

机器学习在糖尿病视网膜病变中的应用研究进展[▲]

李苏蓝 戴霞 吴妮晓 黄峥 徐子岚

(广西医科大学第一附属医院护理部,广西南宁市 530021)

【摘要】 糖尿病视网膜病变是导致糖尿病患者视力下降甚至失明的重要原因,因此,有效、准确地诊断糖尿病视网膜病变至关重要。机器学习作为人工智能的一个重要子领域,已经在疾病检测、生存预测和风险评估等方面展现出巨大潜力。本文就机器学习在糖尿病视网膜病变中的应用研究现状进行综述,旨在为更精准的医疗决策和综合个性化的慢性病服务管理提供参考。

【关键词】 机器学习;糖尿病视网膜病变;综述

【中图分类号】 R 587.2;R 774.1 **【文献标识码】** A **【文章编号】** 0253-4304(2025)08-1196-04

DOI: 10.11675/j.issn.0253-4304.2025.08.17

2021年全球糖尿病调查结果显示,全球约有5.37亿人患有糖尿病^[1]。在中国,糖尿病患病率从1980年的不足1%上升至2019年的11.6%^[2]。糖尿病视网膜病变(diabetic retinopathy, DR)作为糖尿病的严重并发症,是一种影响视力的神经血管疾病,是不可逆转失明的主要病因之一^[3]。在全球糖尿病患者中,DR预估患病率为22.27%,约60%的2型糖尿病患者会在发病后20年内并发DR^[4-5]。随着糖尿病患病率的增加,更多的患者面临视网膜病变的风险。DR造成的视力障碍和失明严重影响患者的生活质量,并增加其他并发症的发生风险,给个人、家庭、社会带来沉重的负担^[6]。及早的筛查、治疗对预防和控制DR的发生与发展具有重要意义。

机器学习是人工智能的一个重要子领域。当机器学习接触到复杂的大规模生物数据集时,可利用其在图像处理、图像分析和生物信息学分析等方面的优势,快速、精准地学习与识别这些数据中的特定模式,并通过各种数据挖掘算法进行数据分析和预测^[7]。机器学习算法主要分为监督学习、无监督学习、半监督学习和强化学习。其中,监督学习是通过分析带有标签的训练数据来推导出一个能够预测未知数据的函数;无监督学习是通过分析不带有标签的训练数据来发现数据中的模式或结构,从而对数据进行分类或聚类;半监督学习是监督学习和非监督学习的组合,既适用于标签数据,也适用于无标签

数据;强化学习是基于奖励和惩罚的机制,通过与环境的交互学习来优化行为,以达到最大化长期奖励的目标^[8]。随着计算能力呈指数级增长和海量数据处理技术的发展,机器学习已越来越广泛地应用于临床医学领域^[9]。近年来,机器学习在DR中的应用逐渐受到重视,已在DR的筛查、风险预测、分类等方面取得较大进步^[10]。本文对机器学习在DR中的应用研究现状进行综述,旨在为更精准的医疗决策和综合个性化的慢性病服务管理提供参考。

1 机器学习在DR风险预测中的应用

Tao等^[11]将深度学习(机器学习的一个分支)与正则化诺模图相融合,构建了一种新的深度学习诺模图,挖掘动态血糖监测数据与DR之间的非线性关系,并将患者的糖尿病病程和糖化血红蛋白水平整合进诺模图中,对患者DR发生风险进行评分,结果显示,这种新的深度学习诺模图可以通过动态血糖监测数据识别DR高风险患者,其诊断DR的曲线下面积(area under the curve, AUC)为0.86,特异度为0.81。卷积神经网络是一种深度学习模型,Wang等^[12]利用多分支时间卷积网络与张量数据补全模型,分析糖尿病患者的临床纵向数据(人口学资料、实验室检查、糖尿病并发症种类等)以预测DR的发生,发现该模型不仅可以处理数据集中常见的数据不平衡和缺

