

基于改进 YOLO v8s 的母猪分娩结束识别方法

祝志慧¹ 韩雨彤¹ 侯文烁¹ 黎 煊^{1,2} 徐学文³ 徐迪红^{1,2}

(1. 华中农业大学工学院, 武汉 430070; 2. 生猪健康养殖协同创新中心, 武汉 430070;

3. 华中农业大学动物科学技术学院, 武汉 430070)

摘要: 在母猪分娩自动监测技术中,一般采用识别最后一头仔猪产出为标志的间接方法来判断分娩结束,该方法在实际应用中存在实时性较差、识别精度不高和易受遮挡干扰等问题,难以满足生产需求。针对上述问题,以胎衣为直接识别目标,提出了一种基于改进 YOLO v8s 的分娩结束识别方法。通过引入 Focus 模块、SE 注意力机制、C2f-SCConv 模块及 BiFPN-P2 结构,对模型在特征表达、多尺度信息融合及轻量化设计等方面进行了优化。消融试验结果表明,SE 注意力机制与 C2f-SCConv 模块显著提升了小尺度、低对比度目标的检测精度,BiFPN-P2 结构在保证精度的同时有效降低了参数量和计算复杂度。最终,融合多模块的改进 YOLO v8s 模型精确率、召回率和 mAP 分别达到 98.1%、94.9% 和 98.8%,参数量仅 7.33×10^6 ,推理速度仍保持 61.71 f/s。与 NanoDet、RT-DETR、Faster R-CNN、YOLO v5 及原始 YOLO v8s 等主流模型相比,本方法在检测精度方面表现最优,同时展现了良好的实时性。提出了基于视频帧间隔和帧级计数器的时序判断机制,在真实分娩视频中实现了分娩结束事件的准确识别。在 25 f 间隔条件下,平均时间误差为 5.98 s,在 5 f 间隔条件下进一步缩小至 1.84 s,显著提高了识别时效性与准确性。研究表明,本研究提出的方法突破了依赖仔猪目标的间接推断模式,实现了由图像级检测向视频级时间节点识别的拓展。

关键词: 母猪分娩; 胎衣检测; YOLO v8s; 目标检测; 视频级识别; 智能养殖

中图分类号: S24 **文献标识码:** A **文章编号:** 1000-1298(2026)15-0046-10

OSID:



Method for Identifying End of Sow Farrowing Based on Improved YOLO v8s Model

Zhu Zhihui¹ Han Yutong¹ Hou Wenshuo¹ Li Xuan^{1,2} Xu Xuewen³ Xu Dihong^{1,2}

(1. College of Engineering, Huazhong Agricultural University, Wuhan 430070, China

2. Collaborative Innovation Center for Pig Health Breeding, Wuhan 430070, China

3. College of Animal Science and Technology, Huazhong Agricultural University, Wuhan 430070, China)

Abstract: In automatic monitoring of sow farrowing, the end of parturition is usually determined indirectly by identifying the birth of the last piglet. However, this approach suffers from poor real-time performance, low detection accuracy, and strong susceptibility to occlusion, making it unsuitable for production needs. To address these issues, a farrowing-end detection method that directly identified the placenta was proposed based on an improved YOLO v8s model. By incorporating the Focus module, SE attention mechanism, C2f-SCConv module, and BiFPN-P2 structure, the model was optimized in feature representation, multi-scale information fusion, and lightweight design. Ablation experiments showed that the SE attention mechanism and C2f-SCConv module significantly improved the detection accuracy of small-scale and low-contrast targets, while the BiFPN-P2 structure effectively reduced the number of parameters and computational complexity without sacrificing precision. The improved YOLO v8s model achieved 98.1% precision, 94.9% recall, and 98.8% mAP, with only 7.33×10^6 parameters and an inference speed of 61.71 f/s. Compared with mainstream models such as NanoDet, RT-DETR, Faster R-CNN, YOLO v5, and the original YOLO v8s, the proposed method achieved the

收稿日期: 2026-03-16 修回日期: 2026-04-03

基金项目: 湖北省支持种业高质量发展资金项目(HBZY2023B006-03)

作者简介: 祝志慧(1975—),女,副教授,博士,主要从事无损检测及其自动化、机器视觉研究,E-mail: zzh@mail.hzau.edu.cn

通信作者: 徐迪红(1974—),男,讲师,主要从事畜禽养殖智能装备开发与无损检测研究,E-mail: xudihong@mail.hzau.edu.cn

best detection accuracy while maintaining excellent real-time performance. Furthermore, a temporal judgment mechanism based on video frame intervals and frame-level counters was developed to accurately recognize the end of farrowing in real farrowing videos. Under a 25 f interval, the average time error was 5.98 s, which was further reduced to 1.84 s at a 5 f interval, significantly improving both timeliness and accuracy. The proposed method overcame the limitations of piglet-based indirect inference and extended from image-level detection to video-level temporal event recognition.

Key words: sow farrowing; placenta detection; YOLO v8s; object detection; video-level recognition; intelligent farming

0 引言

随着传感器技术和人工智能的发展,生猪养殖不断向规模化、自动化、智能化方向转变^[1-5]。母猪分娩作为养殖过程中的关键环节,其过程直接关系到仔猪的健康存活率及整个猪场的经济效益^[6]。尤其是在分娩后期阶段,及时识别出分娩结束,便于养殖人员快速介入处理,这对降低猪只的应激反应、减少感染风险具有重要意义^[7]。然而,目前大多数猪场仍依赖人工观察判断母猪是否完成分娩^[8],存在主观性强、劳动强度大、监测不及时等问题。因此,开发一种高精度、实时性的母猪分娩结束自动识别技术,对于推动畜牧养殖的智能升级、保障母猪健康、提高生产效率具有重要的意义与应用前景。

深度学习技术^[9-10]逐步应用至生猪养殖分娩监测领域^[11],近些年的相关研究主要可以概括为以下几类:基于视觉的目标检测与分割方法^[12]、基于行为/姿态的方法^[13]、基于非视觉传感器的多模态监测^[14]。

基于视觉的目标检测与分割方法,主要用于识别母猪、仔猪或群体^[15-17]。有研究通过分析母猪眼部变化,提出了基于眼部评分的异常分娩识别方法,可辅助判断产程是否延长^[18]。此类方法在目标检测上精度较高,但通常以最后一头仔猪产出为分娩结束标志^[19],属于间接推断,需要等待较长时间并易受遮挡影响。

