doi:10.6041/j.issn.1000-1298.2023.06.030

基于改进 YOLO v5s 的轻量级奶牛体况评分方法

黄小平[1,2]　冯　涛[1]　郭阳阳[1,2]　梁　栋[1,2]

(1.安徽大学互联网学院,合肥 230039;2.农业生态大数据分析与应用技术国家地方联合工程研究中心,合肥 230039)

摘要:奶牛体况评分是评价奶牛产能与体态健康的重要指标。目前,随着现代化牧场的发展,智能检测技术已被应用于奶牛精准养殖中。针对目前检测算法的参数量多、计算量大等问题,以 YOLO v5s 为基础,提出了一种改进的轻量级奶牛体况评分模型(YOLO-MCE)。首先,通过 2D 摄像机在奶牛挤奶通道处采集奶牛尾部图像并构建奶牛 BCS 数据集。其次,在 MobileNetV3 网络中融入坐标注意力机制(Coordinate attention,CA)构建 M3CA 网络。将 YOLO v5s 的主干网络替换为 M3CA 网络,在降低模型复杂度的同时,使得网络特征提取时更关注于牛尾区域的位置和空间信息,从而提高了运动模糊场景下的检测精度。YOLO v5s 预测层采用 EIoU Loss 损失函数,优化了目标边界框回归收敛速度,生成定位精准的预测边界框,进而提高了模型检测精度。试验结果表明,改进的 YOLO v5s 模型的检测精度为 93.4%,召回率为 85.5%,mAP@0.5 为 91.4%,计算量为 2.0×10^9,模型内存占用量仅为 2.28 MB。相较原始 YOLO v5s 模型,其计算量降低 87.3%,模型内存占用量减少 83.4%,在保证模型复杂度较低与实时性较高的情况下,实现了奶牛体况的高效评分。此外,改进的 YOLO v5s 模型的整体性能优于 Faster R-CNN、SDD 和 YOLO v3 目标检测模型。本研究为奶牛体况评分商业化提供理论基础和研究思路,为奶牛养殖业提供了智能化解决方案。

关键词:奶牛体况评分;轻量化;目标检测;注意力机制

中图分类号:TP391.4　　**文献标识码**:A　　**文章编号**:1000-1298(2023)06-0287-10　　**OSID**:

Lightweight Dairy Cow Body Condition Scoring Method Based on Improved YOLO v5s

HUANG Xiaoping[1,2]　FENG Tao[1]　GUO Yangyang[1,2]　LIANG Dong[1,2]

(1. *School of Internet*, *Anhui University*, *Hefei* 230039, *China*

2. *National Engineering Research Center for Agro-Ecological Big Data Analysis and Application*, *Hefei* 230039, *China*)

Abstract: Cow body condition score is an important indicator to evaluate the productivity and physical health of cows. At present, with the development of modern farming, intelligent detection technology has been applied to precision farming of dairy cows. In view of the problems of large number of parameters and large calculation of the current detection algorithm, an improved lightweight cow body condition scoring model (YOLO-MCE) was proposed based on YOLO v5s. Firstly, a 2D camera was used to acquire the cow tail images at the cow milking lane, and those images were filtered to obtain the final BCS dataset. Secondly, the coordinate attention (CA) mechanism was integrated into the MobileNetV3 network to build the M3CA network, which was used to replace the YOLO v5s backbone network to reduce the complexity of the model, and make the network feature extraction pay more attention to the location and spatial information of the cow tail area. Finally, the EIoU Loss function was used in the prediction layer of YOLO v5s to optimize the regression convergence speed of the target bounding box and generate a prediction bounding box with accurate positioning. The experimental results showed that the improved YOLO v5s model had a detection precision of 93.4%, a recall rate of 85.5%, an mAP@0.5 of 91.4%, a FLOPs of 2.0×10^9, and a model size of 2.28 MB. Compared with the original YOLO v5s model, the FLOPs and model size of YOLO-MCE were reduced by 87.3% and 83.4%, respectively,

收稿日期:2023-02-21　修回日期:2023-04-05

基金项目:安徽省自然科学基金项目(1908085QF284)和安徽省教育厅自然科学基金项目(KJ2021A0024)

