

基于 SPR - Net 的哺乳母猪分娩前行为识别方法

汪 博^{1,2} 陈德恩³ 许俊华³ 洪林君^{1,2} 张素敏³ 尹 令^{1,3}

(1. 国家生猪种业工程技术研究中心, 广州 510642; 2. 华南农业大学动物科学学院, 广州 510642;
3. 华南农业大学数学与信息学院, 广州 510642)

摘要: 母猪分娩前行为精准识别对于实现分娩预测有着重要意义。提出了一种面向母猪姿态识别的 SPR - Net 网络,该网络使用 HRNet 作为骨干网络,在颈部融合特征金字塔网络与基于令牌的两阶段交叉注意力模块,同时在损失函数中引入关键点头损失作为归纳偏置,并通过非对称加权对分类头施加约束。在测试集上,该网络对母猪 4 种姿态分类 Top - 1 准确率达到 95.6%。网络完成性能验证后,对不同胎次母猪分娩前行为进行探究。研究发现,低胎次组和高胎次组母猪分娩前姿态转变频率呈现阶段性增长趋势;且低胎次组母猪在 Early 阶段、Late 阶段的姿态转变频率显著高于高胎次组母猪,而在 Mid 阶段两组的姿态转变频率差异不显著。此外,低胎次组母猪和高胎次组母猪在分娩前表现出不同的姿态转变频率变化范围,这些行为特征可作为一种临产特征,帮助养殖场及时采取相应措施。SPR - Net 网络在猪只分娩前行为识别中表现出优良性能,并通过视频分析揭示出不同胎次母猪的产前行为规律,为实现哺乳母猪分娩精准预测提供有力的技术支持和理论依据。

关键词: 母猪分娩; 姿态识别; 交叉注意力; 损失加权; HRNet

中图分类号: S24 文献标识码: A 文章编号: 1000-1298(2026)15-0056-08 OSID: 

Recognition of Pre-farrowing Behavior in Sows Based on SPR - Net

Wang Bo^{1,2} Chen Deen³ Xu Junhua³ Hong Linjun^{1,2} Zhang Sumin³ Yin Ling^{1,3}

(1. National Engineering Research Center for Breeding Swine Industry, Guangzhou 510642, China
2. College of Animal Science, South China Agricultural University, Guangzhou 510642, China
3. College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China)

Abstract: Accurate recognition of prepartum behaviors in sows is of great significance for farrowing prediction. An SPR - Net network was proposed for sow posture recognition. The network employed HRNet as the backbone and integrated a feature pyramid network with a token-based two-stage cross-attention module in the neck. Meanwhile, a keypoint head loss was introduced into the loss function as an inductive bias, and asymmetric weighting was applied to constrain the classification head. On the test set, SPR - Net achieved a Top - 1 accuracy of 95.6% for the classification of four sow postures. After performance validation, the network was further used to investigate prepartum behavioral patterns of sows with different parities. The results showed that the posture transition frequency of both low-parity and high-parity sows exhibited a staged increasing trend before farrowing. Moreover, the posture transition frequency of low-parity sows was significantly higher than that of high-parity sows during the early and late prepartum stages, whereas no significant difference was observed between the two groups during the middle stage. In addition, low-parity and high-parity sows showed different ranges of posture frequency variation before farrowing. These behavioral characteristics can serve as indicators of impending parturition, helping farms take timely management measures. Overall, SPR - Net demonstrated excellent performance in prepartum behavior recognition of sows, and video-based analysis revealed parity-related prepartum behavioral patterns. The research result can provide strong technical support and theoretical evidence for achieving accurate farrowing prediction in lactating sows.

Key words: sow farrowing; posture recognition; cross-attention; loss weighting; HRNet

收稿日期: 2026 - 03 - 16 修回日期: 2026 - 04 - 07

基金项目: 农业科技重大项目 (NK20221101)

作者简介: 汪博 (1998—), 男, 硕士生, 主要从事智慧养殖研究, E-mail: 18865212955@stu.scau.edu.cn

通信作者: 尹令 (1977—), 女, 教授, 博士, 主要从事三维点云重建和智能表型获取技术研究, E-mail: yin_ling@scau.edu.cn

0 引言

母猪分娩前常表现出不安、姿态频繁转变等现象^[1]。这些姿态变化不仅是分娩启动的行为标志^[2],也是评估母猪情绪和福利的重要指标^[3]。以站立、坐立、侧卧与俯卧等基础姿态为例,其时间占比与姿态转变频率能够反映母猪在产前的活动水平、健康状态,为分娩临近判断和福利状况提供可观测、量化的行为特征^[4]。

近年来,基于母猪产前行为特征构建分娩预测模型已成为相关研究的重要方向,并取得了一定成效,说明行为特征在分娩预警中具有较好的应用价值。已有研究表明,利用传感器或视频监控技术捕捉母猪分娩前的行为变化,并结合相应算法进行分析,可在分娩前数小时内实现有效分娩预警^[5]。但现有研究多在群体层面统一建模,易忽略不同胎次对母猪产前行为的影响,这可能导致模型将“胎次造成的行为差异”误判为“分娩临近信号”,如低胎次猪只可能因活动水平偏高更易被提前触发预警,增加误报情况;高胎次可能因活动水平偏低而出现漏检,从而影响特征解释性,并增加现场预警的误报与漏报风险。因此,在分娩预测相关研究中应考虑胎次对猪只产前行为特征的影响,明确不同胎次组在产前姿态特征上的差异规律。

猪只行为的表现是一个持久、连续、多变的过程。传统观测手段主要依赖人工观察与录像回放^[6],该方法虽操作简便、设备需求低,但存在主观性强、劳动强度大等弊端,难以满足规模化养殖场的需要。随着深度学习的飞速发展,非接触式的计算机视觉技术成为猪只行为监测的主流趋势^[7-8]。然而,分娩舍环境复杂(如围栏遮挡、光照不均),现有的目标检测算法多侧重于提取目标的全局外观特征^[9-11],难以捕捉精细化的骨骼关节点信息,且在复杂场景下易受环境干扰。