基于行为/姿态分析的方法,可以通过母猪体位变化、活动模式等判断产程阶段。很多研究使用 CNN 架构来识别猪只站立、侧卧、趴卧等姿态^[20]。此外,为了实现更精细的行为追踪,有研究使用 SLEAP 等 pose-tracking 框架识别母猪身体关键点,从而跟踪运动 (locomotion) 并评估如跛行 (lameness) 等问题^[21]。然而,这类方法在识别分娩结束点时存在不确定性。母猪分娩后的行为可能有很大的个体差异,某些行为可能与分娩结束无直接关系,导致难以精确确定分娩真正完成时间。此外,行为分析通常需要高质量、持续的视频流,并依赖复杂

算法对行为模式建模,这在实际部署中可能面临挑战。

为了弥补视觉监测的局限性,研究人员探索了多种非视觉传感器的方法^[22]。但是,传感器在养殖环境中的鲁棒性和长期稳定性是需要考虑的问题。

综上,目前尚缺乏一种在实时性、准确性与部署简便性之间达到良好平衡的“直接判定分娩结束”的解决方案。当前判定分娩结束的难点主要体现在两方面:一方面,基于仔猪目标的间接判定需要等待较长的观察窗口(一般 30 ~ 60 min)以确认是否还有后续出生个体,严重影响实时性;另一方面,仔猪活动频繁、遮挡严重,导致检测不稳定并易产生漏检或误判。相比之下,胎衣在大多数分娩个案中会在产出最后一胎后短时间内排出且呈现一定的形态特征,因此将胎衣作为直接判定标志可缩短判断延迟并降低由遮挡导致的误判风险。本文以此为出发点,提出基于改进 YOLO v8s 的胎衣检测与视频级时序判定方法,以实现更快速可靠的分娩结束识别。

YOLO v8 (You Only Look Once version 8) 是一种端到端的单阶段目标检测模型,YOLO v8s 为该系列的轻量化版本,兼具较高实时性与较低计算开销,适合部署于计算资源受限的边缘设备^[23]。但基础 YOLO v8s 网络在胎衣检测任务中面临如下主要问题:浅层细节在标准下采样/卷积过程中被削弱,导致小目标信息丢失;低对比目标的特征响应弱,单纯卷积难以突出判别信号;多尺度融合对微小目标的响应不足;训练集中胎衣样本相对稀少或不均衡,易造成误检或漏检;实际部署要求模型既要精确又要轻量高速。基于此,本文提出以下改进方向:Focus 用于增强浅层空间信息保留,C2f-SCConv 提升空间感知与边缘刻画,BiFPN-P2 将高分辨率 P2 特征纳入加权融合以强化小目标信息,SE 注意力用于通道重标定以放大低对比目标信号,四尺度 Detect 分支则确保不同尺度目标均被覆盖。同时构建以视频帧为基础的监测体系,将视频帧序列中的检测结果进行聚类与序列判断,实现视频级别的目标事件识别机制,以期为母猪分娩过程的精准监测提供新思

路与新方案,促进生猪养殖业向自动化、智能化方向发展。

1 试验材料

1.1 识别对象

胎衣是胎猪与母体间连接的胎膜组织,通常在最后一胎仔猪分娩完成后 25 min 左右排出^[24],其形态呈不规则团块状,色泽较暗,质地湿润。与仔猪相比,胎衣目标静态性更强,便于模型提取判别特征。同时,系统一经识别到胎衣即可判定母猪产程结束,而无需等待更长时间,方便工作人员快速安排后续事宜。因此本研究以胎衣排出作为分娩结束的识别标志。然而,母猪胎衣检测面临若干实际挑战:①小尺度:胎衣在完整图像中仅占小区域,易在下采样过程中丢失信息。②遮挡与重叠:栏杆、母猪或活跃的仔猪常造成部分遮挡。③低对比度与背景相似:胎衣与产床地面或血迹颜色接近,单纯局部纹理难以区分。④形态多变:胎衣形状不规则,增加判别难度。⑤光照与模糊:夜间或逆光场景、运动模糊均影响检测稳定性。不同挑战场景下的胎衣如图 1 所示。

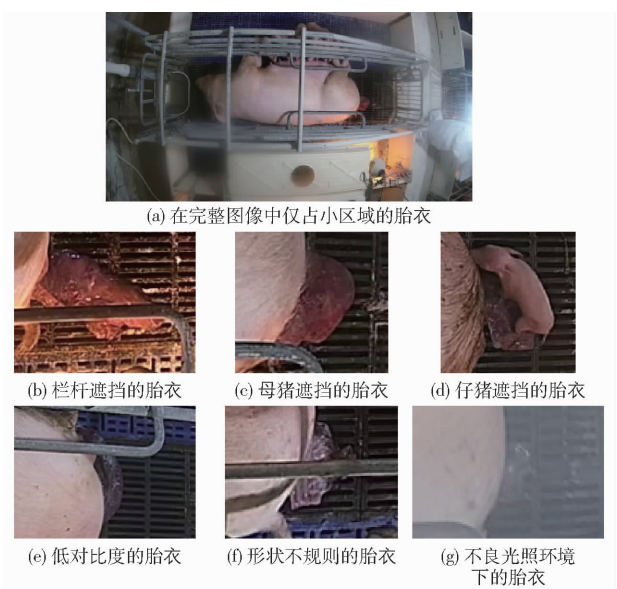


图 1 不同场景下的胎衣
Fig. 1 Placenta in different cases

1.2 视频采集

母猪分娩视频数据采集于武汉诸乐田源生态农业有限公司的生猪养殖场。视频采集设备主要有摄像头、海康威视录像机和机械硬盘,其中摄像头为海康威视 Hikvision DS-2CD3T46WDV3-L 型网络摄像机,该摄像机搭载 400 万像素镜头,F1.6 光圈,适用于光线较暗或无光且对画质有较高要求的拍摄场所,满足 24 h 连续拍摄的任务需求。

视频采集时间为 2024 年 2 月 29 日到 2024 年 9 月 15 日,摄像头 24 h 连续记录视频,最终获得了不同光照条件下来自 60 头母猪俯视角度的产期视频。

1.3 数据集制作

选取采集视频中包含母猪排出胎衣事件的视频段构建数据集。随机抽取 55 头母猪的视频段作为训练集与验证集,剩余 5 头母猪的视频段作为测试集。对所得的视频片段以 0.5 f/s 的速率抽取帧图像,确保数据量足够。对所得图像进行标签工作,最终共完成 5 120 幅图像标注。其中,训练集和验证集按 7:3 的比例进行划分,最终得到训练集(3 226 幅图像)、验证集(1 397 幅图像)和测试集(497 幅图像)3 部分数据。

1.4 数据增强

猪舍场景中光照条件复杂、遮挡因素较多,单一图像数据难以充分训练高性能检测模型,对于胎衣而言,其静态性虽带来了较高的可观测性,但同时也导致来自同一头母猪的数据间差异很小,因此该问题更为突出,而多种现实因素的制约使采集特征足够多样的真实胎衣数据无法实现。为提升模型的鲁棒性与泛化能力,使用图像增强技术来提升训练图像样本的多样性。

