

人工智能在骨肌影像中的诊断应用进展

朱 寻综述, 卢光明审校

【摘要】 目前人工智能技术在医学领域的研究和应用中迅速发展,医学影像也是人工智能在医学领域中的重要应用方向之一。人工智能技术已成功应用于评估儿童骨龄、检测骨折和评估 X 线片上骨关节炎的严重程度等方面,最近的研究也证明了使用人工智能技术在 CT 和 MRI 上识别各种病理异常的可行性,包括转移性疾病、内部紊乱、骨折、感染和关节退化等。文章主要就人工智能在骨肌系统影像的诊断和应用进展进行综述。

【关键词】 人工智能;骨肌影像;诊断;应用

【中图分类号】 R445.2 **【文献标志码】** A **【文章编号】** 1672-271X(2023)01-0063-05

【DOI】 10.3969/j.issn.1672-271X.2023.01.013

Diagnostic and application of artificial intelligence in musculoskeletal system imaging

ZHU Xun reviewing, LU Guangming checking

(Department of Radiology, Jinling Clinical College of Medicine, Nanjing Medical University/General Hospital of Eastern Theater Command, PLA, Nanjing 210002, Jiangsu, China)

【Abstract】 At present, artificial intelligence technology has developed rapidly in the research and application of the medical field. Medical imaging is also one of the important application directions of artificial intelligence in the medical field. AI has been successful in assessing bone age in children, detecting fractures, and assessing the severity of osteoarthritis on X-rays, and recent studies have demonstrated the feasibility of using AI to identify various pathological abnormalities on CT and MRI, including metastatic diseases, internal disorders, fractures, infections and joint degeneration. This paper mainly summarizes the diagnosis and application of artificial intelligence in bone and muscle system imaging.

【Key words】 artificial intelligence; bone muscle imaging; diagnosis; application

0 引 言

人工智能最早是在 20 世纪中叶被提出,被认为是第四次工业革命的代表性技术之一^[1],其中深度学习,是使用最广泛的人工智能形式之一,彻底改变了计算机科学,并迅速扩展到更广泛的科学和工程学科,包括医学影像^[2]。深度学习在不同的医学影像亚专业中有着广泛的应用,如神经、肺、腹部、

乳腺和心脏成像等^[3-4]。人工智能在骨骼肌肉成像的应用也正在兴起,包括组织分割、图像重建和疾病检测^[5-7]。将人工智能技术用于疾病检测对于放射科医师来说尤其重要,因为其可最大程度地提高诊断率,同时减少因分心和疲劳而导致的错误^[8]。本文就人工智能在骨肌系统影像的诊断和应用进展作一综述。

1 人工智能的概念及技术

人工智能是一门研究、开发用于模拟、扩展和延伸人的智能的理论、技术、方法和应用的信息科学^[9],人工智能可指任何执行任务的机器,这些任务通常需要人类的知识,其解决问题的方式与人类类似。深度学习是人工智能的一个子集,以人工

基金项目:科技部科技创新 2030--重大项目“新一代人工智能”专项 (2020AAA0109505)

作者单位:210002 南京,南京医科大学金陵临床医学院(东部战区总医院)医学影像科[朱 寻(现在上海市宝山区中西医结合医院医学影像科工作)、卢光明]

