

双能CT融合图像在人工智能肺结节筛查中检测效能的探索研究

宋冬冬¹, 朱晓明¹, 朱丽娟¹, 顾俊², 伍建林¹, 张清¹

1. 大连大学附属中山医院 放射科, 辽宁 大连 116001; 2. 北京推想科技有限公司 全球临床科研合作学院, 北京 100025

[摘要] 目的 探索双能CT融合120 kVp图像在人工智能 (Artificial Intelligence, AI) 辅助诊断软件肺结节筛查中的检测效能。方法 收集我院行双源CT肺结节筛查的患者381例, 随机分为单源扫描 (A组) 和双能扫描 (B组) 两组。A组183例, 管电压为120 kVp; B组198例, 管电压为双源100/Sn 140 kVp, 两组均采用CareDose 4D技术。记录容积CT剂量指数 (CTDI_{vol}) 值、剂量长度乘积 (Dose-Length Product, DLP) 值。使用AI软件对图像进行结节检测, 记录结节的大小、位置、类型 (实性、亚实性)。统计AI软件检测全部结节数、不同大小结节数 (≥ 4 mm、 < 4 mm)、不同类型结节数 (实性、亚实性), 并结合金标准得到相应的真阳数、假阳数、漏诊数, 计算相应的结节检测敏感度、精确度和假阳性率, 比较两组间的差异, $P < 0.05$ 为差异具有统计学意义。结果 B组图像的肺结节检测敏感度高于A组 ($P < 0.05$), 辐射剂量低于A组 ($P < 0.05$), 同时, B组肺结节检测总的假阳性率低于A组。结论 双能融合120 kVp图像在AI软件肺结节检测中的检测效能优于单源扫描120 kVp图像, 其辐射剂量更低, 更加适用于AI软件肺癌筛查。

[关键词] 肺癌筛查; 计算机断层扫描; 人工智能; 检测效能; 辐射剂量; 双源CT; 融合图像

Investigation on the Detection Capability of Artificial Intelligence for Pulmonary Nodules Using Dual-energy Fusions Image

SONG Dongdong¹, ZHU Xiaoming¹, ZHU Lijuan¹, GU Jun², WU Jianlin¹, ZHANG Qing¹

1. Department of Radiology, Affiliated Zhongshan Hospital of Dalian University, Dalian Liaoning 116001, China;

2. Institute of Global Clinical Research Collaboration, Infervision Technology Co., Ltd, Beijing 100025, China

Abstract: **Objective** This study explores the pulmonary nodules detection capability of an artificial intelligence diagnostic system using dual-energy CT fusion with 120 kVp image. **Methods** 381 lung cancer screening patients underwent dual-energy CT scans in our hospital were prospectively enrolled in this study and randomly divided into two groups: group A (120 kVp single energy scan, 183 cases) and group B (100/Sn 140 kVp dual-energy scan, 198 cases). Both groups were treated with CareDose 4D technology. Volumetric CT Dose index (CTDI_{vol}) value and dose-length Product (DLP) value were recorded. AI software was used to detect the nodules in the image, and the size, location and type of nodules (solid and subsolid) were recorded. By comparing with the golden standard, total number of detected nodules was calculated, as well as nodules with different size (≥ 4 mm and < 4 mm), nodules with different density (solid and subsolid), and TPF, FPF, FNF. Finally, the sensitivity, precision and false positive rate of the AI system were obtained. Difference comparison studies were conducted, and $P < 0.05$ for the difference was statistically significant. **Results** The sensitivity of pulmonary nodules detection in group B was higher than that in group A ($P < 0.05$), and the radiation dose was lower than that in group A ($P < 0.05$). Meanwhile, the total false positive rate of pulmonary nodules detection in group B was lower than that in group A. **Conclusion** It is concluded from this study that dual-energy fusion 120 kVp image is more effective than single-source scan 120 kVp image in detection of lung nodules, and its radiation dose is lower, which is more suitable for detection of lung cancer by AI software.

