

Research progresses of convolutional neural network in musculoskeletal radiology

WANG Chunjie, YUAN Huishu*

(Department of Radiology, Peking University Third Hospital, Beijing 100191, China)

[Abstract] Artificial intelligence has been gradually applied in medical image diagnosis, showing good efficiency and diagnostic accuracy. As a recent innovation in artificial intelligence, convolutional neural network (CNN) displayed the ability to interpret medical images with accuracy at or near that of skilled clinicians for some applications, indicating overwhelming clinical application prospects. The research progresses of CNN in musculoskeletal radiology were reviewed in this article.

[Keywords] skeleton; muscles; artificial intelligence; neural networks, computer; diagnostic imaging

DOI:10.13929/j.issn.1003-3289.2020.09.024

卷积神经网络在骨骼肌肉放射学中的研究进展

王春杰, 袁慧书*

(北京大学第三医院放射科, 北京 100191)

[摘要] 人工智能逐渐应用于医学图像诊断, 显现出良好的效率和诊断准确率。卷积神经网络(CNN)是人工智能领域的一项突破, 在某些方面已可与放射科医师相媲美, 展现出巨大的临床应用前景。本文就 CNN 在骨骼肌肉放射学中的研究进展进行综述。

[关键词] 骨骼; 肌肉; 人工智能; 神经网络, 计算机; 诊断显像

[中图分类号] R681; R445 **[文献标识码]** A **[文章编号]** 1003-3289(2020)09-1375-04

卷积神经网络(convolutional neural network, CNN)是人工智能(artificial intelligence, AI)的一个快速发展的子集, 为放射学领域的最新技术之一, 适用于解决图像识别和分类问题^[1-2], 目前在分析某些医学图像方面已达到媲美放射科医师的水平^[3], 如检出肺结节^[4]、对胶质瘤进行级别分类^[5]和诊断骨折^[6]等, 对放射学的未来发展产生了深远影响。本文就 CNN 在骨骼肌肉放射学中的研究进展进行综述。

1 CNN 相关概念及发展

AI 是计算机科学的子领域, 是为计算机编程, 便于其学习和执行复杂任务的学科。机器学习(machine

learning, ML)是 AI 的一个子集, 通过计算机模拟人类的学习过程, 使计算机能够在数据中学习。传统 ML 算法虽可随着执行任务积累经验而逐步改善, 但对于图像问题通常仍需人类专家手动进行特征选择, 进而确定哪些图像特征更为重要。深度学习(deep learning, DL)是 ML 新发展的子集^[7], 其最大优点在于计算机不再需要人类专家手动进行特征选择, 算法本身可以确定最佳图像特征, 进而回答针对图像的各种问题。

CNN 是实现图像 DL 最常用的算法模型, 由输入层、卷积层、池化层、全连接层及输出层组成^[8], 经输入

[基金项目] 国家自然科学基金(81871326)。

[第一作者] 王春杰(1996—), 男, 山东临沂人, 在读硕士。研究方向: 骨肌影像学。E-mail: chunjiewang@pku.edu.cn

[通信作者] 袁慧书, 北京大学第三医院放射科, 100191。E-mail: huishuy@bjmu.edu.cn

[收稿日期] 2019-12-30 **[修回日期]** 2020-08-18

层输入数据,由卷积层和池化层对输入数据进行特征提取,全连接层将提取的特征进行组合分类,最后由输出层输出分类结果。

对 CNN 的研究可追溯至 20 世纪 90 年代。1989 年 LECUN 在论述中首次使用“卷积”一词对算法结构进行解释,其构建的 LeNet-5 模型定义了 CNN 的基本结构,成功用于解决识别手写数字的问题。2012 年 KRIZHEVSKY 等^[9]使用 CNN 模型获得了 ImageNet 大规模视觉识别挑战赛的胜利。此后,快速发展的图形处理器和多个开源库^[10]的建立使得 CNN 取得了突破性进展,涌现出一些 CNN 经典模型,如 VGG16/VGG19^[11]、GoogLeNet^[12]、ResNet^[13]和 U-Net^[14]等。伴随硬件设备更新和数据样本不断扩增,CNN 模型的图像识别与分类精度不断提高。时至今日,CNN 图像分析结果代表图像分析的最高水平^[15]。

