

基于YOLO-V5l与ResNet50的农田害虫检测

柳春源, 陈洪建, 曾小辉, 向滔, 寇喜鹏

重庆科技学院, 数理与大数据学院, 重庆

收稿日期: 2022年6月17日; 录用日期: 2022年7月27日; 发布日期: 2022年8月5日

摘要

目前, 我国农田受虫害影响日渐严重, 虫情分析可以针对不同区域的农田虫情状况, 制定不同的治理农田害虫方案。传统的虫情分析靠人工收集与统计, 耗时耗力, 随着深度学习技术在计算机视觉领域的发展, 本文提出结合YOLO-V5l目标检测与ResNet50神经网络搭建农田害虫检测模型。昆虫在图像数据中呈现时具有体态多样、鳞片缺失、肢体脱落等特点, 对目标检测与分类的影响较大, 因此本文将28种害虫按照体态, 颜色等进行粗分类为A~G七种后, 利用YOLO-V5l模型对其进行检测与计数, 再将检测结果代入ResNet50识别模型中确定其种类。这种方法极大降低了农田害虫检测的误检率。并且, 本文提出一种预测增强算法, 对待检测害虫图像进行增强后, 分别带入识别模型中, 对识别的结果取其加权平均, 得到最终结果。单一的YOLO-V5l模型的mAP.5:.95为71.4%, 平均精确率80.91%, 漏检率5.39%。本文提出的虫情检测模型其平均精确率为89.56%, 提升了对农田害虫的识别准确率。该模型将改善原始人工统计的缺点, 推进我国智慧农业的发展。

关键词

害虫检测, YOLO-V5l, ResNet50, 目标检测

Farmland Pest Detection Based on YOLO-V5l and ResNet50

Chunyu Liu, Hongjian Chen, Xiaohui Zeng, Tao Xiang, Xipeng Kou

School of Mathematics and Big Data, Chongqing University of Science and Technology, Chongqing

Received: Jun. 17th, 2022; accepted: Jul. 27th, 2022; published: Aug. 5th, 2022

Abstract

At present, China's farmland is increasingly affected by insect pests. Insect situation analysis can formulate different plans to control farmland pests according to the insect situation in different regions. Traditional pest situation analysis relies on manual collection and statistics, which is

time-consuming and labor-consuming. With the development of deep learning technology in the field of computer vision, this paper proposes to build a farmland pest detection model by combining YOLO-V5l target detection and ResNet50 neural network. Insects have the characteristics of diverse body shapes, missing scales and falling limbs in the image data, which have a great impact on the target detection and classification. Therefore, this paper roughly classifies 28 pests into seven A~G species according to their body shapes and colors, uses YOLO-V5l model to detect and count them, and then substitutes the detection results into ResNet50 recognition model to determine their species. This method greatly reduces the false detection rate of farmland pest detection. Moreover, this paper proposes a predictive enhancement algorithm. After the pest images to be detected are enhanced, they are brought into the recognition model respectively, and the recognition results are weighted average to get the final results. mAP.5:.95 of single YOLO-V5l model was 71.4%, the average accuracy rate was 80.91%, and the missed detection rate was 5.39%. The average accuracy of the pest detection model proposed in this paper is 89.56%, which improves the recognition accuracy of farmland pests. The model will improve the shortcomings of original artificial statistics and promote the development of Intelligent Agriculture in China.

Keywords

Pest Detection, YOLO-V5l, ResNet50, Target Detection

Copyright © 2022 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

目前,我国农田受虫害影响日渐严重,虫情分析可以针对不同区域的农田虫情状况,制定不同的治理农田害虫方案[1]。传统的虫情分析靠人工收集与统计,耗时耗力。目前,虫情测报灯在虫情采集中应用较为广泛[2]。它主要依靠灯光诱捕装置捕获农业害虫,实现自动诱集、杀虫、虫体分散、拍照等作业,并实时地将虫情信息上传至云平台。虫情测报灯的投入使用可帮助植保人员高效地进行虫情分析,提高测报工作效率和准确率[3]。在统计害虫种类和数量时,存在着耗时耗力和准确性差等不足。随着深度学习的发展,通过虫情测报灯所采集到的图像文件,利用图像检测和识别算法对其统计害虫种类与数量或许可以解决人工统计上耗时耗力的问题。

2. 数据探究与预处理

深度学习技术中,需要相对较高的数据质量,它在一定程度上影响了所训练的模型效果,所以数据探究与预处理是在深度学习技术中不可或缺的一部分。经过对数据探究可以更好地了解数据集,更好地针对数据集的特点做不同的预处理操作,提高模型的识别与检测效果。

2.1. 数据探究

2.1.1. 数据资料

本文的灯诱害虫数据集来源是由第十届“泰迪杯”数据挖掘挑战赛 A 题的官方赛事数据,分为灯诱害虫图像数据与害虫位置坐标数据,图像尺寸为 5472 × 3648 px,以 jpg 格式存储,共 28 种害虫图像数据。

