

人工智能在颈椎疾病诊断与治疗中应用的相关研究进展

Advances in research on the application of artificial intelligence in diagnosis and treatment of cervical diseases

庞华程, 欧阳汉强, 李危石

(北京大学第三医院骨科 骨与关节精准医学工程研究中心 脊柱疾病研究北京市重点实验室 100191 北京市)

doi: 10.3969/j.issn.1004-406X.2023.11.10

中图分类号: R681.5 文献标识码: A 文章编号: 1004-406X(2023)-11-1043-04

人工智能(artificial intelligence, AI)是近年来发展迅速的一种多学科交叉的尖端技术,可以通过算法对海量数据的学习实现类似 AI 解决问题的能力^[1]。近年来, AI 技术的积累和储备已经完成, AI 赋能传统产业已进入大规模应用的快速发展时期。在医疗领域, AI 与医疗的深度融合也正不断推进,在医学影像识别、疾病诊断、手术规划和预后等方面取得了大量重要成果。

脊柱相关疾病已成为现代社会中的常见病,而脊柱疾病诊断和治疗方案选择难度大,往往依赖医生的临床经验,导致基层医疗机构对脊柱疾病的诊断和治疗能力往往较为有限;针对这一问题,开发和推广医疗 AI 是一种可能的解决方案^[2]。另一方面,脊柱外科的临床实践中,医生往往需要综合患者的多方面信息权衡风险与收益,做出恰当的治疗方案选择,而以深度学习为代表的 AI 技术在影像识别和辅助决策方面具有独特的优势,因此在脊柱外科领域有着广阔的应用前景^[3]。基于以上原因,近年来 AI 与脊柱外科的交叉学科领域蓬勃发展,相关研究日新月异;而在这一领域中,由于颈椎椎体小,结构相对胸椎与腰椎更为精细复杂,以及颈椎退变疾病在欧美国家发病率相对较低等因素, AI 在颈椎区域的应用起步晚,直到近几年才涌现出了一批优秀的研究成果。目前,虽然对 AI 在脊柱整体上的应用已有大量报道,但针对 AI 在颈椎这一区域的应用仍缺乏综述性报道。笔者拟从影像识别、疾病诊断、手术方案选择和预后评估几个方面,综述 AI 在颈椎疾病诊疗中应用的研究现状与发展趋势。

1 颈椎影像识别及处理

随着计算机硬件的进步和 AI 算法的进步,影像识别

和处理已成为 AI 在颈椎疾病诊疗中应用最早、成果较多的领域。颈椎的主要解剖结构包括椎体、脊髓和椎间盘三大部分,对各个解剖结构进行识别和分割是进一步运用 AI 分析颈椎影像,实现影像学参数自动测量、AI 辅助诊断和治疗方案规划,进而直接服务于临床实践的基础,因此也是 AI 在颈椎医学中最早开始探索的领域。2014 年, Daenzer 等^[4]开发了从颈椎磁共振(magnetic resonance, MR)三维图像数据生成二元梯度方向直方图的方法,采用线性分类器、全局多尺度搜索等技术,实现了在颈椎 MRI 上定位 C3~C7 的颈椎椎体,经过在 21 例影像上的测试,在不同大小的不同椎体上取得了良好的效果,这一工作是 AI 在颈椎医学领域内应用最早的成果。近年来, AI 参与颈椎影像识别和处理的研究发展加速,处理的影像种类、功能、精确度都有了长足的进步。2019 年, Huang 等^[5]用视觉词袋模型和 KAZE 算法构建计算机视觉分类器,在前路椎间盘切除减压融合术(anterior cervical discectomy and fusion, ACDF)后患者的颈椎影像上识别手术植入材料的型号,模型给出的首选项与实际植入型号符合率达到(91.5±3.8)%,展示了 AI 技术在颈椎影像识别中的一种新的应用方式。2020 年, Bae 等^[6]采用 2D 卷积神经网络(convolutional neural network, CNN),将 17 名多种颈椎疾病患者和 24 名健康人的颈椎 CT 影像分别作为训练集和验证集,建立了在 CT 影像上对颈椎椎体进行自动分割的算法,再将二者调换进行进一步验证,两种方式下训练的模型都表现出了与专家进行人工标注相近的准确度。2022 年, Wang 等^[7]尝试用机器学习的方法实现对颈椎 X 线平片上椎体顶点等标志的自动标注并测量椎体之间的夹角,运用多输入适应性 U-Net(multi-input adaptive U-Net)学习不同影像之间的相似特征,测试结果中对椎体标志的自动标注的最小点到点误差为 12.988 像素,对椎体间夹角测量的最小对称平均绝对百分比误差为 26.969%。

