

WraNet: 一种基于二维离散小波变换的轻量害虫识别网络

李 晖¹, 吴茜茵^{1*}, 胡欣仪¹, 唐栩燃¹, 罗 伟¹, 赵雪如¹, 谭廷俊², 赵泽华¹, 李超然¹

¹贵州大学, 大数据与信息工程学院, 贵州 贵阳

²贵州大学, 农学院, 贵州 贵阳

收稿日期: 2023年12月8日; 录用日期: 2024年1月12日; 发布日期: 2024年1月19日

摘 要

近年来, 人工智能技术在害虫识别领域得到广泛应用。目前深度网络害虫识别方法仍存在计算量大、对复杂背景下的害虫识别效果差等问题。为了解决计算量大的问题, 本文提出了一种新型轻量网络——WraNet。该网络利用二维离散变换模块对图像进行特征混合, 并学习图像的强先验知识, 例如尺度不变性、平移不变性和边缘稀疏性。这使得单层二维离散小波变换层达到多层深度神经网络的效果, 从而减少了计算量和模型参数的大小。本文还提出了一种新的算法——WraNet-m, 该算法通过软投票集成了WraNet、ResNet50和FPN网络模型, 以进一步提升识别效果。WraNet-m算法在IP102和D0害虫数据集上的准确率分别达到了72.44%和99.52%, 证明了集成方法的有效性和鲁棒性。

关键词

害虫识别, 二维离散小波变换, 计算机视觉, 深度学习

WraNet: An Efficient Pest Recognition Network Based on 2D Discrete Wavelet Transform

Hui Li¹, Xiyin Wu^{1*}, Xinyi Hu¹, Xuran Tang¹, Wei Luo¹, Xueru Zhao¹, Tingjun Tan², Zehua Zhao¹, Caoran Li¹

¹College of Big Data and Information Engineering, Guizhou University, Guiyang Guizhou

²College of Agriculture, Guizhou University, Guiyang Guizhou

Received: Dec. 8th, 2023; accepted: Jan. 12th, 2024; published: Jan. 19th, 2024

*通讯作者。

文章引用: 李晖, 吴茜茵, 胡欣仪, 唐栩燃, 罗伟, 赵雪如, 谭廷俊, 赵泽华, 李超然. WraNet: 一种基于二维离散小波变换的轻量害虫识别网络[J]. 图像与信号处理, 2024, 13(1): 33-46. DOI: 10.12677/jisp.2024.131004

Abstract

In recent years, with the promotion of agricultural informatization, artificial intelligence techniques have been widely applied in the field of pest recognition. However, current deep neural network-based pest recognition methods still face challenges such as high computational complexity and poor performance in complex background scenarios. To address the issue of high computational complexity, we propose a novel network called WraNet. This network employs a two-dimensional discrete transform module for token mixing and learns strong prior knowledge of the image, such as scale-invariance, shift-invariance, and sparseness of edges. It is worth noting that we also propose a new algorithm, WraNet-m, which combines WraNet, ResNet50, and FPN models through soft voting for further performance improvement. The WraNet-m algorithm achieves accuracies of 72.44% on the IP102 pest dataset and 99.52% on the D0 pest dataset, approaching state-of-the-art results on both datasets, thus demonstrating the effectiveness and robustness of the ensemble method.

Keywords

Pest Recognition, 2D Discrete Wavelet Transform, Computer Vision, Deep Learning

Copyright © 2024 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

农业害虫是农业生产中重要的问题，其影响了农作物的产量和质量。害虫识别是指根据植物受害的症状或直接观察到的害虫的外形特征，对害虫进行分类和鉴别的过程。该过程对农业生产至关重要，可以帮助农民快速发现和识别害虫，及时采取有效的防治措施，从而提高作物的产量和质量。

近年来，农业信息化技术的推广，将计算机视觉技术广泛运用于害虫识别领域，并取得一定研究成果。杨红云等[1]提出了一种基于迁移学习和卷积神经网络相结合的害虫识别融合机制。类成敏等[2]提出了一种基于多尺度注意力残差网络的害虫图像识别模型。姚青等[3]建立了运用 ResNet50 的基于双线性注意力网络的农业害虫识别模型。肖小梅等[4]对传统的卷积神经网络 Alexnet 进行优化改进。Wu 等[5]引入大规模害虫数据集 IP102，包含了属于 102 个类别的 75,000 多幅图像，呈现出自然的长尾分布。Liu 等[6]提出了一种新型深度多分支融合残差网络(DMF-ResNet)，创建了深度多分支融合残差网络。Nanni 等[7]应用显著性方法与卷积神经网络之间的融合，提出了一种自动分类器，并在小型数据集和大型数据集 IP102 上测试该方法，准确性达 61.93%。Thenmozhi 等[8]提出了一种有效的深度 CNN 模型，用于在三个公开可用的昆虫数据集上对昆虫种类进行分类，在 DO 数据集上实现分类准确率达 96.0%。在害虫图像识别研究方面，国内外研究人员运用 CNN 网络架构取得了一定的研究成果。但农业害虫存在种类繁多，不同虫态的害虫外观差异大，数据采集难度较大的现象。尤其目前常用的害虫识别架构在图像分析任务上泛化能力的提高是以增加参数和层数、数据集大小、训练和测试计算以及 GPU 内存为代价的，缺少同等性能的轻量型害虫识别机制。在能达到高准确率的同时，轻量化模型所需参数，是本文的研究方向。

2. 本文算法

2.1. WraNet 网络模型

本文提出的新型神经网络 WraNet 由改进的 2D-DWTM 模块[9]和残差注意力模块等组成。WraNet 网络框架如图 1 所示,输入图像首先通过卷积层创建特征映射,然后使用 2D-DWTM 进行处理,以达到参数减约的目的。此外,为了进行图像识别,还添加了一个残差注意力模块、一个全局平均池化层和一个线性层。其中,2D-DWTM 和残差注意力模块是 WraNet 的核心模块,下文将对其进行详细介绍。

