

杜粤猛,史慧,高峰,等.基于姿态估计和关键点特征向量的奶牛跛行识别方法[J].华中农业大学学报,2023,42(5):251-261.  
DOI:10.13300/j.cnki.hnlkxb.2023.05.028

# 基于姿态估计和关键点特征向量的奶牛跛行识别方法

杜粤猛,史慧,高峰,邓红涛

石河子大学机械电气工程学院,石河子 832003

**摘要** 针对目前养殖场中自动检测奶牛跛行效率低、准确率不高等问题,设计了一种基于姿态估计和膝关节角度特征向量的奶牛跛行识别方法。鉴于奶牛行为具有随机停留的特点,制作奶牛在不同远近视场尺度和观测角度等条件下奶牛姿态估计数据集。将Faster RCNN卷积神经网络模型引入到奶牛关键点检测中提高跛行识别的可靠性;以ResNet101网络作为特征提取网络,构建奶牛姿态估计网络,并采用超参数微调训练方法,对网络模型进行迁移训练。通过视频中的奶牛姿态信息和关键点坐标信息,计算出奶牛行走时膝关节的角度特征,并利用1-D Convolution分类模型实现奶牛的跛行识别。实验结果显示:以ResNet101网络模型为基础的奶牛姿态估计网络的PCK@0.1值可以达到0.925 0;使用1-D Convolution模型对奶牛行为分类识别的准确率为97.22%,与LSTM、Bi-LSTM、GRU模型相比,分别提高5.55、2.78、11.11个百分点。以上结果表明,所提方法对自然环境下奶牛跛行有较好的检测效果,可用于奶牛智能化养殖并为养殖管理提供技术参考。

**关键词** 奶牛;深度学习;姿态估计;时序特征;行为识别

**中图分类号** TP391 **文献标识码** A **文章编号** 1000-2421(2023)05-0251-11

奶牛的健康状态是提高奶量和优质生产的重要保证,同时也是牧场和个体养殖户追求的目标。圈养饲喂环境下,奶牛因各种蹄病出现的跛行行为引起养殖场的严重经济损失<sup>[1-2]</sup>。奶牛蹄部出现损伤时,会对奶牛的行走行为产生影响,包括肢体关节的弯曲角度<sup>[3]</sup>、甩腿程度<sup>[4]</sup>、弓背参数<sup>[5-6]</sup>、牛蹄跟随性<sup>[7]</sup>以及行走速度<sup>[8]</sup>等。因此,可以通过识别奶牛跛行的这些特征及早发现奶牛蹄病。

传统的奶牛跛行检测采用人工方法,对奶牛跛行的判定依赖观察者的主观判断,不仅费时费力,且结果也会产生偏差。接触式方法大多使用电子项圈、电子脚环等可穿戴设备,通过这些设备获取奶牛的步幅等行为特征参数<sup>[9]</sup>,但接触式方法容易造成奶牛应激反应,可能影响奶牛的正常生产。相比上述方法,由视频图像处理技术和机器视觉等技术组成的非接触式检测方法应用愈加广泛。基于机器视觉的奶牛行为识别方法,包括躯体关键点识别<sup>[10-11]</sup>和目标检测结合行为识别<sup>[12-14]</sup>、行为分类<sup>[15-16]</sup>等。如付涛<sup>[17]</sup>以农场拍摄的奶牛视频作为研究对象,提出一种基于卷积神经网络模块级联的牛体姿态估计方

法,在验证集上的准确率最高可达83.91%;Wu等<sup>[13]</sup>使用YOLOv3网络提取奶牛腿部、头部和躯干图像,利用跛足奶牛相对步长呈现不规律变化的特征,构建相对步长特征向量,输入基于LSTM的分类器识别跛足。

尽管基于机器视觉的奶牛跛行识别取得了一些成果,但仍然存在以下问题:一是奶牛具有体型大、四肢相似度高、对称性高的特点,基于目标检测方法难以精准检测到奶牛行走特征;二是现有的特征分类方法,对奶牛跛行特征不敏感,分类效果不佳。针对以上问题,本研究基于Faster RCNN和ResNet101网络模型,提出了一种基于姿态估计和膝关节角度特征向量的奶牛跛行识别方法,以实现自然养殖环境下奶牛跛行的有效识别。

## 1 材料与方法

### 1.1 实验数据获取及预处理

研究的实验数据采集于新疆石河子市天路牧场,为了高度还原奶牛真实的生活环境,对行走和站立状态的奶牛进行拍摄得到视频数据集。采集设备

收稿日期:2022-11-17

基金项目:新疆维吾尔自治区岗位体系专家项目(YTHSD2022-19);石河子大学科研能力提升项目(KX01230305)

杜粤猛,E-mail:245321481@qq.com

通信作者:邓红涛,E-mail:denghtshzu@163.com

主要参数为:1 200万像素摄像头,IMX703传感器,帧率为30帧/s,分辨率为1 920像素×1 080像素,每个视频约100~200帧,采集数据时闪光灯保持关闭状态。

研究共采集了228段奶牛行为视频,其中随机选取40段跛行、40段站立、25段正常行走视频用于视频分帧和训练奶牛姿态估计网络,剩余的123段视频用于跛行识别。

为加强模型的鲁棒性和防止模型出现过拟合问题,人工挑选视频分帧后的1 050幅图像数据,并采用水平镜像数据增强的方式扩充样本容量。在深度学习中,对图像的数据增强操作大多是对图像进行平移、旋转、增加噪声、增减亮度等方式来扩充实验数据集。但在奶牛的姿态估计研究方面,上述的数据增强方法对奶牛的姿态变换并无影响。本研究数据增强则采用水平镜像的方式,此方法主要改变了牛体的左右肢体骨架结构<sup>[18]</sup>,扩充后的样本数量为2 100幅完整的奶牛牛体图像。

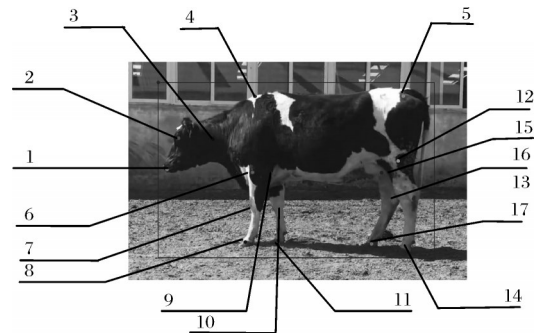
## 1.2 奶牛姿态估计数据集制作

本研究中的奶牛姿态估计数据集需要对牛体的位置信息和牛体关键点信息进行标注。奶牛出现跛行后,奶牛头部和四肢部位相较于奶牛背部、颈部和尾椎等部位对行走特征影响较大,因此,本研究忽略奶牛的尾巴区域,仅对奶牛的头部、颈部、脊椎和四肢等部分进行标注。本次标签制作使用的标注工具为LabelMe软件。主要标注牛体的位置信息和牛体的17个关键点信息,对于被遮挡但能够估计出其位置的关键点,在进行标注时把该关键点认为是可见的并进行标注,具体标注信息如图1所示。本研究的奶牛姿态估计数据集已在SneakerDuYM/cow\_data(github.com)共享。

