

基于改进 YOLO - MAO 检测框架的笼养白羽肉鸡行为检测方法

夏元天 寇旭鹏 薛洪成 李林

(中国农业大学信息与电气工程学院, 北京 100083)

摘要: 在大规模的肉鸡养殖场中,肉鸡行为通常由饲养员或专业兽医观察和分析,以确定肉鸡健康状况和养殖环境状态。然而这种方法耗时且主观。此外,在笼养环境中,由于鸡只的高密度和严重的互相遮挡,行为的视觉特征不明显,传统的检测算法不能准确地识别鸡只的行为特征。因此,本文提出一种改进的笼养白羽肉鸡行为检测的目标检测算法。所提出的算法由2个模块组成:多尺度细节特征融合模块(MDF)和目标关系推理模块(ORI)。多尺度细节特征模块充分利用和提取网络浅层特征映射中包含的多尺度细节特征,并将它们融合到负责相应尺度检测的特征映射中,实现细节特征的有效传输和补充。目标关系推理模块充分利用对象之间的位置关系进行推理和判断,使模型能更充分地利用对象之间的潜在关系来辅助检测。为验证所提出算法的有效性,在目标检测领域具有权威性的 COCO 公共数据集以及真实的大规模笼养白羽肉鸡养殖环境中自建的行为检测数据集上进行大量对比实验。实验结果表明,与其他最先进的模型相比,本文所提出的改进算法在 COCO 数据集和自建数据集上均达到最佳识别准确率;对喂食、饮水、移动和张嘴等影响肉鸡健康状况较为重要的行为进行检测,识别精度分别达 99.6%、98.7%、99.2% 和 98.3%。

关键词: 白羽肉鸡; 行为识别; 目标检测; 多尺度细节特征融合模块; 关系推理模块

中图分类号: TP391 文献标识码: A 文章编号: 1000-1298(2024)11-0103-09

OSID:



Behavior Detection Algorithm for Caged White-feather Broilers Based on Improved YOLO Detection Framework

XIA Yuan-tian KOU Xu-peng XUE Hong-cheng LI Lin

(College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China)

Abstract: In large-scale broiler farms, the behavior of broilers is usually observed and analyzed by feeders or professional veterinarians to determine their health status and breeding environment status. However, this method is time-consuming and subjective. In addition, in caged environments, due to the high density of chickens and serious mutual occlusion, the visual features of behavior are not obvious, and traditional detection algorithms cannot accurately identify the behavior characteristics of chickens. Therefore, an improved object detection algorithm for behavior detection of caged white-feather broilers was proposed. The proposed algorithm consisted of two modules: multi-scale detail feature fusion module (MDF) and object relation inference module (ORI). The multi-scale detail feature module fully utilized and extracted the multi-scale detail features contained in the shallow feature maps of the feature extraction network, and integrated them into the corresponding feature maps responsible for detection at the corresponding scale, achieving effective transmission and supplementation of detail features. The relational reasoning module fully utilized the positional relationships between objects for inference and judgment, enabling the model to more fully utilize the potential relationships between objects to assist in detection. To verify the effectiveness of the proposed algorithm, a large number of comparative experiments on both authoritative public datasets in the field of object detection and self-built behavior detection datasets in real large-scale caged white-feather broiler breeding environments was conducted.

收稿日期: 2023-09-08 修回日期: 2024-07-01

基金项目: 国家科技创新 2030—新一代人工智能重大项目(2021ZD0113701)

作者简介: 夏元天(1990—),男,博士生,主要从事深度学习机和计算机视觉研究,E-mail: xiayuantian@cau.edu.cn

通信作者: 李林(1963—),女,教授,博士生导师,主要从事农业工程和计算机视觉研究,E-mail: lilinli@cau.edu.cn

The experimental results showed that the proposed improved algorithm achieved the best detection accuracy compared with other state-of-the-art models, both in the COCO dataset and the self-built dataset. For the detection of behaviors such as feeding, drinking, moving, and opening the mouth, which were crucial for the health status of broiler chickens, the algorithm achieved accuracy rates of 99.6%, 98.7%, 99.2%, and 98.3% respectively.

Key words: white-feather broilers; behavior recognition; object detection; multi-scale detail feature fusion module; relation inference module

0 引言

不同的行为模式反映了动物本身的不同状态,这是评估动物健康状况和福利的重要指标^[1]。在早期阶段,对动物行为的识别主要通过人工观察进行,这不仅耗时耗力,而且高度主观。随着技术升级,一些具有行为识别功能的可穿戴设备被应用于动物养殖行业^[2]。但这些设备会分散动物的注意力,影响它们的自然行为表现。近年来,随着基于深度学习的计算机视觉技术不断进步,基于目标检测的动物行为识别技术被广泛应用于农业养殖场景中。翟亚红等^[3]对 YOLO v5n 模型进行改进,提出一种绵羊行为识别方法,对舍养绵羊进食、躺卧以及站立行为进行识别,平均识别精度达到 95.8%。王旺等^[4]提出一种基于改进 YOLO v8s 的羊只行为识别方法,平均识别精度达到 90.01%。袁洪波等^[5]提出一种融合三维残差卷积神经网络、双向长短期记忆网络和注意力机制的深度学习网络模型,对羊只站立、躺卧、进食、行走和反刍 5 种行为进行识别,平均精度达到 98.72%。李恩泽等^[6]提出一种融合时间和空间注意力信息的多分支并行 CAFNet 奶牛行为识别模型,实现对奶牛进食、躺卧、站立、行走和甩尾行为精确识别,平均准确率为 91.06%。段青玲等^[7]提出一种基于 SNSS-YOLO v7 的肉牛行为识别方法,以提高肉牛行为识别效率和准确性,平均精度为 95.2%。李林等^[8]提出 K-SSD-F 算法,对东亚飞蝗个体及行为进行检测计数,平均检测精度为 88.94%。

