

DOI: 10.13733/j.jcam.issn.2095-5553.2025.01.029

高天赐, 王克俭, 陈晨, 等. 基于 DCP-ShuffleNetV2 的轻量级森林害虫识别方法[J]. 中国农机化学报, 2025, 46(1): 190-197

Gao Tianci, Wang Kejian, Chen Chen, et al. Identification method of lightweight forest pest based on DCP-ShuffleNetV2 [J]. Journal of Chinese Agricultural Mechanization, 2025, 46(1): 190-197

基于 DCP-ShuffleNetV2 的轻量级森林害虫识别方法^{*}

高天赐^{1, 2}, 王克俭^{1, 2}, 陈晨^{1, 2}, 韩宪忠^{1, 2}, 王超^{1, 2}, 李会平^{2, 3}

(1. 河北农业大学信息科学与技术学院, 河北保定, 071001; 2. 河北省城市森林健康技术创新中心, 河北保定, 071001; 3. 河北农业大学林学院, 河北保定, 071001)

摘要:针对现有害虫识别模型复杂度高、计算量和参数量巨大的问题,提出一种基于 DCP-ShuffleNetV2 的轻量级森林害虫识别模型。该模型主要从特征提取、特征融合、轻量化方面进行改进。首先通过引入金字塔分割注意力模块 PSA 提取多尺度的空间信息和跨通道依赖关系,有效地学习上下文信息;其次将基准网络模型 ShuffleNetV2 的 Stage 模块修改为 CSP 结构,增强特征融合能力;将模型的普通卷积替换为动态卷积,压缩模型参数量和计算量。试验以雄安新区“千年秀林”害虫为研究对象,构建 30 类常见害虫数据集。结果表明,改进后的 DCP-ShuffleNetV2 模型在自制的 Forest30 数据集上的害虫识别准确率是 92.43%,模型参数量、计算量和内存大小分别是 0.13 M、24.53 M 和 9.53 MB,相比于基准网络模型,识别准确率提升 3.11%,参数量、计算量和内存大小分别减少 62.83%、42.48%和 15.13%。与目前常用的分类模型相比,识别准确率平均提高 5.39%,模型参数量、计算量和内存大小平均减小 14.32 M、1 035.80 M 和 35.98 MB。

关键词:害虫识别;DCP-ShuffleNetV2;注意力机制;CSP 结构;特征提取

中图分类号:S763; TP391.4 **文献标识码:**A **文章编号:**2095-5553 (2025) 01-0190-08

Identification method of lightweight forest pest based on DCP-ShuffleNetV2

Gao Tianci^{1, 2}, Wang Kejian^{1, 2}, Chen Chen^{1, 2}, Han Xianzhong^{1, 2}, Wang Chao^{1, 2}, Li Huiping^{2, 3}

(1. College of Information Science and Technology, Hebei Agricultural University, Baoding, 071001, China;
2. Hebei Provincial Innovation Center of Urban Forest Health Technology, Baoding, 071001, China;
3. College of Forestry, Hebei Agricultural University, Baoding, 071001, China)

Abstract: This paper proposes a pest recognition model based on lightweight DCP-ShuffleNetV2 to solve the problems of high complexity of model, large amount of computation and reference. The model is improved from feature extraction, feature fusion and lightweight. Firstly, to learn context information effectively, the Pyramid Split Attention (PSA) module is introduced to extract multi-scale spatial information and cross-channel dependency. Secondly, to enhance the feature fusion capability, the Stage module of the benchmark network model ShuffleNetV2 is modified to Cross Stage Partial (CSP) structure. Finally, to compress the number of parameters and computation, the regular convolution is replaced by dynamic convolution for the model. In the experiment, a data set of Forest 30 was constructed in the “Millennium Xiulin” of Xiong’an New Area. The experimental results show that the pest identification accuracy of the DCP-ShufflenetV2 model is 92.43%, and the number of parameters, computation amount and memory size of the improved model are 0.13 M, 24.53 M and 9.53 MB, respectively. Compared with the ShufflenetV2 network model, the pest identification accuracy of the improved model increased by 3.11%, and the reference number, computation amount and memory size were reduced by 62.83%, 42.48% and 15.13%, respectively. Compared with the current commonly used classification model, the average recognition accuracy is increased by 5.39%, the number of parameters, computation amount and memory size of the improved model are reduced by 14.32 M, 1 035.80 M and 35.98 MB on average.

收稿日期:2023 年 4 月 13 日 修回日期:2023 年 6 月 5 日

^{*} 基金项目:国家自然科学基金项目(32171799);河北省林业和草原局科学计划项目(2001023)

第一作者:高天赐,男,1998 年生,河北保定人,硕士研究生;研究方向为农业领域图像识别。E-mail: 1137618513@qq.com

通讯作者:王克俭,女,1971 年生,河北保定人,硕士,教授;研究方向为农业信息化。E-mail: wkj71@163.com

Keywords: pest identification; DCP-ShuffleNetV2; attention mechanism; CSP structure; feature extraction

0 引言

我国林业有害生物全年发生面积达 12 784.5 km²，为近 10 年发生面积最大，其中，森林虫害占比 61.84%，林业虫害一旦发生将严重危害林业的健康发展^[1]。因此，科学防治虫害，精准分类、精准消杀对林业生态系统十分重要，关系生态系统的平衡和稳定^[2]。