## 1.3 奶牛姿态估计网络和膝关节角度特征向量的奶牛跛行识别

1)总体技术路线。研究设计一种基于姿态估计和膝关节角度特征向量的奶牛跛行识别方法,其技术路线如图2所示。通过获取奶牛行走、跛行和站立3类行为的视频数据,构建奶牛图像关键点数据集和制备奶牛行为数据集,并结合迁移学习方法对本研究提出的奶牛姿态估计网络模型进行训练,实现奶牛的前腿腿根、膝关节、牛蹄等关键点检测,获取其特征信息;根据检测到的关键点位置坐标信息计算出前腿膝关节角度,将所得到的角度按视频帧的顺

序构造膝关节角度的一维特征向量;将该特征向量送入1-D Convolution分类模型,最终能够实现对奶牛的正常行走、跛行和站立3种行为的识别分类。



1.牛嘴 Cow mouth;2.额头 Forehead;3.颈部 Neck;4.脊椎 Spine;5.尾椎 Tailbone;6.前左腿根 Front left leg root;7.前左膝 Front left knee;8.前左蹄 Front left hoof;9.前右腿根 Front right leg root;10.前右膝 Front right knee;11.前右蹄 Front right hoof;12.后左腿根 Rear left leg root;13.后左膝 Rear left knee;14.后左蹄 Rear left hoof;15.后右腿根 Rear right leg root;16.后右膝 Rear right knee;17.后右蹄 Rear right hoof.

图1 奶牛牛体位置标定和关键点标定

Fig.1 Dairy cow body position calibration and key point calibration

2)奶牛姿态估计网络。研究采用自顶向下的热力图方法来实现奶牛的姿态估计,基于Faster RCNN网络和ResNet101网络,构建奶牛姿态估计网络,其技术路线如图3所示。首先,利用Faster RCNN网络完成对奶牛目标位置的检测;其次,以ResNet101模型为主干网络,获取图像中相应位置奶牛的特征信息,基于热力图法实现关键点坐标的预测,最终可得到大小为 $K \times H' \times W'$ 的热力图,其中 $K$ 为关键点数量, $H'$ 和 $W'$ 分别为原始图像长和宽的1/4;最后在每个热力图上找到最大值点,即为关键点的位置,将每个关键点进行连接,得到完整的奶牛姿态,如图4所示。

3)关节坐标数据提取。奶牛无论在行走还是站立过程中后腿膝关节角度仅有细微的变化,而前腿膝关节角度这一特征有着明显的变化,因此,本研究采用检测前腿膝关节角度的变化情况来识别奶牛行为。奶牛正常行走时,前、后腿的关节有规律地变化,体现在前腿膝关节角度在 $120^\circ \sim 180^\circ$ 较为规律的变化;奶牛跛行时,蹄部疼痛导致步态不平衡,会出现2个前腿的膝盖关节都不能正常弯曲活动的情况,体现在前腿膝关节角度在 $140^\circ \sim 180^\circ$ 无规律的变化,在膝关节角度不能正常弯曲和行走时所站立的时间

更长;奶牛站立时,4条腿伸直以支撑身体,前腿膝关节角度在 $170^{\circ}\sim 180^{\circ}$ 变化。

前腿膝关节角度是检测到的腿根、膝关节和牛蹄的3个坐标信息,代表了奶牛的真实运动情况,通过记录腿根、膝关节和牛蹄的3个坐标位置的变化,计算得到前腿膝关节角度的变化曲线。采用奶牛姿

态估计网络检测牛蹄的位置坐标,符合构建前腿膝关节角度特征向量的基本要求;获取腿根、膝关节和牛蹄的位置坐标得到前腿膝关节角度的变化曲线;使用分类算法进而判断出奶牛是正常行走、跛行还是站立状态。使用前腿膝关节角度作为特征向量符合本研究的实际研究情况。

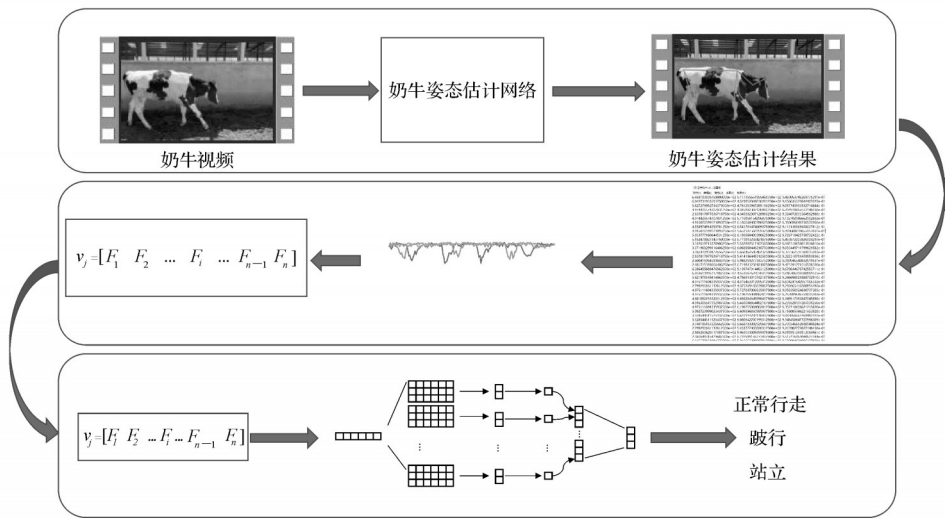


图2 基于姿态估计和膝关节角度特征向量的奶牛行为识别技术路线

Fig.2 A technical route to cow behavior recognition based on posture estimation and knee joint angle feature vectors



