

# 轻量化卷积神经网络在智能垃圾分类识别中的优化与验证

杨娜 付丹 要郁凯

蒙古国民族大学, 蒙古国乌兰巴托, 999097;

**摘要:** 随着垃圾分类成为生态文明建设的重要环节, 智能垃圾分类识别技术凭借高效、精准的优势逐渐替代传统人工分拣模式。然而, 传统卷积神经网络模型参数量大、计算复杂度高, 难以适配终端设备的部署需求, 轻量化卷积神经网络为此提供了有效解决方案。本文围绕轻量化卷积神经网络在智能垃圾分类识别中的应用痛点, 从网络结构优化、模型压缩策略、特征增强机制三个维度展开深入研究。通过改进深度可分离卷积的分组方式、设计动态通道注意力模块、结合知识蒸馏与量化技术, 在保证识别精度的前提下, 显著降低模型参数量与计算成本。同时, 提出基于场景适配性的验证框架, 从模型泛化能力、实时性、鲁棒性三个层面验证优化效果, 为轻量化模型在智能垃圾分类终端设备中的落地应用提供理论支撑与技术参考。

**关键词:** 轻量化卷积神经网络; 智能垃圾分类; 深度可分离卷积; 注意力机制

**DOI:** 10.64216/3080-1516.25.12.092

## 引言

垃圾分类是实现资源循环利用、缓解环境压力的关键举措, 而智能识别技术是提升垃圾分类效率的核心支撑。近年来, 卷积神经网络在图像识别领域取得突破性进展, 但其庞大的模型规模与高昂的计算开销, 使得其无法在嵌入式设备、移动终端等资源受限场景中部署——在智能垃圾分类设备中, 终端设备除了要快速响应识别需求外, 还要控制硬件成本与能耗。轻量化卷积神经网络通过减少网络结构的冗余, 在低资源环境下实现高效推理, 是解决这一矛盾的核心技术方向。目前, MobileNet, EfficientNet-Lite 等轻量化模型已初步应用于垃圾分类识别, 但是在遮挡、污渍、光照变化等复杂场景下, 仍然存在识别精度不够, 特征提取不够等缺陷。因此, 深入研究轻量化卷积神经网络的优化策略, 构建适配智能垃圾分类场景的模型, 有着重要的理论价值与实际意义。

## 1 轻量化卷积神经网络在智能垃圾分类识别中的痛点分析

### 1.1 特征提取和模型轻量化的矛盾

智能垃圾分类场景中, 垃圾种类繁多(如可回收物、厨余垃圾、有害垃圾等), 且同类垃圾形态差异大(如不同材质的塑料瓶、不同状态的厨余垃圾), 需要模型具备较强的细粒度特征提取能力。传统轻量化模型为降低计算量, 通常简化卷积操作(如采用  $1 \times 1$  卷积降维), 导致特征图通道数减少, 关键细节信息丢失。例如, MobileNet 系列采用的深度可分离卷积虽将参数量降低 8-9 倍, 但在处理相似外观垃圾(如纸类与塑料薄膜)时, 易因特征区分度不足导致误判。此外, 垃圾图像中常存

在背景干扰(如垃圾桶、包装袋), 轻量化模型的浅层特征提取模块难以有效过滤噪声, 进一步加剧了特征提取的难度。

### 1.2 模型泛化能力与场景适配性不足

智能垃圾分类设备的应用场景多样, 包括家庭厨房、社区垃圾站、商场等, 不同场景的光照条件(如强光、弱光)、拍摄角度(如俯视、侧视)差异显著, 要求模型具备较强的泛化能力。现有轻量化模型多在标准数据集(如 TrashNet)上训练, 数据分布与真实场景存在偏差——标准数据集图像背景单一、垃圾形态规整, 而真实场景中垃圾常存在遮挡、堆叠、污渍等问题。模型在训练集上表现出较高精度, 但迁移到真实场景后, 精度大幅下降, 出现“过拟合”现象。同时, 终端设备的计算资源有限, 无法通过增加模型复杂度提升泛化能力, 进一步限制了轻量化模型的应用范围。

