

文章编号:1007-2780(2022)09-1216-12

## 基于深度学习的林业害虫检测优化

赵 严, 刘应安\*, 业巧林, 周小亮

(南京林业大学 信息科学技术学院, 江苏 南京 210037)

**摘要:**目前林业害虫检测研究多数基于传统机器学习算法,且存在精度低、效果差的问题。对此,提出了一种基于深度学习模型 YOLOv4 的林业害虫检测模型——Pest-YOLOv4。采用 K-means++ 算法聚类先验框,提高了先验框 avg-IoU 值。将 ECA(Efficient Channel Attention)和 CBAM(Convolutional Block Attention Module)结合,构成 ECA-CBAM 注意力机制,使网络更多关注有利于林业害虫检测的特征信息。重新组织网络颈部,构成 SPP-PANet,融合多重感受野捕获的特征信息。利用 Focal Loss 思想改进损失函数,在平衡正负样本比例的同时关注难区分样本的学习。实验结果表明, Pest-YOLOv4 林业害虫检测模型 mAP 达到 90.4%,相较于 YOLOv4 提高 4.2%,FPS 保持在 33.4 f/s,满足林业害虫检测任务的检测精度与实时性要求。

**关键词:**林业害虫检测;深度学习;Pest-YOLOv4;注意力机制;Focal Loss

**中图分类号:**TP391.4 **文献标识码:**A **doi:**10.37188/CJLCD.2022-0077

## Forestry pest detection optimization based on deep learning

ZHAO Yan, LIU Ying-an\*, YE Qiao-lin, ZHOU Xiao-liang

(College of Information Science and Technology, Nanjing Forestry University, Nanjing 210037, China)

**Abstract:** At present, most of the research on forestry pest detection is based on traditional machine learning algorithms, and there are some problems such as low accuracy and poor effect. Therefore, a new forestry pest detection model based on YOLOv4, Pest-YOLOv4, is proposed. K-means++ algorithm is used to cluster anchor boxes to obtain a higher avg-IoU value. Combining ECA (Efficient Channel Attention) and CBAM (Convolutional Block Attention Module) to form ECA-CBAM attention mechanism, the network can pay more attention to the characteristic information. Reorganizing the network neck to SPP-PANet, the feature information can be fused with multiple receptive fields effectively. Using Focal Loss to improve the loss function, the learning of samples that are difficult to distinguish is paid attention, and the proportion of positive samples and negative samples are balanced. The experimental results show that Pest-YOLOv4 mAP reaches 90.4%, 4.2% higher than YOLOv4, while the FPS remains at 33.4 f/s, which meets the detection accuracy and real-time requirement of forestry pest detection tasks.

**Key words:** forestry pest detection; deep learning; Pest-YOLOv4; attention mechanism; Focal Loss

收稿日期:2022-03-08;修订日期:2022-03-30.

基金项目:国家自然科学基金(No. 62072246)

Supported by National Natural Science Foundation of China(No. 62072246)

\*通信联系人, E-mail: lyastat@163.com

## 1 引言

林业害虫泛指危害森林及林产品的昆虫。林业害虫的侵袭或寄生会导致林木生长发育不良,产品质量下降,生态环境恶化。2020年中国生态环境状况显示,全国主要林业有害生物发生面积为1 278.45万公顷,比2019年上升3.4%,虫害发生面积为790.62万公顷<sup>[1]</sup>。林业害虫检测是虫情监测的关键环节,其方法研究对于林业发展有重要意义。

近年来,深度学习发展迅速,卷积神经网络(Convolutional Neural Network, CNN)<sup>[2]</sup>凭借自动提取特征的优势逐渐替代人工特征提取方法<sup>[3-5]</sup>。研究者开始将深度学习同昆虫检测相结合。2018年,Z. Liu等人<sup>[6]</sup>首次将深度学习目标检测模型应用于识别储粮害虫稻象甲和玉米象。结果表明,Inception-ResNet-v2作特征提取网络,Faster R-CNN<sup>[7]</sup>作检测器,mAP达到85.76%。可是该研究需要额外制作预训练数据集;采用两步预测模型,计算开销大。2020年,J. Liu等人<sup>[8]</sup>利用图像金字塔改进YOLOv3<sup>[9]</sup>,实现多尺度特征的融合,并在不同图片尺寸下进行番茄病虫害检测。2021年,袁哲明等人<sup>[10]</sup>提出基于深度学习的轻量化田间昆虫检测模型,以YOLOv3为基础融合深度可分离卷积和联合交并比。2021年,J. W. Chen等人<sup>[11]</sup>利用多种目标检测方法对3类害虫进行基于移动平台的检测研究,结果表明YOLOv4<sup>[12]</sup>取得了最高的检测精度。2021年,M. Zha等人<sup>[13]</sup>提出一种轻量化YOLOv4林业害虫检测方法,该方法以添加CA注意力机制<sup>[14]</sup>的MobileNetv2作特征提取网络,利用BiFPN和ASFF进行特征融合,在北京林业大学制作的数据集上mAP达到88.93%。

本文以深度学习模型YOLOv4为基础进行林业害虫检测优化研究,无需制作预训练数据集且为端到端单阶段模型,相较于Faster R-CNN等双阶段模型,具备一定的检测实时性。为进一步提高林业害虫检测精度,本文给出优化的林业害虫检测模型Pest-YOLOv4:

(1)为CSPDarknet-53特征提取网络引入ECA-CBAM注意力机制。使网络更多地关注有利于检测的特征信息,在一定程度上缓解特征提取信息丢失问题。

(2)重新组织网络颈部,构成SPP-PANet。融合多重感受野捕获的特征信息,在缓解梯度消失的同时增强特征提取信息的表达能力。

(3)利用Focal Loss思想改进损失函数,缓解小样本检测任务存在的正负样本不平衡问题,加强网络对难区分样本的学习。

此外,采用K-means++算法重新聚类先验框,提高了avg-IoU值。

## 2 YOLOv4

### 2.1 YOLOv4模型结构

YOLOv4网络结构分为3个部分:

(1)特征提取网络。相较于YOLOv3的特征提取网络,YOLOv4引入CSP(Cross Stage Partial connections)结构<sup>[15]</sup>,构成CSPDarknet-53特征提取网络。

(2)网络颈部。采用SPP<sup>[16]</sup>模块和PANet<sup>[17]</sup>。YOLOv4采用优化的SPP模块进行多尺度池化,获取了更丰富的语义信息。PANet在FPN<sup>[18]</sup>后添加了自底向上融合部分,从而缩短底层和顶层特征之间的信息路径。

