

基于多级Logit知识蒸馏的植物病害识别研究

谭颖铃, 邢荣桓, 马滨滨, 李 升*

广东海洋大学数学与计算机学院, 广东 湛江

收稿日期: 2026年4月29日; 录用日期: 2026年7月8日; 发布日期: 2026年7月17日

摘 要

植物病害是威胁全球粮食安全的重要因素, 传统人工检测方式难以满足现代智慧农业对精准化、自动化病害识别的需求。深度学习虽为此提供了解决思路, 但常规卷积神经网络参数量大、计算复杂度高, 而轻量化网络又面临分类精度不足的问题。为此, 本文提出一种基于多级Logit知识蒸馏的植物病害分类方法。该方法以ResNet34为教师模型, 通过联合实例级、批次级与类别级蒸馏损失, 从多个维度将教师网络的知识迁移至轻量学生模型中, 在不增加推理开销的前提下提升其分类性能。基于Kaggle公开的多源植物病害分类综合数据集进行实验。结果表明, 蒸馏后学生模型的Top-1准确率最高可达95.19%, 超越教师模型的94.20%。CAM可视化进一步验证了蒸馏模型对病害区域的识别能力显著增强。该方法在精度与效率间取得了良好平衡, 适合边缘设备部署, 可为农业现代化提供有效技术支撑。

关键词

图像分类, 知识蒸馏, 病害识别

Plant Disease Recognition Based on Multi-Level Logit Knowledge Distillation

Yingling Tan, Ronghuan Xing, Binbin Ma, Sheng Li*

College of Mathematics and Computer, Guangdong Ocean University, Zhanjiang Guangdong

Received: April 29, 2026; accepted: July 8, 2026; published: July 17, 2026

Abstract

Plant diseases are a major threat to global food security. Traditional manual inspection methods cannot meet the demands of modern smart agriculture for accurate and automated disease identification. Deep learning offers a promising solution, yet conventional convolutional neural networks suffer from large parameter sizes and high computational complexity, while lightweight networks often face insufficient classification accuracy. To address this issue, this paper proposes a plant disease

*通讯作者。

文章引用: 谭颖铃, 邢荣桓, 马滨滨, 李升. 基于多级 Logit 知识蒸馏的植物病害识别研究[J]. 人工智能与机器人研究, 2026, 15(4): 1059-1068. DOI: 10.12677/airr.2026.154095

classification method based on multi-level logit knowledge distillation. Taking ResNet34 as the teacher model, the method combines instance-level, batch-level, and class-level distillation losses to transfer knowledge from the teacher network to lightweight student models from multiple dimensions, thereby improving their classification performance without increasing inference overhead. Experiments are conducted on the publicly available multi-source Plant Disease Classification Merged Dataset from Kaggle. The results show that the distilled student model achieves a Top-1 accuracy of up to 95.19%, surpassing the teacher model's 94.20%. CAM visualization further confirms that the distilled model's ability to recognize disease regions is significantly enhanced. The proposed method achieves a favorable balance between accuracy and efficiency, is suitable for deployment on edge devices, and can provide effective technical support for agricultural modernization.

Keywords

Image Classification, Knowledge Distillation, Plant Disease Identification

Copyright © 2026 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

近年来,植物病害已成为威胁全球农业生产效率与粮食安全的重要因素之一。据联合国粮食及农业组织统计,全球每年因植物病虫害造成的作物损失高达 20%~40%,其中部分主要粮食作物的减产幅度甚至可达到 40%。随着全球气候变化加剧、国际贸易往来频繁以及生态环境变化,病原菌和害虫向新区域传播的速度明显加快,病害发生呈现出高频化、复杂化和区域扩散化趋势。这不仅严重影响农作物产量与品质,也给农业可持续发展和全球粮食供应链带来了巨大挑战。当前,多数植物病害检测仍主要依赖人工巡田、专家经验判断或实验室检测等传统方式,但这些方法普遍存在效率低、成本高、主观性强以及难以大规模实时监测等问题,已难以满足现代智慧农业对精准化、自动化病害监测的实际需求。而深度学习技术的快速发展为植物病害自动识别提供了新的解决思路。其中,以卷积神经网络(Convolutional Neural Network, CNN)为代表的视觉模型在图像分类任务中表现突出,并已广泛应用于农业病害识别领域。但传统卷积神经网络往往存在参数量过大,计算复杂度高等问题,从而导致难以将模型部署到手机、无人机等边缘设备上。

