

基于时空注意力反演的猪只排便行为长视频分析框架

杨秋妹<sup>1,2</sup> 吴曼婷<sup>1,2</sup> 肖德琴<sup>1,2</sup> 陈泽中<sup>1,2</sup> 吴纪沿<sup>1,2</sup> 洪其伟<sup>1,2</sup>

(1. 华南农业大学数学与信息学院, 广州 510642; 2. 农业农村部华南热带智慧农业技术重点实验室, 广州 510642)

**摘要:** 猪只排便频率是评估猪群健康与环境卫生的重要指标,但现有方法难以在多头猪只的饲养环境和长时序场景中实现个体排便行为的精准识别与统计。为解决这些挑战,本文提出了一种自动化识别和统计猪只排便行为的时空注意力反演推理(Spatiotemporal attention-based inverse reasoning, SAIR)框架。设计了一种基于空间注意力增强的YOLO v8模型,实现猪只个体、臀部及粪便的实时检测,并结合改进的BoT-SORT算法进行轨迹跟踪。提出了一种基于自适应加权多层感知机融合网络,对受背景干扰和低分辨率影响的疑似粪便图像进行精细化重识别。以粪便识别帧为起点回溯收集历史视频帧数据,并采用K-Means算法对猪群行为轨迹中心进行聚类,计算粪便与猪只个体的空间相关性。同时,基于时序注意力机制计算粪便和猪只个体的时间相关性。融合粪便与猪只个体的空间、时间相关性,反演推理确定发生排便行为的猪只个体。结果表明,在短视频测试中,本文方法分类准确率、召回率和精确率分别为97.7%、97.2%和98.1%;在长视频场景下,其F1值、召回率和精确率分别达到88.5%、90.0%和87.1%。该方法有效降低了猪只单次排泄多坩粪便和长时序轨迹误判带来的识别误差,为猪只排便行为识别及健康评估提供了可靠支持。

**关键词:** 猪只排便; 行为分析; 计算机视觉; 时空注意力; 多目标跟踪

中图分类号: TP391 文献标识码: A 文章编号: 1000-1298(2026)10-0295-13

OSID:



Spatiotemporal Attention-based Inverse Reasoning Framework  
for Pig Defecation Behavior Analysis in Long Videos

YANG Qiumei<sup>1,2</sup> WU Manting<sup>1,2</sup> XIAO Deqin<sup>1,2</sup> CHEN Zezhong<sup>1,2</sup> WU Jiyang<sup>1,2</sup> HONG Qiwei<sup>1,2</sup>

(1. College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China

2. Key Laboratory of Smart Agricultural Technology in Tropical South China,  
Ministry of Agriculture and Rural Affairs, Guangzhou 510642, China)

**Abstract:** Defecation frequency serves as a crucial indicator for assessing pig health and environmental hygiene. However, existing methods struggle to accurately identify and quantify individual defecation behaviors in multi-pig housing environments and long-duration scenarios. To address these challenges, a spatiotemporal attention-based inverse reasoning (SAIR) framework for automated analysis of pig defecation was proposed. The framework first designed a spatial attention-enhanced YOLO v8 model for detecting pigs, pig hindquarters, and feces. It incorporated an improved BoT-SORT algorithm for pig tracking. Subsequently, an adaptive weighted fusion multi-layer perceptron network was proposed for refined recognition of suspected feces with significant background interference and low resolution. Starting from the feces detection frame, the framework backtracks historical trajectories using K-Means clustering to analyze spatial correlations between pigs and feces. Simultaneously, temporal correlations between feces and pigs were computed by using a temporal attention mechanism. Finally, by integrating the spatial and temporal correlations, the model identified the specific pig involved in defecation. In short video tests, the proposed method achieved classification accuracy, recall, and precision of 97.7%, 97.2%, and 98.1%, respectively. In long video scenarios, the F1-score, recall, and precision reached 88.5%, 90.0%, and 87.1%, respectively. The proposed approach effectively reduced recognition errors caused by multiple fecal discharges in a single defecation event and trajectory misjudgments in

收稿日期: 2025-08-15 修回日期: 2025-12-30

基金项目: 广东省重点领域研发计划项目(2023B0220140001)和国家重点研发计划项目(2021YFD2000802)

作者简介: 杨秋妹(1983—),女,讲师,博士,主要从事智慧农业研究,E-mail: yqmbegonia@163.com

通信作者: 肖德琴(1970—),女,教授,博士生导师,主要从事智慧农业研究,E-mail: deqinx@scau.edu.cn

long-term sequences, providing reliable support for pig defecation behavior recognition and health assessment.

**Key words:** pig defecation; behavior analysis; computer vision; spatiotemporal attention; multi-object tracking

0 引言

随着深度学习的发展,人工智能技术在猪只行为监测中的工程应用价值越发显著<sup>[1-3]</sup>。基于视觉的行为监测可自动获取排便频率与持续时长等关键指标,为评估猪只营养负荷能力和消化系统健康状况提供重要依据<sup>[4]</sup>。此外,粪便残留不仅显著影响猪舍卫生并诱发猪只蹄部疾病<sup>[5-6]</sup>,也是养殖环境中氨气排放的重要来源,与养殖人员呼吸系统健康密切相关<sup>[7-8]</sup>。因此,对猪只排便行为进行识别和统计,对猪只健康评估和养殖环境管理均具有重要意义。

然而,传统人工观察猪只排便行为的方法存在人力消耗大和效率低等问题<sup>[9]</sup>。随着计算机视觉视频监控技术在猪只检测、跟踪和行为识别等领域广泛应用<sup>[10-11]</sup>,基于视觉技术实现猪只排便行为的自动化识别与统计已成为研究热点。DING等<sup>[12]</sup>提出在猪舍排泄区域安装摄像头,结合排便姿态与异常粪便检测断奶仔猪的腹泻情况。杨灵星等<sup>[13]</sup>通过使用 EfficientDet 对热红外视频图像中呈现排泄姿态且臀部后方有排泄物的猪只进行检测,从而识别出正在排泄的猪只。但由于猪只排便姿态与其行走或站立时的姿态高度相似,当猪只在其他猪只的粪便上行走或站立时,易被误识别为正在排泄。这种姿态的高度相似性显著影响了排泄行为的精准识别,限制了上述方法的适用性和检测效果。另外,ZHU等<sup>[14]</sup>通过预先设定排泄区并监测猪只是否进入该区域统计排便时间和次数。但猪只并不总是在指定区域排便,导致误差较大。因此,另一种研究思路是不依赖排泄区域,而是基于猪只自身的行为特征进行判断。如吴世海等<sup>[15]</sup>采用 CCD 相机获取猪只视频,采用多层卷积神经网络提取并融合图像深度特征,以识别猪只呼吸和排泄等行为。然而,该方法仅适用于画面中单头猪只的场景,而在以群养为主的实际养殖环境中,猪只间相互遮挡会对识别效果产生干扰。此外,该方法局限于短视频识别,难以满足现实中对长时间监控的需求。针对该问题,LI等<sup>[16]</sup>通过在模型中引入长短期记忆网络(LSTM),提出了一种融合排便行为与粪便图像的模式(HCC-YOLO v8-LSTM)。然而,该方法仍依赖于猪只排泄时的异常行为作为判断依据,对

于姿态变化和行为特征不显著的情况,其识别能力有限。综上所述,尽管基于计算机视觉的方法在猪只排便行为识别上取得了一定成果,但在实际应用中,这些方法普遍难以应对排便姿态不明显以及猪只相互遮挡导致的识别准确率下降问题;同时难以在长时间场景下对猪只排便行为进行有效识别与统计。

为解决上述问题,本文提出一种基于时空注意力的反演推理(SAIR)框架,对猪只排便行为进行自动识别与统计,以期为猪群健康管理提供可靠数据支持。

1 数据来源与数据集构建

1.1 数据来源

所使用数据采自广州萝岗试验猪场,选取一栏4头长白猪作为试验对象。猪舍尺寸为4 m×5 m,摄像机安装在猪舍中央,距地面约5 m高的位置。摄像机为海康红外网络摄像机(DS-2CD3345-I型,杭州海康威视数字技术股份有限公司),镜头焦距为2.8 mm,图像分辨率为2 560像素×1 440像素,视频帧率为25 f/s。拍摄的视频在猪舍本地采用海康网络硬盘录像机(DS-7804N-K1/C型,杭州海康威视数字技术股份有限公司)存储。试验周期共13 d,监测时间为24 h/d,实现全日连续观测,白天利用自然光照明,夜间使用红外夜视功能进行成像。由于夜晚猪只多处于睡眠状态,故选取猪只较为活跃的时间段进行研究,具体为每日08:00—18:00。

1.2 数据预处理与数据集构建

将收集的13 d数据分为2部分,即前6 d和后7 d数据。这一划分旨在保证数据之间相互独立,避免同一时段样本同时用于训练和测试,从而提高结果的客观性与可靠性。其中,前6 d数据用于构建图像数据集,命名为 $D_{img}$ ,而后7 d数据则用于构建视频数据集,命名为 $D_{vid}$ 。