(3)网络检测器。沿用YOLOv3的YOLO Head,利用CIoU损失<sup>[19]</sup>改进边界框回归损失。

### 2.2 YOLOv4损失函数

YOLOv4损失函数由置信度损失、类别损失和边界框回归损失3个部分构成。3部分及YOLOv4总损失分别如式(1)、(2)、(3)、(6)所示。

$$\begin{aligned} \text{Loss}_{\text{conf}} = & -\lambda_{\text{obj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} [\hat{C}_i^j \ln(C_i^j) + \\ & (1 - \hat{C}_i^j) \ln(1 - C_i^j)] - \\ & \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{noobj}} [\hat{C}_i^j \ln(C_i^j) + \\ & (1 - \hat{C}_i^j) \ln(1 - C_i^j)], \end{aligned} \quad (1)$$

$$\begin{aligned} \text{Loss}_{\text{cls}} = & -\lambda_{\text{class}} \sum_{i=0}^{S^2} I_{ij}^{\text{obj}} \sum_{c \in \text{classes}} [\hat{P}_i(c) \ln(P_i(c)) + \\ & (1 - \hat{P}_i(c)) \ln(1 - P_i(c))], \end{aligned} \quad (2)$$

$$\text{Loss}_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{\text{gt}})}{k^2} + \alpha\nu, \quad (3)$$

$$\nu = \frac{4}{\pi^2} \left( \arctan \frac{w^{\text{gt}}}{h^{\text{gt}}} - \arctan \frac{w}{h} \right)^2, \quad (4)$$

$$\alpha = \frac{\nu}{1 - \text{IoU} + \nu}, \quad (5)$$

$$\text{Loss} = \text{Loss}_{\text{conf}} + \text{Loss}_{\text{cls}} + \text{Loss}_{\text{CIoU}}, \quad (6)$$

式中,  $\lambda_{\text{obj}}$ 、 $\lambda_{\text{noobj}}$  和  $\lambda_{\text{class}}$  分别为正样本置信度权重、负样本置信度权重和类别损失权重;  $S^2$  为划分的网格数量;  $B$  为先验框数量;  $I_{ij}^{\text{obj}}$  和  $I_{ij}^{\text{noobj}}$  用于判断第  $i$  网格的第  $j$  先验框内样本的正负性, 若为正样本则  $I_{ij}^{\text{obj}}$  和  $I_{ij}^{\text{noobj}}$  分别为 1 和 0, 否则分别为 0 和 1;  $\text{classes}$  集合为类别集合;  $\hat{C}_i$  和  $C_i^t$  分别为第  $i$  网格存在正样本的实际置信度和预测置信度;  $\hat{P}_i(c)$  和  $P_i(c)$  分别为第  $i$  网格类别为  $c$  的实际概率和预测概率;  $\text{IoU}$  为真实框和预测框的交并比;  $(b^{\text{gt}}, w^{\text{gt}}, h^{\text{gt}})$  和  $(b, w, h)$  分别为真实框和预测框的中心坐标、宽、高;  $\rho^2(b, b^{\text{gt}})$  表示  $b$  和  $b^{\text{gt}}$  两点间欧式距离;  $k$  为同时覆盖真实框和预测框的最小包围框的对角线长度。

### 3 Pest-YOLOv4

#### 3.1 K-means++先验框聚类

为了匹配实验数据集样本尺寸, 需要重新聚类先验框的宽高比。Pest-YOLOv4 采用 K-means++ 算法进行聚类。K-means++ 相较于 K-means 而言, 采用了使初始聚类中心间的相互距离尽可能远的策略。Pest-YOLOv4 结构图如图 1 所示。算法流程如下:

输入: 集合  $S, K$

- (1) 从集合  $S$  中随机选取一个点作为初始聚类中心  $C_i$ ;
  - (2) 计算  $S$  中每个点  $x$  至其最近聚类中心  $C_x$  的距离  $D(x)$ , 依据  $D(x)$  生成各点成为下一聚类中心的概率  $P(x)$ 。依据各点  $P(x)$  使用加权随机抽样选取一个新的聚类中心  $C_i$ ;
  - (3) 重复步骤 (2), 直到找到  $K$  个聚类中心;
  - (4) 计算集合  $S$  中每个点至  $K$  个聚类中心的距离  $D(x)$  并将其划入  $D(x)$  最小的聚类中心所在簇。重新生成每个簇的聚类中心  $C_i$ ;
  - (5) 重复步骤 (4), 直至聚类中心  $C_i$  不再变化。
- 输出:  $K$  组聚类中心

集合  $S$  为林业害虫宽高集合;  $K$  为需要生成的先验框数量。距离度量  $D(x)$  采用样本间 IoU 值。距离度量  $D(x)$ 、聚类中心概率  $P(x)$ 、重新生成聚类中心  $C_i$  的计算公式分别如式 (7)、(8)、(9) 所示。

$$D(x) = 1 - \text{IoU}(x, C_x), \quad (7)$$

$$P(x) = \frac{D(x)^2}{\sum_{x \in S} D(x)^2}, \quad (8)$$

$$C_i = \frac{\sum_{x \in C_i} x}{|C_i|}, \quad (9)$$

式中,  $\text{IoU}(A, B)$  为真实框  $A$  和真实框  $B$  的交并比。

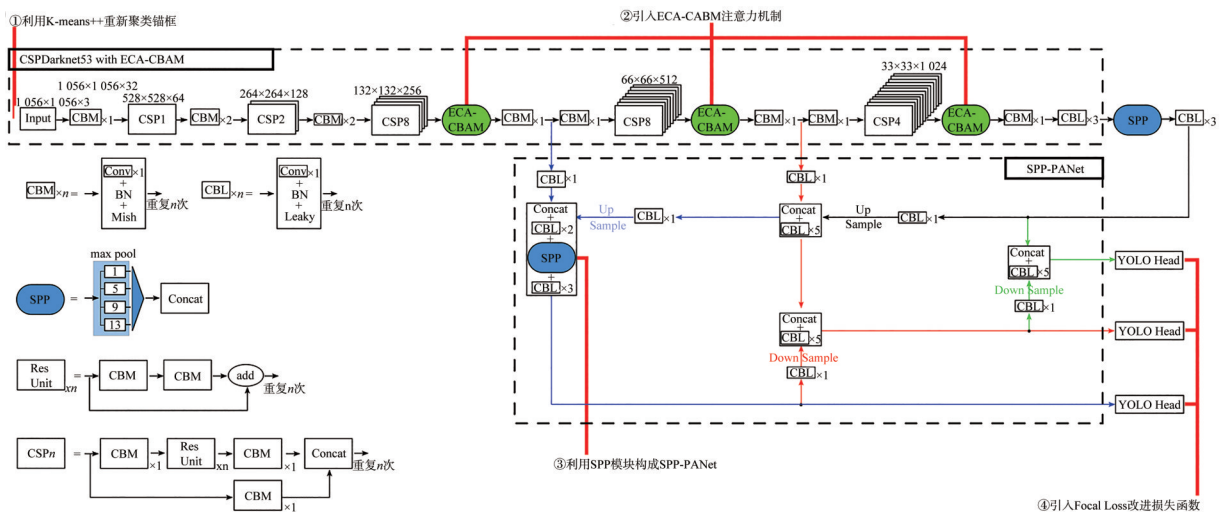


图 1 Pest-YOLOv4 结构图

Fig. 1 Structure of Pest-YOLOv4

### 3.2 ECA-CBAM

为深度学习模型增加注意力机制可使模型更多地关注有利于检测的特征信息,从而提升模型检测精度<sup>[20-21]</sup>。Pest-YOLOv4提出ECA-CBAM注意力机制,该注意力机制基于CBAM(Convolutional Block Attention Module)<sup>[22]</sup>模块和ECA(Efficient Channel Attention)<sup>[23]</sup>模块实现。