因此,为满足农业田间作业实时检测以及移动终端部署需求,研究人员提出采用 MobileNet、ShuffleNet 等轻量化网络,这些网络通过采用深度可分离卷积、通道重组等特殊架构设计,有效降低模型参数量和计算复杂度,使病害识别系统能够部署在无人机、智能手机、边缘设备等资源受限平台上。然而,这类轻量化模型的识别分类精度往往低于传统的深度卷积神经网络,如何在不增加网络计算开销的同时有效提升轻量化模型的分类精度成为当前研究的核心挑战。

2. 相关工作

知识蒸馏作为主流的模型压缩技术,核心原理是将大型教师模型的深层特征知识迁移至轻量化学生模型,在控制模型体积的前提下,保障识别精度。知识蒸馏主要分为基于 logit 输出的蒸馏[1]与基于中间特征的蒸馏[2]两类。前者结构简单、易于实现,但知识迁移能力有限;后者优化效果更优,却需要额外提取网络中间特征,在实际应用场景中当所有模型的中间层都可用的时候容易导致数据泄露等风险。因此金颖等人[3]提出多级对数蒸馏策略,通过多维度预测对齐方式,从样本维度、批次维度与类别维度,

统一教师模型与学生模型的输出分布，让学生模型同步学习样本特征、数据关联特征与类别关联特征。该方法无需调用教师模型中间层特征，在保证蒸馏效果的基础上，也保证了数据安全性。

近年来，知识蒸馏在植物病害识别领域逐渐成为提升模型精度与轻量化部署的重要技术路线。现有研究主要围绕模型压缩、复杂场景识别以及小样本学习等方向展开，并取得了较好的应用效果。刘媛媛等[4]提出了一种融合知识蒸馏与模型剪枝的轻量化植物病害识别方法。该研究以改进的 ResNet 为基础，在蒸馏过程中引入辅助教师模型向学生模型传递知识，并结合稀疏训练、结构化剪枝以及学习率倒带重训练策略，在显著压缩模型规模的同时保持识别性能。周罕觅等[5]针对田间复杂背景下植物病害识别准确率低、模型参数量大的问题，提出了基于知识蒸馏和轻量级卷积神经网络的 LWCR-Net 模型。该方法以 ShuffleNetV2 为骨干网络，在特征提取模块中加入残差连接，并引入 CBAM 注意力机制增强病斑区域特征表达，同时采用 DenseNet121 作为教师模型对学生网络进行蒸馏训练。聂远等[6]则进一步将知识蒸馏应用于小样本植物病害识别任务，提出增强型上下文感知知识蒸馏框架。该方法构建了教师-学生协同网络，通过指数移动平均机制实现参数更新，并融合全局特征对齐、局部特征对齐及监督式掩码图像建模等多视图蒸馏策略，以提升模型在样本稀缺条件下的泛化能力。

3. 方法

3.1. 初步介绍

本文提出采用多级 Logit 知识蒸馏提升轻量模型的植物病害分类性能。首先用规模较大的 ResNet34 网络在植物病害数据集上完成预训练，以获得一个能力更强，分类精度更高的模型作为教师模型。随后选取 ResNet18、MobileNetV2 及 ShuffleNetV1 作为学生模型，通过蒸馏策略将教师模型中的知识迁移至这些轻量化的网络中，从而提高小模型的判别能力。

教师模型选用 ResNet34，该网络通过残差连接有效缓解深层网络退化问题，具有较强的特征表达能力，在图像分类任务中表现优异，适合作为知识提供者。学生模型包括 ResNet18、MobileNetV2 与 ShuffleNetV1 三种结构。其中，ResNet18 与教师模型同属残差网络，可验证同构网络蒸馏效果；MobileNetV2 采用深度可分离卷积与注意力机制，具有较低参数量；ShuffleNetV1 通过通道重排提升计算效率，适用于边缘设备部署。

3.2. 方法描述

传统知识蒸馏通常采用教师模型与学生模型输出概率分布之间的 KL 散度作为监督信号，仅关注单一样本预测结果，忽略样本间关系及类别结构信息。为此，本文采用多层级 Logit 蒸馏方法，从实例级、批次级和类别级三个层面进行知识迁移。首先是实例级蒸馏，该层级的蒸馏约束学生模型输出接近教师模型预测分布，使其学习教师的软标签信息。设教师模型与学生模型输出的 logits 分别为 z_t 与 z_s ，温度系数为 T ，则实例级蒸馏损失定义为：

$$L_{ins} = T^2 \cdot \text{KL} \left(\sigma \left(\frac{z_t}{T} \right), \sigma \left(\frac{z_s}{T} \right) \right), \quad (1)$$

