

引用格式: 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. NIE Yuan, ZHOU Houkui, ZHANG Guangqun, *et al.* Few-shot learning for plant disease recognition with enhanced knowledge distillation[J]. *Journal of Zhejiang A&F University*, 2025, 42(4): 667–676.

## 基于增强知识蒸馏的小样本植物病害识别方法

聂远<sup>1</sup>, 周厚奎<sup>1,2,3</sup>, 张广群<sup>1,2,3</sup>, 何涛<sup>1,2,3</sup>, 胡军国<sup>1,2,3</sup>

(1. 浙江农林大学 数学与计算机科学学院, 浙江 杭州 311300; 2. 浙江农林大学 浙江省林业智能监测与信息技术重点实验室, 浙江 杭州 311300; 3. 浙江农林大学 国家林业与草原局林业感知技术与智能装备重点实验室, 浙江 杭州 311300)

**摘要:** 【目的】针对植物病害识别中大规模标注数据依赖性强、新病害适应性差的问题, 本研究旨在提升小样本方法在复杂环境下的识别能力, 为小样本病害识别提供理论依据。【方法】本研究提出一种结合小样本转换器和监督式掩码知识蒸馏的增强型上下文感知知识蒸馏框架 (ECKD), 构建包含教师网络与学生网络的协同结构, 并结合全局特征对齐、局部特征对齐和监督式掩码图像建模视图策略, 引入通道注意力残差模块和上下文感知模块, 设计新型 ViT-EC 编码器以增强模型的特征提取与语义理解能力。基于 PlantVillage 数据集, 采用原型分类器进行 5-way 1-shot 和 5-way 5-shot 的小样本任务评估。【结果】在 5-way 1-shot 和 5-way 5-shot 小样本任务中, ECKD 的原型分类器平均准确率分别达 74.98% 和 88.28%, 在多个小样本任务上明显优于现有的主流方法。消融实验表明: 注意力残差模块、上下文感知模块和监督式掩码图像建模视图策略对性能有正向贡献, 在数据不完整场景下能增强模型对全局语义的理解与局部特征的重建能力。【结论】ECKD 通过多视图增强策略与知识蒸馏机制的融合, 有效缓解了植物病害中小样本的识别难题, 提升了识别的精度和模型的稳定性。该方法为农业智能感知系统提供了高效、可行的解决方案, 具有良好的推广潜力。图 8 表 4 参 29

**关键词:** 植物病害识别; 小样本学习; 知识蒸馏; 视图策略; 原型分类器

中图分类号: S126 文献标志码: A 文章编号: 2095-0756(2025)04-0667-10

## Few-shot learning for plant disease recognition with enhanced knowledge distillation

NIE Yuan<sup>1</sup>, ZHOU Houkui<sup>1,2,3</sup>, ZHANG Guangqun<sup>1,2,3</sup>, HE Tao<sup>1,2,3</sup>, HU Junguo<sup>1,2,3</sup>

(1. College of Mathematics and Computer Science, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China; 2. Zhejiang Provincial Key Laboratory of Forestry Intelligent Monitoring and Information Technology, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China; 3. Key Laboratory of Forestry Perception Technology and Intelligent Equipment, National Forestry and Grassland Administration, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China)

**Abstract:** [Objective] This study aims to address the strong dependence of plant disease recognition on large-scale annotated data and its poor adaptability to novel diseases, and enhance the performance of few-shot learning methods under complex environments, so as to provide a theoretical basis for few-shot plant disease recognition. [Method] A novel recognition framework, enhanced context-aware knowledge distillation (ECKD), was proposed by integrating a few-shot transformer with supervised masked knowledge distillation. The framework adopts a collaborative architecture composed of teacher and student networks, and introduces

收稿日期: 2025-04-07; 修回日期: 2025-07-02

基金项目: 浙江省自然科学基金资助项目 (LY24F020005)

作者简介: 聂远 (ORCID: 0009-0003-3392-0794), 从事人工智能与大数据研究。E-mail: [2023611031024@stu.zafu.edu.cn](mailto:2023611031024@stu.zafu.edu.cn)。通信作者: 周厚奎 (ORCID: 0000-0001-7915-8684), 副教授, 博士, 从事深度学习在农林领域的应用研究。E-mail: [zhouhk@zafu.edu.cn](mailto:zhouhk@zafu.edu.cn)

strategies such as global feature alignment, local feature alignment, and supervised masked image modeling views. To enhance feature extraction and semantic understanding, a new encoder ViT-EC was developed using a channel attention residual module and a context-aware module. The model was evaluated using prototype-based classifiers on the PlantVillage dataset under 5-way 1-shot and 5-way 5-shot tasks. **[Result]** The prototype classifier based on the ECKD framework achieved average accuracies of 74.98% and 88.28% in the 5-way 1-shot and 5-way 5-shot tasks, respectively, significantly outperforming several existing methods. Ablation studies confirmed the positive contributions of the attention residual module, context-aware module, and the supervised masked image modeling strategy in enhancing global semantic understanding and local feature reconstruction, especially under incomplete data conditions. **[Conclusion]** The ECKD framework effectively alleviates the challenges of few-shot learning in plant disease recognition by integrating multi-view augmentation strategies and a knowledge distillation mechanism. It significantly improves recognition accuracy and model stability, offering an efficient and practical solution for intelligent agricultural perception systems with promising applicability. [Ch, 8 fig. 4 tab. 29 ref.]

**Key words:** plant disease identification; few-shot learning; knowledge distillation; view strategy; prototype classifier

植物病害严重威胁作物的产量和品质, 如何高效监测作物健康状况是保障农业生产的关键<sup>[1-3]</sup>。传统病害识别依赖农民经验或农业专家现场判断, 但受限于经验不足、专家稀缺和成本高等问题, 难以及时准确地应对新发病害。因此, 研发高效智能的植物病害识别算法具有重要意义。

近年来, 深度学习技术在农业图像分析领域取得了显著进展, 深度学习算法逐步取代机器学习算法, 成为植物病害分类的主流方法<sup>[4-6]</sup>。尽管采用 AlexNet<sup>[7]</sup>、VGGNet<sup>[8]</sup> 等卷积神经网络在植物病害识别中取得了良好效果, 但深度卷积神经网络在处理跨域、异质病斑图像以及应对新类别病害时表现出明显的适应性不足。因此, 如何在样本稀缺的条件下实现对植物病害的准确识别, 已成为当前研究关注的重点。