利用 $D_{img}$ 构建 $D_{img}^{(1)}$ 和 $D_{img}^{(2)}$ 。 $D_{img}^{(1)}$ 用于训练猪只个体、臀部和粪便的目标检测模型。具体来说,从 $D_{img}$ 中选取存在粪便的视频片段,通过提取这些片段中的每1帧共生成27 846幅图像。使用标注软件“X-AnyLabeling”对图像中猪只个体、臀部和粪便进行人工标注,形成数据集 $D_{img}^{(1)}$ ,并按照比例4:1

划分为训练集和测试集。 $D_{img}^{(2)}$  则用于训练粪便精细化识别模型。具体为利用上述  $D_{img}^{(1)}$  中的粪便标注框原图像进行裁剪,最终获得 33 762 幅粪便图像,另外再随机裁剪 33 762 幅非粪便图像,最终形成  $D_{img}^{(2)}$ ,并同样按照比例 4:1 划分为训练集和测试集。 $D_{img}^{(1)}$  和  $D_{img}^{(2)}$  具体划分细节如表 1 所示。

表 1 图像数据集

Tab.1 Image dataset

| 数据集             | 目的     | 类型  | 数量           |
|-----------------|--------|-----|--------------|
| $D_{img}^{(1)}$ | 目标检测   | 训练集 | 22 277       |
|                 |        | 测试集 | 5 569        |
| $D_{img}^{(2)}$ | 精细化重识别 | 训练集 | 27 010 (粪便)  |
|                 |        |     | 27 010 (非粪便) |
|                 |        | 测试集 | 6 752 (粪便)   |
|                 |        |     | 6 752 (非粪便)  |

利用  $D_{vid}$  构建  $D_{vid}^{(1)}$ 、 $D_{vid}^{(2)}$  和  $D_{vid}^{(3)}$ 。 $D_{vid}^{(1)}$  用于测试猪只多目标跟踪算法的性能。从  $D_{vid}$  中随机选取 2 段时长分别为 10 min (视频 1) 和 1 h (视频 2) 的片段构成  $D_{vid}^{(1)}$ 。 $D_{vid}^{(2)}$  用于验证本文所提猪只排便行为识别框架在短视频中的表现。基于此,从  $D_{vid}$  中截取了多个时长为 4 ~ 30 s 的猪只排便行为短视频,并通过水平翻转、亮度和对比度调整等数据增强手段扩展到 110 个短视频。同时也截取 110 个时长为 4 ~ 30 s 不包含排便行为的短视频,最终构成  $D_{vid}^{(2)}$ 。 $D_{vid}^{(3)}$  用于评估本文所提猪只排便行为识别与统计框架在长视频中的性能。为此,从  $D_{vid}$  中选取 7 个时长均为 10 h 的视频片段构成  $D_{vid}^{(3)}$ 。 $D_{vid}^{(1)}$ 、 $D_{vid}^{(2)}$  以及  $D_{vid}^{(3)}$  划分如表 2 所示。

表 2 视频数据集

Tab.2 Video dataset

| 数据集             | 目的    | 视频   | 描述             |
|-----------------|-------|------|----------------|
| $D_{vid}^{(1)}$ | 多目标跟踪 | 视频 1 | 持续 10 min      |
|                 |       | 视频 2 | 持续 1 h         |
| $D_{vid}^{(2)}$ | 短视频测试 | 排便   | 110 个 4 ~ 30 s |
|                 |       | 非排便  | 110 个 4 ~ 30 s |
| $D_{vid}^{(3)}$ | 长视频测试 | 长视频  | 7 个 10 h       |

## 2 猪只排便行为识别与统计算法

猪只排便行为识别与统计主要由 3 个关键模块组成,分别为:用于目标检测的基于空间注意力增强的 YOLO v8 模型 (Spatial attention-enhanced - YOLO v8, SAE - YOLO v8);用于多目标跟踪的基于 ID 重绑定机制的 BoT - SORT 算法 (ID rebind BoT - SORT, IDR - BSORT);用于精细识别的自适应加权融合多层感知机 (Adaptive weighted fusion - multi-layer perceptron, AWF - MLP)。

基于上述 3 个模块设计 1 个基于时空注意力的反演推理 (SAIR) 算法。图 1 展示了该算法的整体框架。

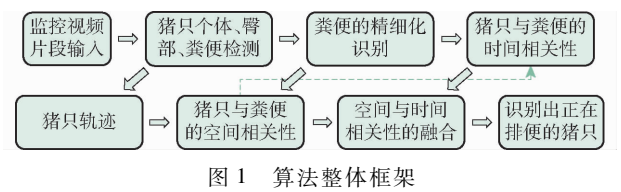


Fig.1 Framework of proposed method

### 2.1 猪只检测模型 (SAE - YOLO v8)

猪只个体、臀部和粪便的检测效果直接影响排便行为识别与统计的准确性。然而,由于上述目标在尺寸、颜色等特征上差异显著,且猪舍环境复杂,易引入冗余背景信息,降低模型对目标的注意力。为此,在 CBAM 结构<sup>[17]</sup>基础上提出一种空间注意力增强 (Spatial attention-enhanced, SAE) 机制,并将其融入目标检测模型 YOLO v8<sup>[18]</sup> 中的第 9 层与第 10 层之间,以增强特征图的空间表达能力,构建出检测性能更优的 SAE - YOLO v8 模型。其中 SAE 结构如图 2 所示。

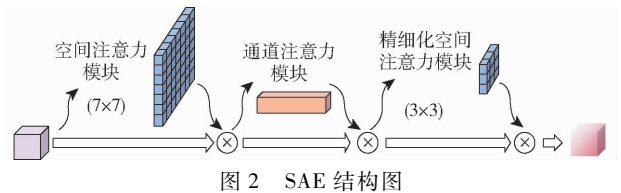


Fig.2 Structure diagram of SAE

由图 2 可知,SAE 由大感受野空间注意力、通道注意力和精细化空间注意力模块依次组成。通过引入 1 个卷积核较大的空间注意力模块 (7 × 7),以增强模型在复杂背景环境中的感知能力,捕捉更广泛的空间信息,提高模型对背景噪声的鲁棒性;通道注意力模块对特征通道进行加权,以自适应地突出目标的相关特征并有效抑制冗余信息;引入 1 个卷积核为 3 × 3 的空间注意力模块,进一步细化目标的空间感知,特别是针对小尺寸目标 (如粪便) 的检测进行精细化处理。通过多尺度融合的注意力机制 SAE 有效提升了模型对猪舍复杂环境中猪只个体、臀部及粪便目标的空间关注度,降低了环境干扰对检测性能的影响。

### 2.2 猪只多目标跟踪算法 (IDR - BSORT)

在精准检测猪只个体、猪只臀部和粪便后,需确定粪便对应的目标猪只以支撑后续反演推理。BoT - SORT 是一种优化 SORT 系列算法<sup>[19]</sup> 的多目标跟踪方法<sup>[20]</sup>,通过引入优化的卡尔曼滤波器状态向量<sup>[21]</sup> 等策略提升跟踪性能。然而,BoT - SORT 仍存在身份丢失等问题。常用的解决方案是集成基于外

观或运动特征的 Re-ID 模块,但其计算开销较大<sup>[22]</sup>,难以满足本研究对连续视频实时分析的要求。

为此,提出一种 ID Rebind (IDR) 机制并将其与 BoT-SORT 融合形成新的多目标跟踪算法 IDR-BSORT。IDR 的设计初衷并非取代 Re-ID,而是在固定猪只数量、遮挡模式相对稳定的养殖场景中,提供一种足够有效、计算开销低、易于部署的 ID 管理方式。IDR 机制使得当目标在长时间遮挡后重新出现时,能够将其重新绑定回原来的 ID。具体而言,首先构建包含所有可能目标 ID 的集合  $ID_{des}$ 。例如,如果猪舍中有 4 头猪只, $ID_{des}$  为  $\{1, 2, 3, 4\}$ 。通过  $ID_{cur}$  集合来存储当前帧中跟踪器分配的 ID,计算  $ID_{des}$  和  $ID_{cur}$  之间的差集,找出尚未分配的 ID 集合  $ID_{unu}$ 。在 ID 重新分配中,该机制通过检查每个跟踪目标的 track id 是否在  $ID_{des}$  中,若不在,则将其顺序替换为  $ID_{unu}$  中的一个未使用的 ID。这种方式限制了每个目标的 ID 值都在  $ID_{des}$  中,并尽量保证所有猪只个体与 ID 值的正确绑定,而不会因为遮挡或检测误差而生成新的 ID。IDR 机制的伪代码如下

输入: 跟踪目标列表 track\_list,最大目标数量 max\_objects

输出: 更新后的跟踪目标列表 updated\_track\_list

- 1: 初始化目标 ID 集合  $ID_{des}$  为  $\{1, 2, \dots, \max\_objects\}$
- 2: 遍历 track\_list,获取当前帧中每个跟踪目标的 track id,得到  $ID_{cur}$
- 3: 基于  $ID_{des}$  和  $ID_{cur}$ ,计算未分配的 ID 集合  $ID_{unu}$
- 4: 初始化未使用的 ID 索引为 0
- 5: 遍历 track\_list,对每个跟踪目标进行 ID 检查与分配
- 6: 更新跟踪目标列表 track list
- 7: 输出更新后的跟踪目标列表 updated\_track\_list

