

基于改进 GoogLeNet 微核细胞图像分类

刘 勃¹ 何秉蔚¹ 蒋一波¹ 陈 坤¹ 魏凯斌¹ 赵玉祥¹

LIU Qing HE Bingwei JIANG Yibo CHEN Kun WEI Kaibin ZHAO Yuxiang

摘 要

微核的出现与染色体损伤或环境污染有关,为解决传统检测方法效率低且易受人为因素的影响,文章提出了一种基于改进 GoogLeNet 的微核图像分类深度学习方法,首先进行数据增强来减少类别不平衡过拟合的现象,其次引用迁移学习的方式,最后引入注意力机制和可变形卷积网络,增强模型对微核细胞图像中关键区域的关注能力。通过对比其他深度学习卷积网络,改进后的 GoogLeNet 在微核细胞图像分类任务中表现优异,分类平均准确率达到 96.96%,优于传统方法。

关键词

微核;卷积网络;注意力机制;可变形卷积

doi: 10.3969/j.issn.1672-9528.2025.08.017

0 引言

微核(micronucleus, MN)是一种细胞内未被正常分配的染色体或染色体碎片,通常是在细胞分裂过程中出现异常时产生的。微核的出现是细胞基因突变、染色体损伤、染色体不稳定性等生物学异常的重要标志^[1]。由于其与基因突变、癌症以及环境毒素有关,微核的检测在癌症早期筛查、遗传病研究和环境污染监测等领域具有重要意义。传统的微核检测^[2]方法多依赖显微镜下的人工观察,这种方法不仅耗时长,而且容易受到人为因素的影响,尤其是在大规模筛查时,效率和准确性面临着巨大的挑战。

随着计算机视觉和深度学习技术的发展,越来越多的研究开始探索如何利用自动化技术进行微核细胞图像的分类与检测。卷积神经网络(CNN)作为一种强大的深度学习方法,已经在图像分类领域取得了突破性进展^[3]。Alaif 等人^[4]通过在 GoogLeNet 的基础上引用迁移学习,并在自主策划的细胞核和微核图像数据集,达到了 80.06%的平均准确率。陶洪^[5]提出了融合局部注意力机制和 AlexNet 的微核细胞分类模型,达到了 93.2%的准确率。李竞宇^[6]融合多层次感知注意力机制提出了 MoblieVit-MN 轻量级网络,其准确率达到 93.3%。Wei 等人^[7]利用 AlexNet 和视觉注意力改进了一种识别微核的计算机辅助诊断方法,其以较少的参数取得了优异的性能。

然而,受限於微核自身的特殊形态特性:尺寸微小、形态不规则、染色对比度低,以及显微图像中复杂的背景干扰,现有卷积神经网络在微核分类任务中仍面临显著挑战^[8]。当前主流 CNN 模型(如 VGG、ResNet)多基于固定几何结构的卷积核,难以适应微核因制备差异或成像条件导致的几何形变,如边缘扭曲、染色质分布不均^[9]。此外,微核与背景的弱对比性要求模型具备精准的空间注意力聚焦能力,而传统通道注意力机制在复杂噪声干扰下易出现误激活^[10]。同时,医学图像标注成本高昂,小样本学习下的模型泛化性还待优化。

针对以上问题,本文提出一种基于改进 GoogLeNet 的微核图像分类网络,通过多阶段特征优化机制实现高效形变感知与注意力引导的协同建模。首先,在 Inception 模块中引入轻量化可变形卷积(DCNv4),利用其动态偏移预测能力捕捉微核的不规则边缘与局部形变特征;其次,融合卷积注意力模块(CBAM),通过通道、空间双维度加权抑制背景噪声,增强对微核核心区域的聚焦能力;进一步结合迁移学习与数据增强策略,缓解小样本场景下的过拟合风险。旨在提升模型在微核细胞图像分类中的准确性与鲁棒性。通过实验验证了改进后的 GoogLeNet 在微核细胞图像分类任务中的优越性能,并与传统方法及原始 GoogLeNet 进行了对比。

