

研究论文 ORIGINAL PAPERS

基于伪标签和YOLOv4的
野生动物检测方法孟继森¹, 马玉明¹, 杨紫合², 孙茜³, 巨友娟⁴, 谢将剑², 张军国^{2*}

1. 天津市规划和自然资源局林业事务中心, 天津 300191;
2. 北京林业大学工学院, 北京 100083;
3. 天津商业大学信息工程学院, 天津 300134;
4. 互助土族自治县林木种苗站, 互助土族自治县 810500)

稿件运行过程

收稿日期: 2024-08-23

修回日期: 2024-10-31



关键词: 野生动物;
目标检测;
红外相机;
运动检测;
伪标签

Keywords: Wildlife;

Object detection;
Camera-trapping;
Motion detection;
Pseudo-labels

中图分类号: Q958.1; TP391.4

文献标志码: A

文章编号:

2310-1490(2025)-03-0523-10

DOI: 10.12375/ysdwxb.20250306

摘要

红外相机是监测野生动物的常用方法,但存在数据量大、背景信息复杂等问题,导致监测数据标注和检测困难。针对以上问题提出一种基于伪标签和YOLOv4的野生动物检测方法。本方法首先基于运动检测的伪标签标定方法,通过背景差分法和形态学操作实现对视频数据集的自动快速标注,降低监测环境中复杂背景的不利影响;然后通过跨阶段局部卷积块,减少YOLOv4中路径聚合网络所需的计算量;最后在密集卷积区域引入Swish激活函数,提高模型在深层区域的特征提取能力,以浙江江山仙霞岭自然保护区的6种野生动物监测视频作为数据集进行实验。结果表明:本方法在平均精度均值和帧率指标上达到了86.41%和18.93帧/s,相比于YOLOv4、RFCN、YOLOv8x算法分别提高1.62、3.43、7.11个百分点,证明所提出算法可以有效克服现有方法标注和检测困难的问题,提升了野生动物监测数据检测平均精度均值和帧率,有助于野生动物监测数据分析的自动化和智能化。

Wildlife Detection Method
Based on Pseudo-labels and YOLOv4MENG Jisen¹, MA Yuming¹, YANG Zihe²,
SUN Qian³, JU Youjuan⁴, XIE Jiangjian², ZHANG Junguo^{2*}

基金项目: 中央财政林业科技推广示范项目(津[2021]JTG02号);国家自然科学基金项目(32371874)

第一作者简介: 孟继森(1979—),男,高级工程师;主要从事林业信息化可持续研究。E-mail:653327403@qq.com

*通信作者: 张军国, E-mail:zhangjunguo@bjfu.edu.cn

- (1. Forestry Affairs Center of Tianjin Municipal Bureau of Planning and Natural Resources,
Tianjin 300191, China;
2. School of Technology, Beijing Forestry University, Beijing 100083, China;
3. School of Information Engineering, Tianjin University of Commerce, Tianjin 300134, China;
4. Forest Seed and Seeding Station of Huzhu Tu Autonomous County,
Huzhu Tu Autonomous County 810500, China)

Abstract: The camera traps is a common method to monitor wildlife. A large amount of data and complex background information make it difficult to label and detect. In order to solve the above problems, a wild animal detection method based on pseudo-labels and YOLOv4 was proposed. Firstly, a pseudo-labeling calibration method based on motion detection was proposed to realize automatic and fast calibration of video datasets by background difference method and morphological operation, which solved the problem of difficult automatic labeling caused by complex background in monitoring data. Then, a cross stage partial block was proposed to reduce the amount of computation required for the path aggregation network in YOLOv4. Finally, the Swish activation function was introduced in the dense convolution area, which improves the feature extraction ability of the model in the deep region. In this study, the monitoring videos of six kinds of wild animals in Zhejiang Jiangshan Xianxialing Provincial Nature Reserve were used as data sets for experiments. The results show that the method proposed in this study achieves 86.41% mean average precision (mAP) and a frame rate of 18.93 frames per second. These metrics represent improvements of 1.62 percentage points, 3.43 percentage points, and 7.11 percentage points over the YOLOv4, RFCN, and YOLOv8x algorithms, respectively. This demonstrates that the proposed algorithm effectively addresses the labeling and detection challenges faced by existing methods. Balanced improvement of detection mAP and frame rate contributes to automated and intelligent analysis of wildlife monitoring data.

野生动物保护是维持生态系统稳定性的基础,有效的动物监测使研究人员能够更好地了解它们的生存状况,并制定适当的保护策略^[1]。红外相机是目前常用的野生动物监测工具,其收集的大量相机监测数据已被广泛用于研究物种多样性、行为模式和种群结构估计^[2]。通过该方式可以积累大量的野生动物监测数据,但人工处理需要耗费大量的人力、物力,使得数据处理的时效性降低。引入计算机视觉技术自动处理监测图像,逐渐成为提升处理效率的主流方案。然而,野外环境复杂多变,导致监测图像的背景复杂,准确检测图像中的野生动物仍是一项具有挑战性的任务。

早期主要使用数字图像处理方法来识别图像中的野生动物。如 Steen *et al.*^[3]提出了一种基于灰度分布直方图的动物识别方法;初末萌^[4]基于边缘检测方法对野生动物图像边缘信息进行分析,并通过 Hough 变换的方式完成目标检测。图像中野生动物的特征信息会随着环境变化而变化,仅凭简单的特征很难准确识别野生动物。随后,传统机器学习方