### 1.3 模型压缩与精度保留的平衡难题

为适配终端设备, 需对轻量化模型进一步压缩(如量化、剪枝), 但压缩过程易导致模型精度损失。量化技术将浮点型参数转换为整型参数, 虽能减少内存占用与计算延迟, 但会造成参数精度下降, 尤其在低比特量化(如 4 比特、2 比特)时, 模型对特征的表征能力显著减弱。剪枝技术通过移除冗余参数与通道, 简化网络结构, 但过度剪枝会破坏网络的特征传播路径, 导致模型无法提取有效特征。例如, 对 MobileNetV2 的瓶颈结构进行通道剪枝时, 若剪枝比例超过 50%, 模型在垃圾分类任务中的精度会下降 10% 以上。如何在深度压缩模型的同时, 最大程度保留识别精度, 是轻量化模型落地应用的关键难题。

## 2 轻量化卷积神经网络的优化策略

### 2.1 基于深度可分离卷积的结构优化

针对传统深度可分离卷积特征提取能力不足的问题,本文提出改进型深度可分离卷积(Modified Depthwise Separable Convolution, MDSC),通过优化分组方式与卷积核设计,提升特征区分度。传统深度可分离卷积分为深度卷积(Depthwise Convolution)与点卷积(Pointwise Convolution):深度卷积对每个输入通道单独卷积,仅提取单通道特征;点卷积通过 $1\times 1$ 卷积融合多通道特征,但易丢失通道间的关联信息。MDSC在深度卷积阶段引入“交叉分组卷积”机制,将输入通道分为多个组,每组内采用 $3\times 3$ 卷积提取局部特征,组间通过共享卷积核实现信息交互,增强通道间的关联性。同时,在点卷积阶段加入“通道拆分-融合”模块,将输入特征图拆分为高频特征与低频特征两个分支,高频分支采用 $1\times 1$ 卷积保留细节信息,低频分支采用 $3\times 3$ 卷积增强全局特征,最后通过拼接操作融合两类特征,提升模型对细粒度特征的提取能力。此外,针对垃圾图像的背景干扰问题,在MDSC的输出端设计“背景抑制模块”(Background Suppression Module, BSM)。BSM基于注意力机制,通过计算特征图中每个像素的“垃圾置信度”,生成注意力权重图——对垃圾区域赋予高权重,对背景区域赋予低权重,实现特征图的自适应加权。具体而言,BSM首先对MDSC输出的特征图进行全局平均池化,得到通道级特征向量;然后通过全连接层与Sigmoid激活函数生成通道注意力权重,突出关键特征通道;最后将通道注意力权重与原特征图相乘,结合空间注意力机制(通过 $3\times 3$ 卷积与Softmax激活函数生成空间权重图),实现对背景噪声的有效过滤。

### 2.2 结合知识蒸馏的模型压缩优化

为解决模型压缩与精度保留之间的核心矛盾,本文引入“教师-学生”知识蒸馏框架,通过高精度复杂模型(教师模型)的特征迁移与知识引导,在训练轻量化模型(学生模型)的过程中,同步实现模型规模缩减与关键特征提取能力的保留。这一框架的核心逻辑在于,利用教师模型在长期训练中习得的复杂特征映射规律,为学生模型提供更丰富的学习信号,避免轻量化模型因结构简化而陷入特征学习不充分的困境。在教师模型与学生模型的设计上,需充分考虑垃圾分类任务的特征需求与终端部署的资源约束。教师模型选用特征表征能力强的EfficientNet-B4,它的核心优点是复合缩放策略,通过将网络的深度,宽度与输入分辨率同步调整,在不显著增加计算量的情况下,高效捕捉垃圾图像中细粒度

