

基于改进 YOLO 11n – Pose 和 Jetson Orin NX 的多目标小鼠行为分析方法

梁秀英^{1,2} 刘 岩¹ 何 磊³ 赵文瑞¹ 祝梓涵⁴ 张恩帅¹ 杨万能^{2,4}

(1. 华中农业大学工学院, 武汉 430070; 2. 华中农业大学作物遗传改良国家重点实验室, 武汉 430070;
3. 华中农业大学信息学院, 武汉 430070; 4. 华中农业大学植物科学技术学院, 武汉 430070)

摘要: 针对人工观察小鼠行为存在费时费力、主观性强的问题,提出了一种适用于 Jetson Orin NX 平台的多目标小鼠行为分析方法。经比较,YOLO 11n – Pose 在多个姿态关键点模型中效果最优,精确率为 94.1%、召回率为 94.5%、平均精度均值为 96.4%,单幅图像推理用时 36.04 ms。提出 WIoU 损失函数改进,精确率提升 2.1 个百分点,平均精度均值提升 0.7 个百分点。基于结构简化和量化模型并部署到 Jetson Orin NX,压缩后模型的精确率、召回率和平均精度均值分别为 93.93%、94.18% 和 95.88%,单幅图像推理用时 11.01 ms,缩短 69.46%。对比 4 个多目标跟踪模型,结果表明 OC Sort 的效果优于其他模型,单幅图像推理用时 2.18 ms。基于 OC Sort 多目标跟踪模型进行改进,改进后的 IDF1、MOTA 和 MOTP 分别为 83.5%、92.0% 和 27.6%,ID 跳变比原模型减少 41.18%,单幅图像推理用时 2.32 ms,结果表明改进 OC Sort 在不影响实时性的同时对于准确性有明显提升;采用基于阈值法的小鼠行为预测,将人工值和预测值对比,结果表明蜷缩、弯曲和伸长行为判断准确率分别达到 95.12%、81.71% 和 80.49%;采用该方法对比青年鼠与老年鼠 24 h 的行为活动,结果表明青年鼠的活跃度明显高于老年鼠,青年鼠更多时候处于休息状态,而老年鼠则更多处于睡觉状态。研究表明,改进 YOLO 11n – Pose 和 OC Sort 相结合的多目标小鼠行为分析方法,在提升跟踪与关键点准确性的同时,经量化部署后可实现便携、实时的多目标小鼠行为分析。

关键词: 小鼠;改进 YOLO 11n – Pose;多目标跟踪算法;Jetson Orin NX;模型轻量化;行为分析

中图分类号: TP391.4; S24 **文献标识码:** A **文章编号:** 1000-1298(2026)14-0091-11 **OSID:** 

Multi-object Mice Behavior Analysis Method Based on Improved YOLO 11n – Pose and Jetson Orin NX

Liang Xiuying^{1,2} Liu Yan¹ He Lei³ Zhao Wenrui¹ Zhu Zihan⁴ Zhang Enshuai¹ Yang Wanneng^{2,4}

(1. College of Engineering, Huazhong Agricultural University, Wuhan 430070, China
2. National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China
3. College of Informatics, Huazhong Agricultural University, Wuhan 430070, China
4. College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China)

Abstract: Aiming to address the drawbacks of manual observation of mouse behavior, which is time-consuming, labor-intensive, and highly subjective, a multi-object mice behavior analysis method adapted for the Jetson Orin NX embedded platform was proposed. Through comparative experiments among multiple pose keypoint estimation models, YOLO 11n – Pose achieved the optimal comprehensive performance, with a precision of 94.1%, recall of 94.5%, the mean average precision was 96.4%, and inference latency of 36.04 ms per single image. To further optimize the model, an improvement scheme with WIoU loss function was proposed, which increased the precision by 2.1 percentage points and the mAP by 0.7 percentage points. Subsequently, model structure simplification and quantization were performed for edge deployment on the Jetson Orin NX platform. The compressed model achieved a precision of 93.93%, recall of 94.18%, and mAP of 95.88%, with the inference latency per single image reduced to 11.01 ms, a 69.46% reduction compared with that of the original model. Furthermore,

收稿日期: 2025 – 05 – 04 修回日期: 2025 – 06 – 22
基金项目: 国家自然科学基金项目 (U21A20205)、湖北洪山实验室重大项目 (2022hszd024) 和华中农业大学自主科技创新基金项目
作者简介: 梁秀英 (1976—), 女, 副教授, 博士, 主要从事动物表型、智慧农业技术与装备研究, E-mail: nancy@mail.hzau.edu.cn
通信作者: 杨万能 (1984—), 男, 教授, 博士, 主要从事作物表型组研究, E-mail: ywn@mail.hzau.edu.cn

comparative experiments were conducted on four mainstream multi-object tracking models, and the results showed that OC Sort outperformed other models with an inference latency of 2.18 ms per single image. On this basis, an improved OC Sort model was proposed. The optimized model achieved an IDF1 score of 83.5%, MOTA of 92.0%, and MOTP of 27.6%. The number of ID switches was reduced by 41.18% compared with that of the original OC Sort, while the inference latency was only 2.32 ms per single image. These results indicated that the improved OC Sort model significantly enhanced tracking accuracy without compromising real-time performance. For mouse behavior recognition, a threshold-based method was adopted. By comparing manual annotation results with model recognition outputs, the recognition accuracy of curving, shrinking and elongation behaviors reached 95.12%, 81.71% and 80.49%, respectively. The proposed method was applied to compare the 24 h behavioral activities of young mice and aged mice. The results showed that the activity level of young mice was significantly higher than that of aged mice; young mice stayed in a resting state for more time, while aged mice were mostly in a sleeping state. It demonstrated that the proposed multi-object mice behavior analysis method, which combined the improved YOLO 11n - Pose and optimized OC Sort models, can not only improve the accuracy of keypoint estimation and multi-object tracking, but also can realize portable and real-time multi-object mice behavior analysis after quantization and edge deployment.

Key words: mouse; improved YOLO 11n - Pose; multi-object tracking algorithm; Jetson Orin NX; model lightweight; behavioral analysis

0 引言

小鼠是医学试验最常见的动物,小鼠行为是大脑底层神经计算最重要的外在体现形式,反映了动物情绪、学习记忆、运动协调和其他高级中枢功能^[1]。小鼠行为中蕴含着诸多健康信息,直接或间接反映其生理、病理以及营养等内在状态^[2-3]。研究小鼠行为对精神类疾病的治疗和药物开发有重要作用^[4]。观察和有效识别小鼠行为具有重要意义。传统的小鼠行为观察主要依赖于人工观察,但由于工作效率低、人员培训成本高逐渐被取代。动物行为分析方法的发展经历了人工观察、传感检测技术、视频跟踪技术等阶段^[5-10]。

近年来,随着计算机视觉技术和深度学习的发展,基于视觉特征实现动物行为识别作为一种非接触式、自动化且经济高效的方法成为当前的主流方法^[11]。卷积神经网络能够实现图像的特征提取、目标检测、姿态估计及目标跟踪。Sheppard 等^[12]利用姿态估计中的 HRNET 提取单只小鼠的姿态关键点,能够从上而下的角度测量步态和全身姿态参数。Nath 等^[13]开发了一种动物行为分析工具 DeepLabCut,基于 ResNet 能够对小样本数据集进行精准的学习和分类,利用 Open Pose 的 PAF (Part affinity fields)^[14] 装配方法,从图像中预测关键点。设计多任务卷积神经网络 DLCRNET,得到得分概率图、位置偏移、部分亲和力和动物身份^[15]。Lauer 等^[16]对 DeepLabCut 的主干网络 ResNet 进行调整,引入 DLCRNet_ms5 提取特征,多个并行反卷积层预测关键点。Liu 等^[17]在 DeepLabCut 基础上加入 1