法被引入到野生动物识别中,如 Matuska *et al.*^[5]采用 SVM 分类器对全局区域算子提取的野生动物特征进行分类,在对 5 种野生动物的分类实验中,取得了 86% 的准确率;Kamencay *et al.*^[6]利用主成分分析对野生动物图像中提取的核心特征降维后,输入到多种算法融合的分类模型中,达到了 90% 的识别准确率。但传统机器学习方法需要手动设计输入特征,无法保证特征的有效性,识别性能有待进一步提升。随着深度学习技术的发展,深度学习已经成为当前野生动物检测的主流方法。其中,目标检测方法利用深层次的卷积神经网络(convolutional neural network, CNN)能自动提取图像特征,并定位目标^[7-10],大幅减少了对人工设计特征的依赖,提升了复杂环境下野生动物的检测能力^[11-16],但这类方法在训练模型时需要大量的标记数据。目标检测模型的训练数据不仅需要标记类别,还需要给出野生动物的位置。人工标注足够的训练数据会耗费大量人力^[17-18],亟需一种新的标注方法,提升标记数据的效率。

针对以上问题,本研究在分析浙江江山仙霞岭省级自然保护区采集的野生动物监测数据的基础上,提出基于伪标签和YOLOv4的野生动物目标检测方法(wildlife detection model, WDM)。该方法首先利用背景差分法和形态学操作将视频中的野生动物自动快速标注为伪标签数据,降低监测环境中复杂背景的消极影响;然后通过跨阶段局部卷积块,改进YOLOv4模型中的路径聚合网络(path aggregation network, PANet),减少模型的参数量和运算量;最后在密集卷积区引入Swish激活函数,提高模型的特征提取能力,实现监测数据的快速准确检测。

1 野生动物监测数据的伪标签标定

采用浙江江山仙霞岭省级自然保护区的野生动物监测视频作为数据集。该数据集包含黄腹角雉

(*Tragopan caboti*)、白鹇(*Lophura nycthemera*)、灰胸竹鸡(*Bambusicola thoracicus*)、野猪(*Sus scrofa*)、亚洲狗獾(*Meles leucurus*)和小麂(*Muntiacus reevesi*)6个物种。为了减少野生动物监测图像标定的时间成本,提出一种基于人机协作的伪标签标定方法。该方法结合专家提供的物种分类标签与基于运动检测的目标框自动生成技术,实现了对视频数据集自动快速标定。

为了自动生成视频数据的目标框,首先将视频中每5帧抽出1帧作为图像数据,并将其转化为灰度图像。选择每个视频的最后一帧作为背景,对该视频中的其他图像进行差分,得到背景差分图。由于在野生动物的监测视频中存在复杂环境因素以及受动物保护色的影响,致使背景差分图中经常出现噪声和空洞现象,如图1所示。

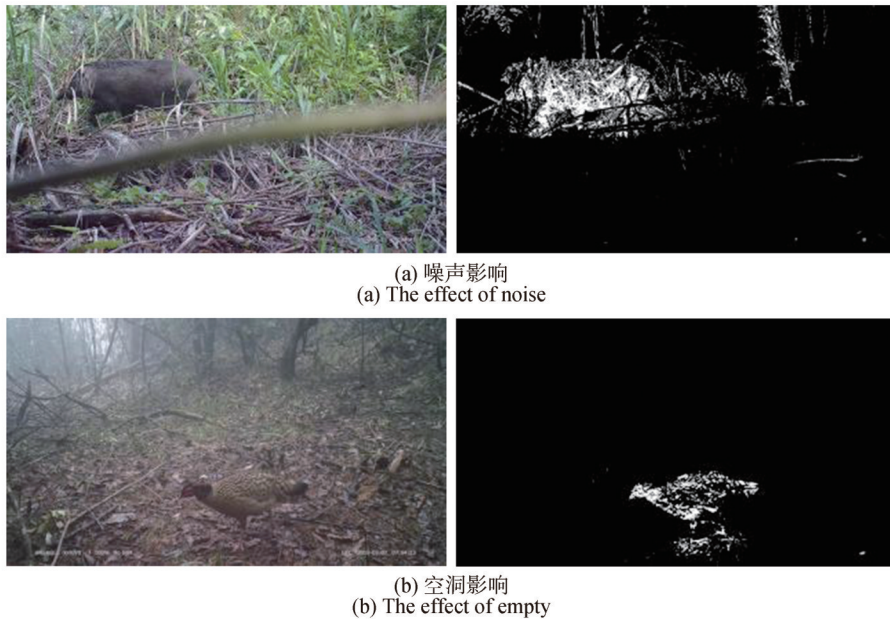


图1 背景差分示例

Figure 1 Example of background difference

为了减少背景噪声与空洞现象对标定框的影响,利用形态学的腐蚀和膨胀操作对图像去噪^[19],其处理效果如图2所示。

形态学腐蚀可以把小于结构元素的物体去除或分开,它的作用在于消除物体边界点或去除小的噪声。设 a 、 b 分别是原始图像和结构元素, $a \ominus b$ 表示 a 被 b 形态学腐蚀,其结果为所有使 b 被 x 平移后包含于 a 的点 x 的集合,定义如式(1)所示。

$$a \ominus b = \{x | (b)_x \subseteq a\}. \quad (1)$$

式中: b 为 3×3 的核。形态学腐蚀的处理效果如图

2(b)所示。之后对腐蚀后的图像进行形态学闭操作,闭操作可以平滑和磨光图像中较大物体的边界,是具有延伸性的运算,如式(2)和式(3)所示。

$$a \cdot b = (a \oplus b) \ominus b, \quad (2)$$

$$a \oplus b = \{x | (\hat{b})_x \cap a \subseteq a\}. \quad (3)$$

式中: b 为 20×20 的核, \oplus 为膨胀运算符,其形态学闭操作的处理效果如图2(c)所示。最后,通过判定图像中的最大连通区域,得到动物的边界框,并生成野生动物伪标签数据集(图3)。

