

## 人工智能助力皮肤罕见病诊疗

王煜坤, 刘 洁

中国医学科学院北京协和医院皮肤科 疑难重症及罕见病国家重点实验室  
国家皮肤与免疫疾病临床医学研究中心, 北京 100730

通信作者: 刘 洁, E-mail: Liujie04672@pumch.cn

**【摘要】** 在临床实践中, 皮肤罕见病的早期诊断、准确评估和有效治疗均存在一定困难。随着大数据时代的到来, 图像数据、多组学数据及电子病历等生物医疗数据呈指数级增长, 以机器学习为代表的人工智能 (AI) 在处理大量复杂信息中极具优势, 已有研究将 AI 引入到皮肤罕见病的辅助诊疗中。本文分别对基于图像数据、多组学数据和文本数据的 AI 在辅助皮肤罕见病诊疗及 AI 辅助皮肤罕见病药物探索等方面的研究进行了简述、探讨和展望, 以期提高皮肤科医师对该领域的认识并积极推动皮肤罕见病人工智能研究发展。

**【关键词】** 人工智能; 深度学习; 卷积神经网络; 皮肤罕见病

**【中图分类号】** R-1; R44; R75 **【文献标志码】** A **【文章编号】** 2097-0501(2023)02-0157-07

**DOI:** 10.12376/j. issn. 2097-0501. 2023. 02. 003

## Artificial Intelligence Applications in Rare Skin Diseases

WANG Yukun, LIU Jie

Department of Dermatology, State Key Laboratory of Complex Severe and Rare Diseases,  
National Clinical Research Center for Dermatologic and Immunologic Diseases, Peking Union Medical College Hospital,  
Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100730, China

Corresponding author: LIU Jie, E-mail: Liujie04672@pumch.cn

**【Abstract】** In clinical practice, early diagnosis, accurate assessment and effective management of rare skin diseases are difficult. The Big Data gives rise to the exponential growth of biomedical data, including medical images, multi-omics information and electronic health records. Artificial intelligence (AI), particularly machine learning, has its advantage in processing complex and abundant information. Researches have applied AI in the field of rare skin diseases. In this paper, we briefly describe, discuss, and foresee the research on AI based image data, multi-omics data & text data and AI in assisting rare skin disease drug development, in order to improve the awareness of dermatologist understanding of this field and actively promote the development of AI usage on rare skin diseases.

**【Key words】** artificial intelligence; deep learning; convolutional neutral network; rare skin diseases

基金项目: 中央高水平医院临床科研业务费 (2022-PUMCH-B-092); 国家自然科学基金 (82173449); 北京市自然科学基金 (7232114); 中国医学科学院医学与健康科技创新工程 (2022-I2M-C&T-A-007)

引用本文: 王煜坤, 刘洁. 人工智能助力皮肤罕见病诊疗 [J]. 罕见病研究, 2023, 2 (2): 157-163. doi: 10.12376/j. issn. 2097-0501. 2023. 02. 003.

**Funding:** National High Level Hospital Clinical Research Funding (2022-PUMCH-B-092); National Natural Science Foundation (82173449); Beijing Natural Science Foundation (7232114); Chinese Academy of Medical Sciences (CAMS) Innovation Fund for Medical Sciences (2022-I2M-C&T-A-007)

*J Rare Dis*, 2023,2(2):157-163

罕见病一般指在某个特定区域内患病人数相对较少的疾病,由于各个国家和地区的社会情况、经济水平和卫生资源等因素存在差异,罕见病目前尚无国际统一的定义<sup>[1]</sup>。中华医学会医学遗传学分会曾建议将罕见病定义为患病率低于1/50万或新生儿发病率低于1/万的疾病,由此可预估中国罕见病总患病人口约为1680万<sup>[2]</sup>。皮肤罕见病或存在皮肤表现的系统性罕见病是其中重要的组成部分,在这些疾病中,皮肤的各种组织结构均可受累,并由不同的病理生理机制导致形态各异的皮肤病变和功能受损。目前,根据国际疾病分类(International Classification of Disease, ICD)-10标准,已知的皮肤罕见病有450余种,包括罕见的皮肤角化异常性疾病、色素异常性疾病、代谢性疾病、自身免疫性疾病和皮肤肿瘤等<sup>[3]</sup>,而中国国家罕见病注册系统(National Rare Diseases Registry System, NRDRS)收录的相关病种也已达50余种,如鱼鳞病、掌跖角化病、遗传性大疱性表皮松解症(epidermolysis bullosa, EB)、泛发性脓疱型银屑病(generalized pustular psoriasis, GPP)、恶性黑色素瘤(malignant melanoma, MM)及原发性皮肤淋巴瘤等,并且尚在不断新增和完善。由于罕见病患者稀少而分散,相关教学培训也相对缺乏,多数皮肤科医师对罕见病的识别能力有限,且因其临床表型复杂多变,即使在医疗资源相对丰富的大型医学中心,也常需要耗费大量时间和经济成本进行多学科联合会诊才能明确诊断及治疗方案。

