

· 综述 ·

人工智能辅助 X 线识别对骨科的应用价值研究进展

薛冬¹, 徐海林², 王伟^{1,3}

(1. 锦州医科大学附属第一医院骨科, 辽宁 锦州 121000; 2. 北京大学人民医院骨科, 北京 100000; 3. 锦州医科大学骨外科学研究所, 辽宁 锦州 121000)

【摘要】 成像作为评估肌肉骨骼病情的重要工具, 从对疾病的患病风险评估到对疾病及病情进展的判断以及预后评分等均起到重要作用, 伴随人工智能(artificial intelligence, AI)在图像检测和图像解释领域中的迅速发展, 一些涉及肌肉骨骼 X 射线成像的 AI 辅助识别研究已经检验并显示了很高的潜在价值, 可增强 X 射线成像价值链的各个部分, 可通过提高成像效率、成像质量和诊断准确性对临床医生起指导性作用。目前, AI 辅助成像识别技术发展仍处在早期阶段, AI 算法需要进一步的提升和发展, 影像数据仍不足且质量存在较大异质性, 技术性能的长期准确性和稳定性需要进一步观察研究。

【关键词】 人工智能; 图像处理, 计算机辅助; X 线; 综述
中图分类号: R603

DOI: 10.12200/j.issn.1003-0034.2020.09.019

开放科学(资源服务)标识码(OSID):



Progress on artificial intelligence assisted X-ray film recognition in orthopedics XUE Dong, XU Hai-lin*, and WANG Wei. *Department of Orthopaedics, Peking University People's Hospital, Beijing 100000, China

ABSTRACT As an important tool for assessing musculoskeletal conditions, imaging plays an important role in assessing the risk of disease, judging disease and the progress of disease, and prognosis scores. Accompanied with the rapid development of artificial intelligence(AI) in the field of image detection and interpretation, some AI-assisted recognition studies involving musculoskeletal X-ray imaging have been examined and shown a high potential value, which can enhance various parts of the X-ray imaging value chain and guide clinicians by improving imaging efficiency, imaging quality, and diagnostic accuracy. At present, the development of AI-assisted imaging recognition technology is still at an early stage. AI algorithms need to be further improved and developed. Image data is still insufficient and the quality is relatively heterogeneous. The long-term accuracy and stability of technical performance require further observation and research.

KEYWORDS Artificial intelligence; Image processing, computer-assisted; X-rays; Review

进入 21 世纪后, 各种形式的数字成像在医学和医疗保健领域变得越来越强大和不可或缺^[1], 影像归档和通信系统(picture archiving and communication systems, PACS)对影像数据质量的提升提高了数据整合的可行性和数据分析的有效性, 以及深度学习(deep learning, DL)算法对异构数据集的适应性, 均促进了人工智能(artificial intelligence, AI)在医学成像中的应用研究, 可替代人类医生大量的重复性工作, 减轻工作压力, 协助提高诊断率、降低误诊率、减少漏诊、降低患者看病成本、协助改善工作流程^[2-3]。X 射线成像作为肌肉骨骼系统诊断成像领域的领导者, 是除病史和临床检查外, 对骨科疾病诊疗的重要支柱, AI 辅助 X 平片识别对骨科的应用研究具有很

高的潜在价值。

1 人工智能

人工智能被定义为通过研究人工的方法和技术, 模仿、延伸和扩展人的智能, 实现机器智能。机器学习(machine learning, ML)作为人工智能主要研究领域之一, 通过算法使得机器从大量数据中挖掘数据特征、学习规律, 从而拥有对新数据分析和研究的能力, 并不断获取新的知识或技能来改善自身的性能。其中, 所提取的医学图像特征要能够抵抗噪声, 强度和旋转角度的变化, 比如像素值和边缘强度等^[4]。机器学习根据任务类型分为 3 大类: 监督学习、无监督学习和强化学习, 监督学习侧重于分类^[5], 是对患者和健康受试者进行区分的主要应用。

深度学习是机器学习、神经网络研究中一个新领域, 是包含多个隐藏处理层的人工神经网络算法的子集, 具有很强的泛化能力和非线性映射能力^[6],

通讯作者: 徐海林 E-mail: hailinxu66@qq.com

Corresponding author: XU Hai-lin E-mail: hailinxu66@qq.com

且深度学习神经网络还可纳入放射学文本报告^[7-8]、实验室值和先前成像结果等医疗记录,可给出更适当成像检查方法的选择意见,可更好地识别图像表现,也可能会简化临床实践的工作流程,成为目前在医学成像领域中应用最多的算法。

