: The incidence of thyroid cancer is increasing year by year, ultrasound-guided fine needle puncture is a relatively simple, minimally invasive, and highly accurate diagnostic method. In recent years, artificial intelligence (AI) has been widely used in many medical fields, significantly improving the efficiency and accuracy of disease diagnosis. This paper focuses on the latest literature on the application of AI in thyroid cytopathological diagnosis, reviews the research status, analyzes the difficulties and challenges faced by AI in thyroid cytopathology, and points out the future application prospects of AI in thyroid cytopathology.

Keywords: artificial intelligence; thyroid; pathology; cytology; digital images
Cited as: Lin Y, Luo YK, Li J, et al. Research advances in artificial intelligence in cytopathologic diagnosis of thyroid［J］. Acad J Chin PLA Med Sch, 2023, 44（11）: 1297-1300, inside back cover.

甲状腺癌是内分泌系统中常见的恶性肿瘤，占全球所有癌症的 3%[1]。超声引导下甲状腺细针穿刺 (fine needle aspiration，FNA) 是美国甲状腺协会 (American Thyroid Association，ATA) 指南推荐的诊断甲状腺结节最准确、有效的方法[2]。近年来，人工智能 (artificial intelligence，AI) 在解决各种复杂图像的识别和分析方面取得了令人瞩目的进展[3-4]，这项先进的技术已经开始应用于病理图像的分析与诊断。在精准医疗时代，使用人工智能检测和量化医疗数据可以帮助病理医师提高诊断的准确性和精确度[5]。本文就人工智能在甲状腺细胞病理学领域的应用进展予以综述，为临床应用研究提供参考。

1　可用于甲状腺细胞病理学诊断的人工智能系统

1.1　基于非图形输入的 AI　甲状腺细胞病理学 Bethesda 报告系统 (the Bethesda system for reporting thyroid cytopathology，TBSRTC) 是标准的甲状腺细针穿刺活检诊断报告系统。TBSRTC 分为 6 个诊断类别：标本无法诊断/不满意 (Ⅰ类)、良性 (Ⅱ类)、意义不明确的细胞非典型病变或滤泡性病变 (Ⅲ类)、滤泡性肿瘤/可疑滤泡性肿瘤 (Ⅳ类)、可疑恶性肿瘤 (Ⅴ类) 和恶性肿瘤 (Ⅵ类)，每个类别都有其相对应的恶性风险度，对临床有着重要的指导意义[6]。Ⅲ、Ⅳ、Ⅴ类不确定结节存在一定的恶性风险，临床管理中可选择手术治疗，患者需承担手术带来的身心负担和术后并发症风险[7]。

收稿日期：2023-02-07
基金项目：北京市科委其他项目 (Z221100003522001)；解放军总医院临床科研扶持基金 (ZH19021)
作者简介：林宇，女，在读硕士，医师。研究方向：临床病理。Email: ly2020xgm@163.com
通信作者：王宏伟，男，主任医师，主任，硕士生导师。Email: whw6626@sina.com

Jassal 等[8]将基于 4 种分类算法 (k-近邻、支持向量机、决策树、朴素贝叶斯) 的 AI 模型应用于 1 055 个甲状腺结节的细胞学涂片，以切除标本的组织学诊断作为金标准，利用细胞学专家描述的微观特征并整合血清促甲状腺激素水平、超声特征等临床资料评估了模型对甲状腺结节良恶性的诊断性能，结果显示支持向量机在预测结节良恶性方面表现最好，敏感度为 89%，特异度为 83%，准确率为 89%，AUC=0.860。因此，研究者认为该模型有希望被用于资源贫乏地区，可作为非甲状腺专家的决策支持工具。

