

基于小样本学习的鳞翅目害虫图像识别方法

杨信廷^{1,2} 周子洁^{1,2} 李文勇^{2,3} 陈晓^{2,4} 王慧^{2,5} 于合龙¹

(1. 吉林农业大学信息技术学院, 长春 130118; 2. 北京市农林科学院信息技术研究中心, 北京 100097;
3. 国家农业信息化工程技术研究中心, 北京 100097; 4. 上海海洋大学信息学院, 上海 201306;
5. 山东农业大学信息科学与工程学院, 泰安 271018)

摘要: 针对面对害虫数据稀缺的实际场景时, 现有害虫图像识别方法容易出现过拟合导致模型表达能力不足的问题, 本研究提出了一种结合度量学习和迁移学习的小样本田间害虫图像分类识别方法。首先, 使用 ECA - Pyramid - ResNet12 模型在 mini - ImageNet 数据集上进行预训练; 其次, 在度量模块中添加 ECA 通道注意力机制, 通过捕捉通道间的依赖关系来增强害虫的图像特征表示; 然后, 使用特征金字塔结构来捕获害虫图像的局部特征和害虫的多尺度特征; 最后, 利用 20 类自建鳞翅目害虫图像作为元数据集, 对模型进行元训练和元测试。实验结果表明, 在 3 - way 5 - shot 和 5 - way 5 - shot 条件下, 本文模型准确率分别达到 91.16% 和 87.26%, 比 SSFormers、DeepBDC 方法分别提高 4.58、1.35 个百分点。提出的模型有效提升了小样本学习中目标图像特征的表达能力, 能够为数据稀缺场景下的田间害虫自动识别提供方法参考。

关键词: 害虫识别; 小样本学习; 迁移学习; 度量学习; 通道注意力机制; 金字塔全卷积网络

中图分类号: S24; TP39

文献标识码: A

文章编号: 1000-1298(2025)02-0402-09

OSID:



Image Recognition Method for Lepidoptera Pests Based on Few-shot Learning

YANG Xinting^{1,2} ZHOU Zijie^{1,2} LI Wenyong^{2,3} CHEN Xiao^{2,4} WANG Hui^{2,5} YU Helong¹

(1. College of Information Technology, Jilin Agricultural University, Changchun 130118, China

2. Information Technology Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China

3. National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China

4. College of Information Technology, Shanghai Ocean University, Shanghai 201306, China

5. School of Information Science and Engineering, Shandong Agricultural University, Tai'an 271018, China)

Abstract: In real-world scenarios where pest data is scarce, existing pest image recognition methods are prone to overfitting, resulting in insufficient model expressiveness. To address this issue, a novel few-shot field pest image classification method that integrated metric learning with transfer learning was proposed. Firstly, the ECA - Pyramid - ResNet12 model was pretrained on the mini - ImageNet dataset. Subsequently, PN was chosen as the classifier, and cosine similarity was selected as the distance metric. The ECA channel attention mechanism was then incorporated into the metric module to enhance pest image feature representation by capturing inter-channel dependencies, with a kernel size of 3. Additionally, a feature pyramid structure was employed to capture the local and multi-scale features of pest images. After evaluating different pooling combinations, the $2 \times 2 + 4 \times 4$ pooling combination was selected. Finally, a meta-dataset comprising 20 self-built categories of Lepidoptera pest images was utilized for meta-training and meta-testing of the model. Experimental results demonstrated that under 3 - way 5 - shot and 5 - way 5 - shot conditions, the proposed method achieved accuracy rates of 91.16% and 87.26%, respectively, surpassing the most relevant works of the past two years, SSFormers and DeepBDC, by 4.58 percentage points and 1.35 percentage points. The proposed model effectively enhanced the feature representation of target images in few-shot learning, providing a methodological reference for the automatic identification of field pests in data-scarce scenarios.

Key words: pest recognition; few-shot learning; transfer learning; metric learning; ECA; PyramidFCN

收稿日期: 2024-09-03 修回日期: 2024-09-23

基金项目: 国家重点研发计划项目(2022YFD2001801)和北京市农林科学院协同创新中心建设专项

作者简介: 杨信廷(1974—),男,研究员,主要从事农业信息化关键技术研究, E-mail: yangxt@nercita.org.cn

通信作者: 于合龙(1974—),男,教授,博士生导师,主要从事智慧农业和光学农业研究, E-mail: yuhelong@jlau.edu.cn

0 引言

农作物虫害是农业生产中的重大问题之一, 严重威胁我国粮食生产安全、农作物增产及农民收入^[1]。识别、监测和预警虫害有助于减少农药过度使用, 降低生产成本并保护生态环境, 对现代农业发展具有重要意义。传统的虫害识别主要依赖人工现场勘察, 存在主观性强、效率低等局限性^[2]。相比之下, 机器视觉识别方法在效率、稳定性和覆盖范围等方面具有显著优势^[1]。传统机器识别方法通常使用图像识别技术, 如 K-means 聚类^[3-4]、支持向量机^[5-6]等方法对害虫样本进行分类。然而, 由于传统方法的特征提取能力有限, 在处理复杂多变的田间环境图像时, 性能显著下降^[7]。

近年来, 深度学习技术因具备特征自主学习优势, 在农作物病虫害识别领域得到了广泛的研究与应用^[8-11]。然而, 由于深度学习数据驱动的特性, 当缺乏足够的标注数据时, 往往难以获得识别性能较好的模型^[12]。尤其是在复杂的农业种植环境下, 收集和标注数据的成本高昂, 构建大规模数据集非常困难。相反, 人类却能通过少量数据快速学习。受到这一点启发, 小样本学习 (Few-shot learning, FSL) 技术应运而生, 使深度学习更接近人类思维方式^[1, 13]。