2 AI 辅助 X 平片识别在骨科中的应用研究

2.1 骨折

在实际临床工作中,X 线平片是针对疑似或明确骨折患者必需的检查手段,目前 AI 辅助骨折识别研究在骨科领域的 AI 技术应用中最为广泛,依据解剖部位包括股骨颈骨折^[9-10]、转子间骨折^[10-11]、腕部骨折^[12-15]、桡骨远端骨折^[16]、肱骨近端骨折^[17]和踝关节骨折^[18]等,研究通过基于 AI 算法构建模型使其能够对特定骨骼部位的 X 线片进行自动化检测,判断是否存在骨折,且在性能测试中 AUC 值均达 0.95 以上,平均准确率在 90%~97%,接近人类专家水平,可为基层医生及急诊科医生提供准确、及时的决策支持,为急症患者节省抢救时间。另外,Chung 等^[17]在研究中依据 Neer 分型尝试对肱骨骨折类型进行判断,虽然仅获得 86% 的最优模型准确率,整体表现不佳,但证明了网络模型对骨折类型判断的可行性和发展潜力。研究中所构建网络模型包括 VGG、DenseNet、ResNet、GoogleNet、AlexNet 等,多数研究对数据进行了增强处理以提高系统模型的性能,包括随机裁剪、随机翻转、随机旋转和仿射变换等技术。

2.2 骨龄评估

骨龄评估可以反映个体的生长发育水平和成熟程度,常用于评价儿科某些内分泌疾病。目前,骨龄评估多采用 Greulich 和 Pyle (GP) 或 Tanner-Whitehouse (TW) 方法,过程繁琐、耗时,且观察者间的差异性降低了评估的准确性。一些研究^[19-21]使用各年龄段左手腕部 X 线片数据集研究了 AI 算法对骨骼成熟度进行自动化评估的表现,并采用均方根或平均绝对误差进行性能测试,获得较高准确率,误差范围仅为 0.5~0.8 年之间不等,但目前研究中模型对骨龄的评估值缺乏统一的参考标准,均以所纳入的放射学文本报告或限定的几位医生的主观性评估值为依据,如若对其进行交叉验证,准确率可能会下降,但 Tajmir 等^[21]验证了 AI 与人类医生相结合获得的性能要高于 AI 或人类医生单独诊断的性能,可更好地降低变异性、提高准确性,证实了 AI 辅助骨龄评估的潜在价值。

2.3 骨关节炎

骨关节炎 (osteoarthritis, OA) 是最常见的关节疾病,常导致慢性疼痛并造成运动受限,对全世界 10%~15% 成人的日常生活产生影响^[22],该病发生受

多种因素的影响,包括生活习惯、性别、肥胖、慢性病等,Lim 等^[23]基于这些健康行为信息及医疗资源利用的统计数据,开发了一种深度学习模型来检测 OA 患病率,确定 OA 发病高危因素,以期早期发现 OA 高风险患者。当 OA 中晚期,骨质结构发生变化,如骨赘和硬膜下硬化,关节间隙也会变窄,X 线片表现较为明显,Tiulpin 等^[24]基于 AI 算法应用膝关节 X 线片成像构建网络模型,对膝关节是否患有 OA 进行自动化识别,并依据 Kellgren-Lawrence 分级量表对 OA 严重程度进行概率分布推断,可能无法明确的将严重程度归类为某一个等级,但与医生在临床实践中的表现没有太大的区别。另外,Hirvasniemi 等^[25]对髋关节炎的骨质变化进行特征提取分析,实现对 OA 发病率的自动预测,并对需行全髋置换治疗的病情进行评估。此上 3 个研究,贯穿了 OA 发展的全部过程,未来,可将 3 个方面研究进行整合并进一步完善,比如扩大应用范围,包括脊柱关节、手指间关节等,全方位覆盖 OA 的预防和诊治,实现对 OA 的精准治疗,改善预后。

2.4 脊柱相关疾病

伴随 AI 辅助成像识别的应用研究进展,实现了基于 X 线片对脊柱形状的自动评估能力^[26-27],Horng 等^[26]基于脊柱矢状位图像通过对椎体骨轮廓检测、分割,进而测量 Cobb 角大小以计算脊柱弯曲度,而 Galbusera 等^[27]在研究中基于脊柱矢状位和冠状位双平面图像通过识别端板中心、髋关节中心和 S1 端板边缘等 78 个解剖标志点,除对 Cobb 角进行测量外,评估了 T₄-T₁₂ 脊柱后凸角、L₁-L₅ 脊柱前凸角、骨盆入射角和骶骨、骨盆倾斜度等参数,实现对脊柱畸形更全面的全自动放射学分析。这些研究可消除医生手动测量所产生的误差,更快地给出更为客观的测量结果,更好地辅助诊断、治疗和随访青少年特发性脊柱侧凸、成人畸形和椎管狭窄等脊柱相关疾病。