Key words: lung cancer screening; computed tomography; artificial intelligence; detection efficiency; radiation dose; dual-energy CT; fusion image

[中图分类号] R318; R814.4

doi: 10.3969/j.issn.1674-1633.2021.02.018

[文献标识码] A

[文章编号] 1674-1633(2021)02-0073-04

引言

肺癌是我国目前发病率和死亡率最高的恶性肿瘤^[1], 并

仍呈逐年增长趋势。多层螺旋 CT 肺癌的早期筛查, 对肺癌的预防以及早期治疗意义重大^[2], 能有效提高早期肺癌患者的 5 年生存率^[3]。随着人工智能 (Artificial Intelligence, AI) 深度学习技术在医疗领域的推广应用, 肺 CT 图像的 AI 辅助结节筛查, 在临床上得到广泛开展^[4], 在提高工作

收稿日期: 2020-04-08

基金项目: 大连市科技局基金项目 (2015E12SF120)。

通信作者: 张清, 主任医师, 主要研究方向为呼吸及神经系统临床研究与应用。

通信作者邮箱: zhangqingsmile@163.com

效率的同时能有效降低肺结节的漏诊率^[5]。目前关于 AI 对肺结节检测效能的研究正逐步深入,但这些研究多集中在胸部 CT 单源扫描以及 AI 临床应用方法及价值^[6-8]的探讨上,而双能扫描对人工智能辅助肺结节检测效能的研究尚无人开展。本文旨在探讨临床中,胸部双能 CT 扫描结合 AI 行肺结节筛查,其检测效能与常规胸部 CT 单源 120 kVp 扫描的对比,同时比较两种模式的辐射剂量。

1 资料与方法

1.1 一般资料

研究通过了本院伦理委员会批准,所有研究对象检查前签署知情同意书。收集 2018 年 12 月至 2019 年 2 月行双源 CT 肺结节筛查的患者。纳入标准:18 岁以上,临床行胸部 CT 平扫检查的患者。排除标准:由于胸部手术史、胸腔积液、弥漫性间质性病变、肺不张、肺部炎性病变及 CT 图像中呼吸运动伪影等,影响结节确认及标注者。381 例患者纳入研究。按照扫描模式不同将患者分为两组:A 组(183 例)采用单源 120 kVp 扫描,B 组(198 例)采用双能模式扫描并收集融合 120 kVp 图像。两组患者的年龄、性别、结节大小及结节类型差异均无统计学意义(表 1)。

表1 两组患者的临床资料比较结果

组别	例数	年龄 /岁		性别 /例		结节大小 /mm		结节 类型	
		≥40	<40	男	女	≥4	<4	实性	亚实性
A 组	183	140	43	77	106	280	984	975	289
B 组	198	144	54	80	118	235	769	791	213
P 值		0.398		0.740		0.479		0.347	

1.2 扫描参数与方法

所有患者均在双源 CT 机(Siemens Somatom Definition Flash, 德国)上进行。扫描范围为胸廓入口到膈肌水平,包括整个肺野。单源扫描方式参数:探测器准直器 128×0.6 mm,机架旋转时间为 0.5 s/周,螺距为 1.2,管电压 120 kVp,参考管电流 130 mAs,开启 CARE Dose 4D 技术;双能扫描方式参数:探测器准直器 64×0.6 mm,机架旋转时间为 0.28 s/周,螺距为 0.7, A 管球管电压 100 kVp, B 管球管电压 sn140 kVp,融合图像为 120 kVp(两者按 4:6 的比例融合图像),参考管电流 A 管球 110 mAs、B 管球 94 mAs,开启 CARE Dose 4D 技术。

1.3 图像重建与结节检测

A、B 两组图像均行 1 mm、骨算法重建,得到单源 120 kVp 和双能融合 120 kVp 两组标准 CT 薄层图像。使用推想科技的肺结节薄层训练模型检测软件(Infer Read CT Lung Research, Infervision, Beijing, China),该系统基于 40 万训练数据集且数据来源于全国多家大型三甲医院,能将疑似肺结节的部位准确标记^[9]。对上述两组 CT 图像进行肺结节检测,记录检测出所有结节的数量、位置,并按大小(直径≥4 mm 或<4 mm)、类型(实性或亚实性)分类;三名不同年资的影像诊断医生结合 AI 对两组图像的肺结节进行分析,均依据《肺结节诊治中国专家共识(2018 年)》。首先由两名有 10 年以上经验的医师对两

组图像所有结节的数量、位置分别进行标注,并按大小(直径≥4 mm 或<4 mm)、类型(实性或亚实性)分类,标注结果不同之处由二人讨论后统一结果,最后再由一名主任医师对以上二人制定结果进行审阅,确立最终金标准结节。将人工智能软件自动检测结果与金标准比对,并统计检出结节的真阳数、假阳数,计算出检测敏感度、精确度和假阳性率。按公式 $ED=DLP \times k$ (k 为转换系数)计算有效辐射剂量。

