

基于YOLOv5改进的粘蝇纸家蝇识别算法

王亚辉

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2025年4月30日; 录用日期: 2025年5月23日; 发布日期: 2025年5月31日

摘要

文章针对粘蝇纸上家蝇目标体积小、形态多变且背景复杂的检测难题, 提出了一种基于YOLOv5改进的检测算法。为充分反映实际应用场景, 本研究搜集和采集并标注了500张粘蝇纸图像, 其中400张用于训练, 100张用于测试。改进工作主要集中在检测头部分, 通过引入Zoom_cat模块实现多尺度特征的对齐与融合、采用ScalSeq模块增强特征序列化处理能力, 并结合注意力机制提升目标区域的显著性, 从而优化小目标的特征提取和定位效果。实验结果表明, 改进后的模型在mAP、精确率和召回率等关键指标上均显著优于原始YOLOv5m模型, 充分验证了所提方法在家蝇检测中的有效性和鲁棒性。该研究为粘蝇纸上家蝇的数量监测提供了一种高效、准确的识别方法, 同时为小目标检测问题的进一步探索提供了新的思路。

关键词

YOLOv5, 家蝇检测, 图像处理

Enhanced Sticky Fly Paper Housefly Recognition Algorithm Based on YOLOv5

Yahui Wang

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 30th, 2025; accepted: May 23rd, 2025; published: May 31st, 2025

Abstract

In this paper, we propose an enhanced detection algorithm, termed YOLOv5, for identifying housefly targets on sticky fly paper. These targets exhibit characteristics such as small size, variable morphology,

and complex background. To thoroughly reflect the actual application scenario, a comprehensive data set was collected, acquired, and labeled. This included 500 images of sticky fly paper, with 400 utilized for training and 100 for testing purposes. The primary focus of the improvement work is on the detection head part, which optimizes the feature extraction and localization effect of small targets by introducing the Zoom_cat module to achieve multi-scale feature alignment and fusion, adopting the ScalSeq module to enhance the feature serialization processing capability, and combining with the attention mechanism to enhance the saliency of the target region. The experimental results demonstrate that the enhanced model significantly outperforms the original YOLOv5m model in terms of key metrics such as mAP, precision, and recall, thereby validating the efficacy and reliability of the proposed method for housefly detection. This study proposes an efficient and accurate identification method for the population monitoring of houseflies on sticky fly paper and concomitantly offers a novel approach for further exploration of the small target detection problem.

Keywords

YOLOv5, Housefly Detection, Image Processing

Copyright © 2025 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

近年来,随着人工智能与深度学习技术的快速发展,计算机视觉在目标检测领域取得了显著进展。这些技术的广泛应用使得自动化监测与识别成为可能,为公共卫生、农业防控、环境治理等领域提供了有力支持。家蝇作为一种常见且具有传播疾病风险的昆虫,其精准检测在保障食品安全、环境卫生以及公共健康方面具有重要意义。然而,传统的家蝇监测方法主要依赖人工检查或基于传统图像处理算法的手段[1],不仅效率低下,而且在复杂背景下难以保证较高的检测准确率。Sharif 开发了基于模糊逻辑的数字图像处理算法,对水稻田间稻纵卷叶、直纹稻弄蝶、稻暮眼蝶等 6 种常见害虫的图像进行识别、计数,取得了较好的识别效果[2]。Fukatsu 通过融合机器视觉与无线网络,基于背景差分技术的图像算法实现了田间害虫稻缘蝽的远程自动监测计数,计数准确率达 89.1% [3]。沈佐锐应用计算机视觉技术对温室白粉虱进行自动计数,计数准确率达 92% [4]。邱道尹设计了基于机器视觉的农田害虫实时检测系统,基于神经网络分类器对农田常见的 9 种害虫进行分类,准确率达 93.5% [5]。Wen 研究了基于图像的果园害虫自动鉴定与分类方法,通过结合局部与全局特征,对苹果蠹蛾、红带卷蛾、玫瑰色卷蛾等 5 种果园害虫进行分类,分类准确率达 86.6% [6]。韩瑞珍研究了基于计算机视觉对于大田害虫的识别研究,其所采用支持向量机分类器进行实验,所得的实验结果对于用图形处理进行害虫的监测起到了积极的作用,但得到平均准确率只有 87.4%,且识别速度较慢[7]。