1 GoogLeNet 网络

1.1 GoogLeNet 网络结构

GoogLeNet 是由 Google 团队于 2014 年提出的深度卷积神经网络,GoogLeNet 的一个核心特点是其 Inception 模块,这种模块通过并行使用不同尺寸的卷积和池化操作来提取图像的多尺度特征,使得网络能够高效地处理复杂和多样化的

1. 天水师范大学电子信息与电气工程学院 甘肃天水 741001
[基金项目] 天水师范大学研究生创新项目(TCXM2531);
国家自然科学基金项目(62363031);甘肃省科技计划项目
(23JRRE0740 25JRRE004);天水市秦州区科技计划项目
(2023-SHFZG-6497)

图像信息。Inception 模块通过多分支路径对输入特征进行处理，并在通道维度上将输出结果进行串联合并，作为下一层的输入。多个 Inception 模块可逐层堆叠构成深层网络，从而构建出一种稀疏连接的结构，模拟了生物视觉皮层中的信息处理方式。这种结构不仅提高了网络的非线性表达能力，而且有效模拟了传统卷积神经网络的最优拓扑结构。为了进一步提高效率并控制参数量，GoogLeNet 在每个 Inception 模块中引入了 1×1 卷积操作，用作特征通道的降维工具。通过在较高维度的特征图输入前施加 1×1 卷积，网络可以显著减少计算复杂度，降低内存开销，同时避免信息丢失。此外，这种结构设计赋予网络一定程度上的“自动结构选择能力”，即网络在每一层通过不同尺寸的卷积与池化操作，自动捕捉适合当前输入特征的空间尺度，而非由人工预先设定卷积核大小或结构路径。这不仅简化了网络设计的工程复杂度，也提升了模型在实际应用中的适应性和可扩展性。因此，GoogLeNet 在保持较高准确率的同时，成功控制了模型复杂度和参数量。GoogLeNet 的总体结构图如图 1 所示。

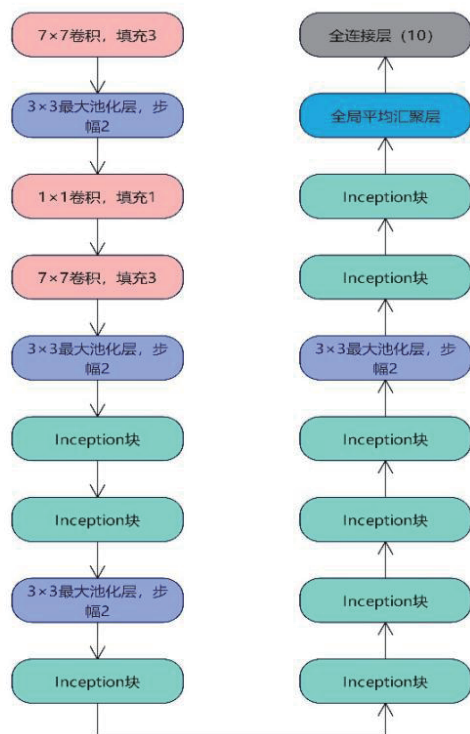


图 1 GoogLeNet 网络结构

1.2 Inception 网络结构

Inception 模块包含 4 条并行路径，各自处理输入数据。如图 2 所示，第一路径直接应用 1×1 卷积；第二、三路径先用 1×1 卷积减少通道数，随后分别从 3×3 和 5×5 的卷积层提取不同尺度特征；第四路径先进行 3×3 最大池化操作，再通过 1×1 卷积调整通道维度。所有路径都采用适当填充保持输出与输入在高宽维度上一致，最终将 4 条路径的输出在通道

