doi:10.6041/j.issn.1000-1298.2025.12.053

基于 DAM - ResNet 的广视野下牛舍奶牛个体识别方法

司永胜[1,2]　鲁文柯[1,2]　王克俭[1,2]　马亚宾[3]　袁　明[3]　王　斌[1,2]　王振存[4]

(1.河北农业大学信息科学与技术学院,保定 071001;2.河北省农业大数据重点实验室,保定 071001;
3.河北省畜牧良种工作总站,石家庄 050061;4.唐山市锦程牧业有限公司,唐山 064000)

摘要:针对牛舍内广视野下奶牛图像存在遮挡、变形较大而造成个体识别困难等问题,提出了一种基于 DAM - ResNet 模型的广视野下牛舍奶牛个体识别方法。采用 Mask R - CNN 模型对奶牛背部、躯干左侧、躯干右侧以及臀部 4 类奶牛部位进行分割。在 ResNet34 残差网络的基础上引入第二代可变形卷积,以加强对分割结果中奶牛花纹变形的图像特征提取;将 AFF 注意力特征融合模块融入到残差结构中,实现对远距离小目标奶牛图像的准确识别;采用细粒度分类损失函数互通道损失 MC - Loss 提高模型对花纹相似奶牛的识别准确率。利用奶牛分割后的图像构建奶牛多部位数据集,并训练 DAM - ResNet 模型。对 57 头奶牛共 12 864 幅图像数据集进行了奶牛个体识别测试,结果表明,奶牛背部、躯干左侧、躯干右侧、臀部识别准确率分别为 96.13%、96.56%、96.94%、93.14%,比 ResNet 原模型分别提高 2.76、2.88、2.92、4.25 个百分点。本文方法在 10 ~ 20 m、20 ~ 30 m 和 30 ~ 40 m 范围内识别准确率分别为 97.54%、90.72% 和 82.17%。研究结果可为奶牛智能化养殖提供技术支持。

关键词:奶牛个体识别;广视野;可变形卷积;注意力特征融合;细粒度分类

中图分类号:TP391.41　　**文献标识码**:A　　**文章编号**:1000-1298(2025)12-0581-10　　**OSID**:

Individual Identification Method of Dairy Cows in Cowsheds under Wide Field of View Based on DAM - ResNet

SI Yongsheng[1,2]　LU Wenke[1,2]　WANG Kejian[1,2]　MA Yabin[3]　YUAN Ming[3]　WANG Bin[1,2]　WANG Zhencun[4]

(1. *College of Information Science and Technology*, *Hebei Agricultural University*, *Baoding* 071001, *China*
2. *Key Laboratory of Agricultural Big Data of Hebei Province*, *Baoding* 071001, *China*
3. *Hebei Provincial Station for Livestock Varieties Producing and Spreading*, *Shijiazhuang* 050061, *China*
4. *Tangshan Jincheng Animal Husbandry Co.*, *Ltd.*, *Tangshan* 064000, *China*)

Abstract: Aiming at the problems such as occlusion and large deformation of cow images in the cowshed under a wide field of view, which cause difficulties in individual identification, a method for individual identification of cows in the cowshed under a wide field of view based on the DAM - ResNet model was proposed. The Mask R - CNN model was used to segment four types of dairy cow parts: the back, the left side of the trunk, the right side of the trunk and the buttocks. Based on the ResNet34 residual network, the second-generation deformable convolution was introduced to enhance the extraction of deformed image features of cow patterns in the segmentation results. Integrating the AFF attention feature fusion module into the residual structure to achieve accurate recognition of images of small target cows at a distance. The fine-grained classification loss function, interchannel loss (MC - Loss), was adopted to improve the recognition accuracy of the model for cows with similar patterns. The multi-part dataset of dairy cows was constructed by using the segmented images of dairy cows, and the DAM - ResNet model was trained. Individual cow recognition tests were conducted on a dataset of 12 864 images of 57 dairy cows. The results showed that the recognition accuracy rates of the back, left trunk, right trunk and buttocks of dairy cows were 96.13%, 96.56%, 96.94% and 93.14%, respectively, which were 2.76, 2.88, 2.92 and 4.25 percentage points higher than those of the original ResNet model. The recognition accuracy rates of the method proposed within the ranges of 10 ~ 20 m, 20 ~ 30 m and 30 ~ 40 m were

收稿日期:2024 - 08 - 05　修回日期:2024 - 09 - 15

基金项目:河北省自然科学基金项目(C2025204216)

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和无损检测研究,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

97.54%, 90.72% and 82.17%, respectively. The research results can provide technical support for intelligent dairy cow breeding.

Key words: cow individual identification; wide-field; deformable convolution; attentional feature fusion; fine-grained classification

0 引言

奶牛个体识别是实现奶牛精准管理的前提[1-3],对奶牛健康监测、自动行为分析、体况评分和产品溯源等方面具有重要作用[4]。

基于计算机视觉的奶牛个体识别具有自动化程度高、成本低、对奶牛无应激等优点[5-7]。部分研究利用奶牛虹膜、口鼻、面部特征识别奶牛个体[8-10],得到了较高的准确率,但此类方法采集图像困难、抗干扰性弱,不利于推广[11]。

荷斯坦奶牛身体花纹更易采集,多项研究利用花纹特征实现了奶牛的个体识别[12]。何东建等[13]在挤奶厅通道采集奶牛背部图像,使用基于改进的 YOLO v3 模型识别奶牛个体,识别准确率为 95.91%。LI 等[14]计算尾根部的 Zernike 不变矩特征识别奶牛个体。HU 等[15]在挤奶厅通道拍摄奶牛侧视图像,提取头部、躯干和腿部 3 个部位的特征并融合,最后使用支持向量机分类奶牛个体。

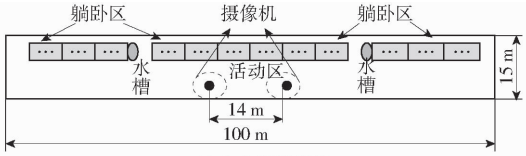
上述研究均为特定场景下奶牛个体识别,摄像机和奶牛间位置不变,奶牛花纹在图像中形变较小,此外,拍摄画面仅出现 1 头奶牛。为尝试在非限制条件下奶牛个体识别,XIAO 等[16]采集非限制性环境下奶牛背部图像,利用 SVM 算法对奶牛识别分类。司永胜等[17]提取奶牛的深度特征和传统特征,然后特征融合进行奶牛个体识别。上述方法是对非限制场景下的多头奶牛进行个体识别,但摄像机和地面的角度接近垂直,拍摄范围较小。TASSINARI 等[18]对范围较大的牛舍内奶牛进行个体识别,准确率为 66%,准确率较低,且研究对象只有 4 头奶牛,样本较少。奶牛大部分时间生活在牛舍中,奶牛在牛舍中个体识别是躺卧率自动计算和行为识别等智慧养殖环节的基础。在牛舍中布设过多摄像机,将降低方法的实用性,不利于推广,因此,类似文献[18]采集广视野图像的方案更有利于牛场的实际应用。在牛舍广视野情况下,如何处理由于奶牛姿态、拍摄角度变化导致的花纹形变较大及部分奶牛花纹相似,仍是奶牛个体识别的难题[19]。