近年来,以 YOLO 系列为代表的实时目标检测算法因其优秀的检测速度和较高的精度而受到广泛关注。尤其是 YOLOv5,凭借其轻量化的网络结构和高效的特征提取能力,已经在多种目标检测任务中展现出较强的应用潜力。然而,由于家蝇体积较小、姿态多变且易受复杂背景干扰,直接采用 YOLOv5 进行家蝇识别仍存在一定局限性。例如,模型在处理小目标时可能会出现特征提取不足、检测漏检率较高等问题。现有方法在速度和准确性方面仍有局限性,尤其是在室外不受控制的光照条件下[8]。

在本文中,我们针对粘蝇纸上家蝇的检测任务,对 YOLOv5 模型进行了改进。通过在检测头部分引入 Zoom_cat 模块,实现多尺度特征的对齐与融合。采用 ScalSeq 模块增强特征序列化处理能力。并结合

注意力机制(Attention_model 模块)提升目标区域的显著性,从而优化小目标的特征提取和定位效果。并通过在模型改进中进行消融实验,并与 YOLOv8、YOLOv10 分别对比训练模型评价指标后,证明了所提方法在家蝇检测中的有效性和鲁棒性。同时,在数据集层面上,本文构建了一个包含多样场景下家蝇图像的数据集,并采用了多种数据增强技术以提高模型的泛化能力。

2. 改进 YOLOv5 算法设计

针对粘蝇纸上家蝇目标较小、易受背景干扰的特点,本文对原始 YOLOv5 模型在检测头部分进行了如下改进,网络结构如图 1 所示。

Zoom_cat 模块: Zoom_cat 模块是一个特征融合模块,主要接收三个不同尺度的特征图(大、中、小),并将其融合到一个共同的尺度上。对于大尺度特征图,使用自适应最大池化和自适应平均池化将其尺寸调整到中尺度特征图的尺寸。通过将两种池化的结果相加,可以同时保留特征图的局部最大值和平均值信息,从而增强特征的鲁棒性。对于小尺度特征图,使用最近邻插值将其尺寸上采样到中尺度特征图的尺寸。特征融合将调整尺寸后的三个特征图沿着通道维度拼接起来,形成一个新的特征图,从而融合了不同尺度的特征信息。Zoom_cat 模块通过将不同尺度的特征图融合到一个共同的尺度上,能够使模型在处理多尺度目标时更加有效。在处理包含密集小目标的图像时,Zoom_cat 模块通过融合小尺度特征图的信息,能够显著提升模型对小目标家蝇的检测能力。通过自适应池化和插值操作,Zoom_cat 模块能够有效地保留不同尺度特征图的重要信息,避免因直接裁剪或简单插值导致的信息丢失。

ScalSeq (Scale-Aware Sequential Module)模块的引入: ScalSeq 模块接收三个特征图,首先对每个特征图分别应用单独的卷积操作,这些卷积操作可能用于提取特征图的局部特征信息,并对特征图进行一定的特征转换和降维。然后对这些经过卷积处理后的特征图进行尺寸调整,使它们的尺寸能够匹配,以便后续的融合操作。接下来将调整尺寸后的特征图转换成三维张量,这是为了将不同尺度的特征图在通道维度上进行拼接,形成一个包含多尺度信息的三维张量结构。最后通过三维卷积对拼接后的三维张量进行融合,三维卷积能够在空间和通道维度上同时捕捉特征之间的关联,从而实现对多尺度特征的有效融合。融合后的特征图再经过最大池化操作进一步提炼特征,以突出重要信息。ScalSeq 模块的核心工作原理是通过多尺度特征的融合来增强模型对不同尺度目标的检测和分割能力。它先对不同尺度的特征图分别进行处理,提取各自的特征信息,然后通过三维卷积将这些特征图在空间和通道维度上进行融合,使模型能够同时捕捉到大、中、小不同尺度的目标特征。最大池化操作则进一步增强了模型对重要特征的提取能力,通过选择性地保留局部区域中的最大值,能够突出特征图中的关键信息,去除一些冗余或不重要的细节,使模型更加关注于目标的关键特征,从而提高模型的鲁棒性和性能。ScalSeq 模块通过对多尺度信息的进一步序列化处理,强化了各尺度特征之间的联系,使模型能够更好地捕捉到粘蝇纸上家蝇的尺度变化特征,提升了对目标边界的定位精度。

