

本文引用格式: 王锦俊,蔡延光.一种改进 ResNet34 模型的乳腺图像识别方法[J].自动化与信息工程,2025,46(3):30-36.

WANG Jinjun, CAI Yanguang. An improved ResNet34 model for mammographic image recognition method[J]. Automation & Information Engineering, 2025,46(3):30-36.

一种改进 ResNet34 模型的乳腺图像识别方法*

王锦俊¹ 蔡延光^{1,2}

(1.广东工业大学自动化学院, 广东 广州 510006

2.广州理工学院人工智能学院, 广东 广州 510540)

摘要: 为了提高乳腺图像的识别准确率, 提出一种改进 ResNet34 模型的乳腺图像识别方法。该方法以 ResNet34 模型为基础, 引入平行注意力残差块 (PARB) 模块来增强乳腺图像通道间的关联性, 进一步提取乳腺图像的重要特征信息, 从而提高模型的识别准确率; 利用科尔莫戈洛夫-阿诺尔德网络 (KAN) 替代传统的多层感知器 (MLP), 减少模型参数, 提高模型的识别速度。实验结果表明, 改进的 ResNet34 模型比 ResNet34 模型的准确率、精确率、召回率和 F1-Score 分别提升了 4.0%、0.6%、8.0%和 4.7%, 表明该方法对乳腺图像具有更好的识别效果。

关键词: 乳腺图像识别; ResNet34; 平行注意力残差块; 科尔莫戈洛夫-阿诺尔德网络

中图分类号: TP391.41; TP183

文献标志码: A

文章编号: 1674-2605(2025)03-0005-07

DOI: 10.12475/aie.20250305

开放获取

An Improved ResNet34 Model for Mammographic Image Recognition Method

WANG Jinjun¹ CAI Yanguang^{1,2}

(1.College of Automation, Guangdong University of Technology, Guangzhou 510006, China

2.School of Artificial Intelligence, Guangzhou Institute of Science and Technology, Guangzhou 510540, China)

Abstract: To enhance the recognition accuracy of mammographic images, an improved ResNet34 model for mammographic image recognition method is proposed. Building upon the ResNet34 model, this method introduces a parallel attention residual block (PARB) module to strengthen inter-channel correlations in mammographic images, further extracting critical feature information to improve recognition accuracy. Additionally, it replaces the traditional multilayer perceptron (MLP) with Kolmogorov-Arnold networks (KAN) to reduce model parameters and increase recognition speed. Experimental results demonstrate that the improved ResNet34 model achieves enhancements of 4.0%, 0.6%, 8.0%, and 4.7% in accuracy, precision, recall, and F1-Score respectively compared to the original ResNet34 model, indicating superior recognition performance for mammographic images.

Keywords: mammographic image recognition; ResNet34; parallel attention residual block; Kolmogorov-Arnold networks

0 引言

乳腺癌是女性常见的恶性肿瘤之一, 其发病率逐年增加^[1]。数字乳腺 X 射线影像因具有设备成本低、检查速度快、图像分辨率高、放射剂量低等特点, 被广泛应用于临床诊断^[2]。然而, 数字乳腺 X 射线影像质量参差不齐, 不仅影响医生诊断, 还可能增加误诊风险^[3]。随着深度学习技术在医学图像分析领域的广

泛应用, 如何快速、准确地识别不完整乳腺图像已成为研究热点, 其不仅为医生诊断提供了重要参考, 还推动了医学图像技术的发展。

近年来, 利用深度学习技术识别不完整图像受到学术界的广泛关注。文献[4]提出一种基于场效应的双线性深度网络, 用于评估不完整图像中的缺失特征; 但模型参数较多, 训练时间较长, 不适用于实时性要

求高的场景。文献[5]基于低秩矩阵的恢复思想,提出一种不完整人脸图像识别深度学习算法,通过矩阵恢复不完整的人脸图像;但当图像损坏较严重时,识别准确率有所下降。文献[6]提出一种基于自注意力机制的不完整图像精细化识别模型,通过互增强操作,利用生成对抗网络特征来加强不完整图像的特征信息,实现不完整图像的快速识别;但模型计算较复杂,需要消耗大量的资源,导致硬件成本增加。文献[7]提出一种基于残差网络和挤压激励层的深度学习分类网络模型,可定位并识别不完整的头颈部磁共振图像;但其识别准确率仅有 78.8%。文献[8]提出一种基于分块统计的模板匹配算法,用于对完整和不完整目标的准确定位;但无法有效识别经过缩放或旋转的图像。