本研究选择以下 3 种 YOLO v8s 模型的数据增强方法:①图像旋转(Rotation):设置超参数 degrees 为 30,表示图像在训练过程中将以 $\pm 30^\circ$ 范围内的角度进行随机旋转,用于模拟摄像头拍摄角度的自然偏移,提高模型对不同视角下目标结构的适应能力。②图像混合(Mix Up):设置 mixup 为 0.2,即以 20% 的概率对两幅图像及其标签进行线性融合。该操作可有效增强模型对模糊边界与遮挡区域的识别能力,同时起到抑制过拟合的作用。③水平翻转(Flip Left-Right):设置 flipplr 为 0.2,即以 20% 的概率对图像进行左右镜像翻转,用以模拟胎衣目标在不同方位下的视觉变化,提升模型在实际部署时对不同朝向样本的识别准确性。YOLO v8s 在训练过程中自动调用上述增强策略对训练图像进行动态处理,无需手工扩展图像数据集,从而有效扩张了样本的分布范围,提升了模型对未知场景的泛化能力。数据增强效果如图 2 所示。



图 2 数据增强效果图

Fig. 2 Data enhancement effect diagrams

2 母猪分娩结束识别方法

2.1 改进的 YOLO v8s 目标识别算法

相比 YOLO v5、YOLO v7 等前代系列,YOLO v8 在结构设计和检测性能方面均有显著提升^[25],在保持较高检测精度的同时有效降低了计算量与模型参数量,具备更优的实时性与可部署性,其在生猪畜牧业的多个领域发挥了重要的作用,包括生猪个体追踪^[26]、行为识别^[27-29]、动物福利^[30]与健康评估^[31]等,其基础网络由 4 个部分组成:输入端(Input)、主干网络(Backbone)、特征融合网络(Neck)和检测头(Head)。

本文在 YOLO v8s 基础上进行了 4 方面结构优

化,从特征提取、信息融合与通道注意层面进行了针对性优化。多模块协同作用,以使改进后的网络在弱边界、小尺寸胎衣的识别中表现优越。

改进后的整体网络结构如图 3 所示,改进后的整体网络检测流程包括:①输入处理:对图像进行数据增强处理(旋转、混合、翻转等)。②Backbone 特征提取:依次经过 Focus、多个 C2f-SCConv 和 SPPF 模块,生成多层特征图。③Neck 多尺度融合:采用 BiFPN-P2 结构实现浅层细节与高层语义的加权融合,在输出前经过 SE 注意力模块增强关键通道。④Head 预测输出:由 4 个 Detect 分支分别对 P2~P5 特征进行目标检测,输出边界框、类别与置信度。

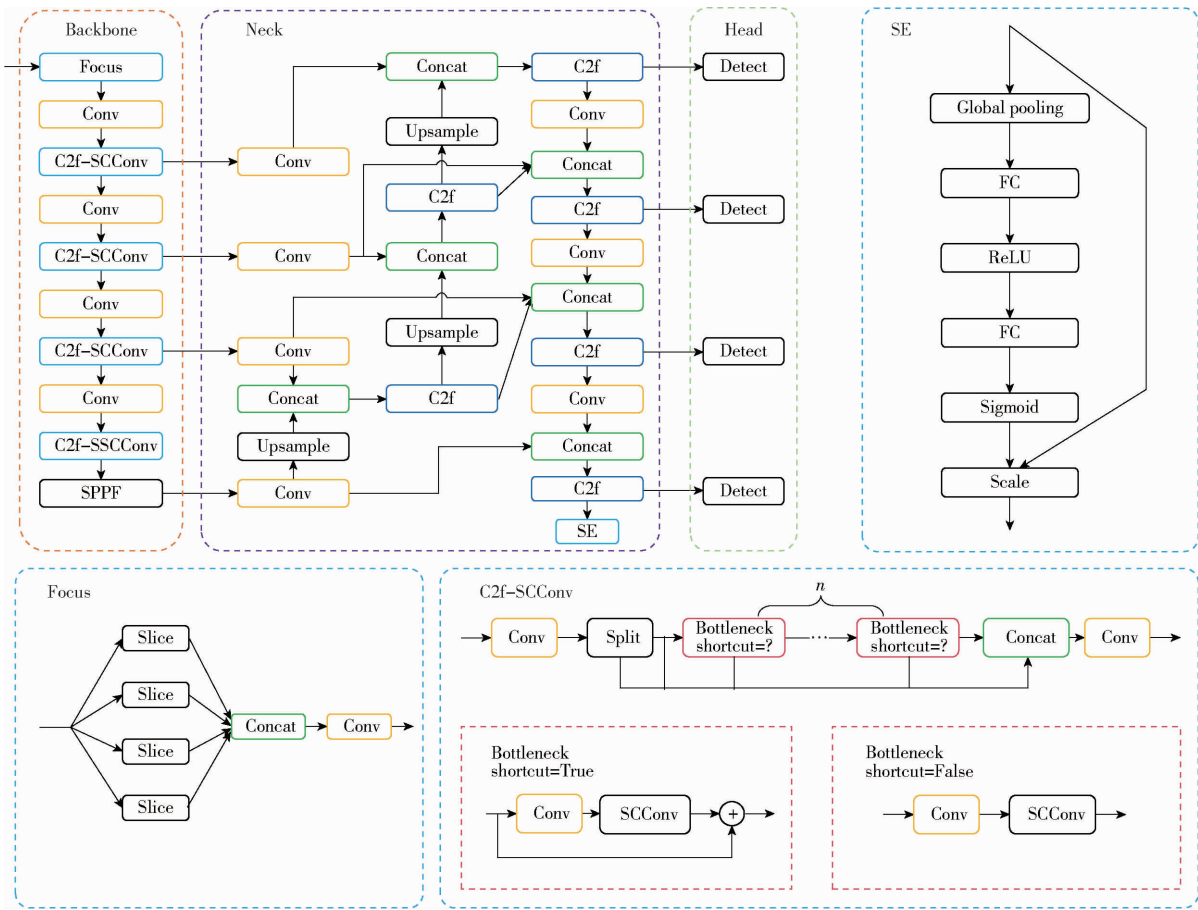


图 3 改进的 YOLO v8s 网络结构图

Fig. 3 Network structure of improved YOLO v8s

2.1.1 引入 Focus 模块

在目标检测网络中,高效的特征提取和下采样结构对于模型的检测精度与速度至关重要。母猪胎衣在图像中通常具有边缘模糊、颜色近似背景、形态多变等特点,若在网络初始阶段采用常规下采样操作,容易导致细粒度空间信息的过早丢失,从而影响后续检测性能。