个 CNN 层,以帧差图像为输入以减少颜色变化的影响,克服低视频质量和重遮挡问题。Graving 等^[18]提出 DeepPoseKit 动物行为分析工具,限制在单目标姿态估计,采用自顶而下,使用一种高效多尺度深度学习模型 StackedDenseNet,以及一种基于图形处理器 (Graphics processing unit, GPU) 的快速峰值检测算法。Pereira 等^[19-20]提出 Social LEAP 同时实现了自顶向下和自底向上的多目标姿态估计和跟踪,解决了 LEAP^[21] 在复杂环境下造成的算法限制。SLEAP 利用动物形态和皮毛纹理作为区分特征,解决了多目标跟踪过程中由于多动物遮挡造成的身份跟踪失败的问题。基于 U-net^[22] 作为网络框架,支持 30 多种先进的神经网络主干。随着技术的不断创新,Schweihoff 等^[23]提出了 DeepLabStream 可以用于实时的姿态估计和跟踪,基于 TensorFlow^[24] 框架,将卷积神经网络 ResNet^[25] 作为基础结构框架。但 DeepLabStream 需要进行大量的数据计算和处理,需要一定的计算机性能支持。

基于以上传统和现有方法的问题,本文设计一种便携式多目标小鼠行为实时分析装置,研究基于改进 YOLO 11n - Pose^[26] 和改进 OC Sort 的多目标小鼠行为分析方法,经 Jetson Orin NX 边缘部署^[27],对多目标小鼠行为进行实时分析跟踪,以期小鼠行为学分析提供一种便捷装置。

1 材料与方法

1.1 试验材料与数据获取

试验采用自主研发小鼠饲养笼,试验对象为 3 笼 C57BL/6 雌鼠,每笼内饲养 3 只。从 2021 年 9

月起,饲养在湖北省武汉市华联科生物技术有限公司。笼体顶部安装工业级 RGB 摄像头,摄像机采集视频帧率设置为 20 f/s,分辨率为 2 560 像素 × 1 440 像素,采集原始图像如图 1a 所示,试验提取 6 月龄和 19 月龄时期的视频数据。由于摄像机镜头结构问题,采集到的视频存在桶形畸变,本文采用张氏标定法^[28]获取相机内外参,并基于光学畸变模型对像素坐标进行几何变换校正,恢复为无畸变图像。通过边缘裁剪去除冗余区域,将有效区域调整为 1 440 像素 × 1 440 像素,并采用双线性插值法将图像统一缩放至 720 像素 × 720 像素以提升处理效率。最终预处理效果如图 1b 所示,图像畸变得得到有效矫正。

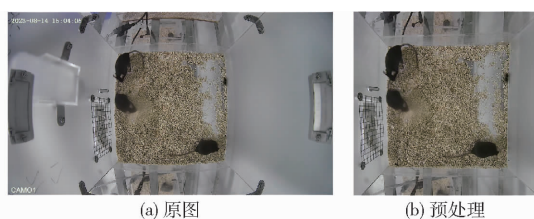


图 1 图像预处理
Fig. 1 Image preprocessing

1.2 数据集构建

基于饲养笼内小鼠环境变化,构建具有时空多样性的多目标检测与姿态估计数据集。采集 3 ~ 19 月龄小鼠视频序列,覆盖饲养笼内不同空间位置及多种社交交互状态,包含小鼠分开、靠近、遮挡场景,制作目标检测、姿态估计数据集,共筛选 1 135 幅有效图像样本。按照比例 8:2 划分训练集和测试集。为了改善模型的泛化能力和抗干扰能力,减少过拟合,并提高模型的性能和鲁棒性,采用添加噪音、翻转、降低光照等图像增强方法将训练集图像扩增到 3 960 幅。使用 Labelme 图像标注工具进行双盲标注,构建的标注体系包含:基于最小外接矩形的目标检测标注(不标注尾巴);6 个关键点标注(鼻尖、左耳、右耳、颈部、身体中心及尾巴根部)。最后将标注得到的 Json 文件转为 txt 格式,以用于 YOLO 11-Pose 的模型训练。

1.3 基于 YOLO 11n-Pose 的 IoU 损失函数改进

YOLO 11^[29] 核心结构包括 3 部分:骨干网络、颈部网络以及检测头。在 YOLO v8 的基础上,提出了 C3K2 模块替换 C2f 模块,C3K2 通过引入多种尺寸的卷积核并融入通道分割策略,显著增强了特征提取的效率与能力。具体而言,该模块将输入特征一分为二,一部分直接通过标准卷积层传递,另一部分经由多个 C3K 或 Bottleneck 结构进行深入的特征提取。随后,这 2 部分特征被拼接起来,并通过 1 个

1 × 1 卷积层进行融合,以生成更具表现力的特征。C3K 作为一种多尺度卷积核,其尺寸 K 可灵活调整,包括 3 × 3、5 × 5 等,有效扩展了模型的感受野,提升了模型在复杂场景下的特征捕捉能力。当超参数 C3K 被设置为 True 时,原本的 Bottleneck 模块会被 C3K 所替代,否则使用 C2f 模块。提出 C2PSA 机制,是 C2 内部嵌入 1 个多头注意力机制,并在原本的分类检测头中增加 2 个 DW-Conv,可以大幅减少参数数量和浮点运算量^[30]。骨干部分用来提取图像特征的网络。颈部网络实现来自不同网络层级特征图融合,增强特征表示能力。检测头部负责执行最终的目标检测、分类、关键点检测任务,如图 2 所示。

对 YOLO 11n-Pose 模型进行改进,以适应本文应用场景。采用 WIoU 替换原 IoU,提高模型对遮挡目标的关注度,提高模型精度。传统 IoU 损失函数只考虑了预测框与真实框的重叠部分,而未考虑两者之间的区域,导致在评估结果时可能存在偏差。基于这一问题,提出一种新的损失函数,名为 Wise-IoU (WIoU)^[31],它结合 IoU 的思想,并引入动态非单调的特性。WIoU (v1) 通过距离注意力机制实现了 2 层注意力机制,以增强对不同区域的关注。为了有效降低简单样本对损失的贡献,借鉴 Focal Loss 的理念,设计了针对交叉熵的单调聚焦机制的 WIoU (v2)。进一步地,WIoU (v3) 构建了非单调聚焦系数,动态调整损失函数。凭借其梯度增益分配策略,WIoU 显著提升了模型性能。具体公式为

$$L_{WIoUv3} = rL_{WIoUv1}$$

$$\text{其中 } L_{WIoUv1} = R_{WIoU} L_{IoU} \quad r = \frac{\beta}{\delta \alpha^{\beta-\delta}}$$

$$R_{WIoU} = \exp \left(\frac{\rho^2(c, c^{gt})}{W_g^2 + H_g^2} \right)$$

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

式中 $\rho^2(c, c^{gt})$ ——真实框和预测框中心点距离
 W_g, H_g ——最小外接矩形框宽、高
 $\bar{L}_{IoU}$ ——滑动平均值
 δ, α ——超参数,将 W_g, H_g 从计算图中分离

1.4 基于改进 OC Sort 的多目标跟踪

多目标跟踪算法 Sort 采用 TBD 范式,核心为卡尔曼滤波和匈牙利算法。卡尔曼滤波用于根据前一帧的状态预测当前帧目标的位置,而匈牙利算法则负责将预测信息与检测结果进行匹配。本文采用 OC Sort 多目标跟踪算法,并结合 YOLO 11n-Pose

