

# 基于改进 YOLOv4-tiny 的节肢动物目标检测模型

余咏<sup>1</sup>, 吴建平<sup>1,2</sup>, 何旭鑫<sup>1</sup>, 韦杰<sup>1</sup>, 高雪豪<sup>1</sup>

(1. 云南大学信息学院, 云南昆明 650504;

2. 云南省电子计算中心, 云南昆明 650223)

**摘要:**针对自然环境下节肢动物背景复杂、形态万千、遮挡目标和目标尺度多样等因素,导致模型检测效率不高、边界框定位不准确的情况,提出一种基于改进 YOLOv4-tiny 的节肢动物目标检测模型。首先,结合空间、通道卷积注意力机制(CBAM),抑制背景噪声;其次,引入可变形卷积(DCN)以及改进的加权双向特征金字塔,重塑卷积和特征融合方式进行多尺度预测;最后,在FPN网络中引出一层Feat@3,嵌入空间金字塔池化结构,有效提取节肢动物的各种显著特征,使模型泛化能力更强,将改进后的模型命名为YOLOv4-tiny-ATO。实验结果表明,该模型在大小仅为54.6 Mb的前提下,很好地平衡了检测速度和检测精度,检测精度为0.725,检测速度达到89.6帧·s<sup>-1</sup>,召回率为0.585,较改进前相比YOLOv4-tiny模型,检测精度提高0.426,模型在模型大小、检测速度上更适用于移动端部署,模型检测精度也能达到应用标准,满足对节肢动物的检测需求。

**关键词:**节肢动物;目标检测;可变形卷积;YOLOv4-tiny;双向特征金字塔

中图分类号:TP312

文献标识码:A

文章编号:1673-629X(2024)01-0114-07

doi:10.3969/j.issn.1673-629X.2024.01.017

## Arthropod Object Detection Model Based on Improved YOLOv4-tiny

YU Yong<sup>1</sup>, WU Jian-ping<sup>1,2</sup>, HE Xu-xin<sup>1</sup>, WEI Jie<sup>1</sup>, GAO Xue-hao<sup>1</sup>

(1. School of Information Science & Engineering, Yunnan University, Kunming 650504, China;

2. Yunnan Provincial Electronic Computing Center, Kunming 650223, China)

**Abstract:** Aiming at the situation that the model detection efficiency is not high, and the bounding box prediction is wrong caused by the complex background, variety of morphology, occlusion target and diverse target scale of arthropods in the natural environment, an arthropod target detection model based on improved YOLOv4-tiny is proposed. Firstly, combining spatial and channel convolutional attention mechanism (CBAM), the background noise is suppressed. Secondly, deformable convolution (DCN) and an improved weighted bidirectional feature pyramid are introduced to reshape the convolution and feature fusion methods for multiscale prediction. Finally, a layer of Feat@3 is extracted in the FPN network, and a spatial pyramid pool structure is embedded to effectively extract various significant features of arthropods, so as to enhance the generalization ability of the model. The improved model is named YOLOv4-tiny-ATO. The experimental results show that the proposed model balances detection speed and accuracy well with a size of only 54.6 Mb. The detection accuracy is 0.725, the detection speed reaches 89.6 frames per second, and the recall rate reaches 0.585, which is 0.426 higher than that of the YOLOv4-tiny model before the improvement. The model is more suitable for mobile deployment in terms of model size and detection speed, and the model detection accuracy can also meet the application standards to meet the detection needs of arthropods.

**Key words:** arthropods; object detection; deformable convolution; YOLOv4-tiny; bidirectional feature pyramid

## 0 引言

节肢动物在目前已知的物种中约有三分之二, Arthropoda, 其门下包括常见的蜘蛛、甲虫、蚊子、蝉、蚂蚁、蜜蜂、飞蛾、豆娘等。现阶段,对节肢动物的检测手段相对局限,耗时耗力。因此,提高昆虫识别的准确率

和效率,进而促进精准农业节肢动物的鉴定与识别,对生态环境监测<sup>[1-2]</sup>、减少病虫害损失等方面有着重大的意义。

目前,常见的卷积神经网络检测算法可划分为单阶段目标检测算法和两阶段目标检测算法。只进行一

收稿日期:2023-03-28

修回日期:2023-07-31

基金项目:国家自然科学基金项目(62172354)

