

3D ResNet deep learning model for automatically identifying sequences of prostate multi-parametric MRI: A multicenter study

SUN Zhaonan¹, WANG Kexin², HUANG Wenpeng³, WU Pengsheng⁴,
ZHANG Xiaodong¹, WANG Xiaoying^{1*}

(1. Department of Radiology, 3. Department of Nuclear Medicine, Peking University First Hospital, Beijing 100034, China; 2. School of Basic Medical Sciences, Capital Medical University, Beijing 100069, China; 4. Beijing Smart Tree Medical Technology Co. Ltd, Beijing 100011, China)

[Abstract] **Objective** To construct a 3D ResNet deep learning model based on multi-parametric prostate MRI (mpMRI), and to observe its value for automatically identifying the main MR sequences. **Methods** Totally 1 153 sets pre-biopsy prostate mpMRI data of 1 086 patients who underwent ultrasound-guided prostate biopsy in 3 hospitals were collected and divided into different image datasets, *i. e.* T2WI, diffusion-weighted imaging (DWI) and apparent diffusion coefficient (ADC) maps with a total of 5 151 images. Then the images were categorized into non-fat-suppressed T2WI (T2WI_{nan}, $n=1\ 000$), fat-suppressed T2WI (T2WI_{fs}, $n=1\ 188$), high b-value DWI (DWI_{High}, $b\text{-value} \geq 500\ \text{s/mm}^2$, $n=1\ 045$), low b-value DWI (DWI_{Low}, $b\text{-value} < 500\ \text{s/mm}^2$, $n=1\ 012$) or ADC map ($n=906$), also divided into training set ($n=4\ 122$), verification set ($n=513$) and test set ($n=516$) at the ratio of 8:1:1. After preprocessing and augmentation, a 3D ResNet model for automatically identifying image categories was trained and optimized in the training and verification sets, and its classification efficiency was evaluated in the test set. **Results** The identifying accuracy, sensitivity, specificity, positive predictive value, negative predictive value, F1 score and Kappa value of the obtained model for automatically identifying categories of images in the test set was 0.995—1.000, 0.990—1.000, 0.998—1.000, 0.990—1.000, 0.998—1.000, 0.995—1.000 and 0.994—1.000, respectively. **Conclusion** The obtained 3D ResNet deep learning model could effectively and automatically identify the main sequences of prostate mpMRI.

[Keywords] prostatic neoplasms; magnetic resonance imaging; artificial intelligence

DOI:10.13929/j.issn.1003-3289.2024.05.028

3D ResNet 深度学习模型自动甄别前列腺多参数 MR 扫描序列:多中心研究

孙兆男¹,王可欣²,黄文鹏³,吴鹏升⁴,张晓东¹,王霄英^{1*}

(1. 北京大学第一医院医学影像科, 3. 核医学科, 北京 100034; 2. 首都医科大学基础医学院, 北京 100069; 4. 北京赛迈特锐医疗科技有限公司, 北京 100011)

[摘要] **目的** 构建基于前列腺多参数 MRI(mpMRI)自动甄别其主要扫描序列的 3D ResNet 深度学习模型,并评估其价值。**方法** 收集于 3 个医疗中心接受超声引导下前列腺穿刺的 1 086 例患者穿刺前 1 153 次前列腺 mpMRI 资料,并按不同扫描序列加以拆分,分别将 T2WI、弥散加权成像(DWI)及表观弥散系数(ADC)图归入相应数据集,共获得 5 151

[第一作者] 孙兆男(1993—),女,山东烟台人,在读博士。研究方向:前列腺癌人工智能辅助诊断系统的临床应用。E-mail: zhaonan_sun@163.com

[通信作者] 王霄英,北京大学第一医院医学影像科,100034。E-mail: wangxiaoying@bjmu.edu.cn

[收稿日期] 2023-08-08 **[修回日期]** 2023-12-05

组图像,并将归类为非脂肪抑制 T2WI(T2WI_{nan}, $n=1\ 000$)、脂肪抑制 T2WI(T2WI_{fs}, $n=1\ 188$)、高 b 值 DWI(DWI_{High}, b 值 $\geq 500\text{ s/mm}^2$, $n=1\ 045$)、低 b 值 DWI(DWI_{Low}, b 值 $< 500\text{ s/mm}^2$, $n=1\ 012$)及 ADC 图(ADC map, $n=906$)。按 8:1:1 比例将全部图像分为训练集($n=4\ 122$)、验证集($n=513$)和测试集($n=516$)。行预处理及扩增后,采用 3D ResNet 于训练集及验证集训练及优化自动甄别图像类别模型,以测试集评估模型分类效能。结果 所获模型分类测试集不同序列图像的准确率、敏感度、特异度、阳性预测值、阴性预测值、F1 值及 Kappa 值分别为 0.995~1.000、0.990~1.000、0.998~1.000、0.990~1.000、0.998~1.000、0.995~1.000、0.994~1.000。结论 3D ResNet 深度学习模型能有效自动甄别前列腺 mpMRI 所涉主要扫描序列。