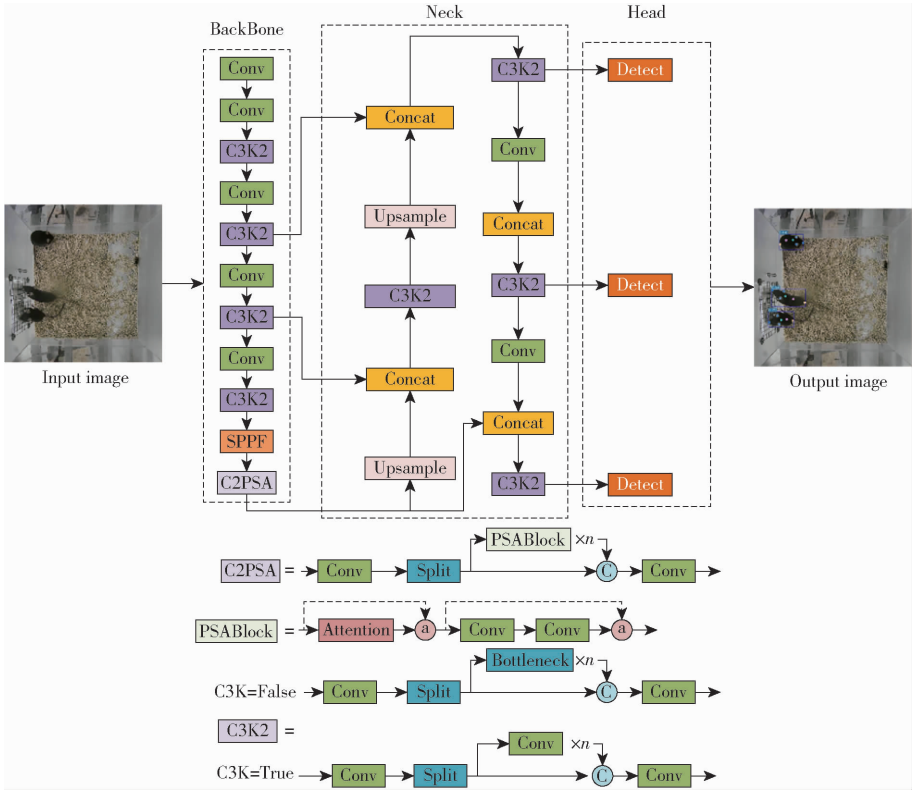


图 2 YOLO 11n 整体结构

Fig. 2 Overall structure of YOLO 11n

目标检测网络作为检测器。首先,通过 YOLO 11n-Pose 获取每只小鼠检测框,并将这些框输入到 OC Sort 算法中。随后,自适应卡尔曼滤波预测小鼠的运动信息,并结合运动特征进行目标匹配,从而实现有效的多目标跟踪。

为提升 OC Sort 算法的效果,特别针对小鼠聚集现象进行了改进。由于检测到的小鼠坐标框数量可能不等于 3,将上一帧的 3 只小鼠与当前帧的检测框进行匈牙利匹配,依据目标框之间的欧氏距离进行判断。当检测结果少于 3 只时,将上一帧未匹配的小鼠框添加到当前帧;当检测结果多于 3 只时,则去除当前帧未匹配的坐标框。最后,将预处理后的检测框输入网络进行跟踪。改进 OC Sort 工作流程如图 3 所示。

1.5 模型训练、部署

1.5.1 环境配置

模型训练的软件环境配置为 Windows 10、Python 3.8、Cuda 11.2、ultralytics 8.3.79 和 Pytorch 1.12;硬件环境为 CPU Intel Xeon w3-2423 处理器, GPU 采用 NVIDIA RTX A4000 显卡,显存为 16 GB。批次为 16 倍,工作进程为 4 倍,迭代次数为 100。

1.5.2 模型优化

模型量化是常见的深度学习模型压缩技术。进行模型量化之前,采用基于常量叠加以及计算图优化的方式,进行模型简化。基于常量叠加是将计算

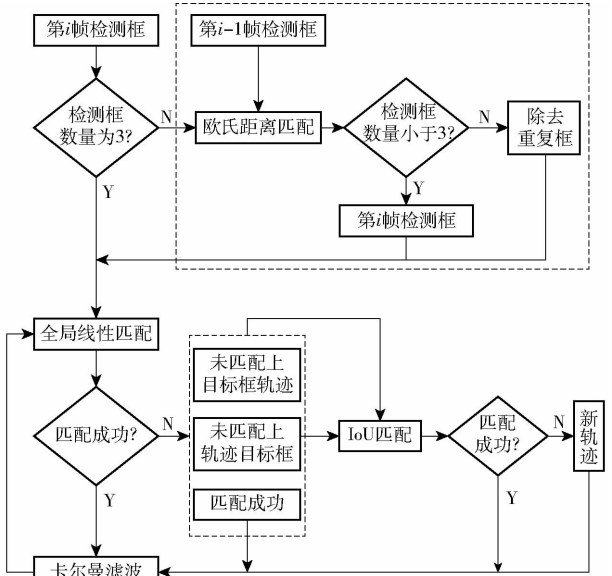


图 3 基于改进 OC Sort 工作流程图

Fig. 3 Diagram based on improved OC Sort workflow

图中静态节点,通过预计算其输出结果并替换为常量值,从而移除冗余计算步骤;计算图优化是将连续的同类型操作合并为单层,减少计算复杂度。模型量化是通过将 32 位浮点运算数转换成 8 位整数降低计算内存,提高模型的推理速度。本文采用 simplify 方法对模型进行结构简化,随后采用量化进行模型压缩。

1.5.3 模型部署

采用 Jetson Orin NX 进行边缘部署,该平台根据

内存大小有 8 GB 和 16 GB 版本,采用 16 GB 版本的 Jetson Orin NX。平台主要硬件环境为 6 核 ARM Cortex - A78 的 CPU 和搭载 32 个 Tensor Core 的 1024 核 NVIDIA Ampere 架构的 GPU。软件环境主要由 Python 3.8、Jetpack 5.1.1、CUDA 11.4、Cudnn 8.6、Tensor Rt 8.5 搭建,操作系统为 Ubuntu 20.04。相较于 Jetson Orin Nano 提供的 40 TOPS 浮点运算性能,Jetson Orin NX(16 GB)在处理整数运算时能够提供高达 100 TOPS 的性能,约相当于前者的 2.5 倍。这种显著的性能提升使得 Jetson Orin NX 能执行复杂的计算任务,能更高效地满足用户需求。Jetson Orin NX 边缘部署系统如图 4 所示。

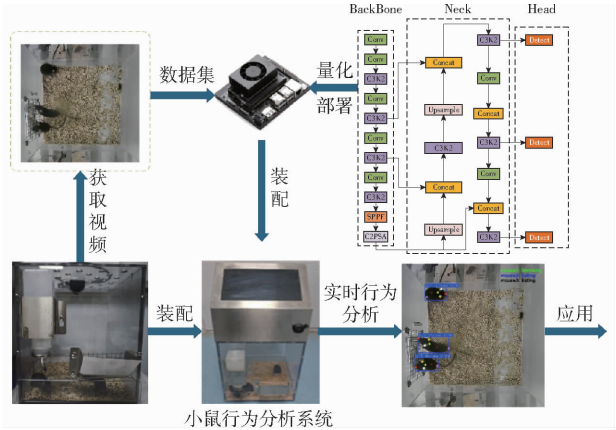


图 4 小鼠行为分析系统
Fig. 4 Mouse behavior analysis system

为进一步加快推理速度,采用 Tensor Rt 进行推理加速。Tensor Rt 对网络结构进行整合,在运行时对网络结构进行优化,大大减少浮点运算量,从而实现模型推理加速:①进行垂直整合。在深度学习网络中 Conv - Bn - ReLU 结构十分常见,Tensor Rt 将网络结构中的 Conv 层、BN 层和 ReLU 层进行合并,形成 1 个 CBR 层,将模型推理到该结构时的 3 次运算减少到 1 次。②再水平整合,将相同张量和相同操作的部分进行融合,减少 CUDA 核心资源的消耗,从而提高显存的利用率。③融合 Concat 层,将 Contact 层的输出直接输入后续的操作中,不用单独进行 Concat 计算,减少了传输次数。网络结构优化前后示意图如图 5 所示。

