光电技术及应用

DOI：10.16818/j.issn1001－5868.2024062401

基于 LID-YOLO 的小目标昆虫轻量化检测算法

陈中举*，李和平，许浩然

（长江大学 计算机科学学院，湖北 荆州 434023）

摘　要： 针对复杂背景下新疆棉田昆虫识别误检率高、小目标昆虫难以检测等问题，提出了一种基于 YOLOv5s 改进的轻量化昆虫检测模型（Lightweight Insect Detection-YOLO，LID-YOLO）。首先，主干网络使用 GhostNet 网络替换原 CSPDarknet53 网络，并采用 Slim-Neck 模块对颈部网络进行改进，以实现模型轻量化；其次，引入 BottleNet Transformer 融合模块，减少模型参数量并增强网络特征提取能力，更好检测小目标；最后，加入 NAM 注意力机制，通过应用权重稀疏性惩罚抑制不显著权重来提取细节特征，提高模型准确率。实验结果表明，LID-YOLO 模型在参数量、计算量、模型权重大小方面，相比 YOLOv5s 模型，其分别减少了 30.9%，45.6% 和 29.7%。LID-YOLO 模型的准确率达到了 97.4%，检测速度为 55.25 FPS，与原 YOLOv5s 模型相比，提高了 1 个百分点和 2.62 FPS。LID-YOLO 模型在保证轻量化的同时进一步提高了检测精度，更好满足农作物昆虫实际检测的需要。

关键词： YOLOv5s；BottleNet Transformer；NAM；轻量化网络；目标检测

中图分类号： TP391　**文章编号：** 1001－5868(2024)06－0945－08

Light Weight Detection Algorithm for Small Target Insects Based on Lightweight Insect Detection-YOLO

CHEN Zhongju，LI Heping，XU Haoran

（School of Computer Science，Yangtze University，Jingzhou 434023，CHN）

Abstract： Aiming at the problem of low insect identification accuracy and difficult detection of small target insects against the complex background of cotton fields in Xinjiang，a lightweight insect detection model based on YOLOv5s（LID-YOLO）is proposed. First，the GhostNet network is used to replace the original cross-stage partial CSPDarknet53 network in the backbone，and the Slim-Neck module is used to improve the neck network，to achieve a lightweight model. Second，the fusion module BottleNet Transformer is introduced to reduce the number of model parameters and enhance the capability of network feature extraction to better detect small targets. Finally，the normalization-based attention module（NAM）is added to extract detail features by applying a sparse weight penalty to suppress non-significant weights and improve model accuracy. The experimental results show that compared with YOLOv5s，the LID-YOLO model reduces the number of parameters，calculations，and model weight by 30.9%，45.6%，and 29.7% respectively. The accuracy rate of LID-YOLO model reached 97.4%，and detection speed was 55.25 FPS，which is 1% point and 2.62 FPS higher than that of the original YOLOv5s model. The LID-YOLO model not only ensures lightweight，but also improves

收稿日期：2024－06－24.

基金项目：中国高校产学研创新基金项目(2023IT269).

* **通信作者：**陈中举

detection accuracy to better meet the requirements of crop insect detection.

Key words: YOLOv5s; BottleNet Transformer; NAM; lightweight network; target detection

0 引言

棉花是我国重要的经济作物之一，棉花产业在国家经济中占据重要地位[1]。新疆是我国重要的棉花生产基地，棉花在整个生长周期中会受到棉铃虫、臭虫等昆虫的危害，严重影响其产量。因此，随着棉花种植面积的扩大，能够快速准确地识别出棉田中的昆虫种类对棉花生产至关重要。近年来，众多学者针对昆虫检测方法进行研究，传统的计算机视觉方法在农作物昆虫识别领域取得了重要研究成果。如刘岩等[2]为准确识别玉米图像中的有害昆虫区域，通过中值滤波和 K-means 算法对图像进行去噪和分割处理，并采用 Gentle AdaBoost 算法筛选特征，构建分类器，排除背景值，利用 SIFT 识别目标。肖志云等[3]为快速、准确识别马铃薯中的有害昆虫，利用高频协方差阵特征值、HELM 的纹理特征和空间域 Hu 的形状特征，进行支持向量机的昆虫识别。但这种基于人工的特征提取方法容易受到光照不均匀和背景复杂等问题的干扰，其检测速度和精度无法满足大田作物昆虫识别的要求。