CBAM为卷积块注意力模块,包括通道域(Channel Attention Module, CAM)和空间域(Spatial Attention Module, SAM)两个部分。CAM结构、SAM结构、CBAM连接方式如图2所示。

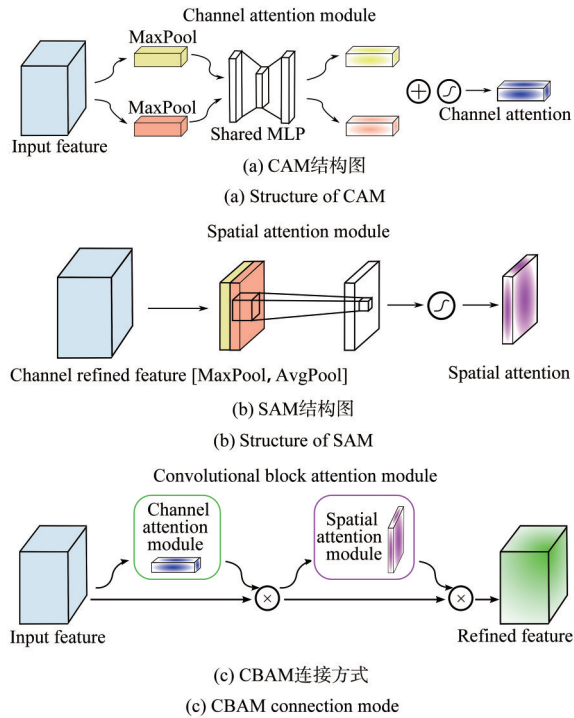


图2 CBAM结构图

Fig. 2 Structure of CBAM

CBAM内CAM采用的全连接层被设计为捕捉非线性跨通道交互,涉及降维以控制模型复杂度。降维对通道注意力权重预测产生了负面影响,并且没有必要捕获所有通道之间的依赖关系<sup>[23]</sup>。ECA模块在不降低维数的通道级全局平均池化后,通过尺寸为 $k$ 的一维卷积,利用每个通道及其 $k-1$ 个邻居来捕获局部跨通道交互信息。该方法保证了模型效率,提高了计算效果。ECA模块如图3所示,其中 $k=3$ 。

$k$ 代表参与一个通道注意力权重预测的相邻

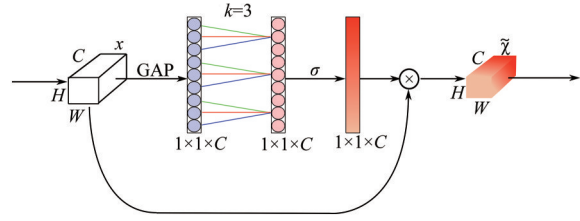


图3 ECA模块结构图

Fig. 3 Structure of ECA

通道数。 $k$ 通过通道维度 $C$ 的映射自适应地确定,计算公式如式(10)所示:

$$k = \left\lceil \frac{\log_2 C + b}{\gamma} \right\rceil_{\text{odd}}, \quad (10)$$

式中, $\lceil t \rceil_{\text{odd}}$ 为距离 $t$ 最近的奇数; $\gamma$ 和 $b$ 分别为2和1。

ECA模块可有效降低CAM参数量和计算量,并避免CAM降维操作带来的通道注意力权重预测效果损失,其结构如图3所示。Pest-YOLOv4基于ECA模块和CBAM内的SAM部分,构成ECA-CBAM模块,结构如图4所示。

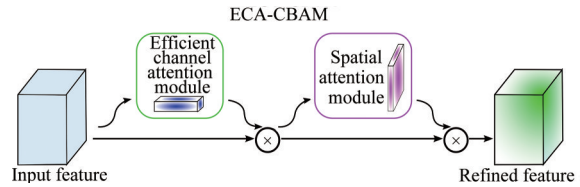


图4 ECA-CBAM模块结构图

Fig. 4 Structure of ECA-CBAM

ECA-CBAM改进CBAM原有CAM存在的问题,并保留CBAM串联合通道注意力模块和空间注意力模块的策略。ECA-CBAM对特征图依次进行通道注意力加权和空间注意力加权,操作步骤如式(11)、(12)所示。

$$F' = M_c(F) \otimes F, \quad (11)$$

$$F'' = M_s(F') \otimes F', \quad (12)$$

式中, $F$ 为ECA-CBAM的输入特征图, $M_c(\cdot)$ 为ECA通道注意力权重计算公式, $M_s(\cdot)$ 为SAM空间注意力权重计算公式, $\otimes$ 为元素乘法。

$M_c(\cdot)$ 和 $M_s(\cdot)$ 计算方式如式(13)、(14)所示。

$$M_c(F) = \sigma(f^{1 \times k}(\text{AvgPool}(F))), \quad (13)$$

$$M_s(F) = \sigma(f^{3 \times 3}([\text{AvgPool}(F); \text{MaxPool}(F)])), \quad (14)$$

式中,  $\sigma(\cdot)$  为 Sigmoid 激活函数,  $f^{a \times b}(\cdot)$  为核尺寸  $a \times b$  的卷积操作,  $k$  由式 (10) 确定,  $\text{AvgPool}(\cdot)$  为平均值池化操作,  $\text{MaxPool}(\cdot)$  为最大值池化操作。

在 ECA-CBAM 与网络结合的过程中, 本文分析 YOLOv4 网络框架, 在特征提取网络内添加注意力模块时, 合适的位置是在 CSP 模块后及特征图引入 PANet 和 SPP 前。该处的 ECA-CBAM 可以增强 CSP 模块提取的特征信息表达能力, 提高 PANet 的特征融合效果, 使 SPP 模块融合的多尺度特征信息向有利于林业害虫检测的方向发展。共计选择 3 个添加位点, 如图 1 中的 ② 所示。