2.1.2. 数据噪声问题

在所给的数据集中,许多害虫存在重叠,以及一些未给标签的害虫种类,还有许多害虫的残肢、残

翅和落叶树枝等等，如所给数据集中图 1 所示。

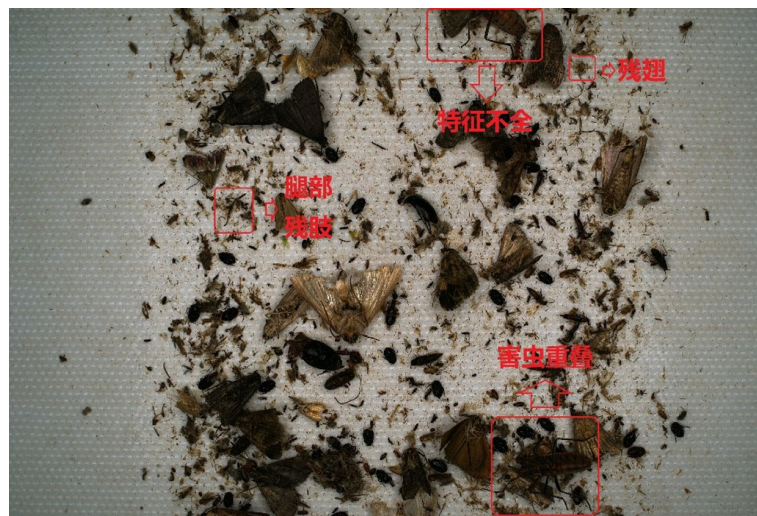


Figure 1. Noise problem of lamp-induced pest data
图 1. 灯诱害虫数据噪声问题

该数据集的噪声较为复杂，如果直接使用原图进行训练的话，可能导致之后的识别结果中，会把一些残肢，残翅等等也认为是一些体型较小的昆虫，对结果影响较为严重。本文选择利用所给标签进行分割，尽量使特征占到图像的较大比例，减少影响。

2.1.3. 昆虫相似度较大

训练数据集中包含了螟、蛾、虱属等。同属下的特征相似度较大，带入模型中很容易将其互相混淆，导致模型的准确率大大降低。一般不同属类的相似度差异较大，但螟属和蛾属的大多种类从颜色、大小、翅膀等特征上来看较为相似，大螟与线委夜蛾的颜色与身体特征相似度较大如图 2。



Figure 2. Borer (left) and noctuid moth (right)
图 2. 大螟(左)与线委夜蛾(右)

不仅是大螟与线委夜蛾相似度较大，根据数据集发现，螟类和蛾类的差别都不是很大，在数据质量不高的状况下，并且昆虫的特征受身体姿态等影响，会导致训练出模型的准确率极低的情况，为此，本文利用已有数据，根据科属与颜色以及身体特征最终粗分类出 7 类害虫。如表 1 所示。那么可以先对 28 种害虫进行一个粗分类，分出 A~G 类后在从分别确定种类，这样可以提高模型的准确率。

Table 1. Rough classification of 28 pests
表 1. 28 种害虫粗分类

害虫名称	类别	害虫名称	类别	害虫名称	类别	害虫名称	类别
二化螟	A	大螟	A	二点委夜蛾	A	线委夜蛾	A
甜菜白带野螟	A	歧角螟	A	石蛾	A	甘蓝夜蛾	A
瓜绢野螟	A	豆野螟	A	干纹冬夜蛾	A	紫条尺蛾	A
草地螟	A	甜菜夜蛾	A	水螟蛾	A	稻纵卷叶螟	A
八点灰灯蛾	B	黄毒蛾	C	褐飞虱属	D	白背飞虱	D
蟋蟀	E	蟋蟀	E	黄足猎蝽	E	大黑鳃金龟	F
粘虫	G	稻螟蛉	G	地老虎	G	棉铃虫	G

2.2. 数据预处理

数据预处理技术路线图如下图 3。

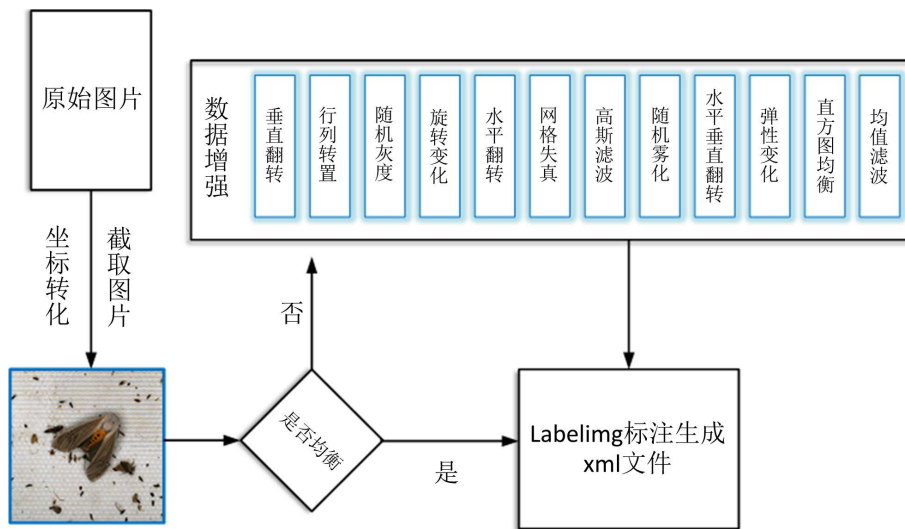


Figure 3. Data preprocessing technology roadmap
图 3. 数据预处理技术路线图

由于图像中可用信息占原图比例太少，有太多的数据冗余以及噪声，因此本文根据所需特征对原图进行切割。如果直接按照害虫大小进行切割，则切割的出来的图片尺寸大小不一，虫子也会直接充斥在整个图像中，当使用 YOLO-V5 模型对图像进行训练时，则锚定框很难做出适应，导致模型的检测效果极差。因此本文以给定的中心点坐标为中心，重新设置计算切割框大小对图像切割，使得害虫图像尽可能在图像中心且不能占据图像过于大的面积，然后统一缩放到 550×550 像素，如图 4 所示。