2.5 骨质疏松

骨质疏松患者骨密度下降或骨形成不足^[28],骨强度和韧性下降,可加速脊柱退变性侧凸畸变^[29],增加骨折风险。伴随人口老龄化,骨质疏松愈加被重视,骨质疏松性骨折或伴有骨质疏松的骨折的治疗理念转变^[30],使得对骨质疏松的诊断至关重要,Singh 等^[31]从骨质疏松患者跟骨的 X 线片中提取了与骨质微结构变化相关的包括纹理强度、对比度等在内的 6 个统计特征,训练构建机器学习模型,运用监督分类的方法对骨质疏松症患者和健康受试者进行区分。另外,机器学习被应用于髌部骨折,通过对髌部骨折的危险因素及其相互作用进行统计分析,对骨质疏松性髌部骨折进行风险评估,改善对髌部

骨折的预测能力,相关危险因素包括骨密度、骨折史、药物使用以及生化胆固醇和生化葡萄糖含量变化等^[32-34],这些研究表明,骨小梁微结构特征与 AI 算法的结合可更准确的预测骨小梁的生物力学强度,有助于骨质疏松症疾病进展和治疗反应的自动化客观评估。

所有上述运用深度学习神经网络算法的研究中,学习训练过程基本相同,均将数据集分为训练数据集、验证数据集和测试数据集,对训练数据集进行注释,然后将注释的数据输入神经网络模型,对模型进行学习培训,使其能够对输入的数据点输出相应的数据标签;然后运用验证数据集对层宽度(每层的单位数)、激活函数、二次函数的使用类型和数据增强的程度、正规化水平等超参数进行调整,以优化网络模型性能;最后运用独立于训练和验证数据集的测试数据集^[35]对模型进行性能测试并将其与人类专家的诊断性能作对比分析。

3 AI 继续发展面临的挑战

3.1 AI 算法瓶颈

(1)黑匣子。到目前为止所描述的所有机器学习方法,权重和激活函数中使用的值均不能被提取为可由人类释的某种形式信息。因此,当人类医生与系统模型对图像的判断产生分歧时,差异无法通过讨论来解决^[36],可能会对医生的诊疗措施选择造成负面影响。(2)过度拟合。研究所构建的网络模型性能与纳入数据量之间具有相关性,数据量越大越易实现算法稳定的收敛和高分类精度,数量过少会造成训练中图像的重复使用,尤其当模型过于复杂时,过度拟合使模型在训练数据中表现良好,而在新数据中表现不佳^[37],限制了其泛化能力,而当前 AI 算法发展无法满足在小数据量时实现高性能。

3.2 数据质量的异质性

目前研究所纳入的数据源较单一,而不同医疗机构因设备不同、拍摄标准的选取不同以及影像技师间差异等因素导致图像的质量参差不齐,从而限制了 AI 技术的泛化和推广能力。因此,需要在实际临床实践中使用 AI 技术时注意他们的表现如何受到医院、成像模式及扫描或重建条件之间差异的影响,以寻求更稳健的方法对共享图像进行质量控制并确保图像数据的完整性,比如建立 PACS 系统数据储存模式统一性、完善 PACS 系统储存能力等。

3.3 研究设计的局限性

目前为止,研究均是对先前生成的影像数据执行回顾性分析,虽然在性能测试中获得较满意的 ROC 和 AUC 表现,但两者不一定表示临床效用,且研究对模型的评估时间太短,缺乏前瞻性分析,不能

反映长期的准确性和稳定性,致使 AI 技术能否与临床实践相适应及所面对的患者人群选择等问题缺乏考证。

3.4 医生面临的挑战

AI 医学技术性能有望很快超过人类专家,可能会加重部分医师对智能报告的依赖性而逐步丢失自身的专业知识,弱化专业技能,降低专业水平,虽有作者对 AI 技术持怀疑态度^[38-39],也有作者提出 AI 医疗技术将会在不久的将来取代放射科医生和解剖及病理学专家^[40],面对 AI 的迅速发展,人类医生需要学习 AI 相关理论知识,了解机器学习工具的相关属性,获得指导、监督和监视 AI 在临床中安全且有效应用的能力^[41-42],而这些跨领域、跨学科的知识对人类医生来说都是全新的开始,尤其对于较年老的医生更是极大的挑战。