作者简介:黄小平(1984—),男,讲师,博士,主要从事畜牧养殖信息化研究,E-mail:hxping@mail.ustc.edu.cn

通信作者:郭阳阳(1991—),男,讲师,博士,主要从事智能化监测技术研究,E-mail:guoyangyang113529@ahu.edu.cn

which further showed that the proposed method can achieve efficient scoring of cow body conditions under the condition of low model complexity and high real-time performance. In addition, the overall performance of the improved YOLO v5s model was superior to that of the Fast R-CNN, SDD and YOLO v3 object detection models. The research result can provide a theoretical basis and research ideas for the commercialization of dairy cow body condition scoring, and offer a research direction for the application of intelligent algorithms.

Key words: body condition score; lightweight; target detection; attentional mechanism

0 引言

近年来,畜牧养殖业逐渐成为我国农业的支撑产业之一,规模化养殖逐渐取代了小型零散型养殖方式[1]。在规模化养殖中,动物的身体健康信息不仅有利于智慧养殖、提高动物福利,还有利于提高产品的品质。其中,奶牛作为畜牧养殖中的代表,其体态健康状况与牛奶品质直接相关[2]。体况评分(Body condition score, BCS)作为评价动物体态健康的重要措施,是对奶牛脂肪沉积程度和能量代谢状况的综合评估,也是检验和评价牧场饲养管理水平的一项实用指标[3-5]。

传统的奶牛体态监测往往通过人工观察测量的手段,易引起牛只应激反应,且主观性强、误差大,无法满足规模化养殖实时高效的发展需求[6-7]。随着深度学习技术的发展,其在视频图像处理过程中实现了快速、批量化处理,可满足现代化牧场的需求[8]。目前,在规模化养殖中,将基于深度学习框架以及衍生的改进方法应用于实际饲养环境中已成为研究趋势[9-10]。在动物体况评价研究中常采用3D或2D传感器设备来获取奶牛躯体图像,通过深度学习技术提取与体态相关区域特征,并构建和生成评价模型。文献[11]在通道上方安装 Kinect v2来获取奶牛行走的3D图像,采用 SqueezeNet 模型来估计 BCS,结果表明,在实际 BCS 的0.50点偏差内实现了97%的整体准确度。文献[12]基于3D特征图提出了一种 EfficientNet 网络的奶牛体况评分方法,试验结果表明,该模型的识别分类准确率为73.12%,识别误差在0.25内的图像占比为98.6%。上述相关研究中实现了奶牛体况评价,但3D设备成本高,3D数据处理耗时且对设备性能要求较高,这导致在实际农场中推广有较大难度。基于2D设备的体况评分工作取得了重大进展。文献[13]采用 R-CNN 算法构建奶牛体况评价模型,对奶牛尾部图像进行评估,试验结果表明,采用VGG19作为 R-CNN 特征提取层的性能最高,总体准确率为67.39%。文献[14]采用 Faster R-CNN 网络对奶牛体况相关的尾部图像进行识别定位与分类,相比 SSD 算法其检测精度略高,检测准确率为84%,BCS 平均准确率为70%。在奶牛体况评分中采用2D摄像机虽降低了成本,但传统深度学习模型复杂度较高,故对计算设备的算力有较高的要求。此外,实际环境复杂,存在光照变化、运动模糊、遮挡等因素,模型的实际应用仍需要进一步评估。故在饲养环境中,实现实时高效检测奶牛身体区域并获取 BCS 值仍具有难度,且检测算法部署于低算力设备的可行性和应用性仍需要进一步探索。

模型轻量化处理可减小网络模型的占比和模型参数量,提高模型的执行效率,因此得到了广泛关注[15]。其中相关学者在 YOLO 系列模型中进行轻量化处理,实现了果实目标区域的检测[16-18],大幅降低了计算量和模型体积。此外,为提高复杂环境下目标区域的检测准确率,区域特征的有效表达是重要前提。针对图像样本中含有大量无关信息问题,目前常采用引入注意力机制方式,以提高模型对目标区域的关注。该模式在畜牧目标检测中得到了应用[19]。综上可知,对深度学习模型轻量化处理并引入注意力机制的方案是可行的,但针对于奶牛体况评价模型仍需进一步研究可部署于实际饲养环境并推广的模型。