随着智能信息化技术和深度学习的发展，利用卷积神经网络（Convolutional Neural Network，CNN）和 YOLO 算法对农作物昆虫识别的研究不断加深[4]。如 Jiao 等[5]针对小目标昆虫，提出一种改进 Fast R-CNN 端到端的检测方式，具有较好的检测精度。赵康迪等[6]为实现复杂背景下玉米草地贪夜蛾的图像识别，提出一种基于实例分割的卷积神经网络结合迁移学习的方法，测试准确率达到 96.32%。梁文哲等[7]针对低信噪比条件下昆虫检测性能差，通过短时傅里叶变换丰富图像特征，利用图像特征对图像识别，提出基于 YOLOv3 的昆虫目标检测算法。郑果等[8]为解决水稻有害昆虫体型小而难以识别的问题，提出一种将卷积块注意力和特征金字塔引入 YOLOv7 算法。高家军等[9]针对复杂场景下多幼虫个体图像识别分割困难问题，提出一种融合 Swin Transformer 图像分割方法。由上可知，深度学习用于检测农作物昆虫具有较好效果。为满足当前棉花昆虫识别的现实需求，提出一种轻量型 LID-YOLO（Lightweight Insect Detection-YOLO）网络结构，以 YOLOv5s 为基准网络模型，通过引入 ChostNet 网络和 Slim-Neck 模块，分别轻量化主干网络和颈部网络，减小模型参数量并且提高检测速度；引入 NAM 注意力机制和 BoTC3 融合模块，提升感知目标信息的能力，提高模型精确度。

1 YOLOv5 算法以及改进

1.1 网络模型

YOLOv5 是一个单阶段目标检测模型，按照网络的深度和宽度分为 YOLOv5n，YOLOv5s，YOLOv5m，YOLOv5l，YOLOv5x 五种模型，其中 YOLOv5s 模型具有简洁、高精度和低消耗的特点，可作为昆虫检测任务的基础模型。YOLOv5s 主要结构包括输入端（Input）、主干网络（Backbone）、颈部网络（Neck）和头部网络（Head）[10]。输入端对输入的图像数据进行增强，并进行自适应锚框计算；主干网络使用 CSPDarknet53 作为特征提取网络，获取多特征图；颈部网络通过结合特征金字塔网络（Feature Pyramid Network，FPN）[11]和路径聚合网络（Path Aggregation Network，PAN）加强特征提取；头部网络将检测任务分为 3 个不同尺度，通过 1×1 卷积输出。

针对 YOLOv5s 模型检测复杂背景下新疆棉田昆虫误检率高、小目标昆虫不容易检测等问题，提出了一种基于改进 YOLOv5s 的昆虫检测算法 LID-YOLO。改进点主要分为三点：(1)在主干网络中使用 C3Ghost 和 GhostConv 分别替换掉部分 C3 及部分 Conv，构成轻量化 GhostNet 网络，在颈部网络中使用 GSConv 和 VoVGSCSP 构成的 Slim_Neck 对网络进行轻量化改进；(2)在主干网络中使用 BoTC3 模块替换最后一个 C3，获取更多图像上下文信息，提高对昆虫目标的检测精确度；(3)在颈部网络引入 NAM 注意力机制。LID-YOLO 网络模型如图 1 所示。

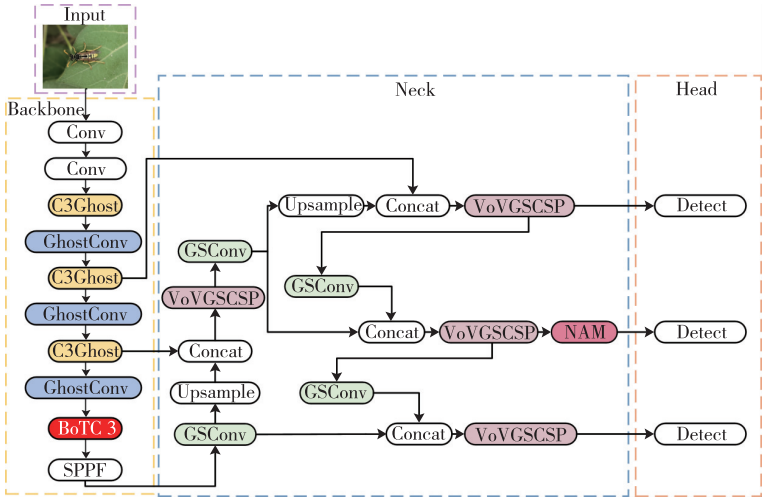


图 1　LID-YOLO 模型结构图

1.2　GhostNet 网络

为了降低 YOLOv5s 模型参数量，本研究以 Ghost 卷积为基础构建 GhostConv 和 C3Ghost 模块分别替换主干网络的 Conv 和 C3 模块，以达到模型轻量化目的。相比于 C3 模块，C3Ghost 模块中的 GhostBottleNeck 是一个轻量级的卷积块，可以提高模型的计算效率，结构如图 2 所示。

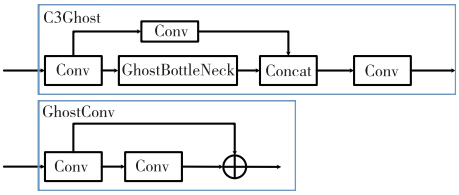


图 2　C3Ghost 和 GhostConv 结构图

Ghost 卷积是 Han[12] 等利用特征图的冗余性并通过一系列的线性变换，提出的一种新型的轻量级卷积模块。Ghost 卷积工作原理如图 3 所示。

与传统卷积的不同在于 Ghost 卷积首先在输入特征层使用标准的卷积生成一些真实的特征层；然后通过一系列廉价的线性操作形成另一部分特征层，并将这两个特征层结合在一起形成输出特征层。通过这种方式，Ghost 卷积能够使用更少的计算生成更多的特征映射，在减少参数量的同时保持了模型的特征提取能力。

