

基于改进轻量级MobileNetV2的辣椒病虫害图像识别

李艳美

兰州交通大学数理学院, 甘肃 兰州

收稿日期: 2025年8月29日; 录用日期: 2025年9月23日; 发布日期: 2025年9月30日

摘要

针对真实环境下辣椒病虫害识别准确率不高以及深度卷积网络参数多、模型内存大等问题, 本文提出了一种基于改进MobileNetV2的辣椒病虫害图像识别算法。首先在基线模型的基础上引入通道注意力和空间注意力机制, 提升模型对特征信息的敏感度; 同时将L2正则化加入到损失函数中, 平滑损失函数的梯度, 以缓解模型过拟合。最后实验结果表明, 改进后模型识别准确率达到94.43%, 相较于基线模型, 改进后模型精确率提升了4.38%, 召回率提升了3.38%, F1值提升了3.88%, 同时改进后模型参数量仅为2.43 M, 与基线模型相比只增加了约0.2 M。本研究方法在保持较低模型参数量的同时达到了较高的识别准确率, 具有一定的优势。

关键词

辣椒病虫害, MobileNetV2, 注意力机制, 深度学习

Identification of Pepper Diseases and Pests Based on Improved Lightweight MobileNetV2

Yanmei Li

School of Mathematics and Physics, Lanzhou Jiaotong University, Lanzhou Gansu

Received: August 29, 2025; accepted: September 23, 2025; published: September 30, 2025

Abstract

To address the challenges of low recognition accuracy for pepper diseases and pests in real-world environments, as well as the large parameter size and high memory consumption of deep convolu-

tional networks, this paper proposes an improved MobileNetV2-based image recognition algorithm for pepper disease and pest identification. First, channel attention and spatial attention mechanisms are introduced into the baseline model to enhance its sensitivity to feature information. Additionally, L2 regularization is incorporated into the loss function to smooth the gradient and mitigate model overfitting. Experimental results demonstrate that the improved model achieves an accuracy of 94.43%, with a 4.38% increase in precision, a 3.38% improvement in recall, and a 3.88% boost in F1-score compared to the baseline model. Furthermore, the enhanced model has only 2.43 M parameters, an increase of merely 0.2 M over the original model. The proposed method maintains a low parameter count while achieving high recognition accuracy, demonstrating significant advantages in practical applications.

Keywords

Pepper Pests and Diseases, MobileNetV2, Attention Mechanism, Deep Learning

Copyright © 2025 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

辣椒是我国种植面积最广的蔬菜之一，其种植面积达到约 223 万公顷，随着居民生活条件的持续改善，以及全球市场对调味品业务和蔬菜需求的快速增长，对辣椒批量化生产加工的需求也日益增长[1]。然而，在辣椒种植和生长过程中，极易受到各种病害的侵扰，这些都给辣椒的产量和质量带来了重大影响。因此，辣椒病害的早期识别和防治对于保障辣椒的高产和质量极其重要[2]。

传统的作物病害识别主要依靠专家肉眼观察或者依据经验判断，这些方法耗费时间精力，主观性强容易造成误判，导致错过最佳治疗时间。早期图像识别主要依赖手工设计特征(如 SIFT、HOG、LBP)结合传统分类器(如 SVM、随机森林)，从而实现图像识别的目的。针对传统虫害识别方法的不足，Ling 等[3]提出一种融合深度学习与机器学习的番茄害虫智能识别方案。通过构建专业图像数据集采用 CNN 模型进行特征提取，结合 DA、SVM 和 KNN 三种分类器，并引入贝叶斯优化调参。实验表明，经图像增强后 VGG16 模型准确率达 94.95%，而 ResNet50 + DA 组合模型最优准确率达 97.12%。该混合方法显著提升了识别效率与精度，为农业生产提供可靠的技术支持。李玲等[4]利用灰度共生矩阵法来实现纹理特征的快速提取，结合支持向量机分类器，实现了 90.88%的土地覆被分类。魏丽冉等[5]基于 Lab 色彩空间的 a、b 分量实现了叶片的高效分割与特征提取，在此基础上结合 K 均值聚类算法优化分割效果，同时运用核函数支持向量机完成 4 类病害的检测识别与分类，识别准确率达到 89.5%。Han 等[6]提出基于模糊识别理论的机器视觉优化算法，设计农业害虫识别系统。实验表明该算法识别率最高达 98.06%，错误率低至 5.83%，显著优于传统方法，可有效提升病虫害监测能力，助力农业减损增效。针对农业害虫图像的长尾分布问题，Chen 等[7]提出基于实例的数据增强(IDA)和约束特征调优(CFT)方法。IDA 通过重采样和混合增强提升稀有类特征提取，CFT 实现精准特征分类。实验表明，优化后的 ViT 模型在多个数据集上表现优异，最高准确率达 98.21%，较基线模型显著提升，有效解决了长尾分布下的害虫识别难题。然而早期机器学习识别方法依赖于人工特征工程，分类性能受特征质量制约，泛化能力差。