小样本学习 (FSL) 能够在样本有限的条件下, 快速适应新类别识别需求。近年来, 小样本学习已成为农业视觉领域解决数据稀缺问题的重要研究方向。基于度量的方法因其直观的理论基础和优异的性能, 成为小样本学习领域的主流研究方向<sup>[9]</sup>。研究人员探索了多种度量学习方法, 如基于类均值表示的欧氏距离<sup>[10]</sup>、脊回归方法<sup>[11]</sup> 以及图神经网络<sup>[12]</sup> 等。在农业领域, 也有部分小样本分类的研究。这些方法通过生成对抗网络<sup>[13]</sup>、度量学习损失<sup>[14]</sup>、聚合损失函数<sup>[15]</sup> 等策略, 在特定场景下取得了一定成效。然而, 这些方法仍存在特征表达不足、跨域泛化能力弱等问题, 尤其在应对复杂背景与异质数据时性能下降明显。基于转换器 (Transformer) 架构的视觉模型, 因其强大的全局特征建模能力, 逐渐成为农业图像识别领域的新趋势。现有方法主要通过融合迁移学习与小样本学习降低数据依赖<sup>[16]</sup>, 利用轻量级注意力机制结合高效 Transformer 架构<sup>[17]</sup>, 进行局部与全局特征协同建模<sup>[18]</sup>, 以及引入掩码自编码器 MAE<sup>[19]</sup> 等方式, 增强模型的小样本适应能力, 从而提升植物病害识别的性能与鲁棒性。尽管上述研究在植物病害小样本识别方面取得了一定进展, 但对数据预处理依赖性强、计算资源需求高以及特征表达能力受限等问题仍然存在。鉴于此, 本研究基于视觉转换器 (Vision Transformer), 将小样本转换器<sup>[20]</sup> 和知识蒸馏思想引入到植物病害识别当中, 设计出一种全新的增强型上下文感知知识蒸馏框架 (ECKD), 通过改进监督式掩码知识蒸馏思想, 融合全局特征信息与局部特征信息, 解决模型识别效率低和应用可行性的问题, 以期能为植物病害小样本智能识别提供理论依据和方法支持。

## 1 研究方法

### 1.1 小样本学习任务

小样本学习任务中数据集通常划分为源数据集 ( $D_s$ ) 和目标数据集 ( $D_t$ )。  $D_s$  用于模型训练, 提供大量标注样本以学习通用特征表达;  $D_t$  用于模型评估, 包含仅有少量样本的新类别。在给定的  $N$ -way  $K$ -shot

(5-way 5-shot 和 5-way 1-shot) 任务中, 每个任务包含支持集样本 ( $D_{\text{sup}}$ ) 和查询集样本 ( $D_{\text{qv}}$ ), 任务表示为  $D_{\text{task}}=(D_{\text{sup}}, D_{\text{qv}})$ 。其中,  $D_{\text{sup}}$  和  $D_{\text{qv}}$  分别从  $D_t$  的  $N$  个类别中采样。 $D_{\text{sup}}$  包含每个类别的  $K$  个样本, 用于估计类别原型,  $D_{\text{qv}}$  则用于评估模型的性能。在小样本学习任务中, 目标是将  $D_{\text{qv}}$  中的样本正确分类到这  $N$  个类别中。因此, 本研究的重点是在  $D_s$  上训练 1 个具有良好泛化能力的特征提取器, 在  $D_t$  上使用特征提取器进行小样本分类任务, 使其能够在  $D_t$  的小样本任务中实现出色的分类性能。

## 1.2 增强知识蒸馏框架

本研究提出的 ECKD 旨在提升模型在植物病害识别任务中的表现, 图 1 为 ECKD 的框架结构。该框架包含学生网络和教师网络 2 个部分, 网络之间通过指数移动平均 (EMA) 保持协同。如图 2 所示: ECKD 融合了全局特征对齐 (CA)、局部特征对齐 (PA) 和监督式掩码图像建模 (SMIM) 3 种视图策略, 这些策略用于提升学生网络在小样本学习中的泛化能力和特征表达能力。ECKD 实现过程如下: 首先, 从同一种类别标签中输入 2 张不同的图像, 图像 1 在进行分块处理后采用遮蔽增强策略输入学生网络, 图像 2 则在分块处理后保持完整输入教师网络。随后, 学生网络和教师网络中的图像块均通过 ViT-EC 编码器转化为全局特征 (CLS) 和局部特征 (Patch), 学生网络与教师网络通过 EMA 实现参数更新。最后, 通过图像 1 和图像 2 的全局特征和局部特征计算 CLS 和 Patch 的损失, 实现全局特征和局部特征对齐。ECKD 通过 ViT-EC 编码器和视图策略实现增强知识蒸馏, 保证学生网络能够在不完整输入数据的情况下, 从教师网络中学习更丰富的全局和局部特征表示, 最终提升模型在小样本学习场景下的特征表达能力和泛化能力。

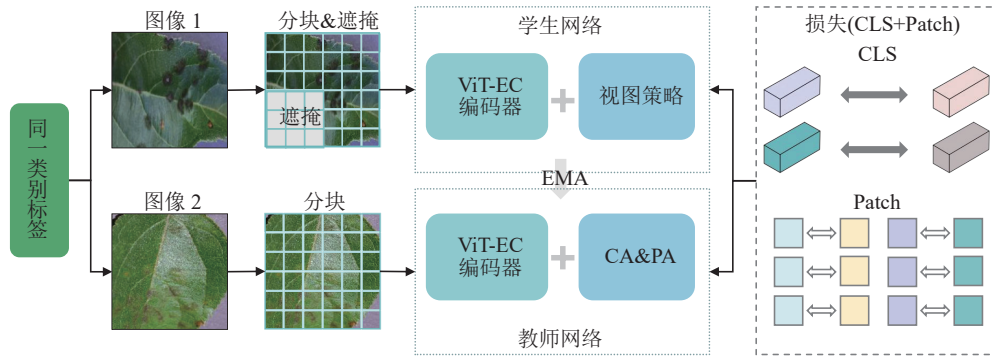


图 1 增强型上下文感知知识蒸馏框架 (ECKD) 结构流程

Figure 1 Enhanced context-aware knowledge distillation (ECKD) structure flow

**1.2.1 全局特征对齐** 模型对全局特征的捕捉能力决定其在稀缺训练数据下的泛化能力和性能表现。全局特征对齐策略的核心思想是: 虽然同一类别不同样本的局部特征存在差异, 但其全局特征表征应保持高度的一致性, 这种一致性可以通过跨视图知识蒸馏来实现。首先, 从同一类别的图像集合中采样 2 张不同的图像 ( $x_a, x_b$ ), 并生成对应的增强视图  $V_a$  和  $V_b$ 。随后, 通过模拟图像局部遮挡的场景对视图  $V_a$  施加块状掩码从而生成损坏的视图  $\hat{V}_a$ , 并将掩码区域替换为可学习向量集合。最后, 教师网络接收未掩码的视图  $V_b$ , 通过 ViT-EC 编码器生成 CLS 的投影分布 ( $P_{\text{CLS}}^t$ )。学生网络处理掩码视图  $\hat{V}_a$ , 通过 ViT-EC 编码器输出投影分布 ( $P_{\text{CLS}}^s$ )。公式 (1) 为基于交叉熵的监督对比损失 ( $\mathcal{L}_{\text{CLS}}$ ),  $C$  表示同一类别的内部交叉视图对。通过全局特征对齐, 学生网络会在 CLS 上学习到更准确的全局类别信息, 明显提升模型在小样本学习任务中的分类性能。

