

基于 CBCW - YOLO v8 的猪只行为识别方法研究

仝志民¹ 徐天哲¹ 石传淼¹ 李盛章¹ 谢秋菊² 荣丽红¹

(1. 青岛农业大学机电工程学院, 青岛 266109; 2. 东北农业大学电气与信息学院, 哈尔滨 150030)

摘要: 随着现代生猪养殖业快速发展,对猪只行为精准识别需求日益增长。针对猪只行为多样性、特征相似性、相互遮挡和堆积等问题,提出一种基于改进 YOLO v8 模型的猪只行为识别方法。首先,引入 ConvNeXt V2 作为主干特征提取网络,以增强对检测目标的语义信息提取能力。其次,在特征融合网络中添加加权双向特征金字塔网络 (BiFPN),强化模型特征融合能力。此外,结合上采样算子 CARAFE,进一步提升模型在行为识别过程中特征提取能力。最后,使用 WIoUv3 作为损失函数,优化模型检测精度。经实验验证,改进后模型准确率、召回率、平均精度均值和 F1 值分别达到 89.6%、88.0%、91.9% 和 88.8%,与 TOOD、YOLO v7 和 YOLO v8 模型相比,平均精度均值分别提高 10.9、6.3、3.7 个百分点,显著提高猪只行为识别精度。消融实验表明,各项改进均对模型的识别性能有提升效果,ConvNeXt V2 主干特征提取网络对模型的提升效果最明显。综上所述,CBCW - YOLO v8 模型在猪只行为识别任务中展现出优良的综合性能,为猪只健康管理和疾病预警提供有力的技术支持。

关键词: 猪只行为识别; YOLO v8; 特征提取网络; CARAFE; WIoUv3; 目标检测

中图分类号: TP391

文献标识码: A

文章编号: 1000-1298(2025)02-0411-09

OSID:



Pig Behavior Recognition Based on CBCW - YOLO v8 Model

TONG Zhimin¹ XU Tianzhe¹ SHI Chuanmiao¹ LI Shengzhang¹ XIE Qiuju² RONG Lihong¹

(1. College of Mechanical and Electrical Engineering, Qingdao Agricultural University, Qingdao 266109, China

2. College of Electrical and Information, Northeast Agricultural University, Harbin 150030, China)

Abstract: With the rapid development of modern pig breeding industry, the demand for precise recognition of pig behaviors is increasing. Aiming to address the issues of diversity of pig behaviors, similarity of features, mutual occlusion and stacking, a pig behavior recognition method based on the improved YOLO v8 model was proposed. Firstly, the ConvNeXt V2 was introduced as the backbone feature extraction network to enhance the ability to extract semantic information of the detection target. Secondly, the bi-directional feature pyramid network (BiFPN) was added to the feature fusion network to enhance the feature fusion ability of the model. Thirdly, combined with the CARAFE up-sampling operator, the feature extraction ability of the model in the process of behavior recognition was further improved. Finally, the WIoUv3 was used as the loss function to optimize the detection accuracy of the model. The experimental results showed that the precision rate, recall rate, mean average precision and F1 value of the improved model reached 89.6%, 88.0%, 91.9% and 88.8%, respectively. Compared with TOOD, YOLO v7 and YOLO v8 models, the mean average precision was increased by 10.9, 6.3 and 3.7 percentage points, respectively, which significantly improved the accuracy of pig behavior recognition. The ablation experiments showed that all the improvements improved the recognition performance of the model, and the ConvNeXt V2 backbone feature extraction network had the most obvious improvement effect on the model. In summary, the CBCW - YOLO v8 model demonstrated excellent overall performance in pig behavior recognition tasks and provided powerful technical support for pig health management and disease early warning.

Key words: pig behavior recognition; YOLO v8; feature extraction network; CARAFE; WIoUv3; target detection

收稿日期: 2024-09-30 修回日期: 2024-10-24

基金项目: 青岛农业大学博士启动基金项目(1121005)、农业农村部华南热带智慧农业技术重点实验室开放课题(HNZHNY-KFKT-202206)、山东省科技型中小企业创新能力提升工程项目(2023TSGC0741)和国家自然科学基金面上项目(32372934)

作者简介: 仝志民(1977—),男,教授,博士,主要从事农业信息传感与物联网技术研究,E-mail: leicahit@hotmail.com

通信作者: 荣丽红(1976—),女,副教授,博士,主要从事农业信息系统、复杂系统故障检测及容错控制研究,E-mail: ronglh@qau.edu.cn

0 引言

随着生猪养殖从小规模散养转向现代工业化养殖,对猪只健康管理的要求日益提高。研究表明,动物行为可有效反映其健康状态和环境适应能力^[1-2]。因此,准确获取猪只行为信息,分析健康状态,是猪只健康管理的关键。

目前,动物行为识别方法主要分为两大类:传感器监测^[3-4]和计算机视觉识别^[5-6]。传感器监测需要将传感器安装在动物身上,这可能导致动物应激反应和设备磨损。相比之下,计算机视觉识别无需直接接触动物,不会对其日常生活造成干扰,因此在现代工业化养殖环境下更具应用前景。

传感器监测通过获取各种数据来监测动物行为^[7-9]。深度学习作为计算机视觉的一个重要技术,用于训练和优化计算机视觉模型,处理计算机视觉任务。在计算机视觉识别领域,深度学习因其出色的泛用性和精准度得到广泛应用。尽管目前研究在猪只单一行为识别方面取得了进展^[10-13],但实际生产中往往需要识别多种行为。在多行为识别方面,目前国内外学者已经利用传感器和深度学习对猪只的多种行为进行一定研究^[14-19],但传感器监测对猪只生活造成影响使其难以满足养殖需求,深度学习因无需直接接触,更适合当前需要。