目前,目标检测算法模型可以分为两种类型:以 Faster-RCNN 系列^[9-13]为代表的两阶段目标检测器和以 YOLO 系列^[14-21]、SSD 系列^[22-23]为代表的单阶段目标检测器。由于单阶段检测器具有检测速度快、准确性高和易于部署等优点,因此在计算机视觉的下游任务中被广泛使用。在笼养白羽肉鸡养殖场景中,检测任务主要面临以下关键问题:目前,关于动物行为识别的研究主要集中在大型动物如猪上,对于笼养白羽肉鸡的研究较少,缺乏相关视频和图像数据库;在圈养条件下,笼子的墙壁会对个体鸡造成大范围遮挡;为最大化利润,鸡通常在有限的笼

子空间中被尽可能地饲养,导致大量目标的粘连和遮挡。

本文对基于单阶段目标检测算法的通用框架结构进行改进:基于真实笼养养殖生产场景,构建一个笼养白羽肉鸡的行为识别数据集,以解决相关研究缺乏相关视频和图像数据的问题;提出一种多尺度细节特征融合方法,以解决鸡只互相遮挡导致的细节特征丢失问题;提出一种目标关系推理方法,以解决模糊目标类别的边界判别问题。

1 数据采集与预处理

1.1 数据采集

本文所用笼养白羽肉鸡图像于 2022 年 7 月 15 日—8 月 26 日,及 2022 年 10 月 31 日—12 月 10 日,在山东省青岛市莱西市正大养殖基地采集。采集设备为两台大华 DH-IPC-HFW3433T-ZAS 型 400 万像素红外变焦枪式网络摄像机,连续拍摄相邻的顶层鸡笼。图 1 为鸡舍结构和设备布局,图 2 为捕捉鸡行为图像过程。

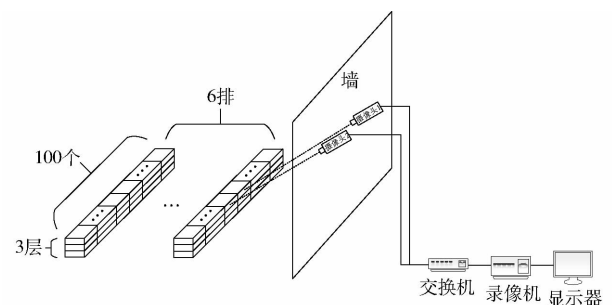


图 1 鸡舍结构和设备布局

Fig. 1 Structure of chicken house and equipment arrangement



图 2 鸡只行为图像采集

Fig. 2 Collection of chicken behavior images

1.2 数据预处理

从图 3a 可看出,原始图像中目标鸡舍周围存在其他鸡舍干扰。为获取准确对焦的目标鸡舍图像和

统一格式的目标鸡舍图像,使用 Python 和 OpenCV 对原始图像进行处理,调整后图像如图 3b 所示。



(a) 原始图像 (b) 调整后图像

图 3 原始图像和调整后的图像

Fig. 3 Original and adjusted image

经处理,笼养白羽肉鸡行为识别数据集共包括 35 681 幅图像,这些数据包括不同日龄的白羽肉鸡图像,同一日龄段的不同时间段图像,以及两批次中在喂食、饮水、移动和张嘴行为时有、无隔热膜覆盖的图像。为避免训练过程中由于数据不平衡导致的准确性问题,各种行为在总数据中的比例相同。将总数据的 80% 作为训练集,10% 作为验证集,10% 作为测试集。最终,得到由 28 545 幅图像组成的训练集,由 3 568 幅图像组成的验证集,由 3 568 幅图像组成的测试集。

使用开源应用程序 Labeling 来标注训练集和验证集中白色肉鸡的喂食、饮水、移动和张嘴行为。行为标签名称分别为 eat、drink、move 和 open_mouth。对每个图像中的饮水槽和喂食槽进行标注,用于在对象关系推断过程中辅助区分类别。此外,采用 PASCAL VOC^[24] 的数据格式存储带有标注的行为标签和图像数据。

2 改进方法

2.1 多尺度细节特征融合算法

目前,单阶段目标检测模型的框架主要以 SSD 和 YOLO 为代表,并衍生出一系列具有出色检测性能的改进模型。尽管不同版本的模型具有不同的网络设计和改进方案,但它们通常可以通过其整体结构进行分析。整体结构由特征提取骨干网络 Backbone、特征融合网络 Neck 和检测头部网络 Head 组成。特征提取骨干网络负责从输入图像中提取特征,并通过连续的下采样将提取的特征图进一步传递到网络中。特征融合网络负责融合不同尺度的特征图,以提高特征图的多尺度特征表达能力。检测头部网络负责检测特征融合网络所融合的特征图,并最终提供目标的位置和类别预测结果。图 4 为单阶段目标检测模型的一般框架结构。

如图 4 所示,随着网络深度增加和感受野扩大,目标的高级语义特征逐渐取代其细节特征。因此,如何充分利用网络低层特征图中的细节特征来弥补丢失的细节特征,成为增强模型预测能力的关键。

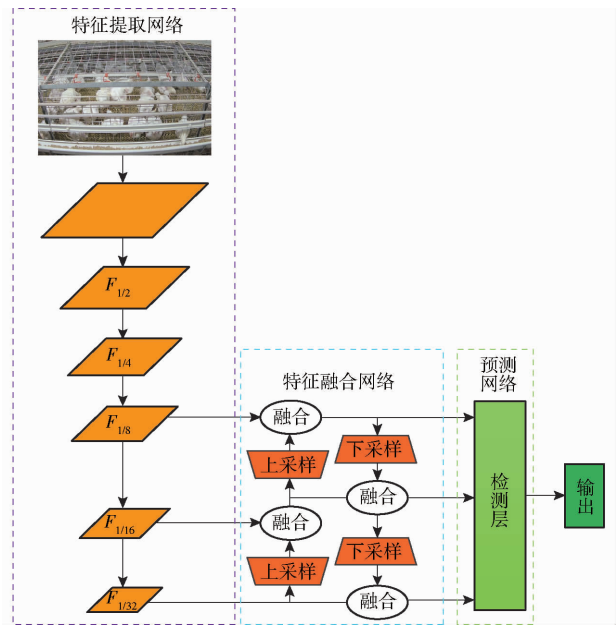


图 4 单阶段目标检测框架结构图

Fig. 4 Structure of one-stage object detector

为此,本文在原始特征提取骨干网络的基础上设计一个独立的多尺度细节特征提取网络,并对特征提取网络进行修改,去除原有基于上采样特征映射的下采样路径,仅保留自顶向下的特征传播路径。修改后,网络中所有自底向上的特征传播路径都由多尺度详细特征提取模块完成,如图 5 所示。