$$\mathcal{L}_{\text{CLS}} = -\frac{1}{|C|} \sum P_{\text{CLS}}^t(\hat{V}_a) \ln P_{\text{CLS}}^s(V_b). \quad (1)$$

**1.2.2 局部特征对齐** 在全局特征对齐的基础上进一步引入局部特征对齐策略, 有利于充分挖掘图像的细节信息。局部特征对齐任务假设: 同一类别的不同图像在局部细节上可能有所不同, 但仍有一些相似的语义信息, 通过在这些共享区域进行蒸馏, 模型能学习更具泛化能力的特征。教师网络从未损坏视图  $V_b$  中提取第  $j$  个分块 (Patch) 的嵌入特征 ( $F_j^t$ )。然后, 在学生网络的损坏视图  $\hat{V}_a$  中找到与其相似程度最高的第  $k$  个 Patch 嵌入特征 ( $F_k^s$ )。公式 (2) 用于计算同一类别的内部图像中的 Patch 相似度  $\text{Sim}(F_j^t, F_k^s)$ , 通过交叉熵损失对匹配成功的 Patch 对进行知识蒸馏。 $P$  表示所有匹配 Patch 对的集合,  $w_{jk}$  表示匹配

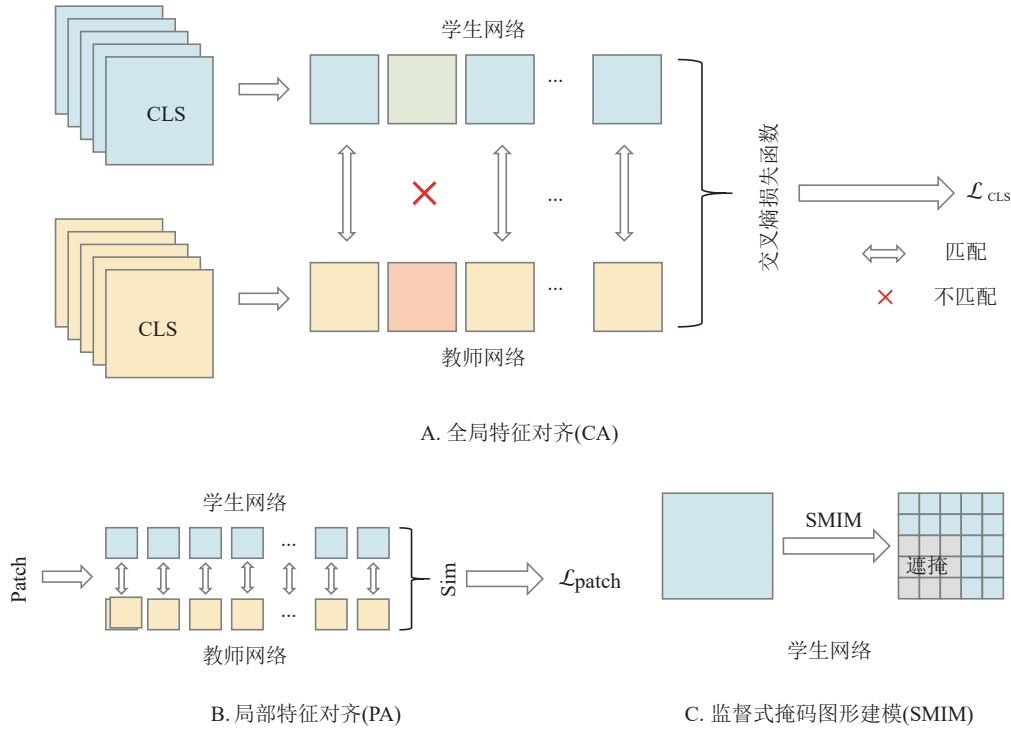


图2 增强型上下文感知知识蒸馏框架 (ECKD) 的3种视图策略

Figure 2 Three view strategies of enhanced context-aware knowledge distillation (ECKD)

Patch 的权重。最终通过公式 (3) 计算 Patch 的损失  $\mathcal{L}_{\text{Patch}}$ 。

$$\text{Sim}(F_j^t, F_k^s) = \frac{F_j^t F_k^s}{\|F_j^t\| \|F_k^s\|}; \quad (2)$$

$$\mathcal{L}_{\text{Patch}} = -\frac{1}{|P|} \sum_{j,k=1}^P w_{jk} F_j^t \ln F_k^s. \quad (3)$$

**1.2.3 监督式掩码图像建模** 监督式掩码图像建模利用图像的遮蔽增强,提升学生网络在不完整数据情况下处理特征的能力。输入学生网络的视图,通过随机选择图像块进行掩码,掩码比例在 0.1~0.5 之间随机选择,以增加任务的难度和多样性。具体来说,首先将输入图像  $X$  分割为块,并转换为嵌入向量序列,分块令牌表示为  $X = \{x_1, x_2, \dots, x_i, i = 1, 2, \dots, n\}$ ,再给定随机采样的掩码序列  $M$  进行掩码,  $|M|$  表示图像的分块数量,用共享的可学习令牌向量集合  $E$  替换这些掩码位置的原始嵌入,可学习令牌序列在位置编码添加后输入网络训练,模型基于未掩码块的上下文信息预测被掩码块的特征表示。损坏的图像  $\hat{X}$  包含未遮蔽和遮蔽的分块输入学生网络,  $\odot$  表示逐元素乘法,公式 (4) 用于计算损坏的图像  $\hat{X}$ 。接着使用  $P_i^s$  表示学生网络在掩码分块上的分类分布,使用  $P_i^t$  表示教师网络在未掩码图像上的分类分布,公式 (5) 通过教师网络和学生网络的分类分布计算交叉熵损失 ( $\mathcal{L}_{\text{SMIM}}$ )。最后,通过最小化  $\mathcal{L}_{\text{SMIM}}$  的损失值,使学生网络能够学习到更有意义的特征表示。

$$\hat{X} = X \odot (1 - M) + E \odot M; \quad (4)$$

$$\mathcal{L}_{\text{SMIM}} = -\frac{1}{|M|} \sum_{i=1}^M P_i^t \ln P_i^s. \quad (5)$$