▲基金项目:国家自然科学基金(82160440);广西自然科学基金(2022JJA141061)

第一作者简介:李苏蓝,在读硕士研究生,研究方向为糖尿病教育与管理。

通信作者简介:戴霞,本科,主任护师,研究方向为糖尿病管理。

失问题,还能够捕捉临床纵向数据中变量之间的时间相关性和复杂的交互作用,该模型预测DR的AUC为0.949,相较于传统的卷积神经网络模型具有更好的预测性能。Islam等^[13]训练和优化人工神经网络、支持向量机、随机森林和极端梯度增强4个应用广泛的机器学习模型,将其用于预测DR,其中基于极端梯度增强的机器学习模型具有最高的预测性能,AUC为0.850,准确率为90.01%,精确率为91.80%。有学者利用包含118 868张多种族糖尿病患者眼底图像的数据集,开发并验证了一种深度学习模型进行特征提取,以预测糖尿病患者5年内DR进展的个体化时间,结果显示,该模型预测DR进展的一致性指数较高(0.754~0.846),综合Brier指数较低(0.153~0.241),表明该模型可用于糖尿病患者并发DR的风险评估,促进糖尿病患者个性化DR管理^[14]。

2 机器学习在DR诊断中的应用

Shoaib等^[15]利用卷积神经网络模型识别不同眼部疾病数据集(共包含8种眼部疾病)中的DR病例,结果显示该模型对DR的诊断准确率达到98.3%。Xu等^[16]提出一种多视图联合学习的DR诊断模型,通过对全眼底图像特征和血管图像的局部细节特征进行整合,可以减少单眼底图像学习的局限性,其诊断DR的AUC为0.98,该模型出色的分类能力可以有效帮助早期发现和治疗DR病变,为医生提供更可靠的诊断依据。Bilal等^[17]使用卷积神经网络结合改进支持向量机的创新模型对威胁视力的糖尿病视网膜病变(vision-threatening diabetic retinopathy, VTDR)进行检测,该模型通过分析单个像素的细节、空间关系,优化了视网膜图像处理,增强了模型的稳健性和准确性,其检测VTDR的准确率为99.18%、敏感度为98.15%、特异度为100%。Khan等^[18]使用灰狼优化-极限学习机算法与全卷积编码器-解码器网络相结合的新模型,对眼底图像进行更精确的图像分割和疾病分类,并基于准确度、召回率、精确度、特异度、F度量、G均值和马修斯相关系数等评估指标来评估检测DR的性能,结果表明该模型诊断DR的准确度达到96.21%。