注意力机制的整合(Attention_model 模块): Attention_model 模块包括通道注意力机制和空间注意力机制两部分。通道注意力机制部分首先会对输入的特征图进行全局平均池化或全局最大池化操作,以获取特征图在通道维度上的全局信息。然后通过一维卷积或全连接层等网络结构对池化后的结果进行处理,生成通道注意力权重。这些权重通常通过激活函数进行归一化,使其值在 0 到 1 之间,表示不同通道的重要性。最后将这些通道注意力权重与原始特征图的对应通道相乘,实现对特征图通道的加权,增强重要通道的特征,抑制不重要通道的特征。空间注意力机制部分先对特征图在宽度和高度方向上分别进行均值计算和最大值计算,得到两个描述特征图空间信息的向量。然后将这两个向量在通道维度上进行拼接,再通过一系列卷积操作生成空间注意力图。空间注意力图的每个元素值表示对应位置的空间重要性,同样通过激活函数进行归一化后,与原始特征图相乘,突出特征图中的关键空间区域,抑制不重要的区

域。**Attention_model** 模块的工作原理是模拟人类视觉注意力机制，使模型能够自适应地聚焦于输入数据中最有信息量的部分，从而提高模型对重要特征的感知能力和表达能力。通道注意力机制通过为每个通道分配权重，能够使模型根据任务需求动态地调整对不同通道特征的关注程度。空间注意力机制则能够让模型在空间维度上突出图像中的关键区域，在目标检测任务中，模型会将注意力集中在目标物体所在的区域，而对背景区域的注意力则相对较低，这样能够减少背景噪声对模型的干扰，提高目标检测的精度。通过将通道注意力和空间注意力结合起来，**Attention_model** 模块能够同时在通道和空间两个维度上对特征进行优化和调整，使模型能够更加精细地处理输入数据，生成更有区分度和代表性的特征表示，从而有效提高了家蝇的检测准确率和鲁棒性。

