

基于改进 ResNet 的乳腺癌病理图像分类研究

高 谦¹
GAO Qian

摘 要

乳腺癌是一种发病率和致死率较高的恶性肿瘤,及时诊断对于治疗乳腺癌具有重要意义。当前计算机对乳腺癌病理图像分类研究虽有进展,但传统机器学习依赖人工特征提取,深度学习模型存在特征表达不足和样本分布不均衡等问题。针对以上问题,文章提出了一种基于改进 ResNet50 的乳腺癌病理图像多分类方法。为弥补传统模型在处理高分辨率图片时的不足,使用 RFACnv 卷积替代传统卷积,同时加入 EMA (efficient multi-scale attention) 注意力模块,实现对精细图片的上下文信息融合,利用跨通道特征重组策略,实现对图片特征信息的优化提取。

关键词

乳腺癌病理图像分类;残差网络;卷积;多尺度特征融合;注意力机制

doi: 10.3969/j.issn.1672-9528.2026.03.004

0 引言

根据世界卫生组织于 2020 年发布的全球癌症流行病学监测数据,在全球范围内,乳腺癌的新发病例数量大约为 226 万例,这些病例占女性确诊癌症总数的 11.7%,乳腺癌是女性群体中发病率最高的恶性肿瘤类型之一,当年其新发病例数超过了肺癌(约 220 万例)的新发病例数。尽管乳腺癌的确诊基数庞大,但在癌症导致的死亡人数中,乳腺癌的致死率排在第五位,这一结果表明,与其他癌症相比,乳腺癌转移和恶化的速度相对较慢。进一步的分析结果显示,处于 I 期的乳腺癌患者,5 年生存率为 99%,在临床上已接近治愈的标准。从这一看似矛盾的现象中,可知在乳腺癌的诊断治疗过程中,早期确诊和分期十分重要,尽早进行治疗干预对于延长患者的生命周期至关重要。因此,目前医生和研究人员正在攻关的难题是,如何借助高分辨率的影像诊断技术提高对早期癌症的识别能力。

目前学术界针对乳腺癌病理图像的自动分类,存在两种不同的发展路线:一种是借助人工特征工程的机器学习框架,如 Kabir 等^[1]基于超声图像提取统计与纹理特征,结合时序建模技术,提取图片信息,进而利用 KNN^[2]和 SVM 等^[3]经典分类算法构建分类器;另一种则是基于深度学习的端到端分类模式,通过卷积神经网络(convolutional neural network, CNN)实现图像特征的自动编码和语义映射。在这一技术基础上, Sammut 等^[4]打破了常规单一模态分析的局限性,在数据输入端,将病理图片、临床参数和基因组不同结构的数据整合,构建了一个多模态融合的输入。同时,在输出端设

计了融合 Logistic 回归和随机森林的混合集成系统。这种不再局限于单一模态数据的分类器,具有较强的生物学可解释性。这两种方法分别从特征工程优化和多源信息融合的角度推动了乳腺癌病理诊断智能化的发展。

卷积神经网络在图像处理领域具有优秀的上下文信息整合能力,且已在医学影像处理领域得到广泛应用并推动了多项研究成果的诞生。Xu 等^[5]提出了一种改进的卷积神经网络架构,有效提高了乳腺癌二分类的精度。他们根据应用的场景、SE 模块的位置和数量,设计了三种 MFSCNET 变体,并且验证了三种变体在不同的环境中各有其优越性。与此同时, ElAghouri 等^[6]比较了 ResNet 和 Xception 架构在 Moroccan H&E 染色切片数据集上的分类效果。

在对病理图像进行特征提取时,无论是传统机器学习模型还是目前流行的深度学习框架,普遍存在着特征提取不够充分的问题。针对上述问题,本研究提出了一种基于改进残差网络的乳腺癌病理图像多分类方法。该方法以 ResNet50 为架构基础,引入 RFACnv 卷积,并创造性地嵌入 EMA 注意力机制至每个残差单元中,以优化特征提取并提升分类性能。

