

基于YOLOv5s的蝴蝶种类检测

赵凯旋^{1,2*}, 覃林^{1,2}, 谢本亮^{1,2,3}

¹贵州大学大数据与信息工程学院, 贵州 贵阳

²教育部功率半导体器件可靠性工程中心, 贵州 贵阳

³贵州大学公共大数据国家重点实验室, 贵州 贵阳

收稿日期: 2023年10月20日; 录用日期: 2023年11月22日; 发布日期: 2023年11月29日

摘要

蝴蝶对周围环境敏感, 能作为反应生态环境的指示物种, 因此对其进行识别研究对研究生态稳定性具有重大意义。但蝴蝶分类细致, 相似度高, 传统识别方法效率低。为解决上述问题, 本文以野外蝴蝶图像的种类自动识别为目标, 提出了一种基于YOLOv5s的改进的目标检测方法。为了减少信息丢失, 提高精度, 在YOLOv5s的主干特征提取网络上设计了CSandGlass模块来代替残差模块; 并加入了SE注意力机制和对损失函数进行改进。实验结果表明, 改进后模型平均精度为92.6%, 相比原模型平均精度提升2%, 且具有较强的鲁棒性和稳定性, 可满足自然环境下的蝴蝶种类识别需求。

关键词

蝴蝶识别, 目标检测, YOLOv5s, CSandGlass, SE注意力

Butterfly Species Detection Based on YOLOv5s

Kaixuan Zhao^{1,2*}, Lin Qing^{1,2}, Benliang Xie^{1,2,3}

¹College of Big Data and Information Engineering, Guizhou University, Guiyang Guizhou

²Power Semiconductor Device Reliability Engineering Center of the Ministry of Education, Guiyang Guizhou

³State Key Laboratory of Public Big Data, Guizhou University, Guiyang Guizhou

Received: Oct. 20th, 2023; accepted: Nov. 22nd, 2023; published: Nov. 29th, 2023

Abstract

Butterflies are sensitive to their surroundings and can be used as indicator species responding to

*通讯作者。

文章引用: 赵凯旋, 覃林, 谢本亮. 基于YOLOv5s的蝴蝶种类检测[J]. 建模与仿真, 2023, 12(6): 5834-5842.

DOI: 10.12677/mos.2023.126529

the ecological environment, so their identification studies are of great significance for studying ecological stability. However, butterfly classification is highly detailed with high similarity, and the traditional recognition methods are inefficient. In order to solve the above problems, this paper proposes an improved object detection method based on YOLOv5s with the goal of automatic species identification of butterfly images. In order to reduce the information loss and improve the accuracy, the CSandGlass module is designed on the backbone feature extraction network of YOLOv5s to replace the residual module; and the SE attention mechanism and the loss function are added and improved. The average accuracy of the improved model is 92.6%, compared with the original model, the average accuracy is improved by 2%, and has strong robustness and stability, which can meet the demand of butterfly species recognition in natural environment.

Keywords

Butterfly Recognition, Object Detection, YOLOv5s, CSandGlass, SE Attention

Copyright © 2023 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

蝴蝶是动物界的一种昆虫，在生态系统中起着重要的作用，其对周围环境敏感，可作为地方生态环境的指示物种[1]，研究生态学的专家们会通过蝴蝶的种类识别来判断当地的生态环境[2]，因此研究蝴蝶种类识别对研究生态系统的稳定十分重要。地球上所有的蝴蝶都有相似的外表，科学家根据蝴蝶翅膀的形状和行为将蝴蝶分为不同的类别，世界上已记载的有两万多种，在中国有 2100 多种[3]。值得注意的是，蝴蝶幼虫大部分是农作物和树木的害虫，与农作物和人类和动物的环境有着直接的关系。因此，可靠的物种鉴定方法在害虫防治、边境防治和环境保护等方面具有非常重要的实际应用价值。蝴蝶的种间差异很小，而种内变异很大，这使得昆虫学家很难确定蝴蝶的种类。此外，采集、识别和存档标本图像对昆虫学家来说也是一项繁琐而昂贵的工作。蝴蝶种类的自动准确识别具有重要的现实意义，可以为生物多样性研究和物种分析提供帮助。