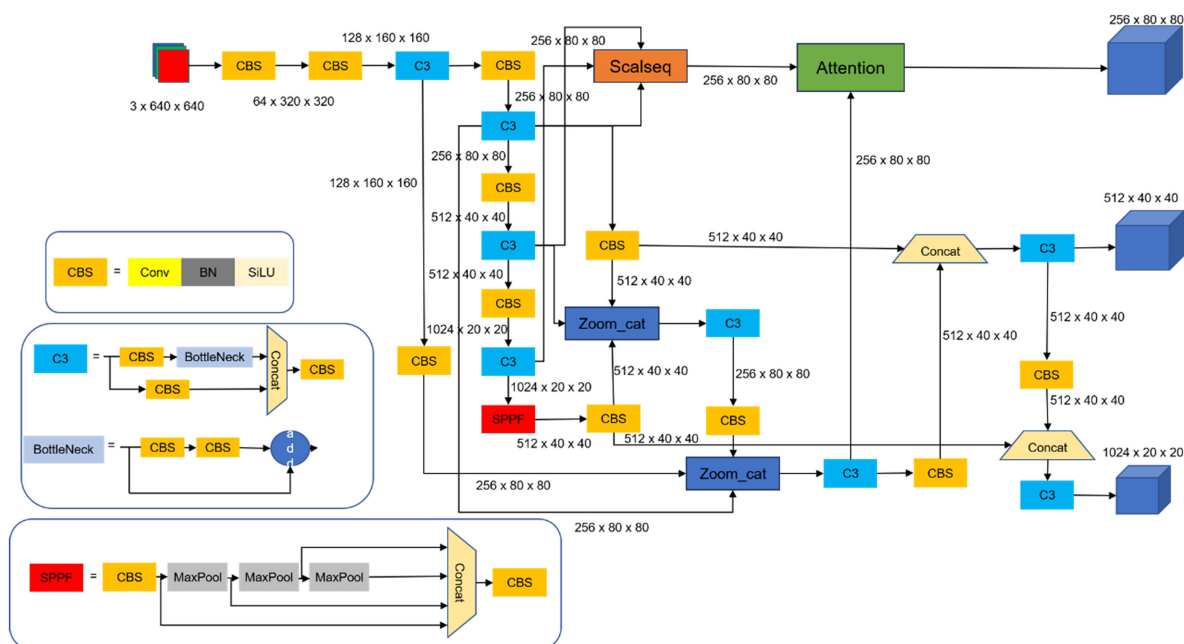


Figure 1. Network structure of the improved YOLOv5 algorithm

图 1. 改进的 YOLOv5 算法网络结构图

3. 实验设计与结果分析

3.1. 数据集

本研究数据集较难获取，通过网络搜寻和实地拍摄共采集了 500 张粘蝇纸图像，这些图像在不同光照、角度和环境条件下拍摄，以尽可能涵盖家蝇在实际应用中出现的多样情况。每张图像经过人员仔细标注，主要标记出粘蝇纸上家蝇的位置信息，确保目标框精确覆盖目标区域。数据标注过程中，重点对家蝇这种小目标的特征进行捕捉，保证在复杂背景下也能体现出目标细节。数据集被划分为训练集和测试集，分别包含 400 张和 100 张图像，为模型的训练与评估提供了充足而具有代表性的样本。

3.2. 实验结果与分析

在实验中，采用 Windows 系统端进行训练，学习率设置为 0.01，训练轮数为 100 轮。硬件配置是 GPU 为 3090Ti 和 CPU 为 AMD R7745 的处理器。本文的实验配置环境 Python 版本为 3.8.8，PyTorch 版本为 2.1.0 + cu121，CUDA 版本为 11.1，利用划分好的 400 张训练集和 100 张测试集进行模型训练与评估。

YOLOv5 算法的模型评估指标主要包括精度(Precision)、召回率(Recall)、平均精确度(AP)和所有类别

的平均准确率(mAP)。

精度(Precision)的定义是在所有预测目标中正确预测目标的百分比。其表达式如下：

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (1)$$

召回率(Recall)的定义是，在所有被标注为正确的目标中，被正确预测的目标所占的比例。其表达式如下：

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

在上述公式中，真阳性(TP)表示预测正确的阳性示例的数量，假阳性(FP)表示预测错误的阴性示例的数量，假阴性(FN)表示预测错误的阳性示例的数量。

平均精确度(AP)是衡量训练有素的网络模型在单一类别中有效性的重要指标。其表达式如下：

$$\text{AP} = \int_0^1 \text{P(R)} \text{dR} \quad (3)$$

其中 P(R)表示 P 和 R 形成的曲线。

所有类别的平均准确率(mAP)是评估训练有素的网络模型在每个类别中整体性能的重要指标。以下表达式说明了 mAP 的计算方法：

$$\text{mAP} = \frac{\sum_{i=1}^N \text{AP}_i}{N} \quad (4)$$

改进后的 YOLOv5 模型在多个评价指标上均优于 YOLOv5m、YOLOv8m、YOLOv10m 模型，如表 1 所示。具体表现在：关于 mAP(平均精度均值)，改进模型在各个尺度目标上均获得更高的检测精度，充分体现了多尺度特征融合和注意力机制的优势。改进模型的 mAP 值较原始 YOLOv5m 提升了 8 个百分点。关于 Precision(精确率)，得益于精细化的特征提取与特征融合策略，模型在抑制误检上取得了显著改善。关于 Recall(召回率)，增强的小目标检测能力使得模型在检测粘蝇纸上的家蝇时，漏检率明显降低。对于各模块我们进行了消融实验，基于 YOLOv5m 模型分别加入各模块进行模型训练，结果如表 1 所示，基于 YOLOv5m 来对比，分别加入各模块后各评价指标均发现有一定的提升效果。Zoom_cat 模块对性能提升有一定影响，尤其是在多尺度目标检测方面。ScalSeq 模块对性能提升也有显著贡献，尤其是在融合多尺度特征方面。Attention_model 模块对性能提升有显著贡献，尤其是在处理复杂背景时。如图 2 所示可以看出，改进的算法通过引入 Zoom_cat 模块、ScalSeq 模块以及注意力机制(Attention_model 模块)，实现了多尺度特征的高效融合和关键特征的突出，相较于原 YOLOv5m 模型的监测效果，对于更小目标的家蝇监测有了明显提升。