2.2.1. 数据增强

数据增强是一种从现有的训练样本中生成新的训练样本的技术，是一种在数据约束环境下提高机器学习模型性能和准确性的低成本和有效的方法。增强方法包括：均值滤波、水平反转、水平垂直翻转、行列转置、随机灰度、随机旋转、网格失真、弹性变换、高斯滤波、自适应直方图均衡、随机雾化等[4]。

Mosaic 数据增强方法是 YOLOv4 论文中提出来的，主要思想是将四张图片进行随机裁剪，再拼接到一张图上作为训练数据[5]。这样做的好处是丰富了图片的背景，并且四张图片拼接在一起变相地提高批

尺寸(batch_size), 在进行批标准化(batch normalization)的时候也会计算四张图片, 所以对本身 batch_size 不是很依赖。如图 5 所示。

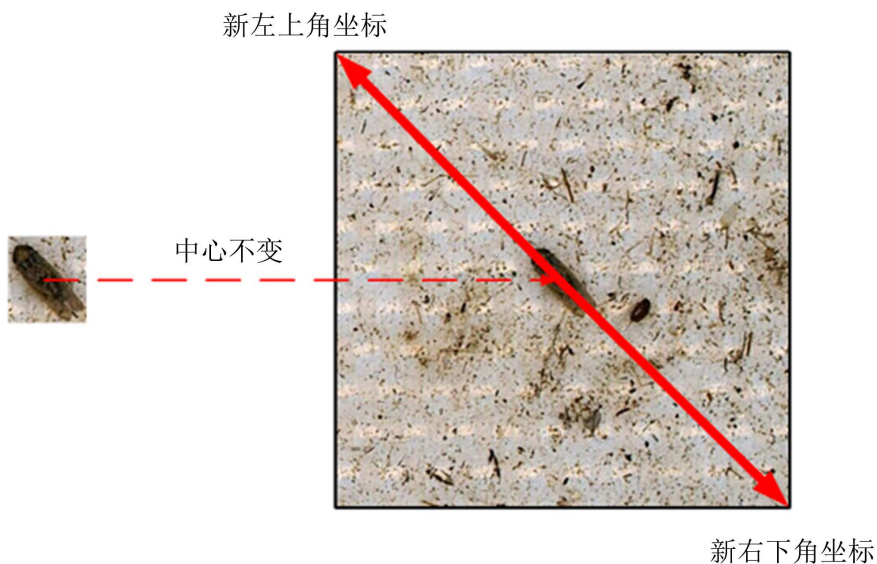


Figure 4. Diagram of coordinate transformation and cutting of brown planthopper
图 4. 褐飞虱属坐标转化切割示意图

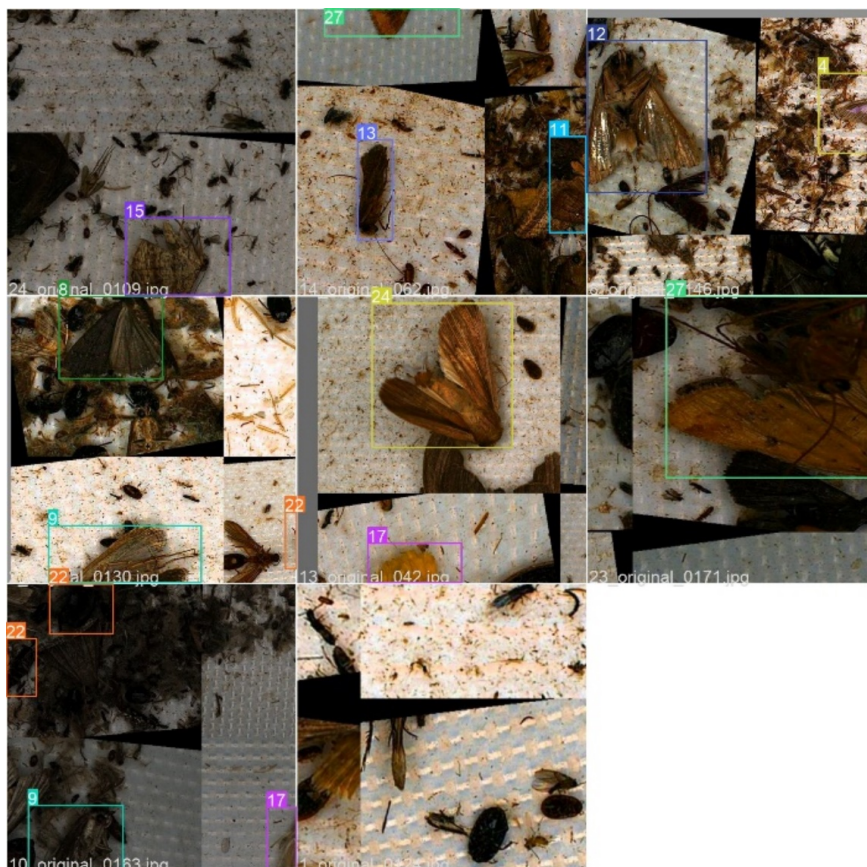


Figure 5. Mosaic data enhancement for training sets
图 5. 训练集的 Mosaic 数据增强

2.2.2. 数据标记

当使用 YOLO-V5 进行检测时, 需要传入数据在图中的坐标以及标签, 所以需要对处理好的数据进行标注, Labeling 是一个图形图像注释工具, 本文即采用该工具对数据进行标记[6]。由人工完成农业灯诱目标害虫坐标信息标注, 以 xml 格式保存, xml 文件中包含了改图中害虫的位置坐标、种类、图片路径等信息。