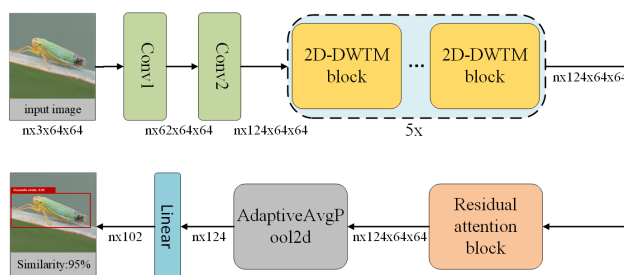


Figure 1. WraNet network architecture

图 1. WraNet 网络架构

2.2. 基于 2D-DWT 的参数减约模块

二维离散小波变换模块(2D Discrete Wavelet Transform Module, 2D-DWTM)是一种新型的神经网络模块,可像 transformers 和卷积神经网络一样进行推广。与传统卷积神经网络相比,2D-DWTM 利用二维离散小波变换来处理图像像素中的空间信息,从而能够更有效地利用图像的尺度不变性和边缘稀疏性,并能够利用 2D-DWT 来提取特征的空间维度以降低所需的计算量,具体的结构图如图 2 所示。

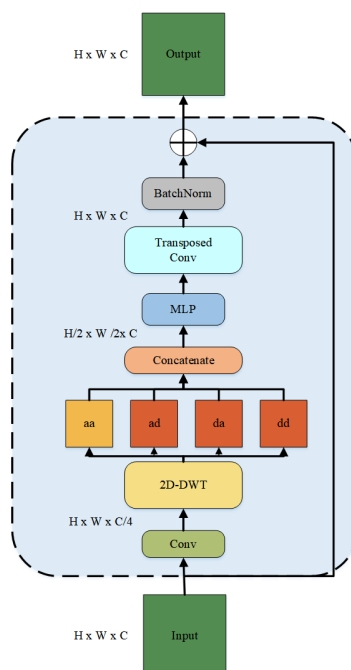


Figure 2. Structure of 2D-DWTM module

图 2. 2D-DWTM 模块结构

具体而言, 2D-DWTM 采用小波分解对输入的特征图进行多尺度分解, 并利用不同尺度的小波系数进行特征交换, 以促进特征之间的信息交流。将对应的输入和输出张量表示为 x_{in} 和 x_{out} ; 四个小波滤波器以及他们的下采样分别表示为 w_{aa} , w_{ad} , w_{da} , w_{dd} (a 代表近似值, d 代表细节); 卷积、多层感知机、转置卷积(上卷积)和批归一化操作分别表示为 c 、 m 、 t 和 b ; 它们对应的可训练参数集表示为 ε 、 φ 、 ϕ 和 γ ; 通道维度上的串联操作表示为 \oplus , 点对点加法表示为 $+$, 则 2D-DWTM 内部对应的操作可以用以下公式表示:

$$x_0 = c(x_{in}, \varepsilon); x_{in} \in R^{H \times W \times C}, x_0 \in R^{H \times W \times C/4} \quad (1)$$

$$x = [w_{aa}(x_0) \oplus w_{ad}(x_0) \oplus w_{da}(x_0) \oplus w_{dd}(x_0)]; x \in R^{H/2 \times W/2 \times 4C/4} \quad (2)$$

$$\tilde{x} = b(t(m(x, \varphi), \phi), \gamma); \tilde{x} \in R^{H \times W \times C} \quad (3)$$

$$x_{out} = \tilde{x} + x_{in}; x_{out} \in R^{H \times W \times C} \quad (4)$$

在公式(1)中, 本文首先使用卷积层来提取可学习的、空间不变的特征, 并将嵌入维度降低至四分之一。接着, 本文通过 2D-DWT 进行空间多尺度特征融合和下采样(参见公式(2)), 从而实现尺度不变的特征提取。在这一过程中, 本文通过计算一个近似子带和三个细节子带来降低空间分辨率, 将输出通道数增加到输入通道数的四倍。通过这种适用于图像且不会损失信息的 2D-DWT 下采样, 可以有效地降低计算量。最后, 本文使用可学习的 MLP 层进行通道融合, 并使用转置卷积层恢复特征映射的空间分辨率(参见公式(3)和(4))。其中 2D-DWTM 中的 MLP 与主干网络中的 MLP 的设计相同, 在代码中都是通过 `ff_chanel`、`mult` 两个超参数来设计 MLP 前馈过程中的输出通道数以及通道的扩展数。

因此, 2D-DWTM 模块可以利用更多的图像先验信息, 降低计算和模型参数的大小。将该模块应用于害虫识别中, 有助于进一步减小网络模型的尺寸, 这对于在害虫识别的部署阶段提高识别设备的效率具有重要意义。

