

机器学习在脑卒中诊断与治疗中的应用进展

高小波 舒畅 李祥飞 骆达 葛乐

【摘要】 随着人工智能技术的迅速发展,机器学习算法广泛应用于脑卒中诊断、治疗、预后各阶段,并表现出巨大的应用潜力。本文从机器学习算法研究,机器学习算法在脑卒中诊断与分类诊断、结局预测、风险因素识别等方面总结机器学习算法在脑卒中诊断与治疗中的应用现状并展望未来发展方向。

【关键词】 卒中; 机器学习; 综述

Advances in machine learning for the diagnosis and treatment of stroke

GAO Xiao-bo¹, SHU Chang², LI Xiang-fei¹, LUO Da³, GE Le³

¹School of Economics and Management, Tiangong University, Tianjin 300387, China

²Tianjin Neurosurgical Institute, Tianjin Key Laboratory of Cerebral Vascular and Neurodegenerative Diseases, Tianjin Huanhu Hospital, Tianjin 300350, China

³Tianjin Huanhu Hospital, Tianjin 300350, China

GAO Xiao-bo and SHU Chang contributed equally to the article

Corresponding author: GE Le (Email: huanhu_data@126.com)

【Abstract】 With the rapid development of artificial intelligence (AI), machine learning (ML) algorithms has been widely used in stroke diagnosis, treatment and prognosis, and has shown great application potential. This paper summarizes the application status of ML algorithms in stroke research from the aspects of ML algorithms research, ML algorithms in stroke diagnosis and classification diagnosis, outcome prediction, risk factor identification and so on, and looks into the future development direction.

【Key words】 Stroke; Machine learning; Review

This study was supported by Tianjin Science and Technology Plan Project (No. 20JCZDJC00540).

Conflicts of interest: none declared

机器学习(ML)由 Samuel 于 1959 年首次提出,系一门涉及多领域的交叉学科,通过研究计算机模拟或实现人类学习行为,以获取新知识或新技能^[1],因其在大数据处理中的优势,近 10 年在医学领域迅速发展^[2]。脑卒中作为临床常见的脑血管病,具有高发病率、高病死率和高病残率,特别是病死率和病残率仅次于缺血性心脏病,位居第二^[3],严重危害人类健康。随着人工智能(AI)的迅速发展,机器学习算法应用于脑卒中研究已进入蓬勃发展阶段,涉

及脑卒中诊断、治疗、预后各阶段,表现出巨大的应用潜力,一方面,为诊断与治疗提供新的方法,具有较人工更高的敏感性^[4],显著提高脑卒中诊断与治疗的准确性^[5],如基于机器学习算法的图像处理可以帮助临床医师识别脑出血的特定亚型,以改善临床结局^[6];另一方面,还可对脑卒中严重程度和症状发作进行预测^[7-8],为临床决策的制定提供重要依据。本文拟综述机器学习算法研究、机器学习算法在脑卒中诊断与分类诊断、结局预测以及风险因素识别中应用进展,总结其应用现状并展望未来发展方向。

一、机器学习算法研究

脑卒中机器学习算法研究主要集中于深度学习(DL)、支持向量机(SVM)、人工神经网络(ANN)、随机森林(RF)、Logistic 回归(LR)等,并广泛应用于脑卒中诊断与分类诊断、治疗、预后各阶段。(1)诊断与分类诊断:深度学习是 Hinton 等^[9]于 2006 年研

doi: 10.3969/j.issn.1672-6731.2023.01.002

基金项目:天津市科技计划项目(项目编号:20JCZDJC00540)

作者单位:300387 天津工业大学经济与管理学院(高小波,李祥飞);300350 天津市环湖医院 天津市神经外科研究所 天津市脑血管与神经变性重点实验室(舒畅);300350 天津市环湖医院(骆达,葛乐)