通信作者:卢光明, E-mail: cjr_luguangming@vip. 163. com

神经元层的组合为代表,每一层都包含许多单元,其中每个单元都是神经元细胞的简化表示,其灵感来自人脑中的结构^[10]。深度学习旨在自动从数据中提取模式或特征,并不需要人为指定规则。深度学习的核心思想是使用称为神经网络的特定算法架构。卷积神经网络(CNN)是主要的神经网络架构。CNN由一系列卷积层组成,其中散布着一系列池化层。CNN将多个卷积层连接到一个深度网络中,该网络可检测越来越抽象和复杂的特征层次结构。每个卷积层由一个或多个过滤器或内核组成,分析输入图像中区域体素的特征,继而生成特征图。CNN通常经过监督训练,其中大量成对的图像和参考分类标签被分为单独的训练、验证和测试组^[8]。为创建有效CNN模型,需要有大量的图像来训练模型,即先建立训练组。训练深度学习模型需要大量的标记数据,其中图像收集和注释是最耗时的任务,这需要有经验的医师来进行图像注释,这种标记也是后期训练和测试的黄金标准^[11]。一旦CNN模型在足够数量的图像上进行了训练,该模型的诊断性能就可以在一个单独的数据集上进行测试,称为验证组。最后将该模型应用到实际的临床工作中,形成相应的测试组^[8]。由于任务的复杂性和新的方法的出现,网络的层次越来越高,其模型的数量可达到上千个层次,这使深度学习能够完成更加复杂的推理和决策^[12]。

2 人工智能在骨骼肌肉成像中的临床应用

2.1 骨折检测中的应用 许多研究已经证明了人工智能在骨折检测中的应用,大多数研究使用开源CNN和大型训练数据集来检测多个身体部位的骨折,包括肩部、手腕、脚踝等^[13-15]。急诊科经常碰到疑似骨折的患者,而X射线成像是临床医师用于评估患者骨折的主要诊断工具。在X线片中遗漏骨折通常会给患者带来严重后果,导致治疗延迟和功能恢复不良。Lindsey等^[16]使用改进的U-Net分类CNN来检测11个身体部位的骨折,并利用由135845张X线片组成的训练组,使用由300张随机选择的X线片组成的保留测试组,该机器检测骨折的AUC为0.99,灵敏度和特异性分别为94%和95%。临床医师在使用机器辅助解读X线片时,误判率平均降低了47%。

目前多种深度学习方法也应用在CT上检测骨

折。Raghavendra等^[17]使用定制设计的分类CNN分析了100名有骨折的受试者和60名没有胸腰椎椎体骨折的受试者的CT扫描,最后机器检测椎体骨折的敏感性和特异性分别为100%和98%。Pranata等^[18]使用联合CNN和CAD的方法分析跟骨CT骨折,其中有683幅有骨折的CT图像和1248幅没有骨折的CT图像,该机器对跟骨骨折的准确率为98%。临床工作中,各大医院目前应用最多的是肋骨CT骨折的智能识别,人工智能在骨折检测中的应用具有很大的前景,极大的提高了医师的准确率,降低了漏诊率,避免了一些医疗事故的发生。然而,所有这些研究的局限性是,CNN模型必须在评估的特定身体部位进行训练,每个训练集都需要大量正确标记的X线图片、CT图片,通常是数万张。与人类不同,CNN识别一个身体部位骨折的能力不会延续到其他新的身体部位上。此外,所有报道的骨折检测研究的输出都是二元的:即骨折存在或不存在。目前还缺乏对骨折形态或指导特定治疗的影像学表现描述^[19]。

2.2 在X线片上评估骨龄 骨龄检测即利用X线摄片来观察被检测者左手掌骨、指骨、腕骨以及桡尺骨下端的骨化中心的发育程度,以此来确定骨龄,继而了解被检测者的生长发育情况有无异常。Larson等^[20]发表了最早和最大的研究之一,该研究使用14036张手部X线片来训练定制设计的分类CNN,以经验丰富的放射科医师的解释作为参考标准来估计儿科骨龄。使用由200张手部X线片组成的保留测试组,机器提供的骨龄估计值与人类阅片提供的骨龄估计值之间的均方根和平均绝对差分别为0.63岁和0.50岁。相较于临床常规使用Greulich-Pyle(GP)法或Tanner-Whitehouse(TW)法对X线片进行骨龄评估,机器算法更加高效且正确,减少了人为因素导致的主观性,缩短了阅片时间。目前一些算法已被美国食品和药物管理局(USFDA)批准用于临床实践。