人工智能(artificial intelligence, AI)为计算机科学的分支,通过对多模态数据的特征进行提取和学习,并以神经网络等方式进行处理,能够根据待解决问题的类型高效地对大量复杂数据进行分析、识别和分类等。随着计算机软硬件的更新迭代,AI的性能得到了长足发展,与医疗健康的结合也愈发紧密,其中尤其以深度学习为代表的机器学习应用最为突出。皮肤病具有可视性强和影像资料丰富等特点,卷积神经网络(convolutional neural network, CNN)等深度学习技术在多种皮肤疾病的皮损图像分割、分类识别及病情评估等方面的诊疗水平已达到甚至超过皮肤科医师,成为近年来皮肤科领域的研究热点<sup>[4]</sup>。近

年来,图像数据、多组学数据和电子病历(electronic health records, EHRs)等不同来源的数据信息不断积累,基于知识和数据驱动的AI辅助临床诊断和治疗决策系统也得到迅速发展,其中一些工具已在罕见病诊疗中得到初步研究和应用,如基于临床摄影、皮肤镜和组织病理等各种类型图像进行皮肤罕见病的辅助诊断、病情评估和预后预测,通过分析基因组学数据,可辅助识别致病变异与疾病表型之间的联系,以进行罕见病药物探索与研发,以及对EHRs等文本数据进行信息抽取和分析等,显示出AI在辅助皮肤罕见病诊疗方面的巨大潜能。本文分别对基于图像数据、多组学数据和文本数据的AI在辅助皮肤罕见病诊疗及AI辅助皮肤罕见病药物探索等方面的研究进行了简述、探讨和展望,以期提高皮肤科医师对该领域的认识并积极推动皮肤罕见病AI研究发展。

## 1 基于图像数据的皮肤罕见病人工智能辅助诊疗研究

### 1.1 基于临床和皮肤镜图像的人工智能辅助恶性黑色素瘤诊疗

AI在皮肤疾病研究中最突出的领域为计算机视觉,也是具有较好前景的临床实践应用场景之一。目前,在皮肤图像分类中最大、使用频率最高的公开数据集为国际皮肤成像合作组织(International Skin Imaging Collaboration, ISIC)所提供的ISIC数据集。自2016年起,ISIC与其他领先的计算机视觉会议共同组织关于图像分割、特征提取及疾病分类的年度挑战赛,各年度积累的数据信息使得ISIC数据集的规模和复杂度不断提升,而在中国发病率较低的MM及需与之鉴别的皮肤肿瘤(如色素痣)是其中最主要的病种<sup>[5]</sup>。2017年,Esteva等<sup>[6]</sup>基于近13万张皮肤疾病图像(以临床图像为主,部分源自ISIC数据集)构建CNN,其在鉴别MM与色素痣的测试中已达到皮肤科专家的平均分类水平,受试者工作特征(receiver operating characteristics, ROC)曲线下面积(area under the curve, AUC)为0.96,该研究发表于*Nature*,成为皮肤科领域AI研究的里程碑。由于皮肤镜图像采

集更易标准化、背景噪声较临床图像显著降低，且皮肤镜对皮损有放大及去除表面反光作用，皮肤镜图像可显示皮损更多的亚宏观细节特征，从而为 CNN 所识别和学习，因此，在 CNN 辅助诊断 MM 等皮肤肿瘤中，多数研究采用皮肤镜图像数据。Haenssle 等<sup>[7]</sup> 研究使用皮肤镜图像训练的 CNN 与 58 位皮肤科医师的诊断能力对比，结果显示医师鉴别 MM 和色素痣的平均敏感性和特异性分别达 86.6% 和 71.3%，如获取临床信息，医师的平均敏感性和特异性分别提升至 88.9% 和 75.7%，而在相同的敏感性下，CNN 的特异性为 82.5%，CNN 的 AUC 值 (0.86) 亦高于医师 (0.79)，显示出 AI 鉴别黑素细胞源性肿瘤的良好性能具有极佳准确性。Marchetti 等<sup>[8]</sup> 的三分类 (MM、色素痣和脂溢性角化病) 研究中，将 CNN 的分类结果用于替换皮肤科医师诊断自信度较低的皮损分类结果，以模拟 CNN 辅助医师诊断的场景，最终使高年资医师的诊断准确性由 73.4% 提升至 75.4%，住院医师则由 69.4% 提升至 72.6%。亚洲人种最常见的 MM 类型为肢端雀斑样痣黑素瘤 (acral lentiginous melanoma, ALM)。Lee 等<sup>[9]</sup> 基于 1072 张 ALM、肢端色素痣和介于两者之间的交界性肿瘤皮肤镜图像构建了一个能够准确识别 ALM 的 CNN，即 ALMnet，其 AUC 值达 0.976，准确性、敏感性和特异性分别为 92.5%、90.0% 和 95.0%。该研究中，60 位皮肤科医师分别仅基于皮肤镜图像及同时参考临床病史对皮损作出是否活检的决策，准确性为 74.7% 和 79.0%，而在参考 ALMnet 的分类结果及其可能性后，准确性显著提升至 86.9%，体现出 AI 辅助皮肤科医师对 MM 进行诊断和活检决策的重要应用前景。