采用 AI 对影像进行自动识别和标注的目的是辅助医师测量具有临床意义的颈椎运动参数,如 Cobb 角、SVA 值、椎体滑移距离等,同时判断颈椎受压等情况,这一方面也已经出现许多有价值的工作。2020 年, Shin 等^[8]采用能

基金项目:国家自然科学基金(82102638);北京大学第三医院临床重点项目(BYSYZD2021040);北京大学第三医院队列建设项目(BYSYDL2022007);北京大学教学新思路 2.0 项目(2023YB01);国家卫生健康委医疗质量(循证)管理研究项目(YLZLXZ23G027)

第一作者简介:男(2002-),本科生在读,研究方向:骨科

电话:(010)82266626 E-mail:phc@bjmu.edu.cn

通讯作者:欧阳汉强 E-mail:ouyanghanqiang@bjmu.edu.cn

对颈椎侧位 X 线平片进行自动分割的深度学习系统,对 2006~2018 年中的 13691 例颈椎侧位平片上的颈椎曲度进行测量,发现 13 年内 70 岁以下的群体中颈椎正常生理曲度有显著的丢失,体现了 AI 在大规模的临床和流行病学研究中的价值。2021 年,Merali 等^[9]用退行性颈椎病患者的 MRI 影像建立了检测脊髓受压迫的深度学习模型。该模型用 201 例影像进行训练,在 88 例影像中测试,结果中 ROC 曲线下面积(area under curve,AUC)达到 0.94,敏感性为 0.88,特异性为 0.81,F1 分数为 0.82。

2 颈椎疾病诊断

为了提前规划手术,选择恰当的术式、器械和入路,对颈椎疾病进行精确的定位和定性诊断是临床上一项重要的工作。由于涉及脊髓和神经根的受压和变性,颈椎疾病的躯体运动障碍和肢体麻痹等症状复杂多变,临床上需要结合专科查体、影像学检查和肌电图等其他辅助检查明确诊断,这一过程高度依赖医生的临床经验。因此,AI 在提高对颈椎疾病的诊断效率,降低误诊率和漏诊率方面具有潜在的应用价值。2019 年,Cui 等^[10]用随机森林分类(random forests classification)的方法处理 C4~6 脊髓受大鼠模型的躯体感觉触发电位(somatosensory evoked potentials,SEP)并对其进行时频分析,进而模拟对脊髓型颈椎病的定位诊断,模型总体的准确性为 84.72%。2021 年,Small 等^[11]评估了 Aidoc 公司开发的卷积神经网络模型 C-spine 在 CT 图像上识别颈椎骨折的能力,在 665 个样本上的准确率为 92%,敏感性和特异性分别为 76%和 97%,达到了与影像学专业相近的水准,并且模型所忽略的骨折所在的层面和解剖结构与人工阅片的结果相似,作者认为这一结果证明了这一模型的辅助阅片的实用价值;同年,Voter 等^[12]对同一模型评估了诊断准确性并进行了失效模式分析(failure mode analysis),发现模型对颈椎骨折的识别敏感性受患者年龄、骨折的部位和严重程度等因素影响,并且容易将颈椎退行性改变识别为颈椎骨折;研究者认为这一模型在未来推广中仍然存在许多难题,并建议未来类似的 AI 辅助决策系统(decision support system)的开发和评估需要标准化。2022 年,Lee 等^[13]用卷积神经网络在颈椎侧位 X 线平片上建立识别颈椎病的算法,准确性达到 87.1%,AUC 为 0.864;作者认为,这一 AI 算法可以辅助医生决定患者是否需要进一步行 MRI 检查以进一步诊断颈椎病。2023 年,Tamai 等^[14]采用卷积神经网络模型 EfficientNetB2 训练了在颈椎 X 线平片上识别由退行性椎管狭窄引起的脊髓压迫的算法,AUC 为 0.81,准确率达到 81.0%,超过了 10 名作为对比的医生。