其中， $\sigma(\cdot)$ 表示 softmax 函数， $\text{KL}(\cdot)$ 表示 Kullback-Leibler 散度。然后计算批次级蒸馏损失。利用一个批次内样本预测关系矩阵，使学生学习样本间相似性结构信息。设一个 batch 中共有 N 个样本，教师与学生预测输出分别构成矩阵 $P_t, P_s \in \mathbb{R}^{N \times C}$ ，则样本关系矩阵可表示为：

$$R_t^b = P_t P_t^\top, \quad R_s^b = P_s P_s^\top. \quad (2)$$

批次级蒸馏损失定义为：

$$L_{batch} = \frac{1}{N^2} \|R_t^b - R_s^b\|_F^2 \quad (3)$$

其中, $\|\cdot\|_F$ 表示 Frobenius 范数。最后计算类别级蒸馏: 构建类别相关矩阵, 使学生捕获类别之间潜在联系, 提高细粒度病害识别能力。类别关系矩阵定义为:

$$R_t^c = P_t^\top P_t, \quad R_s^c = P_s^\top P_s, \quad (4)$$

则类别级蒸馏损失为:

$$L_{class} = \frac{1}{C^2} \|R_t^c - R_s^c\|_F^2. \quad (5)$$

因此总的蒸馏损失可以表示为:

$$L_{KD} = L_{ins} + L_{batch} + L_{class}. \quad (6)$$

在实际训练过程中, 学生模型最终优化目标通常结合真实标签监督损失共同进行训练, 即:

$$L = \alpha L_{CE} + \beta L_{KD} \quad (7)$$

其中, L_{CE} 为交叉熵损失, α , β 为权重参数。通过上述联合优化策略, 可有效提升轻量化学生模型在植物病害分类任务中的识别精度与泛化能力。

4. 实验

4.1. 数据集描述及数据预处理

我们采用来自 Kaggle 平台的公开多源植物病害分类综合数据集(Plant Disease Classification Merged Dataset)进行实验训练。该数据集整合了 14 个公开植物病害数据源构建而成, 旨在为植物病害识别任务提供更大规模、更高多样性的图像资源。数据集包含实验室环境与真实田间环境采集的叶片图像, 覆盖苹果、玉米、葡萄、辣椒、咖啡、木薯等多种重要农作物, 同时包含健康叶片与多种常见病害类别。数据集总共约有 7.9 万张图像, 共 88 个分类类别, 总规模约 18 GB。具体而言, 各类别样本量差异显著, 从超过 2000 张(如番茄晚疫病)到不足 100 张(如特定的罕见锈病)不等, 这种分布真实反映了现实农业场景中的内在挑战, 即不同病害的发生频率存在巨大差异。如图 1 所示, 本文可视化了样本数量最多的前五个类别, 和样本数量最少的五个类别。且该数据集的图像具备场景复杂、类别丰富、样本量大的优点, 适合作为农业病虫害识别的训练数据集。为提升模型泛化能力并减少过拟合现象, 在训练前对原始图像进行统一预处理。首先, 将所有图像缩放至 224×224 像素, 以满足 ResNet50、ResNet18、MobileNetV3 及 ShuffleNetV2 等网络输入要求。随后, 对训练集采用数据增强策略, 包括: 随机水平翻转, 随机垂直翻转, 随机裁剪以及色彩抖动。验证集与测试集仅进行尺寸缩放与归一化处理, 不采用随机增强操作。数据集按 7:3 的比例划分为训练集与测试集, 用于模型训练、参数调优及最终性能评估。

4.2. 实验环境设置

本文实验基于 Ubuntu 22.04 操作系统, 在 PyTorch 2.1.2 深度学习框架下实现, Python 版本为 3.10, CUDA 版本为 11.8。硬件平台采用单卡 NVIDIA GeForce RTX 4090 GPU (24 GB 显存)进行模型训练, CPU 为 16 vCPU Intel Xeon Gold 6430, 系统内存为 120 GB。实验环境能够满足教师模型与轻量化学生模型的训练需求, 并保证模型训练效率与稳定性。在训练过程中, 采用 GPU 加速完成前向传播与反向传播计算, 数据读取及预处理由 CPU 协同完成。预训练教师模型将 batch 大小设置为 32, 共训练了 60 轮。其余的轻量学生模型训练时的 batch 大小都设置为 64, 总训练轮数为 100 轮初始学习率均设置为 0.1, 蒸馏温度默认都设置为 2.0。蒸馏过程中停止了教师的梯度, 防止参数回传影响模型训练。

