

人工智能医学影像应用:现实与挑战

金征宇

【关键词】 人工智能; 计算机辅助诊断; 深度学习; 医学影像**【中图分类号】** R05; R395 **【文献标识码】** A **【文章编号】** 1000-0313(2018)10-0989-03

DOI:10.13609/j.cnki.1000-0313.2018.10.001

开放科学(资源服务)标识码(OSID):



伴随着人工智能的方兴未艾,医学影像与人工智能的结合被认为是最具发展前景的领域^[1]。医学影像人工智能肇始于上世纪 60 年代,上世纪 80 年代,随着人工神经网络和基于人工智能的计算机辅助诊断(CAD)软件的开放应用,人工智能开始逐渐整合到放射科日常工作流程中。进入 21 世纪以来,随着人工智能技术的飞速发展,人工智能在医学影像中的应用日新月异,在肿瘤检出、定性诊断、自动结构化报告、肿瘤提取、肿瘤放疗靶器官勾画等方面已经有较多的临床研究和临床应用^[2]。

由于医学影像临床工作的复杂性,伴随计算机技术的发展,直到近期人工智能技术才能进行精准的医学影像图像分析^[3]。目前学术界普遍接受的观点是人工智能应用于医学影像日常工作中,可以减少放射科医生的重复简单工作并降低人为错误,提高医生的工作效率,提高诊断准确率,促进精准医疗在影像医学的应用,但即使人工智能技术能达到更高的技术水平并能控制成本应用于临床工作,人工智能也不能取代放射科医生的全部临床工作,尤其是需要与人沟通交流的相关工作^[4]。

在探讨如何将人工智能应用于医学影像之前,要先了解在无人工智能的情况下,医学影像所面临的 3 个问题。第一,医疗数据中超过 90% 的数据来自于医学影像,但这些数据几乎全部需要人工分析。人工分析的缺点显而易见:①不够精确,医生仅能凭借经验去辨别,经常缺乏量化的标准,容易造成误判;②不可避免地会出现人眼视力产生的误差及视力疲劳;③海量的图像信息量容易产生漏诊。第二,医疗从业人员短缺。目前我国医学影像数据年增长率约为 30%,而放射科医师数量年增长率约为 4%,其间的差距为 26%^[5]。放射科医师数量增长远不及影像数据增长,且医师从业需要较长时间的培训和学习,这意味着放射科医师在未来处理影像数据的压力会越来越大,难以承担巨大的负荷。

第三,在多学科相互渗透交叉的时代,放射科医师缺乏强有力的方法或武器参与竞争以稳定学科发展的方向。而人工智能与医学影像的融合恰好可在学科发展方面拾遗补缺。

人工智能在医学影像的应用主要分为图像识别和深度学习两个部分。这两部分均是基于医学影像大数据所进行的数据上的挖掘和应用,其中深度学习是人工智能应用的核心环节。如果这两部分均得以实现将改写医学影像学在学科发展中的方向。

人工智能在医学影像应用最成熟的领域为肿瘤影像,其中以下三个方面已较广泛应用于医学影像诊疗临床工作中:

1. 肺部结节和肺癌筛查

2016 年进行的一项大规模研究,在 50 台 CT 扫描设备上使用了 4 种不同的 CAD 软件用以检出放射科医生漏诊的肺部结节,发现 CAD 软件可以检出漏诊的 56%~70% 结节,其中包括 17% 的 3mm 以下肿瘤和 69%~78% 的 3~6mm 肿瘤,这种大小的肿瘤经常被医生漏诊^[6]。使用更加先进的人工智能数据处理技术如多视点卷积网络可进一步降低 CAD 筛查肺部结节的假阳性率^[7]。如果不同放射科医生对于肺部结节性质有争议时,CAD 软件的诊断意见有助于结节性质的判定。肺部结节和肺癌筛查的人工智能 CAD 软件可以帮助放射科医生准确检出早期小肿瘤,降低医生工作强度和人为错误的发生率^[8]。CAD 软件在低剂量 CT 肺部结节筛查中可显著提高实习医师的诊断敏感性,以共同读片模式可以明显缩短实习医师的阅片时间^[9]。这表明人工智能技术在肺部结节 CT 筛查中可以切实应用于低年资医师教学和临床诊断工作中。

2. 乳腺癌筛查

乳腺癌影像筛查是人工智能机器学习较早应用的领域,目前 CAD 已经较好的融入放射科医生对乳腺癌进行影像诊断的日常工作流程中,在 X 线断层成像^[10]、乳腺钼靶^[11]、超声^[12]、MRI^[13]等不同检查方法中,基于人工智能的 CAD 筛查乳腺结节、诊断乳腺癌的准确性均较高。

作者单位: 100730 北京,中国医学科学院/北京协和医学院/北京协和医院放射科

作者简介: 金征宇(1960-),男,江苏南通人,主任医师,教授,博士生导师,主要从事临床医学影像诊断及介入治疗工作。

乳腺 CAD 目前被广泛应用在 X 线摄影对乳腺癌的筛查上,相关研究主要集中在提高钙化灶和肿块检出的准确性方面^[14],其中 X 线摄影对微钙化灶的检出率较高,而对肿块的检出率则会受到腺体密度的影响^[15]。最新开发的自然语言处理软件算法可准确获得了 543 例乳腺癌患者乳腺 X 线摄影的关键特征,并与乳腺癌亚型进行了关联,其诊断速度是普通医师的 30 倍,且准确率达 99%^[16]。CAD 系统可以辅助乳腺 MRI 的视觉评估并提供有用的附加信息。研究发现 CAD 系统对于 MRI 评估浸润性乳腺癌对新辅助化疗的反应具有高特异度,可达 100%,但由于敏感度较低,仅为 52.4%,因此尚不能取代视觉成像评估^[17]。CAD 系统对于 MRI 评估浸润性乳腺癌的多灶性具有明显优势,但对评估淋巴结的转移状态效果不佳^[18]。

3. 前列腺癌影像诊断

多参数磁共振在前列腺癌的影像检出、定位和分期方面发挥着重要作用,但是前列腺多参数磁共振成像技术较为复杂,序列较多,结果判读时间较长,难度较大。研究发现基于多模态卷积神经网络的人工智能 CAD 技术,在多参数磁共振中可以较准确的自动诊断前列腺癌^[19]。使用 CAD 软件也可以准确进行前列腺区域自动分割和肿瘤体积测定^[20]。

此外,在肿瘤影像诊断临床工作中,基于人工智能的机器学习技术,在肿瘤患者影像学检查方法和流程制定、影像成像、自动解析影像和结构化报告、图像质量分析、检查放射剂量预估等方面,也有相关的研究报道,在今后的肿瘤影像实际临床工作中可能会起到积极的作用^[21]。需要放射科肿瘤影像医生注意的是随着人工智能技术的发展以及和临床数据的紧密结合,肿瘤影像数据结合临床数据在人工智能辅助下转换成临床决策是今后肿瘤临床诊疗路径中的重要发展方向^[22]。

随着人工智能在医学影像领域应用的进一步拓展,在非肿瘤影像诊断的实际临床工作中整合了越来越多的人工智能相关产品。研究发现 AI 成像优化技术联合图像后处理迭代算法技术可在主动脉 CTA 成像中保证图像质量并明显降低辐射剂量^[23]。人工智能技术辅助冠脉 CT 血管成像对冠心病的诊断效能可以达到高年资放射科医生水平^[24]。这些研究进一步拓展了人工智能在医学影像的实际应用,具有很好的临床和科研应用前景。

现阶段,人工智能在医学领域的探索可谓五花八门,但最终应用到临床却微乎其微。主要原因之一是人工智能系统在复杂的临床应用中不能准确搜集到高质量可应用的数据即有效数据,所得出的结论缺乏可靠性,难以保持测试数据集上的高准确率。其次是临

床医学数据的搜集和预处理不够完善,未将医生的工作流程纳入考虑范围。医生对于疾病的诊断很重要的一点是依靠科学的思维和临床经验。医生的思维模式难以复制,而医生基于临床诊断作出的处理决定是融合了科学基础和人文关怀的综合考量。但作为医生的有力助手,人工智能结合医学影像具有诸多优势,患者、放射科医师、医院均能从人工智能的应用中受益。人工智能不仅能帮助患者更快速地完成健康检查包括 X 线、超声、磁共振成像等,同时也可以帮助影像医生提升读片效率,降低误诊概率,并通过提示可能的副作用来辅助诊断。随着人工智能和医学影像大数据在医学影像领域的普及和应用,医学影像所面临的诊断准确性和医生缺口等问题便可迎刃而解,两者的融合将成为医学影像发展的重要方向。

在医学影像领域,人工智能技术的介入结合大数据挖掘,使得医学影像大数据在人工智能的筛选、梳理和提取后,可能转换成有效的临床决策^[25]。2017 年 7 月 8 日国务院印发的《新一代人工智能发展规划》提出了我国研发人机协同临床智能诊疗方案的计划。医学影像技术的发展、人工智能技术的进步和医疗大数据的不断积累促使智能医疗的发展进入新时代。

参考文献:

- [1] Kahn CE Jr. From images to actions: opportunities for artificial intelligence in radiology[J]. Radiology, 2017, 285(3): 719-720.
- [2] Gyawali B. Does global oncology need artificial intelligence? [J]. Lancet Oncol, 2018, 19(5): 599-600.
- [3] Siegel E. Artificial intelligence and diagnostic radiology: not quite ready to welcome our computer overlords[J]. Appl Radiol, 2012, 41(4): 8.
- [4] Amato F, López A, Peña-Méndez EM, et al. Artificial neural networks in medical diagnosis[J]. J Appl Biomed, 2013, 11(2): 47-58.
- [5] 2018-2024 年中国人工智能+医疗影像行业市场研究及投资前景预测报告[R]. <http://www.chyxx.com/research/201710/578114.html>
- [6] Liang M, Tang W, Xu DM, et al. Low-dose CT screening for lung cancer: computer-aided detection of missed lung cancers[J]. Radiology, 2016, 281(1): 279-288.
- [7] Setio AA, Ciompi F, Litjens G, et al. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks[J]. IEEE Trans Med Imaging, 2016, 35(5): 1160-1169.
- [8] Ciompi F, Chung K, van Riel SJ, et al. Towards automatic pulmonary nodule management in lung cancer screening with deep learning[J]. Sci Rep, 2017, 7: 46479.
- [9] 胡琼洁, 陈冲, 王玉锦, 等. 实习医师以共同阅片和第二阅片者模式使用计算机辅助检测系统在低剂量 CT 中的应用研究[J]. 放射学实践, 2018, 33(10): 1022-1028.
- [10] Samala RK, Chan HP, Hadjiiski L, et al. Mass detection in digital breast tomosynthesis: deep convolutional neural network with

- transfer learning from mammography[J]. Med Phys, 2016, 43(12):6654-6666.
- [11] Huynh BQ, Li H, Giger ML. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks[J]. J Med Imaging (Bellingham), 2016, 3(3):034501.
- [12] Gu P, Lee WM, Roubidoux MA, et al. Automated 3D ultrasound image segmentation to aid breast cancer image interpretation[J]. Ultrasonics, 2016, 65:51-58.
- [13] Bickelhaupt S, Paech D, Kickingereder P, et al. Prediction of malignancy by a radiomic signature from contrast agent-free diffusion MRI in suspicious breast lesions found on screening mammography[J]. J Magn Reson Imaging, 2017, 46(2):604-616.
- [14] 李欣, 梁森, 黄正南, 等. 乳腺 X 线 AI 智能病灶检测[J]. 放射学实践, 2018, 33(10):1029-1032.
- [15] Mahersia H, Boulehmi H, Hamrouni K. Development of intelligent systems based on Bayesian regularization network and neuro-fuzzy models for mass detection in mammograms; a comparative analysis[J]. Comput Methods Programs Biomed, 2016, 126:46-62.
- [16] Patel TA, Puppala M, Ogunti RO, et al. Correlating mammographic and pathologic findings in clinical decision support using natural language processing and data mining methods[J]. Cancer, 2017, 123(1):114-121.
- [17] Böttcher J, Renz DM, Zahm DM, et al. Response to neoadjuvant treatment of invasive ductal breast carcinomas including outcome evaluation; MRI analysis by an automatic CAD system in comparison to visual evaluation[J]. Acta Oncol, 2014, 53(6):759-768.
- [18] Song SE, Seo BK, Cho KR, et al. Computer-aided detection (CAD) system for breast MRI in assessment of local tumor extent, nodal status, and multifocality of invasive breast cancers; preliminary study[J]. Cancer Imaging, 2015, 15:1.
- [19] Le MH, Chen J, Wang L, et al. Automated diagnosis of prostate cancer in multi-parametric MRI based on multimodal convolutional neural networks[J]. Phys Med Biol, 2017, 62(16):6497-6514.
- [20] Turkbey B, Fotin SV, Huang RJ, et al. Fully automated prostate segmentation on MRI; comparison with manual segmentation methods and specimen volumes[J]. AJR, 2013, 201(5):W720-W729.
- [21] 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.
- [22] 许晶晶, 谭延斌, 张敏鸣. 影像学在肿瘤精准医疗时代的机遇和挑战[J]. 浙江大学学报(医学版), 2017, 46(5):455-461.
- [23] 王明, 王怡宁, 于敏, 等. AI 成像优化联合迭代算法在“双低”主动脉 CTA 的初步应用[J]. 放射学实践, 2018, 33(10):1009-1016.
- [24] 黄增发, 王翔. 人工智能冠状动脉 CT 血管成像在冠心病诊断中的应用[J]. 放射学实践, 2018, 33(10):1017-1021.
- [25] Robertson SP, Quon H, Kiess AP, et al. A data-mining framework for large scale analysis of dose-outcome relationships in a database of irradiated head and neck cancer patients[J]. Med Phys, 2015, 42(7):4329-4337.

(收稿日期:2018-10-12)

· 人工智能与医学影像专题 ·

客观看待人工智能在医学影像中的作用

萧毅, 刘士远

【摘要】 人工智能(AI)近年来发展迅猛,对各行各业产生了深远影响。近年来 AI 在医疗行业的应用有了众多的进展, AI 技术确实到了一个能够在医疗上逐步实现应用的阶段,但是也不能过度神话。现在的 AI 技术在医学的应用还处于初级阶段,距离临床生根发芽,开花结果还有很长一段路要走。正确认识、客观对待,才能促进医学影像 AI 的持续健康发展。

【关键词】 人工智能; 医学影像; 深度学习

【中图分类号】 R05; R395 **【文献标识码】** A **【文章编号】** 1000-0313(2018)10-0992-04

DOI:10.13609/j.cnki.1000-0313.2018.10.002

开放科学(资源服务)标识码(OSID):



人工智能(artificial intelligence, AI)近年来发展迅猛,从 IBM Watson 在一个知识类的电视节目上击败人类冠军,到 Alpha Go 击败围棋世界冠军,现在几乎各行各业都快速进入了一个 AI 为王的时代。人工智能在医疗的应用更是涉及了医学影像、疾病风险预测、药物研发、健康管理、医疗机器人等等领域。

人工智能的概念和层次

1956 年在著名的达特茅斯会议上首次提出“人工智能”这一概念,标志着这门新兴学科的正式诞生^[1]。人工智能是研究计算机模拟人的某些思维过程和智能行为的学科,涉及计算机科学、心理学、哲学和语言学等,已成为一门广泛交叉的新兴前沿学科^[2-3]。经过近 70 年的多次热潮和寒冬,随着算法、算力和大数据的发展,人工智能尤其是机器学习的算法迅速发展,其中深度学习作为机器学习中的一个领域发展迅猛。卷积神经网络(convolutional neural network, CNN)或循环神经网络(recurrent neural network, RNN),结合现在强大的计算能力和大数据训练出来的模型,在某些领域达到甚至超越了人类专家的水平,让人惊叹不已^[4]。

人工智能按照实现的能力可以分为 3 个层次:①弱人工智能(artificial narrow intelligence, ANI),擅长于单个方面或任务的人工智能;②强人工智能(artificial general intelligence, AGI),是指在各方面都能模仿人类甚至和人类比肩的人工智能;③超人工智能

(artificial superintelligence, ASI)是指牛津哲学家及知名人工智能思想家 Nick Bostrom 定义的“在几乎所有领域都比最聪明的人类大脑都聪明很多的人工智能,包括科学创新、通识和社交技能”^[5]。由此看来,目前 AI 技术比如下棋、人机对话、人脸识别、医疗影像识别,都属于“弱”人工智能,即便有些 AI 程序或者机器人在某个方面超越人类非常多,比如下围棋,它也只是执行一个闭环的任务,本身并不能像人一样全方位的思考,也没有意识。

目前人工智能在医疗领域的进展

深度学习、计算机视觉、自然语言理解及大数据存储等关键技术的突破带来了人工智能技术新一轮的发展,而这些技术的发展已经推动了以数据、知识、脑力劳动密集为特征的医疗产业与人工智能的深度融合。随着人们生活条件的改善,人们对于身体健康的意识也越发强烈;与此同时,人们面临医疗资源分配不均,药物研制周期长、费用高,医务人员培养成本高等问题,都给 AI 的发展带来了社会需求和机遇。

AI 可以快速且不间断的读取海量文献,拥有最全面的基础知识和最新进展,结合“学习”专家医生的医疗知识,模拟医生的思维和诊断推理,从而可以给出可靠诊断和治疗方案建议^[6]。智能诊疗场景也给 AI 在智能健康管理上的应用提供了可能,可以实现风险识别、虚拟护士、精神健康、在线问诊、健康干预以及基于各种医疗数据的健康管理等^[7]。

AI 技术在计算机视觉上的发展,尤其是深度神经网络的应用,使 AI 在视觉领域达到了新的高度^[8]。有的 AI 产品可以不需通过数据和逻辑得出结论,只需根据患者的面容、心率、呼吸等状况迅速反应出疾病的类型,这种能力已经很接近人类医生^[9]。

计算机视觉基于图像识别,可以将医学影像进行

作者单位: 200003 上海,海军军医大学长征医院影像科
作者简介: 萧毅(1971—),女,博士,副主任医师,副教授,主要从事心胸影像诊断、影像新技术精确成像与术前评估以及疾病早期筛查与人工智能研究工作。

通讯作者: 刘士远, E-mail: liushiyuan@smmu.edu.cn

基金项目: 上海市科学技术委员会基金(17411952400);国家重点研发计划政府间合作(2016YFE0103000);上海市卫计委智慧医疗(2018ZHYL0101);科技部国家重点研发计划(2018YFC0116404)

深入分析,获取有意义的信息。通过数据量的日积月累,能力指数级提升,有望为精准诊断提供极大的帮助^[10]。

2016 年,《JAMA》杂志上发表了一篇美国加州大学的 Gulshan 团队的文章^[11],描述了人工智能在 10 万多张视网膜眼底照片中进行糖尿病视网膜病变的诊断,并将其诊断结果与 54 位持有医生执照的眼科医师、资深住院医师进行比较,最终结果是 AI 的敏感性 & 特异性均高于人类的诊断成绩单。

2017 年,荷兰内梅亨大学 Bejnorid 教授的团队在《JAMA》杂志发表的相关研究表明^[12],在评估乳腺癌患者腋窝淋巴结有无少量癌细胞扩散方面,人工智能的表现要优于病理医师用显微镜的常规做法,尽管其结果还不能完全代替病理学家,但已经大大提高了诊断效果。

2018 年,发表在《Lancet Oncology》上的一项研究首次证实,人工智能可以处理医学图像以提取生物学和临床信息^[13],通过设计算法并将其开发用于分析 CT 扫描图像,研究者们创建了一组放射学特征,该特征定义了肿瘤的淋巴细胞浸润水平,并提供了患者免疫治疗功效的预测评分。

国内近两年 AI 研究如火如荼,各个器官疾病诊断的模型都有研究和尝试,基于深度学习的 AI 应用目前已经覆盖病灶检测、病理诊断、放疗规划和术后预测等各临床阶段^[14-16]。其中基于 X 线的肺部筛查、乳腺钼靶筛查^[17-18];基于 CT 影像的肺结节检测模型^[14]显示出较好的临床使用潜力。冠心病的智能化诊断模型也已经在一些医院落地^[19]。甚至部分已经成为临床工作中医生不可分割的好帮手。以肺结节为例,AI 算法模型能自动分割胸腔区域,快速准确定位疑似肺结节的病灶,从大数据集学习所得到的算法模型可以避免主观偏差,虽然部分模型筛选的结果中包含了一些假阳性结节,但明显降低了假阴性的发生,大大减轻了影像科医师的工作量^[10]。除此之外,AI 算法模型不仅能提取肺结节的位置形态信息,还能进一步提供肺结节分类(实性、亚实性及钙化等)乃至肿瘤良恶性分级等一些决策意见供医师参考。

医疗行业需要人工智能,但不能过度神话

2018 年,美国食品药品监督管理局(food and drug administration,FDA)已经批准了世界人工智能医疗设备 IDx-DR,该设备可以在没有医生帮助的情况下诊断疾病。IDx-DR 自此成为了第一个获得 FDA 批准进行独立检测而不需要医生解释结果的人工智能设备。IDx-DR 类 AI 产品的批准通过,意味着不仅仅是让普通人可以更方便自主的完成常规诊断,同时还

极大的节约了社会成本,因此很多业内人士都认为,AI 诊断爆炸时代已然来临。这同样也是 AI 应用的一个重要的里程碑。

2017 年 8 月 31 日,国家食品药品监督管理总局(CFDA)发布了新版《医疗器械分类目录》,新《分类目录》对原来的 43 个子目录做了大量精简、优化,调整为 22 个,其中尤为引人注目的就是新的分类目录特别增加了与人工智能辅助诊断相对应的类别^[20]。相关法规、准入制度的更改,表明在深度学习越来越风靡的高科技行业中,AI 在辅助医生进行诊断与拟定治疗方案等方面的应用日渐普及。可以预见的是无论是国内还是国外,未来的几年内将会有大量 AI 诊断系统通过审批,越来越多 AI 医疗应用在国内外落地。

AI 技术确实到了一个能够在医疗上实现应用的阶段,而且表现可圈可点,但是也不能过度神话。目前,有些宣传为了吸引眼球进行夸大,貌似 AI 无所不能,很快可以替代人类医生。但现在的 AI 技术在医学的应用还处于非常初级阶段,远远没有达到能自主学习分析的强人工智能水平。夸大的宣传只会误导政府和公众,透支行业发展潜力,不利于 AI 的有序健康发展。

现阶段,人工智能更多是围绕某单一影像任务研发解决方案,在单一任务方面人工智能有明显优势。但诊断/分类也仅仅是人工智能医学影像应用场景中的一小部分,医学影像临床工作交织着多种不同任务。除了医学影像的分类(诊断)上,研究者还需继续探索人工智能技术在医学影像检测(发现异常)、分割(量化测量)、配准(随访跟踪)中的应用。人工智能也不止应用于影像分析,深度神经网络能就不同模态的信息(如文本、语音)产出有价值的应用成果,应该对医院全数据链进行完整分析,并嵌入到医疗流程中,这样才能最终生产出更加有意义的人工智能模型。

现有模型的鲁棒性和有效性也还有待于更加客观的检验。各个企业用于训练的数据集质量参差不齐,导致产品性能不稳定,大范围的推广受到一定的制约,需要进一步提高产品普适性。各个企业的数据模型的诊断效果多为自我评价,具体的评估过程是未知的,评估指标不统一,需要第三方公正平台对各个数据模型进行评价,保证产品质量。

此外,值得大家深思的是人工智能是目的,实现这一目的的技术有很多种,深度学习只是其中一种技术。现在深度学习或者深度神经网络的快速发展俨然成了人工智能的化身,能够解决一切的问题。在医学影像领域中有很多种类的问题和场景,并不是所有的场景需要用深度学习的技术去实现,也并不是所有的场景都适合用深度学习技术去实现,我们应该注重的是目

标,为了目标问题去使用技术,使用各种最适合的技术。切忌为了深度学习而深度学习,为了蹭热度舍本逐末地消耗人力物力去做可以用简单技术或者传统技术解决的问题。

人工智能在医疗中应该扮演的角色

AI 技术可以帮助医生做大量重复性的劳动,提高工作效率,改变医生的工作模式,提供更加精准的预后判断。哈佛大学医学院病理系教授 Golden 博士说:“人工智能可能正是病理医师所期待的。医疗卫生行业出现人工智能、大数据的成本下降、可以获得数字化图像,这些都在助力数字病理的成功。就像人工智能之前的电镜、免疫组化、分子诊断一样,病理医师被取代的风险极低。但其工作模式可能会改变,病理医师对患者诊疗的贡献仍将十分关键^[21]”。

让机器先实现人类医生的专业水平,从而帮助医生去做一些重复性工作,让医生有更多的精力解决更多疑难问题,积累更多经验,而新的经验和知识又会传递给 AI,以此来形成一个良性的循环和生态。所以在当下,AI 的一个重要作用并不是解决疑难杂症,解决高精尖问题,而是在临床最需要的领域,用机器来把好的医生经验复制到医疗资源缺乏的地区去。AI 技术可以为“基础薄弱、人才缺失”的基层地区提高诊断率、进而缓解三甲医院的就诊压力。

人工智能技术还应该能够提供肉眼难以识别的疾病特征,使得医生能够使用 AI 成像来识别位于身体任何部位的肿瘤的生物现象,而无需进行活组织检查。还可以在疾病诊断、预后判断、疗效评估、用药指南、疾病随访等方面发挥重要的作用,真正融入到医疗的每一个环节,人机协作,共同为人类的健康做出应有的贡献。

总之,人工智能已进入一个快速发展期,给医学影像带来可见的变化,改变着我们的工作效率和流程。虽然目前仍处于弱人工智能阶段,不能为医生分担大部分工作,但伴随着技术的不断进步,弱人工智能过渡到强人工智能时代是值得期待的。

参考文献:

- [1] McCarthy J, Minsky ML, Rochester N, et al. A proposal for the dartmouth summer research project on artificial intelligence[J]. J Molecular Biology, 2006, 278(1): 279-289.
- [2] Nilsson NJ. Artificial intelligence: a new synthesis[M]. San Francisco: Morgan Kaufmann Publishers Inc, 1998: 37-48.
- [3] 熊瑶, 陈敏. 人工智能在医疗领域应用现状探讨[J]. 医学信息学杂志, 2018, 39(4): 24-28.
- [4] He K, Zhang X, Ren S, et al. Delving deep into rectifiers: surpassing Human-level performance on image net classification[J]. arXiv, 2015: 1502. 01852.

- [5] Pohl J. Artificial superintelligence: extinction or nirvana? [C]. California, Intersymp-International Conference on Systems Research, Informatics and Cybernetics, 2015: 107-127.
- [6] Litjens G, Sánchez CI, Timofeeva N, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis [J]. Sci Rep, 2016, 6: 26286.
- [7] Griffin J, Treanor D. Digital pathology in clinical use: where are we now and what is holding us back? [J]. Histopathology, 2017, 70(1): 134-145.
- [8] Lecun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436.
- [9] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[J]. arXiv, 2015: 1512. 03385.
- [10] JJha S, Topol EJ. Adapting to artificial intelligence: radiologists and pathologists as information specialists[J]. Jama, 2016, 316(22): 2353-2354.
- [11] Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs[J]. Jama, 2016, 316(22): 2402-2410.
- [12] Ehteshami BB, Veta M, Johannes VDP, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer [J]. Jama, 2017, 318(22): 2199-2210.
- [13] Roger Sun, Elaine Johanna Limkin, Maria Vakalopoulou, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study[J]. Lancet Oncology, 2018, 19(9): 1180-1191.
- [14] Ciompi F, Chung K, Riel SJV, et al. Towards automatic pulmonary nodule management in lung cancer screening with deep learning[J]. Sci Rep, 2017, 7: 46479.
- [15] Wang H, Zhou Z, Li Y, et al. Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from ¹⁸F-FDG PET/CT images[J]. Ejnmmi Research, 2017, 7(1): 11-20.
- [16] Rahul P, Samuel HH, et al. Deep feature transfer learning in combination with traditional features predicts survival among patients with lung adenocarcinoma[J]. Tomography A Journal for Imaging Research, 2016, 2(4): 388-395.
- [17] Rajpurkar P, Irvin J, Zhu K, et al. Che X-Net: radiologist-level pneumonia detection on chest X-rays with deep learning[J]. arXiv, 2017: 1711. 05225.
- [18] Mohamed AA, Berg WA, Peng H. A deep learning method for classifying mammographic breast density categories [J]. Med Phys, 2018, 45(1): 314-321.
- [19] Slomka PJ, Dey D, Sitek A, et al. Cardiac imaging: working towards fully-automated machine analysis & interpretation [J]. Expert Rev Med Devices, 2017, 14(3): 197-212.
- [20] 麻良, 黄军斌, 王菲, 等. CFDA2017 新《医疗器械分类目录》浅析 [J]. 中国医学装备, 2017, 14(10): 130-133.
- [21] Golden JA. Deep learning algorithms for detection of lymph node metastases from breast cancer: helping artificial intelligence be seen[J]. Jama, 2017, 318(22): 2184-2186.