针对上述问题,本文采集广视野下非限制牛舍内的奶牛视频,构建包含多种姿态的奶牛部位图像数据集,提出一种基于 DAM-ResNet 分类网络和广视野下奶牛个体识别方法,以期为奶牛智能化养殖提供技术支持。

1 实验材料

1.1 视频图像采集

视频图像数据采集于河北省保定市弘达牧业股份有限公司,采集时间为 2023 年 6 月 8—12 日 08:00—18:00。共采集 57 头荷斯坦奶牛在牛舍自由活动的图像。牛舍总长 100 m,宽 15 m,为使拍摄视角足够广,覆盖牛舍更大的面积,本文将 2 个广角摄像机(DH-IPC-HFW8449K-ZFR 型,浙江大华技术股份有限公司)安装在牛舍中部,如图 1a 所示,安装高度为 6 m,两两相对的摄像机间距离为 14 m,摄像机俯视角(水平向下倾斜角度)约为 15°。摄像机焦距为3.6 mm,拍摄视频帧率为 25 f/s,分辨率为 2 688 像素×1 520 像素。图 1b 为图像采集现场环境。



(a) 牛舍图像采集



(b) 图像采集现场环境

图 1 数据采集示意图

Fig. 1 Data acquisition diagrams

1.2 图像预处理与数据集制作

将每个摄像机采集的视频逐帧转化成图像并进行去重处理,采用结构相似性指数算法[20]去除相似度大于 0.85 的图像,随机选取 1 500 幅图像,2 个摄像机共选取 3 000 幅图像作为 Mask R-CNN[21-22]模型的数据集,按比例 7∶2∶1 划分为训练集、验证集、测试集。使用 Labelme 标注软件对数据集进行标注。当拍摄距离超过 30 m 时,由于拍摄角度和距离原因,奶牛在图像中易相互遮挡且目标太小,造成人工标注难度较大,因此本文标注距离为 30 m 内的奶牛(设定拍摄距离为摄像机垂直投影到地面的点

与奶牛的直线距离)。在广视野条件下,奶牛姿态和尺度变化较大,即分类的类内差异较大,如果将奶牛整体图像直接作为分类数据,腿部姿态变化等将降低识别准确率。将奶牛图像按特征固定的几个部位分别输入模型,有利于模型聚焦特定的特征,提高识别效果。国内最常见荷斯坦奶牛黑白花纹是良好的个体识别特征。奶牛背部、躯干左右侧面积较大,易采集,具有稳定的花纹特征,多项研究利用上述部位实现了奶牛个体识别[13－14,23]。牛脸也是奶牛个体识别的主要部位,但是在广视野环境下,很难采集到清晰牛脸,本文在摄像机斜视拍摄角度下,可以拍摄到牛脸时,往往是采集到花纹更为清晰的背部。此外,当奶牛远离摄像头方向时,易采集到的是臀部图像。综上,将奶牛图像按照背部、臀部和躯干两侧作为模型的识别依据。标注奶牛的背部、躯干左侧、右侧和臀部共 4 个部位,生成 4 类标签,选取的 4 类部位能包含所有姿态下奶牛图像。图 2 为 Mask R－CNN 奶牛分割结果。



图 2　奶牛部位实例分割结果

Fig. 2　Instance segmentation results of cow body parts

利用分割后的各个部位图像制作奶牛个体识别数据集,包含奶牛在牛舍各种姿态图像。数据集分为奶牛背部、躯干左侧、右侧和臀部图像共 4 个子数据集,每个子数据集包含 57 头奶牛对应的身体部位图像。整个数据集共 12 864 幅图像,每头奶牛背部、躯干左侧和躯干右侧每类图像约 60 幅。监控图像中出现奶牛背部、躯干左侧和躯干右侧部位的几率大于臀部,每头奶牛臀部图像 45 幅左右。将各子数据集均按比例 7∶3划分训练集和验证集。

图 3 为部分奶牛的个体识别数据集示例。同一个子数据集中同 1 头奶牛图像,形变较大且差异明显,部分奶牛距离摄像机较远,在图像中目标较小,此外部分不同奶牛花纹相似,这些都给网络模型提取有效特征增加了难度。

2　研究方法

本文提出的广视野下无约束牛舍奶牛个体识别方法流程如图 4 所示。首先,将牛舍视频帧输入 Mask R－CNN 模型,对奶牛背部、单侧躯干和臀部进行实例分割,然后将奶牛部位图像分别输入对应 DAM－ResNet 模型,通过融入 DCN 和 AFF 的 ResNet 残差网络提取有效特征,利用 MC－Loss 学习到更有区分度的特征指导模型参数更新,最终输出奶牛个体识别结果。

2.1　DAM－ResNet 网络模型设计

2.1.1　ResNet 网络结构

深层网络能够捕捉到奶牛图像中的复杂细节,提取更丰富的特征信息。ResNet 网络利用残差学习思想,逐层计算残差,用更简单清晰的框架训练深度网络,解决深层网络训练中梯度消失和梯度爆炸问题。在残差结构中,残差块将输入直接加到非线性变换的输出中,通过不断堆叠网络层数,获得深度较大的网络模型和更优越的网络性能。本文选择 ResNet34 作为原模型进行训练和调优,模型共有 5 个阶段,16 个残差结构,包括 33 个卷积层和 1 个全连接层。

2.1.2　第二代可变形卷积

针对图像中奶牛花纹形变问题,引入第二代可变形卷积。对不同尺度或感受野自适应调整,更好地提取输入的特征。可变形卷积在标准卷积规则网络采样点上添加可学习的偏移量,实现采样网格的自适应变形。

标准卷积运算过程是使用规则网格 R 对输入特征图 x 进行采样,然后在每个采样点处乘以权值 W 并求和。第一代可变形卷积在传统卷积操作上加入了一个采样点的偏移量 Δp_n,以调整关键元素的采样位置。第二代可变形卷积在其基础上对偏移量添加一个权重系数 Δm_n,增大网络对于卷积操作