作者简介:余咏(1998-),男,硕士研究生,研究方向为图像处理;通信作者:吴建平(1965-),男,高级工程师,硕士,研究方向为智能化、信息化技术。

次特征提取的目标检测算法以YOLO(You Only Look Once)<sup>[3]</sup>,SSD(Single Shot multibox Detector)<sup>[4]</sup>,RetinaNet<sup>[5]</sup>系列算法为代表,提升检测速率同时兼顾了较好的准确性。两阶段目标检测算法以R-CNN为代表,后续衍生出Fast R-CNN以及Faster R-CNN<sup>[6]</sup>,检测精度得到大幅度提升,但因网络参数量较大,导致检测速度较慢;单阶段目标检测算法在一定程度上比两阶段目标检测算法的检测效率更高,使得实时检测成为了可能。

近年来,基于深度学习的发展,为识别方向的研究奠定了坚实的基础,对节肢动物检测提供了可能。2019年Lin等<sup>[7]</sup>通过改进灰狼优化算法,利用主旋转局部二值模式(Dominant Rotated Local Binary Patterns,DRLBP)的纹理特征提取算法作为提取器,基于概率协同表示的分类器(Probabilistic Collaborative Representation Based Classifier,PROCRC)实现鳞翅目昆虫识别;2020年黄世国等<sup>[8]</sup>使用一种新的特征选择技术以准确、快速地对鳞翅目进行分类;2021年谢娟英等<sup>[9]</sup>提出2种新的注意力机制,DSEA(direct squeeze-and-excitation with global average pooling)和DSEM(direct squeeze-and-excitation with global max pooling),改进单阶段目标检测典型算法RetinaNet,对自然环境鳞翅目(Lepidoptera)其下蝴蝶科进行自动识别;唐万梅等<sup>[10]</sup>使用新的区域候选框建议提取网络Imp-RPN进行融合,在改进损失函数的基础上,提取特征再筛选,通过迁移学习的方式实现蜻蜓目检测;2022年Guo等<sup>[2]</sup>设计了基于YOLO v5的昆虫识别方法,解决小样本情况下识别少量类别精确度的问题,实现对四种有害昆虫的识别。

当前算法大多只对某一种或单一目的节肢动物进行自动识别,在自然环境复杂多样的条件下难以满足实际的检测需求。因此,该文提出了基于轻量化、易部署的YOLOv4-tiny目标检测框架,针对自然生态环境下存在的密集目标、复杂背景及多尺度等情况提出改进算法。该算法将提取的浅层特征和深层特征进行融合,针对性改变原有模型结构,通过类激活图对提取特征有效性分析后,验证了该算法的可行性。

## 1 YOLOv4-tiny-ATO

自然环境下由于节肢动物背景复杂、形态万千、遮挡目标和目标尺度多样等因素,对节肢动物特征进行提取时,容易出现边界不清晰,噪声过多,导致模型检测效率不高,边界框预判错误的情况。该方法从硬件设备的配置要求出发,选取具有易部署、多尺度、端到端等特点的轻量级模型YOLOv4-tiny<sup>[11]</sup>进行节肢动物检测的研究。引入可变形卷积(DCN)以及改进的

加权双向特征金字塔,重塑卷积和特征融合方式进行多尺度预测;结合空间、通道卷积注意力机制(CBAM),抑制背景噪声;在FPN网络中引出一层Feat@3,嵌入空间金字塔池化结构,有效提取节肢动物的各种显著特征。改进后模型泛化能力和鲁棒性更强,将其命名为YOLOv4-tiny-ATO。

### 1.1 可变形卷积(Deformable Convolution)

可变形卷积(DCN)的核心思想是在原有卷积的基础上,通过增加空间采样点数量,利用额外的偏移量实现对目标形态学习感知(如图1所示),新的卷积方式不需要额外的监督学习,而是直接进行反向传播,便能完成端到端的正则化学习。

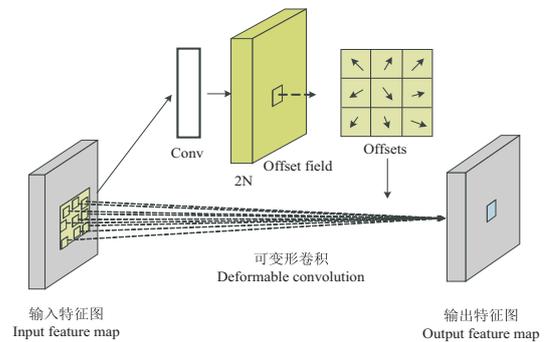


图1 可变形卷积过程

常规定义卷积公式如式1:

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n) \quad (1)$$

其中, $p_n$ 是由卷积核规定对于 $p_0$ 的偏移量, $p_0$ 是输出feature map中的采样点。

可变形卷积与式1相比,在原有基础上多出的偏移量,则是由输入feature map通过第二层卷积对每一个点进行卷积生成,绝大部分非整数。

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \quad (2)$$

由于偏移量 $\Delta p_n$ 大部分为非整数,无法对在feature map上的采样点进行求导,便利用双线性插值(Bilinear)的方法获得间断点采样值,公式表示如下:

$$\begin{aligned} x(p) &= \sum_q G(q,p) \cdot x(q) = \\ &= \sum_q g(q_x,p_x) \cdot g(q_y,p_y) \cdot x(q) = \\ &= \sum_q \max(0,1 - |q_x - p_x|) \cdot \\ &= \max(0,1 - |q_y - p_y|) \cdot x(q) \end{aligned} \quad (3)$$

双线性插值的主要思想是通过将采样点进行归一化处理后,使其与feature map上临近4个像素点,根据采样点横纵坐标来规定处理每个点的权重,利用线性函数进行求值,最终得到目标像素的值。从上式可看出,函数中规定了对间断点插值的权重距离,最大值小于1个像素距离 $\max(0,1-\dots)$ 。

两种卷积采样位置对比如图2所示。

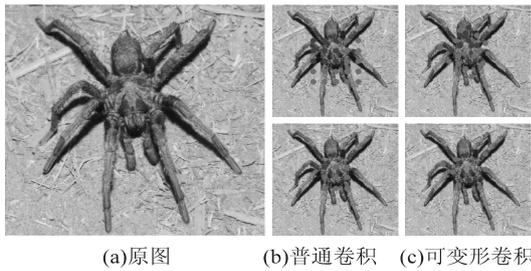


图 2 两种卷积采样位置对比

在可变形卷积结构中,其偏移量是通过另一层卷积结构进行计算得到,属于同一个网络结构,不同卷积层进行处理的结果,均可在网络中进行梯度反向传播训练。可变形卷积核权重通过目标形状特征进行自适应动态处理,将偏移量结合学习后,实现对复杂的不规则图像内容进行特征提取。

### 1.2 激活函数改进

在 YOLOv4-tiny 网络中的激活函数采用 LeakyReLU,其函数表达如下:

$$\text{LeakyReLU}(x) = \begin{cases} x, & x \geq 0 \\ \alpha x, & \text{otherwise} \end{cases} \quad (4)$$

选用更优的 ELU 激活函数代替 YOLOv4-tiny 网络中的 LeakyReLU 函数,其函数表达式如下:

$$\text{ELU}(x) = \begin{cases} x, & x > 0 \\ \alpha(e^x - 1), & x \leq 0 \end{cases} \quad (5)$$

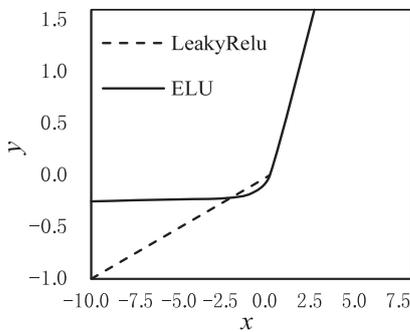


图 3 激活函数 ELU 和 LeakyReLU 对比

式中,  $x$  为输入图像的特征,  $\alpha$  数值可变, 负值部分饱和程度由  $\alpha$  控制。由式 5 能够看出 ELU 在  $x \leq 0$  时能够更好的收敛, 是因为此时具有指数函数特性。图 3 为激活函数 Leaky RELU 与 ELU 的对比, 从图中能够看出, 在负区间, ELU 比 LeakyRELU 对噪声的过滤效果更好, 收敛速度更快。

### 1.3 基于注意力机制的多尺度预测

基于节肢动物尺度差异较大的特性, YOLOv4-tiny 模型仅对两级有效特征层进行预测, 再进行特征聚合, 在模型训练过程中, 检漏、误检情况较多, 实际效果并不理想。增加一级浅层有效特征层, 位于网络颈部, 尺寸为  $52 \times 52$ , 扩大特征提取感受野。引入注意力机制, 能够有效弥补添加浅层特征层带来的问题, 减轻提取过多背景噪声带来的负面影响。

前馈神经网络卷积块注意力机制, 是一种简单有效的轻量级双独立维度 (Spatial and Channel) 注意力机制模块 (CBAM), 仅需要较少的计算资源, 便可无缝集成到 YOLOv4-tiny 模型中, 具体网络结构如图 4 所示。

### 1.4 引入改进 BiFPN 和 SPP 空间金字塔池化

YOLOv4-tiny 网络模型采用 FPN 结构, 但由于其融合特征的方式是单向的, 导致提取的不同尺度信息利用率不高。于是, 在 FPN 基础上, 改变那些只有一个输入的节点, 自顶向下对不同特征进行融合, 引入 PANet (Path aggregation network)<sup>[12]</sup> 实现双向路径聚合, 其次, 将同一级输出与输入节点进行链路聚合, 引入 BiFPN 双向特征金字塔结构, 减轻网路负担的同时, 能够有效提取目标中更具有代表性的特征信息。根据训练实际情况, 由于过多的链路聚合的方式存在, 融合了大量节肢动物的无效特征信息, 从而移除横向尺度链接, 降低了过拟合。BiFPN 和改进 BiFPN 的网络结构如图 5 所示。