小样本学习利用有限的训练数据来快速适应和泛化新任务, 从而实现高效的模型训练和准确预测^[14]。大部分研究^[15-20]主要针对骨干网络或者损失函数进行优化以提升特征的有效表达能力, 但难

以从根本上解决小样本学习中特征区分能力不足的问题。为此, 本文提出一种结合度量学习和迁移学习的田间害虫图像分类识别新方法。在度量模块中引入高效通道注意力机制^[21] (Efficient channel attention, ECA), 该机制能够捕捉通道间的依赖关系, 增强害虫图像特征表示, 从而基于类别间的差异性提升害虫图像的辨别能力。通过在主干特征提取网络 ResNet12 之后引入特征金字塔全卷积网络^[22] (Pyramid full convolution layer, PyramidFCN), 与 ECA 注意力机制结合来提取不同尺度的害虫特征与局部特征, 以期在小样本条件下实现鳞翅目虫害的精准识别。

1 材料与方法

1.1 数据采集与数据处理

1.1.1 数据采集

预训练数据: 在预训练阶段, 本文采用 mini-ImageNet 公开数据集对特征提取网络进行预训练。mini-ImageNet^[23] 广泛应用于小样本学习和元学习研究。该数据集从原始 ImageNet 中筛选出 100 个类别, 每个类别包含 600 幅彩色图像, 总计 60 000 幅图像。

元训练数据: 本文使用的自建害虫数据集涵盖 20 类鳞翅目成虫, 每类包含 100 幅图像, 共 2 000 幅图像。为确保数据的多样性, 图像数据通过田间相机拍摄和网络抓取 2 种方式获取。具体害虫类别及其示例如图 1 所示。



图 1 mini-ImageNet 部分类别和自建害虫数据集

Fig. 1 Selected categories from mini-ImageNet and self-built pest dataset

1.1.2 数据处理

预训练阶段: mini-ImageNet 按照标准分区方案^[23] 设置, 其中 64/16/20 类别分别用于训练、验证

和测试。预训练阶段采用颜色抖动、随机裁剪和随机水平翻转处理的数据增强策略, 整体样本数量保持不变。

元训练与元测试阶段:基于小样本分类的现有研究经验,将20类自建害虫数据集按照10/5/5类别以图中顺序进行分类,分别表示训练、验证和测试。元训练阶段采用随机裁剪(84×84)和随机水平翻转的数据增强方法。