1.2 基于基因层面的 AI 分子检测给术前未明确诊断的甲状腺结节提供了一种非侵入性的检查方法。二代测序可以在一次操作中对多个基因高速并行分析，结果输出的准确性很高。AI 作为一种大数据算法，可以在不同的学习任务中集成多组数据，并自动实现高层特征的检测或分类。一些遗传分类器，基于基因突变的分类器 (ThyroSeq) 和 microRNA 的分类器在甲状腺结节术前诊断中发挥各自的优势[9-11]。有学者使用 ThyroSeq V3 鉴别 FNA 中不确定结节的良恶性，敏感度为 98.0%、特异度为 81.8%，准确率为 90.9%，AUC=0.906。研究结果显示该基因表达分类器改善术前风险评估的巨大潜力。Patel 等[11]发现 RNA 基因测序分类器 (genomic sequencing classifier，GSC) 在 Bethesda Ⅲ类和Ⅳ类甲状腺结节良恶性鉴别中表现出较高的测试敏感度 (91.00%) 和特异度 (68.00%)，并在嗜酸细胞肿瘤的鉴别上也具有一定的特异度 (58.80%)，因此认为 GSC 作为临床辅助诊断的工具，有望减少不必要的诊断性手术，降低医疗成本，提高患者的生活质量。

1.3 基于数字图像的 AI 数字切片扫描仪的引入促进了 AI 在细胞病理学中的应用。Yao 等[12]在 40 个甲状腺细胞学涂片的全切片数字图像 (whole slide imags，WSIs) 上生成 400 张低倍 (100×) 和 400 张高倍 (400×) 图像，利用特征工程和监督机器学习的数字图像分析方法，对 Bethesda Ⅲ类结节进行良恶性鉴别。最终该方法在低倍图像的 AUC 为 0.750，高倍图像的 AUC 为 0.740，准确率高于病理医师 (0.71 *vs* 0.65)。Hirokawa 等[13]利用卷积神经网络模型 (EfficientNetV2) 对 393 个甲状腺结节的 14 895 张 WSIs 通过 5 次交叉验证进行良恶性鉴别，最终模型诊断的敏感度、特异度、阳性预测值和阴性预测值分别为 94.4%、15.4%、60.7% 和 66.7%，AUC＞0.950，然而该模型难以将甲状腺乳头状癌、甲状腺髓样癌和滤泡癌进行区分。Elliott Range 等[14]使用 WSIs 作为卷积神经网络 (convolutional neural network，CNN) 的输入，采用 2 个基于视觉几何组 (visual geometry group，VGG) 框架的 CNN 组成的半监督方法，第 1 个 CNN 识别感兴趣区 (regions of interest，ROI)，即病理专家标注的含有滤泡上皮细胞的区域；第 2 个 CNN 确定 ROI 的 TBSRTC 分类和最终组织病理 (良/恶性)。AI 在区分 ROI 和非 ROI 时所获得的 AUC 为 0.985，诊断最终组织病理恶性的敏感度和特异度分别为 92.0% 和 90.5%。AI 和病理学家诊断的 AUC 分别为 0.932 和 0.931。遗憾的是，AI 在根据 TBSRTC 分类的良恶性病例中，对恶性结节的诊断性能较差。因此，Elliott Range 等对算法的操作进行了修改，结节 TBSRTC 良/恶性分类以病理医师的诊断为主，其他 TBSRTC 分类则以 AI 诊断为主。最后 AI 和病理医师共同诊断甲状腺结节恶性的特异度从 90.5% 升至 92.9%，AUC 从 0.931 升至 0.962。相比 AI 或病理医师单独诊断取得了更好的结果。Dov 等[15]在 Elliott Range 的研究基础上通过 WSIs 提出了一种能够更方便快速识别 ROI 的 AI 算法，并且利用 CNN 以近似甲状腺细胞病理学报告系统的方式预测结节的 TBSRCT 类别，最后获得的 AUC 与病理专家的 AUC (0.932) 相当，且诊断的准确率有所提高 (0.872 *vs* 0.808)。研究者指出，他们的算法对初级病理医师来说是一个很好的工具，可以在没有病理专家的情况下帮助其快速诊断不确定的病例。

2 人工智能辅助甲状腺细胞病理学诊断现状

2.1 甲状腺滤泡性和嗜酸性病变的鉴别 当影像学检查提示结节有恶性征象时，建议行 FNA，通过细胞学特征结合基因检测来判断结节的良恶性，但容易在观察者之间产生差异。为了解决这些问题，已经有一些 AI 的模型应用于甲状腺细胞学病理诊断，并具有一定的准确性。1996 年，Karakitsos 等[16]尝试使用学习向量量化 (learning vector quantization，LVQ) 对滤泡性和嗜酸性病变进行四分类，即区分良/恶性滤泡性病变和嗜酸性病变，他们对 51 例患者的 Giemsa 染色细胞学涂片中约 100 个细胞核的面积和密度特征进行测量，利用这些测量结果来训练和测试 LVQ 神经网络。结果显示 LVQ 神经网络二分类的性能要优于四分类，即能够很好地区分良性和恶性病变，对良性和恶性细胞核的检测准确率分别为 91.12% 和 89.64%。随后，Gopinath 和 Shanthi[17]用类似的方