尽管已有研究证实深度学习在猪只行为检测中的可行性,但现有方法仍存在猪只行为多样、特征相似、喜堆积、易遮挡的问题,导致检测困难、精度低。针对以上问题,本研究提出一种基于 CBCW-YOLO v8 的猪只行为识别方法,利用 ConvNeXt V2 特征提取网络中特有的大卷积核,同时在特征融合层中替换 BiFPN 结构,去除冗余特征信息,使网络模型更侧重学习关键特征信息,提升网络学习能力。结合 CARAFE 上采样算子,聚合上下文信息对特定内容进行感知;使用 WIoUv3 损失函数,更多地关注普通质量的样本,以期提高网络模型泛化能力和整体性能。

1 实验数据获取

1.1 图像采集

视频数据采集在山东省东营市某养殖场猪舍内

进行。采用海康威视摄像头,以向下倾斜 60° 进行拍摄。此角度设置旨在最大限度地捕捉到猪只行为特征,减少由于猪只相互遮挡而造成的数据损失。考虑到夜间和恶劣天气条件可能导致环境光线不足,影响数据清晰度和准确性,采集时间定在晴天 08:00—16:00 之间,选取行为特征清晰的视频片段进行特征提取。

重点观察猪只 5 种基本行为:站立、行走、侧卧、趴卧和进食,这些行为能够有效反映猪只健康状况,其具体描述见表 1,行为示例见图 1。使用 OpenCV 计算机视觉库进行视频关键帧的截取。分析发现,在无外界干扰的情况下,猪只行动较慢,动作差异较小。为避免相邻帧间产生大量冗余图像和重复样本,将取帧间隔设置为 5 s,增加图像间猪只行为差异性,提高后续行为识别准确性和效率。

表 1 猪只行为描述
Tab. 1 Description of pig behavior

行为	行为描述	标签
站立	猪只保持前双腿、后双腿或四腿呈现平行姿态,身体不接触地面	stand
行走	猪只前腿和后腿同时呈现交替姿态,腿部向上抬起,脚掌离开地面	walk
侧卧	猪只以躯干一侧为支撑,躯干一侧完全接触地面,四肢伸直	side
趴卧	猪只四肢和腹部同时接触地面,四肢呈弯曲姿态,头部向前伸展	lie
进食	猪只走到固定食槽,颈部向前伸展,头部伸到食槽中	eat

1.2 数据集构建

首先对视频截取的图像进行人工筛选,去除模糊和相似图像,利用优质图像构建数据集。使用图像标注工具 LabelImg 对数据集中 5 种行为进行标注,标准为猪只所能容纳的最小矩形框,若猪只被遮挡超过 50% 则不进行标注。数据集构建完成后,以比例 8:1:1 划分为训练集、验证集和测试集,并采取翻转、旋转、增强亮度和增加高斯噪声等方法进行图像增强。增强后的图像包含更多背景与多尺度信息,有助于加快模型收敛速度,提升模型泛化能力。

经过上述预处理,共生成 1 944 幅图像用于猪



图 1 猪只行为示例图

Fig. 1 Example diagrams of pig behavior

只行为识别,行为标注数量共 5 316 个。这些图像涵盖 5 种不同的行为类别,各类别的标注数量存在差异,具体标注数量见表 2。由于猪只大部分时间处于站立状态,因此站立行为标注数量最多。行走