### 3.3 SPP-PANet

YOLOv4 引入优化后的 SPP 模块, 以小的计算开销为代价, 大幅提高了 mAP。该模块结构如图 5 所示。

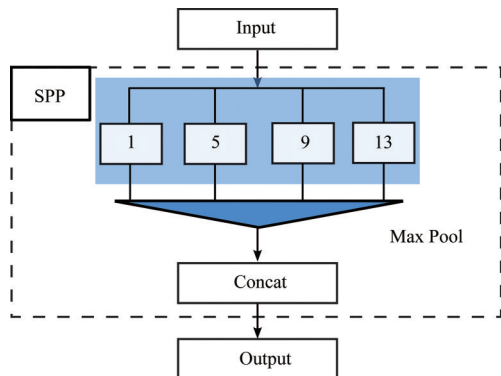


图 5 SPP 模块结构图

Fig. 5 Structure of SPP

特征图输入后并联进入核尺寸为  $\{13, 9, 5, 1\}$ 、步长为 1 的最大值池化, 拼接融合, 即得到 SPP 模块的输出。SPP 模块利用 4 个不同尺度的最大值池化进行特征提取并拼接, 实现多尺度感受野信息的融合, 在显著增加网络感受野的同时分离出重要的上下文特征。

YOLOv4 仅对深层特征使用一次 SPP 模块以丰富特征图信息。此次信息丰富的效果直接作用于  $25 \times 25$  的 YOLO Head 并通过 PANet 深层特征上采样融合路径间接作用于  $50 \times 50$  和  $100 \times 100$  的 YOLO Head。文献 [24] 提出利用多 SPP 模块优化目标检测模型, 构建 YOLOv4-SPP3 模型, 在 VEDAI 小目标数据集上

mAP 提高 2.3% (mAP50 为 75.4%), 可见该策略具备提高检测效果的能力。但是 SPP 模块会将原特征图通道数扩增 4 倍。YOLOv4-SPP3 为每个 YOLO Head 都配置直接作用的 SPP 模块, 共计存在 3 个 SPP 模块并且引入额外的卷积操作, 极大地增加了模型计算开销。

YOLOv4 模型  $100 \times 100$  的 YOLO Head 距 SPP 模块位置远、受影响程度小, 且其前存在大量卷积操作, 为其引入直接作用的 SPP 模块将极大丰富该处特征图信息并缓解梯度消失现象。在网络第 134 层增加 SPP 模块, 构成 SPP-PANet。此时网络仅存在两个 SPP 模块并且未额外引入卷积操作, 有效控制了网络模型的计算开销。两 SPP 分别直接作用于  $25 \times 25$  和  $100 \times 100$  的 YOLO Head, 并且通过上采样和下采样路径共同间接作用于  $50 \times 50$  的 YOLO Head。SPP-PANet 结构如图 1 中的 ③ 所示。

### 3.4 Focal Loss 优化损失函数

单阶段检测模型为了保证算法的实时性, 在整张图片上回归预测物体坐标, 较双阶段模型减少了生成目标候选区域的过程。林业害虫检测存在小目标的特点, 背景像素居多。在此情况下, 大部分检测框位于背景区域, 正负样本不平衡问题显著。同时, 林业害虫检测存在难区分问题, 部分类别的害虫颜色、外观不易区分。Pest-YOLOv4 利用 Focal Loss<sup>[25]</sup> 缓解上述问题。

YOLOv4 置信度损失函数和类别损失函数采用二元交叉熵形式构成, 如式 (1)、式 (2) 所示。二元交叉熵一般形式如式 (15) 所示:

$$L_{ce} = \begin{cases} -\ln y, & \text{when } \hat{y} = 1 \\ -\ln(1 - y), & \text{when } \hat{y} = 0 \end{cases}, \quad (15)$$

式中,  $\hat{y}$  表示真实值,  $y$  表示预测值。

Focal Loss 在二元交叉熵形式上引入平衡因子  $\alpha$  和调制因子  $\gamma$ , 可以改善正负样本不平衡问题, 同时减少简单样本对网络的改变量, 增强网络对复杂样本的学习能力。

仅添加调制因子  $\gamma$  时, 其取值大于 0。损失函数形式如式 (16) 所示:

$$L_{fl-\gamma} = \begin{cases} -(1 - y)^\gamma \ln y, & \text{when } \hat{y} = 1 \\ -y^\gamma \ln(1 - y), & \text{when } \hat{y} = 0 \end{cases}. \quad (16)$$

当  $\gamma=2$  时, 预测值为 0.9 的正样本 (简单样本) 的损失值将降低至原 1/100, 预测值为 0.5 的

正样本(复杂样本)的损失值将降低至原1/4,实现了增加复杂样本的损失值占比,削弱了简单样本对模型的影响。

仅添加平衡因子 $\alpha$ 时,其取值范围为0.5~1。损失函数形式如式(17)所示:

$$L_{\text{fl}-\alpha} = \begin{cases} -\alpha \ln y, & \text{when } \hat{y} = 1 \\ -(1-\alpha) \ln(1-y), & \text{when } \hat{y} = 0 \end{cases} \quad (17)$$

当 $\alpha=0.75$ 时,正样本损失值将降低至原3/4,负样本损失值将降低至原1/4。此时,增加了正样本损失值占比,削弱了负样本对模型的改变量。

同时使用平衡因子 $\alpha$ 和调制因子 $\gamma$ 时,平衡因子 $\alpha$ 取值0~0.5。损失函数形式如式(18)所示:

$$L_{\text{fl}} = \begin{cases} -\alpha(1-y)^{\gamma} \ln y, & \text{when } \hat{y} = 1 \\ -(1-\alpha)y^{\gamma} \ln(1-y), & \text{when } \hat{y} = 0 \end{cases} \quad (18)$$

此时,调制因子 $\gamma$ 已实现对简单负样本影响的削弱,平衡因子 $\alpha$ 的作用在于保证简单样本的损失值不至于过小。训练中,平衡因子取值0.25,调制因子取值 $2^{[26-27]}$ 。

改进后的置信度损失、类别损失如式(19)、(20)所示。

$$\text{Loss}_{\text{conf-fl}} = -\lambda_{\text{obj}} \sum_{i=0}^{s^2} \sum_{j=0}^B I_{ij}^{\text{obj}} [\hat{C}_i^j \alpha (1 - C_i^j)^{\gamma} \ln(C_i^j) + (1 - \hat{C}_i^j) \ln(1 - C_i^j)] - \lambda_{\text{noobj}} \sum_{i=0}^{s^2} \sum_{j=0}^B I_{ij}^{\text{noobj}} [\hat{C}_i^j \ln(C_i^j) + (1 - \hat{C}_i^j)(1 - \alpha)(C_i^j)^{\gamma} \ln(1 - C_i^j)], \quad (19)$$

$$\text{Loss}_{\text{cls-fl}} = -\lambda_{\text{class}} \sum_{i=0}^{S^2} I_{ij}^{\text{obj}} \sum_{c \in \text{classes}} [\hat{P}_i(c) \alpha (1 - P_i(c))^{\gamma} \ln(P_i(c)) + (1 - \hat{P}_i(c)) P_i(c)^{\gamma} (1 - \alpha) \ln(1 - P_i(c))] \quad (20)$$