图3 奶牛姿态估计网络技术路线

Fig.3 Dairy cow posture estimation network technology route

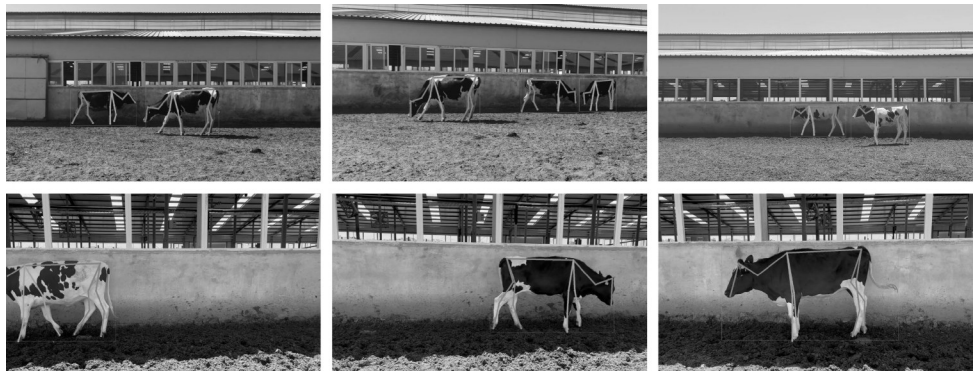


图4 奶牛姿态估计结果

Fig.4 Results of dairy cow stance estimates

设奶牛的前腿腿根、前腿膝盖、前蹄的位置分别为 $A(x_1,y_1)$ 、 $B(x_2,y_2)$ 、 $C(x_2,y_2)$ ,则前腿膝关节角度计算公式为:

$$L_1 = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{1}$$

$$L_2 = \sqrt{(x_2 - x_3)^2 + (y_2 - y_3)^2} \tag{2}$$

$$L_3 = \sqrt{(x_1 - x_3)^2 + (y_1 - y_3)^2} \tag{3}$$

$$\alpha = \arccos\left(\frac{L_1^2 + L_2^2 - L_3^2}{2L_1 \cdot L_2}\right) \cdot \frac{180}{\pi} \tag{4}$$

式(1)~(4)中, $L_1$ 、 $L_2$ 和 $L_3$ 分别表示腿根到膝盖的距离、膝盖到牛蹄的距离和腿根到牛蹄的距离, $\alpha$

表示前腿膝关节角度。

计算所有视频序列帧的前腿膝关节角度,并构造出对应的特征向量 $v_j$ ,其公式如下:

$$v_j = [F_1 \ F_2 \ \dots \ F_i \ \dots \ F_{n-1} \ F_n] \quad (5)$$

式(5)中,视频中 $F_i$ 表示第 $i$ 帧的前腿膝关节角度, $n$ 为视频的帧数,将所有视频构建得到的特征向量记为 $v_j$ ,特征向量 $v_j$ 中 $j$ 代表视频的个数。 $v_j$ 用作分类器的输入。

4)基于1-D Convolution的跛行识别。传统的神经网络由于各层节点不连通,在没有构造时序特征的情况下不能通过动态的记忆过程学习到时序特性,而卷积神经网络的特点主要在于它能够捕捉局部特征,并将这些局部特征逐渐融合成整体特征。使用一维卷积神经网络可对时间序列数据进行分类,将一维数据输入到一维卷积层,在一维卷积层中通过滑动卷积滤波器来学习特征。相较于LSTM等模型,一维卷积神经网络的优点是在处理长时间序列时能够并行且速度更快。

本研究构建了1-D Convolution分类模型用于检测奶牛是否跛行,1-D Convolution模型结构如图5所

示,将奶牛视频得到的特征向量 $v_j$ 输入到1-D Convolution分类模型中,输入的数据通过一维卷积,再通过ReLU激活函数、正则化后得到每个区域大小的特征图,特征图通过全局平均池化的结果输入到全连接层,再将结果通过SoftMax层后,得到正常奶牛、跛行和站立所对应的概率值,并显示出测试结果。

为实现基于1-D Convolution分类模型的奶牛跛行的识别和分类,采取以下步骤:

第一步,整理数据,将原始的123段视频样本通过水平镜像的方式扩充到246段视频样本,使用这246段视频样本进行训练,每个视频持续100帧到200帧左右以保证提取的前腿膝关节角度特征向量的维度一致性。246个视频样本按7:3的比例分为训练集和测试集。训练集中有174个样本(跛行70个样本、站立72个样本和正常行走32个样本),测试集中有72个样本(跛行28个样本、站立30个样本和正常行走14个样本)。

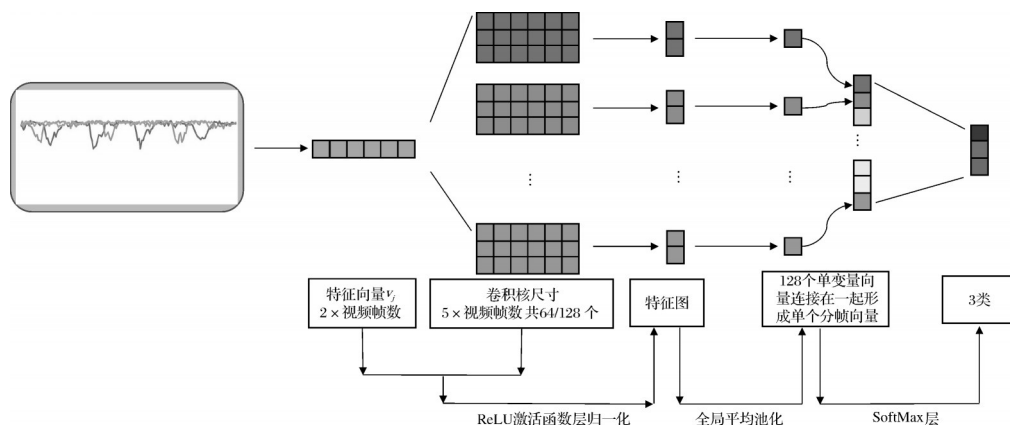


图5 1-D Convolution模型结构

Fig.5 1-D Convolution model structure

第二步,提取246个视频样本的膝关节角度特征向量并建立标签。奶牛姿态估计网络用于检测视频中每帧奶牛的前腿腿根、膝关节、牛蹄3个跟随特性特征参数,并计算每帧中奶牛前腿膝关节角度;结合每帧膝关节角度变化情况,构造出与视频帧数长度相对应的膝关节角度特征向量;最后,为训练视频对应的膝关节角度特征向量建立标签,奶牛的跛行、正常行走和站立视频膝关节角度特征向量对应的标签分别为0、1和2。

第三步,标准化246个视频样本的246个前腿膝关节角度特征向量。

第四步,设计并训练1-D Convolution分类模型。在批次大小为35、迭代次数为1000的行为分类实验中,本研究设计了一个双层的一维卷积分类模型。该模型使用了大小为 $5 \times 5$ 的卷积核,分别包含了64和128个卷积核,最后基于测试集评估模型的识别效果。

#### 1.4 评价指标

本研究评价奶牛姿态估计网络的性能采用PCK(probability of correct keypoint)关键点检测评价指标。当某个被检测的关键点与真实的关键点坐标之间的距离小于某个长度的某个比例时,则认为该

检测是有效的。PCK是以关键点为单位计算的指标,其计算公式为:

$$PCK_s^p = \frac{1}{|\tau|} \sum_{\tau} \delta(\|x_p^f - y_p^f\|_2 < \sigma) \quad (6)$$

式(6)中, $\sigma$ 表示关键点是否和ground truth匹配的阈值。 $\tau$ 表示测试集合, $x_p^f$ 表示图像帧中预测位置, $y_p^f$ 表示图像帧中ground truth。PCK度量的是包围盒宽度的归一化公差下的检测精度,例如,包围盒大小为100像素的PCK@0.2,是指检测误差小于20像素的检测精度。