4 展望

AI 在医疗领域的应用发展目标已经明确,未来 AI 算法将会伴随数据的增加而得到扩展,性能会进一步改善,其发展空间也将更为广阔,在智能诊疗、医疗机器人、健康管理等方面都将呈现质的飞跃,可协助解决医疗资源供需失衡、医疗质量分配不均、人口老龄化等重要问题,也可帮助实现优质资源下沉,为基层医生提供及时、准确的决策支持。在骨科领域,AI 将会辅助诊疗更为复杂的病变,比如肌腱损伤,骨肿瘤、小儿骨创伤等。我们要在确保数据隐私性和安全性的基础上积极参与 AI 医疗技术的研究与开发,促使 AI 与医疗更好的融合。

参考文献

[1] Naylor CD. On the prospects for a (deep) learning health care system[J]. JAMA, 2018, 320(11): 1099-1100.

[2] Thrall JH, Li X, Li Q, et al. Artificial intelligence and machine learning in radiology: opportunities, challenges, pitfalls, and criteria for success[J]. J Am Coll Radiol, 2018, 15(3 Pt B): 504-508.

[3] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence[J]. Nat Med, 2019, 25(1): 44-56.

[4] Fu GS, Levin-Schwartz Y, Lin QH, et al. Machine learning for medical imaging[J]. J Health Eng, 2019, 2019: 9874591.

[5] Deo RC. Machine learning in medicine[J]. Circulation, 2015, 132(20): 1920-1930.

[6] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.

[7] Hosny A, Parmar C, Quackenbush J, et al. Artificial intelligence in radiology[J]. Nat Rev Cancer, 2018, 18(8): 500-510.

[8] van Ginneken B. Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning[J]. Radiol Phys Technol, 2017, 10(1): 23-32.

[9] Adams M, Chen W, Holcdorf D, et al. Computer vs human: deep learning versus perceptual training for the detection of neck of femur fractures[J]. J Med Imaging Radiat Oncol, 2019, 63(1): 27-32.