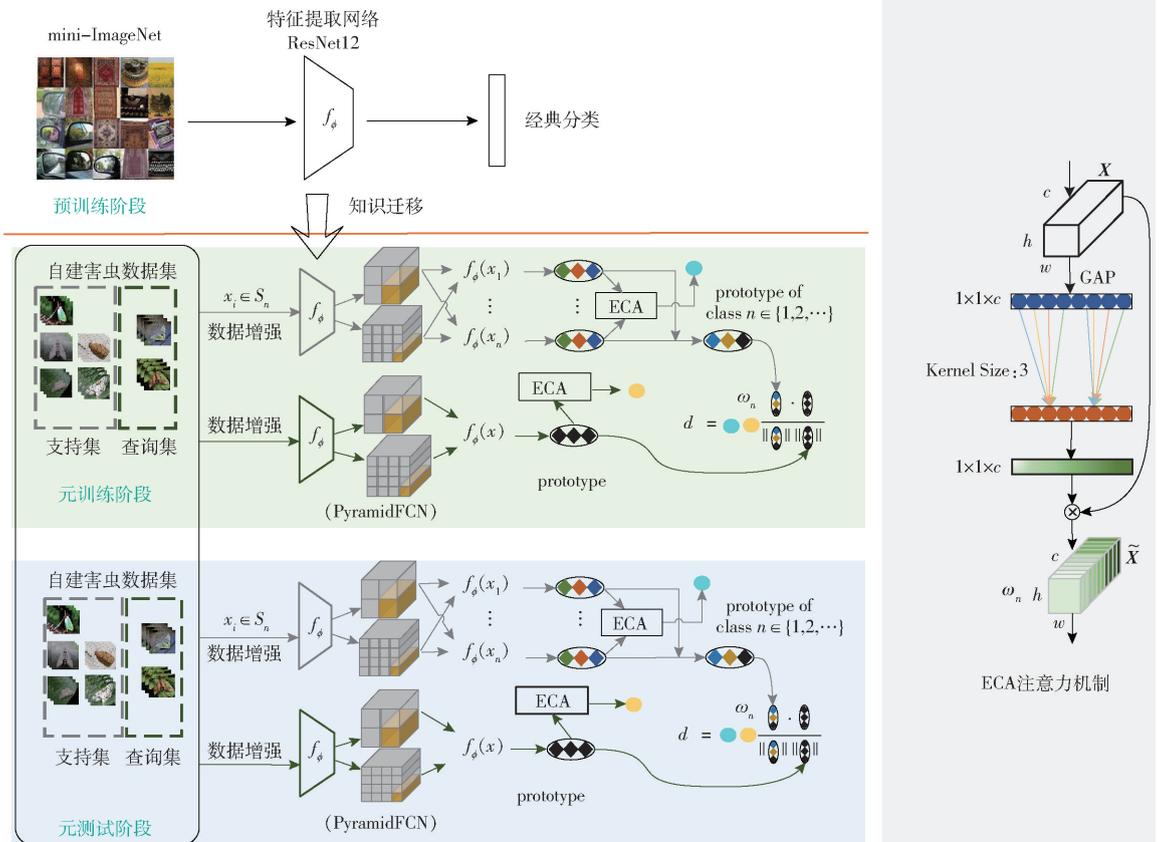


图2 ECA-Parasit-ResNet12 整体框架图

Fig. 2 ECA-Parasit-ResNet12 experimental framework diagram

PyramidFCN、ECA 注意力机制和分类器等。

在预训练阶段,利用 mini-ImageNet 数据集对特征编码器 ResNet12 进行训练,使模型能够学习到丰富且通用的特征表示^[24]。

在元训练阶段,害虫图像被划分为支持集和查询集两部分。与预训练阶段不同的是,采用特征增强处理后的害虫图像作为输入,并且对 ResNet12 后的全局特征进行 PyramidFCN 的池化组合操作以及将 ECA 注意力机制添加到原型网络(Prototypical network, PN)分类器^[25]中。在此过程中, N -ways K -shot 任务被优化以满足快速学习自建害虫数据集特征的需求。

在元测试阶段,模型具有与元训练阶段相同的网络结构。为了确保模型在新类别上的性能,模型会不断更新参数,适应新的特征分布,以便新类数据集不断快速学习新类别支持集害虫特征,从而对查询集害虫图像进行分类。

算法细节的伪代码为:

1.2 模型构建

本文融合迁移学习和度量学习构建基于小样本的害虫识别模型,其详细结构如图2所示,构建过程主要包括3个阶段:预训练、元训练和元测试。网络结构包含输入数据、特征编码器、金字塔全卷积网络

data; Support set (N -ways, K -shot)、query set (N -ways, q -shot)

Result: pest class

Start

Step 1: Pre-training

model ← ResNet12 (backbone) /* Initialise base */

Step 2: Meta-training

ECA ← feature attention network

PyramidFCN ← PyramidFCN (backbone)

While $i \leq \text{epoch}$ do

train_stats ← train (data_loader, model, criterion, optimizer, ...)

While $j \leq \text{batch}$ do

Support_set, Support_label, x, y ← batch;

Support_set ← Augmentation (Support_set);

Prototypes ← mean (PyramidFCN (model (Support_set)));

ω_n ← ECA (Support_set) /* for each class */

```

outopts ← model(x);
logits ← dECA(Prototypes, outputs, ωn);
loss ← CrossEntropy(logits, y);
optimize(loss.backward());
ACC ← evaluate(query_set, model);
if ACC ≥ ACC then
    save model
end
end
end
Step 3: Meta-test
model ← load(saved_model) /* Best version */
While t ≤ Steps do
    Support_set ← Augmentation(Support_set);
    Prototypes ← mean(PyramidFCN(model(Support_set)));
    ωn ← ECA(Support_set) /* for each class */
    Query_outputs ← model(Query set);
    logits ← dFA(Prototypes, Query_outputs, ωn);
    loss ← CrossEntropy(logits, y);
    optimize(loss.backward());
    ACC ← evaluate(Query_set, model);
end

```

1.2.1 预训练

在源域上训练深度网络并在目标域中运用已有

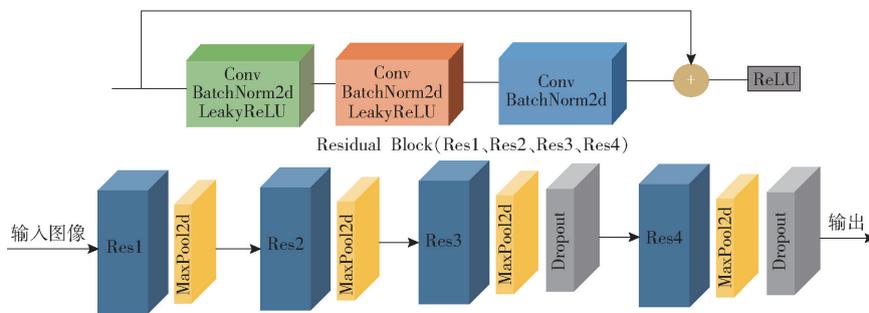


图3 ResNet12 框架图和 Residual Block (Res) 残差块

Fig. 3 ResNet12 framework diagram and Residual Block (Res) residual block

这种设计不仅使得网络更容易训练,还提高了模型的性能和稳定性。

各残差块中的卷积核数目(Num output)如表1所示。

1.2.2 元训练

元学习算法旨在学习跨任务的可迁移知识,使得预训练任务上学习的知识能够应用于仅使用少量训练数据的新任务中^[28-29]。在元训练阶段,本文采用了原型网络的分类方式。具体来说,通过计算查询样本与支持集样本之间的距离从而实现害虫图像的分类任务。在该阶段,模型通过元学习方法进行

的知识来学习新的知识称为迁移学习^[26]。在预训练阶段,本文选取 ResNet12^[22]作为网络结构以便模型学习 mini-ImageNet 数据集的特征。选用 ResNet12 作为主干网络的原因是,相比更深的网络(如 ResNet50 或 ResNet101),ResNet12 在小样本学习和迁移学习任务中表现出色且存储需求较低^[24,27]。