根据现代牧场发展需求,本文采用2D摄像机获取奶牛尾部区域,并基于 YOLO v5s 模型提出一种轻量级奶牛体况评分方法。首先,针对轻量化的需求,采用融入坐标注意力机制的 MobileNetV3 网络来替代 YOLO v5s 的主干网络,在降低模型复杂度的同时增强模型对牛尾区域特征的提取能力。其次,YOLO v5s 预测层采用高效交并比损失,以提升模型回归速度和目标预测框精度,实现模型整体性能的优化。为了进一步验证本研究方法的鲁棒性和应用性,在构建的复杂多样性的样本集下进行验证分析,并在实际牧场中进行应用评估。

1 材料与方法

1.1 数据采集

图像视频采集于安徽省六安市华好生态养殖有限公司与江苏省淮安市盱眙卫岗牧业有限公司,采集平台如图1所示。将海康威视网络相机安装在挤奶通道上方,距地面2.4 m,记录奶牛通过通道的俯



(a) 挤奶通道　(b) 计算中心　(c) 样本获取

图 1　数据采集平台

Fig. 1　Data acquisition platform

视视频。采集视频分辨率为 1 297 像素 ×720 像素，存储格式为 MP4。通道宽度为 1 m，有利于奶牛排队通过，避免了视频中奶牛拥挤、关键部位遮挡等问题。采集样本示例如图 1c 所示。

1.2　数据集构建

相关研究表明，奶牛尾根区域是其体膘储备的主要部位，该部位的体况评估有利于判断奶牛的脂肪储备和健康状况[20]。故以奶牛尾部区域为研究区域，结合规模养殖环境中奶牛体况生长发育周期，构建了奶牛 BCS 数据集。泌乳期的奶牛最优体况评分在 3.25 ~ 3.75 之间[21]。由于泌乳期的奶牛会消耗其体内能量储备来增加产奶量，导致体况评分出现 0.5 ~ 0.75 的浮动，故将 BCS 间隔设置为 0.25 进行分类。

评分为 1.0 和 2.0 的奶牛，其营养缺乏且体况极差，在当下规模化养殖场中很难找到；此外，体况评分在 4.5 ~ 5.0 之间的奶牛属于高危牛群，易引发多种疾病，严重降低牧场经济效益。故规模化养殖场中的奶牛体况评分数值通常分布在 3.25 ~ 4.25 之间。通过对视频数据筛选，最终获取包含 300 头奶牛在自然环境下的体况图像，共 8 972 幅。结合动物医学专家的评分经验，采用 LabelImg 软件对奶牛尾部区域进行标注，按 BCS 为 3.25、3.50、3.75、4.00、4.25 共 5 个样本类别进行分类，样本类别数据集如图 2 所示。

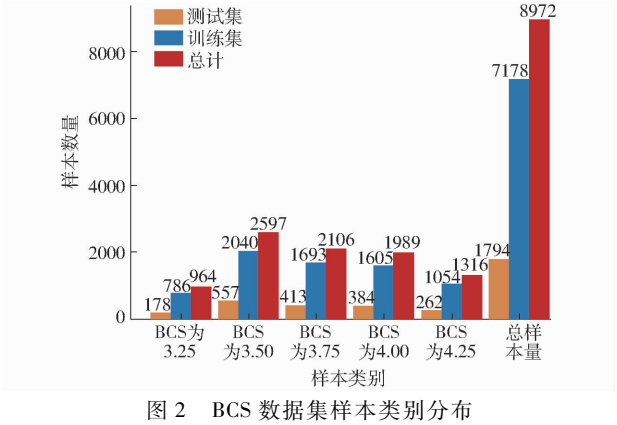
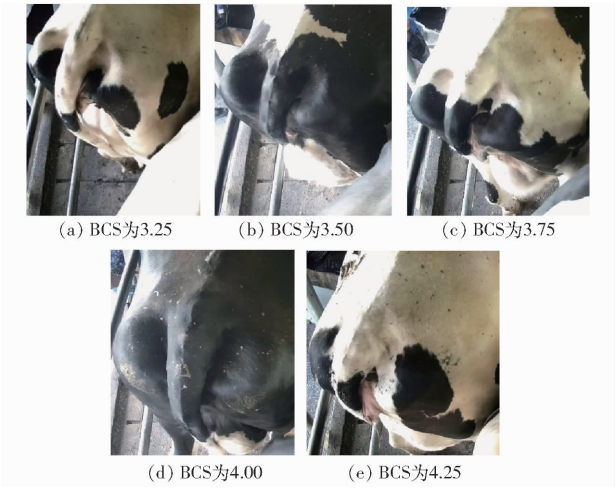