法鉴别甲状腺结节的良恶性，研究者采用数学形态学分割方法从自动裁剪的子图像中提取含有甲状腺滤泡上皮细胞的区域；然后用二级小波分解方法提取滤泡上皮细胞的纹理特征，分别使用决策树、k-近邻、Elman 神经网络和支持向量机分类器对甲状腺结节进行良恶性分类，最后利用多数投票和线性组合规则将四个独立分类器的输出融合在一起，以提高诊断性能。Elman 神经网络和支持向量机分类器的分类结果显示，总体诊断准确率为 90%，敏感度为 85% 和 100%，特异度为 90% 和 90%。而采用多数投票规则和线性组合规则的多分类器融合诊断的准确率达到 96.66%。实验结果表明，该方法可以作为甲状腺癌诊断的辅助工具，为病理医师的诊断决策提供第二意见。

2.2　滤泡腺瘤和滤泡性癌的鉴别　滤泡腺瘤和滤泡癌的主要鉴别点是通过手术切除标本判断有无包膜和血管侵犯，在细胞学涂片上区分滤泡腺瘤和滤泡癌是一项具有挑战性的任务，仅借助于细胞形态很难区分。Savala 等[18] 和 Alabrak 等[19] 尝试借助 AI 方法解决这个问题。Savala 等[18] 基于人工神经网络模型 (artificial neural network，ANN)，利用甲状腺滤泡上皮细胞的核特征 (核面积、直径、形态等) 成功区分了滤泡腺瘤和滤泡癌。最终取得的准确率为 100%，AUC=1，但此研究纳入的样本量只有 57 例 。而 Alabrak 等[19] 则是基于 CNN 模型，利用从 BethesdaⅣ类结节中提取的 886 张 (滤泡腺瘤 527 例，滤泡癌 359 例) 数字图像训练 CNN 模型，以 256×256 的彩色图像作为输入，通过卷积层和汇集层自动提取图像特征，最后输出滤泡腺瘤与滤泡癌的二值化分类。结果显示，该模型鉴别滤泡腺瘤与滤泡癌的准确率为 78.0%，敏感度为 88.4%，特异度为 64.0%，AUC=0.870。因此研究者认为 CNN 模型可作为 BethesdaⅣ类甲状腺结节亚分类的辅助诊断技术。

2.3　乳头状癌和非乳头状癌的鉴别　甲状腺乳头状癌 (papillary thyroid carcinoma，PTC) 是最常见的甲状腺癌，区分 PTC 和良性病变是开发用于甲状腺细胞学分类和诊断的 AI 的非常重要一步。Guan 等[20] 利用两种基于深度卷积神经网络的模型 (VGG-16 和 Inception-v3)，从 279 例患者的结节样本中提取出 887 张图像，使用其中的 407 张 PTC 患者图像和 352 张良性结节患者图像作为训练集，69 张 PTC 患者图像和 59 张良性结节患者图像作为测试集，以 PTC 患者的组织学病理结果以及良性结节患者的临床数据、实验室和影像学结果为金标准。结果显示，VGG-16 和 Inception-v3 在 PTC 和良性结节细胞学涂片的鉴别中均表现出较优的性能，VGG-16 检测 PTC 的敏感度为 100.00%，特异度为 94.91%，Inception-v3 检测 PTC 的敏感度为 98.55%，特异度为 86.44%。然而，这项研究并没有将 AI 模型的诊断性能与病理医师的诊断性能进行比较。Sanyal 等[21] 也做了类似的研究，他们使用了 184 张 PTC 图像和 186 张非 PTC 图像，放大倍数为 40 ×或 10 ×。评估采用两个标准：(1) 病变必须在任何放大倍数 (10 ×或 40 ×) 中标记为 PTC，才可诊断为 PTC；(2) 病变必须在 10 ×和 40 ×放大倍数下标记为 PTC，才能诊断为 PTC。结果显示两种算法的敏感度分别为 90.48% 和 33.33%，特异度分别为 83.33% 和 98.48%，准确率分别为 85.06% 和 82.76%。这为 PTC 的诊断开辟了一个新视角。但此项研究未与病理医师的诊断进行比较。与滤泡腺瘤和滤泡癌类似，具有乳头状核特征的非浸润性甲状腺滤泡性肿瘤 (non-invasive follicular thyroid neoplasm with papillary-like nuclear features，NIFTP) 与 PTC 的鉴别，需要手术切除后行组织病理明确诊断。Maleki 等[22] 使用了基于语义和文本的分析方法，提出了一个基于甲状腺 FNA 显微镜下描述的模型来鉴别 PTC 和 NIFTP，他们通过模型处理关键词和短语训练支持向量机，该模型成功将 PTC 和 NIFTP 区分开来，诊断的敏感度为 72.6%，特异度为 81.6%。虽然 NIFTP 的诊断不是细胞学诊断，并且其诊断依赖于手术切除标本除外包膜和血管侵犯以及乳头状结构的存在，但此研究在细胞学上对 NIFTP 的鉴别也有一定帮助。