2.3 在MRI上检测膝关节内部紊乱

2.3.1 软骨病变 深度学习方法已被用于在MRI上检测膝关节内的软骨病变。Liu等^[21]使用编码器-解码器VGG-16CNN在矢状面脂肪抑制T2加权二维快速自旋转回波(FSE)图像上分割软骨,并提取175名受试者股骨和胫骨关节面上约100个感兴趣区域的软骨图像。以经验丰富的放射科医师的

解释作为参考标准,确定软骨病变的存在或不存在。由 1310 个软骨图像组成的独立测试组进行了两项单独评估。该检测软骨病变的 ROC 曲线下面积(AUC)为 0.92,灵敏度和特异性分别为 84%和 85%。Pedoia 等^[22]使用 U-Net CNN 在 1478 名受试者的矢状脂肪抑制质子密度加权 3D FSE 图像上分割髌骨软骨,然后通过定制的分类 CNN 分析分割的软骨,该机器的 AUC 为 0.88,灵敏度和特异性为 80%。

2.3.2 半月板撕裂 半月板损伤非常常见,是由于发育不良、慢性劳损和急性扭伤等各种情况引起半月板完整性破坏,并伴有一系列临床症状,如疼痛和功能障碍,严重影响患者的活动能力和生活质量。一旦诊断出半月板损伤,大多数病例需要手术治疗。准确及时的术前诊断具有重要意义^[23]。Roblot 等^[24]使用快速区域 CNN 在 1123 名受试者的矢状面脂肪抑制 T2 加权二维 FSE 图像上分割半月板,然后使用经验丰富的放射科医师的解释作为参考标准来确定半月板撕裂的存在与否。该机器检测半月板撕裂时的 AUC 为 0.94。Couteaux 等^[25]使用类似的深度学习方法来检测脂肪抑制 T2 加权二维 FSE 图像上的半月板撕裂,一个由 1128 名受试者的图像组成的训练组和一个由 700 名受试者的图像组成的测试组,该机器检测半月板撕裂的 AUC 为 0.89,灵敏度和特异性分别为 90%和 82%。

2.3.3 前交叉韧带(ACL)撕裂 前交叉韧带撕裂是膝关节最常损伤的韧带之一,通常是由于膝关节扭转或过伸暴力超过其强度所致。前交叉韧带撕裂会加速关节退化并导致骨关节炎的产生。因此,需要在更短的时间内准确地诊断出有无 ACL 损伤^[26]。Liu 等^[27]使用由 LeNet-5 和 YOLA 组成的耦合 CNN,在 175 例 ACL 撕裂患者和 175 例无 ACL 撕裂患者的矢状质子密度加权和脂肪抑制 T2 加权二维 FSE 图像上分离 ACL。然后用 DenseNet 分类 CNN 分析分离的 ACL,以关节镜作为参考标准确定是否存在 ACL 撕裂。最后机器检测 ACL 撕裂的 AUC 为 0.98,灵敏度和特异性分别为 96%和 98%。Chang 等^[28]使用 U-Net CNN 在 130 名 ACL 撕裂患者和 130 名未撕裂患者的冠状质子密度加权二维 FSE 图像上分离出 ACL,然后用 ResNet 分类 CNN 对分离的 ACL 进行分析,以经验丰富的放射科医师的解释作为参考标准,确定 ACL 撕裂的存在或不

存在。该机器检测 ACL 撕裂的 AUC 为 0.97,灵敏度和特异性分别为 100%和 93%。对于膝关节内部紊乱病变,机器检测的准确率越高,越有利于帮助医师诊断,以便于临床及早进行相应的手术治疗。