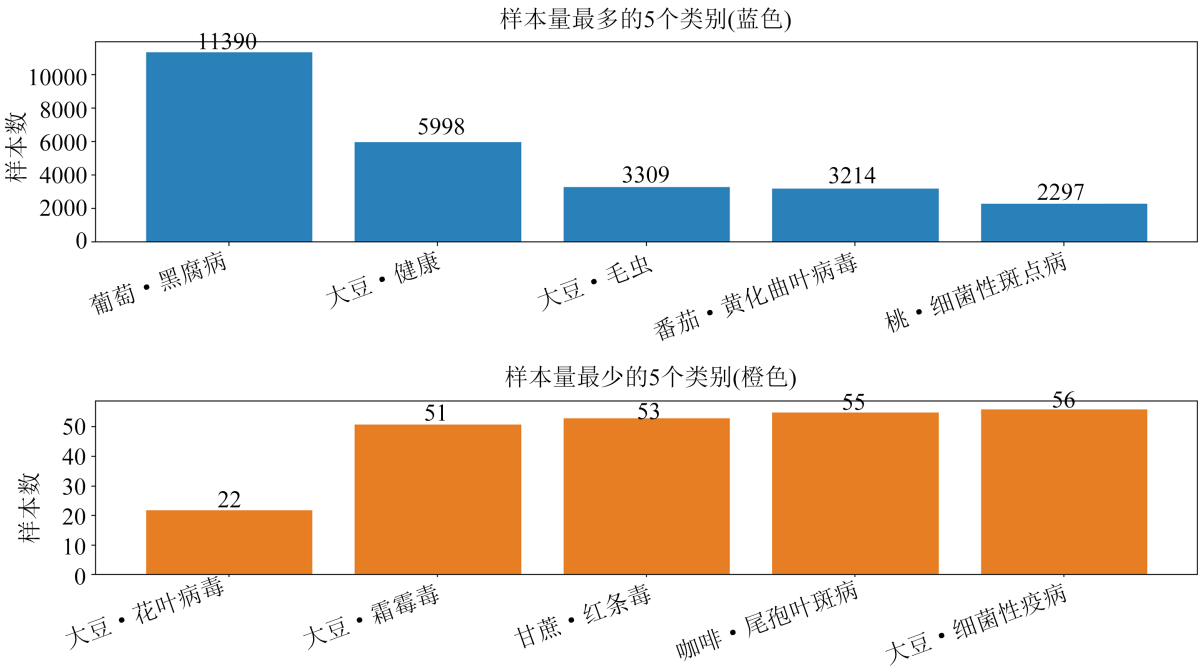


Figure 1. The top five categories with the largest sample size and the bottom five categories with the smallest sample size
图 1. 样本数量最多的五个类别与样本数量最少的五个类别

4.3. 实验分析

表 1 展示了教师模型与各学生模型在植物病害分类数据集上的性能对比结果，其中教师模型采用 ResNet34，学生模型包括 ResNet18、MobileNet 和 ShuffleNet，并分别给出了基线训练(baseline)与知识蒸馏(KD)后的实验结果。

Table 1. Multi-level logit distillation effect comparison experiment table
表 1. 多级 logit 蒸馏效果对比实验表