### 1.3 ViT-EC 编码器

本研究提出一种基于多层感知机 (MLP) 和上下文转换模块 (CoT)<sup>[21]</sup> 的编码器优化思路,用于提升编码器在小样本学习场景下的特征表达能力和上下文建模能力。本研究在原始编码器中引入增强注意力残差模块 (EARM) 和 CoT 模块构建新型编码器 ViT-EC。传统编码器的 MLP 层升级为 MLP with EARM 结构,通过 SENetsqueeze-and-excitation (SENet) 通道注意力机制,在特征提取初期增强模型对重要特征的



表达能力；CoT 模块插入 ViT-EC 编码器后半部分用于深层语义优化，在抽象特征层面建立全局空间和局部空间关联，增强对目标物体整体结构的感知能力。在特征提取阶段，EARM 模块对输入特征标准化后，依次通过 SENet 通道注意力机制和增强型 MLP 处理，实现特征跨层传递时的动态校准。在模型的后续阶段，CoT 模块利用空间维度和通道维度构建三维注意力机制，实现静态与动态上下文的结合。CoT 模块首先通过  $1 \times 1$  卷积生成 Key Map 和 Value Map，并将原始特征投影为 Query 向量，利用自注意力机制计算特征块间的空间依赖权重，再与 Value Map 进行加权融合，最后通过  $1 \times 1$  卷积重组通道语义与原输入双重残差连接。图 3 为 ECKD 中 ViT-EC 编码器架构，EARM 模块在浅层网络完成通道维度的特征筛选，CoT 模块在深层网络构建空间和通道的联合注意力。ViT-EC 通过 EARM 模块和 CoT 模块的协同工作，不仅提升模型在小样本学习中的特征表达能力与上下文建模能力，同时还加快模型的训练速度和训练效率。

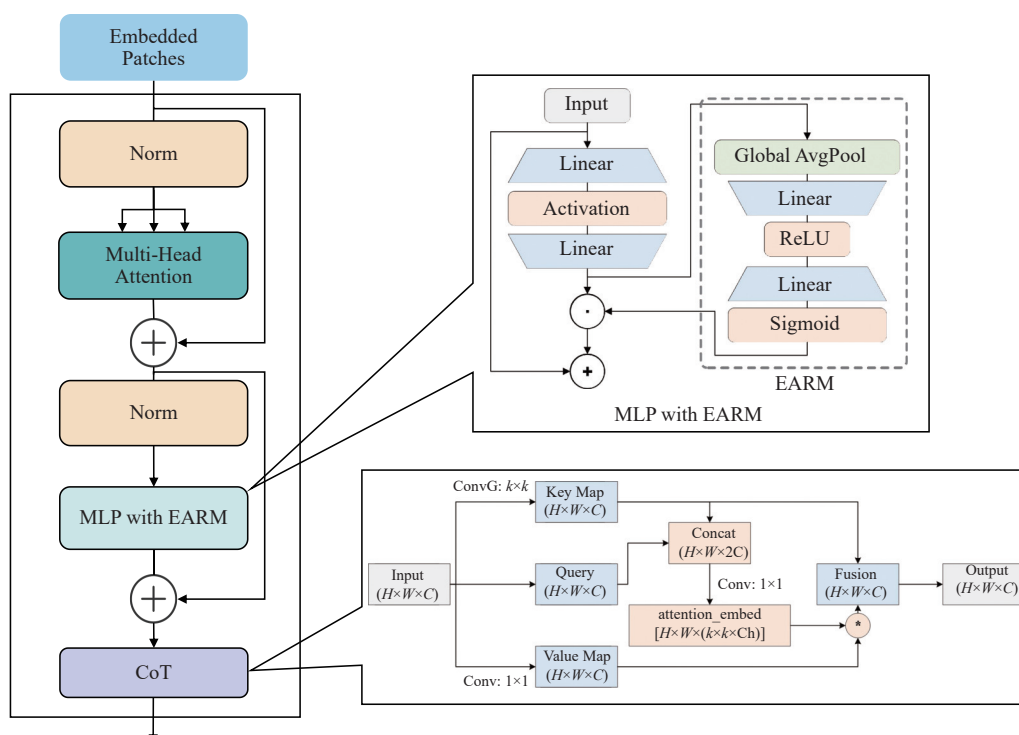


图 3 ViT-EC 编码器

Figure 3 ViT-EC encoder

#### 1.4 小样本评估分类器

本研究采用原型分类器进行实验分析。原型分类器通过采取支持集的每个类别特征的均值作为该类别的原型，进而计算查询样本与各类别原型之间的相似度，从而完成类别的分类(图 4)。首先从所有类别中随机选择若干类别，每个类别选取  $N$  个支持集样本 ( $n_{\text{sup}}$ ) 和查询集样本 ( $n_{\text{qy}}$ )，通过公式 (6) 计算支持集样本的特征向量 ( $z_{i,c}$ ) 的均值，并作为类别的原型 ( $p_c$ )，紧接着计算查询样本与所有类别原型之间的余弦相似度 [ $\text{CosSim}(z_q, p_c)$ ]，如公式 (7) 所示。最后，通过相似度最大化确定查询样本所属类别，并计算分类准确率。

$$p_c = \frac{1}{n_{\text{sup}}} \sum_{i=1}^{n_{\text{sup}}} z_{i,c}; \quad (6)$$

$$\text{CosSim}(z_q, p_c) = \frac{n_{\text{qy}} p_c}{\|n_{\text{qy}}\| \|p_c\|}. \quad (7)$$

#### 1.5 数据集

为验证所提方法的有效性，本研究选择公开的 PlantVillage 数据集进行植物病害识别的实验。PlantVillage 数据集对每幅图像都进行严格的筛选和标注，包含 54 303 幅植物叶片图像，涵盖 14 种植物和 26 种叶病，结合植物和叶病，形成 38 个类别(包含健康类别)。图 5 为 PlantVillage 数据集的图像类别个例，表 1 为 PlantVillage 数据集的所有作物类型和相关病害。按照数量比例为 6:4 随机划分数据集类

