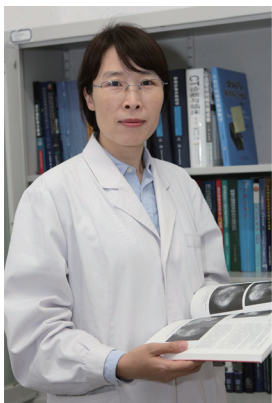




· 专题论著 ·



王霄英, 北京大学第一医院医学影像科, 主任医师, 教授。中华医学会放射学分会青年委员会委员, 中华医学会放射学分会磁共振成像学组成员, 中国医学影像技术研究会放射学分会委员。《中国医学影像技术》副主编, 《当代医学-中国肿瘤影像学》执行主编, 《中华放射学杂志》特邀编委, 《实用放射学杂志》编委。研究方向: 泌尿生殖系统影像学诊断, 影像学新技术临床应用, 人工智能辅助诊断。

基于深度学习和阈值分割方法对乳腺MRI中背景实质强化自动分类的研究

马明明¹, 张耀峰², 王祥鹏², 张晓东¹, 秦乃姗¹, 王霄英¹

1. 北京大学第一医院医学影像科, 北京 100034;
2. 北京赛迈特锐医学科技有限公司, 北京 100011

[摘要] 目的: 探讨基于深度学习和阈值分割方法对乳腺磁共振成像(magnetic resonance imaging, MRI)背景实质强化(background parenchymal enhancement, BPE)自动分类的可行性。方法: 回顾并分析2010年1月—2018年10月北京大学第一医院行乳腺MRI检查的547例乳腺癌患者的影像学资料。由2名高年资放射科医师共同评估, 依据乳腺影像报告和数据系统(Breast Imaging Reporting and Data System, BI-RADS)对健侧乳腺的BPE进行人工BPE的4分类作为金标准。先用训练好的深度学习U-Net 3D模型分割乳腺纤维腺体组织(fibroglandular tissue, FGT)区域, 分割效果的评价指标为Dice相似性系数(Dice similarity coefficient, DSC)。再利用阈值分割技术获取BPE区域, 计算BPE体积与FGT体积的比值得到自动BPE分类, 并生成到结构化报告中。将自动BPE分类与人工BPE分类进行比较, 以混淆矩阵进行效能分析, 计算总准确度、F1值和Kappa值。结果: U-Net 3D模型对547个数据FGT分割的平均DSC为0.902。通过阈值分割技术计算得到自动BPE分类结果的整体准确率为0.95(95% CI: 0.88~0.93), 宏平均F1为0.84, 微平均F1为0.90, Kappa为0.81(95% CI: 0.76~0.86)。结论: 基于深度学习和阈值分割技术对乳腺MRI背景强化的自动化分类是可行的, 有潜在的临床应用价值。

[关键词] 乳腺癌; 深度学习; 阈值分割; 磁共振成像; 背景实质强化; 结构化报告

DOI: 10.19732/j.cnki.2096-6210.2021.05.003

中图分类号: R737.9; R445.2 文献标志码: A 文章编号: 2096-6210(2021)05-0332-07

Automated classification of breast MRI background parenchymal enhancement using deep learning and thresholding segmentation method MA Mingming¹, ZHANG Yaofeng², WANG Xiangpeng², ZHANG Xiaodong¹, QIN Naishan¹, WANG Xiaoying¹ (1. Department of Radiology, Peking University First Hospital, Beijing 100034, China; 2. Beijing Smart-Imaging Technology Co. Ltd., Beijing 100011, China)