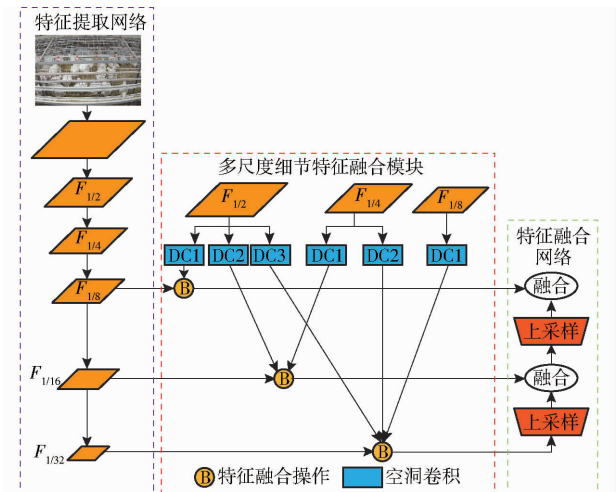


图 5 多尺度详细特征融合模块结构图

Fig. 5 Structure of multi-scale detail feature fusion module

如图 5 中特征提取网络所示,输入图像经过一系列步长为 2 的卷积下采样操作,直到获得 32 倍下采样的特征图。在下采样过程中,细节特征逐渐减少,而语义特征逐渐增加。本文使用具有丰富细节特征的低级特征图,例如 $F_{1/2}$ 、 $F_{1/4}$ 和 $F_{1/8}$ 来弥补深层特征图中丢失的细节特征。

如图 5 中多尺度详细特征融合模块所示,将 $F_{1/2}$ 、 $F_{1/4}$ 和 $F_{1/8}$ 映射到多尺度详细特征模块中。 $F_{1/2}$

通过3个特征提取分支DC1、DC2和DC3生成细节特征图 $F_{1/2 \rightarrow 1/8} - F_{1/2 \rightarrow 1/32}$ ，这些特征图与特征图 $F_{1/8} - F_{1/32}$ 具有相同的感受野和特征区域。 $F_{1/4}$ 通过两个特征提取分支DC1和DC2生成细节特征图 $F_{1/4 \rightarrow 1/16}$ 和 $F_{1/4 \rightarrow 1/32}$ ，这些特征图与特征图 $F_{1/16}$ 、 $F_{1/32}$ 具有相同的感受野和特征区域。 $F_{1/8}$ 通过一个特征提取分支DC1生成一个细节特征图 $F_{1/8 \rightarrow 1/32}$ ，该特征图与特征图 $F_{1/32}$ 具有相同的感受野和特征区域。在特征融合之前，需确保所有特征图具有相同的感受野和特征区域，否则可能导致不一致的特征。因此，需要计算特征提取分支DC1、DC2和DC3的卷积核尺寸和步长。

不同层级特征图感受野可通过递归公式获得，公式为

$$\begin{cases} f_1 = k \\ f_i = kf_{i-1} - (f_{i-1} - 1)(k - s) \end{cases} \quad (1)$$

式中 f_i ——两个特征图之间间隔为 i 的感受野
 k ——卷积核尺寸
 s ——步长

根据式(1)可得不同层级特征图的感受野。特征图 $F_{1/2 \rightarrow 1/32}$ 、 $F_{1/4 \rightarrow 1/32}$ 、 $F_{1/8 \rightarrow 1/32}$ 和 $F_{1/32}$ 中的每一个特征点对应特征图 $F_{1/2}$ 中的一个 31×31 的感受野。特征图 $F_{1/2 \rightarrow 1/16}$ 、 $F_{1/4 \rightarrow 1/16}$ 和 $F_{1/16}$ 中的每一个特征点对应特征图 $F_{1/2}$ 中的一个 15×15 的感受野。特征图 $F_{1/2 \rightarrow 1/8}$ 和 $F_{1/8}$ 中的每一个特征点对应特征图 $F_{1/2}$ 中的一个 7×7 的感受野。因此，可以得到DC1卷积核尺寸为7，步长为4，DC2的卷积核尺寸为15，步长为8，DC3卷积核尺寸为31，步长为16。

为避免较大卷积核引起的计算量激增及过于密集采样导致的冗余信息，本文使用扩张卷积来稀疏采样特征图。扩张率计算公式为

$$k' = k + (k - 1)(d - 1) \quad (2)$$

式中 k' ——等效卷积核尺寸
 d ——扩张率

根据式(2)可得DC1、DC2和DC3扩张率分别为3、7和15。具体融合过程如图6所示。

在多尺度细节特征融合后，负责目标检测的特征图中包含所有低层特征图中的多尺度细节特征。此外，所有的融合路径都是直接独立连接的，有效避免特征逐层传递所引起的特征丢失和混淆的问题。

2.2 目标关系推理算法

目标关系推理模块位于检测层之后，它能够充分利用高置信度的目标类别来辅助识别周围置信度较低的目标。对于给定的数据集，构建一个对称的目标矩阵 $C = (c_1, c_2, \dots, c_n) \in \mathbf{R}^{n \times n}$ ，其中元素 c_{ij} 表示类别 i 和 j 之间的相关性。 c_{ij} 计算公式为

