

基于改进 YOLO v4 的肉鸽行为检测模型研究

郭建军^{1,2} 何国煌^{1,2} 徐龙琴^{1,3} 刘同来^{1,4} 冯大春^{1,2} 刘双印^{1,2}

(1. 仲恺农业工程学院信息科学与技术学院, 广州 510225; 2. 仲恺农业工程学院智慧农业创新研究院, 广州 510225;
3. 仲恺农业工程学院广东省农产品安全大数据工程技术研究中心, 广州 510225;
4. 仲恺农业工程学院省级现代农业(农产品质量安全溯源)产业技术研发中心, 广州 510225)

摘要: 肉鸽行为表现与鸽舍环境舒适度和肉鸽健康状况密切相关。为实现肉鸽行为精准检测、及时掌握肉鸽健康状况,提出了基于改进 YOLO v4 模型的肉鸽行为检测方法。由于肉鸽社交等行为特征相似性程度高,为了在复杂环境下准确识别肉鸽行为,本文采用自适应空间特征融合(Adaptively spatial feature fusion, ASFF)模块改进 YOLO v4 模型,在特征金字塔网络中增加 ASFF 模块,根据特征权值自适应融合多层特征,充分利用不同尺度特征信息,并且 ASFF 模块能有效过滤空间冲突信息、抑制反向梯度不一致问题、改善特征比例不变性以及降低推理开销。基于多时段的肉鸽清洁和社交行为数据集,自制 5 类肉鸽行为图像数据库,采用 OpenCV 工具进行模糊、亮度、水雾和噪声等处理扩充图像数据集(共 10 320 幅图像),增加数据多样性和模拟不同识别场景,提升模型泛化能力。本文按照比例 8:2 划分训练集和验证集,训练总共迭代 300 个周期,对不同时段、角度、尺寸的肉鸽数据集进行检测。检测结果表明,在阈值 0.50 和 0.75 时 YOLO v4-ASFF 检测精度比 YOLO v4 的 mAP_{50} 和 mAP_{75} 提高 14.73、14.97 个百分点。对比 Faster R-CNN、SSD、YOLO v3、YOLO v5 和 CenterNet 模型验证本文模型检测性能,在测试集中 mAP_{50} 分别提高 13.98、14.00、18.63、14.16、10.87 个百分点。视频检测速度为 8.1 f/s,在推理速度相当情况下,本文改进模型识别准确率更高,复杂环境泛化能力更强,且对相似度高的行为误检和漏检情况更少,可为智能化肉鸽养殖和科学管理提供技术参考。

关键词: 肉鸽行为检测; 改进 YOLO v4; 多尺度特征; 自适应空间特征融合

中图分类号: TP391 文献标识码: A 文章编号: 1000-1298(2023)04-0347-09

OSID:



Pigeon Behavior Detection Model Based on Improved YOLO v4

GUO Jianjun^{1,2} HE Guohuang^{1,2} XU Longqin^{1,3} LIU Tonglai^{1,4} FENG Dachun^{1,2} LIU Shuangyin^{1,2}

(1. College of Information Science and Technology, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China
2. Institute of Intelligent Agriculture Innovation, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China
3. Guangdong Big Data Engineering Research Center for Agricultural Product Safety, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China
4. Provincial Modern Agriculture (Agricultural Product Quality and Safety Traceability) Industrial Technology R&D Center, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China)

Abstract: Pigeon whole behavior is closely related to the loft environmental comfort and pigeon whole health. For human observation and recording the pigeon whole behavior is time-consuming, sampling limited, subjective and other issues, to timely meet the pigeon whole precision detection and pigeon whole behavior and health, based on the YOLO v4 pigeon whole behavior detection method was proposed. In this method, CSPDarkNet53 was used as the Backbone network to extract feature maps covering shallow semantic information of pigeons, and then PANet was used to transfer the bottom features and stack features to the top. Aiming at the high similarity degree of pigeon social behavior features, in order to achieve accurate identification of pigeon behavior in complex environment. The adaptively spatial

收稿日期: 2022-07-18 修回日期: 2022-08-28

基金项目: 国家自然科学基金项目(61871475)、广东省基础与应用基础研究基金项目(2022B1515120059)、广州市重点研发计划项目(202103000033)、广东省普通高校创新团队项目(2021KCXTD019)、广东省企业科技特派员项目(GDKTP2021004400)和广州市增城区农村科技特派员项目(2021B42121631)

作者简介: 郭建军(1982—),男,副教授,博士,主要从事农业信息化与自动化和机器学习研究,E-mail: glgxbaobao@163.com

通信作者: 刘双印(1977—),男,教授,博士,主要从事机器学习、人工智能和智能信息处理研究,E-mail: hdlstxylq@126.com

feature fusion (ASFF) module was adopted to improve the YOLO v4 model, and the ASFF module was added to the feature pyramid network, which can adaptively fuse multi-layer features according to the feature weights and make full use of the features information of different scales. Moreover, ASFF can effectively filter spatial conflict information and suppress reverse gradient inconsistency, improve feature proportion invariance and reduce inference overhead. Based on the cleaning and social behaviors of meat pigeons in multiple periods, a database of five kinds of meat pigeon behavior images was made. OpenCV tool was used to process blur, brightness, haze and noise to expand the image data set (totally 10 320 images), increase data diversity and simulate different recognition scenes, and improve the generalization ability of the model. A 8:2 ratio was used to divide the training and validation sets. The training iterated 300 epochs in total, and the detection was carried out through meat pigeon data sets of different time periods, angles and sizes. The detection results showed that the detection accuracy of improved YOLO v4 - ASFF model was 14.73 percentage points and 14.97 percentage points higher than that of mAP₅₀ and mAP₇₅ of original YOLO v4 model at the threshold of 0.50 and 0.75. Compared with Faster R - CNN, SSD, YOLO v3, YOLO v5 and CenterNet model, mAP₅₀ of the YOLO v4 - ASFF was improved by 13.98 percentage points, 14.00 percentage points, 18.63 percentage points, 14.16 percentage points and 10.87 percentage points in test set, respectively. The video detection speed was 8.1 f/s, and the improved model had higher recognition accuracy under the condition of the same inference speed, strong generalization ability in complex environment, and less misdetection and omission of behaviors with high similarity. The research on meat pigeon behavior detection can provide technical reference for intelligent meat pigeon breeding and scientific management.