在蝴蝶识别领域，早期大多采用传统的机器学习算法，如局部二进制模式(LBP)、K-最近邻(KNN)和支持向量机(SVM)，被广泛用于解决分类问题，特别是在小数据集上表现良好。G.M 等使用二值化等图像处理工具对蝴蝶进行分割，并将 RGB 图像转换为 HIS 颜色模型，对 9 种蝴蝶进行分类取得了良好的识别效果[4]。但对于更复杂的数据集，深度学习的表现优于传统的机器学习。谢娟英等基于他们的数据集提出了一种新的数据分区方法和数据扩展技术，并在 RetinaNet 模型上实现了最高 79.7% 的 mAP [5]。李飞等人提出 LDResNet 网络对自然环境蝴蝶图像进行识别，引入可变形卷积，注意力机制和改进的深度可分离卷积，取得了 87.61% 的平均识别准确率[6]。目标检测识别算法具有高效，准确地自动学习、检测分类能力。其主要分为单阶段和两阶段目标检测算法。两阶段目标检测算法会先在第一阶段生成一些候选框，然后在第二阶段进行对生成的候选框进行分类和回归。而单阶段算法则直接在整个图像上预测目标的位置和类别。双阶段算法对目标检测的精度较高，而单阶段算法的效率更高。YOLO 系列算法是目前目标检测的主流算法[7] [8] [9]。Liang 等综合不同训练机制下的 YOLO 模型的结果，提出了一种基于 YOLO 算法的蝴蝶自动检测分类算法。极大地提高了 YOLO 算法的泛化能力，使其具有更好地解决小样本问题的能力[10]。

上述研究在蝴蝶识别检测上已取得一定进展，但仍可以进一步优化：1) 对同科的识别精度不够高，采用 CSandGlass 模块和 SE 注意力机制。2) 优化损失函数，提高蝴蝶检测的精确度和鲁棒性。

2. YOLOv5s 网络结构

YOLOv5s 在结构上可以分为主干网络，颈部和头部。主干部分使用 CSPDarknet 作为特征提取网络从输入图像中提取特征。颈部使用 PANet 生成一个特征金字塔网络，对特征进行聚合，并将其传递给头部进行预测。头部会根据锚框生成预测的层，用于对象检测。YOLOv5 还采用了 Mosaic 数据增强方法，该方法会对四张图片进行随机裁剪与拼接，以增强数据的多样性，同时增加更多的小目标。它还使用二元交叉熵损失函数，sigmoid 作为激活函数，解决了损失函数权重更新过慢的问题。同时 YOLOv5s 还使用了 Focal Loss 技术，能有效缓解类别不平衡问题。YOLOv5s 结构如图 1 所示。

