

论 著

基于深度学习算法的CT自动诊断系统在NSCLC患者TNM分期中的建立与应用

王 敏* 阮振平 沈 霖

安徽医科大学第二附属医院放射科/安徽医科大学医学影像研究中心 (安徽 合肥 230601)

【摘要】目的 建立基于深度学习算法的计算机断层扫描(CT)自动诊断系统,并探讨其在非小细胞肺癌(NSCLC)患者肿瘤淋巴结转移(TNM)分期中的应用。

方法 回顾性分析本单位2020年1月~2022年6月收治的245例NSCLC患者的临床资料作为训练集,提取CT检查特征信息和TNM分期数据,基于深度学习算法构建CT自动诊断系统。另选取本单位2022年7月~2023年6月收治的102例NSCLC患者作为验证集,经该系统诊断TNM分期,将病理学检查结果作为“金标准”,评价该系统对TNM分期的诊断价值。

结果 基于深度学习算法构建了CT自动诊断系统,该系统诊断各期的灵敏度、特异度、准确度与约登指数均高于常规CT诊断;基于深度学习算法的CT自动诊断系统诊断TNM分期与病理学诊断一致性高(Kappa=0.846, $P<0.001$),常规CT诊断TNM分期与病理学诊断也有一致性(Kappa=0.721, $P<0.001$)。

结论 基于深度学习算法的CT自动诊断系统诊断TNM分期优于常规CT。

【关键词】 深度学习算法; 计算机断层扫描; 自动诊断系统; 非小细胞肺癌; 分期

【中图分类号】 R734.2

【文献标识码】 A

DOI:10.3969/j.issn.1672-5131.2025.03.021

Establishment and Application of a CT Automatic Diagnosis System based on Deep Learning Algorithm in TNM staging of NSCLC Patients

WANG Min*, RUAN Zhen-ping, SHEN Ji.

Radiology Department of the Second Hospital of Anhui Medical University/ Medical Imaging Research Center of Anhui Medical University, Hefei 230601, Anhui Province, China

ABSTRACT

Objective To establish a computer tomography (CT) automatic diagnosis system based on deep learning algorithms and explore its application in staging of tumor node metastasis (TNM) in non-small cell lung cancer (NSCLC) patients. **Methods** The clinical data of 245 NSCLC patients admitted to our department from January 2020 to June 2022 were retrospectively analyzed as a training set, which was extracted CT examination feature information and TNM staging data, and a CT automatic diagnosis system based on deep learning algorithms was constructed. 102 NSCLC patients admitted to our department from July 2022 to June 2023 were selected as the validation set. TNM staging was diagnosed using this system, and pathological examination results were used as the "gold standard" to evaluate the diagnostic value of this system for TNM staging. **Results** A CT automatic diagnosis system was constructed based on deep learning algorithms. The sensitivity, specificity, accuracy, and Youden index of the system for each stage of diagnosis were higher than those of conventional CT diagnosis. The CT automatic diagnosis system based on deep learning algorithm had high consistency between TNM staging and pathological diagnosis (Kappa=0.846, $P<0.001$), and conventional CT diagnosis of TNM staging and pathological diagnosis also had consistency (Kappa=0.721, $P<0.001$). **Conclusion** The CT automatic diagnosis system based on deep learning algorithm is superior to conventional CT in diagnosing TNM staging.

Keywords: Deep Learning Algorithm; Computer Tomography; Automatic Diagnostic System; Non Small Cell Lung Cancer; Stages