随着人工智能的迅速发展，基于计算机视觉和深度学习算法的技术被广泛应用于医学[8] [9]、金融[10]和农业[11] [12]等各种领域。深度学习采用包含大量处理层的人工神经网络架构，通过有监督或自监督学习进行训练，利用反向传播算法优化网络参数，显著提升了模型在图像识别等任务中的性能[13]。赵

法川等[14]利用层级金字塔结构,在结合深度可分离残差和循环全连接残差的基础上提出一种融合多头注意力机制的轻量级网络,该模型在实现轻量化的基础上,同时在 PlantVillage 病害数据集上实现了大于 99.7%的 Top5 准确率以及大于 95.9%的 Top1 准确率。戴敏等[15]通过引入 CBAM 注意力机制和 L2 正则化,提出了轻量化 CBAM-GoogLeNet 模型在保持高精度(准确率为 99.5%)的同时,显著减小模型体积和缩短训练时间,实现了辣椒病虫害的高效识别,为防治提供技术支持。Ferentinos 等[16]开发了一种用于植物病害检测与诊断的卷积神经网络,在 87,848 张公开图像数据集上经过训练,达到了 99.53%的准确率,可直接应用于真实种植环境。谢琬等[17]利用无参数注意力模块和空洞卷积改进 EfficientNet V2 模型,对玉米 4 种病害 Top1 准确率达到 95.99%。相较于原模型提高 1.52 个百分点。

本研究选取轻量级 MobileNetV2 模型作为主干提取网络,结合 CBAM 注意力机制使网络更加关注于病虫害特征,避免嘈杂背景的影响,同时在损失函数中加入 L2 正则化,平滑损失函数的梯度,以缓解模型过拟合。本研究提出的模型便于在移动设备端部署,同时为复杂背景下辣椒病虫害的快速识别提供帮助。

2. 数据和方法

2.1. 数据及预处理

本研究所使用的辣椒图像数据均来自于科大讯飞开放平台,这些图像均由人用手机在复杂环境下真实拍摄,选取其中 4 种类型辣椒作为研究对象:疫病、用药不当、病毒病以及温度不适。同时为了验证模型的泛化能力,本文也使用了 Plant Village 公开数据集的苹果病害图像数据,包括苹果健康叶片 1645 张,黑星病叶片 1000 张,锈病叶片 1000 张,疮痂病叶片 1000 张,各类别图像示例如图 1 所示。



Figure 1. Partial data display
图 1. 部分数据展示

为了能够对模型更好的训练,将模糊、变易等图像剔除,剩余图像按照 7:1:2 的比例划分为训练集、验证集和测试集。为避免数据量较少导致模型训练效果差等问题,本研究通过亮度增强、对比度增强和随机旋转角度等多种数据增强方法对训练集数据进行扩充处理。同时将所有图像尺寸调整为 $224 \times 224 \times 3$,并对输入图像进行归一化处理。表 1 为扩充前后各个类别的数量。

Table 1. Types and data distribution of chili pepper diseases and pests
表 1. 辣椒病虫害种类及数据分布

病虫害 类型	图像数量(张)			
	训练集	验证集	测试集	扩充后训练集
疫病	274	39	78	822
用药不当	230	33	66	690
病毒病	312	45	89	936
温度不适	283	40	81	849

2.2. 辣椒病虫害图像分类模型

2.2.1. MobileNetV2 模型

MobileNetV2 [18]模型由 Google 团队于 2008 年提出，其采用深度可分离卷积替代标准卷积，将计算复杂度降低为传统 CNN 的 1/8~1/9；同时该网络使用倒残差结构，如图 2 所示，对输入到网络中的特征先升维后降维，在更高维空间进行深度卷积和特征变换，最后压缩回低维，有效提升特征多样性；最后将倒残差结构中逐点卷积降维后的 ReLU6 激活函数替换为线性激活函数，有效提升了降维后的特征信息保留率[19]。

