

川金丝猴轻量化检测

徐浩然^{1,2}

(1. 陕西理工大学机械工程学院, 陕西 汉中 723000;
2. 陕西省工业自动化重点实验室, 陕西 汉中 723000)

摘要: 为实现复杂环境下川金丝猴的快速检测, 提出了一种基于模块重构的川金丝猴检测模型 PCB-YOLOv5n。采用轻量化卷积 PConv 构建残差分支改进模块 CPB, 以减少模型计算量和内存访问量, 引入轻量化注意力模块 ECA 对 YOLOv5n 改进特征提取网络, 自适应为通道特征加权, 实现局部跨通道信息交互。在自建川金丝猴数据集上, PCB-YOLOv5n 与 YOLOv5n 模型相比, 推理速度提升明显。为川金丝猴轻量化检测设备部署提供了更小的计算代价和更高的检测速度。

关键词: 川金丝猴检测; 轻量化; YOLOv5; 模型重构

川金丝猴(*Rhinopithecus roxellanae*)作为中国独有的保护动物, 深受中国人民的喜爱。但是由于自然资源过度开采、气候变化等影响, 川金丝猴的生存、繁衍仍然面临严峻的威胁和挑战。为更好地保护川金丝猴, 维护物种多样性, 对川金丝猴的观察研究变得尤为重要。

随着计算机视觉领域蓬勃发展, 可以利用轻量化设备结合目标检测算法代替人工巡查, 对川金丝猴出现区域进行智能化监测。近几年, 由于卷积神经网络(CNN)的加入, 川金丝猴检测识别步入了智能化阶段。本文采用 YOLOv5 6.0 中参数量最小的 YOLOv5n 作为基准模型, 使用轻量化模块 CPB 和注意力机制 ECA 对模型进行重构。在满足检测精度的同时大幅降低模型计算复杂度, 为川金丝猴智能监测提供技术支撑。

一、数据集与方法

(一) 数据集制作

本实验所用数据集为陕西佛坪熊猫谷(33° 40' 6.183" N, 107° 58' 23.92" E)自媒体拍摄数据, 使用 Labelimg 标注工具对所构建的 1160 张图片中的川金丝猴面部和身体进行 YOLO 格式标注, 形成初始数据集。通过对初始数据集进行数据增强, 使数据集扩增到 5800 张。其中包含身体目标 14027 个, 猴脸目标 6528 个。对增强后数据集设置随机数种子 123, 按照 8: 1: 1 的比例划分为训练集、验证集和测试集(图 1)



图 1 数据集样本

(二) CPB 模块设计

Faster_Block 是一种用于目标检测的模块结构, 它可以将特征图通过一系列的卷积层进行特征提和变换, 这些卷积层有助于网络学习到更具有区分性的特征。本文在 YOLOv5n 的 C3 模块中

引入 Faster_Block 中的部分卷积模块(PConv), 卷积核大小设置为 3×3, 同时将 Bottleneck 分支中卷积核大小为 3×3 的 CBS 换成卷积核大小为 1×1 的 CBS, 构成 PB(PConv-Bottleneck)残差分支, 分支数量设为 3 的倍数, 与模型深度倍数乘积输出。PB 分支与卷积核大小为 1×1 的 CBS 进行张量拼接后再输入卷积核大小为 1×1 的 CBS 实现通道维度的改变, 构成 CPB 模块。

(三) ECA 模块

为减小剪枝对模型精度的影响, 在网络训练时融合注意力机制。本试验采用高效通道注意力 ECA, 为每个通道分配权重, 通过引入可学习的一维卷积层捕捉通道间的依赖关系来实现。该卷积层将每个通道的特征图全局平均池化实现通道的全局描述。ECA 模块根据通道数 C 自适应确定卷积核大小 K:

$$K = \left\lceil \frac{\log_2(C)}{\lambda} + \frac{b}{\lambda_{\text{odd}}} \right\rceil$$

其中, 参数 λ 和 b 是 ECA 注意力机制的超参数。 λ 参数用于调整注意力权重, b 参数用于转移注意力权重。 $\lceil t \rceil_{\text{odd}}$ 是取绝对值并向下取整到 t 最近的奇数, 确保卷积核大小是奇数。得到核大小 K 后, ECA 模块将一维卷积应用于输入特征上, 从而学习每个通道相对于其他通道的重要性。这个过程可以用以下公式表示:

$$\omega = \sigma(\text{CID}_K(y))$$

其中, CID_K 表示一维卷积操作(卷积核大小为 K), y 表示通道, σ 表示 Sigmoid 激活函数, 通道维数越多, 局部跨通道相互作用的范围越大。融合 ECA 注意力模块在不增加显著计算负担的情况下, 实现了性能的提升。

(四) 评价指标

为综合评价模型性能, 试验采用计算性能指标包括参数数量或模型大小、网络计算量, 识别性能指标包括平均精度均值(mAP)、精确率(Precision, P)、召回率(Recall, R)对实验模型的有效性进行评价。

平均精度均值(mean Average Precision, mAP)是 2 个类别(川

金丝猴面部和川金丝猴身体)的 AP (Average Precision) 的平均值, AP 用来衡量单类别的模型平均精度。mAP 值越高, 代表模型检测性能越好。AP 和 mAP 的计算公式如下:

$$AP = \int_0^1 P(R)dR \quad mAP = \frac{1}{F} \sum_{k=1}^F AP_k$$

其中, F 为类别数。为了对川金丝猴更好的检测, 本文选择对川金丝猴面部和川金丝猴身体的检测精度取均值作为川金丝猴最终的检测精度, 从两种类别中学习更加丰富的特征, 减小目标局部缺失的漏检率。