图像识别算法可分为传统的图像识别方法^[3]和基于深度学习的图像识别方法^[4]。传统的图像识别方法首先将采集到的害虫数据进行预处理，然后提取害虫的图像特征，比如纹理、形状、大小和颜色等，提取方法包括小波域变换、局部二值模式 LBP、颜色共生矩阵 CCM 和灰度共生矩阵 GLCM 等，然后将提取的特征送入训练好的机器学习模型进行分类识别^[5]。肖志云等^[6]提取马铃薯害虫的小波域纹理特征和空间域颜色特征以及形状特征组合成特征向量，使用 SVM 分类器进行分类，相比传统纹理特征提取方法，在特征计算量不增加的同时，平均识别率提高了 17 个百分点。邹修国等^[7]将稻飞虱害虫的图像进行灰度化处理之后，使用数学形态学滤波，通过 4 种不变矩提取特征值后进行分类，平均识别率为 91.7%。上述方法在特征提取过程中过度依赖算法设计者的主观判断，手动特征的选择往往依赖于数据集，当数据来源发生变化时，特征可能需要重新设计^[8]。

相比于传统的机器学习算法，随着数据量的爆炸式增长和计算机性能的提升，深度学习算法受到学者青睐，通过使用卷积神经网络 CNN 自动提取图像底层特征，以端到端的方式自动完成学习，具有更高的鲁棒性^[9]。孔建磊等^[10]提出了一种多流高斯概率融合网络，挖掘害虫细粒度特征，对自制的 181 种病虫害的平均精度高达 93.81%。Wei 等^[11]提出了一种基于多尺度特征融合(MFFNet)的农作物害虫识别方法，对于 12 种农作物害虫的分类准确率达到 98.2%。Khanramaki 等^[12]集成了 AlexNet、VGG16、ResNet50，使用迁移学习进行预训练，利用集成学习投票法得到最终结果，试验表明，对于三种常见柑橘害虫的准确率达到 99.04%，优于其他 CNN 方法。Yang 等^[13]通过较小的计算成本有效地改进 SqueezeNet 模型，在数据集 IP102 上的害虫识别精度比原始模型提高 2.3%，并部署在移动端。

尽管上述深度学习算法在害虫识别任务中有较好的表现，但是考虑到林业害虫识别的应用场景大多对嵌入式或者移动设备的性能要求比较高，硬件资源受限，但目前大量神经网络模型参数量大，计算量高，因此需

要在保障准确率的情况下尽量压缩模型的复杂度^[14]。另一方面在自然环境中多数害虫的保护色与背景环境极其相似，并且类间差异小，模型需要在众多输入信息中聚焦关键信息，并降低其他信息的关注度。因此，如何在保证准确率的情况下，设计轻量化卷积神经网络模型应用到实际环境中已经成为亟待解决的问题。

针对这些问题，本文提出一种基于 DCP-ShuffleNetV2 的轻量级害虫识别方法，能够在保持较高的识别准确率的情况下进一步减小模型参数量和计算量。

1 试验数据获取与处理

1.1 害虫图像数据集制作

由于缺乏样本，深度学习在林业害虫识别中的应用受到严重限制，构建 30 类常见林业害虫数据集 Forest30。图像采集地点为雄安新区“千年秀林”（39.005 156°N, 116.047 255°E），在自然光环境下使用自主开发的 APP“雄安新区千年秀林生物调查系统”共采集 30 类害虫，各类别约 100~150 张并自动上传至服务器，如图 1 所示。

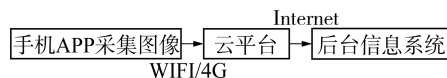


图 1 数据采集流程

Fig. 1 Data acquisition process

数据采集人员共分为 5 组，通过使用不同型号移动设备来提高数据集的多样性。采集时间包括早晨 8:00—10:00、中午 12:00—1:00、傍晚 5:00—6:00。为使试验样本丰富，充分利用网络资源，从 Google、Baidu、Bing 等搜索引擎上搜集数据，对样本进行补充。在林业专家的指导下，共采集星天牛、美国白蛾、黄刺蛾等 30 类常见林业害虫，并标记害虫图像类别，建立样本数据集。部分林业害虫图像如图 2 所示。



图 2 部分林业害虫样本

Fig. 2 Sample of some forest pests

从图 2 可以看出，害虫图像拥有复杂的背景，并且同类害虫形态差异比较小。其中星天牛、美国白蛾等大部分害虫图像的识别受背景因素干扰较大，并且

褐边绿刺蛾和黄刺蛾以及斑须蝽和麻皮蝽等同类害虫相似度较高,增大了整体识别难度。

1.2 图像预处理

1.2.1 数据扩充

在开始训练之前,Forest30 被分为 3 组:训练集、验证集和测试集。首先从采集的图像中选择 70%作为训练集,在剩下的 30%中,选择 70%作为验证集,剩余部分作为测试集。Thenmozhi 等^[15]也使用了同样的划分方法。在深度学习中,卷积神经网络的训练需要大量的数据,并且样本分布不均匀也会影响模型识别的准确率,因此对数据集进行扩充十分必要^[16]。在 pytorch 框架下结合 openCV 完成数据的扩充,主要采用方式包括平移变换(shift)、旋转(rotation)、噪声(noise)、水平翻转(horizontal)等数据扩充方法,进而提升网络模型的泛化能力。扩充后的数据训练集为 10 540 张,验证集为 3 133 张,测试集为 1 348 张。Forest30 数据集各类别详细数量如表 1 所示。