2.4 在 CT 和核医学上检测骨转移性疾病 深度学习方法已被应用在脊柱 CT 和 PET-CT 上去定位椎体转移病灶。Chmelik 等^[29]使用定制设计的 CAD 方法在 31 例受试者的矢状面 CT 图像上分割整个脊柱,其中有 1046 个溶骨性病变和 1135 个硬化性病变。由定制设计的基于体素的分类 CNN 进行分析,以确定是否存在转移性疾病。该机器检测溶骨性病变的 AUC 为 0.80,检测硬化性转移病变的 AUC 为 0.72。机器在定位直径<1.5 mm 的小转移病灶时,每个受试者有 45.6 个假阳性(92%的敏感度),在定位直径>3 cm 的大转移病灶时,每个受试者有 5.9 个假阳性(99%的敏感度)。Xu 等^[30]使用两个增强的 V-Net CNN 级联构建 W 型框架来学习骨骼的体积特征表示,并使用 PET-CT 图像区分正常骨和骨转移瘤。训练数据包括 2000 个有转移性病变的图像块和 2000 个没有转移性病变的图像块,该机器具有 73%的敏感性和 99%的特异性。上述研究的局限性是对于大的肿瘤病灶有很好的敏感性,小病灶的敏感性低,要进一步提高诊断效能。

2.5 在 MRI 上检测退行性椎间盘疾病和脊柱感染 椎间盘疾病临床中尤为常见,使用机器学习会更有效的提高医师的诊断,为疾病的分级、治疗方案的选择提供帮助。有研究使用定制设计的 CAD 方法在 2009 个受试者的腰椎矢状 T2 加权二维 FSE 图像上分离椎体和椎间盘节段。然后用 VGG-M 分级 CNN 对分离的椎体和椎间盘节段进行分析,以确定是否存在退行性椎间盘疾病的各种表现,包括 Pfirrmann 分级、椎间盘狭窄分级、中央管狭窄、椎体终板缺损、椎体滑脱和骨髓改变,使用有经验的放射科医师的解释作为参考标准。机器和医师阅片之间的一致性在 70%之间(用于确定 Pfirrmann 等级)和 93%(用于确定椎体骨髓变化的存在或不存在)之间^[31-32]。Kim 等^[33]使用定制设计的分类 CNN 分析了 80 例结核性脊柱炎患者和 80 例化脓性脊柱炎患者的非分割轴向 T2 加权二维 FSE 图像,并以活检为参考标准。通过 4 倍交叉验证,该机器用于区分结核性和化脓性脊柱炎的 AUC 为 0.80,敏感性和特异性分别为 85%和 68%,该机器的诊断性能

与医师阅片评估相似。

2.6 X 线片和 MRI 上骨关节炎的检测 骨关节炎是一种广泛存在且具有破坏性的疾病,会导致疼痛、生活质量下降并产生相应的高昂医疗保健费用。因此要有术前准确的诊断,并对其进行评估,尽早干预,来达到具有成本效益的治疗^[34]。在临床实践和研究中 Kellgren-Lawrence (KL) 系统被广泛用于评估膝和髋关节 X 线片上骨关节炎(OA)的严重程度^[35]。Tiulpin 等^[36]利用膝关节对称性,用 Siamese 分类 CNN 评估膝关节骨性关节炎的严重程度。该模型使用来自 MOST 的图像进行训练,并在由 5960 张有骨关节炎的膝关节 X 线片组成的测试组上进行评估。该机器的平均多级精度为 67%。Norman 等^[37]使用改进的 DenseNet CNN 在膝关节 X 线片上分配 KL 等级。使用由 621 张有骨关节炎的膝关节 X 线片组成的保留测试组,该机器的检测结果示无 OA (KL 0 级、1 级)、轻度 OA (KL 2 级)、中度 OA (KL 3 级)、重度 OA (KL 4 级)的灵敏度分别为 84%、70%、69%、86%,对应的特异性分别为 86%、84%、97%、99%。上述研究对于骨关节炎严重程度的智能分级具有一定的指导意义。