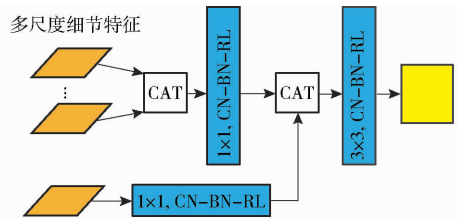


图6 多尺度特征融合操作

Fig. 6 Multi-scale fusion operation

$$c_{ij} = \sum_{n=1}^N R_n(i, j) \quad (3)$$

$$R_n(i, j) = \begin{cases} 0 & (d_{ij} = None) \\ e^{-\frac{d_{ij}^2}{2\sigma^2}} & (d_{ij} \neq None) \end{cases} \quad (4)$$

式中 N ——样本数量

$R_n(i, j)$ ——第 n 个样本中类别 i 和类别 j 之间的目标关系

d_{ij} ——两个目标中心点之间的距离

σ ——方差

当 $d_{ij} = None$ 时，表示类别 i 和类别 j 没有出现在同一幅样本图像中，此时 $R_n(i, j)$ 为 0。否则，两个目标之间的关系由一个均值为 0、方差为 σ 的高斯散射核决定，将两个目标之间的关系限制在 $0 \sim 1$ 范围内。

根据样本图像高度 h 和宽度 w 可得可能存在于样本中的两个目标之间的最大距离 $d = \sqrt{w^2 + h^2}$ 。为更准确地描述两个目标之间的相关程度，定义当 $d_{ij} \leq \frac{d}{2}$ 时， $R_n(i, j) \geq 0.5$ ，当 $d_{ij} > \frac{d}{2}$ 时， $R_n(i, j) < 0.5$ 。因此，可以得到满足上述条件的 σ 值，公式为

$$\sigma = \frac{d}{2\sqrt{2\ln 2}} \quad (5)$$

本文目标是重新评估在具有较高置信度的目标周围可能存在的具有类别模糊性目标的置信度值。当目标的置信度大于 0.6 时，认为检测结果可靠。当置信度在 0.4、0.6 之间时，检测结果存在模糊性。对于具有模糊性的目标，其置信度将进行重新评估，公式为

$$s_i(x) = s_{i-1}(x) + \lambda \frac{e^{p_i(x)} - e^{-p_i(x)}}{e^{p_i(x)} + e^{-p_i(x)}} \quad (6)$$

$$p_i = \sum_{j \in S} c_{ij} - T \quad (7)$$

式中 $s_i(x)$ ——目标 x 属于 i 类的置信度

p_i —— x 属于 i 的概率

λ ——权衡参数

T —— $c_{ij} = 0$ 的次数

如果 $p_i > 0$ ，则 $s_i(x)$ 的值将增加，如果 $p_i < 0$ ，

$s_i(x)$ 的值将减少。具体过程如图 7 所示。

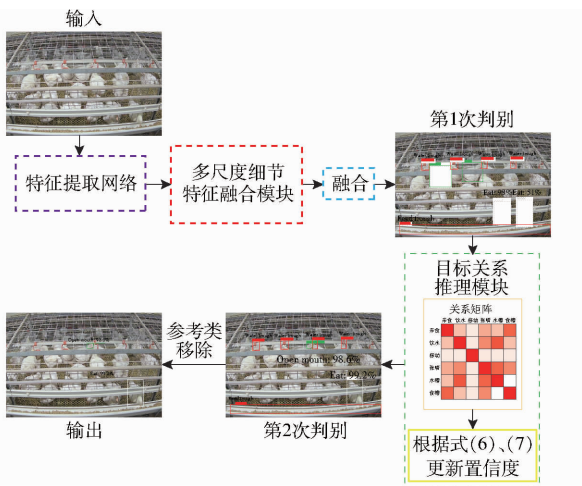


图 7 目标关系推理过程

Fig. 7 Process of object relationship inference

如图 7 所示,关系矩阵中的颜色深度表示不同类别之间的关联程度,它是根据式(4)在训练过程中动态建立。在第 1 次和第 2 次判别图中出现的红色位置框是预先校准的参考类别,即饲料槽和水槽,用于辅助检测其他类别。在第 2 次判别后,参考类别将被移除,最终输出只包含主要类别的检测结果。

3 实验与结果分析

3.1 实验细节和环境

本文使用 DSSD、SSD、YOLO v2 ~ YOLO v7 以及它们相应的主干网络作为本文改进的基准。

对于 COCO 2017^[25] 数据集,使用训练集和验证集进行训练,使用测试集进行测试。将输入尺寸设置为 512 像素 × 512 像素,动量设置为 0.9,权重衰减设置为 0.000 5,批量大小设置为 32。使用随机梯度下降(SGD)算法对模型进行 160 个周期的训练,初始学习率设置为 0.000 25。学习率使用 warm-up 和余弦退火函数进行动态调整。其他优化和数据增强方法与相应的基准方法一致。

对于自建的笼养肉鸡行为识别数据集,在训练集和验证集上训练模型,并在测试集上进行测试。使用 SGD 算法对模型进行 100 个周期的训练,初始学习率为 0.000 25。其余参数设置与在 COCO 数据集上的训练一致。

考虑到当目标置信度低于 0.4 时几乎不存在正样本,而当目标置信度高于 0.6 时几乎不存在负样本,存在歧义的目标置信度最有可能处于 0.4 ~ 0.6 之间,因此,对于关系推理模块,将置信度模糊区间设置为 0.4 ~ 0.6。当检测到的目标类别置信度在此区间内时,关系推理模块将进行二次判别并更新

置信度。

本文模型在 Intel Core i9 - 9900k、NVIDIA 3090、64 GB RAM、Python 3.8、PyTorch 1.10.0 和 CUDA 11.3 的环境下运行。

采用 AP (Average precision) 和 mAP (mean average precision, mAP) 精度指标进行评估。在 mAP 指标体系中,AP₅₀ 和 AP₇₅ 分别对 IoU 阈值为 50 和 75 条件下的模型进行评估,AP_S、AP_M 和 AP_L 等指标则分别对像素面积小于 32 × 32 的小型目标、像素面积在 32 × 32 到 96 × 96 之间的中等目标以及像素面积大于 96 × 96 的大型目标进行评估。