Correspondence to: WANG Xiaoying E-mail: wangxiaoying@bjmu.edu.cn

[Abstract] **Objective:** To explore the feasibility of automatic classification of breast magnetic resonance imaging (MRI) background parenchymal enhancement (BPE) based on deep learning and thresholding segmentation. **Methods:** Breast MRI data

of 547 breast cancer patients were retrospectively collected from January 2010 to October 2018 in Peking University First Hospital. According to the Breast Imaging Reporting and Data System (BI-RADS), each image was classified into four categories by two senior radiologists, which was used as the reference standard. Three steps were performed to acquire the BPE category automatically. Firstly, the region of fibroglandular tissue (FGT) was segmented by using a U-Net 3D model. The Dice similarity coefficient (DSC) was used to evaluate the efficacy of the model. Secondly, the region of BPE was segmented by the thresholding technology. Thirdly, the ratio of the BPE volume to the FGT volume was calculated and the BPE category was obtained. The BPE category was automatically input into the structured report. The multi-class confusion matrix was used to analyze the performance of the calculated BPE, with the metrics of accuracy, F1 score, and Kappa. **Results:** The average DSC of the U-Net model was 0.902. The Macro accuracy, F1 Macro, F1 Micro, and Kappa was 0.95 (95% CI: 0.88-0.93), 0.84, 0.90, and 0.81 (95% CI: 0.76-0.86), respectively. **Conclusion:** The automated classification of breast MRI background parenchymal enhancement is feasible and has potential clinical application value.

[**Key words**] Breast cancer; Deep learning; Thresholding segmentation; Magnetic resonance imaging; Background parenchymal enhancement; Structured reporting

女性乳腺癌现已超过肺癌成为2020年全球发病率最高的癌症，每年估计有230万例新增病例，占有癌症病例的11.7%^[1]。磁共振成像（magnetic resonance imaging, MRI）可用于乳腺癌高危患者的诊断、术前分期、化疗效果监测、乳腺假体植入物评估等^[2]。乳腺MRI背景实质强化（background parenchymal enhancement, BPE）是乳腺动态对比增强MRI（dynamic contrast enhanced MRI, DCE-MRI）时正常乳腺纤维腺体组织（fibroglandular tissue, FGT）的强化区域，BPE水平与乳腺癌的发生密切相关^[3]，也可用于新辅助化疗治疗效果的预测^[4]。为了实现BPE的标准化定量，在临床工作中放射科医师根据乳腺影像报告和数据系统（Breast Imaging Reporting and Data System, BI-RADS）对BPE进行分类^[5]，将BPE分为极小、轻度、中度、显著4类。BPE的主观视觉评估容易出现观察者之间和观察者内的差异，因此需要更客观、简便的方法来评估BPE。

目前，深度学习技术已应用于乳腺MRI中FGT的分割^[6]和强化体积的计算^[7]，但关于BPE分类的研究很少。北京大学第一医院已经完成了基于U-Net深度学习模型对乳腺DCE-MRI图像中乳房和乳腺癌强化病灶的自动分割^[8]，本研究进一步探讨基于深度学习和阈值分割方法对BPE进行自动分类的可行性，并将BPE分类结果自动接入到结构化报告系统中，未来拟自动完成BPE的客观定量，并将BPE定量应用于乳腺癌化

疗效果预测的研究。

1 资料和方法

1.1 研究对象

收集2010年1月1日—2018年10月30日于北京大学第一医院接受乳腺DCE-MRI检查的患者的影像学资料，共纳入547例患者，患者均为女性，年龄22~82岁，平均年龄（52.3 ± 11.6）岁。纳入标准：① 所有患者均使用相同的标准MRI扫描方案，图像质量合格；② 乳腺肿瘤标本（活检或手术）病理学检查结果为恶性；③ 患者未接受任何治疗（包括手术、化疗、放疗等）。排除标准：① 假体植入术后；② 乳腺内极少FGT（几乎全是脂肪）；③ 双侧乳腺癌患者。由2名高年资放射科医师（A和B）共同读片，根据第5版BI-RADS对健侧BPE进行4分类并作为金标准。BPE分类标准如下^[9]：① 极小，<25%FGT的强化；② 轻度，25%~50%FGT的强化；③ 中度，50%~75%FGT的强化；④ 显著，>75%FGT的强化。