### 1.2 基于组织病理图像的人工智能辅助恶性黑素瘤诊疗

Brinker 等<sup>[10]</sup> 基于 HE 染色的组织病理图像构建了可从病理水平辅助诊断 MM 的 CNN，无论是否经专家标识肿瘤区域，该 CNN 的识别准确性 (94.0% 和 88.0%) 均与 18 名皮肤病理医师的平均诊断准确性 (88.88%) 相仿，两者的不一致率仅 13.45%。Lodha 等<sup>[11]</sup> 曾报道不同皮肤病理医师在病理诊断 MM 和色素痣时不一致率为 25%，提示 AI 可为经验不足的皮肤病理医师有效进行 MM 辅助诊断，以及为皮肤病理专家进行黑素细胞源性皮损组织病理切片的高效筛检。Comes 等<sup>[12]</sup> 基于来自公开数据集的 MM 组织病理图像构建了可预测 I ~ III 期患者 1 年无病生存率 (disease-free survival, DFS) 的深度神经网络模型，

该模型可通过识别 MM 组织病理图像上定量的生物标志物来进行预后预测，在验证中该模型的 AUC 为 0.667，准确性为 72.7%。

### 1.3 基于图像数据的人工智能在其他皮肤罕见病中的应用

除研究最为深入的 MM 外，基于图像数据的 AI 也应用于其他罕见的皮肤肿瘤和非肿瘤皮肤病的辅助诊断研究中。蕈样肉芽肿 (mycosis fungoides, MF) 为常见的原发性皮肤淋巴瘤，本研究团队<sup>[13]</sup> 基于 959 张皮肤镜图像构建了可对早期 MF 及其临床表现与之相似的炎症性皮肤病进行二分类的 CNN，并与 13 名不同经验的皮肤科医师诊断结果进行对比，结果显示医师鉴别早期 MF 和炎症性皮肤病的敏感性、特异性和 Kappa 值分别为 70.19%、94.74% 和 0.677，而 CNN 分别为 87.50%、93.85% 和 0.920，体现出该模型对早期 MF 具有较好识别能力，总体优于医师的平均诊断水平。复旦大学和南京邮电大学的团队<sup>[14]</sup> 则构建了可在组织病理图像中快速、准确识别乳房外 Paget 病 (extramammary Paget's disease, EMPD) 肿瘤细胞的 CNN，在 40 倍放大的图像中识别准确性达 95.52%，对于提高基层皮肤病理医师对 EMPD 的诊断准确性方面具有潜在价值。红皮病型银屑病和脓疱型银屑病为银屑病的罕见亚型，Aijaz 等<sup>[15]</sup> 基于临床图像构建了可分类银屑病不同亚型的 CNN 模型，总体分类准确性达 84.2%，其中，CNN 对红皮病型银屑病的分类敏感性和特异性均为 85%，而对脓疱型银屑病的分类敏感性和特异性分别为 73% 和 75%。2022 年，中国医学科学院北京协和医院皮肤科推出了中国首个罕见皮肤病 AI 辅助识别工具，该程序可于手机端基于患者或医生上传的皮损图像进行 GPP 的初步疾病风险评估，识别准确率可达 85% 以上，此外，还为识别为高风险的患者提供导诊服务，引导其至全国 14 个城市 30 家具备 GPP 规范诊疗能力的医院就诊，帮助患者及时获得规范化诊疗方案<sup>[16]</sup>。Akay 等<sup>[17]</sup> 基于 1888 张正常皮肤和系统性硬化症 (systemic sclerosis, SSc) 患者皮肤的组织病理图片构建了轻量级网络 MobileNetV2，该模型鉴别正常皮肤、早期 SSc 和晚期 SSc 病变的准确性为 94.8%，可作为 SSc 的组织病理辅助筛检工具。Shi 等<sup>[18]</sup> 基于 106 张直接免疫荧光图像构建了可辅助筛检获得性 EB 的 CNN，其可自动识别 u 形锯齿样抗体沉积模式，诊断准确性达 89.3%。