Key words: pigeon behavior detection; improved YOLO v4; multi-scale characteristics; adaptively spatial feature fusion

0 引言

我国肉鸽养殖产量和消费位居世界之首,2020年全国肉鸽存栏量达 5.3×10^7 对,其中广东省饲养量和消费量位居全国前列,肉鸽产业对产业结构调整,增加农民收入,提供优质蛋白发挥了重要作用^[1]。肉鸽行为是反映肉鸽养殖环境和健康状态关键因素之一,通过肉鸽行为及时精准调控环境和饲喂,确保肉鸽健康养殖,提高养殖效益;但在传统肉鸽养殖中,人工观测和记录肉鸽行为,存在因专业知识匮乏和养殖经验贫乏导致肉鸽行为漏检和误判,且存在采样有限、效率低下和检测不及时等问题,容易错失肉鸽异常行为检测时机,对肉鸽养殖造成危害。因此,为减少人工因素影响、降低人力成本、预防肉鸽疾病、掌控肉鸽健康状况,对肉鸽行为进行精准检测研究,有利于准确把握肉鸽健康状况,及时防治肉鸽病害,提高肉鸽产能,为科学管理肉鸽养殖提供依据^[1-5]。

传统肉鸽养殖和行为研究主要通过人工观测和记录^[3-5],研究取得一定成效,但存在耗时费力、效率低、周期长、工作量繁重、主观因素影响、采样有限等问题,难以满足健康养殖中对肉鸽行为精准检测的要求。而基于计算机视觉技术的行为目标检测具有非接触、实时、不间断、精确、高效、自动检测等优势,能解决传统方法存在的问题。目标检测是计算机视觉基本任务^[6-9],可实现肉鸽行为检测、辅助人

力提高检测效率、及时掌握肉鸽健康数据等。

近年来,国内外学者采用计算机视觉方法对动物行为检测和健康状态判断进行大量研究,并在鸚鵡、蛋鸡、野鸡等动物行为检测方面取得了进展^[10-15]。汤一平等^[10]采用计算机视觉识别鸚鵡形态和跳动行为,分析虎皮鸚鵡行为与地震先兆数据,探究鸚鵡异常行为与地震先兆之间内在关系。劳凤丹等^[11]采用深度图像分析蛋鸡群体分布、活跃度和群体中个体的采食、躺、站和坐行为,蛋鸡喜欢分布于容易采食和饮水区域,根据蛋鸡行为状态监控不同时段分布和活跃程度,便于生产者更好对蛋鸡生产进行分析和管理工作。WANG等^[12]采用YOLO v3识别和目标定位自制6类蛋鸡行为,通过迁移学习提高模型检测精度和适用性,有利于提高蛋鸡养殖福利水平。王凯等^[13]针对蛋鸡养殖行为检测操作复杂、分类单一和实时性差等问题,提出TD - YOLO v3网络结构,实现轻量级网络实时自动检测笼养蛋鸡行为,为蛋鸡养殖行为检测提供参考。方鹏等^[14]采用Faster R - CNN框架提取蛋鸡行为特征,通过多个教师网络知识蒸馏训练结构简单的ResNet34学生网络识别蛋鸡采食和休息行为,实现行为识别检测有利于评估动物福利和监测健康状态。易诗等^[15]提出了一种增强型目标检测网络(Enhanced tiny - YOLO, ET - YOLO),增大特征提取网络深度和检测尺度以提高目标检测精度,实现野外复杂环境野鸡快速识别。上述行为检测方法取得了一定成

效,但对复杂环境中相似程度高的肉鸽行为图像适用性有待进一步研究。

YOLO v4 是单阶段检测算法,在 YOLO 算法基础上优化了数据处理、主干网络、激活函数等,均衡了速度和检测性能,使得 YOLO 算法在众多领域取得广泛应用^[16-18]。但目前将目标检测方法应用到肉鸽养殖中进行行为检测的研究较少。

由于肉鸽养殖环境复杂,相比梳理和摄食行为,采集的肉鸽行为数据中亲吻和互动等行为出现频率较少,导致数据不均衡,且社交行为中互动、亲吻和互相梳理羽毛等行为特征相似性高。本文基于 YOLO v4 模型采用数据增强策略使数据均衡,结构上采用自适应空间特征融合 (Adaptively spatial feature fusion, ASFF) 模块有效融合行为特征,以提高模型识别准确率,减少人力物力消耗,增加肉鸽养殖福利,为肉鸽行为研究提供参考^[19]。

1 数据采集与处理

1.1 数据采集

本文肉鸽数据拍摄于广东省梅州市兴宁市肉鸽养殖基地。数据采集设备使用海康威视 DS-2CD3T47EDWD-L (4mm) 型摄像机,分辨率为 2 560 像素×1 440 像素,帧率为 25 f/s,对实际养殖场景下 4 对米玛斯肉鸽进行拍摄跟踪,视频采集时段为 2021 年 8—12 月,图 1 为现场设备图,单个摄像头拍摄的视野涵盖 12 个鸽笼,基于视角和观察,本次主要采用正对摄像机的 4 个鸽笼作为实验数据,图 2 为设备示意图。



