doi:10.6041/j.issn.1000-1298.2026.23.006

基于改进 YOLO 11n－Pose 与有效姿态帧筛选的肉牛体尺自动测量方法

李子若[1,2]　李晓聪[1,2]　张亚丹[1,2]　曾雪婷[1,2]　尹文龙[1,2]　姚　冲[1,2]　刘　刚[1,2]

（1.中国农业大学智慧农业系统集成研究教育部重点实验室，北京 100083；
2.中国农业大学农业农村部农业信息获取技术重点实验室，北京 100083）

摘要：针对肉牛姿态变化对体尺自动测量精度影响较大的问题，本文提出一种基于改进 YOLO 11n－Pose 与有效姿态帧筛选的肉牛体尺自动测量方法。首先，以 YOLO 11n－Pose 为基线模型，嵌入 SimAM 注意力机制，采用 ADown 模块替换 YOLO 11n－Pose 骨干网络中的部分步长卷积下采样结构，并引入 Pose_LQE 定位质量评估分支，增强体尺关键点局部特征表达和定位质量评估能力，降低模型复杂度；其次，针对采集的 RGB－D 图像数据筛选有效姿态帧，定义头部俯仰角、头部偏航角、前肢步幅角和后肢步幅角 4 项姿态参考量，构建有效姿态帧筛选策略，剔除抬头、低头、扭头及前后肢站姿偏差较大的高误差风险帧；最后，结合深度信息恢复关键点三维坐标，自动测量体高、十字部高、胸宽、腰角宽和髋宽等体尺参数。为验证方法有效性，采集了 130 头育肥期西门塔尔肉牛的双视角 RGB－D 图像数据，并构建了 2 000 组 RGB－D 图像对数据集。试验结果分析表明，改进 YOLO 11n－Pose 模型的 mAP@0.5 和 mAP@0.5:0.95 分别为 96.8% 和 93.8%，参数量和计算量分别为 2.22×10^6 和 5.73×10^9，帧率和模型平均推理时间为 134.77 f/s 和 7.42 ms。经有效姿态帧筛选后，体高、十字部高、胸宽、腰角宽和髋宽的相对误差分别降至 1.95%、1.32%、1.73%、1.76% 和 2.14%。研究结果表明，本文方法能够减少非理想姿态对肉牛体尺自动测量精度的影响，为固定通道场景下肉牛体尺非接触式自动测量提供技术支持。

关键词：肉牛；体尺测量；姿态筛选；关键点检测；YOLO 11n－Pose

中图分类号：TP391.41　　**文献标识码**：A　　**文章编号**：1000-1298(2026)23-0069-12　　**OSID**：

Automatic Body Size Measurement Method for Beef Cattle Based on Improved YOLO 11n－Pose and Effective Posture Frame Screening

Li Ziruo[1,2]　Li Xiaocong[1,2]　Zhang Yadan[1,2]　Zeng Xueting[1,2]　Yin Wenlong[1,2]　Yao Chong[1,2]　Liu Gang[1,2]

(1. *Key Laboratory of Smart Agriculture System Integration*, *Ministry of Education*, *China Agricultural University*, *Beijing* 100083, *China*
2. *Key Laboratory of Agricultural Information Acquisition Technology*, *Ministry of Agriculture and Rural Affairs*, *China Agricultural University*, *Beijing* 100083, *China*)

Abstract: Aiming to address the problem that posture variations in beef cattle have a considerable influence on the accuracy of automatic body size measurement, an automatic body size measurement method for beef cattle was proposed based on improved YOLO 11n－Pose and effective posture frame screening. Firstly, YOLO 11n－Pose was used as the baseline model. The SimAM attention mechanism was embedded, the ADown module was adopted to replace part of the stride convolution downsampling structures in the YOLO 11n－Pose backbone network, and a Pose_LQE localization quality estimation branch was introduced to enhance the local feature representation and localization quality evaluation capability of body size keypoints while reducing model complexity. Secondly, effective posture frames were screened from the collected RGB－D image data. Four posture reference variables, namely head pitch angle, head yaw angle, forelimb stride angle, and hindlimb stride angle, were defined, and an effective posture frame screening strategy was constructed to remove high-error-risk frames caused by head

收稿日期：2026－06－26　修回日期：2026－07－07

基金项目：国家重点研发计划项目(2024YFD1300603－3)、中国农业大学 2115 人才工程项目和中央高校基本科研业务费专项资金项目(2025TC013)

作者简介：李子若(2000—)，女，硕士生，主要从事智慧农业研究，E-mail：liziruo123@cau.edu.cn

通信作者：刘刚(1966—)，男，教授，博士生导师，主要从事电子信息技术在农业中的应用研究，E-mail：pac@cau.edu.cn

lifting, head lowering, head turning, and large deviations in forelimb and hindlimb stance. Finally, the three-dimensional coordinates of keypoints were recovered by using depth information, and body height, rump height, chest width, hook width, and hip width were automatically measured. To verify the effectiveness of the proposed method, dual-view RGB - D image data were collected from 130 fattening Simmental beef cattle, and a dataset containing 2 000 pairs of RGB - D images was constructed. The experimental results showed that the improved YOLO 11n - Pose model achieved mAP@0.5 and mAP@0.5:0.95 values of 96.8% and 93.8%, respectively, with 2.22×10^6 parameters and 5.73×10^9. The frame rate and average inference time were 134.77 f/s and 7.42 ms, respectively. After effective posture frame screening, the relative errors of body height, rump height, chest width, hook width, and hip width were reduced to 1.95%, 1.32%, 1.73%, 1.76%, and 2.14%, respectively. The results indicated that the proposed method can reduce the influence of non-ideal postures on automatic body size measurement accuracy and provide technical support for non-contact automatic body size measurement of beef cattle in fixed-channel scenarios.

Key words: beef cattle; body size measurement; posture screening; key point detection; YOLO 11n - Pose

0 引言

体尺参数是表征肉牛生长发育、体型结构和生产性能的重要表型指标,在选种育种、生长监测和精准饲养管理中具有重要应用价值[1]。传统体尺测量主要依赖测杖、卷尺等量具进行人工接触式操作,存在劳动强度大、测量效率低、主观误差明显和重复性差等问题,且极易引起牛只应激反应[2]。因此,开发快速、准确的肉牛体尺非接触式自动测量方法,对提升肉牛表型信息获取效率和养殖管理智能化水平具有重要意义。

随着机器视觉技术的发展,较多学者开展了基于二维图像、RGB - D 图像和三维点云的牲畜体尺自动测量方法的研究[3-8]。综合比较可见,RGB - D 图像在固定场景下更适合兼顾测量精度、处理效率和系统实现可行性,为肉牛体尺自动测量提供了较为适宜的技术路线。

体尺自动测量的核心在于体尺测量关键点的准确、稳定提取。刘刚等[9]基于改进 YOLO v5n - pose 实现了群养生猪体尺自动测量;Yang 等[10]提出面向奶牛体尺测量的一阶段关键点检测网络;Deng 等[11]将 CREStereo 与 MobileViT - Pose 融合用于牛体尺快速测量,提高了测点检测与体尺参数计算流程的一体化水平。曹子卓等[12]以 YOLO 11n - Pose 为基础提出轻量化肉牛体尺关键点检测策略;Weng 等[13]构建 YOLO v8n - DMC 模型实现牛体三维体尺测量,说明关键点检测已成为畜体测点自动提取的重要手段。Peng 等[14]采用改进关键点检测结合单侧深度成像实现牛体尺寸测量;Yue 等[15]进一步利用轻量化 YOLO v8n - Pose 与三维点云实现牛体自动测量;Hou 等[16]通过关键区域识别开展肉牛体尺测量,表明关键点检测正在由单一测点提取向测量流程一体化方向发展。上述研究表明,关键点检测已成为畜体测点自动提取的重要手段,并推动了体尺测量流程向高效化和自动化的方向发展。

除关键测点提取外,姿态变化也是影响畜体自动测量结果的重要因素。初梦苑等[17]通过关键帧提取与头颈部去除提高了奶牛俯视深度图像体尺测量效率与精度,说明关键帧筛选及头颈部干扰处理会影响体尺测量结果。Li 等[18]针对肉牛点云提出基于姿态的测量调整方法,证实姿态偏差会影响体尺测量精度。Luo 等[19]采用统计形状模型进行姿态归一化,表明自由运动条件下姿态变化会导致体尺测量结果不一致。Sun 等[20]提出面向实际应用场景的多姿态山羊体尺测量方法,说明多姿态适应已成为体尺自动测量研究的重要内容。Han 等[21]进一步提出面向羊体点云的非刚性姿态矫正方法,表明姿态校正有助于提升复杂姿态下的体尺测量稳定性。上述研究说明,姿态相关处理对于提高畜体自动测量精度和稳定性具有重要作用。