为了快速、准确地识别不完整乳腺图像,以提高医学图像评估的准确率^[9],本文提出一种改进 ResNet-34 模型的乳腺图像识别方法。该方法在 ResNet34 模型^[10]的基础上,引入了平行注意力残差块(parallel attention residual block, PARB)模块,并利用科尔莫戈洛夫-阿诺尔德网络(Kolmogorov-Arnold networks, KAN)替代传统的多层感知器(multilayer perceptron, MLP),以提高乳腺图像的识别准确率,降低医生误诊率,提升工作效率。

1 相关内容

1.1 压缩和激励模块

压缩和激励(squeeze and excitation, SE)模块^[11]是一种通道注意力模块,能对输入特征图进行通道特征加强,以提高卷积神经网络的表征能力,其结构如图 1 所示。

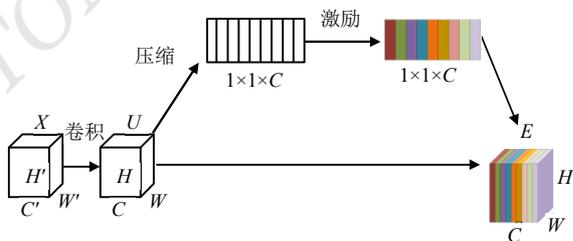


图 1 SE 模块结构图

SE 模块的操作流程如下:

- 1) 对输入特征图 X 进行卷积操作,产生新的特征图 U ;
- 2) 对新的特征图 U 进行压缩操作,提取通道间的全局特征信息,生成一个通道权重向量;
- 3) 对通道权重向量进行激励操作,确定每个通道的相对重要性;
- 4) 利用通道权重来调整新的特征图 U ,通过元素乘法得到 SE 模块的输出 E 。

1.2 S2 注意力机制模块

S2 注意力机制是一种改进的注意力机制^[12],主要用于增强序列建模中不同位置之间的关联性^[13-15]。其通过优化注意力计算方式,能够更有效地捕捉序列中远距离的依赖关系,从而提升模型对序列数据的建模能力。S2 注意力机制模块通过层次化的金字塔结构特性,可捕获更精细的视觉特征,提高模型的识别精度。S2 注意力机制模块的结构如图 2 所示。

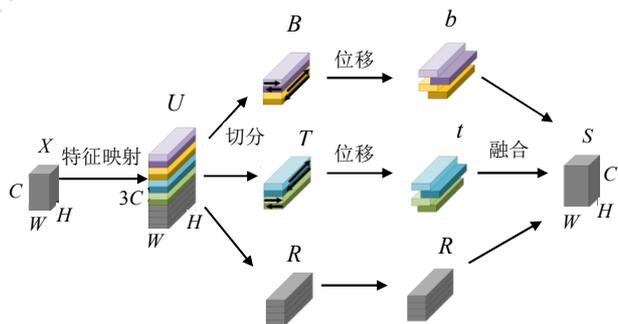


图 2 S2 注意力机制模块结构图

S2 注意力机制模块的操作流程如下:

- 1) 对输入特征图 X 进行特征映射,产生新的特征图 U ;
- 2) 展开新的特征图 U ,并将其切分成 3 个部分(特征图 B 、特征图 T 和特征图 R);
- 3) 将特征图 B 、 T 分别进行不同的空间位移操作,得到特征图 b 、 t ; 特征图 R 保持不变;
- 4) 融合特征图 b 、 t 、 R ,得到 S2 注意力机制模块的输出 S 。