1 材料和方法

1.1 方法

1.1.1 ResNet 模型

在卷积神经网络模型中,随着网络层数不断增多,有可能出现模型过拟合以及退化等状况。2015 年, He 等^[7]参考了 LSTM 中门控机制的想法,在卷积神经网络的层与层之间引入了恒等映射,使得输入数据可以直接跟经过非线性变换后的输出相加,通过残差映射, ResNet 能直接学习输入与输

1. 湖北工业大学 湖北武汉 430068

出之间的信息差,保证了网络每一层都是在优化,解决了深层架构导致的性能退化问题。

ResNet 是由多个残差模块组成的,这些模块借助跳跃连接实现了输入与输出的元素级相加,这种简单的运算方式避免了额外参数的引入以及计算负担的加重,还提升了模型训练的效率以及最终的表现。根据堆叠的残差块数量不同,ResNet 模型有 5 种不同深度的版本,本文选择 ResNet50 为基础架构。其具体的网络配置如图 1 所示。

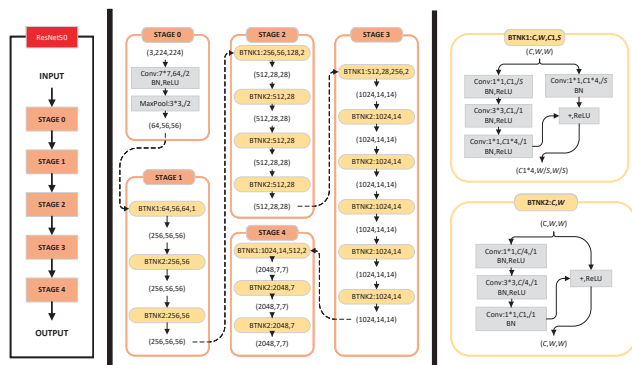


图 1 网络配置图

1.1.2 RFACnv 卷积块

在传统的卷积神经网络中,卷积核参数共享,这一操作虽然大幅降低了模型的复杂程度和计算成本,但同时也导致感受野特征同质化等问题,Zhang 等^[8]提出基于感受野注意力机制的新型卷积模块 RFACnv。该方法通过动态生成与卷积核尺寸匹配的注意力权重,将空间注意力与感受野特征解耦,解决了大卷积核在传统注意力机制下的参数共享问题。同时,通过 GroupConv 技术,快速提取感受野内特征,实现特征差异化的高效建模。其核心设计用公式表示为:

$$F = \text{Softmax}(g^{1 \times 1}(\text{AvgPool}(X))) \times \text{ReLU}(\text{Norm}(g^{k \times k}(X))) = A_{1 \times 1} \times F_{k \times k} \quad (1)$$

式中: $g^{k \times k}$ 指分组卷积。

与传统固定窗口大小的卷积不同,RFACnv 卷积借助解耦达成参数共享约束,促使每个感受野里的卷积核权重依据空间特征的变化进行动态调整,在具体的实现过程中,此方法运用分组卷积提取局部特征,接着把多个局部特征依靠全局平均池化来压缩整合,最后利用轻量级 1×1 卷积与 Softmax 函数生成空间注意力图。在改进 ResNet50 时,将 Bottleneck 结构里的 3×3 标准卷积替换成 RFACnv 模块,在具体实现中,输入特征先依靠 4 组分组卷积提取局部感受野特征,接着利用通道压缩生成注意力权重,再将其与归一化特征进行加权融合。这种设计保留了残差学习框架的优势,还凭借动态权重调整提高了局部特征的代表能力,模块内部采用分组卷积和参数共享策略,使得计算开销的增加低于 0.1 GFLOPs,并且在 ImageNet-1K 数据集上实现了 Top-1 准