2 CNN 在骨骼肌肉放射学中的应用

目前 CNN 在骨骼肌肉放射学中的应用主要包括骨骼肌肉病变的检出、分级和分割解剖结构三个方面。

2.1 检出病变 2017 年 RAJPURKAR 等^[16]在包含 36 808 幅图像的骨骼肌肉影像(musculoskeletal radiographs, MURA)数据集上训练 CNN 模型 DenseNet,以检测包括肩部、肱骨、肘部、前臂、腕、手掌和手指在内的上肢骨折。针对 556 幅图像的测试集,DenseNet 模型的 AUC 为 0.92,其检测腕部骨折的能力与放射科医师相当,二者检测结果与诊断金标准之间的一致性较高(Kappa 系数均为 0.93);而对于上肢骨折,该模型的总体表现比放射科医师略差,二者与诊断金标准之间的 Kappa 系数分别为 0.71 与 0.78。该研究将大型数据集 MURA 公开发布,以推动骨折自动检测领域的发展。截至 2019 年 12 月,排名前 8 的 CNN 模型 Kappa 系数介于 0.80~0.84,均优于原始研究中放射科医师的表现(Kappa 系数 0.78)。同样,在检测脊柱和下肢骨折方面,CNN 模型也展现出可以比拟放射科医师的诊断能力。

TOMITA 等^[17]应用 CNN 模型检测椎体压缩性骨折,其准确率为 89%。PRANATA 等^[18]以基于 CNN 的 ResNet 模型在 CT 图像中自动检测跟骨骨折,诊断准确率高达 98%。CNN 亦可用于腕部骨折^[19-20]。2019 年 CHENG 等^[20]在 25 505 幅四肢 X 线片上对 CNN 模型进行预训练,而后在 3 605 幅骨盆正位 X 线片上进一步训练,最后对 100 例腕部骨折患者进行测试,结果显示其在 X 线片上检测腕部骨折的准确率为 91%,AUC 为 0.98。BIEN 等^[21]在包含 1 130

幅图像的训练数据集上训练基于 CNN 的 MRNet 模型,以检测半月板撕裂,并在 120 幅图像上进行测试,得到该模型的 AUC 为 0.85。2019 年 ROBLOT 等^[22]及 LSSAUA 等^[23]应用 CNN 模型检测半月板撕裂,其 AUC 分别为 0.90、0.91。LIU 等^[24]报道了基于 CNN 的软骨损伤自动检测模型,使用 660 幅图像组成的训练集和 1 320 幅图像组成的测试集,该模型检测软骨损伤的 AUC 高达 0.92。2019 年 PEDOIA 等^[25]建立了基于 CNN 的 DenseNet 模型,用以检测软骨异常并诊断骨关节炎,其 AUC 为 0.83。

CNN 用于脊柱退行性病变领域包括自动检测椎间隙狭窄、椎管狭窄及椎体滑脱,准确率分别达 75%、94%及 95%^[26]。此外,基于 CNN 的骨龄自动检测在 DL 领域已有较多研究^[27-31]。北美放射学会儿童骨龄 ML 挑战赛共收入 105 种竞赛模型,根据骨龄估计值与真实参考值之间的绝对差值的平均值对提交模型进行排序,其中 BILBILY 和 CICERO 开发的 CNN 模型最终以平均绝对差值为 4.27 个月取得挑战赛冠军^[29]。YUNE 等^[32]报道了一种基于 CNN 的 VGG16 模型,可通过手掌和手腕部 X 线片评估性别,准确率可达 96%;类激活图显示,该模型主要通过集中于桡骨远端、第二和第三掌骨基底部及第三掌指关节的特征实现性别区分。

2.2 病变分级 ANTONY 等^[33]利用 CNN 模型创建了膝关节骨关节炎自动量化系统,将膝关节骨关节炎分为轻度、中度和重度。2018 年 TIULPIN 等^[34]使用 Siamese CNN 模型在多中心骨关节炎研究(multicenter osteoarthritis study, MOST)数据集共 18 376 幅图像上进行训练,并在骨关节炎创始(osteoarthritis initiative, OAI)数据集 5 960 幅图像上进行测试,结果显示,根据骨关节炎凯尔格伦-劳伦斯(Kellgren-Lawrence, K-L)分级,CNN 模型的 AUC 为 0.93,提示该模型具有学习相关骨关节炎特征的能力。上述报道中同样包括了类激活图,可显示模型做出决策的依据主要来自于哪些区域,使得模型的自动决策过程具有一定透明性。