综上所述,本文提出一种母猪姿态识别网络 SPR - Net。该网络以 HRNet 作为骨干网络,在颈部融合特征金字塔网络并引入基于令牌的两阶段交叉注意力模块以增强关键特征交互,同时在损失函数中引入关键点定位作为归纳偏置,引导网络聚焦骨骼拓扑与有效姿态信息,从而实现对 4 类姿态的精准分类。在此基础上,本文以胎次为分组变量,将母猪划分为低胎次组与高胎次组,基于姿态识别结果对产前行为特征进行对比分析,进一步探索胎次差异对分娩预测特征的影响,以期为实行精准分娩预测提供理论依据。

1 数据采集与数据集构建

1.1 实验平台搭建与数据采集

本研究在广东某种猪试验场的分娩舍设计并搭建一套母猪分娩全过程行为监测系统。系统主要由海康威视摄像机、海康威视网络录像机、交换机、PC 组成。摄像机安装在距离地面 2.2 m 处,采用俯视的角度,便于拍摄完整的猪只姿态,视频帧率为 25 f/s,视频分辨率为 1 920 像素 × 1 080 像素。视频录制自待产母猪进入分娩栏开始,录制完成的视频数据可通过互联网进行远程上传。

2025 年 6—10 月,通过母猪分娩监测系统对 40 头待产母猪行为进行数据采集。实验对象为 1~6 胎的待产母猪。在预产期前 3 d,将待产母猪转移到分娩单元中,每头母猪安置在 1.8 m × 2.2 m 的分娩栏中。为保证实验的严谨性,本实验采集了两批不同日期的视频数据。第 1 批视频数据采集自 2025 年 6 月,采集 32 头母猪数据用作训练集和验证集。第 2 批视频数据采集自 2025 年 10 月,共 8 头母猪数据用作测试集,全部数据均为母猪分娩前 1 d 的完整视频段。本研究参考文献[12],将母猪的行为定义为 4 种基本姿态,如图 1 所示。分别为站立、坐立、侧卧、俯卧,具体定义见表 1。

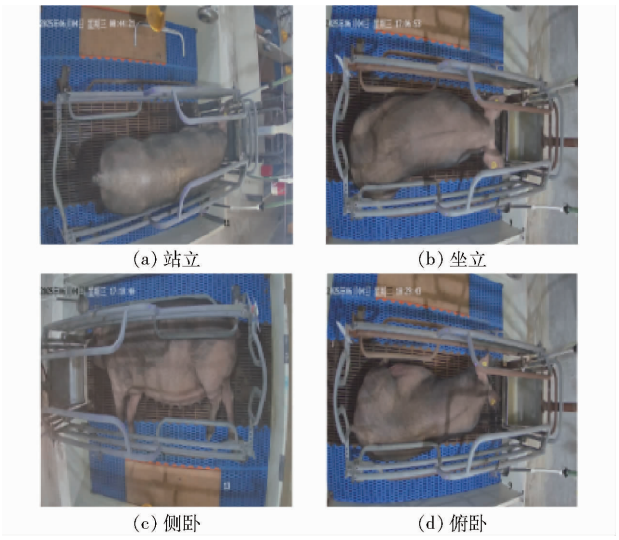


图 1 猪只姿态示意图

Fig. 1 Schematics of pig postures

表 1 猪只姿态定义

Tab. 1 Definitions of pig postures

姿态	姿态描述
站立	身体直立、腿伸直,仅蹄部接触地面
坐立	前腿伸直,尾部接触地面,上半身躯干直立
侧卧	身体侧卧在地面,四肢可见,乳房完全露出
俯卧	胸腹贴于地面,前后腿可折叠于身下,也可伸出,乳房侧露或完全不可见

1.2 数据预处理及数据集构建

本实验使用 FFmpeg 以每秒取 1 帧的方式进行抽帧处理,由于监控视频中存在大量重复画面,为此本研究使用 SSIM 算法对抽取的帧进行去重处理。具体而言,计算相邻帧之间的 SSIM 值,设定相似度阈值 0.9,剔除 SSIM 值高于阈值的高度相似帧。通过算法去重后,得到 6 月的视频共 23 040 帧,10 月的视频共 5 760 帧。

抽帧完成后,为提高数据集质量,进行严格的质量筛选,去除模糊、损坏的图像,并保持 4 种姿态的数量平衡。筛选完成后,6 月视频共计 12 555 帧,以 9:1 的比例划分为训练集、验证集;10 月份视频共计 2 001 帧,用作测试集。

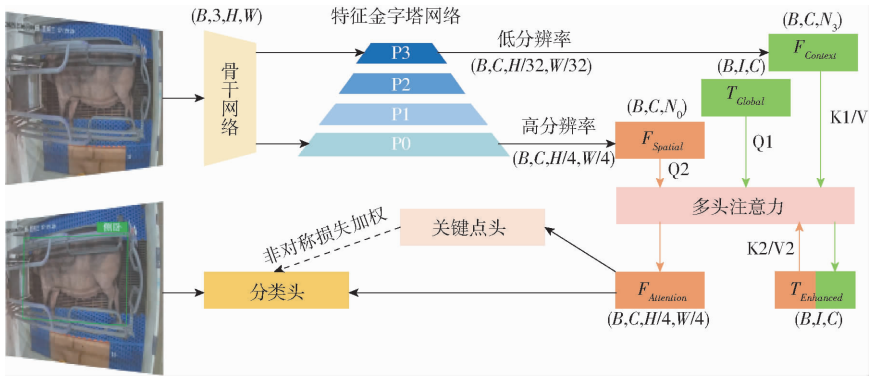


图 2 SPR-Net 网络结构图

Fig. 2 SPR-Net network architecture diagram

本文引入一种基于令牌的多头交叉注意力模块。该模块通过利用非对称注意力机制,将冗余的输入特征迭代提炼为紧凑的潜在语义瓶颈^[15]动态整合细节信息与全局空间信息。

最后,网络采用并行的双头部结构(即关键点头与分类头)进行联合预测,并提出非对称损失加权融合策略。在该策略下,辅助性的关键点头损失被用作主任务分类损失的结构化归纳偏置。该方法能有效正则化共享特征空间,强制网络学习同时具备几何感知和语义判别的特征表示,提升最终的分类精度和模型的泛化能力。