的特征——无论是不同材质可回收物的纹理差异,还是厨余垃圾由于状态的变化产生的形态差异,都可通过多维度缩放的网络结构进行区分。而学生模型则以改进的MDSC构建,在继承MDSC对通道的关联性和特征细节的保留能力的同时,通过简化网络层数,优化通道数量,将参数量控制在教师模型的较低比例,保证轻量化特性。在蒸馏过程中,为避免单一尺度特征传递造成的知识丢失,本文采用“多尺度特征蒸馏”策略:将教师模型与学生模型的神经网络均划分为多个对应阶段,每个阶段输出适配该层级特征抽象程度的特征图,教师模型各阶段的特征图作为“软标签”,不仅为学生模型提供类别概率信息,更引导其学习特征在不同抽象层级的分布规律——从浅层的边缘,纹理特征,到深层的类别语义特征,在知识蒸馏奠定模型精度基础后,应进一步结合量化与剪枝技术,实现模型深度压缩,满足终端设备的资源限制。量化环节采用“量化感知训练”(QAT),区别于传统的后量化方法,QAT在模型训练过程中模拟参数从浮点型到整型转换的量化误差,通过对量化敏感的参数进行针对性校准,减少量化操作对模型精度的冲击。在具体量化策略上,并非采用单一的量化标准,而是根据不同网络层的功能重要性进行差异化处理:对于MDSC的点卷积层与BSM的注意力层这类直接影响特征融合与关键区域识别的关键层,采用更高精度的量化方式,以保留其对细粒度特征的处理能力;对于其他对精度影响较小的普通卷积层,则采用常规精度量化,在精度与压缩率之间实现平衡。剪枝环节则采用“结构化剪枝”策略,避免非结构化剪枝导致的模型部署困难。剪枝前需先通过计算各层参数的L1范数,评估参数对模型输出的贡献度,以此确定参数的重要性;剪枝过程中,严格保留教师模型通过蒸馏传递给学生模型的关键特征通道,防止因冗余通道移除而破坏特征传播路径。

### 2.3 基于数据增强的特征增强优化

针对模型泛化能力不足的问题,本文从数据层面提出“场景自适应数据增强”策略,通过模拟真实垃圾分类场景的复杂条件,扩充训练数据的多样性,提升模型对不同场景的适配能力。

#### 2.3.1 动态场景数据增强

传统数据增强方法(如随机翻转、裁剪、颜色抖动)参数固定,无法针对性模拟真实场景的差异。本文提出“动态增强参数调整”机制,基于场景特征(如光照强度、背景复杂度)动态调整增强参数:在弱光场景下,增加亮度调整范围(如 $\pm 30\%$ ),同时加入高斯噪声模拟低光成像噪声;在遮挡场景下,采用“随机遮挡增强”,生成不同大小、形状的遮挡区域(如圆形、矩形),模

拟垃圾堆叠导致的遮挡；在背景复杂场景下，采用“背景替换”技术，将垃圾图像与不同场景的背景（如厨房、社区垃圾站）进行融合，增强模型对背景干扰的鲁棒性。

### 2.3.2 特征级数据增强

除图像级增强外，在模型训练过程中引入“特征级增强”，通过对特征图进行扰动，提升模型的特征学习能力。具体而言，在 MDSC 的输出特征图上，随机添加微小的高斯扰动（方差为 0.01），模拟真实场景中图像噪声对特征提取的影响；同时，采用“特征混合”技术，将不同垃圾类别的特征图按一定比例融合，生成新的特征样本，增强模型对相似类别垃圾的区分能力。特征级增强与图像级增强协同作用，有效扩充了训练数据的特征空间，提升了模型的泛化能力。