3 机器学习在DR分类中的应用

研究表明,生成式对抗网络-卷积神经网络模型可以对异常曝光的图像进行预处理,修复图像的异常曝光,从而更好地显示眼底图像的细节信息,即使

在数据质量不佳的情况下,分类网络仍然可以产生准确的分类结果,将DR按照疾病严重程度分为转诊授权的糖尿病视网膜病变(referral-warranted diabetic retinopathy, RWDR)和非转诊的糖尿病视网膜病变(non-referral diabetic retinopathy, NRDR),该模型辨别NRDR和RWDR的准确率分别达到94.5%和94.4%,能够较好地解决社区筛查中图像质量差导致DR分类准确率不高的问题^[19]。Roberto-Oraó等^[20]采用基于注意力机制的卷积神经网络模型分离眼底图像的亮结构和暗结构,从图像上区分红色病变(微动脉瘤和出血等)和明亮病变(硬渗出物等)的病理区域,并自动进行DR分级(5个严重程度),结果显示该模型对DR进行分级的准确率达到83.7%,可作为DR筛查和分级的有效方法。Chilukoti等^[21]使用转移学习和二次加权Kappa综合学习的集合方法对DR进行分类,采用对比度受限自适应直方图均衡和高斯模糊等数据增强技术进行图像增强和降噪,并开发了一个三层分类器(前两层从图像中提取特征,最后一层对DR等级进行分类),以减少过拟合并提高泛化能力,结果显示该模型对3个数据库进行DR分类均可获得较高的Quadratic Weighted Kappa分数(该评分能反映实际分类与模型预测分类之间的匹配程度),分别为0.901、0.967和0.944。Riazi-Esfahani等^[22]使用支持向量回归机器学习方法来识别视网膜层的边界,通过量化分析视网膜内外丛状层平滑指数和面积等生物标记物,在鉴别非增殖性糖尿病视网膜病变(non-proliferative diabetic retinopathy, NPDR)和增殖性糖尿病视网膜病变(proliferative diabetic retinopathy, PDR)方面展现出良好的效果,尤其在分析水平扫描中央凹区内核层面积、水平扫描鼻区内丛状层平滑指数、水平扫描鼻区外丛状层顶部边界平滑指数及水平中央凹颞区内丛状层平滑指数时,该方法区分NPDR和PDR的精确度可达87.6%、97.2%、84.1%和89.8%。

4 机器学习在发掘DR相关标志物中的应用

Wang等^[23]通过分析PDR患者和健康人群的单细胞RNA测序数据,鉴定出的氧化应激相关基因,然后基于氧化应激相关基因通过12种机器学习算法构建PDR预测模型,构建的模型预测PDR的AUC为0.989,具有良好的预测效能和应用前景。铁下垂在DR的发病机制中起着重要作用。Liu等^[24]通过生物信息学方法筛选铁下垂相关中枢基因,利用最小绝对收缩和选择算子回归与随机森林算法进一步筛选及确

定5个在诊断DR中表现良好的铁下垂相关中枢基因,该研究为铁下垂的发病机制提供新的见解,亦为探索DR的新诊断标志物和治疗策略提供理论基础。Cao等^[25]使用人工神经网络、Logistic回归、随机森林、支持向量机和极端梯度增强5种机器学习算法,分析血浆细胞因子的变化以预测NPDR,结果显示随机森林机器学习算法的预测性能最好,其灵敏度为92.3%,特异度为75%,血浆血管生成素1、血小板衍生生长因子BB和血管内皮生长因子受体2与NPDR的发生有关,可作为NPDR筛查的新诊断标志物,纵向跟踪这些标志物的变化,可为医生及时根据疾病进展制订或改良治疗方案。Yagin等^[26]利用可解释人工智能模型、极端梯度增强、自然梯度增强概率预测和解释性基准模型对DR进行分类,发现解释性基准模型与极端梯度增强结合的方法对DR的预测效果较好,准确率、精确度分别为91.25%、89.33%,这种联合方法同时具有极端梯度增强的预测能力和解释性基准模型的可解释性,在模型准确性和可解释性之间实现平衡,因此该方法是发掘DR生物标记物和预测DR病程的有效方法之一。此外,该研究还确定了DR分类的6个重要生物标志物,分别是色氨酸、磷脂酰胆碱C42:2、丁酰基肉碱、酪氨酸、十六酰基肉碱和总二甲基精氨酸^[26]。这些生物标志物不仅可以帮助临床医生更有针对性地评估DR的严重程度,还有助于优化治疗措施。

5 机器学习在DR并发症风险预测中的应用

曹长玲等^[27]分别使用LightGBM模型与Logistic回归模型建立DR合并冠心病、肾病、下肢动脉病变的风险预测模型,结果显示LightGBM模型预测上述并发症发生的AUC分别为0.92、0.91、0.81,相比Logistic回归模型表现出更佳的性能。Betzler等^[28]开发了一个针对DR患者视网膜图像进行糖尿病肾病预测的深度学习算法模型,该模型预测糖尿病肾病的AUC为0.826,可作为糖尿病肾病筛查的一种非侵入性、低成本工具。Mellor等^[29]通过对DR患者视网膜照片进行深度学习,获取深度学习评分,以预测心血管疾病的发病风险和危险因素,结果显示,与心血管疾病关联最强的危险因素是糖化血红蛋白和糖尿病病程,深度学习评分与心血管事件独立相关。