1.6 基于关键点的小鼠行为识别方法

1.6.1 小鼠运动特征和姿态特征

小鼠行为由运动特征和姿态特征组成,运动特征能够反映小鼠行为在时间序列上的连续程度,而姿态特征可以反映小鼠在这一帧的身体状态。

(1)运动特征。如图 6 所示,通过小鼠关键点检测模型及多目标跟踪模型,可以得到每只小鼠的运动信息,及前后帧小鼠身体中心的像素变化量,像

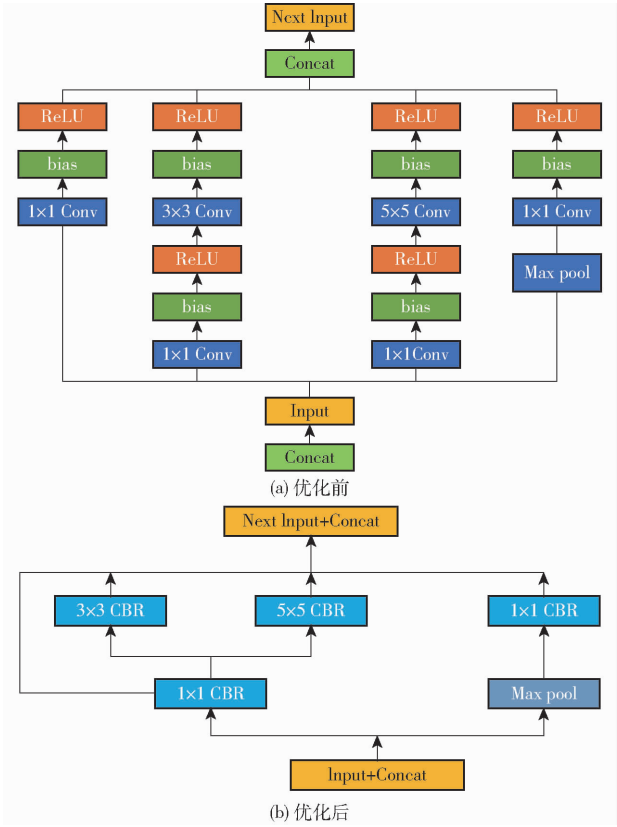


图 5 Tensor Rt 网络优化结构
Fig. 5 Tensor Rt network optimization structure

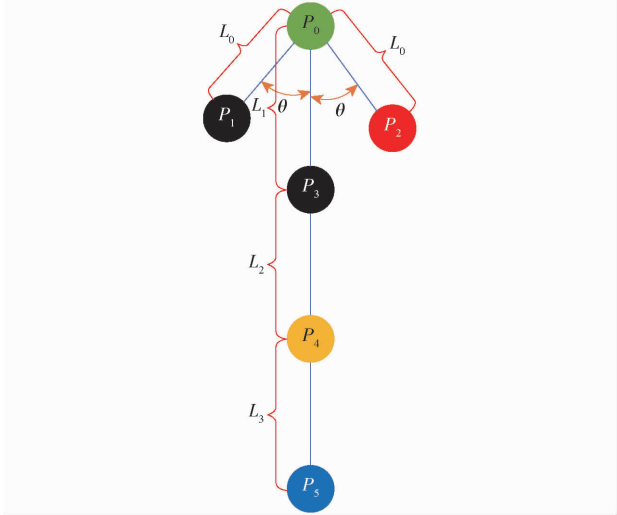


图 6 小鼠姿态骨架
Fig. 6 Mouse posture skeleton

素变化量除以帧率得到小鼠瞬时速度 v 。

(2)姿态特征。基于小鼠的姿态关键点,构建小鼠骨骼模型,获取精准姿态特征,分析行为。如图 6 所示,具体关键点包括鼻尖 P_0 、左耳 P_1 、右耳 P_2 、脖子 P_3 、身体中心 P_4 和尾巴 P_5 ,小鼠姿态骨架参数如表 1 所示。

1.6.2 小鼠行为识别

对饲养笼中小鼠进行行为分析有助于深入了解每只小鼠的生活习惯和个性特征,从而获取更

表 1 小鼠姿态骨架参数

参数	定义
L_0	鼻尖与左、右耳距离
L_1	鼻尖与脖子距离
L_2	脖子与身体中心距离
L_3	身体中心与尾巴距离
θ	头部朝向和身体方向的夹角

准确的行为数据。为了确保数据的普遍性,对小鼠的体态特征进行归一化处理。基于小鼠运动速度将行为分为动态行为和静态行为,通过分析小鼠的姿态特征,静态行为可进一步细分为进食、饮水、爬网、伸长、弯曲、蜷缩、站立、静止、休息和睡觉,动态行为为行走,如表 2 所示。小鼠行为如图 7 所示。

表 2 小鼠行为特征阈值

行为	体态阈值	速度阈值/ (像素·s ⁻¹)	类别
进食	$P_0 R$	$v < 200$	瞬时行为
饮水	$P_0 Y$	$v < 200$	瞬时行为
爬网	$P_4 \in R$	$v < 200$	瞬时行为
站立	$P_4 \in R_1$	$v < 200$	瞬时行为
伸长	$L/L_{\max} > 0.95, \theta < 30^\circ$	$v < 200$	瞬时行为
弯曲	$30^\circ < \theta < 80^\circ$	$v < 200$	瞬时行为
蜷缩	$L/L_{\max} < 0.6$	$v < 200$	瞬时行为
静止	$L/L_{\max} < 0.95, \theta < 30^\circ$	$v < 200$	瞬时行为
行走		$v > 200$	瞬时行为
睡觉	$L_c < 10$ 像素, $t > 40$ s	$v < 200$	持续行为
休息	$L_c < 10$ 像素, 20 s $< t < 40$ s	$v < 200$	持续行为

注: L 表示体长, Y 表示饮水区域, R 表示食物区域, v 表示速度, R_1 表示站立区域, $L_{\max}$ 表示体长最大值, L_c 表示身体中心前后帧像素变化, t 表示持续时间。