3　人工智能在甲状腺细胞病理学诊断中的展望

3.1　价值与优势　在过去的几年里，AI 在风险分层、筛选、诊断、预测和治疗计划等问题上的表现可以与临床专家相媲美[23-25]。数字病理的出现使病理图像电子化储存得以实现，随着数字切片逐渐整合到临床病理工作流程中，先进的算法和 AI 辅助诊断技术将病理学家的视野扩展到显微镜载玻片之外，并能够真正利用并整合超越人类极限和界限的知识[26]。数字病理和 AI 在肿瘤学和精准医疗领域具有巨大的潜力。一方面借助相关深度学习的技术，使许多手工和主观的任务变得自动化和更加标准化，从而提高诊断的客观性和准确性；另一方面数字切片扫描仪的引入使病理医师能够更方便和灵活地管理数字切片，减少了病

理医师的周转时间，提高了工作效率。AI 在病理学应用中的独特优势有可能改变病理学诊断的方式，实现共享图像和数据，提高效率和综合诊断效能，使病理工作流程逐渐数字化。

3.2 面临的挑战 (1) 数据有效性：多数 AI 研究是在单中心、小数据集上进行的，虽然有些研究取得了令人欣慰的结果，但它们只在较少的样本中进行了验证。因此，更需要关注的是如何对 AI 算法的性能做出更有说服力的研究。收集跨区域、多中心和样本量更大的数据集是验证 AI 算法能力的关键一步，然而由于各个实验室组织固定、染色流程、成像设备、扫描方案的不同导致输入数据的异质性，这给多中心数据的训练带来了挑战[27]。因此，有必要制定并推广统一的扫描规则，并且采用领域自适应的方法减少多源数据的迁移，提高模型的泛化能力。此外目前的方法主要是通过监督学习来训练的，这种学习依赖于精细的人类标记数据。然而，在医疗任务中，由于各领域专家有限，难以获得大量高质量的标记数据，更多的未标记数据尚未被利用[28-29]。近年来，人们提出了各种各样的自监督学习和半监督学习方法来挖掘被忽略的未标记数据，可以解决标记数据有限的问题。(2) 模型训练与验证：在训练 AI 模型的过程中，可能会出现过拟合、系统偏差、性能漂移等问题[5]。这些问题非常具有挑战性，并且会对 AI 系统的性能产生不利影响，甚至会导致 AI 在很长一段时间内出现不能被发现的错误。虽然很多关于 AI 算法的研究正在进行，但由于缺乏统一的临床验证标准和机构验证，大多数研究仍使用自主标准，所以即使深度学习算法表现良好，得到了临时许可，并且在后续的训练过程中进行了升级，也很难确认其在使用时是否也具有同样的性能[30]。(3) 医工结合的难点：将 AI 应用到甲状腺病理诊断中需要获得监管机构的批准，最重要的方面之一是让临床医师了解其工作原理[31]。然而由于 AI 过度参数化的黑盒特性，临床医师很难准确理解它的功能；并且由于深度学习人工神经网络难以开放，巨量参数如何同时交互作用难以理解，导致 AI 的可解释性有限[32]。有研究者试图将深度学习算法变成“玻璃盒子”，以人们可以理解和分析的方式，理清输入和计算输出[33-34]。(4) 人工智能环境：要发展一个稳定的人工智能环境，必须有一定程度的监管创新；在医疗保健中 AI 的监管应考虑到对人类的整体影响，而不是 AI 单独做出结论。

