

基于3D CNN-BiLSTM-ATFA网络和步态特征的奶牛个体识别方法

司永胜¹ 宁泽普¹ 王克俭¹ 马亚宾² 袁明²

(1. 河北农业大学信息科学与技术学院, 保定 071001; 2. 河北省畜牧良种工作站, 石家庄 050061)

摘要: 针对基于花纹的奶牛个体识别中纯色或花纹较少的奶牛识别准确率较低的问题, 本文提出一种基于步态特征的奶牛个体识别方法。首先, 将 DeepLabv3 + 语义分割算法的主干网络替换为 MobileNetv2 网络, 并引入基于通道和空间的 CBAM 注意力机制, 利用改进后模型分割出奶牛的剪影图。然后, 将三维卷积神经网络(3D CNN)和双向长短期记忆网络(BiLSTM)构建为 3D CNN-BiLSTM 网络, 并进一步集成自适应时间特征聚合模块(ATFA)生成 3D CNN-BiLSTM-ATFA 奶牛个体识别模型。最后, 在 30 头奶牛的共 1 242 条视频数据集上进行了奶牛个体识别实验。结果表明, 改进后 DeepLabv3 + 算法的平均像素准确率、平均交并比、准确率分别为 99.02%、97.18% 和 99.71%。采用 r3d_18 作为 3D CNN-BiLSTM-ATFA 的主干网络效果最优。基于步态的奶牛个体识别平均准确率、灵敏度和精确度分别为 94.58%、93.47% 和 95.94%。奶牛躯干和腿部不同部位进行加权特征融合的个体识别实验表明识别准确率还可进一步提高。奶牛跛足对步态识别效果影响较为明显, 实验期间由健康变为跛足和一直跛足的奶牛个体识别准确率分别为 89.39% 和 92.61%。本文研究结果可为奶牛的智能化个体识别提供技术参考。

关键词: 奶牛; 个体识别; 步态特征; 3D CNN; BiLSTM

中图分类号: TP391.41

文献标识码: A

文章编号: 1000-1298(2024)07-0315-10

OSID:



Individual Identification Method of Cows Based on 3D CNN-BiLSTM-ATFA Network and Gait Feature

SI Yongsheng¹ NING Zepu¹ WANG Kejian¹ MA Yabin² YUAN Ming²

(1. College of Information Science and Technology, Hebei Agricultural University, Baoding 071001, China)

(2. Hebei Provincial Station for Livestock Varieties Producing and Spreading, Shijiazhuang 050061, China)

Abstract: Aiming at the low identification accuracy of cows with solid color or less pattern in pattern-based individual identification of cows, an individual identification method was proposed based on cow gait features. Firstly, the backbone network of DeepLabv3 + semantic segmentation algorithm was replaced by MobileNetv2 network. The channel and space based CBAM attention mechanism was introduced into this segmentation algorithm. The improved model was used to segment the silhouette of the cow. Then the 3D convolutional neural network (3D CNN) and the bidirectional long short-term memory network (BiLSTM) were constructed as the 3D CNN-BiLSTM network. The adaptive temporal feature aggregation module (ATFA) was further integrated into the above network to generate the 3D CNN-BiLSTM-ATFA cow individual identification model. Finally, individual identification experiments were conducted on a total of 1 242 video datasets from 30 cows. The results showed that the MPA, MIOU and Accuracy of the improved DeepLabv3 + algorithm were 99.02%, 97.18% and 99.71%, respectively. Individual recognition was optimal when r3d_18 was used as the backbone network of 3D CNN-BiLSTM-ATFA. The average accuracy, sensitivity and precision of individual identification based on cow gait were 94.58%, 93.47% and 95.94%, respectively. Individual identification experiments with weighted feature fusion for torso and legs showed that identification accuracy can be further improved. Lameness in dairy cows had a significant effect on gait identification,

收稿日期: 2024-02-20 修回日期: 2024-04-24

基金项目: 河北省重点研发计划项目(22327404D、22326609D)

作者简介: 司永胜(1973—), 男, 副教授, 主要从事农业领域机器视觉应用及无损检测研究, E-mail: siyongsheng@hebau.edu.cn

通信作者: 马亚宾(1975—), 男, 研究员, 主要从事奶牛生产性能测定和高效养殖研究, E-mail: dhimyb@163.com

the individual identification accuracies were 89.39% and 92.61% for cows that changed from healthy to lame and cows that remained lame during the experiment, respectively. The results can provide technical reference for intelligent individual identification of dairy cows.

Key words: dairy cows; individual identification; gait feature; 3D CNN; BiLSTM

0 引言

在推进奶牛养殖业规模化和精细化过程中,奶牛的个体识别尤为关键^[1-3]。准确识别个体奶牛是实现健康监测、疾病防控、产品溯源、行为分析等养殖管理工作的前提和基础^[4-6]。

部分养殖企业选择采用基于射频识别技术(RFID)的电子标签进行奶牛个体识别。电子标签存在成本较高、受电磁干扰和易脱落等问题^[7]。

基于计算机视觉的个体识别^[8-9]在不需要人为干预情况下进行,具有非接触、成本低、不易引起牲畜应激反应等优势^[10-12]。随着深度学习技术的发展,奶牛的个体识别研究取得了较大的进展。KUMAR等^[13]提出运用卷积神经网络(CNN)和深度信念网络(DBN)方法提取牛只口鼻纹理特征进行个体识别,识别准确率达98.99%。LI等^[14]基于CNN设计了一种轻量级神经网络,在10 239幅牛脸图像数据集上进行了实验,准确率为98.37%。上述方法图像采集难度大、容易引起奶牛的应激反应,不易推广。

部分研究采集奶牛的侧面和背部等部位图像来进行个体识别。何东健等^[15]提出了一种基于改进的YOLO v3网络,利用奶牛背部花纹进行奶牛识别的方法,奶牛的个体识别准确率达到95.91%。HU等^[16]使用YOLO目标检测算法对奶牛个体进行检测,然后将牛躯干分为头、躯干、腿部3部分,使用CNN对3部分分别进行特征提取,将3个部位提取到的特征进行特征融合,最后使用SVM算法对奶牛进行个体识别,准确率为98.36%。XIAO等^[17]利用Mask R-CNN识别奶牛背部,并提取背部花纹特征,然后使用SVM对奶牛进行个体识别,准确率为98.67%。司永胜等^[18]利用Mask R-CNN识别站立和躺卧姿态下的奶牛,然后对提取的奶牛花纹的传统特征和深度特征进行融合,最后使用SVM对站立和躺卧姿态下的奶牛进行个体识别,识别率分别达到98.66%和94.06%。

上述方法取得了较高的识别准确率。荷斯坦奶牛中存在一定比例的花纹较少或纯色的奶牛。当样本中存在花纹较少或纯色无花纹奶牛时,识别准确率会明显降低。在奶牛步态识别领域,MIMURA等^[19]提出了一种使用奶牛行走的RGB视频进行奶