本研究经过医院伦理审查委员会的批准，研究工作按照北京大学第一医院人工智能（artificial intelligence, AI）模型训练规范进行。

1.2 用例定义

根据北京大学第一医院AI训练管理方法，首先定义研发乳腺MRI图像上BPE分类模型的用例。包括模型的ID、临床问题、场景描述、模型

在临床工作中的调用流程、模型输入及输出数据结构等。AI模型返回结果为BPE分类“极小、轻度、中度、显著”。

BPE分类模型预测结果返回到结构化报告中的“背景强化”模块(图1)。

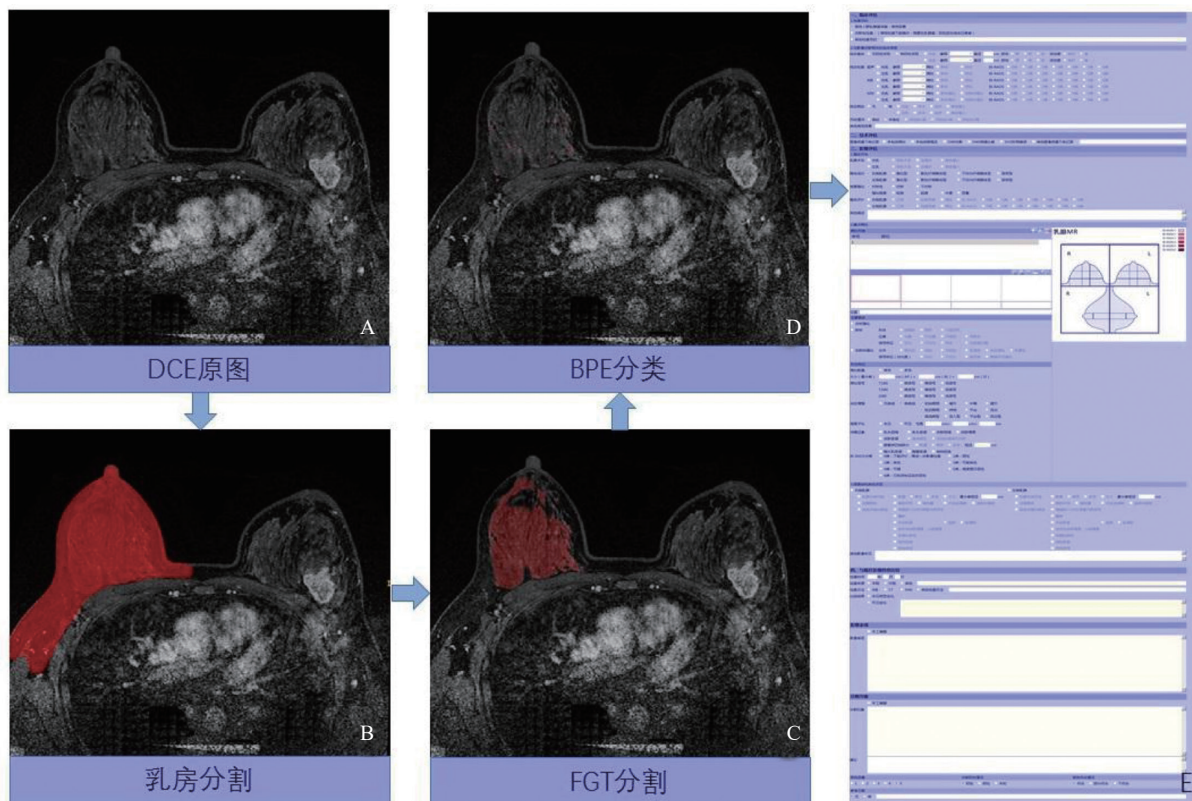


图1 本研究简要流程图

DCE-MRI图像(A)经过乳房分割(B)、FGT分割(C)、BPE分类(D),程序将BPE分类结果自动填写到结构化报告(E)中。

1.3 图像采集

采用美国GE公司的Signa Excite HD 1.5 T MRI扫描仪,使用4通道双穴乳腺线圈行乳腺MRI检查。扫描序列包括轴位T1加权成像(T1-weighted imaging, T1WI)、压脂T2加权成像(T2-weighted imaging, T2WI)、弥散加权成像(diffusion-weighted imaging, DWI)及DCE。其中DCE的参数如下:重复时间(repetition time, TR)为3.9 ms,回波时间(echo time, TE)为1.7 ms,层厚1.4 mm,层间距0 mm,视野36 cm × 36 cm,矩阵320 × 320。静脉注射钆对比剂0.2 mmol/kg,注射速率2 mL/s,随后使用