1.3 PARB 模块

PARB 模块融合了 SE 模块和 S2 注意力机制模块的特点，加强了不同位置间的关联性，可进一步提取通道间的重要特征信息，其结构如图 3 所示。

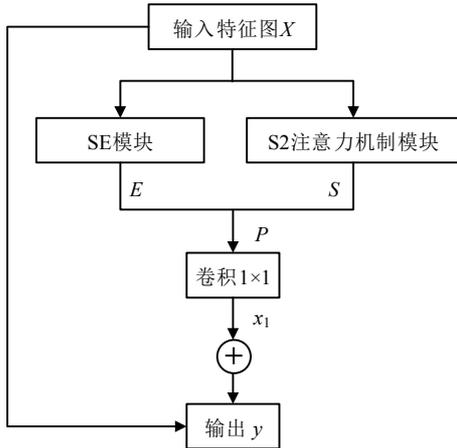


图 3 PARB 模块结构图

PARB 模块的操作流程如下：

1) 利用 SE 模块和 S2 注意力机制模块分别对输入特征图 X 进行特征权重提取，其中，SE 模块通过卷积、SE 操作获取通道的重要特征信息，生成 SE 模块特征 E ；S2 注意力机制模块通过对输入特征图 X 进行特征映射、切分、位移和融合操作，加强对输入特征图不同位置的关联性，生成 S2 注意力特征 S ；

2) 将 SE 模块特征 E 与 S2 注意力特征 S 相加，得到新的特征 P ；

3) 对 P 进行 1×1 卷积操作，生成卷积特征 x_1 ，即为调整输入特征图的输入通道数；

4) 将卷积特征 x_1 与输入特征图 X 进行残差相加，生成 PARB 模块输出 y 。

1.4 KAN

在深度学习中，传统的 MLP 存在参数众多、调参过程复杂、学习效率较低等问题。本文利用 KAN^[16] 替代 MLP^[17]。KAN 结合了 MLP 和样条曲线的优点，通过 Kolmogorov-Arnold 表示定理，将输入的高维函数转换为低维函数进行分析。KAN 结构如图 4 所示。

KAN 的操作流程如下：

1) 将前层（如平均池化层）的输出图像特征作为 KAN 的输入；

2) 利用多个非线性激活函数，对输入的每个特征维度进行处理，生成多组基函数结果；

3) 通过可学习的双线性权重，将第一层激活后的基函数结果两两加权组合，生成中间交互特征；

4) 先对中间交互特征应用激活函数，再经过线性层加权求和，得到 KAN 的输出结果。

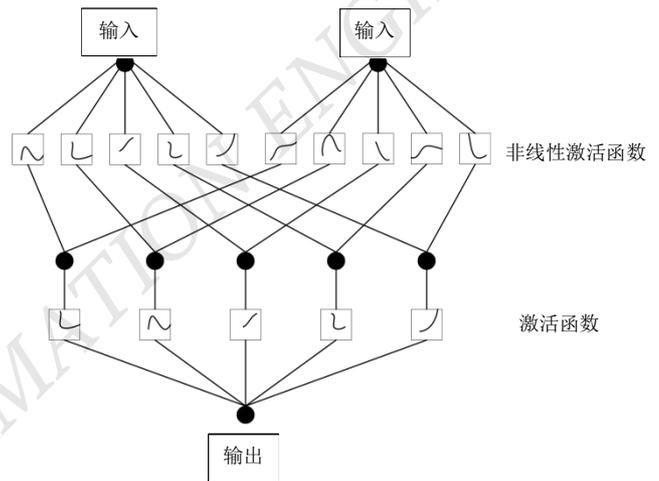


图 4 KAN 结构图

1.5 改进的 ResNet34 模型

本文以 ResNet34 模型为基础，在卷积层后引入 PARB 模块，以增强乳腺图像通道间的特征联系，进一步获取乳腺图像的重要特征信息，提高了模型的识别准确率；利用 KAN 替代 MLP，减少模型参数，提升模型的识别速度。改进的 ResNet34 模型结构如图 5 所示。

改进的 ResNet34 模型操作流程如下：

1) 利用 1 个 7×7 的卷积层 1 初步提取乳腺图像特征；

2) 通过批量归一化层和激活函数解决模型计算过程中的梯度消失和梯度爆炸问题；

3) 通过最大池化层进一步提取乳腺图像特征，并减少计算量；

- 4) 经过卷积层 2, 提取乳腺图像的深层特征;
- 5) 利用 PARB 模块, 加强乳腺图像的深层特征, 并进一步增强乳腺图像不同通道间的联系;
- 6) 重复步骤 4)、5) 3 次;