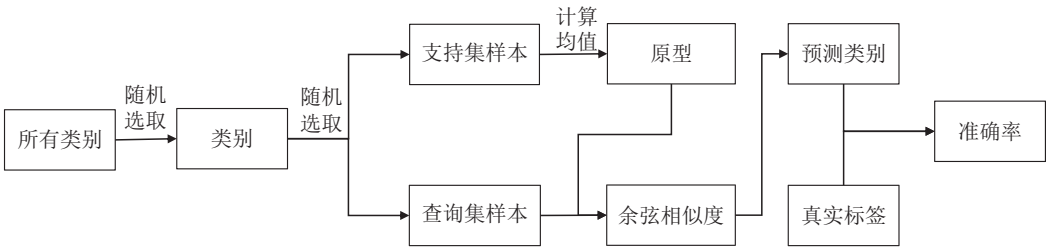


图 4 原型分类器评估流程

Figure 4 Evaluation process of prototype classifier

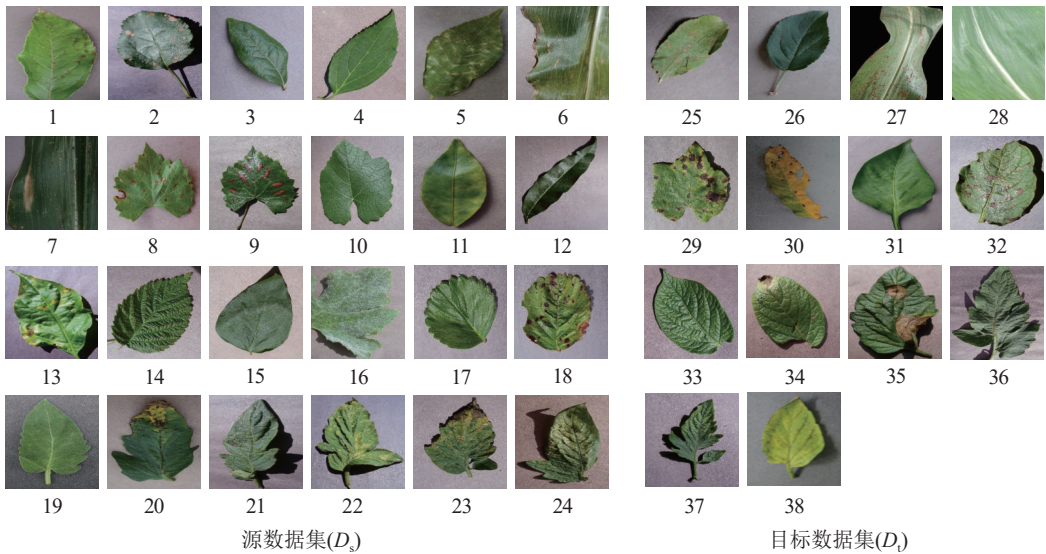


图 5 PlantVillage 数据集的图像类别个例

Figure 5 Examples of image categories in the PlantVillage dataset

表 1 PlantVillage 数据集图像类别信息

Table 1 Image category information of the PlantVillage dataset

植物	病害名称	数据集	序号	植物	病害名称	数据集	序号
苹果 <i>Malus pumila</i>	黑点病	$D_s$	1	番茄	早疫病	$D_s$	20
苹果	黑腐病	$D_s$	2	番茄	健康	$D_s$	21
蓝莓 <i>Vaccinium uliginosum</i>	健康	$D_s$	3	番茄	叶霉病	$D_s$	22
樱桃 <i>Cerasus pseudocerasus</i>	健康	$D_s$	4	番茄	斑枯病	$D_s$	23
樱桃	白粉病	$D_s$	5	番茄	二斑叶螨病	$D_s$	24
玉米 <i>Zea mays</i>	灰斑病	$D_s$	6	苹果	锈病	$D_t$	25
玉米	枯叶病	$D_s$	7	苹果	健康	$D_t$	26
葡萄 <i>Vitis vinifera</i>	黑腐病	$D_s$	8	玉米	锈病	$D_t$	27
葡萄	黑痘病	$D_s$	9	玉米	健康	$D_t$	28
葡萄	健康	$D_s$	10	葡萄	叶枯病	$D_t$	29
橘子 <i>Citrus reticulata</i>	黄龙病	$D_s$	11	桃子	细菌斑点病	$D_t$	30
桃子 <i>Prunus persica</i>	健康	$D_s$	12	胡椒	健康	$D_t$	31
胡椒 <i>Piper nigrum</i>	细菌斑点病	$D_s$	13	土豆 <i>Solanum tuberosum</i>	早疫病	$D_t$	32
树莓 <i>Rubus idaeus</i>	健康	$D_s$	14	土豆	健康	$D_t$	33
大豆 <i>Glycine max</i>	健康	$D_s$	15	土豆	晚疫病	$D_t$	34
南瓜 <i>Cucurbita moschata</i>	白粉病	$D_s$	16	番茄	晚疫病	$D_t$	35
草莓 <i>Fragaria ananassa</i>	健康	$D_s$	17	番茄	轮斑病	$D_t$	36
草莓	叶焦病	$D_s$	18	番茄	花叶病	$D_t$	37
番茄 <i>Solanum lycopersicum</i>	细菌斑点病	$D_s$	19	番茄	黄曲叶病	$D_t$	38

说明：序号所对应的植物病害见图5。 $D_s$ 表示植物病害的源数据集； $D_t$ 表示植物病害的目标数据集。

别, 其中 24 个类别的样本作为训练集, 剩余 14 个类别的样本作为评估集。

## 1.6 实验细节

本研究的所有工作均基于 PyTorch 框架实现, 所有训练和评估实验均在配备 CUDA 11.3 的 NVIDIA GTX-3090 GPU 上进行。CPU 处理器为 12th Gen Intel(R) Core(TM) i9-12900KF(3.20GHz), 搭配 32 GB 内存, batch size 为 16, 学习率初始设置为 0.000 5, 采用余弦退火调度, 优化器设置为 AdamW 优化器。软件环境为 64 位 Windows 10 专业版操作系统, Python 3.8 和 PyTorch 1.11。

