

文章编号: 1001-6325(2020)05-0707-04

短篇综述

人工智能及机器学习在心血管疾病中的应用

宋绮蕊, 蔡 军*

(中国医学科学院 北京协和医学院 国家心血管病中心 阜外医院 高血压中心, 北京 100037)

摘要: 人工智能(AI)及机器学习(ML)因其独特的优势逐渐在医学领域得到了较为广泛的应用。在心血管疾病中,该技术在处理电子病历记录中繁杂的数据,预测分析疾病发展及预后,自动分析和识别心血管影像学及心律失常,发现疾病新亚型等方面已经取得了一定进展。AI及ML在心血管疾病研究中潜力巨大,将会为心血管领域带来全新的突破。

关键词: 人工智能;机器学习;心血管疾病

中图分类号:R714.252 文献标志码:A

Application of artificial intelligence and machine learning in cardiovascular diseases

SONG Qi-ru, CAI Jun*

(Hypertension Center, Fuwai Hospital, National Center for Cardiovascular Diseases, CAMS & PUMC, Beijing 100037, China)

Abstract: Artificial intelligence(AI) and machine learning(ML) have gradually been widely used in the medical field due to their unique advantages. In cardiovascular diseases, this technology has made some progress for processing complex data from electronic medical record, analyzing disease mechanisms and for predicting prognosis, automatically analyzing and identifying cardiovascular imaging and arrhythmia, and discovering new subtypes of disease. The potential of AI and ML in cardiovascular disease research will bring a new breakthrough in the cardiovascular field.

Key words: artificial intelligence; machine learning; cardiovascular disease

人工智能(artificial intelligence, AI)是一个广泛的术语,指的是能够模拟人的某些思维过程和智能行为的技术或系统。机器学习(machine learning, ML)是AI的子集,它是指机器独立学习并做出准确预测的能力。近年来,AI技术发展迅速,开始渗透到医学及心血管医学领域^[1]。过去10年,关于ML和深度学习(deep learning, DL)应用于心血管领域的文章显著增多^[2-3],这两种方法被证明有助于包

括超声心动图的解释、舒张性心功能衰竭的分级等一系列复杂领域的分析^[4-5]。心血管疾病是全世界非传染病死亡原因的首位^[6],给患者及患者家庭造成了巨大的负担和压力。尽管将其应用于临床实践仍面临许多障碍和挑战,AI技术毋庸置疑将会引领心血管医疗实践走向更加个体化、更加精准的方向^[7]。本篇综述目的在于阐述AI及ML在心血管疾病研究中的应用进展,探讨其应用于医学领域的

收稿日期:2020-02-13 修回日期:2020-03-19

基金项目:国家自然科学基金(81630014,81825002);国家重点研究计划(2018YFC1312703);北京市优秀青年科学家计划(BJJWZYJH01201910023029);中国医学科学院创新基金(2016-12M1-006)

*通信作者(corresponding author):caijun7879@126.com

局限性以及未来的展望。

1 AI 及 ML 概述

AI 指的是可模拟和模仿人类智能的计算程序, ML 是应用较广泛的 AI 技术之一。在 ML 中, 机器从数据中学习并基于所学习的模型执行任务, 在处理大数据方面表现良好。按照对数据的学习方式, ML 主要分为监督学习和无监督学习。监督学习主要用于解决分类和回归问题, 分类的目标是正确分配二进制或多类标签, 回归的目标是正确预测结果。无监督学习主要用于发现数据中的隐藏结构, 探索变量之间的关系, 主要包括聚类和降维, 无监督学习在心血管疾病中最有前途的用途之一是发现疾病新亚型, 将疾病分成更精准的、个体化的亚型^[7-8]。DL 是 ML 的子领域, 它模仿人脑的工作模式, 使用多层神经网络对数据集进行训练并形成自动预测, 在传统的 ML 中, 算法运行需要分析人员一定程度的判断, 如特征的选择, 而 DL 更加的具有自我驱动性。DL 是一种强大的方法, 主要包括卷积神经网络 (convolutional neural network, CNN) 和递归神经网络 (recursive neural network, RNN)^[9]。DL 相比于其他 ML 算法, 在分析图像等复杂数据时表现良好, 主要用于图像识别。