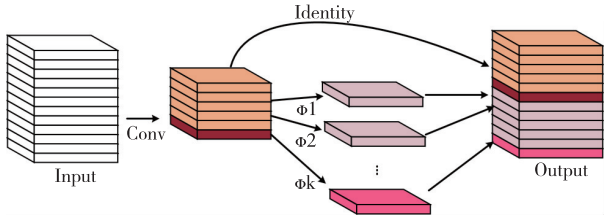


图 3　Ghost 卷积工作原理图

1.3　Slim-Neck 模块

深度可分离卷积（Depth-wise Separable Convolution，DSC）是一种常见的轻量化卷积，可减少模型计算量，但也会使检测精度大幅下降。为使模型更加轻量化并保持精度，将 Slim-Neck 引入到 YOLOv5s 的颈部网络中。Slim-Neck 具有较快的速度和较高的准确率，在目标检测领域有较好的表现[13]。在 Slim-Neck 模块中首先使用轻量级的卷积 GSConv 代替传统卷积 Conv。GSConv 是由标准卷积、DSC 和随机排序所构成的卷积结构，GSConv 的网络结构如图 4 所示。

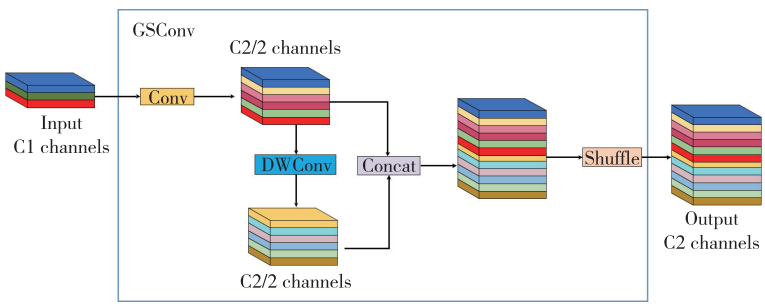


图 4　GSConv 模块结构图

GSConv 模块将输入的特征图像进行标准卷积实现跨通道信息交互，接着进行 DWConv 操作，使结构变得高效，然后将标准卷积和 DWConv 特征图 Concat 拼接，最后进行 Shuffle 通道混洗操作，使标准卷积生成的特征信息渗透到 DSC 内部，提高特征丰富程度。

Slim-Neck 是在 GSConv 和 VoVGSCSP 的基础上进一步扩展形成的，在 Neck 层中使用 GSConv 代替标准卷积 Conv，使用如图 5 所示的 VoVGSCSP 模块替换其 C3 结构，组成 Slim-Neck。Slim-Neck 模块是一种适用于目标检测的特征融合模块，可以有效减少模型的参数量和计算量[14]。

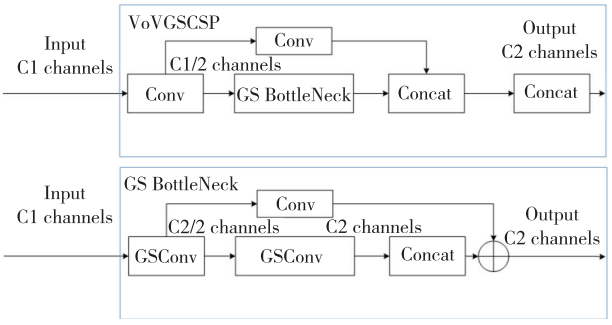


图 5 VoVGSCSP 模块结构图

1.4 BottleNet Transformer 模块

BottleNet Transformer(BoT)[15]模块是由卷积神经网络和 Transformer 组合成的模型，该结构将自注意力机制应用于目标检测任务中。BoT 模块是在 ResNet 基础上改进的模型，将 ResNet BottleNeck[16]中的 3×3 卷积替换为多头注意力机制(Multi-Head Self Attention，MHSA)。在主干网络末端 C3 模块基础上引入 BoT 模块从而形成 BoTC3 模块，不仅减少了参数量和计算量，而且提高了网络的上下文感知能力。ResNet BottleNeck 和 BottleNet Transformer 结构如图 6 所示。