3 结语与展望

人工智能在骨骼肌肉成像中目前有许多新兴应用,利用人工智能技术为疾病自动诊断提供了很多帮助,时间更短,效率更高,准确性更高。深度学习方法已成功用于肌肉骨骼成像的病变识别和严重程度评估,例如骨折、膝关节病变、骨龄评估、骨关节炎和脊柱退行性病变等。利用深度学习方法,可以大大降低医师的误诊率和漏诊率。同时深度学习需要大量的影像图片,也需要专业的医师进行图片的标注,需要耗费很多的时间,而且各医院的数据量和病例不同,也会影响相关研究的开展进行。在此研究过程中也会产生很多的假阳性和假阴性,图片的质量、解剖结构的异常等也会对研究结果产生一定的影响,只有不断的优化才能达到最佳的效果。因此,人工智能不可能完全取代放射科医师,但两者可齐头并进。只有充分利用好人工智能技术,并凭借放射科医师本身的经验水平,才能达到更好的诊断水准。

【参考文献】

- [1] Hricak H. 2016 New Horizons Lecture: beyond imaging-radiology oftomorrow[J]. Radiology, 2018,286(3):764-775.
- [2] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015,521: 436-444.
- [3] Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis[J]. Med Image Anal, 2017,42:60-88.
- [4] Shen D, Wu G, Suk HI. Deep learning in medical image analysis[J]. Annu Rev Biomed Eng, 2017,19:221-248.
- [5] Liu F, Zhou Z, Jang H, et al. Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging[J]. Magn Reson Med, 2018, 79(4):2379-2391.
- [6] Zhou Z, Zhao G, Kijowski R, et al. Deep convolutional neural network for segmentation of knee joint anatomy[J]. Magn Reson Med, 2018, 80(6):2759-2770.
- [7] Norman B, Pedoia V, Majumdar S. Use of 2D U-Net convolutional neural networks for automated cartilage and meniscus segmentation of knee MR imaging data to determine relaxometry and morphometry[J]. Radiology, 2018, 288(1):177-185.
- [8] Kijowski R, Liu F, Caliva F, et al. Deep Learning for Lesion Detection, Progression, and Prediction of Musculoskeletal Disease[J]. J Magn Reson Imaging, 2020, 52(6): 1607-1619.
- [9] 卢光明,张志强. 人工智能医学影像[J].医学研究生学报, 2018,31(7):683-687.
- [10] Kalmet PHS, Sanduleanu S, Primakov S, et al. Deep learning in fracture detection: a narrative review[J]. Acta Orthop, 2020, 91(2): 215-220.
- [11] Zhou XW, Wang H, Feng CY, et al. Emerging Applications of Deep Learning in Bone Tumors: Current Advances and Challenges[J]. Front Oncol, 2022, 12: 908873.
- [12] Olczak J, Fahlberg N, Maki A, et al. Artificial intelligence for analyzing orthopedic trauma radiographs[J]. Acta Orthop, 2017, 88(6): 581-586.
- [13] 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.
- [14] 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.
- [15] Kitamura G, Chung CY, Moore BE. 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.
- [16] 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.
- [17] Raghavendra U, Bhat NS, Gudigar A, et al. Automated system for the detection of thoracolumbar fractures using a CNN architecture[J]. Futur Gener Comput Syst, 2018,85:184-189.
- [18] Pranata YD, Wang KC, Wang JC, et al. Deep learning and