维度上合并，形成 Inception 模块的完整输出。这种设计的主要可调节超参数是各层的输出通道数，通过调整这些参数可以控制模型复杂度。

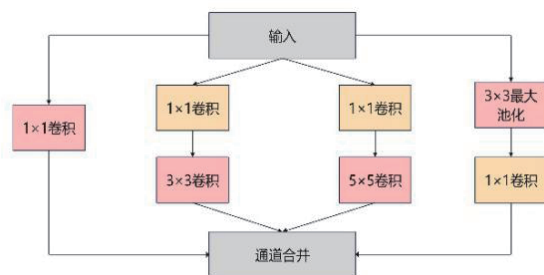


图 2 Inception 网络结构

1.3 全局平均池化层

Lin 等人^[10]提出全局平均池化（global average pooling, GAP）的策略，GAP 的核心思想是将每个卷积层输出的特征图通过对每个通道上的所有像素值进行平均，转换成单一数值，最终将每个通道的平均值汇聚成一个一维的向量。全局平均池化的整个过程没有任何参数需要学习，它将空间信息总结出来，强调了特征图与类别之间的对应关系。相比传统的全连接层，GAP 不需要显式地引入大量的参数，避免了模型在训练时容易发生拟合的情况，尤其适用于小样本数据集，此外，GAP 通过捕捉图像的全局信息而非局部信息，能够更好地对整个图像进行特征总结，从而提高模型的泛化能力。CNN 全局平均池化如图 3 所示。GoogLeNet 使用 GAP 代替全连接层，在使网络更加轻量化的同时避免了过拟合的发生。

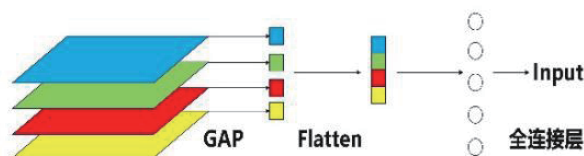


图 3 全局平均池化层

2 改进的 GoogLeNet

2.1 混合注意力机制

卷积注意力模块（convolutional block attention module, CBAM）是混合注意力机制的代表模型^[12]。图 4 呈现的 CBAM 架构采用级联方式组合这两种注意力：通道注意力部分计算各特征通道的权重系数，突出重要通道信息；空间注意力部分则对不同位置分配权重，使模型聚焦于关键区域。这种双重注意力机制能够有效增强特征图中重要信息的表达，动态调整特征的权重分布，从而显著提升模型对微核等细微特征的识别能力。

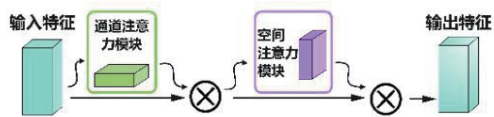


图4 CBAM结构

2.2 可变形卷积模块

可变形卷积网络（deformable convolutional networks version, DCNv4）是基于前几代卷积的进一步优化版本^[11]，目的是利用高效的动态形变建模能力，提升模型对复杂几何形变目标的适应性。相较于DCNv3而言^[12]，DCNv4取消了Softmax归一化操作，DCNv4的每个卷积核都能自适应地调整其聚合权重，对每个像素的响应不再受到固定范围的限制，使得偏移量的生成更加精细和高效，从而更好地适应各种目标的形状和大小。在计算效率方面，DCNv4通过改进的算法和实现方式，DCNv4能够在保持灵活性的同时，减少了计算量和内存消耗，提高处理速度。这使得它能够处理更大规模的输入数据或更复杂的任务。

2.3 改进的GoogLeNet网络

GoogLeNet的核心构造为Inception模块，该模块通过并行整合多种尺度的卷积核与池化操作实现多尺度特征提取。本研究基于此架构进行优化改进，其改进后的网络拓扑结构如图5所示。

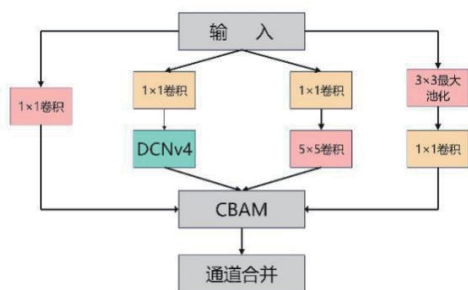


图5 改进后的Inception结构

网络通过添加DCNv4并融合CBAM模块来提高特征提取和分类的准确性。具体操作为：将DCNv4模块集成到GoogLeNet的Inception模块中，替换3x3卷积为DCNv4，同时在各分支后引入CBAM注意力机制增强特征表达。此外，采用LeakyReLU激活函数取代传统ReLU函数，有效缓解神经元失活现象。DCNv4通过动态调整卷积核的采样点来适应几何形变，而CBAM通过通道和空间注意力机制增强重要特征。通过两个模块的协同工作，来解决微核细胞图像低对比度、小目标和复杂背景问题。