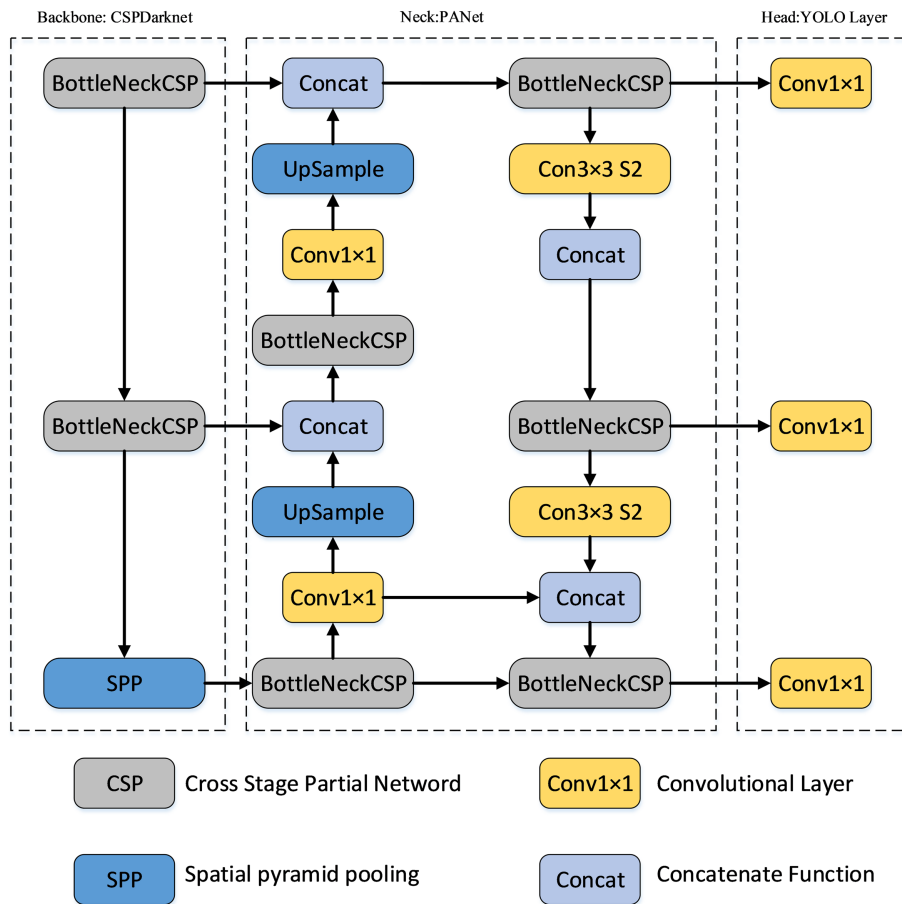


Figure 1. YOLOv5s network architecture
 图 1. YOLOv5s 网络结构

3. YOLOv5s 网络改进

3.1. 主干网络改进

3.1.1. CSandGlass 模块

YOLOv5s 中的 CSPNET 结构将基础层的特征层分为两部分，然后使用跨阶段层次结构将两者合并，从而使网络能实现更丰富的梯度组合信息，但这也更容易造成信息丢失和梯度混淆。因此，本文借鉴了

MobileNext [11]的思想,使用沙漏状模块 CSandGlass 来替代 YOLOv5s 网络中的 Res 单元模块。CSandGlass 模块的结构如图 2 所示。与中间有深度卷积的瓶颈结构不同,本文用高维表示将 3×3 深度卷积层(Dwise)移动到残差路径的两端,并将 YOLOv5s 的两个基本组件 CBL 在中间。两个深度卷积可以编码更多的空间信息,并使更多的梯度在多层间传播,减少信息丢失。使用 CSandGlass 改进的特征有较好的提取效果,背景信息和特征信息也更加清晰。

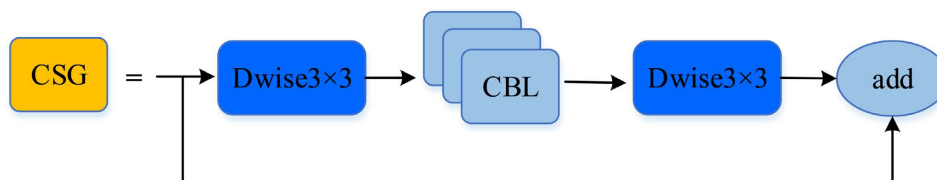


Figure 2. Structure diagram of CSandGlass
图 2. CSandGlass 结构图

3.1.2. SE 注意力机制

SE 注意力机制是通道注意力模式下一种确定权重的方法,SE 注意力机制[12]旨在研究通道特征的相互依赖性。而非引入新的空间维度 为了融合通道上的特征图,该模块使用新的特征重新校准策略。SE 注意力机制通过学习获得每个通道特征的重要性,然后据此提升有用的特征或抑制无用的特征重要性。它通过在不同通道间分配权重达到主次优先的目的。图 3 为 SE 注意力机制的结构图。

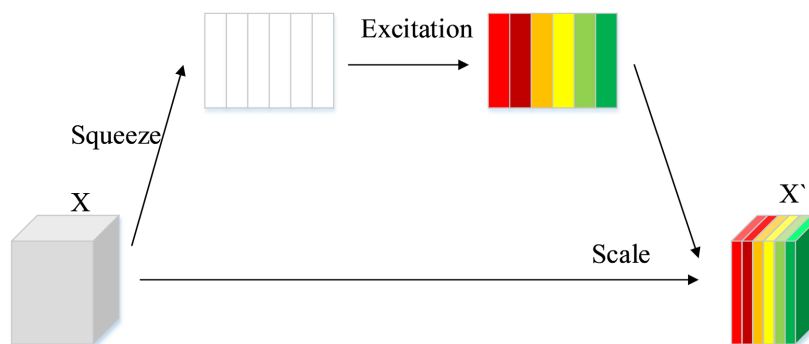


Figure 3. SE attention module
图 3. SE 注意力模块

先经过全局平均池化将每个通道进行特征压缩,从而使每个通道由一个数值所表示。压缩的操作如式(1)所示。

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (1)$$

上式中: z_c 为压缩操作后的输出结果; u_c 为输入矩阵的第 c 个通道数; H 和 W 分别为通道空间的长和宽。

得到特征图的每个通道特征权重后,由于其间并没有相关性,因此进行两次全连接进行激励操作获得不同通道间的相关性。第一层全连接采用了 ReLU 激活函数,目的是压缩通道数量,降低计算量;第二层全连接采用了 Sigmoid 激活函数,将通道特征缩放到 0 到 1 之间,使网络具有更多非线性,加强通道间相关性的拟合,并把通道升到原始的维度。然后再将通道的权重加权到输入特征图的每个通道上,形成新的特征图。具体操作如式(2) (3)所示。