表 1 Forest30 数据集类别及数量

Tab. 1 Types and quantity of the Forest30 dataset 张

编号	害虫名称	数量	编号	害虫名称	数量
1	星天牛	521	16	乌柏癭皮瘤蛾	475
2	桑天牛	470	17	褐边绿刺蛾	483
3	黄足黄守瓜	480	18	斑衣蜡蝉	503
4	角蜡蚧	476	19	松墨天牛	495
5	大青叶蝉	495	20	菜粉蝶	486
6	稻棘缘蝽	504	21	台湾黄毒蛾	546
7	黄刺蛾	490	22	霜天蛾	502
8	桃蛀螟	484	23	点蜂缘蝽	535
9	悬铃木方翅网蝽	528	24	丝带凤蝶	486
10	斑须蝽	516	25	人纹污灯蛾	501
11	栗瘿蜂	500	26	尘污灯蛾	473
12	小绿叶蝉	522	27	菜蝽	536
13	麻皮蝽	477	28	赤条蝽	522
14	茶翅蝽	500	29	斜纹夜蛾	514
15	美国白蛾	531	30	柳蓝叶甲	495

1.2.2 图像增强

由于部分图片数据来自网络资源,导致清晰度低于使用移动设备拍摄的原始图像,为解决这一问题,使用一种多尺度的图像细节提升算法^[17],来改善图像质量,该方法使用 3 个不同尺度的高斯模糊对原图进行滤波,然后将滤波结果与原图做减法,由此获得不同程度的图像细节信息,然后将这些细节信息融合到原图中,从而提升图像细节。该方法的增强效果如图 3 所示,可以看到,害虫背部的纹理、腿部等各细节得到明显增强。



图 3 图像增强对比

Fig. 3 Image enhancement contrast

2 轻量级害虫识别模型构建

2.1 DCP-ShuffleNetV2 害虫识别模型

ShuffleNetV2 卷积神经网络^[18]是 2018 年提出的轻量化模型,提出 G1~G4 准则,通过通道拆分、通道混洗、分组卷积和深度可分离卷积,该网络在计算复杂度和精度上取得了优异的表现。但是在资源受限的林业害虫具体的识别任务中,对于自然环境背景相对复杂的害虫,识别效果仍然不佳,为兼顾模型复杂度与识别精度,选用 ShuffleNetV2 为基准网络模型,从特征提取、特征融合和轻量化三个方面进行改进。在轻量化方面,引入动态卷积(Dynamic Convolution);在特征融合方面,引入 CSP(Cross Stage Partial)结构;在特征提取方面,引入金字塔分割注意力模块 PSA(Pyramid Split Attention)对 ShuffleNetV2 模型进一步优化改进。综上,将本文模型命名为 DCP-ShuffleNetV2 网络模型,网络结构如图 4 所示。

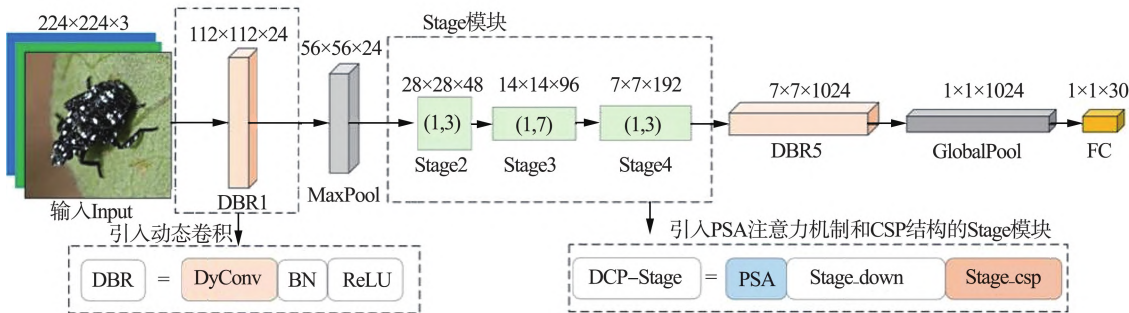


图 4 DCP-ShuffleNetV2 网络结构

Fig. 4 Network structure for DCP-ShuffleNetV2

注:DBR 由动态卷积(DyConv)、批量归一化(BN)、激活函数(ReLU)构成;MaxPool 代表最大池化;Stage 由金字塔分割注意力模块(PSA)、下采样单元模块(Stage_down)和修改为 CSP 结构的基本单元模块(Stage_csp)构成,其中(m,n)代表 stride 为 2 时 Stage_down 模块堆叠 m 次, stride 为 1 时 Stage_csp 模块堆叠 n 次;GlobalPool 代表全局平均池化;FC 代表全连接层。

