

基于临床 - 影像多模态融合的乳腺MRI复发风险预测

郝凯亮

青岛大学数学与统计学院, 山东 青岛

收稿日期: 2026年3月3日; 录用日期: 2026年3月27日; 发布日期: 2026年4月8日

摘要

乳腺癌复发风险评估对于术后随访管理、辅助治疗调整及高危患者的早期干预具有重要意义。为比较临床模型、影像模型及临床-影像多模态融合模型在乳腺MRI复发风险预测中的表现,并验证融合策略在低阳性率场景下的应用价值,本文基于Duke-Breast-Cancer-MRI公开队列,完成临床变量清洗、编码与筛选以及病灶ROI三通道输入(Pre/Post/Sub)构建,并采用固定随机种子进行分层划分,得到训练集、验证集和测试集,样本数分别为588、148和184。临床变量筛选在训练集内依次进行低方差剔除、高相关特征过滤及单变量相关性筛选,其中低方差阈值设为0.01,特征间绝对相关系数阈值设为0.90,最终保留67个临床变量进入建模。针对阳性样本比例约9.46%的类别不平衡问题,在训练阶段结合加权采样、Focal Loss与类别权重进行优化,在推理阶段采用“召回优先+特异度下限0.75”的阈值选择策略。结果显示,在测试集上,Fusion-DualMeta模型的AUC、AUPRC、Sensitivity、Specificity、F1值和MCC分别为0.8633、0.2801、0.9412、0.7844、0.4638和0.4667;与Clinical-XGBoost模型相比,其AUC、F1值、Sensitivity和MCC分别提高0.0673、0.1480、0.4118和0.2253;与Image-EmbROI模型相比,上述指标分别提高0.0902、0.1951、0.4118和0.2818,且假阴性病例数由8例降至1例。结合样本规模、类别不平衡程度及模型可解释性需求,本文采用基于元学习的后融合策略,并对其与特征拼接、注意力机制融合及张量积融合等方法适用性进行了讨论。研究表明,在保持特异度约束的前提下,临床-影像多模态融合模型能够显著增强乳腺癌复发高风险患者的识别能力,可为乳腺癌随访筛查与辅助决策提供更具实用价值的技术支持。

关键词

乳腺MRI, 复发预测, 多模态融合, 类别不平衡, XGBoost

Recurrence Risk Prediction on Breast MRI via Clinical-Image Multimodal Fusion

Kailiang Hao

School of Mathematics and Statistics, Qingdao University, Qingdao Shandong

Abstract

Assessment of breast cancer recurrence risk is of great importance for postoperative follow-up management, adjuvant treatment adjustment, and early intervention in high-risk patients. To compare the performance of clinical, imaging, and clinical-image multimodal fusion models for recurrence prediction on breast MRI, this study was conducted on the Duke-Breast-Cancer-MRI public cohort. Clinical variables were cleaned, encoded, and filtered, and tumor-centered ROI three-channel inputs (Pre/Post/Sub) were constructed. The data were split into training, validation, and test sets with sizes of 588, 148, and 184, respectively. Clinical variable screening was performed on the training set only, including low-variance removal (threshold = 0.01), high-correlation filtering ($|r| > 0.90$), and univariate relevance screening, resulting in 67 retained variables. To address the strong class imbalance with a recurrence-positive rate of about 9.46%, weighted sampling, Focal Loss, and class weighting were adopted during training, while a recall-prioritized thresholding strategy with a minimum specificity constraint of 0.75 was applied during inference. On the test set, the Fusion-DualMeta model achieved an AUC of 0.8633, an AUPRC of 0.2801, a sensitivity of 0.9412, a specificity of 0.7844, an F1 score of 0.4638, and an MCC of 0.4667. Compared with Clinical-XGBoost, the improvements in AUC, F1 score, sensitivity, and MCC were 0.0673, 0.1480, 0.4118, and 0.2253, respectively. Compared with Image-EmbROI, the corresponding gains were 0.0902, 0.1951, 0.4118, and 0.2818, with false negatives reduced from 8 to 1. Considering sample size, class imbalance, and interpretability, a meta-learning-based late-fusion strategy was adopted and discussed against other multimodal fusion approaches. The results indicate that under a specificity-constrained setting, clinical-image multimodal fusion can substantially improve the identification of high-risk recurrence patients and may provide useful support for follow-up screening and decision-making in breast cancer care.