**2.3 猪只粪便精细化识别网络(AWF-MLP)**

粪便与排便行为识别的精度与效率至关重要,但仅依赖检测网络往往易导致较高的粪便误检率。多层感知机(Multi-layer perceptron, MLP)<sup>[23]</sup>作为一种前馈神经网络,能够通过线性变换和非线性激活学习复杂判别特征,常用于分类任务。然而,在实际猪舍环境中,由于粪便形态多样、颜色差异显著,单层 MLP 难以充分表征粪便特征。为此,提出一种自适应加权融合(Adaptive weighted fusion, AWF)与多层 MLP 相结合的 AWF-MLP 精细识别网络,通过构建 3 个针对正负样本具备不同感知能力的独立 MLP 分类层,对粪便样本(正样本)和非粪便样本(负样本)进行精细化重识别,从而有效降低误检率,如图 3 所示。

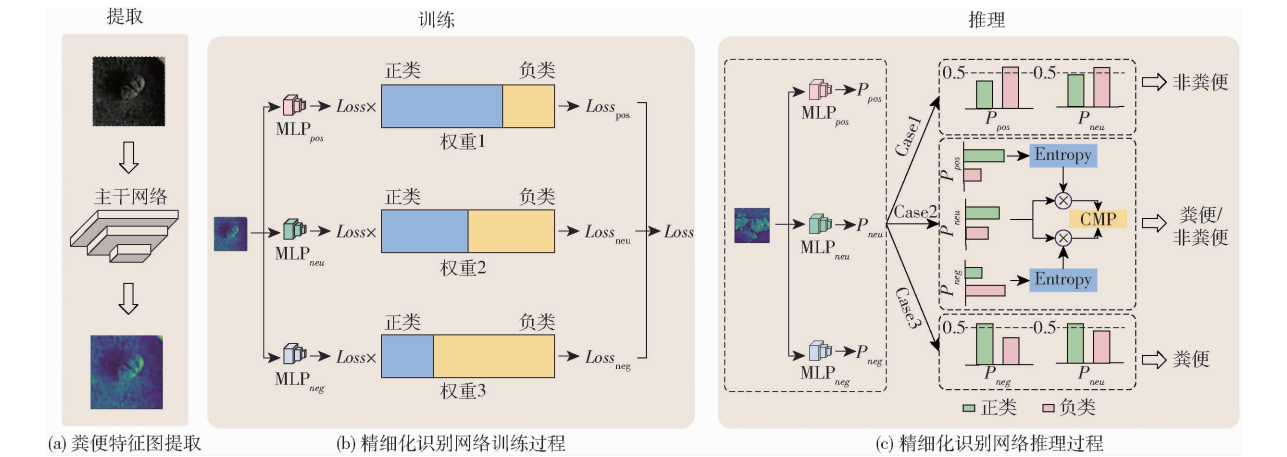


图3 自适应加权多层感知机融合网络

Fig.3 Adaptive weighted multilayer perceptron fusion network

由图 3a 可知,为兼顾重识别模块的实时性,采用 MobileNetV2<sup>[24]</sup>作为主干网络提取粪便的特征图。AWF-MLP 网络的训练过程和推理过程分别如图 3b、3c 所示。在训练过程中,3 个 MLP 基于正负损失偏差进行独立训练,从而学习到对正负样本的差异化判别能力。具体而言,MLP<sub>pos</sub>通过增大对正样本的分类损失权重从而增强对正样本的识别能力,损失函数公式为

$$Loss_{pos} = -\frac{1}{n} \sum_{i=1}^n [(1 + \theta)y_i \lg(p_i) + (1 - \theta)(1 - y_i) \lg(1 - p_i)] \tag{1}$$

式中  $n$ ——样本数量  
 $p_i$ ——第  $i$  个样本预测为正样本的概率  
 $\theta$ ——损失数值偏置  
 $y_i$ ——第  $i$  个样本的真实标签值  
类似地,MLP<sub>neg</sub>通过增大对负样本的分类损失权重

从而增强对负样本的识别能力,其损失函数为

$$Loss_{neg} = -\frac{1}{n} \sum_{i=1}^n [(1-\theta)y_i \lg(p_i) + (1+\theta)(1-y_i) \lg(1-p_i)] \quad (2)$$

为了更平衡地融合各 MLP 分支的预测结果,训练  $MLP_{neu}$  作为基线预测,即正负样本损失权重均设为 1,即

$$Loss_{neu} = -\frac{1}{n} \sum_{i=1}^n [y_i \lg(p_i) + (1-y_i) \lg(1-p_i)] \quad (3)$$

最后将 3 个分支 MLP 损失融合得到整体损失函数  $Loss_{total}$ ,即

$$Loss_{total} = Loss_{pos} + Loss_{neg} + Loss_{neu} \quad (4)$$

随后,通过反向传播机制<sup>[25]</sup>对网络参数进行更新,模型得以训练出 3 个具备不同分类优势的 MLP:  $MLP_{pos}$  更擅长正样本的学习;  $MLP_{neu}$  具有均衡的正负样本学习能力;  $MLP_{neg}$  更擅长负样本的学习。

在推理过程中,该模型综合考虑 3 个对正负样本具有不同识别能力的 MLP 的预测结果( $P_{pos}, P_{neu}, P_{neg}$ ),并结合信息熵进行辅助判断,得出最终结果。由于这 3 个 MLP 在预测结果上的差异相互约束,模型最终可能产生以下 3 种输出结果之一,即

$$Y = \begin{cases} 0 & (P_{pos}(0) > 0.5 \text{ 且 } P_{neu}(0) > 0.5) \\ 1 & (P_{neg}(1) > 0.5 \text{ 且 } P_{neu}(1) > 0.5) \\ \frac{0}{1} & (\text{其他}) \end{cases} \quad (5)$$

式中  $P_{pos}(0), P_{neu}(0)$ —— $MLP_{pos}$  和  $MLP_{neu}$  输出结果中正样本概率

$Y$ ——模型最终预测结果  
 $P_{neg}(1), P_{neu}(1)$ —— $MLP_{neg}$  和  $MLP_{neu}$  输出结果中负样本概率

对于第 1 种情况,如果擅长学习正样本的  $MLP_{pos}$  和具有均衡学习能力的  $MLP_{neu}$  对负样本的预测概率均大于 0.5,则结果为负样本;对于第 2 种情况,如果擅长学习负样本的  $MLP_{neg}$  和  $MLP_{neu}$  对正样本的预测概率均大于 0.5,则结果为正样本。特别地,对于第 3 种情况,其输出结果需结合信息熵进行计算,具体计算过程为

$$Y = \begin{cases} 1 & (P_{neu}(1) \exp(-H_{pos}) > P_{neu}(0) \exp(-H_{neg})) \\ 0 & (\text{其他}) \end{cases} \quad (6)$$

其中  $H_{pos} = \sum_{i=1}^n (-P_{pos}(i) \ln(P_{pos}(i) + 1e-5))$

(7)

式中  $H_{pos}, H_{neg}$ —— $MLP_{pos}$  和  $MLP_{neg}$  预测结果概率分布信息熵

将式(7)中  $P_{pos}$  替换为  $P_{neg}$ ,可得  $H_{neg}$ 。由式(6)可知,通过对  $H_{pos}$  和  $H_{neg}$  进行负指数变换便可得  $H_{pos}$  和  $H_{neg}$  置信度  $\exp(-H_{pos})$  和  $\exp(-H_{neg})$ <sup>[26]</sup>。基于置信度和  $MLP_{neu}$  的正负类别预测概率的乘积对比分析便可得出最终预测类别。通过上述这种融合策略,模型能够在不同情况下更准确地分类样本。