牛识别的算法,但这种方法使用纹理的运动相关性来比较视频,未能直接捕捉到奶牛步态的本质特征。王萌萌^[20]利用能量图和肢体角度提取奶牛步态特征,提出了一种融合步态特征与纹理的奶牛个体识别方法,融合后步态识别准确率为72.73%。FUMIO等^[21]利用步态能量图提取奶牛步态特征,提出了一种结合步态和花纹特征的奶牛个体识别方法,单独利用步态特征的个体识别准确率为75.6%。以上研究都是使用传统方法进行识别,相较于深度学习方法,识别准确率较低。

本文以荷斯坦奶牛为研究对象,提出一种基于3D CNN-BiLSTM-ATFA网络和奶牛步态的个体识别方法,以期对奶牛养殖管理智能化提供技术支持。

1 数据采集与预处理

1.1 视频图像的采集

本文视频数据于2022年8月15日至9月18日采集自河北省保定市弘达牧业股份有限公司,共采集35 d,采集时间为07:00—10:00、14:00—17:00。采集了经过挤奶通道的30头荷斯坦奶牛的视频数据,图1为数据采集环境。

奶牛每天会进行2次挤奶工作,分批次经过一条长约40 m、宽约4 m的挤奶通道走至挤奶厅挤奶,并在完成挤奶后依次原路返回牛舍,选择在奶牛挤奶完成离开挤奶厅时进行视频数据采集。将一台相机固定安装在过道内侧,距离奶牛约3 m,调整相机距离地面高度约0.95 m,使奶牛位于相机视野正中间位置,拍摄角度与奶牛行走方向垂直。相机(MOKOSE型,C100-3.2 mm)的图像传感器型号是Sony Exmor R CMOS 1/2.3,帧率为30 f/s,空间分辨率为1 920像素×1 080像素。

选取实验牛舍的30头奶牛作为实验对象,其中包含常规花纹奶牛、纯色无花纹奶牛、花纹较少奶牛、花纹相似奶牛。所采集视频包含不同天气和光照,以保证模型的泛化能力和适应性。由于好奇或应激,存在奶牛奔跑或行走停滞、重叠等现象,本文剔除了上述情况视频数据,选择正常行走数据。

将采集的30头奶牛数据集按照个体的类别标签COW1~COW30进行编号,每头奶牛包含30~50个完整的视频步态周期。1个步态周期是指从4只牛蹄同时着地到下一次4只牛蹄同时着地的时间

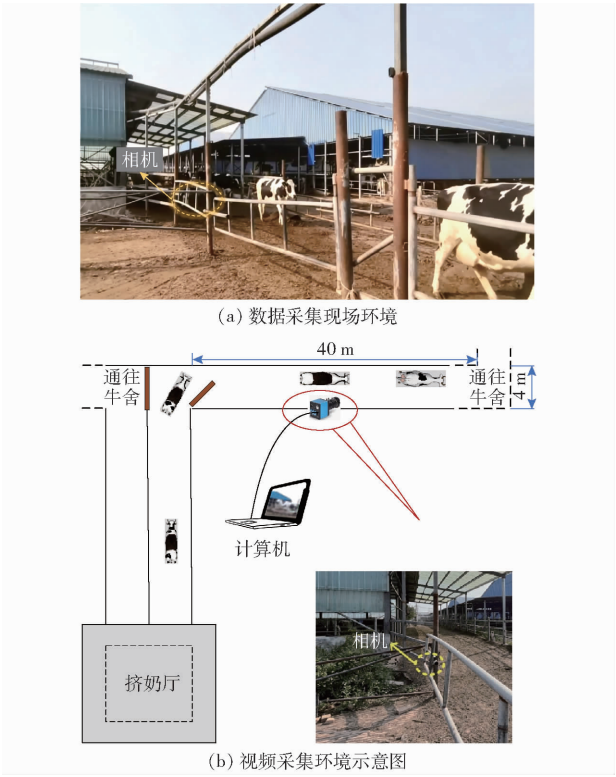


图 1 数据采集环境

Fig. 1 Data acquisition environment

段。将上述视频数据手工剪辑为单步态周期的视频,剪辑后数据集为 1 242 条视频数据。然后使用视频分解技术将视频数据分割为图像序列,用于奶牛步态个体识别。

本文实验平台的配置为 Intel (R) Xeon (R) Silver 4208 CPU @ 2.10 GHz, 运行内存为 64 GB (RAM), GPU 为 NVIDIA GeForce RTX 3090。软件环境为 Windows 10 系统,网络训练和测试均基于 Pytorch 深度学习框架,Python 版本为 3.8。

1.2 数据预处理

视频数据背景较为复杂,且存在花纹和阴影等

干扰因素。本研究使用改进的 DeepLabv3 + 模型对奶牛图像序列进行分割。

从 30 头奶牛的整个数据集中随机抽取 1 500 幅图像用于训练 DeepLabv3 + 模型,按照比例 7:2:1 划分为训练集、验证集、测试集。本文使用 Labelme 对训练集和验证集图像进行标注。选择 MobileNetv2 作为 DeepLabv3 + 模型的主干网络,并使用训练后的模型对所有图像进行分割,分割效果如图 2a 所示,输出奶牛身体轮廓的黑白剪影图,如图 2b 所示。然后将所有的剪影图序列按照比例 7:3 随机划分为训练集和验证集。

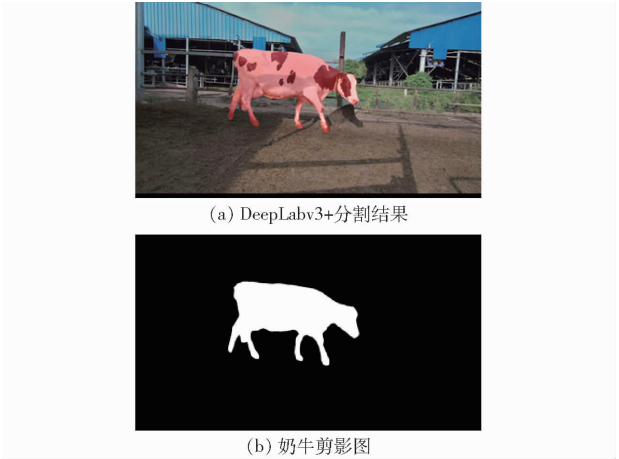


图 2 DeepLabv3 + 分割结果及分割后的剪影图

Fig. 2 Segmentation result and silhouette of DeepLabv3 +

2 研究方法

基于步态特征的奶牛个体识别过程如图 3 所示。利用 DeepLabv3 + 模型对图像序列中奶牛进行分割,输出黑白剪影图。然后将剪影图序列输入到本文提出的 3D CNN - BiLSTM - ATFA 网络中提取深度步态特征,并输出识别结果。