高小波与舒畅对本文有同等贡献

通讯作者:葛乐,Email:huanhu_data@126.com

究人工神经网络时提出的一种机器学习算法,是经典人工神经网络的扩展,目前已成为应用最广泛的机器学习算法之一^[10]。相较于经典的人工神经网络,深度学习的最大优势在于可应用更多的隐藏层处理具有各种结构的复杂数据^[11],从而极大地促进机器学习算法的临床应用^[12],特别对于脑卒中,深度学习算法可以显著提高诊断与分类诊断的准确性^[13-14]。此外,其他机器学习算法在脑卒中诊断与分类诊断中的应用也日益增多,如支持向量机、随机森林、Logistic 回归、人工神经网络等,通过大规模数据提取进行脑卒中诊断,诊断准确性较高^[15]。Maier 等^[16]对比分析 9 种机器学习算法诊断缺血性卒中的准确性,结果显示,随机森林和卷积神经网络(CNN)的骰子度量(dice metric)值均为 0.67 ± 0.18 ,并认为这两种算法可以实现总体最佳分割结果,较其他算法具有更高的分类精度和诊断准确性。(2)治疗与预后:机器学习算法在脑卒中治疗与预后中的应用深度和广度也不断拓展。急性缺血性卒中静脉溶栓治疗“时间窗”是发病后 4.5 小时,但在临床实践中有很大一部分患者对缺血性卒中的确切发病时间不甚清楚而被排除在静脉溶栓“时间窗”外。基于此,Zhu 等^[17]提出一种基于 DWI 和 FLAIR 成像自动识别脑卒中发病时间(TSS)的机器学习算法,该算法首先需开发一个跨模态卷积神经网络,以便可以从 DWI 和 FLAIR 上准确分割病变;其次从 DWI 和 FLAIR 提取兴趣区(ROI)特征;最后将这些特征提供给机器学习模型以识别脑卒中发病时间。Arac^[18]认为,机器学习算法主要在三维数据的获取和分析上帮助临床医师进行神经康复训练三维运动分析,并表现出良好的应用前景。总之,各种机器学习算法广泛应用于脑卒中各领域,无论是在早期的疾病诊断检测还是后期的治疗与预后均表现出巨大的应用潜力^[19]。

二、机器学习算法在脑卒中诊断与分类诊断、结局预测、风险因素识别中的应用

1. 脑卒中诊断与分类诊断 脑卒中分类诊断与病因分型对临床选择针对性治疗方案和预防措施以及进行临床研究具有重要意义^[20]。现有分类系统如哈佛大学卒中登记处(The Harvard Cooperative Stroke Registry)^[21],美国国立神经病学、语言障碍和卒中研究所(NINCDS)卒中数据库^[22]等,对脑卒中各亚型的诊断准确性和可靠性均较低,这是由于数据库包含大量病例数据,在大数据背景下进行人工

分类诊断不仅耗时费力,且极易出现错误。脑卒中分类诊断和病因分型方法有多种,但临床尚无统一标准。缺血性卒中占脑卒中的 70% 以上,病死率略低于出血性卒中,但发病率远高于出血性卒中^[23-24],因此,机器学习算法研究主要集中于预测缺血性卒中发病时间和症状^[25-26],以及制定缺血性卒中治疗方案^[27]。Garg 等^[28]将电子病历(EHR)自然语言处理与机器学习算法相结合,实现缺血性卒中病因分型的自动化,且分类诊断效果优于人工分类。机器学习算法的分类运算性能和优势在脑卒中诊断中表现出巨大潜力,不同类型算法根据输入数据提取不同脑卒中亚型特征,分类准确性和精确性均较高,但各算法之间因纳入数据标准不同,并未显示出何种算法的分类效果更佳^[29]。