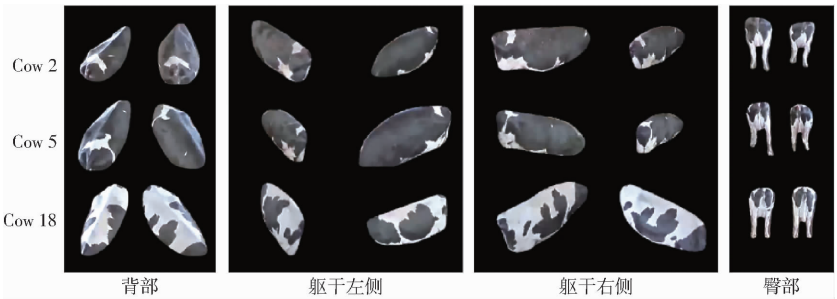


图 3　奶牛个体识别数据集示例

Fig. 3　Example of dairy cow individual identification dataset

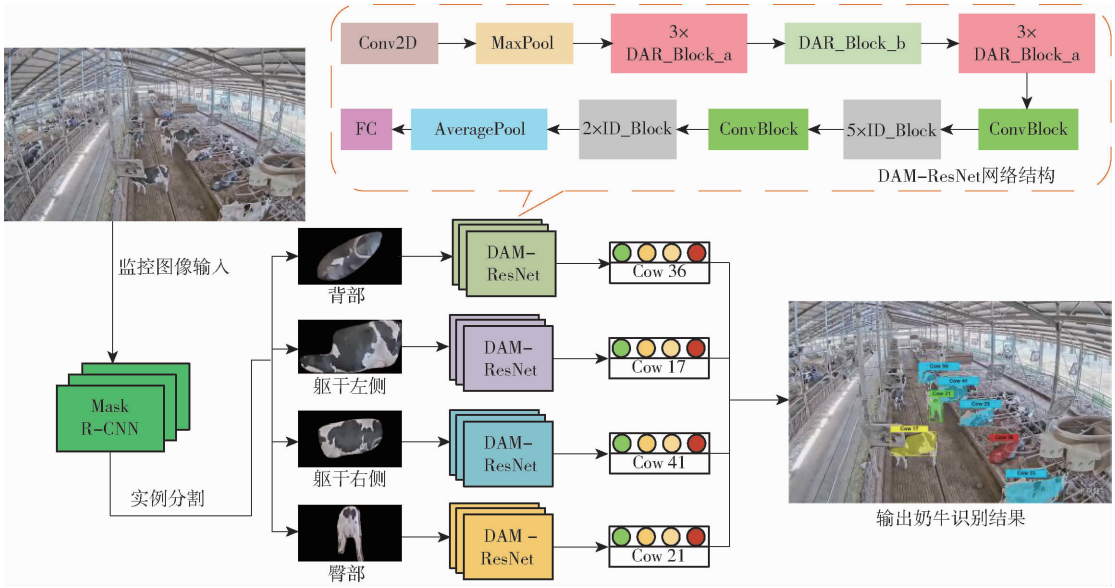


图 4 广视野下奶牛个体识别方法流程

Fig. 4 Flowchart of individual identification method for dairy cows with a wide field of view

的自由度,在学习中降低无关的采样点权重[24],下文中可变形卷积指第二代变形卷积。在卷积运算过程中,对于中心采样点位置p_0,可变形卷积为

$$y(p_0)=\sum_{p_n\in R}\boldsymbol{w}(p_n)x(p_0+p_n+\Delta p_n)\Delta m_n \quad (1)$$

式中 y——输出特征图

R——感受野区域

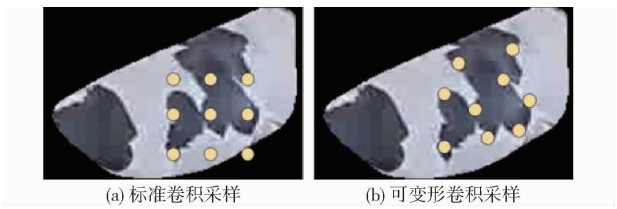
p_n——R 中所有采样位置

$\boldsymbol{w}(p_n)$——p_n位置权重矩阵

Δp_n——采样点偏移量

Δm_n——不同偏移量权重系数

以 3×3 卷积采样为例,可变形卷积与标准卷积在奶牛图像采样方式如图 5 所示。可变形卷积通过增加可训练的偏移量,对花纹形变有适应能力,增加偏移量后自适应感受野不再是方形,在采样时可更贴近奶牛花纹形状,提高模型提取能力。



(a) 标准卷积采样 (b) 可变形卷积采样

图 5 卷积采样对比

Fig. 5 Comparison of convolutional sampling

2.1.3 注意力特征融合

图像中奶牛尺度不一,目标较小的奶牛特征提取难度较大。特征融合方法能更好地提取不同尺度的特征,传统特征融合是通过组合不同卷积核、分组或网络层的特征实现,组合时不考虑是否适合特定对象,这会削弱目标区域的精细特征信息,影响网络模型的识别效果[25-26]。

注意力特征融合[27](Attentional feature fusion, AFF)以上下文规模感知的方式动态地、自适应地融合接收到的特征。为解决尺度变化导致的小目标识别困难问题,AFF 模块聚合来自不同尺度、不同接受野的上下文信息,通过沿通道维度聚合多尺度上下文信息,解决特征融合不同尺度特征不一致的问题。如图 6 所示,H 和 W 表示每个特征通道高度和宽度,C 表示特征通道个数,X 和 Y 表示两个不同特征映射。AFF 模块在关注全局通道的基础上增加分支提取局部特征,采用逐点卷积获取多尺度特征上下文信息,保留和突出底层特征中的精细细节,并在关注模块中完成局部特征与全局特征的融合。

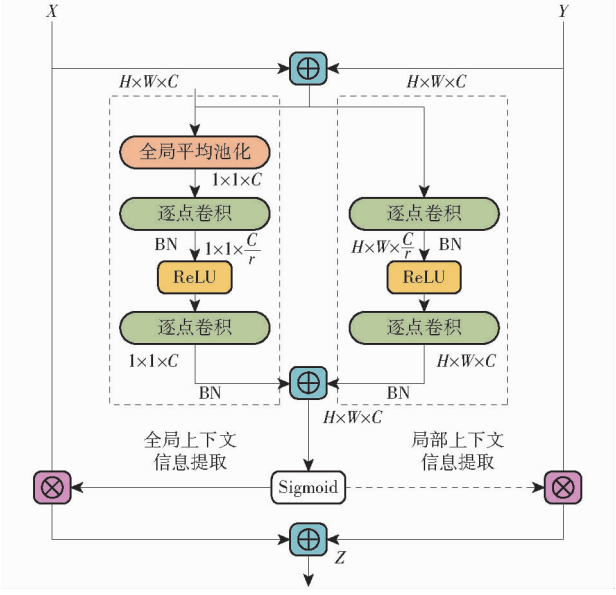


图 6 AFF 注意力特征融合模块

Fig. 6 AFF attention feature fusion module

全局通道上下文信息 $G(X)$ 和局部通道上下文信息 $L(X)$ 的计算式为

$$G(X)=\mathrm{BN}(\mathrm{PWConv}_2(\delta(\mathrm{BN}(\mathrm{PWConv}_1(GAP(X))))))\tag{2}$$

$$L(X)=\mathrm{BN}(\mathrm{PWConv}_2(\delta(\mathrm{BN}(\mathrm{PWConv}_1(X)))))\tag{3}$$