现有研究已为畜体体尺自动测量提供了较好的技术基础,但在固定通道限位自然站立条件下,肉牛体尺测量仍面临两方面问题:一是缺乏兼顾检测精度、模型轻量化和部署效率的关键点检测方法;二是对体尺测量前非理想姿态帧的识别与筛除关注不足。针对上述问题,本文以固定通道内处于育肥期的西门塔尔肉牛为研究对象,提出一种基于改进 YOLO 11n - Pose 与有效姿态帧筛选的 RGB - D 体尺自动测量方法。首先,对 YOLO 11n - Pose 进行轻量化改进,构建体尺关键点检测模型;其次,构建基于多姿态条件联合约束的有效姿态帧筛选策略,剔除高误差风险帧;最后,结合顶部 RGB - D 图像的深度信息恢复关键点三维坐标,实现体高、十字部高、胸宽、腰角宽和髋宽的自动测量。

1　数据获取

1.1　数据采集

试验数据于 2024 年 12 月在河南省鼎元种牛育种有限公司养殖场采集，试验对象为 130 头育肥期西门塔尔肉牛。试验牛场每月定期开展肉牛称量作业，肉牛需逐头进入养殖场原有称重通道。为结合牛场现有称量管理流程，本文选取该通道作为数据采集区域，架设双视角 RGB－D 采集设备。通道长约 3 m、宽约 1 m、高约 2.5 m，两侧设有围栏，前后设有限位栅栏门。采集时，单头肉牛进入通道后关闭前后栅栏门，限制其前后及横向活动，使其处于限位自然站立状态。每次仅采集 1 头肉牛，以避免多目标遮挡干扰。

数据采集平台采用固定通道双视角 RGB－D 采集方案，由固定通道、限位栅栏、顶部相机、侧面相机、计算机和本地存储单元组成。固定通道双视角 RGB－D 数据采集平台示意图如图 1 所示。顶部相机和侧面相机均采用奥比中光 Gemini 336L 型双目立体视觉相机。顶部相机安装于通道正上方，安装高度约 3.0 m，采集肉牛背部及躯干区域 RGB－D 图像；侧面相机安装于通道一侧围栏外，镜头中心距地面约 1.6 m，与通道中心线水平距离约 2.8 m，采集肉牛侧面轮廓 RGB－D 图像。图像分辨率为 1 280 像素×800 像素，采集帧率为 15 f/s。试验现场相机布设示意图如图 1 所示。

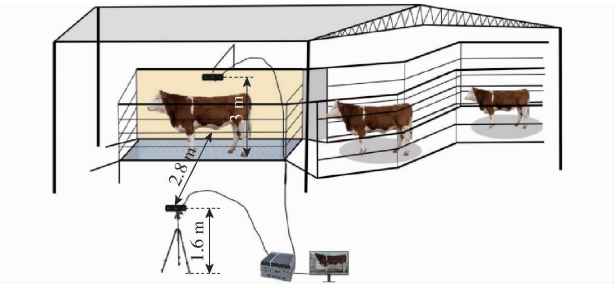


图 1　固定通道双视角 RGB－D 数据采集平台示意图

Fig. 1　Schematic of dual-view RGB－D data acquisition platform in fixed channel

1.2　数据集构建

本文基于固定通道双视角 RGB－D 连续图像构建数据集，剔除目标缺失、个体严重遮挡、图像模糊和深度信息大面积缺失的样本，选取肉牛体表轮廓及关键解剖部位清晰的图像进行人工标注。最终选取 130 头肉牛，体质量范围为 150～900 kg，可覆盖实际生产中常见体型差异肉牛质量分布区间，如图 2 所示。并从 130 头肉牛采集图像中筛选得到 2 000 组 RGB－D 图像对，以肉牛个体为单位按比例 8∶1∶1 划分为训练集、验证集和测试集，其中训练集 104 头、验证集 13 头、测试集 13 头，训练集包含 1 600 组图像对，验证集包含 200 组图像对，测试集包含 200 组图像对。其中训练集在保持关键点相对几何关系不变的前提下进行数据增强，样本数量由 1 600 组扩充至 2 600 组，验证集和测试集不做数据增强，用来真实评价模型效果。数据增强效果示意图如图 3 所示。

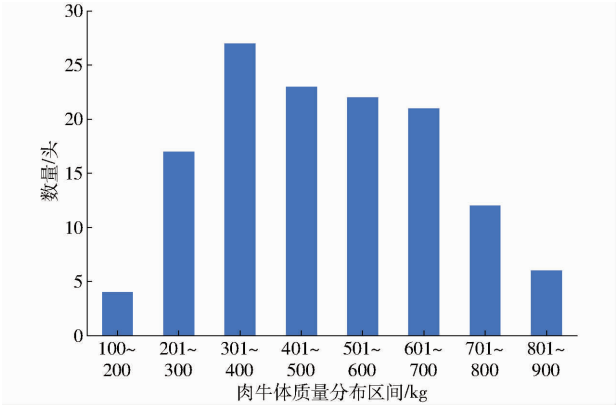


图 2　肉牛体质量分布

Fig. 2　Body weight distribution of beef cattle

根据体尺测量和姿态筛选任务需求，分别构建体尺关键点数据集和姿态参考点数据集。本文利用 Labelme 工具对 RGB 图像进行人工标注。体尺关键点标注时，参考 NY/T 4443—2023《种牛术语》、NY/T 5482—2026《肉牛体型线性鉴定技术规范》和《肉牛性能测定技术手册》中相关测量部位定义，选取 8 个体尺测量关键点，记为 P1～P8，其中，P1 为鬐甲最高点，P2 为十字部最高点，P3、P4 为胸部左右外缘测点，P5、P6 为左右腰角外缘测点，P7、P8 为左右髋部外缘测点，如图 4 所示。

姿态参考点标注时，根据头部俯仰角、头部偏航角、前肢步幅角和后肢步幅角计算需求选取姿态参考点。侧面图像标注肩胛骨上缘点、头顶最高点、前躯参考点、后躯参考点及前后蹄辅助点，用于计算头部俯仰角、前肢步幅角和后肢步幅角；顶部图像标注头部基准点、前躯参考点和后躯参考点，用于计算头部偏航角。标注时仅对体表轮廓清晰、测点位置可判定的图像进行关键点标注；对于左右成对测点，当且仅当两个测点均可见且位置明确时进行标注，否则不标注该组测点。

2　测量方法

2.1　肉牛体尺自动测量方法流程

肉牛体尺测量主要包括数据获取、有效姿态帧筛选、体尺关键点检测与体尺参数计算 4 个环节，流程如图 5 所示。首先，利用固定通道双视角 RGB－D 采集平台获取限位自然站立状态下肉牛顶部和侧