## 2 结果与分析

### 2.1 预训练结果分析

本研究采用可视化手段对 ECKD 模型与基线 (Baseline) 模型的性能进行比较。图 6 为 ECKD 的预训练情况和不同视图策略学习时设置的学习率, 表明在训练的早期阶段, ECKD 能够更快地达到稳定状态, 在预训练时的效果比基线更好。图 7 为 ECKD 和基线在原型分类器上的准确率和训练轮次 (epoch) 的对比。在原型分类器方法验证中, ECKD 大约在 200 个训练轮次就可以达到基线的效果, 这比基线的预训练次数少 100 个训练轮次, 表明 ECKD 方法的特征提取方式在捕捉全局信息和关键细节方面更具优势。实验设置重复运行 10 次小样本分类评估, 计算每次运行的准确率均值和标准差。如图 8 所示: 在 2 类小样本任务评估中, 无论采用何种模型, 增加样本数量 (从 5 way 1-shot 提升至 5 way 5-shot) 都能带来明显的性能增益, 且在相同的样本条件下, ECKD 模型的表现始终优于基线模型 ( $P < 0.001$ )。

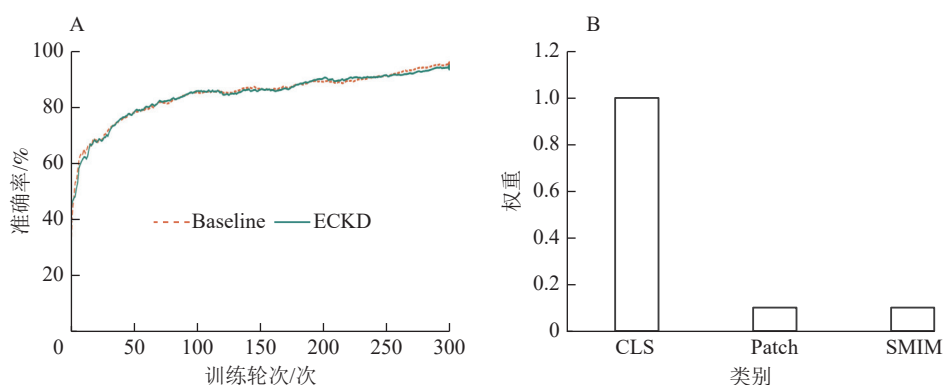


图 6 模型的预训练准确率 (A) 和不同损失函数的权重值 (B)

Figure 6 Training accuracy of ECKD and Baseline (A) and weight values of different loss functions (B)

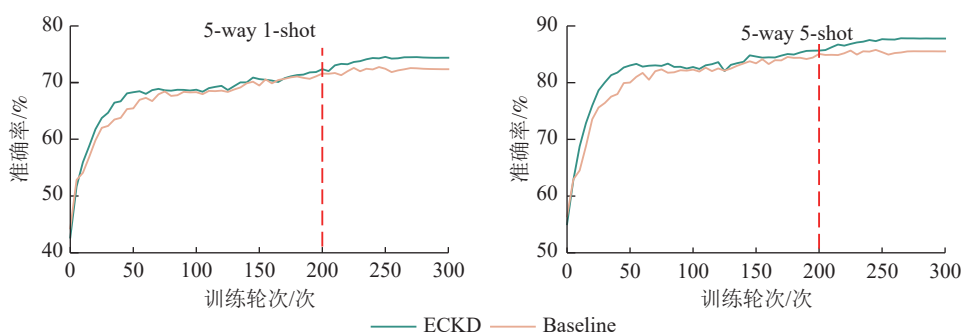


图 7 原型分类器评估方法在 ECKD 和 Baseline 模型中的准确率

Figure 7 Accuracy of prototype classifier evaluation method in ECKD and Baseline models

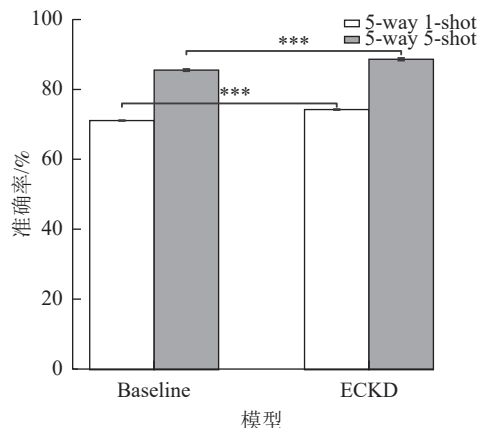
### 2.2 消融实验

本研究采用 ViT-S 作为基线, 编码器层数为 12 层, MLP 扩展比例为 4.0, 输入补丁大小为  $16 \times 16$ , 采用 GELU 激活函数, 所有实验均采用相同数据预处理与训练策略, 确保对比结果的公正性与可重复性。如表 2 所示: ViT-S 在引入 EARM 模块后, 模型准确率有小幅提升, 表明 EARM 模块对局部特征的强化有助于提高目标分类的效果。在保持主干网络不变的前提下, CoT 模块集成到 ViT 网络中进行

训练,结果显示 CoT 模块在加强全局语境和增强特征之间的交互方面发挥了关键作用。最后,将 EARM 模块和 CoT 模块结合构成新的编码器 ViT-EC,模型在 5-way 1-shot 和 5-way 5-shot 的小样本任务中准确率明显提升,表明 ViT-EC 在局部特征的精细表达与全局上下文的丰富性之间找到平衡,使得模型能够更好地关注关键特征。

### 2.3 与小样本方法比较

由表 3 可见:与现有的小样本学习方法相比,ECKD 明显提高了小样本学习任务中的性能。匹配网络和关系网络以及原型网络作为小样本基础方法<sup>[22]</sup>,虽然在 5-way 1-shot 和 5-way 5-shot 任务中取得不错的结果,但在复杂数据的特征表达和上下文建模能力上存在局限性。ECKD 基于更强大的 Vision Transformer 架构,相比最佳的基础方法,在 5-way 1-shot 任务中的识别结果达 74.98%,超越了小样本基础方法中的最佳结果。在 5-way 5-shot 任务中,ECKD 的识别结果达 88.28%,总体上与小样本基础方法保持持平。Iterative SS<sup>[23]</sup> 和 Mahalanobis<sup>[24]</sup> 方法在小样本任务上准确率仅达 34.00% 和 46.60%,难以应对复杂的小样本场景。CMSFF+CA<sup>[25]</sup> 和 Frequency+GC<sup>[26]</sup> 以及 FREN<sup>[27]</sup> 利用改进后的 ResNet12 作为特征提取器,在 5-way 1-shot 和 5-way 5-shot 任务中虽然取得较好的结果,但仍然难以充分建模小样本情境下的上下文关系和特征表达。基于 ResNet12 度量学习框架的 Prune-FSL<sup>[28]</sup> 虽然展示出良好的识别性能,但 Prune-FSL 缺乏对图像中结构性区域的显式建模,在处理复杂病斑语义关系、增强跨样本一致性及结构表征方面,ECKD 展现出更强的竞争力与实用性。PMF+FA 方法<sup>[29]</sup> 采用 ViT 作为主干网络,虽然表现出良好的识别性能,但未考虑病斑在空间结构层面的完整性与边缘连续性,ECKD 引入的 SMIM 策略通过掩码学习增强模型对局部结构区域的感知能力,从而在背景干扰严重时依然能有效聚焦病斑核心区域。虽然 Transformer 系列(PMF+FA 和 ECKD)在参数量方面略高,但其计算量与其他方法基本保持在同一量级,而在 5-way 5-shot 场景下,Transformer 系列对准确率的提升明显优于 CNN 系模型,说明 ViT-EC 在模型规模与计算代