颈椎后纵韧带骨化(ossification of posterior longitudinal ligaments,OPLL)是一种由颈椎后纵韧带发生骨化,压迫脊髓和神经根造成一系列躯体运动和感觉症状,或内脏神经功能紊乱的疾病。目前认为,OPLL 进展可能导致退行性颈椎病和脊髓损伤的发生,因此,如何在早期无症状阶

段识别 OPLL 患者并评估其风险,从而尽早干预、避免患者的不良结局,是目前颈椎领域一大研究课题。目前,OPLL 的诊断主要依靠影像学上发现后纵韧带骨化的表现,因此 AI 在识别 OPLL 方面有潜在的应用价值。2021 年,Murata 等^[15]采用残差神经网络(residual neural network,RNN)在 672 例患者的 2318 张颈椎侧位 X 线平片中训练识别颈椎后纵韧带骨化的模型,分类模型的准确度为 98.9%,灵敏度为 97.0%,特异性为 99.4%,假阳性率为 2.2%,假阴性率为 1.0%,总体 AUC 达到 0.99。2022 年,Ogawa 等^[16]在 50 例通过 CT 影像确诊为后纵韧带骨化的患者和 50 例因腰椎手术导致无症状的颈椎改变的患者的颈椎屈、伸和中立位 X 线平片中用 CNN 训练了分类模型,识别后纵韧带骨化的准确性为 90%,敏感度和特异性分别为 80%和 100%,尽管使用的样本规模较小,但取得了比普通骨科医生和脊柱外科医生更好的表现。2022 年,Tamai 等^[17]收集了 486 例颈椎侧位 X 线和 CT 影像,构建识别后纵韧带骨化的 CNN 模型,识别准确度为 0.88,AUC 为 0.94,诊断准确性同样超过了作为对比的脊柱外科医生。

3 手术方案与预后

手术治疗可以直接解除颈椎病理改变对脊髓和神经的压迫,改善颈椎的力学关系,恢复颈椎生理曲度和椎间高度,是颈椎病治疗的重要手段^[8]。但是,颈椎病手术方法复杂,入路、术式和器材的选择需要在前期充分采集病史和辅助检查,尤其是影像资料,综合分析后做出决策。若能将 AI 运用于预测患者手术预后,从而为这一决策过程提供更多参考,将有助于高水平中心的诊疗经验惠及基层的更多患者。相较于前述影像识别和疾病诊断领域,这一方面研究者已经取得了相对丰富的成果。ACDF 手术是临床上十分常见的颈椎病外科治疗手段,如何通过术前的评估和影像学检查,预测行 ACDF 术患者术后的脊髓功能恢复情况和并发症的发生,是研究者十分关心的问题。2018 年,Arvind 等^[18]进行了一项多中心的研究,用人工神经网络(artificial neural network,ANN)、Logistic 回归(logistic regression,LR)、支持向量机(support vector machine,SVM)和随机森林决策树(random forest decision tree)四种模型,分别分析了 20879 例行 ACDF 术颈椎病患者的年龄、性别、种族、生活习惯、术前评分等信息,并据此预测 ACDF 术后死亡和手术并发症的发生,结果表明 ANN 和 LR 对 ACDF 术后并发症的预测能力超过了美国麻醉师协会制定的体格情况分级,并且预测死亡和伤口并发症时 ANN 比 LR 表现出更好的敏感性。2021 年,Wong 等^[20]用 SVM 模型分析了 ACDF 术前患者的颈椎两侧深部肌群的 MRI 影像,通过影像参数和特征预测术后早期邻近节段病的发生,在 62 个样本中,预测的准确率达到 96.7%,AUC 为 0.97,研究者认为,对术前颈深部肌群的横截面积、组成和不对称性的测量对于预测 ACDF 术后早期邻近节段病有一定价值,可以用于患者的预后评估和个性化干预;同