20 mL生理盐水冲管。注射同时开始扫描,每个时相扫描时间为55 s,共扫描9个时相。

1.4 数据处理和图像分割

将DICOM图像转换为NiFTI格式。由1名高年资放射科医师使用ITK-SNAP软件标注100个数据的FGT区域,训练分割模型。以本单位既往训练好的U-Net 3D模型分割乳房作为掩膜(mask)^[11],进一步训练FGT分割模型。使用547个数据进行外部验证,分别为262个左侧FGT和285个右侧FGT(图2)。以DCE-MRI第一期图像为基线,使用上述模型分割得到基线FGT的体积。

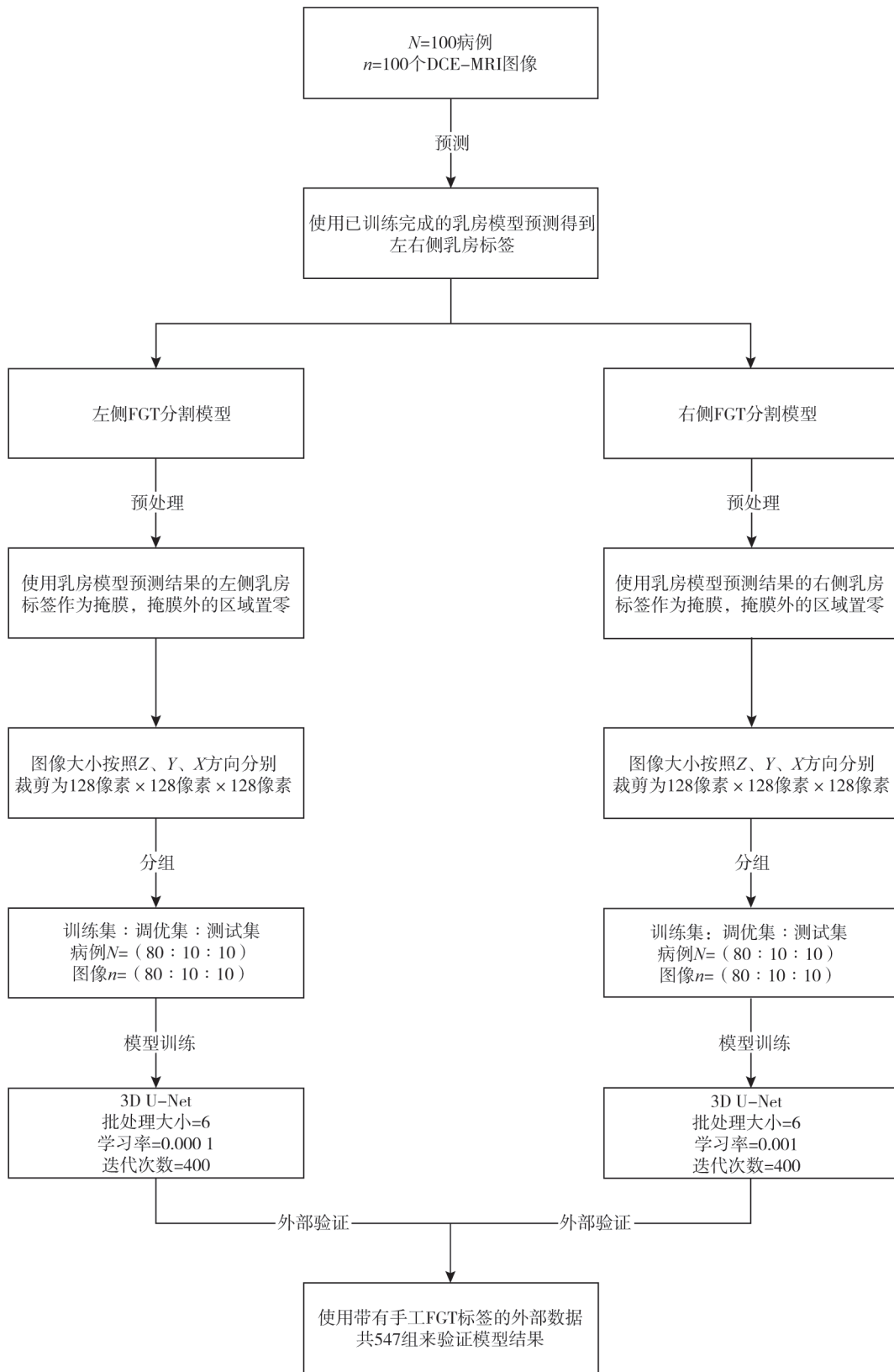


图2 训练乳腺FGT U-Net 3D分割模型的流程及主要参数

1.5 BPE分类模型计算

以增强扫描第三期图像和基线图像做减影, 得到DCE-sub图像。在DCE-sub图像上对547个数据的BPE进行阈值分割, 取全图中信号强度 >50%区域为BPE (图3)。最后计算BPE体积

与基线FGT体积的比值, 公式: $BPE \text{ 比值} = \frac{BPE \text{ 体积}}{\text{基线FGT体积}} \times 100\%$ 。根据BPE比值分为4类: 极小 (<25%); 轻度 (25%~50%); 中度 (50%~75%); 显著 (>75%)。