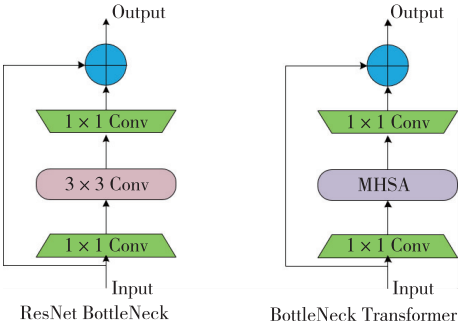


图 6 ResNet BottleNeck 和 BottleNeck Transformer 结构对比图

BoT 模块的核心为 MHSA 结构，其 MHSA 结构与传统 Transformer 中的 MHSA 非常相似，区别在于 BoT 模块的位置编码部分多了两个可学习的相对位置编码 R_h 和 R_w，分别表示高度和宽度 2 个维度。BoT 模块的 MHSA 结构如图 7 所示。

BoT 模块的 MHSA 计算过程如下：

$$Attention(q,k,v,r)=SoftMax(qk^{T}+qr^{T})v \quad (1)$$

式中，q，k，v，r 分别为查询、键、值和位置编码。通过引入 r 可以加速模型的收敛，更关注位置信息。

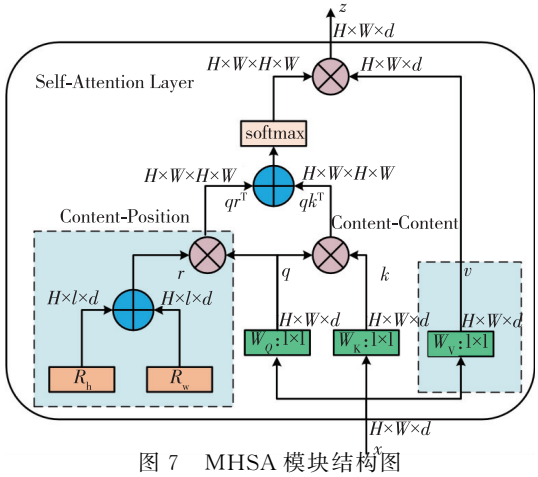


图 7 MHSA 模块结构图

1.5 NAM 注意力机制

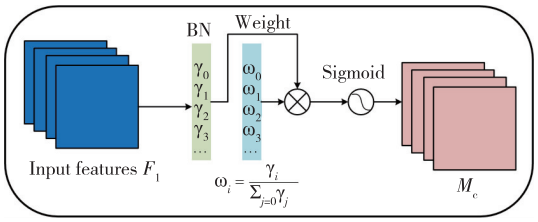
YOLOv5s 模型轻量化后，在大幅减少其参数量和计算量的同时，也容易导致其忽略特征图的细节信息，使模型精度下降。为了提高模型的检测精度，将 NAM 注意力[17]模块嵌入到检测模型的颈部网络中，使模型更好地提取图像的细小特征。

NAM 是一种轻量级且高效的注意力机制，采用 CBAM 的模块集成方式，重新设计了通道注意力和空间注意力子模块，利用权重的贡献因子来提升注意力的性能。对于通道注意力子模块，使用批量归一化(BN)中的比例系数，并由比例系数计算通道的方差来测量权重的重要程度，如公式(2)所示为：

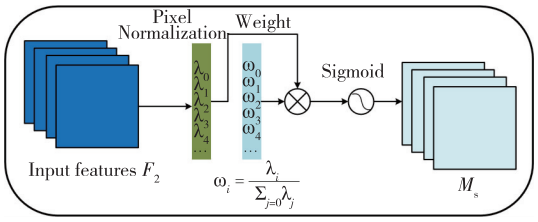
$$B_{out}=BN(B_{in})=\gamma\frac{B_{in}-\mu_B}{\sqrt{\sigma_\beta^2+\varepsilon}}+\beta \quad (2)$$

式中，μ_B 表示小批量 B 的值；σ_B 表示小批量 B 的标准差；γ，β 分别表示可训练的放射变化参数(尺度和位移)；ε 表示较小的常数，防止分母为 0。对于空间注意子模块，将 BN 的比例系数在空间维度上测量像素重要程度，权重 w_i 反应空间的变化程度，表示空间的重要性。NAM 的通道注意力模块如图 8(a)所示，其中 F_1 表示通道注意力输入特征图，W_γ 表示每个通道的权值，M_c 表示通道注意力输出的特

征图；空间注意力模块如图 8(b)所示，其中 F_2 表示空间注意力输入特征图，W_λ 表示空间像素的权值，M_s 表示空间注意力输出特征图。



（a）通道注意力机制



（b）空间注意力机制

图 8　NAM 的通道和空间注意力模块

$$W_\gamma=\gamma_i\Big/\sum_{j=0}^{n}\gamma_j \tag{3}$$

$$W_\lambda=\lambda_i\Big/\sum_{j=0}^{n}\lambda_j \tag{4}$$

$$M_c=Sigmoid(W_\gamma(BN(F_1))) \tag{5}$$

$$M_s=Sigmoid(W_\lambda(BN_s(F_2))) \tag{6}$$

2　实验数据与设置

2.1　实验环境和数据来源

本研究基于 Windows 11 系统，该系统搭载的型号为 NVIDIA GeForce RTX 3060 的 GPU 和型号为 AMD Ryzen 7 6800H 的 CPU，深度学习框架选择 PyTorch 2.0.0，运行环境为 Python 3.11。输入图片的尺寸为 640×640，采用马赛克数据增强，优化器采用 SGD 算法，并将权重衰减设置为 0.000 5。设置训练时的迭代次数为 300，初始学习率为 0.01，动量因子为 0.937，Batch Size 为 8。