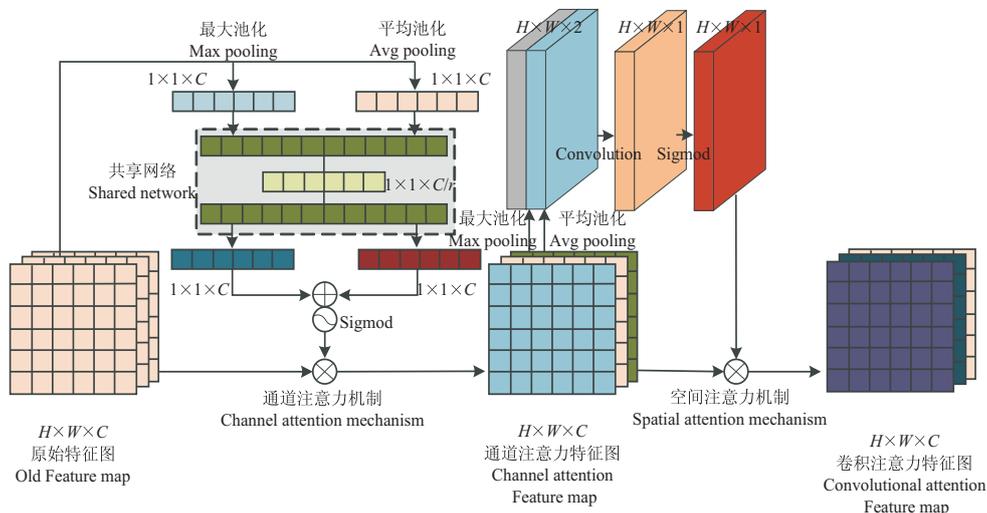


图 4 CBAM 网络结构

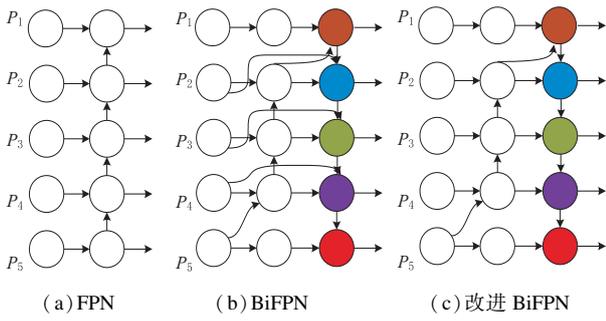


图 5 3 种特征金字塔结构

SPP 能解决因全连接层特征数固定,导致输入图像尺寸与输入标准不统一的问题,对多方面特征进行融合,丰富提取特征的语义表达能力。为针对节肢动物图像尺寸大小不一导致语义信息模糊的情况,在特

征融合网络中嵌入空间金字塔池化 SPP 结构,使模型检测性能得到提升。

### 1.5 YOLOv4-tiny-ATO 网络结构

结合空间、通道卷积注意力机制 (Convolutional Block Attention Module, CBAM)<sup>[13]</sup>、可变形卷积 (Deformable Convolution, DCN)<sup>[14]</sup>,在其颈部网络引出 52×52 的大尺度特征层,引入改进的加权双向特征金字塔 (Bidirectional Feature Pyramid Networks, BiFPN)<sup>[15]</sup>,重塑卷积和特征融合方式进行多尺度预测,并在 FPN 前加入空间金字塔池化结构 (Space Pyramid Pooling, SPP),有效提取节肢动物的各种显著特征,使模型泛化能力更强。改进后的模型命名为 YOLOv4-tiny-ATO,模型结构如图 6 所示。