2.1 骨干网络

在监控视频中,猪只体表纹理单一且常与背景

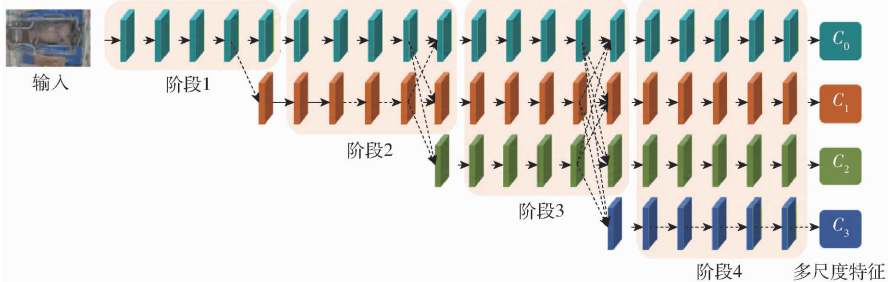


图 3 骨干网络 HRNet 结构图

Fig. 3 HRNet backbone network architecture diagram

本实验参考 AP-10K 数据集^[13],设置了 17 个身体部位的关键点,使用 Labelme 工具对数据集中图像可见关键点与姿态分类进行标注,并将其转换为符合 COCO 数据集格式的标注文件。

2 SPR-Net 哺乳母猪姿态识别模型

为结合关键点检测和目标检测的优势,提高真实养殖场场景下猪只姿态分类的鲁棒性,本文提出一种端到端的深度学习网络 SPR-Net,网络结构如图 2 所示。该网络采用 HRNet^[14]作为特征主干,以提取包含丰富细节和空间信息的多尺度特征表示。随后,特征金字塔网络通过重复的上采样和残差连接,实现多尺度特征的高效聚合与语义对齐。

同质化,传统的深度学习网络如 ResNet^[16]特征提取能力受到限制。这是因为 ResNet 采用的是级联下采样架构,会导致在逐层下采样过程中,空间信息的逐渐衰减,使得网络难以捕捉到猪只体表细节及肢体位置等关键特征。

HRNet 在网络全流程中并行维护多分辨率分支,并通过跨分辨率的特征融合。该设计使得即便在猪只体表纹理简单且与背景相似的情况下,HRNet 也能够提取到有效的空间和结构特征(网络结构见图 3)。

在本架构中,HRNet 的输出是一个多尺度的特征图元组 $\{C_0, C_1, C_2, C_3\}$, C_0, C_1, C_2, C_3 分别对应输入图像分辨率的 1/4、1/8、1/16、1/32 共 4 个层级。

2.2 特征金字塔网络

HRNet 的输出已具备多分辨率优势,但其各分支的通道维度仍存在异构性,且高分辨率的浅层特征(如 C_0)的语义层级较低。为弥合不同层级间的语义鸿沟并实现通道维度统一,本文引入特征金字塔网络用于多尺度特征的融合,具体见图 4。

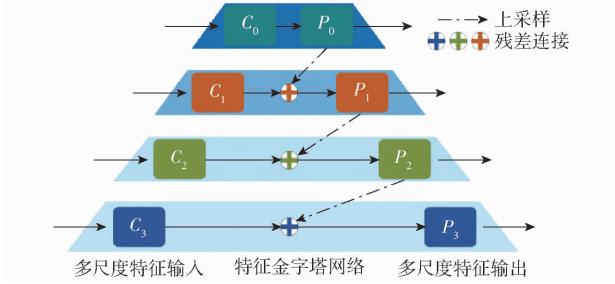


图 4 特征金字塔网络结构图

Fig. 4 Feature pyramid network architecture diagram

FPN 通过自顶向下的路径和横向连接进行特征融合。该过程将深层、强语义的特征(如 C_3)通过上采样,逐层注入到浅层、高分辨率的特征中。利用深层特征的语义信息,提升模型对猪只本体特征的捕捉能力,减少了背景噪声的干扰,进而提高在复杂养殖环境中的检测效果。

FPN 的输出是一个新的特征金字塔 $\{P_0, P_1, P_2, P_3\}$,该金字塔中的所有层级均具有统一的通道维度,且每一层都同时富含精确的空间细节和高级的全局语义。

2.3 基于令牌的多头交叉注意力机制

FPN 使用固定的上采样和横向连接进行静态特征融合,并未在关键点定位任务和分类任务之间建立显式的联系,权重也不随输入内容自适应改变。为实现两大任务的主动协同推理,本文设计了两阶段的基于令牌的多头交叉注意力模块(具体结构见图 2),该模块的优势是一个可学习的全局令牌,其功能是充当信息瓶颈和协同枢纽。该机制使得模型能够利用未被遮挡的猪只躯干信息,推断出被遮挡猪只肢体的潜在状态,提高模型在复杂场景下的性能。

本研究定义 $MHA(Q, K, V)$ 为多头交叉注意力函数, $LN(\cdot)$ 为层归一化操作。

首先,使用全局令牌 $T_{Global} \in \mathbf{R}^{B \times 1 \times C}$ 作为查询向量(Query),对 FPN 输出的最低分辨率语义特征图 $F_{Context} \in \mathbf{R}^{B \times C \times N_3}$ (在空间维度展平)作为键值(Key/Value)执行多头交叉注意力,并设置残差连接与层归一化。此步骤将 $F_{Context}$ 中零散的全局空间语义信息浓缩至 T_{Global} 中,形成一个整体姿态表征 $T_{Enhanced} \in \mathbf{R}^{B \times 1 \times C}$,具体公式为

$$T_{Enhanced} = LN(T_{Global} + \hat{T}) \quad (1)$$

其中 $\hat{T} = MHA(Q = T_{Global}, K = F_{Context}, V = F_{Context})$ (2)