Keywords

Breast MRI, Recurrence Prediction, Multimodal Fusion, Class Imbalance, XGBoost

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

乳腺癌复发风险评估直接关系到术后随访频率、辅助治疗方案调整及高危患者的早期干预，因此建立准确、稳定的复发预测模型具有重要临床意义。已有研究表明，基于乳腺 MRI 的影像特征分析能够从病灶形态、强化模式及肿瘤异质性中提取与预后相关的信息，为复发风险预测提供重要依据[1]。同时，公开影像数据库的建设也为可重复、可验证的模型研究提供了数据基础，其中 Duke-Breast-Cancer-MRI 队列因同时提供 MRI 影像与临床信息而具有较高研究价值[2]。在建模方法方面，XGBoost 等集成学习模型在结构化临床数据处理中表现稳定，适用于处理中小样本、缺失值较多及特征维度较复杂的二分类问题[3]；而 EfficientNet 等卷积网络则能够较好地提取医学图像中的深层视觉表征[4]。此外，针对不平衡分类任务，Focal Loss 能够降低易分样本主导效应，从而提升模型对少数类的关注[5]。在评价层面，MCC、AUPRC 等指标被认为比单纯准确率更适合用于阳性率较低的医学预测问题[6]-[8]。

基于上述研究基础，乳腺癌复发预测逐步从单一临床变量分析扩展到影像特征建模，再进一步发展到临床-影像多模态联合分析。已有研究提示，临床病理因素可提供较稳定的群体层面风险信息，而MRI视觉表征更有助于刻画肿瘤内部异质性和强化行为，两者可能具有互补性[9]-[11]。近年来，面向乳腺癌复发风险评估的多模态深度学习研究也开始增多，提示融合模型有望进一步提升预测表现[12][13]。

现有复发预测研究仍存在两方面不足：一是部分研究依赖私有数据或缺乏统一的数据处理流程，尤其在临床变量筛选步骤中，未明确说明低方差别除、高相关特征过滤及单变量筛选所采用的具体阈值，影响结果的可复现性；二是在多模态建模中，虽然融合模型通常优于单模态模型，但对不同融合策略的适用场景讨论仍不充分，尤其缺少对特征拼接、注意力机制融合、张量积融合及后融合等方法优缺点的说明。与此同时，在复发样本占比较低的条件下，若仅关注总体准确率，也容易掩盖对高风险病例的漏检问题，而临床筛查场景通常更强调对阳性病例的优先识别[6]-[8]。

基于此，本文围绕 Duke-Breast-Cancer-MRI 公开队列构建临床分支、影像分支与融合分支三类模型，在统一数据划分、统一评价指标和统一阈值决策原则下开展对比研究。本文旨在建立从数据清洗到模型评估的可复现实验流程，明确临床变量筛选标准，比较不同模态模型在 AUC、AUPRC、F1 值、MCC 及混淆矩阵等方面的差异，并进一步讨论为何在本研究中选用基于元学习的后融合策略，以及该策略在低阳性率乳腺癌复发筛查场景中的应用价值。

2. 资料与方法

2.1. 数据来源与研究终点

本研究采用 Duke-Breast-Cancer-MRI 公开队列作为研究数据来源[2]。该队列包含乳腺癌患者的MRI影像资料及相应临床信息，能够为乳腺癌复发风险预测提供影像表型和结构化变量两类数据支持。结合本研究任务，对原始数据进行样本筛选、变量整理及标签对齐后，构建用于复发风险分类的研究队列。

根据既定纳入标准，最终获得可用于建模分析的样本共 920 例，其中训练集、验证集和测试集分别为 588 例、148 例和 184 例。数据集划分采用固定随机种子下的分层抽样方法，以保证不同子集中阳性与阴性样本分布基本一致。总体样本中复发阳性比例约为 9.46%，呈现明显类别不平衡特征，因此在模型训练与评价过程中需要重点关注少数类样本的识别能力。

2.2. 临床数据预处理

临床分支以公开临床特征表为输入，首先统一列名、异常值和缺失值标记；随后保留可被数值化编码的结构化变量，并在训练集内部完成缺失填补、特征筛选与数值转换，以避免信息泄露。

临床变量筛选共分三步进行。第一步为低方差别除：对经编码后的候选变量计算训练集内方差，删除方差小于 0.01 的特征，以去除几乎不提供区分信息的变量。第二步为高相关特征过滤：对保留变量计算两两相关系数，当绝对相关系数大于 0.90 时，删除其中与结局单变量相关性较弱的变量，以减少多重共线性影响。第三步为单变量相关性筛选：计算各变量与复发标签之间的单变量相关性，按照绝对相关性从高到低排序，最终保留前 67 个变量进入建模。所有筛选阈值均在训练集内确定后固定应用于验证集和测试集。