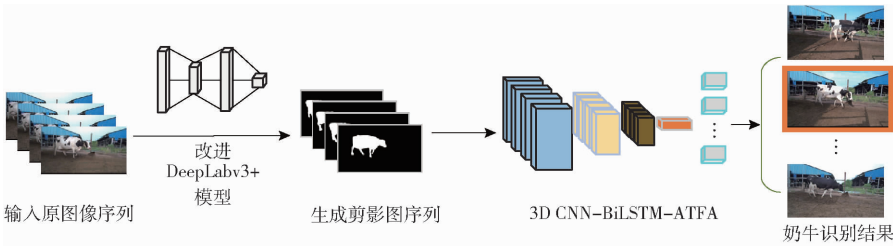


图 3 基于步态特征的奶牛个体识别过程

Fig. 3 Gait feature-based cow individual identification process

2.1 奶牛身体轮廓分割

2.1.1 分割模型选取

奶牛步态个体识别首先需要从背景中分割出奶牛个体。轮廓分割细节的准确性直接影响步态识别效果^[22]。分割算法主要分为语义分割和实例分割 2 类。本文为对单头奶牛进行分割,选取简单快速

的语义分割。

DeepLabv3 + 模型已经被验证可以较好地适用于复杂背景下的奶牛分割任务^[23]。该模型参考了目标检测中常见的特征融合策略,使网络保留比较多的浅层信息,同时加入深度可分离卷积对分割网络的速度进行优化,实现精度和时间的平衡。

2.1.2 主干网络

原始的 DeepLabv3 + 语义分割模型以 Xception 作为主干网络,其网络较深,参数量较大,分割速度较慢。采用轻量化的 MobileNetv2 网络作为语义分割模型的主干网络。MobileNetv2 在使用深度可分离卷积的基础上引入了倒置残差模块和线性瓶颈层,减少了模型参数量,使网络收敛更快,分割效率更高^[24]。

2.1.3 添加 CBAM 注意力机制

DeepLabv3 + 存在目标边缘分割不精确现象^[25-26],引入用于计算机视觉任务的 CBAM 注意力机制^[27],提高模型的训练与预测性能。CBAM 由 2 个子模块组成:通道注意力模块和空间注意力模块。通过同时关注通道和空间信息来提高卷积神经网络的性能,实现奶牛轮廓的准确分割。本文将一层 CBAM 添加在 DeepLabv3 + 网络的 ASPP 模块之后,ASPP 模块提取了多尺度的特征信息,CBAM 模块可以在这些特征上进一步提取和增强信息。

2.2 步态个体识别网络构建

2.2.1 3D CNN - BiLSTM - ATFA 网络模型

奶牛步态是指奶牛行走时的方式,步态个体识别技术主要通过对个体的体型和行走姿态、模式进行分析和辨认,进而确定个体的身份,其核心是对步态特征的提取^[28-29]。采用三维卷积神经网络(3D CNN)从图像序列中提取特征,然后使用双向长短期记忆网络(BiLSTM)和自适应时间特征聚合(ATFA)来进一步处理提取的特征。3D CNN 可以从图像序列中学习步态的关键空间和时间信息,3D CNN 和 BiLSTM 以及 ATFA 的结合能够进一步提取图像序列中潜在的时序信息。

2.2.2 3D CNN 网络模型

为了从视频中学习奶牛步态个体识别中关键的空间和时间信息,采用 3D CNN 对步态特征进行提取。Residual 3D(R3D)采用 ResNet 中的残差结构,这一架构在视频领域取得了较好的性能。其中 r3d_18 就是在 ResNet18 的基础上用 3D 卷积核取代 2D 卷积核,以适应视频数据的处理。其 3D 卷积核大小为 $3 \times 3 \times 3$,步长为 $1 \times 1 \times 1$ 。除了 pool1 的核大小为 $1 \times 2 \times 2$,步长为 $1 \times 2 \times 2$,目的是在前期保留时间信息,其余 3D 池化层都为 $2 \times 2 \times 2$,步长为 $2 \times 2 \times 2$ 。每个全连接层具有 4 096 个输出单元。

2.2.3 BiLSTM 网络模型

奶牛的步态特征与步态前后动作之间的关系相关,采用 BiLSTM 来提取序列数据中的时序特征。该网络是一种递归神经网络(RNN),包含 2 层 LSTM,分别处理输入序列的正向和反向序列信息,

能够捕获序列中过去和未来上下文信息^[30]。本系统中使用的 LSTM 单元定义为

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_c[h_{t-1}, x_t] + b_c)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

$$h_t = o_t \tanh(c_t)$$

式中 σ ——sigmoid 函数

W_i, W_f, W_c, W_o ——输入门、遗忘门、单元状态和输出门的权重矩阵

b_i, b_f, b_c, b_o ——对输入门、遗忘门、单元状态和输出门偏置矢量

h_t ——隐藏状态

x_t ——时间步长 t 的输入

c_t ——单元状态

i_t, f_t, o_t ——输入门、遗忘门、输出门

其中时间步长根据最短步态周期长度设置为 $t = 30$ 。

2.2.4 ATFA 模块

在奶牛步态个体识别过程中,为了使模型更加关注重要的时间特征,更好地适应不同时间段的数据特征,本研究提出了一种自适应时间特征聚合(Adaptive temporal feature aggregation, ATFA)模块。该模块通过学习时间权重来自适应地聚合不同时间尺度的特征,以增强时间序列数据的特征表示能力。如图 4 所示,在该结构中定义一个 Transformer-based Attention 模块^[31-32],该模块包含 query、key、value 3 个线性层,用于计算注意力权重,这些注意力权重表示在时间序列中各个时间步的重要性,并使用 Softmax 函数对权重进行归一化。然后,使用这些注意力权重对特征进行加权聚合,以获取聚焦在关键时间步上的特征表示,从而使得模型能够在关键时间步上有更强的关注度。ATFA 模块中的 2 个多层感知机分别用来计算 key、value,之后将 query 和 key

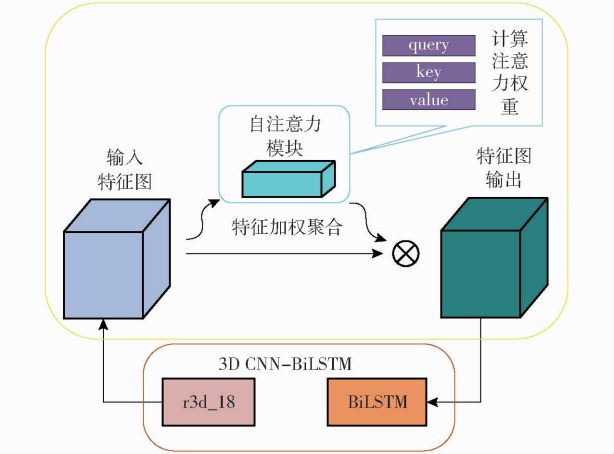


图 4 ATFA 模块结构

Fig. 4 ATFA module structure

组合用来计算权重,然后结合 value 进行加权求和,动态分配不同时间步的权重,提取最优特征。

ATFA 被添加在 3D CNN 网络之后,聚合来自 3D CNN 的特征,然后将聚合后的特征传递给 BiLSTM 网络进一步提取时序特征。上述设计可以使模型通过 ATFA 充分利用 3D CNN 的特征提取能力,并保留时间信息。