图 1 现场设备图
Fig. 1 Site equipment diagram

采集到的视频数据中,包含肉鸽繁殖、孵化、哺育等不同时段乳鸽和成鸽状态^[20]。本文关注的肉鸽养殖阶段 5 种行为分别为自梳理羽毛、互梳理羽毛、展翅、亲吻和互动行为,图 3 是对应行为不同形态,肉鸽行为定义如表 1 所示^[4]。由图 3 可以看出,肉鸽梳理翅膀状态时,羽毛张开幅度大,自梳理羽毛与展翅行为相似;肉鸽亲吻阶段会存在互动行为和互梳理羽毛行为,亲吻行为和互动行为形态较为接

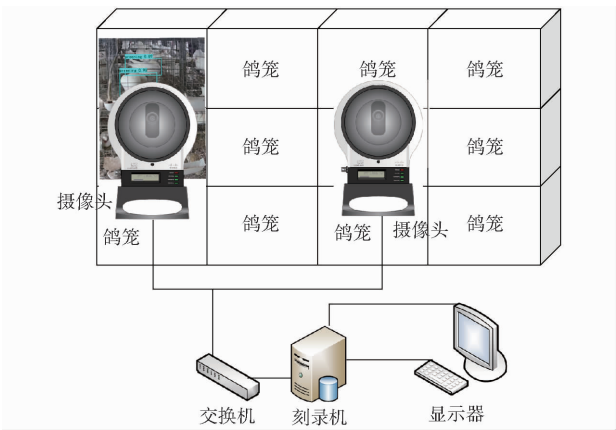


图 2 设备示意图
Fig. 2 Equipment diagram



图 3 各类行为图
Fig. 3 Behavior graphs

表 1 肉鸽行为描述

Tab. 1 Behavior description of meat pigeon	
行为	描述
自梳理羽毛	用喙触碰自身羽毛,站立或者趴卧状态
互梳理羽毛	用喙触碰其他肉鸽羽毛,站立或趴卧状态
展翅	翅膀单侧或双侧张开,站立、趴卧或者悬空
亲吻	肉鸽互相触碰对方的喙
互动	肉鸽之间头部相互触碰,喙未触碰对方

近,且养殖场所环境复杂,进一步增大了行为的检测难度。

1.2 数据处理

本文主要针对正对视场角的 4 个鸽笼进行实验,使用 OpenCV 按 25 f/s 进行视频截取,取得的 JPG 格式图像通过筛选形成肉鸽行为图像数据集。为增加数据多样性,随机裁剪上下左右鸽笼形成原始图像 3 120 幅。为了避免训练过程出现过拟合现象,随机抽选 624 幅组成测试集和 1 800 幅图像按 8:2 划分为训练集和验证集训练和验证,利用数据增强技术对样本进行水雾、随机翻转、增加噪声、模糊处理等,扩充数据后共 10 320 幅。根据 Pascal VOC 数据集格式,采用开源工具 LabelImg 进行人工标注,形成 xml 格式数据集文件。数据增强效果如图 4 所示。

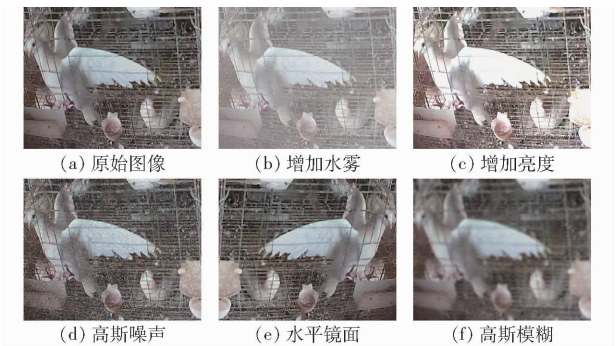


图 4 数据增强效果图

Fig.4 Data enhancement rendering

2 研究方法

本文提出一种改进 YOLO v4 的肉鸽行为检测模型。先对养殖环境中的肉鸽数据进行采集和处理,并定义和自制 5 类肉鸽行为数据集;该模型在行

为形态相近和复杂环境问题中,通过加入自适应空间特征融合模块结构进一步融合底层和高层语义特征,并有效过滤空间冲突信息,抑制反向梯度不一致问题,从而提高模型检测效果。

2.1 YOLO v4 模型

在 YOLO v4 目标检测模型中,主要包括 5 部分:数据输入、骨干网络 (CSPDarkNet^[21])、Neck 网络 (SPP^[22]、FPN^[23] 和 PAN^[24])、检测头 (Head^[25]) 和结果输出,原始网络结构如图 5 所示。输入部分主要是训练阶段数据和验证集数据;骨干网络主要是通过卷积神经网络进行特征提取;Neck 部分主要是对特征进行融合,处理不同尺度特征以便融合特征;检测头 Head 对特征进行分类预测和边框检测,从而得到模型预测输出结果。

2.2 骨干网络

YOLO v4 使用的骨干网络是 CSPDarkNet53^[21], CSPDarkNet53 是在 CSPNet 基础上进行改进,改进结构解决了密集块和过渡层的信息流、处理速度低和内存占用率高等问题,从而加快模型推理速度。CSPDarkNet53 的优点包括增加卷积神经网络学习能力,保证模型准确性的同时进行轻量化,降低计算瓶颈,减小内存使用成本,减小梯度消失以及训练中反向传播路径优化等^[21]。

2.3 颈部网络

颈部网络由空间金字塔池化网络 (Spatial pyramid pooling, SPP)^[22]、特征金字塔网络 (Feature pyramid networks, FPN)^[23] 和路径聚合网络 (Path aggregation network, PAN)^[24] 组成。SPPNet 结构由