非小细胞肺癌(NSCLC)是常见的肺恶性肿瘤,在所有肺癌患者中占比约80%~85%,而肺癌是全球癌症死亡的首位原因^[1]。调查指出^[2],我国每年约有80万新发肺癌患者,其中NSCLC占比达90%,且每年肺癌死亡人数达70万。目前肺癌患者常用的治疗方案包括手术、放化疗、靶向治疗、免疫治疗等,而肿瘤淋巴结转移(TNM)分期是确定治疗方案的重要依据。计算机断层扫描(CT)是临床常用的影像学检查方法,该技术利用精确准直的X线束、γ射线等围绕人体某个部位做断层扫描,可准确显示疾病信息,扫描速度快、图像清晰、可及时获取检查结果,在肺癌诊断中十分常用^[3]。有报道^[4]显示,CT可反映恶性肿瘤的直径、数量、位置等,并且还可评价淋巴结转移情况,因此对癌症患者TNM分期诊断具有指导价值。但是常规利用CT检查信息判断NSCLC患者的TNM分期具有主观性,容易导致判断有误,从而影响治疗方案的选择^[5]。深度学习算法是人工智能的产物,包括卷积神经网络、循环神经网络、生成对抗网络、强化学习等技术,理论上该技术可通过大量数据分析解决复杂问题,因此在大量数据支持下该机器学习方法的潜力方可充分释放和发挥。深度学习算法在癌症诊断中的作用已得到认可^[6-7],但在CT诊断NSCLC患者TNM分期中的应用尚未见报道。故而本研究基于深度学习算法构建CT自动诊断系统,用于诊断NSCLC患者的TNM分期,并验证该系统的价值,报道如下。

1 资料与方法

1.1 一般资料 纳入标准:符合NSCLC诊断标准^[8],且治疗前均实施CT检查;有明确的TNM分期病理学诊断结果;临床资料完整。排除标准:CT检查与病理学检查间隔>1周;NSCLC复发患者;其它部位原发性恶性肿瘤转移至肺部者;CT检查图像不清晰者;合并其它类型肺部疾病者,如肺结节等。回顾性分析本单位2020年1月~2022年6月收治的245例NSCLC患者的临床资料作为训练集,其中男性183例、女性62例;年龄34~78岁,平均(62.50±10.13)岁;有吸烟史169例;病理类型:腺癌145例、鳞癌82例、其它18例。

另选取本单位2022年7月至2023年6月收治的102例NSCLC患者作为验证集。纳入标准:符合NSCLC诊断标准,且治疗前均实施CT检查;实施病理学检查明确TNM分期;签署本研究知情同意书。排除标准同上。102例患者中男性78例、女性24例;年龄32~79岁,平均(63.25±10.30)岁;有吸烟史70例;病理类型:腺癌60例、鳞癌34例、其它8例。

1.2 方法 基于深度学习算法的CT自动诊断系统构建:(1)采用CT扫描仪(Discovery 750型)实施全肺CT扫描,检查前指导患者吸气、屏气训练。从肺尖至肺底扫描,两侧胸

【第一作者】王 敏,男,主要研究方向:医学影像技术CT、MRI成像质量控制。E-mail: dixjjaz@163.com

【通讯作者】王 敏

壁、腋窝、锁骨上窝及胸腔、脑、骨、肾上腺等均需扫描。扫描参数：层厚5mm、层间距5mm、螺距1.375，管电压120kV、管电流80mAs，行1.25mm薄层重建，并将扫描信息上传至系统工作站。(2)利用CT检查信息构建基于深度学习算法的CT自动诊断系统，获取训练集每例患者的病灶区域，由工作经验≥10年的放射科医师利用ITK-SNAP3.6.0软件手动勾画病灶区域，然后对所有图像进行逐一标注，包括肿瘤直径、淋巴结转移和其它部位肿瘤情况，比较每例患者的最大截面，选取最大截面的肿瘤设为中央层面、机器自动测量肿瘤长径设为最大肿瘤直径；利用Python3.6编程从CT图像中提取特征，包括图像一阶统计、形状、灰度区域大小矩阵、灰度共生矩阵、灰度形成纹理矩阵、CT评分等，采用最小绝对值收敛和选择算子(LASSO)算法筛选与淋