首先输入图像,图像被送到卷积模块 DBR1,该卷积层修改普通卷积为动态卷积 DyConv,经最大池化 MaxPool 层之后,进入 Stage 模块。将该模块改进为 DCP-Stage 模块,改进后的模块分为三部分,第一部分是 PSA 金字塔分割注意力模块,该模块同样使用动态卷积,且拥有 4 种不同的感受野,卷积核大小分别 1×1 、 3×3 、 5×5 、 7×7 。第二部分为步长为 2 时下采样单元 Stage_down。第三部分为步长为 1 的 Stage_CSP 单元,该单元将原模型的基本单元修改为 CSP 结构。最后经过卷积层 DBR5、GlobalPool 全局平均池化和全连接层 FC 处理之后,输出 30 类害虫的预测结果。

2.2 基于动态卷积的金字塔分割注意力模块 PSA

在传统卷积神经网络的卷积操作中,一旦训练结束,所有的卷积核参数不再发生变化,所有的卷积核对输入的特征图进行相同处理,因此为提高模型的性能,往往会增加卷积层的数量或者卷积层的通道数,导致网络模型参数和计算量的大幅增长。为压缩模型,满足轻量化、易部署的需求,将 ShuffleNetV2 中的普通卷积更改为动态卷积。动态卷积^[19]的卷积核具有自适应能力,主要借鉴注意力机制的方法,是一种具有注意力机制的卷积核,会随着输入特征图变化而动态地发生变化。

如图 5 所示,输入的特征图 F ,通过注意力 attention 模块计算出 k 个卷积核的权重 π_i ,对卷积核进行加权,其中 $\pi_1 + \dots + \pi_k = 1$,通过非线性方式叠加多个卷积核,采用分组卷积的形式,最终形成动态卷积,再经过批量归一化处理以及激活函数,输出特征图 F' ,减少参数数量和计算量的同时带来更强的表示能力。

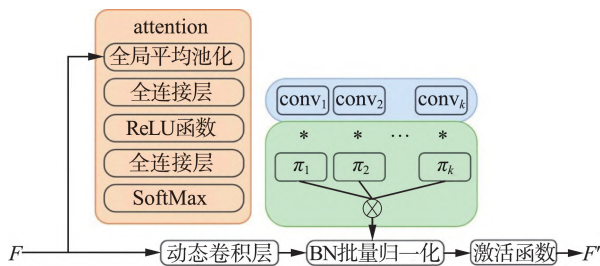


图 5 动态卷积结构

Fig. 5 Structure of the dynamic convolution

由于害虫识别的自然场景比较复杂,林木中大部分害虫存在保护色,例如大青叶蝉整体呈绿色,与叶片很容易混为一体。并且同类害虫差异也很小,比如松墨天牛、星天牛和桑天牛,基本拥有一样的形态特征,不同之处在于身体颜色以及纹理细节特征,捕捉这些细节纹理相对比较困难,然而颜色的差异和细微的纹理变化是区分不同害虫的关键。针对此问题在 ShuffleNetV2 的 Stage 模块中引入将普通卷积修改为动态卷积的金字塔分割注意力模块 PSA,通过提取多尺度的空间信息和跨通道依赖关系,

有效地学习上下文信息,从而提升模型对细粒度的特征信息的学习能力。PSA^[20]模块结构如图 6 所示。

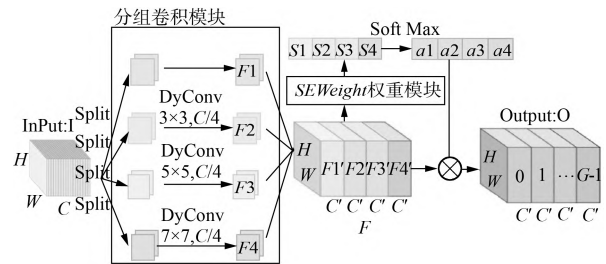


图 6 PSA 模块结构图

Fig. 6 Structure of PSA module

注: H 代表特征图高度; W 代表特征图宽度; C 代表特征图通道数; $\otimes$ 代表逐元素相乘。

首先将原始特征图分为 4 部分,分别进行动态分组卷积,对于给定的输入特征图 $I \in R^{C \times W \times H}$,通过分组卷积模块得到多尺度空间特征图 $F_i \in R^{C \times W \times H}$,其中 $i=1,2,3,4$ 。

$$F_i = \text{DyConv}(I, K_j \times K_j, C/4) \quad (1)$$

式中: $\text{DyConv}()$ ——动态分组卷积操作;

$K_j \times K_j$ ——卷积核大小, $j=1,3,5,7$ 。

对于不同比例的输入特征图 F_i ,通过权重模块获取多尺度通道注意力权重向量

$$s_i = \text{SEWeight}(F_i) = \{s_1, s_2, s_3, s_4\} \quad (2)$$

式中: $\text{SEWeight}()$ ——SE 权重函数。

再经过 SoftMax 函数进行归一化处理,将结果映射到 $[0,1]$ 之间,得到 $a_i = \{a_1, a_2, a_3, a_4\}$ 。

$$a_i = \text{SoftMax}(s_i) = \frac{\exp(s_i)}{\sum_{i=0}^{C-1} \exp(s_i)} \quad (3)$$

将多尺度空间特征图 F_i 与通道权重向量 a_i 相乘得到多尺度通道注意权重的特征图 $Y_i = F_i \otimes a_i$ 。最终使用拼接操作 Concat 得到最终的输出特征图 $O = \text{Concat}([Y_1, Y_2, Y_3, Y_4])$ 。