表 2 猪只行为标注数据
Tab. 2 Pig behavior annotation data

行为类别标签	行为数量
stand	2 235
walk	1 253
side	862
lie	647
eat	319

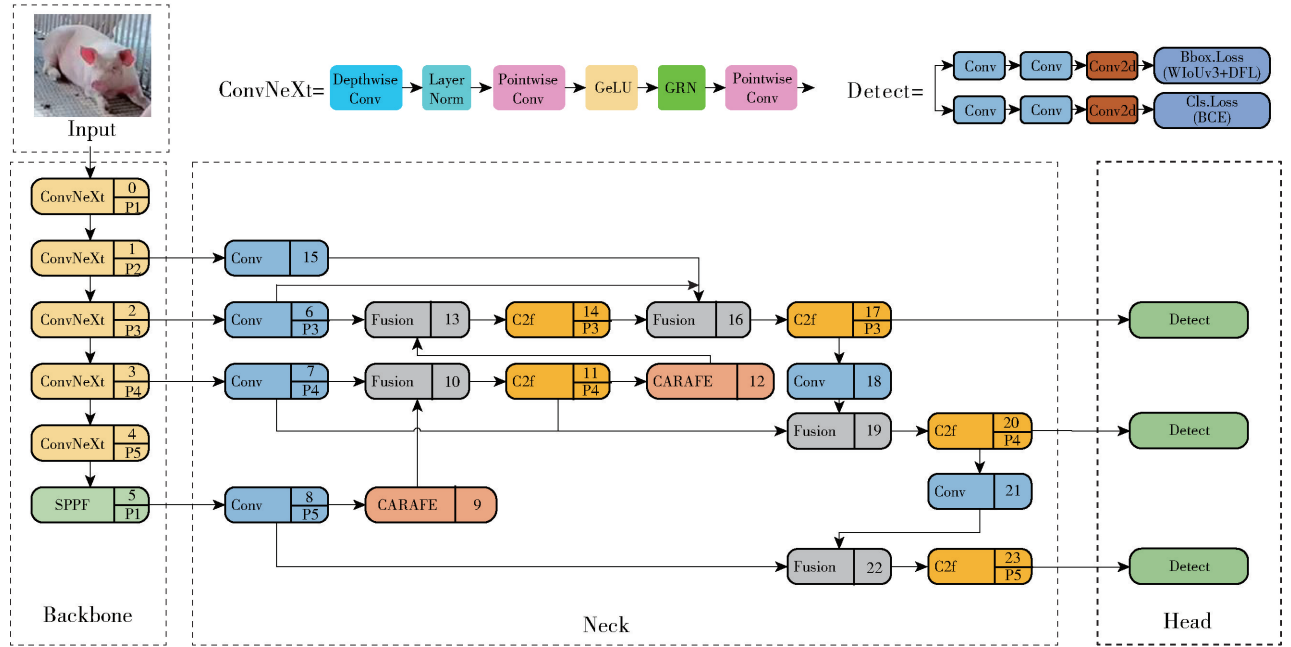


图 2 CBCW - YOLO v8 网络模型结构图
Fig. 2 CBCW - YOLO v8 network model structure diagram

2.2 YOLO v8 模型改进

2.2.1 ConvNeXt V2 特征提取网络

为了满足猪只健康管理的高精度需求,不仅要准确识别猪只行为特征,还需面对猪只堆积和遮挡情况,通过实验发现,原始特征提取网络面对上述情况出现特征提取不完整的问题,使得部分遮挡的猪只行为无法识别。为此,引入 ConvNeXt V2^[20] 替换原始特征提取网络。ConvNeXt V2 是基于 ConvNeXt^[21] 的改进模型,它采用大卷积核的深度可分离卷积进行逐通道卷积操作。大卷积核提供更广阔的感受野,可以审视更为宽广区域的特征图,从而获得更好的全局特征表示。为了解决较高的参数量和浮点运算量,运用深度可分离卷积。深度可分离卷积是输入通道和输出通道以及分组数都相同的卷积形式,这种卷积形式在提高特征提取能力的同时,可显著减少参数量。此外,ConvNeXt V2 还引入

和躺卧行为持续时间相近,标注数量也相似。进食行为持续时间较短且频率较低,标注数量最少。各类行为标注数量与猪只日常行为分布情况相符,这有助于模型更好地学习和识别猪只行为特征。

2 研究方法

2.1 改进 YOLO v8 模型结构

为了在 YOLO v8 模型基础上进一步提高猪只行为识别精度,引入 4 个模块进行改进:ConvNeXt V2 特征提取网络、加权双向特征金字塔网络 (BiFPN)、CARAFE 上采样和 WIoUv3 损失函数。改进后 CBCW - YOLO v8 模型结构如图 2 所示。

GeLU 激活函数替代常见的 ReLU 激活函数,并引入全局响应归一化 (Global response normalization, GRN),增强通道对比度和选择性。

首先进行特征提取,采用 7 × 7 大卷积核的设计,利用层归一化使梯度更稳定。随后,通过 1 × 1 卷积将特征维度提升 4 倍,以保留更多特征信息。接着利用 GeLU 激活函数和全局响应归一化将维度还原,最后进行特征融合并输出。ConvNeXt V2 结构如图 3 所示。通过这些改进,ConvNeXt V2 特征提取网络能够更有效地处理猪只行为识别中的挑战,为 YOLO v8 模型提供强大的特征提取能力,从而提高模型在猪只行为识别任务中的性能。

2.2.2 加权双向特征金字塔网络

在识别图像过程中,由于猪只处于运动状态,与摄像头的距离会发生变化,造成图像中存在尺寸较

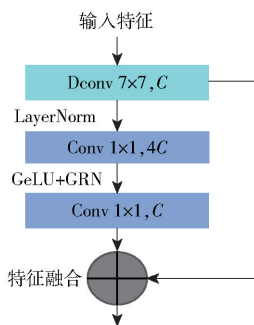


图3 ConvNeXt V2 结构图

Fig. 3 ConvNeXt V2 structure diagram

小的小目标。针对小目标特征图分辨率较低,导致猪只图像细节丢失,造成误检和漏检。为此,需提高模型的特征融合能力,实现对不同尺度的特征有效融合。然而,传统的特征金字塔网络 (Feature pyramid networks, FPN) 通常只采用自上而下的路径,这限制了特征的融合效率。为了克服这一限制,引入加权双向特征金字塔网络 (Bi-directional feature pyramid network, BiFPN) [22]。

BiFPN 结构是对 PA-FPN 结构的改进,模型结构如图 4 所示。该结构在同一水平的原始输入与输出节点之间增加一条额外的边,形成多次双向路径,从上向下和从下向上传递特征信息,保留更多上下文信息,提升特征信息传递效率,实现更高层次的特征融合。在猪只行为识别过程中,采集的猪只图像大小不一、行为姿态多样,使得不同尺度特征的重要性难以准确区分。BiFPN 通过引入双向特征融合和动态网络设计,使特征融合更全面且适应性更强,有效利用含有多个小目标特征的中低维特征图,提高小目标检测精度和重要特征的代表能力,从而提升整体检测精度与速度。