1.4 统计学方法

所有统计学处理均应用 SPSS 20.0 版本进行。计量资料使用均数±标准差($\bar{x} \pm s$)表示,比较采用卡方检验,分别对敏感度、DLP、CTDI_{vol}、ED 进行组间方小差分析, $P<0.05$ 差异具有统计学意义。

2 结果

2.1 两组检测效能比较

两组图像对于 AI 的肺结节检测效能比较,差异有统计学意义($P<0.005$),见表 2。

表2 两组图像对于AI的肺结节检测效能

组别	金标准 结节数	算法检出 总结节数	真阳性 结节数	敏感 度/%	精确 度/%	假阳 性率
A 组	1264	1330	1047	82.8	78.7	1.6/CT
B 组	1004	1187	919	91.5	77.4	1.4/CT
χ^2 值	—	—	—	36.705	—	—
P 值	—	—	—	<0.001	—	—

2.2 两组对不同大小肺结节 AI 检测效能比较

两组对于 AI 对直径≥4 mm 和直径<4 mm 肺结节的检测效能比较,差异有统计学意义($P<0.005$),见表 3~4。

表3 两组图像对于AI对直径≥4 mm肺结节的检测效能

组别	金标准 结节数	算法检出 总结节数	≥4 mm 真 阳性结节数	敏感 度/%	精确 度/%	假阳 性率
A 组	280	435	255	91.1	58.6	1.0/CT
B 组	235	382	230	97.9	60.2	0.8/CT
χ^2 值	—	—	—	7.892	—	—
P 值	—	—	—	0.003	—	—

表4 两组图像对于AI对直径<4 mm肺结节的检测效能

组别	金标准 结节数	算法检出 总结节数	<4 mm 真 阳性结节数	敏感 度/%	精确 度/%	假阳 性率
A 组	984	895	792	80.5	88.5	0.6/CT
B 组	769	805	690	89.7	85.7	0.6/CT
χ^2 值	—	—	—	28.847	—	—
P 值	—	—	—	<0.001	—	—

2.3 两组对不同类型肺结节 AI 检测效能比较

两组图像对于 AI 对实性肺结节和 AI 对亚实性肺结节的检测效能比较,差异有统计学意义($P<0.005$),见表 5~6。

表5 两组图像对于AI对实性肺结节的检测效能

组别	金标准 结节数	算法检出 总结节数	实性真阳 性结节数	敏感 度/%	精确 度/%	假阳 性率
A 组	975	1019	842	86.4	82.6	1.0/CT
B 组	791	866	726	91.8	83.8	0.7/CT
χ^2 值	—	—	—	13.124	—	—
P 值	—	—	—	<0.001	—	—

表6 两组图像对于AI对亚实性肺结节的检测效能

组别	金标准 结节数	算法检出 总结节数	亚实性真 阳性结节数	敏感度 /%	精确度 /%	假阳 性率
A组	289	311	205	70.9	65.9	0.6/CT
B组	213	318	191	89.3	60.1	0.6/CT
χ^2 值	—	—	—	17.352	—	—
P值	—	—	—	<0.001	—	—

2.4 两组辐射剂量比较

A组CTDI_{vol}、DLP、ED均高于B组，差异有统计学意义($P<0.05$)，见表7。

表7 两组扫描辐射剂量

组别	CTDI _{vol} /mGy	DLP/mGy·cm	ED/mSv
A组	7.79 ± 2.2	285 ± 80.3	4.0 ± 1.1
B组	7.0 ± 1.9	231.4 ± 70.1	3.2 ± 1.0
P值	<0.001	<0.001	<0.001

3 讨论

基于深度学习^[10]的AI肺结节辅助检测在临床CT图像上的广泛应用^[11]，极大地提高了医生对肺结节的检出率^[12]。目前临床多在单源扫描下采集CT薄层图像，进行AI辅助肺结节筛查，其检测效能还有不足之处，尤其在临床上极易漏诊的最大直径<4 mm和亚实性结节的检出上^[13]，敏感度还有待提高^[14]。