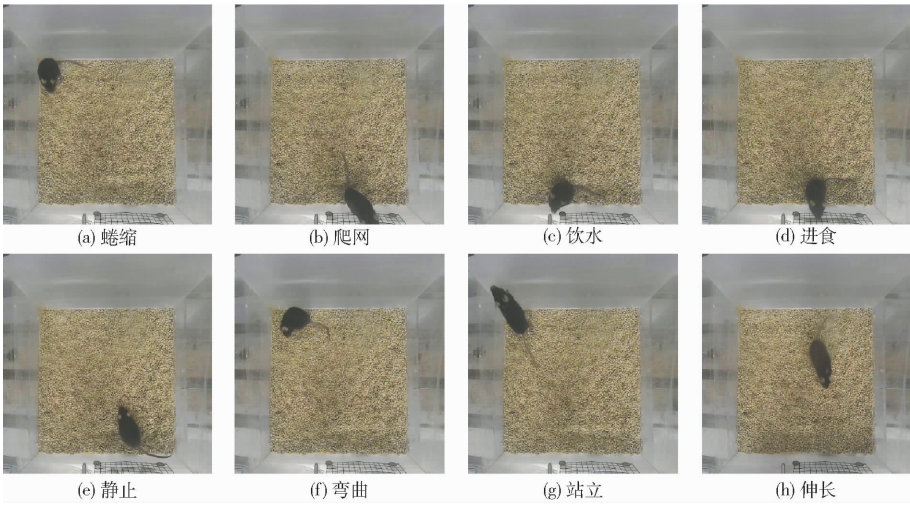


图 7 小鼠个体行为
Fig. 7 Individual behavior of mouse

1.7 评价指标

为评估目标检测模型的精度,采用精确率 (Precision)、召回率 (Recall) 和平均精度均值 (mean Average Precision, mAP) 作为性能评价指标。采用单幅图像推理用时衡量模型的推理速度,作为判断实时性的评估指标。为评估模型在 Jetson Orin NX 平台上的适合性,选用模型参数量和运行时模型内存占用量作为模型尺寸评估指标。为评估多目标跟踪模型的效果,基于多目标跟踪性能指标 MOTA (Multiple object tracking accuracy)、MOTP (Multiple object tracking precision)、IDF1 (Identification F1 score) 和 IDSW (Identity switch) 进行评估。此外,通过对姿态关键点进行行为预测,将人工标注的真实值与预测结果进行对比,以此作为行为预测准确性的指标。

2 结果与分析

2.1 不同目标检测模型性能测试

针对小鼠多目标检测与姿态估计任务,通过测试当前主流姿态估计模型的综合性能,验证 YOLO 11n - Pose 算法适用性。将 YOLO v8s - Pose、YOLO v8n - Pose、YOLO v9 - Pose、YOLO v10n - Pose 和 YOLO 11n - Pose 5 类模型的综合性能进行测试,结果如表 3 所示。由表 3 可知,YOLO 11n - Pose、YOLO v8s - Pose、YOLO v8n - Pose 和 YOLO v9 - Pose 4 个模型精确率、召回率、mAP 均为 98% 以上,说明 4 个模型在小鼠检测效果均满足需求。在关键点检测上,YOLO v8s - Pose 精确率、召回率、mAP 均高于其他模型,说明关键点的预测效果更好。但考虑到边缘部署,推理用时、模型参数量、内存占用量是更重要的指标,经 Jetson Orin NX 测试推理用时、模型参数量和内存占用

量, YOLO v9 - Pose、YOLO v8s - Pose 推理用时明显小于 YOLO v8n - Pose 和 YOLO 11n - Pose, 且内存占用量和模型参数量也大于 YOLO v8n - Pose 和 YOLO 11n - Pose。关键点检测 3 个模型各指标均高于 94%, 因此不选用 YOLO v8s - Pose、YOLO v9 - Pose。YOLO 11n - Pose 相比 YOLO v8n - Pose

推理用时分别减少 3.5%, 模型参数量减少 13.63%, 内存占用量提高 17.51%, Jetson Orin NX 内存为 16 GB 能够满足内存占用量的增长, 模型参数量和推理用时的减少能更好满足实时性的需求, 因此选用 YOLO 11n - Pose 为基础模型, 并对其进行轻量化改进。

表 3 不同目标检测模型性能测试结果

Tab.3 Performance test results of different object detection models

检测模型	目标检测指标			关键点指标			模型推理指标		
	精确率/	召回率/	mAP/%	精确率/	召回率/	mAP/%	推理用时/	参数量	内存占用量/MB
	%	%		%	%		ms		
YOLO v8s - Pose	99.2	99.4	99.4	96.2	95.8	97.3	63.90	1.142×10^7	516.73
YOLO v8n - Pose	98.9	99.3	99.5	95.2	94.4	96.5	37.40	3.080×10^6	259.68
YOLO 11n - Pose	98.5	99.3	99.5	94.1	94.5	96.4	36.04	2.660×10^5	305.17
YOLO v9 - Pose	98.7	98.3	99.5	95.4	94.6	96.9	169.85	3.234×10^7	1699.84
YOLO v10n - Pose	94.0	94.8	97.7	90.5	90.8	94.0	41.06	2.780×10^6	370.34

2.2 不同 IoU 损失函数模型性能测试

为检验 WIoU 损失函数对 YOLO 11n - Pose 性能的提升, 选用 507 幅图像作为测试集, 对不同 IoU 损失函数的 YOLO 11n - Pose 模型进行性能测试。由表 4 可知, 基于 WIoU 损失函数的 YOLO 11n - Pose 在精确率、召回率、mAP 上表现均最优。相比原模型精确率提升 2.1 个百分点, mAP 提升 0.7 个百分点, 召回率无明显提升。由于改用 WIoU 损失函数对 YOLO 11n - Pose 在精度上的不足有补充, 选用 WIoU 替换 YOLO 11n - Pose 的 IoU。

表 4 不同损失函数的 YOLO 11n - Pose 模型性能结果

Tab.4 Performance results of YOLO 11n - Pose models with different loss functions

模型	精确率	召回率	mAP
YOLO 11n	94.1	94.5	96.4
YOLO 11n - WIoU	96.0	94.8	97.1
YOLO 11n - EIoU	94.8	94.9	96.8
YOLO 11n - GIoU	95.0	94.7	96.6
YOLO 11n - Shape - IoU	95.4	95.5	97.0
YOLO 11n - SIoU	94.8	94.9	96.8

表 5 YOLO 11n - Pose 不同量化精度测试结果

Tab.5 Test results of YOLO 11n - Pose with different quantization precisions

量化处理	PC 端				Jetson Orin NX 平台					
	精确率/%	召回率/%	mAP/%	推理用时/ms	精确率/%	召回率/%	mAP/%	推理用时/ms	参数量	内存占用量/MB
量化前	94.51	94.73	96.37	4.12	94.60	94.72	96.36	36.04	2.66×10^6	305.17
Fp32	94.75	94.12	96.26	2.85	94.75	94.01	96.25	25.60	2.71×10^6	270.64
Fp16	94.76	94.12	96.26	2.13	94.75	94.06	96.25	14.81	2.71×10^6	270.64
Int8	94.12	94.12	95.84	2.07	93.93	94.18	95.88	11.29	2.65×10^6	231.17
Int8 - simplify	94.37	94.25	95.94	1.91	93.93	94.18	95.88	11.01	2.65×10^6	231.17

2.4 不同多目标跟踪算法性能测试

为验证改进 OC Sort 跟踪模型的性能优势,本试验选取连续 601 幅图像作为测试集。为充分体现跟踪模型的效果,图像包含多只小鼠遮挡、重叠、模糊的图像,且存在人工判断每只小鼠较为困难的图像。选用现阶段主流的多目标跟踪模型 OC Sort、Byte Track、Deep Sort、Strong Sort 和 C-BIoU Track 作为对比分析。由表 6 可知,OC Sort、Byte Track 和 C-BIoU Track 在推理用时和 IDSW 指标表现更为优秀,表明其在实时性和 ID 稳定性优于另 2 种模型。主要原因可能是 Deep Sort 和 Strong Sort 都使用了外观特征,对黑鼠而言,外观十分相似,外观特征反而影响 ID 识别的准确性,且增加跟踪推理时间。进一步对比 OC Sort、Byte Track 和 C-BIoU Track,OC Sort 的 IDF1 和 MOTA 较 C-BIoU Track 高 8.27% 和 74.61%,比 Byte Track 高 2.80% 和 11.04%,综合考虑模型的准确性和实时性,OC Sort 更合适。为进一步提升跟踪的准确性,对 OC Sort 进行改进。改进的 OC Sort 相比原模型,ID 切换减少 41.18%,表明在跟踪过程中目标对象的 ID 标识更加稳定,有效降低身份混淆的概率;IDF1 和