随后,使用 FPN 输出的高分辨率空间特征图 $F_{Spatial} \in \mathbf{R}^{B \times C \times N_0}$ (在空间维度展平)作为查询向量(Query)对上述生成的 $T_{Enhanced}$ 令牌作为键值(Key/Value)执行第 2 次多头交叉注意力,并设置残差连接与层归一化。此步骤将 $T_{Enhanced}$ 携带的全局空间信息高效地融合到 $F_{Spatial}$ 的每一个像素上,有助于分类头解决因遮挡或模糊导致的局部歧义,最终输出具备判别力的特征图,并通过 Reshape 操作还原特征图 $F_{Attention} \in \mathbf{R}^{B \times C \times \frac{H}{4} \times \frac{W}{4}}$,具体公式为

$$F_{Attention} = LN(F_{Spatial} + \hat{F}) \quad (3)$$

其中 $\hat{F} = MHA(Q = F_{Spatial}, K = T_{Enhanced}, V = T_{Enhanced})$ (4)

2.4 多预测头与加权损失函数

经过 FPN 和多头注意力融合后,网络颈部输出的 $F_{Attention}$ 特征图同时输入至以下两个任务头部。

关键点头接收高分辨率的 $F_{Attention}$,并采用 SimCC^[17] 关键点解码器进行高精度的坐标回归,产生损失 L_{kpt} 。

分类头接收高分辨率的 $F_{Attention}$,经过全局平均池化操作,将空间特征 $\mathbf{R}^{C \times H \times W}$ 压缩为全局语义向量 $\mathbf{R}^{C \times 1 \times 1}$,随后送入全连接层以输出猪只 4 种姿态分类的概率分布,产生分类损失 L_{cls} 。

网络的优化目标定义为两项加权损失之和,即

$$L_{Total} = \lambda_{kpt} L_{kpt} + \lambda_{cls} L_{cls} \quad (5)$$

其中 λ_{kpt} 和 λ_{cls} 分别为关键点头损失权重和分类头损失权重。

对骨架拓扑结构的正确理解是精确分类姿态的前提,在保证分类损失 L_{cls} 为主要损失来源的前提下,适度提高关键点损失权重(本文使用 $\lambda_{kpt} = 2.0$, $\lambda_{cls} = 1.0$)。这种加权方式强化 L_{kpt} 对共享特征(Backbone 层和 Neck 层)的梯度贡献,使得 L_{kpt} 充当强大的结构化归纳偏置,迫使网络学习猪只的肢体关节的几何拓扑,而非仅依靠猪身的纹理与颜色,提高分类任务的鲁棒性与泛化性。

3 实验

3.1 实验环境

模型训练所使用的硬件配置为 CPU Intel(R) Core(TM) i9-14900K、GPU NVIDIA GeForce RTX 4080。所有训练均在 Ubuntu 22.04 系统下进行,并使用 CUDA 12.4、PyTorch 2.0.0、MMPose 1.3.2。

3.2 评估指标

本文遵循标准的 COCO 评估协议,使用基于 OKS 的平均精度均值(mAP)评估关键点检测性能,

并采用 Top - 1 准确率、宏平均精确率、精确率作为分类任务的评估指标。

本文使用 OKS 指标评价猪只关键点估计的准确性,公式为

$$OKS = \frac{\sum_{i=1}^N \exp\left(-\frac{d_i^2}{2s^2\kappa_i^2}\right)\delta(v_i > 0)}{\sum_{i=1}^N \delta(v_i > 0)} \quad (6)$$

式中 N ——关键点总数
 i ——关键点索引
 d_i ——预测关键点与真实值欧几里得距离
 s ——目标缩放因子,定义为检测边界框面积的平方根
 κ_i ——预定义的归一化常数
 v_i ——真实关键点的可见性标志

$\delta(\cdot)$ ——指示函数

若预测实例的 OKS 值超过预设阈值,则归类为真阳性(TP),否则归类为假阳性(FP)。该分类标准是计算平均精度均值(mAP)的基础。

3.3 实验超参数设定

本文对输入图像使用数据增强技术,包括随机水平翻转、随机半身裁剪、随机边界框变换以及仿射变换。所有模型均基于 MMPreTrain 公开预训练。

3.4 网络对比实验

本文基于目前主流的分类模型^[18-21]在测试集上进行系统性评估。主要评价指标选用 Top - 1 准确率、宏平均精确率、精确率为主要评价指标,并以关键点准确率(mAP)为辅助评价指标,详细结果如表 2 所示。

SPR - Net 在各项评价指标上均取得最佳性能。

表 2 网络对比实验结果

Tab.2 Comparison experiments results of networks

骨干网络	参数量	浮点数运 算量	关键点 mAP/%	分类		精确率/%			
				Top - 1 准确率/%	宏平均精确率/%	站立	坐立	侧卧	俯卧
MobileNetV3	$8.713\,0\times10^6$	$1.285\,0\times10^9$	—	83.90	84.74	90.90	71.23	93.59	83.23
ResNet	$3.197\,9\times10^7$	$8.362\,0\times10^9$	—	84.10	84.16	84.60	78.01	92.62	81.39
EfficientNetV2	$2.778\,2\times10^7$	$6.808\,0\times10^9$	—	88.85	89.34	85.76	84.04	96.10	91.46
CSPNeXt	$1.430\,4\times10^7$	$5.097\,0\times10^9$	—	88.00	88.97	94.57	78.67	88.93	93.69
SwinTransformer	$2.757\,9\times10^7$	$5.149\,0\times10^9$	—	80.85	81.29	91.81	66.77	90.94	75.64
SPR - Net(本文)	$3.691\,7\times10^7$	$1.177\,8\times10^{10}$	87.68	95.60	95.60	96.33	95.12	97.30	93.63