式中　GAP——全局平均池化

BN——批归一化

δ——ReLU 激活函数

PWConv_1、PWConv_2——逐点卷积

对于给定的 2 个特征映射 X、$Y\in R^{H\times W\times C}$，注意力特征融合可以表示为

$$Z=M(X\oplus Y)\otimes X+(1-M(X\oplus Y))\otimes Y\tag{4}$$

式中　Z——加权特征融合后的输出

$\oplus$——广播加法　　$\otimes$——元素乘法

$M(X\oplus Y)$——AFF 生成的融合信道权值，用于对输入的特征映射进行加权

2.1.4　Mutual－Channel Loss 损失函数

针对类内差异性和类间相似性，引入细粒度分类任务的互通道损失函数[28]（Mutual－Channel Loss，MC－Loss），其由判别性组件和多样性组件 2 个协同作用的组件组成。判别性组件通过约束每个类别对应的特征通道挖掘对应其类别信息的细粒度特征区域，多样性组件使同一组特征通道内的多个特征互斥以学习不同特征区域。

MC－Loss 通过超参数 μ 将互通道损失函数与交叉熵损失函数进行聚合，形成本文模型最终的损失函数

$$Loss=L_{CE}+\mu L_{MC}\tag{5}$$

式中　L_{CE}——传统的交叉熵损失函数，使网络专注于全局性判别区域

L_{MC}——互通道损失函数，使网络专注于局部判别区域

将卷积神经网络提取到的深度特征图表示为 $F\in R^{C\times W\times H}$，$L_{MC}$ 计算式为

$$L_{MC}(F)=L_{CE}\left(y',\underbrace{\frac{\left[\mathrm{e}^{g(F_0)},\mathrm{e}^{g(F_1)},\cdots,\mathrm{e}^{g(F_{c-1})}\right]^{\mathrm{T}}}{\sum_{i=0}^{c-1}\mathrm{e}^{g(F_i)}}}_{\mathrm{Softmax}}\right)+\lambda\frac{1}{c}\sum_{i=0}^{c-1}h(F_i)\tag{6}$$

式中　F_i——第 i 个类别一组特征图

y'——样本期望输出

c——分类任务的类别数

λ——标量超参数

Softmax——归一化指数函数

$g(F)$——判别性组件函数

$h(F)$——多样性组件函数

2.1.5　DAM－ResNet 模型

提出的 DAM－ResNet 模型整体结构如图 7 所示。Stage 表示 ResNet34 中由卷积层组成的一组块，它们输出相同空间分辨率特征，但通道数不同。将 Stage 2 和 Stage 3 中 2 个基础残差模块 ID_Block 和 ConvBlock 中的传统 3×3 卷积替换成可变形卷

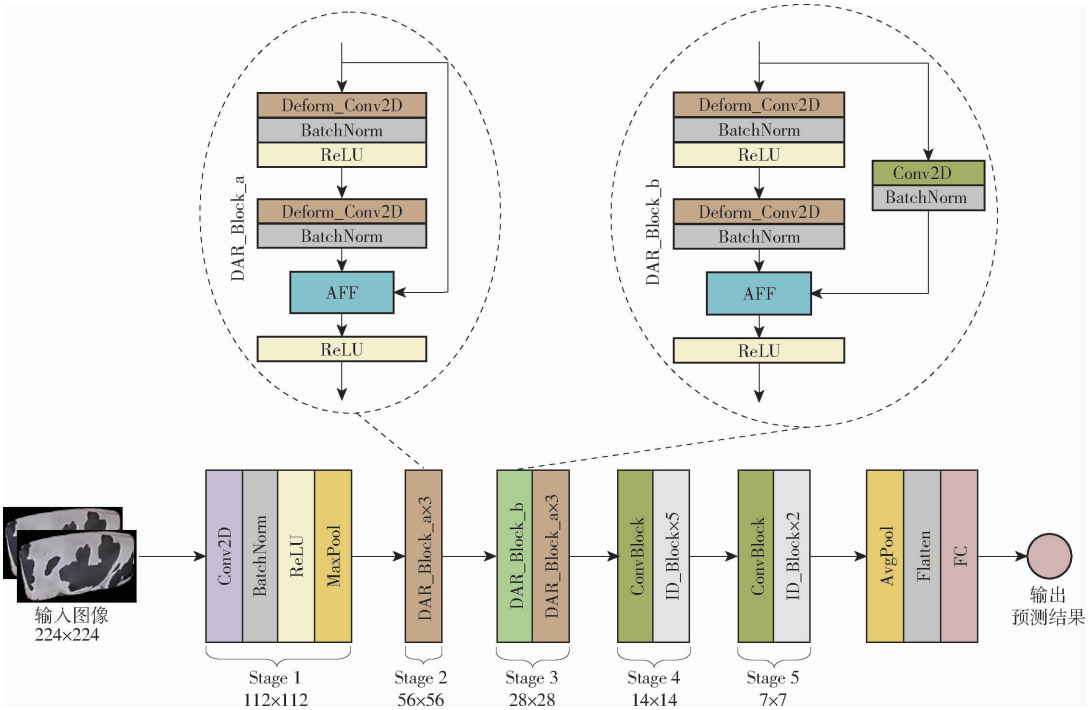


图 7　DAM－ResNet 结构图

Fig. 7　DAM－ResNet network architecture diagram

积,并在每个可变形卷积层后增加批归一化(BatchNorm)层,以加快模型的收敛速度,减轻梯度消失情况。引入 AFF 模块,通过融合不同网络层的语义信息,生成网络映射和残差的融合权值。得到 2 种新的网络模块 DAR_Block_a 和 DAR_Block_b。网络输入为 224 像素×224 像素的 3 通道 RGB 奶牛图像。图像进入网络首先经过一个步距为 2,卷积核为 7×7 的卷积层,再经过批归一化处理、ReLU 激活函数以及 1 个 3×3 最大池化层,输入图像被转换为 56 像素×56 像素特征图像。Stage 2~5 为添加了残差结构的卷积块,4 个卷积模块组进行浅层至深层的特征提取。Stage 2 和 Stage 3 中新的残差模块通过 DCN 和 AFF 模块,针对不同尺寸和形状的花纹自适应地提取有效特征。随后经过连续融入 AFF 残差卷积块进行特征提取,得到尺寸为 7×7、通道数为 512 的特征图。将这组特征图分为 2 个分支计算损失,一个分支直接输入 MC-Loss 模块计算互通道损失,另一分支通过全连接层后计算交叉熵损失。奶牛数据集有 57 种类别,将 512 通道特征图分为 57 组,其中 56 类每类都用 9 个通道表征,另外一类用 8 个通道表征,以计算互通道损失。通过超参数 μ 将互通道损失与交叉熵损失聚合计算损失,指导模型参数的更新。最终将数据依次输入至展平层和全连接层整合卷积层的特征,通过 Softmax 分类器输出对应奶牛个体识别结果。