2.3 奶牛个体识别及评价指标

本文将奶牛步态个体识别准确率 (Accuracy)、灵敏度 (Sensitivity)、精确度 (Precision) 作为模型性能的评价指标。

3 结果与分析

3.1 奶牛身体轮廓分割模型选取

为了选取最佳奶牛分割模型,本文对 U - Net、PSPNet、RefineNet、DeepLabv3 + 等进行比较,均使用 MobileNetv2 作为主干网络,结果如表 1 所示。从分割准确率与速度两方面综合考虑,本文选用 DeepLabv3 + 模型。

表 1 不同模型分割奶牛轮廓结果

Tab.1 Results of segmenting cow contours using different models

模型	平均像素 准确率/%	平均交 并比/%	准确 率/%	单帧预测 时间/ms	参数量
U - Net	98.92	96.85	99.60	56.24	1.077×10^7
PSPNet	96.69	93.72	98.79	23.85	4.380×10^6
RefineNet	95.78	93.58	98.82	83.52	1.634×10^7
DeepLabv3 +	98.53	96.95	99.43	24.96	5.810×10^6
CBAM +	99.02	97.18	99.71	24.98	5.830×10^6
DeepLabv3 +					

在 DeepLabv3 + 模型基础上增加 CBAM 注意力机制,相比于原模型,在保持原速度的基础上,平均像素准确率 (Mean pixel accuracy, MPA)、平均交并比 (Mean intersection over union, MIOU)、准确率分别提高了 0.49、0.23、0.28 个百分点。

3.2 3D CNN 模型的选取

传统的 2D CNN - BiLSTM (以下称为 CNN - BiLSTM) 中 CNN 主要是提取图像的二维特征^[33-34],与常规的采用 CNN - BiLSTM 处理时序任务不同,本文采用 3D CNN - BiLSTM 网络进行奶牛步态个体识别。结果表明,基于 r3d_18、mc3d_18 和 r2 + 1d 的 3D CNN - BiLSTM 模型相较基于 VGG19、ResNet18、MobileNetv2 和 DenseNet201 的 CNN - BiLSTM 模型具有明显优势。如图 5 所示,基于上述 CNN 主干网络的平均识别准确率分别为 73.18%、70.75%、69.56% 和 85.54%,改进后基于 3D CNN

主干网络准确率分别为 92.36%、90.27% 和 88.74%。

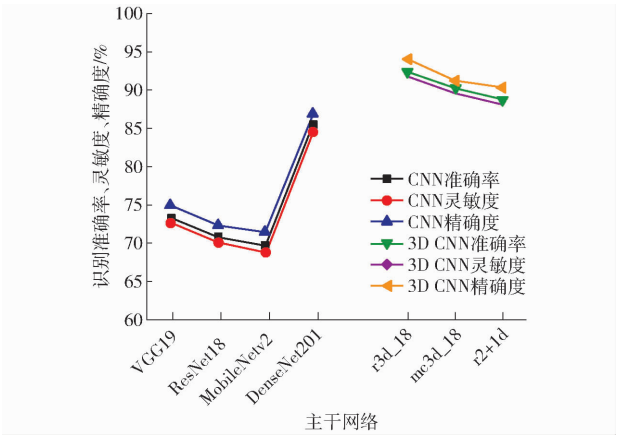


图 5 基于不同主干网络的步态个体识别结果
Fig.5 Results of individual gait identification based on different backbone networks

奶牛步态具有明显的时序特征,其中每个时间步都包含了重要的运动信息和动态变化。CNN 只能提取空间特征,而 3D CNN 能够在其基础上进一步提取时间特征,捕获奶牛运动的动态信息,其与 BiLSTM 的结合,能够有效地提取时序特征和序列中的动态变化,实现较高准确率的个体识别。

3D CNN 中,r3d_18 模型在识别奶牛步态个体方面表现最佳,达到了 92.36% 的准确率,对比其他 2 个主干网络分别高 2.09、3.63 个百分点。另外,灵敏度和精确度,r3d_18 分别达到 91.87% 和 94.04%,表现最好,体现了该模型对于正例准确的预测和识别能力。r3d_18 模型基于 ResNet18 架构,该模型的高准确率归因于其较深的网络结构和较强的特征提取能力。综上,将 r3d_18 作为主干网络。

3.3 ATFA 模块效果分析

ATFA 模块显著提升了步态个体识别效果,较改进前准确率提高 2.22 个百分点,达到 94.58%,灵敏度和精确度也分别提高 1.6、1.9 个百分点。为了直观展示 ATFA 模块的作用,采用 t-SNE (t-distributed stochastic neighbor embedding) 降维算法对奶牛高维特征进行降维,将特征进行可视化展示。从训练集中随机选取 9 个类别的特征进行可视化。如图 6 所示,在添加 ATFA 模块后,不同类别之间特征的距离更远,同种类别特征之间的距离更近,表明改进后模型具有较强的步态特征提取能力。

性能提升的原因为:改进前方法中简单堆叠时间特征可能导致信息的部分丢失,而 ATFA 模块通过自适应聚合方式减少了这种信息丢失。ATFA 模块的自适应性允许模型根据不同数据自动调整时间权重,从而更好地提取奶牛步态中的时间信

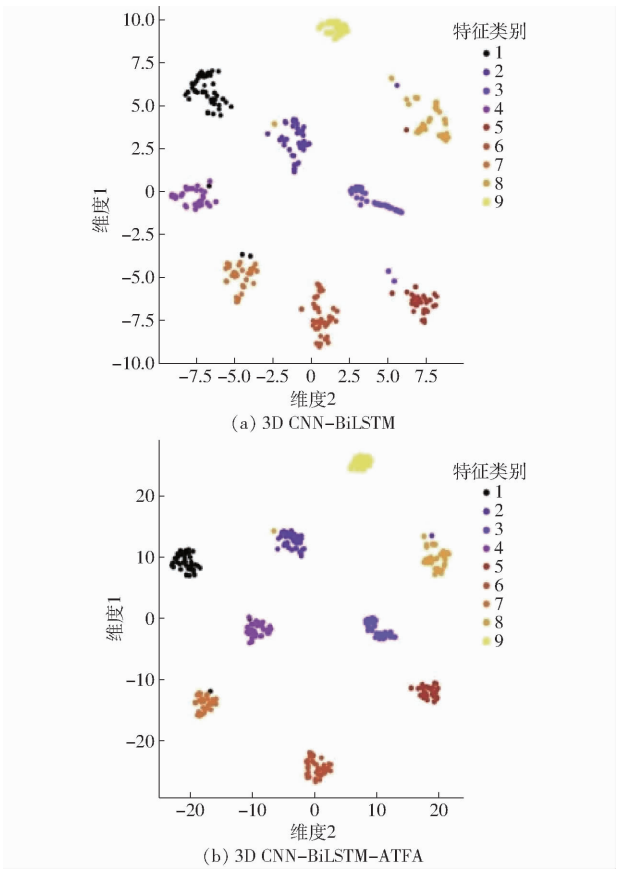


图 6 添加 ATFA 前后特征提取效果对比
Fig.6 Comparison of feature extraction effect before and after adding ATFA