2. 脑卒中结局预测 早期研究主要关注机器学习算法在脑卒中发病时间和诊断时间预测中的作用,可显著提高预测准确性^[5,30]。随着机器学习算法应用广度和深度的不断拓展,其在缺血性卒中结局预测方面也表现出较高的准确性,对指导治疗至关重要。脑卒中结局包括放射学离散结局(最终梗死体积、出血性转化等)、合并症(脑卒中相关肺炎)、病死率及各种功能独立性指标[包括改良 Rankin 量表(mRS)评分、Barthel 指数(BI)、认知功能和语言功能等]^[5]。一项对比分析机器学习算法与回归模型预测缺血性卒中功能结局的研究显示,随机森林、分类回归树(CART)、支持向量机和 LASSO 回归模型等机器学习算法预测脑卒中后功能障碍的准确性较高(65%~72%)^[31]。Matsumoto 等^[32]总结单中心住院治疗的 4237 例急性缺血性卒中患者的临床数据,对比分析 6 种传统脑卒中预后模型[院前合并症、意识水平、年龄、神经功能缺损(PLAN)评分,缺血性卒中预测风险评分(IScore),洛桑急性脑卒中登记分析(ASTRAL),休斯顿动脉内血管再通治疗(HIAT)评分,血管事件总体健康风险(THRIVE)评分,年龄和美国国立卫生研究院卒中量表评分总和(SPAN-100)]与新型预测模型(线性回归或决策树集成模型)预测不良功能结局和医院内病死率的效能,结果显示,线性回归或决策树集成模型对不良功能结局预测效能略高于传统模型(曲线下面积为 0.88~0.94 对 0.70~0.92),但对医院内病死率预测效能相当(曲线下面积为 0.84~0.88 对 0.87~0.88)。因此认为,机器学习预测模型如支持向量机、随机森林和人工神经网络模型等的预测

准确性均较高,可为临床决策提供信息支撑^[33]。但是由于患者个体差异和临床症状不同,上述机器学习算法对脑卒中的结局预测也不尽相同,因此准确预测脑卒中结局仍存挑战;加之各类算法所纳入的数据不同,应用场景也有较大差异,未来研究需进一步统一各类机器学习算法的数据纳入标准、规范应用场景,从而提高脑卒中结局预测的准确性和可靠性。

3. 风险因素识别 脑卒中风险因素可以分为不可变因素和可变因素两大类,不可变因素包括性别、年龄、种族、民族等;可变因素包括高血压、高胆固醇血症、糖尿病、动脉粥样硬化等,此外,不良生活习惯如久坐、肥胖、营养不良、吸烟、饮酒也可诱发脑卒中^[34-36]。疾病早期如果未识别出可变因素,有可能增加脑卒中发病率和病死率。随着机器学习算法在脑卒中研究领域的广泛应用,其对脑卒中风险因素识别的准确性逐渐提高,增强临床医护人员对脑卒中的预警意识,从而进行有效预防并及时选择治疗措施,对缩短住院时间、改善患者预后、减轻家庭和社会经济负担具有重要意义。传统的弗莱明翰卒中量表(FSP)难以精确识别各种风险因素,给临床决策提出挑战^[37]。机器学习算法则可通过计算机高速运算快速识别各种脑卒中风险因素,例如,Yang等^[38]基于机器学习算法构建出高精度的高血压发病后3年脑卒中风险预测模型,该模型较易在电子病例中实施,对脑卒中风险因素的筛查更普遍,有效提高疾病早期预防和干预效率;Scrutinio等^[39]采用机器学习算法预测重度脑卒中患者3年后病死率,发现该算法可以更精确地识别有死亡风险的脑卒中患者;Bacchi等^[40]采用机器学习算法预测脑卒中预后风险,从而为出院后康复训练计划的制定提供依据。上述研究均表明相较传统风险评估方法,机器学习算法对脑卒中风险因素识别具有更高的准确性和更大的潜力。应注意的是,机器学习算法基于大数据,而临床实践中缺乏大样本、高质量数据集,有可能使结果产生偏差,无法得出准确的诊断和预测结果^[5,41],影响临床决策;此外,机器学习算法存在黑盒效应,使算法决策和运行过程缺乏可解释性,应重点关注机器学习算法的自身风险,如数据传输的安全性和隐私性等^[42],研究显示,通过机器学习算法提取敏感和机密信息时极易受病毒攻击,导致机密信息外泄,无法保证数据的安全性和隐私性^[43],因此,在医疗数据的提取和应用