2.4 基于 SAIR 的猪只排便行为识别

以某次排便行为识别和统计过程为例,进行详细解释,流程如图 4 所示。

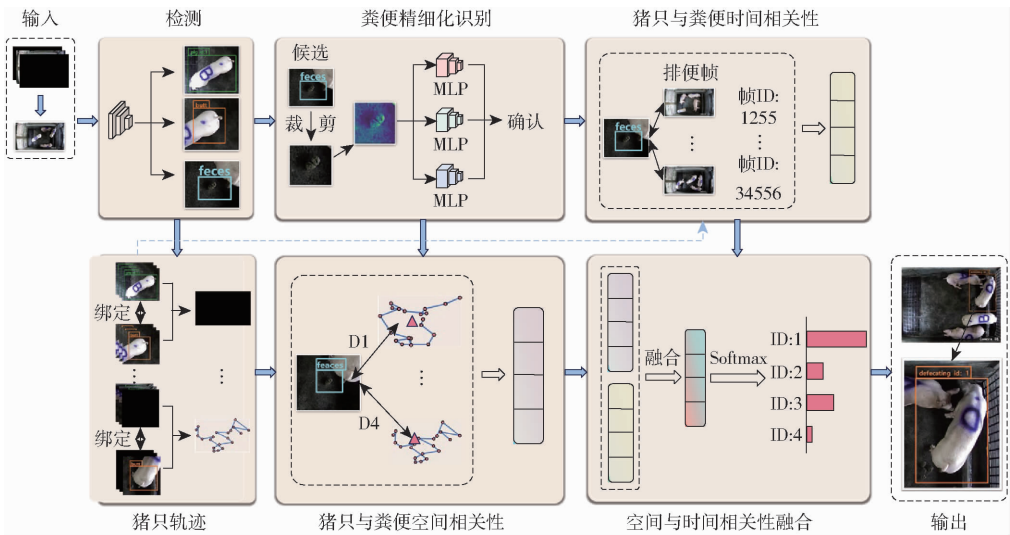


图 4 猪只个体排便行为识别流程

Fig.4 Process of identifying pig defecation behavior

以频率 1 f/s 从视频序列中提取帧,并采用 SAE-YOLO v8 对猪只个体、臀部及粪便进行检测。

对检测到的猪只个体使用 IDR-BSORT 进行多目标跟踪,并基于猪只轨迹与猪只臀部检测框的空间关

系,实现猪只臀部与对应猪只 ID 的绑定,从而间接获得猪只臀部轨迹。对检测到的粪便,首先基于欧氏距离与历史粪便记录进行空间筛选,以避免重复统计;若新检测到的粪便与历史记录中的粪便之间的欧氏距离大于阈值  $\eta$ ,则该粪便被判定为新目标并进入精细化重识别流程。随后采用 AWF-MLP 网络对其进行重识别,过程可参考 2.3 节。若确认其为粪便,则将其存入粪便列表。其中,为合理设定粪便新旧判定阈值  $\eta$ ,对数据集中粪便样本进行统计分析。统计结果显示,粪便检测框的平均宽度与高度分别为 31 像素和 39 像素,作为单个粪便目标的平均尺寸。在最易混淆的极端情况下,即两坨粪便检测框仅在角点相接时,其中心点欧氏距离为 48.6 像素。为预留一定容错空间并采用整十刻度,向上扩展至 50 像素作为阈值  $\eta$ 。当两坨粪便中心距离大于 50 像素时,判定其更属于不同排泄事件。

计算新检测粪便与各猪只个体之间的空间和时间相关性。空间相关性是指粪便与猪只个体在空间距离上的关联,时间相关性则衡量了粪便在时间序列中属于特定猪只个体的可能性。当检测到新粪便时,算法向前回溯 50 帧,获取这 50 帧内所有猪只臀部轨迹数据,并通过 K-means 聚类算法<sup>[27]</sup> 计算得到每个猪臀部的轨迹聚类中心。聚类中心可以通过公式  $\mu_k = \left( 1/t \sum_{j=1}^t x_j, 1/t \sum_{j=1}^t y_j \right)$  得到,其中  $\mu_k$  表示第  $k$  个猪只臀部的轨迹聚类中心, $t$  表示猪只臀部的轨迹数据点的数量, $x_j$  和  $y_j$  分别表示该猪只臀部的第  $j$  个轨迹数据点的横坐标和纵坐标。接着,计算新粪便检测框中心点与每个猪只臀部的轨迹聚类中心点之间的欧氏距离,即

$$dist_{new,k} = \sqrt{(x_{new} - \mu_k^x)^2 + (y_{new} - \mu_k^y)^2} \quad (8)$$

式中  $\mu_k^x, \mu_k^y$ ——第  $k$  个猪只臀部轨迹聚类中心点横坐标和纵坐标

$dist_{new,k}$ ——新粪便检测框中心点  $(x_{new}, y_{new})$  与第  $k$  个猪只臀部的轨迹聚类中心点  $(\mu_k^x, \mu_k^y)$  之间的欧氏距离

基于上述计算得到的距离,评估新粪便与各猪只之间的空间相关性,可通过公式  $a_{new,k} = [1 - (dist_{new,k}/dist_{max})]$  得到,其中  $dist_{max}$  表示当前新粪便与所有猪只臀部轨迹聚类中心中最大的欧氏距离, $a_{new,k}$  表示当前新粪便与第  $k$  头猪只之间的空间相关性。

然后构建新粪便与猪只的空间相关性向量,即

$$\mathbf{A}_{new} = (a_{new,1}, a_{new,2}, \cdots, a_{new,n})^T \quad (9)$$

式中  $\mathbf{A}_{new}$ ——新粪便与所有猪只空间相关性向量

接着计算新粪便与各猪只的时间相关性。历史记录粪便序列与各猪只的空间相关性向量矩阵  $\mathbf{M}$ ,

即

$$\mathbf{M} = [\mathbf{A}_1 \quad \mathbf{A}_2 \quad \cdots \quad \mathbf{A}_m] = \begin{bmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{bmatrix} \quad (10)$$

式中  $\mathbf{A}_1$ ——已记录的粪便中第 1 个粪便与所有猪只的空间相关性向量

$m$ ——已记录的历史粪便总数

$a_{m,n}$ ——第  $m$  个历史记录粪便与  $n$  号猪只空间相关性

接着,根据猪只排泄具备时间间隔性这个特点,为矩阵  $\mathbf{M}$  中每个历史粪便的空间相关性向量赋予不同时序权重得到时序相关性向量矩阵  $\tilde{\mathbf{M}}$ ,即

$$\tilde{\mathbf{M}} = \mathbf{M} \left( \frac{m}{m}, \frac{m-1}{m}, \cdots, \frac{1}{m} \right) = \begin{bmatrix} \mathbf{A}_1 \frac{m}{m} & \mathbf{A}_2 \frac{m-1}{m} & \cdots & \mathbf{A}_m \frac{1}{m} \end{bmatrix} \quad (11)$$

由式(11)可看出,若某头猪只在早期已进行过排泄,则当前新检测到的粪便属于该猪只的可能性就越高。最后,综合新粪便与每头猪只个体之间的空间相关性向量  $\mathbf{A}_{new}$  和时序相关性向量矩阵  $\tilde{\mathbf{M}}$ ,计算出猪只得得分向量  $\mathbf{S}$ ,即

$$\mathbf{S} = \mathbf{A}_{new} + \tilde{\mathbf{M}}\mathbf{v} \quad (12)$$

式中  $\mathbf{v}$ —— $m \times 1$  的全 1 向量

$\tilde{\mathbf{M}}\mathbf{v}$ ——对  $\tilde{\mathbf{M}}$  每一行进行求和得到的列向量  
 $\mathbf{S}$ —— $n \times 1$  的向量

$\mathbf{S}$  中最大得分所对应的猪只下标即被判定为发生排便行为的猪只,因此该猪只排便行为次数加 1。在完成当前排泄行为识别与统计后,新粪便与各猪只之间的空间相关性向量  $\mathbf{A}_{new}$  也会更新到历史记录的粪便序列与所有猪只的空间相关性向量矩阵  $\mathbf{M}$  中,更新公式为  $\mathbf{M} = \mathbf{M} \parallel \mathbf{A}_{new}$ ,其中  $\parallel$  表示矩阵拼接操作。

上述猪只排便行为识别过程综合利用猪只个体与粪便之间的空间和时间相关性,不仅能够精准识别正在排便的猪只个体,还能有效进行相关统计分析。这种方法为监测猪只健康状况和行为模式提供了可靠依据。

3 试验结果与分析

3.1 试验环境

通过试验验证所提 SAE-YOLO v8 检测模型、IDR-BSORT 跟踪方法以及 AWF-MLP 网络在

SAIR 框架中的有效性,并在短视频与长视频场景下进行测试评估。所有试验均基于配备 RTX 4070s GPU 的服务器,采用 Pytorch 进行模拟计算。

3.2 评价指标

在目标检测领域,平均精度 (Average precision, AP) 和平均精度均值 (mean Average Precision, mAP) 是衡量模型性能的常用评价指标<sup>[28]</sup>。其中,AP 用于评估模型在单一类别上的检测能力,mAP 则对各类别 AP 取平均,以反映模型的整体检测性能。较高的 AP 和 mAP 表明模型在目标定位与类别判别方面均具有较好的综合表现,是目标检测模型评估中的核心指标。mAP 的计算通常基于 IoU (Intersection over union),并通过设定 IoU 阈值区分有效预测与无效预测。其中,mAP@0.5 表示当 IoU 阈值为 0.5 时计算的平均精度,能够较为直观地反映模型在实际场景中的检测能力。同理,AP@0.5 表示当 IoU 阈值为 0.5 时,模型对单一类别的检测精度。因此,选择 AP@0.5 和 mAP@0.5 作为主要评价指标,对模型检测性能进行分析与验证。

在多目标跟踪任务中,常用评估指标包括 HOTA<sup>[29]</sup>、MOTA<sup>[30]</sup> 和 IDF1<sup>[31]</sup>,分别衡量跟踪的综合性能、检测准确性和目标身份一致性。HOTA 综合考虑检测、跟踪一致性和轨迹关联,提供了对跟踪算法的全面评估,MOTA 评估多目标检测的整体表现,IDF1 则反映目标身份保持能力。

F1 值、召回率 (Recall) 和精确率 (Precision) 用于评估本文提出的 SAIR 框架在长视频中的性能,其中,F1 值是召回率和精确率的调和平均数,兼顾两者的平衡,当二者差距较大时,F1 值偏低,可反映模型在某一方面的不足。