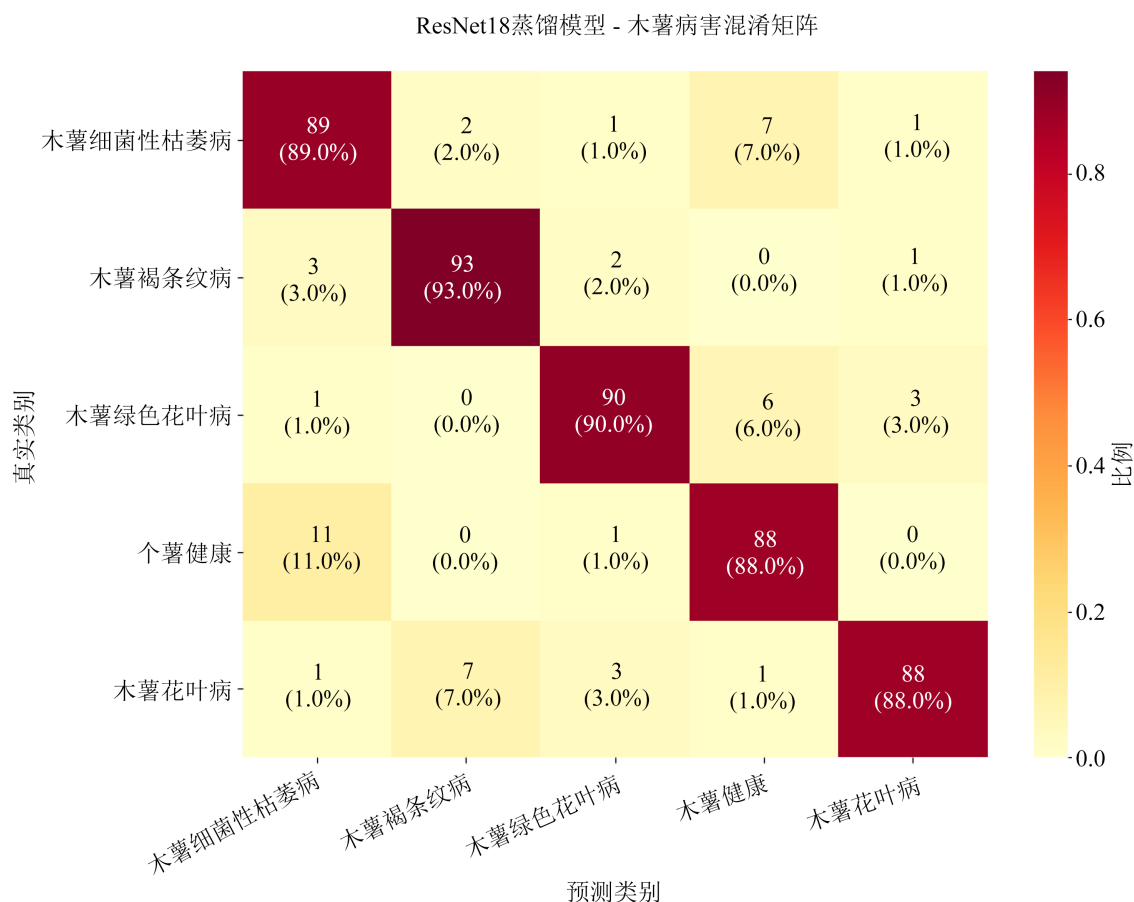
模型	教师模型			学生模型			
	ResNet34	ResNet18		MobileNet		ShuffleNet	
		baseline	KD	baseline	KD	baseline	KD
TOP-1	94.20%	87.12%	95.19%	92.97%	95.07%	92.91%	94.62%
TOP-5	99.41%	98.24%	99.55%	98.82%	99.45%	99.02%	99.34%
recall	0.9420	0.8712	0.9519	0.9297	0.9507	0.9291	0.9462
F1-score	0.9417	0.8663	0.952	0.9299	0.9507	0.928	0.9465
Params	21.32 M	11.22 M	11.22 M	3.30 M	3.30 M	1.35 M	1.35 M
Flops	51.72 G	27.34 G	27.34 G	2.35 G	2.35 G	0.56 G	0.56 G
单样本推理耗时	4.1175 ms	2.6191 ms	2.5638 ms	1.5487 ms	1.6688 ms	1.2063 ms	1.0751 ms
模型大小	81.5 MB	42.9 MB	42.9 MB	12.8 MB	12.8 MB	5.3 MB	5.3 MB

首先,在分类准确率(Top-1 准确率)指标上,所有学生模型在经过蒸馏后均获得明显提升。ResNet18 的 Top-1 准确率由 87.12%提升至 95.19%,提高了 8.07%,甚至超过教师模型 ResNet34 的 94.20%; MobileNet 由 92.97%提升至 95.07%,提升 2.10%; ShuffleNet 由 92.91%提升至 94.62%,提升 1.71%。结果表明,通过教师模型的知识迁移,学生模型能够学习到更具判别性的类别边界信息,从而有效弥补轻量网络表达能力不足的问题。

在 Top-5 Accuracy 指标上,各蒸馏模型同样表现优异。ResNet18 经蒸馏后达到 99.55%,较基线模型提升 1.31%; MobileNet 提升至 99.45%; ShuffleNet 提升至 99.34%。说明蒸馏训练不仅提高了模型对最优类别的判断能力,也增强了其候选类别预测的稳定性。

在召回率(Recall)与 F1-score 指标方面,蒸馏后的模型同样具有明显优势。其中,ResNet18 的 Recall 从 0.8712 提升至 0.9519, F1-score 从 0.8663 提升至 0.9520,提升幅度最为显著,说明蒸馏方法有效改善了模型对病害类别的识别完整性与分类均衡性。并且我们对 ResNet18 网络,随机抽取了 5 个类别绘制了混淆矩阵并且对其准确率、召回率以及 F1 分数进行了可视化,如图 2 所示。此外, MobileNet 和 ShuffleNet 也分别取得了不同程度提升,进一步验证了该方法具有良好的通用性和鲁棒性。

从模型复杂度角度来看,学生模型参数量与计算量远低于教师模型。且在显著提升了模型性能的同时,对比基线模型无额外的计算开销。教师模型 ResNet34 参数量为 21.32 M, FLOPs 为 51.72 G,而蒸馏后的 ResNet18、 MobileNet 和 ShuffleNet 参数量分别仅为 11.22 M、3.30 M 和 1.35 M,计算量分别为 27.34 G、2.35 G 和 0.56 G。且三个模型都在不增加模型复杂度,保持原本模型的极低开销的情况下实现了模型性能的有效提升。