\*\*\*表示同一样本数量不同模型之间差异显著( $P < 0.001$ )。

图 8 ECKD 与 Baseline 模型的  $t$  检验分析

Figure 8  $t$ -test analysis between ECKD and Baseline models

表 2 原型分类器消融实验

Table 2 Prototype classifier ablation experiment

模型	原型分类器准确率/%	
	5-way 1-shot	5-way 5-shot
ViT-S	72.17	85.62
ViT-S + EARM	73.39	86.30
ViT-S + CoT	73.31	86.96
ViT-S + EARM + CoT	74.98(+2.81)	88.28(+2.60)

说明: 括号内数值表示与 ViT-S 相比提升的准确率。

表 3 现有小样本学习方法的 5-way 1-shot 和 5-way 5-shot 实验对比

Table 3 Comparison of 5-way 1-shot and 5-way 5-shot experiment of existing few shot learning methods

方法	主干网络	参数量/M	计算量/G	准确率/%	
				5-way 1-shot	5-way 5-shot
匹配网络	Conv4	1.55	3.87	73.19	88.91
关系网络	Conv4	1.55	3.87	73.01	85.82
原型网络	Conv4	1.55	3.87	72.25	88.30
Iterative SS	CNN	6.75	3.20	34.00	53.10
Mahalanobis	ResNet18	24.88	3.60	46.60	63.50
CMSFF + CA	Resnet12	12.75	4.70	60.70	78.10
Frequency + GC	Resnet12	12.75	4.70	64.50	80.90
FREN	Resnet12	12.75	4.70	66.10	84.20
Prune-FSL	ResNet12	12.75	4.70	77.17	88.86
PMF+FA	ViT-S	21.67	4.24		86.79
ECKD	ViT-S	23.00	4.50	74.98	88.28



价可接受的前提下具备更强的特征建模能力，适合对性能有更高要求的小样本学习任务。从表 4 可见：ECKD 与基线模型均采用 ViT-S 作为骨干网络，相比 Baseline，ECKD 参数量仅增加约 6.0%，FLOPs 参数量增加约 6.1%，在不同样本数下均明显提升了性能。在 5-way 1-shot 和 5-way 5-shot 设置中，ECKD 的性能相较于基线分别提升 2.81% 和 2.66%，表明所提出的原型增强与特征蒸馏机制有效缓解了极少样本条件下的特征不足问题。随着样本数量的进一步增加，ECKD 在 5 way 10-shot 和 5 way 15-shot 设置下准确率分别达 90.97% 和 92.27%，呈现出良好的增长趋势，验证了所提方法在小样本学习场景中的有效性与实用价值。综合对比，ECKD 通过结合 Vision Transformer 的全局和局部上下文建模能力和知识蒸馏框架的特性增强机制，在小样本学习任务中展现出强大的性能优势。

表 4 ECKD 与 Baseline 模型在小样本分类任务中的性能对比

Table 4 Comparison of the few-shot classification performance between ECKD and the Baseline models							
模型	主干网络/M	参数量/G	计算量	准确率/%			
				5-way 1-shot	5-way 5-shot	5-way 10-shot	5-way 15-shot
Baseline	ViT-S	21.67	4.24	72.17	85.62	88.91	90.11
ECKD	ViT-S	23.00	4.50	74.98	88.28	90.97	92.27

3 结论

本研究提出了一种基于 Vision Transformer 和增强知识蒸馏的小样本学习框架 ECKD，通过设计新型编码器和多种视图策略显著提升了模型在小样本任务中的性能。结果表明：ECKD 在 PlantVillage 数据集上明显优于多种现有方法。本研究综合考虑全局特征和局部特征的相互影响，为小样本学习提供了新的思路和解决方案，在植物病害识别等实际应用中具有广泛的潜力。未来应重点探索先进的知识蒸馏策略，扩展数据集规模与多样性，并推进多模态小样本识别研究，进一步促进小样本学习技术在农业智能感知等实际场景中的应用与转化。

4 参考文献

[1] RISTAINO J B, ANDERSON P K, BEBBER D P, *et al.* The persistent threat of emerging plant disease pandemics to global food security[J/OL]. *Proceedings of the National Academy of Sciences of the United States of America*, 2021, **118**(23): e2022239118[2025-03-28]. DOI: [10.1073/pnas.2022239118](https://doi.org/10.1073/pnas.2022239118).

[2] WU Xue, DENG Hongyu, WANG Qi, *et al.* Meta-learning shows great potential in plant disease recognition under few available samples [J]. *The Plant Journal*, 2023, **114**(4): 767–782.

[3] 郑倩. 基于文献计量学和机器学习的小麦生物育种文献分析[J]. *浙江农林大学学报*, 2025, **42**(1): 210–217.  
ZHENG Qian. Literature analysis of *Triticum aestivum* bio-breeding based on bibliometrics and machine learning [J]. *Journal of Zhejiang A&F University*, 2025, **42**(1): 210–217.

[4] MOHANTY S P, HUGHES D P, SALATHÉ M. Using deep learning for image-based plant disease detection[J/OL]. *Frontiers in Plant Science*, 2016, **7**: 1419[2025-03-28]. DOI: [10.3389/fpls.2016.01419](https://doi.org/10.3389/fpls.2016.01419).

[5] 明浩, 苏喜友. 利用特征分割和病斑增强的杨树叶部病害识别[J]. *浙江农林大学学报*, 2020, **37**(6): 1159–1166.  
MING Hao, SU Xiyu. Image recognition of poplar leaf diseases with feature segmentation and lesion enhancement [J]. *Journal of Zhejiang A&F University*, 2020, **37**(6): 1159–1166.