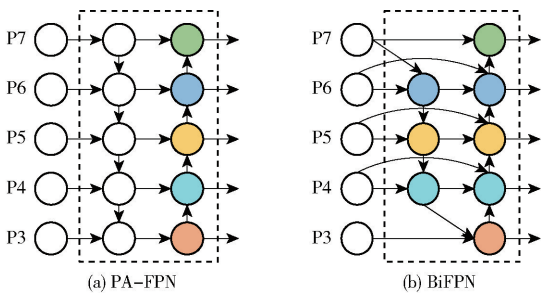


图4 模型结构图

Fig. 4 Model structure diagrams

2.2.3 CARAFE 上采样

在目标检测任务中,增加图像或特征的尺寸和空间分辨率的上采样步骤也十分关键。YOLO v8 一般采用最近邻差值方法,提高检测速度的同时,由于感知域很小,插值结果不够平滑,容易产生锯齿状边缘,图像质量损失较大,降低检测精度[23]。面对猪只行为的特征相似问题,导致输出图像的质量较差。为此,采用具有大感受野的上采样算子 CARAFE[24]。其结构如图 5 所示,分为上采样核预测模块和特征重组模块。在上采样核预测模块中,首先给定一个形状为 $H \times W \times C$ 的输入特征图,经过通道压缩形状变为 $H \times W \times C_m$,目的是减少后续计算量,经过内容预测得到形状为 $\sigma H \times \sigma W \times K_{up}^2$ 的上采样核,经过归一化处理得到权重和为 1 的卷积核,保证特征图幅度不受影响。在特征重组模块中,将输出特征图中信息映射到输入特征图,取出以之为中心的 $K_{up} \times K_{up}$ 的区域,与上采样核进行点积,得到形状为 $\sigma H \times \sigma W \times C$ 的输出特征图[24]。同一位置的不同通道共享同一个上采样核。这种方法使得 CARAFE 能够根据局部区域的内容动态调整上采样过程,更好地保留图像的语义信息和细节。

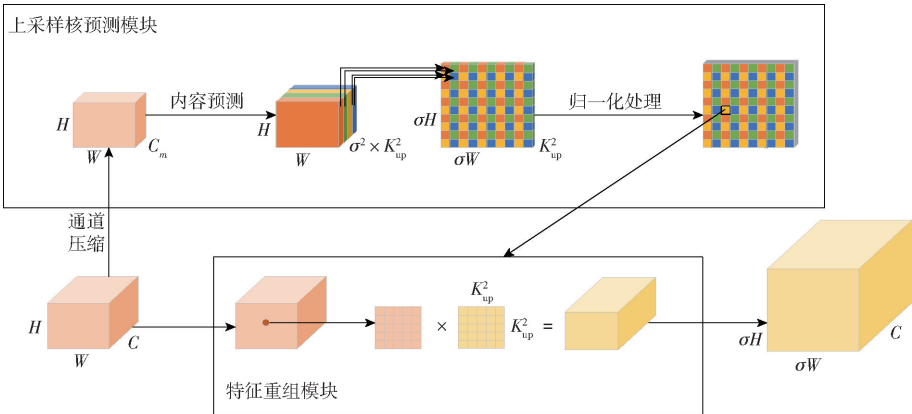


图5 CARAFE 特征融合结构图

Fig. 5 CARAFE feature fusion structure diagram

2.2.4 WIoUv3 损失函数

一个良好定义的损失函数可显著提升模型的检测精度。损失函数参数示意图如图 6 所示,其中 Real box 表示真实框或标签框, Predicted box 表示算

法给出的预测框。 $(b_{cx}^{gt}, b_{cy}^{gt})$ 表示真实框中心点坐标, (b_{cx}, b_{cy}) 表示预测框中心点坐标。 w^{gt} 、 h^{gt} 分别表示真实框宽和高, w 、 h 分别表示预测框宽和高。 c_w 和 c_h 分别表示预测框和真实框形成的最小包围方

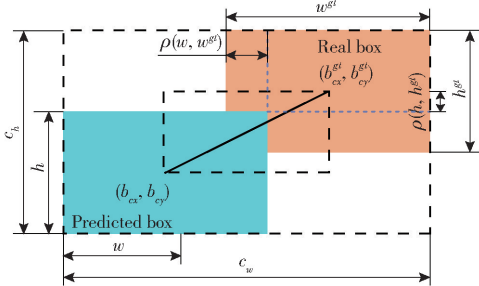


图 6 损失函数参数示意图

Fig.6 Schematic diagram of loss function parameters

框宽和高, $\rho(w, w^g)$ 和 $\rho(h, h^g)$ 分别表示真实框与预测框之间的宽度和高度欧氏距离。

当预测框与真实框具有相同的宽高比,但宽度和高度不相同,原始损失函数都无法优化。为此,使用 WIoUv3^[25] 作为损失函数, WIoUv3 可以使预测框精确定位到猪只边缘,提高模型定位能力,公式为