在预训练过程中,模型通过处理大量的 mini-ImageNet 数据集图像逐步优化其参数,从而形成对不同类别特征的深刻理解。预训练不仅提高了模型的泛化能力,而且使模型学会了学习。如图3所示,本文所采用的 ResNet12 特征提取网络包括4层,每层都包含1个残差块和1个2×2的最大池化层(MaxPool2d)用于下采样。为增强模型的泛化能力,第3层和第4层引入了 DropBlock 正则化机制。每个残差块由3个3×3卷积层(Conv3×3)、3个批归一化层(BatchNorm2d)和2个激活函数(LeakyReLU)组成。残差块核心在于学习输入与输出之间的残差关系,通过有参网络层来实现,计算公式为

$$F(x, \omega_i) = \omega_2(\omega_1 x) \quad (1)$$

式中 ω_1, ω_2 ——权重参数

x ——输入数据

ω_n ——一组权重参数

$F(x, \omega_i)$ ——经过网络处理后的输出

训练,旨在提高其在少样本学习任务中的表现。

在小样本学习中,引入了支持集和查询集的概念。具体而言,每个任务会从数据集 B 中随机抽取 N 个类别和 K 个样本作为支持集。随后,从这 N 个类别的剩余数据中抽取 q 个样本作为查询集。模型通过学习支持集中的 $N \times K$ 个样本,来识别查询集中的 $N \times q$ 个样本的类别。这种任务被称为 N -ways K -shot 任务。每个 episode 由支持集 $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ 和查询集 $Q = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ 构成,模型先在支持集上训练,然后在查询集上测试,这两步过程称为一个 task,计算

表1 各残差块中的卷积核数目

Tab.1 Numbers of convolution kernels in each

residual block		个
残差块	卷积层	卷积核数目
残差块 1	卷积层 1 (Conv3 × 3)	64
	卷积层 2 (Conv3 × 3)	64
	卷积层 3 (Conv3 × 3)	64
残差块 2	卷积层 1 (Conv3 × 3)	160
	卷积层 2 (Conv3 × 3)	160
	卷积层 3 (Conv3 × 3)	160
残差块 3	卷积层 1 (Conv3 × 3)	320
	卷积层 2 (Conv3 × 3)	320
	卷积层 3 (Conv3 × 3)	320
残差块 4	卷积层 1 (Conv3 × 3)	640
	卷积层 2 (Conv3 × 3)	640
	卷积层 3 (Conv3 × 3)	640

公式为

$$S = \{(x_i, y_i)\}_{i=1}^m \quad (m = NK) \quad (2)$$

$$Q = \{(x_i, y_i)\}_{i=1}^n \quad (n = Nq) \quad (3)$$

通过测算查询样本 q 和支持集样本之间的距离来判别查询样本的类别。这不仅简化了计算复杂度,还在保证性能的前提下,提高了模型在小样本情境下的适应能力。

原型网络是一种用于小样本学习的度量学习方法,其通过计算样本与类别中心(原型)之间的距离来实现分类。类原型公式为

$$C_n = \frac{1}{|S_n|} \sum_{(x_i, y_i) \in S_n} f_\phi(x_i) \quad (4)$$

式中 C_n ——backbone 特征嵌入层的 M 维输出向量均值, $C_n \in \mathbf{R}^M$

S_n ——样本集,包含所有属于类别 n 的样本

f_ϕ ——backbone 网络嵌入函数, $\mathbf{R}^D \in \mathbf{R}^M$

ϕ ——网络可学习参数

然后,通过距离函数 d 计算类原型与输入嵌入特征之间的相似性。softmax 函数将相似度分数转换为概率,对查询样本 q 生成一个类别分布。查询样本 q 被分配到类别 n 的概率计算公式为

$$P_\phi(y = N|x) = \frac{\exp(-d(f_\phi(x), C_n))}{\sum_{N'=1}^N \exp(-d(f_\phi(x), C_{N'}))} \quad (5)$$

训练过程中,使用随机梯度下降(SGD)算法来最小化交叉熵损失函数

$$L_{CE} = - \sum_{i=1}^N y_i \ln P_{\phi_i} \quad (6)$$

在元学习中,训练以任务为驱动而非数据为驱动,因为支持集和查询集作为一个任务整体输入网络,用以更新网络参数。这种方法独立于类别数量计算原型,因此可以应用于不同的 few-shot 设置。

1.2.3 元测试

在元学习中,在元训练阶段将数据集分解为不同的元任务,去学习类别变化的情况下模型的泛化能力,在 Meta test 阶段,面对全新的类别,不需要变动已有的模型,就可以完成分类^[28]。在元测试阶段,通过余弦相似度(Cosine similarity)来计算害虫图像识别准确率。

在小样本学习中,常用的距离函数包括欧氏距离(Euclidean distance)和余弦相似度^[13]。余弦相似度的值范围从 -1 到 1,值越接近 1,表示向量之间的夹角越小,相似度越高。

本文中测试结果使用准确率(Accuracy)指标进行评估。

1.2.4 ECA 注意力机制

本文在原型层中应用了一种新的特征级注意模块以提高整个模型的性能。图 2 所示的高效通道注意力 ECA^[21]是一种有效的通道注意力策略,通过捕捉通道间的依赖关系,增强卷积神经网络的特征表示。该模块主要用于提高神经网络在特征提取中的效果,尤其适用于图像分类任务。其核心思想是通过 1D 卷积高效地生成通道权重,从而捕捉不同通道间的局部交互信息。以下是 ECA 注意力机制的流程:

在给定主干网络情况下,小样本学习主干网络将输入变量 X 编码为嵌入特征向量。之后,ECA 注意力机制进行全局平均池化(GAP)以捕捉每个通道的全局上下文信息。GAP 的输出是一个具有 c 维度的向量,反映每个通道的平均响应。

通过 1D 卷积操作对全局特征向量进行处理, k 为 1D 卷积核大小。这个 1D 卷积操作帮助网络动态地确定每个通道重要性,从而更有效地捕捉局部通道间的相关性。1D 卷积核大小 k 可以根据具体任务和经验设置,通常为 3 或 5。选用 k 为 3。