2.3 基于动态卷积 CSP 的模型结构优化

基准网络模型 ShuffleNetV2 的 Stage 模块具有两种分支结构,当卷积操作步长为 1 时,分支结构基本单元如图 7(a) 所示,保持特征图输入输出尺寸大小不变。当卷积操作步长为 2 时,分支结构下采样单元如图 7(b) 所示,包含下采样操作,缩小特征图尺寸,Stage 模块通过通道混洗(Channel Shuffle)以及深度可分离卷积(Depthwise Separable Convolution)^[21]等操作减少计算量的同时,完成了通道之间的信息融合,增强了特征信息的表达能力,然而该操作仅考虑通道之间信息的编码,没有考虑到不同特征层之间信息的融合,在一定程度上影响模型对林业害虫的识别能力。

为进一步提升模型准确率并减小模型参数数量和计算

量,将基本单元结构(Basic unit)改造成 CSPNet 模型^[22]的 CSP 结构,将基准网络模型中的普通卷积替换为卷积核大小可变的动态卷积。如图 8 所示,Stage_csp 单元将原模型的基本单元结构的输入,通过 Channel Split 操作使得通道数减半,一部分经过原来的基本单元路径,另一部分经过卷积操作,最终通过 Concat 方式将两个分支重新合并。CSP 结构将原始害虫特征与经过原路径输出的害虫特征进行跨度连接,促进原始特征信息进行融合,并且通过跨度连接的方式,增加不同特征层之间的信息流动,既降低计算量又丰富梯度信息,进一步提高了模型对害虫纹理、颜色等特征的学习能力。

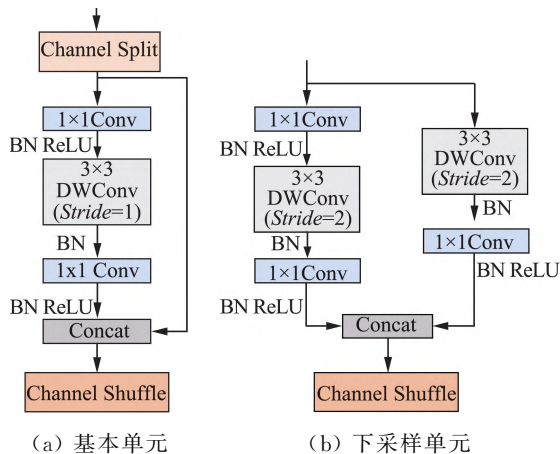


图 7 ShuffleNetV2 模块结构

Fig. 7 Basic unit structure of ShuffleNetV2

注:Channel Split 代表通道拆分;Conv 代表普通卷积;DWConv 代表深度可分离卷积;Concat 代表通道相加;Channel Shuffle 代表通道混洗。

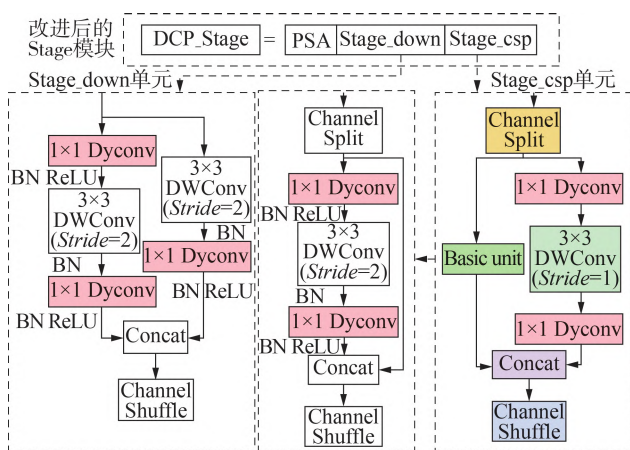


图 8 DCP_Stage 结构

Fig. 8 Structure of the DCP_Stage

2.4 评价指标

选用以下评价指标:参数量、浮点运算量、内存大小、识别准确率 $Accuracy$ 、平均召回率 $MRec$,平均精确率 $MPre$ 以及平均 F1 值 $MF1$ 。

识别准确率是指预测正确的样本量占测试总样本量的比例,计算如式(4)所示。

$$Accuracy = \frac{P}{P_{all}} \times 100\% \quad (4)$$

式中: P ——害虫被正确预测的样本数量;

P_{all} ——测试集中害虫样本总量。

计算每个类别的召回率,取平均值,如式(5)、式(6)所示。 Pre_c 和 $MPre$ 计算如式(7)、式(8)所示。

$$Rec_c = \frac{TP_c}{TP_c + FN_c} \quad (5)$$

$$MRec = \frac{\sum_{c=1}^N Rec_c}{N} \quad (6)$$

式中: Rec_c ——第 c 类的召回率;

TP_c ——第 c 类的真阳性样本数量;

FN_c ——第 c 类的假阴性样本数量;