4 结语

随着 AI 的发展，传统病理学正在向数字病理学发展和演变。将数字病理和 AI 结合在一起可能会给病理诊断带来史无前例的深刻变化。我们认为，病理学家从一开始在开发算法的时候就应该参与进来，以确保这些备受期待的软件包得到改进，甚至被 AI 算法取代。尽管该领域尚未完全成熟，仍需要解决大量技术、伦理和监管等问题，但其会随着研究人员、医师、监管机构和患者倡导团体的创新逐渐成熟并不断为医疗从业者带来新技术。我们期待 AI 在甲状腺细胞病理学的未来发挥更大的支撑作用，使诊断更加高效和准确。

作者贡献 林宇：研究概念，设计，数据收集与分析，文章撰写；王宏伟：研究内容提出，研究设计，论文写作指导与支持；罗渝昆：研究内容提出，研究设计，论文写作指导与支持；李杰：研究内容提出，研究设计，论文写作指导与支持。

利益冲突 所有作者声明无利益冲突。

参考文献

1 Miranda-Filho A，Lortet-Tieulent J，Bray F，et al. Thyroid cancer incidence trends by histology in 25 countries：a population-based study［J］. Lancet Diabetes Endocrinol，2021，9（4）：225-234.

2 Kaliszewski K，Ludwig M，Ludwig B，et al. Update on the diagnosis and management of medullary thyroid cancer：what has changed in recent years?［J］. Cancers，2022，14（15）：3643.

3 De Vries BM，Zwezerijnen GJC，Burchell GL，et al. Explainable artificial intelligence（XAI）in radiology and nuclear medicine：a literature review［J］. Front Med，2023，10：1180773.

4 Cao CL，Li QL，Tong J，et al. Artificial intelligence in thyroid ultrasound［J］. Front Oncol，2023，13：1060702.

5 Van Der Laak J，Litjens G，Ciompi F. Deep learning in histopathology：the path to the clinic［J］. Nat Med，2021，27（5）：775-784.

6 Cibas ES，Ali SZ. The 2017 Bethesda system for reporting thyroid cytopathology［J］. J Am Soc Cytopathol，2017，6（6）：217-222.

7 Patel KN，Yip L，Lubitz CC，et al. Executive summary of the American association of endocrine surgeons guidelines for the definitive surgical management of thyroid disease in adults［J］. Ann Surg，2020，271（3）：399-410.

8 Jassal K，Koohestani A，Kiu A，et al. Artificial intelligence for pre-operative diagnosis of malignant thyroid nodules based on sonographic features and cytology category［J］. World J Surg，2023，47（2）：330-339.

9 Nikiforova MN，Mercurio S，Wald AI，et al. Analytical performance of the ThyroSeq v3 genomic classifier for cancer diagnosis in thyroid nodules［J］. Cancer，2018，124（8）：1682-1690.

10 Lithwick-Yanai G，Dromi N，Shtabsky A，et al. Multicentre validation of a microRNA-based assay for diagnosing indeterminate thyroid nodules utilising fine needle aspirate