巴结转移相关的影像组学特征；确定脑、骨、肾上腺等部位是否有病灶，选取肿瘤截面。为避免数据不足影响该系统诊断TNM分期的准确性，对CT进行数据增强处理、增加数据量，其中所用技术包括平移、随机剪裁、噪声扰动、几何变换等。根据上述数据输入模型利用卷积神经网络、循环神经网络、生成对抗网络、强化学习技术进行反复训练，对数据高度抽象，自动学习TNM分期的有效特征。构建的模型结构包括图像和数据库、待检测模块，二者利用单箱探测器连接，将检测与识别阶段并入一个端-端网络，自动勾勒肿瘤边缘、筛选最大肿瘤界面、测算肿瘤长径，并分析其它部位肿瘤情况，对病灶形态学征象进行分析。将机器学习分析结果与TNM分期标准进行对照，自动传输T、N和M分期。该自动诊断系统操作流程见图1。

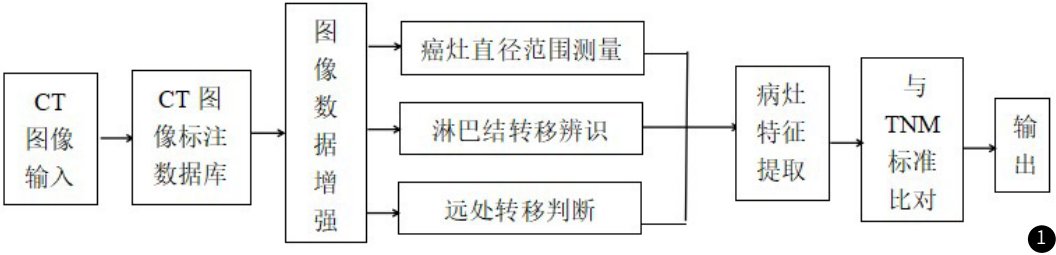


图1 基于深度学习算法的CT自动诊断系统操作流程。

病理学检查：所有患者均经过病理学检查判断TNM分期^[9]，分别代表肿瘤大小、区域淋巴结转移、远处扩散。其中T分为Tis：只限于气道通路内层细胞癌灶；T1：肿瘤<3cm，未扩散到脏层胸膜，未影响到主要支气管；T2：肿瘤>3cm、累及主要支气管但距隆突>2cm或已扩散至脏层胸膜或癌症阻塞气道但未造成肺萎陷或肺炎；T3：扩散至胸痹、膈肌、纵膈胸膜或壁层心包，累及一侧主支气管，距隆突<2cm但不包含隆突，已经长入气道足以造成全肺萎陷或全肺炎；T4：扩散至纵膈、心脏、食管、气管、脊柱或隆突，同一个肺叶有两个或多个独立肿瘤结节，有恶性胸水；N分为N0：癌症未扩散至淋巴结；N1：转移淋巴结仅限于肺内、肺门且是患侧；N2：淋巴结转移至隆突或纵膈且是患侧；N3：淋巴结转移至同侧或对侧锁骨上，或对侧肺门或纵膈；M分为M0：无远处转移；M1：有远处转移，包括其它肺叶、超过N分期中的淋巴结、其它器官或组织等。将TisN0M0记为原位癌；将T1~2N0M0记为I期；将T1~2N0M0记为II期；将T3N0M0或T1~3N2M0或T1~3N3M0或T4N1M0记为III期；将T1~4N1~3M1记为IV期。

1.3 观察指标 (1)统计验证集病理学诊断TNM分期结果；(2)统计验证集经基于深度学习算法的CT自动诊断系统、常规CT诊断

TNM分期结果；(3)分析经基于深度学习算法的CT自动诊断系统、常规CT诊断TNM分期的效能，记录该系统诊断各分期的灵敏度、特异度、准确度，并计算约登指数、诊断结果与“金标准”的一致性系数(Kappa值)；(4)典型病例的CT检查图片及基于深度学习算法的CT自动诊断系统输出结果。

1.4 统计学方法 采用SPSS 20.0软件行统计学分析。基于深度学习算法的CT自动诊断系统诊断TNM分期结果与病理学诊断TNM分期结果的一致性采用Kappa检验，Kappa>0.8认为一致性高。P<0.05为差异有统计学意义。