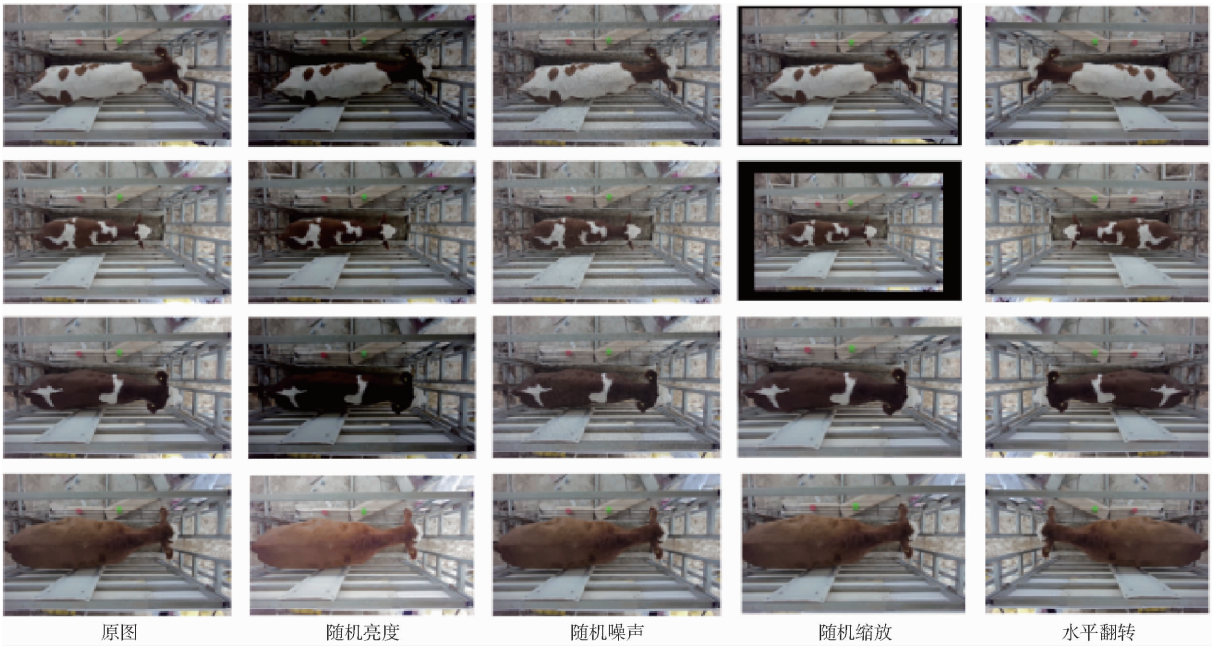


图 3 数据增强效果示意图

Fig. 3 Schematic of data augmentation effects

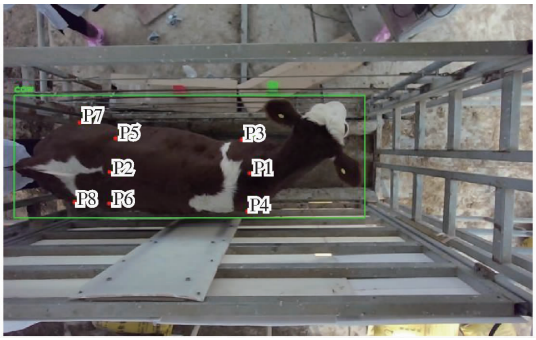


图 4 体尺关键点与姿态参考点定义示意图

Fig. 4 Definition diagram of body size keypoints and posture reference points

面连续图像;其次,采用姿态参考点检测模型提取头部、躯干和四肢姿态参考点,并根据多姿态条件联合约束规则筛选有效姿态帧;再次,采用体尺关键点检测模型提取有效姿态帧中的体尺关键点,结合深度图像和相机参数恢复关键点三维坐标;最后,依据体尺指标几何定义计算体尺参数,并以单牛多帧测量结果的中位数作为最终测量值。

2.2 多姿态条件联合约束的有效姿态帧筛选策略

在固定通道限位自然站立条件下,肉牛前后和横向活动范围虽受到限制,但仍会出现抬头、低头、扭头及前后肢站姿偏差等姿态变化,导致体尺关键点相对位置发生变化,使部分图像帧对应的测量误差显著增大。为减弱姿态变化对后续体尺自动测量精度的影响,本文构建了一种基于姿态参考量与多姿态条件联合约束的有效姿态帧筛选策略。

2.2.1 姿态参考量定义

为定量描述肉牛在采集过程中头颈与肢体姿态

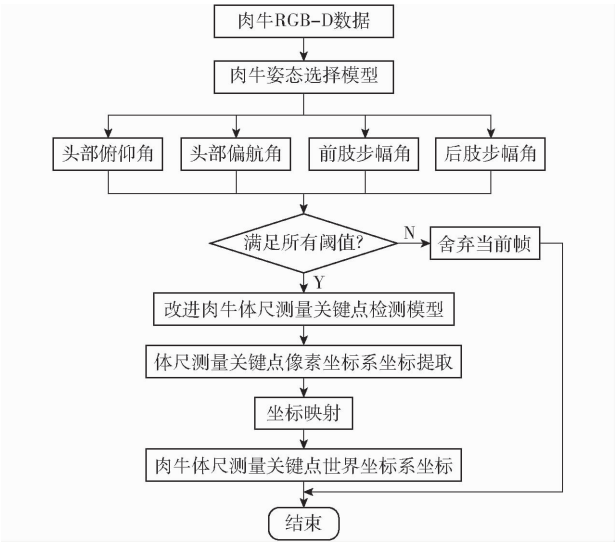


图 5 肉牛体尺自动测量方法总体流程图

Fig. 5 Overall workflow of automatic beef cattle body-size measurement

变化,本文定义头部俯仰角、头部偏航角、前肢步幅角和后肢步幅角 4 个姿态参考量,如图 6 所示。设二维点 $P_i=(x_i,y_i)$ 表示图像中相应关键点坐标,并定义任意两点 P_i、P_j 所连线段的方向角 $\phi_{i,j}$ 为

$$\phi_{i,j}=\operatorname{atan2}(y_j-y_i,x_j-x_i) \tag{1}$$

侧视图像中,肩胛骨上缘点和头顶最高点分别记为 P_s 和 P_h。头部俯仰角 θ_p 定义为线段 P_sP_h 相对于水平方向的夹角,计算公式为

$$\theta_p=\operatorname{atan2}(y_{P_s}-y_{P_h},x_{P_s}-x_{P_h}) \tag{2}$$

顶部图像中,头部基准点、前躯参考点和后躯参考点分别记为 P_b、P_f 和 P_r。头部偏航角 θ_y 定义为头部方向线段 P_bP_f 与躯干中轴线段 P_fP_r 方向角之

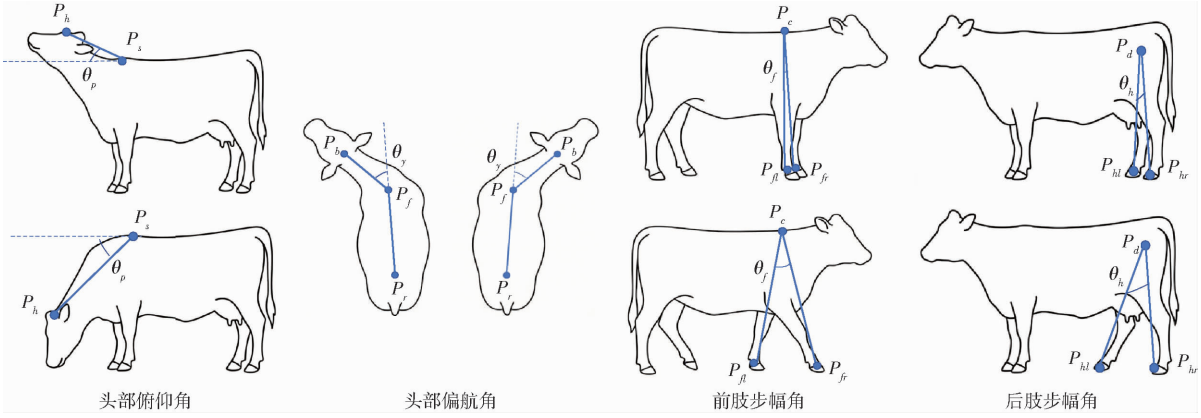


图 6　4 项姿态参考量定义示意图

Fig. 6　Definition of four posture reference variables

差的绝对值，计算公式为

$$\theta_y = |\phi_{P_b,P_f} - \phi_{P_f,P_r}| \tag{3}$$

式中　ϕ_{P_b,P_f}——头部方向线段 P_bP_f 的方向角

ϕ_{P_f,P_r}——躯干中轴线段 P_fP_r 的方向角

侧视图像中，前躯参考点、左前蹄辅助点和右前蹄辅助点分别记为 P_c、P_{fl}和 P_{fr}，后躯参考点、左后蹄辅助点和右后蹄辅助点分别记为 P_d、P_{hl}和 P_{hr}。前肢步幅角 θ_f 和后肢步幅角 θ_h 计算公式分别为

$$\theta_f = |\phi_{P_c,P_{fl}} - \phi_{P_c,P_{fr}}| \tag{4}$$

$$\theta_h = |\phi_{P_d,P_{hl}} - \phi_{P_d,P_{hr}}| \tag{5}$$

式中　$\phi_{P_c,P_{fl}}$——前躯参考点 P_c 与左前蹄辅助点 P_{fl} 构成线段方向角

$\phi_{P_c,P_{fr}}$——前躯参考点 P_c 与右前蹄辅助点 P_{fr} 构成线段方向角

$\phi_{P_d,P_{hl}}$——后躯参考点 P_d 与左后蹄辅助点 P_{hl}构成线段方向角

$\phi_{P_d,P_{hr}}$——后躯参考点 P_d 与右后蹄辅助点 P_{hr}构成线段方向角

2.2.2　有效姿态帧判定规则

为确定姿态筛选阈值，本文基于样本统计结果分析 4 项姿态参考量与体尺测量相对误差之间的关系，并结合高误差样本集中区间和误差变化趋势确定阈值。结果表明，头部俯仰角偏离自然站立范围时，体高和胸宽相对误差增大，取 $-5°\sim32°$作为头部俯仰角筛选参考区间；头部偏航角增大时，体高和胸宽测量误差增加，取 43°作为头部偏航角筛选参考阈值；前肢步幅角增大时，体高和胸宽测量误差增加，取 27°作为前肢步幅角筛选参考阈值；后肢步幅角增大时，十字部高测量误差增加，取 31°作为后肢步幅角筛选参考阈值。