## 2 基于多组学数据的皮肤罕见病人工智能辅助诊疗研究

多组学数据包括基因组学、转录组学、蛋白质组学、代谢组学、脂质组学等多种类型信息，数据量庞大，数据结构复杂，适合作为 AI 分析的数据来源。

### 2.1 基于基因组学的人工智能辅助基因变异识别、注释和疾病关联研究

在基因组学中，多种 AI 技术已成功应用于识别变异、变异分类和基因型-表型整合分析等任务<sup>[19]</sup>。二代测序 (next-generation sequencing, NGS) 的高通量测序程序可能存在技术性误差和数据错误，因此非常依赖于一些生物信息学工具进行数据解释，如 GATK (genome analysis toolkit) 和 SAMtools 等<sup>[20]</sup>。为了克服单纯使用这些传统工具的局限性，Poplin 等<sup>[21]</sup>提出一种基于 CNN 的变异识别算法 DeepVariant，其将 NGS 获取的测序数据拼接和检测任务转化为图像分类任务，可读取大量碱基序列并准确识别其中的单核苷酸多态性 (single nucleotide polymorphisms, SNPs) 和碱基插入/缺失变异，性能优于现有的标准生物信息工具<sup>[22]</sup>。Luo 等<sup>[23]</sup>提出了可用于单分子测序识别变异多重任务的 CNN 模型，即 Clairvoyante，其可识别变异类型 (SNPs 或碱基插入/缺失)、合子型、交替等位基因及碱基插入/缺失的长度，与 DeepVariant 相比，Clairvoyante 能够提供包括变异类型等所识别变异的详细信息，性能优于 DeepVariant。在识别和注释变异后，需要专家根据基因本体 (gene ontology, GO) 数据库、Orphanet 数据库等专业的基因数据库对其进行排序和过滤，将临床意义可能较小的变异筛选后，即可进一步明确导致基因功能改变和疾病表型的致病变异。深度学习的引入可对该过程进行极大的优化和提速。Boudelloua 等<sup>[24]</sup>构建了融合自动推理和深度神经网络的模型 DeepPVP，其可以较传统方法更快速地对全外显子或全基因组测序数据进行变异排序，从而更准确地定位致病变异。此外，传统的全基因组关联研究 (genome-wide association studies, GWAS) 只能检测单个 SNP 与所研究疾病的关系，也无法直接将变异与调控机制进行联系，而 Arloth 等<sup>[25]</sup>构建的 CNN 模型 DeepWAS 将单个变异的调控效应整合进多元 GWAS 分析中，通过结合多种资源和组织中的基因表达和甲基化数量性状位点 (quantitative-trait loci, QTL) 数据，该算法可探测到疾病相

关的变异位点及其对特定组织中染色质性状的影响。

### 2.2 基于基因组学的人工智能辅助基因表达和调控研究

在基因表达和调控方面，深度学习预测调控序列和分子表型之间的联系取得了重大进步。Jaganathan 等<sup>[26]</sup>提出了包含 32 个卷积层的深度神经网络 SpliceAI，其可通过任意前体 mRNA (pre-mRNA) 的序列来预测剪接，从而发现非编码基因变异所致的隐蔽剪接位点。该团队同时通过 RNA 测序验证了 SpliceAI 预测隐蔽剪接变异结果准确性为 75%。据估计，隐蔽剪接可能导致了约 10% 的罕见遗传病致病变异。Agarwal 等<sup>[27]</sup>描述的深度神经网络模型 Xpresso 则可以仅依据启动子 DNA 序列预测 mRNA 的基因表达水平，可帮助推断转录和转录后基因调控机制，该模型的预测能力与基于大量生物化学数据库信息的模型相仿。在表观基因组学领域，深度学习也可揭示在不影响 DNA 序列的情况下调控基因表达的各种原理和现象，如预测单细胞 DNA 甲基化状态<sup>[28]</sup>和组蛋白修饰<sup>[29]</sup>等。上述研究突显出 AI 技术正在加速基因组学的研究，将有助于以遗传性疾病为主的皮肤罕见病诊疗水平的提升。