2 AI 及 ML 在心血管疾病研究中的应用

心血管疾病研究相关数据如电子病历记录、基因组学数据、心脏影像等数量庞大具有异质性, 且变化迅速, 分析和应用具有较大的挑战性。传统的统计学方法不能有效的对其进行分析处理, 导致大量数据未得到利用, AI 具有从海量数据中学习并获取信息的能力, 应用 AI 及 ML 分析心血管研究相关数据将会为临床心血管病诊治带来巨大收益。

2.1 处理电子病历记录中的数据

众所周知, 电子病历记录存储着大量临床数据, 有很大的分析潜力, 但电子病历中的数据通常数量大且形式复杂, 结构不完善, 利用传统的分析方式处理数据受到限制, 且利用电子病历记录进行临床研究时, 研究人员须阅读文本并手动提取有用信息, 给研究者带来负担。在一项关于预测心力衰竭事件发生的巢式病历对照研究中^[10], 运用了 DL 算法对来自一个超过 60 万患者的电子健康记录库中的

216 394 例患者进行分析, 并准确预测了心力衰竭的发作, 证明了 ML 算法可有效处理和分析电子病历记录中庞大且复杂的数据。另外一项研究运用 10 019 例成人先天性心脏病患者的电子病历数据对 DL 算法进行训练^[11], 经训练和验证后的算法可自动从测试数据集中提取诊断及判断预后。由此可见, 经过训练的 ML 模型, 可以自动对电子病历记录中的数据进行提取和分析。利用 ML 算法处理电子病历记录中的数据, 将会使数据得到充分利用, 并节约研究者的时间成本。

2.2 疾病的预测分析

预测分析在心血管疾病临床研究中占重要地位。如高血压、冠状动脉粥样硬化性心脏病等慢性疾病, 对疾病发展和预后进行预测并加以预防, 可有效减少并发症及猝死事件的发生。在心脏外科手术中, 预测手术死亡率及住院时间等同样意义重大。目前较多研究表明, 相对于传统的风险预测模型, ML 模型的预测性能更好。一项通过心肌灌注成像预测严重不良心脏事件的研究表明^[12], DL 算法能较为准确地预测不良事件的发生, 且相比传统的预测模型, 预测准确度更高。另一项研究利用 ML 算法通过冠状动脉 CT 造影 (coronary computerized tomography angiography, CCTA) 对疑似冠脉疾病患者的全因死亡率进行预测^[13], 准确性优于单纯的弗雷明汉风险评分 (Framingham Risk Score, FRS) 及 CCTA 严重性评分。关于心脏外科手术的一项研究也表明^[14], ML 算法在预测开放性心脏手术的院内死亡率方面优于传统预测模型。疾病的发展及预后的预测是 AI 及 ML 应用较广泛的领域, ML 相对于传统的预测评分及预测模型精确度更高, 将革新心血管疾病的预测方式并提高预测准确度。

2.3 自动分析心血管影像

心血管影像包括 CCTA、心脏核磁、超声心动图等, 通常需要影像专业医生对图像进行手动测量分析, 需耗费大量人力及时间, 且主观性较强, 利用 ML 自动识别和分析心血管影像, 可节约成本, 避免主观干扰。DL 可用于自动分割心脏核磁图像^[15], 且其在短轴图像上分割左、右心室及在长轴图像中分割左、右心房的准确度与专业人员相似。一个利用 CNN 建立的 DL 模型^[16], 经过数据训练后, 可以准确识别超声心动图, 自动进行图像分割及测量、计

算射血分数,并可较为准确的识别出肥厚性心肌病,心脏淀粉样变性和肺动脉高压。除可自动测量分析图像外,利用 DL 算法的强大功能可避免患者做不必要或侵入性的检查,DL 算法可自动识别 CCTA 图像中局部心肌的异常^[17],从而推断出支配该区冠状动脉的狭窄情况,其诊断冠脉狭窄的能力不差于侵入性血流储备分数检查,证明该方法可能取代侵入性检查用于判断冠脉功能性狭窄。DL 算法因其强大的图像识别和分析能力,在心血管影像分析领域前景广阔。