注:“—”表示对比实验的基线模型并未使用关键点,所以缺少关键点的 mAP 值。下同。

SPR - Net 的 Top - 1 准确率达到 95.60%,宏平均精确率达到 95.60%,与表现次优的 EfficientNetV2 相比,SPR - Net 有超过 6 个百分点的性能提升。

坐立姿态与俯卧姿态存有较高的类间相似性。而 SPR - Net 对坐立和俯卧姿态分别取得 95.12% 和 93.63% 的高精确率,且对各个姿态的预测效果

较为均衡,表明其对处理母猪复杂姿态识别时具有较好的性能表现。

3.5 消融实验

为验证该网络架构中各个组件的必要性,探究它们对最终性能的独立贡献与协同影响,本文设计了全面的消融实验,具体结果如表 3 所示。

表 3 消融实验结果

Tab.3 Results of ablation experiments

序号	骨干网络	FPN	Cross Attention	Keypoint Loss	关键点 mAP	分类		精确率			
						Top - 1 准确率	宏平均精确率	站立	坐立	侧卧	俯卧
1	HRNet	✓	✓	✓	87.68	95.60	95.60	96.33	95.12	97.30	93.63
2	HRNet	✓			—	86.60	87.66	90.23	75.66	94.01	90.75
3	HRNet		✓		—	86.30	87.65	90.20	74.13	96.40	89.87
4	HRNet			✓	87.56	91.40	91.84	83.89	94.61	95.14	93.70
5	HRNet	✓	✓		—	89.45	89.94	99.08	85.08	93.06	82.55
6	HRNet		✓	✓	86.27	90.65	91.27	87.67	84.76	99.15	93.50
7	HRNet	✓		✓	90.79	92.00	92.16	92.84	87.38	95.36	93.07
8	HRNet				—	88.10	88.69	90.44	84.70	86.13	93.47

注:✓表示运行此模块。

3.5.1 关键点损失作为归纳偏置的有效性
实验结果显示,关键点头作为一种结构化的归

纳偏置,对分类性能起到了较好的正则化与指导作用。

在相同的骨干与颈部网络的配置下,加入关键点损失后能明显提升测试集中分类的 Top - 1 准确率和宏平均精确率(如表 3 中序号 1、5,Top - 1 准确率从 89.45% 提升至 95.60%,宏平均精确率从 89.94% 提升至 95.60%),并明显改善若干难分类的姿态(例如坐立和俯卧)的识别能力。该提升在没有注意力模块的配置下仍可观察到,表明关键点监督本身已是一种强有力的正则化信号。

在本文所提出的两阶段注意力框架中,关键点损失不仅直接约束关键点的输出,而且通过注意力的路径反向传播,引导注意力给更有利于关键点定位与姿态判别的位置赋予更高权重,从全局语义到局部几何之间形成闭环协同。

3.5.2 FPN 与交叉注意力的协同有效性

实验结果表明,FPN 与交叉注意力模块的协同能带来可观的性能提升。

在骨干网络和关键点头的基础上,分别加入 FPN 和交叉注意力机制时,模型在测试集中分类的 Top - 1 准确率和宏平均精确率均有明显提升(如表 3 中序号 1、4,Top - 1 准确率从 91.40% 提升至 95.60%,宏平均精确率从 91.84% 提升至 95.60%)。

FPN 和注意力机制需搭配使用才能发挥最佳的效果,单独使用性能受限(如表 3 中序号 2、3、5)。这表明,本文设计的注意力机制并不适合直接应用于骨干网络输出的原始、未经融合的多尺度输出上。注意力机制在处理语义层级差距过大的特征时,可能引入了噪声或次优的先验,从而干扰了 $F_{Attention}$ 的表征质量。

FPN 的价值在于提供稳定的尺度对齐与语义统一,作为注意力机制的可靠输入;而注意力机制则负责在此基础上,实现动态路由与任务耦合,两者缺一不可。

3.6 模型可视化

3.6.1 混淆矩阵分析与挑战场景归因

为定量评估所提出的网络在猪只 4 种姿态上的实际性能,本文使用混淆矩阵进行详细分析,如图 5 所示。

从混淆矩阵的整体分布来看,模型展现出较好的分类精度,绝大多数样本都被正确分类,非对角线元素的数值较低,有力地证明了本文所提出的网络架构具有强大的特征提取和姿态辨识能力。

尽管网络在整体上表现优秀,但混淆矩阵揭示了对俯卧姿态的识别仍然存在的挑战。在测试集的 509 个俯卧样本中,共有 53 个样本被错误分类,特别是俯卧被误判为坐立(22 例)和站立(19 例)的现象尤为突出。

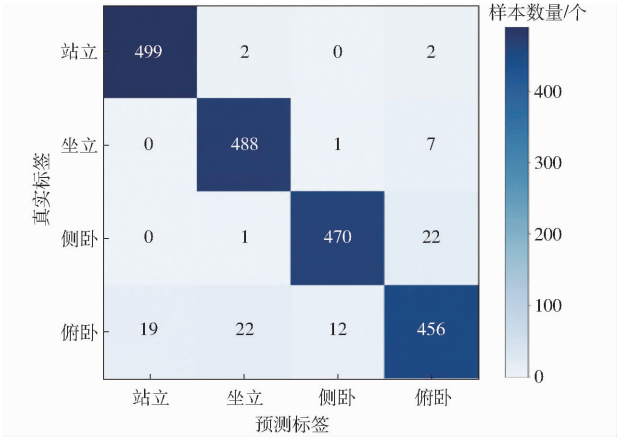


图 5 网络在测试集上的混淆矩阵
Fig. 5 Confusion matrix of network on testing set

从生物学和视觉角度来看,当猪只处于俯卧状态时,其身体轮廓与坐立姿态较为相似,如图 6a、6b 所示。另外,在特定的俯视视角下,猪只在侧卧姿态下背部完整展开,四肢因躯干遮挡而不可见,模型易将其背部特征误判为站立姿态下的特征,如图 6c、6d 所示。