息,增强模型对时间序列数据的理解,提高识别效果。

3.4 数据归一化效果分析

本文设计了相关的对比实验,在初始阶段对语义分割后的奶牛剪影图进行了裁剪、对齐、归一化操作,验证数据归一化是否会对奶牛步态个体识别效果产生影响。图 7 为归一化前后的 2 种数据形式。

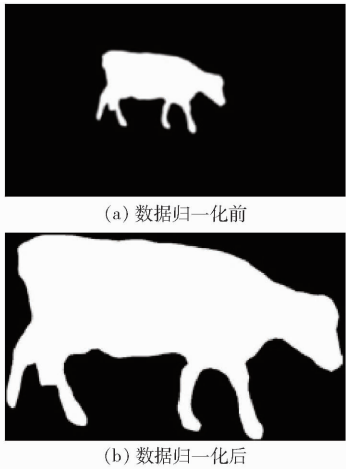


图 7 归一化前后数据

Fig.7 Data before and after normalization

将 r3d_18 作为 3D CNN-BiLSTM 的主干网络,

并结合 ATFA 模块进行了对比和消融实验,图 8 为该实验结果的累计匹配特征 (Cumulative match characteristic,CMC) 曲线。结果显示,基于 3D CNN-BiLSTM 和 3D CNN-BiLSTM-ATFA 模型在数据归一化前原数据的奶牛步态个体识别准确率分别为 92.36% 和 94.58%,高于数据归一化后的 89.95% 和 92.17%。在本文中,数据归一化反而对奶牛个体识别起到了负面作用。

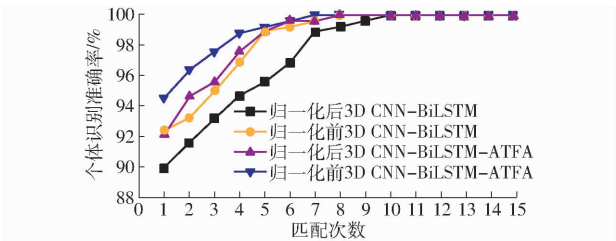


图 8 数据归一化前后对比消融实验
Fig.8 Contrast and ablation experiments before and after data normalization

分析原因,数据归一化通常会将数据的尺度缩放到一个标准范围,以消除尺度差异,然而奶牛的尺度信息是进行个体识别的重要特征。因此本文用未归一化的原始数据进行个体识别。

3.5 不同身体部位在识别中的贡献

在奶牛行走过程中,身体的不同部位动作幅度不同,各个部位对个体识别的影响不同。本文使用 Grad-CAM 算法对提出的模型进行可视化分析,图 9 为步态特征提取过程中可视化热力图,热力图可展示出模型对奶牛不同部位特征的关注情况。奶牛躯干部分的热力覆盖范围更大且颜色更深,可见模型的主要关注区域分布在躯干部分,而腿部的关注区域较小,对腿部的特征提取不充分。

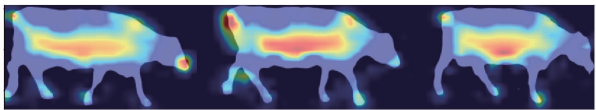


图 9 步态特征热力图可视化

Fig.9 Thermal map visualization of gait features

本研究的假设是由于个体不同,奶牛的步长、步频、步速等会有明显不同,上述结果与该假设不符。进一步分析该问题,有利于识别准确率的提高。

根据热力图关注区域的分布情况将奶牛剪影图分割成躯干和腿部 2 部分。为方便批量自动化的身体分割,本部分研究使用裁剪对齐归一化后的剪影图。按照上下比例 6:4 将归一化的剪影图分割为如图 10 所示的 2 部分。

然后用 3D CNN-BiLSTM-ATFA 对躯干和腿部 2 部分数据重新进行训练和测试,结果显示利用奶牛躯干剪影图进行个体识别的准确率为

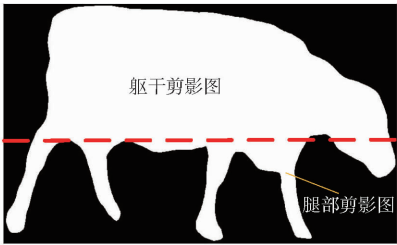


图 10 奶牛分割图
Fig. 10 Cow segmentation map

87.34%，高于腿部的 70.48%。由此可见，躯干部分的运动特征在整个步态个体识别过程中所起作用大于腿部。分析原因是躯干在整个牛身中所占比例大于腿部，包含的信息也更加丰富。另外在奶牛行走过程中，腿部更容易受到外界环境的影响，稳定性不如躯干部位。

对这 2 个身体部位步态特征概率向量进行加权特征融合，来提高步态个体识别的准确率。在使用特征加权求和的融合策略时，为求取最佳的加权系数，对躯干步态特征概率向量和腿部步态特征概率向量权重系数以步长为 0.1 进行不同组合。特征加权求和的躯干步态特征概率向量和腿部步态特征概率向量权重系数比例分别为 1:9、2:8、3:7、4:6、5:5、6:4、7:3、8:2、9:1 时奶牛识别准确率分别为 73.16%、76.48%、80.42%、84.59%、88.78%、93.67%、94.36%、92.43%、88.94%。当躯干步态特征概率向量和腿部步态特征概率向量权重系数比例为 7:3 时奶牛的个体识别准确率最高，为 94.36%，与 3.4 节中归一化后 3D CNN - BiLSTM - ATFA 的 92.17% 相比，融合后比融合前提升 2.19 个百分点。图 11 为特征融合后的热力图，与图 9 相比，腿部热力覆盖范围变大。说明特征融合后模型对躯干和腿部进行了合理的关注。

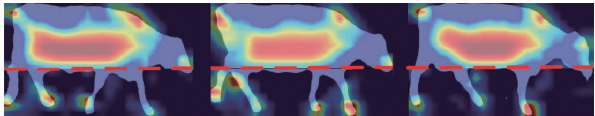


图 11 躯干和腿部步态特征融合热力图可视化
Fig. 11 Thermal map visualization of fusion of trunk and leg gait features