3.3 SAE-YOLO v8 性能

在不同规模的 YOLO v8 上进行试验,结果如表 3 所示。

表 3 YOLO v8 系列不同量级版本试验结果  
Tab.3 Experiment results of different-scale versions within YOLO v8 series %

| 模型           | mAP@0.5 | AP@0.5 |      |      |
|--------------|---------|--------|------|------|
|              |         | 个体     | 臀部   | 粪便   |
| YOLO v8n     | 92.7    | 99.5   | 99.2 | 79.5 |
| YOLO v8s     | 93.0    | 99.5   | 99.2 | 80.3 |
| YOLO v8m     | 93.4    | 99.7   | 99.3 | 81.2 |
| SAE-YOLO v8n | 94.1    | 99.5   | 99.2 | 83.6 |
| SAE-YOLO v8s | 94.3    | 99.6   | 99.2 | 84.2 |
| SAE-YOLO v8m | 94.7    | 99.7   | 99.3 | 85.3 |

从表 3 可看出,较大规模的 YOLO v8 版本比其轻量化版本仅在精度上取得了有限提升,而在引入

SAE 机制后,各版本均实现了一定性能增强。综合考虑检测精度与推理效率,选取 YOLO v8n 作为基线模型。多种同量级的主流目标检测框架上进行对比试验,包括 YOLO v7n、YOLO v8n、YOLO 11n 及 Faster R-CNN,试验结果如表 4 所示。

表 4 不同检测模型及其空间注意力增强版本试验结果  
Tab.4 Experiment results of different detection models and their SAE-enhanced versions %

| 模型               | mAP@0.5 | AP@0.5 |      |      |
|------------------|---------|--------|------|------|
|                  |         | 个体     | 臀部   | 粪便   |
| Faster R-CNN     | 80.0    | 99.5   | 98.2 | 42.4 |
| YOLO v7n         | 91.2    | 99.7   | 99.5 | 74.5 |
| YOLO v8n         | 92.7    | 99.5   | 99.2 | 79.5 |
| YOLO v11n        | 90.5    | 99.5   | 99.2 | 73.0 |
| SAE-Faster R-CNN | 81.2    | 99.7   | 98.2 | 45.8 |
| SAE-YOLO v7n     | 91.8    | 99.8   | 99.4 | 76.2 |
| SAE-YOLO 11n     | 92.2    | 99.5   | 99.3 | 77.9 |
| SAE-YOLO v8n     | 94.1    | 99.5   | 99.2 | 83.6 |

由表 4 可知,YOLO v8n 的性能最优异,其基础版本在 mAP@0.5 达到 92.7%,整体表现均优于 YOLO v7n、YOLO 11n 及 Faster R-CNN 等主流检测网络。值得说明的是,在本研究场景中,YOLO 11n 在超参数变化下的性能波动幅度更大,而 YOLO v8n 始终保持更稳定的检测精度。因此,在统一训练配置与公平对比的前提下,YOLO v8n 具有更稳健的表现,是本研究中作为基线模型的合理选择。为叙述简洁起见,将 YOLO v8n 简称为 YOLO v8。由表 4 可知,所提 SAE 机制在不同检测算法中均能有效提升整体检测性能,如在 YOLO v8 中引入 SAE 机制后,SAE-YOLO v8 的 mAP@0.5 提升至 94.1%,粪便检测精度提升至 83.6%,实现全系列模型中最优性能。上述结果表明,YOLO v8 具备更强的特征表达能力和检测稳定性,是本研究作为基线模型的合理选择,同时验证了引入 SAE 机制能够增强模型对关键目标区域的关注能力,尤其在粪便等小目标检测任务中显著提升检测精度与实用价值。为更直观展示 SAE 的作用效果,基于 GradCAM 算法<sup>[32]</sup>绘制了引入 SAE 前后的特征响应热力图,如图 5 所示。

由图 5 可知,改进前模型在特征关注上存在明显偏差。具体来说,图 5b 中 1 号红框内的猪只热力图明显扩散至目标轮廓之外,表明模型在猪只个体边界识别上存在偏差;图 5e 中,1 号红框内猪只臀部热力图扩散到了猪只腹部,而 2 号红框的热力图错误聚焦于猪只头部,出现明显的语义混淆;图 5h 中,粪便热力图未能有效聚焦于真实粪便位置,而是扩散至地面上的疑似粪便处,这进一步表明了模型的局限性。引入 SAE 机制后,图 5c、5f、5i 所

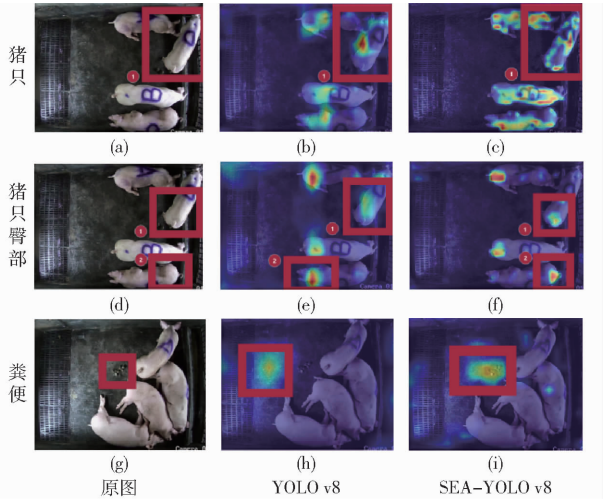


图5 目标检测模型热力图

Fig.5 Heatmaps of object detection model

示的热力图在目标区域内呈现出更加集中的响应。图5c中的猪只个体热力图清晰地显示出猪只的轮廓,而图5f、5i中的猪只臀部及粪便热力图也均集中在正确的位置。通过可视化效果,可以直观地比较模型改进前后的差异,并验证改进措施的有效性。

3.4 IDR-BSORT 性能

选取  $D_{vid}^{(1)}$  数据集测试多目标跟踪算法性能,并将所提 IDR-BSORT 算法与基线算法 BoT-SORT 进行试验对比,结果如表5所示。

表5 改进前后猪只多目标跟踪算法性能

Tab.5 Performance of pig multi-object tracking before and after improvement %

| 指标   | BoT-SORT |       | IDR-BSORT |       |
|------|----------|-------|-----------|-------|
|      | 视频1      | 视频2   | 视频1       | 视频2   |
| HOTA | 87.64    | 71.13 | 90.61     | 76.59 |
| MOTA | 96.61    | 82.02 | 96.62     | 82.03 |
| IDF1 | 94.62    | 82.58 | 98.28     | 90.13 |

由表5可知,改进后猪只多目标跟踪算法 IDR-BSORT 在 HOTA、MOTA 和 IDF1 3 个关键指标上均优于原始的多目标跟踪算法 BoT-SORT。在 HOTA 指标上,IDR-BSORT 在视频1和视频2上分别达到 90.61% 和 76.59%,较 BoT-SORT 提升 2.97、5.46 个百分点,这说明 IDR-BSORT 在目标跟踪的一致性和轨迹关联方面表现更加出色。在 IDF1 指标上,IDR-BSORT 在视频1和视频2上分别提升至 98.28% 和 90.13%,分别提升 3.66、7.55 个百分点,这表明 IDR-BSORT 在目标身份一致性方面表现更佳,特别是在复杂猪舍环境下优势明显。为更直观地展示 IDR-BSORT 的跟踪效果,图6为视频1和视频2中猪只 ID 变化情况。

由图6可知,改进后算法能够保持预测 ID 与真

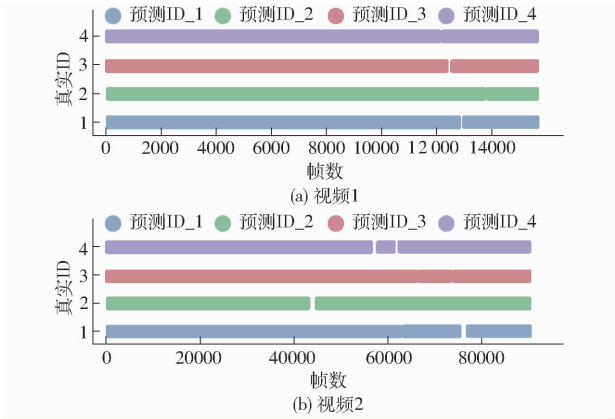


图6 ID变化的可视化分析

Fig.6 Visualization of ID changes

实 ID 的高度一致,断开部分为漏检或匹配错误,可以直观观察到 IDR-BSORT 在复杂环境中保持身份一致性的能力。总体来看,在 3 个评估指标上的显著提升,尤其在更为复杂的视频2环境中,充分体现了 IDR-BSORT 在跟踪一致性、轨迹关联性及身份保持方面的显著优势。

3.5 AWF-MLP 性能

3.5.1 精细化识别模型作用

精细化识别模块旨在降低检测模型误检率。前文已验证 SAE-YOLO v8 在猪只检测任务中表现最优,基于 SAE-YOLO v8 进一步试验,评估 AWF-MLP 模块对粪便精细化识别的提升效果。选取  $D_{img}^{(2)}$  测试数据集,用于测试添加 AWF-MLP 前后粪便检测效果,结果如表6所示。