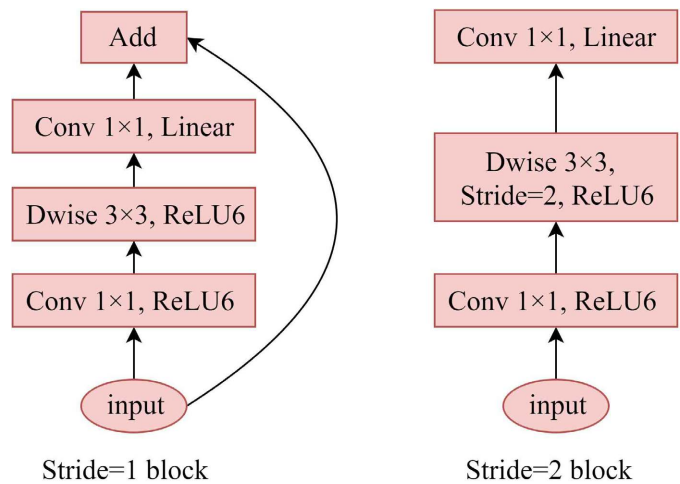


Figure 2. Inverted residual structure diagram
图 2. 倒残差结构图

通过倒残差结构和线性瓶颈结构，MobileNetV2 模型在加深网络层数的同时，每秒浮点运算次数、参数以及 CPU 耗时都比 MobileNet 的各项得到了提升。MobileNetV2 的网络结构如表 2 所示。

Table 2. MobileNetV2 network architecture
表 2. MobileNetV2 网络结构

Input	Operator	t	c	n	s
224 × 224 × 3	Conv2d	-	32	1	2
112 × 112 × 32	Bottleneck	1	16	1	1
112 × 112 × 16	Bottleneck	6	24	2	2
56 × 56 × 24	Bottleneck	6	32	3	2
28 × 28 × 32	Bottleneck	6	64	4	2
14 × 14 × 64	Bottleneck	6	96	3	1

续表

14 × 14 × 96	Bottleneck	6	160	3	2
7 × 7 × 160	Bottleneck	6	320	1	1
7 × 7 × 320	Conv2d 1 × 1	-	1280	1	1
7 × 7 × 1280	AvgPool 7 × 7	-	-	1	-
1 × 1 × 1280	Conv2d 1 × 1	-	k	-	-

注: Bottleneck 代表 MobileNetV2 的倒残差结构; t 代表扩张因子; c 代表输出通道数; n 代表该层重复数; s 代表步距。

2.2.2. 通道与空间注意力模块

卷积块注意力模块(Convolutional Block Attention Module, CBAM)由通道注意力模块和空间注意力模块构成[20], 是一种轻量级的注意力模块, 其模型框架如图 3 所示, 具体表达式如下所示:

$$F' = M_c(F) \otimes F, \quad (1)$$

$$F'' = M_s(F') \otimes F', \quad (2)$$

其中, F 表示输入特征图, M_c 和 M_s 分别表示通道和空间注意力。

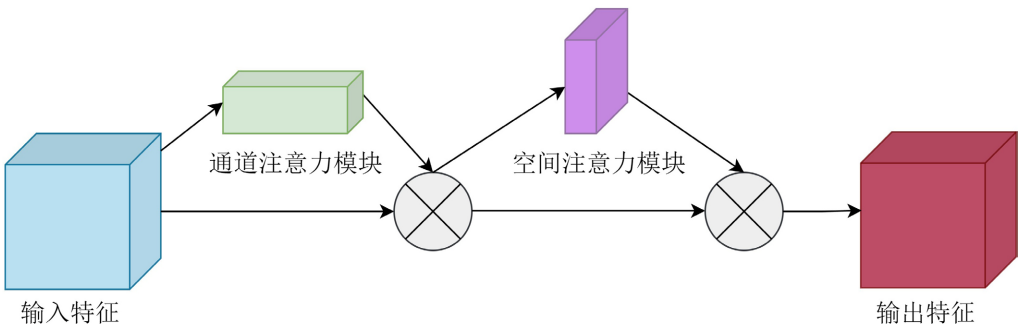


Figure 3. CBAM module structure diagram

图 3. CBAM 模块结构图

通道注意力模块主要是学习不同通道之间的相关性, 增强通道的特征表达, 具体表达式如下所示。

$$\begin{aligned} M_c(F) &= \sigma \left(\text{MLP} \left(\text{AvgPool}(F) \right) + \text{MLP} \left(\text{MaxPool}(F) \right) \right) \\ &= \sigma \left(W_1 \left(W_0 \left(F_{avg}^C \right) \right) + W_1 \left(W_0 \left(F_{max}^C \right) \right) \right) \end{aligned}, \quad (3)$$

式中, AvgPool 、 MaxPool 分别表示平均池化和最大池化操作, MLP 表示一个由 2 层全连接构成的卷积操作, MLP 的权重 W_1 和 W_0 对于两个输入是共享的, σ 表示 Sigmoid 激活函数。