实验采用新疆棉田主要昆虫图像数据集[18]，在不同时间、地点和环境中采集的图像数据共有 3 225 张，主要包含瓢虫、臭虫、棉铃虫、草蛉、食蚜蝇、蜜蜂。将数据集随机按 80%，10%，10%的比例分为训练集、验证集和测试集。

2.2　评价指标

为了衡量模型的检测效果，将平均精度均值（mAP）、参数量、计算量（GFLOPs）、模型大小和检测速度（Frames Per Second，FPS）作为评估改进模型性能的指标。其相关表达式如下：

$$Precision=\frac{TP}{TP+FP} \tag{7}$$

$$Recall=\frac{TP}{TP+FN} \tag{8}$$

$$AP=\int_0^1 P(R)\mathrm{d}R \tag{9}$$

$$mAP=\frac{1}{c}\sum_{i=1}^{c}AP_i \tag{10}$$

$$FPS=\frac{Framenum}{ElapsedTime} \tag{11}$$

公式(7)和(8)中 TP 表示检测目标中正确数量，FP 表示检测目标中错误数量，FN 表示未能检测出的目标。公式(9)中 AP 表示各个类别的检测精度，AP 的值是以 R 为 x 轴，以 P 为 y 轴的 P-R 曲线下的面积。公式(10)中 mAP 表示所有类别的识别平均精度均值，由精确度（P）、召回率（R）和 AP 计算得到，$mAP@0.5$ 是所有类别的识别平均准确率 $AP@0.5$ 相加后取平均值得到的指标。

参数量表示训练过程中总的参数数量。计算量可以衡量模型的复杂程度，单位为 G。模型体积表示模型训练完成后保存的权重文件大小，单位为 MB。FPS 表示模型每秒推理的图像帧数。

3　实验结果与分析

3.1　改进模型前后对比

将改进的模型和原 YOLOv5s 模型在参数量、计算量、检测速度、模型大小和 mAP 方面进行对比。由表 1 可知，LID-YOLO 模型较原 YOLOv5s 模型参数量减少 30.9%，计算量减少 45.6%，检测速度达到 55.25 FPS，模型体积减少 29.7%，mAP 提升 1 个百分点。因此，LID-YOLO 模型在各个方面优于 YOLOv5s。

表 1　改进前后性能对比

模型	参数量	*GFLOPs*	*FPS*	模型大小/MB	*mAP*/%
YOLOv5s	7026307	15.8	52.63	13.7	0.964
LID-YOLO	4852107	8.6	55.25	9.63	0.974

本文还分别对改进前后的 YOLOv5s 模型的检测精度进行对比，如图 9 所示。蜜蜂提升最高，达到了 3.7 个百分点，棉铃虫提升了 3.1 个百分点，草蛉提升了 0.6 个百分点，臭虫提升了 0.4 个百分点，食蚜蝇保持在精度 99.5%，瓢虫精度下降 1.9 个百分点。从实验整体效果来看，改进后的模型识别更加准确，检测精度更高。

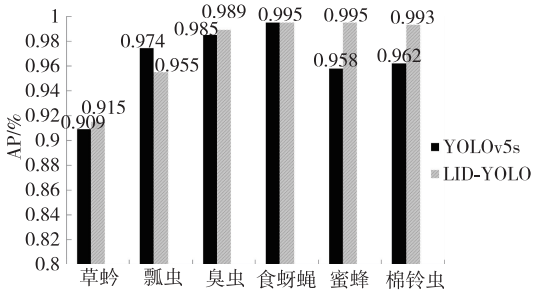


图 9 改进前后检测精度对比图

3.2 模型轻量化对比实验

在 YOLOv5s 模型的基础上，使用不同的轻量化网络替换其主干网络进行实验对比。其中包括 MobileNetV3Small[19]，ShuffleNetV2[20] 和 GhostNet。由表 2 可知，YOLOv5s-ShuffleNetV2 的参数量和计算量最少，但是检测速度较慢，精度较低；YOLOv5s-MobileNetV3Small 的检测速度最慢，精度最低；YOLOv5s-GhostNet 的参数量和计算量较高，但相较原 YOLOv5s 模型已大幅度减少，并且检测速度最快，精度值最高，略微增加模型参数量换取检测精度高的方案是可取的。因此，选用 GhostNet 网络替换 YOLOv5s 主干网络。

表 2 模型轻量化对比

模型	参数量	*GFLOPs*	*FPS*	*mAP*/%
YOLOv5s	7026307	15.8	52.63	0.964
YOLOv5s-ShuffleNetV2	3191111	5.9	47.39	0.916
YOLOv5s-MobileNetV3Small	5031585	11.3	46.3	0.895
YOLOv5s-GhostNet	5100875	11	54.35	0.93