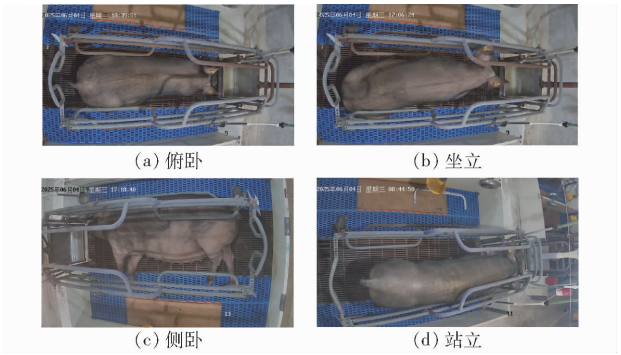


图 6 猪只的相似行为对比
Fig. 6 Comparison of similar behaviors in pigs

本文所提出的网络使用关键点头的损失对分类头进行约束,然而,俯卧姿态在各类姿态中,关键点遮挡情况相对更为突出。当关键点头难以提取有效的骨架拓扑信息时,反向传播至共享特征层的梯度信息将变得模糊甚至含有噪声。关键点几何先验的缺失使得分类头仅能依赖外观特征进行分类,从而导致对相似姿态产生误判。

俯卧类别的混淆并非源于网络的系统性失效,而是反映了该姿态固有的视觉复杂性,以及关键点辅助损失在严重遮挡场景下约束力减弱的局限性。

3.6.2 隐藏层特征演变可视化

为深入解析模型内部的特征处理逻辑,本文对网络的中间层特征图进行可视化分析。观察这 3 个阶段的热图演变,可以清晰地识别出特征响应呈现出“局部感知—上下文扩散—语义聚焦”的趋势。

图 7a 展示骨干网络输出的 C_0 特征图。在此阶段,网络的高响应区域主要集中在猪只的身体轮廓

及部分部位,但由于尚未进行充分的全局上下文融合,特征响应区域伴随着一定的噪声。例如,背景中的栏杆、地板纹理等非目标区域也表现出非零的激活值。这表明 C_0 特征主要编码了底层的视觉信息,缺乏对目标语义的过滤能力。

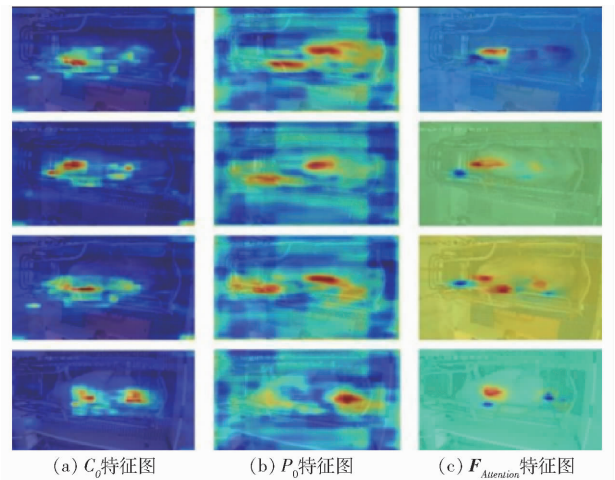


图 7 隐藏层特征可视化

Fig. 7 Visualization of hidden layer features

图 7b 展示经过 FPN 特征金字塔融合后的 P_0 特征图。与上一阶段相比,热图的激活区域发生了明显的扩散现象,覆盖范围明显扩大,甚至延伸至猪只身体周边的区域。这种扩散并非特征退化,而是感受野扩张的体现。FPN 通过扩大关注范围来捕获猪只整体姿态结构以及与周围环境的关系,将深层高级语义信息与浅层细节信息进行融合。这一过程实现从像素级感知向对象级理解的过渡,为后续精确聚焦提供了必要上下文信息。

图 7c 展示了经过交叉注意力机制优化后的 $F_{Attention}$ 特征图。在此阶段,特征分布经历从分散到集中,精准聚焦于猪只具有较强判别力的关键部位(如脊椎线、头部、四肢关节),且背景噪声(如栏杆干扰)被有效抑制。这一现象验证了交叉注意力机制的价值:特征筛选与校准。交叉注意力机制成功剔除冗余的背景信息,从 FPN 的宽泛语境中提取出对姿态分类和关键点定位的高层语义特征,这种聚焦机制保证分类头接收到的是信噪比较高的精炼特征表示。

3.7 分娩前 24 h 低胎次母猪与高胎次母猪行为差异探究

模型的精准性得到有效验证后,本研究将模型应用于视频推理,探究不同胎次母猪分娩前的行为差异。

本研究采用广义线性混合模型(GLMM)进行姿态变化趋势分析:以姿态变化频率为因变量,分娩前时间与胎次(低胎次,25 头;高胎次,15 头)及其交互作用为固定效应,母猪个体编号设为随机效应;模

型采用适用于计数数据的分布形式,存在过度离散时使用负二项分布,并基于模型边际预测绘制趋势曲线及 95% 置信区间(图 8)。为比较不同阶段的胎次差异,对每个阶段内采用非参数检验比较两胎次组差异,并对 3 次检验的 P 值进行校正,显著性水平设为 0.05(图 9),图中用 * 表示差异显著,ns 表示差异不显著。数据的预处理、分析、绘图均使用 R 4.5.1 完成。