空间注意力模块是学习图像中不同区域的重要性, 对不同区域的特征进行加权融合, 空间注意力模块的输入是通道注意力模块的输出特征, 具体表达式如下所示:

$$\begin{aligned} M_s(F) &= \sigma \left(f^{7 \times 7} \left(\left[\text{AvgPool}(F); \text{MaxPool}(F) \right] \right) \right) \\ &= \sigma \left(f^{7 \times 7} \left[F_{avg}^s; F_{max}^s \right] \right) \end{aligned}, \quad (4)$$

式中, $f^{7 \times 7}$ 表示大小为 7×7 的卷积操作。

CBAM 的通道注意力使用瓶颈结构的 MLP , 空间注意力仅使用两层卷积, 增加的参数量极少, 同时针对辣椒病虫害区域大小不一, CBAM 注意力机制可以聚焦关键空间区域, 忽略背景干扰, 增强模型特

征表达能力。

2.2.3. L2 正则化

正则化是机器学习中常用于防止模型过拟合的技术，通过向学习过程引入额外约束或惩罚项，在降低训练误差的同时控制模型复杂度，提升泛化能力。其数学本质是在损失函数中引入模型参数的范数惩罚项：

$$J_{reg}(\theta) = J(\theta) + \lambda \Omega(\theta), \quad (5)$$

其中 $\Omega(\theta)$ 代表正则化项， λ 控制正则化强度， θ 为模型可训练参数。

交叉熵损失函数仅关注预测概率与真实标签的匹配程度，可能导致模型在训练集上表现优异，但在测试集上泛化能力差，L2 正则化通过惩罚模型权重参数的平方和来防止过拟合。同时较大的权重可能导致模型对输入数据的微小变化过于敏感，加入 L2 正则化后，权重值被约束在较小范围内，模型输出更平滑。在交叉熵损失的基础上加入 L2 正则项后，总损失函数为：

$$\mathcal{L}_{total} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log(p_{i,c}) + \lambda \frac{1}{2} \|\theta\|_2^2, \quad (6)$$

在上述公式中，第一部分是原交叉熵损失函数，第二部分是 L2 正则项。

2.2.4. 模型总体框架

本文采用轻量级 MobileNetV2 网络，针对其在复杂背景下注意力分散、识别准确率不高等问题，引入 CBAM 注意力机制对基线模型进行改进，在原模型平均池化操作之前添加 CBAM 操作，加强特征的聚焦。同时在损失函数中引入 L2 范数，对参数权重进行约束，有效防止模型过拟合。模型整体框架结构如图 4 所示。

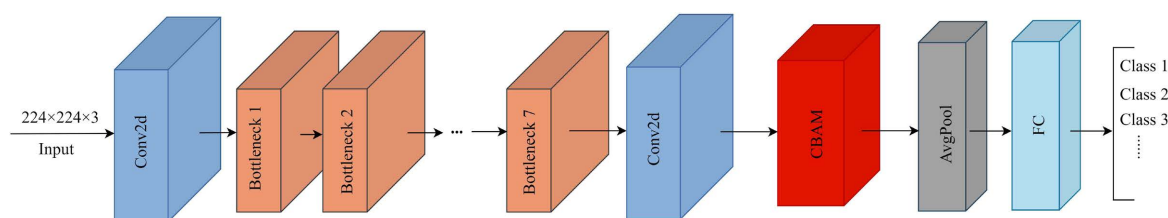


Figure 4. Model framework diagram
图 4. 模型框架图

2.3. 环境配置和超参数设置

本实验训练所用工具包版本和硬件环境参数为 PyTorch 2.0.0，Python 3.8(ubuntu20.04)，CUDA 11.8，编译软件为 Pycharm3.8，操作系统为 Windows 11，GPU 为 RTX 3080x2(20GB) * 1，CPU 为 12 vCPU Intel(R) Xeon(R) Silver 4214R CPU @ 2.40GHz。