2.2 实验设置与评价指标

实验运行平台为 Intel Xeon Silver 4208 CPU,NVIDIA GeForce RTX 3090 GPU,软件环境为 Python 3.8,Windows 10 操作系统,使用 Pytorch 深度学习框架构建模型。在模型训练参数设置方面,训练图像尺寸为 224 像素×224 像素,批量大小(Batch size)为 32,采用 Adam 优化器,学习率设为 0.001。采用识别准确率、精确率和召回率作为分类网络模型评价指标[17]。

3 结果与分析

3.1 不同模型对比

使用本文数据集在相同实验条件下对 Alexnet、Vgg16、MobileNetV3、ResNet34、DAM-ResNet 模型进行了识别准确率对比实验,结果如表 1 所示。对比前 4 种基础网络模型,ResNet34 模型在奶牛 4 个部位的识别准确率均高于其他模型,ResNet 网络使用残差连接可以保留原始特征,在处理本文奶牛数据集复杂的图像特征时增强了网络的特征提取能力,从而实现了较高的准确率。本文模型在 ResNet34 上进行改进,在奶牛 4 个部位均取得了最高的识别准确率,其中臀部准确率最低,主要原因是大部分臀部图像目标较小且相似度较高,同时尾巴摆动也造成特征图不稳定。

表 1 不同模型对比实验结果

Tab. 1 Comparative experiment results of different models %

模型	背部	躯干左侧	躯干右侧	臀部
Alexnet	90.15	91.38	91.73	83.29
Vgg16	92.08	92.21	92.68	85.14
MobileNetV3	91.26	92.42	93.23	81.56
ResNet34	93.37	93.68	94.02	88.89
DAM-ResNet	96.13	96.56	96.94	93.14

3.2 可变形卷积效果

Grad-CAM 热力图[29]可以对网络模型进行可视化分析,直观展示网络模型在进行预测分类时关注的图像区域。本文使用热力图分析可变形卷积模块对奶牛花纹变形的识别效果。如图 8 所示,图 8a 为5 号奶牛的 2 幅躯干左侧图像,图 8b、8c 第 1 行分别为 ResNet 和 DCN-ResNet 2 个模型识别图 8a 图像 1 的热力图,2 个模型关注区域存在差异;ResNet 和 DCN-ResNet 均关注到奶牛图像右下角处的花纹,而 DCN-ResNet 同时关注到右上角的花纹。图 8b、8c 第 2 行分别为 2 个模型识别图 8a 图像 2 时所关注的区域,ResNet 主要关注区域主要在奶牛腹部花纹处,将其错误识别为 13 号奶牛,DCN-ResNet 关注腹部、背部和躯干后侧花纹,识别结果正确。可视化结果说明 DCN-ResNet 通过可变形卷积学习空间偏移量改变网格采样点的位置,当奶牛花纹有形变时,DCN-ResNet 仍可自适应花纹大小和形状,获取了更多有效感受野,从而加强网络对关键特征信息的关注。

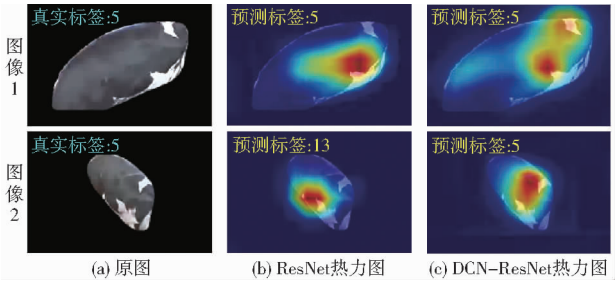


图 8 可变形卷积模块效果可视化对比

Fig. 8 Visualization comparison of deformable convolution module effects

3.3 AFF 模块效果

图 9a 为远距离奶牛臀部图像,尺度较小,清晰度较低。图 9b、9c 为加入 AFF 模块前、后 2 个模型的热力图。加入 AFF 后,热力覆盖奶牛整个花纹,原因是 AFF 模块引入了上下文相关的注意机制,通

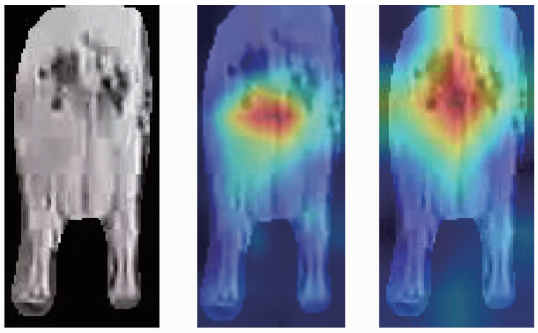


图 9　AFF 模块效果可视化对比

Fig. 9　Visualization comparison of AFF module effects

过聚合来自不同尺度对象的上下文信息，提取全局和局部特征，兼顾浅层纹理和深层语义信息，提升了对小目标奶牛的精细识别效果。

累计匹配曲线（CMC）可以综合反映分类器性能，纵坐标表示匹配度最高的前 k 幅的正确命中率，曲线越高表示模型识别性能越好。图 10 为添加 AFF 模块前、后模型对 4 类部位数据集个体识别准确率累计曲线。增加 AFF 后，模型在 4 类部位数据集的曲线都有提高。曲线 Rank－1 值表征识别准确率，加入 AFF 模块后，模型在 4 类部位数据集的识别准确率分别提升 0.48、1.02、0.74、1.32 个百分点。从多个召回顺序上比较，增加 AFF 模块后，模型在背部数据集上识别率提升相对较小，其原因是 AFF 模块对模型小目标识别能力提升明显。在对远距离奶牛实例分割时，得到大部分为臀部或躯干图像，即小尺度背部图像很少，因此 AFF 模块对背部数据集提升较小。