式中, $\alpha$ 为平衡因子, $\gamma$ 为调制因子,其余符号的解释同式(1)、式(2)。

## 4 实 验

### 4.1 数据集

实验采用北京林业大学制作的小蠹科林业害虫数据集<sup>[28]</sup>。该数据集由室内害虫图像自动采集装置创建。自动采集装置由显示器、诱捕器、拍摄系统构成。采集时用镊子将害虫放入诱捕器的收集器内,模拟野外条件下害虫诱捕环境。利用拍摄系统获取害虫图片,并借助显示器的高清成像捕捉害虫细节,共计获得图片2 183张。其中,1 693张作为训练集,245张作为验证集,245张作为测试集。数据集图片采集情况如图6所示。

数据集共有6类林业害虫,训练集中每类昆虫实例统计结果如表1所示。数据集图片分辨率为1 236×1 236,进行缩放调整至800×800。虫情测报仪利用特殊光谱的光源引诱害虫扑灯,凭借内部远红外线烘干机将虫体烘干。工作人员间隔固定时间通过接虫箱接虫并统计<sup>[29]</sup>。数据集中图片为收集器内拍摄所得,满足实际处理情况。

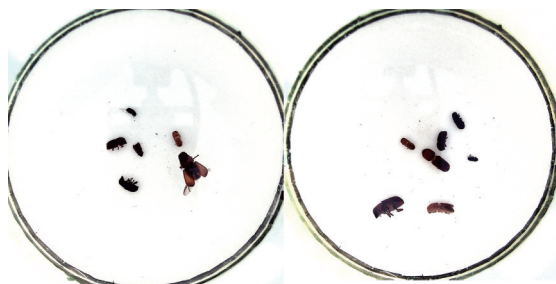


图6 北京林业害虫数据集展示

Fig. 6 Display of BJFU pest data set

表1 训练集昆虫样本统计结果

Tab. 1 Statistics of insects in the training set

| 昆虫名              | 样本数量  |
|------------------|-------|
| 红脂大小蠹(Leconte)   | 2 216 |
| 松十二齿小蠹(Boerner)  | 1 595 |
| 华山松大小蠹(Armandi)  | 1 765 |
| 四眼小蠹(Coleoptera) | 2 091 |
| 六齿小蠹(Acuminatus) | 953   |
| 云杉八齿小蠹(Linnaeus) | 1 727 |

### 4.2 实验平台和参数

实验运行在Ubuntu 18.04.5 LTS操作系统上,CPU为AMD R9 5900X,GPU为NVIDIA RTX

3060,使用Pytorch深度学习框架,编译环境为 cuda 11.0.228, torch 1.90, torchversion 0.10.0, numpy 1.21.1,Python 3.7。实验参数设置见表2。

表 2 实验参数设置

Tab. 2 Experimental parameters setting

| 参数                | 值     |
|-------------------|-------|
| 批大小               | 2     |
| 训练次数              | 250   |
| 动量                | 0.9   |
| 前 100 个 Epoch 学习率 | 0.01  |
| 100 Epoch 后学习率    | 0.001 |

4.3 评价指标

为了验证目标检测模型的有效性,本文主要采用查准率(Precision,  $P$ )、查全率(Recall,  $R$ )、平均准确率(Average Precision, AP)、平均准确率均值(mean Average Precision, mAP)、每秒检测帧数(Frames Per Second, FPS)等评价指标。

查准率用于描述在所有预测为正的样本中预测正确的比例,如式(21)所示。查全率用于描述在所有正例样本中有多少被成功检测出,如式(22)所示。平均准确率指查准率和查全率为坐标轴绘制的  $P$ - $R$  曲线围成面积。平均准确率均值指对所有类别的平均准确率取均值。平均准确率和平均准确率均值的计算公式如式(23)所示。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{21}$$

式中,TP表示预测为正、实际为正的样本;FP表示预测为正、实际为负的样本。

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{22}$$

式中,TP表示预测为正、实际为正的样本;FN表示预测为负、实际为正的样本。

$$\begin{cases} \text{AP} = \int_0^1 P(R) \text{d}R \\ \text{mAP} = \sum_{i=1}^N \text{AP}_i / N \end{cases}, \tag{23}$$

式中, $N$ 表示样本类别数; $P(R)$ 为查准率关于查全率的函数。

4.4 实验结果和分析

4.4.1 Kmeans++先验框优化

Pest-YOLOv4 使用 K-means++ 算法进行

先验框的重新聚类,并采用平均交并比(avg-IoU)评价聚类效果。K-means++算法和K-means算法聚类效果如图7所示。

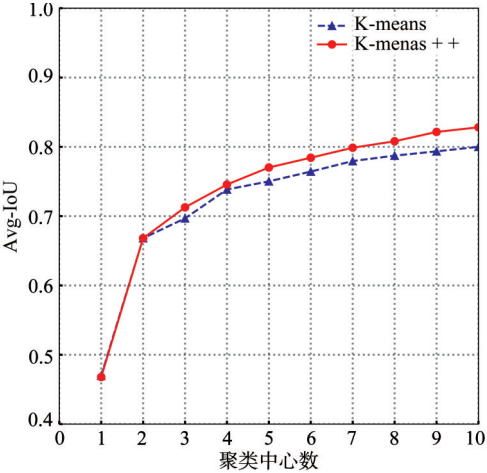


图 7 K-means++和K-means聚类效果图

Fig. 7 Clustering effects of K-means++ and K-means

由图7可得,当聚类中心数为9时,K-means++算法聚类先验框 avg-IoU 值为 82.15%,相较于 K-means 算法的 avg-IoU (78.35%)提高了 3.8%。此时先验框尺寸为(34, 50)、(50, 35)、(44, 72)、(73, 45)、(57, 75)、(70, 59)、(72, 115)、(115, 72)、(107, 108)。

为了进一步探究聚类获得的先验框有效性,分别采用两种聚类算法生成的先验框进行林业害虫检测,实验结果如表3所示。

表 3 先验框优化结果比较

Tab. 3 Comparison of anchor box optimization

| 网络模型   | 先验框<br>聚类策略 | Precision/<br>% | Recall/<br>% | mAP/% |
|--------|-------------|-----------------|--------------|-------|
| YOLOv4 | K-means     | 60.3            | 88.4         | 86.2  |
| YOLOv4 | K-means++   | 60.6            | 88.9         | 86.7  |

由表3可得,两种聚类方式生成的先验框对检测效果产生了影响,显著区别在于查全率和mAP指标。在不改动 YOLOv4 算法的情况下,利用 K-means++ 优化先验框提高了 YOLOv4 模型的检测精度,查全率增加 0.5%,mAP 值增加 0.5%。