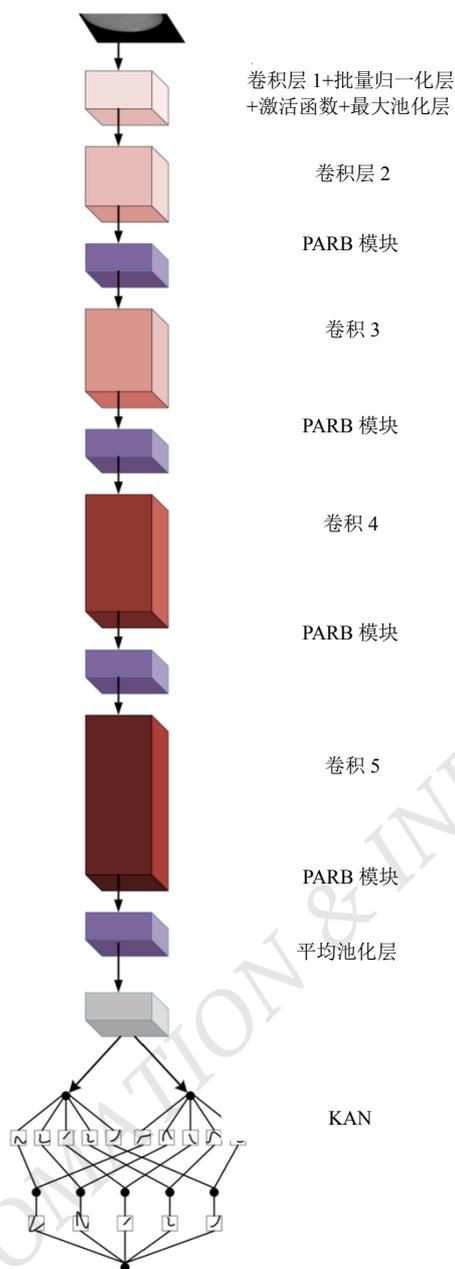


图 5 改进的 ResNet34 模型结构图

- 7) 通过平均池化层降低网络层数、防止过拟合, 并增强乳腺图像的局部特征信息;
- 8) 利用 KAN 对乳腺图像的深层特征进行双线性交互与轻量化聚合, 提升特征判别能力, 减少模型参数, 并输出乳腺图像的识别结果。

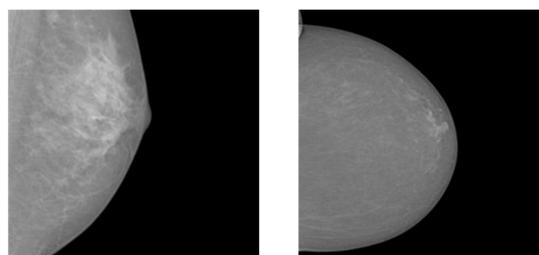
2 实验

2.1 实验环境

本文方法采用 PyTorch 框架实现, 编程语言为 Python, 处理器为 14 vCPU Intel(R) Xeon(R) Gold 6330 CPU@2.00 GHz, 内存为 60 GB, 显卡为 RTX 3090。

2.2 实验数据集

以乳腺 X 射线筛查数字数据库(digital database for screening mammography, DDSM)^[18]为基础制作实验数据集。首先, 从 DDSM 中选取 325 幅乳腺图像; 然后, 通过数据增强方法(旋转、对称变换、平移等)对乳腺图像进行数据扩充; 接着, 依据乳腺 X 射线图像质量标准^[19]裁剪符合实验要求的乳腺图像, 并请专业医生对处理后的乳腺图像进行主观评价, 评价结果包括 545 幅完整乳腺图像(乳腺部位无缺陷, 轮廓清晰完整, 满足诊断要求)和 508 幅不完整乳腺图像(乳腺轮廓不完整, 不满足诊断要求), 如图 6 所示; 最后, 为减少实验数据不均衡带来的误差, 随机选取完整和不完整乳腺图像各 500 幅组成实验数据集。将实验数据集的 1 000 幅乳腺图像按 8:2 的比例划分为训练集和测试集。