跛行识别模型中的评价指标包括准确率(accuracy,  $A$ )、精确率(precision,  $P$ )、召回率(recall,  $R$ )、假阳性率(false positive rate,  $F$ )。 $A$ 、 $R$ 和 $F$ 的计算公式<sup>[19]</sup>为:

$$A = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad (7)$$

$$P = \frac{T_p}{T_p + F_p} \quad (8)$$

$$R = \frac{T_p}{T_p + F_n} \quad (9)$$

$$F = \frac{F_p}{F_p + T_n} \quad (10)$$

式(7)~(10)中, $T_p$ 和 $T_n$ 分别表示当前行为、其他行为的预测和实际结果相同的视频数量, $F_p$ 和 $F_n$ 分别表示当前行为、其他行为的预测和实际结果不相同的视频数量。准确率为测试集中正确预测的视频数量与总视频数量的比值,精确率为测试集中正确预测某类行为的视频数量与预测为该类行为的视频数量的比值,召回率为测试集中正确预测某类行为的视频数量与实际该类行为的视频数量的比值,假阳性率为测试集中错误预测某类行为的视频数量与其他行为的视频数量的比值。

### 1.5 模型训练

本研究的实验训练和测试均在同一台计算机上,操作系统为Windows 10,处理器为Inter(R)Core (TM) i7-9700k CPU@3.2GHz, GPU 为 GeForce RTX 2080,运行内存为32 GB。模型搭建与训练验证的编程语言为Python3.7,深度学习框架采用PyTorch, PyCharm 开发工具,计算框架使用CUDA 10.2版本。训练的Batchsize设置为16,选用Adam优化器,学习率从 $10^{-4}$ 逐渐下降到 $10^{-5}$ ,训练次数设置为210次迭代,在第170和200个Epoch时降低学习率。

按照7:3的比例将实验数据集划分为训练集、验证集。牛体检测框的高宽比设为1:1,并从原图像

中剪裁出检测框,最后本研究中输入图像模型的大小为256像素×256像素。热力图大小设置为64像素×64像素。奶牛姿态估计网络输入的每张图像都经过随机翻转、旋转( $\pm 90^\circ$ )和缩放( $\pm 30\%$ )的数据增强操作,以验证集中关键点的平均PCK值作为最终的评估指标。

## 2 结果与分析

### 2.1 奶牛姿态估计网络性能评估

1)评价指标及结果。为评价牛体关键点的检测性能,本研究使用ResNet101网络模型作为奶牛关键点的特征提取网络,并使用了在horse10数据集上的预训练权重,ResNet101在验证集上的准确率变化曲线如图6所示。网络使用预训练权重在第20个Epoch达到收敛状态,而未使用预训练权重在第80个Epoch达到收敛状态,未使用预训练权重在多个Epoch训练下也能达到收敛,但使用预训练权重能够将网络模型引导朝着更优的方向收敛,从而有效提升模型的识别精度,减少训练所需时间。

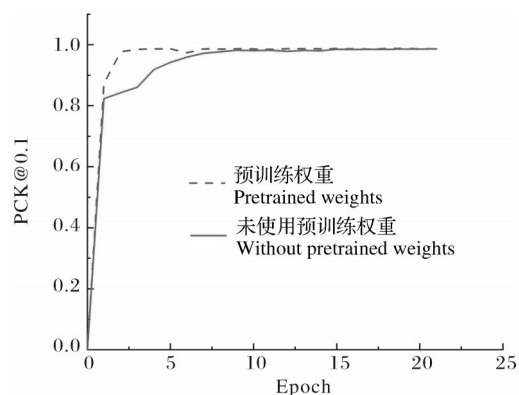


图6 ResNet101的准确率变化曲线

Fig.6 Accuracy curve of ResNet101

数据集中奶牛都处于站立或行走状态,除腿根部分和行走过程中部分帧的牛腿相互交替形成自身遮挡外,绝大部分关键点都能被准确地检测。不同输入尺寸下ResNet101特征提取网络在验证集上的检测结果如表1所示,在实际应用中采集奶牛行为数据大多来自于监控摄像设备,采集的数据受分辨率低的限制,导致网络的特征提取难度增大,由表1可知当输入尺寸为192像素×192像素时,平均PCK@0.1下降0.0560,说明ResNet101网络对奶牛图像的输入尺寸更加敏感,当输入尺寸为192像素×192像素时,热力图大小仅为48像素×48像素,导致网络预测真实关键点的尺寸范围更小,说明在算力的承受范围内,使用较大的输入尺寸有较好的可适性。

不同输入尺寸下 PCK 值取 0.1 时各个关键点的检测结果如表 2 所示。由表 2 可知,在特征明显的关键点奶牛姿态估计网络表现较好,如牛嘴、额头、尾椎、膝和牛蹄,PCK@0.1 均超过了 0.94,而对于奶牛的颈部、脊椎和腿根这 3 类关键点识别准确率相对较低,PCK@0.1 在 0.85 左右,主要是因为牛体黑白色不均匀,导致在关键点的定位过程中产生较大误差。

奶牛姿态估计结果以及图片所预测的热力图如图 7 所示。由图 7 可知,本研究采用的方法能够进一

表 1 ResNet101 特征提取网络在验证集上的检测结果

Table 1 Detection results of ResNet101 feature extraction network on the validation set

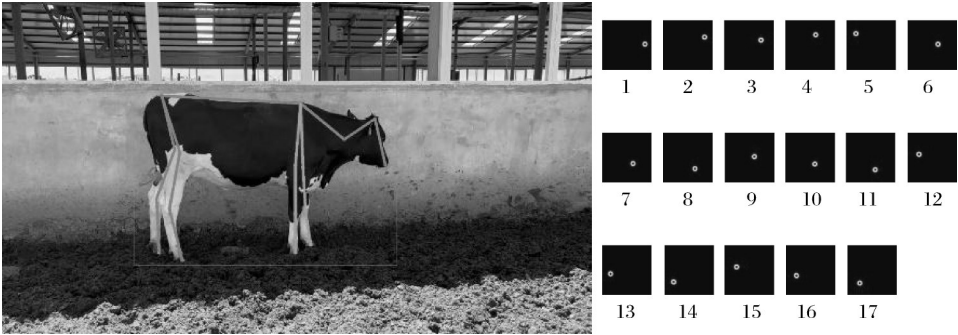
网络模型 Network model	输入尺寸 Input size	FLOPs/G	PCK@0.2	PCK@0.1
ResNet101	192×192	5.76	0.979 4	0.869 0
ResNet101	256×256	10.25	0.982 4	0.925 0

步锁定该关键点所在区域的预测范围,提升关键点的预测精度。

表 2 ResNet101 网络 PCK@0.1 时各个关键点的检测精确度

Table 2 Detection accuracy of each key point of ResNet101 network at PCK@0.1

输入尺寸 Input size	嘴 Mouth	额头 Forehead	颈部 Neck	脊椎 Spine	尾椎 Tailbone	腿根 Leg root	膝 Knee	蹄 Hoof
192×192	0.941 4	0.921 2	0.774 6	0.768 3	0.909 5	0.794 1	0.881 0	0.939 7
256×256	0.969 1	0.940 5	0.844 4	0.834 9	0.946 0	0.876 2	0.955 9	0.965 5



图中数字代表奶牛关键点所对应的编号 The numbers in the figure represent the numbers corresponding to the key points of the cow.