2.3. 改进的残差注意力模块

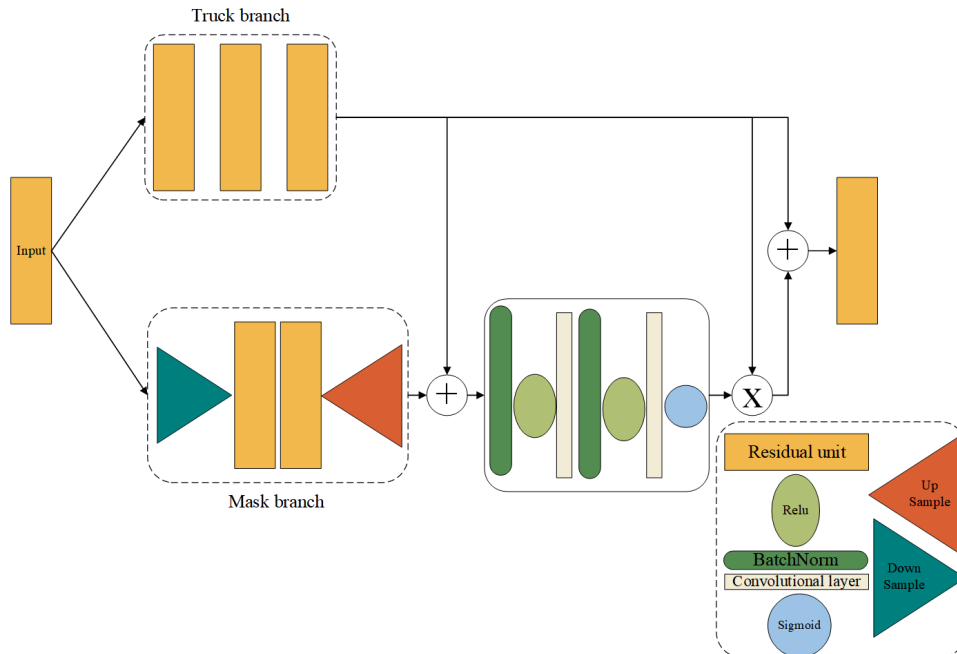


Figure 3. Structure of residual attention module
图 3. 残差注意力模块结构图

由于害虫识别应用需要在自然环境中进行，为了能够使构建的网络能够在复杂的自然背景中准确识别害虫，在网络架构中添加了一个残差注意力模块。该模块由两个关键组件组成，即主干分支和掩膜分支[10]。其中，主干分枝负责图像特征的提取，而掩膜分枝用于学习对输出特征进行软加权的注意力屏蔽。具体而言，该模块的输入经过残差层处理后，被分成两条不同的路径。一条路径直接进入下一层，另一条路径则经过注意力机制进行加权处理后再与第一条路径相乘并相加，最后再进入下一层进行处理。通过这样的设计，剩余注意力模块可以更好地利用图像信息，提高网络的准确性。结合本次实验的具体要求，本文设计了如图3所示的残差注意力模块。

在上图的操作中，本文不妨将残差块变换、下采样(最大池化)、上采样(双线性插值)批归一化、Relu、卷积操作以及 Sigmoid 分别表示为 Re 、 d 、 u 、 b 、 r 、 c 和 s ；对应的可训练参数分别表示为 ε 、 β 、 δ 、 φ 、 ϕ 、 γ 和 η ；点对点加法表示为 \oplus ，点对点乘法表示为 \otimes 。那么本残差注意力模块内部的操作可以用如下公式表示：

$$x_0 = R(x_m, \varepsilon); x_0, x_m \in R^{H \times W \times C} \quad (5)$$

$$x_1 = \text{Re}(\text{Re}(\text{Re}(x_0, \varepsilon), \varepsilon), \varepsilon); x_2 = u(\text{Re}(\text{Re}(d(x_0, \beta), \varepsilon), \varepsilon), \delta)); x_1, x_2 \in R^{H \times W \times C} \quad (6)$$

$$x_3 = x_2 \oplus x_0; x_3 \in R^{H \times W \times C} \quad (7)$$

$$x_4 = s\left(c\left(r\left(b\left(c\left(r\left(b(x_3, \varphi), \phi\right), \gamma\right), \varphi\right), \phi\right), \gamma\right)\eta\right); x_4 \in R^{H \times W \times C} \quad (8)$$

$$x_5 = x_4 \otimes x_0; x_5 \in R^{H \times W \times C} \quad (9)$$

$$x_6 = x_5 \oplus x_0; x_6 \in R^{H \times W \times C} \quad (10)$$

在这个模块中，输入图片首先经过公式(5)中的单层残差模块进行处理。然后，它被分为主干分支和掩膜分支进行处理。主干网络的处理如公式(6)所示，进行了三次残差操作。对于掩膜分支，在公式(6)中，首先通过最大池化下采样将图片的尺寸缩小至一半，然后经过两次残差处理，再利用双线性插值进行上采样，恢复到原始尺寸大小。最后，在公式(7)~(8)中的操作中，通过与主干分支相加并进行一系列的归一化、卷积等操作，最终得到一个位于 0 到 1 之间的张量。在接下来的公式(9)~(10)中，将该张量与主干分支的张量相加和相乘，即通过加权融合，实现了增强重要特征的效果。

通过上述模块的堆叠，成功实现了 WraNet 在识别性能与卷积神经网络相当的情况下，其模型参数的需求大幅降低的优势，同时还实现了性能的提高。这意味着能够在减小网络模型尺寸的同时保持高准确率，这对于网络模型在资源有限的环境下的应用具有非常重要的意义。

3. 实验结果与分析

3.1. 数据集设置

为了探究本文提出的新型轻量网络和相应的集成算法在嵌入式设备上的实用性，本实验分别在 64×64 的 IP102 和 D0 数据集上测试了它们的性能，同时也与其他卷积神经网络进行了比较。此外，为了尝试获得本次提出的 WraNet-m 集成算法的最佳准确性，本文还在 224×224 的 IP102 和 D0 数据集上训练了 ResNet50 和 FPNs [11]，并进行了性能对比。其中，IP102 是当前最大的害虫数据集之一，共包含 75, 222 张图片，其各种类害虫样本数量呈长尾分布，存在内外方差大和数据不平衡的挑战。此外，训练数据样本中还包含受损农作物的照片以及一些呈集中状分布在复杂环境中的小害虫图片，具体如图4所示，这些都对模型的训练具有挑战性。



Figure 4. Image example from IP102 dataset

图 4. IP102 数据集图片示例

D0 [14]数据集包含 4508 张图片, 图片来自 40 种不同的害虫类别, 图片均为处于自然状态下的害虫, 图 5 展示了一些图片的例子。



Figure 5. Example from D0 dataset

图 5. D0 数据集示例

3.2. 实验设置

实验使用了 PyTorch 框架进行网络模型的搭建, 其中 PyTorch 的版本为 1.8.1, Python 的版本为 3.8, 同时也使用了 CUDA 版本为 11.1 的 GPU 进行加速。在一张显存大小为 16 GB 的 RTX3060 上面进行实验。针对 IP102 数据集, 按照 7:1:2 的比例随机划分成训练数据集 45,095 张、验证集 7508 张、测试集 22,619 张图片。而 D0 数据集则划分为训练数据集 3156 张、验证集 405 张、测试集 947 张。详细的网络设置如表 1 所示。