2.4 自动识别心律失常

心律失常是心血管疾病中常见的一组疾病,通常需要常规心电图或 24 h Holter 等检查捕捉心律失常的异常波形。如心房颤动、室性心动过速、房室传导阻滞等心律失常,常常反复发作,且存在发生脑梗死、猝死等严重并发症的风险,许多患者无明显症状或症状轻微,因此心律失常易被忽略。ML 模型可以精确的识别心电图中的不同波形(QRS 波、T 波和 P 波),利用这些信息,可以计算得出心率、轴偏差、间歇长度这些重要的临床参数,从而识别出常见心律失常。目前关于自动识别心房颤动的研究及应用较多,一种基于 CNN 开发的人工智能心电图仪^[18],可在正常窦性心律中自动识别出新发房颤,总体准确性达 83.3%。人工智能心电图仪的出现对房颤患者的识别和管理意义重大,患者可在社区进行检测,早期识别房颤并及时处理。此外,利用 DL 开发的智能手表可对患者进行动态心电图监测^[19],在接受心脏复律的患者中发现心房颤动的敏感性为 98.0%,特异性为 90.2%。虽然该研究样本量较小,但由于智能手表的监测方式更为方便,更具时效性,相信未来的发展中,利用智能手表监测心律失常将得到更多的关注。现有的研究证明 ML 在自动识别心律失常尤其是心房颤动具有较高的准确性,基于 AI 开发的心电图仪及智能手表会为心律失常患者的诊断及管理带来重大意义的突破。

2.5 发现疾病新亚型

心血管疾病本身就是复杂且具有异质性的,病因涉及到遗传、环境、行为等多重因素,原发性高血压、心力衰竭等病因并不明确,治疗困难。聚类分析可利用疾病相关变量对患病人群进行无监督的聚类分析,将其聚类成具有不同特点的亚组并比较各个

亚组患者的预后,更为精确进行疾病分类及诊断,为患者提供更具针对性的防治措施。一项关于射血分数保留心力衰竭(heart failure with preserved ejection fraction, HFpEF)的研究利用聚类分析将 397 例 HFpEF 患者聚类成 3 个临床特点、心脏功能及疾病预后明显不同的亚组^[20]。此外,一项关于 SPRINT 试验(systolic blood pressure intervention trial)的探索性研究利用聚类分析将 SPRINT 试验研究的高血压患者分为了 4 个具有不同表型特点的聚类^[21],各个聚类的心血管预后以及对强化降压的反应不同。以上研究证明利用数据驱动的聚类分析确定疾病新亚型具有可行性,具有不同临床特点及预后的亚型可协助临床医生制定更加个体化的防治方案。

3 问题与展望

AI 及 ML 在心血管疾病中的应用具有巨大的潜力,但具有一定的局限性。在某些特定情况下,AI 不能很好的发挥作用,反而会产生误导性结果,其中一个主要缺陷是容易过拟合数据。AI 用于大数据分析时,具有多项参数的复杂模型容易导致数据过拟合,过拟合的模型在训练数据集中表现良好,但由于该模型对样本数据中的随机噪声同时进行了建模,这使其无法捕获数据中的真实关联,因此可能无法推广到新的数据集。数据质量也是一个重要影响因素,不恰当的数据选择以及不正确的测量可能会导致错误的结果和预测,当数据过于嘈杂或者缺乏重要变量时,将不能应用于 AI 分析。除了技术自身所存在的局限性,将 AI 付诸于医学实践也面临一定的挑战。首先,由于医学领域的特殊性,数据涉及患者隐私,数据的获取会受到一定的限制,因此科研机构需要建立具有完备的个人隐私保护政策的数据库。此外,数据的存储也存在一定的限制,数据往往存储在多个不同服务器中,此时即使 AI 技术能够建立理想的模型,但用于分析的大量参数分散在各个不同系统中,使分析无法进行。高质量的数据以及适当的存储方式对 AI 技术的开展至关重要,发展和建立无缝数据库势在必行。

AI 及 ML 在心血管疾病的应用前景广阔。最近的一份分析报告显示,截至 2025 年,AI 在医疗保健市场中预算将从 21 亿美元增长到 361 亿美元,其在心血管领域的应用也会逐渐增多。随着大数据的

蓬勃发展以及计算机技术的进步,心血管疾病的研究模式将逐渐转变。随着技术的发展,AI及ML会

为心血管领域带来全新的突破,为心血管疾病的防控注入新的力量。

参考文献:

- [1] Johnson KW, Torres Soto J, Glicksberg BS, *et al.* Artificial intelligence in cardiology[J]. *J Am Coll Cardiol*, 2018, 71: 2668-2679.
- [2] Shameer K, Johnson KW, Glicksberg BS, *et al.* Machine learning in cardiovascular medicine: are we there yet? [J]. *Heart*, 2018, 104: 1156-1164.
- [3] Sengupta PP, Shrestha S. Machine learning for data-driven discovery: the rise and relevance [J]. *JACC Cardiovasc Imaging*, 2019, 12: 690-692.
- [4] Lancaster MC, Salem Omar AM, Narula S, *et al.* Phenotypic clustering of left ventricular diastolic function parameters: patterns and prognostic relevance[J]. *JACC Cardiovasc Imaging*, 2019, 12: 1149-1161.
- [5] Gandhi S, Mosleh W, Shen J, *et al.* Automation, machine learning, and artificial intelligence in echocardiography: A brave new world [J]. *Echocardiography*, 2018, 35: 1402-1418.
- [6] Balakumar P, Maung UK, Jagadeesh G. Prevalence and prevention of cardiovascular disease and diabetes mellitus [J]. *Pharmacol Res*, 2016, 113: 600-609.
- [7] Johnson KW, Shameer K, Glicksberg BS, *et al.* Enabling precision cardiology through multiscale biology and systems medicine[J]. *JACC Basic Transl Sci*, 2017, 2: 311-327.
- [8] Antman EM, Loscalzo J. Precision medicine in cardiology [J]. *Nat Rev Cardiol*, 2016, 13: 591-602.
- [9] LeCun YBY, Hinton G. Deep learning [J]. *Nature*, 2015, 521: 436-444.
- [10] Mallya SOM, Srivastava N, Arai T, *et al.* Effectiveness of lstms in predicting congestive heart failure onset [EB/OL]. Available at: <https://arxiv.org/abs/1902.02443>. Accessed August 11, 2019.
- [11] Diller GP, Kempny A, Babu-Narayan SV, *et al.* Machine learning algorithms estimating prognosis and guiding therapy in adult congenital heart disease: data from a single tertiary centre including 10 019 patients [J]. *Eur Heart J*, 2019, 40: 1069-1077.
- [12] Betancur J, Commandeur F, Motlagh M, *et al.* Deep learning for prediction of obstructive disease from fast myocardial perfusion SPECT: a multicenter study [J]. *JACC Cardiovasc Imaging*, 2018, 11: 1654-1663.
- [13] Motwani M, Dey D, Berman DS, *et al.* Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis [J]. *Eur Heart J*, 2017, 38: 500-507.
- [14] Allyn J, Allou N, Augustin P, *et al.* A comparison of a machine learning model with euroSCORE II in predicting mortality after elective cardiac surgery: a decision curve analysis [J]. *PLoS One*, 2017, 12. doi: 10.1371/journal.pone.0169772.
- [15] Bai W, Sinclair M, Tarroni G, *et al.* Automated cardiovascular magnetic resonance image analysis with fully convolutional networks [J]. *J Cardiovasc Magn Reson*, 2018, 20: 65.
- [16] Zhang J, Gajjala S, Agrawal P, *et al.* Fully automated echocardiogram interpretation in clinical practice [J]. *Circulation*, 2018, 138: 1623-1635.
- [17] Zreik M, Lessmann N, van Hamersvelt RW, *et al.* Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis [J]. *Med Image Anal*, 2018, 44: 72-85.
- [18] Attia ZI, Noseworthy PA, Lopez-Jimenez F, *et al.* An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction [J]. *Lancet*, 2019, 394: 861-867.
- [19] Tison GH, Sanchez JM, Ballinger B, *et al.* Passive detection of atrial fibrillation using a commercially available smartwatch [J]. *JAMA Cardiol*, 2018, 3: 409-416.
- [20] Shah SJ, Katz DH, Selvaraj S, *et al.* Phenomapping for novel classification of heart failure with preserved ejection fraction [J]. *Circulation*, 2015, 131: 269-279.
- [21] Yang DY, Nie ZQ, Liao LZ, *et al.* Phenomapping of subgroups in hypertensive patients using unsupervised data-driven cluster analysis: An exploratory study of the SPRINT trial [J]. *Eur J Prev Cardiol*, 2019, 26: 1693-1706.