(a) 不完整的乳腺图像

(b) 完整的乳腺图像

图 6 不同质量的乳腺图像

2.3 参数设置

为提高改进的 ResNet34 模型性能, 经反复训练, 最终确定本文实验的参数如表 1 所示。

表 1 实验参数设置

参数名称	设置
迭代次数/次	50
批次大小	64
学习率	3×10^{-4}
优化器	Adam
损失函数	交叉熵

2.4 评价指标

本文采用准确率(accuracy)、精确率(precision)、召回率(recall)、F1-score 为模型的评价指标。在混淆矩阵^[20]中, TP 表示将正样本预测为正样本的图像数量, FP 表示将负样本预测为正样本的图像数量, FN 表示将正样本预测为负样本的图像数量, TN 表示将负样本预测为负样本的图像数量。

2.4.1 准确率

准确率是指被正确分类的图像数量占总样本图像数量的比例, 计算公式为

$$R_{\text{accuracy}} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

2.4.2 精确率

精确率是指在预测为正样本的图像数量中, 被正确分类的图像数量比例, 计算公式为

$$R_{\text{precision}} = \frac{TP}{TP + FP} \quad (2)$$

2.4.3 召回率

召回率是指在真实为正样本的图像数量中, 被正确分类的图像数量比例, 计算公式为

$$R_{\text{recall}} = \frac{TP}{TP + FN} \quad (3)$$

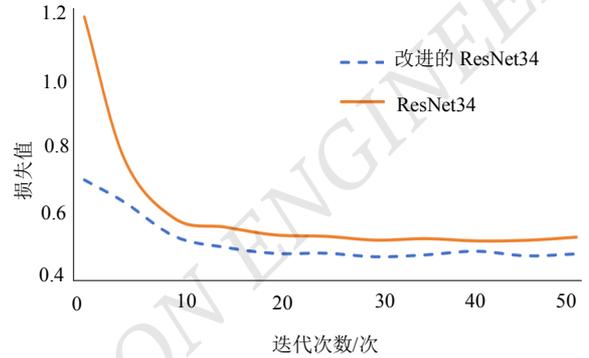
2.4.4 F1-Score

F1-Score 是指精确率和召回率的调和均值, 计算公式为

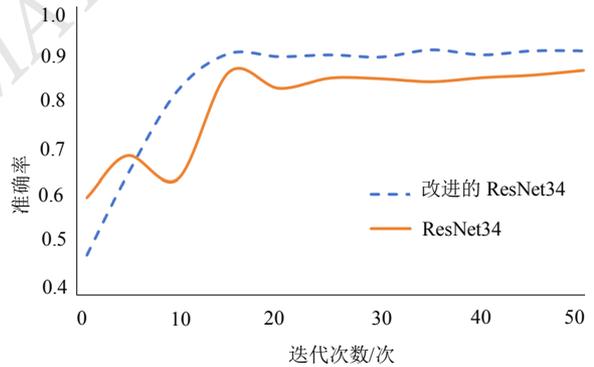
$$F1 = \frac{2TP}{FP + 2TP + FN} \quad (4)$$