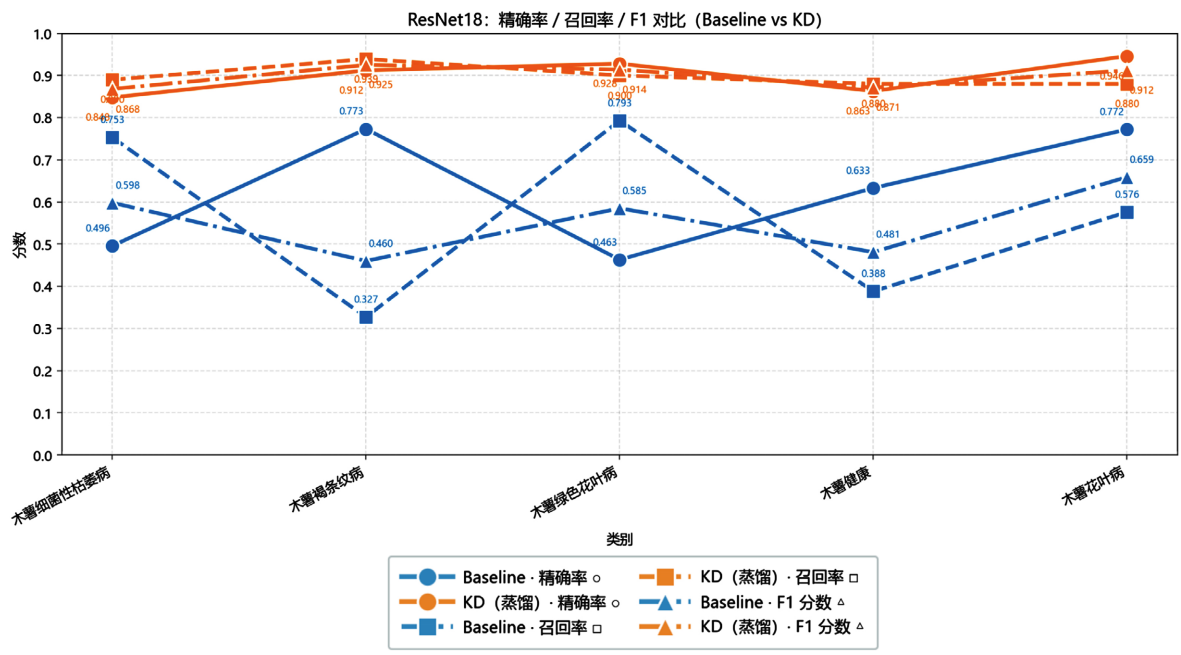


Figure 2. Visualization of confusion matrix, precision, recall and F1-score for partial classes under ResNet18 architecture
图 2. ResNet18 架构下部分类别混淆矩阵与精确率、召回率、F1 分数可视化

综合来看，知识蒸馏方法能够在不增加模型结构复杂度的前提下，显著提升轻量化网络在植物病害识别任务中的性能。其中，ResNet18 蒸馏后表现最佳，在精度上超过教师模型；MobileNet 与 ShuffleNet 则在精度与效率之间取得了更优平衡，更适合移动端或边缘设备部署。因此，该蒸馏策略在农业病虫害识别分类中具有有良好的应用前景。

在本研究中，不同架构的学生模型在接受多级 Logit 蒸馏后表现出的性能增益存在显著差异。其中，ResNet18 的提升最为突出，其 Top-1 准确率达到 95.19%，甚至超越了教师模型 ResNet34 (94.20%)；而 MobileNetV2 与 ShuffleNetV1 虽然也取得了显著提升，但未出现“超越教师”的现象。针对这一现象，本文从以下两个角度进行深度分析。

4.4. 针对不同架构学生模型获益差异的讨论

1) 模型架构的同质性

ResNet18 与教师模型 ResNet34 具有高度的架构相似性，两者均属于残差网络，拥有相同的归纳偏置。这种同质性使得学生模型在特征提取路径和 Logit 分布上与教师模型更为接近，因此在多级蒸馏过程中，ResNet18 能够更轻易地理解并拟合教师模型输出的软标签中的类别相关性。此外，蒸馏损失在此类同质架构中起到了强大的正则化作用，帮助学生模型跳出局部最优解，从而实现了性能的大幅提升。

2) 模型容量差异与信息瓶颈

MobileNetV2 和 ShuffleNetV1 属于极轻量化网络与教师模型 ResNet34 存在较大的容量差距，因此尽管多级 Logit 蒸馏提供了丰富的结构化知识，但由于模型本身的参数容量的限制，轻量化学生模型在吸收复杂特征映射和高维分布知识时存在信息瓶颈，因此模型的性能提升会受到一定的限制。相比之下，ResNet18 拥有足够的参数规模来承载教师模型的知识迁移，因此表现出更强的获益弹性。