$$O_c = \sigma(W_1 \delta(W_0(z_c))) \quad (2)$$

$$Y = u_c O_c \quad (3)$$

上式中 σ 和 δ 分别为 Sigmoid 激活函数和 ReLU 激活函数； W_0 为第一个全连接层，降低通道数量； W_1 为第二个全连接层，上升通道维度； O_c 为激励操作的输出。将 O_c 的通道权重值与输入 u_c 相乘得到 Y ，为经过通道注意力机制后的特征图。

3.1.3. 损失函数改进

YOLOv5 模型共使用三种损失函数，采用分类损失计算锚框与对应的分类是否正确；使用定位损失判定预测框与标定框的误差；采用置信度损失来计算网络置信度，以更好地进行目标检测任务。它采用二元交叉熵作为置信度损失和分类损失的损失函数，提高模型准确性和收敛速度；采用 CIOU 作为定位损失的损失函数，更好处理不同目标间的交叉问题。如式(4) (5)所示。

$$\text{IOU} = \frac{A \cap B}{A \cup B} \quad (4)$$

$$\text{CIOU} = \text{IOU} - \frac{p^2(b, b^{gt})}{c^2} - av \quad (5)$$

A 为预测框面积， B 为真实框面积。 $p^2(b, b^{gt})$ 代表了预测框和真实框中心点之间的欧式距离。 c 代表的是能同时包含预测框和真实框的最小闭包区域的对角线距离。 av 为长宽比。

本文使用 α -CIOU 做定位损失，在 CIOU 基础上引入了一个 α 参数[13]。通过调整参数 α ，可以使检测头具有更高的灵活性。调整合适的 α 可以增加目标边界框的权重，从而可以更多地关注与目标匹配的边界框，加快网络收敛，并且减少噪声边界框对模型性能的影响。 α -CIOU 在训练轻量级模型时表现良好，鲁棒性更强，可以很好地处理噪声边界框，提高检测精度和稳定性。如式(6) (7)所示。

$$L_{\alpha\text{-CIOU}} = 1 - \left(\text{IOU}^\alpha - \frac{p^{2\alpha}(b, b^{gt})}{c^{2\alpha}} - \frac{v^{2\alpha}}{((1 - \text{IOU}) + v)^\alpha} \right) \quad (6)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w^p}{h^p} \right)^2 \quad (7)$$

4. 蝴蝶数据集制作

4.1. 蝴蝶数据集制作

本文使用的数据集来自拍摄、公共数据集和互联网。有 985 张，以 JPG 格式保存，尺寸约为 600×400 像素。数据集包含 5 种蝴蝶，同属粉蝶科，即迁粉蝶，斑粉蝶等，每类图像数量均在 100 张以上。一些蝴蝶样例如图 4 所示。每个类别中的图像数量超过 100 张。本文使用“LabelImg”标注工具对数据集进行手动标注，标注信息存储在 XML 标签文件中，其中每个标注文件逐个对应一个图像文件，每行包含一个目标信息，其中信息按顺序排列：目标类别，检测框中心点的 X 和 Y 轴坐标，以及目标宽度和高度。

4.2. 图像预处理

在目标检测中，如果数据量较少，模型易出现过拟合现象。因此利用图像增强技术生成新的样本数据，扩大数据集大小，提高模型的泛化能力，从而提高目标检测的准确性和稳定性[14] [15]。

本文选择的图像增强方法包括平移、垂直翻转、水平翻转、镜像翻转和缩放增强，随机裁剪等方式。数据增强后生成了带标签的蝴蝶图像共 2995 张。训练集和测试集按照 8:2 的比例划分。