设输入图像帧对应的 4 项姿态参考量分别为 θ_p、θ_y、θ_f 和 θ_h。当头部俯仰角满足 $-5°\leqslant\theta_p\leqslant32°$，头部偏航角满足 $\theta_y\leqslant43°$，前肢步幅角满足 $\theta_f\leqslant27°$，且后肢步幅角满足 $\theta_h\leqslant31°$时，判定该图像帧为有效姿态帧；否则判定为高误差风险帧。该联合约束规则可剔除头部俯仰、头部偏航及前后肢站姿偏差较大的图像帧，减少非理想姿态对后续体尺自动测量的影响。

2.3　改进 YOLO 11n－Pose 体尺关键点检测模型

2.3.1　模型总体结构

体尺关键点定位精度直接影响肉牛体尺自动测量结果。在固定通道限位自然站立条件下，肉牛活动范围受到限制，但头颈姿态变化、前后肢站姿差异仍会影响体尺测点的可见性和定位准确性，同时，肉牛体尺测点多位于体表边缘及骨性标志附近，测点尺度小、局部纹理差异弱，在复杂养殖背景和局部遮挡影响下，易出现测点定位偏移等问题。此外，养殖现场自动测量设备通常受计算资源限制，要求模型在保证检测精度的同时具备较低参数量和计算量。因此，为提高体尺关键点检测精度并降低模型复杂度，本文以 YOLO 11n－Pose 为基线模型，在骨干网络中嵌入 SimAM 注意力机制，增强肉牛体尺测点局部区域特征表达；采用 ADown 模块替换部分步长卷积下采样结构，减小特征图分辨率压缩过程中的空间信息损失；在检测头中引入 Pose_LQE 定位质量评估分支，提高关键点定位结果的可靠性。改进 YOLO 11n－Pose 模型结构如图 7 所示。

2.3.2　SimAM 注意力机制

肉牛体尺测点多位于体表边缘区域，测点尺度小，局部纹理和毛发差异不明显，在复杂养殖背景干扰下，易降低测点区域特征辨识度，影响关键点定位精度。为增强局部测点特征表达能力，本文在 YOLO 11n－Pose 骨干网络中嵌入 SimAM 注意力机制[22]。该模块通过对特征图中关键区域进行自适应重标定，增强鬐甲、十字部及宽度类测点附近的有效信息表达，且无须引入额外可学习参数，可在基本不增加模型复杂度的条件下提高关键点检测精度，

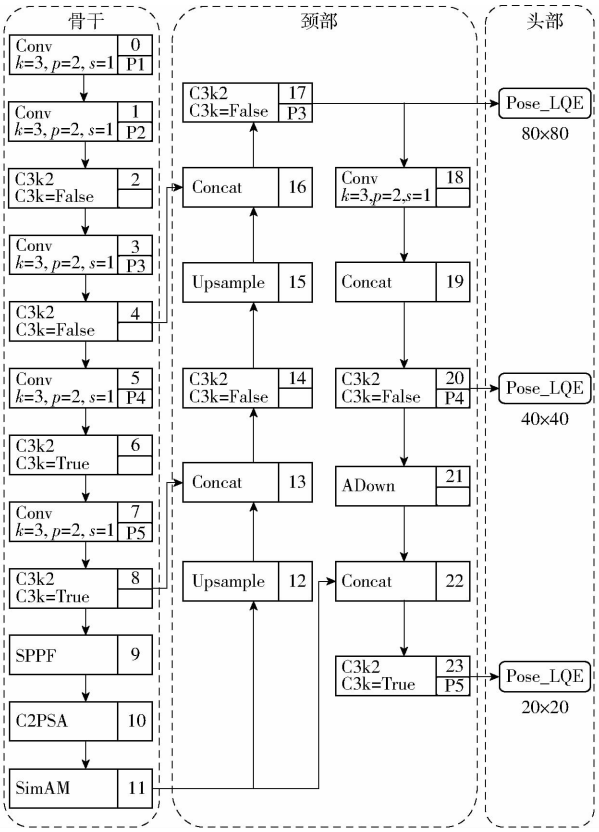


图 7 改进 YOLO 11n－Pose 模型结构图

Fig. 7 Improved YOLO 11n－Pose network structure

SimAM 注意力模块权重生成与作用方式如图 8 所示。

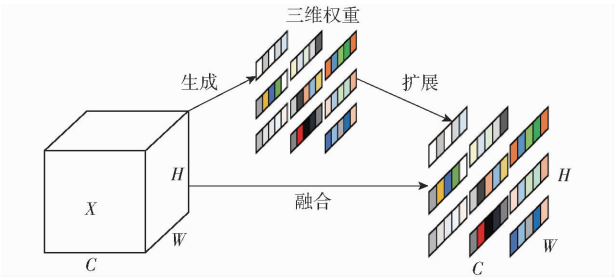


图 8 SimAM 注意力模块权重生成与作用方式

Fig. 8 Weight generation and working mechanism of SimAM attention module

2.3.3 ADown 下采样模块

YOLO 11n－Pose 在特征提取过程中需通过下采样扩大感受野[23]，但常规下采样结构在压缩特征图分辨率时容易造成轮廓边缘和小尺度测点信息损失[24]。为降低模型的参数量和计算量，同时兼顾关键点检测精度，本文采用 ADown 模块[25]替换原始 YOLO 11n－Pose 模型中的部分步长卷积下采样结构。ADown 通过分支特征压缩与融合实现分辨率变换，在减少参数量和计算量的同时保留局部纹理及边缘结构信息，从而提高体尺关键点检测效率和定位准确性，ADown 网络结构如图 9 所示。

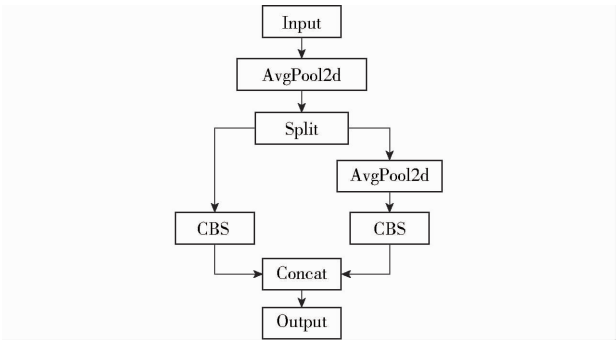


图 9 ADown 网络结构图

Fig. 9 Architecture of ADown network

2.3.4 Pose_LQE 定位质量评估分支

在固定通道采集环境中，部分体尺关键点易受局部遮挡、姿态变化和轮廓模糊影响，出现关键点可被检测但定位精度不足的情况。为增强模型对关键点定位质量的评价能力，本文在 YOLO 11n－Pose 检测头中引入 Pose_LQE 定位质量评估分支[26]。该分支在关键点坐标回归的同时估计关键点定位质量，使模型能够对不同质量的关键点定位结果进行区分，从而提高体尺关键点检测的可靠性，Pose_LQE 内部机制示意图如图 10 所示。

2.4 基于 RGB－D 的体尺参数计算

2.4.1 坐标映射

在完成体尺关键点检测后，本文依据顶部 RGB－D 图像的深度信息和相机内参恢复关键点三维坐标。体高和十字部高由关键点到地面参考平面的垂直距离获得，胸宽、腰角宽和髋宽由左右对应关键点在水平投影平面上的距离获得。设顶部图像中关键点像素坐标为 (u,v)，对应深度为 $D(u,v)$，相机内参分别为 f_x、f_y、c_x 和 c_y。则关键点在顶部相机坐标系下的三维坐标 $\boldsymbol{P}_c=(X_c,Y_c,Z_c)^{\mathrm{T}}$ 可表示为

$$Z_c=D(u,v) \tag{6}$$

$$X_c=\frac{(u-c_x)Z_c}{f_x} \tag{7}$$

$$Y_c=\frac{(v-c_y)Z_c}{f_y} \tag{8}$$

式中 X_c——关键点在顶部相机坐标系下的横向坐标

Y_c——关键点在顶部相机坐标系下的纵向坐标

Z_c——关键点在顶部相机坐标系下的深度坐标

2.4.2 体尺参数计算

根据体尺指标几何定义，体高和十字部高为对应关键点到地面参考平面的垂直距离，胸宽、腰角宽和髋宽为左右对应关键点在水平投影平面上的距离。设地面参考平面为 π，关键点 P_i 到地面参考平面的垂直距离记为 H_i，则高度类体尺参数可表示为