随着双源CT的普及应用^[15-17]，胸部CT双能扫描成为日常临床工作中的一个选项，常规条件下与单源扫描相比，胸部CT双能扫描在不增加辐射剂量的前提下，能够得到融合120 kVp、sn140 kVp与100 kVp不同能量的图像，为诊断与鉴别诊断提供更多的参考依据^[18]。本研究在双源CT常规剂量胸部CT肺结节筛查中，将单源扫描120 kVp与双能扫描融合120 kVp的肺CT图像中，肺结节的AI检测效能进行了比较，结果显示双能扫描融合120 kVp图像的肺结节检测敏感度明显更高，辐射剂量明显更低，而假阳性率却不高，AI检测效能明显优于单源扫描120 kVp图像。其中：无论对于结节直径 ≥ 4 mm（临床需选择性影像随访或12个月影像随访^[19]），还是直径<4 mm（临床需12个月影像随访或6~24个月间不同方式影像随访^[19]），融合120 kVp图像检测敏感度均明显更高，对于直径 ≥ 4 mm的结节融合120 kVp图像检测假阳性率更低；无论实性结节还是亚实性结节，融合120 kVp图像的检测敏感度均明显更高，对于实性结节融合120 kVp图像的假阳性率更低。尽管与单源120 kVp图像相较，融合120 kVp图像在肺结节检测的假阳性率上没有明显的降低，但在AI辅助医生对肺结节的检出上，更高的结节检测敏感度，减少肺结节的漏诊临床意义更大，尤其在临床中容易漏检的小结节（最大直径<4 mm）和亚实性结节上，双能融合120 kVp图像的检测敏感度较单源120 kVp图像有了较大的提高。图1显示了2例临床易漏诊<4 mm亚实性结节在双能融合120 kVp结合AI被检测出来，而在单源120 kVp图像结合AI中漏诊。

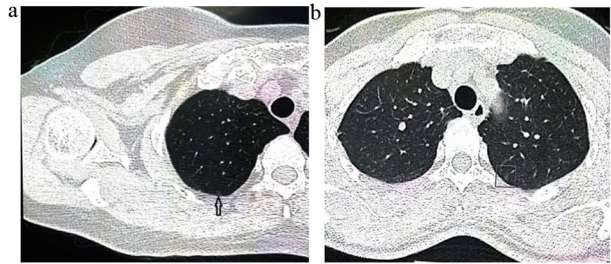


图1 两组图像对临床易漏诊结节的AI检出对比

注：本图为在两名医生单独阅片中均被遗漏的两例结节图像。a. 单源扫描120 kVp图像，箭头所示该肺结节在AI辅助检测中未被检出；b. 双能扫描融合120 kVp图像，方框中肺结节在AI辅助阅片检测中被检出。

最新发表在RSNA旗舰杂志Radiology: Artificial Intelligence的文章^[20]比较了一些扫描参数（比如CT设备厂家、扫描辐射剂量）对于深度学习肺结节检出模型检出效能的影响，结果发现CT设备厂家以及辐射剂量对于检出效能（敏感度、假阳性率）并没有影响，在我们的研究当中，单源120kVp的有效辐射剂量（ 4 ± 1.1 ）mSv虽然略高于双源扫描（ 3.23 ± 0.97 ）mSv，但是这不太可能是影响结节检出率的决定因素。因此我们猜测可能是由于双源融合图像（mixed image）的图像质量优于单源120 kVp图像质量从而导致AI辅助诊断软件的检出效能更高。双源融合图像选取来自低能100 kVp和高能Sn140 kVp的扫描图像，低能（100 kVp）图像具有高对比度，高能图像（Sn140 kVp）具有较低噪声，尤其Sn140是在140 kVp球管上添加了锡板过滤掉低能射线，让射线更“硬”，因此图像噪声得以进一步降低，所以融合图像兼具高对比度和低噪声的特点，往往图像的对比噪声比优于单源120 kVp扫描图像^[21]。相比于单源120 kVp扫描，双能扫描更优的融合图像质量可能是引起双能扫描AI检出效能更高的潜在原因，当然这还有待进一步试验和数据的验证。

本研究也存在一定不足之处，首先两组数据的人口统计学信息没有纳入本研究，其次样本量还需继续加大，我们将在以后的研究中予以完善。

总之，双能扫描融合120 kVp的CT图像较单源120 kVp的CT图像，在结合AI肺结节的检测中，敏感度更高，总体假阳性率也更低，且辐射剂量明显降低，值得在临床中推广应用。

[参考文献]

- [1] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2018, 68(6):394-424.
- [2] Parveen SS, Kavitha C. A review on computer aided detection and diagnosis of lung cancer nodules[J]. *Int J Comput Tech*, 2012, 3(3):393-400.
- [3] Henschke CI, Yankelevitz DF, Libby DM, et al. Survival of