表6 添加 AWF-MLP 模型前后粪便检测效果

Tab.6 Feces detection performance before and after adding AWF-MLP %

| 模型                    | mAP@0.5 | AP@0.5 |
|-----------------------|---------|--------|
| SAE-YOLO v8           | 94.1    | 83.6   |
| SAE-YOLO v8 + AWF-MLP | 95.3    | 87.4   |

由表6可知,引入 AWF-MLP 后,整体 mAP@0.5 提升 1.2 个百分点,粪便检测的 AP@0.5 提升至 87.4%,提高 3.8 个百分点,表明 AWF-MLP 通过自适应加权增强了粪便特征识别能力,有效降低误检,同时保持对猪只个体和臀部的检测性能。

3.5.2 AWF-MLP 有效性

为进一步验证 AWF-MLP 的优越性,选取  $D_{img}^{(2)}$  数据集,对 MobileNetV2、MobileNetV2-MLP (将 MobileNetV2 最后全连接层替换为单层 MLP) 及 AWF-MLP 的 3 种精细化识别模型性能进行测试,结果为 MobileNetV2、MobileNetV2-MLP 和 AWF-MLP 准确率分别为 89.9%、94.1% 和 99.5%。由此结果可知,AWF-MLP 性能存在显著优势。

MobileNetV2 模型准确率为 89.9% ,引入单层 MLP 的 MobileNetV2 – MLP 提升至 94.1% ,但仍难满足精细化识别的需求,可能因单层 MLP 无法充分捕捉形态和颜色多变的粪便特征。相比之下,AWF – MLP 准确率达 99.5% ,多层 MLP 及基于正负损失偏差的独立训练显著增强了模型在复杂场景下的适应能力和鲁棒性。为更直观地体现 AWF – MLP 的优势,选取部分粪便和非粪便图像,并分别使用 MobileNetV2、MobileNetV2 – MLP 和 AWF – MLP 模型对其进行识别,结果如图 7(其中√表示正确,×表示错误)所示。

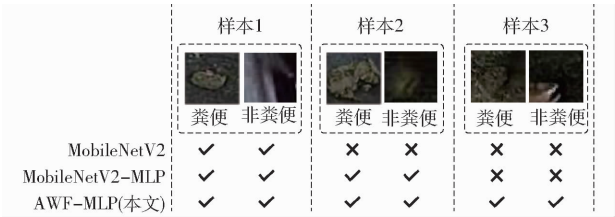


图 7 不同模型的识别结果

Fig. 7 Recognition results across different models

表 7 AWF – MLP 与代表性模型准确率及推理速度评估结果

Tab. 7 Accuracy and inference speed results of AWF – MLP and representative models

| 参数    | ResNet18 | ResNet34 | ResNet50 | VGG19 | ViT   | AWF – MLP (本文) |
|-------|----------|----------|----------|-------|-------|----------------|
| 准确率/% | 90. 5    | 92. 9    | 94. 6    | 97. 0 | 97. 5 | 99. 5          |
| 耗时/s  | 0. 24    | 0. 29    | 0. 41    | 0. 51 | 0. 54 | 0. 38          |

由表 7 可知,AWF – MLP 在识别准确率与推理效率之间实现了优越平衡。其识别准确率达 99.5% ,高于 ViT (97.5% ) ,并显著超越 ResNet50 (94.6% ) 、ResNet34 (92.9% ) 、ResNet18 (90.5% ) 以及 VGG19 (97.0% ) ,显示出卓越的特征表征能力。在推理效率方面,AWF – MLP 可在 0.38 s 内完成对 100 幅图像的推理,优于 ViT(0.54 s)和 VGG19 (0.51 s) ,兼顾高精度与实时性,展现出良好的实际部署潜力。

3.6 短视频下 SAIR 识别猪只排便结果分析

准确率、召回率和精确率用于评估所提 SAIR 框架在短视频数据集上识别猪只排便行为的性能。选取有 110 个含猪只排便行为和 110 个不含猪只排便行为视频片段数据集  $D_{vid}^{(2)}$  进行测试,测试结果混淆矩阵如图 8 所示。

由图 8 可知,在 220 个视频片段中,共有 215 个被正确识别,其中有 3 个含排便行为片段被错误识别为不含排便行为,另有 2 个不含排便行为片段被错误识别为含排便行为。由此计算得准确率、召回率和精确率分别为 97.7% 、97.2% 和 98.1% 。测试结果表明,提出的 SAIR 框架在短视频数据集上性能优异。尽管存在有少量误判,但总体识别效果依

图 7 为 3 个模型对 3 组样本的识别结果,每组样本均包含 1 幅粪便图像和 1 幅非粪便图像。粪便图像具有被踩踏变形以及颜色多变的特征,非粪便图像呈现与粪便相似的轮廓和颜色。3 组样本识别难度由低到高为样本 1、样本 2 和样本 3。MobileNetV2 仅能正确识别样本 1,对样本 2、3 表现不稳定,显示其在复杂背景下的局限性。引入单层 MLP 的 MobileNetV2 – MLP 能够正确识别样本 1、2,但对样本 3 仍有误判,说明单层 MLP 对高相似度样本分类能力仍有限。相比之下,AWF – MLP 在所有样本中均能准确识别,这证明了该模型在面对复杂形态和颜色变化时的优异性能。为进一步验证 AWF – MLP 的性能与适用性,在相同数据集下,将其与 ResNet18/34/50<sup>[33]</sup>、VGG19<sup>[34]</sup> 以及 ViT<sup>[35]</sup> 等模型进行对比。从识别准确率和推理效率 2 个方面进行评估,其中推理效率以处理 100 幅随机选取图像总耗时为衡量标准,结果见表 7。

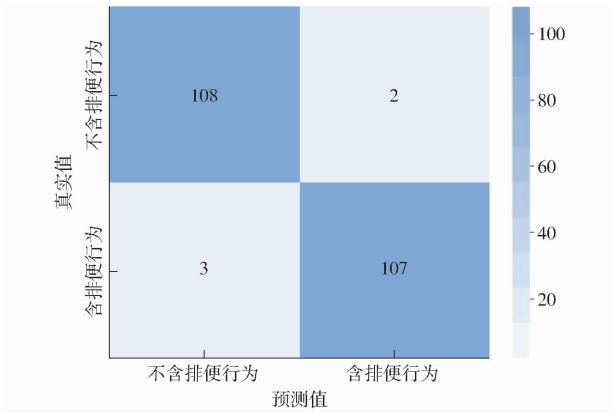


图 8 视频片段排便行为识别结果混淆矩阵

Fig. 8 Confusion matrix of defecation behavior recognition results

然较好。此外,考虑到本文所使用排便行为短视频片段时长为 4 ~ 30 s,良好的识别性能表明该框架具备实时监测排便行为的能力,具备较强实用性。

3.7 长视频下 SAIR 识别猪只排便结果分析

3.7.1 长视频下 SAIR 识别猪只排便的性能

选取  $D_{vid}^{(3)}$  中连续一周白天视频对 SAIR 框架进行测试,结果如图 9 所示。

如图 9 所示,图中蓝色块下方字母对应实际排便猪只编号,粉色块上方字母对应检测到的排便猪

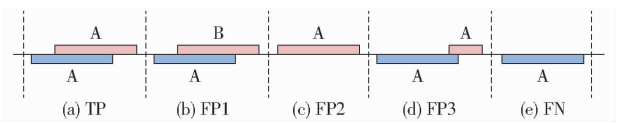
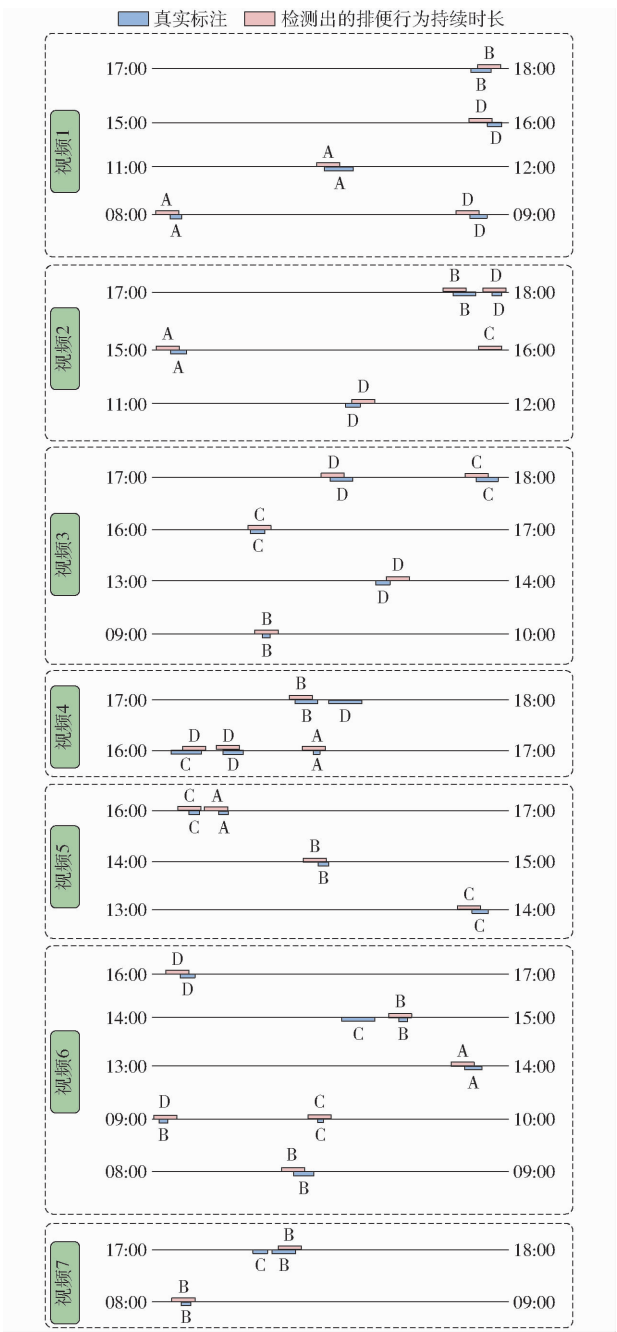