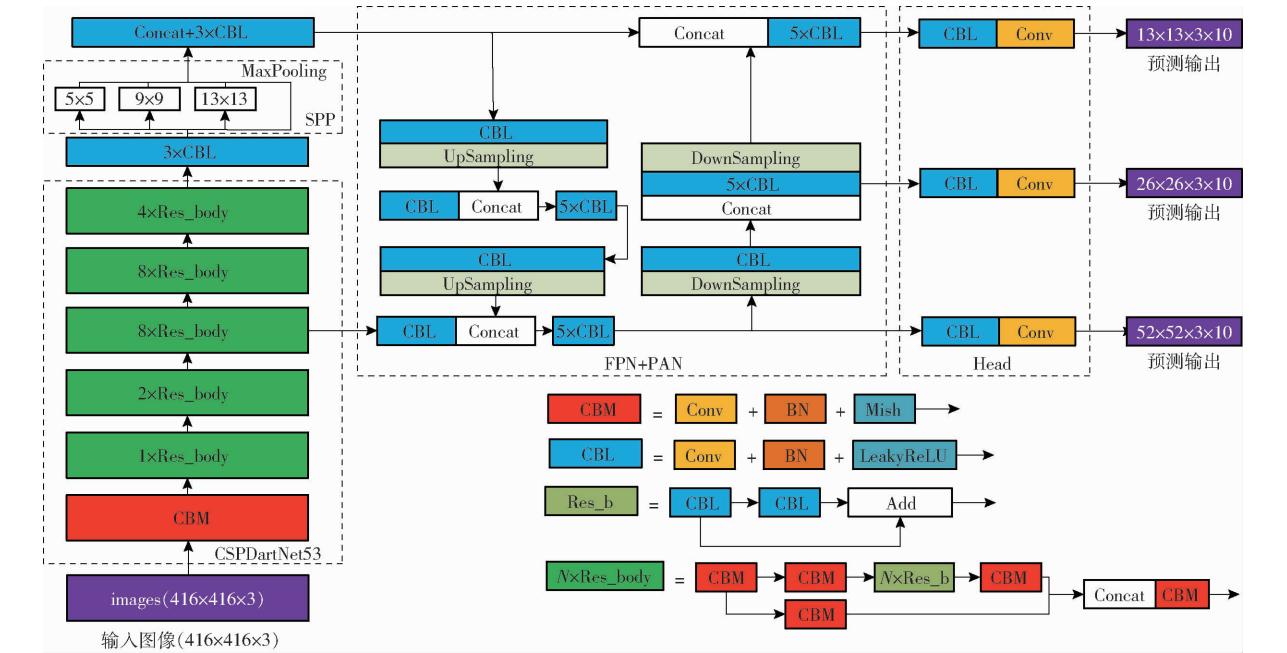


图 5 原始网络结构

Fig.5 Original network structure

3 个卷积和 4 个不同尺寸 (1 × 1、5 × 5、9 × 9、13 × 13) 最大池化层组成,组合成多尺度结构,使用不同大小特征图进行融合增大网络感受野,模块输出是 FPN 网络的一个分支。FPN 主要是在主干网特征提取中连接不同的层进行融合,形成自顶向下的上采样金字塔结构;PANet 根据 FPN 设计思想,包含上、下采样的特征金字塔结构,避免信息丢失,进一步提高底层和高层特征的重用性。

2.4 损失函数

为了使目标模型实现高效检测,使用损失函数对比预测值和真实值之间误差,使误差不断减少,逐渐提升模型准确度。本文采用原始模型损失函数 CIoU (Compleat IoU)。CIoU 在 IoU (重叠度) 基础上进行改进,IoU 损失值计算的是预测框和真实框之间交集和并集比值,CIoU 的损失改进有:①IoU 在边界框非重叠情况下不会存在滑动梯度,基于 IoU 增加了惩罚系数。②采用最小化两个检测框中心点标准化距离,解决损失收敛慢的问题。③考虑检测框重叠面积、中心点距离以及长宽比因素,加入更有效的惩罚项加速逼近真实框^[25]。

2.5 自适应空间特征融合

目标检测模型 YOLO v4^[26]、YOLO v3^[27] 和 RetinaNet^[28] 等网络均采用直接衔接或者相加方式输出多层特征,高层包含低层的特征,直接衔接或者相加不能充分利用不同尺度特征,且会增加特征冗余以及噪声信息干扰,本文采用一种自适应空间特征融合的方法,如图 6 中 ASFF 部分所示。自适应空间特征融合模块主要是为了进行特征缩放和根据权值进行自适应融合,充分利用底层特征、细粒度特征和高层特征语义信息。ASFF 模块的特征图

level1 ~ level3 通过 1 × 1 卷积得到权重参数 α 、 β 和 γ ,Concat 后通过 softmax 函数使 $\alpha + \beta + \gamma = 1$;每层 ASFF 通过每个分支层特性信息分别乘以权重参数再求和。过程计算公式为

$$y_{ij}^l = \alpha_{ij}^l X_{ij}^{1 \rightarrow l} + \beta_{ij}^l X_{ij}^{2 \rightarrow l} + \gamma_{ij}^l X_{ij}^{3 \rightarrow l} \tag{1}$$

y_{ij}^l 为新的特征图, $X_{ij}^{k \rightarrow l}$ 为第 1 层到第 k 层的特征向量, α_{ij}^l 、 β_{ij}^l 和 γ_{ij}^l 为不同特征层的权值。不同尺寸目标具有不同权重值,模型由训练网络学到不同位置的权值,决定特征图中各像素点的激活和抑制,从而在空间上过滤冲突信息和抑制梯度反传不一致性问题,改善特征比例不变性,并降低推理开销^[29]。

2.6 模型改进

本文算法基于 YOLO v4 模型进行改进,在颈部 PANet 结构输出中加入自适应空间特征融合模块,增加特征融合有效性,提升肉鸽模型的检测性能,总体模型结构如图 6 所示。ASFF 主要由特征缩放和自适应融合组成。YOLO v4 中 PANet 层输出尺度大小、通道数不一,特征缩放是不同层之间缩放到适应尺寸;特征自适应融合根据加权值进行融合。同一个位置上不同输出层均会包含同一物体,在进行梯度计算时,某层被分配正样本时,其他两层同一位置会被视为背景,这种不一致性会干扰梯度降低训练效率,ASFF 采用其它特征层相应位置视为忽略区域,即设置梯度为 0,从而消除冲突,增强不同特征层融合效果。