训练过程中均采用 Adam 优化器，一阶动量为 0.9，二阶动量为 0.999，初始学习率为 0.0001，详细参数如表 3 所示。

Table 3. Experiment parameter settings
表 3. 实验参数设置

参数设置	值
批处理大小	32

续表

训练总轮数	100
学习率	0.0001
损失函数	交叉熵损失函数
优化器	Adam
正则化参数	0.00001

2.4. 评价指标

为了评估各个模型对辣椒病虫害图像的分类能力，选择准确率、精确率、召回率和 F_1 值作为模型评价指标，各个指标的计算公式如下所示：

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \times 100\% , \tag{7}$$

$$Precision = \frac{TP}{TP + FP} \times 100\% , \tag{8}$$

$$Recall = \frac{TP}{TP + FN} \times 100\% , \tag{9}$$

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100\% , \tag{10}$$

其中， TP 是真阳性样本数； FP 是假阳性样本数； FN 是假阴性样本数； TN 是真阴性样本数。

3. 实验结果与分析

3.1. 与经典 CNN 模型的结果比较

本研究提出基于注意力机制改进轻量级卷积神经网络的辣椒病虫害识别算法，为验证本文提出改进方法的有效性，将本模型与经典卷积模型进行对比，包括 AlexNet [21]、MobileNetV2 和 VGG16 [22]模型，各个分类模型除主干网络不同外，均使用相同的训练集、验证集、预处理以及优化算法等。同时本文采用迁移学习方法[23]，利用各个模型在 ImageNet 数据集的 1000 个类别上的预训练模型权重，加快模型收敛。利用上述模型进行训练，各模型训练过程中的准确率变化如图 5 所示，实验结果见表 4。

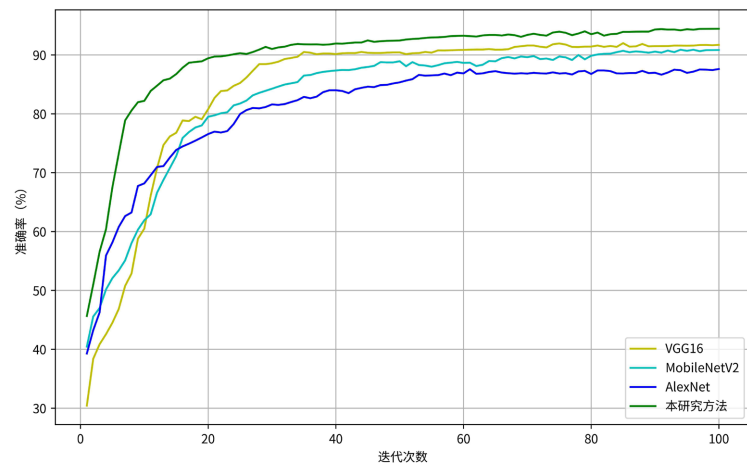


Figure 5. Accuracy variation curves of different models during training
图 5. 各模型训练过程中准确率变化曲线图

上图 5 展示了训练过程中，各个模型的识别准确率随迭代次数的变化。从上图可以发现 VGG16 作为深层网络，其训练效果明显由于其它两个轻量级模型，但其参数量大，导致训练过程收敛缓慢，而基线 MobileNetV2 模型通过轻量化设计实现了后期性能跃升，体现了深度可分离卷积的效率优势，但最终识别效果在该数据集上无明显优势。整体来看，本研究提出的改进模型在训练过程中上升迅速，在迭代 20 次左右趋于稳定，不仅收敛速度快，而且识别准确率高，明显优于其它模型。

下表 4 给出了实验中各个模型的详细结果，本研究方法准确率相较于 AlexNet、MobileNetV2 和 VGG16 模型分别提高了 6.83%、3.56%、2.44%，相较于基线模型，改进后模型精确率提升了 4.38%，召回率提升了 3.38%，F1 值提升了 3.88%，同时可以发现改进模型使用了较少的参数量(2.43M)。综上结果，本文提出的辣椒病虫害图像识别模型在添加少量参数的基础上实现了较好的分类效果。

Table 4. Performance comparison of different models

表 4. 不同模型性能对比

模型	准确率/%	精确率/%	召回率/%	F1 值/%	模型大小/MB	参数量/M
AlexNet	87.60%	87.57%	86.98%	87.27%	55.66MB	14.59M
MobileNetV2	90.87%	88.77%	89.98%	89.37%	8.51MB	2.23M
VGG16	91.99%	90.87%	89.41%	90.13%	512.24MB	134.28M
本研究方法	94.43%	93.15%	93.36%	93.25%	9.28MB	2.43M