· 人工智能与医学影像专题 ·

医学影像和人工智能:了解历史,找准方向,稳步前进

吴山东, 夏黎明

【关键词】 医学影像; 人工智能; 评论

【中图分类号】R05; R395 【文献标识码】A 【文章编号】1000-0313(2018)10-0995-02

DOI:10.13609/j.cnki.1000-0313.2018.10.003

开放科学(资源服务)标识码(OSID):



人工智能(artificial intelligence, AI)目前在健康领域不断渗透,特别在医学影像方面显现出一种改变当前临床诊断流程的潜力。人工智能不是新概念,但是 2012 年开始变得火热的深度学习技术的出现,给人工智能的概念输入了新的燃料,相关的研究和应用落地都得到了极大的推动。学术界的研究现在基本都在向深度学习聚焦,而传统的机器学习迅速降温。工业界出现了一大批 AI 新创公司,积极探索不同国情下的 AI 临床应用场景和落地。由于深度学习的技术门槛并不高,因此也出现了一窝蜂的现象,大大小小的公司都在寻找成功的机会。在表面繁荣火爆的情况下,人工智能在医学影像领域的发展却充满很多挑战,真正的落地并不可能一蹴而就,医疗领域对于新技术的接受也相对缓慢。在开始的激情和热情慢慢褪却后,尊重新技术发展的规律,充分了解实际的临床需求和痛点,找到医学思维和计算技术的学科交叉路径,医生和工程技术人员深度协作,踏实前进,才有机会真正把这波人工智能的潜力转化成生产力,切实提高当前医学影像疾病诊断的精度、读片的效率、图像中信息解读的深度、以及影像在引领整个疾病管理诊治流程中的巨大作用。

本期杂志刊发了一些 AI 在医学影像方面文章,期待从不同的国情、不同的角度、不同的体会等方面给读者就人工智能在医学影像方面的发展提供讨论素材,扩大认知领域,交流实际经验和遇到的问题,碰撞火花,助力 AI 在医学影像相关方面的进一步研究和应用。在檀韬等的文章中,作者就人工智能在医学影像方面发展的历史进行了梳理,这其中最重要的基础是基于计算机的辅助检测和诊断(computer-aided detection/diagnosis, CAD)。CAD 在欧美已研究发展了几十年,一批有物理或计算背景的学者在医疗领域

内倡导和发展了这一概念,并有实际的产品在临床部署。但是客观上 CAD 的临床使用并没有得到期待的效果,这一度使得不少人怀疑 CAD 可达到多高的程度,相关研究也几乎处在瓶颈阶段。深度学习的出现无疑给 CAD 提供了最新的燃料,不少有深度技术背景的专业人士开始憧憬先进的深度学习在技术层面提升传统 CAD 临床效果的前景。这也是目前 AI 应用于医学影像领域的助推力。虽然目前很多学者已经不再纠结于 CAD 的概念,但是了解 CAD 的历史和瓶颈,对目前 AI 技术开发和临床落地是有很大的积极借鉴作用的。当然, CAD 或者 AI 的应用绝对不是单纯的技术问题,至少在可预见的短期内不是,只要有人(也就是医生)的参与,人机的有机协同作业带来的挑战更大,因为人的学习、决策和思维过程很多时候是定性的也很难建模的。在同一篇文章中,作者也探讨了 AI 在手术机器人方面的技术发展和应用场景。人工智能对专家医生的手术经验的学习,某种程度上类似于对放射科医生的读片经验的学习,但是由于手术场景带来的特殊性,这其中蕴含的困难实际上更多。另外,檀韬等作者文中对欧洲 AI 方面的独特视点,也非常有参考价值。

吕乐等的文章着重探讨了 AI 和医学影像在预防医学中的机会。这个视点更多是从美国的医疗情况出发。疾病筛查(如乳腺癌)在美国有国家层面的指南和推动,也形成了事实上的大量人群的筛查体系。筛查过程中产生的大量数据,会切实促进深度学习中数据驱动的学习这一理念,同时创造利用大数据分析的方法发现新的基于影像的标志物。吕乐等在文中更进一步提出了一些 AI 实际应用的场景,这些场景超越了计算机辅助检测和诊断的基本范畴,对整个疾病管理流程中的影像 AI 的角色指出了更多的可能性,也代表了对基于影像数据的精准医学的重要解读。吕乐等更在文中阐明了一些当前的研究难点和机会,比如数据共享、标注、精准分割、建模和可视化等问题,也包括对非结构化数据的使用等问题,特别值得研究人员关注和思考。

中国目前在医学影像 AI 方面的研究、应用以及

作者单位:匹兹堡,宾夕法尼亚州,美国匹兹堡大学医学院放射科(系)和生物医学信息系;工程学院生物工程系;计算和信息学院智能系统系;临床和转化研究院(吴山东)/430030 武汉,华中科技大学同济医学院附属同济医院放射科(夏黎明)
作者简介:吴山东——男,山东人,计算机科学、医学影像双博士,匹兹堡大学临床影像智能计算实验室主任,主要从事计算机和人工智能技术、医学影像转化研究和临床应用。

工业界的积极参与程度方面非常突出,在某些方面处于领先地位。但是当中的过热、跟风、急功近利、不切实际的期望、不够深入的行业理解、对 CAD 或者 AI 不够清晰的角色定位等种种因素,处理不好会对 AI 的发展起到反作用。当前某些 AI 辅助工具的确在临床测试上表现了让人印象深刻的效果,这也是 AI 给人以期待的原由。但是即便如此,这样的 AI 还是在初级甚至比较粗糙的阶段,真正满足于临床诊断要求 AI 的临床研究目前还不到位,没有大规模、正规的临床测试研究的科学数据,以及第三方的客观数据的检验,还很难下确定的结论。即便是有了较为成熟的产品,产品的健壮性、普适性以及临床监管等,都需要大量而艰巨的工作来检验、规范,才能无缝融入临床工作流程并发挥作用。涉及到 AI 和临床医学的共生和交互问题。因此,来自临床一线医生的理性声音,对于 AI 的健康发展至关重要。可赞的是目前相当一部分医生也开始关注人工智能甚至深度学习技术,医院和

工业界的合作更是比比皆是。学科的交叉带来进步,但也要防止一方由于对非自身专业知识的不够了解而产生的冒进甚至错误判断。针对医学影像的分析而言,AI 可以做什么,不能做什么,以及深度学习的本质和优缺点的理解是对行业研究人员、医生以及工业界的共同要求,也是尊重 AI 新技术发展的理性思考。

人工智能如一道曙光,目前吸引了众多行业人员的追随。医学是不同于其它行业的特殊学科,根本目的是对人的关怀。技术的发展应遵循规律,AI 造福人类健康,需要以符合医学本质的角度切入临床流程。目前的 AI 医学影像发展,需要激情,也需要理性,更需要踏踏实实、坚持不懈地把工作做严谨。还有很多问题值得在以后深入探讨和展开。我们也希望共同思考,共同进步,把 AI 医学影像诊断领域的发展推向纵深。

(收稿日期:2018-10-10)

《请您诊断》栏目征文启事

《请您诊断》是本刊 2007 年新开辟的栏目,该栏目以临床上少见或容易误诊的病例为素材,杂志在刊载答案的同时配发专家点评,以帮助影像医生更好地理解相关影像知识,提高诊断水平。栏目开办 8 年来受到广大读者欢迎。《请您诊断》栏目荣获第八届湖北精品医学期刊“特色栏目奖”。

本栏目欢迎广大读者踊跃投稿,并积极参与《请您诊断》有奖活动,稿件一经采用稿酬从优。

《请您诊断》来稿格式要求:①来稿分两部分刊出,第一部分为病例资料和图片;第二部分为全文,即病例完整资料(包括病例资料、影像学表现、图片及详细图片说明、讨论等);②来稿应提供详细的病例资料,包括病史、体检资料、影像学检查及实验室检查资料;③来稿应提供具有典型性、代表性的图片,包括横向图片(X 线、CT 或 MRI 等不同检查方法得到的影像资料,或某一检查方法的详细图片,如 CT 平扫和增强扫描图片)和纵向图片(同一患者在治疗前后的动态影像资料,最好附上病理图片),每帧图片均需详细的图片说明,包括扫描参数、序列、征象等,病变部位请用箭头标明。

具体格式要求请参见本刊(一个完整病例的第一部分请参见本刊正文首页,第二部分请参见 2 个月后的杂志最后一页,如第一部分问题在 1 期杂志正文首页,第二部分答案则在 3 期杂志正文末页)

栏目主持:石鹤 联系电话:027-69378385 15926283035

· 人工智能与医学影像专题 ·

基于纹理分析和深度学习的肝纤维化研究进展

李秋菊, 郭启勇, 陈海斌, 张荣国

【摘要】 肝纤维化的早期诊断和治疗, 对于防治肝硬化、肝癌有重要价值, 肝纤维化的非创伤性影像学诊断一直以来都是肝纤维化研究的重点。近年来, 针对目前影像学检查包含的大量多模态图像信息, 利用影像组学和深度学习等人工智能方法, 通过分析图像中的人眼不能直接识别的海量图像信息实现辅助临床诊断受到社会各界的广泛关注。目前纹理分析和深度学习等方法在肝纤维化无创性诊断方面已经取得有效的临床进展。

【关键词】 肝纤维化; 纹理分析; 深度学习; 人工智能

【中图分类号】 R05; R395; R575.2 **【文献标识码】** A **【文章编号】** 1000-0313(2018)10-0997-05

DOI: 10.13609/j.cnki.1000-0313.2018.10.004

开放科学(资源服务)标识码(OSID):



肝纤维化是指在不同病因的慢性肝脏疾病作用下, 肝脏中胶原蛋白等细胞外基质增生与降解失去平衡, 导致肝脏内纤维结缔组织异常沉积的病理过程。肝纤维化的病因有很多, 临床上常见的有病毒性肝炎、酒精肝、脂肪肝以及自身免疫性疾病、胆汁淤积等。我国以乙型和丙型肝炎病毒导致的肝纤维化占大部分比例。肝纤维化是慢性肝病向肝硬化、肝癌进展的重要中间环节^[1]。大量临床研究表明由于体内存在纤维降解过程, 肝纤维化进程是可以过去除病因、临床干预治疗等得到延缓、甚至逆转的^[2]。因此, 肝纤维化的早期诊断和治疗, 对于防治肝硬化、肝癌有重要价值。

目前肝组织活检仍被作为临床上诊断肝纤维化分期的金标准, 但由于肝脏穿刺标本体积较小, 存在标本量不足或抽样误差; 另外病理组织切片存在观察者内和观察者间的差异所导致肝脏分期不准确^[3]。尤其是肝脏穿刺活检属创伤性检查, 存在出血、胆汁瘤、胆汁性腹膜炎、败血症和菌血症等严重并发症的风险, 使其应用价值受到限制。非创伤性的评价肝纤维化分期的方法成为世界肝脏病学研究的重点。

肝纤维化的非创伤性影像学诊断发展迅速。目前临床应用较多的影像学检查方法有 CT、超声及磁共振检查, 但目前这些临床常规方法对肝纤维化诊断价值有限, 对于肝硬化凭借临床积累经验可以做出诊断, 但对轻、中度肝纤维化的诊断仍缺乏可靠依据。因而, 在肝纤维化影像诊断方法探讨的过程中, 大量研究尝试了很多新的成像技术及科研序列, 但大部分影像检

查方法和技术在早期肝纤维化诊断方面存在着各自的优缺点。基于血流灌注的超声、CT 及 MR 检查能够定量评估血流灌注情况, 但对于早期肝纤维化诊断价值有限^[4-9]。基于弹性成像的超声弹性成像 (ultrasound elastography, UE) 及磁共振弹性成像 (magnetic resonance elastography, MRE) 对肝纤维化分期敏感性和特异性相对其他影像手段较高^[10-13], 结果稳定可信, 但超声弹性成像测量时不能实时显示肝脏二维声像图, 无法避开肝内大血管、胆道系统, 从而影响测量结果, 另外肋骨、肺、腹水、皮下脂肪厚度等也会影响结果的准确性。MRE 与 T₁ 或 T₂ 序列结合, 可以克服解剖定位问题, 但价格昂贵、需要额外的刺激器、MR 机型限制、肝脏内过高的铁含量会导致信噪比受限等特点使其广泛使用受限。基于水分子运动的磁共振弥散加权成像 (diffusion weighted imaging, DWI) 随着肝纤维化程度加重, 表观弥散系数 (apparent diffusion coefficient, ADC) 值降低, 可将轻度肝纤维化与进展期肝硬化进行区分, 但对肝纤维化的早期诊断价值有限^[14]。另外, 新型对比剂钆塞酸二钠增强 MR 检查、磁共振波谱成像、磁共振血氧水平依赖成像、MR 弥散张量成像等检查方法对中重度肝纤维化有一定诊断价值, 但对早期肝纤维化的诊断能力目前仍不能满足临床需要, 并且相对难以实现快速临床转化。

近年来, 基于医学图像的影像组学和深度学习等人工智能方法受到社会各界的广泛关注。针对目前影像学检查包含的多模态的大量图像信息, 大量研究通过纹理分析、深度学习等方法, 分析图像中可能隐藏的海量人眼不能识别的图像信息来辅助临床诊断, 已经取得较好的临床进展。下面将从纹理分析 (texture analysis) 及深度学习 (deep learning) 等方面介绍近年来肝纤维化诊断在人工智能诊断方面的研究进展。

作者单位: 110004 辽宁, 中国医科大学附属盛京医院放射科 (李秋菊、郭启勇); 100000 北京, 推想科技有限公司 (陈海斌、张荣国)

作者简介: 李秋菊 (1985-), 女, 山东菏泽人, 博士研究生, 住院医师, 主要从事肝纤维化影像诊断工作。

通讯作者: 郭启勇, E-mail: guoqy@sj-hospital.org

基金项目: 国家自然科学基金资助 (81471718/81771893)

基于纹理分析的肝纤维化诊断研究

影像组学是指采用高通量、全自动化计算方法将医学图像中病灶感兴趣区域的影像数据转化为具有特征空间的量化数据,并对特征空间区数据进行精准定量分析,以从结构形态、组织细胞、分子和基因遗传等各种表型综合评价疾病,实现精准医学目标。图像纹理特征提取是影像组学研究的重要组成部分,目前影像组学提取的纹理特征包括病灶的几何特征、灰度直方图统计特征、Hu 不变特征、Tammura 纹理特征、灰度共生矩阵(gray level co-occurrence matrix, GLCM)、灰度游程步长矩阵(gray level run-length matrix, GLRLM)以及病灶具有的频域纹理特征。目前大量应用于疾病识别和诊断的纹理分析软件是通过勾画感兴趣,提取几十甚至几百个纹理特征,对比病理等病变诊断标准,经过统计学分析,筛选出对病变诊断贡献相对较大的纹理特征,进而建立疾病诊断和预测模型。目前,图像纹理分析在肝纤维化诊断方面的应用日渐成熟,针对目前现有影像手段均可进行纹理分析。

1. 肝纤维化 CT 图像纹理分析

首先,从形态学角度出发,Hori 等^[15]探讨肝脏形状并定量分析,构建肝脏统计形状模型,发现其对肝纤维化分期有重要价值。其次,在肝实质 CT 图像本身纹理方面,Lubner 等^[16]通过对 212 例肝纤维化患者及 77 例健康对照者进行纹理分析,发现随肝纤维化程度进展,平均灰度强度、熵逐渐增高,而峰度和偏斜度逐渐降低,对 $\geq F2$ 期肝纤维化诊断的受试者工作特性曲线下面积(area under the ROC curve, AUC)值达 0.71~0.73,诊断 F4 期的 AUC 值达 0.86~0.87。童隆正等^[17]也发现能量、熵、逆差矩等纹理特征在不同肝纤维化分期之间有显著性差异。理论上能量和熵值可用来描述区域内图像灰度变化的总体性质,对不同空间的图像纹理反应比较敏感。熵表征了图像空间的复杂性和混乱度,当肝脏发生纤维化改变时,由于结缔组织增生,肝实质不规则变化,图像纹理变粗糙,因而熵升高。能量表征纹理的全局均一度,随肝纤维化进展,肝脏均匀性下降,能量升高。而逆差矩则由于肝纤维化后结缔组织增生而增大,这是由于肝表面颜色变暗,与其周围背景的差别减小,反差特征参量值相对较小,对比度降低。另外,肝纤维化进程中结缔组织不断增生,波动程度增强,方差增大,相关性降低。因而上述这些纹理特征可以用来评价肝纤维化分期。学者还在统计学方法方面进行了大量研究和改进,如利用布朗分形模型和小波变换两种方法来分析正常和纤维化的肝脏 CT 图像,提高肝纤维化诊断效能^[18]。Mala

等^[19]对肝脏 CT 图像提取基于小波的统计纹理特征,应用神经网络的方法,对脂肪肝和肝硬化的分类准确率高达 95%,敏感度(sensitivity)和特异度(specificity)分别为 96%和 94%。另外,考虑到 CT 平扫图像可能包含的图像信息相对较少,Daginawala 等^[20]应用增强 CT 图像,采用更多元化的方法及模型分析,在提取的 41 个纹理特征中,7 个直方图特征、1 个灰度共生矩阵参数、6 个灰度游程步长矩阵参数、1 个法则特征和 4 个灰度梯度矩阵参数共 19 个特征在区分纤维化分组方面存在显著的统计学差异,诊断的 AUC 值达 0.70~0.80。

2. 肝纤维化 MR 图像纹理分析

在形态学方面,Sudhakar 等^[21]评估了重度肝纤维化和肝硬化的肝实质结构、表面结节、肝脏体积变化和门静脉高压症等特征,其中表面结节特征诊断重度纤维化准确度为 80.3%,肝实质纹理诊断肝硬化的准确度为 80.5%。关于形态学及形态定量分析的研究相对较少,大部分研究集中在对 MR 图像中肝实质的纹理特征进行提取及筛选。Yu 等^[22]将 11.7T 磁共振成像扫描仪获取的肝纤维化大鼠 T_1 和 T_2 图像进行纹理分析,提取基于直方图、灰度共生矩阵、灰度游程步长矩阵、灰度梯度矩阵的纹理特征,发现 T_1 图像在直方图四分位数间距和灰度梯度矩阵的方差梯度两个特征方面有非常好甚至优良的分类能力,灰度梯度矩阵中峰度是鉴别肝纤维化分期的最佳纹理特征,AUC 值达 0.90。而对质子密度图像的纹理分析显示灰度共生矩阵相关性和对比度两个特征与肝纤维化分期具有中度至强度相关性^[23]。而在肝实质图像纹理特点的基础上,增加年龄、肝脏脂肪和肝脏 R2 变量等特征,将这些特征纳入广义线性模型,可以大大改善肝纤维化分期诊断效能^[24]。与应用 CT 增强图像进行纹理分析相似,应用双对比剂(超顺磁性氧化铁和 Gd-DTPA)增强 MR 也能增加其诊断效能^[25],Yokoo 等发现应用双对比增强 MR 图像进行纹理分析,诊断 $\geq F2$ 期肝纤维化的 AUC 为 0.89,敏感度和特异度分别为 90%和 78%,诊断肝硬化时 AUC 为 0.98,敏感度和特异度分别高达 100%和 93%。

大量研究验证了 MR 图像纹理分析在肝纤维化分期方面的潜在价值。在此基础上综合肝脏形态学、患者临床信息以及肝脏 MR 增强图像的模型可能对肝纤维化诊断更有帮助。另外,在既往肝纤维化无创性影像诊断研究的过程中,开发了大量特殊 MR 科研序列,如 DWI、MRE 等,由于其图像本身诊断效能相对较好,对于这些灰度图像,也可以通过纹理分析提升其诊断效能。Barry 等^[26]对大鼠 DWI 图像后处理获得的 ADC 图进行纹理分析,提取基于直方图、灰度共

生矩阵和灰度游程步长的特征,验证了 DWI 纹理分析对肝纤维化分期的价值。

3. 肝纤维化超声图像纹理分析

由于超声图像的获取受到临床经验、增益补偿、体位、患者个体差异等多种因素的影响,导致肝实质区域纹理的对比度、光照、角度和旋转等变化较大,也受到多种噪声等干扰因素的影响,因而现阶段基于整块肝实质区域进行肝纤维化分期的纹理分析研究相对较少。首先肝脏图像分割、轮廓识别就是一大难题,部分研究利用贝叶斯分类器不断改进、优化,最终对肝脏超声图像的识别率达到 90%^[27,28]。目前,超声纹理分析对肝硬化识别研究较多,识别准确率能够达到 86%以上^[29,30],而对早期肝纤维化的研究相对较少。陈明丽等^[31]采用 SonoCT 和 XRes 成像技术采集 186 例有肝穿刺病理结果的乙型肝炎患者的标准声像图像,提取声像图纹理的 14 个灰度共生矩阵参数,发现研究诊断早期肝硬化的敏感度高达 93.5%。目前,超声纹理分析对肝纤维化的无创分期研究也逐渐成为研究的热点。

基于深度学习的肝纤维化研究

人工智能在医疗领域应用最为广泛的就是机器学习。机器学习是通过模拟人类认知的过程,通过手动提取目标特征或利用深度学习的方法自动提取特征,从而实现目标的分类和识别问题。其中机器学习以支持向量机^[32]、决策树^[33]、随机森林^[34]等经典的分类器闻名。

Shousha 等^[35]回顾性研究了 427 例慢性丙型肝炎患者,使用数据挖掘分析通过减少误差技术构建决策树,然后使用 Auto-WEKA 工具从 39 种算法中选择最佳分类器来预测晚期纤维化,选择多层感知器神经网络模型作为最佳预测算法,敏感度为 82.5%,特异度为 81.1%,AUC 为 0.880。

深度学习是机器学习的一种,最早是由 Hinton 等^[36]在 2006 年提出,可以使用一种称为贪婪逐层预训练的策略来有效的训练。其中以卷积神经网络(convolutional neural network, CNN)为代表的深度学习^[37]更是以其强大的自动学习和提取特征的能力而受到极大的关注。深度学习的优势在于它把特征抽取和分类器整合到一个框架下,实现了网络框架对数据特征的自主学习,减少了人为干预和人工选取特征的成本。目前基于多模态肝纤维化图像的深度学习,逐渐成为无创性肝纤维化诊断的热点。

Yasaka 等^[38]应用基于深度卷积神经网络模型应用乳酸增强肝胆期磁共振图像诊断肝纤维化,发现深度学习肝纤维化评分与病理分期有显著相关性,其诊

断 S4、S3、S2 期肝纤维化的 AUC 分别为 0.84、0.84 和 0.85。二维剪切波弹性成像(2D shear wave elastography, 2D-SWE)对乙型和丙型肝炎患者肝纤维化的无创分期也具有良好的性能^[39,40]。Wang 等^[41]进行了包含 654 名慢性乙型肝炎患者的前瞻性多中心研究,对其中 398 名患者进行了二维剪切波弹性成像进行深度学习,以评估其诊断肝纤维化分期的准确性,并与 2D-SWE、血清标记物诊断能力进行比较。发现深度学习对肝硬化、 \geq F3 期、 \geq F2 的慢性肝炎患者, AUC 分别为 0.97、0.98、0.85,其诊断效能除了对 \geq F2 期略低于 2D-SWE 外,其余均显著优于 2D-SWE 和血清标记物。另外随着训练学习的图像数量增加,其诊断准确性得到改善。

基于机器学习的肝纤维化预后预测

肝纤维化准确分期的临床意义在于能够指导临床及时、准确地干预肝纤维化的进展进程。对来自 20 个队列的 5817 名丙型肝炎患者的 29 项研究进行荟萃分析,发现随访 2~15 年期间,肝纤维化进展、肝功能失代偿患者甚至死亡比例非常大,由于慢性肝炎的病程呈非线性进展,现有的慢性肝炎疾病进展风险预测模型的准确性有限,因而开发有效的预测模型也是肝纤维化研究的重点内容。Konerman 等^[42,43]总结发现脂肪变性和纤维化基线评分对肝纤维化程度进展的预测价值较大,而血小板计数、谷草和谷丙转氨酶比率、白蛋白、胆红素和年龄等因素则是肝脏失代偿有意义的预测因子。另外,针对肝硬化的丙型肝炎抗病毒长期治疗的患者,应用包含纵向数据的新颖统计方法来改进现有基线模型,使用逻辑回归和随机森林、Boosting 两种机器学习方法构建模型以预测病程进展的结果,发现机器学习方法可以捕捉预测变量和结果之间的复杂关系,从而产生更准确的预测。其他研究者也验证了机器学习在肝纤维化预测方面的重要价值^[44-46]。

小结和展望

在多模态成像方法中,磁共振弹性成像、超声弹性成像及弥散加权成像等有相对较好的肝纤维化诊断价值,但目前仍因为多方面限制,没有广泛应用于临床。常规 CT、MR 及超声对轻中度肝纤维化诊断仍有局限性。而这些多模态的图像中可能隐藏大量人眼不能识别的图像信息,因而我们可以通过影像组学、人工智能等手段提取重要的特征,构建辅助临床诊断的计算机软件,实现无创性肝纤维化分期预测,甚至评估患者的预后,为临床选择最佳治疗时机提供指导。

参考文献:

[1] Afdhal NH, Nunes D. Evaluation of liver fibrosis: a concise review

- [J]. Am J Gastroenterol, 2004, 99(6): 1160-1174.
- [2] Friedman SL. Liver fibrosis-from bench to bedside[J]. J Hepatol, 2003, 38(suppl 1): 38-53.
- [3] Goldin RD, Goldin JG, Burt AD, et al. Intra-observer and inter-observer variation in the histopathological assessment of chronic viral hepatitis[J]. J Hepatol, 1996, 25(5): 649-654.
- [4] 梁明帮, 黎进莲, 安玉松, 等. CT 肝灌注成像在肝纤维化诊断中的价值及影像特点分析[J]. 中西医结合肝病杂志, 2015, 25(5): 295-297.
- [5] Aronhime S, Calcagno C, Jajamovich GH, et al. DCE-MRI of the liver: effect of linear and nonlinear conversions on hepatic perfusion quantification and reproducibility[J]. Magn Reson Imaging, 2014, 40(1): 90-98.
- [6] Ou HY, Bonekamp S, Bonekamp D, et al. MRI arterial enhancement fraction in hepatic fibrosis and cirrhosis[J]. Am J Roentgenol, 2013, 201(4): 596-602.
- [7] 解丽梅, 郭启勇, 刘鑫. 超声造影评价肝纤维化[J]. 中国医学影像技术, 2009, 25(8): 1447-1449.
- [8] Abbattista T, Ridolfi F, Ciabattoni E, et al. Diagnosis of liver cirrhosis by transit-time analysis at contrast-enhanced ultra sonography[J]. Radiol Med, 2008, 113(6): 860-874.
- [9] Liu H, Liu J, Zhang Y, et al. Contrast-enhanced ultrasound and computerized tomography perfusion imaging of a liver fibrosis early cirrhosis in dogs[J]. J Gastroenterol Hepatol, 2016, 31(9): 1604-1610.
- [10] Shi Y, Guo Q, Xia F, et al. MR elastography for the assessment of hepatic fibrosis in patients with chronic hepatitis B infection: does histologic necroinflammation influence the measurement of hepatic stiffness? [J]. Radiology, 2014, 273(1): 88-98.
- [11] Guo Y, Parthasarathy S, Goyal P, et al. Magnetic resonance elastography and acoustic radiation force impulse for staging hepatic fibrosis: a meta-analysis[J]. Abdom Radiol, 2015, 40(4): 818-834.
- [12] Shi Y, Xia F, Li QJ, et al. Magnetic resonance elastography for the evaluation of liver fibrosis in chronic hepatitis B and C by using both gradient-recalled echo and spin-echo echo planar imaging: a prospective study[J]. Am J Gastroenterol, 2016, 111(6): 823-833.
- [13] Chon YE, Choi EH, Song KJ, et al. Performance of transient elastography for the staging of liver fibrosis in patients with chronic hepatitis B: a meta-analysis[J]. PLoS One, 2012, 7(9): e44930.
- [14] Taouli B, Vilgrain V, Dumont E, et al. Evaluation of liver diffusion isotropy and characterization of focal hepatic lesions with two single-shot echo-planar MR imaging sequences: prospective study in 66 patients[J]. Radiology, 2003, 226(1): 71-78.
- [15] Hori M, Okada T, Higashiura K, et al. Quantitative imaging: quantification of liver shape on CT using the statistical shape model to evaluate hepatic fibrosis[J]. Acad Radiol, 2015, 22(3): 303-309.
- [16] Lubner MG, Malecki K, Kloeke J, et al. Texture analysis of the liver at MDCT for assessing hepatic fibrosis[J]. Abdom Radiol (NY), 2017, 42(8): 2069-2078.
- [17] 童隆正, 王磊, 陈海荣, 等. 肝纤维化图像的灰度共生矩阵分析[J]. 首都医科大学学报, 2003, 24(3): 240-242.
- [18] 童隆正, 陈海荣, 贺文. 肝纤维化 CT 图像的小波变换和布朗分形模型分析[J]. 北京生物医学工程, 2003, 22(2): 113-115.
- [19] Mala K, Sadasivam V, Alagappan S. Neural network based texture analysis of CT images for fatty and cirrhosis liver classification[J]. Applied Soft Computing, 2015, 32(1): 80-86.
- [20] Dagainawala N, Li B, Buch K, et al. Using texture analyses of contrast enhanced CT to assess hepatic fibrosis[J]. Eur J Radiol, 2016, 85(3): 511-517.
- [21] Sudhakar K, Venkatesh, Meng Yin, Naoki Takahashi, et al. Non-invasive detection of liver fibrosis: MR imaging features vs MR elastography[J]. Abdom Imaging, 2015, 40(4): 766-775.
- [22] Yu H, Touret AS, Li B, et al. Application of texture analysis on parametric T₁ and T₂ maps for detection of hepatic fibrosis[J]. J Magn Reson Imaging, 2017, 45(1): 250-259.
- [23] Yu H, Buch K, Li B, et al. Utility of texture analysis for quantifying hepatic fibrosis on proton density MRI[J]. J Magn Reson Imaging, 2015, 42(5): 1259-1265.
- [24] House MJ, Bangma SJ, Thomas M, et al. Texture-based classification of liver fibrosis using MRI[J]. J Magn Reson Imaging, 2015, 41(2): 322-328.
- [25] Yokoo T, Wolfson T, Iwaisako K, et al. Evaluation of liver fibrosis using texture analysis on combined-contrast-enhanced magnetic resonance images at 3.0T [J]. Biomed Res Int, 2015: 387653.
- [26] Barry B, Buch K, Soto JA, et al. Quantifying liver fibrosis through the application of texture analysis to diffusion weighted imaging[J]. Magn Reson Imaging, 2014, 32(1): 84-90.
- [27] Wu CM, Chen YC, Hsieh KS. Text features for classification of ultrasonic liver images[J]. IEEE Trans Med Imaging, 1992, 11(2): 145-146.
- [28] 陈菲. 基于 BP 神经网络的超声肝脏图像识别[J]. 网络与信息技术, 2007, 26(8): 43-44.
- [29] Gao S, Peng Y, Guo H, et al. Texture analysis and classification of ultrasound liver images[J]. Biomed Mater Eng, 2014, 24(1): 1209-1216.
- [30] Virmani J, Kumar V, Kalra N, et al. SVM-based characterization of liver ultrasound images using wavelet packet texture descriptors[J]. J Digit Imaging, 2013, 26(3): 530-543.
- [31] 陈明丽, 黄咏红, 王怡, 等. 超声纹理分析方法诊断乙型肝炎早期肝硬化的价值[J]. 肝脏, 2010, 15(6): 414-416.
- [32] Christopher JC, Burges. A tutorial on support vector machines for pattern recognition[J]. Data Mining and Knowledge Discovery, 1998, 2(2): 121-167.
- [33] Quinlan JR. Induction of decision trees[J]. Machine Learning, 1986, 1(1): 81-106.
- [34] Svetnik V, Liaw A, Tong C, et al. Random forest: a classification and regression tool for compound classification and QSAR modeling[J]. J Chemical Information and Computer Sciences, 2003, 43(6): 1947-1958.
- [35] Shousha HI, Awad AH, Omran DA, et al. Data mining and machine learning algorithms using IL28B genotype and biochemical markers best predicted advanced liver fibrosis in chronic hepatitis C[J]. Jpn J Infect Dis, 2018, 71(1): 51-57.
- [36] Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-

- 1554.
- [37] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553):436-444.
- [38] Yasaka K, Akai H, Kunimatsu A, et al. Liver fibrosis; deep convolutional neural network for staging by using gadoxetic acid-enhanced hepatobiliary phase MR images[J]. Radiology, 2018, 287(1):146-155.
- [39] Abe T, Kuroda H, Fujiwara Y, et al. Accuracy of 2D shear wave elastography in the diagnosis of liver fibrosis in patients with chronic hepatitis C[J]. J Clin Ultrasound, 2018, 46(5):319-327.
- [40] Herrmann E, De Lédinghen V, Cassinotto C, et al. Assessment of biopsy-proven liver fibrosis by two-dimensional shear wave elastography; an individual patient data-based meta-analysis [J]. Hepatology, 2018, 67(1):260-272.
- [41] Wang K, Lu X, Zhou H, et al. Deep learning radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B; a prospective multicentre study[J]. Gut, 2018, pii: gutjnl-2018-316204. Epub ahead of print.
- [42] Konerman MA, Yapali S, Lok AS. Systematic review: identifying patients with chronic hepatitis C in need of early treatment and intensive monitoring—predictors and predictive models of disease progression[J]. Aliment Pharmacol Ther, 2014, 40(8):863-879.
- [43] Konerman MA, Zhang Y, Zhu J, et al. Improvement of predictive models of risk of disease progression in chronic hepatitis C by incorporating longitudinal data[J]. Hepatology, 2015, 61(6):1832-1841.
- [44] Chen Y, Luo Y, Huang W, et al. Machine-learning-based classification of real-time tissue elastography for hepatic fibrosis in patients with chronic hepatitis B[J]. Comput Biol Med, 2017, 89:18-23.
- [45] Hashem S, Esmat G, Elakel W, et al. Comparison of machine learning approaches for prediction of advanced liver fibrosis in chronic hepatitis C patients[J]. IEEE/ACM Trans Comput Biol Bioinform, 2018, 15(3):861-868.
- [46] Eslam M, Hashem AM, Romero-Gomez M, et al. International liver disease genetics consortium (ILDGC). FibroGENE; a gene-based model for staging liver fibrosis[J]. J Hepatol, 2016, 64(2):390-398.

(收稿日期:2018-09-20)

关于开放科学标识(OSID)码告《放射学实践》杂志作者和读者书

《放射学实践》杂志自 2018 年 4 月起正式加入 OSID 开放科学计划。通过在杂志每篇论文上添加开放科学(资源服务)标识码(Open Science Identity, OSID),为读者和作者们提供增值服务。每一篇被纳入 OSID 开放科学计划的论文,将匹配一个专属的 OSID 识别码。此码就如同一个具有交互功能的论文“身份证”,给作者提供了更好地与业界同伴交流成果的途径。

OSID 码中包含以下 5 项内容:①作者介绍论文的语音(不超过 1 分钟);②作者与读者在线交流;③作者与读者互动交流精选问答合集;④作者本篇论文的读者圈;⑤论文附加说明(可选择上传论文相关图片或视频)。其中,作者介绍论文的语音是 OSID 识别码必须包含的内容;论文附加说明中,可上传论文相关图片或视频,这为弥补纸刊载体承载内容的局限性提供了一种有效途径。这 5 项互动内容,让作者的论文转换成知识工作者互动、交流的载体平台,使论文变得与众不同,从而提升论文的阅读量、下载量和引用率,并促进学术交流。上传论文的语音介绍,是一种传播作者学术成果的途径,能更好地展现作者的研究成果,提升作者的影响力和学术评价。

同时,我们会为每篇论文的作者开通一个 OSID 开放科学作者账号,并通过邮件告知作者。作者通过微信扫描邮件中的二维码并关注公众号“SAYS 管理平台”,上传对所著论文的 1 分钟语音介绍,以及附加说明(实验过程、推演数据、图像、视频等),完成本刊稿件录取、发表之前的最后一步。登陆作者账号后,作者即刻拥有所著论文的读者圈和问答,可与读者进行交流互动(读者只需在微信上扫 OSID 码,即可直接向作者提问或互动沟通)。

如您有任何疑问,请咨询工作人员 刘琦(电话:18062026009;微信/QQ:249115562)

董盈盈(电话:15623095186;QQ:2368705356;微信号:UED-Test1)

感谢您对本刊的支持,欢迎继续赐稿!

《放射学实践》杂志社

· 人工智能与医学影像专题 ·

医学影像诊断及介入式手术的人工智能应用

檀韬, 喻秉斌, 吴山东

【关键词】 人工智能; 医学影像; 计算机辅助诊断; 手术机器人

【中图分类号】 R05; R395 【文献标识码】 A 【文章编号】 1000-0313(2018)10-1002-04

DOI:10.13609/j.cnki.1000-0313.2018.10.005

开放科学(资源服务)标识码(OSID):



医学影像人工智能的萌芽

上个世纪 60 年代就有一些利用计算机来诊断骨肿瘤的研究应用。但是这些研究都不够成功。最早大规模的关于计算机处理医学图像的研究是上个世纪 80 年代,起源于芝加哥大学的 Kurt Rossmann 实验室。那时 Kunio Doi 教授带领这个实验室开始利用数字化的医学图像来提高诊断精度^[1]。这些课题涉及血管成像,肺癌 X 线的检测(Maryellen Giger)及乳腺 X 线钙化点的检测(Heang-Ping Chan)等。Maryellen Giger 教授后来成为芝加哥大学放射科计算机辅助检测/诊断(computer-aided detection/diagnosis, CAD)研究的领头人,而 Heang-Ping Chan 教授则在密歇根大学把 CAD 研究推向深入。这些计算机诊断的研究主要是模拟放射科医生诊断和检测的过程,同时研究为什么有些肿瘤会被漏诊。基于计算机的辅助诊断的核心是计算机自动找到可疑肿瘤区域或者计算机给指定的区域评出可疑程度(良恶性诊断)。

除了研究本身,芝加哥大学的研究者希望能把 CAD 转化在临床应用上,这是对研究成功最好的证明。因此芝加哥大学积极朝着计算机诊断的商业化努力。他们申请了大量有关计算机诊断的专利。最早的专利是 Doi 教授(第一发明人)提出的关于提高肿瘤在医学图像的自动检测。这个专利目前是 CAD 领域引用次数最多的专利。从上世纪 80 年代开始,芝加哥大学不断展示有关 CAD 的各项研究来激发全世界 CAD 研究。譬如在肺癌,他们试图去除 X 线上的肋骨和在 CT 上检测肿块;乳腺癌上他们试图检测钙化点和肿

瘤。在当时那个年代,研究者对 CAD 的概念是将 CAD 作为第二读片人,也就是说配合人类医生读片,并不是要取代医生;CAD 只是给医生一个建议或者是提醒医生一些可疑的区域。其最主要的原因是 CAD 本身的精度和准确率不够。

第一个成功的商业化医学图像人工智能公司以及产业效应

上个世纪 90 年代 Bob Wang, Wei Zhang, Bob Foley 受到之前学术研究成果的启发,成立 R2 科技公司,致力于世界上第一台商业化乳腺癌 X 线上 CAD 系统开发。R2 向芝加哥大学购买了乳腺 CAD 的专利,同时又向洛克希德马丁实验室购买了神经网络的版权,但 R2 的早期实际临床效果并不理想。R2 公司上个世纪 90 年代末和荷兰 Nijmegen 大学的 Nico Karssemeijer 教授合作,成功利用了 Spiculation 特征算法提升了 CAD 的块状癌症检测水平,并获得美国食品和药品管理局(food and drug administration, FDA)批准。由于针对乳腺癌的 CAD 商业上的成功,R2 又开发了世界上第一个商业化 CT 肺癌自动检测 CAD 系统,获得 FDA 批准。到 2000 年代 R2 累计销售 10 亿美元,并整个公司被 Hologic 收购。R2 可以说是 CAD 在商业上的成功实践。

由于 R2 的成功,世界范围出现了大大小小的 CAD 公司,主要是开发乳腺癌在钼靶、超声、核磁共振;肺癌在 X 线、CT;前列腺癌在核磁共振上的 CAD,或者是开发间接帮助医生的计算机应用,比如计算肿瘤的体积等。现在 CAD 方面的公司有如 iCAD、Parascrip、Mirada Medical 等。而在深度学习出现后,新兴起了一批医学影像人工智能(artificial intelligence, AI)创新型公司,在欧美有 Qview Medical、ScreenPoint Medical、Thirona 等。2016 年谷歌旗下的 DeepMind 在英国进军医疗行业,开始利用人工智能去检查和分析眼底图像,同一时期,国内的阿里、平安和腾讯等重量级大公司开始进入医疗影像领域,未来的竞争将日趋白热化。

作者单位:埃因霍温,荷兰埃因霍温理工大学生物医学工程系(檀韬);不来梅,德国人工智能研究所(喻秉斌);匹兹堡,宾夕法尼亚州,美国匹兹堡大学医学院放射科(系)和生物医学信息系;工程学院生物工程系;计算和信息学院智能系统系;临床和转化研究院(吴山东)

作者简介:檀韬——男,安徽人,助理教授,主要从事医学图像处理算法研究。

通讯作者:吴山东, E-mail: wus3@upmc.edu

基金项目:荷兰国家科学研发项目(629.001.003);荷兰 WBSO 经济部项目(2016YFE0103000);欧盟欧洲之星项目(9714);美国国家卫生研究院/美国国家癌症研究所项目(1R01CA193603, 1R01CA218405)

保险补贴影响了 CAD 公司对算法的开发

在整个 CAD 产业界发展中出现了一个很特殊的现象,以乳腺癌计算机检测为例,从 1994 年到 2002 年,计算机的假阳性个数从每个 X 线图像 10 个左右快速下降到 <1 个。这段时间算法是不断提高的。但是,到了 2004 年,假阳性个数就不再下降了。于此同时,在 2002 年左右,美国医疗保险公司开始补贴对计算机 CAD 系统的使用。这意味者这些 CAD 公司很容易间接地从保险公司手里拿到收入,实现盈亏平衡。由于这些公司已有的源源不断的收入,而提升传统的 CAD 算法又消耗大量的人力财力,因此 CAD 算法的提升有点停滞不前。从某种意义上说,整个工业界都在等待新的催化剂来促进行业的改变。

深度学习刺激了人工智能算法的开发

保险补贴后的 10 年,2012 年以后,由于深度卷积神经网络的兴起、大量数据的累积和基于图像处理器(graphics processing unit, GPU)的计算能力大幅提高,深度学习在医疗图像上的研究空前高涨^[2],而一些 CAD 系统的性能已经接近或者超过放射科医生。如荷兰 Nijmegen DIAG 研究组的 Bram van Ginneken 教授(曾经在 R2 工作)成功应用深度神经网络带领开发了肺癌 CAD,这套系统已经被荷兰飞利浦公司商业化。同时该组在 2017 年发表乳腺病理检测文章^[3]引起广泛关注。值得注意的是由于使用深度神经网络的技术门槛降低,全世界都在大规模应用神经网络。但是深度神经网络目前应用最好的是解决简单任务,比如分类、识别、切割等,关于综合性的诊断、还是需要具有领域知识的专家来参与。另外,需要指出的是卷积神经网络不是最近发明的,之前提到的 R2 公司成立者之一 Wei Zhang,在上个世纪 90 年代就开发了 Shift-Invariant 神经网络^[4]算法并被认为是和 LeCun Yann 同时期独立开发的最早的深度学习原型之一。因此可以说华人对早期深度学习的发展是有启发性贡献的。只不过在 20 年前,基于当时计算机计算能力的限制,卷积神经网络没有得到广泛应用。

人工智能在介入式医疗(手术机器人)的应用

人工智能除了对医学影像的分析,介入式的手术(手术机器人)也是其中一个重要的应用场景,它包括对专家医生手术经验和技巧的学习,手术机器人的技术分解和分析,手术流程的分析与学习以及在智能手术机器人自身的开发和应用。最近几年,人工智能在介入式医疗方面也越来越受到关注。事实上,计算机辅助以及自动化辅助介入式医疗也并不是一个新兴的

概念,从手术机器人诞生的初期开始,比如早期的手术机器人,Unimation Puma 200、ROBODOC、MINERVA 和 Cyberknife,自动化技术就已经被大量的使用。但是相比于传统的工业机器人,介入式手术机器人有着应用场景上的特殊性:人体的组织并不像工业产品一样方便操作,因其复杂的操作环境,柔软易变性的特性,以往的手术机器人非常依赖于对环境的感知(视觉伺服),以及对组织的物理建模。然而,复杂环境下的实时感知往往是比较困难和局限的,比如说应用视觉伺服能较好的解决手术路径的问题,但它并不能很好的控制机械和组织交互的力。而且物理建模依赖于模型的参数,而这些参数的确定往往比较费时和困难。与此相对的,人工智能的数据驱动方式能够直接从感知的数据中学习,它能避免非常复杂的物理和生物力学的建模并且有很好的适用性。

德国人工智能研究中心(DFKI)的研究人员根据应用的场景,对人工智能在介入式医疗的应用分为了“智能辅助医疗”和“智能手术机器人”两方面,并且对目前这两方面的进展做了调查^[5]。

在“智能辅助医疗”方面,研究者们针对手术辅助在 3 个方面尝试了使用人工智能:①学习专家的手术技术和知识;②分解和分析手术机器人的技巧;③分析手术流程。临床手术最关键的莫过于专家医生宝贵的手术经验和手术技巧。因此模仿学习作为监督学习的一种形式被广泛应用于学习模仿专家技巧上。它通过观察专家的手术动作的切换,提取观察到的专家的动作信息,从而学习和提升自身的动作模型和技巧。除此之外,隐行马可夫模型,神经网络和模糊网络等技术也被应用在了学习手术技巧上面。手术技术的分析和评估无论对于医生临床手术还是医疗机器人来说都非常重要。对于医疗机器人来说,评价的指标通常在于手术时间,机器的速度以及任务复杂度等位置信息。对比传统由专家进行技术评估,人工智能通过学习多个专家评估方式和指标,以及通过机器人多个传感器获得的多维度的信息,能够有效的对手术技术进行分级和评估。比如 Reiley 等通过矢量量化(一种无监督学习)和隐行马可夫模型对达芬奇手术机器人的手术数据进行了分析和评估,能够有效分析初级、中级学员和专家的技术区别。对手术流程的分解和分析也非常关键。比如在手术导航过程中,只有当手术流程被分解成为相互关联的多个子任务,手术机器人才能对各个子任务进行建模和控制。目前在手术流程分解方面,应用较多的人工智能方法是对运动学数据进行基于隐行马可夫模型的建模。

“智能手术机器人”作为人工智能在医疗领域上重要的硬件载体,其目的是通过传感器获得多维度的信

息,高效的处理信息并最终通过机器人实体辅助医生更好的实施诊断和手术。在手术机器人的发展过程中,欧洲国家起步非常早。世界上第一台骨科手术机器人就是由英国帝国理工的 BRIAN DAVIES 教授开发的,几乎同时期法国 Edouard Herriot 医院开发了应用在前列腺手术中的 HIFU 手术机器人。然而由于缺乏统一的市场和公司支持,目前在欧洲并没有盈利的手术机器人公司。目前的手术机器人还是由欧盟支持的 H2020(早前的 FP7)项目,以及各国的政府主导的项目,比如德国的 DFG 项目为主。从 2013 年开始人工智能在手术机器人领域有了较大的增长。目前对于智能手术机器人的研究集中在了半自主和全自主机器人上,对此 Moustiris 等^[6]进行了非常详尽的调查和总结。早在 2006 年,德国慕尼黑大学和瑞士 Dalle Molle 人工智能研究所(IDSIA)就使用了递归神经网络学习并使用手术机器人在模型上完成了心脏的缝合手术。卡尔斯鲁厄理工学院的 Weede 等提出了智能内镜导航方案,该方案通过学习之前内镜导航案例的运动轨迹,能够成功预测内镜的现行轨迹。他倡导发展能够长期预测的人工智能,以预计外科医生将在接下来的几分钟内做些什么,以便内镜可以始终被移动到适当的位置。近年来,欧洲国家对手术机器人的智能化非常重视,CASCADE 智能心血管手术机器人项目就是在这一背景下产生的。该项目隶属于欧盟 FP7 框架,由英国帝国理工、比利时鲁文大学和德国人工智能研究中心等几个研究机构参与合作。目的是为了减少现有基于实时 X 线透视操作的心血管手术的辐射,而发展基于自身传感器的高度自主化的心血管柔性机器人。在该项目中科研人员尝试了包括增强学习、模仿学习、神经网络和动态高斯等多种方法,对人工智能在心血管机器人在手术的安全性、智能自主导航、学习专家技巧等多方面有了初步的探索。

人工智能在欧洲的进展

在欧洲,使用人工智能并不能获得医疗保险公司的补贴。因此人工智能公司想真要销售 AI 产品,真正获得收入和盈利,开发出的放射科 AI 产品就需要能真正地增强医疗流程,切实解决临床痛点和难点。这里增强医疗流程可以分成几个方式:第一个是替代医生,做医生可以做的事情;第二个是提升医生的工作效率;第三个是用计算机系统做医生做不到事情。

对于替代医生这一模式,通常会有一些争议,但是在特定的医疗体系和流程中,这是可能的。譬如在欧洲由于实行大规模的乳腺癌筛查,有些欧洲国家(如瑞典)采用双医生读片制,但是人类医生的资源是宝贵并有限的。这时,一个和平均放射科医生水平相当的 AI

系统,就有可能替代双医生读片模式中的其中一个医生。目前在欧洲,我们还没观察到有完全在临床上使用 AI 机器来替代医生读片的现象,但目前这是一个活跃的研究领域。从商业模式来讲,一些欧洲的人工智能厂商开发的一些产品接近了放射科医生的水平,有些产品被安排到第三世界国家开展相关应用。

第二个方式是提升医生工作效率,医生使用智能 AI 产品会节约临床时间,譬如在 Radboud 医学中心,Jan van Zelst 等连续做了关于三维乳腺超声的研究,发现在 AI 系统的帮助下,医生读片的时间最快可以减少 30%。又譬如荷兰 Bram van Ginneken 教授开发的肺结节检测系统,只让医生去读 AI 系统展示的可疑区域,也可以节约读片时间。这里有个前提是在节约时间的同时,AI 辅助读片的精度效果至少需要维持或者提高,但不能下降。

第三个方式是用 AI 系统做医生做不到事情,譬如手术导航、手术机器人,用手术机器人做前列腺穿刺、眼科血管手术等,又譬如用 AI 计算机算法来精确估算骨龄,精确估算乳腺 X 线片的放射剂量等。

未来发展

过去 30 年 CAD 和人工智能在医学领域的发展被市场、技术以及资本驱动,目前在人工智能研发投入上遥遥领先的是美国、中国和日本。欧洲各个国家的投入还比较落后,欧洲各国的市场又远小于美国和日本,同时欧洲国家对医疗的先进技术的产业化也不够重视,事实上一些新兴科技公司发展到成熟阶段后就往往被北美公司收购。因此欧洲在医疗 AI 方面的发展,还是应该往技术创新和创意上聚焦,同时加强产业化支持。AI 技术方面的创新应该不仅仅集中于深度学习,因为深度学习不足够解决 AI 所有问题,未来其中的一个技术方向是解决如何让 AI 像人类一样去轻松的自主学习。同时,积极探索 AI 的落地场景以及实际临床测试和使用,将是使用 AI 改变目前临床工作流程最重要的一步。相比较而言,中国拥有巨大的市场空间和实际需求,医疗 AI 的研究和发展需要遵循规律,务实前进。

参考文献:

- [1] Doi K, Lanzl L, Lin PJ. Recent developments in digital imaging (eds)[M]. AAPM Medical Physics Monograph No. 2, published by AIP, New York, NY, 1985.
- [2] Zhang W, Li R, Deng H, et al. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation[J]. NeuroImage, 108, 214-224.
- [3] Babak Ehteshami Bejnordi, Mitko Veta, Paul Johannes van Diest, et al. Bram van ginneken, nico karssemeijer, geert litjens, jeroen AWM van der laak and the CAMELYON16 consortium, diagnos-

tic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer[J]. JAMA,2017,318(22):2199-2210.

[4] Wei Zhang Kunio Doi Maryellen L. Giger Robert M. Nishikawa Robert A. Schmidt. An improved shift-invariant artificial neural network for computerized detection of clustered microcalcifications in digital mammograms[J]. Medical Physics,1996,23(4):595-601.

[5] Kassahun Y, Yu B, Tibebu AT, et al. Surgical robotics beyond en-

hanced dexterity instrumentation; a survey of the machine learning techniques and their role in intelligent and autonomous surgical actions[J]. International J Computer Assisted Radiology and Surgery,2016,11(4):553-568.

[6] Moustris GP, Hiridis SC, Deliparaschos KM, et al. Evolution of autonomous and semi-autonomous robotic surgical systems; a review of the literature[J]. Int J Med Robot Comput Assist Surg, 2011,7(4):375-392.