【关键词】 前列腺肿瘤; 磁共振成像; 人工智能

【中图分类号】 R737.25; R445.2 【文献标识码】 A 【文章编号】 1003-3289(2024)05-0769-05

多参数 MRI(multi-parametric MRI, mpMRI)可提供多种形态及功能学信息,常用于临床显著前列腺癌(clinically significant prostate cancer, csPCa)和引导穿刺活检^[1-3]。弥散加权成像(diffusion-weighted imaging, DWI)及表观弥散系数(apparent diffusion coefficient, ADC)图是检出外周带 csPCa 的重要序列^[4-5]。目前对于前列腺 MR 扫描尚未形成标准统一,且以不同设备所获 mpMRI 的 DICOM 格式文件命名有所差异^[6],使得 mpMRI 数据呈多样化。人工智能(artificial intelligence, AI)广泛应用于前列腺领域^[7-8]。基于 mpMRI 的 AI 诊断前列腺疾病模型通常以特定 MRI 作为输入,以 T2WI、DWI 及 ADC 图组合最为常用;而以 AI 模型前置模型甄别其主要序列具有必要性。本研究基于多中心前列腺 mpMRI 构建 3D ResNet 深度学习模型,观察其自动甄别 mpMRI 主要序列的价值。

1 资料与方法

1.1 研究对象 收集 2017 年 6 月—2018 年 8 月 1 086 例[男性,年龄 29~94 岁、平均(62.1±9.3)岁]于北京大学第一医院医院、大连医科大学附属第二医院或福建医科大学附属协和医院接受超声引导下前列腺穿刺活检患者穿刺前 1 个月内共 1 153 次 mpMRI;经穿刺病理诊断 478 例为前列腺癌、608 例为非前列腺癌。本研究经医院伦理审查委员会批准(2021 科研 060、大医二院伦审 2021 第 166 号、2021KY157)。检查前患者均知情同意。

1.2 图像分类 将 DICOM 格式 mpMRI 输入北京大学第一医院医学影像科 AI Lab 研究平台并转换为 NIFTI 格式。由 1 名具有 30 年泌尿生殖系统影像学诊断经验的主任医师按不同序列拆分 mpMRI,将全部 5 151 组图像归入 T2WI、DWI 或 ADC 图数据集;将采集参数和空间位置相同图像归为同一类别,再按 b 值对 DWI 进一步分类,分为非脂肪抑制 T2WI(non fat-suppressed T2WI, T2WI_{nan}, $n=1\ 000$)、脂肪抑

制 T2WI(fat-suppressed T2WI, T2WI_{fs}, $n=1\ 188$)、高 b 值 DWI(DWI_{High}, b 值 $\geq 500\text{ s/mm}^2$, $n=1\ 045$)、低 b 值 DWI(DWI_{Low}, b 值 $< 500\text{ s/mm}^2$, $n=1\ 012$)及 ADC 图(ADC map, $n=906$)共 5 类。以 8:1:1 比例将数据集分为训练集($n=4\ 122$)、验证集($n=513$)和测试集($n=516$),且各集所包括患者不同。以上 5 类图像中,训练集各含 799、950、835、808、730 组,验证集为 100、120、105、102、86 组,测试集为 101、118、105、102、90 组。

1.3 构建模型 对数据集行预处理,包括图像强度直方图均衡化、剪裁背景无效区域并设图像大小为 64×128×128,之后以随机旋转、随机噪声、透视变换和随机平移法进行图像扩增。以 3D ResNet 为基础架构构建自动甄别图像序列的深度学习模型,具体网络结构见图 1。应用 PyTorch v1.7.1+cu110 软件(<https://pytorch.org/>)训练模型,模型输入端为不同类别图像(图 2);以 Adam 为梯度下降优化器,设置单次读取图像