为缓解上述问题,本文在网络输入端引入 Focus 模块,通过空间重排操作将原始图像的局部信息映

射至通道维度,在不显著增加计算量的前提下保留更多浅层细节特征。该结构有助于增强模型对小尺度目标及边缘纹理信息的感知能力,为后续特征提取提供更丰富的底层信息支撑,从而提升复杂背景下的目标可分性。

2.1.2 改造 C2f 为 C2f-SCConv 结构

在胎衣检测任务中,由于目标区域常表现为低对比度与弱纹理特征,模型需要具备更强的空间感知能力以提升判别能力。

本研究在 C2f(Cross stage partial with fusing) 模块基础上引入 SCConv(Spatial convolution)^[32], 并构建 C2f-SCConv 模块, 以增强空间建模能力并提升目标表征效果。改进后的 C2f-SCConv 模块不仅继承了 C2f 的残差结构优势, 同时具备更强的感知表达能力, 有助于提高胎衣区域与背景之间的区分度。

C2f-SCConv 模块分别应用于 Backbone 中的 3 个阶段(8 倍、16 倍和 32 倍下采样层), 以在保持参数量不显著增加的前提下, 有效提升中低层特征图对小目标轮廓的敏感度。

2.1.3 在 Neck 部分融合 BiFPN-P2 结构

YOLO v8s 原始结构采用 PANet 作为 Neck 模块, 其浅层特征传递能力仍存在一定不足, 对于小尺度目标的表达能力有限, 针对胎衣目标在图像中尺度较小且细节信息依赖较强的特点, 本文将 Neck 结构进行改进, 引入 BiFPN(Bidirectional feature pyramid network)^[33], 并添加 P2 分支, 构建 BiFPN-P2 特征融合网络结构。

该结构采用了大量 1×1 轻量级卷积核替代原 PAN 结构中的部分 3×3 卷积操作, 并通过权重共享与通道压缩策略减少了特征融合过程中的计算量。其中 BiFPN 的双向加权融合结构能够在保持语义完整性的前提下减少冗余特征通道, 从而进一步降低网络整体参数量, 使模型在性能提升的同时实现轻量化。P2 为 4 倍下采样特征图, 将其引入 Neck 结构对于检测结构模糊、对比度低的胎衣边缘将具有明显提升作用。改进后的 BiFPN-P2 结构如图 4 所示。

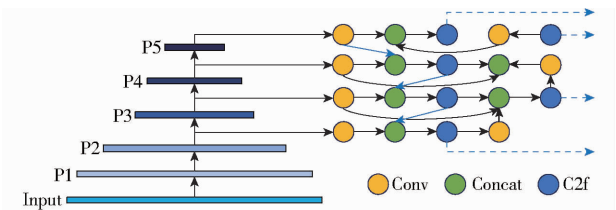


图 4 BiFPN-P2 模块结构图

Fig.4 BiFPN-P2 module block diagram

2.1.4 集成 SE 注意力模块

在复杂猪舍环境中, 胎衣目标常受到地面、血液及仔猪等背景干扰, 导致特征表达不充分甚至产生误检。为增强模型对关键特征的选择能力, 本文在特征融合后引入 SE(Squeeze-and-Excitation) 通道注意力机制。

SE 模块通过“压缩-激励”两个阶段对特征图进行建模: 首先通过全局平均池化(squeeze)提取每个通道的全局语义信息, 然后通过一个轻量的全连接瓶颈结构生成通道权重(excitation), 最终对原始特征进行加权调整^[34], 其结构如图 5 所示。

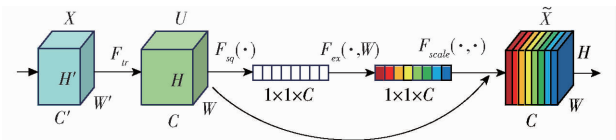


图 5 SE 模块结构图

Fig.5 SE module block diagram

本研究将 SE 模块部署在 BiFPN-P2 的输出端, 对融合后的 P2~P5 尺度的特征图分别进行通道注意力增强。该操作能够在不显著增加计算量的情况下, 有效突出与胎衣相关的语义特征, 并抑制无关背景, 从而提升整体检测鲁棒性, 是一种兼具性能提升与计算效率的有效改进策略。

除以上改进外, Backbone 末端还保留了 SPPF(空间金字塔池化快速版)以增强感受野, 检测头由原 YOLO v8s 的 3 分支扩展为 4 个 Detect 层, 对应 P2、P3、P4、P5 尺度特征图, 实现从浅到深的小目标到大目标的全覆盖检测。

综上所述, 针对胎衣检测任务中“小尺度、低对比度及边界模糊”等关键难点, 本文设计从输入信息保留、空间特征建模、多尺度融合及通道特征选择 4 个方面对 YOLO v8s 进行系统优化。各模块在结构上相互配合, 共同提升模型对复杂场景下弱特征目标的表达能力, 从而为后续检测性能提升提供结构基础。

2.2 模型训练与测试

2.2.1 试验平台与训练参数

本研究基于 YOLO v8s 框架, 结合 Visual Studio Code 深度学习平台进行模型训练, 具体环境配置如下: 计算机处理器采用 12th Gen Intel(R) Core(TM) i9-12900KF; 显卡为 NVIDIA GeForce RTX 4090; 系统环境为 Windows 11; Python 版本采用 3.9.19; CUDA 版本采用 12.8; 训练轮次设为 150 epoch; Batch size 设为 32。

2.2.2 评价指标

为了全面评估模型性能, 试验采用以下指标进行评估: 精确率、召回率、平均精确率、参数量、帧率(每秒处理帧数)。

2.3 视频级推理策略设计

在实际应用场景中, 母猪分娩的全过程以视频形式记录, 而非单帧图像。因此, 为确保本研究模型能够在真实养殖环境中准确、稳定地识别分娩结束时刻, 需基于 YOLO v8s 模型构建一套视频级目标检测策略, 将图像级预测结果提升至时间连续性分析水平。