$$H_i=d(P_i,\pi) \tag{9}$$

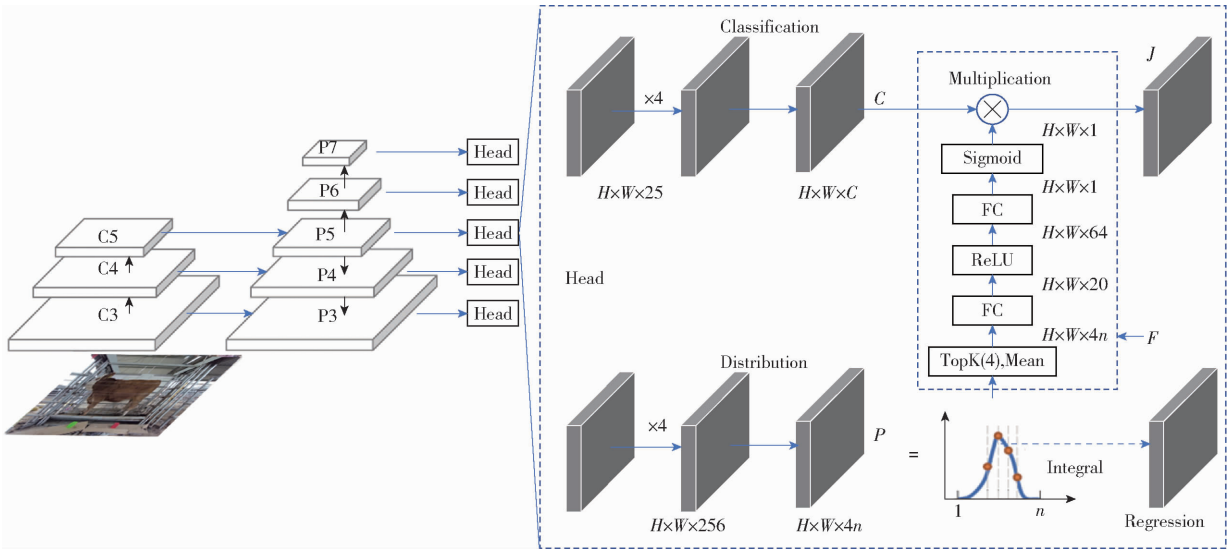


图 10　Pose_LQE 内部机制示意图

Fig. 10　Schematic illustration of internal mechanism of Pose_LQE

式中，$d(\cdot)$表示点到平面的垂直距离。体高由鬐甲最高点 P_1 到地面参考平面的垂直距离表示，十字部高由十字部最高点 P_2 到地面参考平面的垂直距离表示。

对于宽度类指标，设左右对应关键点在顶部相机坐标系水平投影平面上的坐标分别为 $P_j(x_j, y_j)$ 和 $P_k(x_k, y_k)$，则其宽度可表示为

$$W_{j,k} = \sqrt{(x_j - x_k)^2 + (y_j - y_k)^2} \tag{10}$$

式中，$W_{j,k}$为左右对应关键点之间的水平投影距离。胸宽、腰角宽和髋宽分别由 P_3 与 P_4、P_5 与 P_6、P_7 与 P_8 的水平投影距离表示。

采用水平投影距离而非完整三维欧氏距离，可在一定程度上减弱局部深度波动对宽度类参数测量结果的影响。固定通道连续采集过程中，同一头肉牛通常对应多帧有效姿态图像。本文以肉牛个体为统计单元，对保留的多帧有效姿态图像分别进行体尺参数计算，并取中位数作为该牛最终自动测量值，以减小单帧测量波动对结果的影响。

3　试验与结果分析

3.1　试验环境与参数设置

所有试验均在统一的软件与硬件环境下完成，以保证试验结果的可比性与可复现性。试验平台操作系统为 Windows 11，代码运行环境基于 Anaconda 创建的虚拟环境，开发环境为 Python 3.10.16 和 PyTorch 2.2.1，搭配 CUDA 12.6 实现 GPU 硬件加速运算。硬件方面，处理器为 Intel Core i5－14400F，显卡为 NVIDIA GeForce RTX 4060 Ti GPU，显存容量为 8 GB，内存为 24 GB。训练迭代次数（Epochs）设置为 300，批量大小设置为 16，采用 AdamW 优化器，并将初始学习率设定为 0.002，输入图像分辨率统一为 640 像素×640 像素。

3.2　评估指标

为评价所提方法在肉牛体尺自动测量中的有效性，本文分别从体尺关键点检测性能、有效姿态帧筛选效果和体尺自动测量精度 3 个方面进行试验分析。

对于体尺关键点检测模型，采用精确率（Precision，P）、召回率（Recall，R）、mAP@0.5 和 mAP@0.5:0.95 评价检测精度，并以参数量、计算量、帧率（Frames per second，FPS）和平均推理时间（Average inference time，t_{avg}）表征模型复杂度。其中，mAP@0.5 表示 OKS 阈值为 0.5 时的平均精度均值，mAP@0.5:0.95 表示 OKS 阈值从 0.5 变化至 0.95、步长为 0.05 时的平均精度均值。

对于有效姿态帧筛选结果，采用有效姿态帧保留率 R_v 进行评价，计算公式为

$$R_v = \frac{N_v}{N} \times 100\% \tag{11}$$

式中　N——候选图像帧总数

N_v——筛选后保留的有效姿态帧数

对于体尺自动测量结果，采用平均绝对误差（Mean absolute error，MAE）、均方根误差（Root mean square error，RMSE）和平均相对误差（Relative error，RE）进行评价。

3.3　不同模型性能对比

为验证改进 YOLO 11n－Pose 模型在肉牛体尺关键点检测任务中的有效性，选取 YOLO 11n－Pose、YOLO 26n－Pose[26]、YOLO v8n－Pose[15]、RTMPose[27] 和 HRNet[28] 等模型进行对比试验。RTMPose 和 HRNet 在实际应用中采用“目标检测＋

关键点检测”的两阶段流程,需先通过目标检测模型获得肉牛检测框,再输入关键点检测模型进行关键点检测。为反映其实际部署复杂度,本文以 RTMDet 作为前置目标检测模型,并统计 RTMPose + RTMDet 和 HRNet + RTMDet 的整体参数量及计算量,结果如表 1 所示。

表 1 不同关键点检测模型性能对比结果

Tab. 1 Performance comparison results of different keypoint detection models

模型	mAP@0.5/%	mAP@0.5:0.95/%	参数量	计算量	帧率/(f·s^{-1})	平均推理时间/ms
YOLO 11n - Pose	94.5	90.9	2.70×10^{6}	6.72×10^{9}	124.43	8.03
YOLO 26n - Pose	93.6	86.4	2.58×10^{6}	6.02×10^{9}	104.27	9.88
YOLO v8n - Pose	90.8	88.8	3.12×10^{6}	8.50×10^{9}	125.56	7.96
RTMPose	93.0	89.1	3.16×10^{6}	3.4×10^{8}	812.34	1.23
RTMPose + RTMDet	93.0	89.1	1.203×10^{7}	1.509×10^{10}	45.12	22.16
HRNet	98.6	93.2	2.854×10^{7}	7.68×10^{9}	41.75	23.95
HRNet + RTMDet	98.6	93.2	3.741×10^{7}	2.243×10^{10}	18.63	53.68
改进 YOLO 11n - Pose	96.8	93.8	2.22×10^{6}	5.73×10^{9}	134.77	7.42

由表 1 可知,改进 YOLO 11n - Pose 模型在肉牛体尺关键点检测任务中取得了较好的综合性能。与原始 YOLO 11n - Pose 模型相比,改进模型 mAP@0.5 由 94.5% 提高至 96.8%,mAP@0.5:0.95 由 90.9% 提高至 93.8%,分别提高 2.3、2.9 个百分点;参数量由 2.70×10^{6} 降至 2.22×10^{6},计算量由 6.72×10^{9} 降至 5.73×10^{9},分别降低 17.6% 和 14.7%,帧率由 124.43 f/s 提升至 134.77 f/s,平均推理时间由 8.03 ms 降至 7.42 ms,降低了 7.60%。与 YOLO 26n - Pose、YOLO v8n - Pose 和 RTMPose 相比,改进 YOLO 11n - Pose 在检测精度和模型复杂度方面均具有优势;与 HRNet 等高精度模型相比,改进 YOLO 11n - Pose 在保持相近甚至更高 mAP@0.5:0.95 的同时,显著降低参数量和计算量,综合考虑精度、模型复杂度和边缘设备部署成本,更适用于固定通道场景下肉牛体尺关键点检测任务。