年,Goedmakers 等^[21]收集了来自两个中心的,进行 ACDF 术的颈椎病患者的术前 T2 矢状面 MRI 影像,用深度学习的方法构建了对术后邻近节段病的预测模型,准确率达到 95%,且准确率、灵敏度、特异度都超过了影像科和神经外科的医生。2022 年,Rudisill 等^[22]开发了预测 ACDF 术后早期邻近节段退变的机器学习模型,预测准确率达到 0.82, F1 分数为 0.79, AUC 为 0.794, 并且找到了预测早发邻近节段退变最重要的风险因素,包括 C4/5 椎间盘向前或后突出、C6 椎体后方的骨赘、骨赘的出现和 C6/7 椎间盘突出。

除相对热门 ACDF 术预后预测研究外,更广泛的颈椎手术预后预测中 AI 的使用也得到了研究者的关注。2022 年,Valliani 等^[23]用梯度提升树方法预测颈椎融合手术术后住院时间延长的发生,在来自单中心的数据集上的 AUC 为 0.87(SD=0.01),在美国国家住院患者样本数据库(National Inpatient Sample Data Set, NIS)中验证, AUC 为 0.84(SD=0.00)。2021 年,Zhang 等^[24]通过分别采用对 T2 加权 MRI 影像处理得到的传统影像学指标和影像学组学数据,用机器学习的方法构建预测脊髓型颈椎病的患者术后恢复情况的模型,并比较两种来源的模型的优劣;结果表明,基于影像学特征构建的模型能预测术后脊髓功能的恢复情况,且用线性 SVM 预处理后,用 SVM 构建的模型预测性能最优。此外,Zhang 等^[24]基于脊髓型颈椎病患者的 T2 加权 MRI 影像组学构建了极度随机树模型用于预测手术患者术后的恢复情况,构建的模型根据采用数据不同包括影像学指标、影像学组学、临床特征和影像学以及临床特征和影像学组学四种,结果表明基于影像学组学,和基于临床特征与影像学组学的两种模型的 AUC 在 0.7~0.8 之间,有相对更好的预测能力。2022 年,Kim 等^[25]比较了机器学习方法和简单的 Logistic 回归对后纵韧带骨化患者术后发生 C5 神经根麻痹的预测能力,研究者采用了 35 种机器学习预测模型和 5 种 Logistic 回归预测模型,通过不同模型的 AUC 比较模型的预测性能。结果表明,机器学习算法整体上比 Logistic 回归模型有更强的预测能力,其中一种采用自适应强化学习算法的模型有最好的 AUC 表现,为 0.88,高于 Logistic 回归模型中表现最好的 0.55。

精准医学是以个体化医疗为基础,结合 3D 打印、大数据分析 and 基因测序等手段,实现对患者精准治疗的医学模式,而精准外科手术是精准医学的重要组成部分。运用 AI,结合患者个体信息,对脊柱手术中的具体操作进行个性化规划,从而完成精准手术,是近年来较新的研究方向。2022 年,Liu 等^[26]使用离体的猪椎骨构建骨质疏松模型,对比了规划椎弓根螺钉钉道的 AI 模型和传统的人工方法规划结果,发现模型规划的钉道具有更高的抗拔出。同年, Ma 等^[27]开发了一种基于 CT 扫描结果规划椎弓根螺钉钉道的 AI 模型,可以选出相比于 AO 标准具有更高骨密度和抗拔出力的钉道,但这一研究是在腰椎上进行的。目前,颈椎领域的 AI 辅助精准手术研究仍较为缺少,具有广阔

研究前景。

4 总结与展望

随着 AI,尤其是机器学习领域技术的快速发展和迭代,颈椎领域 AI 识别影像、进行预测的性能也在不断提升;在 AI 赋能医疗发展深度推进的潮流中,颈椎疾病的 AI 诊疗目前正迅猛发展。结合以上综述,可将现有研究存在的问题列举如下:

(1)由于病例数据流通不便等原因,现有研究基本为单中心,缺少多中心、大规模的研究。由于单中心研究存在人群较为单一、各医院所使用影像设备等硬件设施存在差异等问题,研究中训练的模型在其他环境应用的效果需要更进一步评估。