本研究设计了一种基于时间帧间隔检测与计数阈值判定的视频识别策略, 整体流程如图 6 所示。

该策略以视频帧为基本处理单位,周期性地调用 YOLO v8s 模型对视频帧进行目标检测,并根据“胎

衣”类别 (placenta) 在连续帧中的出现情况判断是否有效识别,最终确定分娩是否完成。

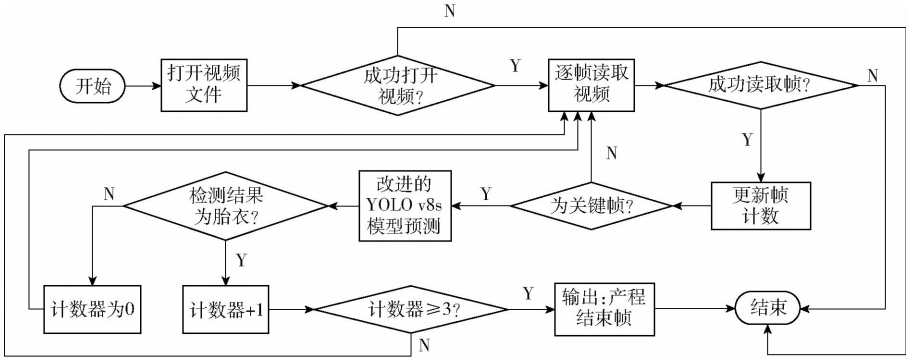


图 6 视频推理流程图

Fig. 6 Video-level inference flowchart

首先,对关键帧进行提取并控制推理间隔。考虑到视频中帧与帧之间信息高度重复,为降低计算负担、提高推理效率,设定固定检测间隔,并进行 interval 为 25 和 interval 为 5 两组试验,即每 25 f/5 f 抽取 1 帧作为关键帧进行目标检测,以确定实际应用中较为合适的参数设置。该操作可有效平衡识别精度与运行速度,在保障连续性判别的前提下提升整体系统实用性。之后对于每一帧输入图像调用训练好的 YOLO v8s 模型进行预测,判断图像中是否含有胎衣目标。

在分娩结束事件的判定策略中,本文采用“连续检测到 3 帧胎衣目标”作为连续帧数判定阈值。该参数设置过小容易受到单帧误检影响而产生错误判定结果,设置过大则会缺失一定程度上对于漏检的容错率,同时会引入时间延迟。相比之下,设置为 3 帧能够在稳定性与响应速度之间取得较优平衡。

从统计角度分析,假设模型在单帧图像上的正确识别率为 p ,则在独立帧条件下连续 3 帧均识别成功的概率为 p^3 ,而连续 3 帧均识别失败的概率为 $(1 - p)^3$,属于典型的二项分布事件。当 $p = 0.93$ 时,其 3 帧连续错误误率仅为 3.4×10^{-4} ,即识别失败率低于 0.04%,能够显著降低偶发误检导致的错误触发。因此,使用 3 帧作为触发阈值在统计意义上可以保证事件判定的稳定性与可靠性,满足智能分娩监测系统的应用需求。

3 结果与分析

3.1 改进 YOLO v8s 的消融试验

为评估本文所提出的各结构模块对胎衣检测性能的影响,设计了包括 Focus 模块、SE 注意力机制、C2f-SCCConv 模块及 BiFPN-P2 结构的多组消融试验,试验结果如表 1 和图 7 所示。为保证对比公平性,所有模型均在相同硬件环境下进行测试。帧率

表 1 消融试验结果

Tab. 1 Ablation study results

SE	Focus	C2f-SCCConv	BiFPN-P2	精确率/%	召回率/%	平均精确率/%	参数数量	帧率/(f·s ⁻¹)
×	×	×	×	95.3	90.2	97.8	1.113×10^7	122.37
√	×	×	×	98.0	94.7	98.2	1.113×10^7	124.52
×	√	×	×	88.5	89.8	95.6	1.113×10^7	124.17
×	×	√	×	97.1	96.0	97.7	1.035×10^7	48.87
×	×	×	√	87.8	90.0	94.5	8.100×10^6	112.31
√	√	×	×	93.4	93.1	96.6	1.113×10^7	140.28
√	×	√	×	87.3	70.1	83.7	7.330×10^6	52.76
√	×	×	√	88.7	87.0	93.8	8.100×10^6	57.67
×	√	√	×	94.1	84.7	93.6	1.036×10^7	65.89
×	√	×	√	91.9	95.8	96.5	8.110×10^6	111.94
×	×	√	√	97.5	93.2	97.6	7.330×10^6	57.85
×	√	√	√	91.0	81.5	93.6	7.330×10^6	29.12
√	×	√	√	89.2	89.6	88.6	7.330×10^6	61.76
√	√	×	√	89.1	88.3	94.5	8.110×10^6	113.90
√	√	√	×	82.4	85.6	78.9	1.036×10^7	64.64
√	√	√	√	98.1	94.9	98.8	7.330×10^6	61.71

注: × 表示不采用该模块;√表示采用该模块。

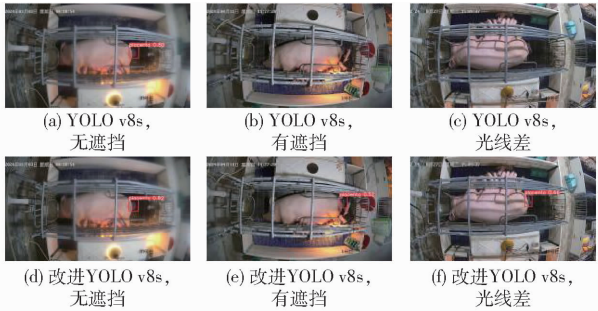


图 7 模型检测效果对比

Fig. 7 Comparison of detection performance across models

统计包含预处理、模型前向推理及 NMS 后处理时间,输入分辨率统一为 640 像素 × 640 像素, batch size 设为 1,采用 FP32 单精度模式,最终结果取测试集上的平均值。

首先,从单模块效果来看,基线模型 YOLO v8s 在精确率、召回率和 mAP 方面分别达到 95.3%、90.2% 和 97.8%。引入 SE 注意力机制后,模型的精确率显著提升至 98.0%,召回率提高至 94.7%,mAP 达到 98.2%,表明动态调整通道权重能够有效突出胎衣关键特征,增强模型的判别能力。在仅引入 Focus 模块的情况下,模型的召回率保持较高水平,而精确率明显下降,其主要原因在于 Focus 的空间重组操作增强了浅层细节与纹理信息,使背景区域中与胎衣形态相似的噪声结构(如血迹、水渍及局部阴影)被过度放大,从而提高了模型对弱纹理区域的敏感性。由于缺乏后续的注意力约束与深层语义融合,网络更容易将部分伪纹理区域误判为胎衣,导致假阳性样本增加、精确率下降。该现象也验证了 Focus 模块需与 C2f-SCConv、SE 等模块协同使用,以在增强细节的同时抑制背景干扰,形成稳定可靠的特征表达。采用 C2f-SCConv 模块后,模型在精确率(97.1%)和召回率(96.0%)方面表现优异,说明空间感知卷积有效加强了对胎衣边缘和纹理的建模能力。仅替换 Neck 为 BiFPN-P2 结构时,虽然精确率和召回率略有下降,但模型参数量大幅减少,帧率保持较高,体现了良好的轻量化特性。SE 模块能够实现对关键语义通道的增强,由于胎衣目标在浅层与深层特征中均存在显著差异,将 SE 模块布置在 Neck 输出位置能够在不增加较大计算量的前提下,进一步强化与胎衣相关的区分性特征,并抑制背景干扰,从而提升检测鲁棒性。由于 SE 模块仅由两层全连接结构组成,其参数规模极小,相比 Backbone 和 Neck 中的大量卷积层,其对总体参数量的影响可以忽略不计,因此加入 SE 模块前后模型总参数量保持不变。