3.4 消融试验

为对比分析不同改进模块对模型检测精度和计算复杂度的影响,开展消融试验,结果如表 2 所示。仅引入 SimAM 注意力机制时,模型 mAP@0.5 和 mAP@0.5:0.95 分别提高 0.4、0.6 个百分点,参数量和计算量基本不变,在不增加运算成本的条件下提升了肉牛体尺关键点检测精度。仅采用 ADown 下采样模块时,模型参数量和计算量分别降低 17.8% 和 14.7%,达到了较好的轻量化效果,同时 mAP@0.5:0.95 提高 1.2 个百分点。仅引入 Pose_LQE 定位质量评估分支时,模型召回率提高 2.2 个百分点,mAP@0.5:0.95 提高 1.6 个百分点,提升了体尺关键点定位质量评估能力。3 个模块同时应用于 YOLO 11n - Pose 模型后,mAP@0.5 和 mAP@0.5:0.95分别达到 96.8% 和 93.8%,较基线模型提高 2.3、2.9 个百分点;参数量和计算量分别降低 17.6% 和 14.7%。结果表明,各改进模块协同作用后,可在降低模型复杂度的同时提高肉牛体尺关键点检测精度,改进前后体尺关键点检测结果对比如图 11 所示。

表 2 各改进模块对 YOLO 11n - Pose 模型性能的影响

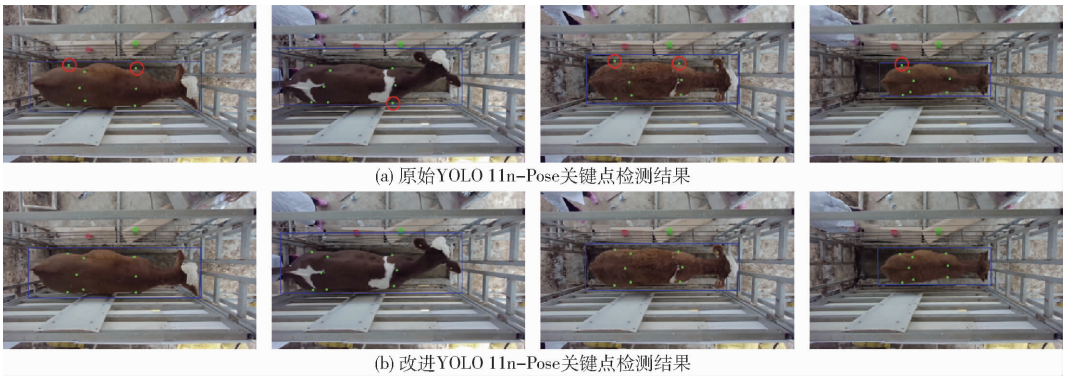
Tab. 2 Effects of different modules on YOLO 11n - Pose performance

SimAM	ADown	Pose_LQE	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%	参数量	计算量
-	-	-	91.6	90.5	94.5	90.9	2.697×10^{6}	6.72×10^{9}
√	-	-	92.3	90.4	94.9	91.5	2.697×10^{6}	6.72×10^{9}
-	√	-	92.5	91.4	95.3	92.1	2.217×10^{6}	5.73×10^{9}
-	-	√	91.4	92.7	95.6	92.5	2.701×10^{6}	6.72×10^{9}
√	√	-	92.9	91.2	95.5	92.6	2.217×10^{6}	5.73×10^{9}
√	-	√	92.1	92.6	95.9	92.9	2.701×10^{6}	6.72×10^{9}
-	√	√	92.6	93.4	96.4	93.3	2.222×10^{6}	5.73×10^{9}
√	√	√	93.4	93.8	96.8	93.8	2.222×10^{6}	5.73×10^{9}

注:“√”表示引入该模块,“-”表示未引入该模块。

3.5 姿态变量对肉牛体尺测量误差的影响分析

在固定通道限位自然站立条件下,肉牛虽然前后和横向活动范围受到一定限制,但仍会出现头部俯仰、头部偏航以及前后肢步幅变化等非理想姿态。



(a) 原始YOLO 11n-Pose关键点检测结果

(b) 改进YOLO 11n-Pose关键点检测结果

图 11　改进前后体尺关键点检测结果对比

Fig. 11　Comparison of keypoint detection results before and after improvement

上述姿态变化会改变头颈部、前躯和后躯区域的空间位置关系，引起体尺关键点定位偏移和深度值波动，进而影响体尺参数计算结果。为进一步分析姿态变化对肉牛体尺测量误差的影响，分别统计头部俯仰角、头部偏航角、前肢步幅角和后肢步幅角与体高、十字部高、胸宽、腰角宽、髋宽相对误差之间的关系，结果如图 12 所示。

由图 12 可知，不同姿态变量对各项体尺指标测量误差的影响程度存在差异。头部俯仰角主要反映肉牛头颈部上下偏转状态，当肉牛头部处于自然站立范围时，体高和胸宽测量误差相对较小，散点分布较为集中；当头部明显低垂或过度抬升时，体高和胸宽相对误差增大，高误差样本数量增加。这主要是由于头部俯仰变化会改变颈部和前躯区域的外观轮廓，影响鬐甲附近关键点定位以及胸部区域测点的深度映射稳定性。

头部偏航角主要表征肉牛头部相对于躯干方向的左右偏转程度。当头部偏航角较小时，肉牛头颈

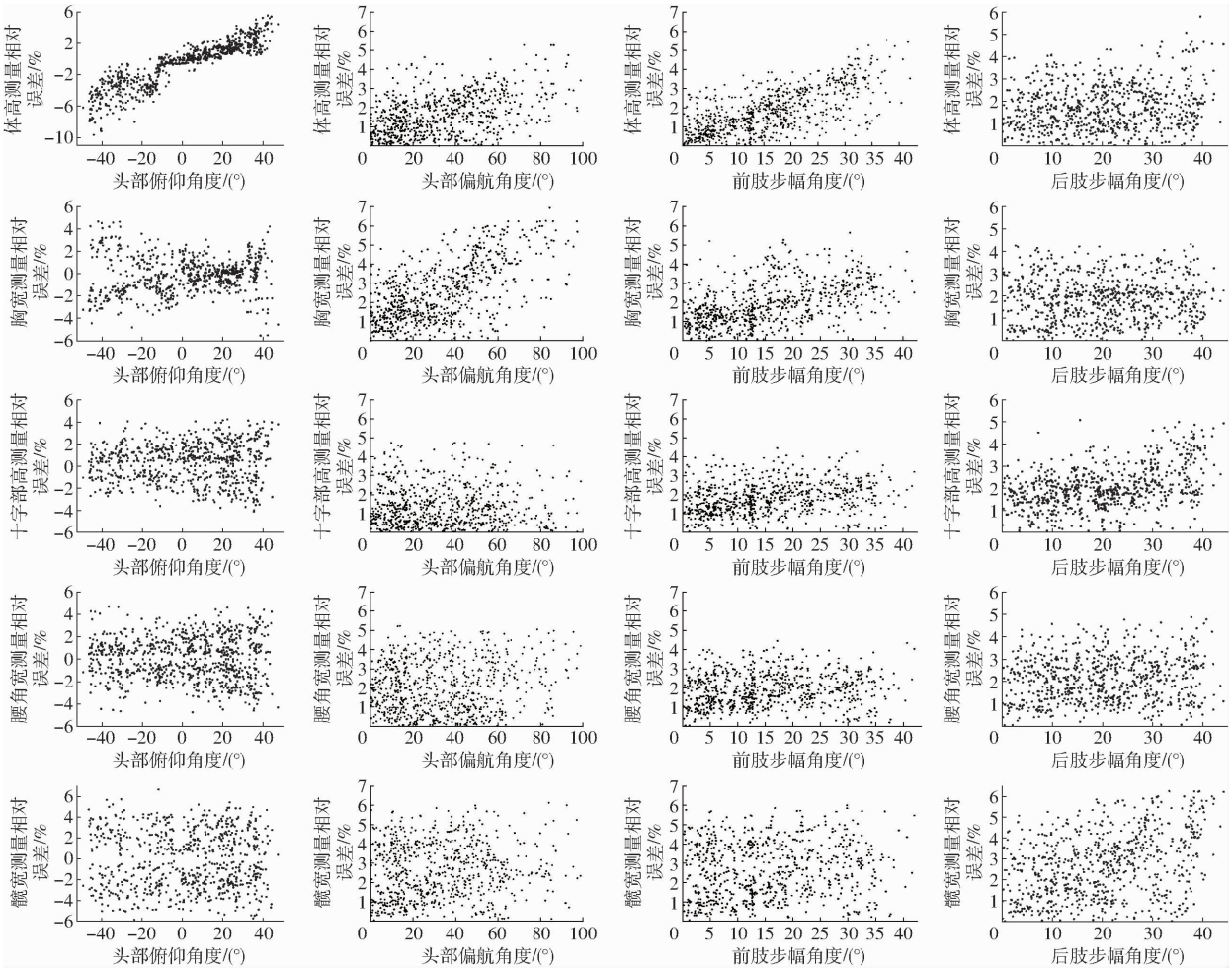


图 12　不同姿态变量对肉牛体尺测量误差的影响

Fig. 12　Effects of different posture variables on body-size measurement errors

部与躯干方向基本一致,体尺关键点空间关系较为稳定;当肉牛发生明显扭头时,侧视图中肩前区域和胸部轮廓发生变化,易造成体高和胸宽测量误差增大。相比之下,十字部高、腰角宽和髋宽对应测点主要位于躯干中后部及骨盆区域,受头部偏航影响相对较小,因此误差变化趋势不如前躯相关指标明显。