JAMALUDIN 等^[35]开发了用于退变椎间盘自动分级的 CNN 模型,模型使用包含 12 018 幅图像的数据集,根据 Pfirrmann 分级标准,准确率可达 70%。LU 等^[36]报道了一种用于椎间孔狭窄和椎管狭窄自动分级的 ResNeXt-50 模型,使用 22 796 幅图片作为训练集,对椎间孔狭窄分级的平均准确率达 78%,对椎管狭窄分级的平均准确率可达 80%。

2.3 分割解剖结构 在 2018 年国际医学图像计算和计算机辅助干预协会 (medical image computing and computer assisted intervention society, MICCAI) 椎间盘分割挑战赛中, GEORGIEV 等^[37] 使用二维集成算法的 CNN 分割模型以戴斯相似系数 (Dice similarity coefficient, DSC) 0.91 获得挑战赛冠军。在此基础上, 2019 年 DOLZ 等^[38] 使用 MICCAI 提供的公开数据集开发并训练了 CNN 模型 IVD-Net, 利用多模态图像信息进行椎间盘分割, 最佳 DSC 可达 0.92。

2018 年 LIU 等^[39] 开发了一种基于 CNN 模型 SegNet 的膝关节骨与软骨自动分割方法, 利用膝关节影像分割 2010 (segmentation of knee images 2010, SKI10) 数据集进行训练和验证, 获得的标准化 SKI10 评分为 64.1 分, 表明其性能优于 U-Net 分割模型 (SKI10 评分为 53.0 分)。

3 小结与展望

CNN 是一种具有优秀图像分析能力的 AI 算法^[40], 在骨骼肌肉放射学中主要用于检测病变、进行分级和分割解剖结构。目前 CNN 应用研究尚仅限于部分器官和组织, 未能用于全身影像学检查, 但过去数年中的多项研究结果表明, CNN 在评估医学图像某些方面已具有专家级表现。随着研究的不断深入, 相信在不久的将来, CNN 将变革放射科现行工作模式, 对放射学发展产生划时代的影响。

[参考文献]

- [1] CHOY G, KHALILZADEH O, MICHALSKI M, et al. Current applications and future impact of machine learning in radiology [J]. *Radiology*, 2018, 288(2):318-328.
- [2] SOFFER S, BEN-COHEN A, SHIMON O, et al. Convolutional neural networks for radiologic images: A radiologist's guide [J]. *Radiology*, 2019, 290(3):590-606.
- [3] LIU X, FAES L, KALE A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis [J]. *Lancet Digital Health*, 2019, 1(6):271-297.
- [4] 侍新, 谢世朋, 李海波. 基于卷积神经网络检测肺结节 [J]. *中国医学影像技术*, 2018, 34(6):934-939.
- [5] 刘征华. 基于深度学习卷积神经网络的胶质瘤级别分类研究 [J]. *实用放射学杂志*, 2020, 36(7):1015-1018.
- [6] GAN K, XU D, LIN Y, 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.
- [7] CHARTRAND G, CHENG P M, VORONTSOV E, et al. Deep learning: A primer for radiologists [J]. *Radiographics*, 2017, 37(7):2113-2131.
- [8] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. *Nature*, 2015, 521(7553):436-444.
- [9] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C]. *Proc of 25th International Conference on Neural Information Processing Systems*, Nevada: Curran Associates Inc, 2012:1097-1105.
- [10] ABADI M, AGARWAL A, BARHAM P, et al. TensorFlow: Large-scale machine learning on heterogeneous distributed systems [J/OL]. [2016-03-16]. <https://arxiv.org/pdf/1603.04467.pdf>.
- [11] SIMONYAN K, ZISSCRMAN A. Very deep convolutional networks for large-scale image recognition [J/OL]. [2015-04-10]. <https://arxiv.org/pdf/1409.1556.pdf>.
- [12] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [J/OL]. [2014-04-17]. <https://arxiv.org/pdf/1409.4842.pdf>.
- [13] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [J/OL]. [2015-12-10]. <https://arxiv.org/pdf/1512.03385.pdf>.
- [14] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [J/OL]. [2015-05-18]. <https://arxiv.org/pdf/1505.04597.pdf>.
- [15] LITJENS G, KOOI T, BEJNORDI B E, et al. A survey on deep learning in medical image analysis [J]. *Med Image Anal*, 2017, 42:60-88.
- [16] RAJPURKAR P, IRVIN J, BAGUL A, et al. MURA: Large dataset for abnormality detection in musculoskeletal radiographs [J/OL]. [2018-05-22]. <https://arxiv.org/pdf/1712.06957.pdf>.
- [17] TOMITA N, CHEUNG Y Y, HASSANPOUR S. Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans [J]. *Comput Biol Med*, 2018, 98:8-15.
- [18] PRANATA Y D, WANG K C, WANG J C, 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] 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.
- [20] CHENG C T, HO T Y, LEE T Y, 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.
- [21] BIEN N, RAJPURKAR P, BALL R L, et al. Deep-learning-assisted diagnosis for knee magnetic resonance imaging: Development and retrospective validation of MRNet [J]. *PLoS Med*, 2018, 15(11):e1002699.