为了深入了解各模型对各类辣椒病虫害的识别情况，利用混淆矩阵分析模型在各个类别的分类效果，结果如下图 6 所示。

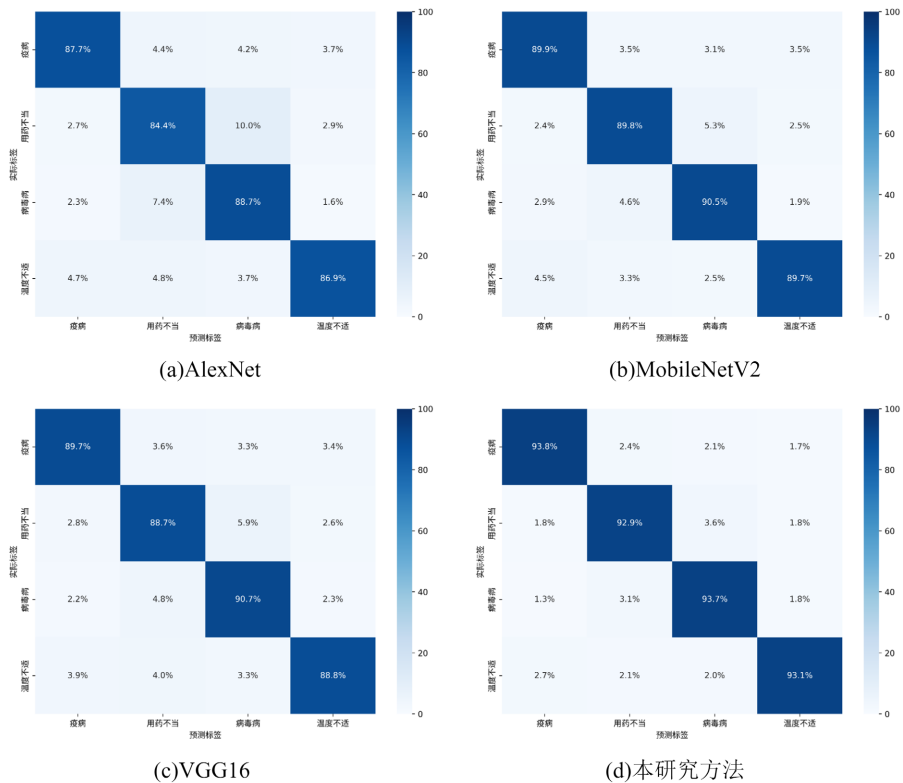


Figure 6. Model confusion matrix
图 6. 模型混淆矩阵

上图 6 清晰地展示了各个模型错分的情况，各经典模型都有一定的概率将用药不当和病毒病两种类别互相错误分类的情况，导致这两个类别的召回率相对偏低，而改进模型在注意力机制的作用下，增强了对病害特征的识别，显著提升了各个病害类别的召回率，可以看出利用 CBAM 和 L2 正则化进行改进的模型有效降低了病虫害误识率。

3.2. 可视化分析

为进一步验证 CBAM 模块在辣椒病虫害识别中的有效性，本研究采用梯度加权类激活图[24] (Gradient-weighted Class Activation Mapping, Grad-CAM)进行可视化分析，结果如下图 7 所示。观察发现，其他模型在病害识别时普遍存在识别分散和受复杂背景干扰的问题，相比之下，本文提出的改进模型通过 CBAM 注意力机制有效解决了上述问题，能够精确定位到病害位置，从而改善病害识别效果。