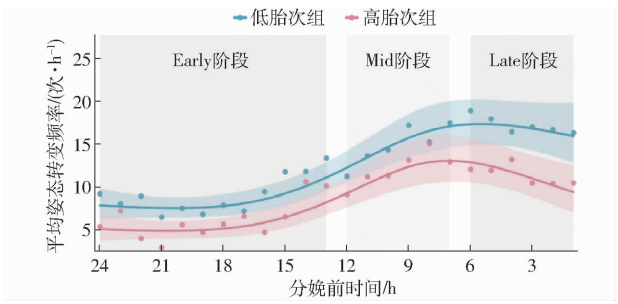


图 8 不同胎次分娩前 24 h 姿态变化趋势图

Fig. 8 Trend chart of postural changes in 24 h before delivery by parity

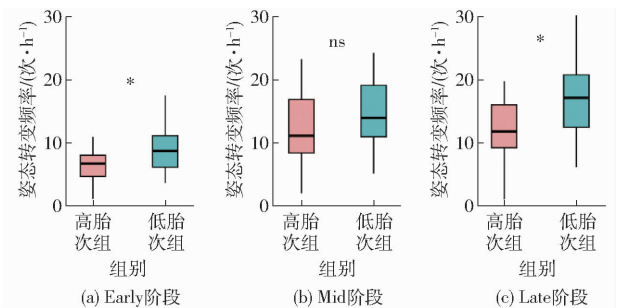


图 9 不同胎次母猪在分娩前 3 个阶段的姿态转变频率差异

Fig. 9 Differences in posture transition frequency of sows with different parities during three prepartum stages

如图 8 所示,母猪在分娩前 24 h 内的低胎次组和高胎次组姿态转变次数总体呈现“随分娩临近逐步升高”的趋势。两组母猪在 Early 阶段转变水平较低且平缓;进入 Mid 阶段后,低胎次组和高胎次组姿态转变次数开始明显上升,意味着此阶段可识别的产前变化已较为明显,可作为启动分娩预警的重要窗口;而在 Late 阶段,高胎次组母猪姿态转变频率呈现下降趋势,而低胎次组母猪则表现出较为稳定的高姿态转换频率水平。与此同时,胎次相关差异主要集中出现在 Early 与 Late 两个阶段,而在 Mid 阶段两组差异不明显,如图 9 所示。具体而言,在 Early 阶段,低胎次组母猪的姿态转变频率变化范围为 6.08 ~ 11.08 次/h,高胎次组母猪为 4.62 ~ 8 次/h;进入 Mid 阶段时,低胎次组母猪的姿态转变频率变化范围为 11 ~ 19.17 次/h,高胎次母猪组为 8.42 ~ 16.92 次/h;最后在 Late 阶段时,低胎次母猪

的姿态转变频率变化范围为 12.50 ~ 20.83 次/h,高胎次组母猪为 9.25 ~ 16.08 次/h。这些姿态转变频率范围可作为临产特征,当母猪的姿态转变频次在相应范围内时,表明待产母猪即将分娩,帮助养殖场及时采取相应措施。

综上所述,姿态转变频率可作为实现分娩预测的行为特征,但该特征在不同胎次之间有明显差异:低胎次母猪在产前整体更易出现较高的姿态转变频率,尤其在 Late 阶段差异更明显。这意味着在实际猪场预警系统中若采用“统一标准”,可能出现两类偏差:一方面低胎次因活动水平偏高更易被提前触发,增加无效警告与人力浪费;另一方面高胎次可能因活动水平偏低而出现漏检,从而错过关键的母猪护理时机。因此在后续的研究中,应将胎次因素纳入到预警建模中,要充分考虑胎次对猪只产前行为差异的影响,从而实现更稳定、适应性更强的分娩预警。

4 结论

(1)本文提出母猪姿态识别网络 SPR - Net:首

先,SPR - Net 利用 HRNet 的并行多尺度特征提取方法,能够精确提取猪只的空间特征并实现关键点定位,定位精度达到 87.68%。其次,网络引入 FPN 进行多层级特征聚合,能够有效整合来自不同尺度的特征,利用深层语义信息来抑制背景噪声,从而增强对猪只本体的检测能力。最后,SPR - Net 采用基于令牌的交叉注意力机制,通过全局语义信息的补充,进一步提升局部细节的表达。尤其是在复杂环境下,如遮挡或背景相似的情况,确保猪只检测的高精度和鲁棒性。并且模型在性能验证中具有良好表现,在测试集上 Top - 1 准确率、宏平均精确率均达到 95.60%,明显优于其他对比网络。此外,通过消融实验,验证了各个模块的有效性,尤其是 FPN、Cross - Attention 和 Keypoint Loss 的联合使用,相较于其他改进方法,展现了明显的优势。

(2)在完成模型性能验证后,对姿态进行量化分析发现:低胎次母猪产前姿态转变频率整体更高,且胎次差异主要出现在 Early 与 Late 阶段,Mid 阶段差异不明显。

参 考 文 献

[1] 洪佳奇,沈飞,邓颖,等. 不同圈栏条件对哺乳母猪行为表达的影响[J]. 云南农业大学学报(自然科学), 2020, 35(5):781 - 788.
Hong Jiaqi, Shen Fei, Deng Ying, et al. Effect of farrowing house on the behaviors of lactating sows[J]. Journal of Yunnan Agricultural University (Natural Science), 2020, 35(5):781 - 788. (in Chinese)

[2] 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.

[3] 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.

[4] Liu L, Zhou J, Zhang B, et al. Visual detection on posture transformation characteristics of sows in late gestation based on libra R - CNN[J]. Biosystems Engineering, 2022, 223:219 - 231.

[5] 刘龙申,柳苹,周杰,等. 母猪精准养殖智能感知技术研究进展[J]. 华南农业大学学报, 2024,45(5):635 - 648.
Liu Longshen, Liu Luo, Zhou Jie, et al. Research progress on precision breeding and intelligent sensing technologies for sows [J]. Journal of South China Agricultural University, 2024, 45(5):635 - 648. (in Chinese)

[6] Wurtz K, Camerlink I, D' Eath R B, et al. Recording behaviour of indoor-housed farm animals automatically using machine vision technology: a systematic review[J]. PLOS ONE, 2019, 14(12):e0226669.