(收稿日期:2018-10-01)

本刊可直接使用的医学缩略语

医学论文中正确、合理使用专业名词可以精简文字,节省篇幅,使文章精炼易懂。现将放射学专业领域为大家所熟知的专业名词缩略语公布如下(按照英文首字母顺序排列),以后本刊在论文中将对这一类缩略语不再注释其英文全称和中文。

- ADC (apparent diffusion coefficient):表观扩散系数
- ALT:丙氨酸转氨酶;AST:天冬氨酸转氨酶
- BF (blood flow):血流量
- BOLD (blood oxygenation level dependent):血氧水平依赖
- BV (blood volume):血容量
- b:扩散梯度因子
- CAG (coronary angiography):冠状动脉造影
- CPR (curve planar reformation):曲面重组
- CR(computed radiography):计算机 X 线摄影术
- CT (computed tomography):计算机体层成像
- CTA (computed tomography angiography):CT 血管成像
- CTPI(CT perfusion imaging):CT 灌注成像
- DICOM (digital imaging and communication in medicine):医学数字成像和传输
- DR(digital radiography):数字化 X 线摄影术
- DSA (digital subtraction angiography):数字减影血管造影
- DWI (diffusion weighted imaging):扩散加权成像
- DTI (diffusion tensor imaging):扩散张量成像
- ECG (electrocardiography):心电图
- EPI (echo planar imaging):回波平面成像
- ERCP(endoscopic retrograde cholangiopancreatography):经内镜逆行胰胆管造影术
- ETL (echo train length):回波链长度
- FLAIR (fluid attenuation inversion recovery):液体衰减反转恢复
- FLASH (fast low angel shot):快速小角度激发
- FOV (field of view):视野
- FSE (fast spin echo):快速自旋回波
- fMRI (functional magnetic resonance imaging):功能磁共振成像
- IR (inversion recovery):反转恢复
- Gd-DTPA:钆喷替酸葡甲胺
- GRE (gradient echo):梯度回波
- HE 染色:苏木素-伊红染色
- HRCT(high resolution CT):高分辨率 CT
- MPR (multi-planar reformation):多平面重组

- MIP (maximum intensity projection):最大密(强)度投影
- MinIP (minimum intensity projection):最小密(强)度投影
- MRA (magnetic resonance angiography):磁共振血管成像
- MRI (magnetic resonance imaging):磁共振成像
- MRS (magnetic resonance spectroscopy):磁共振波谱学
- MRCP(magnetic resonance cholangiopancreatography):磁共振胰胆管成像
- MSCT (multi-slice spiral CT):多层螺旋 CT
- MTT (mean transit time):平均通过时间
- NEX (number of excitation):激励次数
- PACS (picture archiving and communication system):图像存储与传输系统
- PC (phase contrast):相位对比法
- PET (positron emission tomography):正电子发射计算机体层成像
- PS (surface permeability):表面通透性
- ROC 曲线(receiver operating characteristic curve):受试者操作特征曲线
- SPECT (single photon emission computed tomography):单光子发射计算机体层摄影术
- PWI (perfusion weighted imaging):灌注加权成像
- ROI (region of interest):兴趣区
- SE (spin echo):自旋回波
- STIR(short time inversion recovery):短时反转恢复
- TACE(transcatheter arterial chemoembolization):经导管动脉化疗栓塞术
- T₁ WI (T₁ weighted image):T₁ 加权像
- T₂ WI (T₂ weighted image):T₂ 加权像
- TE (time of echo):回波时间
- TI (time of inversion):反转时间
- TR (time of repetition):重复时间
- TOF (time of flight):时间飞跃法
- TSE (turbo spin echo):快速自旋回波
- VR (volume rendering):容积再现
- WHO (World Health Organization):世界卫生组织
- NAA(N-acetylaspartate):N-乙酰天门冬氨酸
- Cho(choline):胆碱
- Cr(creatine):肌酸

· 人工智能与医学影像专题 ·

深度学习和医学影像在预防医学中的机会

吕乐, 吴山东

【关键词】 深度学习; 医学影像; 预防医学

【中图分类号】R05; R395 【文献标识码】A 【文章编号】1000-0313(2018)10-1006-03

DOI:10.13609/j.cnki.1000-0313.2018.10.006

开放科学(资源服务)标识码(OSID):



深度学习在不少领域取得了巨大成功,显现出改变游戏规则潜力。而和医学影像的结合可以促进目前工作流程的改进和改变。本文简要分享深度学习和医学影像在预防医学方面的机会。我们尝试概述现状,指出一些现有的成就,有意义的方向,局限性,以及短期内的机遇及挑战。为了充实我们的观点,我们引用了一些我们最近的关于放射影像研究的例子,也评论了一些最新的基于人工智能的研究成果和工业产品。

医学影像是疾病预防的检查手段之一,是指早期发现疾病,例如肺结节,肿瘤性疾病,肝/骨病变,目的是进行及时的干预和治疗,提高患者的健康管理。疾病的筛查是一种典型的预防医学模式。比如在美国和欧洲一些国家都有国家层面制定的乳腺癌筛查流程和相关指南,有大量的妇女接受定期的影像学检查,比如乳腺钼靶 X 线(digital mammography),乳腺断层(digital breast tomosynthesis),以期早期发现乳腺病变。对于一部分妇女,属于临床定义的高风险发病群体,她们还可以参加额外的基于磁共振的更敏感的筛查。影像学检查加强发病风险的预测和管理,进而支持临床决策,并最终提高疾病治愈率,降低疾病死亡率。传统疾病早期检测的方法是通过非侵入性成像设备的检查来完成,但是最近计算机辅助手段在影像检查和病变解读过程中的作用越来越突出。

深度学习,特别是卷积神经网络(convolutional neural networks, CNNs),这一革命性的新技术已经在各个领域取得重大进展,是目前人工智能技术的基本代表^[1]。在医学影像分析层面,深度学习更是取得了令人印象深刻的效果^[2]。在目前的临床工作流程中,医学影像具有巨大体量的数据,远远超出了其它模态(例如文本)的医学数据,因此非常有利于深度学习

技术的使用。深度学习可以分析各种多模态医学影像,包括 X 线、磁共振、CT、超声、病理和细胞图像等,能够在大量数据中检测到异常表现或者病变。例如,CT 扫描数据对较大的淋巴结和结肠息肉进行分类取得了积极的结果^[3]。在乳腺癌的筛查过程中,用基于深度学习的方法对乳腺密度进行自动分类,可以帮助医生进行发病风险的基本预测和指导进行何种补充性筛查^[4]。另外,用深度学习技术还可以为降低目前较高的复查率提供诊断支持^[5]。甚至在寻找新的定量乳腺癌发病风险标记物研究中,深度学习都取得了显著的效果。

传统的医学图像分析大多是基于临床经验,而传统的计算机医学图像分析大多是基于手动特征,是根据一些预先定义好的计算公式来提取某一类特征,但是手动特征对于描述医学影像是非常困难的,因为临床上很多真正有意义的影像特征都是定性的,基于经验的。而利用基于深度学习的数据驱动的方法,则可以改善其难度,比如 CNN 模型可以根据大规模的医学图像分析和标记,自动、自主地提取和组织有效的影像特征。基于 CNN 的各种研究还建立了大量的经验和证据,表明低级别的影像特征可以在不同成像设备所训练的神经网络模型之间进行共享和微调(也就是通过所谓的迁移学习的方法)。基于这一机制,医学影像建模甚至可以从用大量的自然图像训练的网络中受益。例如,很多研究工作都是利用 ImageNet(一个有超过百万张有标记的自然图像数据集)来预先训练深度学习模型,并取得了可观的效果。

深度学习也重新点燃并加强了工业界对医学成像和临床应用的兴趣。目前,有一群健康产业的初创公司专注于医学影像的智能分析和信息学开发,主要的推动力还是深度学习。这些公司包括 Heart-Flow、Enlitic、Arterys、Viz. ai、Zebra-Me 和 Butterfly Network 等等。这些新创公司的努力和在细分领域的深度探索很好的补充了大型企业(如 GE、西门子、飞利浦、Google、IBM 和腾讯等)的研发领地。目前工业界和风险投资覆盖了医学影像分析的各个方面,也积极推动深度学习向临床转化。尽管目前还存在很多技术、业务和临

作者单位:贝塞斯特,马里兰州,平安科技美国东部研究院;巴尔地摩,马里兰州,美国约翰霍普金斯大学计算机系(吕乐)/匹兹堡,宾夕法尼亚州,匹兹堡大学医学院放射科(系)和生物医学信息系;工程学院生物工程系;计算和信息学院智能系统系;临床和转化研究院(吴山东)

作者简介:吕乐——男,河南人,计算机哲学博士,主要从事医学图像、医疗信息学的研究和管理工作。

通讯作者:吴山东, E-mail: wus3@upmc.edu

基金项目:美国国家卫生研究院/美国国家癌症研究所项目(1R01CA193603, 1R01CA218405)

床方面的挑战,有效且可扩展的深度学习产品将在未来几年内继续进步,大力推动高性能和可部署的医学成像分析、决策支持以及临床信息学方面的应用。

深度学习的出现对于疾病早期筛查的发展将是一个极大的推动。对大量的筛查影像数据的深度挖掘,对生物标志物的研究将会产生巨大的机会。深度学习也会提高对大规模数据的分析和解读能力。在深度学习广泛应用之前,对影像的解读研究主要集中在计算机辅助检测/诊断上(computer-aided detection/diagnosis, CAD)上,如通过乳腺 X 线检测乳腺病变/肿块,在 CT 检测结肠息肉、肺结节或 CT 血管造影发现血管病变和肺栓塞等。来自几个供应商的此类商业软件产品被开发并已经部署到临床工作流程中,但广泛的成功和读片效果并没有达到预期值。根据美国 FDA 批准的协议, CAD 产品通常以第二读片人的方式工作,这意味着放射科医师首先在没有 CAD 的情况下进行独立图像诊断,然后结合 CAD 的分析发现再做最终诊断。该方法旨在借助 CAD 软件提高病变检测的灵敏度,但过多的假阳性造成了医生工作量的增加。事实上很多诊断工作对于有经验的放射科医师来说并不太困难,而 CAD 所增加的灵敏度并不能补偿工作量的增加,这是一个很现实的问题。美国在 CAD 方面的研究已经进行了几十年,但是临床实践效果却很难让人满意。CAD 的概念和想法让人充满憧憬,但是临床工作的效果却取决于多个因素。首先,具体的 CAD 实现在计算机算法层面是有相当挑战性的,特别是考虑到医学影像的差异性和多样性。传统的机器学习方法有其局限性。其次,要改变放射科医生的思维模式也是相当困难的。因此,利用计算工具进行辅助诊断决策是一个涉及多方面因素的研究课题,其效果评价也应真正的临床成效为唯一事实标准。

为了使 CAD 在预防性系统和临床应用真正起飞,需要解决真正的临床痛点,即传统的第二读片人方法或额外的工作负担。在这些方面,深度学习都可以带来积极的效果。同时, AI 技术应该更好地被应用于提高对患者的临床护理能力。下面我们简单介绍几个有前景的 AI 应用场景和例子。

美国食品和药品管理局(food and drug administration, FDA)最近通过了一款作为第一读片人的软件(来自 Viz. ai 的工作),该软件的本意是对中风患者提高检测大血管闭塞(large vessel occlusion, LVO)的机率并早期治疗。目前手动 LVO 检测的工作流程效率低下,并延迟了对患者的治疗,而通过计算机进行辅助分诊,实现基于 AI 软件的通知,节省卒中专家的时间,缩短了诊断时间窗。另一项类似的工作是低剂量全身 CT 扫描动脉粥样硬化血管钙化检测和分割,这

是一项非常耗时的任务,很容易被人工诊断遗漏。人工诊断的缺点和深度学习的高性能创新方案意味着这些任务适合 CAD 方法。其关键还是需要识别和抓住第一读片软件的机会,以进一步提高 CAD 解决方案的成效。

胸部 X 线是最常见的医学成像检查手段,用于社区诊所和医院筛选健康和亚健康人群(如年度健康体检时)。一个性能可靠、价格合适的自动化胸部 X 线检查和转诊工具,将是一个革新性的应用程序。这种软件程序可以部署在大量人群聚集或者居住的区域,特别是那些离大医院较远的地区。当然这样的任务也面临很多挑战(这些挑战也并不局限于胸部 X 线成像):①无论是对于放射科医生还是计算机辅助诊断系统,胸部 X 线检测与其它成像方式相比,比如用胸部 CT 检查肺结节,存在更高层次的诊断不确定性;②预防性筛查需要降低假阳性率;③疾病知识图谱的建模和整合对于提高图像分类器的诊断结果和推理至关重要;④打造临床上高性能的 CAD 系统需要人类医生可解释、可重复以及可验证的结果。克服这些挑战仍然需要很多努力,但最近的研究使该领域的发展更进一步,例如疾病部位的弱监督分析等工作^[6]。

全身预防性癌症筛查是正常或者无症状人群筛查的重要任务。理想的初始筛查手段是使用更经济和非侵入性的技术,例如“CancerSeek”血液测试^[7],筛查所有参加年度健康体检。接下来,使用高性能、高精度的自动医学成像手段和病理学分析工具来定位和验证初始发现。最后,结合所有非成像和成像测试结果的临床决策融合模块可以向专家出具一份报告,并将已识别的“高风险”患者与绝大多数的“平均风险”患者区分。虽然最近检测和分割复杂器官及异常解剖的工作初步取得了很好的进展^[8],对于一般人群的预防性筛查,基于大规模数据集进行深度学习模型训练的方法仍需要进一步研究。

从历史上看,由于机器学习技术的不足、缺少临床合作伙伴以及临床试验和/或数据上的不开放,开展表征和定量生物标记研究所需的数据档案缺乏有力的支持,定量和精准成像面临障碍。然而,随着深度学习能力的提高和精准医学政策方面的推动,我们看到了巨大的机遇。我们应该努力聚焦医院层面上的数据共享,用先进的计算技术探索精准医学生物标记,解决医生希望做到但是靠人类自身能力无法实现的问题。这些研究应该集中在有较高发病率的疾病,特别是癌症上。目前需要进一步开发的关键是如何在现有数据源(如医院档案)上训练深度学习系统,这些数据源的规模往往非常大,但也非常混乱和非结构化。

在基于医学影像的临床研究中,特别是大规模的

定量分析和有标记的深度学习,先决条件之一是从医学扫描中精确且稳健地分割解剖结构,即将每个像素或体素分类为语义。基于 CNN 的深度分割方法展现出优越的性能,现在占主导地位。使用影像分割技术的价值在于可以将原始图像扫描转换为语义和人类可解释的特征,例如左心室的体积或患者胰腺的形状等。这些基于器官/解剖学的形状、体积和外观的特征可以从 2D/3D/4D 图像计算获得,以辅助个性化诊断/治疗以及大数据分析。其重要的挑战是搜集足够多的具有标定的数据用于深度学习模型训练,并确保任何分割解决方案都具有更高的适应性和强壮性。

最后还有很重要的一点,精准医学的关键目标之一是对临床数据进行回顾性分析,以探索和发现与发病相关的影像标志物。需利用一个医院或多家医院的数据,对患者的图像和非图像数据进行标记、挖掘和分析。比如,可以在图论基础上对多模态成像和非图像的患者数据进行建模和可视化,构建患者相似性,自动发现肿瘤类型簇等^[9]。这种建模和高阶表示可以提供可索引的患者数据视图,超越普通分类的分析。重要的是鉴于许多疾病的长尾分布,如果没有强大的计算机化技术(如深度学习)有效利用大数据,这种分析即使不是不可能,也很难实施。

总之,最近深度学习的发展对于医学影像的分析正在产生巨大的影响,甚至可以使一些通过非深度学习方法不可能完成的重要任务(例如,通过第一读片进行数据分流)达到临床可以接受的水平。筛查影像大数据结合深度学习将是一个充满希望但具有挑战性。我们预期在不远的将来,这方面的研究和临床转化工

作将会蓬勃发展。

参考文献:

- [1] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553):436-444.
- [2] Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning[J]. Cell, 2018, 172(5):1122-1131.
- [3] Roth HR, Lu L, Liu J, et al. Improving computer-aided detection using convolutional neural networks and random view aggregation [J]. IEEE Trans Med Imaging, 2016, 35(5):1170-1181.
- [4] Mohamed AA, Berg WA, Peng H, et al. A deep learning method for classifying mammographic breast density categories[J]. Medical Physics, 2018, 45(1):314-321.
- [5] Aboutalib SS, Mohamed AA, Berg WA, et al. Deep learning to distinguish recalled but benign mammography images in breast cancer screening[J]. Clinical Cancer Research, 2018. In press.
- [6] Wang X, Peng Y, Lu L, et al. ChestX-ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 3462-3471.
- [7] Cohen JD, Li L, Wang Y, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test[J]. Science, 2018, 359(6378):926-930.
- [8] Cai J, Tang Y, Lu L, et al. Accurate weakly supervised deep lesion segmentation on CT scans; self-paced 3D mask generation from RECIST[J]. arXiv:1801.08614 [cs.CV]. 2018.
- [9] Yan K, Wang X, Lu L, et al. Deep lesion graphs in the wild: relationship learning and organization of significant radiology image findings in a diverse large-scale lesion database[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

(收稿日期:2018-10-02)

· 人工智能与医学影像专题 ·

AI 成像优化联合迭代算法在“双低”主动脉 CTA 的初步应用

王明, 王怡宁, 于敏, 王沅, 王曼, 金征宇

【摘要】 目的:探讨人工智能(AI)成像优化技术联合迭代算法(ClearView+)对“双低”(低辐射剂量及低碘负荷)主动脉 CTA 图像质量的影响。**方法:**前瞻性连续纳入 2018 年 2~5 月在本院 NeuViz 128 CT 行主动脉 CTA 检查的 40 例患者,随机分为 A、B 两组,每组 20 例。A 组为“双低”组(管电压 80 kVp,对比剂 40 mL);B 组为常规组(管电压 120 kVp,对比剂 80 mL)。A 组图像采用 ClearView+ 迭代算法分别按照 6 个不同重建档位(0%、10%、30%、50%、70%及 90%;0%为滤波反投影)重建图像,选出图像质量最佳重建档位作为 A1 组,进一步对 A1 组图像采用 AI 成像优化技术进行图像优化,所得图像作为 A2 组;B 组则采用滤波反投影重建算法。对 A1、A2 和 B 3 组图像进行客观评价,分别以测量 3 组主动脉感兴趣区的 CT 值及噪声(SD)值,计算信噪比(SNR)及对比噪声比(CNR)作为图像质量客观评价指标,主观评价由 2 名高年资放射科医生采用双盲法进行 5 分制评价(5 分—优良,1 分—极差)。记录和比较 A 组与 B 组的对比剂用量和辐射剂量(CTDIvol, DLP 和 ED)。**结果:**ClearView+ 90%档位分别与 0%、10%、30%、50%和 70%档位比较,主动脉 SD 显著降低、SNR 及 CNR 显著升高($P < 0.05$),因此 Clearview+90%作为最佳图像质量的迭代档位。三组图像定量评价,SD、SNR、CNR 两两比较结果:A2 组 $SD(12.21 \pm 2.66) < A1$ 组 $SD(17.67 \pm 2.75) < B$ 组 $SD(21.83 \pm 4.64)$ ($P < 0.05$),A2 组 $SNR(36.77 \pm 10.13) > A1$ 组 $SNR(29.57 \pm 7.41) > B$ 组 $SNR(19.09 \pm 5.65)$ ($P < 0.05$),A2 组 $CNR(47.92 \pm 12.36) > A1$ 组 $CNR(38.50 \pm 8.96) > B$ 组 $CNR(23.37 \pm 6.61)$ ($P < 0.05$)。三组图像之间主观评分无统计学差异($P > 0.05$)。“双低”组与常规组的辐射剂量评价指标 CTDIvol(2.82 ± 0.36 vs 13.34 ± 1.84) mGy, DLP(181.21 ± 21.98 vs 871.17 ± 134.26) mGy·cm 和 ED(2.54 ± 0.31 vs 12.20 ± 1.88) mSv 均有显著性差异。“双低”组有效辐射剂量 ED 较常规组显著降低 79.18%,对比剂用量降低 50%。**结论:**AI 成像优化技术联合 ClearView+ 迭代算法,使 NeuViz 128 CT 主动脉“双低”扫描可以获得和常规扫描同样的图像质量且有效地降低了辐射剂量和对比剂用量。

【关键词】 人工智能;图像处理,计算机辅助;主动脉;体层摄影术,X 线计算机**【中图分类号】** R05; R814.42; R814.3; R322.121 **【文献标识码】** A**【文章编号】** 1000-0313(2018)10-1009-08

DOI:10.13609/j.cnki.1000-0313.2018.10.007

开放科学(资源服务)标识码(OSID):



Preliminary application of artificial intelligence-based Image optimization and iterative algorithm in "double low" aorta CTA WANG Ming, WANG Yining, YU Min, et al. Department of Radiology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China

【Abstract】 Objective: To investigate the value of artificial intelligence (AI)-based image optimization technique and iterative algorithm (ClearView+) in the image quality of "double low" aorta CTA. **Methods:** Forty patients who were referred to perform aorta CTA examination between February and May 2018 in NeuViz128 CT of our hospital were prospectively collected and randomly divided into two groups. Group A was scanned using a low tube voltage of 80kVp and a reduced contrast medium volume of 40mL ioversol (320mg I/mL), and group B was scanned with a standard 120kVp tube voltage and an injection of 80mL contrast. Group A was reconstructed using ClearView+ technique of six percentages including 0% (filtered back projection), 10%, 30%, 50%, 70% and 90%. After comparing

作者单位:100730 北京,中国医学科学院/北京协和医学院,北京协和医院放射科(王明、王怡宁、王沅、王曼、金征宇);
100193 北京,东软医疗设备有限公司(于敏)

作者简介:王明(1985—),男,山东莱芜人,技师,主要从事心血管影像学 and 分子影像学。

通讯作者:王沅, Email: wangyun8637@163.com

these different percentages, the best percentage level of ClearView+ was selected as group A1. Both best percentage of iterative algorithm and further AI-based image optimization were used in group A2. Group B was reconstructed by using filtered back projection (FBP). The mean attenuation of contrast-enhancement values, image noise, signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) were calculated to evaluate image quality objectively. In addition, subjective image analysis was performed using five-point scales independently by two radiologists (5-excellent, 1-undiagnostic). The estimated radiation dose in terms of CT dose index (CTDIvol), dose length product (DLP), and effective dose (ED) was recorded and compared between group A and group B. **Results:** Compared with other percentage of ClearView+, the image noise of Clearview+ 90% was decreased and whereas SNR and CNR increased significantly ($P < 0.05$). Therefore, 90% was the best percentage level of Clearview+ without over-smoothing. The comparison results of SD, SNR and CNR among the three groups were below: group A2 SD (12.21 ± 2.66) < group A1 SD (17.67 ± 2.75) < group B SD (21.83 ± 4.64) ($P < 0.05$), group A2 SNR (36.77 ± 10.13) > group A1 SNR (29.57 ± 7.41) > group B SNR (19.09 ± 5.65) ($P < 0.05$), A2 group CNR (47.92 ± 12.36) A1 group CNR (38.50 ± 8.96) > B group CNR (23.37 ± 6.61) ($P < 0.05$). No statistical differences were observed among the three groups in the scores of subjective image quality ($P > 0.05$). There were significant differences between the "double low" group and the conventional group in the evaluation indexes of radiation dose including CTDIvol (2.82 ± 0.36 vs 13.34 ± 1.84) mGy, DLP (181.21 ± 21.98 vs 871.17 ± 134.26) mGy · cm and ED (2.54 ± 0.31 vs 12.20 ± 1.88) mSv. Compared with the conventional group, the effective dose and contrast medium volume of the "double low" group were reduced by 79.18% and 50% respectively. **Conclusions:** With the combination of ClearView+ and AI-based image optimization technique, the "double low" aorta CTA scan of NeuViz 128 CT could provide the consistent image quality with the regular scan, while magnificently reducing the radiation dose and contrast medium intake.

[Key words] Artificial intelligence; Image processing, computer assisted; Aorta; Tomography, X-ray computed

近年来,主动脉疾病伴随着较高的发病率及致死率,且由于症状隐匿等特征,在早期很难被觉察^[1]。早期诊断对于指导治疗和改善预后都具有重要意义。CT 主动脉血管造影(computed tomography angiography, CTA)检查因其良好的空间和时间分辨率、无创及检查快速等特点,已经成为临床上主动脉疾病的常规诊断手段。然而,全主动脉 CTA 检查由于扫描范围大,尤其是术后随访患者需要多次进行主动脉 CTA 检查,其带来的高辐射剂量和对比剂用量,可能增加致癌风险,或导致心血管反应、对比剂肾病等并发症^[2]。最近有学者提出了“双低”概念,即在保证临床诊断质量的基础上,降低 X 线辐射剂量和碘摄入量,进而降低对受检者损害的风险^[3]。目前最常用的降低辐射剂量的方法有降低管电压及管电流、大螺距扫描、减少扫描范围及时间等。由于辐射剂量与管电压的平方成正比,因此降低管电压会更有效的降低辐射剂量^[4]。管电压降低, X 线光子能量比碘的 K 层电子结合能稍大或等于时,光电效应发生几率增加,导致含碘对比剂的血管腔与周围组织对比度增大,从而间接降低对比剂的使用量^[5]。但降低管电压会使得图像质量下降,比

如噪声增加,伪影增多等。如何通过改进图像重建算法来改善图像质量已成为新的研究方向。传统的滤波反投影(filtered back projection, FBP)算法简单、重建速度快,被广泛用于 CT 图像重建,但它也存在一些局限性,包括图像噪声增加、对比度分辨率降低等。与该算法相比,迭代重建可通过多次迭代校正循环计算来提高图像质量,能够有效降低图像噪声,抑制伪影产生,但这类方法往往需要较长的处理时间以及足够的投影数据^[6]。伴随人工智能(artificial intelligence, AI)的迅速发展,医学影像与 AI 的结合被认为是最有发展前景的领域^[7]。最近,一些机器学习技术已经被改进用来处理低剂量 CT 中常见的噪声增加和细节丢失等问题^[8-10]。本研究将 AI 成像优化技术及迭代重建算法(ClearView+)应用在“双低”主动脉 CTA 成像,评价其图像质量和辐射剂量,并初步探讨其临床应用价值。

材料与方法

1. 病例资料

前瞻性连续搜集 2018 年 2~5 月于本院 128 层

CT 行主动脉 CTA 检查的患者。40 例病患纳入研究,男 22 例、女 18 例,年龄 18~81 岁,平均年龄(51.5 ± 13.0)岁,BMI 为(23.68 ± 3.59) kg/m^2 ;纳入标准:临床诊断或疑似大动脉炎、主动脉夹层、主动脉瘤者。入组标准符合临床诊断、完成血清肌酐检查。排除标准:年龄 <18 岁;肾功能异常(血清肌酐 $\geq 120 \mu\text{mol/L}$);含碘对比剂过敏,或其他严重过敏史;确诊或疑似甲状腺功能亢进或嗜铬细胞瘤;血液动力学状态不稳定;呼吸无法自主配合检查;妊娠期妇女。采用随机数字表法分为 A、B 两组,每组各 20 例。A 组(管电压 80 kVp),年龄 28~71 岁,平均年龄(47.5 ± 10.8)岁,BMI 为(23.36 ± 3.55) kg/m^2 ;B 组(管电压 120 kVp),年龄 25~81 岁,平均年龄(55.55 ± 13.8)岁,BMI 为(23.99 ± 3.60) kg/m^2 。本研究经院伦理委员会批准;所有患者检查前均签署知情同意书。

2. CT 检查

采用沈阳东软医疗 NeuViz 128 CT 行全主动脉 CTA 扫描。患者采取仰卧位,扫描范围从锁骨至耻骨联合。CT 扫描参数:实验组管电压 80 kVp,对照组管电压 120 kVp,采用管电流自动调节技术(ref mA:150),旋转时间 0.5 s,螺距 1.2,扫描视野 $360 \text{ mm} \times 360 \text{ mm}$,重建层厚 1.0 mm,重建层间距 0.5 mm,重建矩阵 512×512 ,准直 $128 \times 0.625 \text{ mm}$,滤波参数 F20。实验组采取迭代算法 ClearView+分别重建出 6 个不同权重(0%、10%、30%、50%、70%、90%;0%为 FBP)的图像,并将图像质量最优组设为 A1 组;在东软医疗 AVW1.0.8 专业工作站上用 AI 成像优化技术对 A1 组进行再处理,重建图像为 A2 组;对照组的原始图像则采取 FBP 算法进行重建。对比剂 AB 组均采用含碘量为 320 mg I/mL 的碘佛醇(江苏恒瑞医药股份有限公司)。增强扫描采用双筒高压注射器(美国 Mallinckrodt Optivantage V5)及 18G 套管针经肘静脉注射。注射方案:实验组 40 mL 对比剂,注射流率为 3.0 mL/s,以 4.5 mL/s 流率续 45 mL 的生理盐水;对照组为 80 mL 对比剂,注射流率为 5.0 mL/s,后以同样流率续 50 mL 生理盐水。应用对比剂智能追踪技术触发扫描,监测点位于降主动脉(膈肌平面),触发阈值设置为 180 HU,达阈值延迟 6 s 后触发扫描。

3. 图像质量分析

在东软医疗 AVW1.0.8 专业工作站进行后处理重建,重建图像包括 MPR、MIP、VR 图像。对图像质量进行主观及客观评价,并对辐射剂量进行分析。选择感兴趣区(region of interest,ROI),并测量相应的 CT 平均值。ROI₁ 位于升主动脉(肺动脉干平面),ROI₂ 位于主动脉干,ROI₃ 位于降主动脉(肺动脉干平面),ROI₄ 位于降主动脉(肠系上膜动脉的根部平面),