3.2 权衡参数设置

为获得关系推理模块中的权衡参数 λ ,使用 YOLO v7 作为改进的基准模型,在 COCO 数据集上进行参数调优实验。如表 1 所示,当 $\lambda = 0.4$ 时,模型达到最佳准确率 54.1%。因此,在后续实验中采用 $\lambda = 0.4$ 。

表 1 权衡参数设置

Tab. 1 Trade-off parameter settings

输入/(像素 × 像素)	数据集	λ	AP/%
512 × 512	COCO 2017	< 0.4	< 50.0
		0.4	54.1
		0.5	53.1
		0.6	53.5
		> 0.6	< 50.0

3.3 COCO 2017 实验结果

表 2 为本文的方法与其他检测器在 COCO 2017 数据集上的比较结果。改进模型的检测结果是通过使用 3.1 节中设置的参数进行训练得到的,而其他检测器的结果则来自于文献。

如表 2 所示,当 DSSD 和 ResNet101 作为基准模型时,改进模型在 AP 指标上达到 35.7%,比基准模型高 2.5 个百分点。在 AP₅₀ 和 AP₇₅ 指标下,改进模型分别提高 4.5、2.6 个百分点。在 AP_S、AP_M 和 AP_L 指标下,改进模型分别提高 3.2、3.0、0.6 个百分点。

当 SSD 和 VGG16 作为基准模型时,相比基准模型,改进模型在 AP 上提高 4.7 个百分点。在 AP₅₀ 和 AP₇₅ 指标下,改进模型分别提高 5.5、5.3 个百分点。在 AP_S、AP_M 和 AP_L 指标下,改进模型分别提高 6.8、5.2、2.1 个百分点。

当 YOLO v2 作为基准模型时,相比基准模型,改进模型在 AP 上提高 4.2 个百分点。在 AP₅₀ 和 AP₇₅ 指标下,改进模型分别提高 4.6、4.7 个百分点。在 AP_S、AP_M 和 AP_L 指标下,改进模型分别提高 3.7、3.5、3.1 个百分点。

表2 COCO 2017 检测结果

Tab.2 Experimental results on COCO 2017

方法	特征提取网络	AP/%	AP ₅₀	AP ₇₅	AP _s	AP _M	AP _L	%
R-FCN ^[12]		29.9	51.9		10.8	32.8	45.0	
SSD513 ^[23]	ResNet101	31.2	50.4	33.3	10.2	34.5	49.8	
DSSD513 ^[23]		33.2	53.3	35.2	13.0	35.4	51.1	
DSSD513 + MAO		35.7	57.8	37.8	16.2	38.8	51.7	
Faster R-CNN ^[11]		24.2	45.3	23.5	7.7	26.4	37.1	
SSD512 ^[22]	VGG16	28.8	48.5	30.3	10.9	31.8	43.5	
SSD512 + MAO		33.5	54.0	35.6	17.7	37.0	45.6	
YOLO v2 ^[15]	Darknet19	21.6	44.0	19.2	5.0	22.4	35.5	
YOLO v2 + MAO		25.8	48.6	22.9	8.7	25.9	38.6	
YOLO v3 ^[16]	Darknet53	33.0	57.9	34.4	18.3	35.9	41.9	
YOLO v3 + MAO		35.2	59.6	37.5	20.6	37.5	43.7	
YOLO v4 ^[17]	CSPDarknet53	43.0	64.9	46.5	24.3	46.1	55.2	
YOLO v4 + MAO		45.5	66.1	48.7	25.9	47.6	57.1	
YOLO v5-L ^[19]	Modified CSP v5	49.6	67.9	54.2	30.8	53.5	63.7	
YOLO v5-L + MAO		51.8	69.5	56.3	33.1	55.9	66.1	
YOLO v6-L ^[20]	EfficientRep	52.5						
YOLO v6-L + MAO		54.6	69.5	58.6	34.7	58.1	68.4	
YOLO v7 ^[21]	v7 backbone	51.2	69.7	55.5	35.2	56.0	66.7	
YOLO v7 + MAO		54.1	71.6	58.3	36.5	57.8	68.0	

当以 YOLO v3 作为基准模型时,相比基准模型,改进模型在 AP 上提高 2.2 个百分点。在 AP₅₀ 和 AP₇₅ 指标下,改进模型分别提高 1.7、3.1 个百分点。在 AP_s、AP_M 和 AP_L 指标下,改进模型分别提高 2.3、2.1、1.8 个百分点。

当 YOLO v4 作为基准模型时,相比基准模型,改进模型在 AP 上提高 2.5 个百分点。在 AP₅₀ 和 AP₇₅ 指标下,改进模型分别提高 1.2、2.2 个百分点。在 AP_s、AP_M 和 AP_L 指标下,改进模型分别提高 1.6、1.5、1.9 个百分点。

当 YOLO v5-L 作为基准模型时,相比基准模型,改进模型在 AP 上提高 2.2 个百分点。在 AP₅₀ 和 AP₇₅ 指标下,改进模型分别提高 1.6、2.1 个百分点。在 AP_s、AP_M 和 AP_L 指标下,改进模型分别提高 2.3、2.4、2.4 个百分点。

当 YOLO v6-L 作为基准模型时,改进模型在 AP 指标上达到 54.6%,比基准模型高 2.1 个百分点。

当 YOLO v7 作为基准模型时,相比基准模型,改进模型在 AP 上提高 2.9 个百分点。在 AP₅₀ 和 AP₇₅ 指标下,改进模型分别提高 1.9、2.8 个百分点。在 AP_s、AP_M 和 AP_L 指标下,改进模型分别提高 1.3、1.8、1.3 个百分点。

上述结果表明,改进模型显著提高检测准确性,尤其是在更严格的 AP₇₅ 指标和最具挑战性的小目标检测评估指标 AP_s 上。这充分证明本文所提出方

法的有效性。

3.4 笼养白羽肉鸡行为识别数据集实验结果

表 3 为本文方法与其他检测算法在笼养白羽肉鸡行为识别数据集检测结果比较。

表 3 在笼养白羽肉鸡行为识别数据集上的实验结果

Tab.3 Experimental results on captive white-feathered chicken behavior recognition dataset