3. 灯诱害虫智能检测与识别模型

3.1. 基于 YOLO-V5l 模型的害虫粗分类检测

YOLOv5 是一种单阶段目标检测算法, 该算法在 YOLOv4 的基础上添加了一些新的改进思路, 使其速度与精度都得到了极大的性能提升[7]。

根据 YOLO-V5 模型官方提供的模型在 COCO 数据集上的表现, YOLO-V5 的预训练权重较好的有 YOLO-V5x, YOLO-V5l, YOLO-V5m 这 3 个权重文件。其中最好的是 YOLO-V5x, 但是该权重文件收敛最慢, 时间成本较高, 所以我们选择相对较好的 YOLO-V5l 预训练权重文件进行迁移学习。

本文在昆虫相似度较高的情况下, 提出的 YOLO-V5l 模型的粗分类检测, 先对数据集的害虫利用 YOLO-V5l 模型根据表 1 所划分的 A~G 类七种作为标签, 对害虫进行一次检测, 以此来提高模型的精确度。

3.2. 基于 ResNet50 模型的害虫细分类识别

针对害虫种类之间的相似度较高, 导致 YOLO-V5 检测的结果可能同科属下分类错误的状况, 本文提出在 YOLO-V5 检测结果基础上, 对其进行二次分类, 来确保模型的准确率。

ResNet 残差神经网络的主要思想是在网络中增加了直连通道, 即 Highway Network 的思想。此前的网络结构是性能输入做一个非线性变换, 而 Highway Network 则允许保留之前网络层的一定比例的输出。[8] ResNet 残差神经网络的思想 and Highway Network 的思想也非常类似, 允许原始输入信息直接传到后面的层中。ResNet 网络结构如图 6 所示。

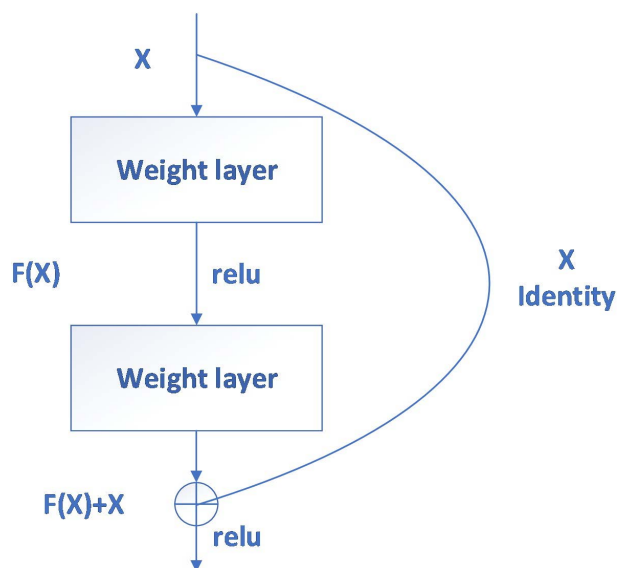


Figure 6. Residual structure diagram

图 6. 残差结构示意图

ResNet 原始网络结构首先经过第一层卷积后进行池化层处理，池化层后先经过三层卷积，然后将经过三层卷积的神经元与直接经过卷积核为 1×1 ，步长 $s = 1$ 的卷积操作后的神经元进行叠加。然后传入下一层卷积层中，将经过三层卷积的神经元与下采样结果(经过卷积核为 1×1 ，步长 $s = 2$ 的卷积操作)进行叠加，之后继续传入卷积层，重复该卷积与下采样叠加操作。该下采样操作会使得输入特征图丢失部分信息。所以本文改进该下采样操作，将每层卷积层之间的下采样部分卷积操作修改为 3×3 ，步长设置为 $s = 2$ 。如图 7 所示。改进下采样的灯诱害虫特征提取，如图 8 所示。