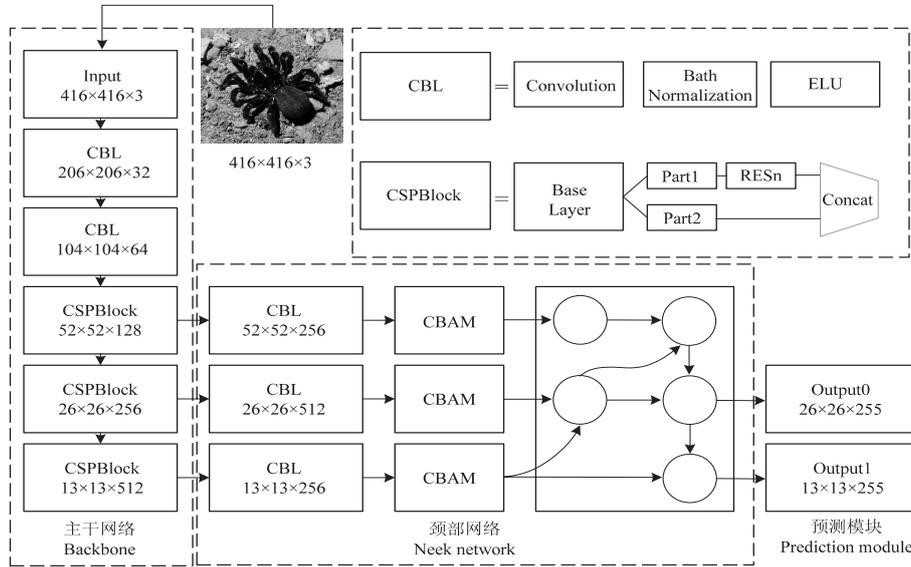


图 6 YOLOv4-tiny-ATO 节肢动物检测模型

## 2 实验方法与分析

### 2.1 模型训练环境与方法

所有模型使用深度学习框架 Pytorch1. 10. 0 + cu113,加速环境为 CUDA11. 0, GPU 型号为 NVIDIA GTX3060,显存 6 GB, CPU 型号为 AMD Ryzen 7 5800H CPU @3. 2 GHz。batchsize 设置为 16,图片输入尺寸 416×416,使用余弦退火算法进行训练,最小学习率为 0. 001,最大学习率为 0. 01,每个实验均迭代 70 个 epoch。

### 2.2 数据集

使用公开带标签的节肢动物图像数据集,称其为节肢动物数据集 (Arthropod Taxonomy Orders Object Detection Dataset, ArTaxOr),所描述的节肢动物种类超过 130 万种,总共有 15 374 张图。包括双翅目 (Diptera) (蚊子、蠓等)、鳞翅目 (Lepidoptera) (蝴蝶、飞蛾)、蜘蛛目 (Araneae) (成蛛、幼蛛)、膜翅目 (Hymenoptera) (蚂蚁、蜜蜂、黄蜂)、鞘翅目 (Coleoptera)

(甲虫)、蜻蜓目 (Odonata) (蜻蜓,豆娘等)、半翅目 (Hemiptera) (蚜虫、蝉、飞虱、盾蝽等)七类,部分数据集图像如图 7 所示。将其按比例 8 : 2,随机划分为训练集 (12 234 张)和验证集 (3 140)张。

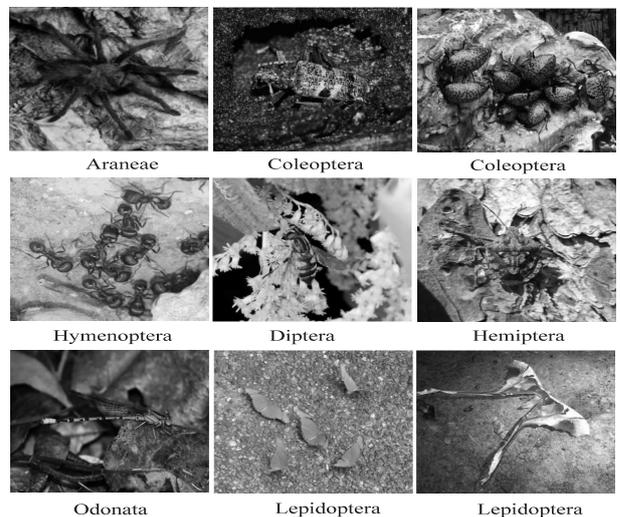


图 7 部分数据集图像

### 2.3 模型评价指标

该文规定  $IoU=0.5 \sim 0.95$  作为判断阈值,采用平均精度(Average Precision, AP)、平均精度均值(mean Average Precision, mAP)、检测速度(FPS)、模型大小(Model size)和召回率(Recall)5种性能指标对各实验进行评价。其中,召回率为被检测出的正样本,如式6所示。AP是针对一组IoU阈值分别为0.5~0.95时计算的,当前类别下模型Recall为0~1时的平均精度,即以精确率、召回率绘成的PR曲线下的面积,计算方法如式7所示。FPS指每秒检测的图像数量,单位为帧·s<sup>-1</sup>。mAP是每类别平均精度AP的均值,如式8所示。