3.3 不同模型对比实验

为进一步验证 LID-YOLO 模型的检测性能，将其与目前主流的目标检测模型 YOLOv3，YOLOv5s[21]，YOLOv7-tiny，YOLOv7 和 YOLOv8n 进行比较，各个模型的检测结果如表 3 所示。其中最新的 YOLOv8n 模型检测精度较低，YOLOv7 模型的检测精度最高，但参数量、计算量和模型体积最大；YOLOv7-tiny 模型相较 YOLOv7，虽然在参数量、计算量等方面有所下降，但精度值也大幅下降。YOLOv3 模型在参数量上最少，模型体积最小，但计算量最大，达到 154.6，mAP值最低，为 91.6%。LID-YOLO 模型，相较 YOLOv3，YOLOv5s，YOLOv7-tiny 以及 YOLOv8n，检测精度分别提升 5.8，1，0.8，1.5 个百分点。对于昆虫检测问题，应该在保持高检测精度的前提下，尽可能减小模型的参数量、计算量和模型体积。实验结果表明，LID-YOLO 模型在实现轻量化的同时提高了检测精度，是可行的方案。

表 3 不同模型对比

模型	参数量	*GFLOPs*	模型大小/MB	*mAP*/%
YOLOv3	2176850	154.6	4.23	0.916
YOLOv5s	7035811	15.8	13.7	0.964
YOLOv7-tiny	6028518	13.2	11.7	0.966
YOLOv7	37223526	105.2	71.3	0.988
YOLOv8n	3006818	8.1	5.97	0.959
LID-YOLO	4852107	8.6	9.63	0.974

如图 10 所示为各个模型的 mAP 的训练曲线变化图，训练时每个模型采用统一的参数，经过 300 轮次的训练，LID-YOLO 模型具有较好的精度值和收敛速度。

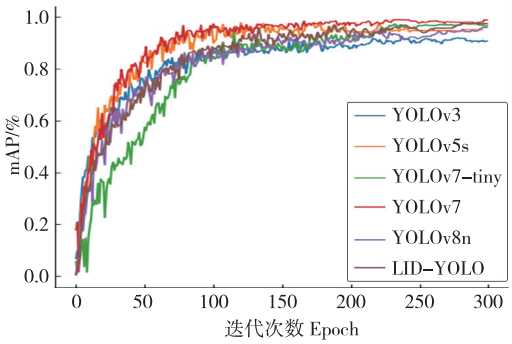


图 10 模型训练对比图

3.4 多种注意力机制对比实验

为了更好地验证 NAM 注意力机制的有效性，模型轻量化后，在其颈部网络嵌入四种目前主流的注意力机制，分别为 SE[22]，CA[23]，CBAM[24]，NAM。根据表 4 可知，通过嵌入 SE，CA，CBAM 注意力机制后 mAP 值均有不同程度的下降，而在嵌入 NAM 注意力机制后 mAP 提升了 1.7 个百分点。NAM 注意力在参数量和 mAP 上均优于其他注意力机制。因此，在轻量化模型上嵌入 NAM 注意力机制能够有效提高模型检测精度。

表 4 多种注意力机制对比

网络	参数量	*GFLOPs*	*mAP*/%
YOLOv5s+GhostNet+Slim-Neck+BoTC3	4560587	8.4	0.957
YOLOv5s+GhostNet+Slim-Neck+BoTC3+SE	4852107	8.6	0.947
YOLOv5s+GhostNet+Slim-Neck+BoTC3+CA	4858275	8.6	0.954
YOLOv5s+GhostNet+Slim-Neck+BoTC3+CBAM	4893569	8.6	0.942
YOLOv5s+GhostNet+Slim-Neck+BoTC3+NAM	4852107	8.6	0.974

3.5　消融实验

以 YOLOv5s 为基准模型，通过引入 GhostNet 网络和 Slim-Neck 轻量化模块、引入 BottleNet Transformer 模块和嵌入 NAM 注意力机制方法对其进行改进。为了验证各部分改进模块的有效性，进行了消融实验并与 YOLOv5s 原模型进行对比，实验结果如表 5 所示。

表 5　消融实验

模型	参数量	*GFLOPs*	*mAP*/%
YOLOv5s	7026307	15.8	0.964
＋GhostNet	5100875	11	0.93
＋GhostNet＋Slim-Neck	4565467	8.5	0.94
＋GhostNet＋Slim-Neck＋BoTC3	4560587	8.4	0.957
＋GhostNet＋Slim-Neck＋BoTC3＋NAM	4852107	8.6	0.974