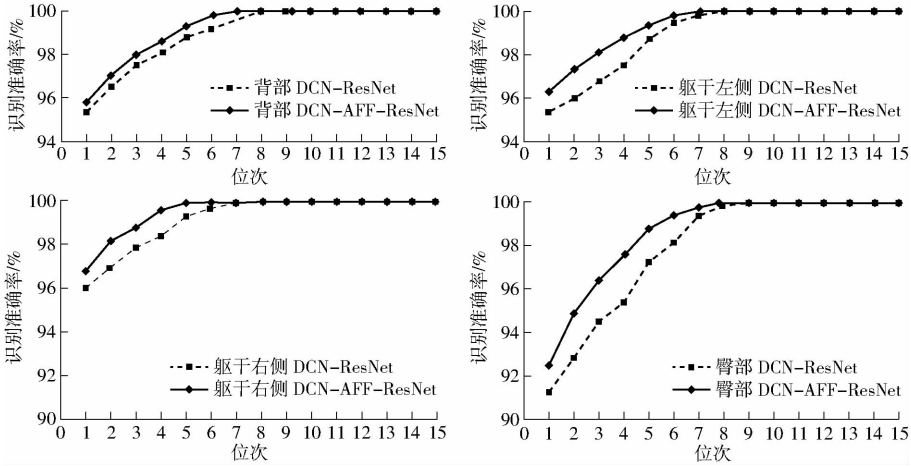


图 10　加入 AFF 模块前后对比曲线

Fig. 10　Comparison curves before and after adding AFF module

3.4　消融实验

为量化分析模型各部分效果，开展消融实验。在 ResNet34 原模型中分别引入 DCN 模块、AFF 模块以及 MC－Loss，不同模型准确率结果如表 2 所示。相比原模型，DCN－ResNet 在采用可变形卷积后，背部、躯干左侧、躯干右侧、臀部准确率分别提升 1.98、1.64、2.06、2.36 个百分点，DCN 改变了卷积操作的采样方式，使模型在特征提取阶段能够有效地学习奶牛花纹的形状特征，自适应花纹的变化并且提高了特征提取能力，使模型性能提升明显。DCN－AFF－ResNet 在 DCN－ResNet 上添加了 AFF 模块，通过全局和局部特征的融合来加强对奶牛图像形态大小差异的关注，增强了模型学习不同尺度图像特征的能力，使得模型对于一些细小特征的变化更加敏感。DCN－MC－ResNet 在 DCN－ResNet 上引入 MC－Loss，在特征分类阶段进一步优化了特征空间的分布，使得不同类别的特征更加易于区分。4 个奶牛部位的识别准确率均有提升，其中图像相似度较高的奶牛臀部的准确率提升 1.06 个百分点，较其他 3 个部位提升较为明显。与其他 3 个部位相比，奶牛臀部图像差异较小，这说明模型对花纹相似、个体差异较小的奶牛识别表现更好。

表 2　消融实验结果

Tab. 2　Ablation experiment results　%

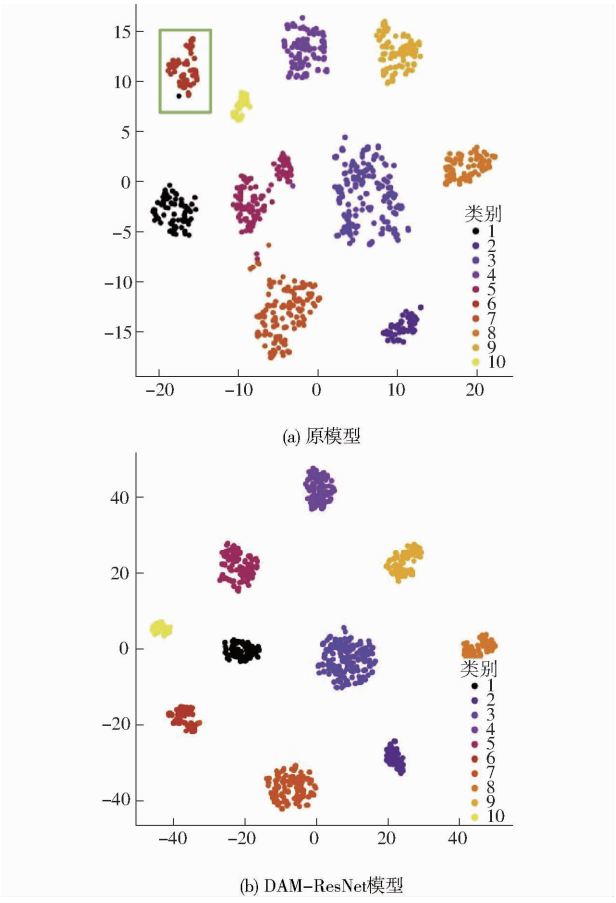
模型	DCN	AFF	MC－Loss	背部	躯干左侧	躯干右侧	臀部
ResNet34	×	×	×	93.37	93.68	94.02	88.89
DCN－ResNet	√	×	×	95.35	95.32	96.08	91.25
DCN－AFF－ResNet	√	√	×	95.83	96.34	96.82	92.57
DCN－MC－ResNet	√	×	√	95.89	95.94	96.56	92.31
DAM－ResNet	√	√	√	96.13	96.56	96.94	93.14

注：√表示采用此模块，×表示未采用此模块。

将上述 3 种改进方法同时融入原模型形成最终的 DAM－ResNet 模型，4 个奶牛部位准确率均最高，分别为 96.13%、96.56%、96.94%、93.14%。最终平均准确率为 95.86%。相较于原模型，背部、躯干左侧、躯干右侧、臀部识别准确率分别提升

2.76、2.88、2.92、4.25 个百分点。DCN、AFF 和 MC-Loss 结合使用时,通过特征提取、融合和判别的全方位优化,使得模型在复杂特征和图像变形时表现出显著的性能提升,充分发挥了各模块的优势,达到综合更好的识别效果。

为展示 DAM-ResNet 模型相对于原 ResNet 模型的改进效果,采用 t-SNE 降维算法对高维特征降维,进行可视化展示。从数据集中随机选取 10 个类别数据,使用原模型和 DAM-ResNet 模型分别对提取的特征进行可视化。如图 11 所示,尽管原模型也取得了较好的分类效果,但出现了部分错误,极个别黑色点分类到了棕红色点类中(图 11a)。本文模型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,最终分类效果更优。



(a) 原模型

(b) DAM-ResNet模型

图 11 模型特征降维效果可视化

Fig. 11 Visualization of feature dimensionality reduction effects

3.5 拍摄距离对奶牛识别影响

为研究拍摄距离对奶牛个体识别的影响,本文补充标注 30~40 m 范围内能够人工识别的奶牛图像,并重新训练分割模型。随机选取 100 幅图像利用新训练的分割模型进行实例分割,最后使用本文模型对其进行个体识别。按照拍摄距离分段统计识别结果,如表 3 所示。

表 3 不同拍摄距离下奶牛识别结果

Tab. 3 Recognition results of dairy cows at different shooting distances

拍摄距离/m	准确率/%	精确率/%	召回率/%
0~10	99.63	99.38	99.14
10~20	97.54	97.67	97.02
20~30	90.72	91.14	90.57
30~40	82.17	84.11	79.73