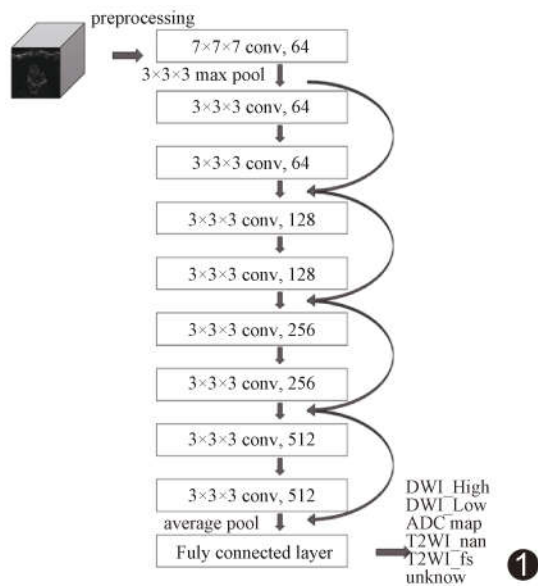


图 1 自动甄别前列腺 mpMRI 序列模型 3D ResNet 网络结构图

数量为 4, 学习率为 0.000 1, 循环次数为 400; 模型输出端采用单一输出通道, 输出结果为图像类别, 共分为 T2WI_nan、T2WI_fs、DWI_High、DWI_Low、ADC map 及 unknown(预测结果并非前 5 个序列中的任意 1 个)。以验证集数据调整学习率等超参数, 以提高模型泛化能力, 并采用早停机制避免模型在训练集中过度拟合; 发现模型针对验证集数据性能不再提升时, 提前停止训练过程, 以防止模型对噪声过度敏感。应用交叉熵损失函数缩小模型对各类图像的分类误差。以测试集数据评估模型分类效能。

以临床医师分类结果为参考, 采用混淆矩阵评估模型分类效能; 使用 scikit-learn 机器学习库中 sklearn.metrics 包的 classification_report 函数 (https://scikit-learn.org/stable/modules/model_

evaluation.html) 评价模型分类准确率、敏感度、特异度、阳性预测值、阴性预测值、F1 及 Kappa 值。

2 结果

图 3 为所获模型对各数据集图像的多分类混淆矩阵。模型针对测试集各序列图像进行分类的准确率、敏感度、特异度、阳性预测值、阴性预测值、F1 及 Kappa 值分别为 0.995 ~ 1.000、0.990 ~ 1.000、0.998 ~ 1.000、0.990 ~ 1.000、0.998 ~ 1.000、0.995 ~ 1.000 及 0.994 ~ 1.000。见表 1。测试集中, 模型误分类 1 组 T2WI_fs 为 DWI_Low、1 组 DWI_Low 为 ADC map、5 组 DWI_Low 为 DWI_High、6 组 DWI_High 为 DWI_Low。见图 4。