图 7 奶牛姿态估计图及热力图

Fig.7 Dairy cow stance estimation diagram and heat map

2)不同特征提取网络。本研究综合比较 HRNet 网络模型和 ResNet 网络模型在奶牛数据集上的表现,以寻求适合本研究的特征提取网络模型。在特征提取网络中分别配置了 HRNet32、HRNet48、ResNet50 网络模型,输入尺寸为 256 像素×256 像素,使用在 ImageNet 上的预训练权重。表 3 展示了不同特征提取网络对奶牛姿态估计的识别结果。

由表 3 可知,对于 HRNet32、HRNet48、ResNet50、ResNet101 网络模型,HRNet48 比 HRNet32 网络模型在 PCK@0.1 值要略高 0.005 2,但计算量和参数量要高出 1 倍以上,说明 HRNet 网络在奶牛数据集上均有较好的识别结果,与网络复杂程度相关不大。ResNet101 比 ResNet50 网络模型在 PCK@0.1 要高 0.0297;而 ResNet101 与 HRNet48 网络的 PCK@0.1 相差不大,且 ResNet101 网络有着更低的计算量和参数量,说明采用 ResNet101 网络更适用于本研究中的

表 3 不同特征提取网络对奶牛姿态估计的识别结果

Table 3 Identification results of different feature extraction networks for cow posture estimation

网络模型 Network Model	FLOPs/G	Params/M	PCK@0.2	PCK@0.1
HRNet32	10.25	28.54	0.982 1	0.914 2
HRNet48	21.00	63.59	0.982 3	0.919 4
HRNet48*	21.00	63.59	0.962 4	0.853 7
ResNet50	5.38	25.56	0.978 2	0.895 3
ResNet101	10.25	44.55	<b>0.982 4</b>	<b>0.925 0</b>
ResNet101*	10.25	44.55	0.979 9	0.906 1

注:\* 代表未使用预训练权重。Note: \* means no pre-training weights are used.

奶牛关键点进行特征提取。

使用 HRNet48 和 ResNet101 网络奶牛跛行和正常行走状态下进行姿态估计,结果如图 8 所示,本研究所使用的 ResNet101 网络模型相较于 HRNet48 网络表现更好,当奶牛跛行时,该网络能够准确地识别

出由于自身遮挡的奶牛前腿关键点;当奶牛正常行走时,ResNet101 比 HRNet48 网络对关键点的定位更精准。

3)奶牛目标检测对姿态估计的影响。在奶牛站立、行走 2 种状态下,本研究使用未添加 Faster RCNN 的姿态估计网络和本研究方法对其进行姿态

估计,结果如图 9 所示,可以明显看出相较于对整个图像进行识别,在本研究中加入目标检测算法对牛体进行检测缩小了奶牛关键点的定位范围,进一步提高了关键点的准确率。使用目标检测算法可以提高本研究中奶牛姿态估计模型的性能,证明了本研究方法的有效性。

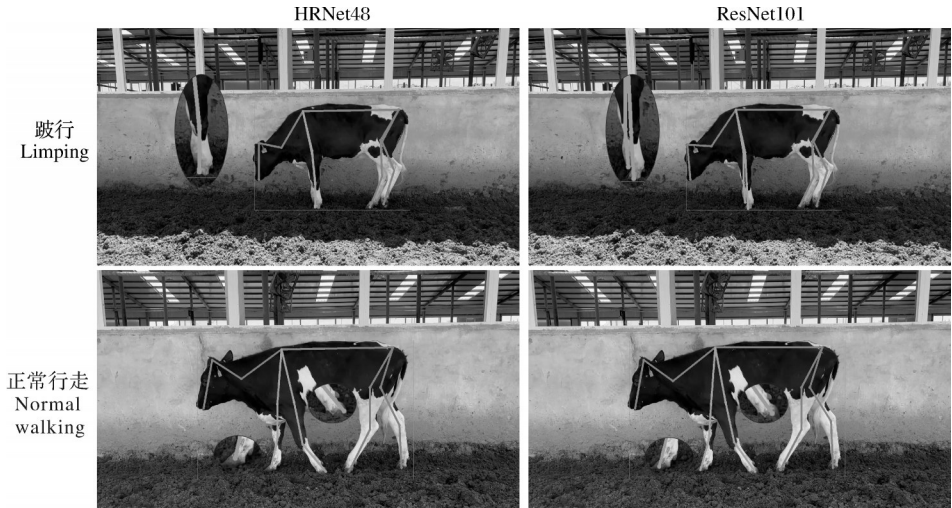


图 8 HRNet48 和 ResNet101 网络在奶牛跛行、正常行走状态下的姿态估计结果

Fig.8 Pose estimation results of HRNet48 and ResNet101 networks for cows in lame , normal walking condition

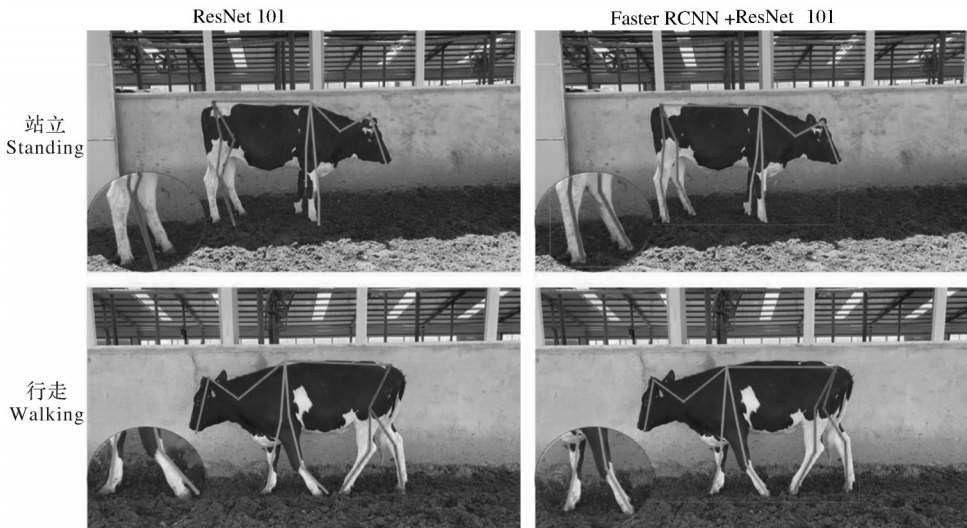


图 9 未使用 Faster RCNN 的姿态估计网络和本文方法对奶牛站立、行走的姿态估计结果

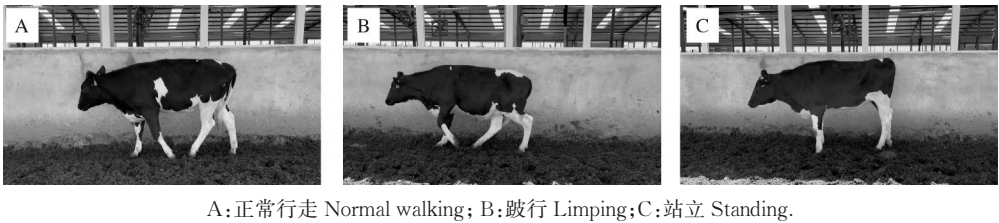
Fig.9 Posture estimation network without Faster RCNN and the results of this study's method for estimating the posture of cows standing and walking