Table 1. A comparison of performance metrics for different model training

表 1. 不同模型训练的性能指标对比

	mAP0.5 (%)	mAP0.5:0.95 (%)	Recall (%)	Precision (%)
本文模型	93.3	60.1	94	92
YOLOv5m + Zoom_cat	89.7	53.5	90	90
YOLOv5m + ScalSeq	90.5	56.8	91	90
YOLOv5m + Attention_model	89.8	55.9	88	90
YOLOv5m	88.5	52.3	89	90
YOLOv8m	86.1	48.2	87	86
YOLOv10m	91.1	54.4	90	89

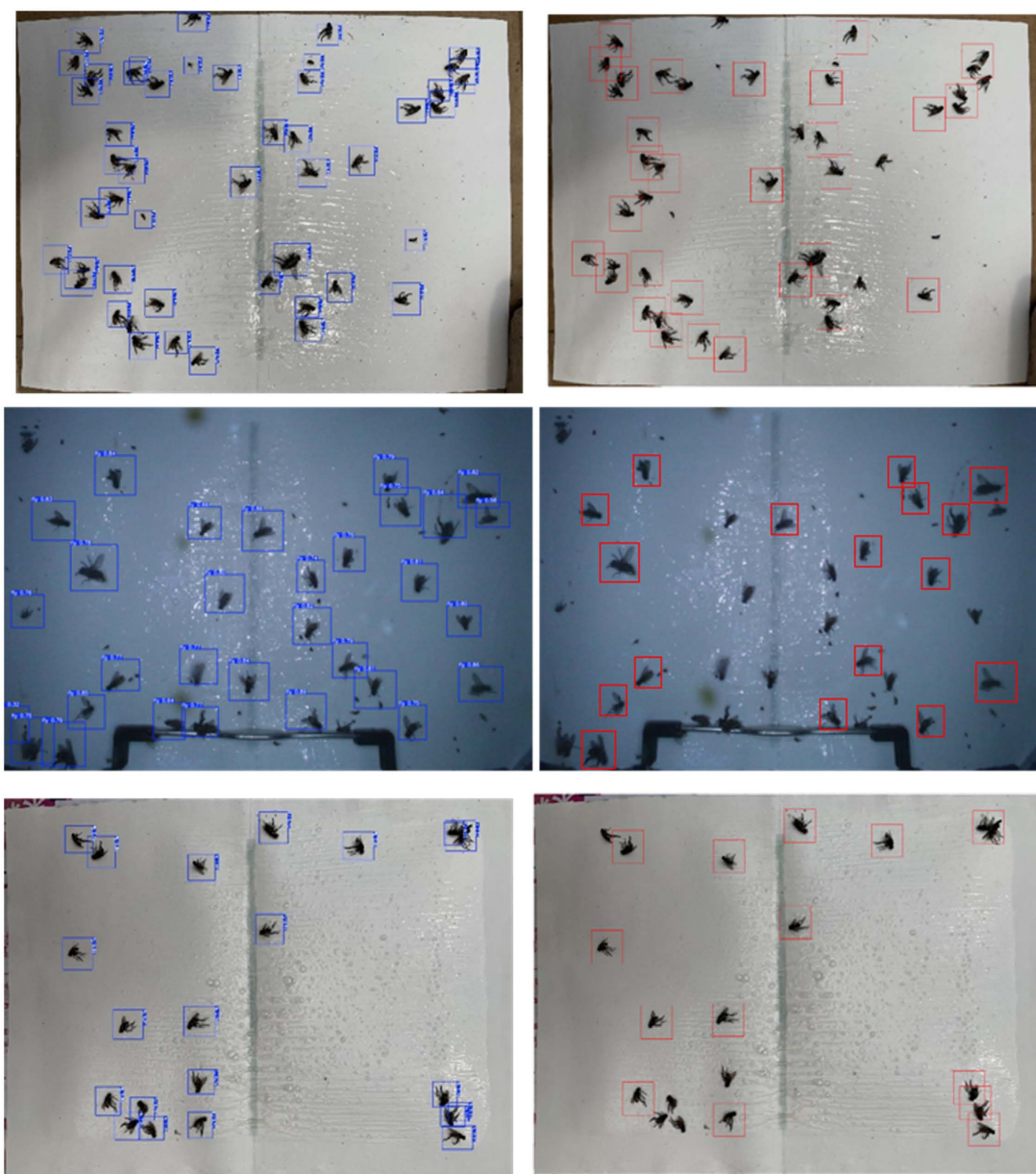


Figure 2. Detection comparison between the improved algorithm and the original YOLOv5 algorithm