根据表 5 可知，在主干网络引入 GhostNet 轻量化网络后，模型的参数量下降 27.4%，计算量下降 30.4%，同时 mAP 值下降 3.4 个百分点。在颈部网络加入 Slim-Neck 模块后，模型的参数量和计算量分别继续下降 10.4% 和 22.7%，检测精度 mAP 值有所提高。加入 BoTC3 融合模块后，不仅模型精度提高了 1.7 个百分点，同时参数量和计算量减少。引入 NAM 注意力机制后，虽然参数量和计算量略微增加，但 mAP 值得到了较高的提升，LID-YOLO 模型检测精度达到了 97.4%。由此可见，实验所采用的改进方法是有效的。

为了验证 LID-YOLO 模型的性能，将改进前后的模型在测试集上测试，部分检测结果如图 11 所示。可以看出，YOLOv5s 在复杂背景下识别昆虫图像存在误检情况，对小目标昆虫识别存在漏检问题。这是因为 YOLOv5s 主干网络更关注局部特征，导致模型特征提取能力较差，模型未能利用输入特征的其他细节信息。而 LID-YOLO 模型在轻量化的同时引入 BottleNet Transformer 模块，加强上下文信息的联系，嵌入 NAM 注意力机制，更加关注提取的细节特征，解决了以上问题。综上，LID-YOLO 模型检测效果优于原模型，误检率低、置信度高，同时对小目标昆虫有更好的检测效果。检测结果验证了 LID-YOLO 模型的有效性。

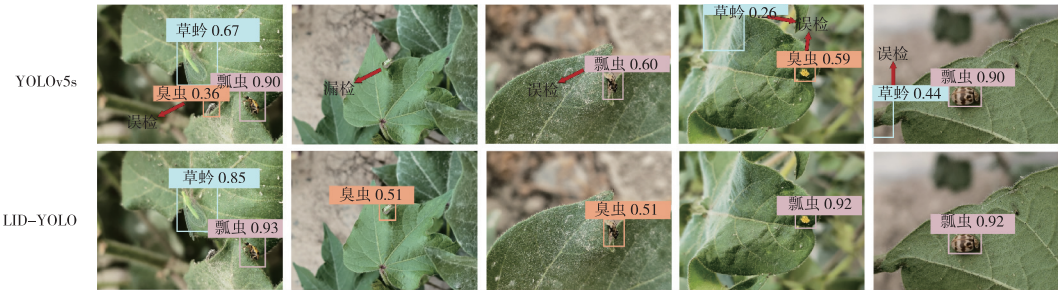


图 11　检测结果

4　结论

针对在复杂背景下新疆棉田昆虫图像识别存在误检、漏检等问题，提出了一种轻量化棉田昆虫检测方法 LID-YOLO，对常见的棉田昆虫进行研究、实验，得出如下结论：(1)通过融合 GhostNet 网络和 Slim-Neck 模块，大幅度减少参数量和计算量，使模型轻量化，引入 BoTC3 模块和 NAM 注意力机制，提升模型对重要特征的关注，提高模型的检测速度和准确率；(2)为验证 LID-YOLO 模型的优势，将其与 YOLOv3，YOLOv5s，YOLOv7-tiny，YOLOv7 和 YOLOv8n 等主流模型进行对比实验，结果表明，LID-YOLO 模型综合性能最佳。综上所述，本文提出的 LID-YOLO 模型实现了模型轻量化并提高了检测精度，能较好地满足实时性和精确度要求，具有明显的应用价值，可为农业昆虫识别提供参考。

参考文献：

[1]　赵立新，侯发东，吕正超，等. 基于迁移学习的棉花叶部病虫害图像识别[J]. 农业工程学报，2020，36(7)：184-191.
Zhao Lixin，Hou Fadong，Lü Zhengchao，et al. Image recognition of cotton leaf diseases and pests based on transfer learning[J]. Transactions of the Chinese Society of Agricultural Engineering，2020，36(7)：184-191.

[2]　刘　岩，张宁宁，海　玲，等. 基于机器视觉的玉米虫害区域 SIFT 识别仿真[J]. 计算机仿真，2023，40(7)：215-219.
Liu Yan，Zhang Ningning，Hai Ling，et al. Research on maize pest detection based on instance segmentation[J]. Computer Simulation，2023，40(7)：215-219.

[3]　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.

[4] 谭鑫平，高志辉，韩航迪，等. 基于改进 YOLOv5 的荧光图像细胞智能检测研究[J]. 半导体光电，2023，44(5)：709-716.
Tan Xinping，Gao Zhihui，Han Hangdi，et al. Intelligent detecrion of cells in fluorescence images based on improved YOLOv5[J]. Semiconductor Optoelectronics，2023，44(5)：709-716.

[5] Jiao L，Dong S，Zhang S，et al. AF-RCNN：An anchor-free convolutional neural network for multi-categories agricultural pest detection[J]. Computers and Electronics in Agriculture，2020，174：105522.

[6] 赵康迪，单玉刚，袁 杰，等. 基于实例分割的玉米虫害检测研究[J]. 河南农业科学，2022，51(12)：153-161.
Zhao Kangdi，Shan Yugang，Yuan Jie，et al. Research on maize pest detection based on instance segmentation [J]. Journal of Henan Agricultural Sciences，2022，51(12)：153-161.