2.2 奶牛跛行分类识别实验

奶牛行走过程中,正常奶牛后蹄着地点应接近或超过同侧前蹄着地点,如图 10A 所示;奶牛跛行程度越大,后蹄越难以达到同侧前蹄着地点,如图 10B 所示;奶牛在行走时会采用站立的方式进行短暂休息,如图 10C 所示。此外,通过观察奶牛前腿膝关节角度变化曲线,可以较为准确地识别出跛行的奶牛,

有利于跛足的诊断和准确治疗。

1)不同分类器准确率比较。为了比较不同分类算法的有效性,本研究基于 246 个测试视频片段(98 个跛行、102 个站立和 46 个正常行走视频)进行前腿关键点的检测,并构建了 246 个前腿膝关节角度特征向量。实验中设置学习率为 0.001,训练轮数为 1 000 轮。在 LSTM、Bi-LSTM、GRU、1-D Convolution 分



A:正常行走 Normal walking; B:跛行 Limping;C:站立 Standing.

图 10 3 种行为状态下的奶牛侧视图

Fig.10 Side view of a cow in 3 states of behavior

类网络完成训练后,结合前腿膝关节角度特征向量,分类网络得出各类行为相对应的置信度值,并根据该值完成行为分类。

实验结果表明,LSTM、Bi-LSTM、GRU、1-D Convolution 均可完成奶牛行为识别的任务,其识别准确率分别为 91.67%、94.44%、86.11%、97.22%。相较于其他模型,1-D Convolution 模型实现了最高的分类精度,分别高出 LSTM、Bi-LSTM 和 GRU 5.55、2.78 和 11.11 百分点。相较于 LSTM 网络,Bi-LSTM 网络的获得了更高的准确率,但在相同的输入大小和隐藏元素下,Bi-LSTM 的网络结构比 LSTM 更复杂,计算参数更多。1-D Convolution 模型简单,采用 1-D Convolution 用于检测奶牛跛行,在保证检测精度的同时,可以有效降低对硬件的要求,在一定程度上降低实际应用的成本。因此,本文方法对奶牛的检测跛行行为具有更高的实用性。

2)不同分类器的精确率、召回率和假阳性率。表 4 显示了 4 种分类模型在跛行、正常行走和站立这 3 类行为的精准率、召回率和假阳性率的测试结果。

由表 4 可知,在相同实验数据下,奶牛站立时相

较于奶牛行走过程有较大的区别,4 种模型在奶牛站立分类时都有着很高的精确率和召回率以及非常低的假阳性率;奶牛跛行时,1-D Convolution 模型的精确率分别比 LSTM、Bi-LSTM、GRU 网络高 0.107 1、0.038 5、0.172 4,召回率分别比 LSTM、Bi-LSTM、GRU 网络高 0.035 7、0.035 7、0.071 5;而在正常行走情况下 1-D Convolution 模型的精确率分别比 LSTM、Bi-LSTM、GRU 网络高 0.051 5、0.062 5、0.187 5,召回率分别比 LSTM、Bi-LSTM、GRU 网络高 0.176 5、0.071 4、0.214 3。使用 1-D Convolution 模型相比于 RNN 网络有着更好的结果,这表明前者更合适本研究提出的相对分类器的跛行识别。轻微跛足的奶牛和正常的奶牛在行走时的差别不大,即使是精确率最高的 1-D Convolution 网络的分类模型也不能够完全区分,导致轻微跛足的奶牛被分错类别。

结果表明,本研究所构建的 1-D Convolution 分类方法对奶牛跛行的识别具有较高的精确率、召回率和较低的假阳性率,该算法能很好地识别奶牛跛行、正常行走和站立这 3 类行为。

表 4 不同分类器精确率、召回率和假阳性率

Table 4 Precision(P), recall(R) and false positive rate(F) values of different classifiers

模型 Model	跛行 Limping			正常行走 Normal walking			站立 Standing		
	P	R	F	P	R	F	P	R	F
LSTM	0.892 9	0.892 9	0.063 8	0.823 5	0.823 5	0.051 7	1	1	0
Bi-LSTM	0.961 5	0.892 9	0.022 7	0.812 5	0.928 6	0.051 7	1	1	0
GRU	0.827 6	0.857 1	0.113 6	0.687 5	0.785 7	0.086 2	1	0.9	0
1-D Convolution	1	0.928 6	0	0.875 0	1	0.034 5	1	1	0

3)基于前腿膝关节角度特征向量方法对奶牛跛行识别的分析。随机选取 3 类不同行为的奶牛视频并检测其前腿膝关节角度变化,如图 11 所示。本研究采集数据时没有按照奶牛的初始运动周期(以 4 个牛蹄全部抬起并落下 1 次为 1 个运动周期)作为起点,因此在每个特征向量中的起点位置不一样。由图 11 可知,在奶牛站立时前腿膝关节角度不会发生较大的变化,而前腿膝关节角度在 170°~180°波动变

化,主要是因为奶牛站立时会发生轻微的动作,如奶牛舔毛时会影响姿态估计结果;图 11 中第 13~51 帧表示了正常奶牛的 1 个运动周期,在奶牛正常行走时前腿膝关节角度在 120°~180°较为规律地变化;第 6~61 帧表示了奶牛跛行的 1 个运动周期,跛行的奶牛前腿膝关节角度在 140°~180°变化的规律性较低,体现在膝关节角度不能正常弯曲导致膝关节角度最小弯曲只能达到 140°左右;图 11 中第 8~115 帧是跛