$$L_{WIoUv3} = \gamma R_{WIoU} L_{IoU} \tag{1}$$

其中
$$\gamma = \frac{\beta}{\delta \alpha^{\beta - \delta}} \tag{2}$$

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \in [0, +\infty) \tag{3}$$

式中 L_{WIoUv3} ——使用 WIoUv3 损失函数计算的边界框回归的损失值

γ ——非单调聚焦因子

R_{WIoU} ——惩罚项 β ——离群度

δ ——梯度增益 α ——学习参数

L_{IoU} ——基本边界框损失值

L_{IoU}^* ——单调聚焦系数

$\bar{L}_{IoU}$ ——运算平均值

β 值越大表示锚点框质量越低。 γ 值越小, 分配给

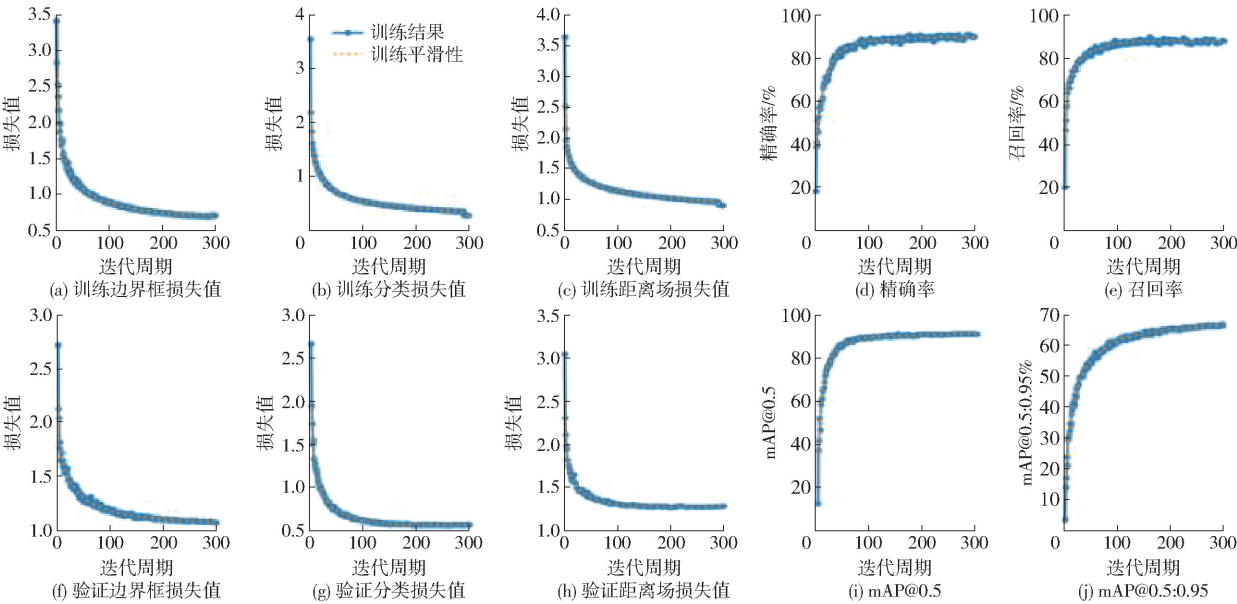


图 7 训练效果变化曲线

Fig.7 Training effect change curves

它的梯度增益越小,从而减少低质量锚点框产生的有害梯度。WIoUv3 损失函数通过动态非单调聚焦机制,能有效避免对标注质量较低样本的过度惩罚。

3 实验与结果分析

3.1 实验设置

训练和测试所用的计算机配置为: CPU 型号为 Intel(R) Core(TM) i5-10400, GPU 型号为 NVIDIA GeForce RTX3060。程序编译环境为 Window 11 系统,使用 Pytorch 1.11.0、Python 3.9 和 CUDA 12.1,主要参数设置见表 3。

表 3 参数设置

Tab.3 Parameter settings

参数	数值
训练图像尺寸/(像素×像素×字节)	640×640×3
批大小	8
迭代运行次数	300
初始学习率	0.01
周期学习率	0.01
学习率动量	0.973
权重衰减系数	0.000 5
禁用马赛克增强	10

3.2 性能评价指标

采用目标检测算法常用的 4 个性能评价指标: 精确率(Precision, P)、召回率(Recall, R)、平均精度均值(Mean average precision, mAP)和 F1 值来评估模型性能。

3.3 结果与分析

3.3.1 训练性能分析

为验证 CBCW-YOLO v8 模型在训练过程中的性能,通过观察损失函数值和性能指标的变化趋势来评估训练效果。模型训练效果变化曲线如图 7 所

示。其中边界框损失函数用于计算预测边界框与真实标注框之间的差异,分类损失函数用于计算预测类别与真实类别之间的差异,距离场损失函数用于处理类别不平衡问题。

由图 7a、7f 可知,随着训练的深入,边界框损失值逐渐减小,表明模型在边界框定位精度上持续改进。由图 7b、7g 可知,分类损失值下降趋势表明模型在类别判别能力上逐渐增强。由图 7c、7h 可知,确保模型对少数类别的样本也具有较高的识别能力。其他性能指标随着训练的进程而逐渐提高,直至达到一个稳定值。这一平稳状态的曲线表明,模型训练过程中没有出现拟合或欠拟合现象,在训练过程中学习到了有效的特征表示,显示出模型在完成识别任务时具备良好的稳定性和泛化能力。

3.3.2 损失函数对比实验

为验证 WIoUv3 损失函数对模型检测性能的提升效果,分析不同损失函数对模型性能的影响。通过将 YOLO v8 模型中的损失函数替换为其他类型进行对比实验,实验结果如表 4 所示。其中 YOLO v8 模型使用的是 CIoU 损失函数。

由表 4 可知,与 CIoU 相比,使用 DIoU 和 GIoU

表 4 损失函数对比实验结果

Tab.4 Loss function comparison experiment %				
损失函数	<i>P</i>	<i>R</i>	F1 值	mAP
CIoU	88.9	88.2	88.6	91.6
DIoU	88.9	86.8	87.8	91.3
GIoU	87.1	87.6	87.4	91.1
SIoU	90.2	86.8	88.5	92.0
EIoU	89.5	86.1	87.8	91.1
WIoUv3	89.6	88.0	88.8	91.9