Table 1. Parameter settings of the network

表 1. 网络的参数设置

网络模型	学习率	批量大小	优化器	学习率调整机制	权值衰减	丢弃率	训练最大轮次	输入图片大小	数据集
ResNet-50	0.0001	64	0.9/0.999	0.96	0.00001	0.3	100	64 × 64/224 × 224	IP102, D0
FPNs	0.0001	32	0.9/0.999	0.96	0.00001	0	100	64 × 64/224 × 224	IP102, D0
WraNet (IP102)	0.01	128	0.9/0.999	0.96	0.00001	0.3	150	64 × 64	IP102
WraNet (D0)	0.001	128	0.9/0.999	0.96	0.0001	0.3	100	64 × 64	D0
SqueezeNet1_1	0.0001	64	0.9/0.999	0.96	0.0002	0.5	100	64 × 64	IP102, D0
MobileNetV3	0.0001	64	0.9/0.999	0.96	0.00004	0.2	100	64 × 64	IP102, D0
shufflenet_v2	0.0001	64	0.9/0.999	0.96	0.0001	0.1	100	64 × 64	IP102, D0
EfficientNet-B0	0.0001	64	0.9/0.999	0.96	0.0001	0.3	100	64 × 64	IP102, D0

在 WraNet 结构的设计中, 还存在 `final_dim` 和 `depth` 两个超参数, 分别控制的是进入二维离散小波变换层的通道维度以及堆叠的 2D-DWTM 数量。本次实验中本文采用了 `ff_chanel = 124`、`final_dim = 124`、`depth = 5` 的参数组合, 该组合在 IP102 和 D0 数据集中均表现良好。`muti` 参数在 D0 数据集和 IP102 数据集上训练时均设置为 3。本文使用分类交叉熵作为代价函数, 并选择 Adam 作为优化器, 其中 `beta` 参数分别为 0.9 和 0.999。为了调整学习率, 本文采用了指数衰减的策略。每个 `batch` 的大小在训练过程中最小为 32, 最大为 128, 并且本文最多进行了 150 次训练。

3.3. 评价指标

实验中主要采用了 6 种评估指标, 分别为: 参数大小(Param)、宏观平均精度(MPre)、宏观平均召回率(MRec)、宏观平均 F1 值(MF1)、准确率(Acc)和几何平均值(GM)来进行评估。

准确率的计算公式如下:

$$\text{Acc} = \frac{\text{TP}}{N} \quad (11)$$

其中的 TP 代表的是预测正确的数据个数, 而 N 则为参与预测的数据总数。

预测精度(Pre)和宏观平均精度(MPre)的计算公式如下:

$$\text{Pre}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i} \quad (12)$$

$$\text{MPre} = \frac{\sum_{i=1}^I \text{Pre}_i}{I} \quad (13)$$

公式中的 I 代表的是分类总数, 而 TP_i 和 FP_i 代表的就是每个种类当中的正确预测数和误报数。之后将所有类的预测进度做了一个取平均值的操作即得到了宏观平均精度。该指标主要展示了模型在分类是时不会将错误样本标记为正确样本的能力。

类似地, 对于每个类的召回率以及宏观平均召回率的计算公式如下:

$$\text{Rec}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i} \quad (14)$$

$$\text{MRec} = \frac{\sum_{i=1}^I \text{Rec}_i}{I} \quad (15)$$

公式中的 FN_i 即为每个类中的漏报数, 其中宏观平均召回率的计算(MRec)和宏观平均精度(MPre)的计算相同。该指标主要显示了模型在分类时, 召回所有的正确样本的能力。

其次利用上文得到的宏观平均召回率的计算(MRec)和宏观平均精度(MPre), 既可以得到宏观平均 F1 值(MF1), 该指标同时兼顾了 MPre 和 MRec。具体计算公式如下:

$$\text{MF1} = 2 \frac{\text{MPre} \cdot \text{MRec}}{\text{MPre} + \text{MRec}} \quad (16)$$

最后, 根据每个类的召回率作为其预测的敏感度(S_i), 来进行几何平均值(GM)的计算, 其显示了在数据量大的类别和数据量小的类别之间平衡分类的性能。 S_i 和 GM 的具体计算公式如下:

$$S_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i} \quad (17)$$

$$GM = \prod_{i=1}^I \sqrt{S_i} \quad (18)$$

3.4. 模型性能展示

3.4.1. 与轻量级网络比较

根据表 2 和表 3 的比较结果，在 IP102 和 D0 数据集上，WraNet 相对于当前主流的一些轻量级神经网络表现更出色。此外，图 6 展示了 WraNet 在 D0 数据集上的网络参数与其他轻量级网络的对比情况。

Table 2. Performance of various lightweight networks on IP102 (64×64)

表 2. IP102 (64×64)上各轻量网络的性能展示

Model Name	Param Size (MB)	Acc (%)	MPre (%)	MRec (%)	MF1 (%)	GM (%)
SqueezeNet1_1	4.9	41.52	35.91	40.09	37.47	32.67
MobileNetV3	11.9	55.48	49.47	52.3	50.58	46.24
shufflenet_v2	5.4	52.09	44.42	47.94	45.37	39.54
EfficientNet-B0	16	60.42	54.38	56.7	55.27	51.4
WraNet	7.9	57.49	51.19	51.38	51.02	47.86

Table 3. Performance of various lightweight networks on D0 (64×64)

表 3. D0 (64×64)上各轻量网络的性能展示

Model Name	Param Size (MB)	Acc (%)	MPre (%)	MRec (%)	MF1 (%)	GM (%)
SqueezeNet1_1	4.8	85.04	83.84	84.23	83.67	82.13
MobileNetV3	11.6	92.05	91.42	91.85	91.52	91.02
shufflenet_v2	5.1	87.03	83.39	87.4	82.93	76.46
EfficientNet-B0	15.7	94.14	93.27	94.49	93.69	92.84
WraNet	7.9	95.71	95.1	95.97	95.39	94.89