2 结果

2.1 验证集病理学诊断TNM分期结果 102例NSCLC患者中经验证集诊断TNM分期显示，I期7例、II期31例、III期46例、IV期18例。

2.2 验证集经基于深度学习算法的CT自动诊断系统诊断TNM分期结果 102例NSCLC患者中经验证集诊断TNM分期显示，I期7例、II期32例、III期46例、IV期18例；常规CT诊断TNM分期结果显示，I期9例、II期29例、III期48例、IV期16例。该结果与病理学诊断结果的对比见表1。

表1 验证集病理学诊断及基于深度学习算法的CT自动诊断系统、常规CT诊断TNM分期结果统计

方法	病理学诊断				
		I 期(n=7)	II 期(n=31)	III 期(n=46)	IV 期(n=18)
基于深度学习算法的CT自动诊断系统	I 期(n=6)	6	1	0	0
	II 期(n=32)	1	29	1	0
	III 期(n=46)	0	1	44	1
	IV 期(n=18)	0	0	1	17
常规CT诊断	I 期(n=9)	5	4	0	0
	II 期(n=29)	2	22	4	1
	III 期(n=48)	0	5	40	3
	IV 期(n=16)	0	0	2	14

表2 基于深度学习算法的CT自动诊断系统诊断TNM各期的效能

诊断TNM分期		灵敏度(%)	特异度(%)	准确度(%)	约登指数
I 期	自动诊断系统	85.71(6/7)	98.95(94/95)	98.04(100/102)	0.847
	常规CT	71.43(5/7)	95.79(91/95)	94.12(96/102)	0.672
II 期	自动诊断系统	93.55(29/31)	95.77(68/71)	95.10(97/102)	0.893
	常规CT	70.97(22/31)	90.14(64/71)	84.31(86/102)	0.611
III期	自动诊断系统	95.65(44/46)	96.43(54/56)	96.08(98/102)	0.921
	常规CT	86.96(40/46)	85.71(48/56)	86.27(88/102)	0.727
IV期	自动诊断系统	94.44(17/18)	98.81(83/84)	98.04(100/102)	0.933
	常规CT	77.78(14/18)	97.62(82/84)	94.12(96/102)	0.754

2.3 经基于深度学习算法的CT自动诊断系统诊断TNM分期的效能分析 基于深度学习算法的CT自动诊断系统诊断验证集各期的灵敏度、特异度、准确度与约登指数均高于常规CT, 具体见表2。基于深度学习算法的CT自动诊断系统诊断TNM分期与病理学诊断一致性高(Kappa=0.846, $P<0.001$), 且常规CT诊断TNM分期与病理学诊断也有一致性(Kappa=0.721, $P<0.001$)。

2.4 典型病例图片及基于深度学习算法的CT自动诊断系统输出结果对比 典型病例1, 何某某, 男, 43岁, 肿瘤最大径7mm, 分期cT1aN0M0 I A期, 见图2-3。

典型病例2, 倪某某, 男, 76岁, 肿瘤大小63×57 mm, 分期T3N1M0 II B期, 左肺上叶支气管截断闭塞, 无明显转移, 见图4-5。

典型病例3, 代某某, 男, 78岁, 肿瘤大小115×85×107mm, 分期cT4N2M0 IIIB期, 多发纵隔淋巴结、右下肺转移, 见图6-7。

典型病例4, 黄某某, 男, 72岁, 肿瘤大小61×57mm, 分期cT2bNxM1a IVA期, 胸膜转移, 无淋巴转移, 见图8-9。

典型病例5, 胡某某, 男, 56岁, 肿瘤大小30×18mm, 分期cT1bNxM1c IVB期, 左肺腺癌伴左侧胸膜、骨转移, 多发小淋巴结转移, 见图10-11。