4.4.2 Pest-YOLOv4

Pest-YOLOv4 和 YOLOv4 训练损失情况如图8所示。

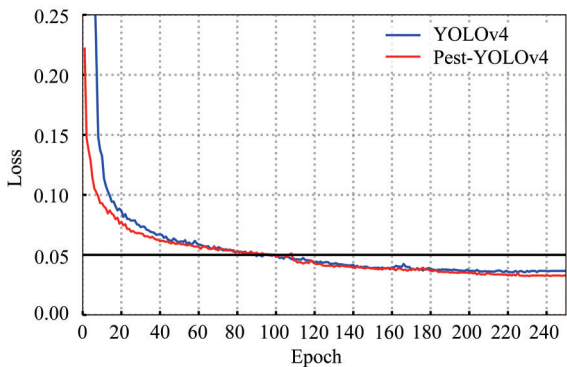


图 8 损失情况图  
Fig. 8 Loss curves

Pest-YOLOv4 模型在 109 个 Epoch 时,总损失值达到 0.045 59,低于 0.05。在 230 个 Epoch 时网络总损失值趋于稳定。比较 Pest-YOLO4 和 YOLOv4 模型收敛曲线可知,Pest-YOLO4 收敛速度略快、收敛情况更加平稳。

为了验证所提 Pest-YOLOv4 对林业害虫检测的有效性和先进性,采用目前目标检测中性能较优的 Faster R-CNN、SSD<sup>[30]</sup>、YOLOv3 及 YOLOv4 进行对比实验。实验结果如表 4 和图 9 所示。

Pest-YOLOv4 相较于 YOLOv4 在查准率、查全率和 mAP 指标上皆有不同程度的提高,最明显的提高指标为查准率(Precision 值 67.5%),提高了 7.2%。Pest-YOLOv4 和 Faster R-CNN 查全率接近。在查准率和 mAP 指标上,Pest-YOLOv4 分别高出 4.5% 和 0.9%。相较于 SSD 和 YOLOv3,Pest-YOLOv4 检测效果改善明显。同时,Pest-YOLOv4 具备检测实时性,FPS 指标达到

表 4 5 种模型林业害虫检测结果比较

Tab. 4 Comparison of five models for forestry pest detection

| 网络模型                    | Precision/<br>% | Recall/<br>% | FPS  |
|-------------------------|-----------------|--------------|------|
| Faster R-CNN(ResNet101) | 63.0            | 92.4         | 4.6  |
| SSD                     | 57.2            | 83.9         | 29.3 |
| YOLOv3                  | 58.4            | 85.7         | 34.5 |
| YOLOv4                  | 60.3            | 88.4         | 39.2 |
| Pest-YOLOv4             | 67.5            | 92.8         | 33.4 |

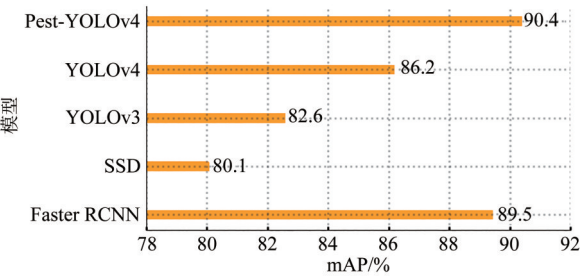


图 9 5 种模型林业害虫检测 mAP 比较

Fig. 9 mAP comparison of 5 models for forestry pest detection

33.4,远高于双阶段检测模型 Faster R-CNN,并高于 SSD 模型,同 YOLOv3 模型 FPS 接近。

Pest-YOLOv4 与 YOLOv4 对每类林业害虫的检测结果如表 5 所示。Pest-YOLOv4 在多数林业害虫种类上达到很好的检测效果,尤其是 Leconte 类别(AP 值 96.2%),但 Coleoptera 类别的识别效果一般(AP 值 86.2%)。这两类昆虫情况如图 10 所示。

表 5 Pest-YOLOv4 和 YOLOv4 检测结果

Tab. 5 Detection results of Pest-YOLOv4 and YOLOv4

| 昆虫名称       | YOLOv4      |          |      | Pest-YOLOv4 |          |      |
|------------|-------------|----------|------|-------------|----------|------|
|            | Precision/% | Recall/% | AP/% | Precision/% | Recall/% | AP/% |
| Leconte    | 71.5        | 94.0     | 95.4 | 81.9        | 95.3     | 96.2 |
| Boerner    | 57.5        | 89.1     | 88.9 | 81.3        | 96.1     | 94.5 |
| Armandi    | 65.9        | 86.2     | 84.8 | 68.6        | 93.4     | 87.6 |
| Coleoptera | 39.2        | 87.2     | 80.5 | 36.5        | 95.4     | 86.2 |
| Acuminatus | 52.5        | 82.9     | 79.8 | 51.2        | 94.9     | 87.8 |
| Linnaeus   | 75.4        | 91.1     | 88.0 | 85.3        | 81.9     | 90.3 |
| 均值         | 60.3        | 88.4     | 86.2 | 67.5        | 92.8     | 90.4 |

由图 10 可得,Coleoptera 类别相较于 Leconte 类别,尺寸更小、颜色外观更加不易区分。但是

Pest-YOLOv4 对此类林业害虫检测效果仍有明显提高(AP 值提高 5.7%),说明 Pest-YOLOv4

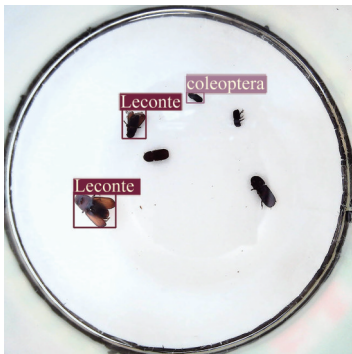


图 10 Leconte 和 Coleoptera 采集情况  
Fig. 10 Samples of Leconte and Coleoptera

在处理小目标和难区分样本问题上有很大进步。

在部分样本密集的检测图片上 , Pest-YOLOv4 和 YOLOv4 检测结果如图 11 所示。

由图 11 可得 , Pest-YOLOv4 和 YOLOv4 的边界框拟合皆表现良好。对比图 11(i-a) 与图 11(i-b) , Pest-YOLOv4 检测出的样本置信程度较 YOLOv4 更高且更稳定。图 11(ii-a)、图 11(ii-b) 和图 11(iii-a) 显示 YOLOv4 对尺寸相对较小的 Acuminatus 类别林业害虫存在严重漏检 , 此与表 6 显示的 YOLOv4 对该类害虫检测查全率低一致。然而 Pest-YOLOv4 能够顺利检出该类害虫 , 未出现严重漏检。图 11(iii-a)、图 11(iii-b) 显示 , YOLOv4 对位置分布靠近的样