确率 1.6% 的提升(从 69.59% 提升至 71.23%)。

1.1.3 插入 EMA 注意力机制

Ouyang 等^[9]提出一种新的卷积注意力模块 EMA。为高效地聚合多尺度空间结构信息,EMA 模块引入了并行部署的 1×1 和 3×3 卷积分支,建立短距离和长距离依赖关系。其中, 1×1 卷积分支用于实现快速通道变换,而 3×3 卷积分支则专注于捕捉更广泛的空间上下文。两个分支的输出特征图通过跨空间整合策略进行融合,确保了不同尺度信息的有效结合。此外,EMA 模块的设计特别考虑了特征分组的重要性,旨在通过优化特征间的依赖关系来增强模型性能。与现有的注意力模块相比,EMA 模块不仅提高了模型的准确率,而且使参数更加高效。

EMA 模块采用并行子网络结构,含两个分支: 1×1 卷积分支可减少通道数、加快信息传递并保持特征多样性,功能类似坐标注意力模块; 3×3 卷积分支能扩大感受野,增强对局部细节和全局上下文的感知,以捕获多尺度空间结构信息。

各并行分支中,EMA 模块会重新校准通道权重。通过融合两分支输出特征图以捕获像素级成对关系,该模块不仅有效整合多尺度信息,还提升了特征表达的丰富性与准确性,显著增强模型性能。

1.2 数据集与实验方案

1.2.1 BreakHis 数据集

本研究采用 2016 年公开的 BreakHis 数据集验证模型。该数据集含 82 位患者的 7 909 张乳腺组织病理学图像,涵盖 4 种放大倍率,分别是 $40 \times$ 、 $100 \times$ 、 $200 \times$ 、 $400 \times$,包括 2 480 张良性、5 429 张恶性病变图像,并按 WHO 分类标准标注详细病理亚型:良性含腺瘤(A)、纤维腺瘤(F)、叶状瘤(PT)、管状腺瘤(TA);恶性含导管癌(DC)、小叶癌(LC)、黏液癌(MC)、乳头状癌(PC)。其精确到亚型的特性,不仅适用于二分类,也为多分类提供了良好条件。

1.2.2 参数设置及评价指标

为了与现有研究进行对比,采用准确率、精确度、召回率、 F_1 分数、AUC 值及混淆矩阵作为评估指标。其中,准确率是衡量分类模型正确分类图像数量占总观测值比例的经典指标,其计算公式为:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (2)$$

精确度计算公式为:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

模型召回率计算公式为:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

式中: TP (真阳性) 指真实阳性案例中被正确预测为阳性的病例数; FN (假阴性) 指真实阳性案例中被错误预测为阴性的病例数; FP (假阳性) 指真实阴性案例中被错误预测为阳性的病例数; TN (真阴性) 指真实阴性案例中被正确预测为阴性的病例数。

F_1 是结合精确度和召回率对测试正确率的度量, 用公式表示为:

$$F_1 = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (5)$$

AUC 是一种评估分类模型性能的指标, 在评估模型在多个类别之间的分类能力方面非常有用, 用公式表示为:

$$\text{AUC} = \frac{\sum_{i \in \text{正样本}} \text{rank}_i - (M(M+1))/2}{M \times N} \quad (6)$$

式中: M 表示正样本总数; N 表示负样本总数; rank_i 表示第 i 条样本的排序序号 (按概率得分从小到大排序, 对应第 rank 个位置)。

通过可视化神经网络模型的混淆矩阵, 观察模型各个类别具体分布情况。混淆矩阵的对角线表示正确分类的样本, 每一列代表预测类别为此类的概率。每一行代表数据真实所属的类别。

2 结果

为了验证改进模型的有效性, 在 BreakHis 数据集 4 种不同放大倍数下利用 Accuracy、Precision、Recall 和 F_1 四项指标做定量分析, 对比结果如表 1 所示最优数据已加粗。