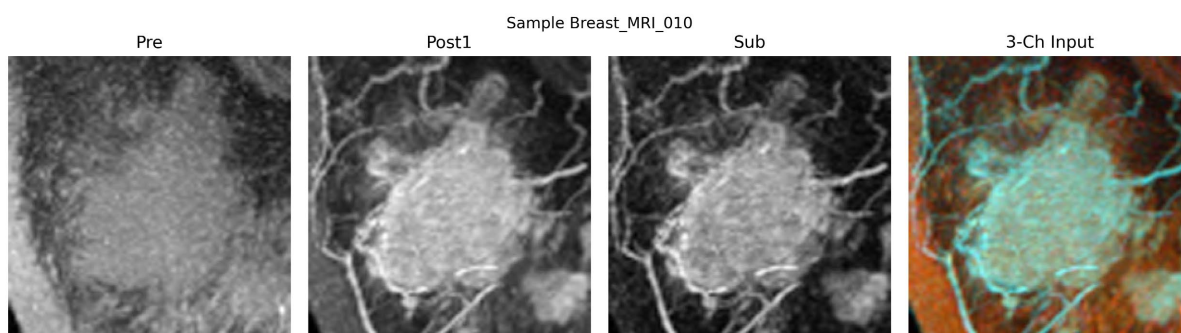
在变量表示方面，连续变量采用训练集内中位数进行缺失填补，并进行标准化处理；分类变量统一编码，并对缺失或未定义类别设置显式的“Unknown”映射，以保持表示完整性。最终得到的 67 个临床变量覆盖人口学信息、生物标志物、肿瘤分期分级、影像元数据、治疗信息及肿瘤部位等主要风险相关维度(见表 1)。

Table 1. Grouping statistics of clinical input variables**表 1.** 临床输入变量分组统计

Group	FeatureCount	AvgMissingRate
Biomarker	28	15.6%
ImagingMeta	12	5.9%
Demographics	7	34.6%
Other	7	0.4%
Tumor	7	18.8%
Outcome/Time	6	11.9%

2.3. 影像输入构建

影像分支以病灶区域为核心输入单元。根据已有病灶定位信息提取感兴趣区域(region of interest, ROI), 并构建三通道输入, 包括治疗前增强图像(Pre)、增强后图像(Post)及差值图像(Sub)。其中, 差值图像用于突出强化变化信息, 以增强模型对病灶动态增强特征的表征能力[1] [9]。为保证模型输入一致性, 所有 ROI 图像在送入网络前均进行尺寸统一及数值归一化处理(见图 1)。

**Figure 1.** Example of Pre/Post/Sub channels and 3-channel fusion input for lesion ROI**图 1.** 病灶 ROI 的 Pre/Post/Sub 及三通道融合输入示例

2.4. 数据划分与类别不平衡处理

本文采用固定随机种子(seed = 42)进行分层划分, 训练集、验证集和测试集规模分别为 588、148 和 184 例, 各子集阳性比例基本一致, 见表 2。针对阳性样本稀缺的问题, 图像训练阶段采用 WeightedRandomSampler 以提升少数类采样概率, 损失函数采用 Binary Focal Loss 以降低易分类负样本的主导效应[5]; 临床树模型则通过 scale_pos_weight 引入类别权重[3] [7]。在模型选择阶段, 本文采用“召回优先 + 特异度下限 0.75”的阈值决策原则, 以更贴近复发高危筛查场景。

Table 2. Dataset split and positive rate**表 2.** 数据集划分与阳性比例

Split	N	Positive	Negative	PosRate
Train	588	56	532	9.52%
Val	148	14	134	9.46%
Test	184	17	167	9.24%
Total	920	87	833	9.46%

2.5. 模型构建

2.5.1. Clinical-XGBoost

临床最终模型采用 XGBoost 二分类框架[3]，核心超参数设置为 $n_estimators = 300$ 、 $max_depth = 5$ 、 $learning_rate = 0.01$ 、 $colsample_bytree = 0.75$ ，并根据训练集阳性比例动态计算 $scale_pos_weight$ 。模型输出为复发概率，并通过验证集阈值搜索确定测试阶段的最终决策点。

2.5.2. Image-EmbROI

图像最终模型采用“双分支融合”策略：其一为迁移学习深度嵌入分支，使用 EfficientNet-B0 提取视觉表征，再经标准化、PCA 与 Logistic 回归得到概率输出[4]；其二为 ROI 统计分支，在病灶与背景区域上构建强度均值、标准差、分位数及差值/比值等统计特征，并使用 XGBoost 建模。两分支概率通过验证集搜索得到的权重进行线性融合，形成 Image-EmbROI 输出。