N ——害虫类别数量。

$$Pre_c = \frac{TP_c}{TP_c + FP_c} \quad (7)$$

$$MPre = \frac{\sum_{c=1}^N Pre_c}{N} \quad (8)$$

式中: FP_c ——第 c 类的假阳性样本数量。

$MF1$ 为 $MRec$ 和 $MPre$ 的调和均值,计算如式(9)所示。

$$MF1 = 2 \times \frac{MPre \times MRec}{MPre + MRec} \quad (9)$$

3 试验结果与分析

试验平台为矩池云深度学习云平台,基于 pytorch 1.8 深度学习框架,python3.8 API 训练环境,在 Pycharm 软件上编译,在显存大小为 11 G 的 Tesla K80 GPU 上训练。超参数配置如下:初始学习率为 0.01,损失函数采用交叉熵损失函数,选择 Adam 作为学习率优化器,使用 ExponentialLR 学习率衰减策略调整学习率。为避免训练次数过多所导致的过拟合问题,试验设置最大 $epoch$ 为 100,当验证集上的分类准确率在 10 个 $epoch$ 之后依然没有提高时,训练阶段将停止。

3.1 DCP-ShuffleNetV2 模型消融试验

为实现高性能、轻量化的卷积神经网络模型,以 ShuffleNetV2 为基准网络模型,验证 DCP-ShuffleNetV2 模型的有效性。采用 5 种消融试验方案,在 Forest30 数据集上进行训练和评估。方案 1 仅采用原模型 ShuffleNetV2 进行试验。方案 2 在原模型的基础上,引入金字塔分割注意力 PSA 模块进行试验。方案 3 在原模型基础上,修改 Stage 模块为 CSP 结构进行试验。方案 4 在原模型基础上,引入 PSA 模块,并修改 Stage 模块为 CSP 结构进行试验。方案 5 在方

案 4 改进的基础上,将普通卷积修改为动态卷积,形成本文的 DCP-ShuffleNetV2 模型。试验结果如表 2 所示。

比较方案 1 和方案 2 可知,相比于原模型,在基准模型中引入 PSA 模块后,准确率提升 1.18%,同时参数量和计算量分别增长 17.97%和 69.89%。比较方案 1 和方案 3 可知,将 ShuffleNetV2 的 Stage 模块,修改为 CSP 结构,相比于原模型,准确率提高 1.56%,参数量减少 16.39%,浮点运算量减少 19.95%。并且所占内存减小 1.7 MB,说明 CSP 结构可以有效的增强特征融合能力的同时减小计算成本。综合方案 2、方

案 3 和方案 4 的试验结果可知,方案 4 结合 PSA 模块和 CSP 结构,在 3 组试验方案中,达到最高的识别准确率 92.23%。相较于单独使用两个改进策略,通过 PSA 模块提取的害虫细节特征,再经过 CSP 结构的跨阶段融合后,能够达到更好的效果。比较方案 4 和方案 5 可知,将普通卷积替换为动态卷积之后,模型的准确率虽然只增长 0.2%,但是参数量和浮点运算量分别减少 63.41%和 61.64%。表明动态卷积得益于自适应生成的小卷积核尺寸和动态卷积内部所使用的分组卷积,使得模型参数量和计算量大幅下降。

表 2 DCP-ShuffleNetV2 模型消融试验结果

Tab. 2 Ablation study results for DCP-ShuffleNetV2 model

改进策略	平均精确率/%	平均召回率/%	平均 F1 值/%	准确率/%	参数量	浮点运算量/M	内存大小/MB
Base	89.28	89.35	89.34	89.32	372 542	42.65	11.23
Base+PSA	90.53	90.65	90.54	90.50	439 494	72.46	11.74
Base+CSP	90.94	90.63	90.89	90.88	311 498	34.14	9.53
Base+PSA+CSP	92.33	92.10	92.25	92.23	378 450	63.95	10.22
Base+DCP	92.69	92.49	92.51	92.43	138 458	24.53	9.53

注:Base 代表基准网络模型 ShuffleNetV2(0.5×)。

此外,方案 1 和方案 5 的试验结果表明,相对于改进前的 ShuffleNetV2 模型,改进后的 DCP-ShuffleNetV2 模型在 Forest30 数据集上的害虫识别准确率提升 3.11%,参数量、计算量和内存大小分别减少 62.83%、42.48%和 15.13%。综上所述 DCP-ShuffleNetV2 模型是一种高性能、轻量化的网络模型。

3.2 同类卷积神经网络模型性能对比试验

为评估 DCP-ShuffleNetV2 网络对林业害虫识别性能的优越性,选取同类优秀网络模型进行比较,包括 AlexNet、ResNet-18、ResNet-50 以及轻量级网络 MobileNetV2、MobileNetV3、GhostNet、ShuffleNetV2。为遵循单一变量原则,排除其他因素的干扰,试验均采用相同的试验参数配置,保持模型基本架构不变的情况下,将最后一个全连接层的维度修改为 30,以适应的 30 个害虫类别。试验结果如表 3 所示。

表 3 不同模型性能比较

Tab. 3 Performance comparison of different models

模型	参数量/M	浮点运算量/M	内存大小/MB	平均 F1 值/%	准确率/%
AlexNet	57.13	711.57	4.19	83.52	83.60
ResNet-18	11.19	1 820.00	25.65	89.91	89.84
ResNet-50	23.57	4 120.00	109.68	88.37	91.39
MobileNetV2	2.26	318.99	74.00	89.24	89.52
MobileNetV3	2.71	262.15	45.14	77.89	77.89
GhostNet	3.94	146.99	48.70	87.58	87.72
ShuffleNetV2(0.5×)	0.37	42.65	11.23	89.84	89.32
Our method	0.13	24.53	9.53	92.51	92.43