3 实验

3.1 实验环境

实验训练平台使用 Ubuntu 16.04 LTS 操作系统,Intel(R) Xeon(R) Gold 6146 CPU @ 3.20 GHz、

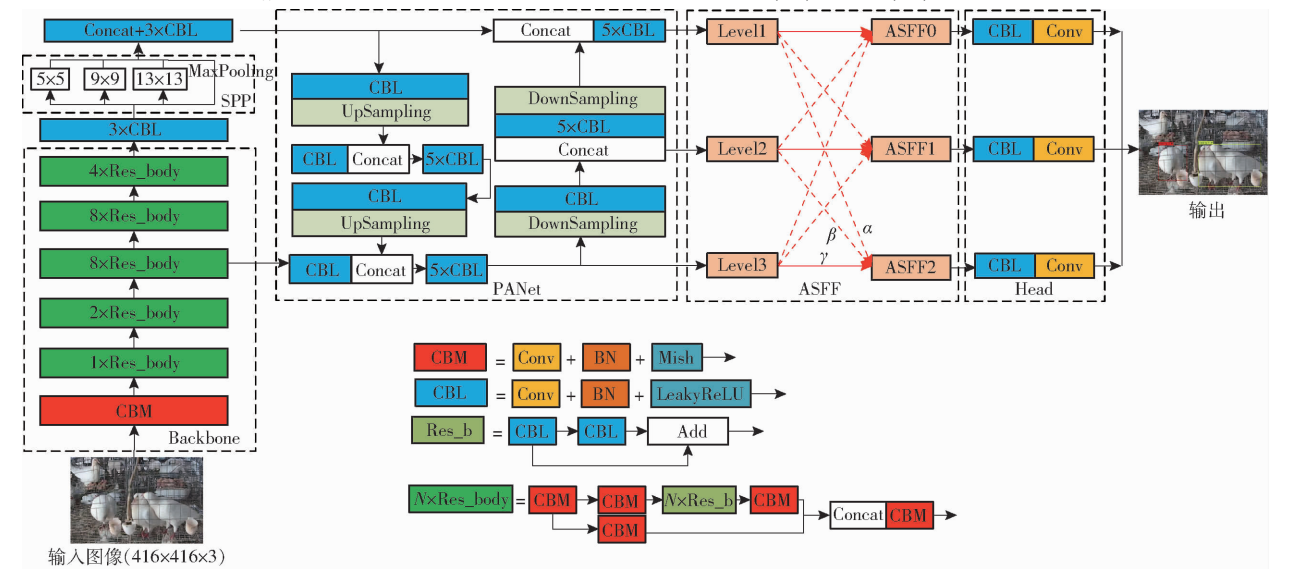


图 6 改进模型网络结构
Fig. 6 Improved model network structure

128 GB 内存、8 TB 硬盘、3 个 NVIDIA TITAN RTX 显卡,CUDA 为 11.1,加速库 Cudnn 为 8.2 版本,采用开源深度学习框架 Pytorch 和 Python 3.7 实现。

3.2 参数设置

实验根据比例 8:2 划分训练集和验证集,训练集 7 757 幅图像、验证集 1 939 幅图像、测试集 624 幅图像,用于模型训练、验证和测试。模型输入尺寸为 416 像素×416 像素,批量大小 (batchsize) 为 16,模型总共迭代 300 个周期,学习率为 0.001,动量设置为 0.92,非极大值阈值为 0.6,采用 Adam 作为优化器,其他参数采用 YOLO v4 模型默认设置。肉鸽行为数据集检测模型在 NVIDIA TITAN RTX 上进行训练和测试。

3.3 改进模型有效性

在训练过程中,为了增加模型训练性能和检测效果,采用 COCO 数据集预训练的 Backbone 网络权重进行迁移学习训练,其中 1~150 个周期采用冻结骨干网络训练方式,150~300 个周期采用解冻骨干网络进行训练。主干网络作为目标检测底层特征提取部分,物体底层特征相似性高,冻结底层权重网络能够加快网络训练速度,并且防止训练初期权重被破坏,有效防止训练过程中模型过拟合,训练过程中对验证集进行验证。图 7 为两种模型的训练集和验证集损失值。因为前几个迭代周期损失值较大,图中显示第 5~300 个周期数据。由图 7 可知,随着迭代周期的增大,损失值均在下降,模型能充分有效学习;在第 150 个周期解冻骨干网络训练,损失值进一步下降,证明采用预训练权重训练的有效性。模型训练集和验证集损失值曲线均未出现过拟合,但在相同情况下,原始模型损失值曲线均位于上方,网络收敛较慢,模型拟合数据较弱,收敛效果较差,训练集和验证集损失值较大,达到 3.23 和 3.05,而改进算法在迭代周期为 30 左右时损失值达到 4 左右,最终稳定在 2 左右,表明改进模型特征融合较优,在相同情况下模型收敛速度和训练效果较佳。

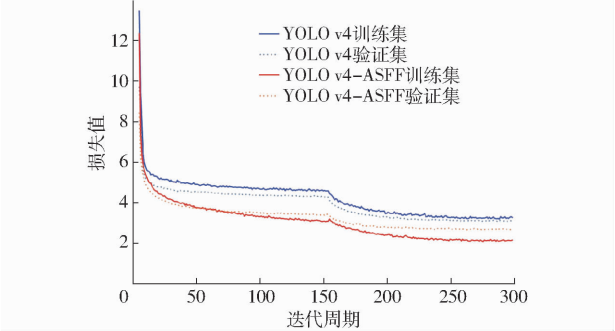


图 7 模型训练集和验证集损失值

Fig. 7 Model training and verification of loss values

本文采用 AP (平均精度)、mAP (平均精度均值)、检测速度作为评价指标。为验证 ASFF 模块特征融合有效性,设置模型训练参数一致,分别在模型迭代周期为 45、105、150、180、255 和 300 时进行实验,计算 mAP₅₀ (阈值为 0.50 时的平均精度均值) 和 mAP₇₅ (阈值为 0.75 时的平均精度均值) 结果如表 2 所示。