后,模型的 *P*、*R* 和 mAP 均有所下降。使用 SIoU 和 EIoU 后,精确率分别提升 1.3、0.6 个百分点,召回率分别下降 1.4、2.1 个百分点,下降幅度较大。使用 SIoU 和 WIoUv3 后,mAP 均有提升。除了采用 WIoUv3 损失函数,其余模型 F1 值均下降。WIoUv3 模型的 *P* 和 mAP 分别提高 0.7、0.3 个百分点,WIoUv3 采用合理的梯度增益分配策略,有利于提高算法检测性能,减少错误预测。

3.3.3 行为识别实验

为验证本文模型在不同行为上的提升效果,分析对于不同行为识别效果的差异性和可能的原因,对原始模型与 CBCW – YOLO v8 模型进行数据对比,实验结果如表 5 所示。

表 5 行为识别实验结果

Tab.5 Experimental results of behavior recognition							%
mAP				<i>P</i>	<i>R</i>	F1 值	mAP
行走	侧卧	趴卧	进食				
95.8	82.1	80.7	88.5	84.6	81.9	83.2	88.2
96.0	89.5	86.1	91.1	89.6	88.0	88.8	91.9

由表 5 可知,CBCW – YOLO v8 模型对猪只行为识别的 *P*、*R* 和 mAP 达到 89.6%、88.0% 和 91.9%,实现对猪只行为准确识别。具体来说,模型对站立和行走行为的识别效果最好,其 mAP 达到 97% 和 96%,由于侧卧和趴卧行为特征相似,进食行为较少,导致模型对这 3 种行为的识别精度较低。引入 ConvNeXt V2 特征提取网络,可以有效改善猪只堆积和遮挡造成的目标特征提取不完整问题,提升检测精度。使用 BiFPN,提升不同尺度的特征融合能力,改善由于猪只进食行为较少导致的识别准确率低的问题。结合 CARAFE,采用大感受野改善侧卧和趴卧行为由于特征相似导致的识别精度低的问题。使用 WIoUv3 避免模型对标注质量较低样本的过度惩罚。实验证明 CBCW – YOLO v8 满足模型改进的预期目标,总体表现良好。

3.3.4 不同模型对比实验

为验证本文模型的性能优势,首先对 5 种不同尺寸 YOLO v8 模型进行对比实验,实验结果如表 6

所示。对于猪只行为识别,更高的参数量和更复杂的模型结构不一定具有更显著的检测效果。相反,由于模型复杂度的提高,模型计算时间也会大幅增加。综合考虑下,虽然 YOLO v8s 拥有最高的检测精度,相较于 YOLO v8n,参数量增加 8.09×10^6 ,浮点运算量增加 2.03×10^{10} ,*P* 减少 0.4 个百分点,模型深度增加和计算量增大并未对模型检测性能带来实质性的提升,故选择 YOLO v8n 作为本研究对比模型。

表 6 YOLO v8 不同尺寸模型对比实验结果

Tab.6 Comparative experiments of YOLO v8 models with different sizes					
模型	<i>P</i> /%	<i>R</i> /%	mAP/%	参数量	浮点运算量
YOLO v8n	85.6	83.4	88.8	3.01×10^6	8.10×10^9
YOLO v8s	85.2	83.9	89.2	1.11×10^7	2.84×10^{10}
YOLO v8m	85.9	77.8	84.7	2.58×10^7	7.87×10^{10}
YOLO v8l	86.1	77.1	83.6	4.36×10^7	1.65×10^{11}
YOLO v8x	88.3	68.0	78.4	6.81×10^7	2.57×10^{11}

为进一步验证本文模型实用性,在实验环境和数据集保持不变的前提下,使用 CBCW - YOLO v8 模型与其他几种主流模型对猪只行为进行识别,实验结果如表 7 所示。

表 7 不同模型对比实验结果

Tab.7 Comparison experiments of different models

模型	<i>P</i>	<i>R</i>	F1 值	mAP
Faster R - CNN		58.6		81.9
Cascade R - CNN		57.7		81.9
TOOD		63.0		81.0
YOLO v7	78.6	81.7	80.1	85.6
YOLO v8n	85.6	83.4	84.4	88.8
CBCW - YOLO v8	89.6	88.0	88.8	91.9

由表 7 可知,CBCW - YOLO v8 模型在检测性能上显著优于 Faster R - CNN、Cascade R - CNN、TOOD、YOLO v7 和 YOLO v8n 等模型。作为主流的双阶段目标检测器,Faster R - CNN 和 Cascade R - CNN 的性能相对较低,其中 Cascade R - CNN 的性能最差,mAP 相同。尽管单阶段目标检测器 TOOD 有所改进,*R* 最高提高 4.4 个百分点,但 mAP 降低 0.9 个百分点。YOLO 系列模型的性能指标显

著提升,其中 YOLO v8n 的性能高于 YOLO v7,而 CBCW - YOLO v8 模型则优于 YOLO v8n,*P* 分别提高 11、4 个百分点,F1 值分别提高 8.7、4.4 个百分点。相较于其他算法,CBCW - YOLO v8 模型的 *R* 分别提高 29.4、30.3、25、6.3、4.6 个百分点,mAP 分别提高 10、10、10.9、6.3、3.1 个百分点。综上所述,CBCW - YOLO v8 模型在猪只行为识别任务中展现了卓越的性能,其高精确率、高召回率以及高 F1 值表明该模型具备更强的识别性能和适应性。这些结果证实 CBCW - YOLO v8 模型在处理猪只行为识别中的复杂性和多样性方面的优势,为猪只健康管理提供了一种有效的技术手段。