ROI₅ 位于髂动脉分叉,ROI₆ 位于皮下脂肪(肺动脉干平面)。ROI 层面选择避开有金属支架或运动引起伪影的层面,面积尽可能大并且尽量选择密度均匀区域,避开血管壁、钙化斑块及支架,夹层部位 ROI 位于真腔,并记录 ROI₁~ROI₅ 的标准差(SD1-5)取其平均值作为噪音指数 SD,ROI₁~ROI₅ 的平均值 CT 为对比剂的衰减情况,计算信噪比(signal to noise ratio,SNR)和对比噪声比(contrast to noise ratio,CNR),计算公式为 $\text{SNR} = \text{CT}/\text{SD}$, $\text{CNR} = (\text{CT} - \text{CT}_6)/\text{SD}$ 。

由 2 名放射科医师采用 5 级评分法分别对所有图像进行图像质量评估,意见不一致时进行讨论统一。5 分,图像质量优良,边界锐利,噪声小,基本没有伪影,达到临床诊断要求;4 分,图像质量好,边界锐利,噪声小,有少量伪影,达到临床诊断要求;3 分,图像质量中等,噪声和伪影情况一般,基本达到临床诊断要求;2 分,图像质量较差,噪声和伪影较多,不能达到临床诊断要求;1 分,图像质量极差,噪声和伪影情况严重,无法用于临床诊断。图像主观评分 ≥ 3 分即认为图像可被临床接受^[11]。

4. 辐射剂量分析

记录 2 组容积 CT 剂量指数(CT dose index,CT-DIvol)、剂量长度乘积(dose length product,DLP)、扫描长度、有效辐射剂量(effective dose,ED),其中 $\text{ED} = \text{DLP} \times k$ (推荐全主动脉 k 值为 0.015)。计算 A 组 ED 降低程度, $\text{ED}\% = (\text{EDB} - \text{EDA})/\text{EDB} \times 100\%$ ^[12-13]。

5. 统计方法

所有数据采用 SPSS 19.0 统计学软件进行统计学分析。连续变量用平均数 \pm 方差或中位数(第一四分位数—第三四分位数)表示,分类变量用频数表示。对患者年龄、身高、体重、BMI、扫描长度、主观评分、客观图像质量评价(CT、SD、SNR 及 CNR)以及辐射剂量(CTDIvol、DLP 及 ED)所有数据采用 Shapiro-Wilk W 进行正态分布检验。两组之间比较独立样本 t 检验或 Wilcoxon 秩和检验。对于 3 组之间相互比较,采用单因素方差分析(ANOVA)或 Kruskal-Wallis 检验,若三者之间存在显著差异,两两比较采用 LSD 检验或 All pairwise 检验。临床诊断结果采用 χ^2 检验,取 $P < 0.05$ 作为显著性水准。两位医师对于图像质量的主观评分一致性分析采用 Cohen kappa 检验,kappa 值 1~0.81 优异,0.61~0.80 良好,0.41~0.60 中等,0.21~0.40 一般, <0.20 差。

结 果

1. 临床资料

实验组 A 组与对照组 B 组共 40 例患者均顺利完

成检查,无药物不良反应及身体不适,扫描数据均可以满足诊断要求。实验组与对照组患者年龄、身高、体重、性别比例、BMI 指数及扫描长度均无统计学差异(表 1, $P>0.05$)。

表 1 两组患者临床资料的比较

指标	A 组 (n=20)	B 组 (n=20)	P 值
年龄(岁)	47.5±10.8 28~71	55.5±13.8 25~81	>0.05
性别(男,%)	11(55%)	11(55%)	>0.05
身高(cm)	166.05±8.83	168.80±8.41	>0.05
体重(kg)	64.60±12.55	68.40±12.15	>0.05
BMI(kg/m ²)	23.36±3.55 16.65~30.43	23.99±3.60 17.0~31.24	>0.05
扫描长度 (mm)	548 536.75~549.5	546 533.25~566.25	>0.05

注:A组:“双低”实验组;B组:常规对照组(下表同)。

A 组 20 例的诊断结果中,主动脉夹层患者有 4 例(20%),动脉粥样硬化患者有 3 例,主动脉瘤患者有 2 例,血管内支架术后的患者有 3 例,以及壁内血肿患者有 2 例;而在 B 组 20 例中,主动脉夹层患者有 3 例,动脉粥样硬化患者有 4 例,主动脉瘤患者有 5 例,血管内支架术后的患者有 2 例,以及壁内血肿患者有 0 例,A 组与 B 组两者的疾病占比相比较无统计学差异。

2. 图像质量评价

客观评价:A1 组、A2 组及 B 组的主动脉各个节段管腔内 CT 值差异无统计学意义(表 2, $P>0.05$);同时,3 组的血管 CT 值沿 Z 轴方向(5 个 ROI)均具有一致性(图 1)。

针对不同档位(0%~90%)的 ClearView+ 图像分析,与其他 5 个档位(0%,10%,30%,50%及 70%)的图像相比,90%重建档位的 Clearview+ 重建图像具有最佳主动脉 SD 值降低、SNR 值及 CNR 值逐渐上升。90%重建档位的 ClearView+ 具有最佳图像质量

因此作为 A1 组,即主动脉 SD 最低,SNR 及 CNR 最高(表 3, $P<0.05$),且无蜡块样伪影。与 0%重建档位即传统 FBP 比较,90%重建档位的 ClearView+ 降低 45.19%的噪声以及提升 95.96%的 SNR 和 96.98%的 CNR。

A1 组、A2 组和 B 组三者之间两两比较,图像质量的指标图像噪声($17.67\pm2.75, 12.21\pm2.66, 21.83\pm4.64$)、SNR($29.57\pm7.41, 36.77\pm10.13, 19.09\pm5.65$)及 CNR($38.50\pm8.96, 47.92\pm12.36, 23.37\pm6.61$)均有显著差异(图 2, $P<0.05$)。A2 组图像质量最佳,SD 值低于 A1 组及 B 组,SNR 及 CNR 高于 A1 组及 B 组;A1 组优于 B 组,SD 值低于 B 组,SNR 及 CNR 值高于 B 组。A2 组相比 B 组降低 19.06%的噪声以及提升 54.90%的 SNR 和 64.74%的 CNR。

主观评价:两名诊断医师分别对 A1 组、A2 组及 B 组主动脉图像质量评价(表 4),三组所有患者图像评分均 ≥ 3 分,满足诊断要求(图 3、4)。三组患者之间图像质量评分无统计学差异($P>0.05$),且两位医师对于图像质量的主观评分一致性优异。

表 4 三组图像质量主观评价结果

诊断医师	A1 组	A2 组	B 组	P 值
医师 1	5 (4~5)	5 (4~5)	5 (4~5)	>0.05
医师 2	5 (4~5)	5 (4~5)	5 (4~5)	>0.05
kappa 值	0.886	0.886	0.829	

辐射剂量及对比剂用量比较:与实验组 A 组相比,辐射剂量指标 CTDIvol、DLP 以及 ED 均显著低于对照组 B 组($P<0.05$),其中有效辐射剂量 ED 降低幅度为 79.18%(表 5)。同时,A 组对比剂用量较 B 组降低 50%。

表 2 5 个 ROI 的血管 CT 值结果

ROI 血管 CT 值	A1 组	A2 组	B 组
升主动脉(肺动脉平面)(ROI ₁)	421.4±83.33	420.95±82.82	393.87±56.19
主动脉干(ROI ₂)	408.73±75.81	405.45±74.41	399.66 (364.41~410.01)
降主动脉(肺动脉干平面)(ROI ₃)	405.89±72.24	403.11±70.67	392.98±59.55
降主动脉(肠系膜动的根部平面)(ROI ₄)	390.48±85.42	392.08±85.60	367.45±56.01
髂动脉分叉(ROI ₅)	401.29±91.03	403.83±90.86	403.06 (325.67~432.46)

注:* $P<0.05$,ROI 之间对比。

表 3 不同档位 Clearview+ 重建图像噪声、SNR 与 CNR 的比较结果

项目	Clearview+ 0%	Clearview+ 10%	Clearview+ 30%	Clearview+ 50%	Clearview+ 70%	Clearview+ 90%
图像噪声	27.71±5.31 ^{bdef}	19.69±2.09 ^{nef}	18.64±2.21 ^{af}	18.06±2.39 ^{af}	16.83±2.70 ^{abf}	15.30±3.53 ^{abcde}
SNR	15.09±4.12 ^{bdef}	21.15±4.80 ^{nef}	22.58±5.13 ^{af}	23.55±5.51 ^{af}	25.70±2.09 ^{abf}	29.57±7.41 ^{abcde}
CNR	19.54±4.70 ^{bdef}	27.45±5.31 ^{nef}	29.35±5.79 ^{af}	30.60±6.30 ^{af}	33.42±6.97 ^{abf}	38.49±8.96 ^{abcde}

注:^a $P<0.05$ 与 0%同指标比较;^b $P<0.05$ 与 10%同指标比较;^c $P<0.05$ 与 30%同指标比较;^d $P<0.05$ 与 50%同指标比较;^e $P<0.05$ 与 70%同指标比较;^f $P<0.05$ 与 90%同指标比较。

表 5 两组扫描方案辐射剂量评估比较

指标	A 组	B 组	P 值
CTDIvol(mGy)	2.82±0.36	13.34±1.84	<0.05
DLP(mGy·cm)	181.21±21.98	871.17±134.26	<0.05
ED(mSv)	2.54±0.31	12.20±1.88	<0.05
对比剂用量(mL)	40	80	NA

注:CTDIvol,容积 CT 剂量指数;DLP,剂量长度乘积;ED,有效辐射剂量。

讨 论

研究表明 CT 检查产生 X 线辐射剂量每增加 1 mSv 的有效辐射剂量将会增 0.05% 的恶性肿瘤发病率^[14]。国际放射防护委员会(international commision on radiological protection, ICRP)提出辐射防护与安全最优化理论,推荐获得足够诊断信息及合适图像质量时,尽可能降低辐射剂量(as low as reasonably achievable, ALARA)。此外,碘对比剂会给患者带来潜在危害,高浓度对比剂的碘负荷更大,其高粘度和渗透压会增加对患者毛细血管循环的影响^[15],增加对比剂肾病的风险。

本研究采用 80 kVp 管电压主 动脉 CTA 扫描联合 Clearview+ 迭代重建及 AI 成像优化技术可在明显改善图像质量的同时,有效辐射剂量仅为(2.54±0.31) mSv,对比剂用量仅为 40 mL。与常规 120 kVp 管电压下扫描主 动脉 CTA 相比,大幅度的降低了辐射剂量(79.18%)和对比剂用量(50%)。对比同类研究, Pontana 等^[17]对 80 kVp 和 120 kVp 的条件下的两组患者均注射 100 mL 碘比醇(300 mg I/mL),结果表明 80 kVp 条件下其碘对 X 线的衰减的光电效应增强,从而血管的 CT 值升高,血管强化更为明显^[18]。

与此同时,80 kVp 联合迭代重建算法在保证图像质量的前提下,辐射剂量降低了 50%。本研究的结果也验证了这一结论,但本研究 80 kVp 的实验组仅仅用 40 mL 对比剂也达到高于 250 HU 的 CT 值,满足诊断需求的优质图像的前提,辐射剂量得到了更大程度的降低(79.18%)^[19]。李岩等^[20]发表的《规范化应用“双低”扫描 CT 血管成像》中建议对 BMI≤23 kg/m² 的患者采用 80 kVp, 23 kg/m²<BMI<28 kg/m² 的患者采用 100 kVp, BMI≥28 kg/m² 的肥胖患者采用 120 kVp。Hou 等^[16]利用荷兰飞利浦 Brilliance iCT 256 层 CT 扫描从主动脉弓上 2 cm 水平至耻骨联合水平全主动脉 CTA 检查,该研究表明 80 kVp 的联合迭代重建算法能够得到完全满足诊断图像质量的图像,同时与 120 kVp 联合传统 FBP 重建图像相比,辐射剂量显著降低了 77.6%,仅为(2.8±1.4) mSv。但该研究仅仅对非肥胖的患者得出这样的结论,而肥胖患者由于其更大的降噪需求,研究更具有挑战性和临床价

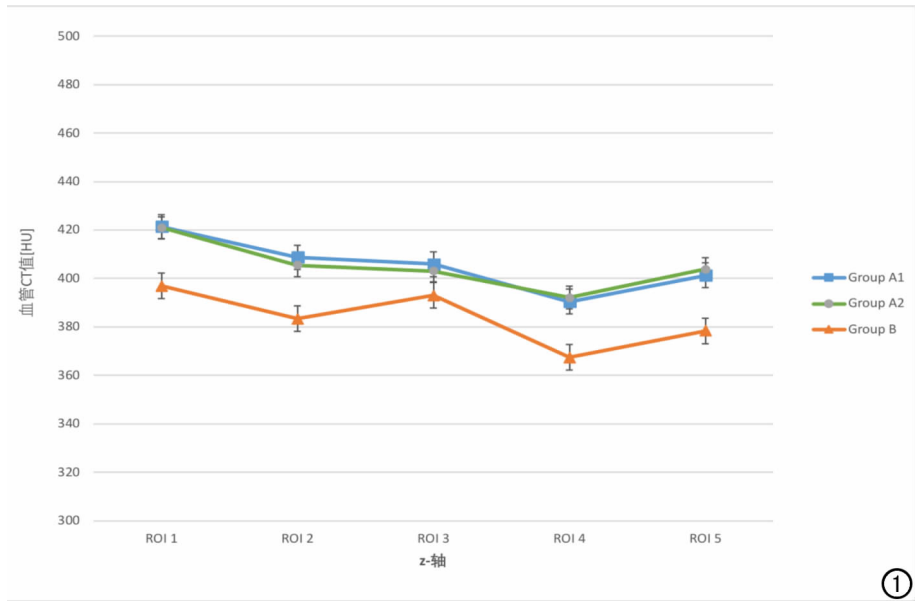


图 1 蓝(方格),绿(圆形)和橙(三角)曲线分别表示 A1 组, A2 组及 B 组沿 Z 轴的血管 CT 值的走势。

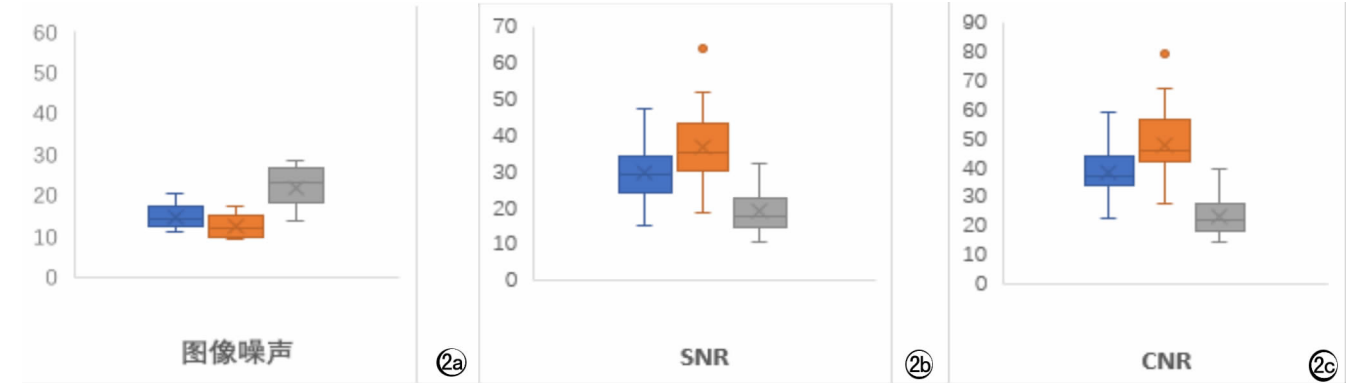


图 2 A1 组, A2 组和 B 组图像质量客观评价指标箱图对比结果。a) 图像噪声; b) SNR; c) CNR。

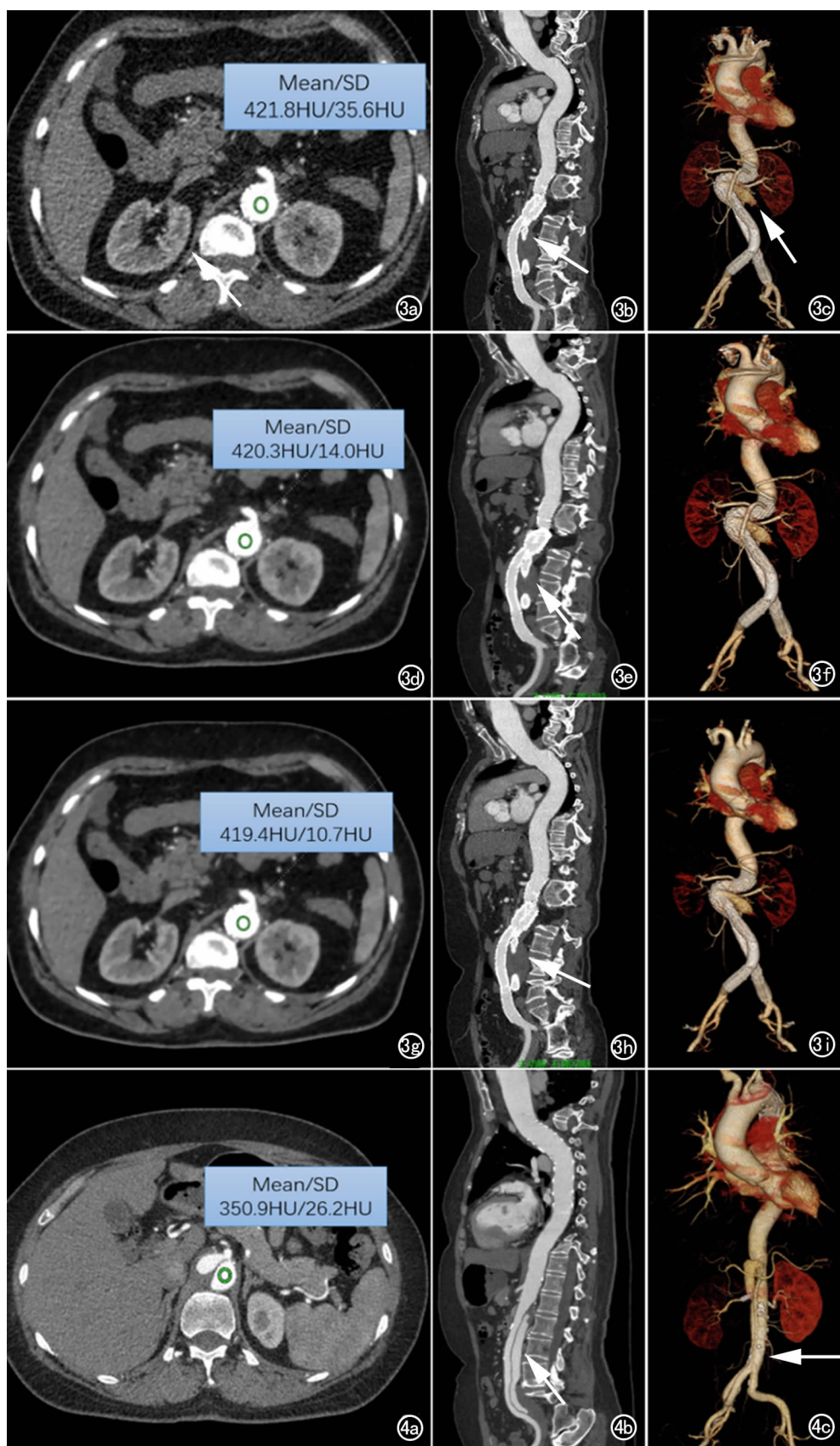


图3 “双低”A组,62岁,女。BMI:30.4 kg/m²,主动脉瘤修复术后,腹主动脉及双侧髂总动脉支架置入后改变,支架内管腔通畅;支架外瘤腔内局部对比剂内漏影(箭)。a) FBP重建轴面图像;b) FBP重建CPR图像;c) FBP重建三维VR图像;d) Clearview+90%重建轴面图像;e) Clearview+90%重建CPR图像;f) Clearview+90%重建三维VR图像;g) Clearview+90%重建+AI成像优化轴面图像;h) Clearview+90%重建+AI成像优化CPR图像;i) Clearview+90%重建+AI成像优化三维VR图像。ClearView+90%迭代重建后主动脉噪声降低60.7%(35.6降低到14);ClearView+90%重建联合AI图像优化后噪声降低69.9%(35.6降低到10.7)。

图4 常规剂量组B组,54岁,女。BMI:31.2 kg/m²,主动脉夹层支架植入术后,支架官腔通畅;腹腔干起始处水平及以下腹主动脉及右侧髂总动脉全程夹层(箭)。a) FBP重建轴面图像;b) CPR图像;c) 三维VR图像。

值。本研究肥胖患者仍作为研究对象,实现“双低”扫描,并且得到了非常显著的图像质量提升(图3)。

然而,以往报道显示传统迭代重建还存在一些缺陷,如过于平滑造成伪自然现象,增加蜡状伪影以及过

长的重建时间等^[21]。本研究 ClearView+迭代算法不仅改善低剂量图像质量,同时由于其基于多模型的双域(多模型即光子分析模型、伪影矫正模型、解剖模型及噪音模型;双域即生数据域与图像域)迭代更好的克

服了上述这些缺陷。ClearView+ 90% 迭代是利用 10% 原始剂量的图像生数据与 100% 的原始剂量进行噪声差异匹配,建立噪声模型,在保留细节的前提下,根据解剖模型修正噪声模型。因此,Cleaview+ 的噪声模型的建立及修正是基于真正低剂量数据图像与原始数据图像的匹配构建的,其获得的去除噪声的图像更为真实和自然。其他厂家的噪声模型校正是将图像数据与基于统计的,考虑到光子和电子噪声的理想噪声模型进行比较(而非原始剂量的真实图像),同时得到的图像按照档位比例在图像域融合,如 90% 档位混合迭代,是将 90% 的迭代重建与 10% 的 FBP 重建图像域进行融合,即理想模型占比更多,更易失真。另外,90% 作为 Clearview+ 重建最佳档位联合 80 kVp 低管电压的扫描图像得到相比传统 FBP 图像降低 45.19% 的噪声以及提升 95.96% 的 SNR 和 96.98% 的 CNR。Clearview+ 的最佳重建档位的效果与其他同类研究的迭代算法效果对于“双低”CT 血管成像改善的相比具有明显优势。Winklehner 等^[22] 研究表明德国西门子迭代技术(sinogram affirmed iterative reconstruction, SAFIRE)在主动脉 CTA 检查中可以降低噪声 30.8%。

人工智能作为新兴技术在 CT 图像低剂量改善方面也成为了热点研究课题。本研究采用新一代迭代技术 Clearview+ 的同时,也进一步比较了联合基于 AI 成像优化技术后的图像质量改善的效果。AI 图像优化技术本质是用卷积神经网络(convolutional neural network, CNN)来学习低剂量 CT 图像的噪声模型,进而实现噪声与图像的分离,并保证在此过程中图像细节不会明显丢失,从而降低患者辐射剂量的同时,所得到的图像依然能够满足临床需求。与此同时,AI 优化避免了手动设置经验参数的过程,而是从数据中直接学习出最优的参数,相比于传统方法采用的通用型、经验噪声模型,这种定制化噪声模型往往会得到更好的结果。此外,此 AI 图像优化技术弥补了迭代算法厂家专一性的缺陷^[9],直接对 CT 图像域进行去噪,解决了使用者难以直接获取的 CT 扫描仪中间投影数据的问题^[23],有望成为改善所有厂商的低剂量图像的有力手段。Hu 等^[9] 基于 CNN 的深度学习算法可以改善胸部及腹部的低剂量 CT 的图像质量,但此研究仅限于平扫,未对三维重建图像进行对比。赵莹等^[7] 研究得出深度学习的像素闪烁算法(pixel shine, PS)可提升高体质质量指数($\text{BMI} \geq 25 \text{ kg/m}^2$)患者低剂量腹部 CT 平扫图像质量。PS 同样是一种基于深度学习的低剂量图像改善的重建算法,将低剂量高噪声的图像和相应的高剂量低噪声进行配对训练并学习其相关性。但此研究仅限于单一体型患者的腹部平扫,实验

组辐射剂量降低仅 28.32%。本研究将 AI 优化技术应用于主动脉 CTA 扫描,不局限患者体形,辐射剂量降低了 79.18%,且 80 kVp 低管电压实验组中 90% 的 Clearview+ 迭代重建算法联合 AI 成像优化技术的图像得到了客观评价指标显著优于传统 120 kVp 管电压的图像,即降噪效果达到 44.08%,SNR 和 CNR 分别提升 92.61% 和 105.05%。这一结果同时也显著优于单独使用 90% 的 Clearview+ 迭代重建算的降噪效果和 SNR 及 CNR 的提升效果。虽然从三者图像质量的主观评分比较并没有显著差异,但客观指标的提升已然说明了改善效果,为改善低剂量图像提供了更多可能。

本研究中尚存在有一些不足之处,本研究的样本量仅有 40 例,其统计结果可能存在偏倚性;其次,研究对于不同 BMI 患者水平采用了统一的对比剂用量;仅探讨对主动脉 CTA 图像质量的影响,未评估其他扫描类型及部位;未评估病灶显示情况或分疾病种类进行评估,将在以后的研究中进一步探讨。

综上所述,NeuViz 128 CT 主动脉“双低”(低辐射剂量和低碘负荷)扫描可以获得和常规扫描同样的图像质量,辐射剂量降 79.18%,对比剂用量降低 50%。ClearView+ 迭代算法的最佳档位为 90%,可以明显改善低剂量图像质量,降低图像噪声,提高图像信噪比及对比噪声比。AI 图像优化可以进一步降低图像噪声,同时保留图像的细节显示。使体重超重患者($\text{BMI} > 28 \text{ kg/m}^2$)仍可实现 80 kVp 低剂量扫描,给予患者更多保护及关爱,同时降低经济成本。

参考文献:

- [1] Hiratzka LF, Bakris GL, Beckman JA, et al. 2010 ACCF/AHA/AATS/ACR/ASA/SCA/SCAI/SIR/STS/SVM Guidelines for the diagnosis and management of patients with thoracic aortic disease. A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, American Association for Thoracic Surgery, American College of Radiology, American Stroke Association, Society of Cardiovascular Anesthesiologists, Society for Cardiovascular Angiography and Interventions, Society of Interventional Radiology, Society of Thoracic Surgeons, and Society for Vascular Medicine[J]. Circulation, 2010, 121(13): e266-e369.
- [2] Rompel O, Glöckler M, Janka R, et al. Third-generation dual-source 70kVp chest CT angiography with advanced iterative reconstruction in young children: image quality and radiation dose reduction[J]. Pediatric Radiology, 2016, 46(4): 462-472.
- [3] 余永强, 李小虎. 正理解双低的概念及其在头颈部 CT 血管成像的初步应用价值[J]. 中华医学杂志, 2014, 94(29): 2243.
- [4] 刘杰, 高剑波, 张永高, 等. 双源 CT 非线性融合技术联合低流率注射对比剂行主动脉的可行性[J]. 中华放射学杂志, 2014, 48(6): 505-506.
- [5] 武永杰, 郑敏文, 赵宏亮, 等. 低浓度对比剂联合应用低 kV 及迭代

- 重建技术的冠状动脉双源 CT 成像可行性[J]. 中华医学杂志, 2014, 8; 2260.
- [6] Fareed A, Vavere AL, Zimmermann E, et al. Impact of iterative reconstruction vs filtered back projection on image quality in 320-slice CT coronary angiography: insights from the CORE320 multicenter study[J]. *Medicine*, 2017, 96(48): e8452.
- [7] Kaln CE Jr. From images to actions: opportunities for artificial intelligence in radiology [J]. *Radiology*, 2017, 285(3): 719-720.
- [8] 赵莹, 刘爱连, 刘静红, 等. 基于深度学习的像素闪烁算法对高体质质量指数患者低剂量腹部 CT 平扫图像质量的影响[J]. *中国医学影像技术*, 2018, 34(3): 434-438.
- [9] Hu C, Yi Z, Mannudeep KK, et al. Low-dose CT with a residual encoder-decoder convolutional neural network[J]. *IEEE Transactions on Medical Imaging*, 2017, 36(12): 2524-2535.
- [10] Kang E, Min J, Ye J. A deep convolutional neural network using directional wavelets for low-dose X-ray reconstruction[J]. *Med Phys*, 2017, 44(10): 360-375.
- [11] Stefanie M, Carlo N, Julian L, et al. Effect of automated tube voltage selection, integrated circuit detector and advanced iterative reconstruction on radiation dose and image quality of 3rd generation dual-source aortic CT angiography, an intra-individual comparison[J]. *Eur J Radiology*, 2016, 85(5): 972-978.
- [12] Bolen MA, Popovic ZB, Tandon N, et al. Image quality, contrast enhancement, and radiation dose of ECG-triggered high-pitch CT versus non-ECG-triggered standard-pitch CT of the thoracoabdominal aorta[J]. *Am J Roentgenol*, 2012, 198(4): 931-938.
- [13] Committee CDDICC. The measurement, reporting, and management of radiation dose in CT. The American Association of Physicists in Medicine report no. 96. AAPM, College Park, MD. 2008
- [14] Moos SI, van Vemde DNH, Stoker J, et al. Contrast induced nephropathy in patients undergoing intravenous (IV) contrast enhanced computed tomography (CECT) and the relationship with risk factors: a meta-analysis[J]. *Eur J Radiol*, 2013, 82(9): e387-399.
- [15] 严亮亮, 杨兴华, 张永高, 等. 270 碘对比剂结合低管电压及 SAFIRE 技术在冠状动脉血管成像中的应用[J]. *实用放射学杂志*, 2014, 30(1): 143-146.
- [16] Hou P, Feng X, Liu J, et al. Low tube voltage and iterative model reconstruction in follow-up CT angiography after thoracic endovascular aortic repair: ultra-low radiation exposure and contrast medium dose[J]. *Acad Radiol*, 2018, 25(4): 494-501.
- [17] Pontana F, Pagniez J, Duhamel A, et al. Reduced-dose low-voltage chest CT angiography with sinogram-affirmed iterative reconstruction versus standard-dose filtered back projection [J]. *Radiology*, 2013, 267(2): 609-618.
- [18] Wang R, Schoepf UJ, Wu R, et al. Image quality and radiation dose of low dose coronary CT angiography in obese patients: sinogram affirmed iterative reconstruction versus filtered back projection[J]. *Eur J Radiol*, 2012, 81(11): 3141-3145.
- [19] Schoepf UJ. CT of heart: principles and applications[M]. New York, Humana Pres, 2005: 377-380.
- [20] 李岩, 束敏, 李坤成. 规范化应用"双低"扫描 CT 血管成像[J]. *中国医学影像技术*, 2018, 34(1): 128-131.
- [21] Singh S, Kalra MK, Hsieh J, et al. Abdominal CT: comparison of adaptive statistical iterative and filtered back projection reconstruction techniques[J]. *Radiology*, 2010, 257(2): 373-383.
- [22] Winklehner A, Karlo C, Puippe G, et al. Raw data-based iterative reconstruction in body CTA: evaluation of radiation dose saving potential[J]. *Eur Radiol*, 2011, 21(12): 2521-2526.
- [23] 章云港, 易本顺, 吴晨玥, 等. 基于卷积神经网络的低剂量 CT 图像去噪方法[J]. *光学学报*, 2018, 38(4): 1-7.