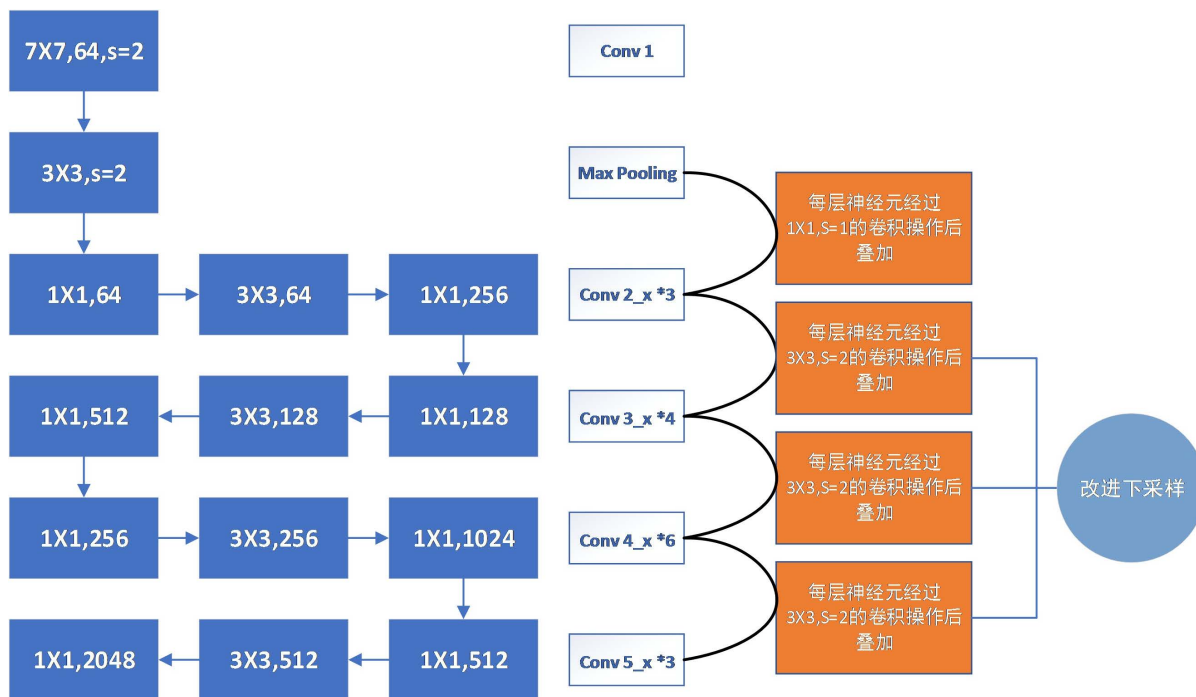


Figure 7. Schematic diagram of down sampling improvement

图 7. 下采样改进示意图

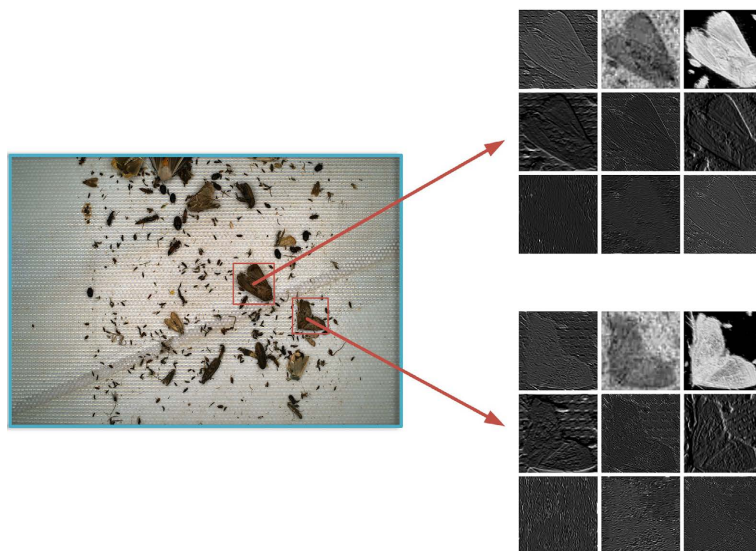


Figure 8. Feature extraction of lamp-induced pests

图 8. 灯诱害虫特征提取

上文, YOLO-V5 检测模型中将 7 类 28 种害虫进行检测, 检测出的结果分为 A~G 七类, 其中 B、C、F 三个种类与其他种类区分度较大, 在上文划分时只有一个小类, 因此害虫检测模型输出的结果不必再进行识别分类过程。但 A、D、E、G 仅是一个粗分类结果, 并不能确定到最终的种类上, 因此本文搭建了 ResNet50 识别模型用于确定最终的种类, 流程总结如下图 9。

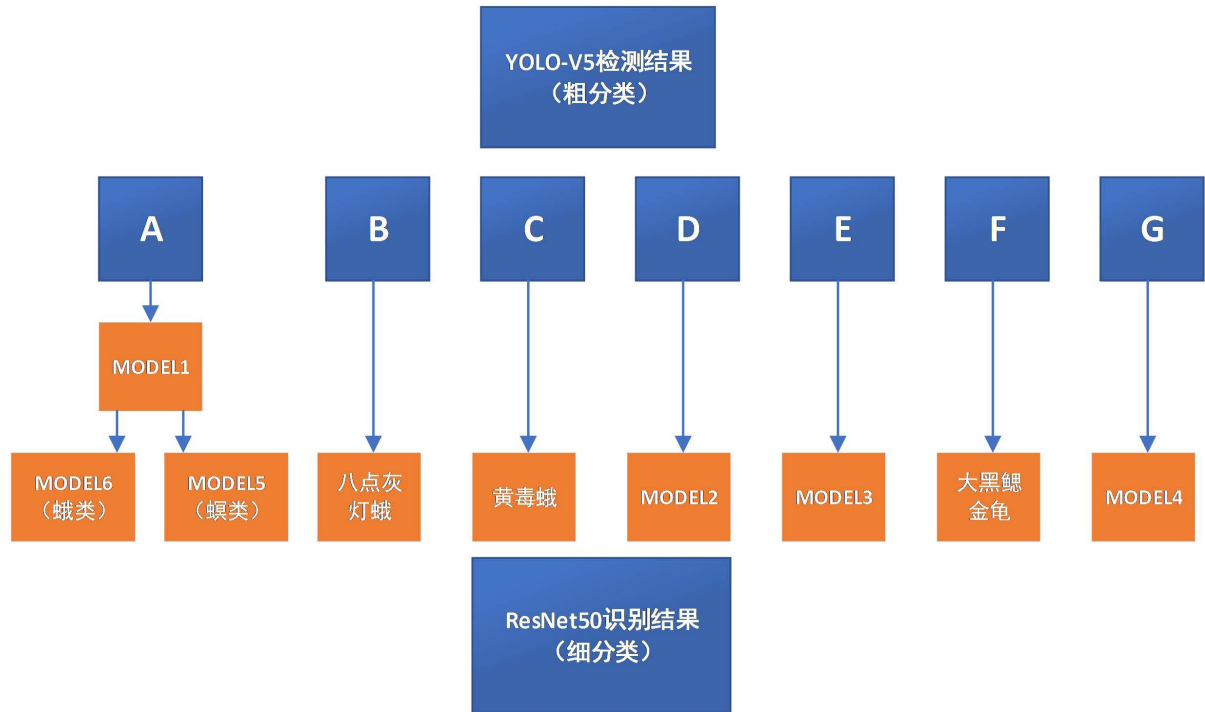


Figure 9. Model building process
图 9. 模型搭建流程

3.3. YOLO-V5I 与 ResNet50 结合的害虫检测与识别模型

在上文基于 YOLO-V5 的害虫检测模型与改进下采样的 ResNet50 害虫识别模型的基础上, 将检测模型与识别模型相结合, 提出基于检测—识别的虫情分析模型。一层检测对灯诱害虫图像进行检测, 得出粗分类结果; 再代入二层识别模型, 得出最终的害虫类别, 如图 10 所示。并对其进行数据统计。