表 2 模型不同置信度阈值的平均精度均值
Tab. 2 Test results of different thresholds for model

平均精 度均值	模型	迭代周期					
		45	105	150	180	255	300
mAP ₅₀	YOLO v4	64.09	67.01	66.84	75.94	76.75	77.15
	YOLO v4-ASFF	52.87	63.62	64.53	84.17	91.66	91.88
mAP ₇₅	YOLO v4	33.51	35.57	36.32	48.79	47.57	50.05
	YOLO v4-ASFF	21.39	27.62	28.20	49.60	62.57	65.02

由表 2 可得,随着迭代周期的增加,在测试集检测中其置信度阈值为 0.50 和 0.75 的 mAP 值均不断提升,前 150 个迭代周期中 YOLO v4-ASFF 模型的 mAP 小于 YOLO v4 模型,但在迭代周期为 180、255、300 时改进模型 mAP₅₀ 和 mAP₇₅ 比原始模型高 8.23、14.91、14.73 个百分点和 0.81、15、14.97 个百分点,证明 150 个迭代周期后解冻主干网进行训练能够增加模型底层特征提取,更好地融合底层和高层语义特征,进一步增强检测效果,模型能更好拟合数据特征。由图 7 和表 2 可得,YOLO v4 和 YOLO v4-ASFF 随着模型迭代训练,模型在 250 个迭代周期后趋于平缓,检测精度提升幅度较小,为进一步验证模型最佳效果,采用迭代训练权重进一步进行测试集测试,结果如表 3 所示。由表 4 可知,迭代周期为 290 时 YOLO v4-ASFF 模型比原始模型在互梳理羽毛、亲吻、互动、自梳理羽毛和展翅方面 AP 分别提升 10.78、56.63、2.82、1.11、3.58 个百分点,而迭代周期为 300 时各类别分别提高 11.02、55.4、2.7、1.12、3.38 个百分点,迭代周期为 290 和 300 时改进模型比原始模型在置信度阈值为 0.50 时 mAP 提升 14.99、14.73 个百分点,置信度阈值为 0.75 时分别提升 16.84、14.97 个百分点。由表 3、4 可得,迭代周期为 290 时检测效果较佳。YOLO v4 模型的亲吻行业检测精度总体上处于最低值 (22.22%),因为亲吻类别数据与互梳理羽毛和互动类别特征相似性较高,且亲吻类别数据量相对少于其他两类。改进模型在各类别检测模型中均能有效提升检测精度。在特征较为相近时,原始模型较难提升类别准确性,YOLO v4-ASFF 能有效根据权值融合特征差异性,提升各类别识别准确率。

表 3 不同迭代周期的 mAP₅₀ 测试结果

Tab. 3 mAP ₅₀ at different epochs								%
迭代周期	270	272	274	276	278	280	282	284
mAP ₅₀	91. 289	91. 231	91. 264	91. 469	91. 273	91. 031	91. 189	90. 990
迭代周期	286	288	290	292	294	296	298	300
mAP ₅₀	90. 572	91. 451	91. 969	91. 137	90. 956	90. 866	91. 572	91. 884

表 4 各类别不同阈值和迭代周期测试结果

Tab. 4 Test results of different thresholds and iterations for different categories								
模型	迭代周期	置信度阈值	AP/%					mAP/%
			互梳理羽毛	亲吻	互动	自梳理羽毛	展翅	
YOLO v4	290	0. 50	81. 71	22. 22	89. 65	96. 70	94. 63	76. 98
YOLO v4 – ASFF			92. 49	78. 85	92. 47	97. 81	98. 21	91. 97
YOLO v4	290	0. 75	55. 08	22. 22	30. 99	78. 59	54. 10	48. 20
YOLO v4 – ASFF			68. 72	72. 29	51. 01	74. 09	59. 08	65. 04
YOLO v4	300	0. 50	82. 53	22. 22	89. 58	96. 68	94. 76	77. 15
YOLO v4 – ASFF			93. 55	77. 62	92. 28	97. 80	98. 14	91. 88
YOLO v4	300	0. 75	58. 77	22. 22	38. 35	75. 72	55. 18	50. 05
YOLO v4 – ASFF			69. 84	75. 44	49. 63	73. 16	57. 04	65. 02

3. 4 模型对比

为了对比本文模型与 SSD、Faster R – CNN、CenterNet、YOLO v3 和 YOLO v4 目标检测模型的检测效果。对比结果见表 5。由表 5 可得,在测试集中 YOLO v4 – ASFF 比 YOLO v4、Faster R – CNN、SSD、YOLO v3、YOLO v5 和 CenterNet 模型的 mAP₅₀ 分别高 14. 82、13. 98、14. 00、18. 63、14. 16、10. 87 个百分点。网络输出权值中 CSPDarkNet53 比其他模型内存占用量大,采用视频对模型帧速率进行测试,本文模型结果与 YOLO v4 相差不大;Faster R – CNN 的帧速率为 5. 27 f/s,为最低值,因为是二阶段算法,误检测较多导致 mAP 值不高;YOLO v4 – ASFF 与 YOLO v4 模型相比,在检测结果上对标签类别预测准确性更高。

表 5 不同模型对比

Tab. 5 Comparison of different models					
模型	骨干网络	内存占 用量/MB	mAP ₅₀ / %	mAP ₇₅ / %	帧速率/ (f·s ⁻¹)
SSD	VGG	92. 6	77. 97	49. 47	7. 59
Faster R – CNN	ResNet50	108. 3	77. 99	60. 38	5. 27
CenterNet	ResNet50	124. 9	81. 10	61. 41	6. 79
YOLO v3	DarkNet53	235. 1	73. 34	44. 60	7. 85
YOLO v4	CSPDarkNet53	244. 4	77. 15	50. 05	8. 57
YOLO v4 – ASFF	CSPDarkNet53	265. 3	91. 97	65. 02	8. 10
YOLO v5	CSPDarkNet	178. 4	77. 81	64. 25	8. 99