一维卷积层输出可以表示为

$$\text{Conv1D}(\text{GAP}(X)) = F_{\text{conv1D}}(\text{GPA}(X)) \quad (7)$$

为了在模型中实现跨通道相关性的自适应学习,一维卷积层的输出通过 Sigmoid 激活函数进行非线性变换,生成每个通道权重向量 ω 为

$$\omega = \sigma(\text{Conv1D}(\text{GAP}(X))) \quad (8)$$

式中 σ ——Sigmoid 激活函数

最后,将生成的权重向量 ω 应用到输入特征图的每个通道上,通过逐元素乘法对每个通道的特征进行加权,重新调整每个通道的特征重要性,产生加权后的输出特征图 \tilde{X} 。在度量模块中,分类器基于每个输入的嵌入特征向量与类别原型之间的距离进行分类。

模型元训练、元测试过程中,除了主干网络 f_ϕ (表示为 ϕ) 的可训练参数外,ECA 注意力机制的权重系数也被添加到原型层的可训练参数中。这些参数通过随机梯度下降 (SGD) 算法进行优化,最小化小样本学习分类任务的交叉熵损失,从而提高模型的泛化能力和分类准确度。

1.2.5 特征金字塔结构 PyramidFCN

为了更全面地捕获图像的局部表示和多个尺度的变化,在 ResNet12 后面引入了一个被称为 PyramidFCN^[22] 的特征金字塔结构。它通过 RoI 池化技术处理 FCN 生成的特征图,并将得到的多尺度特征向量与原始特征向量相融合,形成更丰富的嵌入集,从而提升模型在不同尺度下对目标的感知和识别能力。PyramidFCN 使用了一个额外的自适应平均池化层来为每幅图像获取 20 个密集特征。

如图 4 所示,密集局部表示提取器 f_ϕ 通过不同的池化层将图像均匀地划分为 $h \times w$ 个区域,确保每个区域的特征能够被独立地提取并编码。主干网络 ResNet12 分别对每个图像区域进行编码以生成特征向量,这些特征向量综合起来构成了图像的整体表示。其中,输入图像的特征表示可以表达为 $f_\phi(\cdot) = R^{d \times C}$,其中 C 为通道数, d 为所有局部区域表示数量。采用 $2 \times 2 + 4 \times 4$ 的池化层组合,通过这种多尺度池化层组合,不仅能够捕获图像的全局信息,还能有效地捕捉到图像局部细节,从而提高模型性能。

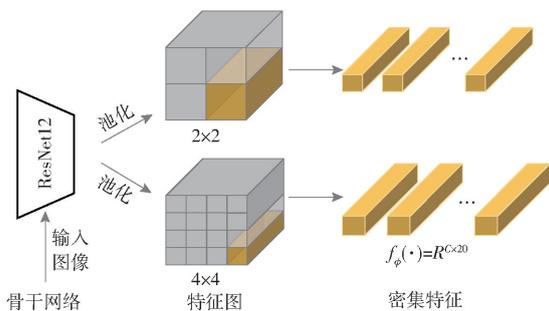


图 4 金字塔全卷积层

Fig. 4 Pyramid full convolution layer (PyramidFCN)

1.3 参数设置与实验环境

在预训练阶段,学习率设定为 0.1,批量大小为 128,训练周期 (epoch) 为 350。在元训练阶段,进行 6 轮训练,学习率为 1×10^{-5} ,训练时 episode 为 2 500,验证和测试时 episode 为 1 000。每个 episode 包含 10 个类别,每个类别在支持集中有 5 幅图像,在查询集中有 15 幅图像。

该方法在 Ubuntu 16.04 LTS 操作系统上运行,硬件配置包括: Intel® Core™ i9 - 9900K CPU @ 3.60 GHz \times 16,内存 62.8 GiB, NVIDIA GeForce RTX

2080 Ti GPU 10.78 GB, CUDA 10.2。使用 Python 3.8 和 PyTorch 1.12.0 深度学习框架。

2 结果

2.1 对比实验

为了验证模型在少样本害虫图像识别任务中的有效性,将本文方法与小样本经典方法和近年来的最先进方法在 5-way 5-shot 任务上进行比较。表 2 展示了不同方法在自建害虫数据集上的识别结果。在自建害虫数据集上,使用 ResNet12 骨干网络时,本文模型在 5-shot 设置下相比 Prototypical Net 准确率提高约 22.7 个百分点。相较而言, SemanticPrompt 在以 ViT 作为主干网络时准确率达到 57.39%,而本文模型在使用 ResNet12 时准确率达到 87.26%。值得注意的是,本文模型在自建害虫数据集上使用 ResNet12 骨干网络时,在 5-way 5-shot 设置中,比 SSFormers、DeepBDC 分别提高 4.58、1.35 个百分点。该性能提升的主要原因在于所提出方法能够识别与当前任务相关的图像区域,并基于层次化表示执行通道注意力机制。

表 2 自建害虫数据集上不同小样本学习方法对比结果

Tab. 2 Comparison of different few-shot learning methods on self-built pest dataset

方法	主干网络	5-way 5-shot 准确率/%
Prototypical Network ^[25]	Conv - 64F	53.79
Prototypical Network ^[25]	ResNet12	64.55
Matching Network ^[23]	BidirectionalLSTM	66.64
MTL ^[26]	ResNet12	79.92
MCL ^[22]	ResNet12	67.84
SemanticPrompt ^[30]	ViT	57.39
SSFormers ^[27]	Conv - 64F	62.35
SSFormers ^[27]	ResNet12	82.68
DeepBDC ^[31]	ResNet12	85.91
本文模型	ResNet12	87.26