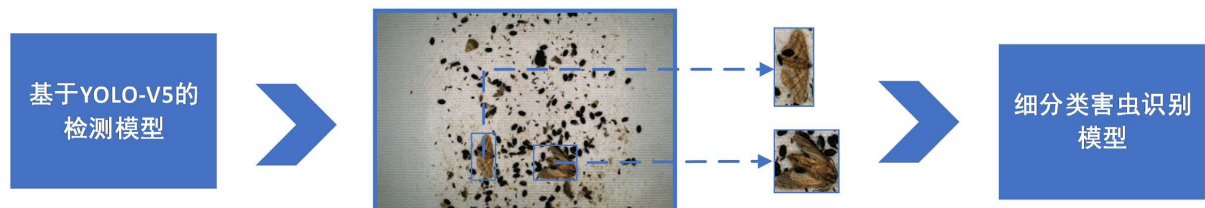


Figure 10. Pest detection and identification model process based on the combination of YOLO-V5I and ResNet50
图 10. YOLO-V5I 与 ResNet50 结合的害虫检测与识别模型流程

3.4 模型预测增强算法

由于害虫的识别受身体姿态、重叠状况、形状、鳞片特征等影响较大[9]。不同的待识别图片传入细分类害虫识别模型中可能影响较大, 可能做某些图像操作例如: 随机旋转、均值滤波、水平反转、水平

垂直翻转等操作后，再次传入细分类模型会得出不同的分类结果，导致模型预测结果的错误。因此本文提出对待训练图片进行数据增强后，再带入细分类害虫识别模型中，对数据增强检测结果与原始图片的检测结果进行加权计算如式(1)。得出最种预测结果，流程如下图 11 所示。

$$Class = \max(\alpha_i Class_i) \quad (1)$$

$Class$ 表示最终预测种类， α_i 表示权重因子， $Class_i$ 表示每种数据增强与原图所检测的结果，最终取 $Class_i$ 最多的结果作为最终检测种类。

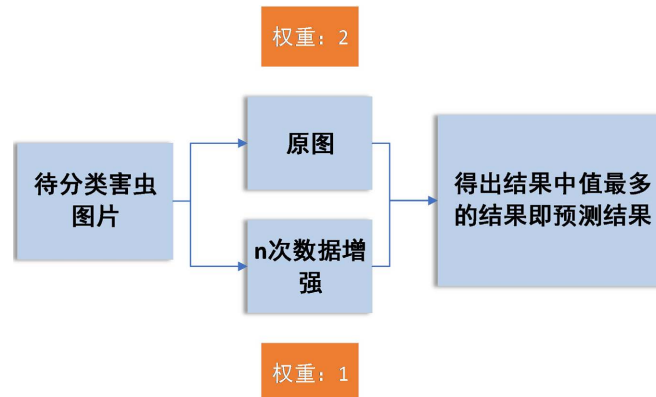


Figure 11. Prediction enhancement algorithm
图 11. 预测增强算法

4. 评价指标

判断一个算法、模型是否优秀需要一个评判标准，常用的评价指标有以下几种：准确率(Accuracy)，精准率(Precision)，召回率(Recall)等[10]。

准确率是指正确分类的样本量占总样本量的比例，其计算公式如下式(2)：

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

召回率是指实际正样本在预测正样本中所占的比例，反映了分类器发现全部正样本的能力，其计算公式如下式(3)：

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

式中： TP 表示正例分类正确的数量； TN 表示负例分类正确的数量； FP 表示正例分类错误的数量； FN 表示负例分类错误的数量。

本模型的精确率如图 12(左)与召回率如图 12(右)在训练过程中的变化可以从 Tensorboard 中显示出来，横坐标为 Epoch 次数，纵坐标为精确率与召回率在不同迭代次数时的精度，图中浅色的线表示真实的数据图，而深色的线表示平滑之后的数据图。

本模型在精确率与召回率上有着较好的表现。但在目标检测中，单纯使用精确率与召回率在目标检测中并不足以体现模型的性能。 AP 是不同 Recall 下 Precision 的均值，也就是对应 PR 曲线的面积。对于一个 PR 曲线来说，每一个阈值对应着一个 P 和 R。当阈值从 0 到 1 走过，绘制出每个阈值对应的 P 和 R 值，得到的就是 PR 曲线[11]。而 m 则代表各类别的平均。 mAP 值(Mean Average Precision)，即平均 AP 值。