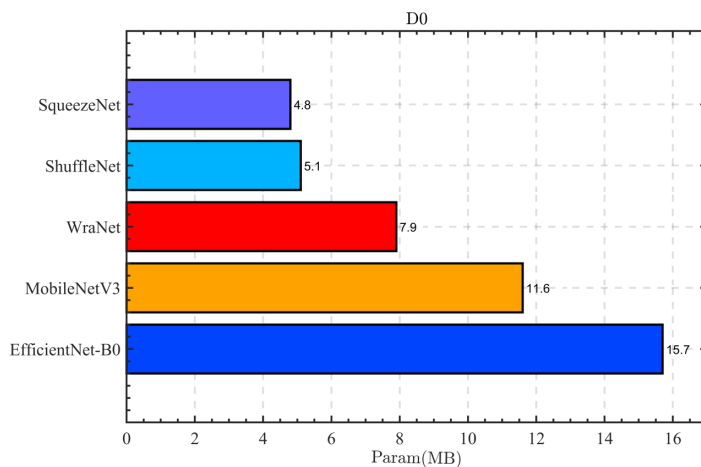


Figure 6. Parameter comparison between WraNet and mainstream lightweight neural networks

图 6. WraNet 与主流轻量神经网络的参数对比

从上述数据可以看出，WraNet 的模型参数在几个轻量级网络中处于中间位置，这表明 WraNet 的参数大小适中，适用于轻量级网络。与当前主流的轻量神经网络相比，WraNet 的表现相当出色。因此，本文有信心在未来的实验中通过进一步调整网络结构和超参数设置，提升 WraNet 在害虫识别方面的性能，同时保持其参数优势。

3.4.2. 与非轻量型网络比较

在 IP102 与 D0 上, WraNet 与 ResNet 系列网络以及 FPNs 表现如表 4、表 5 所示。

Table 4. Recognition performance of IP102 dataset (64×64)

表 4. IP102 数据集(64×64)的识别性能

Model name	Param Size (MB)	Acc (%)	MPre (%)	MRec (%)	MF1 (%)	GM (%)
ResNet18	42.9	57.12	51.29	55.42	52.91	48.61
ResNet34	81.5	57.98	51.85	55.72	53.22	49.02
ResNet50	90.8	60.38	54.18	57.28	55.3	51.37
ResNet101	163.5	59.71	53.81	57.23	55.03	50.91
FPNs	111	60.23	53.65	56.38	54.6	50.72
WraNet	7.9	57.49	51.19	51.38	51.02	47.86
WraNet-m	-	64.10	56.75	61.82	58.65	53.81

Table 5. Recognition performance of D0 dataset (64×64)

表 5. D0 数据集(64×64)的识别性能

Model name	Param Size (MB)	Acc (%)	MPre (%)	MRec (%)	MF1 (%)	GM (%)
ResNet18	42.8	93.77	92.89	94.09	93.29	92.48
ResNet34	81.4	93.88	93.34	94.56	93.76	93.1
ResNet50	90.3	96.41	96.07	96.68	96.25	95.96
ResNet101	163	96.51	95.97	96.86	96.34	95.83
FPNs	110.7	95.56	95.08	95.97	95.4	94.89
WraNet	7.9	95.71	95.10	95.97	95.39	94.89
WraNet-m	-	97.68	97.14	97.99	97.49	97.05

在本次实验中, 通过表 4、5 的对比不同模型参数大小以及其他性能指标可以看到, WraNet 在模型参数大小方面具有显著的优势, 仅需其他网络的 1/5 到 1/20 的模型参数大小。这样的优势在实际应用中能够显著降低模型的存储和计算资源消耗, 进一步提高模型的运行效率和应用范围。图 10 进一步直观地展示了 WraNet 在模型参数大小方面的优越性。图 7 展示了 IP102 数据集上各个模型的参数大小比较情况(由于在两个数据集上使用的模型大小大致相同, 这里仅展示 IP102)。可以看出, WraNet 在两个数据集上相对于其他网络具有更小的模型参数大小。这也证明了 WraNet 所具有的高效性和适用性, 使其成为解决资源有限场景下的理想选择, 例如移动设备和嵌入式设备等。

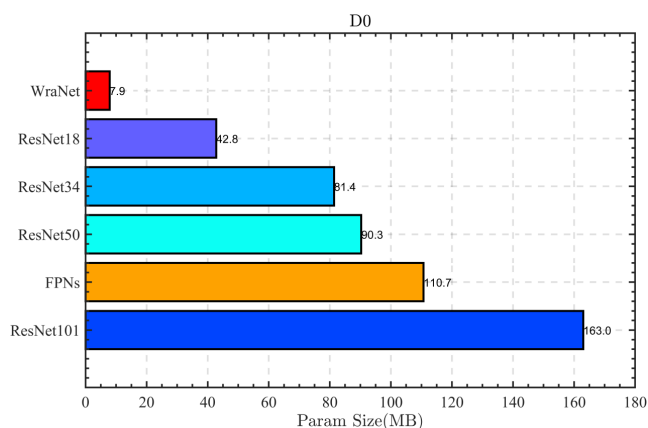


Figure 7. Parameter comparison between WraNet and non-lightweight neural networks

图 7. WraNet 与非轻量神经网络的参数对比

此外, 根据表 4 和表 5 的结果显示, WraNe-m 算法在 D0 和 IP102 (64×64) 两个数据集上表现最佳, 分别达到 97.68% 和 64.1% 的准确率。这表明软投票集成有效地融合了每个模型的优点, 从而取得了更好的性能。综合以上数据表现来看, WraNet 仍然具备成为一款优秀的轻量害虫识别网络的潜力。因此, 在未来的实验中, 本文将继续探索更多超参数设置以及优化 WraNet 结构的方法。

3.4.3. 不同图像尺寸对 WraNet-m 的影响

为了进一步评估 WraNet-m 算法的性能, 本文在 IP102 和 D0 数据集上采用了 224×224 的图像尺寸重新对 ResNet50 和 FPNs 进行训练。本文记录了它们以及 WraNet-m 算法的结果, 并将其详细列于表 6。通过增大输入图像尺寸, 本文观察到 ResNet50、FPNs 和 WraNet-m 算法的性能都有显著提升。在 IP102 数据集上, ResNet50 和 FPNs 的准确率均提高了约 10%, 而 WraNet-m 的准确率提高了约 8%。对于 D0 数据集, ResNet50、FPNs 和 WraNet-m 的准确率均提高了约 3%。这表明, 当对识别速度要求不高或设备内存较大时, 通过使用更大的图像尺寸可以提高识别性能, 具体对比情况如图 8 所示。