进一步分析多模块组合效果,加入 SE 与 Focus 模块后,模型的推理速度能力得到了提升,然而精确率和召回率均有所回落,表明浅层细节增强与通道注意力机制在缺乏有效融合路径的情况下,可能导致特征响应偏向局部纹理,从而引入一定程度的误检与漏检。SE 与 C2f-SCConv 的结合提升了精确率,但召回率下降较为明显,说明空间结构建模与通道注意力在联合过程中更倾向于强化显著特征,从而可能对部分弱目标产生抑制效应,反映出特征选择机制在不同层级之间仍存在一定的不均衡性。C2f-SCConv 与 BiFPN-P2 的结合带来了较为平衡的性能表现,精确率达到 97.5%,召回率 93.2%,同时模型参数量表现良好,表明空间结构建模与多尺度特征融合具有良好的互补性:前者强化局部边缘与纹理信息,后者提升跨尺度语义表达能力,从而共

同增强模型对小尺度目标的检测能力。

最终,当所有改进模块共同作用时,模型实现了最佳检测性能,精确率达 98.1%,召回率 94.9%,mAP 达到 98.8%,且参数量控制在 7.33×10^6 ,帧率仍保持在 61.71 f/s,充分展示了改进网络在提升胎衣检测精度与召回率的同时,兼顾了轻量化和实时推理的需求。

综上所述,消融试验验证了本文提出的各模块对提升胎衣等小尺度、低对比度目标检测性能的有效性。特别是 SE 注意力机制和 C2f-SCConv 模块的引入,大幅增强了关键特征的表达能力;BiFPN-P2 结构则通过多尺度加权融合优化了信息流通,综合提升了模型的检测准确率和泛化能力。虽然 Focus 模块单独表现有限,但在整体网络中协同作用下有助于浅层细节的保留。该系列结构优化方案显著改进了 YOLO v8s 在复杂背景下对胎衣目标的检测效果,满足了实际应用的实时性和精度需求。

图 8 展示了模型在训练与验证过程中的损失曲线变化情况。可以看出,损失值在 20 轮内快速下降,至 100 轮后趋于稳定,说明模型已有效收敛,无明显过拟合或欠拟合。

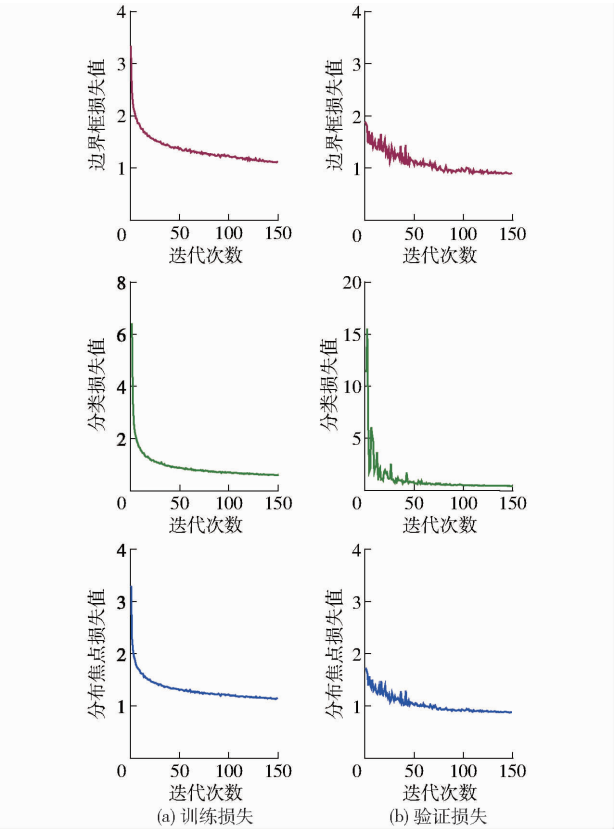


图 8 改进 YOLO v8s 模型损失曲线

Fig. 8 Loss curves of improved YOLO v8s model

3.2 改进的 YOLO v8s 与主流检测模型对比试验

为验证所提出方法的有效性,将改进的 YOLO v8s 与 NanoDet、RT-DETR^[35]、Faster R-CNN^[36]、

YOLO v5 及 YOLO v8s 等主流目标检测模型进行了对比分析,结果如表 2 所示。整体来看,轻量化模型 NanoDet 在胎衣检测中的表现较差,其精确率和召回率仅为 25.4% 和 2.8%,难以满足实际应用需求。RT-DETR 与 Faster R-CNN 的检测精度有所提高,mAP 分别达到 97.5% 和 88.1%,但受限于较大的参数量,推理速度偏低,难以实现高效的实时检测。YOLO v5 在检测精度和速度之间取得了一定平衡,mAP 达到 96.5%,帧率为 78.96 f/s。然而,YOLO v8s 在各项指标上进一步提升,其精确率和召回率分别为 95.3% 和 90.2%,mAP 达到 97.8%,且在保证较高精度的同时实现了 122.37 f/s 的高速推理,具备良好的实用价值。

表 2 不同模型检测性能对比

Tab. 2 Comparison of detection performance across different models

模型	精确率/%	召回率/%	mAP/%	参数量	帧率/(f·s ⁻¹)
NanoDet	25.4	2.8	50.9	9.400×10^5	24.80
RT-DETR	60.9	65.7	97.5	3.281×10^7	12.04
Faster R-CNN	56.4	47.1	88.1	4.135×10^7	24.59
YOLO v5	93.8	87.2	96.5	9.110×10^6	78.96
YOLO v8s	95.3	90.2	97.8	1.113×10^7	122.37
改进 YOLO v8s	98.1	94.9	98.8	7.330×10^6	61.71

表 3 不同模块的轻量化效果

Tab. 3 Lightweight effects of different modules

模型变体	参数量	浮点运算量	内存占用量/MB	帧率/(f·s ⁻¹)	说明
YOLO v8s (基线模型)	1.113×10^7	2.84×10^{10}	44.52	122.37	原始效果
+ Focus	1.113×10^7	2.92×10^{10}	44.80	124.17	Focus 保留空间信息,参数增长小
+ C2f-SCConv	1.035×10^7	2.62×10^{10}	41.40	48.87	C2f-SCConv 替代原 C2f,稍微调整通道/层结构
+ BiFPN-P2	8.100×10^6	5.49×10^{10}	32.40	112.31	BiFPN-P2 使用轻量化融合减少总体参数
+ SE	1.113×10^7	2.84×10^{10}	44.52	124.52	SE 引入瓶颈 FC,参数影响小
最终改进(全部)	7.330×10^6	5.35×10^{10}	29.32	61.71	最终组合