MOTA 分别提高 51.82% 和 1.5 个百分点,说明跟踪的准确性得到显著提升。由图 8 知,OC Sort 在检测错误多出 1 个目标时,在后面的帧赋予 1 个新的目标框。但改进的 OC Sort 会根据上一帧信息将检测错误的目标删除,并将上一帧的信息重新匹配赋值到当前帧,保证了跟踪的准确性,并且防止 ID 跳变。由于改进后结构进行优化,推理用时(2.32 ms)虽较 OC Sort(2.18 ms)增加 6.4%,但并未对实时性造成影响,说明改进的 OC Sort 模型可很好地适用于本文的多目标小鼠跟踪。改进 OC Sort 的 IDF1、MOTA、MOTP 分别为 83.5%、92%、27.6%。

表 6 不同多目标跟踪模型性能测试结果
Tab.6 Performance test results of different multi-object tracking models

跟踪模型	IDF1/ %	MOTA/ %	MOTP/ %	IDSW/ 次	推理 用时/ms
改进 OC Sort	83.5	92.0	27.6	10	2.32
Deep Sort	38.6	87.6	27.3	42	18.50
OC Sort	55.0	90.5	27.5	17	2.18
C-BIoU Track	50.8	51.3	28.3	16	0.62
Strong Sort	34.9	82.7	27.2	49	18.02
Byte Track	53.5	81.5	27.1	17	1.96

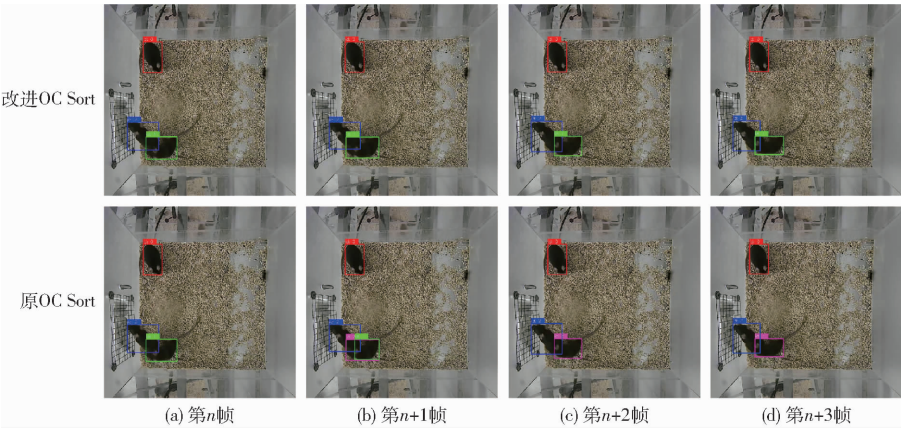


图 8 改进 OC Sort 与原 OC Sort 跟踪效果
Fig.8 Improved tracking effect of OC Sort and original OC Sort

2.5 小鼠行为预测试验

采用基于阈值法的图像分析对 246 幅图像进行行为学分类验证,结果如表 7 所示。基于阈值法行为预测对于小鼠饮水、进食、爬网以及站立均是根据区域范围划分,行为的判断依赖于关键点识别的准确度。对于区域划分没有问题时,能保证判断不会出现错误。小鼠休息、睡觉是持续行为,并不依赖于单幅图像行为判定。因此小鼠主要判定的行为:弯曲、伸直、蜷缩。

由表 7 可知,基于角度、体长进行判断的小鼠行为准确率均高于 80%,其中蜷缩、弯曲和伸长准确率分别为 95.12%、81.71% 和 80.49%。弯曲、伸长

表 7 不同行为预测试验结果
Tab.7 Test results of prediction for different behaviors

行为	人工结果	预测结果	准确率/%
蜷缩	82	78	95.12
弯曲	82	67	81.71
伸长	82	66	80.49

的预测由于姿态关键点的位置预测存在误差,因此基于阈值法对于这 2 种行为的判断准确率存在一定误差。但总体行为判断准确率符合试验要求。

2.6 青年鼠与老年鼠行为测试

选用 6 月龄青年鼠与 19 月龄老年鼠的 24 h 行为学视频分析。由图 9 可知,小鼠 1 h 的运动轨迹

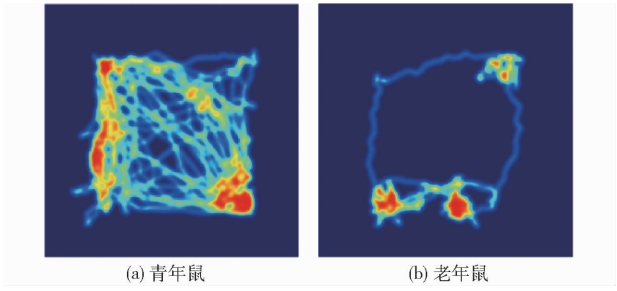


图 9 运动轨迹热图

Fig. 9 Motion trajectory heatmaps

定量分析进一步证实,青年鼠在 1 h 内的运动较老年鼠更高,且青年鼠会在饲养笼内到处走动,而老年鼠更倾向于待在 4 个角落,表明青年鼠对周围环境更加好奇,倾向于探索周围环境。这些行为差异表明青年鼠比老年鼠运动能力更强,对周围环境更加

好奇,更喜欢探索环境。行为分析结果如图 10 所示,其中使用 t 检验进行差异显著性分析,当 $P < 0.001$ 时,为高度显著(标记为 ***);当 $0.001 \leq P < 0.01$ 时,为极显著(标记为 **);当 $0.01 \leq P < 0.05$ 时,为显著(标记为 *)。由图 10a 可知,老年鼠比青年鼠的爬网次数更多,但无显著差异;由图 10e 可知,青年鼠比老年鼠在行为上更喜欢伸长,存在高度显著差异;由图 10f 可知,青年鼠一天的行走路程与老年鼠存在高度显著差异,青年鼠的行走路程远远多于老年鼠;由图 10g 可知,青年鼠比老年鼠更喜欢站立,存在极显著差异;由图 10d、10h 可知,睡眠模式分析表明,老年鼠倾向于进入睡眠状态,而青年鼠倾向休息状态,休息状态存在极显著差异。