3.3.5 消融实验

为全面探究 CBCW - YOLO v8 模型中各组件对检测性能的具体贡献,进行一系列消融实验。在保持数据集和训练参数不变的情况下,将各个改进部分作为变量进行控制,以确定其对最终性能的影响。基于原始 YOLO v8 模型,依次加入 ConvNeXt V2、BiFPN、CARAFE 和 WIoUv3。对所有改进进行组合测试,最终结果见表 8,“√”表示使用此改进,“×”表示未使用此改进。

表 8 消融实验结果

Tab.8 Results of ablation experiment

网络模型	ConvNeXt V2	BiFPN	CARAFE	WIoUv3	<i>P</i>	<i>R</i>	F1 值	mAP
模型 1	×	×	×	×	84.6	81.9	83.2	88.2
模型 2	√	×	×	×	89.4	86.7	88.0	91.6
模型 3	×	√	×	×	82.2	80.5	81.3	85.8
模型 4	×	×	√	×	84.9	82.8	83.8	88.8
模型 5	√	√	×	×	89.3	87.7	88.4	91.4
模型 6	√	×	√	×	89.4	87.8	88.6	91.5
模型 7	√	√	√	×	88.9	88.2	88.5	91.6
模型 8	√	√	√	√	89.6	88.0	88.8	91.9

由表 8 可知,模型 1 为原始 YOLO v8 算法,以它为基准,其 *P* 为 84.6%,*R* 为 81.9%,mAP 为 88.2%,F1 值为 83.2%。模型 2 在 YOLO v8 的基础上,单独引入 ConvNeXt V2,*P*、*R*、mAP、F1 值分别提高 4.8、4.8、3.4、4.8 个百分点,表明网络能够提取更丰富的上下文语义信息和全局特征,提升检测精度。模型 3 在 YOLO v8 的基础上,单独添加 BiFPN 结构,但数值有所下降,说明该结构不宜单独使用。模型 4 在 YOLO v8 的基础上,单独结合 CARAFE 上采样算子,*P*、*R*、mAP、F1 值分别提高 0.3、0.9、0.6、0.6 个百分点,表明 CARAFE 上采样算子能扩大感知域,提升模型对边缘特征的提取能力。模型 5 结合 ConvNeXt V2 和 BiFPN,*R* 提升 1 个

百分点,*P* 和 mAP 基本保持不变,表明两者结合有效降低模型复杂度。模型 6 结合 ConvNeXt V2 和 CARAFE,*R* 和 F1 值提高 1.1、0.6 个百分点,*P*、mAP 基本保持不变。模型 7 结合 ConvNeXt V2、BiFPN 和 CARAFE,*P*、*R*、mAP、F1 值分别提高 4.3、6.3、3.4、5.3 个百分点,表明模型整体精度显著提升。模型 8 是最终改进,结合 ConvNeXt V2、BiFPN、CARAFE 和 WIoUv3 4 个模块,性能最佳,*P*、*R*、mAP、F1 值达到 89.6%、88.0%、91.9% 和 88.8%,分别提升 5、6.1、3.7、5.6 个百分点,所有数值均优于原模型,模型性能达到最佳。如图 8 所示,消融实验结果直观地展示了各个改进部分对模型性能的具体贡献。改进模型漏检和误检显著减少,各行为的识别



图 8 消融实验效果

Fig. 8 Effect diagrams of ablation experiment

精度均有所上升,这些结果证明 CBCW - YOLO v8 模型在猪只行为识别任务中的优越性和改进部分的必要性。

4 结 论

(1)重点观察猪只 5 种行为,提出了一种猪只行为识别模型 CBCW - YOLO v8。为改善猪只堆积和遮挡问题,引入 ConvNeXt V2 有效捕获重要特

征信息。为解决猪只图像大小不一问题,添加 BiFPN 模块增强特征融合能力。为增强模型对侧卧、趴卧相似特征的识别能力,结合 CARAFE 模块更加关注图像细节。为解决标注质量存在差异的问题,使用 WIoUv3 损失函数。改进后的模型对站立、行走、侧卧、趴卧和进食行为平均精度均值分别达到 97%、96%、89.5%、86.1% 和 91.1%。

(2)CBCW - YOLO v8 模型对于猪只行为识别有较高的检测性能。消融实验表明,各项改进对模型识别性能均有提升效果,其中 ConvNeXt V2 主干网络提升效果最为显著。改进后模型 P 、 R 、 mAP 、 $F1$ 值分别提升 5、6.1、3.7、5.6 个百分点,相比原始 YOLO v8 模型有显著提升。对比实验结果表明,与 Faster R - CNN、Cascade R - CNN、TOOD、YOLO v7、YOLO v8n 模型相比, mAP 分别提高 10、10、10.9、6.3、3.1 个百分点,平均精度均值有显著提高,CBCW - YOLO v8 在猪只行为数据集上取得了较优效果。

参 考 文 献

[1] 张宏鸣,孙扬,赵春平,等. 反刍家畜典型行为监测与生理状况识别方法研究综述[J]. 农业机械学报,2023,54(3):1 - 21.
ZHANG Hongming, SUN Yang, ZHAO Chunping, et al. Review on typical behavior monitoring and physiological condition identification methods for ruminant livestock[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(3): 1 - 21. (in Chinese)

[2] MLUBA H S, ATIF O, LEE J, et al. Pattern mining-based pig behavior analysis for health and welfare monitoring[J]. Sensors, 2024, 24(7): 2185.