图 10 考虑 tICR 的 TP、FP 和 FN  
Fig. 10 Consideration of TP, FP, and FN in tICR

主流研究亦普遍沿用这一设定<sup>[37]</sup>。图 10a 的 TP 表示检测出的排便行为在持续时长上与真实标注的排便行为具有  $tICR > 0.5$ , 且对应的猪只编号一致; 图 10b 的 FP1 表示检测结果与真实标注在时间上具有  $tICR > 0.5$ , 但猪只编号不一致; 图 10c 的 FP2 表示检测到排便行为, 但该时间段内不存在真实标注的排便行为; 图 10d 的 FP3 表示检测结果与真实标注在时间上具有  $tICR < 0.5$ ; 图 10e 的 FN 表示在真实标注的排便行为时间范围内未检测到对应的排便行为。根据上述对 TP、FP、FN 的定义, 统计了各长视频检测到排便行为的 TP、FP、FN, 如表 8 所示。

表 8 从 7 个长视频中获得的排便行为识别结果  
Tab. 8 Recognition results of defecation behavior obtained from seven long videos

| 类别 | 视频 |   |   |   |   |   |   | 总计 |
|----|----|---|---|---|---|---|---|----|
|    | 1  | 2 | 3 | 4 | 5 | 6 | 7 |    |
| TP | 4  | 4 | 4 | 4 | 4 | 5 | 2 | 27 |
| FP | 1  | 1 | 1 | 0 | 0 | 1 | 0 | 4  |
| FN | 0  | 0 | 0 | 1 | 0 | 1 | 1 | 3  |

由表 8 可计算得到 F1 值、召回率和精确率分别为 88.5%、90.0% 和 87.1%。由表 8 可知, 在长视频中 SAIR 框架能够较准确地识别猪只排便行为, TP 数量较高, 表明大多数真实排便事件被成功检测。例如, 视频 1~5 中 TP 均为 4, 视频 6 中 TP 达到 5。与此同时, FP 数量相对较少, 视频 4、5 和 7 中未出现误检, 说明模型具备较好的判别能力。FN 仅在少数视频中出现, 且次数较低, 反映出系统具有较高的召回性能。综合来看, 较高的 F1 值表明该方法在检测准确性与完整性之间取得了良好平衡, 验证了 SAIR 框架在猪只排便行为识别任务中的有效性与鲁棒性。

### 3.7.2 在长视频下 SAIR 时间性能

为评估在长视频场景下 SAIR 框架实时性能, 对 7 个长视频进行测试, 如图 11 所示。由图 11 可知, 在所有长视频上 SAIR 框架运行时长均低于其真实时长的一半。其中视频真实时长为 10 h, 运行时长为 4.69~4.92 h。该高效性主要源于 2 方面优化设计: ①重识别模型 AWF-MLP 以 MobileNetV2 作为主干网络, 有效降低了计算复杂度。②在跟踪阶段以轻量化的身份重绑定 (IDR) 机制替代 BoT-SORT 原有的 ReID 模块, 减少了额外的特征匹配开

只编号。由于本文框架在识别到粪便时才开始判定此时是否发生排便行为, 检测结果具有滞后性。结合试验视频长期观察, 猪只排便平均持续约为 40 s, 因此将检测持续时长统一设置为检测帧前 28 s 和后 12 s 的时间范围, 该划分基于不同猪舍环境下的观察结果, 具有一定普适性。为进行定量分析, TP、FP、FN 定义如图 10 所示。

将检测出的排便行为持续时长与实际标注的排便行为的持续时长的交集覆盖率 ( $tICR$ ) 阈值设定为 0.5, 主要是因为目标检测和视频行为识别领域, 0.5 已成为一个广泛接受的标准阈值<sup>[36]</sup>, 相关

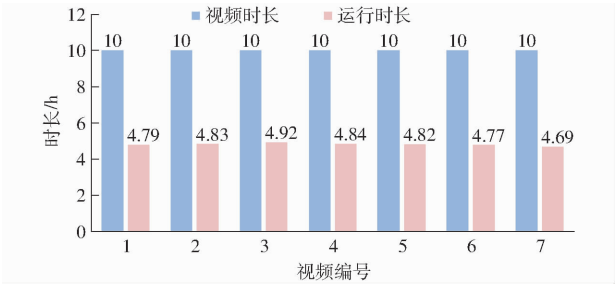


图 11 7 个长视频的测试运行时间

Fig. 11 Testing runtime of seven long videos

销。由此可见,SAIR 框架不仅能够准确识别猪只排便行为,还具备较高的处理效率,可满足实时监控猪只排便需求。

3.7.3 猪只排便行为时间分布

对长视频中检测到的真实排便行为进行时间统计分析,如图 12 所示。

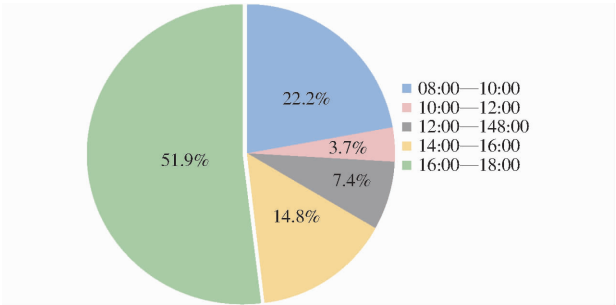


图 12 排便行为的时间统计结果

Fig. 12 Temporal statistical analysis of defecation behavior

由图 12 可知,排便行为在不同时段内呈现明显差异。08:00—10:00 时段占比 22.2%,表明早晨为排便相对活跃时段,可能与夜间休息后活动和进食增加有关。10:00—14:00 期间排便行为减少(占比 3.7%~7.4%),可能与猪只活动降低及午休习惯相关。14:00—16:00 排便占比回升至 14.8%,表明下午活动逐渐增强。16:00—18:00 为排便行为最集中时段,占比高达 51.9%,可能与下午活动高峰及日间消化过程接近尾声有关。总体来看,猪只排便行为在早晨和下午呈双峰分布特征,尤其在傍晚最显著,为合理安排猪舍清洁与管理时序提供了依据。

3.7.4 猪只排便行为空间分布

为分析猪只排便习性,对长视频测试中检测到的真实排便行为进行了空间统计,并在猪舍平面图上展示了粪便分布情况,如图 13 所示。

由图 13 可知,排泄区(绿色框 1)内粪便分布较为稀疏,仅检测到少量样本,表明猪只并未严格在设定的排泄区内排便;相比之下,进食区(绿色框 2)附近粪便分布更集中,显示猪只倾向于在食槽周边发生排便行为。这一现象可能与进食后生理状态变化

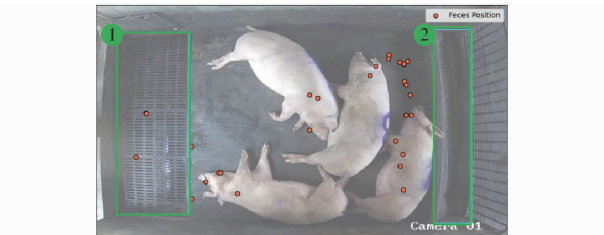


图 13 粪便在猪舍地图上的分布情况

Fig. 13 Distribution of feces on pigpen

有关,如进食后血糖上升<sup>[38]</sup>导致的慵懒状态,使得猪只倾向于在食槽附近停留并排便。该发现为猪舍空间布局优化提供了参考,可考虑在喂食槽附近合理设置排泄区,以提升空间利用效率并改善粪便管理水平。此外,猪舍墙边及中部区域亦存在零散粪便分布,说明猪只排便行为具有一定随机性,并不完全受限于固定区域。上述空间统计结果揭示了猪只排便行为的空间偏好与随机特征,为猪舍功能分区优化及粪污集中管理提供了有价值的参考,有助于提升养殖环境卫生水平与管理效率。

3.7.5 猪只排便行为排便频率分析

为评估所提 SAIR 框架在个体层面的识别与统计准确性,采用表格对长视频中各猪只的检测排便次数与真实排便次数(Ground truth)进行了对比分析,如表 9 所示。其中,真实排便次数由人工逐帧观察视频并进行统计得到,也是相关研究中常用的真实标注方法。