方法	特征提取网络	AP				mAP
		喂食	饮水	移动	张嘴	
SSD512 ^[22]	VGG16	69.8	67.9	68.5	65.7	69.0
SSD512 + MAO		72.3	70.6	71.4	68.9	71.6
DSSD513 ^[23]	ResNet101	80.4	79.3	78.6	75.7	79.5
DSSD513 + MAO		83.1	82.3	81.5	79.5	82.4
YOLO v2 ^[15]	Darknet19	65.3	64.5	65.1	61.4	64.9
YOLO v2 + MAO		70.2	68.9	67.5	65.8	68.8
YOLO v3 ^[16]	Darknet53	78.4	77.2	77.9	73.5	77.5
YOLO v3 + MAO		82.2	80.2	80.6	77.6	81.1
YOLO v4 ^[17]	CSPDarknet53	87.8	87.0	87.3	82.4	86.9
YOLO v4 + MAO		90.1	90.4	90.5	86.5	90.3
YOLO v5-L ^[19]	Modified CSP v5	93.1	92.3	92.7	87.4	91.7
YOLO v5-L + MAO		96.5	95.8	96.0	92.4	95.8
YOLO v6-L ^[20]	EfficientRep	96.1	94.7	95.3	92.9	95.2
YOLO v6-L + MAO		99.6	98.5	98.9	97.7	98.6
YOLO v7 ^[21]	v7 backbone	96.2	95.0	95.7	93.6	95.7
YOLO v7 + MAO		99.6	98.7	99.2	98.3	99.4

如表 3 所示,对于所选的基准模型,改进模型在识别准确率方面有显著提升。当 DSSD 作为基准时,改进模型平均精度均值(mAP)提高 2.9 个百分

点。当 SSD 作为基准时,改进模型 mAP 提高 2.6 个百分点。当 YOLO v2 作为基准时,改进模型 mAP 提高 3.9 个百分点。当 YOLO v3 作为基准时,改进模型 mAP 提高 3.6 个百分点。当 YOLO v4 作为基准时,改进模型 mAP 提高 3.4 个百分点。当 YOLO v5-L 作为基准时,改进模型 mAP 提高 4.1 个百分点。当 YOLO v6-L 作为基准时,改进模型 mAP 提高 3.4 个百分点。当 YOLO v7 作为基准时,改进模型 mAP 提高 3.7 个百分点。

从以上分析可知,当 YOLO v6 和 YOLO v7 作为基准时,改进模型不同行为识别准确率均超过 98%。其中,当 YOLO v7 作为改进基准时,喂食、饮水、移动和张嘴的识别精度分别达到 99.6%、98.7%、99.2% 和 98.3%,mAP 达到 99.4%,为所有比较模型中最高的准确率。

为进一步说明本文的改进模型在白羽肉鸡养殖过程中不同日龄、同一日龄不同时段,以及有无隔热膜覆盖时的检测性能,使用表 3 中取得最佳检测性能的未改进模型 YOLO v7 作为对比基线,在上述几种不同的养殖场景中与改进模型进行对比。

表 4 为在不同日龄下基线模型与改进模型结果对比。如表 4 所示,将白羽笼养肉鸡的整个饲养周期分为 3 个阶段,分别为 1~15 日龄、16~30 日龄和 31~45 日龄。改进模型在这 3 个日龄段上所获得的 mAP 分别比基线模型 YOLO v7 高 6.0、1.8、1.9 个百分点。由于在 1~15 日龄段时,鸡只体积较小,体型发育速度较快,导致小目标检测和多尺度目标检测等问题,增加了模型检测难度,因此,从以

上对比数据可以看出,本文所提出的算法能够以较高的准确率有效地检测出处于较早日龄时的鸡只行为,提升模型应对小型目标检测和多尺度目标检测的能力。此外,无论是基线模型还是改进模型,对于 16~30 日龄段下的检测效果均优于前后两个日龄段。这是因为处于该日龄段内的鸡只处于发育平稳期,并且体型适中,避免了 1~15 日龄存在的体型过小、尺度变化大等问题,也避免了 31~45 日龄期间由于体型过大而导致的遮挡严重问题。

表 4 不同日龄下实验结果对比

Tab. 4 Comparative experimental results at different ages

方法	日龄/ d	AP/%				mAP/ %
		喂食	饮水	移动	张嘴	
YOLO v7 ^[21]	1~15	92.8	91.7	93.1	92.2	92.4
	16~30	98.3	97.0	98.7	97.2	97.8
	31~45	97.4	96.1	97.3	96.8	96.9
YOLO v7 + MAO	1~15	98.5	97.6	98.4	97.8	98.4
	16~30	99.8	99.1	99.3	98.4	99.6
	31~45	98.9	98.2	98.7	97.7	98.8

表 5 为同一日龄不同时段下基线模型与改进模型之间的对比。如表 5 所示,选择处于发育平稳期的第 22 日龄的鸡只进行实验,并将该日龄以 6 h 为一个间隔分为 4 个时间段。由于处于同一日龄内鸡只的变化几乎可以忽略不计,模型对于同一日龄下的鸡只行为检测几乎具有相同的准确率。并从表 5 可以看出,改进模型的 mAP 以及喂食、饮水、移动、张嘴平均精度分别比基线模型 YOLO v7 高 1.2、1.6、2.2、0.9、1.4 个百分点。

表 5 同一日龄不同时段实验结果对比

Tab. 5 Comparative experimental results at different time within the same age

方法	日龄/d	时间段	AP/%				mAP/%
			喂食	饮水	移动	张嘴	
YOLO v7 ^[21]	22	00:00—06:00					
		06:00—12:00					
		12:00—18:00	97.9 ± 0.4	97.1 ± 0.2	98.5 ± 0.2	97.0 ± 0.3	98.3 ± 0.3
		18:00—24:00					
YOLO v7 + MAO	22	00:00—06:00					
		06:00—12:00					
		12:00—18:00	99.5 ± 0.3	99.3 ± 0.2	99.4 ± 0.3	98.4 ± 0.1	99.5 ± 0.4
		18:00—24:00					