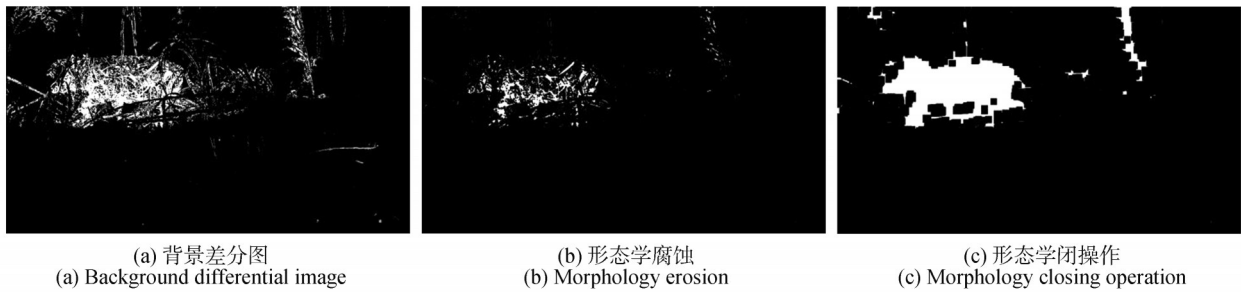


图2 形态学操作示意图

Figure 2 Schematic representation of morphology

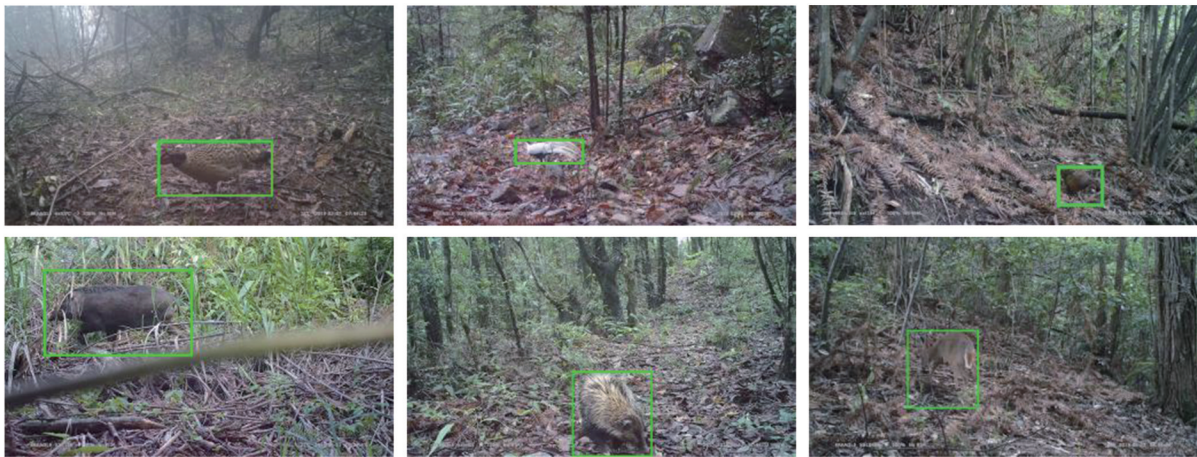


图3 自动标注伪标签示例

Figure 3 Pseudo-labels example of automatically labeled

本研究共标注6种动物的115个监测视频,获得10 915张标注图像,其中训练集包含90个视频的8 726张图像,测试集包含25个视频的2 189张图像,比例为8:2(表1)。

表1 数据集统计情况
Table 1 Statistics of the dataset

物种 Species	图像数量/张 Image count	
	训练集 Training set	测试集 Test set
黄腹角雉 <i>Tragopan caboti</i>	1 650	387
白鹇 <i>Lophura nycthemera</i>	1 219	322
灰胸竹鸡 <i>Bambusicola thoracicus</i>	1 405	367
野猪 <i>Sus scrofa</i>	1 725	380
小鹿 <i>Muntiacus reevesi</i>	1 434	463
亚洲狗獾 <i>Meles leucurus</i>	1 293	270

2 野生动物检测模型

为提高模型检测性能,提出了包含Swish激活函数的跨阶段局部卷积块(cross stage partial block,

CSP_Block)来改进YOLOv4模型中的PANet结构^[20],以下为模型的具体结构。

2.1 跨阶段局部卷积块

PANet是YOLOv4模型的特征聚合网络,该网络的密集卷积区包含4个深层卷积块,拥有较多的参数量,需要较大的计算量,是影响模型帧率的关键^[21]。因此,本研究提出了跨阶段局部卷积块用于改进PANet结构中的密集卷积区。跨阶段局部卷积块的特征提取包括两条支路:一条支路利用CBS_5卷积块特征提取,另一条支路通过卷积得到调整通道数后的特征,然后对两条支路的特征进行融合(图4(a))。

YOLOv4中不同的检测头对应不同的特征图尺寸与通道数,图4(b)展示了其中一个CBS_5的结构,其为一个包含5层CBS模块的卷积块。CBS_5结构中的第1层CBS模块用来压缩通道数,减少为最后输出需求通道数的二分之一。第2至第5层CBS模块分为2个沙漏型结构,每个沙漏型结构中的第1层用于将通道数缩减一半,第2层再进行卷积核为 3×3 的特征提取,并返回原通道数。