图 5 为本文模型在 5-way 5-shot 任务中的 2 个关键参数变化曲线。由图 5 可知,在训练初始阶段,特别是前 1 至 3 个训练轮次内,损失值迅速下降,随后下降速率减缓并伴有轻微波动。尽管如此,整体趋势依然是下降,表明模型在训练数据集上的学习能力逐渐提升。准确率在前 4 个训练轮次内显著增加,之后增速减缓。这一趋势表明模型在验证数据集上的性能提升,且在未见过的数据上具有良好的泛化能力。因此可知,本文模型不仅能够有效学习训练数据,还具备良好的泛化能力。

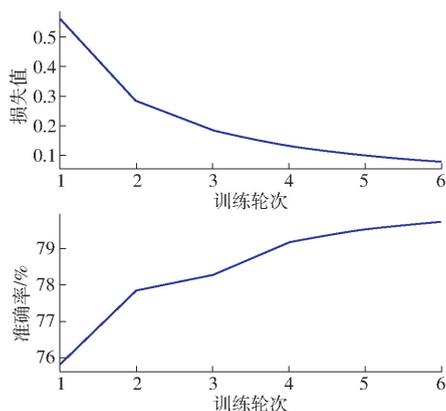


图5 训练过程中的损失曲线以及验证过程的准确率

Fig.5 Curves of loss during training and accuracy during validation

2.2 迁移学习

本文进一步评估了迁移学习与从零开始进行元训练2种方式下的模型性能。在迁移学习的设置中,首先在 mini-ImageNet 数据集上对 ResNet12 进行预训练。在此基础上,进入元训练和元测试阶段时,除 PyramidFCN、ECA 注意力机制和分类器的参数外,其他骨干网络的权重参数被迁移至小样本学习流程中。而在从零开始训练的设置中,采用一个随机初始化的网络代替预训练的网络,然后从零开始进行元训练。

对于 ResNet12 主干网络,使用相同的方法和超参数进行元训练和微调,以比较使用预训练网络的影响。迁移学习显著优于从零开始的训练。在 ResNet12 主干网络下,迁移学习准确率为 87.26%,而从零开始元训练的准确率为 61.27%。这表明迁移学习能够有效利用预训练模型的知识,从而显著提升模型在小样本学习任务中的性能表现。

2.3 度量方式

在元学习任务中,当使用 ResNet12 作为主干网络时,比较欧氏距离与余弦相似度两种度量方法的分类效果。采用欧氏距离度量准确率为 73.50%,而采用余弦相似度的准确率在元测试中显著提升至 85.47%。这可能是因为余弦相似度能够更有效地利用样本间的相似性进行分类,从而提升分类性能。

2.4 消融实验

2.4.1 ECA 注意力机制

ECA 注意力机制是全局平均池化、1D 卷积和 Sigmoid 激活函数的组合,实现了对每个通道的动态重要性评估和调整。为探讨 ECA 注意力机制对本文模型分类性能的影响,进行了消融实验研究。实验结果如表 3 所示,当 kernel 为 3 时,ECA 模块同时被添加到度量模块的支持集特征和查询集特征之后,模型表现出最佳分类性能,分类准确率显著高于

表3 ECA 注意力机制对模型识别精度的影响

Tab.3 Impact of ECA attention mechanism on model recognition accuracy

主干网络	ECA 注意力机制	5-way 5-shot 准确率/%
ResNet12	无	85.47
	kernel 为 3	85.93
	kernel 为 5	85.38

当 kernel 为 5 时的模型配置。结果表明,选取适当的 kernel 能够有效增强 ECA 模块对支持样本和查询样本特征的表示能力,从而基于类别间的差异性提升害虫图像的辨别能力。

2.4.2 特征金字塔结构 PyramidFCN

由于单一尺度的池化层可能无法充分捕捉图像的局部细节和上下文信息,因此采用不同的池化层组合(如 $2 \times 2 + 4 \times 4$ 、 $2 \times 2 + 3 \times 3 + 4 \times 4$ 等)来生成图像的局部表示,结果如表 4 所示。使用池化层组合为 2、3、4 的 ResNet12 在元测试准确率中达到 86.33%,相较于无池化结构提升 0.86 个百分点。由此可知,添加特征金字塔模块结构后,能够有效增强骨干网络特征提取的多尺度性能和局部细节捕捉能力。

表4 不同池化组合对模型识别精度的影响

Tab.4 Impact of different pooling combinations on model recognition accuracy

主干网络	池化大小	5-way 5-shot 准确率/%
ResNet12 (FCN)	无	85.47
	$2 \times 2 + 3 \times 3$	86.15
	$2 \times 2 + 4 \times 4$	86.15
	$2 \times 2 + 5 \times 5$	85.68
	$3 \times 3 + 4 \times 4$	86.26
ResNet12 (PyramidFCN)	$3 \times 3 + 5 \times 5$	85.45
	$4 \times 4 + 5 \times 5$	86.09
	$2 \times 2 + 3 \times 3 + 4 \times 4$	86.33
	$2 \times 2 + 3 \times 3 + 5 \times 5$	85.60
	$3 \times 3 + 4 \times 4 + 5 \times 5$	85.95

2.4.3 ECA 和 PyramidFCN

表 5 展示了在不同配置下,PN 分类器在 5-way 5-shot 任务中的识别准确率。当仅使用 PN 分类器时,准确率为 85.47%。引入 PyramidFCN(2,3,4)后,准确率提升至 86.33%。当仅使用 ECA 注意力机制($k=3$)时,准确率为 85.93%。