前肢步幅角用于描述肉牛前肢站姿偏差程度。当前肢步幅角较小时,前躯支撑状态较稳定,体表关键点的空间映射结果较可靠;当前肢步幅角增大时,肉牛前躯支撑关系发生变化,肩部、胸部及背部局部姿态随之改变,导致体高和胸宽测量误差增加。对于十字部高、腰角宽和髋宽等位于躯干中后部的体尺指标,前肢步幅变化对其影响相对有限,误差散点整体未表现出明显发散趋势。

后肢步幅角主要反映肉牛后肢站姿及后躯支撑状态变化。当后肢步幅角增大时,十字部高测量误差更容易增大,这是因为十字部高对应关键点位于后躯上方,后肢跨步会改变骨盆和后躯支撑状态,使十字部最高点的空间位置及深度映射结果发生波动。相比之下,体高、胸宽、腰角宽和髋宽受后肢步幅角影响相对较弱,误差变化整体较为平缓。

综上所述,头部俯仰角、头部偏航角、前肢步幅角和后肢步幅角均会在一定程度上影响肉牛体尺自动测量结果,且不同姿态变量对各体尺指标的影响程度存在差异。基于上述分析,采用多姿态条件联合约束规则筛选有效姿态帧,以剔除抬头、低头、扭头以及前后肢站姿偏差较大的图像帧,从而为后续体尺关键点检测和体尺参数计算提供更加稳定的输入。

3.6 有效姿态帧筛选结果分析

依据多姿态条件联合约束规则,对测试集肉牛连续 RGB-D 图像进行有效姿态帧筛选。测试集共包含 13 头肉牛,每头肉牛选取 900 帧连续图像作为候选图像帧,候选图像帧总数为 11 700 帧。经姿态筛选后,保留有效姿态帧 4 001 帧,有效姿态帧保留率为 34.2%。单头肉牛平均有效姿态帧数为 307.8 帧,最少为 246 帧,最多为 369 帧。结果分析表明,经多姿态条件联合筛选后,各测试集个体均保留了较充足的有效姿态帧,能够满足后续单牛多帧体尺测量结果聚合需求。

有效姿态帧与典型无效姿态帧示例如图 13 所示,其中有效姿态帧中肉牛头部、躯干和四肢姿态较为稳定,顶部视角下鬐甲、十字部、胸部、腰角和髋部等体尺测点位置清晰,关键区域深度信息连续性较好。被剔除图像帧主要表现为抬头、低头、扭头以及前后肢站姿偏差较大等类型,上述姿态变化会改变头颈部、前躯和后躯区域的空间关系,会引起鬐甲、十字部、胸部、腰角和髋部等测点位置及深度值波动,从而增加关键点定位误差和体尺测量误差。试验表明基于 4 项姿态参考量的联合约束规则,可有效剔除姿态偏差较大的高误差风险帧,降低非理想姿态对后续体尺关键点检测和体尺参数计算的影响。

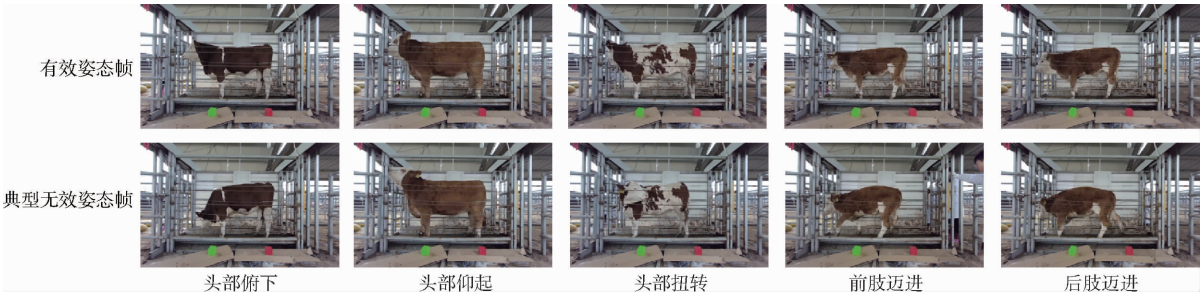


图 13 有效姿态帧与典型无效姿态帧示例

Fig. 13 Examples of valid and typical invalid posture frames

3.7 姿态筛选前后体尺测量精度对比分析

为验证有效姿态帧筛选对肉牛体尺自动测量精度的影响,以人工测量值作为参考值,分别在未施加姿态筛选和施加有效姿态帧筛选两种条件下进行体尺测量,对同一头肉牛的多帧测量结果取中位数,作为该牛最终的自动测量值。未施加姿态筛选时,候选图像帧全部参与单牛多帧测量结果聚合;施加有效姿态帧筛选时,仅保留满足多姿态条件联合约束规则的图像帧参与体尺关键点检测和体尺参数计算,筛选前后体尺自动测量误差统计结果如表 3 所示。

表 3 姿态筛选前后体尺自动测量误差对比结果

Tab. 3 Comparison results of automatic body size measurement errors before and after posture filtering

参数	筛选前			筛选后		
	MAE/mm	RMSE/mm	RE/%	MAE/mm	RMSE/mm	RE/%
体高	40.78	48.32	3.30	23.88	25.74	1.95
十字部高	31.58	37.14	2.43	17.04	20.05	1.32
胸宽	11.29	14.01	2.67	7.12	7.75	1.73
腰角宽	10.51	13.35	2.77	6.44	7.06	1.76
髋宽	14.31	17.81	3.00	9.67	10.78	2.14

由表 3 可知，施加有效姿态帧筛选后，5 项体尺指标的 MAE、RMSE 和 RE 均有所降低。体高、十字部高、胸宽、腰角宽和髋宽的 RE 分别由 3.30%、2.43%、2.67%、2.77% 和 3.00% 降至 1.95%、1.32%、1.73%、1.76% 和 2.14%，均小于 2.5%。其中，体高和十字部高的 MAE 分别由 40.78 mm 和 31.58 mm 降至 23.88 mm 和 17.04 mm，胸宽、腰角宽和髋宽的 MAE 分别由 11.29、10.51、14.31 mm 降至 7.12、6.44、9.67 mm。各指标 RMSE 同步下降，说明姿态筛选后较大误差样本对最终测量结果的影响减小。

施加有效姿态帧筛选后，5 项体尺指标的 RE 均明显降低。其中，体高和十字部高的 RE 分别降低 1.35、1.11 个百分点，胸宽、腰角宽和髋宽的 RE 分别降低 0.94、1.01、0.86 个百分点。5 项体尺指标总体平均 RE 由 2.83% 降至 1.78%，降低 1.05 个百分点，降幅达 37.1%。结果分析表明，有效姿态帧筛选能够显著减小非理想姿态对肉牛体尺自动测量结果的影响。

经有效姿态帧筛选后，肉牛 5 项体尺指标的自动测量误差均有不同程度降低。其主要原因是姿态筛选剔除了抬头、低头、扭头及前后肢站姿偏差较大的高误差风险帧，减少了非理想姿态对体尺关键点定位和深度值获取的影响。体高和十字部高对头颈姿态及后躯支撑状态变化较敏感，因此误差降低幅度相对较大；胸宽、腰角宽和髋宽主要由顶部视角下左右测点水平投影距离计算，姿态筛选后躯干投影关系更加稳定，测量误差也相应降低。结果分析表明，基于改进 YOLO 11n－Pose 与有效姿态帧筛选的肉牛体尺自动测量方法能够提高固定通道限位自然站立条件下体尺测量精度。

4　结论

(1)建立了一种基于改进 YOLO 11n－Pose 的肉牛体尺关键点检测模型。在原始 YOLO 11n－Pose 模型基础上嵌入 SimAM 注意力机制，采用 ADown 下采样模块替换部分下采样结构，并引入 Pose_LQE 定位质量评估分支，提高了体尺测点局部特征表达能力和关键点定位质量。改进后模型的 mAP@0.5 和 mAP@0.5:0.95 分别达到 96.8% 和 93.8%，较原始模型提高了 2.3、2.9 个百分点；参数量和计算量分别为 2.22×10^{6} 和 5.73×10^{9}，较原始模型降低了 17.6% 和 14.7%；帧率为 134.77 f/s，较原始模型提升了 8.31%，平均推理时间为 7.42 ms，较原始模型降低了 7.60%。说明改进 YOLO 11n－Pose 在肉牛体尺关键点检测中具有较好的精度和轻量化优势。