2.5 实验结果与分析

为验证改进的 ResNet34 模型的有效性, 将其与 ResNet34 模型进行对比实验, 其损失曲线、准确率曲线如图 7 所示。



(a) 损失曲线图



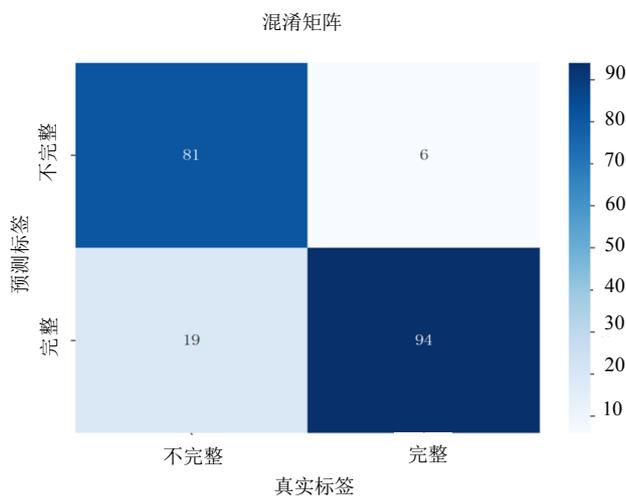
(b) 准确率曲线图

图 7 ResNet34 和改进的 ResNet34 模型曲线对比图

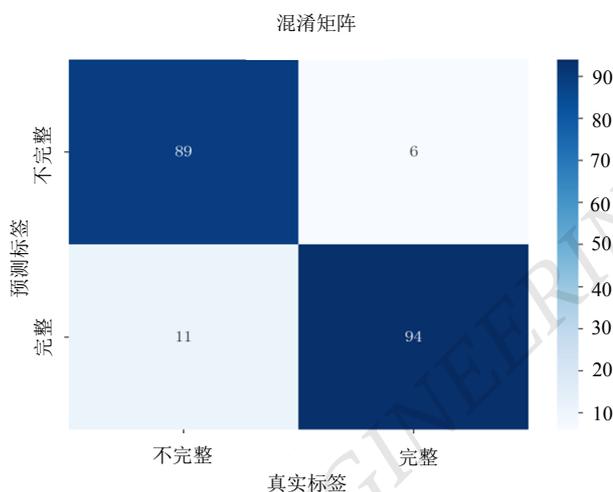
由图 7 可知, 虽然 ResNet34 和改进的 ResNet34 模型均趋于收敛状态, 但改进的 ResNet34 模型的准确率明显高于 ResNet34 模型, 说明改进方法有效地提高了模型的识别能力。

2.6 消融实验

为了直观地了解各模块对改进的 ResNet34 模型性能的影响, 将 ResNet34、ResNet34+PARB、ResNet34+KAN 和改进的 ResNet34 模型进行消融实验。根据实验最优模型结果得到的混淆矩阵如图 8 所示。



(a) ResNet34 模型



(d) 改进的 ResNet34 模型

图 8 混淆矩阵

根据混淆矩阵计算得到的评价指标如表 2 所示。

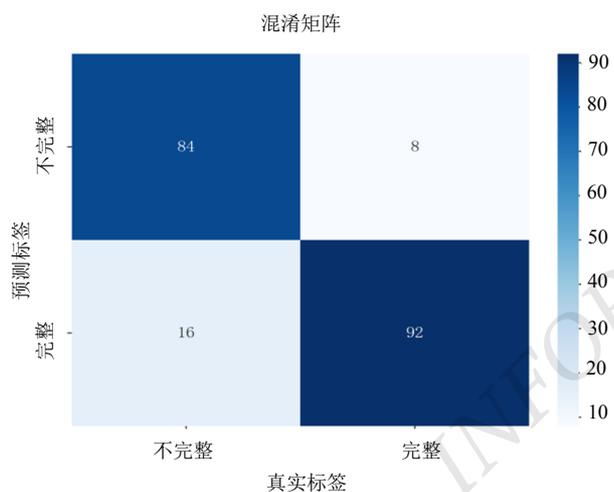
表 2 评价指标

模型	准确率/ %	精确率/ %	召回率/ %	F1-Score/%
ResNet34	87.5	93.1	81	86.6
ResNet34+PARB	88.0	91.3	84	87.4
ResNet34+KAN	89.5	92.5	86	88.9
改进的 ResNet34	91.5	93.7	89	91.3

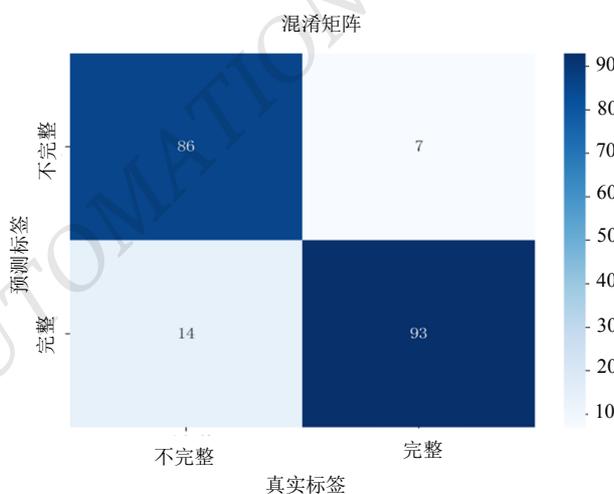
由表 2 可知，改进的 ResNet34 模型比 ResNet34 模型的准确率、精确率、召回率和 F1-Score 分别提升了 4.0%、0.6%、8.0%和 4.7%，表明改进的 ResNet34 模型具有更好的识别效果。