(收稿日期: 2018-09-10)

· 人工智能与医学影像专题 ·

人工智能冠状动脉 CT 血管成像在冠心病诊断中的应用

黄增发, 王翔

【摘要】 目的:探讨人工智能(artificial intelligence, AI)冠状动脉 CT 血管成像(CCTA)在冠心病诊断中的应用价值。**方法:**选取本院 2018 年 3—8 月,临床疑诊为冠心病的患者 50 例,对所有患者常规行 CCTA 扫描,对上传至后处理 Sigovia 工作站的图像分别由两位高年资医生进行手动冠状动脉分析软件的图像处理、完成报告和由冠心病智能辅助诊断工作站的自动识别图像中的血管和病变并生成报告。评价 AI 在冠心病中的诊断价值。**结果:**冠心病诊断中,与常规高年资医生相比,AI 对冠心病诊断的阳性预测值为 80%,阴性预测值为 70%,敏感度为 80%,特异度为 70%。AI 对冠脉病变定位识别的敏感度和特异度分别为 73.91%,82.69%。AI 对冠脉病变斑块性质识别的敏感度和特异度分别为 54.35%,81.73%。在冠脉病变管腔狭窄程度的诊断中,AI 和高年资医生对冠脉狭窄的总检出率分别为 30.67%,32%,差异无统计学意义($P>0.05$),AI 对冠脉病变管腔狭窄程度识别的敏感度和特异度分别为 60.87%,80.77%。**结论:**与高年资医生相比,AI 辅助 CCTA 对冠心病的诊断具有一定的价值。

【关键词】 人工智能;冠状动脉 CT 血管成像;冠心病;诊断

【中图分类号】 R05; R814.42; R541.4; R543.3 **【文献标识码】** A

【文章编号】 1000-0313(2018)10-1017-05

DOI:10.13609/j.cnki.1000-0313.2018.10.008

开放科学(资源服务)标识码(OSID):



Artificial intelligence-based coronary computed tomography angiography in the evaluation of coronary artery disease HUANG Zeng-fa, WANG Xiang. Department of Radiology, the Central Hospital of Wuhan, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430014, China

【Abstract】 Objective: To evaluate the clinical value of artificial intelligence (AI)-based coronary computed tomography angiography (CCTA) in the diagnosis of coronary artery disease (CAD). **Methods:** We selected 50 patients with suspected CAD from March to August 2018 in our hospital. After routinely scanning by CCTA, the images were uploaded to the post-processing sigovia workstation and processed by two senior doctors with manual coronary artery analysis software and finished the reports. Then, the AI diagnosis workstation of CAD automatically identified the coronary artery and lesions and generated reports. At last, the diagnostic value of AI in CAD was evaluated. **Results:** Compared with senior doctors, the positive predictive value, negative predictive value, sensitivity and specificity of AI for diagnosis of coronary heart disease were 80%, 70%, 80% and 70% respectively. The sensitivity and specificity of AI for detection of coronary lesions were 73.91% and 82.69% respectively. The sensitivity and specificity of AI for the identification of coronary atherosclerotic plaques were 54.35% and 81.73% respectively. The detection rates of AI and senior doctors for coronary stenosis were 30.67% and 32%, respectively. There was no significant difference between them ($P>0.05$). The sensitivity and specificity of AI for diagnosis the degree of coronary artery stenosis were 60.87% and 80.77% respectively. **Conclusion:** AI-based CCTA had certain value in the diagnosis of CAD compared with senior doctors.

【Key words】 Artificial intelligence; Coronary computed tomography angiography; Coronary artery disease; Diagnosis

作者单位: 430014 武汉, 华中科技大学同济医学院附属武汉中心医院影像科

作者简介: 黄增发(1987—), 男, 湖北武汉人, 博士, 住院医师, 主要从事心血管影像学研究工作。

通讯作者: 王翔, E-mail: wangxiang385@aliyun.com

基金项目: 湖北省自然科学基金面上项目(2013CFB376)

冠状动脉粥样硬化性心脏病,简称冠心病,是指冠状动脉粥样硬化使血管腔狭窄或闭塞,或(和)因冠状动脉功能改变(痉挛)导致心肌缺血缺氧或坏死而引起的心脏病。冠心病严重危害人类的健康,2018 年 5 月,国家心血管中心发布的《中国心血管病报告 2017》指出:近年来我国城乡冠心病的发病率和死亡率总体呈现上升态势,冠心病死亡率继续呈现快速上升趋势,农村地区冠心病的死亡率和发病率已高于城市地区。因此,对冠心病患者尽早做出诊断,并采取科学的有效治疗措施,具有非常重要的现实意义。目前,诊断冠心病的金标准是冠状动脉造影,但是此种方法属于有创检查,且价格较贵,并且会存在并发症等问题,不适合常规体检,而且在基层医院推广相对困难^[1]。相比之下,冠状动脉 CT 血管成像(coronary computed tomography angiography, CCTA)的诊断方法安全可靠无创,不但能准确观察病变管腔狭窄的程度,还能对斑块性质进行准确评估,已在临床广泛应用^[2]。然而, CCTA 诊断冠心病依赖于医生对扫描图像的后处理,需要大量的医疗资源和人力资源的支持。近年来,大数据和人工智能等前沿技术在逐渐应用到医疗领域^[3],而随着深度学习的兴起,医疗辅助诊断领域也开始与这门技术结合,将人工智能应用于 CCTA 冠心病诊断中,对于缓解医疗资源紧张具有重大现实意义。

材料与方法

1. 一般资料

选取本院 2018 年 3—8 月临床疑似冠心病患者 50 例,其中男 21 例,女 29 例,年龄 33~76 岁,平均(58.3±8.9)岁,平均 BMI 为(23.9±2.9) kg/m²(表 1)。纳入标准:患者均为自由心率,心律变异相≤5 次/分;排除标准:严重心律不齐,检查时不能配合者,碘剂过敏患者。

表 1 AI 与医生对患者冠心病总体诊断结果的比较

AI	医生		总计
	病变	无病变	
病变	24	6	30
无病变	6	14	20
总计	30	20	

注: χ^2 值=0.091, $P=0.763$ 。

2. 检查方法

向患者介绍及解释冠状动脉 CTA 检查的过程,缓解患者压力,询问碘过敏史,并行对比剂过敏皮试,研究对象右侧肘前静脉预埋 18 号留置针,预推 10~20 mL 生理盐水保证置管通畅。让患者平卧于双源 CT(Somatom Definition, Germany, Siemens)检查床,留置针接通双筒高压注射器(MALLINCKRODT, USA),再次预推 10~20 mL 生理盐水保证通畅,并告

知研究对象推药过程中一过性发热等情况,对其进行严格的屏气呼吸训练,患者双侧锁骨上及肋弓下粘贴电极,正确连接到心电监护仪检测心率,观察到稳定的心电图后,定位,准备开始扫描。先行定位相扫描确定扫描范围(气管隆突下至心脏膈面),再行增强扫描,选用非离子型对比剂碘(370 mg I/mL, Bayer Schering Pharma, Germany)60~70 mL,生理盐水 30 mL,流率 5.0~5.5 mL/s;选用对比剂示踪法选择主动脉根部感兴趣区(ROI)监测 CT 值(触发扫描 CT 值设为 100 HU),延迟 8 s,采用回顾性心电门控法进行扫描,扫描参数:管电压 100 kV,管电流 280 mAs,层厚 0.75 cm,总扫描时间约 4.0 s。

3. 图像处理和分析

对最佳收缩期、舒张期获取的图像结果与数据,上传至 Sigovia 工作站,借助冠状动脉分析软件对图像进行处理,包括最大密度投影、曲面重建和容积再现,以显示冠状动脉各节段,分析有无斑块及狭窄,由冠心病智能辅助诊断工作站(已经过>2000 例高质量 CCTA 的数据训练学习)从 Sigovia 工作站中调取图像信息,自动识别 CCTA 图像中的血管和病变生成报告,根据标准投照角度自动排版出片(图 1)。参照美国心脏病协会关于冠心病分段法标准判断右冠状动脉(RCA)、左前降支(LAD)和回旋支(LCX)是否存在狭窄、狭窄部位、狭窄程度以及斑块性质(钙化斑块、非钙化斑块和混合斑块)。管腔狭窄的计算参照 Austen 等^[4]的研究,以邻近狭窄处正常冠状动脉内径为参照。狭窄程度的判断分级按照 CAD-RADS 的参考^[5]分无狭窄、轻微狭窄(0%~25%)、轻度狭窄(25%~50%)、中度狭窄(50%~75%)、重度狭窄(>75%)和闭塞。所有图像由 1 名>5 年 CCTA 工作经验的主治医师来分析评价,并由 1 名>10 年 CCTA 工作经验的副主任医师进行分析审核。

4. 统计学分析

研究采用 SPSS 18.0 统计软件包对数据进行统计学分析,计数资料率用%表示,采用配对 χ^2 检验(McNemar's test)比较两组间的差异,计量资料用 $\bar{x} \pm s$ 表示。 $P<0.05$ 为差异有统计学意义。

结 果

1. AI 冠脉图像后处理和冠心病诊断的评价

基于 AI 的 CCTA 的图像能清晰的显示患者冠脉主干及主要分支血管,高年资医生平均后处理时间约 6 min, AI 后处理图像所需时间平均约 10 s(图 2、3)。在冠心病的诊断中,与常规高年资医生相比, AI 对冠心病诊断的阳性预测值为 80%(24/30),阴性预测值为 70%(14/20),敏感度为 80%(24/30),特异度为

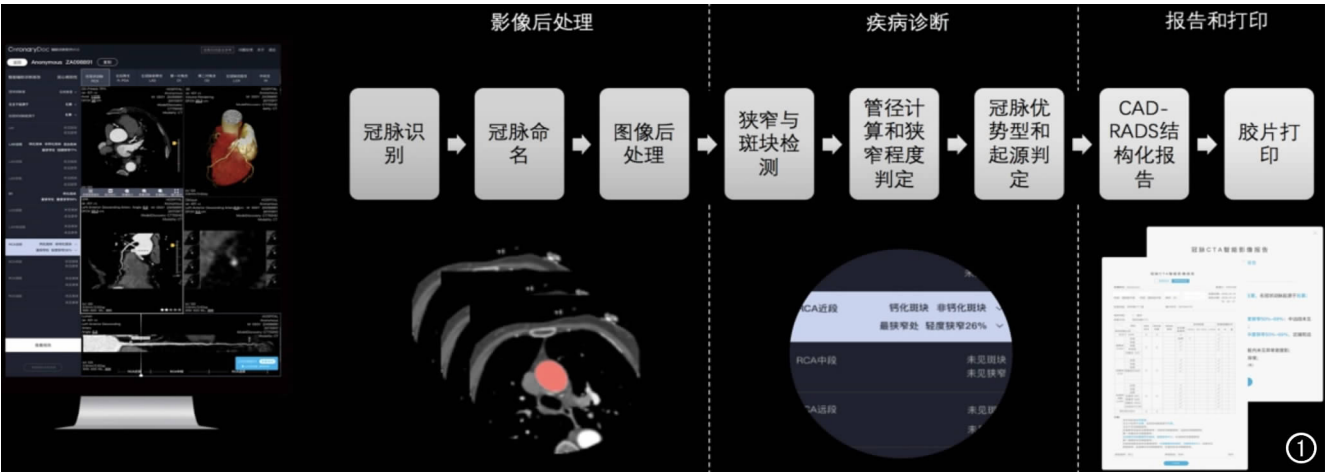


图 1 冠心病智能辅助诊断工作站 CCTA 后处理流程图

70%(14/20)。两者对冠心病总体诊断的差异无统计学意义(表 2)。

2. AI 冠脉病变定位的评价

在冠脉病变定位的诊断中,与高年资医生相比, AI 对冠脉病变定位的识别总体阳性预测值,阴性预测值,敏感度和特异度分别为 65.38%(34/52),87.76%(86/98),73.91%(34/46),82.69%(86/104)。AI 对冠脉病变 3 大支中病变定位的识别敏感度和特异度分别为 69.23%(9/13)和 81.08%(30/37);75%(18/24)和 80.77%(21/26);77.78%(7/9)和 85.37%(35/41)。两者对冠心病患者冠脉定位诊断的差异无统计学意义(表 2)。

3. AI 冠脉斑块性质的评价

对于冠脉病变斑块性质的诊断,相比高年资医生,

AI 对冠脉病变斑块性质的识别总体阳性预测值,阴性预测值,敏感度和特异度分别为 56.82%(25/44),80.19%(85/106),54.35%(25/46),81.73%(85/104)。

4. AI 冠脉管腔狭窄程度的评价

在冠脉病变管腔狭窄程度的诊断中, AI 和高年资医生对冠脉狭窄的总检出率分别为 30.67%、32%;二者对 3 大支(RCA、LAD、LCX)的检出率分别为 26%,48%,18%;30%,44%,22%,两者间差异均无统计学意义(表 3)。与高年资医生相比, AI 对冠脉病变管腔狭窄程度识别的总体阳性预测值,阴性预测值,敏感度和特异度分别为 58.33%(28/48),82.35%(84/102),60.87%(28/46)和 80.77%(84/104)。

表 2 AI 与医生对主要冠脉病变定位诊断结果的比较

AI	RCA			LAD			LCX		
	医生		总计	医生		总计	医生		总计
	病变	无病变		病变	无病变		病变	无病变	
病变	9	7	30	18	5	23	7	6	13
无病变	4	30	20	6	21	27	2	35	37
总计	30	20	—	24	26	—	9	41	—
χ ² 值	0.437			0.04			0.932		
P	0.509			0.841			0.334		

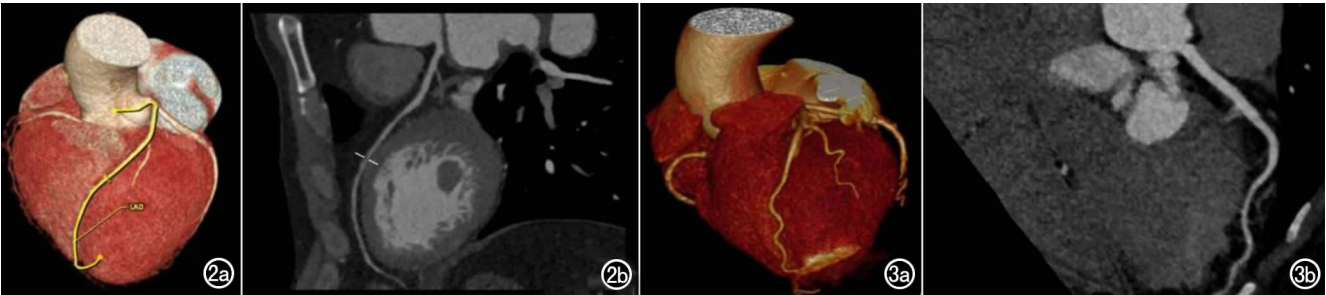


图 2 Sigovia 工作站冠脉软件后处理图。a) 冠脉 VR 图;b) LAD。图 3 相同患者 AI 自动识别和分割后处理图。a) 冠脉 VR 图;b) LAD。

表 3 AI 与医生对各冠脉狭窄检出率

狭窄	总狭窄	RCA	LAD	LCX
AI	30.67%	26%	48%	18%
医生	32%	30%	44%	22%
P 值	0.871	0.774	0.791	0.774

讨 论

心血管疾病是近年来威胁大众生命健康的主要疾病之一,在中老年人群中发病率极高,目前有年轻化的趋势,由于冠心病的发病原因主要表现为冠状动脉的狭窄,所以及早的发现狭窄,精准的定位,以及对斑块性质的准确判断是临床医师为患者给予针对性治疗的前提。双源 CCTA 图像的真实性较高,不断可以清晰地反应冠脉血管数量和分段解剖结构,而且能提供较清晰的血管和组织的对比度,诊断精确性较高^[6]。

随着医疗影像数据的不断增长、计算机硬件计算能力的提高、神经网络算法的进一步革新以及影像工作量的剧增驱动了人工智能在影像诊断领域的应用。目前,基于深度学习的 AI 的医学影像的研究应用呈不断增长趋势,涉及 X 线、CT、MRI、超声、PET 和病理等多种医疗影像,AI 辅助诊断所涉及的疾病种类也逐渐扩大^[7-12]。AI 在冠心病影像领域的研究也逐渐增多^[10-12]。Arsanjani 等^[10]通过研究机器学习方法从 SPECT 心肌灌注图像中提取特征,对疑似冠心病患者的血运重建进行预测,其敏感度和特异度分别为 73.6%和 74.7%。Motwani 等^[11]通过机器学习方法预测疑似冠心病 5 年全因死亡情况的准确性明显高于临床或者单独 CCTA 预测指标。一项超声影像的多中心临床研究表明超声心动图图像的全自动分析能快速和重复的评估左心室射血分数和纵向应变^[12]。

我国在冠心病识别模型、多参数智能诊断的研究较早^[13-15],近年来基于机器学习的神经网络的算法和硬件支持的快速发展,使得智能诊断进一步兴起。本研究的方法是基于卷积神经网络,与其他图像分类算法相比的主要优点是使用非常少的预处理^[16]。在图像分割中使用的是基于区域配准的算法,通常在肺结节和肿瘤领域应用的是基于特征配准的方法,需要检测匹配不同特征,如边缘、点、曲面等,由于冠状动脉的连续性,需要识别连续区域,因此应用了基于区域配准的算法^[16]。本研究经过 AI 自动识别 50 例 CCTA 图像并得到后处理的图像,能较好的显示冠脉主干及各分支血管,并能较准确的显示病变部位,而且能明显减少医生后处理时间。研究表明基于 AI 的 CCTA 诊断冠心病的敏感度为 80%,特异度为 70%。Kang 等^[17]基于深度学习的算法,在 CCTA 图像上诊断冠心病患者冠脉阻塞性病变和非阻塞性病变的敏感性和特

异性分别为 93%,95%。我们的研究与 Kang 的差别的原因可能是本研究的 AI 训练程度还不够高。Takx 等^[18]的研究表明与标准方法相比,应用基于机器学习的监督分类自动评估钙化积分有较好的可靠性和一致性。AI 与医生对各冠脉狭窄检出率的差别,分析可能原因是 AI 深度学习的数据来自多家医院的多个医师处理的数据,而我们挑选的医师与 AI 学习的数据库仍有不同,后期需要进一步增大 AI 学习数据库,并进行多中心的不同医师的验证。目前基于 AI 的冠心病的研究已从结构评价向功能及预后评估的研究发展^[19,20],并且已经取得了一定效果。本研究后续的基于 AI 的心功能分析,冠脉血流储备分数(FFR)分析,管腔内密度衰减梯度(TAG)以及校正的管腔内对比度衰减梯度(CCO)分析正在进一步研究中,随着训练集不断扩大,学习病例质量的不断提高,以及多中心的开展,相信基于 AI 的 CCTA 冠心病结构和功能的诊断效能会得到进一步提升。

本研究显示 AI 诊断冠心病的敏感度和特异度低于报道的 AI 肺结节筛查及肿瘤诊断价值,其一可能由于 AI 训练集的数据来自多个中心的不同设备的数据。本研究结果的验证集主要来自我院 50 个患者的数据,训练和验证的数据量还不够大;其次,多中心的数据学习,单中心的验证也会产生一定的偏差。

本研究样本量尚较小,尚需要进一步进行前瞻性、大样本量、多中心的随机对照临床研究,对此系统进行更深入的系统验证和优化,以评估其临床应用价值。

参考文献:

- [1] Kalisz K, Halliburton S, Abbara S, et al. Update on cardiovascular applications of multienergy CT [J]. Radiographics, 2017, 37(7): 1955-1974.
- [2] Douglas PS, Hoffmann U, Patel MR, et al. Outcomes of anatomical versus functional testing for coronary artery disease [J]. N Engl J Med, 2015, 372(14): 1291-1300.
- [3] Diprose W, Buist N. Artificial intelligence in medicine: humans need not apply? [J]. N Z Med J, 2016, 129(1434): 73-76.
- [4] Austen WG, Edwards JE, Frye RL, et al. A reporting system on patients evaluated for coronary artery disease. Report of the Ad Hoc Committee for Grading of Coronary Artery Disease, Council on Cardiovascular Surgery, American Heart Association [J]. Circulation, 1975, 51(4 Suppl): 5-40.
- [5] Cury RC, Abbara S, Achenbach S, et al. CAD-RADS: coronary artery disease-reporting and data system: an expert consensus document of the society of cardiovascular computed tomography (SC-CT), the American college of radiology (ACR) and the north American society for cardiovascular imaging (NASCI) [J]. Endorsed by the American College of Cardiology. JACR, 2016, 13(12 Pt A): 1458-1466.
- [6] 马俊锋, 狄海庭, 郑永宏. 双源 CT 冠状动脉血管成像在冠心病诊断中的临床应用价值 [J]. 中国血液流变学杂志, 2017, 27(4): 469-

471.

[7]

Cicero M, Bilbily A, Colak E, et al. Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs[J]. Invest Radiol, 2017, 52(5):281-287.

[8]

Setio AA, Ciompi F, Litjens G, et al. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks[J]. IEEE Trans Med Imaging, 2016, 35(5):1160-1169.

[9]

Havaei M, Davy A, Warde-Farley D, et al. Brain tumor segmentation with deep neural networks[J]. Med Image Anal, 2017, 35:18-31.

[10]

Arsanjani R, Dey D, Khachatryan T, et al. Prediction of revascularization after myocardial perfusion SPECT by machine learning in a large population[J]. J Nucl Cardiol, 2015, 22(5):877-884.

[11]

Motwani M, Dey D, Berman DS, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis [J]. Eur Heart J, 2017, 38(7):500-507.

[12]

Knackstedt C, Bekkers SC, Schummers G, et al. Fully automated versus standard tracking of left ventricular ejection fraction and longitudinal strain: the FAST-EFs multicenter study [J]. J Am Coll Cardiol, 2015, 66(13):1456-1466.

[13]

叶学松. 基于人工神经网络及心音小波分析的冠心病诊断方法的研究[J]. 浙江大学学报(自然科学版), 1999, 33(2):123-128.

[14]

徐冠, 夏克文, 徐乃勋. 基于神经网络的冠心病智能诊断系统[J]. 计算机工程与应用, 2005, 41(35):197-199.

[15]

姜燕, 李亢. 人工神经网络在冠心病中的仿真研究[J]. 计算机仿真, 2011, 28(6):243-246.

[16]

Massalha S, Clarkin O, Thornhill R, et al. Decision support tools, systems, and artificial intelligence in cardiac imaging [J]. Can J Cardiol, 2018, 34(7):827-838.

[17]

Kang D, Dey D, Slomka PJ, et al. Structured learning algorithm for detection of nonobstructive and obstructive coronary plaque lesions from computed tomography angiography [J]. J Med Imaging (Bellingham), 2015, 2(1):014003.

[18]

Takx RA, de Jong PA, Leiner T, et al. Automated coronary artery calcification scoring in non-gated chest CT: agreement and reliability [J]. PLoS One, 2014, 9(3):e91239.

[19]

Dey D, Gaur S, Ovrehus KA, et al. Integrated prediction of lesion-specific ischaemia from quantitative coronary CT angiography using machine learning: a multicentre study [J]. Eur Radiol, 2018, 28(6):2655-2664.

[20]

Frizzell JD, Liang L, Schulte PJ, et al. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: comparison of machine learning and other statistical approaches [J]. JAMA Cardiol, 2017, 2(2):204-209.