## 3 轻量化模型的验证框架构建

### 3.1 验证维度与指标设计

为全面评估优化后轻量化模型的性能，构建“三维验证框架”，从泛化能力、实时性、鲁棒性三个层面设计验证指标：

#### 3.1.1 泛化能力验证

泛化能力衡量模型在不同场景下的识别精度稳定性，采用“跨场景测试集”进行验证：测试集分为标准场景（与训练集数据分布一致）、弱光场景（光照强度低于 50lux）、遮挡场景（遮挡比例 20%~50%）、复杂背景场景（包含多种干扰物体）四类。采用平均精度均值（mAP）作为核心指标，同时计算各类别垃圾的识别精度（如可回收物精度、厨余垃圾精度），评估模型对不同类别垃圾的适配能力。

#### 3.1.2 实时性验证

实时性衡量模型在终端设备上的推理速度，采用“嵌入式平台测试”（如 NVIDIA Jetson Nano、树莓派 4B），测试指标包括模型推理时间（单张图像的平均推理时间）、帧率（每秒处理的图像数量）、内存占用（模型加载时的内存消耗）。实时性验证需模拟终端设备的实际运行环境，关闭冗余进程，确保测试结果的真实性。

#### 3.1.3 鲁棒性验证

鲁棒性衡量模型对噪声、扰动的抵抗能力，采用“对抗性测试”与“噪声测试”：对抗性测试通过添加微小的对抗性扰动（如 FGSM 攻击），评估模型在恶意干扰下的精度变化；噪声测试通过添加高斯噪声、椒盐噪声，评估模型在图像质量下降时的稳定性。采用鲁棒性得分（抗扰动后的精度与原始精度的比值）作为核心指标，

得分越高，模型鲁棒性越强。

### 3.2 验证流程与对比实验设计

#### 3.2.1 验证流程

验证流程分为三个阶段：第一阶段，在标准数据集上训练优化后的轻量化模型，确定基准精度；第二阶段，将模型部署到嵌入式平台，进行实时性测试；第三阶段，在跨场景测试集、对抗性测试集、噪声测试集上进行测试，评估泛化能力与鲁棒性。每个阶段均记录详细数据，形成完整的验证报告。

#### 3.2.2 对比实验设计

为突出优化方案的有效性，设计对比实验：选择传统轻量化模型（MobileNetV2、EfficientNet-Lite4）、未优化的 MDSC 模型作为对比对象，在相同测试环境下，对比各类验证指标。例如，在泛化能力测试中，对比优化后模型与 MobileNetV2 在弱光场景下的 mAP 差异；在实时性测试中，对比优化后模型与 EfficientNet-Lite4 的推理时间与内存占用。通过对比实验，验证结构优化、模型压缩、特征增强策略的实际效果。

## 4 结论

本文围绕轻量化卷积神经网络在智能垃圾分类识别中的应用痛点，从结构优化、模型压缩、特征增强三个维度提出解决方案：改进型深度可分离卷积（MDSC）与背景抑制模块（BSM）提升了特征提取能力，知识蒸馏与量化剪枝的协同优化实现了模型的深度压缩，场景自适应数据增强了模型的泛化能力。同时，构建多维度验证框架，从泛化能力、实时性、鲁棒性三个层面确保模型的实际应用价值。研究表明，优化后的轻量化模型在保证识别精度的前提下，显著降低了参数量与计算成本，能够适配终端设备的部署需求。

### 参考文献

- [1] 资政, 程慧, 何佳. 基于深度学习的垃圾智能分类技术[J]. 科技与创新, 2023, (21): 46-49.
- [2] 林亮. 基于卷积神经网络的垃圾图像智能识别研究[D]. 中北大学, 2023.
- [3] 邓浩, 江旺. 基于卷积神经网络的智能分类垃圾系统设计[J]. 湖南科技学院学报, 2022, 43(03): 28-32.
- [4] 刘晏铭, 丛子健, 王通, 黄子悦. 基于卷积神经网络的智能分类垃圾桶设计[J]. 智能城市, 2022, 8(06): 66-68.
- [5] 朱伟军. 基于卷积神经网络的生活垃圾智能分类研究[D]. 华东师范大学, 2022.