随着拍摄距离增加,识别准确率、精确率和召回率均明显下降。拍摄距离在 20 m 内奶牛的各项指标均在 97% 以上,此范围内模型可准确识别。在 20~30 m 范围内,随着奶牛目标在图像上逐渐变小,提取特征难度增大,识别准确率、精确率和召回率降低到 90% 左右。随着拍摄距离增大到 30~40 m 范围,识别准确率下降到 82% 左右。为直观地展示识别效果,随机选取 1 幅奶牛图像进行分割和个体识别,并给出模型 Softmax 层输出预测标签类别的概率。如图 12 所示,图 12b 中第 1 行前 4 头奶牛,对应图 12a 中右侧躺卧区由近到远依次躺卧的奶牛(蓝色分割框),属于躯干右侧数据集,均以高预测概率正确识别。图 12b 第 1 行和第 2 行最右侧距离分别为 29 m 和 37 m 的奶牛,对应图 12a 中右侧远处蓝色分割框中奶牛。模型将 29 m 处奶牛 Cow 11 错误识别为 Cow 5,识别错误的原因是由于奶牛距离较远导致目标较小,且花纹图案过少,提取特征难度较大,37 m 处的奶牛尽管正确识别,但预测概率较低。图 12a 右侧示例的 27 m 和 42 m 的 2 头奶牛由于距离过远且图像模糊,在实例分割阶段未被成功分割。类似图像由于人工也难以识别,因此在模型训练标记阶段无法使用。

遮挡问题是个体识别一个难点。如图 12b 中第 1 行的奶牛,多头奶牛被栏杆遮挡,但本文模型仍以较高预测概率正确识别。类似地,第 2 行 Cow 31 奶牛(对应图 12a 中左侧最远端黄色分割框中奶牛)被风扇遮挡严重,模型仍正确识别,说明模型在局部遮挡情况下鲁棒性较好。

本文模型在实验环境 20 m 范围内识别准确率、精确率和召回率都超过 97%,精度较高,在 20~30 m 范围内识别准确率、精确率和召回率超过 90%。适当增大镜头的焦距可以使远距离的奶牛更清晰,但会缩小拍摄视场。规模化奶牛场牛舍长度一般为 100 m[30-31],因此本文方法难以利用一个摄像机实现整个牛舍范围的奶牛个体识别。下一步工作拟在本研究基础上,通过多个摄像机实现整个牛舍的覆盖和奶牛个体识别。

3.6 与现有研究方法对比

与现有代表性相关研究结果对比如表4所示。

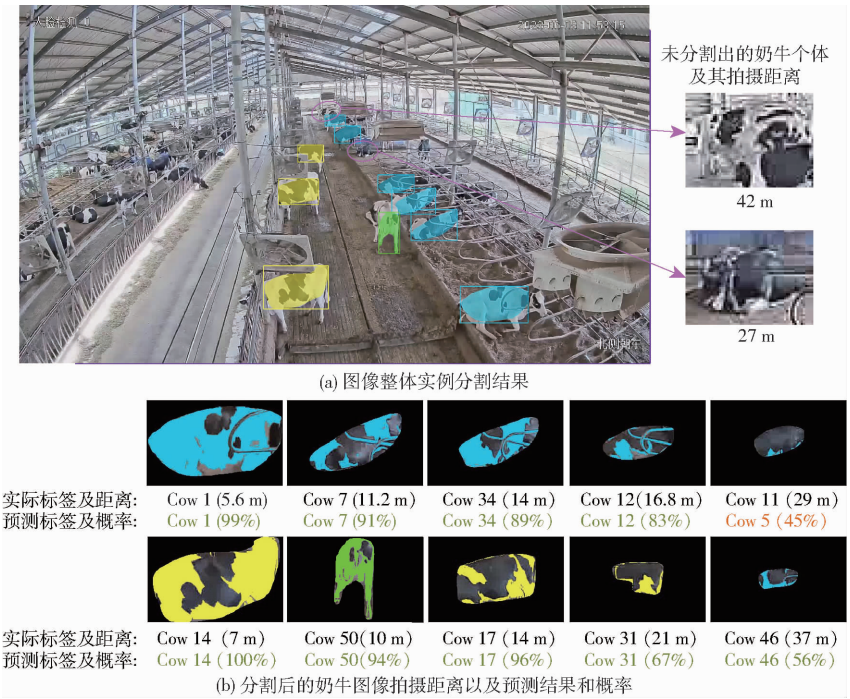


图 12　不同拍摄距离下奶牛识别结果

Fig. 12　Cow recognition results at different shooting distances

表 4　与现有方法结果对比

Tab. 4　Comparison with results of existing methods

识别方法	识别区域	奶牛数量	识别准确率/%
YOLO v3[18]	臀部左右侧	4	66.00
RetinaNet + CNN[32]	背部	46	93.75
ShuffleNet－Triplet[33]	奶牛整体	87	82.93
Mask－RCNN + DAM－ResNet(本文)	背部、躯干、臀部	57	95.86

文献[18]利用 YOLO 算法提取 4 头奶牛臀部的左侧和右侧图像特征进行个体识别，识别准确率为 66%，识别准确率较低。和文献[18]类似，文献[32]利用无人机获取奶牛背部图像进行个体识别，识别准确率为 93.75%。奶牛背部特征较为稳定，是理想的奶牛个体识别特征。然而，如果获取广视野的奶牛背部图像，需要摄像机与奶牛背部之间有较远的拍摄距离，这在牛舍环境下难以实现。文献[33]主要聚焦于模型轻量化和群体中进入新奶牛时模型再训练问题，基于 ShuffleNet－Triplet 模型对整体奶牛图像进行识别，识别准确率为 82.93%。由于输入图像中有奶牛腿部等干扰，各类差异较大，影响了识别准确率。本文方法较好地降低了牛舍广视野条件下奶牛个体识别中各类差异较大的影响，实现了较高的识别准确率。

4　结论

(1)提出了先通过 Mask R－CNN 分割不同奶牛部位，再利用改进的 DAM－ResNet 模型对奶牛进行识别的方法，实现了较为准确的奶牛个体识别。

(2)DAM－ResNet 模型通过引入第二代可变形卷积，提高了模型对奶牛花纹变化的识别能力，新增 AFF 注意力融合模块，加强了模型对小目标的识别能力，同时结合细粒度损失函数 MC－Loss，增强了模型对相似奶牛的分类效果。DAM－ResNet 模型对奶牛背部、躯干左侧、躯干右侧、臀部图像识别准确率分别达到 96.13%、96.56%、96.94%、93.14%。优于其他对比模型，且相较原模型识别准确率分别提高 2.76、2.88、2.92、4.25 个百分点。

(3)本文方法在牛舍环境下 20 m 范围内识别准确率、精确率和召回率均超过 97%，可以实现可靠的准确个体识别，在 20～30 m 范围内识别准确率超过 90%。本文方法对遮挡问题和花纹变形具有较高的鲁棒性，为现代牛场的智能奶牛个体识别提供了新思路。

参　考　文　献

[1]　BERCKMANS D. Precision livestock farming technologies for welfare management in intensive livestock systems[J]. Rev. Sci. Tech., 2014, 33(1): 189－196.