(2)目前机器学习的研究主要单独考虑病历文字、X 线、CT 或 MRI 等信息中的一种,缺少多模态相结合、综合给出诊断和预后的探索,如能充分利用多模态信息,将提高诊断和预测的准确性,同时更切合临床需求。

(3)由于作为当下本领域内主流的神经网络本身的“黑箱”特点,目前的研究对算法识别和预测缺乏因果分析,对模型的准确率难以解释,往往停留在对个别病例中机器识别与人工标注结果出现分歧的原因的简单猜测。

(4)训练 AI 进行影像识别,通常需由相关专业的医生对影像进行人工标注;关于标注中所使用的工具软件,目前尚未具备统一的标准,造成影像标注文件格式不互通、在不同终端兼容性差等问题,影响了学术交流和研究成果最终服务临床的进程;对此,一些学会和组织给出了标准或建议,如美国放射学会在其网站上给出的信息化指引 (<https://www.acr.org/Practice-Management-Quality-Informatics/Informatics>),但目前并未形成广泛的行业内共识。

未来颈椎疾病智能医疗的进一步发展,需要通过广泛的学术交流,由相关组织和机构牵头,形成行业共识和标准。在具体的技术层面,需要紧跟 AI 领域的新发展、新技术,不断优化改进使用的算法;推进可用于少见疾病和特殊生理构造识别的小样本训练;融合多模态信息,打通从早期发现疾病、影像识别和智能辅助诊断、治疗方案规划、手术到术后康复的链条;探索可解释的 AI 系统,将机器单纯的“给出答案”变为“给出解释”,提高广大医疗从业者和患者对 AI 的信任程度。此外,在医学伦理和医疗法律层面,机器学习模型训练中使用的患者资料的隐私保护需要重视,任何研究都应当取得患者的知情同意、并对敏感数据进行严格管理;AI 参与医疗行为,影响患者和医生的医疗决策从而产生的责任归属等法律问题,也需要更多的探索。

5 参考文献

1. Mintz Y, Brodie R. Introduction to artificial intelligence in medicine[J]. Minim Invasive Ther Allied Technol, 2019, 28(2):