过程中,应充分考虑数据的安全性以及数据使用的隐私性和伦理性等问题。未来研究或许需要更多关注机器学习算法自身的风险因素,进一步探讨临床实践中机器学习算法的安全性。

三、小结与展望

机器学习算法应用于脑卒中研究是当前热点,各类算法广泛应用于脑卒中的诊断与分类诊断、结局预测以及风险因素识别等各领域,并表现出巨大的应用潜力,有助于临床医师做出更好的临床决策。其中,机器学习算法评估脑卒中风险因素是国内外学者关注的重点,不仅用于脑卒中风险因素识别,而且对机器学习算法自身存在的风险和伦理道德提出担忧,这或许成为未来研究方向。临床医师也应意识到机器学习算法的固有缺陷,在临床实践中不能完全依赖该算法,而是将其作为一种辅助工具。随着机器学习算法在不同研究领域应用广度和深度的拓展,其应用于脑卒中诊断与治疗领域的文献呈井喷式增长,不难预测未来该研究领域必将迎来繁荣发展的阶段。

利益冲突 无

参 考 文 献

- [1] Samuel AL. Some studies in machine learning using the game of checkers[J]. IBM J Res Dev, 2000, 44:206-226.
- [2] Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis[J]. Cancer Inform, 2007, 2:59-77.
- [3] GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019 [J]. Lancet Neurol, 2021, 20:795-820.
- [4] Lee H, Lee EJ, Ham S, Lee HB, Lee JS, Kwon SU, Kim JS, Kim N, Kang DW. Machine learning approach to identify stroke within 4.5 hours[J]. Stroke, 2020, 51:860-866.
- [5] Mainali S, Darsie ME, Smetana KS. Machine learning in action: stroke diagnosis and outcome prediction [J]. Front Neurol, 2021, 12:734345.
- [6] Arbabshirani MR, Fornwalt BK, Mongelluzzo GJ, Suever JD, Geise BD, Patel AA, Moore GJ. Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration [J]. NPJ Digit Med, 2018, 1:9.
- [7] Ho KC, Speier W, Zhang H, Scalzo F, El-Saden S, Arnold CW. A machine learning approach for classifying ischemic stroke onset time from imaging[J]. IEEE Trans Med Imaging, 2019, 38:1666-1676.
- [8] Park E, Lee K, Han T, Nam HS. Automatic grading of stroke symptoms for rapid assessment using optimized machine learning and 4-limb kinematics: clinical validation study [J]. J Med Internet Res, 2020, 22:e20641.
- [9] Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets[J]. Neural Comput, 2006, 18:1527-1554.
- [10] Saba L, Biswas M, Kuppili V, Cuadrado Godia E, Suri HS, Edla