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$AP = \int_0^1 P(r) dr \quad (7)$$

$$mAP = \frac{\sum_{q=1}^Q AP(q)}{Q} \quad (8)$$

### 2.4 YOLOv4-tiny-ATO 模型收敛性能

YOLOv4-tiny-ATO 模型,加载 YOLOv4-tiny 预训练权重,实验损失曲线如图8所示。损失在0~10

表 1 消融实验模型结构及结果

YOLOv4-tiny	Feat@3	CBAM	改进 BiFPN	DCN	mAP	Recall	FPS
√					0.299	0.311	132.7
√	√				0.357	0.332	127.5
√	√	√			0.371	0.345	124.3
√	√	√	√		0.529	0.406	115.4
√	√	√	√	√	0.725	0.585	89.6

根据实验数据来看,在 YOLOv4-tiny 模型基础上,从特征融合网络中引出 Feat@3,但对模型性能的提升并不明显,主要原因在于浅层特征虽然能提取更多的特征信息,同时也会忽略更多特征信息中的无效信息及噪声。YOLOv4-tiny\_Feat@3\_CBAM 模型嵌入注意力机制能够在不增加计算量的前提下,一定程度上改善对目标的关注度,快速获取高语义特征信息。YOLOv4-tiny\_Feat@3\_CBAM\_BiFPN 模型引入改进的特征金字塔结构,仅占用较少计算资源,有效融合不同层级的信息,从而进行多尺度特征表示。YOLOv4-tiny\_Feat@3\_CBAM\_BiFPN\_DCN 模型在引入 DCN 后,模型通过反向传播更好的自适应学习目标形态特征,显著提高了模型的检测性能。

### 2.6 YOLOv4-tiny 改进前后的对比结果

改进前后模型进行实验对比,效果如图9所示。YOLOv4-tiny 模型出现较多的误检、漏检以及预测锚框位置不准确等情况,改进后的 YOLOv4-tiny-ATO

epoch 明显下降,后续快速趋于平稳,说明余弦退火算法匹配性强,改进后的模型训练效果较好。当 epoch 为 50 时, val loss 和 train loss 接近收敛,在第 epoch 为 70 时模型具有良好的稳定性,训练损失基本收敛,模型性能达到最佳,实验表明该改进方法是可行的。

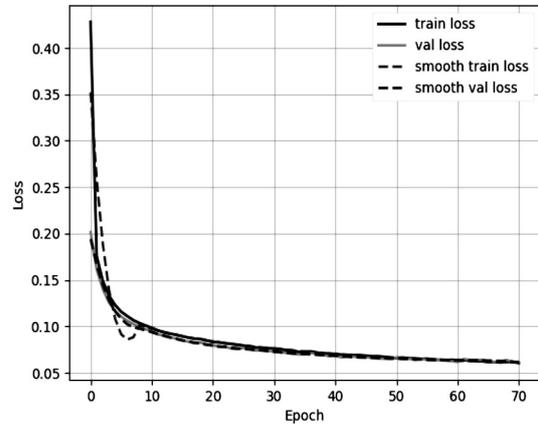


图 8 训练损失和验证损失曲线

### 2.5 YOLOv4-tiny-ATO 模型的消融实验

消融实验采用余弦退火算法进行训练,为了量化 YOLOv4-tiny 模型,引入以上改进方法对节肢动物检测,逐一增添机制与模型结合,实验结果如表 1 所示。

模型对于不同环境下的目标检测效果明显提升,弥补了原模型对密集目标检测性能差的缺陷,同时提高了对拟态动物的检测效率,预测锚框定位也更加准确。

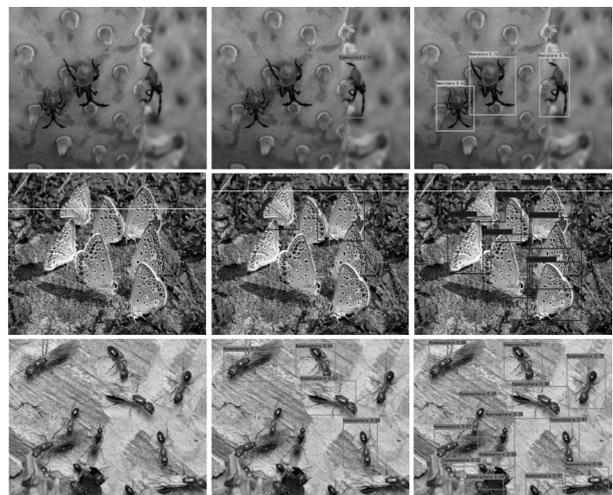


图 9 改进前后模型检测效果

图 9 改进前后模型检测效果

### 2.7 不同模型的对比实验

为进一步验证提出的 YOLOv4-tiny-ATO 算法的优越性,对比了目前主流单阶段 end-to-end 的全卷积神经网络模型 RetinaNet、YOLO 系列目标检测算法 YOLOv4 和 YOLOv5-l<sup>[16]</sup>以及现阶段节肢动物检测效果最优模型 AROD RCNN<sup>[17]</sup>。实验数据如表 2 所示,利用各类 AP、Model size、mAP、recall 和 FPS 对模型进行评估。