[7] 梁文哲，冯阳凯，王 锐，等. 低信噪比下基于 YOLOv3 的昆虫目标检测[J]. 信号处理，2022，38(1)：109-117.
Liang Wenzhe，Feng Yangkai，Wang Rui，et al. Insect target detecction based on YOLOv3 under low SNR[J]. Journal of Signal Processing，2022，38(1)：109-117.

[8] 郑 果，姜玉松，沈永林. 基于改进 YOLOv7 的水稻害虫识别方法[J]. 华中农业大学学报，2023，42(3)：143-151.
Zheng Guo，Jiang Yusong，Shen Yonglin. Recognition of rice pests based on improved YOLOv7[J]. Journal of Huazhong Agricultural University，2023，42(3)：143-151.

[9] 高家军，张 旭，郭 颖，等. 融合 Swin Transformer 的虫害图像实例分割优化方法研究[J]. 南京林业大学学报(自然科学版)，2023，47(3)：1-10.
Gao Jiajun，Zhang Xu，Guo Ying，et al. Research on the optimized pest image instance segmentation method based on the Swin Transformer model[J]. Journal of Nanjing Forestry University：Natural Sciences Edition，2023，47(3)：1-10.

[10] 储 鑫，李 祥，罗 斌，等. 基于改进 YOLOv4 算法的番茄叶部病害识别方法[J]. 江苏农业学报，2023，39(5)：1199-1208.
Chu Xin，Li Xiang，Luo Bin，et al. Identification method of tomato leaf diseases based on improved YOLOv4 algorithm[J]. Jiangsu Journal of Agricultural Sciences，2023，39(5)：1199-1208.

[11] Lin T Y，Dollar P，Girshick R，et al. Feature pyramid networks for object detection [C]// Proc. of the IEEE Conference on Computer Vision and Pattern Recognition，2017：2117-2125.

[12] Han K，Wang Y，Tian Q，et al. Ghostnet：More features from cheap operations [C]// Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition，2020：1580-1589.

[13] Li H，Li J，Wei H，et al. Slim-neck by GSConv：A better design paradigm of detector architectures for autonomous vehicles[J]. arXiv preprint arXiv：2206.02424，2022.

[14] 李尚平，卞俊析，李凯华，等. 基于改进 YOLOv5s 的复杂环境下蔗梢分叉点识别与定位[J]. 农业机械学报，2023，54(11)：247-258.
Li Shangping，Bian Junxi，Li Kaihua，et al. Identification and height localization of sugarcane tip bifurcation points in complex environments based on improved YOLOv5s[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(11)：247-258.

[15] Srinivas A，Lin T Y，Parmar N，et al. Bottleneck transformers for visual recognition[C]// Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition，2021：16519-16529.

[16] He K，Zhang X，Ren S，et al. Deep residual learning for image recognition[C]// Proc. of the IEEE Conference on Computer Vision and Pattern Recognition，2016：770-778.

[17] Liu Y C，Shao Z R，Teng Y Y，et al. NAM：normalization based attention module[EB/OL]. [2022-02-03].

[18] 杨满仙，陈燕红，李雨晴，等. 新疆棉田主要昆虫图像数据集 CottonInsect [DS/OL]. V2. Science Data Bank，2023 [2023-12-20].
Yang Manxian，Chen Yanhong，Li Yuqing，et al. Image dataset of major insects in cotton fields of Xinjiang CottonInsect [DS/OL]. V2. Science Data Bank，2023[2023-12-20].

[19] Howard A，Sandler M，Chu G，et al. Searching for mobilenetv3[C]//Proc. of the IEEE/CVF Inter. Conf. on Computer Vision，2019：1314-1324.

[20] Zhang X，Zhou X，Lin M，et al. Shufflenet：An extremely efficient convolutional neural network for mobile devices [C]// Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition，2018：6848-6856.

[21] 代少升，曾 奇，黄 炼，等. 基于 S3-YOLOv5s 的矿井人员防护设备检测算法研究[J]. 半导体光电，2023，44(1)：153-160.
Dai Shaosheng，Zeng Qi，Huang Lian，et al. Research on detection algorithm of mine personnel protection equipment based on S3-YOLOv5s [J]. Semiconductor Optoelectronics，2023，44(1)：153-160.

[22] Hu J，Shen L，Sun G. Squeeze-and-excitation networks [C]//Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition，2018：7132-7141.

[23] Hou Q，Zhou D，Feng J. Coordinate attention for efficient mobile network design[C]// Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition，2021：13713-13722.

[24] Woo S，Park J，Lee J Y，et al. Cbam：Convolutional block attention module [C]// Proc. of the European Conf. on Computer Vision (ECCV)，2018：3-19.

作者简介：

陈中举(1976—)，男，副教授，硕士生导师，研究方向为计算机网络、人工智能、计算机视觉。