由表 3 可知,DCP-ShuffleNetV2 模型在所有待比较模型中达到最好的性能,与目前常用的分类模型相比,识别准确率平均提高 5.39%,模型参数量、计算量和内存大小平均减小 14.32 M、1 035.80 M 和 35.98 MB。因此,该模型具有高性能、轻量化的特点。后续工作可尝试将其应用于嵌入式设备,以满足林业害虫识别任务的真实需求,促进智慧林业的发展。

3.3 不同注意力机制性能对比试验

为验证金字塔分割注意力 PSA 的性能,将 DCP-ShuffleNetV2 模型中的 PSA 模块替换成通道注意力 ECA(Efficient Channel Attention)^[23]、坐标注意力 CA(Coordinate Attention)^[24]、通道注意力 SE(Squeeze-and-Excitation)^[25]进行对比试验,试验结果如表 4 所示。

表 4 不同注意力机制的模型性能对比

Tab. 4 Comparison results of models using different attention mechanisms

模型	参数量	浮点运算量/M	平均 F1 值/%	准确率/%
SE	135 618	65.22	89.90	89.89
CA	137 682	16.62	90.06	90.33
ECA	124 498	30.52	90.23	90.65
PSA	138 458	24.53	92.51	92.43

比较方案 1 和 3 可知,SE 和 ECA 都为通道注意力机制,ECA 性能明显优于 SE 模块,这表明在通道注意力机制中,ECA 模块更有效,ECA 将原来 SE 模块的全连接层直接去掉,在经过全局平均池化之后的特

征上,通过一个 1D 卷积进行学习,而 SE 模块捕获所有通道的依赖关系,徒增了参数量和计算量,是低效且没有必要的。比较方案 2、方案 3 和方案 4 可知,CA 运算量最低,但识别准确率与 PSA 相比减少 2.1%,ECA 参数量最低,识别准确率与 PSA 相比减少 1.78%,而 PSA 只增长少量参数量和计算量,取得更高的识别准确率,满足提高准确率的前提下,尽量减少模型参数量和计算量的需求。

由于卷积神经网络具有不可解释性,为进一步探究 PSA 的有效性,使用激活热图(Grad-CAM)^[26]对不同模型进行可视化分析比较。从图 9 可以看出,对于输入的原始害虫图像,SE、CA 和 ECA 受到背景信息的干扰。而 PSA 既考虑空间信息,也考虑通道信息,可以更加精确的聚焦到图像中害虫的重要区域,并且降低复杂背景的关注度。表明 PSA 模块可以根据多尺度的上下文信息有效提取林业害虫的关键特征,因此具有更强的学习能力。

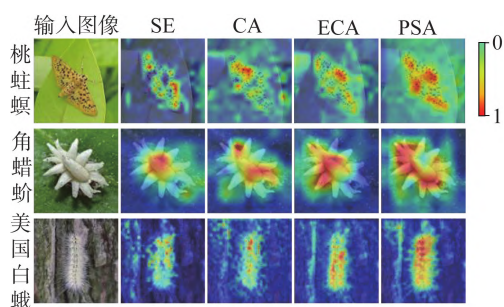


图 9 各注意力机制热力激活图对比

Fig. 9 Comparison of thermal activation diagrams with different attention mechanisms

3.4 害虫识别模型部署

基于 DCP-ShuffleNetV2 模型,设计并实现 30 类林业害虫小程序(图 10)。首先,用户通过相册或者摄像头拍摄上传一幅害虫图像,通过小程序将待识别的害虫图像发送给 Web 应用框架 Flask 搭建的系统后台,并将图像分辨率转换为 $224 \times 224 \times 3$ 进行归一化处理,输入到部署好的模型中进行识别,最终将识别结果返回到小程序展示给用户。



图 10 害虫识别小程序应用效果图

Fig. 10 Effect diagrams of pest identification applet application

由图 10 可知,DCP-ShuffleNetV2 模型在林业害

虫识别任务中具有较好的应用效果,后续工作将会进一步采集害虫数据,扩大数据集,并考虑在资源受限的嵌入式设备中开展试验与部署工作。

4 结论

1) 选用轻量级卷积神经网络 ShuffleNetV2 为基准网络模型,设计并提出基于 PSA 模块、CSP 结构和动态卷积的 DCP-ShuffleNetV2 网络模型。构建包含 30 类虫害数据集 Forest30,通过数据扩充增强策略,平衡各类别虫害图像的分布。

2) 针对林业害虫复杂背景以及类间差异小的问题,引入金字塔分割注意力 PSA 模块,提取多尺度融合的空间信息以及跨通道信息,增强模型的特征提取能力。试验结果表明,相比于原模型,引入 PSA 模块后,准确率提升 1.18%。

3) 将基准模型的 Stage 模块修改为 CSP 结构,将原始害虫特征与经过原路径输出的害虫特征进行跨度连接,增加不同特征层之间的信息流动,有效地增强特征融合能力,并且降低计算量。试验结果表明,将 ShuffleNetV2 的 Stage 模块,修改为 CSP 结构,相比于原模型,准确率提高 1.56%,参数量减少 16.39%,浮点运算量减少 19.95%。

4) 为进一步压缩模型参数量和计算量,引入动态卷积,形成 DCP-ShuffleNetV2 林业害虫识别模型。该模型参数量仅为 138 458,模型大小仅为 9.53 MB,与基准模型相比,在自制的 Forest30 数据集上的害虫识别准确率提升 3.11%,参数量、计算量和内存大小分别减少 62.83%、42.48%和 15.13%有效的平衡识别准确率和模型的复杂度。

参 考 文 献

- [1] 国家统计局. 2021 中国统计年鉴 [M]. 北京: 中国统计出版社, 2021.
- [2] 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.
- [3] 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 Chinese Society of Agricultural Engineering, 2014, 30 (14): 154-162.