本文还通过采取迁移学习的方式^[13]，将从ImageNet数据集上训练所得到的权重应用于微核图像分类当中。通过使

用预训练模型，不仅提高了训练效率，减少时间和资源成本，还减轻了过拟合的发生，增强模型的泛化能力。

3 实验结果与分析

3.1 数据集

本文所使用的细胞数据来自某省医院。对所收集的图片进行人工整理分类得到一个微核细胞图像分类的数据集，部分样本示例如图6所示。

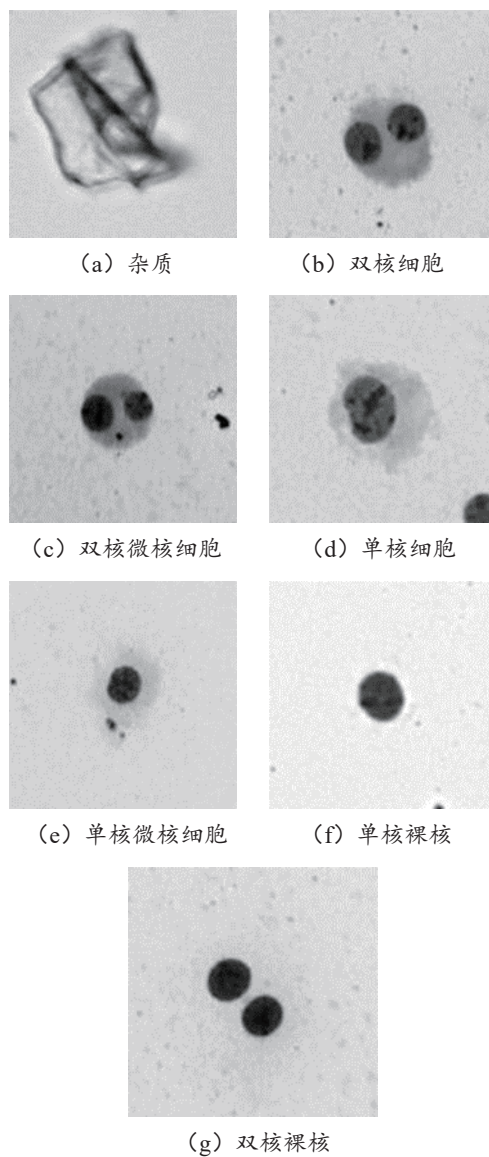


图6 部分样本数据集展示

由于微核细胞数据不足以及与别的类型数量高度不平衡问题，需要对数据集进行增强来避免发生过拟合，因此本文采用随机水平翻转、随机垂直翻转，调整亮度将数据集扩充来达到数据间的平衡。该数据集共有6303张图像，分为7种不同类别的微核细胞图像，其次将所有微核细胞图像9:1的比例随机划分为训练集和测试集，如表1所示。再将所有图像的大小统一调整为224 px×224 px。

表 1 微核细胞图像分类及数量

类别名称	训练样本数 / 张	测试样本数 / 张
杂质	3 774	419
双核细胞	342	38
双核微核	387	43
单核细胞	351	39
单核微核	252	28
单核裸核	288	32
双核裸核	279	31

3.2 参数设置及评价指标

网络使用深度学习框架 PyTorch 部署并在 ImageNet 上进行预训练。采用 focal loss 损失函数环节不完美数据集带来的影响。使用带有动量的随机梯度下降法对网络训练 60 轮次。为了使网络在训练前期迅速收敛以及在后期更好地得到最优解，使用学习率衰减策略，初始学习率设为 0.01，每隔 20 轮学习率变为原来的 1/10。batch_size 设置为 32。本文的实验环境为 Windows 系统，处理器为 Intel(R) Core(TM) i7-13500H CPU，GPU 为 NVDIAGeForce RTX4060，软件环境为 CUDA11.1，深度学习框架为 PyTorch1.8.1。