3 讨论

为目标疾病诊断模型部署适当的前置 MR 序列

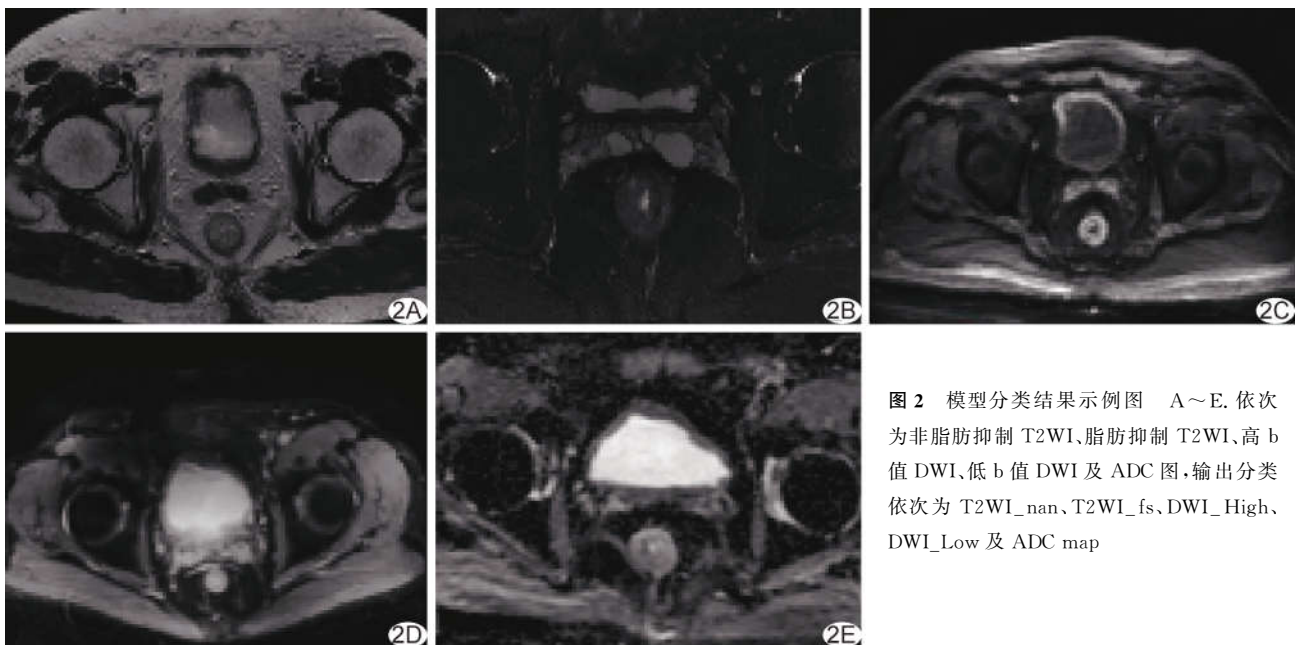


图 2 模型分类结果示例图 A~E. 依次为非脂肪抑制 T2WI、脂肪抑制 T2WI、高 b 值 DWI、低 b 值 DWI 及 ADC 图, 输出分类依次为 T2WI_nan、T2WI_fs、DWI_High、DWI_Low 及 ADC map

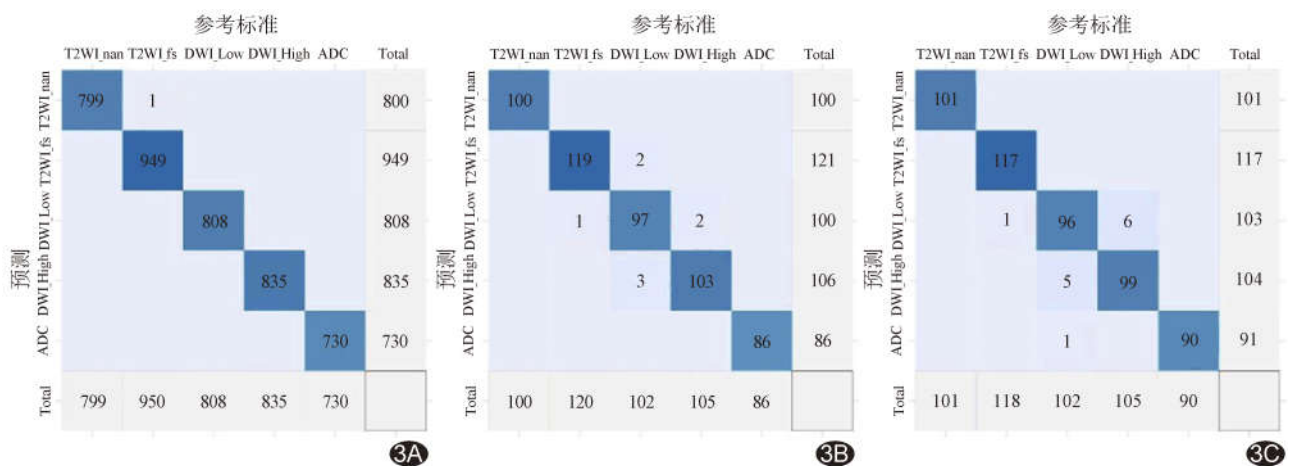


图 3 自动甄别前列腺 mpMRI 序列模型分类数据的混淆矩阵 A. 训练集; B. 验证集; C 测试集 (沿对角线排列数据为模型正确分类数量)