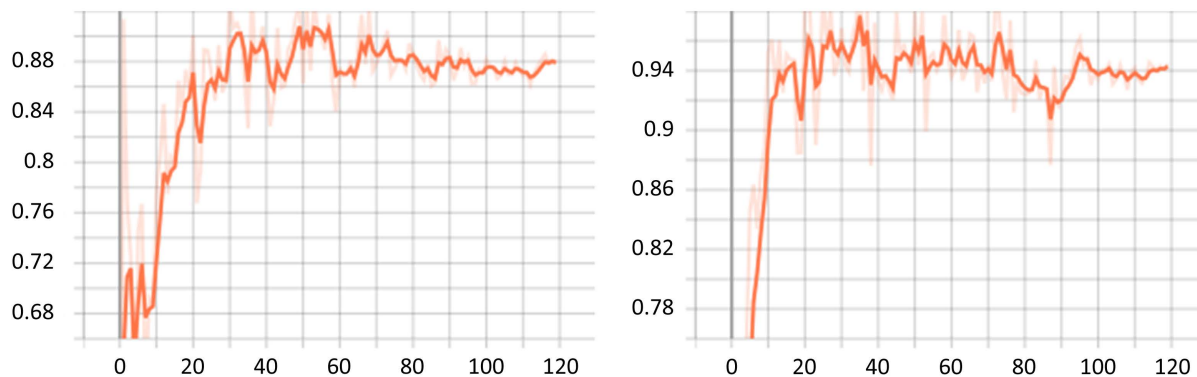


Figure 12. Accuracy (left) and recall (right)

图 12. 精确率(左)与召回率(右)

$$mAP = \frac{1}{|Q_R|} \sum_{q \in Q_R} AP(q) \quad (4)$$

其中 Q_R 表示全部种类, q 表示每个种类, $AP(q)$ 表示每个种类的 AP 值。

mAP@0.5: 即将 IoU 设为 0.5 时, 计算每一类的所有图片的 AP , 然后所有类别求平均, 即 mAP 。本害虫检测模型的 $mAP@0.5$ 如下图 13(左)所示。**mAP@.5:.95 (mAP@[.5:.95])**: 表示在不同 IoU 阈值(从 0.5 到 0.95, 步长 0.05) (0.5、0.55、0.6、0.65、0.7、0.75、0.8、0.85、0.9、0.95)上的平均 mAP 。如图 13(右)所示, 图中浅色的线表示真实的数据图, 而深色的线表示平滑之后的数据图。 $mAP@0.5$ 指标与 $mAP@.5:.95$ 指标分别最终达到了 0.9491 与 0.7043, 具有较好的结果。

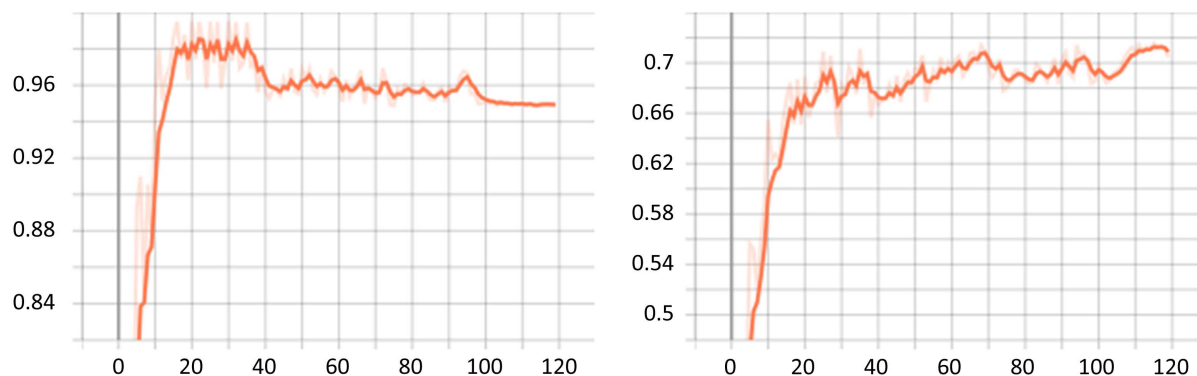


Figure 13. mAP@0.5 (left) and mAP@.5:.95 (right)

图 13. mAP@0.5 (左) and mAP@.5:.95 (右)

5. 实验环境与训练参数

本项目在内存为 64 G, CPU 为 i9-9900KF, GPU 为 NVIDIA GeForce RTX 2080 的设备上进行模型的训练, 使用 Keras 2.3.1, Opencv-python 4.4.4.64, Pytorch 1.11.0, Python 3.7.12 的训练环境, 训练参数 Batch Size 为 16, Epoch (YOLO-V5l)为 120, Epoch (ResNet50)为 50, Image Size 为(500, 500), IOU 为 0.5。

6. 结果分析

基于上文的害虫检测模型, 将虫情采集图像上传至模型中, 检测图像中的害虫类别与数量, 结果如下图 14, 有部分害虫漏检, 这张图像的漏检率为 6.97%, 其中可以看到有部分的透明残翅, 与一些较小的残叶与树枝被误检为虱类害虫, 原因可能是因为虱类害虫在图中占比太小, 导致识别不佳, 准确率为 88.4%。



Figure 14. Test results
图 14. 检测结果

对其所有预测结果求其平均误检率与平均漏检率结果如下表 2。

Table 2. Detection result
表 2. 检测结果

检测图像数量	使用方法	平均漏检率	平均精确率
521	YOLO-V5l	5.39%	80.91%
521	YOLO-V5l + ResNet50	5.39%	84.56%
521	YOLO-V5l + ResNet50 + 预测增强算法	5.39%	89.56%

对于平均精确率，主要由于许多残肢与残叶等细小噪声，容易被误判为虱属，导致结果的平均精确率并不完美，平均漏检率主要是由于一些肢体的重叠和遮挡，导致会出现漏检，其中八点灰灯蛾，由于背部和腹部颜色差异极大，一定程度上的导致了它的漏检率较高。