6 小 结

机器学习的发展为DR的早期筛查和诊断提供新的可能,具有节省时间、提高诊疗效率及减轻DR危害的潜力。未来应将机器学习与临床实践进行整合创新,使其简化为临床可用的应用程序或设备,用于临床支持决策系统、可穿戴设备和糖尿病人群远程管理程序等领域,以提高医疗服务的质量和效率,为患者提供更加精准和个性化的治疗方案^[30-31],从而推动糖尿病管理服务迈向新的高度。

参 考 文 献

- [1] Sun H, Saeedi P, Karuranga S, et al. Erratum to "IDF diabetes atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045" [Diabetes Res. Clin. Pract. 183 (2022) 109119] [J]. Diabetes Res Clin Pract, 2023, 204: 110945.
- [2] 周伍明, 崔璇璇, 陆宗庆, 等. 1990-2019年中国2型糖尿病疾病负担分析 [J]. 中华疾病控制杂志, 2023, 27(6): 650-654, 690.
- [3] Cheung N, Mitchell P, Wong TY. Diabetic retinopathy [J]. Lancet, 2010, 376(9735): 124-136.
- [4] 刘雪立. 全球糖尿病视网膜病变患病率和2045年负担预测——美国眼科学会报告要点摘编 [J]. 眼科新进展, 2024, 44(1): 80-81.
- [5] Cho NH, Shaw JE, Karuranga S, et al. IDF diabetes atlas: global estimates of diabetes prevalence for 2017 and projections for 2045 [J]. Diabetes Res Clin Pract, 2018, 138: 271-281.
- [6] Song PG, Yu JY, Chan KY, et al. Prevalence, risk factors and burden of diabetic retinopathy in China: a systematic review and meta-analysis [J]. J Glob Health, 2018, 8(1): 010803.
- [7] Greener JG, Kandathil SM, Moffat L, et al. A guide to machine learning for biologists [J]. Nat Rev Mol Cell Biol, 2022, 23(1): 40-55.
- [8] Shaye I, Dushi P, Banafaa M, et al. Handover management for drones in future mobile networks—a survey [J]. Sensors (Basel), 2022, 22(17): 6424.
- [9] Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023 [J]. N Engl J Med, 2023, 388(13): 1201-1208.