表 1 自动甄别前列腺 mpMRI 序列模型分类图像类别结果

图像类别	准确率	敏感度	特异度	阳性预测值	阴性预测值	F1	Kappa
训练集							
T2WI_nan(<i>n</i> =799)	1.000	1.000	1.000	0.999	1.000	0.999	0.999
T2WI_fs(<i>n</i> =950)	0.998	0.997	0.999	0.998	0.999	0.997	0.997
DWI_High(<i>n</i> =835)	0.996	0.994	0.998	0.991	0.998	0.992	0.990
DWI_Low(<i>n</i> =808)	0.992	0.987	0.998	0.991	0.997	0.989	0.986
ADC map(<i>n</i> =730)	1.000	1.000	1.000	0.999	1.000	0.999	0.999
验证集							
T2WI_nan(<i>n</i> =100)	1.000	1.000	1.000	1.000	1.000	1.000	1.000
T2WI_fs(<i>n</i> =120)	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DWI_High(<i>n</i> =105)	0.989	0.978	1.000	1.000	0.995	0.989	0.987
DWI_Low(<i>n</i> =102)	0.998	1.000	0.995	0.981	1.000	0.990	0.988
ADC map(<i>n</i> =86)	1.000	1.000	1.000	1.000	1.000	1.000	1.000
测试集							
T2WI_nan(<i>n</i> =101)	1.000	1.000	1.000	1.000	1.000	1.000	1.000
T2WI_fs(<i>n</i> =118)	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DWI_High(<i>n</i> =105)	0.995	0.990	1.000	1.000	0.998	0.995	0.994
DWI_Low(<i>n</i> =102)	0.999	1.000	0.998	0.990	1.000	0.995	0.994
ADC map(<i>n</i> =90)	1.000	1.000	1.000	1.000	1.000	1.000	1.000



图 4 自动甄别前列腺 mpMRI 序列模型分类错误示例图 A. 误将 T2WI_fs 分类为 DWI_Low; B. 误将 DWI_Low 分类为 DWI_High; C. 误将 DWI_High 分类为 DWI_Low

甄别模型及图像质量控制模型有其必要性^[9-13]。VERGANI 等^[14]开发的肝脏 MR 图像质量控制系统可在分析图像前自动甄别不同电影序列图像,包括短轴位、四腔心长轴位、左心室两腔心长轴位及左心室流入流出道位图像,整体准确率达 0.989~0.998。额·图娅等^[15]利用 3D-ResNet 将上腹部 MR 动态对比增强图像自动分类为平扫期、动脉早期、动脉晚期、门静脉-延迟期图像,整体准确率为 0.975~1.000。

前列腺 mpMRI 所涉序列众多,无论是医师或以 AI 诊断前列腺疾病,均需首先正确甄别扫描序列,以提升下游诊断模型^[16]效能。有研究^[9,17]采用深度学习技术基于前列腺 mpMRI 开发前置序列甄别模型,但均系基于前列腺 mpMRI 全部序列进行分类,涉及序列类别较多,准确性欠佳。方俊华等^[17]利用 2D ResNet18 深度学习框架分类前列腺 mpMRI 不同序

列图像,包括轴位 DWI、冠状位 T2WI、轴位灌注图像、矢状位 T2WI、轴位 ADC 图、轴位 T1WI 及轴位 T2WI,分类准确率分别为 100%、77.5%、96.7%、100%、100%、100% 及 100%。既往研究^[9]利用 3D-ResNet 将来自单中心的前列腺 mpMRI 分类为 9 个类别,其整体准确率均大于 0.99。

3D ResNet 在时间维度上引入卷积操作,可捕捉图像中的时空关系,相比其他网络处理体积数据更具优势,可学习图像高级特征,更有利于分类前列腺 3D mpMRI。本研究纳入多中心以 15 台不同 MR 扫描仪采集的 mpMRI 数据,基于诊断前列腺癌所需重要序列图像,即 T2WI、DWI 及 ADC,再根据 b 值及有无脂肪抑制将全部图像分为 5 个不同类别,协同应用传统图像处理 and 深度学习技术,以 3D ResNet 构建自动甄别前列腺 mpMRI 序列模型,结果显示该模型对各序

列的分类准确率均大于 0.99,提示该模型可用于分类以不同 MR 仪、不同协议所获 MRI,协助减轻阅片压力;但模型甄别 DWI_High 与 DWI_Low 稍逊于其他,可能由于 DWI 的 b 值呈梯度分布, b 值接近 500 s/mm² 图像数据特征并不明显,导致模型分类混淆。针对该模型输出错误图像及性能分析,发现图像分辨率低或存在空气伪影致图像质量不合格,或因盆腔存在明显软组织水肿、盆腔积液、多发骨转移、多发淋巴结转移等显著病变时, MRI 特征发生改变而导致模型错误分类。

综上,本研究构建的 3D ResNet 深度学习模型能有效自动甄别前列腺 mpMRI 主要序列。但本研究为回顾性分析,数据量在不同设备间不均衡,且未观察模型与影像科医师分类结果的一致性,有待后续进一步优化模型并通过外部验证提升其泛化能力。