尽管单独使用 PyramidFCN(2,4)时未达到最佳效果,但当 PyramidFCN(2,4)和 ECA 注意力机制($k=3$)结合使用时,准确率达到最高值 87.26%。结果表明,PyramidFCN 和 ECA 注意力机制的结合能够在一定程度上解决害虫特征提

表 5 PN 分类器在不同配置下的准确率

Tab. 5 Recognition accuracy for PN classifier in different configurations

配置	5-way 5-shot 准确率/%
PN	85.47
PN + PyramidFCN(2,3,4)	86.33
PN + ECA($k=3$)	85.93
PN + PyramidFCN(2,3,4) + ECA($k=3$)	86.84
PN + PyramidFCN(2,4) + ECA($k=3$)	87.26

取、分类不够精确的问题,展现出其在害虫识别任务中的潜力与有效性。

2.4.4 way 和 shot 对结果的影响

为了评估模型在不同 way 和 shot 设置下的识别能力,进行了多组实验。具体设置包括 way 值为 3 和 5,以及 shot 值分别为 1、5、10、15、20 来观察识别准确率的变化。测试结果如图 6 所示。

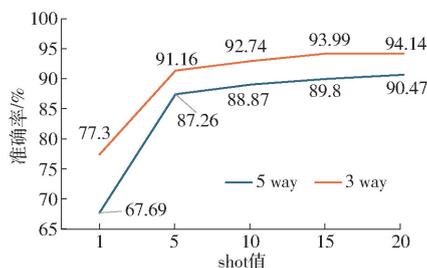


图 6 不同 shot 值下本文模型测试结果

Fig. 6 Test results of proposed model under different shot values

从图 6 可以看出,在 way 值相同的情况下,随着 shot 值的增加,模型识别准确率显著提高,尤其是在 shot 值较小时提升幅度较大。这表明,在相同 way 条件下,shot 值与识别效果呈正相关关系,但这种关系在较高的 shot 值时会趋于稳定。

当保持 shot 值为 5,分别设置 way 值从 2 到 5,并在自建害虫数据集上使用本文模型进行识别测试,结果如图 7 所示。图中显示,在相同 shot 条件下,随着 way 值的增加,模型识别精度呈平缓下降趋

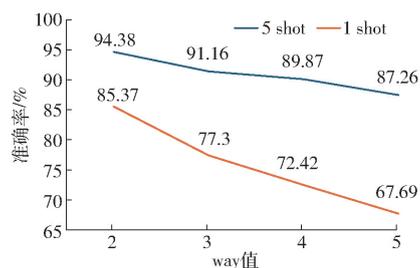


图 7 不同 way 值下本文模型测试结果

Fig. 7 Test results of proposed model under different way values

势,说明在固定 shot 值情况下,way 值与识别效果呈负相关关系。

综上所述,实验结果表明,当固定 way 值时,增加 shot 值有助于提升模型识别性能;而在固定 shot 值时,增加 way 值会导致识别效果下降。因此,在害虫图像数据极度稀缺情况下,为提高识别精度可以考虑在构建识别模型时选取较小的 way 值。

3 结论

(1) 针对害虫图像数据稀缺场景,提出了一种融合 ECA- Pyramid- ResNet12 结构的害虫图像识别方法,该方法结合迁移学习和度量学习两种技术,以 ResNet12 作为骨干网络,并在度量模块中引入 ECA 注意力机制以及在 FCN 全卷积网络中引入特征金字塔结构,从而有效缓解了害虫特征提取不精确所导致的识别精度下降问题,提升了模型性能。实验结果表明,在 3-way 5-shot 以及 5-way 5-shot 条件下,该方法准确率分别为 91.16%、87.26%,相较于 SSFormers、DeepBDC 方法,分别提高 4.58、1.35 个百分点,验证了其在数据稀缺场景中的有效性。

(2) 实验验证了迁移学习的有效性,相较于从零开始训练的模型,本文模型显著提升了性能。本文方法不仅有效应对了实际场景中害虫数据稀缺问题,还提升了鳞翅目害虫识别准确率和效率,为农业害虫识别提供了一个可行的解决方案。

参 考 文 献

- [1] 翟肇裕, 曹益飞, 徐焕良, 等. 农作物病虫害识别关键技术研究综述[J]. 农业机械学报, 2021, 52(7): 1-18. ZHAI Zhaoyu, CAO Yifei, XU Huanliang, et al. Review of key techniques for crop disease and pest detection[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 1-18. (in Chinese)
- [2] 张萌, 钱蓉, 朱静波, 等. 农作物害虫图像识别研究进展与展望[J]. 安徽农业科学, 2018, 46(34): 11-12, 15. ZHANG Meng, QIAN Rong, ZHU Jingbo, et al. Research progress and prospect of crop pest image recognition[J]. J. Anhui Agric. Sci., 2018, 46(34): 11-12, 15. (in Chinese)
- [3] 赵新苗, 冯向萍, 赵涛. 农业信息分类中 K-means 与 SVM 的混合算法研究[J]. 计算机技术与发展, 2017, 27(6): 178-182. ZHAO Xinmiao, FENG Xiangping, ZHAO Tao. Investigation on K-means and SVM mixed algorithm for agriculture information classification[J]. Computer Technology and Development, 2017, 27(6): 178-182. (in Chinese)
- [4] GARCÍA J, POPE C, ALTIMIRAS F. A distributed K-means segmentation algorithm applied to lobesia botrana recognition[J]. Complexity, 2017(1): 5137317.
- [5] PATTNAIK G, PARVATHI K. Machine learning-based approaches for tomato pest classification [J]. Telkommika