图 2　BCS 数据集样本类别分布

Fig. 2　Distribution of sample categories of BCS dataset

奶牛 BCS 类别示例如图 3 所示，可见随着 BCS 的递增，其尾部的形状逐渐趋于饱满。数据集中的 5 个类别数据分别按照 8∶2 的比例随机划分为训练集和测试集，其中训练集样本为 7 178 幅，测试集样本为 1 794 幅。



(a) BCS为3.25　(b) BCS为3.50　(c) BCS为3.75

(d) BCS为4.00　(e) BCS为4.25

图 3　奶牛 BCS 类别示例

Fig. 3　Example of BCS category for dairy cow

为进一步验证本文模型的鲁棒性和泛化性，从测试集中的 5 个类别中分别按 50% 比例随机抽取图像做模糊化处理（模糊核 $K=32$），常规与模糊效果如图 4 所示。



(a) 原图　(b) 运动模糊

图 4　图像变换效果对比

Fig. 4　Comparison of image transformation effects

2　奶牛体况评分方法

2.1　YOLO v5s 目标检测模型

YOLO v5 为一种单阶段目标检测算法，主要包括 4 部分：① Input 输入层：采用 Mosaic 数据增强，提高训练样本量，增加样本多样性。② Backbone主干网络层[22-24]：采用多种特征提取模块对奶牛区域进行特征提取并生成目标的语义信息等。③ Neck 特征融合层[25]：降低网络深度同时加强奶牛尾部特征信息的融合能力。④ Prediction输出层：采用边界框回归损失和非极

大值抑制算法来提高模型对牛尾的识别准确度。其中,YOLO v5s 网络是 YOLO v5 系列中深度最小、特征图宽度最小的网络,可应用于轻量化研究。YOLO v5s 网络结构如图 5 所示。