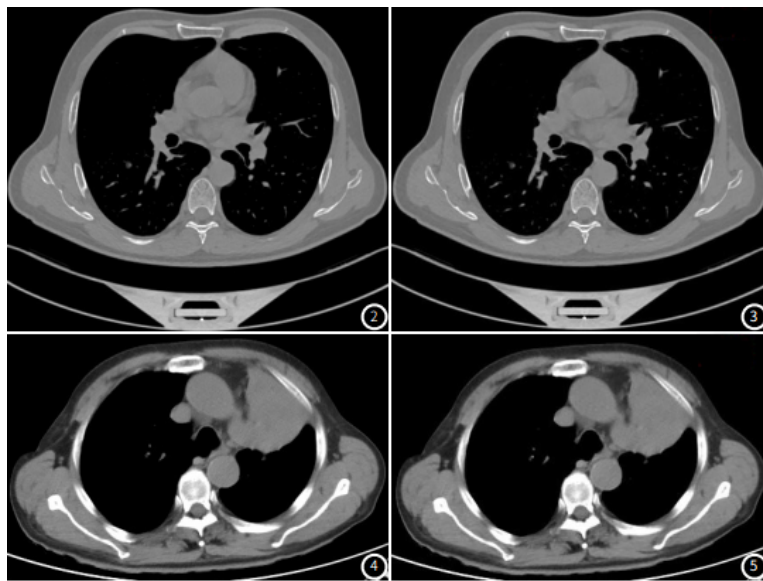


图2 典型病例1。CT图片。图3 典型病例1基于深度学习算法的CT自动诊断系统输出结果。
图4 典型病例2。CT图片。图5 典型病例2基于深度学习算法的CT自动诊断系统输出结果。

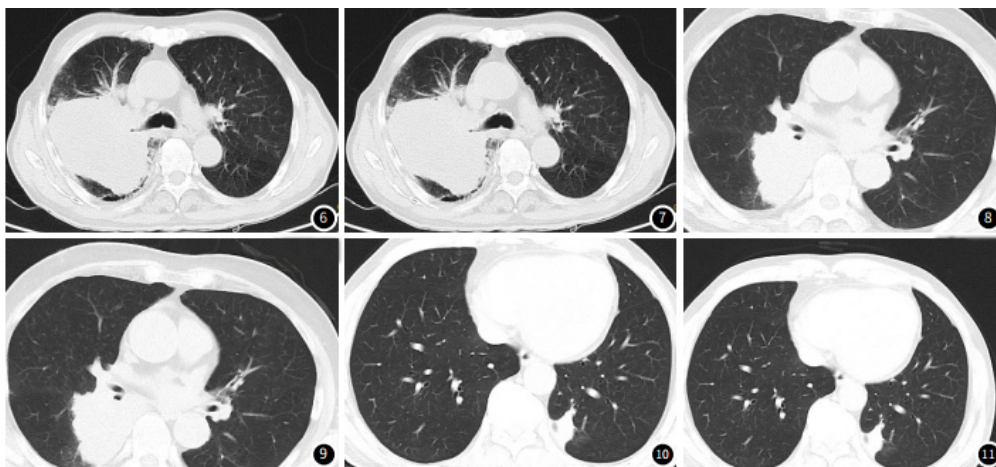


图6 典型病例3。CT图片。图7 典型病例3基于深度学习算法的CT自动诊断系统输出结果。
图8 典型病例4。CT图片。图9 典型病例4基于深度学习算法的CT自动诊断系统输出结果。
图10 典型病例5。CT图片。图11 典型病例5基于深度学习算法的CT自动诊断系统输出结果。

3 讨 论

NSCLC病情进展迅速、预后差,临床医师需要及时对其实施针对性的抗肿瘤治疗以控制病情进展、延长生存期。而TNM分期是临床医师针对NSCLC患者选择合理抗肿瘤治疗方案的依据。根据相关文件^[10],TNM分期主要是根据肿瘤直径、淋巴结转移及远处转移情况判断的,需要实施病理学检查方可确定,但对患者有创、且操作复杂、需等待检查结果。因此探讨其它准确的NSCLC患者TNM分期替代诊断方案十分必要。另外NSCLC中晚期常因为错过最佳手术时机而需要放化疗,但总体疗效并不理想^[11-12]。