2.5.3. Fusion-DualMeta

融合模型采用轻量级元学习策略，不直接拼接高维原始特征，而是以前序模型输出概率为核心输入，包括临床概率 p_clin 、图像旧分支概率 p_img_old 以及图像新分支概率 p_img_new ，并进一步构造平方项、乘积项和两两绝对差等阶交互特征[12][13]。最终以 Logistic 回归完成融合判别，使模型在保持解释性的同时增强泛化稳定性。

本文采用该后融合策略，主要基于以下考虑。首先，特征拼接属于早融合方法，虽然实现简单，但在本研究样本规模有限、影像与临床特征维度差异较大的情况下，易引入维度膨胀和冗余噪声，增加过拟合风险。其次，注意力机制融合能够自适应学习不同模态的重要性，但通常需要更大样本量和更复杂的端到端训练过程，对当前 920 例样本且阳性率不足 10% 的场景未必最优。再次，张量积融合能够显式建模跨模态高阶交互，但参数量增长较快，对数据规模和训练稳定性要求更高。相比之下，本文采用的元学习后融合方法仅以各单模态模型输出概率及其有限高阶交互作为输入，一方面能够保留临床分支与影像分支各自已经学到的有效判别信息，另一方面能够在低维空间中完成跨模态整合，更适合中小样本、不平衡分类及强调可解释性的应用场景。

2.6. 评价指标

为全面评价不同模型在类别不平衡场景下的预测性能，本文采用受试者工作特征曲线下面积(area under the receiver operating characteristic curve, AUC)、精确率 - 召回率曲线下面积(area under the precision-recall curve, AUPRC)、敏感度(Sensitivity)、特异度(Specificity)、F1 值及 Matthews 相关系数(MCC)作为主要评价指标[6]-[8]。

其中，AUC 用于衡量模型整体区分阳性与阴性样本的能力；AUPRC 更适用于阳性样本比例较低的任务，可更直观地反映模型对少数类样本的识别表现；Sensitivity 和 Specificity 分别反映模型对阳性样本的检出能力和对阴性样本的排除能力；F1 值综合考虑精确率与召回率的平衡；MCC 则能够在类别不平衡条件下较全面地评价分类性能[6]。此外，本文还结合混淆矩阵分析假阳性与假阴性样本分布，以进一步评估模型在临床筛查场景中的应用价值。

3. 结果

3.1. 三模型主结果对比

测试集结果见表 3。整体来看，Clinical-XGBoost 和 Image-EmbROI 均表现出较为稳定的分类能力，可作为乳腺癌复发风险预测任务中的单模态基线模型。其中，Clinical-XGBoost 反映了结构化临床变量在

风险分层中的基础价值,说明年龄、分期及相关临床特征对复发风险判断具有一定贡献; Image-EmbROI 则体现了 MRI 病灶区域视觉表征在捕捉肿瘤内部异质性方面的优势,表明影像信息同样能够为复发预测提供有效支持[9]-[11]。在此基础上, Fusion-DualMeta 在 AUC、AUPRC、Sensitivity、F1 值和 MCC 等多个关键指标上均取得最佳结果,显示出临床信息与影像特征具有明显的互补性[12][13]。相较于单一模态模型,多模态融合不仅提高了整体判别能力,也增强了模型在类别不平衡条件下对阳性样本的识别效果,说明融合策略能够更全面地表征影响乳腺癌复发的多维信息。

尤其值得注意的是, Fusion-DualMeta 在召回率方面达到 0.9412,明显高于 Clinical-XGBoost 和 Image-EmbROI 的 0.5294。这表明融合模型能够识别出绝大多数复发病例,在高危筛查任务中具有更高的实际应用价值。对于乳腺癌复发风险评估而言,漏检往往意味着高风险患者未能被及时纳入强化随访或进一步干预,因此模型对阳性样本的检出能力具有优先意义。从这一角度看,融合模型在 Sensitivity 上的显著提升,不仅是统计指标上的改善,也意味着其在临床应用场景中更有利于实现对复发高危人群的前置识别。与此同时, Fusion-DualMeta 在 F1 值和 MCC 上也取得最高水平,说明该模型并非单纯通过牺牲整体平衡性换取高召回,而是在正负样本综合判别上实现了更优结果(见图 2)。

Table 3. Main comparison of the three final models

表 3. 三个最终模型的主结果对比

Method	AUC	AUPRC	Sensitivity	Specificity	F1	MCC
Clinical-XGBoost	0.7961	0.2612	0.5294	0.8144	0.3158	0.2414
Fusion-DualMeta	0.8633	0.2801	0.9412	0.7844	0.4638	0.4667
Image-EmbROI	0.7732	0.2288	0.5294	0.7545	0.2687	0.1848