从对比数据可看出,文中模型对节肢动物检测有针对性的改善,提出的 YOLOv4-tiny-ATO 模型在大小仅为 54.6 Mb 的前提下,很好地平衡了整体的性能,检测精度和召回率分别为 0.725 和 0.585,检测速度达到 89.6 帧·s<sup>-1</sup>。因此,在模型大小、检测速度上更适用于移动端部署,检测精度也满足对节肢动物的检测需求。

### 2.8 类激活图可视化分析

基于梯度定位的 Grad-CAM<sup>[18]</sup>网络可视化结构,通过类激活图呈现出可视化图像,能观察出模型对图像中感兴趣的区域。模型改进前后的热力图对比结果如图 10 所示,两个模型进行对比后发现,YOLOv4-tiny 模型对单一目标关注度比 YOLOv4-tiny-ATO 更高,YOLOv4-tiny-ATO 模型对密集目标和重叠目标

均可较好地定位关注。从热力图分析可看出,改进的模型具有以下特性:

- (1)在复杂背景前提下,能有效区分噪声边界,提取出具有高语义的特征信息;
- (2)对于复杂背景,被遮挡目标的图像,模型特征提取效率高,目标边界清晰;
- (3)模型关注度匹配于节肢动物多重特性,有效区分出待检测目标。

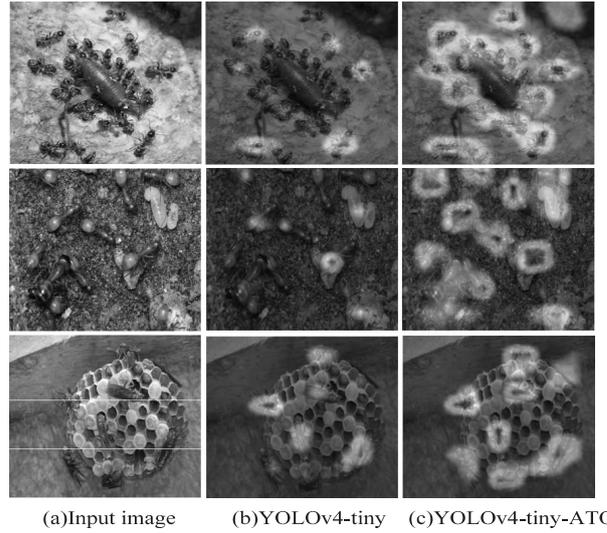


图 10 改进前后模型类激活图

表 2 对比主流检测模型结果

模型	蜘蛛目	鞘翅目	双翅目	半翅目	膜翅目	鳞翅目	蜻蜓目	Model size	mAP	recall	FPS
Fast R-CNN	0.349	0.202	0.211	0.239	0.111	0.259	0.476	256.3	0.264	0.484	21.3
RetinaNet	0.518	0.421	0.451	0.502	0.225	0.397	0.746	80.0	0.465	0.477	19.7
Yolov4	0.722	0.617	0.579	0.672	0.516	0.605	0.811	244.3	0.646	0.526	83.2
Yolov4-tiny	0.297	0.221	0.197	0.201	0.153	0.276	0.513	22.6	0.299	0.311	132.7
Yolov5-l	0.767	0.602	0.591	0.691	0.522	0.631	0.801	91.0	0.659	0.620	147.2
AROD RCNN	0.808	0.693	0.661	0.753	0.561	0.689	0.857	326.5	0.717	0.787	7.2
Yolov4-tiny-ATO	0.825	0.697	0.674	0.758	0.571	0.683	0.869	54.6	0.725	0.585	89.6

### 3 结束语

基于轻量型易部署的 YOLOv4-tiny 目标检测框架,提出一种自然环境复杂背景下改进的节肢动物检测算法。通过替换原模型特征融合网络为改进的 BiFPN,有效融合高语义信息,并在其网络颈部引出 Feat@3,提高模型特征提取感受野,嵌入 CBAM 减少感受野扩大带来的背景噪声干扰,改变原有卷积方式为 DCN,自适应融合临近形状和位置特征,使检测感受野随特征的变化而变化。针对节肢动物门下 7 类物种,在模型大小、检测速度和检测精度达到平衡的前提下,能够一定程度上解决背景复杂、数量密集及重叠等