本研究构建了基于深度学习算法的CT自动诊断系统,目的在于为NSCLC患者提供一种准确可行的TNM分期诊断方案。深度学习算法是一种新型的人工智能概念,与传统的计算机辅助诊断模式不同,该方法可通过多重非线性变换使数据产生高层抽象,从而可以无限逼近复杂函数,学习图像或数据表征,实现机器的自主学习。深度学习算法优点多,包括高速运算、深层次分析、表达能力强等,对图像、语音等难以直观判断、意义模糊的数据分析有明显优势^[13-14]。当前基于深度学习算法的计算机辅助诊断模型在多种疾病诊断中已得到推广应用,且其优势也得到认可^[15-17]。本研究中基于深度学习算法的CT自动诊断系统诊断NSCLC患者TNM分期的灵敏度分别为85.71%、93.55%、95.65%、94.44%,诊断各期的特异度和准确度均>95%,其诊断效能指标均高于常规CT诊断,且其诊断结果与病理学诊断结果高度一致,证实本研究构建的基于深度学习算法的CT自动诊断系统有确切价值。本研究利用深度学习算法的无监督、自主学习特征,对原始图像和数据逐层进行分析和特征变换并获得新的特征向量,对TNM特征分类特征实现自动化和可视化。对图像的预处理包括噪声消除和数据增强,从有限的病例数据库中提取肿瘤大小、淋巴结转移及远处转移的特征向量,制成训练数据集。将训练数据输入深度学习算法模型中训练,利用卷积神经网络对数据模型进行优化,然后测试数据集验证模型,综合影像表征的权重,获得高于临床诊断的水准。另外,本研究为避免数据不足,采用平移、随机剪裁等技术增强处理、增加数据量,提高对影像表征的准确性,增加对TNM分期判断的灵敏度与准确性。顾佳毅等^[18]针对直径≤5cm胃的胃肠间质瘤鉴别构建了一种基于增强CT的深度学习算法系统模型,不仅提高了鉴别准确性,也提高了诊断效率;袁紫旭等^[19]构建了一个基于CT图像的深度学习算法模型可以准确判断结直肠癌同时性腹膜转移发生情况;吴健等^[20]构建了一个基于多模态深度学习的胃癌TNM分期预测系统。本研究与上述报道相符,均证实深度学习算法技术在癌症诊断及TNM分期诊断中可行,且准确度高。

本研究中基于深度学习算法的CT自动诊断系统诊断NSCLC患者I期的灵敏度稍低,为85.71%,诊断II、III和IV期的灵敏度及诊断各期的特异度、准确度均较高,提示该自动诊断系统仍有不足,存在误诊可能。分析其中原因可能为:I期患者肿瘤特征不明显,且也受CT检查图像清晰度、检查设备灵敏度等限制,均可影响诊断的准确性。除了加强医师CT检查培训、做好CT仪器维护校准外,还应当更深入挖掘I期患者的肿瘤特征对基于深度学习算法的CT自动诊断系统进行优化。

综上,基于深度学习算法的CT自动诊断系统诊断NSCLC患者TNM分期较常规CT更有优势,诊断各期的灵敏度、特异度、准确度和约登指数均较高,与病理学诊断结果高度一致,建议推广。然而本研究仍存在不足,即基于深度学习算法的CT自动诊断系统诊断TNM分期仍有误诊风险,后续应进一步对该系统进行优化以提高诊断准确度。