3 结论

为了快速、准确地识别不完整的乳腺图像，提高医生的工作效率，本文提出一种改进 ResNet34 模型的乳腺图像识别方法。该方法通过引入 PARB 模块，获取通道间的重要特征信息，提高了模型的识别性能；利用 KAN 替代 MLP，减少了模型参数，提高模型的识别速度。通过实验结果可知，改进的 ResNet34 模型具有更好的识别效果，其准确率、精确率、召回率和 F1-Score 均有提升。在未来的研究中，将进一步探究更高效、更准确的图像识别技术，使其能够更准确地识别乳腺图像。



(b) ResNet34+PARB 模型



(c) ResNet34+KAN 模型

参考文献

- [1] 刘佩芳,鲍润贤.乳腺 X 线检查用于乳腺癌筛查有效性的争论[J].中华放射学杂志,2014,48(10):797-799.
- [2] AKIN O, BRENNAN S B, DERSHAW D D, et al. Advances in oncologic imaging: Update on 5 common cancers[J]. CA: A Cancer Journal for Clinicians, 2012,62(6):364-393.
- [3] 梁永刚,付丽媛,钟群,等.全数字化乳腺 X 射线摄影系统质量控制检测方法探讨[J].医疗卫生装备,2020,41(5):61-64.
- [4] ZHONG S H, LIU Y, HUA K A. Field effect deep networks for image recognition with incomplete data[J]. ACM Transactions on Multimedia Computing, Communications, and Applications, 2016,12(4):1-22.
- [5] ZHAO J, LV Y, ZHOU Z, et al. A novel deep learning algorithm for incomplete face recognition: Low-rank-recovery network[J]. Neural Networks, 2017,94:115-124.
- [6] 孙浩强.面向复杂场景的不完整车辆图像精细化识别研究[D].合肥:安徽大学,2022.
- [7] 易音巧.基于深度学习的医学影像质量评估[D].上海:华东师范大学,2021.
- [8] 邓泽峰,熊有伦,黄小鹏.适应不完整目标的快速模板匹配[J].光电工程,2010,37(5):7-11.
- [9] MÅNSSON L G. Methods for the evaluation of image quality: A review[J]. Radiation Protection Dosimetry, 2000,90(1-2):89-99.
- [10] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:770-778.
- [11] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018:7132-7141.
- [12] YU T, LI X, CAI Y, et al. S²-MLPv2: Improved spatial-shift MLP architecture for vision[J]. arXiv Preprint arXiv:2108.01072, 2021.
- [13] 崔海朋,姜英昌.基于注意力机制的海上小目标重识别方法[J].机电工程技术, 2022,51(7):100-103.
- [14] SHAW P, USZKOREIT J, VASWANI A. Self-attention with relative position representations[J]. arXiv Preprint arXiv:1803.02155, 2018.
- [15] 黄辉,吴建强,肖豪,等.基于注意力机制的接线端子文本检测与识别[J].机电工程技术,2023,52(6):202-206.
- [16] LIU Z, WANG Y, VAIDYA S, et al. KAN: Kolmogorov-arnold networks[J]. arXiv Preprint arXiv:2404.19756, 2024.
- [17] TAUD H, MAS J F. Multilayer perceptron (MLP)[M]. Geomatic Approaches for Modeling Land Change Scenarios, 2018: 451-455.
- [18] LEE R S, GIMENEZ F, HOOGI A, et al. A curated mammography data set for use in computer-aided detection and diagnosis research[J]. Scientific Data, 2017,4(1):1-9.
- [19] 陈燕.全视野数字化乳腺摄影质量控制与影像质量评价分析[J].青海医药杂志,2019,49(4):65-66.
- [20] TOWNSEND J T. Theoretical analysis of an alphabetic confusion matrix[J]. Perception & Psychophysics, 1971,9(1): 40-50.

作者简介:

王锦俊,男,1999年生,在读硕士研究生,主要研究方向:控制与优化。E-mail: wangjinjun320@163.com

蔡延光,男,1963年生,博士研究生,教授,主要研究方向:网络控制与优化、组合优化、智能优化、智能交通系统等。E-mail: caiyg99@163.com