- [10] Cheng CT, Ho TY, Lee TY, et al. Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs[J]. *Eur Radiol*, 2019, 29(10): 5469–5477.
- [11] Urakawa T, Tanaka Y, Goto S, et al. Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network[J]. *Skeletal Radiol*, 2019, 48(2): 239–244.
- [12] Jakub O, Niklas F, MAKI A, et al. Artificial intelligence for analyzing orthopedic trauma radiographs[J]. *Acta Orthop*, 2017, 88(6): 581–586.
- [13] de Bruijn B, Cranney A, O'Donnell S, et al. Identifying wrist fracture patients with high accuracy by automatic categorization of X-ray reports[J]. *J Am Med Inform Assoc*, 2006, 13(6): 696–698.
- [14] Kim DH, MacKinnon T. Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks[J]. *Clin Radiol*, 2018, 73(5): 439–445.
- [15] Lindsey R, Daluiski A, Chopra S, et al. Deep neural network improves fracture detection by clinicians[J]. *Proc Natl Acad Sci USA*, 2018, 115(45): 11591–11596.
- [16] Gan KF, Xu DL, Lin YM, et al. Artificial intelligence detection of distal radius fractures: a comparison between the convolutional neural network and professional assessments[J]. *Acta Orthop*, 2019, 90(4): 394–400.
- [17] Chung SW, Han S, Lee JW, et al. Automated detection and classification of the proximal humerus fracture by using deep learning algorithm[J]. *Acta Orthop*, 2018, 89(4): 468–473.
- [18] Kitamura G, Chung CY, Moore BE 2nd. Ankle fracture detection utilizing a convolutional neural network ensemble implemented with a small sample, de novo training, and multiview incorporation[J]. *J Digit Imaging*, 2019, 32(4): 672–677.
- [19] Lee H, Tajmir S, Lee J, et al. Fully automated deep learning system for bone age assessment[J]. *J Digit Imaging*, 2017, 30(4): 427–441.
- [20] Mutasa S, Chang PD, Ruzal-Shapiro C, et al. MABAL: a novel deep-learning architecture for machine-assisted bone age labeling[J]. *J Digit Imaging*, 2018, 31(4): 513–519.
- [21] Tajmir SH, Lee H, Shailam R, et al. Artificial intelligence-assisted interpretation of bone age radiographs improves accuracy and decreases variability[J]. *Skeletal Radiol*, 2019, 48(2): 275–283.
- [22] Ezzat AM, Li LC. Occupational physical loading tasks and knee osteoarthritis: a review of the evidence[J]. *Physiother Can*, 2014, 66(1): 91–107.
- [23] Lim J, Kim J, Cheon S. A deep neural network-based method for early detection of osteoarthritis using statistical data[J]. *Int J Environ Res Public Health*, 2019, 16(7): PMID: 30974803.
- [24] Tiulpin A, Thevenot J, Rahtu E, et al. Automatic knee osteoarthritis diagnosis from plain radiographs: a deep learning-based approach[J]. *Sci Rep*, 2018, 8(1): 1727.
- [25] Hirvasniemi J, Gielis WP, Arabi S, et al. Bone texture analysis for prediction of incident radiographic hip osteoarthritis using machine learning: data from the Cohort Hip and Cohort Knee (CHECK) study[J]. *Osteoarthritis Cartilage*, 2019, 27(6): 906–914.
- [26] Horng MH, Kuok CP, Fu MJ, et al. Cobb angle measurement of spine from X-ray images using convolutional neural network[J]. *Comput Math Methods Med*, 2019, 2019: 6357171.
- [27] Galbusera F, Niemeyer F, Wilke HJ, et al. Fully automated radiological analysis of spinal disorders and deformities: a deep learning approach[J]. *Eur Spine J*, 2019, 28(5): 951–960.
- [28] Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures[J]. *BMJ*, 1996, 312(7041): 1254–1259.
- [29] 黄道余, 沈亚骏, 王飞, 等. 退行性腰椎侧凸与骨质疏松症的相关性分析[J]. *中国骨伤*, 2019, 32(3): 244–247.
HUANG DY, SHEN YJ, WANG F, et al. Correlative analysis of degenerative lumbar scoliosis and osteoporosis[J]. *Zhongguo Gu Shang/China J Orthop Trauma*, 2019, 32(3): 244–247. Chinese with abstract in English.
- [30] 伍骥, 陈渲宇, 郑超. 再度认识骨质疏松性脊柱骨折的治疗[J]. *中国骨伤*, 2019, 32(7): 587–590.
WU J, CHEN XY, ZHENG C. Re-recognizing the treatment of osteoporotic spine fractures[J]. *Zhongguo Gu Shang/China J Orthop Trauma*, 2019, 32(7): 587–590. Chinese.
- [31] Singh A, Dutta MK, Jennane R, et al. Classification of the trabecular bone structure of osteoporotic patients using machine vision[J]. *Comput Biol Med*, 2017, 91: 148–158.
- [32] Kruse C, Eiken P, Vestergaard P. Machine learning principles can improve hip fracture prediction[J]. *Calcif Tissue Int*, 2017, 100(4): 348–360.
- [33] Ho-Le TP, Center JR, Eisman JA, et al. Prediction of hip fracture in post-menopausal women using artificial neural network approach[J]. *Conf Proc IEEE Eng Med Biol Soc*, 2017, 2017: 4207–4210.
- [34] Tseng WJ, Hung LW, Shieh JS, et al. Hip fracture risk assessment: artificial neural network outperforms conditional logistic regression in an age- and sex-matched case control study[J]. *BMC Musculoskelet Disord*, 2013, 14: 207.
- [35] Carin L, Pencina MJ. On deep learning for medical image analysis[J]. *JAMA*, 2018, 320(11): 1192–1193.
- [36] Yasaka K, Abe O. Deep learning and artificial intelligence in radiology: current applications and future directions[J]. *PLoS Med*, 2018, 15(11): e1002707.
- [37] Lee JG, Jun S, Cho YW, et al. Deep learning in medical imaging: general overview[J]. *Korean J Radiol*, 2017, 18(4): 570–584.
- [38] Aminololama-Shakeri S, López JE. The doctor-patient relationship with artificial intelligence[J]. *AJR Am J Roentgenol*, 2019, 212(2): 308–310.
- [39] Wang Y, Yan FG, Lu XF, et al. IILS: Intelligent imaging layout system for automatic imaging report standardization and intra-interdisciplinary clinical workflow optimization[J]. *EBioMedicine*, 2019, 44: 162–181.
- [40] Obermeyer Z, Emanuel EJ. Predicting the future -big data, machine learning, and clinical medicine[J]. *N Engl J Med*, 2016, 375(13): 1216–1219.
- [41] Stead WW. Clinical implications and challenges of artificial intelligence and deep learning[J]. *JAMA*, 2018, 320(11): 1107–1108.
- [42] Verghese A, Shah NH, Harrington RA. What this computer needs Is a physician: humanism and artificial intelligence[J]. *JAMA*, 2018, 319(1): 19–20.

(收稿日期: 2020-02-19 本文编辑: 李宜)