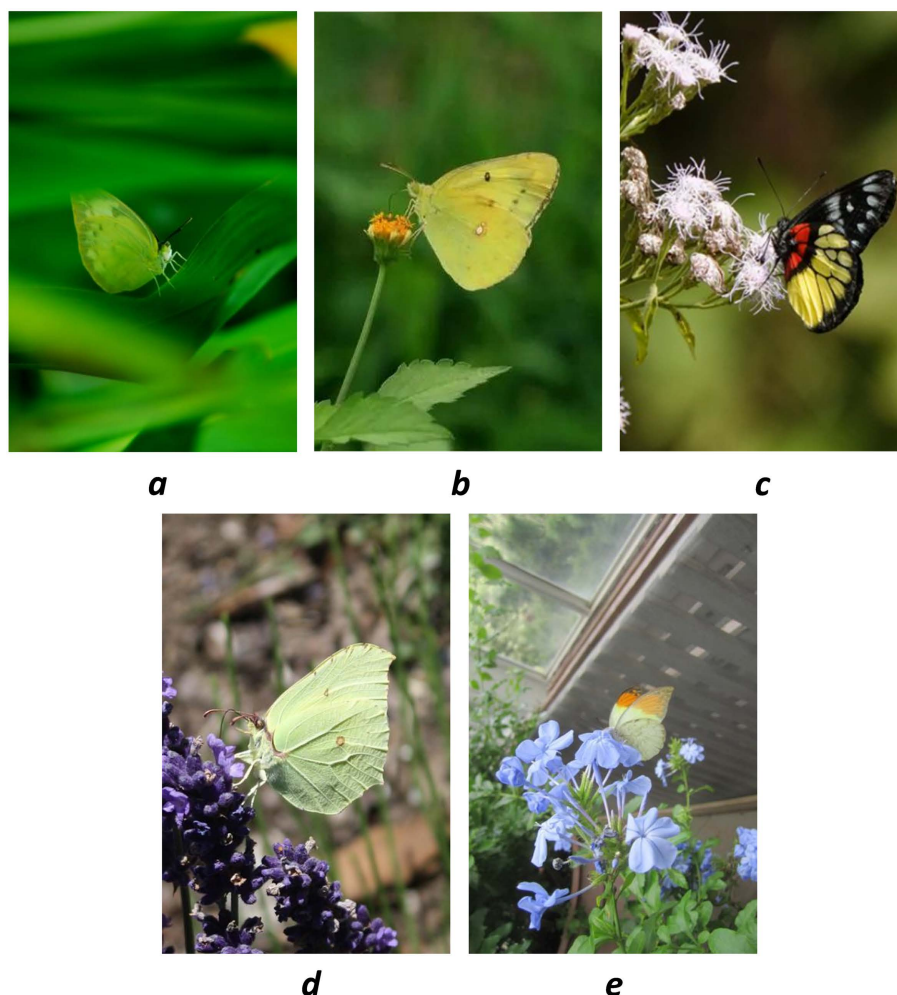


Figure 4. Butterfly samples. (a) *Catopsilia_pomona*; (b) *Colias_erate*; (c) *Delias_pasithoe*; (d) *Gonepteryx_rhamni*; (e) *Hebomoia_glaucippe*

图 4. 蝴蝶样例。(a) 迁粉蝶；(b) 豆粉蝶；(c) 斑粉蝶；(d) 钩粉蝶；(e) 鹤顶粉蝶

5. 实验结果与分析

5.1. 实验参数

本文实验的操作系统为 Windows10, 使用 Anaconda 创建虚拟环境。其中, python 版本为 3.8.1, pytorch 版本为 1.7.1, GPU 为 NVIDIA RTX3080, CUDA 版本为 11.0。

YOLOv5 的输入端可以对输入图片的大小调整, 预设定图片大小为 256×256 。Batchsize 设定为 8, work 设定为 1。训练轮数为 300 轮, 采用随机梯度下降作为优化算法, 初始学习率为 0.01。