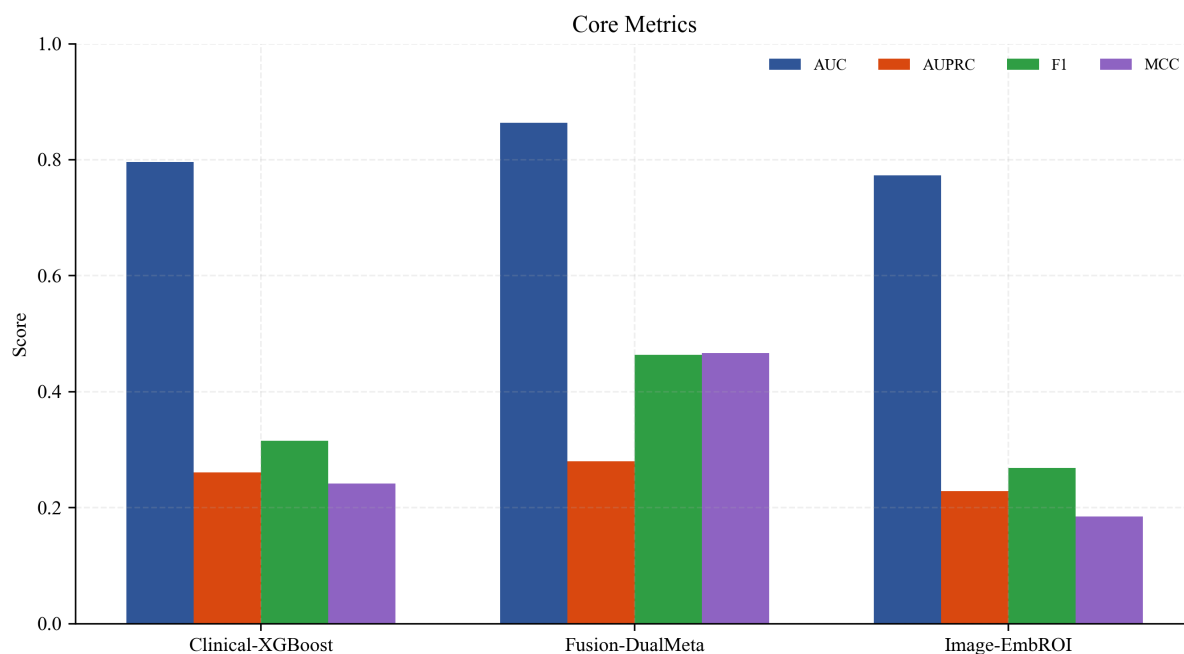


Figure 2. Comparison of core metrics among the three models

图 2. 三模型核心评价指标对比

3.2. ROC 与 PR 曲线分析

ROC 曲线如图 3 所示。可以看出, Fusion-DualMeta 的 ROC 曲线整体位于两种单模态模型之上,对

应 AUC 达到 0.8633，优于 Clinical-XGBoost 和 Image-EmbROI。这表明在不同分类阈值下，融合模型均表现出更强的区分能力，即其对复发与非复发样本的排序能力更优。从判别分析角度看，AUC 的提升说明融合模型在更广泛的阈值范围内保持了较好的稳定性，不依赖单一阈值点即可展现优势。对于临床应用而言，这种特性意味着模型在面对不同风险控制要求时具有更好的适应性，既可用于偏重高召回的筛查场景，也可在一定程度上支持更审慎的风险分层判断。