- [10] Nadeem MW, Goh HG, Hussain M, et al. Deep learning for diabetic retinopathy analysis: a review, research challenges, and future directions[J]. *Sensors (Basel)*, 2022, 22(18):6780.
- [11] Tao R, Yu X, Lu J, et al. A deep learning nomogram of continuous glucose monitoring data for the risk prediction of diabetic retinopathy in type 2 diabetes[J]. *Phys Eng Sci Med*, 2023, 46(2):813-825.
- [12] Wang ZK, Chen SH, Liu TM, et al. Multi-branching temporal convolutional network with tensor data completion for diabetic retinopathy prediction[J]. *IEEE J Biomed Health Inform*, 2024, 28(3):1704-1715.
- [13] Islam MM, Rahman MJ, Rabby MS, et al. Predicting the risk of diabetic retinopathy using explainable machine learning algorithms[J]. *Diabetes Metab Syndr*, 2023, 17(12):102919.
- [14] Dai L, Sheng B, Chen TL, et al. A deep learning system for predicting time to progression of diabetic retinopathy[J]. *Nat Med*, 2024, 30(2):584-594.
- [15] Shoaib MR, Emara HM, Zhao J, et al. Deep learning innovations in diagnosing diabetic retinopathy: the potential of transfer learning and the DiaCNN model[J]. *Comput Biol Med*, 2024, 169:107834.
- [16] Xu X, Liu D, Huang G, et al. Computer aided diagnosis of diabetic retinopathy based on multi-view joint learning[J]. *Comput Biol Med*, 2024, 174:108428.
- [17] Bilal A, Imran A, Baig TI, et al. Improved support vector machine based on CNN-SVD for vision-threatening diabetic retinopathy detection and classification[J]. *PLoS One*, 2024, 19(1):e0295951.
- [18] Khan AQ, Sun GM, Khalid M, et al. A novel fusion of genetic grey wolf optimization and kernel extreme learning machines for precise diabetic eye disease classification[J]. *PLoS One*, 2024, 19(5):e0303094.
- [19] Yang K, Lu YF, Xue LY, et al. URNet: system for recommending referrals for community screening of diabetic retinopathy based on deep learning[J]. *Exp Biol Med (Maywood)*, 2023, 248(11):909-921.
- [20] Romero-Oraú R, Herrero-Tudela M, López MI, et al. Attention-based deep learning framework for automatic fundus image processing to aid in diabetic retinopathy grading[J]. *Comput Methods Programs Biomed*, 2024, 249:108160.
- [21] Chilukoti SV, Shan LQ, Tida VS, et al. A reliable diabetic retinopathy grading *via* transfer learning and ensemble learning with quadratic weighted kappa metric[J]. *BMC Med Inform Decis Mak*, 2024, 24(1):37.
- [22] Riazi-Esfahani H, Jafari B, Azimi H, et al. Assessment of area and structural irregularity of retinal layers in diabetic retinopathy using machine learning and image processing techniques[J]. *Sci Rep*, 2024, 14(1):4013.
- [23] Wang J, Sun H, Mou L, et al. Unveiling the molecular complexity of proliferative diabetic retinopathy through scRNA-seq, AlphaFold 2, and machine learning[J]. *Front Endocrinol (Lausanne)*, 2024, 15:1382896.
- [24] Liu J, Li X, Cheng Y, et al. Identification of potential ferroptosis-related biomarkers and a pharmacological compound in diabetic retinopathy based on machine learning and molecular docking[J]. *Front Endocrinol (Lausanne)*, 2022, 13:988506.
- [25] Cao B, Zhang N, Zhang YY, et al. Plasma cytokines for predicting diabetic retinopathy among type 2 diabetic patients *via* machine learning algorithms[J]. *Aging (Albany NY)*, 2020, 13(2):1972-1988.
- [26] Yagin FH, Yasar S, Gormez Y, et al. Explainable artificial intelligence paves the way in precision diagnostics and biomarker discovery for the subclass of diabetic retinopathy in type 2 diabetics[J]. *Metabolites*, 2023, 13(12):1204.
- [27] 曹长玲, 翁郁华, 李晓琼, 等. 基于机器学习的糖尿病视网膜病变合并症风险预测模型[J]. *生命科学仪器*, 2023, 21(2):56-63.
- [28] Betzler BK, Chee EYL, He F, et al. Deep learning algorithms to detect diabetic kidney disease from retinal photographs in multiethnic populations with diabetes[J]. *J Am Med Inform Assoc*, 2023, 30(12):1904-1914.
- [29] Mellor J, Jiang W, Fleming A, et al. Can deep learning on retinal images augment known risk factors for cardiovascular disease prediction in diabetes? A prospective cohort study from the national screening programme in Scotland[J]. *Int J Med Inform*, 2023, 175:105072.
- [30] Contreras I, Vehi J. Artificial intelligence for diabetes management and decision support: literature review[J]. *J Med Internet Res*, 2018, 20(5):e10775.
- [31] 闫朝霞, 陈小平, 程团结, 等. 基于5G网络的复合型糖尿病AI管理平台的设计与应用[J]. *护理学杂志*, 2024, 39(5):6-9.

(收稿日期:2025-02-10 修回日期:2025-04-12)