表 1 与现有部分乳腺癌八分类研究结果对比

文献	放大倍数	Accuracy/%	Precision/%	Recall/%	F_1 %
Boumaraf	40×	94.49	93.81	94.78	94.15
	100×	93.27	92.94	91.59	92.23
	200×	91.29	91.18	88.28	89.47
	400×	89.56	87.97	87.97	87.77
Zaalouk	40×	97.01	96.85	96.17	96.47
	100×	95.17	95.08	94.02	94.37
	200×	91.54	90.08	90.16	89.91
	400×	90.22	90.99	89.87	89.97
Sarker	40×	95.56	96.16	95.85	95.99
	100×	94.34	94.40	94.87	94.59
	200×	86.83	89.65	81.33	84.75
	400×	93.48	92.75	92.02	91.90
本方法	40×	97.16	96.30	98.01	97.14
	100×	96.92	96.77	97.04	96.90
	200×	96.38	96.56	95.32	95.89
	400×	95.09	94.22	95.16	94.66
	Mixed	95.89	95.86	95.73	95.79

本研究提出的改进型残差卷积神经网络在四种不同放大倍率下进行八分类任务时, 分类准确率达到 94.2%~97.0%。在混合放大倍率数据集上执行相同任务时, 准确率达到 95.9%。

此外, 在基于 BreakHis 数据集的实验中, 本模型在准确率和 F_1 得分方面均优于 Boumaraf 等^[10]、Zaalouk 等^[11]和 Sarker 等^[12]此前提出的针对乳腺病理图像的多分类网络模型。这表明所设计的网络架构在处理不同放大倍率下的乳腺组织病理学图像分类问题时, 具有更高的效能和鲁棒性。

在图 2 中, 对比展示了原始与优化后的 ResNet50 模型在 100 个训练周期后的损失值和准确率变化。

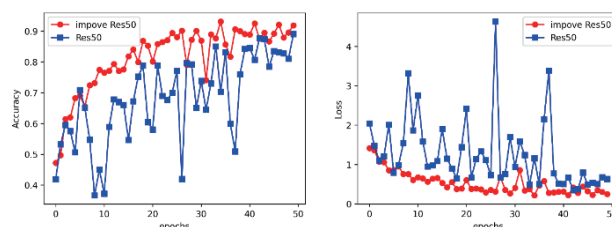


图 2 准确率及损失值变化对比图

图 3 展示了改进模型在多分类任务中的混淆矩阵, 以直观反映方法的有效性, 其中纵轴为实际标签, 横轴为预测标签。这些图表清晰地评估了模型改进策略对分类性能的提升效果。

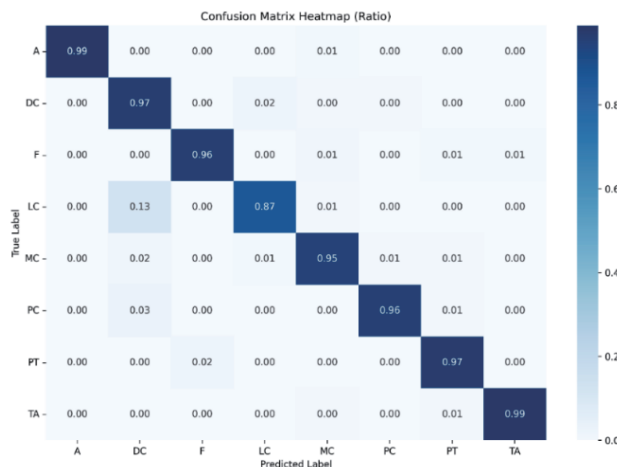


图 3 改进模型的混淆矩阵

3 结论

本研究根据原始 ResNet50 架构以及其改进版本, 对两者的参数规模与性能表现展开了对比分析, 从实验得出的结果可看出, 改进模型的参数总量由 2 352 万稍微增加至 2 407 万, 增长幅度大概是 2.3%, 更为关键的是, 在计算开销的增幅被严格控制在 3% 以内的情形下, 优化后的模型于 ImageNet 基准测试中, Top-1 准确率提高了 1.8%, 这足以说明所采用的模型改进策略是有成效的。