弓背曲率是奶牛跛行和个体识别的重要特征，为研究背部在步态识别中的贡献，按比例 4:6 将躯干分割为脊柱和腰部 2 部分。对脊柱和腰部数据进行训练和测试，结果显示利用奶牛脊柱剪影图进行个体识别的准确率为 71.69%，高于腰部的 65.06%。图 12 为对躯干进一步分割后的热力图，与图 11 相比，脊柱部位热力覆盖范围变大，表明躯干分割后模型更专注于脊柱部分的特征。

对脊柱和腰部的步态特征进行加权特征融合，

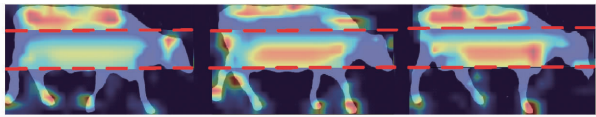


图 12 奶牛被分为 3 部分的步态特征热力图
Fig. 12 Thermal map of gait features of cows segmented into three parts

权重系数比例为 6:4 时准确率达到最高，为 75.53%，低于融合前整个躯干的 87.34%。可能的原因是将躯干分割为 2 部分，都丢失了一些步态识别中的重要信息，如脊柱和腰部之间的相互作用、躯干整体的动态特征等。上述结果表明，脊柱在步态识别中具有较高贡献，但将奶牛分割为整个躯干和腿部 2 部分，识别效果更优。

3.6 步态周期长度对识别的影响

奶牛的步态周期会对步态特征提取及识别效果产生影响。本文对数据集中步态周期长度进行了统计，并将步态周期分为 6 个等级。如图 13 所示，1 个步态周期对应的视频帧数最短的为 30 帧，最长的为 58 帧，而约 90% 的奶牛帧数在 35 ~ 50 帧之间。由柱状图可见，在 30 ~ 55 帧数范围内，奶牛个体识别平均准确率为先增后减的趋势。可能的原因是步态周期对应的帧数较低时，奶牛行进过快，对步态特征提取不足。随着步态周期变长，奶牛行进更稳定，更容易提取步态特征，识别准确率增高。轻度跛足奶牛行进较慢，步态特征不稳定，其步态周期一般分布在 50 ~ 55 帧之间，步态周期在上述范围附近的出现识别准确率下降的情况。步态周期帧数高于 55 帧时，准确率有明显提升，通过分析发现此范围为重度跛足奶牛，其数量较少且具有明显的异常步态，更容易识别。

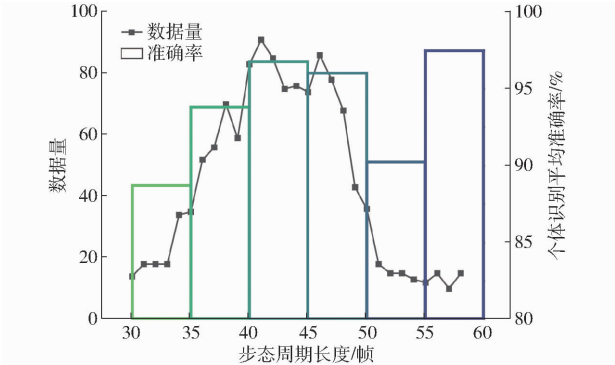


图 13 帧数统计及其识别准确率
Fig. 13 Frame count statistics and identification accuracy

3.7 跛足对步态识别的影响

通过 3.6 节分析可知，跛足会影响基于步态的奶牛个体识别，对此进行单独分析。在畜牧专家指导下，根据奶牛跛足情况将数据集划分为数据采集期间健康、由健康变为跛足、跛足 3 个类别。各个类

别数据集容量及实验结果如表 2 所示,健康奶牛个体识别准确率最高,为 96.55%。由健康变为跛足奶牛在跛足后识别准确率为 89.39%,跛足奶牛识别准确率为 92.61%。

如图 14 所示,采用 t-SNE 降维算法将 3 类奶牛特征进行可视化展示,健康奶牛的特征距离最紧凑,跛足奶牛的次之,由健康变为跛足的奶牛特征之间的距离最远。表明跛足对奶牛的步态识别的效果影响较为明显。

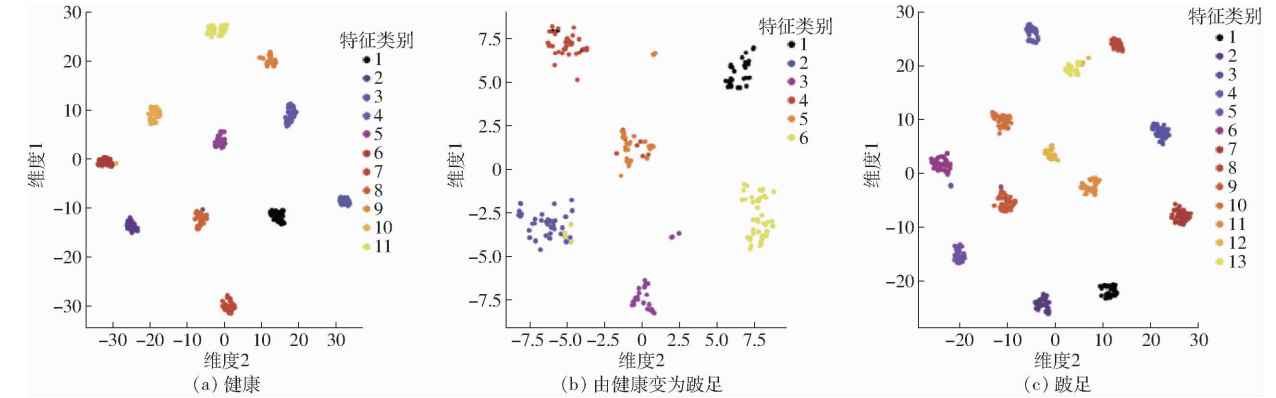


图 14 不同健康情况奶牛步态特征提取效果可视化

Fig. 14 Visualization of gait feature extraction results for cows with different health conditions

由图 14 可知,健康奶牛的步态特征最为稳定,步态识别效果最优。与健康奶牛相比,跛足奶牛步态特征不稳定,可能的原因是跛足程度变化而造成步态特征的改变。由健康变为跛足的奶牛行走姿势变化,步态特征改变最大,但仍能正确聚类大部分样本,其原因是奶牛跛足步态变化主要发生在腿部,而稳定性较高的躯干部分在步态识别中作用占比高于腿部,这和 3.5 节的结果相符。

4 结论

(1) 针对纯色或花纹较少奶牛识别准确率低的问题,提出了一种基于 3D CNN – BiLSTM – ATFA 网络和步态特征的奶牛个体识别方法,实现了对奶牛个体的准确识别。