足奶牛的2个运动周期,第17~86帧是正常奶牛的2个运动周期,这是因为跛足奶牛相较于正常奶牛由于蹄部疼痛导致在行走过程中所花费的站立时间更长。

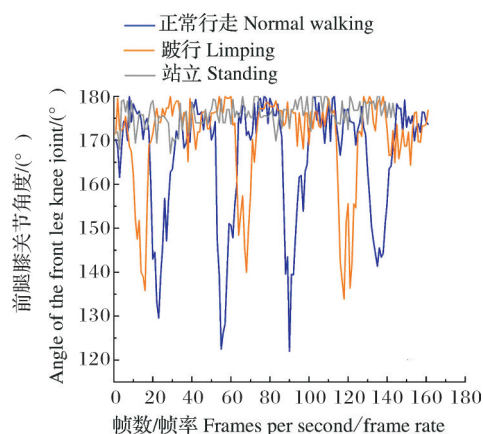


图11 奶牛3类行为的前腿膝关节角度运动变化曲线

Fig.11 Change curve of foreleg knee angle movement for 3 types of behavior in dairy cows

综上所述,可以看出跛牛的前腿膝关节角度变化幅度较小,相较于正常奶牛,跛足奶牛的变化规律性较低。

### 3 讨论

本研究针对奶牛跛行行为识别任务,提出了一种基于姿态估计和膝关节角度特征向量的奶牛跛行识别方法,并通过迁移学习实现自然养殖环境下奶牛正常行走、跛行和站立行为的有效识别,能够实现视频数据下的奶牛行为分类,提升检测精度的同时降低奶牛养殖的人工成本。在奶牛姿态估计网络构建中,有自顶向下和自底向上2种方法<sup>[20-21]</sup>,自顶向下的姿态估计方法使用图像的背景和特征点来估计目标的姿态;自底向上的姿态估计方法则直接检测不同的关键点,并通过聚类等方式连接成骨架。与自底向上的方法相比,自顶向下的姿态估计方法拥有简单的后处理方法且能提供更准确的结果。经典的关键点定位方法主要分为回归关键点方法和热力图法,后者比前者的扩展性和准确率都更高<sup>[22]</sup>。因此,本研究采用自顶向下的热力图方法来实现奶牛的姿态估计。迁移学习的目的是将源域学习到的知识应用到目标域的学习过程中,使用该方法将提高网络模型的泛化性和收敛速度<sup>[23-24]</sup>。奶牛与马体型大小相似,且奶牛与马的关键点具有高度相关性,因此,本研究采用基于实例的迁移方法<sup>[25-26]</sup>对模型进行训练。

为了实现多目标姿态估计任务,本研究采用Faster RCNN来得到目标检测框,在此基础上使用ResNet101进行姿态估计,不仅减少了背景因素的干扰,有利于提升骨架信息的精度;同时以目标框为基础,个体姿态信息已经得到分组,免除了多目标时姿态信息复杂的后处理过程;ResNet101凭借深度网络,不仅可以实现深层次信息理解与提取,而且残差连接的方式可以使得网络在学习过程中自动将不利于训练的隐层权重降低,有利于网络的收敛。对于得到的骨架信息,1-D Convolution拥有较少的参数量及较高的泛化性,基于此对建立的膝关节角度特征向量集进行分类得到较高精度的分类结果,实现自然养殖环境下奶牛正常行走、跛行和站立行为的有效识别。研究的主要结论如下:

1)本研究引入Faster RCNN目标检测网络模型对牛体进行识别,缩小奶牛关键点检测范围,奶牛关键点检测的精度明显提高。

2)建立基于ResNet101的奶牛姿态估计模型,从Faster RCNN模型的检测框中提取奶牛关键点特征,其PCK@0.1值可以达到0.925 0,与HRNet32、HRNet48、ResNet50网络模型相比,能更准确地反映出奶牛关键点特征。

3)使用迁移学习将在大数据集上训练完成的模型迁移至奶牛侧视数据集。结果表明,在参数量和计算量一致的情况下,使用迁移学习后模型的PCK@0.1均有较大的提升。

4)通过奶牛的前腿关键点坐标构建前腿膝关节角度特征向量集,并利用1-D Convolution分类模型实现奶牛的跛行、站立和正常行走的检测,其分类准确率为97.22%。相较于LSTM、Bi-LSTM、GRU模型,准确率分别提高5.55、2.78和11.1百分点。本研究设计的膝盖关节角度特征向量不受限于拍摄距离远近的影响,有较好的实用性,可为奶牛跛行自动检测提供参考。

### 参考文献 References

- [1] 邵军,林为民,孙新文,等. 石河子地区奶牛肢蹄病类型和发病情况调查分析[J]. 中国奶牛, 2017(1): 30-33. SHAO J, LIN W M, SUN X W, et al. Investigation and analysis on the types and incidence of limb and hoof disease in dairy cows in Shihezi area [J]. China dairy cattle, 2017(1): 30-33 (in Chinese).
- [2] BOOTH C J, WARNICK L D, GRÖHN Y T, et al. Effect of lameness on culling in dairy cows [J]. Journal of dairy science,