[6] 王冠, 王建新, 孙钰. 面向边缘计算的轻量级植物病害识别模型[J]. *浙江农林大学学报*, 2020, **37**(5): 978–985.  
WANG Guan, WANG Jianxin, SUN Yu. Lightweight plant disease recognition model for edge computing [J]. *Journal of Zhejiang A&F University*, 2020, **37**(5): 978–985.

[7] BRAHIMI M, BOUKHALFA K, MOUSSAOUI A. Deep learning for tomato diseases: classification and symptoms visualization [J]. *Applied Artificial Intelligence*, 2017, **31**(4): 299–315.

[8] FERENTINOS K P. Deep learning models for plant disease detection and diagnosis [J]. *Computers and Electronics in Agriculture*, 2018, **145**: 311–318.

[9] LI Xiaoxu, YANG Xiaochen, MA Zhanyu, *et al.* Deep metric learning for few-shot image classification: a review of recent developments[J/OL]. *Pattern Recognition*, 2023, **138**: 109381[2025-03-28]. DOI: [10.1016/j.patcog.2023.109381](https://doi.org/10.1016/j.patcog.2023.109381).

- [10] JI Zhong, CHAI Xingliang, YU Yunlong, *et al.* Improved prototypical networks for few-shot learning [J]. *Pattern Recognition Letters*, 2020, **140**: 81–87.
- [11] BERTINETTO L, HENRIQUES J F, TORR P H S, *et al.* Meta-learning with differentiable closed-form solvers [C]//*Proceedings of the 7th International Conference on Learning Representations*. San Diego: ICLR, 2019.
- [12] ZHAO Kangkang, ZHANG Ziyang, JIANG Bo, *et al.* LGLNN: label guided graph learning-neural network for few-shot learning [J]. *Neural Networks*, 2022, **155**: 50–57.
- [13] HU Gensheng, WU Haoyu, ZHANG Yan, *et al.* A low shot learning method for tea leaf's disease identification[J/OL]. *Computers and Electronics in Agriculture*, 2019, **163**: 104852[2025-03-28]. DOI: [10.1016/j.compag.2019.104852](https://doi.org/10.1016/j.compag.2019.104852).
- [14] LI Yang, YANG Jiachen. Few-shot cotton pest recognition and terminal realization[J/OL]. *Computers and Electronics in Agriculture*, 2020, **169**: 105240[2025-03-28]. DOI: [10.1016/j.compag.2020.105240](https://doi.org/10.1016/j.compag.2020.105240).
- [15] GARG S, SINGH P. An aggregated loss function based lightweight few shot model for plant leaf disease classification [J]. *Multimedia Tools and Applications*, 2023, **82**(15): 23797–23815.
- [16] USKANER HEPSAĞ P. Efficient plant disease identification using few-shot learning: a transfer learning approach [J]. *Multimedia Tools and Applications*, 2024, **83**(20): 58293–58308.
- [17] BOULILA W. An approach based on performer-attention-guided few-shot learning model for plant disease classification [J]. *Earth Science Informatics*, 2024, **17**(4): 3797–3809.
- [18] THAKUR P S, KHANNA P, SHEOREY T, *et al.* Explainable Vision Transformer Enabled Convolutional Neural Network for Plant Disease Identification: PlantXViT[EB/OL]. 2022-07-16[2025-03-28]. <https://doi.org/10.48550/arXiv.2207.07919>.
- [19] WANG Yuzhi, YIN Yunzhen, LI Yaoyu, *et al.* Classification of plant leaf disease recognition based on self-supervised learning[J/OL]. *Agronomy*, 2024, **14**(3): 500[2025-03-28]. DOI: [10.3390/agronomy14030500](https://doi.org/10.3390/agronomy14030500).
- [20] CHEN Xuanchi, ZHENG Xiangwei, SUN Kai, *et al.* Self-supervised vision transformer-based few-shot learning for facial expression recognition [J]. *Information Sciences*, 2023, **634**: 206–226.
- [21] LI Yehao, YAO Ting, PAN Yingwei, *et al.* Contextual transformer networks for visual recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, **45**(2): 1489–1500.
- [22] 肖伟, 冯全, 张建华, 等. 基于小样本学习的植物病害识别研究[J]. *中国农机化学报*, 2021, **42**(11): 138–143.
- XIAO Wei, FENG Quan, ZHANG Jianhua, *et al.* Research on plant disease identification based on few-shot learning [J]. *Journal of Chinese Agricultural Mechanization*, 2021, **42**(11): 138–143.
- [23] LI Yang, CHAO Xuewei. Semi-supervised few-shot learning approach for plant diseases recognition[J/OL]. *Plant Methods*, 2021, **17**(1): 68[2025-03-28]. DOI: [10.1186/s13007-021-00770-1](https://doi.org/10.1186/s13007-021-00770-1).
- [24] NI Jing, YUAN Yichen, LI Yang, *et al.* Few-shot learning in intelligent agriculture: a review of methods and applications [J]. *Journal of Agricultural Sciences*, 2024, **30**(2): 216–228.
- [25] LIN Hong, TSE R, TANG S K, *et al.* Few-shot learning approach with multi-scale feature fusion and attention for plant disease recognition[J/OL]. *Frontiers in Plant Science*, 2022, **13**: 907916[2025-03-28]. DOI: [10.3389/fpls.2022.907916](https://doi.org/10.3389/fpls.2022.907916).
- [26] LIN Hong, TSE R, TANG S K, *et al.* Few-shot learning for plant-disease recognition in the frequency domain[J/OL]. *Plants*, 2022, **11**(21): 2814[2025-03-28]. DOI: [10.3390/plants11212814](https://doi.org/10.3390/plants11212814).
- [27] LIN Hong, QIANG Zhenping, TSE R, *et al.* A few-shot learning method for tobacco abnormality identification[J/OL]. *Frontiers in Plant Science*, 2024, **15**: 1333236[2025-03-28]. DOI: [10.3389/fpls.2024.1333236](https://doi.org/10.3389/fpls.2024.1333236).
- [28] YAN Wenbo, FENG Quan, YANG Sen, *et al.* Prune-FSL: pruning-based lightweight few-shot learning for plant disease identification[J/OL]. *Agronomy*, 2024, **14**(9): 1878[2025-03-28]. DOI: [10.3390/agronomy14091878](https://doi.org/10.3390/agronomy14091878).
- [29] REZAEI M, DIEPEVEEN D, LAGA H, *et al.* Plant disease recognition in a low data scenario using few-shot learning[J/OL]. *Computers and Electronics in Agriculture*, 2024, **219**: 108812[2025-03-28]. DOI: [10.1016/j.compag.2024.108812](https://doi.org/10.1016/j.compag.2024.108812).