表 6 为有、无隔热膜覆盖遮挡时基线模型与改进模型之间的对比。如表 6 所示,当有隔热膜遮挡时,改进模型 mAP 以及喂食、饮水、移动、张嘴平均精度分别比基线模型 YOLO v7 高 9.9、7.1、9.1、9.2、11.3 个百分点,检测精度有较大提升,充分说明本文提出的改进算法能够大幅提高目标遮挡情况

下的检测准确率。

3.5 消融实验

为进一步讨论本文提出的不同组件对检测准确性的影响,构建了 YOLO v7 + MDF、YOLO v7 + ORI 和 YOLO v7 + MDF + ORI 模型。在笼养白羽肉鸡行为识别数据集上进行模型测试,结果如表 7 所示。

表6 有、无隔热膜覆盖时实验结果对比

Tab.6 Experimental comparison results with and without thermal insulation film coverage %

方法	隔热膜	AP				mAP
		喂食	饮水	移动	张嘴	
YOLO v7 ^[21]	有	91.6	89.3	88.6	85.3	88.7
	无	95.9	95.7	96.3	94.5	95.6
YOLO v7 + MAO	有	98.7	98.4	97.8	96.6	98.6
	无	99.8	99.5	99.3	99.1	99.5

表7 不同改进模块对检测准确性的影响

Tab.7 Effect of different improved modules on detection accuracy

方法	图像尺寸/ (像素×像素)	mAP/ %	帧率/ (f·s ⁻¹)
YOLO v7 ^[21]	512×512	95.7	152
YOLO v7 + MDF	512×512	97.9	132
YOLO v7 + ORI	512×512	97.3	146
YOLO v7 + MAO	512×512	99.4	124

如表7所示, YOLO v7 + MDF的mAP达到97.9%, 提高2.2个百分点。YOLO v7 + ORI的mAP达到97.3%, 提高1.6个百分点。当同时使用MDF和ORI模块时, mAP达到99.4%, 提高3.7个百分点。

由于MDF模块增加了额外的细节特征提取分支, 不可避免地增加参数和计算量, 对检测速率产生一定影响。如表7所示, 仅使用MDF时, 帧率降低20 f/s。仅使用ORI时, 帧率降低6 f/s。当同时使用MDF和ORI时, 帧率降低28 f/s。尽管改进模型检测速度有所降低, 但仍能保证实时检测速度, 并提供更高的检测准确性。

3.6 识别结果可视化及定性评估

选择表3中取得最佳检测性能的未改进模型YOLO v7作为基线模型与改进模型在笼养白羽肉鸡数据集上进行识别结果的视觉比较。在复杂的真实饲养场景中, 存在许多复杂情况, 如遮挡、光照不均和密集连接, 这对检测算法的性能提出了更高挑战。如图8所示, 当存在轻度遮挡、密集连接等普通的养殖场景中, 改进模型检测到大部分YOLO v7无法检测或置信度较低的鸡只行为。如图9所示, 当存在重度遮挡、密集连接以及光照不均等更为复杂的典型养殖场景中, 改进模型仍然能够以较高精度完成鸡只行为的检测。在图8和图9中, 目标置信度高于0.5的鸡只行为被显示出来。

此外, 从图9中还可以看出, 如遮挡、密集连接等复杂检测场景通常不是单一存在的, 而是多种复杂场景的叠加, 因此实验结果充分证明所提出

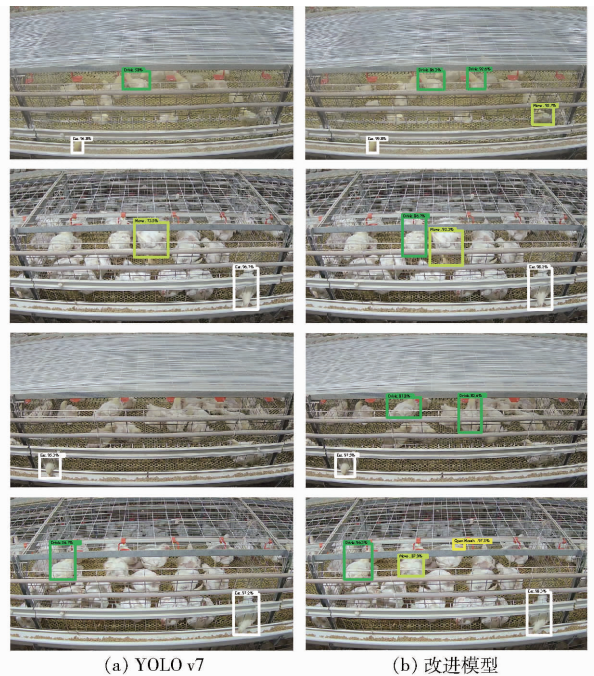


图8 预测结果比较

Fig.8 Comparison of prediction results

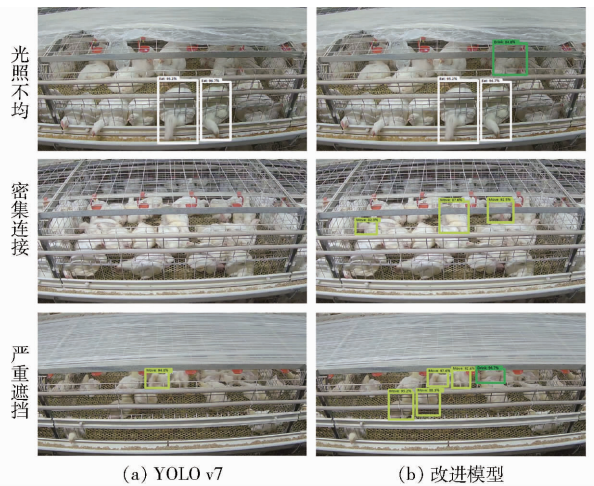


图9 典型情况下检测结果比较

Fig.9 Comparison of detection results under typical scenarios

的方法在复杂场景中能显著提升模型的目标检测能力。

4 结束语

使用基于深度学习的目标检测技术解决了笼养白羽肉鸡在真实饲养场景中的行为识别问题。为更好地适应真实饲养场景, 基于单阶段目标检测算法的框架提出两种改进方法: 多尺度细节特征融合方法和目标关系推理方法。多尺度细节特征模块可以有效提取和融合网络低层中包含的细节特征, 缓解特征提取引起的细节特征丢失问题, 并有效解决鸡只之间、鸡只与笼子之间、鸡只与水槽之间、鸡只与饲料槽之间的遮挡问题。目标关系推理模块可以充