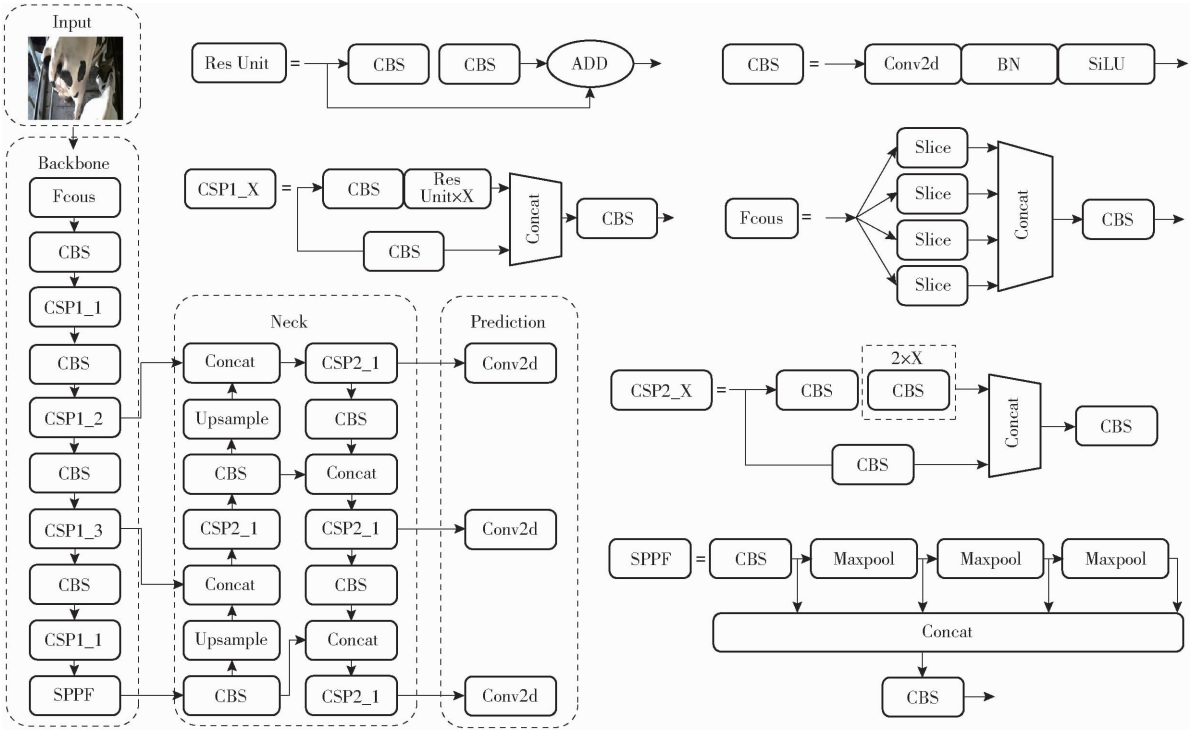


图 5 YOLO v5s 网络结构

Fig. 5 YOLO v5s network structure

2.2 YOLO-MCE 改进模型

针对模型轻量化的需求,本研究对 YOLO v5s 模型进行轻量化处理,以减小网络模型内存占用量、参数量和计算量。此外,由于样本图像中存在背景干扰信息,以及牛只背部部分区域与尾部区域的特征具有相似性,模型实际评估能力较低。故本研究在 MobileNetV3 网络(一种轻量化网络)中引入了坐标注意力机制(Coordinate attention network, CANet)[26]构建 M3CA 网络,并将 M3CA 作为 YOLO v5s 模型的轻量化主干网络,优化了模型处理速度,增强了模型对奶牛尾部区域坐标位置信息的提取能力。其次,YOLO v5s 预测层采用 EIoU Loss(高效交并比损失)来融合更多边界框信息,以优化收敛速度和边界框回归精度。基于 YOLO v5s 改进的 YOLO-MCE 模型网络结构如图 6 所示。

2.2.1 模型轻量化改进

MobileNet 系列网络由 Google 团队在 2017 年提出,是专注于低算力设备的轻量级卷积神经网络。其中,MobileNetV3[27]结合了深度可分离卷积和倒残差结构,删除计算成本较高的网络层,融入了通道注意力机制网络 SENet(Squeeze-and-excitation network)[28]。此外,MobileNetV3 网络采用结构重参数化方法融合了卷积层和批量归一化层,引入易于量化模型的 H_Swish 激活函数,使得模型具有参数少、推理时间短等优势,使网络适用于存储空间和算力受限的设备。主干网络中的 MobileNetV3 结构如表 1 所示。

深度可分离卷积(Depthwise separable convolution, DSC)是 MobileNet 系列的主要特点,网络结构如图 7 所示。深度可分离卷积操作中采用深度卷积(Depthwise convolution, DW)和逐点卷积(Pointwise convolution, PW)对图像进行特征提取。例如,输入特征图尺寸为 $D_F \times D_F \times M$,M 为特征图通道数,卷积核尺寸为 $D_K \times D_K \times 1$,卷积核数量为 N。首先输入特征图经过深度卷积操作(卷积核的数量与输入特征图通道数一致)减少了模型参数量,并生成 M 个特征图。其次经过逐点卷积操作(卷积核尺寸:$1 \times 1 \times M$)将特征图进行组合,生成 N 个特征图,加深模型通道。可见,该卷积运算会将特征图在深度方向上进行加权组合,生成新的特征图。