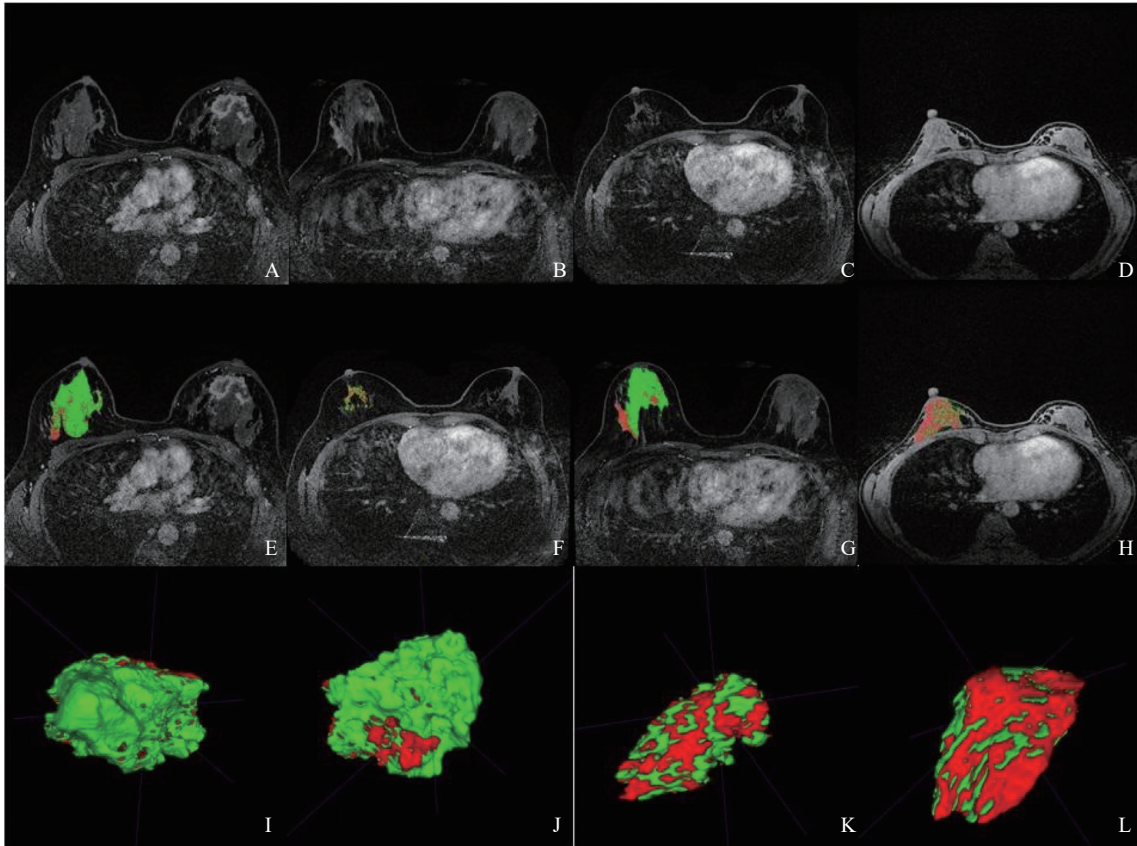


图3 典型病例1~4 (4类BPE强化类型的乳腺癌) MRI图像、FGT和BPE标签

A~D: 4例患者的乳腺MRI增强检查DCE第三期图像, BPE分别为极小、轻度、中度、显著强化; E~H: 绿色和红色区域分别为模型输出的FGT和BPE标签; I~L: 绿色和红色区域分别为分割FGT和BPE的3D图像。

1.6 模型效能评价

对乳腺FGT分割效果的评价指标为Dice相似性系数 (Dice similarity coefficient, DSC)。对乳腺BPE分类结果的评价使用多分类混淆矩阵, 计算其准确度、F1评分和Kappa值。

2 结 果

2.1 FGT分割模型的效果评价

FGT分割模型的平均DSC为0.902, 模型预测FGT区域与高年资医师标注区域一致性高。

2.2 BPE分类模型的定量计算

BPE分类的混淆矩阵和诊断效能见表1、2。

表1 BPE分类混淆矩阵

医师分类	模型分类				合计
	极小	轻度	中度	显著	
极小	307	24	3	0	334
轻度	5	170	22	0	197
中度	0	0	14	0	14
显著	0	0	0	2	2
合计	312	194	39	2	547

表2 BPE分类的诊断效能

评价指标	结果
准确度	0.95
95% CI	0.88~0.93
宏平均F1	0.84
微平均F1	0.90
Kappa系数一致性检验	0.81