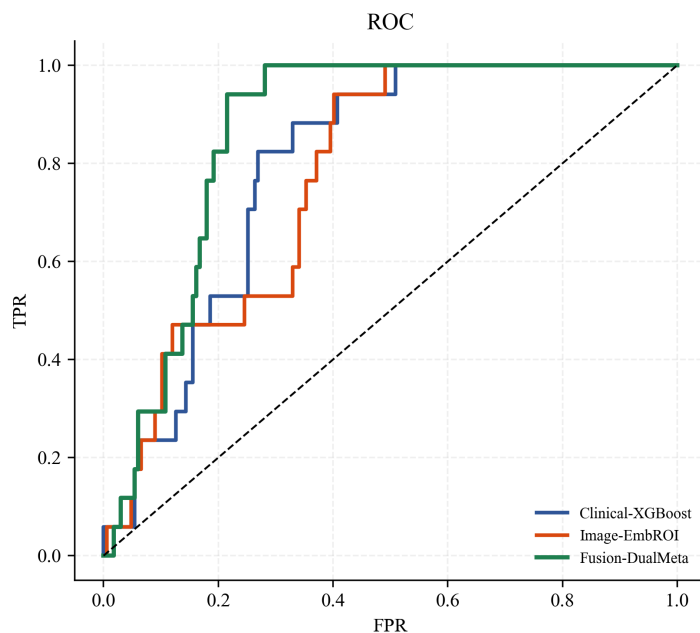


Figure 3. ROC curves of the three models
图 3. 三模型 ROC 曲线对比

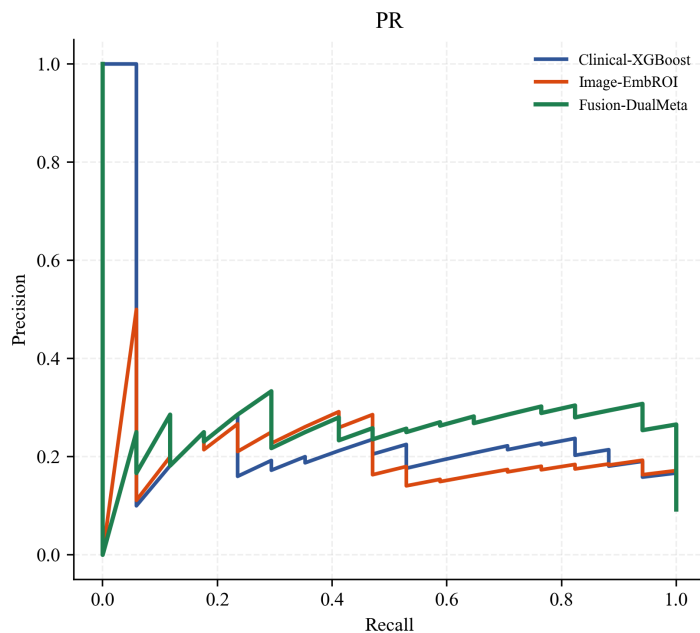


Figure 4. PR curves of the three models
图 4. 三模型 PR 曲线对比

进一步观察 PR 曲线(图 4)可以发现, Fusion-DualMeta 在低召回和中高召回区间总体保持更优的精确率水平, 对应 AUPRC 达到 0.2801, 也高于两种单模态模型。由于本研究中复发阳性样本占比仅约 9.46%, 类别不平衡较为明显, 在这种情况下, PR 曲线较 ROC 曲线更能反映模型对阳性类别的真实识别能力[6][7]。融合模型在 PR 曲线上的优势说明, 其不仅扩大了真阳性样本的覆盖范围, 而且在提升召回率的同时较好控制了无效阳性预测的增长。这意味着模型在识别更多复发病例的同时, 并未出现过度激进的正类判断, 从而在一定程度上抑制了正类识别过程中的噪声累积。换言之, Fusion-DualMeta 在不平衡数据环境下实现了更合理的精确率 - 召回率平衡, 这也是其 AUPRC 和 F1 值得以同步提升的重要原因。

3.3. 漏检控制与融合增益

图 5 给出了三种模型在测试集上的混淆矩阵。结果显示, 与 Clinical-XGBoost 和 Image-EmbROI 相比, Fusion-DualMeta 将假阴性数量由 8 例降至 1 例, 同时真阳性数量由 9 例提升至 16 例, 说明融合模型对于复发样本的识别能力获得了实质性增强。从临床视角看, 假阴性减少具有更为重要的现实意义, 因为假阴性病例意味着真实高风险患者被误判为低风险, 从而可能错失后续强化监测、辅助治疗调整或早期干预机会。融合模型仅保留 1 例假阴性, 表明其在“尽量少漏检”的目标上取得了明显改进。虽然与此相伴, 假阳性数量较单模态模型略有增加, 但其特异度仍保持在 0.7844, 说明模型并未因追求高召回而导致阴性样本识别能力明显失衡。对于复发风险筛查任务而言, 这种以小幅增加误检为代价, 换取显著降低漏检的结果, 更符合临床实际需求, 也更具应用可接受性(见表 4)。

Table 4. Performance gain of the fusion model over single-modality models

表 4. 融合模型相对单模态模型的性能增益

Compare	AUC Gain	F1 Gain	Sensitivity Gain	MCC Gain
Fusion-DualMeta vs Clinical-XGBoost	+0.0673	+0.1480	+0.4118	+0.2253
Fusion-DualMeta vs Image-EmbROI	+0.0902	+0.1951	+0.4118	+0.2818