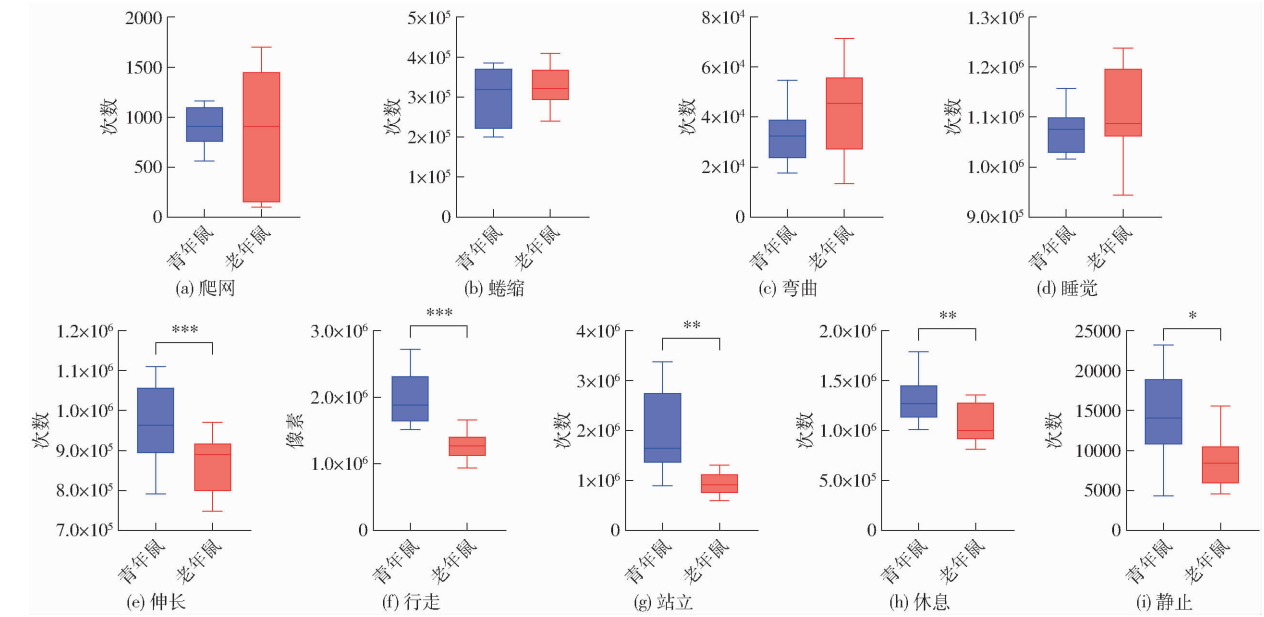


图 10 青、老年鼠行为显著性分析

Fig. 10 Behavioral significance analysis of young and elderly mice

3 结论

(1) 构建了多目标小鼠关键点数据集,选用 YOLO 11n-Pose 作为关键点检测模型,其量化前的推理用时为 36.04 ms。针对原模型精度偏低的问题,将损失函数改为 WIoU,增强对遮挡环境的抗干扰能力,精确率提升 2.1 个百分点。在 Jetson Orin NX 平台部署并模型量化后,模型精确率为 93.93%,召回率为 94.18%,mAP 为 95.88%,推理用时减少至 11.01 ms,缩短 69.46%,运行时内存占用量降低 24.25%,兼具良好准确性与实时性。

(2) 对比 OC Sort、Deep Sort、Strong Sort、Byte Track 和 C-BIoU Track,结果表明 OC Sort 兼顾实时性与准确性,单幅图像推理用时为 2.18 ms。针对 OC Sort 存在的 ID 切换严重的问题,改进匹配机制

并固定跟踪数量,使 ID 切换减少 41.18%,IDF1 和 MOTA 分别提高 51.82% 和 1.5 个百分点。优化后推理用时为 2.32 ms,较 OC Sort 增加 6.4%,最终改进 OC Sort 的 IDF1、IDSW、MOTA 和 MOTP 分别为 83.5%、10 次、92.0% 和 27.6%。

(3) 提出基于阈值法的多目标小鼠个体行为分析方法,该方法可测量小鼠个体行为。结果表明,青年鼠与老年鼠站立存在极显著差异;行走路程存在高度显著差异,青年鼠一天行走路程远远多于老年鼠。青年鼠的行动更活跃,而老年鼠身体退化,四肢不能很好地适应站立及长时间行走。试验结果说明青年鼠比老年鼠的行动更加活跃。

(4) 在 Jetson Orin NX 平台设计了一种多目标小鼠行为实时分析装置,集成改进后 YOLO 11n-Pose 与 OC Sort 算法,为无监督小鼠行为分析及智

能自动化进程提供了一定支撑与推动。摄像头采样帧率为 20 f/s,要求每帧处理时间小于 50 ms,压缩后模型推理用时为 11.01 ms,完全满足实时处理需求。在一些不足处,严重遮挡下关键点识别不准确,ID 跳变,且最多支持 3 只小鼠同时分析。后续可通过个体标记与多视角拍摄进一步改善。

参 考 文 献

[1] 王子元,方璇,陈春花,等. 基于深度学习的动物行为学分析[J]. 解剖学杂志, 2024, 47(6):528 – 532.
Wang Ziyuan, Fang Xuan, Chen Chunhua, et al. Animal behavior analysis based on deep learning[J]. Chinese Journal of Anatomy, 2024, 47(6):528 – 532. (in Chinese)

[2] 赵凯旋. 基于机器视觉的奶牛个体信息感知及行为分析[D]. 杨凌:西北农林科技大学, 2017.
Zhao Kaixuan. Individual information perception and behavior analysis of dairy cows based on machine vision[D]. Yangling: Northwest A&F University, 2017. (in Chinese)

[3] Stricklin W R, Kautz S C C. The role of behavior in cattle production: a review of research[J]. Applied Animal Ethology, 1984, 11(4): 359 – 390.

[4] 金玫蕾. 动物行为学在现代生命科学研究中的应用[J]. 实验动物与比较医学, 2008(1):1 – 3.
Jin Meilei. Application of animal behavior in modern life science research[J]. Laboratory Animal and Comparative Medicine, 2008(1):1 – 3. (in Chinese)

[5] Wei K, Kording K P. Behavioral tracking gets real[J]. Nature Neuroscience, 2018, 21(9): 1146 – 1147.

[6] Xia F, Kheirbek M A. Circuit-based biomarkers for mood and anxiety disorders[J]. Trends in Neurosciences, 2020, 43(11): 902 – 915.

[7] Weber R Z, Mulders G, Kaiser J, et al. Deep learning-based behavioral profiling of rodent stroke recovery[J]. BMC Biology, 2022, 20(1): 232.

[8] Anderson D J, Perona P. Toward a science of computational ethology[J]. Neuron, 2014, 84(1): 18 – 31.

[9] 孙秀萍,王琼,石哲,等. 动物行为实验方法学研究的回顾与展望[J]. 中国比较医学杂志, 2018, 28(3):1 – 7.
Sun Xiuping, Wang Qiong, Shi Zhe, et al. Review and prospect of methodological research on animal behavior experiments[J]. Chinese Journal of Comparative Medicine, 2018, 28(3):1 – 7. (in Chinese)

[10] Tinbergen N. On aims and methods of ethology[J]. Zeitschrift f r Tierpsychologie, 1963, 20(4): 410 – 433.

[11] 吴赛赛,吴建寨,程国栋,等. 基于姿态估计的动物行为识别研究进展[J]. 中国农业大学学报, 2023, 28(6):22 – 35.
Wu Saisai, Wu Jianzhai, Cheng Guodong, et al. Research progress on animal behavior recognition based on pose estimation [J]. Journal of China Agricultural University, 2023, 28(6):22 – 35. (in Chinese)

[12] Sheppard K, Gardin J, Sabnis G S, et al. Gait-level analysis of mouse open field behavior using deep learning-based pose estimation[PP/OL]. BioRxiv (2020 – 12 – 29) [2026 – 04 – 10]https://doi.org/10.1101/2020.12.29.424780.

[13] Nath T, Mathis A, Chen A C, et al. Using DeepLabCut for 3D markerless pose estimation across species and behaviors[J]. Nature Protocols, 2019, 14(7): 2152 – 2176.

[14] Cao Z, Simon T, Wei S E, et al. Realtime multi-person 2D pose estimation using part affinity fields[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7291 – 7299.