- 2004, 87(12):4115-4122.
- [3] PLUK A, BAHR C, POURSAHERI A, et al. Automatic measurement of touch and release angles of the fetlock joint for lameness detection in dairy cattle using vision techniques [J]. Journal of dairy science, 2012, 95(4):1738-1748.
- [4] 吴倩. 基于视频分析的奶牛关键部位提取及跛行检测研究 [D]. 杨凌:西北农林科技大学, 2019. WU Q. Research on cow key part extraction and limping detection based on video analysis [D]. Yangling: Northwest A & F University, 2019 (in Chinese with English abstract).
- [5] POURSAHERI A, BAHR C, PLUK A, et al. Real-time automatic lameness detection based on back posture extraction in dairy cattle: shape analysis of cow with image processing techniques [J]. Computers and electronics in agriculture, 2010, 74(1):110-119.
- [6] 顾静秋, 王志海, 高荣华, 等. 基于融合图像与运动量的奶牛行为识别方法 [J]. 农业机械学报, 2017, 48(6):145-151. GU J Q, WANG Z H, GAO R H, et al. Recognition method of cow behavior based on combination of image and activities [J]. Transactions of the CSAM, 2017, 48(6):145-151 (in Chinese with English abstract).
- [7] 康熙, 张旭东, 刘刚, 等. 基于机器视觉的跛行奶牛牛蹄定位方法 [J]. 农业机械学报, 2019, 50(S1):276-282. KANG X, ZHANG X D, LIU G, et al. Hoof location method of lame dairy cows based on machine vision [J]. Transactions of the CSAM, 2019, 50(S1):276-282 (in Chinese with English abstract).
- [8] ZHAO K, BEWLEY J M, HE D, et al. Automatic lameness detection in dairy cattle based on leg swing analysis with an image processing technique [J]. Computers and electronics in agriculture, 2018, 148:226-236.
- [9] 於少文, 孔繁涛, 张建华, 等. 可穿戴设备技术在奶牛养殖中的应用及发展趋势 [J]. 中国农业科技导报, 2016, 18(5):102-110. YU S W, KONG F T, ZHANG J H, et al. Application and development trend of wearable devices technology in dairy farming [J]. Journal of agricultural science and technology, 2016, 18(5):102-110 (in Chinese with English abstract).
- [10] LI X Y, CAI C, ZHANG R F, et al. Deep cascaded convolutional models for cattle pose estimation [J/OL]. Computers and electronics in agriculture, 2019, 164:104885 [2022-11-17]. <https://doi.org/10.1016/j.compag.2019.104885>.
- [11] LIU H, REIBMAN A R, BOERMAN J P. A cow structural model for video analytics of cow health [DB/OL]. arXiv, 2020:2003.05903 [2022-11-17]. <https://arxiv.org/abs/2003.05903>.
- [12] TSAI Y C, HSU J T, DING S T, et al. An imaging system based on deep learning for monitoring the feeding behavior of dairy cows [J/OL]. American society of agricultural and biological engineers, 2019:1469 [2022-11-17]. <https://www.semanticscholar.org/paper/An-Imaging-System-Based-on-Deep-Learning-for-the-of-Kuan-Tsai/50bbc50d3ba3bac2399018223>.
- c5dd19e263ff8e0. DOI:10.13031/aim.201901469.
- [13] WU D H, WU Q, YIN X Q, et al. Lameness detection of dairy cows based on the YOLOv3 deep learning algorithm and a relative step size characteristic vector [J]. Biosystems engineering, 2020, 189:150-163.
- [14] WU D H, YIN X Q, JIANG B, et al. Detection of the respiratory rate of standing cows by combining the Deeplab V<sup>3+</sup> semantic segmentation model with the phase-based video magnification algorithm [J]. Biosystems engineering, 2020, 192:72-89.
- [15] 刘忠超, 何东健. 基于卷积神经网络的奶牛发情行为识别方法 [J]. 农业机械学报, 2019, 50(7):186-193. LIU Z C, HE D J. Recognition method of cow estrus behavior based on convolutional neural network [J]. Transactions of the CSAM, 2019, 50(7):186-193 (in Chinese with English abstract).
- [16] 王少华, 何东健, 刘冬. 基于机器视觉的奶牛发情行为自动识别方法 [J]. 农业机械学报, 2020, 51(4):241-249. WANG S H, HE D J, LIU D. Automatic recognition method of dairy cow estrus behavior based on machine vision [J]. Transactions of the CSAM, 2020, 51(4):241-249 (in Chinese with English abstract).
- [17] 付涛. 基于卷积热点图回归的牛体姿态估计研究 [D]. 杨凌:西北农林科技大学, 2018. FU T. Research on bovine body pose estimation based on convolution heatmap regression [D]. Yangling: Northwest A & F University, 2018 (in Chinese with English abstract).
- [18] 宋怀波, 李振宇, 吕帅朝, 等. 基于部分亲和场的行走奶牛骨架提取模型 [J]. 农业机械学报, 2020, 51(8):203-213. SONG H B, LI Z Y, LÜ S C, et al. Skeleton extraction model of walking dairy cows based on partial affinity field [J]. Transactions of the CSAM, 2020, 51(8):203-213 (in Chinese with English abstract).
- [19] 周志华. 机器学习 [M]. 北京:清华大学出版社, 2016. ZHOU Z H. Machine learning [M]. Beijing: Tsinghua University Press, 2016 (in Chinese).
- [20] 卢健, 杨腾飞, 赵博, 等. 基于深度学习的人体姿态估计方法综述 [J]. 激光与光电子学进展, 2021, 58(24):69-88. LU J, YANG T F, ZHAO B, et al. Review of deep learning-based human pose estimation [J]. Laser & optoelectronics progress, 2021, 58(24):69-88 (in Chinese with English abstract).
- [21] NIBALI A, HE Z, MORGAN S, et al. Numerical coordinate regression with convolutional neural networks [DB/OL]. arXiv, 2018:1801.07372 [2022-11-17]. <https://arxiv.org/abs/1801.07372>.
- [22] 张飞宇, 王美丽, 王正超. 引入Transformer和尺度融合的动物骨骼关键点检测模型构建 [J]. 农业工程学报, 2021, 37(23):179-185. ZHANG F Y, WANG M L, WANG Z C. Construction of the animal skeletons keypoint detection model based on transformer and scale fusion [J]. Transactions of the CSAM, 2021, 37(23):179-185 (in Chinese with English ab-

- stract).
- [23] 龙满生, 欧阳春娟, 刘欢, 等. 基于卷积神经网络与迁移学习的油茶病害图像识别[J]. 农业工程学报, 2018, 34(18): 194-201. LONG M S, OUYANG C J, LIU H, et al. Image recognition of *Camellia oleifera* diseases based on convolutional neural network & transfer learning[J]. Transactions of the CSAE, 2018, 34(18): 194-201 (in Chinese with English abstract).
- [24] PAN S J, YANG Q A. A survey on transfer learning[J]. IEEE transactions on knowledge and data engineering, 2010, 22(10): 1345-1359.
- [25] SUN K, XIAO B, LIU D, et al. Deep high-resolution representation learning for human pose estimation [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 15-20, 2019, Long Beach, CA, USA. Long Beach: IEEE, 2020: 5686-5696.
- [26] MATHIS A, BIASI T, SCHNEIDER S, et al. Pretraining boosts out-of-domain robustness for pose estimation [C]//2021 IEEE Winter Conference on Applications of Computer Vision (WACV). January 3-8, 2021, Waikoloa, HI, USA. Waikoloa: IEEE, 2021: 1858-1867.

## A method for cow lameness recognition based on posture estimation and keypoints feature vector

DU Yuemeng, SHI Hui, GAO Feng, DENG Hongtao

*School of Mechanical and Electrical Engineering, Shihezi University, Shihezi 832003, China*

**Abstract** To solve the current problems of low efficiency and low accuracy of automatic detection of cow lameness in farms, a cow lameness recognition method based on posture estimation and knee angle eigenvectors was designed. Given the random behavior of dairy cows, a cow posture estimation dataset was produced by combining the imaging characteristics of cows under different conditions such as near and far field of view scales and observation angles. The Faster RCNN convolutional neural network model was introduced into the key point detection of dairy cows to improve the reliability of lameness recognition. Taking ResNet101 network as feature extraction network, the cow posture estimation network was constructed, and the hyperparameter fine-tuning training method was used to train the migration of the network model. Based on the information of cow's posture and key point coordinate in the video, the angle feature of the cow's knee joint when walking were calculated, and the 1-D Convolution classification model was used to realize the cow's lameness recognition. The experimental results showed that the PCK@0.1 value of the cow posture estimation network based on ResNet101 network model can reach 0.925 0. Compared with the LSTM, Bi-LSTM, and GRU models, the accuracy of cow behavior classification and recognition of 1-D Convolution model was 97.22%, which was 5.55, 2.78 and 11.11 percentage points higher, respectively. The above results show that the proposed method has a better detection effect on cow lameness in natural environment, which can provide technical reference for intelligent breeding and management of dairy cows.

**Keywords** dairy cow; deep learning; pose estimation; temporal features; behavior recognition

(责任编辑: 边书京)