Kappa系数95% CI	0.76~0.86

3 讨 论

乳腺MRI增强检查已被广泛应用于乳腺癌诊断、化疗效果评估，其不仅提供了肿瘤的形态学信息，还可提供定量的功能信息。BPE是乳腺癌风险和新辅助化疗（neoadjuvant chemotherapy, NAC）效果评估的主要参数之一^[3-4]。本研究使用U-Net 3D分割模型对乳腺MRI中基线FGT区域进行分割，并利用阈值分割技术得到增强后图像中BPE区域，通过计算二者的体积之比来定量评估BPE的4分类。既往Ha等^[10]已经提出关于使用全卷积神经网络（convolutional neural network, CNN）方法量化乳腺MRI中FGT和BPE，但是BPE采用了手动标注的方法，标注工作量大，且测试集中FGT和BPE的准确度分别为0.813和0.829。本研究与前人比较，相同之处是也使用了深度学习方法。不同之处有两点：第一，本研究中BPE为程序自动分割，操作简便，并且结果优于前人，FGT分割的DSC值达0.902，BPE分类准确度达0.95；第二，本研究的结果可自动分类接入到结构化报告中，对医学影像工作流程优化有益。

本研究提出的BPE自动分类方法整体准确度较高，但是从混淆矩阵中可以看出BPE的4种类型的数据不均衡，这是由本组数据的性质决定的。本组数据BPE自动分类结果提示中度强化类与人工分类差别较大，其错误分类发生在较轻强化类别中，大部分发生在相邻类别的轻度分类。分析其可能原因有以下两个方面。一方面，本研究以人工分类为参考标准，且由2名医师共同读