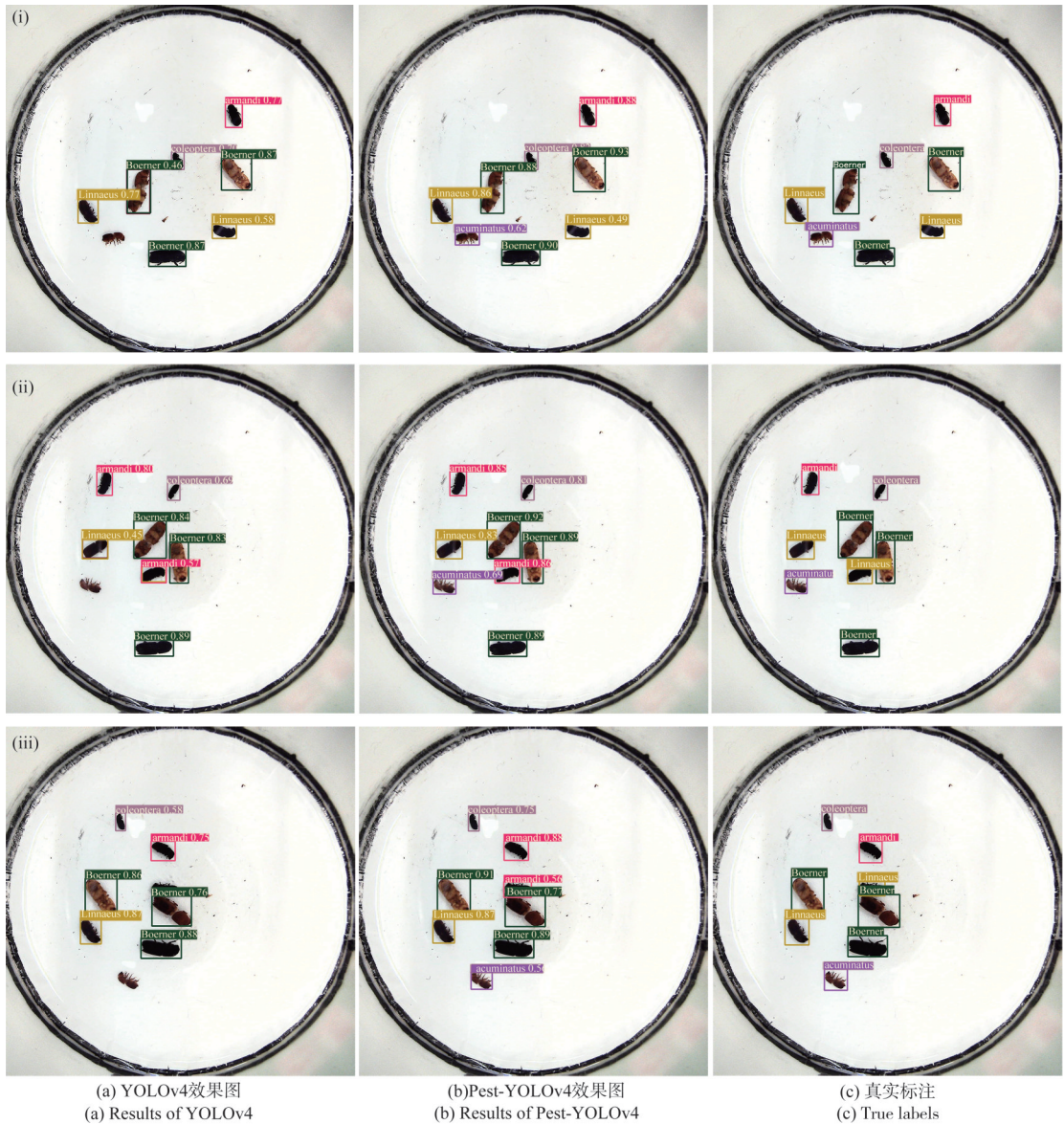


图 11 部分测试图片上 Pest-YOLOv4 和 YOLOv4 效果比较  
Fig. 11 Comparison of Pest-YOLOv4 and YOLOv4 on some test pictures

本存在漏检现象,然而 Pest-YOLOv4 仍可检出该情况样本。

4.4.3 各改进模块对比实验

为探究改进策略对网络检测效果的影响,进行了各改进策略相互组合的检测效果对比实验,实验结果如表 6 所示。

由表 6 可得,YOLOv4 检测 mAP 达到了 86.2%。YOLOv4-A 添加了 ECA-CBAM 注意力模块,相较于 YOLOv4 模型的 mAP 值增加了 2.3%。YOLOv4-B 采用 SPP-PANet 作为网络颈部,mAP 值达到 87.1%,相较于 YOLOv4 增加了 0.9%,说明 SPP-PANet 有利于检测效果的提高。YOLOv4-C 利用 Focal Loss 改进损失函数,mAP

值达到 88.3%,相较于 YOLOv4 提高了 2.1%,说明改进的损失函数可以有效地平衡正负样本比例,提高网络处理复杂样本的能力。YOLOv4-D 在 YOLOv4-B 基础上添加 ECA-CBAM 注意力模块,mAP 值达到 89.0%,相较于 YOLOv4 增加了 2.8%,相较于 YOLOv4-B 增加了 1.9%,其表明 ECA-CBAM 和 SPP-PANet 共同使用可明显提高网络模型的检测效果。Pest-YOLOv4 在 YOLOv4-D 的基础上利用 Focal Loss 改进损失函数,获得了最高的 mAP 值,达到 90.4%,相较于 YOLOv4 提高 4.2%,说明 Pest-YOLOv4 可以明显提高 YOLOv4 模型处理林业害虫检测任务的能力,对林业害虫虫情监测产生助益。

表 6 各改进策略的检测效果  
Tab. 6 Detection effects of various improvement strategies

| 网络模型        | 改进策略      |          |           |            | Precision/% | Recall/% | mAP/% |
|-------------|-----------|----------|-----------|------------|-------------|----------|-------|
|             | K-means++ | ECA-CBAM | SPP-PANet | Focal Loss |             |          |       |
| YOLOv4      |           |          |           |            | 60.3        | 88.4     | 86.2  |
| YOLOv4-A    | ✓         | ✓        |           |            | 58.8        | 92.6     | 88.5  |
| YOLOv4-B    | ✓         |          | ✓         |            | 60.9        | 90.9     | 87.1  |
| YOLOv4-C    | ✓         |          |           | ✓          | 61.6        | 91.7     | 88.3  |
| YOLOv4-D    | ✓         | ✓        | ✓         |            | 63.5        | 91.9     | 89.0  |
| Pest-YOLOv4 | ✓         | ✓        | ✓         | ✓          | 67.5        | 92.8     | 90.4  |

5 结 论

林业害虫检测存在识别难度高、识别精度低的问题。本文进行林业害虫检测优化研究,提出了基于深度学习模型 YOLOv4 优化的 Pest-YOLOv4 网络模型。该模型在 K-means++ 聚类先验框的基础上引入 ECA-CBAM 注意力

机制,重新组织网络颈部构成 SPP-PANet,利用 Focal Loss 改进损失函数,在 1 693 张含有多 种林业害虫的图片上对模型进行训练。结果表明,Pest-YOLOv4 在保证检测实时性的前提下,对多种林业害虫具有良好的检测能力,mAP 值达到 90.4%,相较于 YOLOv4 提高了 4.2%。

参 考 文 献:

[1] 生态环境部. 2020 年中国生态环境状况公报(摘录)[J]. 环境保护,2021,49(11):47-68.  
Ministry of Ecological Environment. China ecological environment status bulletin 2020 (Excerpt) [J]. *Environmental Protection*, 2021, 49(11): 47-68. (in Chinese)

[2] LECUN Y, BENGIO Y. Convolutional networks for images, speech, and time-series [M]//ARBIB M A. *The Handbook of Brain Theory and Neural Networks*. Cambridge: MIT Press, 1995.