smears［J］. J Clin Pathol，2017，70（6）：500-507.
11 Patel KN，Angell TE，Babiarz J，et al. Performance of a genomic sequencing classifier for the preoperative diagnosis of cytologically indeterminate thyroid nodules［J］. JAMA Surg，2018，153（9）：817-824.
12 Yao K，Jing X，Cheng J，et al. A study of thyroid fine needle aspiration of follicular adenoma in the "atypia of undetermined significance" Bethesda category using digital image analysis［J］. J Pathol Inform，2022，13：100004.
13 Hirokawa M，Niioka H，Suzuki A，et al. Application of deep learning as an ancillary diagnostic tool for thyroid FNA cytology［J］. Cancer Cytopathol，2023，131（4）：217-225.
14 Elliott Range DD，Dov D，Kovalsky SZ，et al. Application of a machine learning algorithm to predict malignancy in thyroid cytopathology［J］. Cancer Cytopathol，2020，128（4）：287-295.
15 Dov D，Kovalsky SZ，Feng Q，et al. Use of machine learning-based software for the screening of thyroid cytopathology whole slide images［J］. Arch Pathol Lab Med，2022，146（7）：872-878.
16 Karakitsos P，Cochand-Priollet B，Pouliakis A，et al. Learning vector quantizer in the investigation of thyroid lesions［J］. Anal Quant Cytol Histol，1999，21（3）：201-208.
17 Gopinath B，Shanthi N. Development of an automated medical diagnosis system for classifying thyroid tumor cells using multiple classifier fusion［J］. Technol Cancer Res Treat，2015，14（5）：653-662.
18 Savala R，Dey P，Gupta N. Artificial neural network model to distinguish follicular adenoma from follicular carcinoma on fine needle aspiration of thyroid［J］. Diagn Cytopathol，2018，46（3）：244-249.
19 Alabrak MMA，Megahed M，Alkhouly AA，et al. Artificial intelligence role in subclassifying cytology of thyroid follicular neoplasm［J］. Asian Pac J Cancer Prev，2023，24（4）：1379-1387.
20 Guan Q，Wang YJ，Ping B，et al. Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images：a pilot study［J］. J Cancer，2019，10（20）：4876-4882.
21 Sanyal P，Mukherjee T，Barui S，et al. Artificial intelligence in cytopathology：a neural network to identify papillary carcinoma on thyroid fine-needle aspiration cytology smears［J］. J Pathol Inform，2018，9：43.
22 Maleki S，Zandvakili A，Gera S，et al. Differentiating noninvasive follicular thyroid neoplasm with papillary-like nuclear features from classic papillary thyroid carcinoma：analysis of cytomorphologic descriptions using a novel machine-learning approach［J］. J Pathol Inform，2019，10：29.
23 Esce AR，Redemann JP，Sanchez AC，et al. Predicting nodal metastases in papillary thyroid carcinoma using artificial intelligence［J］. Am J Surg，2021，222（5）：952-958.
24 Djuric U，Zadeh G，Aldape K，et al. Precision histology：how deep learning is poised to revitalize histomorphology for personalized cancer care［J］. NPJ Precis Oncol，2017，1（1）：22.
25 Kather JN，Pearson AT，Halama N，et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer［J］. Nat Med，2019，25（7）：1054-1056.
26 Jahn SW，Plass M，Moinfar F. Digital pathology：advantages，limitations and emerging perspectives［J］. J Clin Med，2020，9（11）：3697.
27 Colling R，Pitman H，Oien K，et al. Artificial intelligence in digital pathology：a roadmap to routine use in clinical practice［J］. J Pathol，2019，249（2）：143-150.
28 Bera K，Schalper KA，Rimm DL，et al. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology［J］. Nat Rev Clin Oncol，2019，16（11）：703-715.
29 Niazi MKK，Parwani AV，Gurcan MN. Digital pathology and artificial intelligence［J］. Lancet Oncol，2019，20（5）：e253-e261.
30 Andreas K，Johan SO，Sepp DR，et al. Designing deep learning studies in cancer diagnostics［J］. Nat Rev Cancer，2021，21（3）：199-211.
31 Kelly CJ，Karthikesalingam A，Suleyman M，et al. Key challenges for delivering clinical impact with artificial intelligence［J］. BMC Med，2019，17（1）：195.
32 Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead［J］. Nat Mach Intell，2019，1（5）：206-215.
33 Guidotti R，Monreale A，Ruggieri S，et al. A survey of methods for explaining black box models［J］. ACM Comput Surv，2019，51（5）：93.1-93.42.
34 Richens JG，Lee CM，Johri S. Improving the accuracy of medical diagnosis with causal machine learning［J］. Nat Commun，2020，11（1）：3923.

（责任编辑：施晓亚）