(2) 基于 DeepLabv3 + 模型,采用 MobileNetv2 作为主干网络提取特征,引入 CBAM 注意力机制,MPA、MIOU 和 Accuracy 达到 99.02%、97.18% 和 99.71%,实现了奶牛身体轮廓的高效准确分割。

表 2 健康和跛足奶牛的识别准确率

Tab. 2 Identification accuracy of healthy and lame cows

健康情况	奶牛数量/头	训练集视频		测试集视频		准确率/%
		视频数/帧	图像数/幅	视频数/帧	图像数/幅	
健康	11	348	13 588	145	5 847	96.55
由健康变为跛足	6	181	7 266	66	3 157	89.39
跛足	13	353	17 292	149	7 375	92.61

注:由健康变为跛足类别中训练集为健康状态,测试集为跛足状态。

(3) 将 CNN – BiLSTM 主干网络替换为 r3d_18 进行改进,生成新的 3D CNN – BiLSTM 网络,并添加了 ATFA 模块进行特征聚合,最终实现对奶牛分类。改进后的模型步态个体识别平均准确率、灵敏度和精确度达到 94.58%、93.47% 和 95.94%,准确率相较 CNN – BiLSTM 模型提高约 9 个百分点,较添加 ATFA 前 3D CNN – BiLSTM 网络提高约 2 个百分点,算法性能较强。

(4) 在将奶牛躯干和腿部不同部位进行加权特征融合的个体识别中,躯干和腿部权重比例为 7:3 时识别准确率最高,达 94.36%,说明如实现腿部和躯干的自动分割,通过特征融合方法可进一步提高奶牛个体识别的准确率。

(5) 奶牛由健康变为跛足后,个体识别准确率降低为 89.39%。奶牛跛足对基于步态的奶牛个体识别准确率影响较为明显,但接近 90% 的识别准确率说明奶牛步态特征具有较强的区分能力和鲁棒性,可以和奶牛花纹特征等结合使用。

参考文献

[1] 司永胜,肖坚星,刘刚,等. 基于 MSRCF 与改进 YOLO v4 的躺卧奶牛个体识别方法[J]. 农业机械学报,2023,54(1):243 – 250,262.

SI Yongsheng, XIAO Jianxing, LIU Gang, et al. Individual identification method of lying cows based on MSRCF and improved YOLO v4 model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 243 – 250, 262. (in Chinese)

[2] 王政,许兴时,华志新,等. 融合 YOLO v5n 与通道剪枝算法的轻量化奶牛发情行为识别[J]. 农业工程学报,2022, 38(23):130 – 140.

WANG Zheng, XU Xingshi, HUA Zhixin, et al. Recognition of estrus behavior in lightweight dairy cows by combining YOLO

v5n and channel pruning algorithm[J]. Transactions of the CSAE, 2022,38(23):130 – 140. (in Chinese)

[3] 刘月峰,边浩东,何滢婕,等. 基于幅值迭代剪枝的多目标奶牛进食行为识别方法[J]. 农业机械学报,2022,53(2): 274 – 281.

LIU Yuefeng, BIAN Haodong, HE Yingjie, et al. Detection method of multi-objective cows feeding behavior based on iterative magnitude pruning[J]. Transactions of the Chinese Society for Agricultural Machinery,2022,53(2):274 – 281. (in Chinese)

[4] AWAD I A. From classical methods to animal biometrics: a review on cattle identification and tracking[J]. Computers and Electronics in Agriculture, 2016, 123: 423 – 435.

[5] 何东健,刘冬,赵凯旋. 精准畜牧业中动物信息智能感知与行为检测研究进展[J]. 农业机械学报,2016,47(5):231 – 244.

HE Dongjian, LIU Dong, ZHAO Kaixuan. Review of perceiving animal information and behavior in precision livestock farming [J].Transactions of the Chinese Society for Agricultural Machinery,2016,47(5):231 – 244. (in Chinese)

[6] DU Yongxing, KOU Yan, LI Baoshan, et al. Individual identification of dairy cows based on deep learning and feature fusion [J]. Animal Science Journal, 2022,93(1): e13789.

[7] WILLIAMS L R, FOX D R, BISHOP-HURLEY G J, et al. Use of radio frequency identification (RFID) technology to record grazing beef cattle water point use[J]. Computers and Electronics in Agriculture, 2019,156:193 – 202.

[8] WANG Yunfei, XU Xingshi, WANG Zheng, et al. ShuffleNet – Triplet: a lightweight RE-identification network for dairy cows in natural scenes[J]. Computers and Electronics in Agriculture, 2023, 205:107632.

[9] KAWAGOE Y, KOBAYASHI I, ZIN T T. Facial region analysis for individual identification of cows and feeding time estimation [J]. Agriculture,2023,13(5):1016.

[10] WENG Z, MENG F, LIU S, et al. Cattle face recognition based on a two-branch convolutional neural network[J]. Computers and Electronics in Agriculture, 2022, 196: 106871.

[11] 杨蜀秦,刘杨启航,王振,等. 基于融合坐标信息的改进 YOLO V4 模型识别奶牛面部[J]. 农业工程学报, 2021, 37(15): 129 – 135.

YANG Shuqin, LIU Yangqihang, WANG Zhen, et al. Improved YOLO V4 model for face recognition of diary cow by fusing coordinate information[J]. Transactions of the CSAE, 2021,37(15): 129 – 135. (in Chinese)

[12] 张宏鸣,孙扬,赵春平,等. 反刍家畜典型行为监测与生理状况识别方法研究综述[J]. 农业机械学报,2023,54(3):1 – 21.

ZHANG Hongming, SUN Yang, ZHAO Chunping, et al. Review on typical behavior monitoring and physiological condition identification methods for ruminant livestock[J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(3): 1 – 21. (in Chinese)

[13] KUMAR S, PANDEY A, SATWIK K S R, et al. Deep learning framework for recognition of cattle using muzzle point image pattern[J]. Measurement, 2018, 116: 1 – 17.

[14] LI Z, LEI X, LIU S. A lightweight deep learning model for cattle face recognition [J]. Computers and Electronics in Agriculture, 2022, 195: 106848.

[15] 何东健,刘建敏,熊虹婷,等. 基于改进 YOLO v3 模型的挤奶奶牛个体识别方法[J]. 农业机械学报,2020,51(4):250 – 260.

HE Dongjian, LIU Jianmin, XIONG Hongting, et al. Individual identification of dairy cows based on improved YOLO v3[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020,51(4):250 – 260. (in Chinese)

[16] HU H, DAI B, SHEN W, et al. Cow identification based on fusion of deep parts features[J]. Biosystems Engineering, 2020, 192(C):245 – 256.

[17] XIAO J, LIU G, WANG K, et al. Cow identification in free-stall barns based on an improved Mask R – CNN and an SVM[J]. Computers and Electronics in Agriculture, 2022, 194:106738.