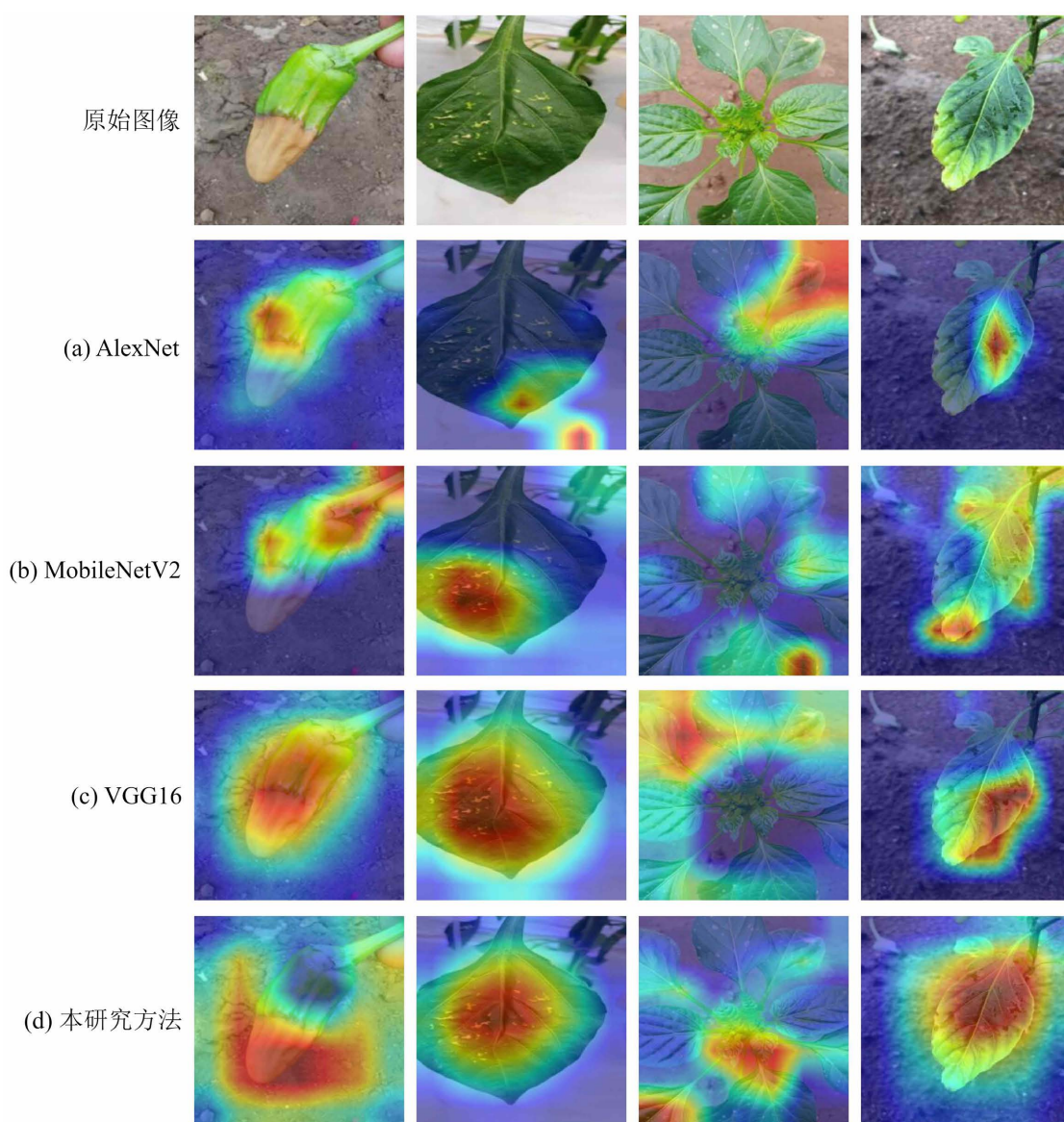


Figure 7. Gradient-weighted class activation mapping visualization results of the model
图 7. 模型梯度加权类激活可视化结果

3.3. 模型鲁棒性分析

为进一步评估本研究方法的鲁棒性，将本方法应用于 PlantVillage 苹果叶病害公开数据集。在此实验中将本研究方法与其他模型进行对比实验，具体试验结果见表 5。与其他模型相比，本方法在测试集的准确率达到 97.52%，模型在该数据集上准确率较高是由于公开数据集中叶片图像均在纯白色背景下拍摄，消除了复杂背景噪声的干扰(如土壤、杂草等)，使模型更专注于病害特征，且每个类别的样本量均衡，避免类别不平衡导致的模型偏见，从而便于模型提取特征。结果表明本研究方法对 PlantVillage 苹果叶病害公开数据集具有较好的识别效果，具有良好的鲁棒性。

Table 5. Experimental results of various models based on the PlantVillage apple leaf disease dataset
表 5. 基于 PlantVillage 苹果叶病害数据集的各模型试验结果

模型	准确率/%
AlexNet	91.26%
MobileNetV2	93.41%
VGG16	94.48%
本研究方法	97.52%

4. 结论

针对真实环境下辣椒病虫害特征较小以及传统模型难以在复杂背景下提取特征等问题，本研究在 MobileNetV2 模型的基础上引入通道注意力和空间注意力机制模块，强化了模型在辣椒病害特征区域的注意力。研究表明，改进后模型识别准确率达到 94.43%，能够更好地识别辣椒病害。相较于基线模型，改进后模型精确率提升了 4.38%，召回率提升了 3.38%，F1 值提升了 3.88%，同时改进后模型参数量仅为 2.43 M，相较于原模型只增加了约 0.2 M，便于在移动设备端部署。同时本文还在 PlantVillage 苹果叶病害公开数据集上进行了实验，准确率达到 97.52%，该模型在 PlantVillage 这一标准化的公开数据集上表现出优异的泛化能力。综上，本研究方法在实现轻量化的同时提高了对辣椒病虫害的特征提取与识别能力，具有一定的优势。

参考文献

[1] 张子峰. 我国辣椒产业发展现状、主要挑战与应对之策[J]. 北方园艺, 2023(14): 153-158.

[2] 林巧, 辛竹琳, 孔令博, 等. 我国辣椒产业发展现状及育种应对措施[J]. 中国农业大学学报, 2023, 28(5): 82-95.

[3] Huang, M., Chuang, T. and Liao, Y. (2022) Application of Transfer Learning and Image Augmentation Technology for Tomato Pest Identification. *Sustainable Computing: Informatics and Systems*, **33**, Article 100646. <https://doi.org/10.1016/j.suscom.2021.100646>

[4] 李玲, 王红, 刘庆生, 等. 基于纹理特征和支持向量机的 ALOS 图像土地覆被分类[J]. 国土资源遥感, 2011(4): 58-63.