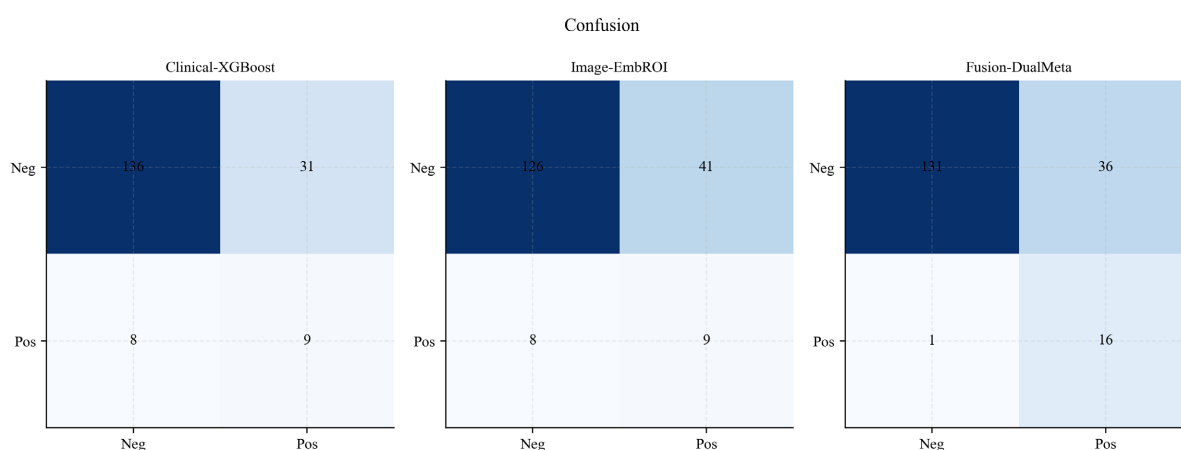


Figure 5. Confusion matrices of the three models on the test set

图 5. 三模型在测试集上的混淆矩阵对比

4. 讨论

本文结果表明, 临床信息与 MRI 视觉证据之间具有较好的互补性。临床分支主要反映年龄、分子标志物、分期分级及治疗信息等结构化风险特征, 能够提供相对稳定的群体统计信息; 影像分支则侧重病

灶强化模式、边界及减影差异等局部表征,更有助于刻画肿瘤异质性[9]-[11]。Fusion-DualMeta 通过对两类信息进行协同建模,在 AUC、F1 值、Sensitivity 及 MCC 等指标上均优于单模态模型,说明多模态融合能够更全面地提取与乳腺癌复发相关的风险信息[12][13]。需要指出的是,本文采用元学习进行后融合,并非默认其在所有多模态任务中均优于其他策略,而是综合本研究数据特征后作出的方法学选择。对于特征拼接方法,其优势在于实现直接、流程简洁,但在本研究中,临床变量与深度影像表征维度差异较大,直接拼接容易带来维度膨胀、特征尺度不一致及过拟合风险。对于注意力机制融合,其优势在于能够动态建模不同模态的重要性分配,但这类方法往往依赖更充足的训练样本和更复杂的端到端优化流程;在阳性样本较少的条件下,训练稳定性和泛化能力可能受到影响。对于张量积融合,其可以更充分地表达跨模态高阶交互,但通常伴随参数规模快速增长,对样本量、算力和正则化设计要求更高。相比之下,本文所采用的后融合元学习方法以单模态输出概率及少量交互项作为输入,一方面降低了融合层参数复杂度,另一方面保留了各单模态模型的独立判别能力,并使最终模型具有更好的可解释性和部署便利性。

从应用场景看,复发筛查任务中漏检代价通常高于误检代价,因此本文采用“召回优先 + 特异度下限”的阈值策略,而非单纯追求 Accuracy 最大化[6]-[8]。结果显示,在 Specificity 仍保持 0.78 以上的情况下,融合模型将假阴性由 8 例降至 1 例,表明该模型更有利于优先识别潜在高危患者。这种以较小误检代价换取明显漏检下降的结果,更符合乳腺癌术后高风险筛查的实际需求。

本文仍存在一定局限。首先,研究仅基于单一公开队列完成验证,缺乏多中心外部测试;其次,影像预处理仍依赖病灶框注释,自动 ROI 提取的稳定性仍有提升空间;再次,本文将复发问题处理为二分类任务,尚未纳入时间到事件信息[10]-[13]。后续可结合外部验证、自动病灶定位以及生存分析建模等方向进一步完善。