利益冲突:全体作者声明无利益冲突。

作者贡献:孙兆男查阅文献、研究实施、数据分析、撰写文章;王可欣统计分析;黄文鹏修改文章;吴鹏升图像处理和分析;张晓东审阅文章;王霄英指导、审阅文章、经费支持。

[参考文献]

- [1] MOTTET N, van den BERGH R C N, BRIERS E, et al. EAU-EANM-ESTRO-ESUR-SIOG guidelines on prostate cancer: 2020 Update. Part 1: Screening, diagnosis, and local treatment with curative intent[J]. *Eur Urol*, 2021, 79(2):243-262.
- [2] SAHA A, HOSSEINZADEH M, HUISMAN H. End-to-end prostate cancer detection in bpMRI via 3D CNNs: Effects of attention mechanisms, clinical priori and decoupled false positive reduction[J]. *Med Image Anal*, 2021, 73:102155.
- [3] 瞿根义,许宁,陈少豪,等.前列腺癌根治术后 Gleason 评分升级与术前多参数 MRI PI-RADS 评分的关系[J]. *中国介入影像与治疗学*, 2016, 13(12):737-741.
- [4] TURKBEY B, ROSENKRANTZ A B, HAIDER M A, et al. Prostate imaging reporting and data system version 2.1: 2019 update of prostate imaging reporting and data system version 2[J]. *Eur Urol*, 2019, 76(3):340-351.
- [5] 姚楠,龙晚生. MR 多参数成像定量参数值评估外周带前列腺癌侵袭性[J]. *中国医学影像技术*, 2020, 36(6):903-907.
- [6] NOGUCHI T, HIGA D, ASADA T, et al. Artificial intelligence using neural network architecture for radiology (AINNAR): Classification of MR imaging sequences[J]. *Jpn J Radiol*, 2018, 36(12):691-697.
- [7] SUN Z, WU P, CUI Y, et al. Deep-learning models for detection and localization of visible clinically significant prostate cancer on multi-parametric MRI[J]. *J Magn Reson Imaging*, 2023, 58(4):1067-1081.
- [8] SUAREZ-IBARROLA R, SIGLE A, EKLUND M, et al. Artificial intelligence in magnetic resonance imaging-based prostate cancer diagnosis: Where do we stand in 2021?[J]. *Eur Urol Focus*, 2022, 8(2):409-417.
- [9] 孙兆男,崔应谱,刘想,等.基于深度学习的前列腺 mpMRI 序列分类的可行性研究[J]. *临床放射学杂志*, 2022, 41(8):1559-1564.
- [10] 孙兆男,崔应谱,刘想,等.利用深度学习实现腹盆部 CT 图像范围及期相分类:临床验证研究[J]. *放射学实践*, 2021, 36(4):551-555.
- [11] KIM T K, YI P H, WEI J, et al. Deep learning method for automated classification of anteroposterior and posteroanterior chest radiographs[J]. *J Digit Imaging*, 2019, 32(6):925-930.
- [12] RAJKOMAR A, LINGAM S, TAYLOR A G, et al. High-throughput classification of radiographs using deep convolutional neural networks[J]. *J Digit Imaging*, 2017, 30(1):95-101.
- [13] WILDEBOER R R, van SLOUN R J G, WIJKSTRA H, et al. Artificial intelligence in multiparametric prostate cancer imaging with focus on deep-learning methods [J]. *Comput Methods Programs Biomed*, 2020, 189:105316.
- [14] VERGANI V, RAZAVI R, PUYOL-ANTON E, et al. Deep learning for classification and selection of cine CMR images to achieve fully automated quality-controlled CMR analysis from scanner to report[J]. *Front Cardiovasc Med*, 2021, 8:742640.
- [15] 额·图娅,郭小超,王可,等.基于深度学习训练上腹部 DCE-MRI 扫描期相分类模型的可行性[J]. *放射学实践*, 2021, 36(8):1052-1058.
- [16] 向雪静,罗媛媛,邓宇,等. MIM 软件自动勾画前列腺癌危及器官[J]. *中国医学影像技术*, 2023, 39(1):79-84.
- [17] 方俊华, LI QIUBAI, 余成新,等.人工智能深度学习对前列腺多序列 MR 图像分类的可行性研究[J]. *中华放射学杂志*, 2019, 53(10):839-843.