### 2.3 基于其他组学数据的人工智能辅助皮肤罕见病诊疗

除基因组学外，其他组学及多组学联合分析也已应用于皮肤病领域<sup>[30]</sup>。目前在皮肤罕见病中也开展了 AI 支持下的相关研究，如转录组学领域，Zhang 等<sup>[31]</sup>利用了一种多任务学习聚类模型 scVDMC (single-cell variance-driven multitask clustering) 来进行单细胞 RNA 测序分析，在隐性营养不良型 EB (recessive dystrophic EB, RDEB) 的单细胞 RNA 测序数据中，除了已知的 *CXCL12/SDF1*、*HMGB1* 等 RDEB 标记基因外，scVDMC 发现了数种新的细胞类型和标记基因，包括 *COL11A1*、*GREM1* 和 *MFAP5*，并在此后的流式细胞术中进行了验证，这些基因所标记的基质细胞亚群可能与 RDEB 患者上皮细胞转化和继发鳞状细胞癌的机制有关。Borcherding 等<sup>[32]</sup>对 1 例 IV A 期 Sézary 综合征 (Sézary syndrome, SS) 患者的外周血单核细胞进行了 RNA 测序，并采用机器学习反向图嵌入算法，识别出了克隆增殖的 SS 细胞 (由 *FOXP3*<sup>+</sup> 逐渐转变为 *GATA3*<sup>+</sup> 或 *IKZF2*<sup>+</sup>) 及其与正常 CD4<sup>+</sup>T 细胞之间的转录水平差异，而使用提升树分类算法依据这种差异进行 SS 早晚期预测的准确性达 80%。Martínez 等<sup>[33]</sup>利用一种非参数无监督算法——

基因集变异分析 (gene set variation analysis, GSVA), 发现皮肤型红斑狼疮 (cutaneous lupus erythematosus, CLE)、银屑病、特应性皮炎和 SSc4 种炎症性疾病的皮损组织基因表达状态有着各自特征, 这种现象经机器学习及分类与回归树 (classification and regression tree, CART) 分析验证, 且上述疾病的非皮损部位皮肤样本也呈现出比皮损部位更显著的基因表达差异, 提示炎症性皮肤病的皮损及非皮损部位皮肤组织均有特征性的异常基因表达谱。

### 3 基于电子病历等其他文本数据的皮肤罕见病人工智能辅助诊疗研究

EHRs 为临床医师在门诊、急诊和住院部等医疗场景通过计算机等电子设备完成的数字化医疗记录, 一般包含患者的病案首页 (基本信息)、主诉、现病史、既往史、体格检查和辅助检查及化验结果等, 对 EHRs 进行合理有效地储存、管理、传输、整理和调阅等处理有助于获取有关皮肤罕见病的大量信息。近年来, 许多医院也已基于大量的 EHRs 等医疗文本数据建立起临床数据仓库 (clinical data warehouses, CDW)。CDW 含有各种结构化和非结构化的文本数据等, 其中, 结构化文本数据具有一定标准性, 如 ICD 编码, 便于后期检索和处理, 而非结构化文本在形成时无一定格式限制, 但包含着大量有关个体疾病表型和治疗相关信息, 如现病史、家族史和用药史等, 尤其对于个体差异较大的罕见病患者诊疗具有重大价值<sup>[34]</sup>。Jamian 等<sup>[35]</sup>分别利用基于规则算法和机器学习算法, 通过识别 EHRs 中患者的 SSc ICD-9、ICD-10-CM 编码、抗核抗体阳性 (滴度 >1:80) 和病历中出现的雷诺现象相关关键词, 能够从大量 EHRs 中发现潜在的 SSc 患者, 最终该团队的随机森林模型阳性预测值、敏感性和 *F* 值分别为 84%、92%、88%。来自四川大学华西医院的研究团队<sup>[36]</sup>近期回顾了该院保存的共 221 例隆突性皮肤纤维肉瘤 (dermatofibrosarcoma protuberans, DFSP) 患者资料, 包括 195 例经典型 DFSP (classic DFSP, C-DFSP) 和 26 例纤维肉瘤型 DFSP (fibrosarcomatous DFSP, FS-DFSP), 结果显示相较于 C-DFSP, FS-DFSP 常发生于老年, 肿瘤细胞不表达 CD34 而有显著更高的 Ki-67 指数, 并基于这些患者数据构建了一个反向传播 (back-propagation, BP) 神经网络模型, 该 BP 神经网络模型对 DFSP 亚型分类的准确率可达 84.1%。