4.5. 可视化分析

为探究蒸馏温度 T 对轻量级学生模型的影响，我们以 ShuffleNet 为基础进行了一系列实验，结果如

表 2 与图 3 所示。可以看出,随着温度升高,Top-1 准确率整体获得明显改善。当温度为 1 时,教师输出接近硬标签,学生网络仅取得 93.66%的准确率;随着温度增加,软化的概率分布提供了更丰富的类别间相似信息,学生网络的性能也随之提升,当 $T = 6$ 时网络的准确率达到了 94.95%。整体而言,在合适的范围内,随着温度的升高学生的准确率也会提升。

Table 2. Performance of ShuffleNet under different distillation temperatures
表 2. ShuffleNet 不同蒸馏温度下的性能表现

温度 T	Top1	Top5	Recall	F1-score
1	93.66%	99.07%	0.9366	0.9365
2	94.62%	99.34%	0.9462	0.9465
3	94.66%	99.48%	0.9466	0.9468
4	94.76%	99.44%	0.9476	0.9477
5	94.63%	99.45%	0.9463	0.9466
6	94.95%	99.42%	0.9495	0.9496

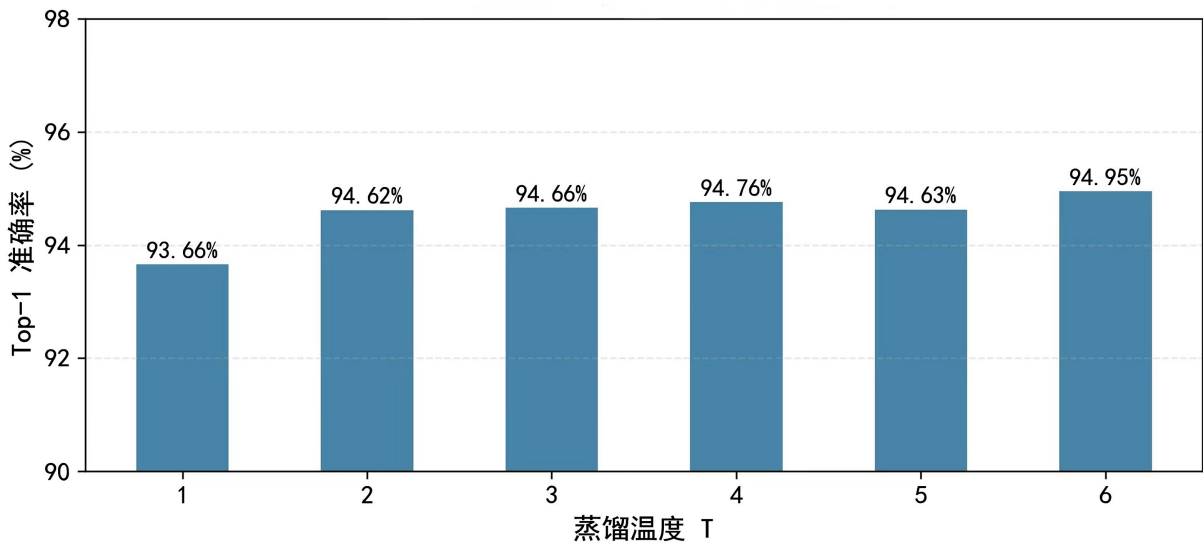


Figure 3. Top-1 accuracy of shufflenet at different distillation temperatures
图 3. 不同蒸馏温度下 ShuffleNet 的 Top-1 准确率

为进一步直观验证知识蒸馏对模型特征提取能力的提升效果,我们选取 ResNet18 在基线训练(Baseline)与知识蒸馏(KD)两种方式下的模型,随机抽取了 4 张病害样本图像进行类激活映射(CAM)可视化,结果如图 4 所示。

图中第一行为原始输入图像,第二行为 Baseline 模型的 CAM 热力图,第三行为 KD 模型的 CAM 热力图。颜色越偏红表示模型对该区域的关注程度越高,越偏蓝则关注度越低。从可视化结果可以看出,Baseline 模型的关注区域相对分散,部分注意力落在了背景或非病害区域,甚至存在偏离主要病斑位置的情况。相比之下,经过知识蒸馏后的 KD 模型,其关注区域更加集中和精准,热力高亮区域与病害部位的轮廓重合度明显更高。这表明通过多级 Logit 蒸馏,学生模型从教师网络中学习到了更具判别性的类别相关区域信息,从而能够更好地定位病害关键特征,抑制背景噪声干扰。