深度可分离卷积参数量 P_D 和计算量 Q_D 的计算公式分别为

$$P_D = D_K^2 M + MN \tag{1}$$

$$Q_D = D_K^2 M(D_K^2 + N) \tag{2}$$

常规卷积参数量 P_R 和计算量 Q_R 的计算公式分别为

$$P_R = D_K^2 MN \tag{3}$$

$$Q_R = D_K^2 D_F^2 MN \tag{4}$$

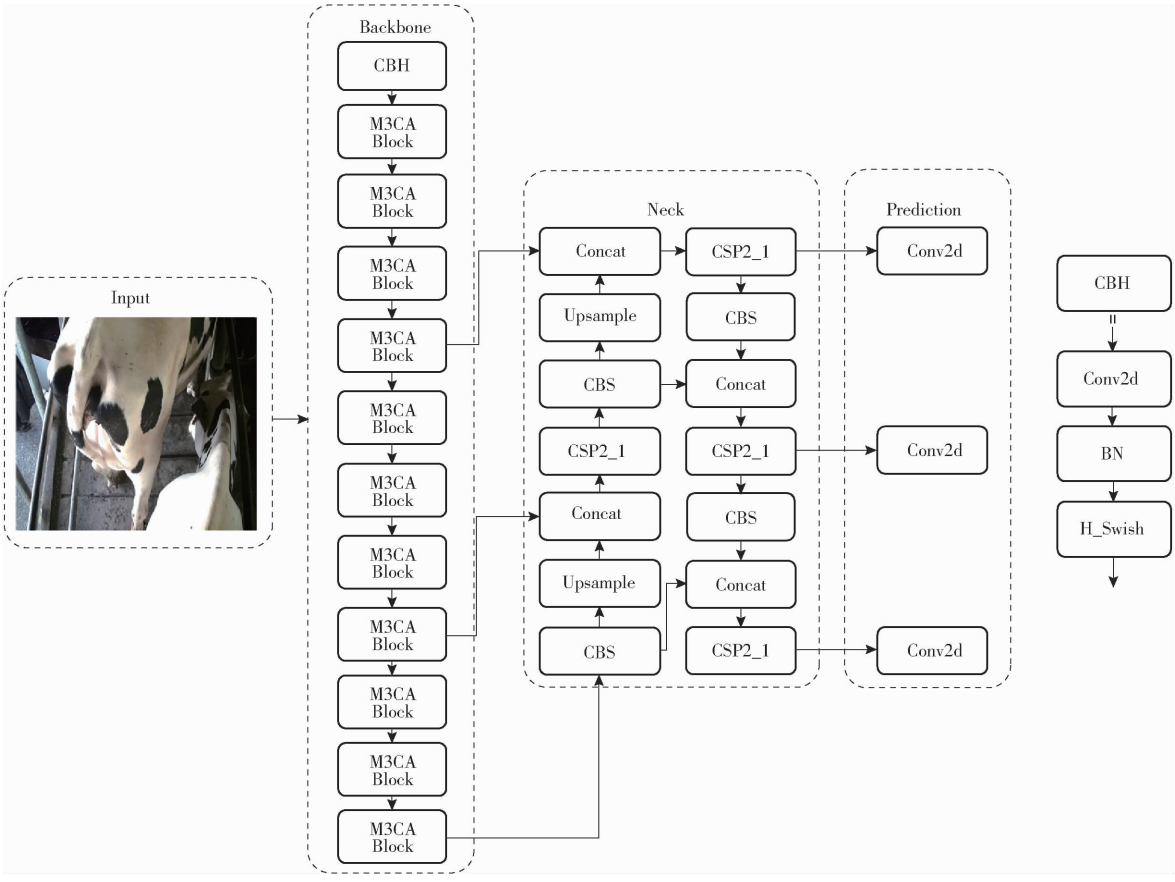


图 6　YOLO－MCE 模型网络结构

Fig. 6　YOLO－MCE network structure

表 1　主干网络中 MobileNetV3 网络结构

Tab. 1　MobileNetV3 network structure in Backbone network

层数	输入	操作	扩张通道	注意力模块	非线性激活函数	步距
1	640×640×3	Conv2d, 3×3			HS	2
2	320×320×16	Bneck, 3×3	16	√	RE	2
3	160×160×16	Bneck, 3×3	72		RE	2
4	80×80×24	Bneck, 3×3	88		RE	1
5	80×80×24	Bneck, 5×5	96	√	HS	2
6	40×40×40	Bneck, 5×5	240	√	HS	1
7	40×40×40	Bneck, 5×5	240	√	HS	1
8	40×40×40	Bneck, 5×5	120	√	HS	1
9	40×40×48	Bneck, 5×5	144	√	HS	1
10	40×40×48	Bneck, 5×5	288	√	HS	2
11	20×20×96	Bneck, 5×5	576	√	HS	1
12	20×20×96	Bneck, 5×5	576	√	HS	1