分利用目标之间的关系,辅助模型进行更准确的检测,并解决鸡只行为类别之间模糊判别的问题。此外,建立了一个笼养白羽肉鸡行为识别数据集,为该

领域增加了新的科研数据。实验结果充分证明本文提出的改进方法在公共数据集和笼养白羽肉鸡行为识别数据集上均取得了最佳检测精度。

参 考 文 献

- [1] PEDEN R S E, TURNER S P, BOYLE L A, et al. The translation of animal welfare research into practice: the case of mixing aggression between pigs[J]. *Applied Animal Behaviour Science*, 2018, 204:1-9.
- [2] MAGHSOUDI O H, TABRIZI A V, ROBERTSON B, et al. Honeybee detection and pose estimation using convolutional neural networks[C]//*Proceedings of the 2017 51st Asilomar Conference on Signals, Systems, and Computers*, 2017: 69-73.
- [3] 翟亚红,王杰,徐龙艳,等.基于改进 YOLO v5n 的舍养绵羊行为识别方法[J].*农业机械学报*,2024,55(4):231-240. ZHAI Yahong, WANG Jie, XU Longyan, et al. Behavior recognition of domesticated sheep based on improved YOLO v5n[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024,55(4):231-240. (in Chinese)
- [4] 王旺,王福顺,张伟进,等.基于改进 YOLO v8s 的羊只行为识别方法[J].*农业机械学报*,2024,55(7):325-335,344. WANG Wang, WANG Fushun, ZHANG Weijin, et al. Sheep behavior recognition method based on improved YOLO v8s[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024,55(7):325-335,344. (in Chinese)
- [5] 袁洪波,曹润柳,程曼.融合 Res3D、BiLSTM 和注意力机制的羊只行为识别方法[J].*农业机械学报*,2024,55(4):221-230. YUAN Hongbo, CAO Runliu, CHENG Man, et al. Fusion of Res3D, BiLSTM and attention mechanism for sheep behavior recognition method[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024,55(4):221-230. (in Chinese)
- [6] 李恩泽,王克俭,司永胜,等.基于改进 ConvNeXt 的奶牛行为识别方法[J].*农业机械学报*,2024,55(5):282-289,404. LI Enze, WANG Kejian, SI Yongsheng, et al. Cow behavior recognition method based on improved ConvNeXt[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024,55(5):282-289,404. (in Chinese)
- [7] 段青玲,赵芷青,蒋涛,等.基于 SNSS-YOLO v7 的肉牛行为识别方法[J].*农业机械学报*,2023,54(10):266-274,347. DUAN Qingling, ZHAO Zhiqing, JIANG Tao, et al. Behavior recognition method of beef cattle based on SNSS-YOLO v7[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023,54(10):266-274,347. (in Chinese)
- [8] 李林,柏召,刁磊,等.基于 K-SSD-F 的东亚飞蝗视频检测与计数方法[J].*农业机械学报*,2021,52(增刊):261-267. LI Lin, BAI Zhao, DIAO Lei, et al. Video detection and counting method of east Asian migratory locusts based on K-SSD-F [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021,52(Supp.):261-267. (in Chinese)
- [9] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//*Proceedings of the Conference on Computer Vision and Pattern Recognition*, 2014:580-587.
- [10] GIRSHICK R. Fast R-CNN[C]//*Proceedings of the IEEE International Conference on Computer Vision*, 2015: 1440-1448.
- [11] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [12] DAI J, LI Y, HE K, et al. R-FCN: object detection via regionbased fully convolutional networks[C]//*Proceedings of the 30th International Conference on Neural Information Processing Systems (NIPS'16)*, 2016: 379-387.
- [13] UIJLINGS J R, SANDE K E, GEVERS T, et al. Selective search for object recognition[J]. *International Journal of Computer Vision*, 2013, 104(2): 154-171.
- [14] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 779-788.
- [15] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 6517-6525.
- [16] REDMON J, FARHADI A. Yolov3: an incremental improvement[J]. *arXiv Preprint, arXiv: 1804.02767*, 2018.
- [17] BOCHKOVSKIY A, WANG C Y, LIAO H Y M, et al. Yolov4: optimal speed and accuracy of object detection[J]. *arXiv Preprint, arXiv: 2004.10934*, 2020.
- [18] 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: 1571-1580.
- [19] JJOCHER G, NISHIMURA K, MINEEVA T, et al. YOLO v5[EB/OL]. (2022-06-26)[2021-06-02]. <https://github.com/ultralytics/yolov5>.
- [20] CHUYI L, LULU L, HONGLIANG J, et al. YOLO v6: a single-stage object detection framework for industrial applications [J]. *arXiv Preprint, arXiv:2209.02976*, 2022.
- [21] WANG C Y, BOCHKOVSKIY A, et al. Yolov7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [J]. *arXiv Preprint, arXiv:2207.02696*, 2022.
- [22] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//*European Conference on Computer Vision*, 2016: 21-37.
- [23] FU C Y, LIU W, RANGA A, et al. DSSD: deconvolutional single shot detector[J]. *arXiv Preprint, arXiv:1701.06659*, 2017.
- [24] EVERINGHAM M, VAN G L, WILLIAMS C K, et al. The pascal visual object classes (voc) challenge[J]. *International Journal of Computer Vision*, 2010, 88(2): 303-338.
- [25] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context[C]//*European Conference on Computer Vision*, 2014: 740-755.