片，虽然分类的把握度较高，但仍可能存在主观偏差。本研究中医师通过在影像存储与传输系统（picture archiving and communication system, PACS）工作站上连续翻动图像获取BPE的三维印象，对于轻度、轻-中度、中度之间的判定可能有一定误差。另一方面，本研究中使用固定阈值分割BPE体积，研究结果表明大部分患者使用固定阈值是可行的，但可能个别患者需要使用有针对性的阈值方能准确地分割BPE体积。上述误差在临床实际工作中是可以纠正的，我们的工作流程是DCE-MRI扫描完成后，程序自动对BPE进行分类并输入到结构化报告中，医师打开报告确认BPE分类是否正确，基于目前的结果可知，仅极少数病例的BPE自动分类结果需要医师修改，绝大多数病例的BPE自动分类结果不需要医师修改，直接接受即可。这种流程既可以提高医师的工作效率，又可以保证最终报告的准确性。

本研究的局限性在于以下几个方面。① 纳入对象仅为乳腺癌初诊患者，有研究^[4]表明，通过定量评估NAC前后DCE-MRI中BPE的变化可反映NAC效果。因此，今后BPE相关研究应同时纳入NAC前后患者，为后续预测NAC效果做准备。② 本研究以2名高年资放射科医师共同对BPE实施的分类结果作为金标准，而医师评估的不一致性很常见^[11]。为了使BPE分类标准更加可靠，将来需要对观察者一致性进行评估。③ 本研究排除了乳腺内极少FGT的数据，由于此类数据乳房内几乎无FGT，BPE计算可能不准确。为了适用于复杂的临床实际工作场景，后续研究应针对不同FGT含量的患者进行分层分析。④ 本研究选择了一个固定阈值分割BPE，这是根据经验选择的，后续研究需要选用个性化的动态阈值进行定量计算BPE。最后，所有研究是在同一机构的磁共振设备进行的，未来应使用多中心的数据对模型进行验证。

总之，基于U-Net深度学习模型和阈值分割技术对BPE进行自动分类的准确度高，有可能应用于临床实践，也有可能用于探索BPE与NAC效果的相关性研究。

[参 考 文 献]

- [1] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. *CA Cancer J Clin*, 2021, 71(3): 209–249.
- [2] MANN R M, BALLEYGUIER C, BALTZER P A, et al. Breast MRI: EUSOBI recommendations for women's information [J]. *Eur Radiol*, 2015, 25(12): 3669–3678.
- [3] KING V, BROOKS J D, BERNSTEIN J L, et al. Background parenchymal enhancement at breast MR imaging and breast cancer risk [J]. *Radiology*, 2011, 260(1): 50–60.
- [4] YOU C, PENG W J, ZHI W X, et al. Association between background parenchymal enhancement and pathologic complete remission throughout the neoadjuvant chemotherapy in breast cancer patients [J]. *Transl Oncol*, 2017, 10(5): 786–792.
- [5] D'ORSI C J, SICKLES E A, MENDELSON E B, et al. ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System [M]. Reston: American College of Radiology, 2013.
- [6] MA X Y, WANG J L, ZHENG X P, et al. Automated fibroglandular tissue segmentation in breast MRI using generative adversarial networks [J]. *Phys Med Biol*, 2020, 65(10): 105006.
- [7] XU X W, FU L, CHEN Y Z, et al. Breast region segmentation using convolutional neural network in dynamic contrast enhanced MRI [C] // 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). July 18–21, 2018, Honolulu, HI, USA. IEEE, 2018: 750–753.
- [8] 马明明, 姜原, 刘义, 等. U-Net深度学习模型对DCE-MRI上乳腺肿块自动分割和定位的准确性分析 [J]. *放射学实践*, 2020, 35(8): 1030–1036.
- [9] RAO A A, FENEIS J, LALONDE C, et al. A pictorial review of changes in the BI-RADS fifth edition [J]. *Radiographics*, 2016, 36(3): 623–639.
- [10] HA R, CHANG P, MEMA E, et al. Fully automated convolutional neural network method for quantification of breast MRI fibroglandular tissue and background parenchymal enhancement [J]. *J Digit Imaging*, 2019, 32(1): 141–147.
- [11] DILORENZO G, TELEGRAFO M, LA FORGIA D, et al. Breast MRI background parenchymal enhancement as an imaging bridge to molecular cancer sub-type [J]. *Eur J Radiol*, 2019, 113: 148–152.

(收稿日期: 2021-09-20 修回日期: 2021-09-25)