(2)构建了一种固定通道内自然站立条件下的肉牛有效姿态帧筛选策略。通过定义头部俯仰角、头部偏航角、前肢步幅角和后肢步幅角 4 项姿态参考量，提出了多姿态条件联合约束规则。在测试集 11 700 帧连续图像中，筛选得到有效姿态帧 4 001 帧，有效姿态帧保留率为 34.2%。该策略可剔除姿态偏差较大的高误差风险帧，为后续体尺关键点检测和体尺参数计算提供满足测量要求的图像帧。

(3)提出了一种基于改进 YOLO 11n－Pose 关键点检测与有效姿态帧筛选的肉牛体尺自动测量方法，通过 RGB－D 图像三维坐标恢复，自动测量肉牛体高、十字部高、胸宽、腰角宽和髋宽等体尺参数。未进行姿态筛选时，5 项体尺指标相对误差分别为 3.30%、2.43%、2.67%、2.77% 和 3.00%；施加有效姿态帧筛选后，相对误差分别降至 1.95%、1.32%、1.73%、1.76% 和 2.14%。5 项体尺指标总体平均相对误差降低了 37.1%。结果分析表明，本文所提方法能够减小非理想姿态对肉牛体尺自动测量结果的影响，为固定通道场景下肉牛体尺非接触式自动测量提供技术支持。

参　考　文　献

[1] Yin Z Y, Wu T, Liu L, et al. Computer vision-based body measurement of cattle and pigs: a systematic review[J]. Computers and Electronics in Agriculture, 2026, 247: 111688.

[2] 谢秋菊，周红，包军，等. 基于机器视觉的畜禽体质量评估研究进展[J]. 农业机械学报，2022，53(10)：1－15.
Xie Qiuju, Zhou Hong, Bao Jun, et al. Review on machine vision-based weight assessment for livestock and poultry[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(10): 1－15. (in Chinese)

[3] Tasdemir S, Urkmez A, Inal S. Determination of body measurements on Holstein cows using digital image analysis and estimation of live weight with regression analysis[J]. Computers and Electronics in Agriculture, 2011, 76(2): 189－197.

[4] Yang G Y, Xu X S, Song L, et al. Automated measurement of dairy cows body size via 3D point cloud data analysis[J]. Computers and Electronics in Agriculture, 2022, 200: 107218.

[5] 赵健，周国源，王智文，等. 基于 SGPointNet++模型的奶牛点云分割与表型自动测定系统设计[J]. 农业机械学报，2025，56(3)：180－187.
Zhao Jian, Zhou Guoyuan, Wang Zhiwen, et al. Design of automatic determination system for point cloud segmentation and

morphology of dairy cows based on SGPointNet + + model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025,56(3):180 - 187. (in Chinese)

[6] Du A, Guo H, Lu J, et al. Automatic livestock body measurement based on keypoint detection with multiple depth cameras [J]. Computers and Electronics in Agriculture, 2022, 198: 107059.

[7] 邓洪兴,许兴时,王云飞,等. 基于双目立体匹配与改进 YOLOv8n - Pose 关键点检测的奶牛体尺测量方法[J]. 华南农业大学学报,2024,45(5):802 - 811.
Deng Hongxing, Xu Xingshi, Wang Yunfei, et al. Dairy cow body size measurement method based on binocular stereo matching and improved YOLOv8n - Pose keypoint detection[J]. Journal of South China Agricultural University, 2024, 45(5): 802 - 811. (in Chinese)

[8] Li R, Wen Y, Zhang S, et al. Automated measurement of beef cattle body size via key point detection and monocular depth estimation[J]. Expert Systems with Applications, 2024, 244: 123042.

[9] 刘刚, 曾雪婷, 刘晓文, 等. 基于改进 YOLO v5 - pose 的群养生猪体尺自动测量方法[J]. 农业机械学报,2025,56(5): 455 - 465.
Liu Gang, Zeng Xueting, Liu Xiaowen, et al. Automatic measurement method of body size of group-raised pigs based on improved YOLO v5 - pose[J]. Transactions of the Chinese Society for Agricultural Machinery,2025,56(5):455 - 465. (in Chinese)

[10] Yang G Y, Qiao Y L, Deng H X, et al. One-stage keypoint detection network for end-to-end cow body measurement[J]. Engineering Applications of Artificial Intelligence, 2025, 146: 110333.

[11] Deng H X, Yang G Y, Xu X S, et al. Fusion of CREStereo and MobileViT - Pose for rapid measurement of cattle body size [J]. Computers and Electronics in Agriculture, 2025, 232: 110103.

[12] 曹子卓, 郭小燕, 李艳梅, 等. 基于分段建模的肉牛体尺关键点检测轻量化策略[J]. 华南农业大学学报, 2026, 47(1): 139 - 149.
Cao Zizhuo, Guo Xiaoyan, Li Yanmei, et al. A lightweight strategy for cattle body size key point detection based on segmented modeling[J]. Journal of South China Agricultural University, 2026, 47(1): 139 - 149. (in Chinese)

[13] Weng Z, Hao W W, Gong C L, et al. YOLOv8 - DMC: enabling non-contact 3D cattle body measurement via enhanced keypoint detection[J]. Animals, 2025, 15(18): 2738.

[14] Peng C, Cao S S, Li S J, et al. Automated measurement of cattle dimensions using improved keypoint detection combined with unilateral depth imaging[J]. Animals, 2024, 14(17): 2453.

[15] Yue C H, Zhang X H, Zhang Y J, et al. Automatic and accurate measurement of cattle body based on a lightweight YOLOv8 - Pose model and 3D point cloud[J]. Machine Learning with Applications, 2025, 20: 100662.

[16] Hou Z X, Zhang Q, Zhang B, et al. CattlePartNet: an identification approach for key region of body size and its application on body measurement of beef cattle[J]. Computers and Electronics in Agriculture, 2025, 232: 110013.

[17] 初梦苑,李孟飞,李前, 等. 基于关键帧提取与头颈部去除的奶牛体尺测量方法[J]. 农业机械学报,2022,53(增刊 2): 226 - 233, 259.
Chu Mengyuan, Li Mengfei, Li Qian,et al. Method of cows body size measurement based on key frame extraction and head and neck removal[J]. Transactions of the Chinese Society for Agricultural Machinery,2022,53(S2):226 - 233, 259. (in Chinese)

[18] Li J W, Ma W H, Bai Q, et al. A posture-based measurement adjustment method for improving the accuracy of beef cattle body size measurement based on point cloud data[J]. Biosystems Engineering, 2023, 230: 171 - 190.

[19] Luo X Y, Hu Y, Gao Z C, et al. Automated measurement of livestock body based on pose normalisation using statistical shape model[J]. Biosystems Engineering, 2023, 227: 36 - 51.

[20] Sun Y, Li Q F, Ma W H, et al. A multi-posture adaptive method for measuring goat bodies dimensions using 3D point clouds in real-world applications[J]. Biosystems Engineering, 2025, 257: 104214.

[21] Han B B, Zhang J R, Xiang Y Y, et al. Shapewarp: a "global-to-local" non-rigid sheep point cloud posture rectification method[J]. Expert Systems with Applications, 2025, 270: 126524.

[22] Yang L, Zhang R Y, Li L, et al. SimAM: a simple, parameter-free attention module for convolutional neural networks[C]// Proceedings of the 38th International Conference on Machine Learning. PMLR, 2021: 11863 - 11874.

[23] Hesse R, Schaub-Meyer S, Roth S. Content-adaptive downsampling in convolutional neural networks[C]// 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2023:4543 - 4552.

[24] 全国畜牧总站. 肉牛性能测定技术手册[M]. 北京:中国农业出版社,2019.

[25] Li X, Wang W, Hu X, et al. Generalized focal loss V2: learning reliable localization quality estimation for dense object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 11627 - 11636.

[26] Jocher G, Qiu J, Liu M, et al. Ultralytics YOLO26: unified real-time end-to-end vision models[PP/OL]. V1. arXiv (2026 - 06 - 02)[2026 - 06 - 26]. https://arxiv.org/abs/2606.03748.

[27] Jiang T, Lu P, Zhang L, et al. RTMPose: real-time multi-person pose estimation based on MMPose[PP/OL]. V2. arXiv (2023 - 07 - 03)[2026 - 06 - 26]. https://arxiv.org/abs/2303.07399.

[28] Sun K, Xiao B, Liu D, et al. Deep high-resolution representation learning for human pose estimation[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 5693 - 5703.