Table 6. Performance demonstration under 224×224 input

表 6. 224×224 输入下性能展示

Dataset	Model Name	Acc (%)	MPre (%)	MRec (%)	MF1 (%)	GM (%)
IP102	ResNet50	71.35	63.62	66.16	64.57	60.87
	FPNs	70.28	63.08	64.77	63.52	60.27
	WraNet-m	72.44	64.24	67.88	65.66	61.28
D0	ResNet50	99.37	99.2	99.37	99.26	99.18
	FPNs	99.37	99.18	99.43	99.28	99.15
	WraNet-m	99.58	99.41	99.6	99.49	99.4

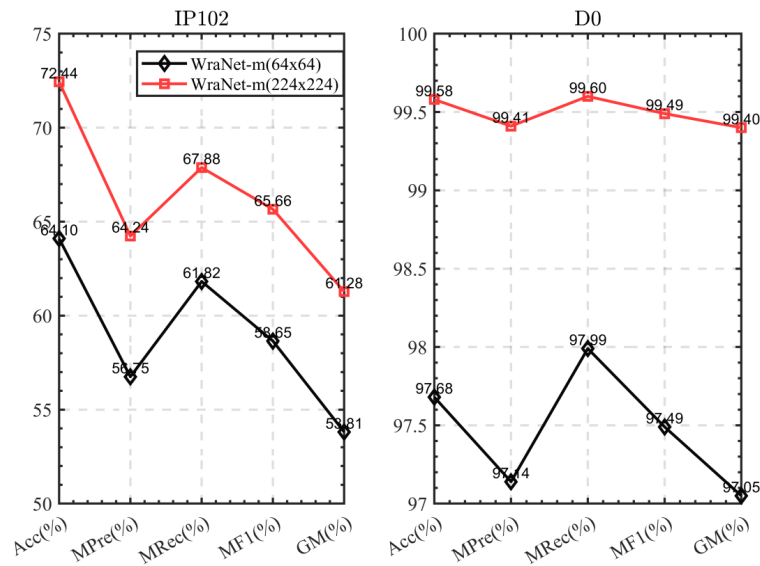


Figure 8. Performance of WraNet-m under different image sizes

图 8. 不同图像尺寸下的 WraNet-m 表现

3.4.5. 与相关工作的比较

为了展示本研究提出的 WraNet-m 算法的可行性, 本文在表 7 中将其与以往工作的结果进行了比较。本文的实验结果在具有挑战性的实际应用场景中得到了验证, 证明了 WraNet 模型的实用性和实际效果。本研究为将来的害虫识别研究提供了新的思路和参考, 即利用 WraNet 模型的集成以及模型集成的方法来

提高模型的性能和准确性。在未来的研究中, 本文将继续改进 WraNet 模型, 并探索如何在其他实际场景中应用该模型。同时, 本文还将研究更好的模型集成方法, 以获得更好的性能。这些研究对于更好地理解图像识别问题, 并为相关领域的应用提供更好的解决方案具有重要意义。

Table 7. Comparison between proposed algorithm and related works

表 7. 本文算法与相关工作的比较

数据集	作者	方法	Acc	MF1	GM
IP102	Wu [5]	ResNet-50	49.4	40.1	31.5
	Ren [12]	FR-ResNet	55.2	54.1	-
	Liu [6]	DMF-ResNet	59.1	58.1	-
	Nanni [7]	Saliency method + CNNs + EM	61.4	-	-
	Peng [13]	CNNS + transformer	74.89	-	-
	本文算法	WraNet	57.07	51.45	47.71
	本文算法	WraNet-m	72.44	65.66	61.28
D0	Xie [14]	MLLF + MKB	89.3	-	-
	Peng [12]	CNNS + transformer	99.47	-	-
	Nanni [15]	ensembles of CNNs	99.81	-	-
	本文算法	WraNet	95.71	95.39	94.89
	本文算法	WraNet-m	99.58	99.49	99.4

3.5. 消融实验

WraNet-m 是由 ResNet50、FPNs 以及 WraNet 进行软投票集成而来。为了确定模型集成中哪个网络占据主导地位, 本研究进行了一系列消融实验。这些实验通过在 64×64 的 IP102 和 D0 数据集上对 WraNet-m 中的每个网络模型进行两两集成的性能测试, 来确定每个模型在本次集成中所发挥的作用。为了保证实验变量的统一性, 本文对不同的网络模型都进行了相同的测试和比较。实验结果见表 8 和表 9, 并且本文还在图 9 和图 10 中对其进行了可视化呈现。

Table 8. Ablation experiment results on D0

表 8. D0 上的消融实验结果

模型集成	Acc	MPre	MRec	MF1	GM
ResNet50 + FPNs	96.39	96.08	96.69	96.31	95.99
ResNet50 + WraNet	97.43	97.07	97.23	97.1	96.98
WraNet + FPNs	97.12	96.49	97.47	96.89	96.36
WraNet-m	97.68	97.14	97.99	97.49	97.05

Table 9. Ablation experiment results on IP102

表 9. IP102 上的消融实验结果

模型集成	Acc	MPre	MRec	MF1	GM
ResNet50 + FPNs	62.36	55.38	59.11	56.83	52.47
ResNet50 + WraNet	62.83	55.74	59.46	57.11	52.89
WraNet + FPNs	62.63	55.28	59.14	56.74	52.26
WraNet-m	64.10	56.75	61.82	58.65	53.81