- patients with stage I lung cancer detected on CT screening[J]. *Eng L J Med*, 2006, 35(17): 1763-1771.
- [4] 郑光远, 刘峡壁, 韩光辉. 医学影像计算机辅助检测与诊断系统综述[J]. 软件学报, 2018, 29(5): 1471-1514.
- [5] 李欣菱, 郭芳芳, 周振, 等. 基于深度学习的人工智能胸部CT肺结节检测效能评估[J]. 中国肺癌杂志, 2019, 22(6): 336-340.
- [6] 高歌, 马帅, 王宵英. 计算机辅助诊断在医学影像诊断中的基本原理和应用进展[J]. 放射学实践, 2016, 31(12): 1127-1129.
- [7] 张鹏, 徐欣楠, 王洪伟, 等. 基于深度学习的计算机辅助肺癌诊断方法[J]. 计算机辅助设计与图形学学报, 2018, 30(1): 90-99.
- [8] 萧毅, 刘士远. 肺结节影像人工智能技术现状与思考[J]. 肿瘤影像学, 2018, 27(4): 249-252.
- [9] 李甜, 李晓东, 刘敬禹. 人工智能辅助诊断肺结节的临床价值研究[J]. 中国全科医学杂志, 2020, 23(7): 828-831.
- [10] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks[J]. *Science*, 2016, (313): 504-507.
- [11] 王杜春, 任龙, 刘宁川, 等. 人工智能+医学影像在肺结节检测中的应用研究[J]. 影像研究与医学应用, 2019, 3(16): 39-41.
- [12] 明佳蕾, 方向明. 基于人工智能的CT肺结节检出临床应用及研究进展[J]. 中华放射学杂志, 2019, 53(6): 522-525.
- [13] Hjelm RD, Calhoun VD, Ruslan S, et al. Restricted Boltzmann machines for neuroimaging: An application in identifying intrinsic network[J]. *Neuroimage*, 2014, (96): 245-260.
- 上接第64页
- cranio-cervical junction[J]. *Chinese J CT MRI*, 2010, 8(4): 6-8.
- [20] Dickman CA, Mamourian A, Sonntag VK, et al. Magnetic resonance imaging of the transverse atlantal ligament for the evaluation of atlantoaxial instability[J]. *J of Neurosurg*, 1991, 75(8): 221-227.
- 上接第68页
- oncolytic adenovirus[J]. *PLoS One*, 2010, 5(11): e13859.
- [27] 叶芬, 王丽萍. 超声显像特征与乳腺癌组织中HER-2表达的关系研究[J]. 实用癌症杂志, 2020, 35(1): 50-52.
- [28] Orel SG, Schnall MD. MR imaging of the breast[J]. *Radiology*, 2001, 220(1): 13-30.
- [29] 彭晖晖, 景丽, 马丽园. 分析乳腺浸润性导管癌超声征象与病理组织学分级及ER、PR表达的关系[J]. 影像研究与医学应用, 2018, 2(16): 25-26.
- [30] Esserman L, Wolverson D, Hylton N. Magnetic resonance imaging for primary breast cancer management: current role and new applications[J]. *Endocr Relat Cancer*, 2002, 9(2): 141-153.
- [31] Zhao X, Li J. Research on the correlation between ultrasonographic features of breast cancer and expressions of ER, CD34 and p53[J]. *J Buon*, 2018, 23(2): 372-377.
- [32] Liu Y, Xiong W, Xu JM, et al. Correlations between the expression of C-erbB-2, CD34 and ER in breast cancer patients and the signs of conventional ultrasonography and ultrasound elastography[J]. *Eur Rev Med Pharmacol Sci*, 2018, 22(17): 5539-5545.
- [33] 王志威, 张晓晓, 王杰, 等. 局部晚期乳腺癌患者腋窝淋巴结转移范围的影响因素分析[J]. 诊断学理论与实践, 2019, 18(2): 189-192.
- [34] 彭微波. 老年乳腺癌患者肿瘤组织ER和PR表达与临床病理特征的相关性分析[J]. 中国现代普通外科进展, 2017, 20(4): 328-329.
- [35] 吕彦利, 康秀梅, 岳胜南, 等. 乳腺癌的彩色多普勒血流特征与ER、PR、HER-2、P53和KI-67表达的相关性研究[J]. 医学影像学杂志, 2019, 29(11): 1919-1922.
- [36] Sandholm J, Lehtimäki J, Ishizu T, et al. Toll-like receptor 9 expression is associated with breast cancer sensitivity to the growth inhibitory effects of bisphosphonates in vitro and in vivo[J]. *Oncotarget*, 2016, 7(52): 87373-87389.

责任编辑 皮志超

责任编辑 皮志超

责任编辑 皮志超