鉴于微核细胞图像样本间存在数量不均衡现象，本文采用平均准确率（AA）作为评估模型性能的指标，它衡量的是模型在所有类别（ K 类）上的整体预测表现，定义为：

$$AA = \frac{1}{K} \sum_{i=1}^K \left(\frac{Y_i}{N_i} \right) \quad (1)$$

式中： N_i 代表第 i 类微核图像的总样本量； Y_i 代表分类正确的样本数量。

在类别不均衡的情况下，平均准确率相较于总体准确率（OA）能够提供一个对整体效果的直观评估，其值越高，表明模型的分类效果越佳。

3.3 微核细胞图像对比实验

本文通过选择一系列主流卷积网络，包括 VGG16、AlexNet、ResNet50、GoogLeNet 以及 MoblieNet。这些网络都采用迁移学习的方式，加载预训练权重，并在相同的实验条件下进行训练并测试。对比结果如表 2 所示。

表 2 在不同网络下微核细胞图像数据集对比实验

网络模型	平均准确率 / %
MobileNet ^[14]	80.12
ResNet50 ^[15]	87.45
VGG16 ^[16]	89.92
AlexNet ^[17]	90.16
GoogLeNet ^[18]	92.12
改进的 GoogLeNet	96.96

从表 2 中可知，在所有对比网络中，MobileNet 最为轻量，拥有最少的参数。同时从实验结果中可以观察到，在微核细胞分类任务中取得了最差的性能，并且准确率都远低于其他 5 个网络。说明其有限的参数无法支撑起强大的学习能力。相比于其他网络，而 VGG-16 和 ResNet50 拥有巨量的参数和复杂的计算复杂度。

但从实验结果看，更大的体量和更高级的训练策略并没有带来性能上的明显提升。这种现象说明，复杂和高级的网络并不是在任何场景都拥有优越的性能，这种结构复杂性和参数规模不仅显著增加了模型训练和推理的计算成本，而且在小样本场景下容易引发过拟合问题，导致模型在验证集和测试集上的泛化性能下降。在微核细胞图像分类任务中，繁杂的结构不仅会增加时间复杂度，还会增加网络过拟合的风险。本研究使用的图像数据为细胞显微图像，其视觉内容相对简单，主要包括细胞核、微核、胞质等关键区域，以及大量无显著诊断信息的背景区域。因此，该任务对模型的表达能力需求相对有限，不需要极其复杂的网络结构。

在此背景下，GoogLeNet 由于其较小的模型规模、高效的 Inception 结构设计以及良好的特征抽取能力，能够以更少的计算资源实现较高的分类性能。实验也验证了其在该医学图像分类任务中的有效性，体现出轻量化深度网络在特定场景下的实用价值。而卷积注意力模块和可变形卷积模块的运用使得网络获得了更为丰富的特征表示，提高了对于细胞微核检测的敏感性，达到了比其他对比网络更好的分类效果。

3.4 消融实验

本文改进的网络主要包含两个模块：DCNv4 可变卷积模块和 CBAM 注意力机制。为了验证改进后模型的性能，消融实验共分为四组。第一组为仅使用预训练后的 GoogLeNet 网络；第二组为单独添加 DCNv4 的 GoogLeNet 网络；第三组为单独添加 CBAM 的 GoogLeNet 网络；第四组则为同时添加 DCNv4 和 CBAM 的网络，对比结果如表 3 所示。

表 3 不同模块在微核细胞图像数据集消融实验

网络模型	平均准确率 / %
基准 GoogLeNet	92.12
GoogLeNet+DCNv4	93.23
GoogLeNet+CBAM	94.41
本文改进的模型	96.96

实验表明，融合 DCNv4 模块，较基线提升 1.11% 准确率，表明可变形卷积有效捕捉微核的不规则边缘特征。融合 CBAM 模块，准确率提升 2.29%，证明注意力机制对背景干扰的滤除作用显著。同时加入两个模块，准确率得到提升，说明形变感知与注意力聚焦的互补性。