问题,有效提升了模型的检测性能。

通过实验结果证明,提出的各项方法在 YOLOv4-tiny 模型上进行改进的有效性,改进后的 YOLOv4-tiny-ATO 模型,识别节肢动物的平均精度达到 0.725,召回率为 0.585,检测速度为 89.6 帧·s<sup>-1</sup>,且模型大小仅为 54.6 Mb,较 YOLOv4-tiny 原始模型,mAP 提高 0.426,Recall 提高 0.274。对比目前节肢动物最优模型 AROD RCNN,大小仅占其 16.7%,mAP 提升 0.008,FPS 提高 82.4 帧·s<sup>-1</sup>。该模型对自然环境下的节肢动物检测性能良好,移动部署性强,可为节肢动物鉴定识别、生态监控和减少病虫害损失等方向的研究提供参考。

## 参考文献:

- [1] NANSEN C, MURDOCK M, PURINGTON R, et al. Early infestations by arthropod pests induce unique changes in plant compositional traits and leaf reflectance[J]. *Pest Management Science*, 2021, 77(11): 5158–5169.
- [2] 郭胜娟, 刘 峰. 基于目标检测技术的农业有害昆虫识别方法研究[J]. *农业开发与装备*, 2022(9): 149–153.
- [3] HSU W Y, LIN W Y. Adaptive fusion of multi-scale YOLO for pedestrian detection[J]. *IEEE Access*, 2021, 9: 110063.
- [4] QIAN Huaming, WANG Huilin, FENG Shuai, et al. FESSD: SSD target detection based on feature fusion and feature enhancement[J]. *Journal of Real-Time Image Processing*, 2023, 20(1): 5–8.
- [5] ZHU M, HU G, ZHOU H, et al. Arbitrary-oriented ship detection based on retinanet for remote sensing images[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 6694–6706.
- [6] LI S, LI D, WANG B. A visual identification method of analog instrument panel based on faster R-CNN[C]//东北大学, 中国自动化学会信息物理系统控制与决策专业委员会. 第34届中国控制与决策会议论文集(5). 出版者不详, 2022.
- [7] 林达坤. 基于灰狼优化算法的鳞翅目昆虫图像识别研究[D]. 福州: 福建农林大学, 2019.
- [8] 林达坤, 黄世国, 张飞萍, 等. 基于改进差分进化算法的鳞翅目昆虫图像识别方法[J]. *林业科学*, 2020, 56(3): 73–81.
- [9] 谢娟英, 侯 琦, 史颖欢, 等. 蝴蝶种类自动识别研究[J]. *计算机研究与发展*, 2018, 55(8): 1609–1618.
- [10] 彭明杰, 唐万梅, 皮家甜, 等. 改进的基于端到端学习的蜻蜓目昆虫识别算法[J]. *重庆师范大学学报: 自然科学版*, 2021, 38(3): 94–99.
- [11] CUI Mingdi, LOU Yunyi, GE Yilin, et al. LES-YOLO: a lightweight pinecone detection algorithm based on improved YOLOv4-Tiny network[J]. *Computers and Electronics in Agriculture*, 2023, 205: 79–85.
- [12] 张 蕊, 孟晓曼, 曾志远, 等. 图卷积神经网络在点云语义分割中的研究综述[J]. *计算机工程与应用*, 2022, 58(24): 29–46.
- [13] CHEN Yantong, ZHANG Xianzhong, CHEN Weinan, et al. Research on recognition of fly species based on improved RetinaNet and CBAM[J]. *IEEE Access*, 2020, 8: 2816–2818.
- [14] YU Hongxia, YUN Lijun, CHEN Zaiqing, et al. A small object detection algorithm based on modulated deformable convolution and large kernel convolution[J]. *Computational Intelligence and Neuroscience*, 2023, 2023: 237–245.
- [15] QU Jiantao, YIN Junru, JIANG Yuhao, et al. A sub-pixel convolution-based improved bidirectional feature pyramid network for pansharpening[J]. *Remote Sensing Letters*, 2023, 14(1): 3–19.
- [16] MOU Xingang, LEI Shuai, ZHOU Xiao. YOLO-FR: a YOLOv5 infrared small target detection algorithm based on feature reassembly sampling method[J]. *Sensors*, 2023, 23(5): 667–675.
- [17] 郭子豪, 董乐乐, 曲志坚. 基于改进 Faster RCNN 的节肢动物目标检测方法[J]. *计算机应用*, 2023, 43(1): 88–97.
- [18] SURESH K, SHILPA G, BISWAJEET P, et al. Explainable AI in scene understanding for autonomous vehicles in unstructured traffic environments on Indian roads using the inception U-Net model with Grad-CAM visualization[J]. *Sensors*, 2022, 22(24): 9714–9725.