此外, AI 也可以利用来自医院外的非 EHRs 文本数据, 例如, 公共卫生机构的疾病登记系统及专业的医疗网站和/或数据库等。Cheraghlou 等<sup>[37]</sup>回顾了美国国家癌症中心 (National Cancer Database, NCDB) 登记的 Merkel 细胞癌 (Merkel cell carcinoma, MCC) 患者信息, 通过改良 CART 算法识别出了影响 MCC 患者预后的因素和分期, 包括 I 期 (局限性疾病)、II 期 ( $\leq 3$  个淋巴结阳性)、III 期 ( $\geq 4$  个淋巴结阳性) 和 IV 期 (出现远处转移), 在验证集中, 上述四期 MCC 患者的 3 年存活率分别为 81.2%、59.6%、38.0% 和 20.2%, 与美国癌症联合委员会 (American Joint Committee on Cancer, AJCC) 提出的分期系统相比, 通过改良 CART 算法建立的预后分期有更好的组内一致性, 能更准确地判断 MCC 患者预后。

### 4 人工智能辅助皮肤罕见病药物探索

在辅助罕见病治疗方面, AI 在综合各类型数据进行药物重定位、有效性、不良反应预测及疾病潜在治疗靶点发现等领域的应用也有望带来重大突破<sup>[38-39]</sup>。Lee 等<sup>[40]</sup>基于机器学习研发了一个用于人类疾病 RNA 样本注释的统一计算架构 URSAHD (Unveiling RNA Sample Annotation for Human Diseases), URSAHD 整合了大量复杂疾病的临床基因表达谱和分子特征信息, 从而准确进行疾病识别和发现疾病相关基因, 这种数据驱动的 AI 架构能够仅凭基因表达数据将罕见病与有着充分药物治疗研究和经验的疾病相关联, 从而推动罕见病的药物重定位。URSAHD 还能根据疾病相关分子信号来预测药物治疗效应, 该团队利用 URSAHD 分别于铁粒幼红细胞性贫血、难治性贫血伴原始细胞增多中完成了药物探索实践, 并于溃疡性结肠炎和银屑病中进行了药物治疗效应预测验证。在皮肤罕见病中, 来自中山大学附属第一医院的团队<sup>[41]</sup>采用文本数据挖掘、GO 富集分析和蛋白质互作网络分析等方法筛选出 DFSP 潜在的治疗靶点基因和候选药物, 然后通过深度学习模型 DeepPurpose 来进行药物靶标相互作用预测, 最终遴选出包括血小板源性生长因子受体  $\beta$  激酶抑制剂等共 10 种潜在的 DFSP 治疗药物。该团队也利用 DeepPurpose 以类似的方法筛选出了 11 种 Erb-B2 受体酪氨酸激酶 2 (erb-B2 receptor tyrosine kinase 2, ERBB2) 靶向药物, 可能作为 MM 的潜在治疗选择<sup>[42]</sup>。

## 5 结语

AI 凭借着其可处理大量复杂信息的优势, 在辅助皮肤罕见病基因型和表型诊断、鉴别诊断、病情评估、预后预测和药物探索等方面充分展示出卓越的效能和广阔的前景。然而, 与 AI 应用已相对成熟的皮肤常见病相比, 由于高质量、标准化的患者数据稀缺, 以及皮肤罕见病及具有皮肤表现的系统性罕见病本身在病因、发病机制、临床表型等方面的复杂性, 皮肤罕见病相关的 AI 研究进展相对缓慢, 提示在国内和国际开展皮肤罕见病多中心数据库建设工作非常必要。

除受到训练数据方面的限制以外, 医学 AI 的研究和临床应用目前还存在其他许多局限性和所面临的挑战, 如 AI 在医学应用中的法律监管及隐私保护策略尚待进一步完善, 亦有很多医学伦理问题尚待各界讨论并达成共识。此外, 目前医学 AI 所达到的高准确性大多基于各自的内部实验数据, 如进行外部验证, 准确性多会有一定程度的下滑, 同时, 囿于多数 AI 固有的可解释性困境, 在实际应用中也充斥着对 AI 可信度的质疑, 因此, 也更加需要强调临床医师对 AI 输出的结果必须结合实际情况理性判断, 而计算机算法也需要进一步优化改良。

2021 年 5 月, 由中国医学科学院北京协和医院皮肤科作为牵头单位, 中国罕见病联盟/北京罕见病诊疗与保障学会皮肤罕见病专业委员会成立, 积极推动着全国皮肤罕见病的诊疗体系建设和研究交流, 为国内的皮肤罕见病患者带来了新的希望, 也为中国皮肤罕见病 AI 研究及相关临床实践提供了有力的学术平台支持。相信在医学界相关组织和社会各界的共同努力下, AI 能够通过辅助基因诊断、赋能远程医疗、助力新药研发和药物重定位探索等方式为皮肤罕见病患者的诊疗现状带来新的突破。

**作者贡献:** 刘洁负责文章选题、审稿和指导; 王煜坤负责文献检索、文章撰写和修订。

**利益冲突:** 所有作者均声明不存在利益冲突。

## 参 考 文 献

[1] 王泽钊, 宋晓琳, 张金子, 等. 国外罕见病治疗保障体系对我国的伦理启示 [J]. 中国医学伦理学, 2022, 35 (10): 1088-1093.