[18] 司永胜,王朝阳,张艳,等. 基于深度与传统特征融合的非限制条件下奶牛个体识别[J]. 农业机械学报,2023,54(6): 272 – 279.

SI Yongsheng, WANG Chaoyang, ZHANG Yan, et al. Individual identification of dairy cows under unrestricted conditions based on fusion of depth and traditional features[J]. Transactions of the Chinese Society for Agricultural Machinery,2023, 54(6):272 – 279. (in Chinese)

[19] MIMURA S, ITOH K, KOBAYASHI T, et al. The cow gait recognition using CHLAC[C] //Proceedings of the 2008 Bio-inspired, Learning and Intelligent Systems for Security. IEEE, 2008: 56 – 57.

[20] 王萌萌. 融合步态特征与纹理的奶牛个体识别研究[D]. 天津:河北工业大学,2020.

WANG Mengmeng. Fusion gait feature and texture of cows individual identification research[D]. Tianjin: Hebei University of Technology, 2020. (in Chinese)

[21] FUMIO O, SAYA I, YASUSHI M, et al. RGB – D video-based individual identification of dairy cows using gait and texture analyses[J]. Computers and Electronics in Agriculture, 2019, 165:104944.

[22] KHAN M H, FARID M S, GRZEGORZEK M. A comprehensive study on codebook-based feature fusion for gait recognition [J]. Information Fusion, 2023, 92: 216 – 230.

[23] WU D, YIN X, JIANG B, et al. Detection of the respiratory rate of standing cows by combining the Deeplab V3 + semantic segmentation model with the phase-based video magnification algorithm[J]. Biosystems Engineering,2020,192: 72 – 89.

[24] LI W, LIU K. Confidence-aware object detection based on MobileNetv2 for autonomous driving[J]. Sensors, 2021, 21(7): 2380.

[25] 杨蜀秦,王鹏飞,王帅,等. 基于 MHSA + DeepLab v3 + 的无人机遥感影像小麦倒伏检测[J]. 农业机械学报,2022,

- 53(8):213–219,239.
- YANG Shuqin, WANG Pengfei, WANG Shuai, et al. Detection of wheat lodging in UAV remote sensing image based on Multi-head Self-attention DeepLab v3 + [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(8):213–219,239. (in Chinese)
- [26] 常哈,郭树欣,张海洋,等. 基于改进 CBAM-DeepLab V3+ 的苹果种植面积提取[J]. 农业机械学报, 2023, 54(增刊2):206–213.
- CHANG Han, GUO Shuxin, ZHANG Haiyang, et al. Apple planting area extraction based on improved DeepLab V3 + [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(Supp.2):206–213. (in Chinese)
- [27] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module [C] // European Conference on Computer Vision, Munich, Germany: Springer, 2018:3–19.
- [28] HERTEM V T, TELLO S A, VIAZZI S, et al. Implementation of an automatic 3D vision monitor for dairy cow locomotion in a commercial farm[J]. Biosystems Engineering, 2018, 173:166–175.
- [29] SCHLAGETER-TELLO A, HERTEM V T, BOKKERS E A, et al. Performance of human observers and an automatic 3-dimensional computer-vision-based locomotion scoring method to detect lameness and hoof lesions in dairy cows[J]. Journal of Dairy Science, 2018, 101(7):6322–6335.
- [30] 李前,初梦苑,康熙,等. 基于计算机视觉的奶牛跛行识别技术研究进展[J]. 农业工程学报, 2022, 38(15):159–169.
- LI Qian, CHU Mengyuan, KANG Xi, et al. Research progress of lameness recognition technology in dairy cows using computer vision[J]. Transactions of the CASE, 2022, 38(15):159–169. (in Chinese)
- [31] HARRER S. Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine[J]. eBioMedicine, 2023, 90:1045122.
- [32] JI Y, ZHOU Z, LIU H, et al. DNABERT: pre-trained bidirectional encoder representations from transformers model for DNA-language in genome[J]. Bioinformatics, 2021, 37(15):2112–2120.
- [33] 姚冲,李前,刘刚,等. 基于 NAS-Res 的局部遮挡荷斯坦奶牛个体识别[J]. 农业机械学报, 2023, 54(增刊1):252–259.
- YAO Chong, LI Qian, LIU Gang, et al. Individual identification of partially occluded Holstein cows based on NAS-Res[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(Supp.1):252–259. (in Chinese)
- [34] WU D, WANG Y, HAN M, et al. Using a CNN-LSTM for basic behaviors detection of a single dairy cow in a complex environment[J]. Computers and Electronics in Agriculture, 2021, 182:106016.

(上接第 279 页)

- [22] 樊阳阳,裴正军,陈俭,等. 基于近红外高光谱成像技术的干制红枣品种鉴别[J]. 光谱学与光谱分析, 2017, 37(3):836–840.
- FAN Yangyang, QIU Zhengjun, CHEN Jian, et al. Identification of varieties of dried red jujubes with near-infrared hyperspectral imaging[J]. Spectroscopy and Spectral Analysis, 2017, 37(3):836–840. (in Chinese)
- [23] 苏军,饶元,张敬尧,等. 基于 GA 优化 SVM 的干制红枣品种分类方法[J]. 洛阳理工学院学报(自然科学版), 2018, 28(4):65–69,93.
- SU Jun, RAO Yuan, ZHANG Jingyao, et al. Study on classification method of jujube varieties based on GA optimized SVM [J]. Journal of Luoyang Institute of Science and Technology (Natural Science Edition), 2018, 28(4):65–69,93. (in Chinese)
- [24] ZHANG R C, YUAN Y C, MENG X, et al. A multitask model based on MobileNetV3 for fine-grained classification of jujube varieties[J]. Journal of Food Measurement and Characterization, 2023, 17(5):4305–4317.
- [25] LIN T Y, ROYCHOWDHURY A, MAJI S. Bilinear CNNs for fine-grained visual recognition[J]. arXiv Preprint arXiv: 1504.07889, 2015.
- [26] FU J, ZHENG H, MEI T. Look closer to see better: recurrent attention convolutional neural network for fine-grained image recognition[C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017:4476–4484.
- [27] ZHENG H, FU J, MEI T, et al. Learning multi-attention convolutional neural network for fine-grained image recognition[C] // 2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017:5219–5227.
- [28] YANG Z, LUO T, WANG D, et al. Learning to navigate for fine-grained classification[J]. arXiv Preprint arXiv: 1809.00287, 2018.
- [29] CHANG D, DING Y, XIE J, et al. The Devil is in the channels: mutual-channel loss for fine-grained image classification[J]. IEEE Transactions on Image Processing, 2020, 29:4683–4695.
- [30] CHEN Y, BAI Y, ZHANG W, et al. Destruction and construction learning for fine-grained image recognition[C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019:5152–5161.
- [31] GAO Y, HAN X, WANG X, et al. Channel interaction networks for fine-grained image categorization[C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2020:10818–10825.