[7] Yang Q, Xiao D. A review of video-based pig behavior recognition[J]. Applied Animal Behaviour Science, 2020, 233: 105146.

[8] Nasirahmadi A, Edwards S A, Sturm B. Implementation of machine vision for detecting behaviour of cattle and pigs[J]. Livestock Science, 2017, 202:25 - 38.

[9] 刘龙申,舒翠霓,李波,等. 基于 EfficientDet 的围产期母猪姿态识别[J]. 农业机械学报, 2022, 53(4):271 - 279.
Liu Longshen, Shu Cuini, Li Bo, et al. Recognition of perinatal sows' posture based on EfficientDet algorithm [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(4):271 - 279. (in Chinese)

[10] 查文文,潘伟豪,陈成鹏,等. 基于边缘计算与改进 YOLO v5 的群养生猪姿态识别及跟踪研究[J]. 东北农业大学学报, 2023, 54(3):83 - 96.
Zha Wenwen, Pan Weihao, Chen Chengpeng, et al. Study on pose recognition and tracking of herd-raised pigs based on edge computing and improved YOLO v5[J]. Journal of Northeast Agricultural University, 2023, 54(3):83 - 96. (in Chinese)

[11] 葛绍娟,冀横溢,詹宇,等. 改进 YOLO v5s 后的轻量化猪只姿态识别方法[J]. 中国农业大学学报, 2025, 30(5): 179 - 189.
Ge Shaojuan, Ji Hengyi, Zhan Yu, et al. Lightweight pig posture recognition method after improving YOLO v5s[J]. Journal of China Agricultural University, 2025, 30(5):179 - 189. (in Chinese)

Chinese)

[17] Ma W, Wang X, Yang S X, et al. Autonomous inspection robot for dead laying hens in caged layer house[J]. Computers and Electronics in Agriculture, 2024, 227:109595.

[18] Yaseen M. What is YOLO v9; an in-depth exploration of the internal features of the next-generation object detector[PP/OL]. V1. arXiv(2024 - 09 - 12) [2026 - 06 - 24]. <https://arXiv.org/abs/2409.07813>.

[19] Wang A, Chen H, Liu L, et al. YOLO v10: real-time end-to-end object detection[J]. Advances in Neural Information Processing Systems, 2025, 37:107984 - 108011.

[20] Khanam R, Hussain M. YOLO v11: an overview of the key architectural enhancements[PP/OL]. V2. arXiv(2024 - 10 - 23) [2026 - 06 - 24]. <https://arXiv.org/abs/2410.17725>.

[21] Tian Y, Ye Q, Doermann D S. YOLO v12: attention-centric real-time object detectors[J]. Advances in Neural Information Processing Systems, 2026, 38: 78433 - 78457.

[22] Sapkota R, Cheppally R H, Sharda A, et al. YOLO26: key architectural enhancements and performance benchmarking for real-time object detection[PP/OL]. V5. arXiv(2026 - 03 - 16) [2026 - 06 - 24]. <https://arXiv.org/abs/2509.25164>.

[23] 杨断利, 齐俊林, 陈辉, 等. 基于改进 YOLO v8n 模型的散养蛋鸡个体行为识别方法与差异分析[J]. 农业机械学报, 2024, 55(11):112 - 123.

Yang Duanli, Qi Junlin, Chen Hui, et al. Individual behavior recognition method and difference analysis of free-range laying hens based on the improved YOLO v8n model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11):112 - 123. (in Chinese)

[24] Song D, Chen Z, Peng Y, et al. JNL - SlowFast: an end-to-end approach for dead rabbits classification using short cage inspection videos in commercial farms[J]. Computers and Electronics in Agriculture, 2025, 240:111190.

[25] 杨断利, 王永胜, 陈辉, 等. 基于 SEEC - YOLO v5s 的散养蛋鸡日常行为识别与统计系统[J]. 农业机械学报, 2023, 54(9):316 - 328.

Yang Duanli, Wang Yongsheng, Chen Hui, et al. Daily behavior recognition and statistics system for free-range laying hens based on SEEC - YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(9):316 - 328. (in Chinese)

[26] Brown J, Nguyen A, Raj N. Effect of camera choice on image-classification inference[J]. Applied Sciences, 2024, 15(1): 246.

(上接第 63 页)

[12] Zheng C, Zhu X, Yang X, et al. Automatic recognition of lactating sow postures from depth images by deep learning detector [J]. Computers and Electronics in Agriculture, 2018, 147:51 - 63.

[13] Yu H, Xu Y, Zhang J, et al. AP - 10K: a benchmark for animal pose estimation in the wild[PP/OL]. V2. arXiv(2021 - 11 - 01) [2026 - 04 - 07]. <https://doi.org/10.48550/arXiv.2108.12617>.

[14] Wang J, Sun K, Cheng T, et al. Deep high-resolution representation learning for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(10):3349 - 3364.

[15] Jaegle A, Gimeno F, Brock A, et al. Perceiver: general perception with iterative attention[C] // Proceedings of the 38th International Conference on Machine Learning(PMLR), 2021:4651 - 4664.

[16] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C] //2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016:770 - 778.

[17] Li Y, Yang S, Liu P, et al. SimCC: a simple coordinate classification perspective for human pose estimation[PP/OL]. V3. arXiv(2022 - 07 - 05) [2026 - 04 - 07]. <https://doi.org/10.48550/arXiv.2107.03332>.

[18] Tan M, Le Q V. EfficientNetV2: smaller models and faster training[C] // International Conference on Machine Learning (ICML), 2021:10086 - 10096.

[19] Wang C Y, Mark Liao H Y, Wu Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2020:1571 - 1580.

[20] Howard A, Sandler M, Chen B, et al. Searching for MobileNetV3 [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019:1314 - 1324.

[21] Liu Z, Lin Y, Cao Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C] // 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021:9992 - 10002.