4 总结

本文针对微核细胞图像分类任务中存在的形态多样性、低对比度及复杂背景干扰等挑战,提出了一种基于改进 GoogLeNet 的深度学习模型。通过将轻量化可变形卷积(DCNv4)与通道-空间注意力机制(CBAM)深度融合,构建了形变感知与特征增强协同优化的双通路架构,显著提升了分类精度与鲁棒性。DCNv4 模块通过稀疏偏移量预测动态捕捉细胞边缘形变特征,CBAM 模块则通过通道-空间注意力加权抑制背景噪声,两者级联使模型在自建数据集上达到了较高的准确率。同时,还使用迁移学习和数据增强策略,不仅加速了训练过程,还避免了过拟合的发生。实验结果表明,该模型相较于其他主流卷积网络,平均准确率明显提升,适用于临床边缘设备部署。未来的工作中,将扩展至 3D 影像分析与多模态数据融合,推动自动化微核检测在遗传毒性评估和精准医疗的应用。

参考文献:

- [1] SOMMER S, BURACZEWSKA I, KRUSZEWSKI M. Micronucleus assay: the state of art, and future directions[J]. International journal of molecular sciences, 2020, 21(4):1534.
- [2] TOOSSI M T B, AZIMIAN H, SARRAFZADEH O, et al. Automatic detection of micronuclei by cell microscopic image processing[J]. Mutation research/fundamental and molecular mechanisms of mutagenesis, 2017, 806:9-18.
- [3] WOO S, PARK J, LEE J-Y, et al. CBAM: convolutional block attention module[C]//Computer Vision – ECCV 2018. Berlin: Springer, 2018: 3-19.
- [4] ALAFIF T, QARI S, ALBASSAM A, et al. Deep transfer learning for nucleus and micronucleus recognition[C]//2020 First International Conference of Smart Systems and Emerging Technologies (SMARTTECH). Piscataway: IEEE, 2020: 21-27.
- [5] 陶洪. 基于全局和局部注意力的细胞微核图像识别研究[D]. 兰州: 西北师范大学, 2023
- [6] 李竞宇. 基于深度学习的细胞微核图像识别[D]. 兰州: 西北师范大学, 2024.
- [7] WEI W Y, TAO H, CHEN W X, et al. Automatic recognition of micronucleus by combining attention mechanism and alexnet[J]. BMC medical informatics and decision making, 2022, 22(1): 138.
- [8] XING F Y, XIE Y P, SU H, et al. Deep learning in microscopic image analysis: a survey[J]. IEEE transactions on neural networks and learning systems, 2017, 29(10):4550-4568.
- [9] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional networks[C/OL]// 2017 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2017 [2024-05-23]. <https://ieeexplore.ieee.org/document/8237351>. DOI:10.1109/ICCV.2017.89.
- [10] LIN M, CHEN Q, YAN S C. Network in network[EB/OL]. (2014-03-04)[2025-01-23]. <https://doi.org/10.48550/arXiv.1312.4400>.
- [11] WANG W H, DAI J F, CHEN Z, et al. InternImage: exploring large-scale vision foundation models with deformable convolutions[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2023: 14408-14419.
- [12] XIONG Y W, LI Z Q, CHEN Y T, et al. Efficient deformable ConvNets: rethinking dynamic and sparse operator for vision applications[EB/OL]. (2024-11-01)[2025-01-03]. <https://arxiv.org/html/2401.06197v1>.
- [13] KRISHNA S T, KALLURI H K. Deep learning and transfer learning approaches for image classification[J]. International journal of recent technology and engineering (IJRTE), 2019, 7: 427-432.
- [14] HOWARD A G, ZHU M L, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17)[2024-11-13]. <https://doi.org/10.48550/arXiv>.
- [15] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 770-778.
- [16] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2015-04-10)[2025-03-26]. <https://doi.org/10.48550/arXiv.1409.1556>.
- [17] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [18] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C/OL]// 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2015 [2024-06-29]. <https://ieeexplore.ieee.org/document/7298594>. DOI:10.1109/CVPR.2015.7298594.

【作者简介】

刘勃(1970—),男,甘肃天水人,博士,教授,研究方向:数字图像处理。

何秉蔚(1998—),男,甘肃天水人,硕士研究生,研究方向:数字图像处理。

(收稿日期:2025-04-15 修回日期:2025-08-05)