5. 结论

本文基于 Duke-Breast-Cancer-MRI 公开队列,建立了从临床变量预处理、ROI 三通道输入构建到临床模型、影像模型及融合模型对比评估的完整流程。结果表明,在类别不平衡且强调漏检控制的条件下,Fusion-DualMeta 相较于 Clinical-XGBoost 和 Image-EmbROI 取得了更优的 AUC、F1 值、Sensitivity 和 MCC,尤其在假阴性控制方面优势明显。

综合方法复杂度、样本规模、不平衡特征及临床可解释性要求,本文认为基于元学习的后融合策略更适合当前研究场景。研究结果说明,临床结构化变量与 MRI 视觉特征具有明显互补性,多模态融合能够更有效地识别乳腺癌复发高风险患者,可为术后随访筛查和辅助决策提供具有应用潜力的技术支持。

致 谢

感谢 TCIA 与 Duke 队列提供公开可用的数据资源,感谢相关开源工具和研究工作为本研究流程复现提供支持。

参考文献

- [1] Saha, A., Harowicz, M.R., Grimm, L.J., Kim, C.E., Ghate, S.V., Walsh, R., *et al.* (2018) A Machine Learning Approach to Radiogenomics of Breast Cancer: A Study of 922 Subjects and 529 DCE-MRI Features. *British Journal of Cancer*, **119**, 508-516. <https://doi.org/10.1038/s41416-018-0185-8>
- [2] The Cancer Imaging Archive (TCIA) (2022) Duke-Breast-Cancer-MRI: Dynamic Contrast-Enhanced Magnetic Resonance Images of Breast Cancer Patients with Tumor Locations.
- [3] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 785-794. <https://doi.org/10.1145/2939672.2939785>

-
- [4] Tan, M. and Le, Q.V. (2019) EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *Proceedings of the 36th International Conference on Machine Learning*, Long Beach, 9-15 June 2019, 6105-6114.
- [5] Lin, T., Goyal, P., Girshick, R., He, K. and Dollar, P. (2017) Focal Loss for Dense Object Detection. *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 2980-2988. <https://doi.org/10.1109/iccv.2017.324>
- [6] Chicco, D. and Jurman, G. (2020) The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation. *BMC Genomics*, **21**, Article No. 6. <https://doi.org/10.1186/s12864-019-6413-7>
- [7] Haibo He, and Garcia, E.A. (2009) Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*, **21**, 1263-1284. <https://doi.org/10.1109/tkde.2008.239>
- [8] Thakur, D., Gera, T., Bhardwaj, V., Mazen, R., Lasisi, A. and Engida, T. (2025) A Comparative Study on Advanced Predictive Modeling of Thyroid Cancer Recurrence Using Multi Algorithmic Machine Learning Frameworks. *Scientific Reports*, **16**, Article No. 3385. <https://doi.org/10.1038/s41598-025-33396-7>
- [9] Li, H., Zhu, Y., Burnside, E.S., Drukker, K., Hoadley, K.A., Fan, C., *et al.* (2016) MR Imaging Radiomics Signatures for Predicting the Risk of Breast Cancer Recurrence as Given by Research Versions of Mammprint, Oncotype DX, and PAM50 Gene Assays. *Radiology*, **281**, 382-391. <https://doi.org/10.1148/radiol.2016152110>
- [10] Koh, J., Lee, E., Han, K., Kim, S., Kim, D., Kwak, J.Y., *et al.* (2020) Three-Dimensional Radiomics of Triple-Negative Breast Cancer: Prediction of Systemic Recurrence. *Scientific Reports*, **10**, Article No. 2976. <https://doi.org/10.1038/s41598-020-59923-2>
- [11] Xu, K., Hua, M., Mai, T., Ren, X., Fang, X., Wang, C., *et al.* (2024) A Multiparametric MRI-Based Radiomics Model for Stratifying Postoperative Recurrence in Luminal B Breast Cancer. *Journal of Imaging Informatics in Medicine*, **37**, 1475-1487. <https://doi.org/10.1007/s10278-023-00923-9>
- [12] Zhang, R., Wang, K., Wang, S., Wang, C., Cao, T., Ci, C., *et al.* (2025) Multimodal Deep Learning Model for Prediction of Breast Cancer Recurrence Risk and Correlation with Oncotype DX. *Breast Cancer Research*, **27**, Article No. 178. <https://doi.org/10.1186/s13058-025-02129-z>
- [13] Yu, Y., Ren, W., Mao, L., Ouyang, W., Hu, Q., Yao, Q., *et al.* (2025) MRI-Based Multimodal AI Model Enables Prediction of Recurrence Risk and Adjuvant Therapy in Breast Cancer. *Pharmacological Research*, **216**, Article ID: 107765. <https://doi.org/10.1016/j.phrs.2025.107765>