[2] 朱以诚, 张抒扬. 我国罕见病诊疗和研究平台建立现状和回顾 [J]. 罕见病研究, 2022, 1 (2): 93-96.

[3] 张学军. 罕见性遗传性皮肤病的研究现状及展望 [J]. 皮肤科学通报, 2020, 37 (1): 1-4.

[4] Rundle CW, Hollingsworth P, Dellavalle RP. Artificial intelligence in dermatology [J]. Clin Dermatol, 2021, 39 (4): 657-666.

[5] 王慧, 戚倩倩, 李雪, 等. 皮肤肿瘤图像自动分类的研究进展 [J]. 计算机工程与应用, 2022, 58 (16): 31-48.

[6] Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks [J]. Nature, 2017, 542 (7639): 115-118.

[7] Haenssle HA, Fink C, Schneiderbauer R, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists [J]. Ann Oncol, 2018, 29 (8): 1836-1842.

[8] Marchetti MA, Liopyris K, Dusza SW, et al. Computer algorithms show potential for improving dermatologists' accuracy to diagnose cutaneous melanoma: results of the international skin imaging collaboration 2017 [J]. J Am Acad Dermatol, 2020, 82 (3): 622-627.

[9] Lee S, Chu YS, Yoo SK, et al. Augmented decision-making for acral lentiginous melanoma detection using deep convolutional neural networks [J]. J Eur Acad Dermatol Venereol, 2020, 34 (8): 1842-1850.

[10] Brinker TJ, Schmitt M, Kriehoff-Henning EI, et al. Diagnostic performance of artificial intelligence for histologic melanoma recognition compared to 18 international expert pathologists [J]. J Am Acad Dermatol, 2022, 86 (3): 640-642.

[11] Lodha S, Saggari S, Celebi JT, et al. Discordance in the histopathologic diagnosis of difficult melanocytic neoplasms in the clinical setting [J]. J Cutan Pathol, 2008, 35 (4): 349-352.

[12] Comes MC, Fucci L, Mele F, et al. A deep learning model based on whole slide images to predict disease-free survival in cutaneous melanoma patients [J]. Sci Rep, 2022, 12 (1): 20366.

[13] 刘兆睿, 张漪澜, 谢凤英, 等. 基于皮肤镜图像智能分析的早期蕈样肉芽肿诊断模型构建 [J]. 协和医学杂志, 2021, 12 (5): 689-697.

[14] Wu H, Chen H, Wang X, et al. Development and validation of an artificial intelligence-based image classification method for pathological diagnosis in patients with extramammary Paget's disease [J]. Front Oncol, 2021, 11: 810909.

[15] Aijaz SF, Khan SJ, Azim F, et al. Deep learning application for effective classification of different types of psoriasis [J]. J Healthc Eng, 2022, 2022: 7541583.