- DR, Omerzu T, Laird JR, Khanna NN, Mavrogeni S, Protogerou A, Sfikakis PP, Viswanathan V, Kitas GD, Nicolaides A, Gupta A, Suri JS. The present and future of deep learning in radiology [J]. *Eur J Radiol*, 2019, 114:14-24.
- [11] Ramos -Murguialday A, Broetz D, Rea M, Läer L, Yilmaz O, Brasil FL, Liberati G, Curado MR, Garcia-Cossio E, Vyziotis A, Cho W, Agostini M, Soares E, Soekadar S, Caria A, Cohen LG, Birbaumer N. Brain - machine interface in chronic stroke rehabilitation: a controlled study[J]. *Ann Neurol*, 2013, 74:100-108.
- [12] Murray NM, Unberath M, Hager GD, Hui FK. Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: a systematic review[J]. *J Neurointerv Surg*, 2020, 12:156-164.
- [13] Tang Z, Xu Y, Jin L, Aibaidula A, Lu J, Jiao Z, Wu J, Zhang H, Shen D. Deep learning of imaging phenotype and genotype for predicting overall survival time of glioblastoma patients[J]. *IEEE Trans Med Imaging*, 2020, 39:2100-2109.
- [14] Li L, Wei M, Liu B, Atchaneeyasakul K, Zhou F, Pan Z, Kumar SA, Zhang JY, Pu Y, Liebeskind DS, Scalzo F. Deep learning for hemorrhagic lesion detection and segmentation on brain CT images[J]. *IEEE J Biomed Health Inform*, 2021, 25:1646-1659.
- [15] Ni Y, Alwell K, Moomaw CJ, Woo D, Adeoye O, Flaherty ML, Ferioli S, Mackey J, De Los Rios La Rosa F, Martini S, Khatri P, Kleindorfer D, Kissel BM. Towards phenotyping stroke: leveraging data from a large - scale epidemiological study to detect stroke diagnosis[J]. *PLoS One*, 2018, 13:e0192586.
- [16] Maier O, Schröder C, Forkert ND, Martinetz T, Handels H. Correction: classifiers for ischemic stroke lesion segmentation: a comparison study[J]. *PLoS One*, 2016, 11:e0149828.
- [17] Zhu H, Jiang L, Zhang H, Luo L, Chen Y, Chen Y. An automatic machine learning approach for ischemic stroke onset time identification based on DWI and FLAIR imaging [J]. *Neuroimage Clin*, 2021, 31:102744.
- [18] Arac A. Machine learning for 3D kinematic analysis of movements in neurorehabilitation [J]. *Curr Neurol Neurosci Rep*, 2020, 20:29.
- [19] Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, Wang Y, Dong Q, Shen H, Wang Y. Artificial intelligence in healthcare: past, present and future[J]. *Stroke Vasc Neurol*, 2017, 2:230-243.
- [20] Yahiya S, Yousif A, Bakri M. Classification of ischemic stroke using machine learning algorithms[J]. *Int J Comput App*, 2016, 149:26-31.
- [21] Mohr JP, Caplan LR, Melski JW, Goldstein RJ, Duncan GW, Kistler JP, Pessin MS, Bleich HL. The harvard cooperative stroke registry: a prospective registry [J]. *Neurology*, 1978, 28: 754-762.
- [22] Sacco RL, Ellenberg JH, Mohr JP, Tatemichi TK, Hier DB, Price TR, Wolf PA. Infarcts of undetermined cause: the NINCDS Stroke Data Bank[J]. *Ann Neurol*, 1989, 25:382-390.
- [23] Washington HH, Glaser KR, Ifejika NL. CE: acute ischemic stroke[J]. *Am J Nurs*, 2021, 121:26-33.
- [24] Inamdar MA, Raghavendra U, Gudigar A, Chakole Y, Hegde A, Menon GR, Barua P, Palmer EE, Cheong KH, Chan WY, Ciaccio EJ, Acharya UR. A review on computer aided diagnosis of acute brain stroke[J]. *Sensors (Basel)*, 2021, 21:8507.
- [25] Ho KC, Speier W, El-Saden S, Arnold CW. Classifying acute ischemic stroke onset time using deep imaging features [J]. *AMIA Annu Symp Proc*, 2018, 2017:892-901.
- [26] Yu Y, Guo D, Lou M, Liebeskind D, Scalzo F. Prediction of hemorrhagic transformation severity in acute stroke from source perfusion MRI [J]. *IEEE Trans Biomed Eng*, 2018, 65:2058 -2065.