参考文献

- [1] Shi Y, Lei Y, Liu L, et al. Integration of comprehensive genomic profiling, tumor mutational burden, and PD-L1 expression to identify novel biomarkers of immunotherapy in non-small cell lung cancer[J]. Cancer Med, 2021, 10(7): 2216-2231.
- [2] 刘枫林, 马伟. 肺癌患者临床流行病学及病理学特点分析[J]. 中国病案, 2021, 22(5): 53-55.
- [3] 王志芳, 张海深, 刘海燕, 等. 小细胞肺癌患者CT表现特点及诊断分析[J]. 中国CT和MRI杂志, 2023, 21(1): 74-76.
- [4] 房刚. 直肠癌临床术前CT诊断分期的临床价值研究[J]. 影像研究与医学应用, 2021, 5(13): 90-91.
- [5] Araujo-Filho JAB, Mayoral M, Zheng J, et al. CT radiomic features for predicting resectability and tnm staging in thymic epithelial tumors[J]. Ann Thorac Surg, 2022, 113(3): 957-965.
- [6] Zurlo IV, Schino M, Strippoli A, et al. Predictive value of NLR, TILs (CD4+/CD8+) and PD-L1 expression for prognosis and response to preoperative chemotherapy in gastric cancer[J]. Cancer Immunol Immunother, 2022, 71(1): 45-55.
- [7] Jensen CT, Gupta S, Saleh MM, et al. Reduced-dose deep learning reconstruction for abdominal CT of liver metastases[J]. Radiology, 2022, 303(1): 90-98.
- [8] Xu Y, Liu H, Chen J, et al. Comparisons between the National Comprehensive Cancer Network (NCCN) non-small-cell lung cancer (NSCLC) Clinical Practice Guidelines (Chinese version), the NCCN original edition, and the European Society for Medical Oncology NSCLC Guidelines in 2009[J]. Thorac Cancer, 2010, 1(2): 83-86.
- [9] 邵楚楚, 王婉莹, 任胜祥. CSCO非小细胞肺癌诊疗指南(2021版)解读[J]. 同济大学学报: 医学版, 2022, 43(1): 1-9.
- [10] 邢力刚, 马晓林. 2021版《中华医学会肿瘤学分会肺癌临床诊疗指南》非小细胞肺癌诊疗更新专家解读[J]. 疑难病杂志, 2022, 21(6): 557-560.
- [11] 贾思思, 黄普超, 马晓艳, 等. 信迪利单抗联合化疗方案对小细胞肺癌患者生活质量及免疫功能的影响[J]. 罕少疾病杂志, 2022, 29(7): 33-35.
- [12] 张喜爱, 白力允, 王胜根. 姑息性放疗联合三阶梯止痛原则对非小细胞肺癌骨转移痛患者疼痛程度及生命质量的影响[J]. 罕少疾病杂志, 2023, 30(5): 25-26.
- [13] Binczyk F, Prazuch W, Bozek P, et al. Radiomics and artificial intelligence in lung cancer screening[J]. Transl Lung Cancer Res, 2021, 10(2): 1186-1199.
- [14] Arabi H, Zaidi H. Deep learning-based metal artefact reduction in PET/CT imaging[J]. Eur Radiol, 2021, 31(8): 6384-6396.
- [15] Sanaat A, Shiri I, Arabi H, et al. Deep learning-assisted ultra-fast/low-dose whole-body PET/CT imaging[J]. Eur J Nucl Med Mol Imaging, 2021, 48(8): 2405-2415.
- [16] Zhang Y, Yue N, Su MY, et al. Improving CBCT quality to CT level using deep learning with generative adversarial network[J]. Med Phys, 2021, 48(6): 2816-2826.
- [17] Oostveen LJ, Meijer FJA, de Lange F, et al. Deep learning-based reconstruction may improve non-contrast cerebral CT imaging compared to other current reconstruction algorithms[J]. Eur Radiol, 2021, 31(8): 5498-5506.
- [18] 顾佳毅, 史豪庭, 杨琳希, 等. 基于增强CT的深度学习算法对直径≤5cm胃的胃肠间质瘤鉴别诊断临床意义[J]. 中华胃肠外科杂志, 2021, 24(9): 796-803.
- [19] 袁紫旭, 徐挺洋, 蔡建, 等. 基于CT图像的深度学习算法在诊断结直肠癌同时性腹膜转移中的应用价值[J]. 中华胃肠外科杂志, 2021, 24(3): 264-268.
- [20] 吴健, 陈满俊, 余日胜, 等. 一种基于多模态深度学习的胃癌TNM分期预测系统: CN20211481494. 6[P]. 2022-03-11.

(收稿日期: 2023-12-05)

(校对编辑: 姚丽娜)