5.2. 对比实验

本文选取 YOLOv3, YOLOv4, SSD, YOLOv5s 和本文改进的 YOLOv5s_CS 模型进行对比, 在超参数完全相同的情况下, 在同个实验设备上运行, 从精确率(P)、F1、平均精度(mAP)三方面进行比较评价, 不同网络实验结果如表 1 所示。使用改进的 YOLOv5s_CS 算法对不同蝴蝶种类检测结果如表 2 所示。

$$P = \frac{TP}{TP + FP} \times 100\% \quad (8)$$

$$R = \frac{TP}{TP + FP} \times 100\% \tag{9}$$

$$F1 = \frac{2TP}{2TP + FP + FN} \times 100\% \tag{10}$$

$$mAP = \frac{1}{n} \sum_{k=1}^n J(P, R)_k \tag{11}$$

其中, TP 表示正确检测的样本数量, FP 表示误检的样本数, FN 表示漏检的样本数, $J(P, R)_k$ 是平均精度函数。

Table 1. Experimental results of different models

表 1. 不同模型实验结果

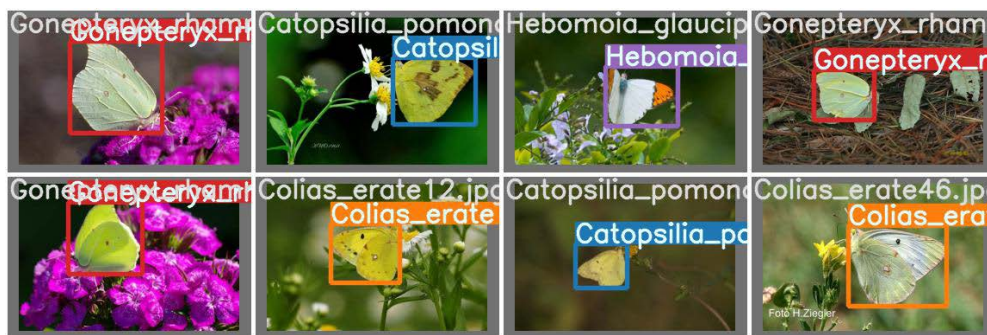
网络模型	精确率%	F1	mAP@0.5%	FPS
YOLOv3	77.1	59.8	74.8	6.8
YOLOv4	82.8	63.2	73.0	2.2
SSD	90.1	87.8	91.6	7.0
YOLOv5s	94.3	90.0	90.6	37.0
YOLOv5s_CS	93.9	91.0	92.6	35.1

Table 2. Different categories of butterfly identification results

表 2. 不同类别蝴蝶识别结果

标签	精确率%	召回率%	mAP@0.5%	mAP@0.95%
Catopsilia_pomona	95.2	64.5	77.0	60.9
Colias_erate	95.8	95.1	94.9	81.6
Delias_pasithoe	97.4	96.4	98.1	79.4
Gonepteryx_rhamni	89.6	92.5	95.5	73.0
Hebomoia_glaucippe	95.2	89.7	97.7	82.9

本文选择了几种目标检测模型与本文提出的 YOLOv5s_CS 在同台机器进行对比。从表 1 的四个方面进行对比分析, 本文提出的 YOLOv5s_CS 模型相比其他模型具有较高的准确率, F1, 和平均精度, 尽管 FPS 有所下降, 但是对识别效率影响不是很大, 这意味着本模型进行蝴蝶检测时具有更高的准确率。图 5 为本模型测试效果, 可以看出本文所提出模型达到了较好的检测效果。



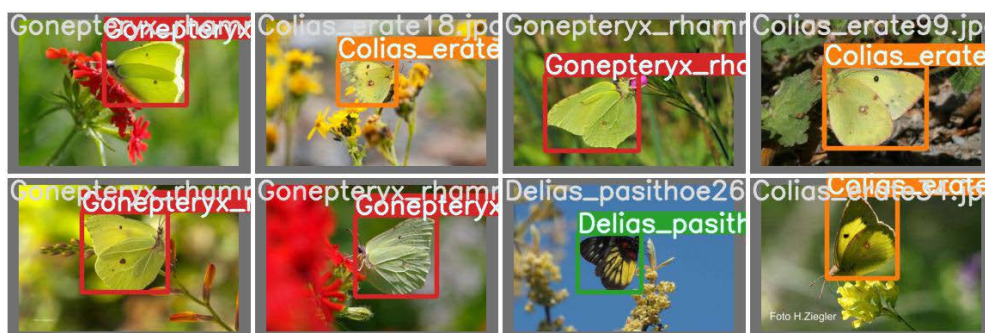


Figure 5. The effect of test

图 5. 测试效果

6. 结论

本文针对粉蝶科下的蝴蝶识别研究,采用 YOLOv5s 作为基础网络,使用 CSandGlass 来替代 YOLOv5s 网络中的 Res 单元,解决信息丢失和梯度混淆的问题;又引入了 SE 注意力机制,将蝴蝶本身特征的权重增强;为增加目标相应的边界框权重,在 CIoU 基础上引入 α 系数来计算定位损失,加快模型的收敛并降低噪声干扰。