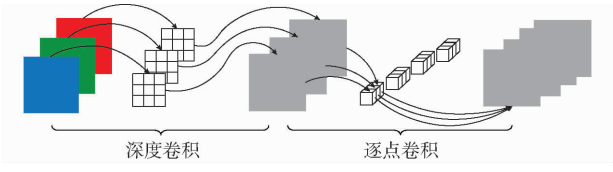


图 7　深度可分离卷积

Fig. 7　Depthwise separable convolution

深度可分离卷积与普通卷积参数量比值 R_P 与计算量比值 R_Q 的计算式为

$$R_P = \frac{P_D}{P_R} = \frac{1}{D_K^2} + \frac{1}{N} \tag{5}$$

$$R_Q = \frac{Q_D}{Q_R} = \frac{1}{D_K^2} + \frac{1}{N} \tag{6}$$

由此得知深度可分离卷积参数量与计算量是普通卷积的$\frac{D_K^2 + N}{D_K^2 N}$。与常规卷积操作相比，其参数量和计算成本都大幅度减少。

2.2.2　坐标注意力机制

通道注意力机制网络仅考虑了通道之间的信息，忽略了位置信息。而坐标注意力机制网络注重特征图像中坐标信息提取。基于以上两点考虑，本方法采用高效的 CANet 替换 MobileNetV3 中的 SENet 构建 M3CA 模块，其结构如图 8 所示。

CANet 通过把坐标位置信息嵌入通道注意力，对输入尺寸为 $C \times H \times W$ 的特征图 $\boldsymbol{X}$，采用平均池化处理生成水平方向尺寸为 $C \times H \times 1$ 和垂直方向尺寸为 $C \times 1 \times W$ 的特征图，C 为特征图通道数，H 为特征图高度，W 为特征图宽度，x_c 为在第 c 个维度的特征输入，即高度 h 和宽度 w 在第 c 个维度的特征输出，计算式分别为

$$z_c^{(h)} = \frac{1}{W} \sum_{0 \leqslant i \leqslant W} x_c(h, i) \tag{7}$$

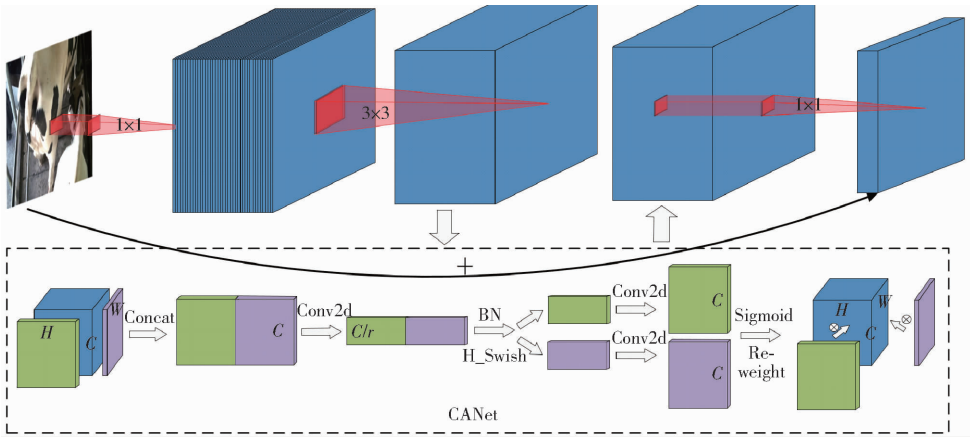


图 8 M3CA 模块

Fig. 8 M3CA block

$$z_c^{(w)} = \frac{1}{H}\sum_{0\leqslant j\leqslant H} x_c(j,w) \tag{8}$$

由式(7)、(8)分别沿两个空间方向进行聚合，可计算输出一对方向可知的特征图。对比 SENet 中全局池化压缩方式，该方式可获取单方向上的长距离关系，同时保留另一个方向上的空间位置信息，有利于网络更准确地定位牛尾区域。