- 73–81.
2. 欧阳汉强, 姜亮, 刘晓光, 等. 人工智能在脊柱外科诊断与治疗中的应用现状和发展趋势[J]. 中华骨科杂志, 2019, 39(24): 1543–1548.
3. Stephens ME, O’Neal CM, Westrup AM, et al. Utility of machine learning algorithms in degenerative cervical and lumbar spine disease: a systematic review[J]. *Neurosurg Rev*, 2022, 45(2): 965–978.
4. Daenzer S, Freitag S, von Sachsen S, et al. VolHOG: a volumetric object recognition approach based on bivariate histograms of oriented gradients for vertebra detection in cervical spine MRI[J]. *Med Phys*, 2014, 41(8): 082305.
5. Huang KT, Silva MA, See AP, et al. A computer vision approach to identifying the manufacturer and model of anterior cervical spinal hardware[J]. *J Neurosurg Spine*, 2019: 1–7.
6. Bae HJ, Hyun H, Byeon Y, et al. Fully automated 3D segmentation and separation of multiple cervical vertebrae in CT images using a 2D convolutional neural network [J]. *Comput Methods Programs Biomed*, 2020, 184: 105119.
7. Wang Y, Huang L, Wu M, et al. Multi-input adaptive neural network for automatic detection of cervical vertebral landmarks on X-rays[J]. *Comput Biol Med*, 2022, 146: 105576.
8. Shin Y, Han K, Lee YH. Temporal trends in cervical spine curvature of south Korean adults assessed by deep learning system segmentation, 2006–2018[J]. *JAMA Netw Open*, 2020, 3(10): e2020961.
9. Merali Z, Wang JZ, Badhiwala JH, et al. A deep learning model for detection of cervical spinal cord compression in MRI scans[J]. *Sci Rep*, 2021, 11(1): 10473.
10. Cui H, Wang Y, Li G, et al. Exploration of cervical myelopathy location from somatosensory evoked potentials using random forests classification[J]. *IEEE Trans Neural Syst Rehabil Eng*, 2019, 27(11): 2254–2262.
11. Small JE, Osler P, Paul AB, et al. CT cervical spine fracture detection using a convolutional neural network[J]. *AJNR Am J Neuroradiol*, 2021, 42(7): 1341–1347.
12. Voter AF, Larson ME, Garrett JW, et al. Diagnostic accuracy and failure mode analysis of a deep learning algorithm for the detection of cervical spine fractures [J]. *AJNR Am J Neuroradiol*, 2021, 42(8): 1550–1556.
13. Lee GW, Shin H, Chang MC. Deep learning algorithm to evaluate cervical spondylotic myelopathy using lateral cervical spine radiograph[J]. *BMC Neurol*, 2022, 22(1): 147.
14. Tamai K, Terai H, Hoshino M, et al. Deep learning algorithm for identifying cervical cord compression due to degenerative canal stenosis on radiography[J]. *Spine (Phila Pa 1976)*, 2023, 48(8): 519–525.
15. Murata K, Endo K, Aihara T, et al. Use of residual neural network for the detection of ossification of the posterior longitudinal ligament on plain cervical radiography[J]. *Eur Spine J*, 2021, 30(8): 2185–2190.
16. Ogawa T, Yoshii T, Oyama J, et al. Detecting ossification of the posterior longitudinal ligament on plain radiographs using a deep convolutional neural network: a pilot study[J]. *Spine J*, 2022, 22(6): 934–940.
17. Tamai K, Terai H, Hoshino M, et al. A deep learning algorithm to identify cervical ossification of posterior longitudinal ligaments on radiography[J]. *Sci Rep*, 2022, 12(1): 2113.
18. 中华外科杂志编辑部. 颈椎病的手术治疗及围手术期管理专家共识(2018)[J]. 中华外科杂志, 2018, 56(12): 881–884.
19. Arvind V, Kim JS, Oermann EK, et al. Predicting surgical complications in adult patients undergoing anterior cervical discectomy and fusion using machine learning[J]. *Neurospine*, 2018, 15(4): 329–337.
20. Wong AYL, Harada G, Lee R, et al. Preoperative paraspinal neck muscle characteristics predict early onset adjacent segment degeneration in anterior cervical fusion patients: a machine-learning modeling analysis[J]. *J Orthop Res*, 2021, 39(8): 1732–1744.
21. Goedmakers CMW, Lak AM, Duey AH, et al. Deep learning for adjacent segment disease at preoperative MRI for cervical radiculopathy[J]. *Radiology*, 2021, 301(3): 664–671.
22. Rudisill SS, Hornung AL, Barajas JN, et al. Artificial intelligence in predicting early-onset adjacent segment degeneration following anterior cervical discectomy and fusion[J]. *Eur Spine J*, 2022, 31(8): 2104–2114.
23. Valliani AA, Feng R, Martini ML, et al. Pragmatic prediction of excessive length of stay after cervical spine surgery with machine learning and validation on a national scale [J]. *Neurosurgery*, 2022, 91(2): 322–330.
24. Zhang MZ, Ou-Yang HQ, Liu JF, et al. Predicting postoperative recovery in cervical spondylotic myelopathy: construction and interpretation of T (2)(*)-weighted radiomic-based extra trees models[J]. *Eur Radiol*, 2022, 32(5): 3565–3575.
25. Kim SH, Lee SH, Shin DA. Could machine learning better predict postoperative C5 palsy of cervical ossification of the posterior longitudinal ligament[J]. *Clin Spine Surg*, 2022, 35(5): E419–E425.
26. Liu JB, Zuo R, Zheng WJ, et al. The accuracy and effectiveness of automatic pedicle screw trajectory planning based on computer tomography values: an in vitro osteoporosis model study[J]. *BMC Musculoskeletal Disorders*, 2022, 23(1): 165.
27. Ma C, Zou D, Qi H, et al. A novel surgical planning system using an AI model to optimize planning of pedicle screw trajectories with highest bone mineral density and strongest pull-out force[J]. *Neurosurg Focus*, 2022, 52(4): E10.

(收稿日期:2023-04-18 末次修回日期:2023-09-17)

(本文编辑 彭向峰)