[3] CORTES C, VAPNIK V. Support-vector networks [J]. *Machine Learning*, 1995, 20(3): 273-297.

[4] EBRAHIMI M A, KHOSHTAGHAZA M H, MINAEI S, *et al.* Vision-based pest detection based on SVM classification method [J]. *Computers and Electronics in Agriculture*, 2017, 137: 52-58.

[5] SETHY P K, BHOI C, BARPANDA N K, *et al.* Pest detection and recognition in rice crop using SVM in approach

- of bag-of-words [C]//*International Conference on Communication and Signal Processing*. Melmaruvathur, India: IEEE, 2017.
- [6] LIU Z C, ZHANG G H, YANG H Y, *et al.* Application of object detection algorithm in identification of rice weevils and maize weevils [C]//*Proceedings of the 2018 2nd International Conference on Deep Learning Technologies*. Chongqing: ACM, 2018: 76-80.
- [7] REN S Q, HE K M, GIRSHICK R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [8] LIU J, WANG X W. Tomato diseases and pests detection based on improved YOLOv3 convolutional neural network [J]. *Frontiers in Plant Science*, 2020, 11: 898.
- [9] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. *arXiv*, 2018: 1804.02767.
- [10] 袁哲明, 袁鸿杰, 言雨璇, 等. 基于深度学习的轻量化田间昆虫识别及分类模型[J]. 吉林大学学报(工学版), 2021, 51(3): 1131-1139.
- YUAN Z M, YUAN H J, YAN Y X, *et al.* Automatic recognition and classification of field insects based on light-weight deep learning model [J]. *Journal of Jilin University (Engineering and Technology Edition)*, 2021, 51(3): 1131-1139. (in Chinese)
- [11] CHEN J W, LIN W J, CHENG H J, *et al.* A smartphone-based application for scale pest detection using multiple-object detection methods [J]. *Electronics*, 2021, 10(4): 372.
- [12] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [J]. *arXiv*, 2020: 2004.10934.
- [13] ZHA M F, QIAN W B, YI W L, *et al.* A lightweight YOLOv4-based forestry pest detection method using coordinate attention and feature fusion [J]. *Entropy*, 2021, 23(12): 1587.
- [14] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021: 13708-13717.
- [15] 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 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Seattle: IEEE, 2020: 1571-1580.
- [16] HE K M, ZHANG X Y, REN S Q, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904-1916.
- [17] LIU S, QI L, QIN H F, *et al.* Path aggregation network for instance segmentation [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 8759-8768.
- [18] LIN T Y, DOLLÁR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection [C]//*Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 936-944.
- [19] ZHENG Z H, WANG P, LIU W, *et al.* Distance-IOU loss: Faster and better learning for bounding box regression [C]//*Proceedings of the 34th AAAI Conference on Artificial Intelligence*. Palo Alto: AAAI Press, 2020: 12993-13000.
- [20] 兰旭婷, 郭中华, 李昌昊. 基于注意力与特征融合的光学遥感图像飞机目标检测[J]. 液晶与显示, 2021, 36(11): 1506-1515.
- LAN X T, GUO Z H, LI C H. Attention and feature fusion for aircraft target detection in optical remote sensing images [J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(11): 1506-1515. (in Chinese)
- [21] 谢书翰, 张文柱, 程鹏, 等. 嵌入通道注意力的 YOLOv4 火灾烟雾检测模型[J]. 液晶与显示, 2021, 36(10): 1445-1453.
- XIE S H, ZHANG W Z, CHEN P, *et al.* Firesmoke detection model based on YOLOv4 with channel attention [J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(10): 1445-1453. (in Chinese)
- [22] WOO S, PARK J, LEE J Y, *et al.* CBAM: Convolutional block attention module [C]//*Proceedings of the 15th European Conference on Computer Vision (ECCV)*. Munich: Springer, 2018: 3-19.
- [23] WANG Q L, WU B G, ZHU P F, *et al.* ECA-Net: Efficient channel attention for deep convolutional neural networks [C]//*2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE,

- 2020.
- [24] 蒋榕圻,彭月平,谢文宣,等. 嵌入 scSE 模块的改进 YOLOv4 小目标检测算法[J]. 图学学报,2021,42(4): 546-555.
- JIANG R Q, PENG Y P, XIE W X, *et al.* Improved YOLOv4 small target detection algorithm with embedded scSE module [J]. *Journal of Graphics*, 2021, 42(4): 546-555. (in Chinese)
- [25] LIN T Y, GOYAL P, GIRSHICK R, *et al.* Focal loss for dense object detection [C]//*Proceedings of the 2017 IEEE International Conference on Computer Vision*. Venice: IEEE, 2017: 2999-3007.
- [26] ROMDHANE T F, ALHICHRI H, OUNI R, *et al.* Electrocardiogram heartbeat classification based on a deep convolutional neural network and focal loss [J]. *Computers in Biology and Medicine*, 2020, 123: 103866.
- [27] 董琪琪,刘剑飞,郝禄国,等. 基于改进 SSD 算法的学生课堂行为状态识别[J]. 计算机工程与设计,2021,42(10): 2924-2930.
- DONG Q Q, LIU J F, HAO L G, *et al.* Student action recognition based on improved SSD algorithm [J]. *Computer Engineering and Design*, 2021, 42(10): 2924-2930. (in Chinese)
- [28] 左彦哲. 基于深度学习的害虫识别系统[D]. 北京:北京林业大学,2018.
- ZUO Y Z. Pest recognition system based on deep learning [D]. Beijing: Beijing Forestry University, 2018. (in Chinese)
- [29] 左天兴,黄桂英,贾丽萍,等. 虫情测报仪在玉溪红塔山自然保护区森林害虫监测与防治中的应用[J]. 林业调查规划,2018,43(3):121-124.
- ZUO T X, HUANG G Y, JIA L P, *et al.* Application of pest detector in forest pests monitoring and control of Hongtashan nature reserve [J]. *Forest Inventory and Planning*, 2018, 43(3): 121-124. (in Chinese)
- [30] LIU W, ANGUELOV D, ERHAN D, *et al.* SSD: Single shot multibox detector [C]//*14th European Conference on Computer Vision*. Amsterdam: Springer, 2016: 21-37.

#### 作者简介:



赵 严(1998—),男,江苏宿迁人,硕士研究生,2020年于南京林业大学获得学士学位,主要从事深度学习及目标检测方面的研究。E-mail:yanzhao\_njfu@163.com



刘应安(1965—),男,安徽寿县人,博士,教授,2003年于东南大学获得博士学位,主要从事林业统计分析及数据挖掘方面的研究。E-mail:lyastat@163.com