- [27] Fatahzadeh M, Glick M. Stroke: epidemiology, classification, risk factors, complications, diagnosis, prevention, and medical and dental management [J]. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*, 2006, 102:180-191.
- [28] Garg R, Oh E, Naidech A, Kording K, Prabhakaran S. Automating ischemic stroke subtype classification using machine learning and natural language processing[J]. *J Stroke Cerebrovasc Dis*, 2019, 28:2045-2051.
- [29] Govindarajan P, Soundarapandian RK, Gandomi AH, Patan R, Jayaraman P, Manikandan R. Classification of stroke disease using machine learning algorithms [J]. *Neural Comput App*, 2020, 32:817-828.
- [30] Asadi H, Dowling R, Yan B, Mitchell P. Machine learning for outcome prediction of acute ischemic stroke post intra -arterial therapy[J]. *PLoS One*, 2014, 9:e88225.
- [31] Alaka SA, Menon BK, Brobbey A, Williamson T, Goyal M, Demchuk AM, Hill MD, Sajobi TT. Functional outcome prediction in ischemic stroke: a comparison of machine learning algorithms and regression models[J]. *Front Neurol*, 2020, 11: 889.
- [32] Matsumoto K, Nohara Y, Soejima H, Yonehara T, Nakashima N, Kamouchi M. Stroke prognostic scores and data - driven prediction of clinical outcomes after acute ischemic stroke[J]. *Stroke*, 2020, 51:1477-1483.
- [33] Lin CH, Hsu KC, Johnson KR, Fann YC, Tsai CH, Sun Y, Lien LM, Chang WL, Chen PL, Lin CL, Hsu CY; Taiwan Stroke Registry Investigators. Evaluation of machine learning methods to stroke outcome prediction using a nationwide disease registry [J]. *Comput Methods Programs Biomed*, 2020, 190:105381.
- [34] Sirsat MS, Fermé E, Câmara J. Machine learning for brain stroke: a review[J]. *J Stroke Cerebrovasc Dis*, 2020, 29:105162.
- [35] Campbell BCV, De Silva DA, Macleod MR, Coutts SB, Schwamm LH, Davis SM, Donnan GA. Ischaemic stroke[J]. *Nat Rev Dis Primers*, 2019, 5:70.
- [36] Campbell BCV, Khatri P. Stroke[J]. *Lancet*, 2020, 396:129-142.
- [37] Pelcher I, Puzo C, Tripodis Y, Aparicio HJ, Steinberg EG, Phelps A, Martin B, Palmisano JN, Vassey E, Lindbergh C, McKee AC, Stein TD, Killiany RJ, Au R, Kowall NW, Stern RA, Mez J, Alosco ML. Revised framingham stroke risk profile: association with cognitive status and MRI - derived volumetric measures[J]. *J Alzheimers Dis*, 2020, 78:1393-1408.
- [38] Yang Y, Zheng J, Du Z, Li Y, Cai Y. Accurate prediction of stroke for hypertensive patients based on medical big data and machine learning algorithms: retrospective study[J]. *JMIR Med Inform*, 2021, 9:e30277.
- [39] Scrutinio D, Ricciardi C, Donisi L, Losavio E, Battista P, Guida P, Cesarelli M, Pagano G, D'Addio G. Machine learning to predict mortality after rehabilitation among patients with severe stroke[J]. *Sci Rep*, 2020, 10:20127.
- [40] Bacchi S, Oakden-Rayner L, Menon DK, Jannes J, Kleinig T, Koblar S. Stroke prognostication for discharge planning with machine learning: a derivation study[J]. *J Clin Neurosci*, 2020, 79:100-103.
- [41] Sakai K, Yamada K. Machine learning studies on major brain diseases: 5-year trends of 2014–2018[J]. *Jpn J Radiol*, 2019, 37:34-72.
- [42] Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: addressing ethical challenges [J]. *PLoS Med*, 2018, 15:e1002689.
- [43] McGraw G, Bonett R, Shepardson V, Figueroa H. The top 10 risks of machine learning security[J]. *Comput*, 2020, 53:57-61.

(收稿日期: 2023-01-19)

(本文编辑: 彭一帆)