根据上述实验和测试的结果,表明本文提出的模型能对同科目下的蝴蝶具有更高的识别精度,适用于蝴蝶分类检测研究,并且模型较轻量化,适合部署到嵌入式设备中进行实时检测。在后续的研究中,将着重于扩充蝴蝶数据集,提升小目标,多种类蝴蝶的识别精度,模型轻量化研究。

基金项目

这项工作得到了国家重点研发计划(编号:2016YFD0201305-07)、贵州省科技项目(No.ZK[2023]060),教育部半导体功率器件可靠性工程中心开放基金项目(No.ERCMEKFJJ2019-06)。贵州大学公共大数据国家重点实验室的计算支持。

参考文献

- [1] Kumar, S., Simonson, S.E. and Stohlgren, T.J. (2009) Effects of Spatial Heterogeneity on Butterfly Species Richness in Rocky Mountain National Park, CO, USA. *Biodiversity and Conservation*, **18**, 739-763. <https://doi.org/10.1007/s10531-008-9536-8>
- [2] 尹晶, 施雯, 王灵敏, 等. 云南省元阳县蝴蝶群落结构与物种多样性研究[J/OL]. 云南农业大学学报(自然科学): 1-8. <http://kns.cnki.net/kcms/detail/53.1044.S.20231010.0927.002.html>
- [3] 武春生, 徐增峰. 中国蝴蝶图鉴[M]. 福州: 海峡出版发行集团, 2017.
- [4] Grajales-Múnera, J.E., and Restrepo-Martinez, A. (2013) Clasificación de Mariposas por Modelos de Color HSI y RGB Usando Redes Neuronales. *Tecno Lógicas*, **669**. <https://doi.org/10.22430/22565337.324>
- [5] 谢娟英, 鲁银圆, 孔维轩, 等. 基于改进 RetinaNet 的自然环境中蝴蝶种类识别[J]. 计算机研究与发展, 2021, 58(8): 1686-1704.
- [6] 李飞, 赵凯旋, 严春雨, 等. 基于残差网络的自然环境下蝴蝶种类识别[J]. 昆虫学报, 2023, 66(3): 409-418.
- [7] Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/CVPR.2016.91>
- [8] Redmon, J., and Farhadi, A. (2017) YOLO9000: Better, Faster, Stronger. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 6517-6525. <https://doi.org/10.1109/CVPR.2017.690>
- [9] Redmon, J.F.A. (2018) Yolov3: An incremental improvement. arXiv: abs/1804.02767.
- [10] Liang, B., Wu, S., Xu, K., and Hao, J. (2020) Butterfly Detection and Classification Based on Integrated YOLO Algo-

rithm. *Genetic and Evolutionary Computing*, Vol. 1107, Springer, Singapore, 500-512.
https://doi.org/10.1007/978-981-15-3308-2_55

- [11] Zhou, D., Hou, Q., Chen, Y., Feng, J., and Yan, S. (2020) Rethinking Bottleneck Structure for Efficient Mobile Network Design. In: Vedaldi, A., Bischof, H., Brox, T., and Frahm, J.-M., Eds., *Computer Vision—ECCV*, Springer International Publishing, Cham, 680-697. https://doi.org/10.1007/978-3-030-58580-8_40
- [12] Hu, J., Shen, L., Sun, G. and Li, Y. (2018) Squeeze-and-Excitation Networks. *IEEE Conference on Computer Vision and Pattern Recognition*, **71**, 32-41. <https://doi.org/10.1109/CVPR.2018.00745>
- [13] 郑远攀, 许博阳, 王振宇. 改进的 YOLOv5 烟雾检测模型[J]. 计算机工程与应用, 2023, 59(7): 214-221.
- [14] 王红尧, 韩爽, 李勤怡. 改进 YOLOv5 的钢丝绳损伤图像识别实验方法研究[J]. 计算机工程与应用, 2023, 59(17): 99-106.
- [15] 杨谢柳, 门国文, 梁文峰, 等. 水下图像增强与复原对深度学习目标检测精度的影响研究[J/OL]. 计算机工程: 1-10. <https://doi.org/10.19678/j.issn.1000-3428.0066610>