- (Telecommunication Computing Electronics and Control), 2022, 20(2): 321 – 328.
- [6] 李文勇, 李明, 陈梅香, 等. 基于机器视觉的作物多姿态害虫特征提取与分类方法[J]. 农业工程学报, 2014, 30(14): 154 – 162.
LI Wenyong, LI Ming, CHEN Meixiang, et al. Feature extraction and classification method of multi-pose pests using machine vision[J]. Transactions of the CSAE, 2014, 30(14): 154 – 162. (in Chinese)
- [7] YU X, ZHENG J. A review on the application of deep learning methods in detection and identification of rice diseases and pests [J]. Computers, Materials & Continua, 2024, 78(1): 197 – 225.
- [8] 翟肇裕, 张梓涵, 徐焕良, 等. YOLO 算法在动植物表型研究中应用综述[J]. 农业机械学报, 2024, 55(11): 1 – 20.
ZHAI Zhaoyu, ZHANG Zihan, XU Huanliang, et al. Review of applying YOLO family algorithms to analyze animal and plant phenotype[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 1 – 20. (in Chinese)
- [9] ALI F, QAYYUM H, IQBAL M J. Faster – PestNet: a lightweight deep learning framework for crop pest detection and classification[J]. IEEE Access, 2023, 11: 104016 – 104027.
- [10] ZHENG T, YANG X, LV J, et al. An efficient mobile model for insect image classification in the field pest management[J]. Engineering Science and Technology, an International Journal, 2023, 39: 101335.
- [11] ZHANG S, WANG H, ZHANG C, et al. JutePest – YOLO: a deep learning network for jute pest identification and detection [J]. IEEE Access, 2024, 12: 72938 – 72956.
- [12] RAGU N, TEO J. Object detection and classification using few-shot learning in smart agriculture: a scoping mini review[J]. Frontiers in Sustainable Food Systems, 2023, 6: 1039299.
- [13] 赵凯琳, 靳小龙, 王元卓. 小样本学习研究综述[J]. 软件学报, 2021, 32(2): 349 – 369.
ZHAO Kailin, JIN Xiaolong, WANG Yuanzhuo. A survey of few-shot learning[J]. Journal of Software, 2021, 32(2): 349 – 369. (in Chinese)
- [14] LIN H, TSE R, TANG S K, et al. The positive effect of attention module in few-shot learning for plant disease recognition[C]// 2022 5th International Conference on Pattern Recognition and Artificial Intelligence (PRAI). IEEE, 2022: 114 – 120.
- [15] LI Y, YANG J. Few-shot cotton pest recognition and terminal realization[J]. Computers and Electronics in Agriculture, 2020, 169: 105240.
- [16] LI Y, YANG J. Meta-learning baselines and database for few-shot classification in agriculture[J]. Computers and Electronics in Agriculture, 2021, 182: 106055.
- [17] PANDEY S, SINGH S, TYAGI V. Meta-learning for few-shot insect pest detection in rice crop[C]// International Conference on Advances in Computing and Data Sciences. Cham: Springer International Publishing, 2022: 404 – 414.
- [18] YANG Z, YANG X, LI M, et al. Small-sample learning with salient-region detection and center neighbor loss for insect recognition in real-world complex scenarios[J]. Computers and Electronics in Agriculture, 2021, 185: 106122.
- [19] GOMES J C, BORGES D L. Insect pest image recognition: a few-shot machine learning approach including maturity stages classification[J]. Agronomy, 2022, 12(8): 1733.
- [20] LIANG Weijian, GUO Qingwen, WANG Chuntao, et al. Few-shot pest classification using spatial-attention-enhanced ResNeSt – 101 network and transfer-based meta-learning[J]. Transactions of the CSAE, 2024, 40(6): 285 – 297.
- [21] WANG Q, WU B, ZHU P, et al. ECA – Net: efficient channel attention for deep convolutional neural networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11534 – 11542.
- [22] LIU Y, ZHANG W, XIANG C, et al. Learning to affiliate: mutual centralized learning for few-shot classification[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 14411 – 14420.
- [23] VINYALS O, BLUNDELL C, LILLICRAP T, et al. Matching networks for one shot learning[J/OL]. Advances in Neural Information Processing Systems, 2016. <https://arxiv.org/abs/1606.04080>.
- [24] ZHANG C, CAI Y, LIN G, et al. Deepemd: few-shot image classification with differentiable earth mover’s distance and structured classifiers[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 12203 – 12213.
- [25] SNELL J, SWERSKY K, ZEMEL R. Prototypical networks for few-shot learning[J]. Advances in Neural Information Processing Systems, 2017, 30: 4077 – 4087.
- [26] SUN Q, LIU Y, CHUA T S, et al. Meta-transfer learning for few-shot learning[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 403 – 412.
- [27] CHEN H, LI H, LI Y, et al. Sparse spatial transformers for few-shot learning[J]. Science China Information Sciences, 2023, 66(11): 210102.
- [28] 黄佳丹, 张志强. 小样本学习研究现状综述[J]. 信息系统工程, 2022(2): 69 – 72.
- [29] 蒋心璐, 陈天恩, 王聪, 等. 农业害虫检测的深度学习算法综述[J]. 计算机工程与应用, 2023, 59(6): 30 – 44.
JIANG Xinlu, CHEN Tianen, WANG Cong, et al. Review of deep learning algorithms for agricultural pest detection[J]. Computer Engineering and Applications, 2023, 59(6): 30 – 44. (in Chinese)
- [30] CHEN W, SI C, ZHANG Z, et al. Semantic prompt for few-shot image recognition[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 23581 – 23591.
- [31] XIE J, LONG F, LV J, et al. Joint distribution matters: deep brownian distance covariance for few-shot classification[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 7972 – 7981.