- [16] 北京协和医院. 中国首个罕见皮肤病人工智能辅助识别工具上线 [J]. 首都食品与医药, 2022, 29 (16): 8.
- [17] Akay M, Du Y, Serhsen CL, et al. Deep learning classification of systemic sclerosis skin using the MobileNetV2 model [J]. IEEE Open J Eng Med Biol, 2021, 2: 104-110.
- [18] Shi C, Meijer JM, Azzopardi G, et al. Use of convolutional neural networks for the detection of u-serrated patterns in direct immunofluorescence images to facilitate the diagnosis of epidermolysis bullosa acquisita [J]. Am J Pathol, 2021, 191 (9): 1520-1525.
- [19] 刘杏, 杨寅, 葛一平, 等. 人工智能在临床基因组学中的应用进展 [J]. 中国医学科学院学报, 2021, 43 (6): 950-955.
- [20] Alharbi WS, Rashid M. A review of deep learning applications in human genomics using next-generation sequencing data [J]. Hum Genomics, 2022, 16 (1): 26.
- [21] Poplin R, Chang PC, Alexander D, et al. A universal SNP and small-indel variant caller using deep neural networks [J]. Nat Biotechnol, 2018, 36 (10): 983-987.
- [22] Kumaran M, Subramanian U, Devarajan B. Performance assessment of variant calling pipelines using human whole exome sequencing and simulated data [J]. BMC Bioinformatics, 2019, 20 (1): 342.
- [23] Luo R, Sedlazeck FJ, Lam TW, et al. A multi-task convolutional deep neural network for variant calling in single molecule sequencing [J]. Nat Commun, 2019, 10 (1): 998.
- [24] Boudelloua I, Kulmanov M, Schofield PN, et al. Deep PVP: phenotype-based prioritization of causative variants using deep learning [J]. BMC Bioinformatics, 2019, 20 (1): 65.
- [25] Arloth J, Eraslan G, Andlauer TFM, et al. DeepWAS: multivariate genotype-phenotype associations by directly integrating regulatory information using deep learning [J]. PLoS Comput Biol, 2020, 16 (2): e1007616.
- [26] Jaganathan K, Kyriazopoulou PS, Merae JF, et al. Predicting splicing from primary sequence with deep learning [J]. Cell, 2019, 176 (3): 535-548.
- [27] Agarwal V, Shendure J. Predicting mRNA abundance directly from genomic sequence using deep convolutional neural networks [J]. Cell Rep, 2020, 31 (7): 107663.
- [28] Angermueller C, Lee HJ, Reik W, et al. DeepCpG: accurate prediction of single-cell DNA methylation states using deep learning [J]. Genome Biol, 2017, 18 (1): 67.
- [29] Yin Q, Wu M, Liu Q, et al. DeepHistone: a deep learning approach to predicting histone modifications [J]. BMC Genomics, 2019, 20 (Suppl 2): 193.
- [30] He J, Jia Y. Application of omics technologies in dermatological research and skin management [J]. J Cosmet Dermatol, 2022, 21 (2): 451-460.
- [31] Zhang H, Lee CAA, Li Z, et al. A multitask clustering approach for single-cell RNA-seq analysis in recessive dystrophic epidermolysis bullosa [J]. PLoS Comput Biol, 2018, 14 (4): e1006053.
- [32] Borcherding N, Voigt AP, Liu V, et al. Single-cell profiling of cutaneous T-cell lymphoma reveals underlying heterogeneity associated with disease progression [J]. Clin Cancer Res, 2019, 25 (10): 2996-3005.
- [33] Martínez BA, Shrotri S, Kingsmore KM, et al. Machine learning reveals distinct gene signature profiles in lesional and nonlesional regions of inflammatory skin diseases [J]. Sci Adv, 2022, 8 (17): eabn4776.
- [34] Escudíé JB, Jannot AS, Zapletal E, et al. Reviewing 741 patients records in two hours with FASTVISU [J]. AMIA Annu Symp Proc, 2015, 2015: 553-559.
- [35] Jamian L, Wheless L, Crofford LJ, et al. Rule-based and machine learning algorithms identify patients with systemic sclerosis accurately in the electronic health record [J]. Arthritis Res Ther, 2019, 21 (1): 305.
- [36] Li Y, Liang J, Xu X, et al. Clinicopathological features of fibrosarcomatous dermatofibrosarcoma protuberans and the construction of a back-propagation neural network recognition model [J]. Orphanet J Rare Dis, 2021, 16 (1): 48.
- [37] Cheraghlou S, Sadda P, Agogo GO, et al. A machine-learning modified CART algorithm informs Merkel cell carcinoma prognosis [J]. Australas J Dermatol, 2021, 62 (3): 323-330.
- [38] Álvarez-Machancoses Ó, Fernández-Martínez JL. Using artificial intelligence methods to speed up drug discovery [J]. Expert Opin Drug Discov, 2019, 14 (8): 769-777.
- [39] Challa AP, Zaleski NM, Jerome RN, et al. Human and machine intelligence together drive drug repurposing in rare diseases [J]. Front Genet, 2021, 12: 707836.
- [40] Lee YS, Krishnan A, Oughtred R, et al. A computational framework for genome-wide characterization of the human disease landscape [J]. Cell Syst, 2019, 8 (2): 152-162. e6.
- [41] Liu M, Yang F, Xu Y. Identification of potential drug therapy for dermatofibrosarcoma protuberans with bioinformatics and deep learning technology [J]. Curr Comput Aided Drug Des, 2022, 18 (5): 393-405.
- [42] Liu M, Xu Y. Gene Identification and potential drug therapy for drug-resistant melanoma with bioinformatics and deep learning technology [J]. Dis Markers, 2022, 2022: 2461055.

(收稿: 2022-12-27 录用: 2023-02-22)

(本文编辑: 郝恺雯)