二、结果与分析

(一) 消融试验

为了将模型轻量化, 本文设计 CPB 模块代替元模型中的 C3 模块, 在减小参数量的基础上, 融合注意力 ECA 模块, 使模型精度得到爬升的同时进一步压缩了参数量, 得到最终模型 PCB-YOLOv5n。采用 AP、参数量、计算量、FPS 对模型进行评估, 评估结果如表 1。

表 1 消融实验

模型 Model	CPB	ECA	mAP (%)	参数量 Parameter (M)	计算量 FLOPs (G)	FPS (b32)
Yolov5n	—	—	98.3	1.77	4.1	534.7
CPB	✓	—	98.2	1.58	3.6	621.0
ECA	✓	✓	98.6	1.57	3.6	633.1

通过对原 YOLOv5n 替换 CPB 模块, 模型计算量减小 0.5G, 在 32 批次大小下, 推理速度在原模型基础上提升 81FPS。融合注意力 ECA 后, 精度提升明显, 且计算量保持不变, 推理速度进一步提高, 达到了 633.1FPS。

(二) 横向模型对比试验

为评价本文模型在川金丝猴检测上的性能, 令模型与 Faster rcnn、YOLOv7-tiny、YOLOv8n、YOLOv9c 在川金丝猴数据集上进行对比试验, 测试结果如表 2 所示。分别在批次大小为 32、单张 640 × 640 的照片进行 150 次进行检测速度测试, 本文模型检测速度较快, 且模型在参数量、计算量和模型大小上与其他模型相比都具有明显优势, 检测精度上也可满足实际需求。

表 2 不同算法模型性能

模型 Model	mAP (%)	参数量 Params (M)	计算量 FLOPs (G)	FPS (b32)	模型大小 Model size (MB)
Faster rcnn	92.0	28.3	474.1	—	113.4
YOLOv7	98.3	6.0	13.2	526.3	12.3
YOLOv8	98.4	3.0	8.1	158.6	6.3
YOLOv9	98.8	50.7	236.6	52.7	102.8
Ours	98.6	1.6	3.6	633.1	3.5

图 2 为 YOLOv5n 与 PCB-YOLOv5n 在不同场景下的金丝猴检

测结果。在不同条件影响下, 两种模型均可完成金丝猴识别的任务, 但当识别目标存在遮挡严重的情况时, 树林场景中, YOLOv5n 存在漏检情况; 当背景颜色和和目标颜色相近时, 草地场景中, YOLOv5n 存在误检的情况。相较于 YOLOv5n 本模型在这两个场景中检测更准确些。

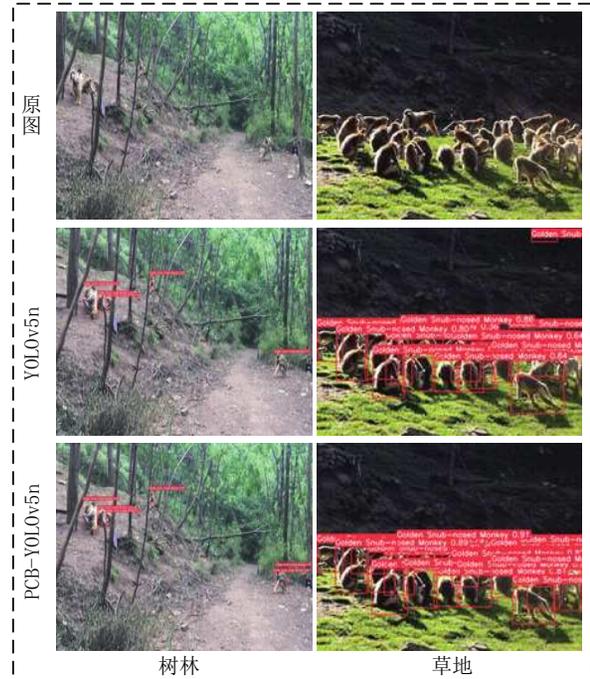


图 2 遮挡检测对比图

三、结论

本研究结合川金丝猴面部与身体特征, 通过对 YOLOv5n 使用基于 Pconv 与残差结构的 CPB 模块, 实现对川金丝猴个体的快速检测。同时, 通过融合注意力机制 ECA, 使模型精度爬升, 达到了 0.986。改进模型同原模型比参数量减小了 0.2M, 计算量减小了 0.5G, 检测速度大幅提升。可帮助致力于川金丝猴保护和研究的工作者减轻人力成本, 促进智慧化林业发展。

参考文献:

[1] 李明, 向左甫. 中国旗舰保护物种: 金丝猴 [J]. 世界环境, 2016 (S1): 12-15.

[2] 王晓宇, 替林森. 秦岭川金丝猴生物学特性及种群保护研究进展 [J]. 陕西林业科技, 2022, 50 (01): 102-107.

[3] 孙蕊, 张旭, 郭颖, 等. 基于 Faster R-CNN 金丝猴优化检测方法 [J]. 激光与光电子学进展, 2020, 57 (12): 259-268.