[3] ALGHARMDI S, ZHAO Z Q, HA S. Improved pig behavior analysis by optimizing window sizes for individual behaviors on acceleration and angular velocity data[J]. Journal of Animal Science, 2020, 100(11):137 - 144.

[4] ZHANG Z T, ZHANG H, HE Y X, et al. A review in the automatic detection of pigs behavior with sensors[J]. Journal of Sensors, 2022, 158: 86 - 103.

[5] 翟肇裕,张梓涵,徐焕良,等. YOLO 算法在动植物表型研究中应用综述[J]. 农业机械学报,2024,55(11):1 - 20.
ZHAI Zhaoyu, ZHANG Zihan, XU Huanliang, et al. Review of applying YOLO family algorithms to analyze animal and plant phenotype[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11):1 - 20. (in Chinese)

[6] OLIVEIRA D A B, PETEIRA L G R, BRESOLIN T, et al. A review of deep learning algorithms for computer vision systems in livestock[J]. Livestock Science, 2021, 253:104700.

[7] SALEH A, ZHAO Z Q, DONG S H, et al. Improved pig behavior analysis by optimizing window sizes for individual behaviors on acceleration and angular velocity data[J]. Journal of Animal Science, 2022, 100(11):293.

[8] 张春慧,宣传忠,于文波,等. 基于三轴加速度传感器的放牧羊只牧食行为研究[J]. 农业机械学报,2021,52(10):307 - 313.
ZHANG Chunhui, XUAN Chuanzhong, YU Wenbo, et al. Grazing behavior of herding sheep based on three-axis acceleration sensor[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(10):307 - 313. (in Chinese)

[9] 李丽华,刘志伟,赵学谦,等. 基于加速度传感器的本交笼种鸡个体行为监测与识别[J]. 农业机械学报,2019,50(12):247 - 254.
LI Lihua, LIU Zhiwei, ZHAO Xueqian, et al. Monitoring and identification of natural mating cage breeding chickens individual behavior based on acceleration sensor[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(12):247 - 254. (in Chinese)

[10] 李菊霞,李艳文,牛帆,等. 基于 YOLO v4 的猪只饮食行为检测方法[J]. 农业机械学报,2021,52(3):251 - 256.
LI Juxia, LI Yanwen, NIU Fan, et al. Pig diet behavior detection method based on YOLO v4[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(3): 251 - 256. (in Chinese)

[11] 杨秋妹,肖德琴,张根兴. 猪只饮水行为机器视觉自动识别[J]. 农业机械学报,2018,49(6):232 - 238.
YANG Qiumei, XIAO Deqin, ZHANG Genxing. Automatic pig drinking behavior recognition with machine vision[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(6):232 - 238. (in Chinese)

[12] 李丹,张凯锋,李行健,等. 基于 Mask R - CNN 的猪只爬跨行为识别[J]. 农业机械学报,2019,50(增刊):261 - 266,275.

- LI Dan, ZHANG Kaifeng, LI Xingjian, et al. Mounting behavior recognition for pigs based on Mask R - CNN [J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(Supp.): 261 - 266, 275. (in Chinese)
- [13] 庄晏榕, 余灵桦, 滕光辉, 等. 基于卷积神经网络的大白母猪发情行为识别方法研究 [J]. 农业机械学报, 2020, 51(增刊1): 364 - 370.
- ZHUANG Yanrong, YU Jionghua, TENG Guanghui, et al. Research on recognition method of estrous behavior of large white sows based on convolutional neural network [J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(Supp. 1): 364 - 370. (in Chinese)
- [14] TU S Q, ZENG Q T, LIANG Y, et al. Automated behavior recognition and tracking of group-housed pigs with an improved DeepSORT method [J]. Agriculture, 2022, 12: 1907.
- [15] 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)
- [16] 漆海霞, 冯发生, 尹选春, 等. 轻量级多场景群养猪只行为识别模型研究 [J]. 农业机械学报, 2024, 55(10): 306 - 317.
- QI Haixia, FENG Fasheng, YIN Xuanchun, et al. Research of lightweight multi-scene group pig behavior recognition model [J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(10): 306 - 317. (in Chinese)
- [17] 董力中, 孟祥宝, 潘明, 等. 基于姿态与时序特征的猪只行为识别方法 [J]. 农业工程学报, 2022, 38(5): 148 - 157.
- DONG Lizhong, MENG Xiangbao, PAN Ming, et al. Recognizing pig behavior on posture and temporal features using computer vision [J]. Transactions of the CSAE, 2022, 38(5): 148 - 157. (in Chinese)
- [18] 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)
- [19] 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)
- [20] WOO S, DEBNATH S, HU R, et al. V2: co-designing and scaling ConvNets with masked autoencoders [J]. arXiv Preprint, arXiv:2301.00808, 2023.
- [21] LIU Z, MAO H, WU C Y, et al. A convnet for the 2020s [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 11976 - 11986.
- [22] TAN M, PANG R, LE Q V. Efficientdet: scalable and efficient object detection [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781 - 10790.
- [23] MAZZINI D. Guided upsampling network for real-time semantic segmentation [J]. Springer Berlin Heidelberg, 2018(6): 19 - 26.
- [24] WANG J, CHEN K, XU R, et al. Carafe: content-aware reassembly of features [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 3007 - 3016.
- [25] TONG Z, CHEN Y, XU Z, et al. Wise - IoU: bounding box regression loss with dynamic focusing mechanism [J]. arXiv Preprint, arXiv:2301.10051, 2023.