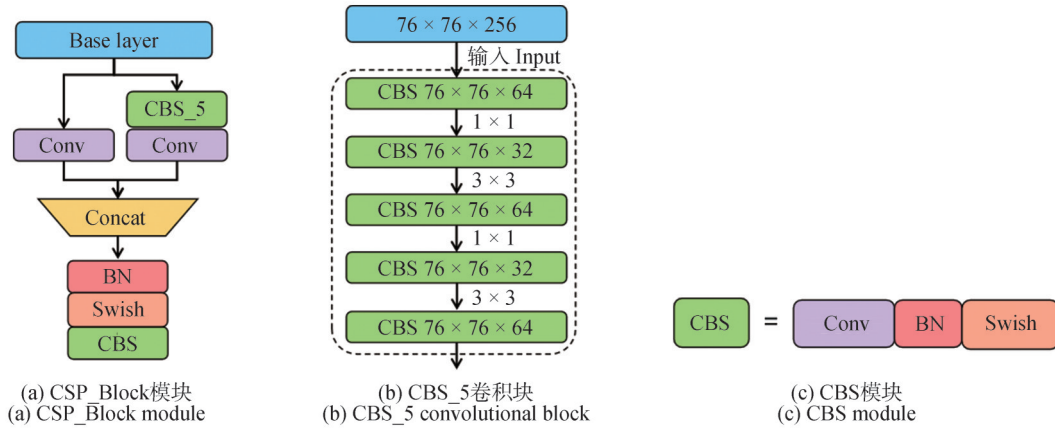


图4 跨阶段局部卷积块结构

Figure 4 Structure of CSP_Block

2.2 激活函数设计

激活函数是卷积神经网络的基本组成部分之一,用于添加到模型中非线性因素,使模型拥有更好的表达能力。本研究提出的CBS模块中包含一个卷积层,一个批量标准化层(batch normalization, BN)和一个Swish激活层,如图4(c)所示。其中,Swish激活函数替代了常用的ReLU激活函数,其函数定义如式(4)和式(5)所示。

$$\text{Swish}(x) = x \cdot \sigma(\beta x), \quad (4)$$

$$\sigma(z) = (1 + e^{-z})^{-1}. \quad (5)$$

式中: $\sigma(z)$ 表示Sigmoid函数, β 为常数。如图5所示,在ReLU激活函数中,只有当输入 $x > 0$ 时才会有输出,输入 $x < 0$ 时输出将被统一抑制为0。然而,在Swish激活函数中,不仅输入 $x > 0$ 时会有输出,在输入 $x < 0$ 时也有负的输出^[22],即模型使用ReLU激活函数时只会关注对物种识别正相关的特征,而模型使用Swish激活函数后则会同时关注对物种识别正相关和负相关的特征,并在训练过程中监督模型提取物种间更加多样性和区分性的特征。

Hard-Swish激活函数^[23]与Swish激活函数在负输入时都会有输出,其函数如式(6)和(7)所示。

$$\text{Hard-Swish}(x) = x \frac{\text{ReLU6}(x + 3)}{6}, \quad (6)$$

$$\text{ReLU6}(x) = \min(6, \max(0, x)). \quad (7)$$

但Hard-Swish激活函数与Swish激活函数相比不够平滑,不利于梯度的稳定流动。而且,Hard-Swish激活函数只在输入 $x > -3$ 时有输出,与Swish激活函数相比限制较大。

在模型的浅层区域Swish激活函数与ReLU激活函数性能没有显著性区别,而在深层区域Swish激活

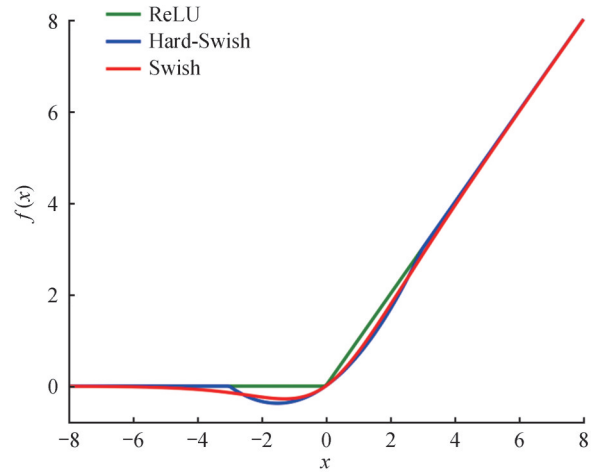


图5 ReLU、Hard-Swish和Swish激活函数曲线图

Figure 5 ReLU, Hard-Swish and Swish activation function curves

函数能表现出更高效的性能^[22]。因此,将模型深层PANet结构中的激活函数替换为Swish激活函数。

2.3 野生动物检测模型搭建

本研究通过跨阶段局部卷积块和Swish激活函数对YOLOv4网络进行改进,并以改进后的YOLOv4网络作为野生动物检测模型的基础(图6)。模型输入图像大小为 608×608 ,采用CSPDarknet53作为主干网络提取图像特征,并将3个尺度的特征输入路径聚合网络中。其中的CSP_Block将多尺度的感受区域特征进行提取和融合,并增加了梯度路径,减少了参数量和计算量。CSP_Block中的激活函数被替换为Swish激活函数,有助于提高模型提取特征的多样性和区分性。CSP_Block提取和融合后的特征图像最终被处理为 76×76 、 38×38 和 19×19 三种尺度,分别用于对大中小3种不同尺寸的动物进行检测。