- SURF for automated classification and detection of calcaneus fractures in CT images [J]. *Comput Methods Programs Biomed*, 2019, 171: 27-37.
- [19] Chea P, Mandell JC. Current applications and future directions of deep learning in musculoskeletal radiology [J]. *Skeletal Radiol*, 2019, 49(2): 183-197.
- [20] Larson DB, Chen MC, Lungren MP, *et al*. Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs [J]. *Radiology*, 2018, 287(1): 313-322.
- [21] Liu F, Zhou Z, Samsonov A, *et al*. Deep learning approach for evaluating knee MR images: Achieving high diagnostic performance for cartilage lesion detection [J]. *Radiology*, 2018, 289(1): 160-169.
- [22] Pedoia V, Norman B, Mehany SN, *et al*. 3D convolutional neural networks for detection and severity staging of meniscus and PFJ cartilage morphological degenerative changes in osteoarthritis and anterior cruciate ligament subjects [J]. *J Magn Reson Imaging*, 2019, 49: 400-410.
- [23] Jie Li, Kun Qian, Jinyong Liu, *et al*. Identification and diagnosis of meniscus tear by magnetic resonance imaging using a deep learning model [J]. *J Orthop Translat*, 2022, 34: 91-101.
- [24] Roblot V, Giret Y, Bou Antoun M, *et al*. Artificial intelligence to diagnose meniscus tears on MRI [J]. *Diagn Interv Imaging*, 2019, 100(4): 243-249.
- [25] Couteaux V, Si-Mohamed S, Nempont O, *et al*. Automatic knee meniscus tear detection and orientation classification with Mask-RCNN [J]. *Diagn Interv Imaging*, 2019, 100(4): 235-242.
- [26] Awan MJ, Rahim MSM, Salim N, *et al*. Improved Deep Convolutional Neural Network to Classify Osteoarthritis from Anterior Cruciate Ligament Tear Using Magnetic Resonance Imaging [J]. *J Pers Med*, 2021, 11(11): 1163.
- [27] Liu F, Guan B, Zhou Z, *et al*. Fully-automated diagnosis of anterior cruciate ligament tears on knee MR images using deep learning [J]. *Radiol Artif Intell*, 2019, 1(3): 180091.
- [28] Chang PD, Wong TT, Rasiej MJ. Deep learning for detection of complete anterior cruciate ligament tear [J]. *J Digit Imaging*, 2019, 32(6): 980-986.
- [29] Chmelik J, Jakubicek R, Walek P, *et al*. Deep convolutional neural network-based segmentation and classification of difficult to define metastatic spinal lesions in 3D CT data [J]. *Med Image Anal*, 2018, 49: 76-88.
- [30] Xu L, Tetteh G, Lipkova J, *et al*. Automated whole-body bone lesion detection for multiple myeloma on 68Ga-pentixafor PET/CT imaging using deep learning methods [J]. *Contrast Media Mol Imaging*, 2018, 2018: 2391925.
- [31] Jamaludin A, Kadir T, Zisserman A. SpineNet: Automated classification and evidence visualization in spinal MRIs [J]. *Med Image Anal*, 2017, 41: 63-73.
- [32] Jamaludin A, Loots M, Kadir T, *et al*. ISSLS prize in bioengineering science 2017: Automation of reading of radiological features from magnetic resonance images (MRIs) of the lumbar spine without human intervention is comparable with an expert radiologist [J]. *Eur Spine J*, 2017, 26: 1374-1383.
- [33] Kim K, Kim S, Lee YH, *et al*. Performance of the deep convolutional neural network based magnetic resonance image scoring algorithm for differentiating between tuberculous and pyogenic spondylitis [J]. *Sci Rep*, 2018, 8(1): 13124.
- [34] Abramoff B, Caldera FE. Osteoarthritis: Pathology, Diagnosis, and Treatment Options [J]. *Med Clin North Am*, 2020, 104(2): 293-311.
- [35] Larson DB, Chen MC, Lungren MP, *et al*. Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs [J]. *Radiology*, 2018, 287(1): 313-322.
- [36] 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.
- [37] Norman B, Pedoia V, Noworolski A, *et al*. Applying densely connected convolutional neural networks for staging osteoarthritis severity from plain radiographs [J]. *J Digit Imaging*, 2018, 32(3): 471-477.

(收稿日期: 2022-11-05; 修回日期: 2022-12-13)

(责任编辑: 刘玉巧; 英文编辑: 吕镗烽)