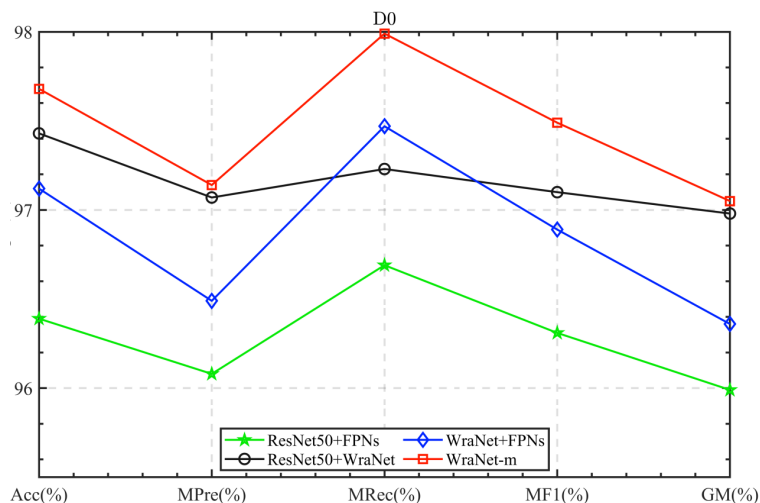


Figure 9. Ablation experiment results demonstration on D0

图 9. D0 上的消融实验结果展示

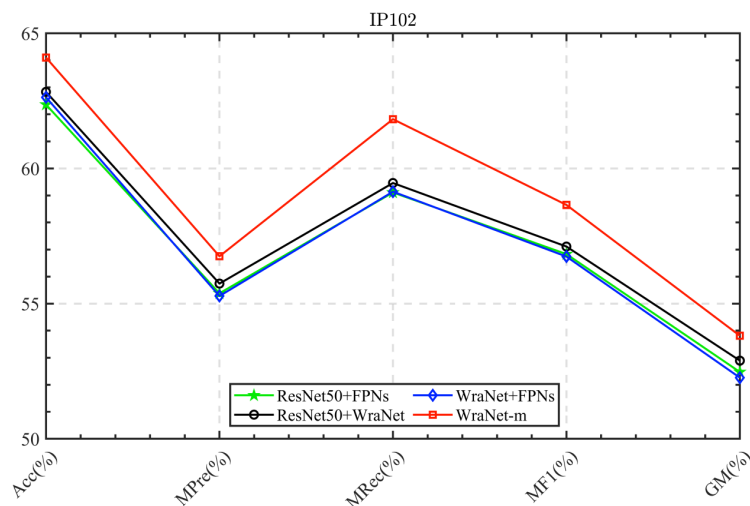


Figure 10. Ablation experiment results on IP102

图 10. IP102 上的消融实验结果

通过消融实验的结果，本文发现 ResNet50 和 WraNet 的集成在 IP102 和 D0 上都取得了最佳性能，即 62.83% 和 97.43%。其次是 WraNet 和 FPNs 的集成，在 IP102 和 D0 上的准确率分别为 62.63%、97.12%。最后是 ResNet50 和 FPNs 的集成，两个数据集上的准确率为 62.36% 以及 96.39%。实验结果表明，WraNet 在模型集成中发挥了最大的作用，其次是 ResNet50 和 FPNs。尽管 WraNet 在性能指标上并没有优于 ResNet50 和 FPNs，但是在消融结果中却占据了主导地位。这说明 WraNet 相比卷积神经网络能够更好地获取图像的额外先验知识，在处理一些对于卷积神经网络来说比较困难的图像时表现更好。因此，在软投票集成时，WraNet 与 ResNet50 以及 WraNet 与 FPNs 的集成能够相互补充，从而超过了 ResNet50 和 FPNs 的集成性能。这为卷积神经网络在应用困难的方面提供了新的解决思路。

总体而言，实验结果表明，采用不同的卷积神经网络结构进行模型集成可以在图像分类任务中获得更好的性能。WraNet 作为一种轻量级网络架构，在保证性能的同时，大大减小了模型参数的大小，具有一定的实际应用价值。

3.6. 网络特征激活图展示

此外, 实验还运用了 GRAD_CAM [16]来对实验中的 ResNet50、FPNs 以及 WraNet 进行可视化展示, 即将模型的最后一层卷积层通过梯度激活来生成激活图。图片为 D0 中的害虫样本, 具体如图 11 所示。