为验证检测模型泛化能力和效果,采用不同位置、角度、类别数量以及尺度不一的图像数据进行测试,不同模型部分测试结果如图 8 所示。预测图中左上角为标签类别和预测该类别的预测置信度,不

同颜色代表不同类别识别效果。

由图 8a、8f、8i 可知,SSD、YOLO v3、YOLO v4 和 Faster R – CNN 存在对互动行为漏检测;图 8c 中 SSD 模型对两只鸽子自梳理羽毛行为存在漏检测情况;图 8e 中亲吻行为与互梳理羽毛和互动行为很相似,本文模型误认为互动行为,而 CenterNet 误检测为互梳理羽毛和互动行为,YOLO v5 和 Faster R – CNN 误识别为互梳理羽毛行为,但其他模型均存在漏检测情况;当情景更加复杂时,图 8f 互动和自梳理羽毛行为只有 SSD 模型存在漏检测;图 8g 中存在展翅和自梳理羽毛行为,SSD 和 YOLO v3 存在漏检测;图 8h 中 SSD 漏检测互梳理羽毛行为,CenterNet 出现误检测互动行为。除了 YOLO v5 和本文模型对亲吻行为检测正确,其他模型均误检测为互梳理羽毛行为或者漏检测情况,且对互动和自梳理羽毛行为存在漏检测情况,相比其他检测模型,本文模型有更强的特征区分能力,在复杂环境下对肉鸽行为的检测准确性较高。

4 结论

(1) 针对肉鸽清洁行为和社交行为特点,基于 YOLO v4 采用自适应空间特征融合模块进行改进,通过自制肉鸽行为数据集进行实验对比,本文模型在不同位置、尺度情况下均能有效检测肉鸽行为,检测效果优于 SSD、Faster R – CNN、CenterNet、YOLO v3 和 YOLO v4 模型。

(2) 采用 ASFF 模块置于 YOLO v4 模型的 PANet 和 Head 之间,在多尺度特征融合之后不同分支上,根据不同物体特征权重进行融合,在空间上过



图 8 模型检测对比

Fig.8 Model detection comparisons

滤冲突信息和抑制梯度反传不一致,进而改善特征比例不变性,有效融合特征。

(3)对比 YOLO v4 模型和 YOLO v4 - ASFF,在计算量和速度相当情况下,本文模型的 mAP_{50} 和 mAP_{75} 提高 14.73、14.97 个百分点,说明本文模型检测准确性更高。

(4)在准确性和泛化能力上,本文模型在测试集中 mAP_{50} 达 91.97%,分别高于 YOLO v4、Faster R - CNN、SSD、YOLO v3、YOLO v5 和 CenterNet 模型 14.82、13.98、14.00、18.63、14.16、10.87 个百分点。在类别相似性更高的复杂环境中,误检测和漏检测更少,有利于复杂环境中肉鸽行为的检测。

参 考 文 献

[1] 周凤珍,郭妍. 不忘初心,36 年坚守鸽业第一线——访广东省肉鸽行业协会会长陈益填研究员[J]. 广东饲料,2020,29(8):8-11.
ZHOU Fengzhen, GUO Yan. Stay true to the original aspiration and stick to the front line of pigeon Industry for 36 years—interview with Researcher Chen Yitan, president of Guangdong Pigeon Industry Association[J]. Guangdong Feed,2020,29(8):8-11. (in Chinese)

[2] 刘旭芳,李华. 中国肉鸽产业发展现状与前景[J]. 农业展望,2021,17(6):57-60.
LIU Xufang, LI Hua. Development status and prospect of meat pigeon industry in China[J]. Agricultural Outlook, 2021, 17(6):57-60. (in Chinese)

[3] 秦卫红,达剑森. 种鸽行为学观察及其对繁殖性能的影响[J]. 中国家禽,2018,40(13):66-68.
QIN Weihong, DA Jiansen. Breeding behavior observation and its effects on reproductive performance[J]. Chinese Poultry, 2018, 40(13):66-68. (in Chinese)

[4] 正旭城,杨海明,张弛,等. 不同繁殖性能的种鸽行为学比较分析[J]. 畜牧与兽医,2020,52(10):21-23.
ZHENG Xucheng, YANG Haiming, ZHANG Chi, et al. Comparative analysis of breeding behavior of pigeons with different breeding performance[J]. Animal Husbandry and Veterinary Medicine,2020,52(10):21-23. (in Chinese)

[5] 程柏丛,王梁,刘康,等. 优化颗粒保健砂对自由采食肉鸽产蛋性能的影响[J]. 饲料研究,2022,45(4):44-47.
CHENG Baicong, WANG Liang, LIU Kang, et al. Effects of optimized granular sand on laying performance of free-feeding pigeons[J]. Feed Research,2022,45(4):44-47. (in Chinese)

[6] 高陈强,陈旭. 基于深度学习的行为检测方法综述[J]. 重庆邮电大学学报(自然科学版),2020,32(6):991-1002.
GAO Chenqiang, CHEN Xu. Review of behavior detection methods based on deep learning[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition),2020,32(6):991-1002. (in Chinese)

[7] 王少华,何东健,刘冬. 基于机器视觉的奶牛发情行为自动识别方法[J]. 农业机械学报,2020,51(4):241-249.
WANG Shaohua, HE Dongjian, LIU Dong. Automatic recognition of estrus behavior of dairy cows based on machine vision[J].

- Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(4): 241 – 249. (in Chinese)
- [8] 李柯泉, 陈燕, 刘佳晨, 等. 基于深度学习的目标检测算法综述[J]. 计算机工程, 2022, 48(7): 1 – 12.
LI Kequan, CHEN Yan, LIU Jiachen, et al. A review of object detection algorithms based on deep learning[J]. Computer Engineering, 2022, 48(7): 1 – 12. (in Chinese)
- [9] 许德刚, 王露, 李凡. 深度学习的典型目标检测算法研究综述[J]. 计算机工程与应用, 2021, 57(8): 10 – 25.
XU Degang, WANG Lu, LI Fan. A review of typical object detection algorithms in deep learning[J]. Computer Engineering and Applications, 2021, 57(8): 10 – 25. (in Chinese)
- [10] 汤一平, 杨昭, 石兴民, 等. 基于计算机视觉的鹦鹉行为分析关键技术的研究[J]. 小型微型计算机系统, 2016, 37(4): 841 – 846.
TANG Yiping, YANG Zhao, SHI Xingmin, et al. Research on key techniques of parrot behavior analysis based on computer vision[J]. Journal of Small Microcomputer Systems, 2016, 37(4): 841 – 846. (in Chinese)
- [11] 劳凤丹, 杜晓冬, 滕光辉. 基于深度图像的蛋鸡行为识别方法[J]. 农业机械学报, 2017, 48(1): 155 – 162.
LAO Fengdan, DU Xiaodong, TENG Guanghui. Behavior recognition method of laying hens based on depth image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(1): 155 – 162. (in Chinese)
- [12] WANG J, WANG N, LI L, et al. Real-time behavior detection and judgment of egg breeders based on YOLO v3[J]. Neural Computing and Applications, 2020, 32: 5471 – 5481.
- [13] 王凯, 姜昊昊, 吕艳, 等. 基于深度学习的笼养蛋鸡行为实时检测方法[J]. 中国农业大学学报, 2019, 24(11): 123 – 133.
WANG Kai, JIANG Wuhao, LÜ Yan, et al. Real-time behavior detection method of cage laying hens based on deep learning[J]. Journal of China Agricultural University, 2019, 24(11): 123 – 133. (in Chinese)
- [14] 方鹏, 郝宏运, 王红英. 基于知识蒸馏的叠层笼养蛋鸡行为识别模型研究[J]. 农业机械学报, 2021, 52(10): 300 – 306.
FANG Peng, HAO Hongyun, WANG Hongying. Behavior recognition model of stacked cage laying hens based on knowledge distillation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(10): 300 – 306. (in Chinese)
- [15] 易诗, 沈练, 周思尧, 等. 基于增强型 Tiny – YOLOV3 模型的野鸡识别方法[J]. 农业工程学报, 2020, 36(13): 141 – 147.
YI Shi, SHEN Lian, ZHOU Siyao, et al. A pheasant identification method based on enhanced Tiny – YOLOV3 model[J]. Transactions of the CSAE, 2020, 36(13): 141 – 147. (in Chinese)
- [16] WU D, LÜ S, JIANG M, et al. Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments[J]. Computers and Electronics in Agriculture, 2020, 178: 105742.
- [17] 李天华, 孙萌, 丁小明, 等. 基于 YOLO v4 + HSV 的成熟期番茄识别方法[J]. 农业工程学报, 2021, 37(21): 183 – 190.
LI Tianhua, SUN Meng, DING Xiaoming, et al. Identification method of mature tomato based on YOLO V4 + HSV[J]. Transactions of the CSAE, 2021, 37(21): 183 – 190. (in Chinese)
- [18] 郭平秀, 李启南, 杨忠鹏. 一种图像增强及改进海洋生物图像检测算法[J]. 计算机工程与应用, 2022, 58(1): 1 – 12.
GUO Pingxiu, LI Qi'nian, YANG Zhongpeng. An image enhancement and improved marine biological image detection algorithm[J]. Computer Engineering and Applications, 2022, 58(1): 1 – 12. (in Chinese)
- [19] LIU S, HUANA D, WANG Y. Learning spatial fusion for single-shot object detection[J]. arXiv preprint arXiv: 1911.09516, 2019.
- [20] 汤青萍, 卜柱, 王志跃, 等. 肉鸽生产性能名词术语的规范与建议[J]. 中国家禽, 2017, 39(11): 64 – 67.
TANG Qingping, BU Zhu, WANG Zhiyue, et al. Pigeon whole production performance and standardize the terms suggest[J]. Chinese Poultry, 2017, 39(11): 64 – 67. (in Chinese)
- [21] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 390 – 391.
- [22] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904 – 1916.
- [23] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759 – 8768.
- [24] REDMON J, FARHADI A. YOLO v3: an incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.
- [25] ZHENG Z, WANG P, LIU W, et al. Distance – IoU loss: faster and better learning for bounding box regression[C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993 – 13000.
- [26] 王巧华, 顾伟, 蔡沛忠, 等. 基于改进 YOLO v4 的群体棉种双面破损检测方法[J]. 农业机械学报, 2022, 53(1): 389 – 397.
WANG Qiaohua, GU Wei, CAI Peizhong, et al. Double sided damage detection method of cotton seed in population based on improved YOLO v4[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(1): 389 – 397. (in Chinese)
- [27] 金耀, 何秀文, 万世主, 等. 基于 YOLO v3 的生猪个体识别方法[J]. 中国农机化学报, 2021, 42(2): 178 – 183.
JIN Yao, HE Xiuwen, WAN Shizhu, et al. Pig individual recognition method based on YOLO v3[J]. Chinese Journal of Agricultural Mechanization, 2021, 42(2): 178 – 183. (in Chinese)
- [28] 于敏, 屈丹, 司念文. 改进的 RetinaNet 目标检测算法[J]. 计算机工程, 2022, 48(8): 249 – 257.
YU Min, QU Dan, SI Nianwen. Improved RetinaNet target detection algorithm[J]. Computer Engineering, 2022, 48(8): 249 – 257. (in Chinese)
- [29] SONG H, DENG B, POUND M, et al. A fusion spatial attention approach for few-shot learning[J]. Information Fusion, 2022, 81: 187 – 202.