3.4 视频级识别试验结果分析

为了验证本文提出的改进 YOLO v8s 模型在视频输入条件下的识别有效性,构建了基于时间帧间隔的推理机制,并结合帧级计数器实现对胎衣目标在视频中的连续检测判断。试验以分娩过程中的监控视频为输入源,分别设定帧间隔为 25 f 和 5 f,当在连续 3 帧中均检测到胎衣目标时,判定为分娩结束事件发生,并记录对应的帧号。识别效果如图 9 所示。

表 4 列出了在 5 段真实分娩视频中的识别结果统计情况。结果表明,不同帧间隔下模型均能够稳定检测分娩结束事件。在 25 f 间隔下,模型在 5 段视频中的平均时间误差约为 5.98 s,最小误差为

相比之下,本文提出的改进 YOLO v8s 模型在精度和鲁棒性上均表现最佳,精确率和召回率分别提高到 98.1% 和 94.9%,mAP 达到 98.8%,较原始 YOLO v8s 有明显提升。虽然推理速度(61.71 f/s)略有下降,但仍能满足实际生产中实时检测的需求。同时,改进 YOLO v8s 的参数量减少至 7.33×10^6 ,仅约为原 YOLO v8s 的 66%,显著降低了模型复杂度和计算成本。综合对比结果表明,改进 YOLO v8s 在胎衣检测任务中兼顾了高精度与较低复杂度,适合在养殖场景下的实时推广应用。

3.3 模型复杂度与轻量化分析

为全面评估所提出改进策略在模型轻量化方面的有效性,本研究对改进前后各网络结构的参数规模、浮点运算量以及推理性能进行了系统分析。首先,以原始 YOLO v8s 为基准模型,通过引入 Focus、C2f-SCConv、BiFPN-P2 以及 SE 等模块,分别构建若干中间模型,并在相同输入尺寸条件下采用 thop 与 torchsummary 工具对各模型的参数量与浮点运算量进行逐项统计,从而量化不同结构调整对整体复杂度的影响。受上述结构优化的综合作用,最终改进模型的参数规模由原始 YOLO v8s 的 1.113×10^7 降至 7.33×10^6 ,在减少约 34.1% 参数量的同时保持较高的推理效率与检测精度。如表 3 所示的对照表,明确了各改进模块对轻量化效果的具体贡献。

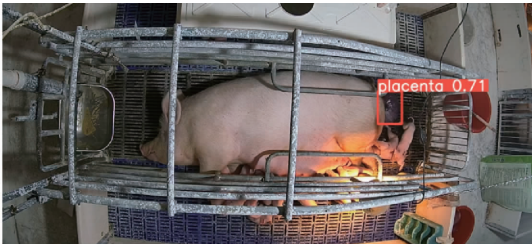


图 9 视频识别效果示意图

Fig. 9 Video recognition performance diagram

1.71 s(V1),最大不超过 12.00 s(V2)。在 5 f 间隔下,检测精度进一步提高,平均时间误差仅约 1.84 s,最小误差为 0.04 s(V1),最大误差为 6.96 s(V4)。

综合对比可见,缩短检测帧间隔能够有效降低

表 4 改进 YOLO v8s 模型视频级识别效果

Tab. 4 Improved YOLO v8s model's video-level recognition performance

帧间隔/ f	视频 编号	视频 总帧数	模型检 测帧数	人工标 注帧数	误差 帧数	时间 误差/s
25	V1	2557	275	234	41	1. 71
	V2	12 041	3 525	3 237	288	12. 00
	V3	8 724	2 975	2 920	55	2. 29
	V4	20 939	9 875	9 618	257	10. 71
	V5	8 201	1 100	1 024	76	3. 17
5	V1	2 557	235	234	1	0. 04
	V2	12 041	3 215	3 237	22	0. 92
	V3	8 724	2 935	2 920	15	0. 63
	V4	20 939	9 785	9 618	167	6. 96
	V5	8 201	1 040	1 024	16	0. 67

时间误差,提升模型在视频级别的时序判断精度;同时,即使在 25 f 间隔的情况下,误差范围仍处于可接受水平,表明改进 YOLO v8s 模型在分娩监测中的实时性和稳定性较强,能够满足实际应用需求。

4 讨论

尽管本文提出的改进 YOLO v8s 模型在胎衣检测任务中取得了较高精度,但在部分特殊场景下仍存在一定局限性。首先,当胎衣被母猪身体、栏杆或仔猪严重遮挡时,其外观特征仅部分暴露,可能导致模型无法提取完整轮廓信息,从而产生漏检现象。其次,在光照极弱或地面污渍较多的情况下,部分背

景区域与胎衣呈现相近的颜色与纹理,易导致模型误将污渍、水渍等伪目标判为胎衣,从而增加误检概率。上述问题说明,在极端环境条件下,仅依赖 RGB 图像的单模态视觉信息仍存在不确定性,需要结合深度信息、红外成像或行为线索等多模态信息进一步增强模型的泛化能力。

5 结论

(1)在结构优化方面,SE 注意力机制与 C2f -SCConv 模块有效提升了小尺度、低对比度目标的特征提取能力,BiFPN - P2 结构在保持检测性能的同时显著降低了模型参数量和计算开销。最终融合多模块的改进 YOLO v8s 模型精确率、召回率和 mAP 分别提升至 98. 1%、94. 9% 和 98. 8% ,参数量降至 $7. 33 \times 10^6$,推理速度仍达 61. 71 f/s,兼顾了精度与实时性。

(2)在模型对比中,改进 YOLO v8s 优于 NanoDet、RT - DETR、Faster R - CNN、YOLO v5 及原始 YOLO v8s 等主流目标检测方法,不仅检测精度更高,而且复杂度更低,显示出较强的工程实用性。

(3)在视频级识别试验中,基于时间帧间隔与帧级计数器的时序机制突破了依赖仔猪目标的间接推断模式,能够在实际分娩视频中稳定识别分娩结束事件。在 25 f 间隔条件下,平均时间误差约为 5. 98 s;在 5 f 间隔条件下,误差缩小至 1. 84 s,显著提升了监测的时效性与鲁棒性。

参 考 文 献

[1] Zhang M, Wang X, Feng H, et al. Wearable internet of things enabled precision livestock farming in smart farms: a review of technical solutions for precise perception, biocompatibility, and sustainability monitoring[J]. Journal of Cleaner Production, 2021, 312: 127712.

[2] Gauthier R, Largouet C, Dourmad J. Prediction of litter performance in lactating sows using machine learning, for precision livestock farming[J]. Computers and Electronics in Agriculture, 2022, 196: 106876.