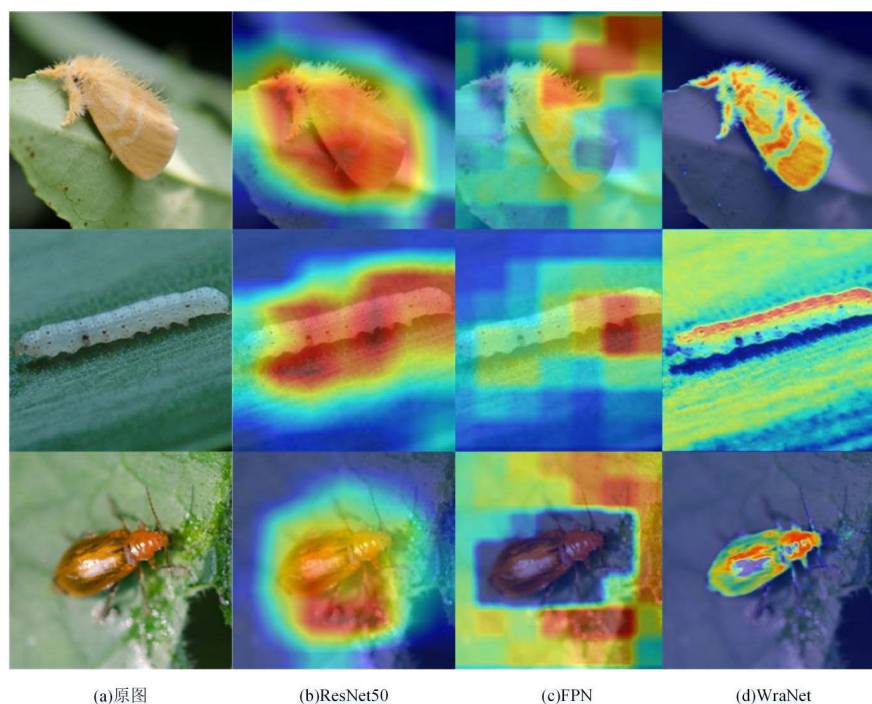


Figure 11. Network feature activation heat map
图 11. 网络特征激活热力图

从图 11 可以看出, ResNet50 和 FPNs 在图像识别中的关注点大多数是正确的, 但是它们的关注范围比较广泛, 不仅关注害虫本身, 还关注害虫周围的环境。然而, 与其它两种模型相比, WraNet 的注意点更加精确, 不会浪费太多的计算量在害虫周围的环境上。通过网络特征激活热力图, 也说明了为什么 WraNet 所需的参数更少以及它相对于当前主流的 CNNs 和 transformers 的优势。此外, 从图 11(c)列中最底部的图片可以观察到, 在某些情况下, 卷积神经网络只关注害虫周围的环境, 相比之下, 图 11(d)列中的 WraNet 能够更精确地关注害虫, 这凸显了其与传统卷积神经网络在获取图像信息方面的差异。通过以上分析和结果展示可知, WraNet 能够与当前主流的卷积神经网络媲美, 在拥有更少参数的同时, 在某些卷积神经网络处理效果差的情况下表现更好, 这表明了 WraNet 在害虫识别领域具有巨大的潜力。

4. 总结与展望

本文提出了一种新型网络 WraNet, 采用二维离散小波变换对图像进行特征混合, 以减少计算量和模型参数的大小。同时, 还提出了 WraNet-m 算法以提升识别效果。这大幅降低了对网络部署模型的性能要求, 为降低深度学习技术在农业方面的应用成本提供了思路。其次, WraNet-m 算法通过集成 WraNet、ResNet 以及 FPN 网络模型的软投票, 使得对于复杂环境中的害虫识别效果得到提升。这为农业害虫识别性能的提高提供了一种简单高效的办法。在未来的研究中, 本文计划进一步探索 WraNet 在害虫识别领域的潜力。本文将投入更多资源, 包括数据集的收集和标注、模型结构的优化以及超参数的调整等方面, 以提高 WraNet 的性能并使其在更广泛的场景下适用。

基金项目

贵州大学大学生创新创业训练计划项目资助(2023 年国家级大学生创新创业训练计划(项目编号 202310657022))。

参考文献

- [1] 杨红云, 肖小梅. 基于卷积神经网络和迁移学习的水稻害虫识别[J]. 激光与光电子学进展, 2022, 59(16): 330045.
- [2] 类成敏, 牟少敏. 基于多尺度注意力残差网络的桃树害虫图像识别[J]. 山东农业大学学报(自然科学版), 2022, 53(2): 253-258.
- [3] 姚青, 姚波. 基于双线性注意力网络的农业灯诱害虫细粒度图像识别研究[J]. 中国农业科学, 2021, 54(21): 4562-4572.
- [4] 肖小梅, 杨红云, 易文龙, 等. 改进的 Alexnet 模型在水稻害虫图像识别中的应用[J]. 科学技术与工程, 2021, 21(22): 9447-54.
- [5] Wu, X., Zhan, C., Lai, Y.K., et al. (2019) Ip102: A Large-Scale Benchmark Dataset for Insect Pest Recognition. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, 15-20 June 2019, 8787-8796. <https://doi.org/10.1109/CVPR.2019.00899>
- [6] Liu, W., Wu, G. and Ren, F. (2020) Deep Multibranch Fusion Residual Network for Insect Pest Recognition. *IEEE Transactions on Cognitive and Developmental Systems*, **13**, 705-716. <https://doi.org/10.1109/TCDS.2020.2993060>
- [7] Nanni, L., Maguolo, G. and Pancino, F. (2020) Insect Pest Image Detection and Recognition Based on Bio-Inspired Methods. *Ecological Informatics*, **57**, 101089. <https://doi.org/10.1016/j.ecoinf.2020.101089>
- [8] Thenmozhi, K. and Reddy, U.S. (2019) Crop Pest Classification Based on Deep Convolutional Neural Network and Transfer Learning. *Computers and Electronics in Agriculture*, **164**, 104906. <https://doi.org/10.1016/j.compag.2019.104906>
- [9] Jeevan, P., Viswanathan, K. and Sethi, A. (2022) Wavemix-Lite: A Resource-Efficient Neural Network for Image Analysis. arXiv: 2205.14375,2019
- [10] Wang, F., Jiang, M., Qian, C., et al. (2017) Residual Attention Network for Image Classification. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, 21-26 July 2017, 3156-3164. <https://doi.org/10.1109/CVPR.2017.683>
- [11] Lin, T.Y., Dollár, P., Girshick, R., et al. (2017) Feature Pyramid Networks for Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, 21-26 July 2017, 2117-2125. <https://doi.org/10.1109/CVPR.2017.106>
- [12] Ren, F., Liu, W. and Wu, G. (2019) Feature Reuse Residual Networks for Insect Pest Recognition. *IEEE Access*, **7**, 122758-122768. <https://doi.org/10.1109/ACCESS.2019.2938194>
- [13] Peng, Y. and Wang, Y. (2022) CNN and Transformer Framework for Insect Pest Classification. *Ecological Informatics*, **72**, 101846. <https://doi.org/10.1016/j.ecoinf.2022.101846>
- [14] Xie, C., Wang, R., Zhang, J., et al. (2018) Multi-Level Learning Features for Automatic Classification of Field Crop Pests. *Computers and Electronics in Agriculture*, **152**, 233-241. <https://doi.org/10.1016/j.compag.2018.07.014>
- [15] Nanni, L., Manfè, A., Maguolo, G., et al. (2022) High Performing Ensemble of Convolutional Neural Networks for Insect Pest Image Detection. *Ecological Informatics*, **67**, 101515. <https://doi.org/10.1016/j.ecoinf.2021.101515>
- [16] Selvaraju, R.R., Cogswell, M., Das, A., et al. (2017) Grad-Cam: Visual Explanations from Deep Networks via Gradient-Based Localization. *Proceedings of the IEEE International Conference on Computer Vision*, Venice, 22-29 October 2017, 618-626. <https://doi.org/10.1109/ICCV.2017.74>