[5] 魏丽冉, 岳峻, 李振波, 等. 基于核函数支持向量机的植物叶部病害多分类检测方法[J]. 农业机械学报, 2017, 48(S1): 166-171.

[6] Han, F., Guan, X. and Xu, M. (2024) Method of Intelligent Agricultural Pest Image Recognition Based on Machine Vision Algorithm. *Discover Applied Sciences*, **6**, Article No. 536. <https://doi.org/10.1007/s42452-024-06224-2>

[7] Chen, S., Gao, Q. and He, Y. (2024) A Pest Image Recognition Method for Long-Tail Distribution Problem. *Frontiers in Environmental Science*, **12**, Article ID: 1391770. <https://doi.org/10.3389/fenvs.2024.1391770>

[8] 武宗渊, 刘振, 张宗明. 人工智能在医学教育领域的现状、未来治理研究[J]. 中国医学伦理学, 2024, 37(9): 1093-1100.

[9] Hirasawa, T., Aoyama, K., Tanimoto, T., Ishihara, S., Shichijo, S., Ozawa, T., et al. (2018) Application of Artificial Intelligence Using a Convolutional Neural Network for Detecting Gastric Cancer in Endoscopic Images. *Gastric Cancer*,

- 21, 653-660. <https://doi.org/10.1007/s10120-018-0793-2>
- [10] 王刚, 陈红, 马敬玲, 等. 基于多尺度 1D-CNN 和注意力机制的汇率多步预测研究[J]. 系统工程理论与实践, 2024, 44(6): 1934-1949.
- [11] 王美华, 吴振鑫, 周祖光. 基于注意力改进 CBAM 的农作物病虫害细粒度识别研究[J]. 农业机械学报, 2021, 52(4): 239-247.
- [12] Juan, C., Palacio, R., Bunn, C., et al. (2025) Factors Affecting Deep Learning Model Performance in Citizen Science-Based Image Data Collection for Agriculture: A Case Study on Coffee Crops. *Computers and Electronics in Agriculture*, **232**, Article 110096.
- [13] Alruwaili, M., Alanazi, S., Abd, S. and Shehab, A. (2019) An Efficient Deep Learning Model for Olive Diseases Detection. *International Journal of Advanced Computer Science and Applications*, **10**, 486-492. <https://doi.org/10.14569/ijacsa.2019.0100863>
- [14] 赵法川, 徐晓辉, 宋涛, 等. 融合多头注意力的轻量级作物病虫害识别[J]. 华南农业大学学报, 2023, 44(6): 986-994.
- [15] 戴敏, 孙文靖, 缪宏. 基于轻量化 CBAM-GoogLeNet 的辣椒病虫害识别[J]. 中国农机化学报, 2025, 46(2): 224-229, 252.
- [16] Ferentinos, K.P. (2018) Deep Learning Models for Plant Disease Detection and Diagnosis. *Computers and Electronics in Agriculture*, **145**, 311-318. <https://doi.org/10.1016/j.compag.2018.01.009>
- [17] 谢琬, 崔艳荣. 基于改进 EfficientNetv2 模型的玉米叶片病害识别[J/OL]. 江苏农业科学, 2025, 2025-04-10. <https://link.cnki.net/urlid/32.1214.S.20250409.1406.004>, 2025-04-15.
- [18] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. and Chen, L. (2018) Mobilenetv2: Inverted Residuals and Linear Bottlenecks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 4510-4520. <https://doi.org/10.1109/cvpr.2018.00474>
- [19] Wang, Z., Zhou, W. and Li, Y. (2025) GFN: A Garbage Classification Fusion Network Incorporating Multiple Attention Mechanisms. *Electronics*, **14**, Article 75. <https://doi.org/10.3390/electronics14010075>
- [20] Woo, S., Park, J., Lee, J. and Kweon, I.S. (2018) CBAM: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Biomedical physics & engineering express*, Springer International Publishing, 3-19. https://doi.org/10.1007/978-3-030-01234-2_1
- [21] Perumal, A., Nithiyantham, J. and Nagaraj, J. (2024) An Improved Alexnet Deep Learning Method for Limb Tumor Cancer Prediction and Detection. *Biomedical Physics & Engineering Express*, **11**, Article 015004. <https://doi.org/10.1088/2057-1976/ad89c7>
- [22] Simonyan, K. and Zisserman, A. (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition. *International Conference on Learning Representations*, San Diego, 7-9 May 2015.
- [23] Shin, H., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I., et al. (2016) Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Transactions on Medical Imaging*, **35**, 1285-1298. <https://doi.org/10.1109/tmi.2016.2528162>
- [24] Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D. (2020) Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *International Journal of Computer Vision*, **128**, 336-359. <https://doi.org/10.1007/s11263-019-01228-7>