- [4] 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.
Wang Meihua, Wu Zhenxin, Zhou Zuguang. Fine-grained identification research of crop pests and diseases based on improved CBAM via attention [J]. Transactions of the Chinese Society of Agricultural Machinery, 2021, 52 (4): 239—247.
- [5] 张永玲, 姜梦洲, 俞佩仕, 等. 基于多特征融合和稀疏表示的农业害虫图像识别方法 [J]. 中国农业科学, 2018, 51 (11): 2084—2093.
- [6] 肖志云, 刘洪. 小波域马铃薯典型虫害图像特征选择与识别 [J]. 农业机械学报, 2017, 48 (9): 24—31.
Xiao Zhiyun, Liu Hong. Features selection and recognition of potato typical insect pest images in wavelet domain [J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48 (9): 24—31.
- [7] 邹修国, 丁为民, 刘德营, 等. 基于 4 种不变矩和 BP 神经网络的稻飞虱分类 [J]. 农业工程学报, 2013, 29 (18): 171—178.
Zou Xiuguo, Ding Weimin, Liu Deying, et al. Classification of rice planthopper based on invariant moments and BP neural network [J]. Transactions of the Chinese Society of Agricultural Engineering, 2013, 29 (18): 171—178.
- [8] 陈娟, 陈良勇, 王生生, 等. 基于改进残差网络的园林害虫图像识别 [J]. 农业机械学报, 2019, 50 (5): 187—195.
Chen Juan, Chen Liangyong, Wang Shengsheng, et al. Pest image recognition of garden based on improved residual network [J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50 (5): 187—195.
- [9] Deng L, Wang Y, Han Z, et al. Research on insect pest image detection and recognition based on bio-inspired methods [J]. Biosystems Engineering, 2018, 169: 139—148.
- [10] 孔建磊, 金学波, 陶冶, 等. 基于多流高斯概率融合网络的病虫害细粒度识别 [J]. 农业工程学报, 2020, 36 (13): 148—157.
Kong Jianlei, Jin Xuebo, Tao Zhi, et al. Fine-grained recognition of diseases and pests based on multi-stream Gaussian probability fusion network [J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36 (13): 148—157.
- [11] Wei D, Chen J, Luo T, et al. Classification of crop pests based on multi-scale feature fusion [J]. Computers and Electronics in Agriculture, 2022, 194: 106736.
- [12] Khanramaki M, Asli-Ardeh E A, Kozegar E. Citrus pests classification using an ensemble of deep learning models [J]. Computers and Electronics in Agriculture, 2021, 186: 106192.
- [13] Yang Z, Yang X, Li M, et al. Automated garden-insect recognition using improved lightweight convolution network [J]. Information Processing in Agriculture, 2021.
- [14] 李江昀, 赵义凯, 薛卓尔, 等. 深度神经网络模型压缩综述 [J]. 工程科学学报, 2019, 41 (10): 1229—1239.
- [15] Thenmozhi K, Reddy U S. Crop pest classification based on deep convolutional neural network and transfer learning [J]. Computers and Electronics in Agriculture, 2019, 164: 104906.
- [16] Liu Y, Liu S, Xu J, et al. Forest pest identification based on a new dataset and convolutional neural network model with enhancement strategy [J]. Computers and Electronics in Agriculture, 2022, 192: 106625.
- [17] Kim Y, Koh Y J, Lee C, et al. Dark image enhancement based on pairwise target contrast and multi-scale detail boosting [C]. International Conference on Image Processing (ICIP). IEEE, 2015: 1404—1408.
- [18] Ma N, Zhang X, Zheng H T, et al. ShuffleNetV2: Practical guidelines for efficient CNN architecture design [C]. Proceedings of the European Conference on Computer Vision, 2018: 116—131.
- [19] Chen Y, Dai X, Liu M, et al. Dynamic convolution: Attention over convolution kernels [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11030—11039.
- [20] Zhang H, Zu K, Lu J, et al. EPSANet: An efficient pyramid squeeze attention block on convolutional neural network [J]. arXiv preprint arXiv: 2105.14447, 2021.
- [21] Sandler M, Howard A, Zhu M, et al. MobileNetV2: Inverted residuals and linear bottlenecks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510—4520.
- [22] Wang C Y, Liao H Y M, Wu Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 390—391.
- [23] Wang Q, Wu B, Zhu P, et al. ECANet: Efficient channel attention for deep convolutional neural networks [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11531—11539.
- [24] Hou Q, Zhou D, Feng J. Coordinate attention for efficient mobile network design [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713—13722.
- [25] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132—7141.
- [26] Selvaraju R R, Cogswell M, Das A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 618—626.