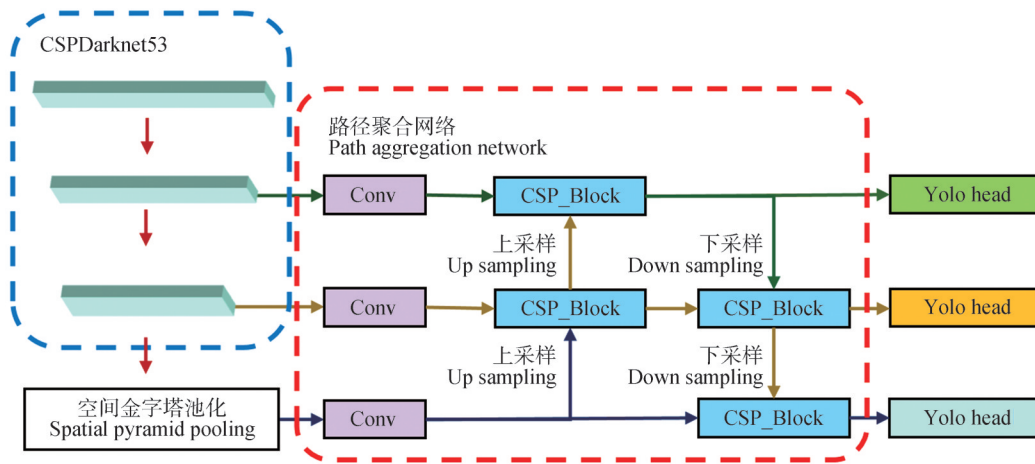


图6 野生动物检测模型结构

Figure 6 Wildlife detection model structure

3 结果与分析

3.1 实验设置

为验证本方法的检测性能,选择 YOLOv4、RFCN^[8]和 YOLOv8 三种常见的目标检测模型进行性能对比。实验中对比模型和本算法使用相同的训练集和测试集,设置相同的超参数:输入图像尺寸为 608×608 , batch_size 为 16, 初始学习率为 0.001, 每 5 个 epoch 学习率衰减为原来的 95%。硬件采用 Intel Core i7-12700KF 处理器, NVIDIA GeForce GTX 2080Ti 显卡;软件环境是 PyTorch 框架, Python 版本为 3.10, GPU 加速库采用 CUDA11.0 和 cuDNN8.3。

3.2 评价指标

评价指标为精确率 (precision, P)、召回率 (recall, R)、平均精度均值 (mean average precision, mAP) 和帧率 (frame rate)。精确率和召回率的计算公式分别如式(8)和式(9)所示。

$$P = \frac{T_p}{T_p + F_p}, \quad (8)$$

$$R = \frac{T_p}{T_p + F_N}. \quad (9)$$

式中: T_p (true positive) 为被正确识别的目标个数, F_p (false positive) 为被错误识别的目标个数, F_N (false negative) 为未被正确识别出来的目标个数。

通过精确率和召回率可计算出每个类别的平均精度 (average precision, AP), mAP 为各类平均精度的平均值, 如式(10)和式(11)所示。

$$P_{mAP} = \frac{\sum_{i=1}^C P_{AP_i}}{C}, \quad (10)$$

$$P_{AP} = \int_0^1 P(R) dR. \quad (11)$$

式中: C 为物种数目, P_{AP_i} 为第 i 个类别的平均精度。本研究中 mAP 使用的交并比 (intersection over union, IoU) 阈值均为 0.5。

3.3 模型对比实验

本方法与 YOLOv4、RFCN 和 YOLOv8 系列模型的性能对比结果如表 2 所示。本方法与 RFCN 模型相比,在 mAP 和帧率方面提高了 3.43 个百分点和 3.52 帧/s。本研究方法与 YOLOv8m 和 YOLOv8x 模型相比,帧率下降了 36.93、8.17 帧/s, mAP 提高了 11.71、7.11 个百分点。这证明本方法相比 RFCN 模型和 YOLOv8 系列模型能更好地适应野生动物监测环境。本方法与原 YOLOv4 模型相比,召回率降低 3.71 个百分点,但精确率、mAP 和帧率分别提高 1.53 个百分点、1.62 个百分点和 2.40 帧/s。可见,本方法在保证检测帧率的同时,实现了更高的检测平均精度均值。

表2 模型对比实验结果

Table 2 Comparison of model experimental results

检测模型 Detection model	精确率 / % Precision	召回率 / % Recall	平均精度 均值 / % mAP	帧率 / (帧·s ⁻¹) Frame rate
RFCN	92.79	49.75	82.98	15.41
YOLOv8m	74.60	72.73	74.70	55.86
YOLOv8x	81.18	78.05	79.30	27.10
YOLOv4	91.40	80.73	84.79	16.53
WDM	92.93	77.02	86.41	18.93