对于 B 类(八点灰灯蛾)、C 类(黄毒蛾)、F 类(大黑鳃金龟)这种不需要进行二次识别的类别，直接使用一层检测模型的 $mAP@.5:.95$ 。结果如表 3。

Table 3. B, C and F $mAP@.5:.95$
表 3. B、C、F 三类 $mAP@.5:.95$

种类	$mAP@.5:.95$
八点灰灯蛾	75.5%
黄毒蛾	76.6%
大黑鳃金龟	54.7%

其中八点灰灯蛾与黄毒蛾的检测效果较好。但是大黑鳃金龟由于与石蛾相似度较高，它们之间存在互相误检，导致 mAP 值较低。

7. 总结

本文介绍了当前我国的农田害虫现状，利用深度学习技术对使用灯诱害虫装置获取的数据进行虫情

分析,可以解决人工统计虫情信息耗时耗力的缺点。灯诱害虫图片数据中,灯诱害虫种类繁多,虫体姿态多样,噪声较大,害虫间的肢体重叠,对于模型的准确性造成了较大的影响。本文从以下几个方面逐步建立虫情分析模型。

数据探索与预处理方面:针对昆虫同科属下相似度较大的情况,提出了数据粗分类,用以提高检测模型时的准确率。针对数据不均衡的情况,使用12种数据增强方式,对数据进行扩充,使数据均衡化。

一层检测模型方面:使用YOLO-V5I模型,将粗分类的数据带入检测模型中,减少由于28分类的分类数量过多,以及类间相似度较大对模型的影响,提高模型的准确率。

二层识别模型方面:将YOLO-V5I检测模型粗分类的结果带入二层识别模型中去做细分类,并提出了一种加权的预测增强算法,以提高模型准确率。

在传统的农业中,往往虫害已经发展开来时才会出现明显特征并实施治理,这样不仅防治过程中产生的成本高,而且由于之前的环保意识不那么强,在害虫的防治上大量用药,导致农业环境污染严重和农业产品的不完全安全。而如今,人们对农业虫害上有了更科学更持续更深入的认知,在农业虫害上更注重预防害虫,这不仅效果好,防治过程中的成本少了,对农业的环境也更科学与友好了。

虫情测报灯便是科学农业的产物,而本文则是利用虫情测报灯收集的图像数据,进一步智能快捷的将虫情分析出结果。因此,关于当前农业的发展,我们要用更发展、更智能、更科学、更环保的眼光来进行思考,不仅要让科技走向越来越智能,也让农业也跟上科学、智能和环保的脚步。

基于YOLO-V5I与ResNet50的农田害虫检测模型在农业灯诱相似害虫检测与识别方法取得了较好的识别精度,为相似害虫识别提供了技术参考。但农业灯诱相似害虫识别精度仍有提升的空间,下一步需要改进与完善的地方包括:

- 1) 对于残肢、残翅、碎叶等的影响需要降低。
- 2) 对于害虫肢体重叠的情况还需进一步加强攻克。
- 3) 对于粗分类时的分类情况,还需进一步定义一个确定的方法。
- 4) 对于昆虫形态、以及同类昆虫不同生长时期的识别还需进一步加强。
- 5) 对于数据采集装置,需要避免对昆虫造成较大的伤害,减少残肢的出现。

参考文献

- [1] 牛霆葳. 基于机器视觉的农田害虫自动识别方法研究[D]: [硕士学位论文]. 天津: 天津科技大学, 2015.
- [2] 张红. 二点委夜蛾使用性诱捕器和虫情测报灯监测效果研究[J]. 农业科技通讯, 2021(1): 231-233.
- [3] 陈烁, 谌晓姝, 赵银娣, 张银松. 智慧农田害虫时空监测系统[J]. 北京测绘, 2019, 33(1): 10-13.
- [4] 陈龙, 丁丹丹. 多残差联合学习的水下图像增强[J]. 中国图象图形学报, 2022, 27(5): 1577-1588.
- [5] 胡嘉沛, 李震, 黄河清, 等. 采用改进YOLOv4-Tiny模型的柑橘木虱识别[J]. 农业工程学报, 2021, 37(17): 197-203.
- [6] 宋怀波, 王亚男, 王云飞, 吕帅朝, 江梅. 基于YOLO v5s的自然场景油茶果检测方法[J/OL]. 农业机械学报: 1-9. <http://kns.cnki.net/kcms/detail/11.1964.S.20220518.1646.038.html>, 2022-08-01.
- [7] 范晓飞, 王林柏, 刘景艳, 周玉宏, 张君, 索雪松. 基于改进YOLO v4的玉米种子外观品质检测[J/OL]. 农业机械学报: 1-16. <http://kns.cnki.net/kcms/detail/11.1964.S.20220518.1551.036.html>, 2022-08-01.
- [8] 谢银成, 黎曦, 李天, 李聪聪. 基于改进ResNet和损失函数的表情识别[J]. 自动化与仪表, 2022, 37(4): 64-69.
- [9] 姚波. 基于深度学习的农业灯诱相似害虫检测与识别方法的研究[D]: [硕士学位论文]. 杭州: 浙江理工大学, 2021.
- [10] 刘伟, 邹伟红, 卢彦杰, 等. 基于迁移学习和模型融合的中药材图像识别研究[J]. 湖南中医药大学学报, 2022, 42(5): 809-814.
- [11] 方伟政. 基于弱监督的昆虫图像细粒度分类技术研究[D]: [硕士学位论文]. 成都: 电子科技大学, 2021.