[15] 申通,王硕,李孟,等. 深度学习在动物行为分析中的应用研究进展[J]. 计算机科学与探索, 2024, 18(3):612 – 626.
Shen Tong, Wang Shuo, Li Meng, et al. Research progress on the application of deep learning in animal behavior analysis [J]. Journal of Frontiers of Computer Science and Technology, 2024, 18(3):612 – 626. (in Chinese)

[16] Lauer J, Zhou M, Ye S, et al. Multi-animal pose estimation and tracking with DeepLabCut[PP/OL]. BioRxiv (2021 – 04 – 30) [2026 – 04 – 10]. https://doi.org/10.1101/2021.04.30.442096.

[17] Liu H, Reibman A R, Boerman J P. Video analytic system for detecting cow structure[J]. Computers and Electronics in Agriculture, 2020, 178: 105761.

[18] Graving J M, Chae D, Naikh, et al. DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning[J]. Elife, 2019, 8: 47994.

[19] Pereira T D, Tabris N, Matsliah A, et al. SLEAP: a deep learning system for multi-animal pose tracking[J]. Nature Methods, 2022, 19(4): 486 – 495.

[20] Pereira T D, Aldarrondo D E, Willmore L, et al. Fast animal pose estimation using deep neural networks[J]. Nature Methods, 2019, 16(1): 117 – 125.

[21] Mihajlovic M, Zhang Y, Black M J, et al. LEAP: learning articulated occupancy of people[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 10461 – 10471.

[22] Cao H, Wang Y, Chen J, et al. Swin-unet: unet-like pure transformer for medical image segmentation[C]//European Conference on Computer Vision, 2022: 205 – 218.

[23] Schweihoff J F, Loshakov M, Pavlova I, et al. DeepLabStream enables closed-loop behavioral experiments using deep learning-based markerless, real-time posture detection[J]. Communications Biology, 2021, 4(1): 130.

[24] Pang B, Nijkamp E, Wu Y N. Deep learning with tensorflow: a review[J]. Journal of Educational and Behavioral Statistics, 2020, 45(2): 227 – 248.

[25] Targ S, Almeida D, Lyman K. Resnet in resnet: generalizing residual architectures[PP/OL]. V1. arXiv (2016-03-25) [2026-04-10]. <https://doi.org/10.48550/arXiv.1603.08029>.

[26] 黄小平,侯现坤,郭阳阳,等. 基于改进 YOLO v7-Pose 的牛脸关键点检测与姿态识别方法[J]. 农业机械学报, 2024, 55(11):84-92, 102.
Huang Xiaoping, Hou Xiankun, Guo Yangyang, et al. Bull face key point detection and pose recognition method based on improved YOLO v7-Pose[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11):84-92, 102. (in Chinese)

[27] 王法安,何忠平,张兆国,等. 基于 PN-YOLO v8s-Pruned 的轻量化三七收获目标检测方法[J]. 农业机械学报, 2024, 55(11):171-183.
Wang Faan, He Zhongping, Zhang Zhaoguo, et al. Lightweight object detection method for *Panax notoginseng* based on PN-YOLO v8s-Pruned[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11):171-183. (in Chinese)

[28] Badgajar C M, Poulouse A, Gan H. Agricultural object detection with You Only Look Once (YOLO) algorithm: a bibliometric and systematic literature review[J]. Computers and Electronics in Agriculture, 2024, 223: 109090.

[29] Khanam R, Hussain M. Yolov11: An overview of the key architectural enhancements[PP/OL]. V1. arXiv (2024-10-23) [2026-04-10]. <https://doi.org/10.48550/arXiv.2410.17725>.

[30] Jegham N, Koh C Y, Abdelatti M, et al. Evaluating the evolution of yolo (you only look once) models: a comprehensive benchmark study of yolo11 and its predecessors[PP/OL]. V4. arXiv (2025-03-17) [2026-04-10]. <https://doi.org/10.48550/arXiv.2411.00201>.

[31] Tong Z, Chen Y, Xu Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[PP/OL]. V3. arXiv (2023-04-28) [2026-04-10]. <https://doi.org/10.48550/arXiv.2301.10051>.

(上接第 31 页)

[21] 颜小飞,孙传京,杜太生,等. 作物茎秆膨胀收缩监测柔性可穿戴传感器研制与试验[J]. 农业工程学报,2024,40(10): 222-227.
Yan Xiaofei, Sun Chuanjing, Du Taisheng, et al. A flexible and wearable sensor to monitor plant stem expansion and contraction[J]. Transactions of the CSAE, 2024,40(10):222-227. (in Chinese)

[22] 王子连. 基于柔性应变传感器的智能设备的研究及应用[D]. 长春:吉林农业大学, 2022.
Wang Zilian. Research and application of intelligent equipment based on flexible strain sensor[D]. Changchun: Jilin Agricultural University, 2022. (in Chinese)

[23] Huang J, Zhou J, Luo Y, et al. Wrinkle-enabled highly stretchable strain sensors for wide-range health monitoring with a big data cloud platform[J]. ACS Applied Materials & Interfaces, 2020, 12(38): 43009-43017.

[24] 孙英,汪忠晟,韩智昊,等. 镓基液态金属柔性应变传感器的设计优化与实验验证[J]. 仪器仪表学报, 2023, 44(1): 16-26.
Sun Ying, Wang Zhongsheng, Han Zhihao, et al. Design optimization and experimental validation of the gallium-based liquid metal flexible strain sensor[J]. Chinese Journal of Scientific Instrument, 2023, 44(1): 16-26. (in Chinese)

[25] Liu J, Zhang J, Liu J, et al. A scalable flexible strain sensor with adaptive response capability to varying microenvironments for digital healthcare[J]. Composites Science and Technology, 2025, 261: 111034.

[26] 赵厚琦. 面向航天器大尺度薄膜结构仿生超敏柔性应变传感元件研究[D]. 长春:吉林大学, 2023.
Zhao Houqi. Research on bionic super-sensitive flexible strain sensor for spacecraft large-scale thin film structure[D]. Changchun:Jilin University, 2023. (in Chinese)

[27] 刘晨阳,张立萍,郑威强,等. 不同含水率对罗布麻力学特性及剥麻效果的影响[J]. 新疆农业科学, 2024, 61(11): 2797-2806.
Liu Chenyang, Zhang Liping, Zheng Weiqiang, et al. Effects of different moisture contents on the mechanical properties and peeling effect of *Apocynum venetum*[J]. Xinjiang Agricultural Sciences, 2024, 61(11): 2797-2806. (in Chinese)

[28] 孟翔宇,刘洋,曹肆林,等. 棉秆在不同弯折方式下的力学性能研究[J]. 安徽农业科学, 2025, 53(19): 175-179,194.
Meng Xiangyu, Liu Yang, Cao Silin, et al. Research on the mechanical properties of cotton stalk under different bending methods[J]. Journal of Anhui Agricultural Sciences, 2025, 53(19): 175-179,194. (in Chinese)

[29] 樊士玉. 夹拔式生姜收获实验平台设计与实验研究[D]. 扬州:扬州大学, 2024.
Fan Shiyu. Design and experimental research of clamp pull ginger harvesting experimental platform[D]. Yangzhou:Yangzhou University, 2024. (in Chinese)

[30] 刘雄,张佳喜,陈明江,等. 棉秆起拔力关键因素的研究与试验[J]. 农机化研究, 2021, 43(9): 157-163.
Liu Xiong, Zhang Jiayi, Chen Mingjiang, et al. Study and experiment on the key factors of pull-up force of cotton stalk[J]. Journal of Agricultural Mechanization Research, 2021, 43(9): 157-163. (in Chinese)