- [22] ROBLLOT V, GIRET Y, ANTOUN M B, et al. Artificial intelligence to diagnose meniscus tears on MRI[J]. *Diagn Interv Imaging*, 2019, 100(4):243-249.
- [23] LSSAUA N, ESTIENNE T, de VONNECOURT P, et al. Five simultaneous artificial intelligence data challenges on ultrasound, CT, and MRI [J]. *Diagn Interv Imaging*, 2019, 100 (4): 199-209.
- [24] LIU F, ZHOU Z Y, 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.
- [25] PEDOIA V, LEE J, NORMAN B, et al. Diagnosing osteoarthritis from T2 maps using deep learning: An analysis of the entire osteoarthritis initiative baseline cohort [J]. *Osteoarthritis Cartilage*, 2019, 27(7):1002-1010.
- [26] JAMALUDIN A, KADIR T, ZISSERMAN A. SpineNet: Automated classification and evidence visualization in spinal MRIs[J]. *Med Image Anal*, 2017, 41:63-73.
- [27] 王永灿,胡勇,申妍燕,等.基于卷积神经网络的骨龄阶段识别研究[J]. *图像与信号处理*, 2018, 7(1):1-15.
- [28] LARSON D B, CHEN M C, LUNGREN M P, 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.
- [29] HALABI S S, PREVEDELLO L M, KALPATHY-CRAMER J, et al. The RSNA pediatric bone age machine learning challenge[J]. *Radiology*, 2019, 290(2):498-503.
- [30] IGLOVIKOV V, RAKHLIN A, KALININ A A, et al. Pediatric bone age assessment using deep convolutional neural networks [J/OL]. [2018-06-19]. <https://arxiv.org/pdf/1712.05053.pdf>.
- [31] TAJMIR S H, 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.
- [32] YUNE S, LEE H, KIM M, et al. Beyond human perception: Sexual dimorphism in hand and wrist radiographs is discernible by a deep learning model[J]. *J Digit Imaging*, 2019, 32(4):665-671.
- [33] ANTONY J, MCGUINNESS K, O'CONNOR N E, et al. Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks [J/OL]. [2016-09-08]. <https://arxiv.org/pdf/1609.02469v1.pdf>.
- [34] 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.
- [35] JAMALUDIN A, LOOTUS 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(5):1374-1383.
- [36] LU J T, PEDEMONTE S, BIZZO B, et al. DeepSPINE: Automated lumbar vertebral segmentation, disc-level designation, and spinal stenosis grading using deep learning[J/OL]. [2018-07-26]. <https://arxiv.org/pdf/1807.10215.pdf>.
- [37] GEORGIEV N, ASENOV A. Automatic segmentation of lumbar spine MRI using ensemble of 2D algorithms [J/OL]. [2019-03-14]. https://link.springer.com/content/pdf/10.1007%2F978-3-030-13736-6_13.pdf.
- [38] DOLZ J, DESROSIERS C, AYED I B. IVD-Net: Intervertebral disc localization and segmentation in MRI with a multi-modal UNet [J/OL]. [2018-11-09]. <https://arxiv.org/pdf/1811.08305.pdf>.
- [39] LIU F, ZHOU Z Y, 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.
- [40] 侍新,谢世朋,李海波.基于卷积神经网络检测肺结节[J]. *中国医学影像技术*, 2018, 34(6):934-939.