(收稿日期:2018-09-20)

下期要目

气管支气管少见原发肿瘤的影像分析	以急性化脓性甲状腺炎首诊的先天性梨状窝瘘影像学表现
儿童球形肺炎的 CT 诊断	DWI 对小胰腺癌(1A 期)和小肿块型局灶性胰腺炎的鉴别
成人孤立性左室心肌致密化不全的少见 MR 表现	诊断
不同重建层厚对定量 CT 椎体骨密度测量值的影响	胎儿颌面颈部淋巴管畸形与梨状窝瘘 MRI 特征
阴道斜隔综合征的 MRI 诊断	腰骶丛 DTI 在慢性炎性脱髓鞘性多发性神经根神经病
心肌炎 CMRI 左心室功能参数分析及与 LGE	患者中的价值
相关性研究	双重病理所致颞叶癫痫边缘系统 MRI 影像学改变

· 人工智能与医学影像专题 ·

实习医师以共同阅片和第二阅片者模式使用计算机辅助检测系统在低剂量 CT 中的应用研究

胡琼洁, 陈冲, 王玉锦, 胡玉娜, 潘月影, 杨朝霞, 艾涛, 孙子燕, 管汉雄, 夏黎明

【摘要】 目的:评估使用计算机辅助检测(CAD)系统辅助在低剂量胸部 CT 图像中,以第二阅片者模式和共同阅片模式对实习医生肺结节检出的敏感性和读片时间的影响。**方法:**前瞻性纳入 60 例胸外恶性肿瘤患者,接受 2 次连续低剂量胸部 CT 扫描。3 名实习生以盲法和随机顺序独立地对低剂量胸部 CT 图像进行三次阅片,一次使用共同阅片模式,第二次不使用 CAD,第三次使用第二阅片模式。记录真阳性率(TP),假阳性率(FP)和阅读时间。以 3 名高年资胸部放射科医师(>20 年)检出的肺结节(包括对 CAD 检出结节认可)的一致意见为真结节参照标准。**结果:**共同阅片模式(259 s)的阅读时间明显短于不使用 CAD 的阅读时间(341 s, $P<0.01$)和第二阅片模式(363 s; $P<0.01$)。低剂量胸部 CT, CAD 作为共同阅片者(CR)的敏感性(67.36%)显著高于不使用 CAD 的敏感性(40.83%, $P<0.001$),而与 CAD 作为第二阅片者(SR)的敏感性(68.34%; $P>0.05$)无统计学差异。CAD 作为 CR 的 FP 率(5.31/CT)高于无 CAD 的 FP 率(3.28/CT, $P<0.0001$),而与 CAD 作为 SR 的 FP 率(5.23/CT, $P>0.05$)无统计学差异。使用 CAD 在共同阅片模式下的 FOM 明显高于不使用 CAD 的 FOM(分别为 0.59 和 0.44, $P<0.001$),使用 CAD 在第二阅片模式的平均 FOM 值为 0.60,与共同阅片模式无统计学意义($P=0.675$)。**结论:**在低剂量 CT 图像,使用 CAD 显著提高实习医师对包括小结节在内的肺结节诊断敏感性;并且以共同读片模式使用 CAD 显著地缩短了实习医师的阅片时间。

【关键词】 体层摄影术, X 线计算机; 辐射剂量; 肺结节; 计算机辅助检测

【中图分类号】 R814.42; R734.2; R05 **【文献标识码】** A **【文章编号】** 1000-0313(2018)10-1022-08

DOI:10.13609/j.cnki.1000-0313.2018.10.009

开放科学(资源服务)标识码(OSID):



A computer-aided detection (CAD) system on low-dose chest CT images in concurrent-reader and second-reader modes: influence on interns HU Qiong-jie, CHEN Chong, Wang Yu-jing, et al. Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China

【Abstract】 Objective: To assess the value of a computer-aided detection (CAD) system in second-reader(SR) and concurrent-reader(CR) models for the detection of pulmonary nodules and time efficiency for interns on low-dose chest CT images. **Methods:** A total of 60 patients with extra-thoracic malignancies, who underwent low-dose CT scan, were prospectively included. 3 interns independently evaluated images of 60 patients in a blinded and randomized manner thrice, once with CR twice without CAD, and third with SR in the images of low-dose chest CT. True positive rates (TP), false positive rates (FP) and reading times were recorded. Free-response receiver operating characteristic analysis was also used for the statistical analysis. The consensus reading of CT images by three senior chest radiologists (>20 years) and those obtained using the CAD system served as the reference standard. **Results:** Reading times were significantly shorter in CR mode (259s) than those without CAD (341s, $P<0.01$) and in SR mode (363s; $P<0.01$). Sensitivity of reading with CAD as CR (67.36%) was significantly higher than reading without CAD (40.83%, $P<0.001$) and not statistically different in CAD as SR (68.34%; $P>0.05$) in the low-dose setting. The FP rates using CAD as CR (5.31/CT) were higher than without CAD (3.28/CT, $P<0.0001$), while not statistically different in CAD as SR

作者单位: 430030 武汉, 华中科技大学同济医学院附属同济医院放射科

作者简介: 胡琼洁(1982—), 女, 湖北黄冈人, 博士, 主治医师, 主要从事胸部影像诊断工作。

通讯作者: 夏黎明, E-mail: cjr.xialiming@vip.163.com

基金项目: 武汉市科技厅重点项目“人工智能深度学习对肺结节检测的多维度分析与智能影像组学研究(2018060401011326)”

(5.23/CT, $P > 0.05$). The FOM for the 3 interns in the CR mode was statistically significant better than that without CAD (0.59 vs 0.44, respectively; $P < 0.001$), and the average FOM value in the SR mode was 0.60, giving not statistically different between the CR and SR modes ($P = 0.675$). **Conclusion:** CAD benefited interns in the low-dose settings including small nodules, which both significantly improved diagnostic performance and shortened the reading times.

【Key words】 Tomography, X-ray computed; Radiation dosage; Pulmonary; Computer-aided detection

肺是恶性肿瘤转移的第三大常见部位,仅次于淋巴结和肝脏。大约 1/3 的患者在疾病过程中会发生肺转移。早期识别肺部转移对于患者至关重要,可通过更积极和多学科综合治疗的方法,如肺转移瘤的切除术或放疗等,治疗后患者可长期存活甚至治愈^[1-3]。计算机断层扫描(CT),而非胸部 X 线摄影检查,是评估肺转移、随访和监测治疗反应的最有效且安全的检查手段^[4]。此外,这类患者需要反复多次接受胸部 CT 检查,其累计辐射剂量增加,低剂量计算机断层扫描(low dose CT, LDCT)被推荐用于此类患者。

随着胸部 CT 在健康体检及肿瘤患者随访复查中越来越广泛的应用,影像科医生的工作量与日俱增。目前我国医学影像数据年增长率约 30%,而放射科医师数量增长率约 4%,两者严重的不匹配意味着放射科医师处理影像数据的压力越来越大,承担巨大的负荷。同时,为了最大限度提高肺结节的检出率和病灶细节的显示,LDCT 筛查需要在薄层、不同对比度的 CT 图像(层厚 1~1.25 mm)上进行观察,加上至少 3 种图像对比(肺窗、纵隔窗、骨窗),每个患者所需观察的图像将近 1000 幅,进一步增加了影像医师的工作负担,容易产生视觉疲劳,降低了工作效率和诊断准确性,导致对肺结节的漏检和误判^[5]。研究表明胸外恶性肿瘤患者肺结节 CT 漏诊约 1/3,其中约 20%的结节是恶性的^[6]。大部分漏诊肺结节常因体积较小而被漏诊。特别对于刚接触影像的实习学生而言,CT 漏诊率更高。

研究表明使用计算机辅助检测系统(computer-aided detection, CAD)明显提高肺结节检出的敏感性^[7-11]。Awai 等^[7]阐述了在标准剂量 CT 扫描中,使用 CAD 的肺结节的检出率高于不使用 CAD 的检出率。Doo 等^[11]研究表明使用体模 CAD 对于肺结节检出率明显提高。目前尚缺乏 CAD 对低剂量 CT 肺结节检出率影响,特别是详细的临床研究的相关报道。Zhao 等^[5]发现 CAD 在肺癌筛查中具有重要价值,可提高肺结节检测的敏感性。可改善评估低剂量肺癌筛查 CT 的阅片者的一致性和随访建议,在排除小结节的情况下^[12,13]。有两篇文献在标准剂量胸部 CT 中探讨 CAD 对胸外恶性肿瘤患者肺结节的检出率^[14,15],

并未见关于 CAD 在低剂量胸部 CT 对胸外恶性肿瘤患者检测结节敏感性的报道。本研究我们假设基于卷积神经网络(CNN)的人工智能 CAD 系统对于实习医生在低剂量两种 CT 扫描剂量中对肺结节的诊断性能显著提高,包括对胸外恶性肿瘤患者的小结节。评估 CAD 对亚实性肺结节检出的敏感性相关研究相对较少^[16],因此本研究根据结节密度特征(实性,亚实性和钙化结节)、结节直径(<4 mm, $4\sim 8$ mm, >8 mm)和结节分布(胸膜下或非胸膜下)分别比较肺结节检测的敏感性。

使用 CAD 发现肺结节的阅片模式主要有 3 种:包括第一阅片者,第二阅片者(second-reader, SR)和共同阅片者(concurrent-reader, CR)。研究表明使用 CAD 作为第二阅片者花费的时间更长,因为需要阅片两次^[17]。许多研究表明了第二阅片者模式的有效性^[7-10,18-23],而迄今仅有少数关于胸部 CT 共同阅片模式的相关研究。先前的研究中 CAD 共同阅片模式的评估结果不一致。如,一项研究中表明共同阅片模式,肺结节检测的灵敏度从 56.5% 提高到 66.5%^[24],但另一项研究中^[17]使用 CAD 共同阅片模式诊断的敏感性无显著性差异。Beyer 等^[17]发现 CAD 共同阅片模式可以显著减少阅片时间,Matsumoto 等^[24]研究未发现 CAD 共同阅片模式对阅片时间的影响。此外,尚未见 CAD 共同读片模式在低剂量胸部 CT 情况下对肺结节检测影响的相关报道。我们假设, CAD 的共同读片模式提高肺结节诊断敏感性或减少阅片时间。

本研究系统地评价实习医师使用 CAD 作为第二阅片者和共同阅片者在低剂量胸部 LDCT 肺结节检出中的作用。

材料与方法

1. 资料

本前瞻性研究获得华中科技大学同济医学院附属同济医院伦理委员会的批准,并且所有受试者签署了书面知情同意书。2017 年 7—12 月,共招募了 129 名胸外恶性肿瘤患者,这些患者计划进行胸部多层螺旋 CT 检查以监测或者评价肺转移情况。招募受试者时排除年龄 <18 岁和孕妇。本研究排除了肺肿块

>100 mm 或肺疾病如严重肺气肿、急性肺部感染、胸腔积液、肺结核或肺结节病的患者。60 名患者被纳入研究。其中,104 例(88.89%)检测出肺结节,13 例(11.11%)未检测出肺结节。

2. CT 扫描方法

所有患者均以仰卧位吸气末屏气时扫描。采用 64 排探测器 CT 扫描(Discovery CT750 HD, GE Healthcare)。扫描参数:准直器宽度 64×0.625 mm;螺距 1.375:1;图像矩阵 512×512。行低剂量 CT(BMI≥22:120 kVp, n=88; BMI<22:100 kVp, n=41; 20 mAs; NI=30)平扫。层厚、层间距分别为 1.25 mm、1.25 mm,使用 ASIR 60%及标准重建方式。

3. CAD 软件

使用 CAD 系统(推想公司)CAD 标记,该系统使用卷积神经网络(CNN)模型。将 CAD 应用于标准剂量和低剂量胸部 CT。软件识别的候选病灶用圆圈标记。每个病变仅一个层面上显示 CAD 标记(即显示病变的最大直径层面)。

4. 检测

阅读疲劳是影响观察者结节诊断敏感性的一个重要因素。并且不同医生诊断敏感性存在差异。为此,随机选择 60 例用于研究在低剂量胸部 CT 中,实习医师在 3 种不同阅片模式下肺结节检测的敏感性。3 名实习医师指定为观察员 A、B、C(工作年限小于 1 年)。3 名实习生在没有 CAD 的情况下独立读片,8 周后以 CR 模式读片,6 周后以 SR 模式读片。观察者以 1 到 4 共 4 级对每个标记结节的进行评级,其中 4 代表最高置信度(确定是结节),得分 1 代表最低置信度(可能不是结节)^[25]。他们根据自己的需要调整窗宽、窗位、放大和缩小、和使用最大密度投影(MIP)厚片图像。记录真阳性率(TP)和假阳性率(FP)和读取时间。通过检测的结节数除以结节总数来计算 TP 比率。FP 率计算为每个 CT 图像中假阳性结节的个数。此外,要求 3 名读片者手动测量结节的平均直径。采用 FROC(free-response receiver operating characteristic)解析(JAFROC 4.2.1, www.devchakraborty.com)分析了独立读片、CR 模式和 SR 模式下的检出性能^[25-28]。肺结节参考标准是由 3 个没有参与读片研究的胸部放射科医师(20 年以上经验)所达成的共识

确定:3 名医生阅片 2 次,两次阅片之间间隔 3 个月。所有高年资胸部放射科医师都认同的 CAD 检出的结节亦纳入参考标准。在研究之前,所有观察者都接受了使用 CAD 系统的培训。此外,根据结节最长直径(<4 mm, 4~8 mm, >8 mm),结节性质(实性,亚实性和钙化结节)和位置(胸膜下或非胸膜下)对肺结节检测的敏感性进行了总体比较和每个亚组的比较。

5. 数据分析

应用 SPSS 18.0 统计学软件进行数据分析。使用 Kolmogorov-Smirnov 检验验证独立读片,第二读片 SR 模式,共同读片 CR 模式下低剂量 CT 图像真阳性率,假阳性率以及读片时间是否满足正态分布。符合正态分布用配对 t 检验或方差分析来比较正态变量,不满足正态分布使用 Wilcoxon 秩和检验或 K-W 检验。FROC 曲线下面积使用 ANOVA 方法。P<0.05 为差异有统计学意义。

结 果

1. 患者特征

首批参加研究受试者 81 例,其中女 44 例,男 37 例,平均年龄(50.63±9.54)岁,年龄范围 29~78 岁; BMI(23.82±1.5)kg/m², BMI 范围 22~28kg/m²; 采用 120 kVp 的低剂量 CT 扫描。第二批参加研究受试者 36 例,其中女 18 例,男 18 例,平均年龄(48.53±13.25)岁,年龄范围 22~76 岁, BMI(20.58±0.87)kg/m², BMI 范围(18.5~22)kg/m²; 采用 100 kVp 的低剂量 CT 扫描。经初步分析,第二批参加研究的 36 例受试者的辐射剂量更低。随机选取 60 个病例纳入本次研究。诊断参考标准:低剂量 CT(LDCT)图像共 338 个结节(98.48%)。

2. 阅片者检测结果

3 名观察者在独立阅读、CR 和 SR 模式下的阅读时间如表 1。低剂量 CT 图像, CR 模式[(259±168) s] 阅读时间明显短于独立阅片模式[(341±206) s, P<0.0001]和 SR 模式[(363±165) s, P<0.0001]。

低剂量 CT, CAD 作为 CR 模式的平均敏感度(67.36%)显著高于不使用 CAD 的平均敏感度(40.83%, P<0.001),而与 CAD 作为 SR 模式的平均敏感度(68.34%, P>0.05)无统计学差异。在亚组分

表 1 低剂量 CT:无 CAD、共同阅片和第二阅片模式检测结节所用时间

实习医师	UR	CR	SR	P1	P2	P3
1	478±133s	351±182s	485±175s	<0.0001	0.802	<0.0001
2	169±66s	133±70s	231±70s	0.014	<0.0001	<0.0001
3	377±237s	292±147s	374±117s	0.061	0.943	<0.0001
1~3 平均	341±206s	259±168s	363±165s	<0.0001	0.207	<0.0001

注:UR=无 CAD 模式(unaided reading mode), SR=第二读片者模式(second-reader mode); P1:UR vs CR; P2:UR vs SR; P3:CR vs SR。

表 2 低剂量 CT 无 CAD 模式、共同阅片和第二阅片者模式结节检测结果(真阳性率和假阳性率)

实习 医师	敏感度						假阳性率					
	UR	CR	SR	P1	P2	P3	UR	CR	SR	P1	P2	P3
1	29.59%	65.09%	72.49%	<0.0001	<0.0001	0.948	1.4/ct	4.7/ct	4.43/ct	<0.0001	<0.0001	1.000
2	38.17%	64.79%	63.02%	0.013	0.002	0.307	2.83/ct	4.58/ct	4.63/ct	0.005	0.005	0.860
3	54.73%	72.19%	69.53%	0.029	0.058	0.531	5.6/ct	6.68/ct	6.63/ct	0.030	0.035	0.310
1~3 平均	40.83%	67.36%	68.34%	<0.0001	<0.0001	0.480	3.28/ct	5.31/ct	5.23/ct	<0.0001	<0.0001	0.215

注:UR=无辅助模式(unaided reading mode), SR=第二读片者模式(second-reader mode);P1:UR vs CR; P2:UR vs SR; P3:CR vs SR。

表 3 低剂量 CT:无 CAD,共同阅片模式和第二阅片者模式肺结节检出结果

阅片 模式	结节密度			结节最大横径(mm)			结节位置	
	实性	亚实性	钙化	≤4	4~8	≥8	胸膜下	非胸膜下
UR	41.87%	27.93%	50.00%	37.40%	64.37%	97.87%	41.85%	40.16%
CR	67.50%	63.96%	71.67%	65.24%	82.76%	100.00%	61.65%	71.06%
SR	67.26%	63.96%	91.67%	66.01%	86.21%	100.00%	65.66%	70.08%
P1	<0.0001	<0.0001	0.017	<0.0001	0.001	0.157	<0.0001	<0.0001
P2	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.157	<0.0001	<0.0001
P3	0.952	1.000	0.001	0.605	0.366	1.000	0.048	0.628

注:UR=无辅助模式(unaided reading mode), SR=第二读片者模式(second-reader mode);P1:UR vs CR; P2:UR vs SR; P3:CR vs SR。

析中,对于<8 mm 的结节,CR 模式的平均敏感性高于不使用 CAD($P<0.001$)。CAD 作为 CR 模式的 FP 率(5.31/CT)高于不使用 CAD (3.28/CT, $P<0.0001$),而与 CAD 作为 SR 模式的 FP 率(5.23/CT, $P=0.263$)无统计学差异(表 2、3)。

观察者平均的 JAFROC FOM 在 CR 模式下为 0.59(95%置信区间:0.53,0.65),在无 CAD 模式下为 0.44(95%置信区间:0.37,0.52),差异有显著性差异($P<0.0001$);SR 模式的平均 FOM 值为 0.60(95%置信区间为 0.51,0.69),CR 模式与 SR 模式的 FOM 值无统计学差异(表 4,图 1、2, $P=0.675$)。

表 4 结节检测的结果:FOM

实习 医师	UR	CR	SR	P1	P2	P3
1	0.41	0.61	0.63	<0.0001	<0.0001	0.4987
2	0.42	0.58	0.57	<0.0001	<0.0001	0.752
3	0.49	0.58	0.60	0.0085	0.003	0.7191
1~3 平均	0.44	0.59	0.60	<0.0001	<0.0001	0.675

注:UR=无辅助模式(unaided reading mode), SR=第二读片者模式(second-reader mode);P1:UR vs CR; P2:UR vs SR; P3:CR vs SR。

讨 论

CT 检查是医疗辐射暴露最重要因素。近年来,随着人们对辐射危害的关注,低剂量胸部 CT 越来越广泛地应用^[29]。美国国家肺癌筛查试验(national lung screening trial,NLST)表明,低剂量 CT 降低了

肺癌死亡率^[30]。此外,累计辐射剂量的增加引起人们的注意,尤其对于肺部感染的监测,评价和监测肺部转移治疗反应,这类患者需要反复多次进行胸部 CT 检查^[4]。因此也被推荐使用低剂量 CT 检查。本研究系统地评估 CAD 作为第二阅片者和共同阅片者两种模式下肺结节检出的效能。尽管实习医师在低剂量胸部 CT 图像上容易漏诊结节,且需要很长的时间来检测结节,但 CAD(在第二读片者和共同读片模式中)显著提高了肺结节诊断敏感性,同时使用 CAD 作为共

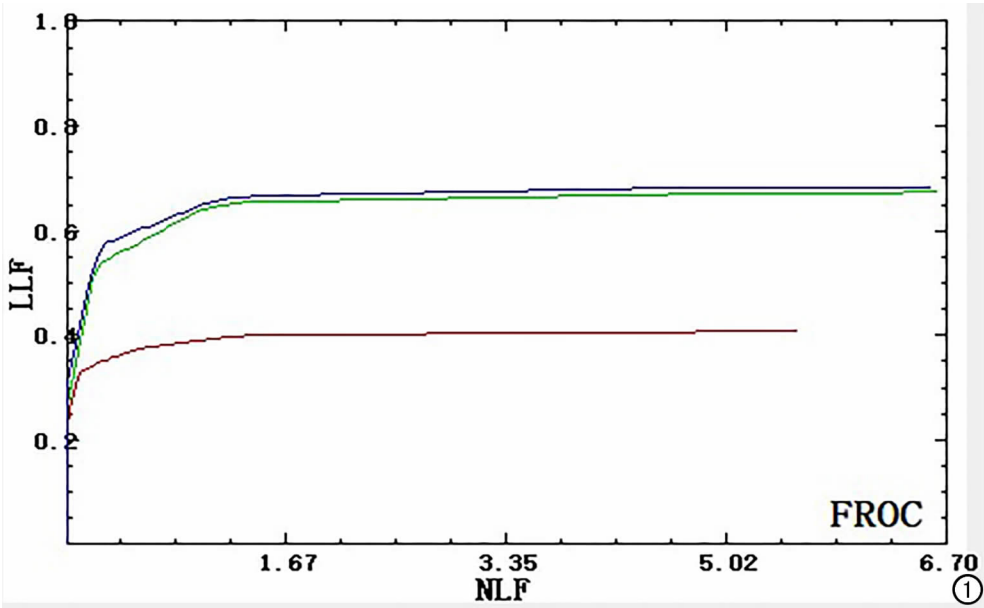


图 1 FROC 曲线代表 3 名实习医师使用共同读片模式(绿线)、第二读片者模式(紫线)和无 CAD 模式(红线)检测肺结节的平均性能。



图 2 48 岁,男,直肠癌患者。a) 实习医师不使用 CAD 时无一人在低剂量 CT 上检测到结节(直径约 2mm); b) 在共同阅片和第二读片者模式中,2 名实习医师正确识别了该结节; c) 7 个月后,肺结节明显增大。

同读片模式缩短了阅片时间。此外,低剂量并没有影响实习医师肺结节检出的假阳性率。

许多研究显示了第二读片者 SR 模式^[7-10,12,13]的有效性。同样,我们发现不论是常规剂量还是低剂量 CT 对于 <8 mm 结节,使用 CAD 作为第二阅片者比较于未使用 CAD,显著提高实习医生的诊断准确度($P < 0.001$)。

使用 CAD 系统作为第二阅片者模式,不可避免地延长读取时间。虽然有许多研究表明使用 CAD 作为第二阅片者模式^[7-10,18-23]提高肺结节敏感性的有存在性,但关于 CAD 作为共同阅片模式的研究很少^[17,24,31]。此外,使用 CAD 作为共同阅片模式在诊断肺结节的敏感性和阅片时间方面存在争议,并且在低剂量胸部 CT,以共同阅片模式使用 CAD 对肺结节诊断的影响还未见报道。Beyer 等的研究使用 CR 模式与未使用 CAD 相比,显著地减少了阅片时间,但对诊断敏感度没有影响^[17]。他们认为共同读片模式的灵敏度没有增加,可能是由于阅片速度的加快,对 CT 图像的视觉分析不太彻底。最近一项研究表明相比独立阅片 CR 模式阅片时间并没有缩短,但诊断敏感度提高^[24]。相反,Foti 等^[31]研究发现 CR 模式既没有缩短阅片时间,也没有显著地提高肺结节检出的敏感性。本研究我们比较了未使用 CAD 与 CAD 作为共同读片模式在低剂量胸部 CT,阅片时间和肺结节诊断敏感度的差异。我们发现与未使用 CAD 模式相比,CR 模式不仅显著地减少了读片时间,而且还增加了实习生在 LDCT 肺结节检测的敏感度。CR 模式下(259 s)的读片时间明显短于无 CAD 模式(341 s, $P < 0.01$)和 SR 模式(363 s, $P < 0.01$)。未使用 CAD 模式的阅片时间在我们的研究中是直接记录的;而先前的研究中^[17,24,31],未使用 CAD 模式下的阅片时间则是从先前的 2 次阅片中提取的,所以本研究的结果可靠性更高。共同阅片模式的敏感度(67.36%)高于未使用

CAD 模式(40.83%, $P < 0.001$),而与第二阅片模型相比,敏感度(68.36%, $P > 0.05$)无统计学差异。在低辐射剂量水平,图像噪声随着辐射剂量减少而增加^[5]。尽管实习医师在低剂量胸部 CT 图像上容易漏诊结节,且需要较长的时间才能检出结节,但使用 CAD 在共同阅片模式可显著提高诊断敏感度,且缩短阅片时间。随着使用 CAD 的敏感度的增加,假阳性率也增加^[32]。我们发现 CR 模式中的平均 FP 率(5.31/CT)比未使用 CAD 模式高(3.28/CT, $P < 0.0001$),而使用 CAD 在 CR 和 SR 两种模式平均 FP 率(5.23/CT, $P = 0.263$)无统计学差异。

JAFROC FOM 于 2004 年首次提出^[26],是研究真阳性率与假阳性率新指标^[25,27,28]。值得注意的是,我们的研究在低剂量 CT 使用 CAD 作为共同阅片者模式的 JAFROC FOM 显著高于未使用 CAD ($P < 0.001$),而 CAD 作为共同阅片者比较于作为第二阅片者 JAFROC FOM 无统计学差异($P > 0.05$)。因此,CR 模式有益于低剂量 CT 实习医师。

本研究局限性在于首先没有可利用的外部参考标准(如组织学证据),而是 3 个高年资胸部放射科医师(超过 20 年的工作时间)两次 CT 阅片的共识作内部参照标准。这一局限性在评价 CAD 效能的研究中非常普遍。其次,真阳性率相对较低,假阳性率相对较高。然而,考虑到我们研究中 90% 的结节直径 < 4 mm,也相对解释为什么本研究中的敏感度相对较低。其次,人类解读图像时的差异不可避免,需要更多中心的研究来证实我们的观察结果。最后,本研究结果仅代表一个特定算法的在特定参数时 CAD 的表现,不能直接用于其它 CAD。然而,本研究的重点是读片者和 CAD 之间的交互,而非 CAD 的性能本身。

总之,我们在临床上证明了尽管实习医师在低剂量胸部 CT 独立阅片时容易漏诊结节,但 CAD 在共同读片和第二读片者模式均显著提高包括小结节在内的

肺结节的诊断敏感性。并且以共同读片模式使用 CAD 显著地缩短了读片时间。

参考文献:

- [1] Hornbeck K, Ravn J, Steinbruchel DA. Outcome after pulmonary metastasectomy; analysis of 5 years consecutive surgical resections 2002—2006[J]. *J Thorac Oncol*, 2011, 6(10):1733-1740.
- [2] Casiraghi M, De Pas T, Maisonneuve P, et al. A 10-year single-center experience on 708 lung metastasectomies; the evidence of the 'international registry of lung metastases'[J]. *J Thorac Oncol*, 2011, 6(8):1373-1378.
- [3] Corona-Cruz JF, Domínguez-Parra LM, Saavedra-Pérez D, et al. Lung metastasectomy; long-term outcomes in an 18-year cohort from a single center[J]. *Surg Oncol*, 2012, 21(3):237-244.
- [4] Mohammed TL, Chowdhry A, Reddy GP, et al. ACR appropriateness criteria screening for pulmonary metastases[J]. *J Thorac Imaging*, 2011, 26(1):W1-W3.
- [5] Zhao Y, de Bock GH, Vliegenthart R, et al. Performance of computer-aided detection of pulmonary nodules in low-dose CT: comparison with double reading by nodule volume[J]. *Eur Radiol*, 2012, 22(10):2076-2084.
- [6] Cerfolio RJ, Bryant AS, McCarty TP, et al. A prospective study to determine the incidence of non-imaged malignant pulmonary nodules in patients who undergo metastasectomy by thoracotomy with lung palpation[J]. *Ann Thorac Surg*, 2011, 91(6):1696-1700.
- [7] Awai K, Murao K, Ozawa A, et al. Pulmonary nodules at chest CT; effect of computer-aided diagnosis on radiologists' detection performance[J]. *Radiology*, 2004, 230(2):347-352.
- [8] Marten K, Seyfarth T, Auer F, et al. Computer-assisted detection of pulmonary nodules; performance evaluation of an expert knowledge-based detection system in consensus reading with experienced and inexperienced chest radiologists[J]. *Eur Radiol*, 2004, 14(10):1930-1938.
- [9] Brown MS, Goldin JG, Rogers S, et al. Computer-aided lung nodule detection in CT; results of large-scale observer test[J]. *Acad Radiol*, 2005, 12(6):681-686.
- [10] Das M, Mühlenbruch G, Mahnken AH, et al. Small pulmonary nodules; effect of two computer-aided detection systems on radiologist performance[J]. *Radiology*, 2006, 241(2):564-571.
- [11] Doo KW, Kang EY, Yong HS, et al. Comparison of chest radiography, chest digital tomosynthesis and low dose MDCT to detect small ground-glass opacity nodules; an anthropomorphic chest phantom study[J]. *Eur Radiol*, 2014, 24(12):3269-3276.
- [12] Zhao Y1, de Bock GH, Vliegenthart R, et al. Performance of computer-aided detection of pulmonary nodules in low-dose CT; comparison with double reading by nodule volume[J]. *Eur Radiol*, 2012, 22(10):2076-2084.
- [13] Jeon KN, Goo JM, Lee CH, et al. Computer-aided nodule detection and volumetry to reduce variability between radiologists in the interpretation of lung nodules at low-dose screening computed tomography[J]. *Invest Radiol*, 2012, 47(8):457-461.
- [14] Schramm A, Wormanns D, Leschber G, et al. Reliability of a computer-aided detection system in detecting lung metastases compared to manual palpation during surgery[J]. *Interact Cardiovasc Thorac Surg*, 2011, 12(1):20-23.
- [15] Vassallo L, Traverso A, Agnello M, et al. A cloud-based computer-aided detection system improves identification of lung nodules on computed tomography scans of patients with extra-thoracic malignancies[J]. *Eur Radiol*, 2018 Jun 15. DOI:10.1007/s00330-018-5528-6[Epub ahead of print].
- [16] Godoy MC, Kim TJ, White CS, et al. Benefit of computer-aided detection analysis for the detection of subsolid and solid lung nodules on thin- and thick-section CT[J]. *AJR Am J Roentgenol*, 2013, 200(1):74-83.
- [17] Beyer F, Zierott L, Fallenberg EM, et al. Comparison of sensitivity and reading time for the use of computer-aided detection (CAD) of pulmonary nodules at MDCT as concurrent or second reader[J]. *Eur Radiology*, 2007, 17(11):2941-2947.
- [18] Goo JM, Kim HY, Lee JW, et al. Is the computer-aided detection scheme for lung nodule also useful in detecting lung cancer? [J]. *J Comput Assist Tomogr*, 2008, 32(4):570-575.
- [19] Hirose T, Nitta N, Shiraishi J, et al. Evaluation of computer-aided diagnosis (CAD) software for the detection of lung nodules on multidetector row computed tomography (MDCT): JAFROC study for the improvement in radiologists' diagnostic accuracy [J]. *Acad Radiol*, 2008, 15(12):1505-1512.
- [20] Yanagawa M, Honda O, Yoshida S, et al. Commercially available computeraided detection system for pulmonary nodules on thin-section images using 64 detectors-row CT; preliminary study of 48 cases[J]. *Acad Radiol*, 2009, 16(8):924-933.
- [21] Sahiner B, Chan HP, Hadjiiski LM, et al. Effect of CAD on radiologists' detection of lung nodules on thoracic CT scans; analysis of an observer performance study by nodule size[J]. *Acad Radiol*, 2009, 16(12):1518-1530.
- [22] White CS, Pugatch R, Koonce T, et al. Lung nodule CAD software as a second reader; a multicenter study[J]. *Acad Radiol*, 2008, 15(3):326-333.
- [23] Teague SD, Trilakis G, Dharaiya E. Lung nodule computer-aided detection as a second reader; influence on radiology residents[J]. *J Comput Assist Tomogr*, 2010, 34(1):35-39.
- [24] Matsumoto S, Ohno Y, Aoki T, et al. Computer-aided detection of lung nodules on multidetector CT in concurrent-reader and second-reader modes; A comparative study[J]. *Eur J Radiology*, 2013, 82(8):1332-1337.
- [25] Vikgren J, Zachrisson S, Svåkvist A, et al. Comparison of chest tomosynthesis and chest radiography for detection of pulmonary nodules; human observer study of clinical cases[J]. *Radiology*, 2008, 249(3):1034-1041.
- [26] Chakraborty DP, Berbaum KS. Observer studies involving detection and localization; modeling, analysis, and validation[J]. *Med Phys*, 2004, 31(8):2313-2330.
- [27] Dobbins JT 3rd, McAdams HP, Sabol JM, et al. Multi-institutional evaluation of digital tomosynthesis, dual-energy radiography, and conventional chest radiography for the detection and management of pulmonary nodules[J]. *Radiology*, 2017, 282(1):236-250.
- [28] Yamada Y, Shiomi E, Hashimoto M, et al. Value of a computer-aided detection system based on chest tomosynthesis imaging for the detection of pulmonary nodules[J]. *Radiology*, 2018, 287(1):333-339.

- [29] Schauer DA, Linton OW. NCRP Report No. 160, Ionizing Radiation Exposure of the Population of the United States, medical exposure—are we doing less with more, and is there a role for health physicists? [J] Health Phys, 2009, 97(1): 1-5.
- [30] National Lung Screening Trial Research Team, Church TR, Black WC, Aberle DR, et al. Results of initial low-dose computed tomographic screening for lung cancer[J]. New England J Medicine, 2013, 368(21): 1980-1991.
- [31] Foti G, Faccioli N, DOnofrio M, et al. Evaluation of a method of computer-aided detection (CAD) of pulmonary nodules at computed tomography[J]. Radiol Med, 2010, 115(6): 950-961.
- [32] Wielpütz MO, Wroblewski J, Lederlin M, et al. Computer-aided detection of artificial pulmonary nodules using an ex vivo lung-phantom: influence of exposure parameters and iterative reconstruction[J]. Eur J Radiol, 2015, 84(5): 1005-1011.

(收稿日期: 2018-09-20)

《放射学实践》杂志微信公众平台开通啦!

2015 年 6 月,《放射学实践》杂志入选北京大学和北京高校图书馆期刊工作研究会共同主持的国家社会科学基金项目“学术期刊评价及文献计量学研究”研究成果——《中国核心期刊要目总览》。这是继 1999, 2008 年之后的第 3 次入选临床医学/特种医学类核心期刊。

《放射学实践》杂志微信公众平台立足于准确地传递医学影像领域的最新信息,致力于为关注医学影像领域的广大人士服务。欢迎大家通过微信平台,以文字、图片、音频和视频等形式与我们互动,分享交流最新的医学影像资讯。您还可以通过微信平台免费阅读及搜索本刊所有发表过的论文,投稿作者可以查询稿件状态等。

您可以通过以下方式关注《放射学实践》杂志微信公众平台:

1. 打开微信,通过“添加朋友”,在搜索栏里直接输入“放射学实践”进行搜索。
2. 在“查找微信公众号”栏里输入“放射学实践”即可找到微信公众号,点击“关注”,添加到通讯录。
3. 打开微信,点击“扫一扫”,手机镜头对准下面的二维码,扫出后点击关注即可。



期待您的加入!

· 人工智能与医学影像专题 ·

乳腺 X 线 AI 智能病灶检测

李欣, 梁森, 黄正南, 夏晨, 张荣国, 吴孝掌, 赖原仲

【摘要】 目的:初步探究使用最新的人工智能(AI)检测方法检测乳腺 X 线病灶,包括肿块,乳腺内淋巴结和钙化,为进一步的乳腺钼靶 X 线 AI 智能系统应用提供初步验证。**方法:**使用深度学习目标检测 Faster R-CNN 算法,训练人工标注的 1892 例乳腺钼靶 X 线数据集,在 400 例测试数据集上验证 AI 病灶检测的性能。**结果:**AI 智能检测出肿块 526 个(共 689 个),乳腺内淋巴结 912 个(共 1098 个),圆形钙化 52 个(共 73 个),环形钙化 519 个(共 692 个),粗糙钙化 353 个(544 个),其敏感度分别为 76.4%,83.1%,71.2%,75.0%,64.9%,假阳率分别为 35.7%,38.6%,0.9%,0.6%,18.4%。**结论:**AI 能较好地检测出乳腺钼靶 X 线影像中的肿块,淋巴结和钙化,为更深入的 AI 智能检测系统研究提供初步验证。

【关键词】 人工智能; 学习; 乳腺肿瘤; 放射摄影术

【中图分类号】 R737.9; R814.41; R446.9 **【文献标识码】** A

【文章编号】 1000-0313(2018)10-1029-04

DOI:10.13609/j.cnki.1000-0313.2018.10.010

开放科学(资源服务)标识码(OSID):



A preliminary investigation of breast mammography AI detection LI Xin, LIANG Shen, HUANG Zheng-nan, et al. Department of Imaging, Tangshan Gongren Hospital, Hebei 063000, China

【Abstract】 Objective: The current study aims to develop a detection model for lesions including masses, lymph nodes, and calcification in mammography with the latest artificial intelligence (AI) algorithms. Our findings provide preliminary verification for the application of deep learning in multi-category lesion detection in mammography. **Methods:** This study used a deep learning object detection algorithm called faster R-CNN to detect different lesions. The training dataset composed of 1892 mammography studies was manually labelled by experienced radiologists. A testing dataset of 400 mammography studies was used to evaluate the performance of AI. **Results:** In the test dataset, our AI algorithm detected 526 masses (689 in total), 912 lymph nodes (1098 in total), 52 round calcifications (73 in total), 519 rim calcifications (692 in total), coarse calcification 353 (544 in total), and sensitivities of each lesion category was 76.4%, 83.1%, 71.2%, 75.0% and 64.9%, respectively. And false positive rates of each lesion were 35.7%, 38.6%, 0.9%, 0.6% and 18.4%, respectively. **Conclusion:** Our AI lesion detection algorithm can detect masses, lymph nodes and calcifications mammography, and lays a research foundation for further multi-category AI detection system for mammography.

【Key words】 Artificial intelligence; Learning; Breast neoplasms; Radiography

乳腺癌是女性群体中发病率最高的一类肿瘤。乳腺癌的早诊早筛有利于提升患者的 5 年生存率,具有重大的临床意义^[1]。目前《中国抗癌协会乳腺癌诊治指南与规范(2017 年版)》建议 40~45 岁的女性进行一年一次的乳腺癌筛查,45~69 岁的女性每 1~2 年进行一次筛查,70 岁以上每 2 年进行一次筛查,并指

定使用钼靶作为主要的影像筛查手段,超声和 MRI 作为辅助方式^[2]。钼靶 X 线对微钙化敏感,对肿块也具有良好的检测效果,而且还具有性价比高的优点。但就高致密乳腺而言,钼靶影像的病灶检出存在一定的局限^[3]。上世纪 80 年代开始,研究者们陆续提出了一些计算机方法对钼靶图像进行辅助诊断^[4-6],此类算法大多基于传统的 CAD 算法。2014 年以来,基于深度学习人工智能逐渐成为计算机视觉的主流方法论,研究者们利用深度学习算法使用大规模数据训练预测模型,在无需手动干预的情况下在包括乳腺钼靶影

作者单位:063000 河北,唐山市工人医院影像科(李欣); 100025 北京,北京推想科技(梁森、黄正南、夏晨、张荣国); 362500 德化县医院放射科(吴孝掌、赖原仲)

作者简介:李欣(1970—),男,吉林伊通县人,硕士,副主任医师,主要从事研究人工智能在影像学诊断中的作用。

通讯作者:赖原仲, E-mail: dhlaiya@126.com

像^[7-9]、眼底影像^[10]、皮肤影像^[11]、CT 影像^[12]、病理影像^[13]等多个临床场景下实现了疾病的准确预测与分类。

本研究我们首先按照最新版 BI-RADS(2013 年)国际标准进行人工数据标记^[14],主要标记的病灶类型包括肿块、乳腺内淋巴结、圆形钙化、粗糙钙化、还有环形钙化。本文采用最新的深度学习病灶检测算法 Faster R-CNN 模型作为主要的算法结构^[15],训练并验证 AI 在乳腺 X 线影像的病灶检测效果。实验结果显示 AI 智能检测系统对肿块、乳腺内淋巴结、圆形钙化、环形钙化、粗糙钙化的检测敏感度分别达到 76.4%,83.1%,71.2%,75.0%和 64.9%。

材料与方法

1. 病例资料

本文回顾性地分析了 2016 年 6 月—2017 年 6 月 1 年内行乳腺钼靶检查的 5765 例病例,病例纳入条件:①女性;②影像报告发现病灶;③术前未经放疗或化疗。共纳入 2292 例病例,年龄 24~85 岁,中位年龄 41 岁。所有病例均行常规双侧乳腺轴面(craniocaudal,CC)和平行于胸大肌斜位(mediolateral oblique,MLO)摄影。

2. 检查方法

采用 GE 乳腺 X 线 Senograph 设备,型号为 ADS 54.10,行常规轴面和斜侧面摄像。设备电压 22~49 kV,电流:小焦点 40 mA,大焦点 100 mA,最大 500 mA。本研究中智能检测算法的开发与训练基于深度学习框架 Mxnet 1.2 版本开发,所有代码基于 Python 2.7 语言,并在 NVIDIA GPU 上运行,该 AI 智能病灶检测系统在 Windows/Linux/Mac 等操作系统上均可使用。

3. 图像处理和人工标注

2292 例乳腺 X 线病例 9168 张 DICOM 影像。每一张影像依次按如下步骤进行图像预处理:①用自适应的方法调整到适合阅片的窗宽窗位;②将图像的像素值归一化到 0~255 范围内;③将医学 DICOM 传输协议图像格式转换为适合计算机处理的 JPEG 图像格式。为了让 AI 可以智能检测出 X 线影像中的病灶,由有经验的影像科医生对影像的病灶进行人工标注。通过用标注好的数据训练 AI 智能病灶检测模型,我们最终得到模型并进行病灶检测效果的验证。

本研究采用最新版的 BI-RADS(2013 年)标准对病例进行人工标注。每例病例的人工标记流程:①3 名放射科副高职称医生分别进行标记;②1 名标记质量控制人员统计不一致的标记数据;③3 名标记人员一起讨论不一致的标记并达成共识。最后,2292 例数

据划分为 1892 例训练数据集和 400 例测试数据集。训练数据集用来训练我们的 AI 智能病灶检测模型,测试数据集用来评估我们 AI 模型的检测效果。

4. 人工智能检测方法

本研究采用深度学习人工智能领域经典的目标检测 Faster R-CNN 算法^[15],相比其它的目标检测算法,它具有高精度的检测优势,目前很多检测算法都是基于该算法发展出来的,比如 R-FCN 检测网络算法^[16]和 DeformbleNet 算法^[17]等。Faster R-CNN 是采用卷积神经网络为主体的两阶段目标检测算法(图 1)。首先对输入的乳腺 X 线图片进行多层卷积层和池化层操作从而实现自动特征提取,得到特征图(Feature Map)后将其输入一个区域提出网络(region proposal network,RPN)网络实现感兴趣病灶区域(ROI)自动提取,并将 ROI 区域映射到特征图上(即 ROI Pooling 操作),最后对映射后的特征图进行预测其病灶的类别和位置。算法流程分为训练阶段和测试阶段。训练阶段使用训练数据集对该算法进行训练,从而得到算法中的参数,此时算法的输入是乳腺 X 线图片和对应的人工标注的病灶标签。在测试阶段,我们使用测试数据集对算法的检测性能进行测试,将算法检测出来的病灶结果与人工标注的结果进行对比,用来评估 AI 智能检测病灶的性能。

5. 评价方法

本研究使用的敏感度和假阳率两个指标对检测的性能进行评价。假设影像中的病灶有 N 个,算法检测出了 M 个病灶,其中 TP 个病灶是真阳,FP 个病灶是假阳,且 TP+FP=M。那么敏感度(Sensitivity)和假阳率(False Positive Rate)的定义为,

Sensitivity=TP/N

False Positive Rate=FP/M

(1)

(2)

结 果

本研究对 1892 例训练数据集进行训练,得到优化后的 AI 智能病灶检测模型,最后基于 400 例测试数据集进行病灶检测的效果验证,得到统计结果(表 1)。

表 1 测试集病灶检测结果

项目	实际病灶数	检测出病灶	敏感度	假阳率
肿块	689	526	0.764	0.357
乳腺内淋巴结	1098	912	0.831	0.386
圆形钙化	73	52	0.712	0.009
环形钙化	692	519	0.750	0.006
粗糙钙化	544	353	0.649	0.184

400 例测试集的实际病灶分布肿块 689 个,乳腺内淋巴结 1098 个,圆形钙化 73 个,环形钙化 692 个,粗糙钙化 544 个。AI 智能病灶检测算法分别检测出肿块 526 个,乳腺内淋巴结 912 个,圆形钙化 52 个,

Faster R-CNN 乳腺X线病灶检测算法

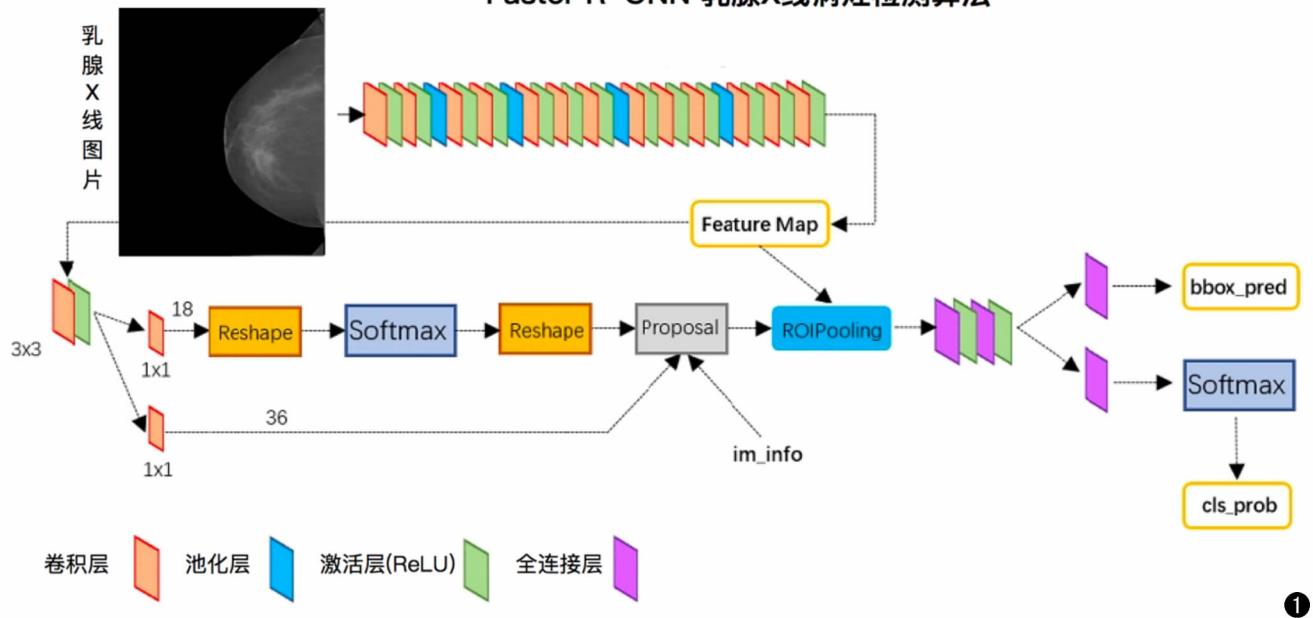


图1 算法的输入为乳腺X线片,然后通过算法的预测,输出病灶的位置(bbox_pred)和对应的类别(cls_prob)。

环形钙化 692 个,粗糙钙化 353 个。AI 智能病灶检测算法敏感度分别为 76.4%, 83.1%, 71.2%, 75%, 64.9%; 假阳率分别为 35.7%, 38.6%, 0.9%, 0.6%, 18.4%(图 2)。

讨 论

目前钼靶 X 线检查为乳腺癌筛查的首选方法。

本文使用最新的人工智能方法对乳腺 X 线影像中的多类病灶进行检测,其中乳腺内淋巴结的检测效果最好,敏感度能达到 83.1%,其次是肿块 76.4%,但两者的假阳率比较高,分别为 35.7% 和 38.6%。就钙化的检测而言,其整体效果没有肿块和乳腺内淋巴结的效果好,最好的是环形钙化敏感度为 75%,最差的是粗糙钙化,敏感度为 64.9%,但钙化的整体假阳率非常

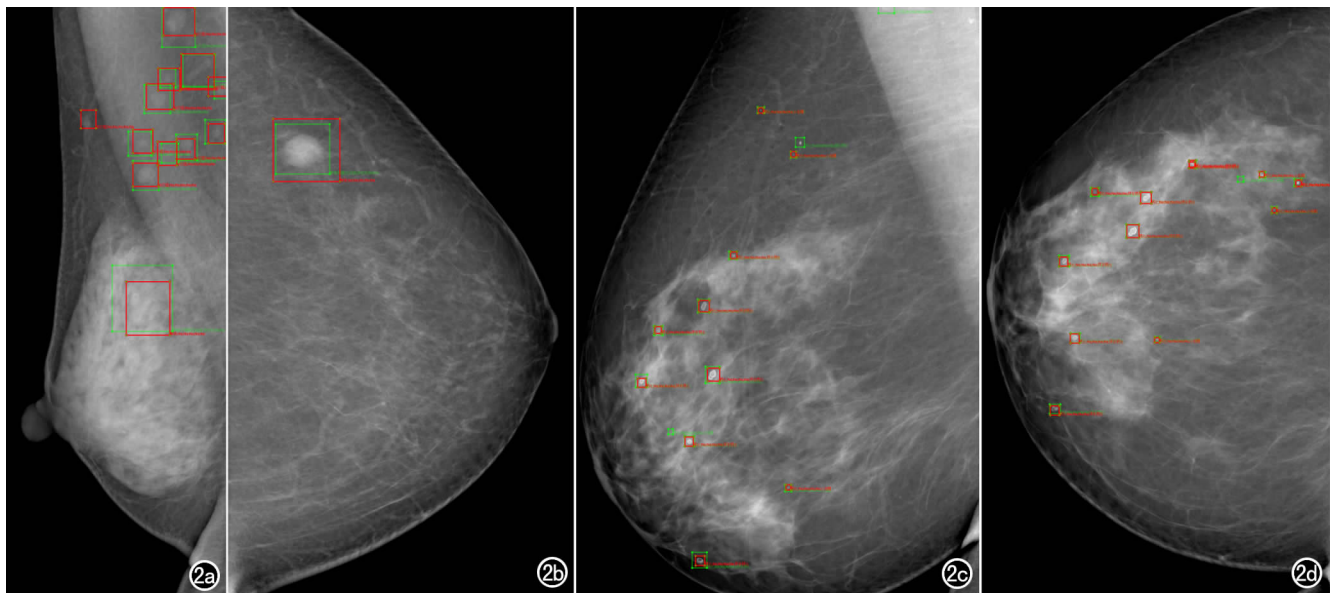


图2 AI 智能病灶检测结果。a~d) 分别为不同患者的乳腺 X 线影像的右乳 MLO 位,左乳 CC 位,右乳 MLO 位和右乳 CC 位。图框表示病灶,绿框代表真实病灶,红框代表 AI 智能检测出的病灶。若绿色框和红色框重合在一起表示真实的病灶被检测出。图中大部分病灶都被检测出;若某个病灶仅有绿框,表示 AI 漏检,c) 漏检了 3 个病灶,d) 漏检了 1 个病灶;若某个病灶仅有红框表示病灶是一个假阳,或者是一个医生漏标记的病灶。

低,环形钙化的假阳率仅为 0.6%。

尝试使用最新的人工智能算法对乳腺 X 线影像进行智能检测,可以同时多类别病灶进行检出和类别预测,并且乳腺 X 线影像可以不需要做复杂的预处理便可以输入模型中进行预测。乳腺内淋巴结检测是本研究的一个亮点,以往的研究基本上都是对钙化或者肿块的检测,很少有对乳腺内的淋巴结肿大情况进行检测。

本研究不足之处:①肿块和乳腺内淋巴结的假阳率较高;②需要大规模的人工标注数据,这需要耗费很多医生时间;③相对传统的 CAD 系统检测敏感度 80%~90%而已。本研究整体的敏感度偏低,但值得注意的是本研究的敏感度评估是针对具体病灶的评估,而传统的 CAD 评估是针对病例评估,本文研究病例量远远大于传统的 CAD 方法。

本研究现阶段仅仅只是用 AI 检测了病灶,然而 AI 在乳腺钼靶 X 线的作用远非如此。研究发现乳腺腺体的致密程度与患乳腺癌的风险成正比^[18],且对于高致密性腺体,因为病灶可能在致密腺体下面,从乳腺钼靶影像中往往很难发现病灶。然而对于人眼不易分辨的病灶,计算机可能具有比较好的分辨力。在未来,AI 可能能在预测乳腺致密程度,以及高致密型腺体的病灶检测上发挥出好的检测效果。

AI 还可通过对病灶进行良恶性鉴别以及识别癌症是否转移等来辅助医生的临床诊断。

参考文献:

- [1] Chen, Wan-qing, Rongshou Zheng, et al. Cancer statistics in China, 2015 [J]. CA: a cancer journal for clinicians, 2016, 66(2): 115-132.
- [2] 中国抗癌协会乳腺癌专业委员会. 中国抗癌协会乳腺癌诊治指南与规范(2017 年版)[J]. 中国癌症杂志, 2017, 27(9): 695-760.
- [3] 王者兵. 乳腺癌在乳腺钼靶检查中的漏误诊成因分析[J]. 中国疗养医学, 2017, 26(10): 1114-1115.
- [4] 张奕昭, 王金花, 谭婉嫦, 等. 微钙化的计算机辅助分析对乳腺导管原位癌及微浸润的诊断价值[J]. 放射学实践, 2016, 31(12): 1196-1200.
- [5] 陈诗慧, 刘维湘, 秦璟, 等. 基于深度学习和医学图像的癌症计算

机辅助诊断研究进展[J]. 生物医学工程学杂志, 2017, 34(2): 314-319.

- [6] 翟蕾蕾. 乳腺 X 线影像中结构紊乱的检测与识别[D]. 武汉大学, 2018.
- [7] Quéllec, Gwenole, Mathieu Lamard, et al. Multiple-instance learning for anomaly detection in digital mammography [J]. IEEE Transactions on Medical Imaging, 2016, 35(7): 1604-1614.
- [8] Ribli, Dezso, Anna Horváth, et al. Detecting and classifying lesions in mammograms with deep learning[J]. Scientific Reports, 2018, 8(1): 4165-4172.
- [9] Kooi T, Litjens G, van Ginneken B, et al. Large scale deep learning for computer aided detection of mammographic lesions[J]. Medical Image Analysis, 2017, 35: 303-312.
- [10] Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learn algorithm for detection of diabetic retinopathy in retinal fundus photographs[J]. JAMA, 2016, 316(22): 2402-2410.
- [11] Esteva A, Kuprel B, Novoa R, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. Nature, 2017, 542(7639): 115.
- [12] Song Q, Zhao L, Luo X, et al. Using deep learning for classification of lung nodules on computed tomography images [J]. J Healthcare Engineering, 2017, 2017: 8314740.
- [13] Khosravi P, Kazemi E, Imielinski M, et al. Deep convolutional neural networks enable discrimination of heterogeneous digital pathology images[J]. EBio Medicine, 2018, 27: 317-328.
- [14] Carl J. D'Orsi. 2013 ACR BI-RADS Atlas: breast imaging reporting and data system[M]. USA: American College of Radiology, 2014: 13-117.
- [15] Ren Shaoqing, Kaiming He, Ross Girshick, et al. Faster R-CNN: towards real-time object detection with region proposal networks [EB]. ArXiv: 1506. 01497 [Cs], <http://arxiv.org/abs/1506.01497>, 2016.
- [16] Dai Jifeng, Yi Li, Kaiming He, et al. R-FCN: object detection via region-based fully convolutional networks [EB]. ArXiv: 1605. 06409 [Cs], <http://arxiv.org/abs/1605.06409>, 2016.
- [17] Dai Jifeng, Haozhi Qi, Yuwen Xiong, et al. Deformable convolutional networks [EB]. ArXiv: 1703. 06211 [Cs], <http://arxiv.org/abs/1703.06211>, 2017.
- [18] Jeffers, Abra, Weiva Sieh, et al. Breast cancer risk and mammographic density assessed with semiautomated and fully automated methods and BI-RADS[J]. Radiology, 2017, 282(2): 348-355.

(收稿日期: 2018-09-20)