表 9 检测与真实排便行为次数对比

Tab. 9 Comparison of detected defecation events and ground truth defecation events 次

| 猪只编号 | 检测值 | 真实值 |
|------|-----|-----|
| A    | 6   | 6   |
| B    | 9   | 10  |
| C    | 6   | 8   |
| D    | 8   | 9   |

表 9 为 4 头猪只(A、B、C、D)在 7 d 内的总排便次数检测结果与真实标注结果。结果表明,猪只 A 检测值与真实值完全一致,均为 0.86 次/d,准确率达到 100%。猪只 B、C、D 检测值与真实值存在一定偏差,其中猪只 B 检测频率为 1.29 次/d,真实频率为 1.43 次/d,准确率为 90%;猪只 C 检测频率为 0.86 次/d,真实值为 1.14 次/d,准确率为 75%,为 4 头猪只中最低,可能与粪便误检或漏检有关;猪只 D 检测频率为 1.14 次/d,真实频率为 1.28 次/d,准确率为 89.0%。总体而言,模型检测结果与真实排便次数具有较高一致性,4 头猪只平均准确率为 88.5%,表明该方法能够较为有效地识别排便行为,但在降低漏检率方面仍有进一步提升空间。

4 结束语

针对猪只排便姿态不明显、个体间遮挡频繁及粪便形态变化显著等问题,提出了一种基于时空注意力反演推理的猪只排便行为自动识别与统计框架(SAIR)。该框架以粪便检测为触发点,融合SAE-YOLO v8检测、IDR-BSORT多目标跟踪与AWF-

MLP精细化识别,结合空间与时间相关性矩阵,实现对排便猪只成功识别。该框架在短视频测试中的准确率、召回率和精确率分别达到97.7%、97.2%和98.1%;在长视频测试中,当tICR大于0.5时,F1值、召回率和精确率分别为88.5%、90.0%和87.1%。这些结果表明该框架具备良好的鲁棒性与实用性。

参 考 文 献

[1] NEETHIRAJAN S. AI in sustainable pig farming: IoT insights into stress and gait[J]. Agriculture, 2023, 13(9):1706.

[2] 全志民,徐天哲,石传森,等. 基于CBCW YOLO v8的猪只行为识别方法研究[J]. 农业机械学报, 2025, 56(2): 411-419.

TONG Zhimin, XU Tianzhe, SHI Chuanmiao, et al. Pig behavior recognition based on CBCW YOLO v8 model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(2): 411-419. (in Chinese)

[3] 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)

[4] ANDERSEN H M L, KONGSTED A G, JAKOBSEN M J. Pig elimination behavior—a review[J]. Applied Animal Behaviour Science, 2020, 222:104888.

[5] HACKER R, OGILVIE J, MORRISON W, et al. Factors affecting excretory behavior of pigs[J]. Journal of Animal Science, 1994, 72(6):1455-1460.

[6] BERGSTEN C. Effects of conformation and management system on hoof and leg diseases and lameness in dairy cows[J]. The veterinary clinics of north america. Food Animal Practice, 2001, 17(1):1-23.

[7] PHILIPPE F X, CABARAUX J F, NICKS B. Ammonia emissions from pig houses: influencing factors and mitigation techniques[J]. Agriculture, Ecosystems & Environment, 2011, 141(3-4):245-260.

[8] REYNOLDS S J, DONHAM K J, WHITTEN P, et al. Longitudinal evaluation of dose-response relationships for environmental exposures and pulmonary function in swine production workers[J]. American Journal of Industrial Medicine, 1996, 29(1): 33-40.

[9] ZHU W, ZHONG F, LI X. Automated monitoring system of pig behavior based on RFID and ARM-LINUX[C]//Proceedings of the 2010 Third International Symposium on Intelligent Information Technology and Security Informatics, 2010:431-434.

[10] LARSEN M L, WANG M, WILLEMS S, et al. Automatic detection of locomotor play in young pigs: a proof of concept[J]. Biosystems Engineering, 2023, 229:154-166.

[11] XU J, ZHOU S, XU A, et al. Automatic scoring of postures in grouped pigs using depth image and CNN-SVM[J]. Computers and Electronics in Agriculture, 2022, 194:106746.

[12] 丁静,沈明霞,刘龙申,等. 基于机器视觉的断奶仔猪腹泻自动识别方法[J]. 南京农业大学学报,2020,43(5):969-978.

DING Jing, SHEN Mingxia, LIU Longshen, et al. Automatic recognition of diarrhea in weaned piglets based on machine vision[J]. Journal of Nanjing Agricultural University, 2020, 43(5):969-978. (in Chinese)

[13] 杨灵星,彭小雪,邹亮,等. 基于EfficientDet的生猪排泄行为识别研究[J]. 河北农业大学学报, 2023, 46(4):125-130.

YANG Lingxing, PENG Xiaoxue, ZOU Liang, et al. Research on pig excretion behavior recognition based on EfficientDet[J]. Journal of Hebei Agricultural University, 2023, 46(4):125-130. (in Chinese)

[14] 朱伟兴,浦雪峰,李新城,等. 基于行为监测的疑似病猪自动化识别系统[J]. 农业工程学报,2010,26(1):188-192.

ZHU Weixing, PU Xuefeng, LI Xincheng, et al. Automatic identification system of pigs with suspected case based on behavior monitoring[J]. Transactions of the CSAE, 2010, 26(1):188-192. (in Chinese)

[15] 吴世海,鲍义东,陈果,等. 基于机器视觉技术的猪行为活动无接触识别系统[J]. 计算机系统应用, 2020, 29(4): 113-117.

WU Shihai, BAO Yidong, CHEN Guo, et al. Contactless identification system for pig behavior based on machine vision[J]. Computer Systems and Applications, 2020, 29(4): 113-117. (in Chinese)

[16] 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.

[17] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018:3-19.

[18] SOHAN M, SAI RAM T, RAMI REDDY C V. A review on YOLOv8 and its advancements[C]//Proceedings of the International Conference on Data Intelligence and Cognitive Informatics, 2024:529-545.

[19] BEWLEY A, GE Z, OTT L, et al. Simple online and realtime tracking[C]//2016 IEEE International Conference on Image

Processing (ICIP), 2016;3464 – 3468.

[20] AHARON N, ORFAIG R, BOBROVSKY B Z. BoT – SORT: robust associations multi-pedestrian tracking[J]. arXiv Preprint, arXiv;2206.14651, 2022.

[21] CHEN L, JIN P, YANG J, et al. Robust Kalman filter-based dynamic state estimation of natural gas pipeline networks[J]. Mathematical Problems in Engineering, 2021, 2021(1): 5590572.

[22] YU E, LI Z, HAN S, et al. Relationtrack: relation-aware multiple object tracking with decoupled representation[J]. IEEE Transactions on Multimedia, 2022, 25:2686 – 2697.

[23] TAUD H, MAS J F. Multilayer perceptron (MLP)[M]//Geomatic Approaches for Modeling Land Change Scenarios, 2017: 451 – 455.

[24] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018;4510 – 4520.

[25] LILICRAP T P, SANTORO A, MARRIS L, et al. Backpropagation and the brain[J]. Nature Reviews Neuroscience, 2020, 21(6):335 – 346.

[26] SHANNON C E. A mathematical theory of communication[J]. The Bell System Technical Journal, 1948, 27(3):379 – 423.

[27] STEINLEY D J B J O M, PSYCHOLOGY S. K-means clustering: a half-century synthesis[J]. British Journal of Mathematical and Statistical Psychology, 2006, 59(1):1 – 34.

[28] YUE Y, FINLEY T, RADLINSKI F, et al. A support vector method for optimizing average precision[C]//Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 2007;271 – 278.

[29] LUITEN J, OSEP A, DENDORFER P, et al. Hota: a higher order metric for evaluating multi-object tracking [J]. International Journal of Computer Vision, 2021, 129:548 – 578.

[30] BERNARDIN K, STIEFELHAGEN R. Evaluating multiple object tracking performance: the clear mot metrics[J]. EURASIP Journal on Image and Video Processing, 2008, 2008:1 – 10.

[31] GUO Q, SUN Y, ORSINI C, et al. Enhanced camera-based individual pig detection and tracking for smart pig farms[J]. Computers and Electronics in Agriculture, 2023, 211:108009.

[32] SELVARAJU R R, DAS A, VEDANTAM R, et al. Grad – CAM: why did you say that? [J]. arXiv Preprint, arXiv;1611.07450, 2016.

[33] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016;770 – 778.

[34] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv Preprint, arXiv;1409.1556, 2014.

[35] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv Preprint, arXiv;2010.11929, 2020.

[36] YANG R, CHEN Z, XU H, et al. Recognizing the rooting action of prepartum sow in free-farrowing pen using computer vision [J]. Computers and Electronics in Agriculture, 2023, 213:108167.

[37] ZHAO Y, XIONG Y, WANG L, et al. Temporal action detection with structured segment networks[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017;2914 – 2923.

[38] CHEN Y J, TSAI Y S, LIN S, et al. Lactobacillus reuteri GKD2 improves blood glucose regulation in high-fat and high-sugar diet animal model[J]. Hans Journal of Biomedicine, 2024, 14:212.