此外,图7进一步展示了6种动物的检测平均精度对比结果。本方法与原始YOLOv4相比,在白鹇、

灰胸竹鸡、野猪和小鹿上分别提高了0.53、10.89、9.34、0.84个百分点,进一步证明本算法的优越性。

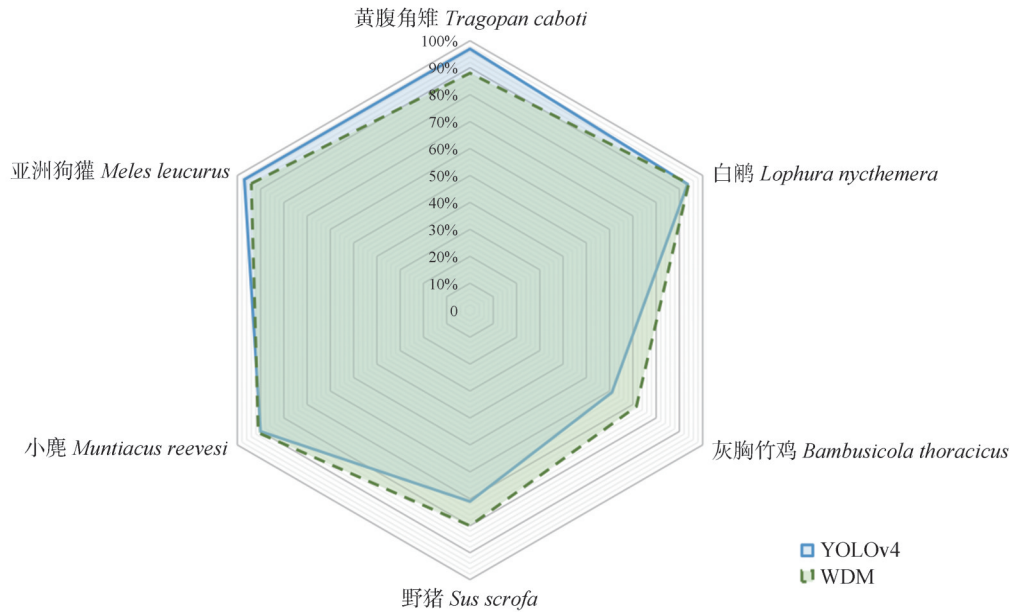


图7 每个类别的检测平均精度

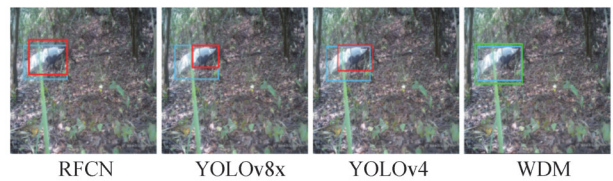
Figure 7 Detection average precision results for each category

本方法与对比算法分别在6种动物上的检测结果如图8所示,其中包括监测中可能遇到的全身、半身、小目标和遮挡等多种情况。图中蓝色框为标定框,绿

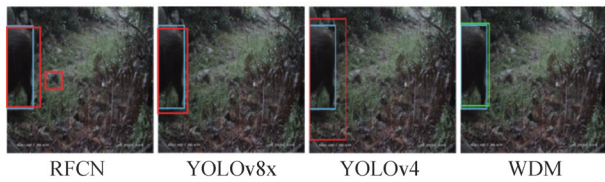
色为本方法的检测框,红色为对比算法的检测框。从图8中可以看出,本算法在各种动物监测情况下的检测都能较好地贴合标定框,验证了本方法的有效性。



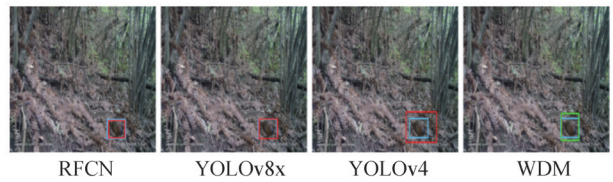
(a) 黄腹角雉
(a) *Tragopan caboti*



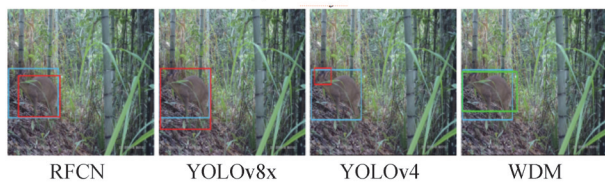
(b) 白鹇
(b) *Lophura nycthemera*



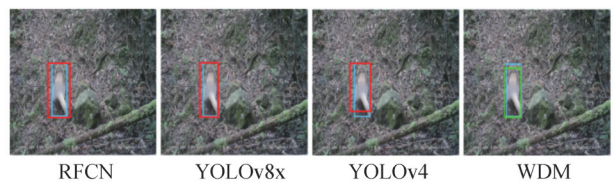
(c) 野猪
(c) *Sus scrofa*



(d) 灰胸竹鸡
(d) *Bambusicola thoracicus*



(e) 小鹿
(e) *Muntiacus reevesi*



(f) 亚洲狗獾
(f) *Meles leucurus*

□. 标定框; □. 本研究方法检测框; □. 对比算法检测框。

□. Calibration box; □. Detection result of our proposed method; □. Detection result of the comparative algorithm.

图8 不同模型对6个物种的目标检测结果对比

Figure 8 Comparison of object detection results of different models for six species

3.4 消融实验

3.4.1 跨阶段局部卷积块对比实验

为验证跨阶段局部卷积块可有效减少YOLOv4中路径聚合网络所需的计算量,将提出的野生动物检测模型与YOLOv4模型进行参数量和浮点运算数

(floating point operations, FLOPs)的对比实验(表3)。本算法与YOLOv4模型相比,在参数量方面减少了 13.84×10^6 ,在FLOPs方面减少了 12.78×10^9 ,验证了跨阶段局部卷积块的有效性。

表3 卷积块实验结果对比

Table 3 Comparison of convolution block experimental results

检测模型 Detection model	平均精度均值 /% mAP	参数量 / 10^6 Params	浮点运算数 / 10^9 FLOPs
YOLOv4	84.79	64.33	64.23
WDM	86.41	50.49	51.45