[3] Wang S, Jiang H, Qiao Y, et al. The research progress of vision-based artificial intelligence in smart pig farming[J]. Sensors, 2022, 22(17): 6541.

[4] Tzanidakis C, Simitzis P, Arvanitis K, et al. An overview of the current trends in precision pig farming technologies[J]. Livestock Science, 2021, 249: 104530.

[5] Li Y, Xiao J, Meng Y, et al. Overview of ecology and automation status of pig breeding[J]. IOP Conference Series: Earth and Environmental Science, 2023, 1171(1): 12060.

[6] van Erp-Van Der Kooij E, de Graaf L F, de Kruijff D A, et al. Using sound location to monitor farrowing in sows[J]. Animals, 2023, 13(22): 3538.

[7] 余华君,孙裕光. 妊娠母猪分娩前后及仔猪饲养管理关键技术[J]. 黑龙江畜牧兽医, 2012(18): 62 - 63.

[8] Garcia R, Aguilar J, Toro M, et al. A systematic literature review on the use of machine learning in precision livestock farming[J]. Computers and Electronics in Agriculture, 2020, 179: 105826.

[9] Archana R, Jeevaraj P S E. Deep learning models for digital image processing: a review[J]. Artificial Intelligence Review, 2024, 57(1): 11.

[10] Kamilaris A, Prenafeta-Boldu F X. Deep learning in agriculture: a survey[J]. Computers and Electronics in Agriculture, 2018, 147: 70 - 90.

[11] Lao F, Brown-Brandl T, Stinn J P, et al. Automatic recognition of lactating sow behaviors through depth image processing[J]. Computers and Electronics in Agriculture, 2016, 125: 56 - 62.

- [12] Chen J, Zhou J, Liu L, et al. Sow farrowing early warning and supervision for embedded board implementations[J]. *Sensors*, 2023, 23(2): 727.
- [13] Tsai Y, Huang Y, Lin E, et al. Monitoring the lactation-related behaviors of sows and their piglets in farrowing crates using deep learning[J]. *Frontiers in Animal Science*, 2024, 5: 1431285.
- [14] Oliviero C, Pastell M, Heinonen M, et al. Using movement sensors to detect the onset of farrowing[J]. *Biosystems Engineering*, 2008, 100(2): 281–285.
- [15] Piyanuch S, Ananta S, Warot B. An embedded system device to monitor farrowing[C]//2018 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA). Krabi: IEEE, 2018:208–213.
- [16] 高云,郭继亮,黎焯,等. 基于深度学习的群猪图像实例分割方法[J]. *农业机械学报*, 2019, 50(4): 179–187.
Gao Yun, Guo Jiliang, Li Xuan, et al. Instance-level segmentation method for group pig images based on deep learning[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 50(4): 179–187. (in Chinese)
- [17] Hansen M E, Smith M L, Smith L N, et al. Towards on-farm pig face recognition using convolutional neural networks[J]. *Computers in Industry*, 2018, 98: 145–152.
- [18] Walls A, Hall E, Lomax S, et al. An investigation of parturient ocular appearance in sows[J]. *Animals*, 2024, 14(18): 2693.
- [19] Rootwelt V, Reksen O, Farstad W, et al. Associations between intrapartum death and piglet, placental, and umbilical characteristics[J]. *Journal of Animal Science*, 2012, 90(12): 4289–4296.
- [20] Rohan A, Rafiq M S, Hasan M J, et al. Application of deep learning for livestock behaviour recognition: a systematic literature review[J]. *Computers and Electronics in Agriculture*, 2024, 224: 109115.
- [21] de Paula T M C G, de Sousa R V, Sarmiento M P, et al. Deep learning pose detection model for sow locomotion[J]. *Scientific Reports*, 2024, 14(1): 16401.
- [22] Oczak M, Maschat K, Baumgartner J. Dynamics of sows' activity housed in farrowing pens with possibility of temporary crating might indicate the time when sows should be confined in a crate before the onset of farrowing[J]. *Animals*, 2019, 10(1): 6.
- [23] 李路,孙超奇,周玉凡,等. 基于 FAS-YOLO v8n 的爬岸上草小龙虾多源图像融合识别方法[J]. *农业机械学报*, 2025, 56(8): 526–534.
Li Lu, Sun Chaoqi, Zhou Yufan, et al. Multi-source image fusion recognition of crawling bank grass crayfish based on FAS-YOLO v8n[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2025, 56(8): 526–534. (in Chinese)
- [24] Yang K Y, Jeon J H, Kwon K S, et al. Effect of different parities on reproductive performance, birth intervals, and tail behavior in sows[J]. *Journal of Animal Science and Technology*, 2019, 61(3): 147–153.
- [25] Rejin V, Sambath M. YOLO v8: a novel object detection algorithm with enhanced performance and robustness[C]//2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS). Chennai: IEEE, 2024:1–6.
- [26] Compte A, Yan Y, Cortés X, et al. Housed pig identification and tracking for precision livestock farming[J]. *Expert Systems with Applications*, 2025, 293: 128466.
- [27] Zhong C, Wu H, Jiang J, et al. YOLO-DLHS-p: a lightweight behavior recognition algorithm for captive pigs[J]. *IEEE Access*, 2025, 13: 58442.
- [28] Luo Y, Lin K, Xiao Z, et al. PBR-YOLO: a lightweight piglet multi-behavior recognition algorithm based on improved YOLO v8[J]. *Smart Agricultural Technology*, 2025, 10: 100785.
- [29] Xia X, Zhan N, Guan Z, et al. PAB-mamba-YOLO: VSSM assists in YOLO for aggressive behavior detection among weaned piglets[J]. *Artificial Intelligence in Agriculture*, 2025, 15(1): 52–66.
- [30] Beecks C, Amalraj A, Groß A, et al. Leveraging YOLO for real-time video analysis of animal welfare in pig slaughtering processes[C]//Cham: Springer Nature Switzerland, 2024:275–281.
- [31] Li H, Li B, Li H, et al. A lightweight decision-level fusion model for pig disease identification using multi-modal data[J]. *Computers and Electronics in Agriculture*, 2025, 231: 109936.
- [32] Jiafeng L, Ying W, Lianghua H. SCConv: spatial and channel reconstruction convolution for feature redundancy[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver: IEEE, 2023:6153–6162.
- [33] Mingxing T, Ruoming P, Quoc V L. EfficientDet: scalable and efficient object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020:10778–10787.
- [34] Jie H, Li S, Samuel A, et al. Squeeze-and-excitation networks[J]. *IEEE Transactions On Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011–2023.
- [35] Yian Z, Wenyu. L, Shangliang X, et al. DETRs beat YOLOs on real-time object detection[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2024:16965–16974.
- [36] Shaoqing R, Kaiming H, Ross G, 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.