图 2. 改进算法与原 YOLOv5 算法检测对比图

综上所述, 本文通过对 YOLOv5m 模型的检测头部分改进, 有效提升了对粘蝇纸家蝇的检测效果。实验结果表明, 改进模型在 mAP、精确率和召回率等关键指标上均优于原始 YOLOv5m, 为粘蝇纸家蝇的数量检测提供了一种更高效、更准确的计数方案。

4. 结论

本文针对粘蝇纸上家蝇的自动检测任务, 提出了一种基于 YOLOv5 改进的检测算法。通过在检测头部分引入 Zoom_cat 模块、ScalSeq 模块以及注意力机制(Attention_model 模块), 实现了多尺度特征的高效融合和关键特征的突出, 从而显著提升了模型对小目标家蝇的检测能力。在构建了包含 500 张粘蝇纸图像的数据集, 并划分为 400 张训练集和 100 张测试集的基础上, 实验结果显示, 改进后的模型在 mAP、

精确率和召回率等指标上均较原始 YOLOv5m 模型有明显提升, mAP 值提升了 8 个百分点。证明了所采用算法在实际场景中具有较好的适应性和鲁棒性。

综上所述, 本文的改进方法不仅提高了家蝇在粘蝇纸图像中的检测准确性, 也为小目标检测问题提供了新的思路。未来工作可以进一步扩充数据集规模、探索更高效的网络结构以及引入更先进的注意力机制, 以期在更复杂的环境下实现更为精准和高效的目标检测。并将算法部署至轻量化设备中进行研究。

参考文献

- [1] 叶耀文. 基于机器学习的蔬菜重大害虫监测计数算法及其应用系统[D]: [硕士学位论文]. 广州: 华南农业大学, 2018.
- [2] Shariff, A., Aik, Y., Hong, W., *et al.* (2006) Automated Identification and Counting of Pests in the Paddy Fields Using Image Analysis. *Proceedings of the 4th World Congress on Computers in Agriculture*, Orlando, 24-26 July 2006, 759-764.
- [3] Fukatsu, T., Watanabe, T., Hu, H., Yoichi, H. and Hirafuji, M. (2012) Field Monitoring Support System for the Occurrence of *Leptocorisa chinensis* Dallas (Hemiptera: Alydidae) Using Synthetic Attractants, Field Servers, and Image Analysis. *Computers and Electronics in Agriculture*, **80**, 8-16. <https://doi.org/10.1016/j.compag.2011.10.005>
- [4] 沈佐锐, 于新文. 温室白粉虱自动计数技术研究初报[J]. 生态学报, 2001, 21(1): 94-99.
- [5] 邱道尹, 张红涛, 刘新宇, 等. 基于机器视觉的大田害虫检测系统[J]. 农业机械学报, 2007, 38(1): 120-122.
- [6] Wen, C. and Guyer, D. (2012) Image-Based Orchard Insect Automated Identification and Classification Method. *Computers and Electronics in Agriculture*, **89**, 110-115. <https://doi.org/10.1016/j.compag.2012.08.008>
- [7] 韩瑞珍, 何勇. 基于计算机视觉的大田害虫远程自动识别系统[J]. 农业工程学报, 2013, 29(3): 156-162.
- [8] Rustia, D.J.A., Lin, C.E., Chung, J., Zhuang, Y., Hsu, J. and Lin, T. (2020) Application of an Image and Environmental Sensor Network for Automated Greenhouse Insect Pest Monitoring. *Journal of Asia-Pacific Entomology*, **23**, 17-28. <https://doi.org/10.1016/j.aspen.2019.11.006>