3.4.2 激活函数对比实验

为进一步验证本算法在密集卷积区所提的Swish激活函数的有效性,在本算法的基础上选择ReLU、Hard-Swish和Swish激活函数进行对比实验(表4)。本研究所提的Swish激活函数与ReLU、Hard-Swish激活函数相比,在精确率方面提高了3.22、1.07个

百分点,在mAP方面提高了4.86、7.21个百分点,可见本研究在密集卷积区域引入Swish激活函数有效提高了野生动物检测模型的特征提取能力;而且在帧率方面分别提高了2.38、1.59帧/s,表明3种激活函数对帧率的影响较小,验证了Swish激活函数拥有更好的综合性能。

表4 激活函数实验结果对比

Table 4 Comparison of activation function experimental results

激活函数 Activation function	精确率 /% Precision	召回率 /% Recall	平均精度均值 /% mAP	帧率 / (帧·s ⁻¹) Frame rate
WDM+ReLU	89.71	77.18	81.55	16.55
WDM+Hard-Swish	91.86	70.10	79.20	17.34
WDM+Swish	92.93	77.02	86.41	18.93

3.4.3 伪标签对比实验

为验证本研究所提的伪标签标记方法能够满足实际需要,以WDM为基础模型,将伪标签数据和人工标签数据分别作为训练集进行对比实验(表5)。

伪标签与人工标签相比,在mAP方面仅降低了2.29个百分点,表明伪标签与人工标签之间差距较小,验证了伪标签标记方法的有效性。

表5 伪标签实验结果对比

Table 5 Comparison of pseudo-labels experimental results

训练集 Training dataset	精确率 Precision	召回率 Recall	平均精度均值 mAP	%
人工标签 Manual labels	90.82	89.74	88.70	
伪标签 Pseudo-labels	92.93	77.02	86.41	

4 讨论

本研究以浙江江山仙霞岭省级自然保护区的6种野生动物作为检测对象,首先分析了复杂环境对伪标签标注带来的噪声和空洞影响,提出了形态

学操作的优化方案,得到了获得高质量伪标签标注方法,再从提高算法检测速度和平均精度均值的角度提出了适合于视频数据处理的野生动物检测模型,其中通过在深层密集卷积区引入跨阶段局部卷积块减少了参数量和浮点运算数,并采用Swish激活

函数优化深度卷积区域的推理过程,提高了模型的检测精度。改进后的模型,mAP为86.41%,平均精度均值提升1.62个百分点;帧率为18.93帧/s,速度提升2.40帧/s。

近年来,先进的人工智能技术已经越来越多地被研究人员改进后运用到野生动物监测图像的自动识别中,这有助于提高野生动物相关监测数据的处理效率,缩短对野生动物相关研究的科研周期^[24]。其中对YOLO系列模型的PANet部分进行改进成为常见且有效的方法。如阎婧宇等^[13]通过在YOLOv7-Tiny算法PANet部分添加高效通道注意力机制ECA-Net来提高检测精度;蒋珏泽等^[14]使用轻量化模块(Ghost)替换了YOLOv5算法PANet部分所有的卷积结构,降低了整体的计算成本;于文华等^[15]使用分组交互卷积模块(GSConv)替换YOLOXs算法PANet部分的普通卷积模块,从而完成轻量化改进。本研究采用了与这些研究相似的思路,通过使用CSP_Block模块替换了YOLOv4算法PANet部分的密集卷积结构,降低了模型的参数量,提高了检测速度。与这些研究不同的是,本研究进一步对CSP_Block模块中的激活函数进行了讨论,使用Swish激活函数替换了常用的ReLU激活函数,并在mAP和帧率指标上提高了4.86个百分点和2.38帧/s。此外,这些研究中的目标检测数据集依靠人工标注边界框,而本研究提出了基于运动追踪的伪标签标注方法节省了人工标注边界框的时间成本,并依靠形态学操作解决了噪声和空洞影响。使用伪标签数据集训练目标检测网络的mAP达到了86.41%,相比人工标注仅下降2.29个百分点,充分体现了该方法的有效性。

本研究不足之处在于目前相机采集到的数据还不够丰富,未来将扩大检测物种的种类和图像数量,并考虑通过更多数据增强的方式生成丰富的数据集。数据集中种类和数量的扩充将显著提高模型的检测效果。

5 结论

本研究以浙江江山仙霞岭省级自然保护区的视频监测数据为研究对象,提出了基于人机协作的伪标签标定方法,依靠专家知识快速生成大量目标检测训练所需的伪标签,并使用跨阶段局部卷积块对原始YOLOv4进行改进,同时在密集卷积区引入Swish激活函数,提高了模型PANet结构的特征提取

能力,从而构建了野生动物检测模型。实验结果表明,改进后的野生动物检测模型平均精度均值、帧率达到86.41%、18.93帧/s,相比原YOLOv4算法提升1.62个百分点和2.40帧/s,有效提高了浙江江山仙霞岭省级自然保护区野生动物监测的自动化和智能化水平。本研究算法适用于在野外复杂环境下,通过野生动物检测模型实现野生动物监测视频和图像的快速检测,为保护野生动物的多样性研究提供了新的途径和技术手段。

参考文献:

- [1] DELISLE Z J, HENRICH M, PALENCIA P, *et al.* Reducing bias in density estimates for unmarked populations that exhibit reactive behaviour towards camera traps [J]. *Methods in Ecology and Evolution*, 2023, 14(12): 3100-3111.
- [2] 李晟. 中国野生动物红外相机监测网络建设进展与展望 [J]. *生物多样性*, 2020, 28(9): 1045-1048.
LI S. Development progress and outlook of the wildlife camera-trapping networks in China [J]. *Biodiversity Science*, 2020, 28(9): 1045-1048.
- [3] STEEN K A, VILLA-HENRIKSEN A, THERKILDSEN O R, *et al.* Automatic detection of animals in mowing operations using thermal cameras [J]. *Sensors*, 2012, 12: 7587-7597.
- [4] 初未萌. 一种基于广义Hough变换的动物图像识别方法研究 [D]. 哈尔滨: 哈尔滨工业大学, 2013.
CHU W M. Research on animal image detection based on generalized Hough transform [D]. Harbin: Harbin Institute of Technology, 2013.
- [5] MATUSKA S, HUDEC R, KAMENCAY P, *et al.* Classification of wild animals based on SVM and local descriptors [J]. *AASRI Procedia*, 2014, 9: 25-30.
- [6] KAMENCAY P, TRNOVSZKY T, BENCO M, *et al.* Accurate wild animal recognition using PCA, LDA and LBPH [C]//2016 ELEKTRO, May 16-18, 2016. Strbske Pleso: IEEE, 2016: 62-67.
- [7] GIRSHICK R, DONAHUE J, DARRELL T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 23-28, 2014. Columbus: IEEE, 2014: 580-587.
- [8] DAI J F, LI Y, HE K M, *et al.* R-FCN: object detection via region-based fully convolutional networks [C]//LEE D D, VON LUXBURG U, GARNETT R, *et al.* NIPS'16: Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona: Curran Associates Inc., 2016: 379-387.
- [9] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: unified, real-time object detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016. Las Vegas: IEEE, 2016: 779-788.

- [10] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13–19, 2020. Seattle: IEEE, 2020: 10778–10787.
- [11] 张雪莹,张浩林,韩莹莹,等. 基于深度学习的野生动物监测与识别研究进展[J]. 野生动物学报, 2022, 43(1): 251–258. ZHANG X Y, ZHANG H L, HAN Y Y, *et al.* Research progress of the wildlife monitoring and identification based on deep learning [J]. Chinese Journal of Wildlife, 2022, 43 (1) : 251–258.
- [12] TUIA D, KELLENBERGER B, BEERY S, *et al.* Perspectives in machine learning for wildlife conservation [J]. Nature Communications, 2022, 13(1): 792.
- [13] 阎婧宇,谢永华. 基于YOLOv7-Tiny算法的无人机实时跟踪野生动物方法[J]. 野生动物学报, 2024, 45(2): 251–261. YAN J Y, XIE Y H. Real time tracking of wild animals using unmanned aerial vehicles based on YOLOv7-Tiny algorithm [J]. Chinese Journal of Wildlife, 2024, 45(2): 251–261.
- [14] 蒋珺泽,谢永华. 应用改进优鲁模型对机载热成像中野生动物种类的识别方法[J]. 东北林业大学学报, 2022, 50(3): 109–112;124. JIANG J Z, XIE Y H. Identification of wild animal species in airborne thermal imaging using improved youlu model algorithm [J]. Journal of Northeast Forestry University, 2022, 50 (3) : 109–112;124.
- [15] 于文华,谢永华. 基于轻量化YOLOX算法的东北虎检测[J]. 野生动物学报, 2023, 44(4): 760–767. YU W H, XIE Y H. Amur tiger detection based on lightweight YOLOX algorithm [J]. Chinese Journal of Wildlife, 2023, 44 (4): 760–767.
- [16] XIE J J, LI A Q, ZHANG J G, *et al.* An integrated wildlife recognition model based on multi-branch aggregation and squeeze-and-excitation network [J]. Applied Sciences, 2019, 9 (14) : 2794.
- [17] SCHNEIDER S, GREENBERG S, TAYLOR G W, *et al.* Three critical factors affecting automated image species recognition performance for camera traps [J]. Ecology and Evolution, 2020, 10 (7): 3503–3517.
- [18] SCHAUS J, UZAL A, GENTLE L K, *et al.* Application of the random encounter model in citizen science projects to monitor animal densities [J]. Remote Sensing in Ecology and Conservation, 2020, 6(4): 514–528.
- [19] 丑永新,张爱华,顾亚,等. 一种快速的数学形态学滤波方法及其在脉搏信号处理中的应用[J]. 仪器仪表学报, 2020, 41(2): 253–262. CHOU Y X, ZHANG A H, GU Y, *et al.* Fast mathematical morphological filtering method and its application on pulse signal processing [J]. Chinese Journal of Scientific Instrument, 2020, 41(2): 253–262.
- [20] BOCHKOVSKIY A, WANG C Y, LIAO H M. YOLOv4: optimal speed and accuracy of object detection [EB/OL]. arXiv (2020-04-23) [2024-08-23]. <https://arxiv.org/abs/2004.10934v1>.
- [21] LIU S, QI L, QIN H F, *et al.* Path aggregation network for instance segmentation [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18–23, 2018. Salt Lake City: IEEE, 2018: 8759–8768.
- [22] RAMACHANDRAN P, ZOPH B, LE Q V. Swish: a self-gated activation function [EB/OL]. arXiv (2017-10-16) [2024-08-23]. <https://arxiv.org/abs/1710.05941v1>.
- [23] HOWARD A, SANDLER M, CHEN B, *et al.* Searching for MobileNetV3 [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27–November 2, 2019. Seoul: IEEE, 2019: 1314–1324.
- [24] XU P F, ZHANG Y Y, JI M H, *et al.* Advanced intelligent monitoring technologies for animals: a survey [J]. Neurocomputing, 2024, 585: 127640.