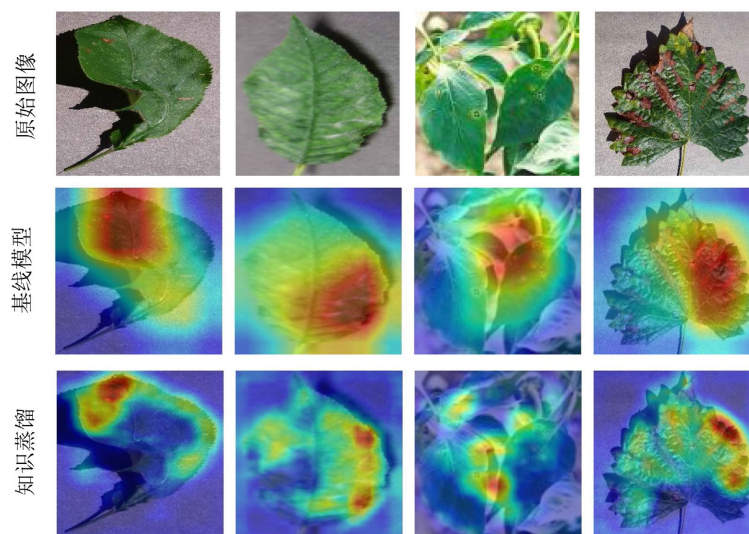


Figure 4. CAM visualization of diseased leaves in different categories based on ResNet18 model

图 4. ResNet18 模型对不同类别病害叶片的 CAM 可视化

此外,对比不同样本的可视化结果还可发现,对于病斑面积较小或病害特征不明显的样本,Baseline模型的定位能力较弱,而KD模型的关注区域依然保持较好的准确性,进一步说明了蒸馏训练在提升模型鲁棒性和泛化能力方面的积极作用。综上,CAM可视化从特征定位角度验证了知识蒸馏方法的有效性,蒸馏后的ResNet18不仅分类精度提升,其对病害区域的感知能力也得到了显著增强。

5. 结论

本研究面向植物病害图像分类任务,提出了一种基于多级Logit知识蒸馏的轻量化识别方法。实验基于Kaggle平台的公开多源植物病害分类综合数据集,该数据集包含约7.9万张图像、88个分类类别,覆盖实验室与田间环境下的多种作物及病害类型。以ResNet34为教师模型,选取ResNet18、MobileNetV2与ShuffleNetV1三种轻量网络作为学生模型,通过联合实例级、批次级与类别级蒸馏损失,将教师模型中的多层级结构化知识迁移至学生网络。结果表明,该蒸馏方法可在不增加推理开销的前提下,显著提升各轻量模型的分类精度。其中,蒸馏后的ResNet18取得95.19%的Top-1准确率,超越教师模型的94.20%;ShuffleNet在仅1.35M参数量和0.56GFLOPs的条件下达到94.62%的准确率,展现出在多类别、复杂场景病害识别任务中良好的精度-效率平衡,适合资源受限的边缘设备部署。

同时,本研究仍存在一些不足之处。例如,数据集类别分布可能存在不均衡现象,部分稀有病害类别的识别效果尚有提升空间;教师模型结构选择较为单一,不同教师网络对多层次蒸馏效果的影响有待进一步对比;此外,蒸馏温度、损失权重等超参数多凭经验设定,未进行系统的自动寻优。未来工作可围绕数据增强与类别均衡策略、更丰富的教师模型结构探索、关键超参数的自动调优等方面展开,从而进一步提升模型的性能与泛化能力。

参考文献

- [1] Hinton, G., Vinyals, O. and Dean J. (2015) Distilling the Knowledge in a Neural Network. <https://arxiv.org/abs/1503.02531>
- [2] Heo, B., Kim, J., Yun, S., Park, H., Kwak, N. and Choi, J.Y. (2019) A Comprehensive Overhaul of Feature Distillation. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 27 October-02 November 2019. <https://doi.org/10.1109/iccv.2019.00201>

-
- [3] Jin, Y., Wang, J. and Lin, D. (2023) Multi-Level Logit Distillation. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, BC, Canada, 17-24 June 2023.
<https://doi.org/10.1109/cvpr52729.2023.02325>
 - [4] 刘媛媛, 王定坤, 邬雷, 等. 基于知识蒸馏和模型剪枝的轻量化模型植物病害识别[J]. 浙江农业学报, 2023, 35(9): 2250-2264.
 - [5] 周罕觅, 陈佳庚, 代智光, 等. 基于知识蒸馏和轻量级卷积神经网络的植物病害识别方法[J]. 南京农业大学学报, 2024, 47(6): 1189-1201.
 - [6] 聂远, 周厚奎, 张广群, 等. 基于增强知识蒸馏的小样本植物病害识别方法[J]. 浙江农林大学学报, 2025, 42(4): 667-676.