[2]　夏雪，侍啸，柴秀娟. 人工智能驱动智慧奶牛养殖的思考与实践[J]. 中国乳业，2020(8): 5－9.
XIA Xue, SHI Xiao, CHAI Xiujuan. Thinking and practice of intelligent cow breeding driven by artifical intelligence[J]. China

Dairy,2020(8): 5 - 9. (in Chinese)
[3] 司永胜,肖坚星,刘刚,等. 基于 MSRCP 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 recognition method of lying cows based on MSRCP and improved YOLO v4 model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 243 - 250, 262. (in Chinese)
[4] 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)
[5] ALVAREZ J R, ARROQUI M, MANGUDO P, et al. Body condition estimation on cows from depth images usingconvolutional neural networks[J]. Computers and Electronics in Agriculture, 2018, 155: 12 - 22.
[6] 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.
[7] WENG Z, MENG F, LIU S, et al. Cattle face recognition based on atwo-branch convolutional neural network[J]. Computers and Electronics in Agriculture, 2022, 196: 106871.
[8] SUN S, YANG S, ZHAO L. Noncooperative bovine iris recognition via SIFT[J]. Neurocomputing, 2013, 120: 310 - 317.
[9] KUMAR S ,PANDEY A ,SATWIK R S K , et al. Deep learning framework for recognition of cattle using muzzle point image pattern[J]. Measurement,2018,116: 1 - 17.
[10] 蔡骋,宋肖肖,何进荣. 基于计算机视觉的牛脸轮廓提取算法及实现[J]. 农业工程学报,2017,33(11):171 - 177.
CAI Cheng,SONG Xiaoxiao,HE Jinrong. Algorithm and realization for cattle face contour extraction based on computer vision [J]. Transactions of the CSAE,2017,33(11):171 - 177. (in Chinese)
[11] 许贝贝,王文生,郭雷风,等. 基于非接触式的牛只身份识别研究进展与展望[J]. 中国农业科技导报,2020,22(7):79 - 89.
XU Beibei,WANG Wensheng,GUO Leifeng,et al. A review and future prospects on cattle recognition based on non-contact identification[J]. Journal of Agricultural Science and Technology,2020,22(7):79 - 89. (in Chinese)
[12] DU Y, KOU Y, LI B, et al. Individual identification of dairy cows based on deep learning and feature fusion[J]. Animal Science Journal, 2022, 93(1): 13789.
[13] 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 method of milking cows based on improved YOLO v3 model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020,51(4):250 - 260. (in Chinese)
[14] LI W, JI Z, WANG L, et al. Automatic individual identification of Holstein dairy cows using tailhead images[J]. Computers and Electronics in Agriculture, 2017, 142: 622 - 631.
[15] HU H, DAI B, SHEN W, et al. Cow identification based on fusion of deep parts features[J]. Biosystems Engineering, 2020, 192: 245 - 256.
[16] 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.
[17] 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)
[18] TASSINARI P, BOVO M, BENNI S, et al. A computer vision approach based on deep learning for the detection of dairy cows in free stall barn[J]. Computers and Electronics in Agriculture, 2021, 182: 106030.
[19] ZHAO J M, LIAN Q S. Compact loss for visual identification of cattle in the wild[J]. Computers and Electronics in Agriculture, 2022, 195: 106784.
[20] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600 - 612.
[21] HUERTAS-TATO J, MARTÍN A, FIERREZ J, et al. Fusing CNNs and statistical indicators to improve image classification [J]. Information Fusion, 2022, 79: 174 - 187.
[22] XU B, WANG W, FALZON G, et al. Automated cattle counting using Mask R - CNN in quadcopter vision system[J]. Computers and Electronics in Agriculture, 2020, 171: 105300.
[23] ZHANG R, JI J, ZHAO K, et al. A cascaded individual cow identification method based on DeepOtsu and EfficientNet[J]. Agriculture, 2023, 13(2): 279.
[24] ZHU X, HU H, LIN S, et al. Deformable convnets v2:more deformable, better results[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,2019: 9308 - 9316.
[25] LI X, WANG W, HU X, et al. Selective kernel networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,2019: 510 - 519.

(下转第 602 页)

learning[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2017, 31(1):4278 - 4284.
[28] BREDIN H. TristouNet: triplet loss for speaker turn embedding[C]// Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). IEEE,2017:5430 - 5434
[29] 宋怀波,华志新,马宝玲,等. 基于 SimCC-ShuffleNetV2 的轻量化奶牛关键点检测方法[J]. 农业机械学报,2023,54(10):275 - 281,363.
SONG Huaibo, HUA Zhixin, MA Baoling, et al. Lightweight keypoint detection method of dairy cow based on SimCC-ShuffleNetV2[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10): 275 - 281, 363. (in Chinese)
[30] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: inverted residuals and linear bottlenecks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510 - 4520.
[31] LIU S, DENG W. Very deep convolutional neuralnetwork based image classification using small training sample size[C]// 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR). IEEE, 2015: 730 - 734.
[32] HAN K, WANG Y, CHEN H, et al. A survey on vision transformer[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022,45(1):87 - 110.
[33] LIU Z, LIN Y, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021:10012 - 10022.
[34] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:770 - 778.

(上接第 590 页)

[26] ZHANG H, WU C, ZHANG Z, et al. Resnest: split-attention networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 2736 - 2746.
[27] DAI Y, GIESEKE F, OEHMCKE S, et al. Attentional feature fusion[C]// Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021: 3560 - 3569.
[28] 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.
[29] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization[C]// Proceedings of the IEEE International Conference on Computer Vision,2017: 618 - 626.
[30] 冯璐,李奇峰,丁露雨,等. 自然通风牛舍环境监测传感器布局策略对比研究[J]. 家畜生态学报,2022,43(2):74 - 81.
FENG Lu, LI Qifeng, DING Luyu, et al. Comparative research on sensor layout strategy of natural ventilation cowshed environment monitoring[J]. Journal of Livestock Ecology, 2022, 43(2): 74 - 81. (in Chinese)
[31] 刘爱瑜,冯曼,李永亮,等. 卷帘舍环境参数与奶牛泌乳性能及行为的相关性分析[J]. 畜牧兽医学报,2023,54(2):617 - 629.
LIU Aiyu, FENG Man, LI Yongliang, et al. Correlation analysis between environmental parameters of rolling shed and lactation performance and behavior of dairy cow[J]. Acta Veterinaria et Zootechnica Sinica, 2023, 54(2): 617 - 629. (in Chinese)
[32] ANDREW W, GAO J, MULLAN S, et al. Visual identification of individual Holstein-Friesian cattle via deep metric learning [J]. Computers and Electronics in Agriculture, 2021, 185: 106133.
[33] WANG Y, XU X, WANG Z, et al. ShuffleNet - Triplet: a lightweight RE-identification network for dairy cows in natural scenes[J]. Computers and Electronics in Agriculture, 2023, 205: 107632.