基于此, 本文将所提出的乳腺癌病理图像分类模型与 Boumaral、Zaalouk、Sarker 等学者的现有方法进行了性能方面的横向比较, 结果说明, 在 4 种不同放大倍率的病理图像数据集上, 本文所提方法在准确率以及 F_1 值等关键指标上, 均呈现出较为十分突出的优势。然而随着图像分辨率的提高, 所有模型的分类性能都会出现系统性的下降, 在高放大倍数时, 乳腺癌组织的细胞形态变得更加复杂, 组织结构的异质性也更为突出, 这致使同类样本在特征空间中的重叠度, 比中低倍率图像要高许多, 而这种类内特征相似性的提高, 直接对卷积神经网络捕捉细微判别特征产生了影响, 在 $400\times$ 高分辨率条件下, 各个模型的特征可分性指标相较于基准水平都有了大幅下降。

为应对上述挑战, 本研究设计改进型多尺度特征融合架构并验证其性能优势。引入 EMA 注意力模块与 RFACnv 卷积后, 虽然网络参数量增加, 但模型多尺度特征提取能力显著提升, 整体分类准确率提高 4.2%。混淆矩阵对比分析显示, 改进架构对 A、DC、F、MC、PC、PT、TA 七类病理组织判别能力出色, 但对 LC 类别的判断存在特异性不足问题: 准确率仅 $87.4\pm 2.1\%$, 且 12.3% 的 LC 样本被误归为 DC。这一误差模式表明, EMA 模块虽能借助跨尺度通道注意力提升全局特征整合能力, 但在区分 LC 与 DC 的细微核异质性特征时仍有局限。针对此问题, 后续研究可尝试融合动态核区域聚焦策略与对比特征解耦机制, 针对性优化 LC 类别的判别边界建模。

参考文献:

- [1] KABIR S M, BHUIYAN M I H, TANVEER M S, et al. RiIG modeled WCP image-based CNN architecture and feature-based approach in breast tumor classification from B-mode ultrasound[J]. Applied sciences, 2021, 11(24): 12138.
- [2] 舒予静. 基于可解释机器学习构建的乳腺癌 HER2 表达三分类预测模型研究[D]. 广州: 南方医科大学, 2024.
- [3] 秦玉华. 基于机器学习对乳腺癌超声图像数据分类[J]. 中国老年保健医学, 2023, 21(3): 140-142.
- [4] SAMMUT S J, CRISPIN-ORTUZAR M, CHIN S F, et al. Multi-omic machine learning predictor of breast cancer therapy response[J]. Nature, 2022, 601: 623-629.
- [5] XU X B, AN M J, ZHANG J D, et al. A high-precision classification method of mammary cancer based on improved DenseNet driven by an attention mechanism[J]. Computational and mathematical methods in medicine, 2022, 2022(1): 8585036.
- [6] EL AGOURI H, AZIZI M, EL ATTAR H, et al. Assessment of deep learning algorithms to predict histopathological diagnosis of breast cancer: first moroccan prospective study on a private dataset[J]. BMC research notes, 2022, 15(1): 66.
- [7] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 770-778.
- [8] ZHANG X, LIU C, YANG D G, et al. RFACnv: innovating spatial attention and standard convolutional operation[EB/OL]. (2024-03-28)[2026-01-22]. <https://doi.org/10.48550/arXiv.2304.03198>.
- [9] OU-YANG D, HE S, ZHANG G Z, et al. Efficient multi-scale attention module with cross-spatial learning[EB/OL]. (2023-06-06)[2026-01-22]. <https://doi.org/10.48550/arXiv.2305.13563>.
- [10] BOUMARAF S, LIU X B, ZHENG Z S, et al. A new transfer learning based approach to magnification dependent and independent classification of breast cancer in histopathological images[J]. Biomedical signal processing and control, 2021, 63: 102192.
- [11] ZAALOUK A M, EBRAHIM G A, MOHAMED H K, et al. A deep learning computer-aided diagnosis approach for breast cancer[J]. Bioengineering, 2022, 9(8): 391.
- [12] SARKER M M K, AKRAM F, ALSHARID M, et al. Efficient breast cancer classification network with dual squeeze and excitation in histopathological images[J]. Diagnostics, 2022, 13(1): 103.

【作者简介】

高谦(2000—), 男, 湖北黄冈人, 硕士研究生, 研究方向: 机器视觉。

(收稿日期: 2025-08-18 修回日期: 2026-03-04)