随后将特征图在空间方向进行拼接，使用 1×1 卷积、BN 进行归一化处理，再通过非线性激活函数 $\delta(\cdot)$ 进行特征转化，即

$$f = \delta(F_1((z^{(h)}, z^{(w)}))) \tag{9}$$

将 f 拆分为两个独立的特征矩阵 $\boldsymbol{f}^{(h)}$ 和 $\boldsymbol{f}^{(w)}$，采用 1×1 卷积和 Sigmoid 激活函数将特征矩阵维度转化为与输入 $\boldsymbol{X}$ 的维度一致。得出横纵 2 个方向上的分量，即

$$\boldsymbol{g}^{(h)} = \sigma(F_h(\boldsymbol{f}^{(h)})) \tag{10}$$

$$\boldsymbol{g}^{(w)} = \sigma(F_w(\boldsymbol{f}^{(w)})) \tag{11}$$

最终将输出的 $\boldsymbol{g}^{(h)}$ 和 $\boldsymbol{g}^{(w)}$ 合并成权重矩阵，计算输出式为

$$\boldsymbol{y}_c(i,j) = x_c(i,j) \times \boldsymbol{g}_c^{(h)}(i) \times \boldsymbol{g}_c^{(w)}(j) \tag{12}$$

$\boldsymbol{y}_c(i,j)$ 强调了水平与垂直坐标编码相关的位置信息，将奶牛尾部特征横纵方向的位置信息精确保留，有助于模型更好定位和识别牛尾区域，使得特征提取网络获取更大区域的信息且避免了更大的计算开销。

2.2.3 边界框回归损失函数

EIoU Loss[29] 将纵横比的影响因子拆开，分别计算目标框和锚框的宽高，考虑了重叠面积、中心点距离以及宽高边长真实差异问题。EIoU Loss 对宽高预测结果进行惩罚的损失函数式为

$$L_{EIoU} = L_{IoU} + L_{dis} + L_{asp} = 1 - IoU + \frac{\rho^2(b,b^{(gt)})}{c^2} + \frac{\rho^2(w,w^{(gt)})}{c_w^2} + \frac{\rho^2(h,h^{(gt)})}{c_h^2} \tag{13}$$

式中 IoU——边界框与真实边界框的交并比

b、$b^{(gt)}$——预测边界框、真实边界框的中心点坐标

$\rho^2(\cdot)$——欧氏距离

c_w、c_h——包围两个边界框的最小的矩形框的宽度、高度

由 3 部分组成，包括重叠损失 L_{IoU}、中心距离损失 L_{dis}、宽高损失 L_{asp}。本方法在预测层采用 EIoU Loss，使模型在回归过程专注于高质量锚框，加速收敛提高回归精度。

3 试验平台及模型评价指标

3.1 试验平台与对比模型

本试验在 Ubuntu 20.04 操作系统和 NVIDIA RTX3080 显卡下，通过 Pytorch 深度学习框架来实现模型的搭建、训练和验证，使用 CUDA 11.3 计算架构，同时将 CUDNN 添加到环境中加速计算机计算能力。改进模型试验过程中样本图像分辨率为 640 像素 × 640 像素，训练批次为 16，初始学习率为 0.01，初始动量为 0.937，所有参照模型均按此参数训练 300 个 epoch。测试阶段参数交并比(IoU)设置为 0.5，置信度阈值设置为 0.5。试验过程中采用 Mosaic 数据增强，扩充试验样本同时增强模型的鲁棒性和泛化能力。同时，为了验证 YOLO - MCE 算法的性能，与 Faster R - CNN、SSD、YOLO v3 - SPP 和 YOLO v5s 进行对比分析。试验平台如表 2 所示。

3.2 评价指标

本研究选择目标检测模型中常用的评价指标：精度(Precision，P)、召回率(Recall，R)、平均精度(Avarge precision，AP)、平均精度均值(Mean avarge precision，mAP)、模型内存占用量、参数量和计算量，对改进后的模型进行综合评价。

表 2　试验平台

Tab. 2　Experimental platform

名称	配置
CPU	Intel(R) Core(TM) i7 - 9700 CPU 3.00 GHz
内存	32 GB
GPU	NVIDIA GeForce RTX 3080
GPU 加速库	CUDA 11.3、CUDNN 8.4.0
操作系统	Ubuntu 20.04 LTS
环境	Python 3.10、Pytorch 1.12.1

精度和召回率衡量了模型对奶牛尾部样本 BCS 分类的能力。AP 为不同 BCS 类别的平均精度；mAP 为 BCS 类别下的 AP 值总和的平均值，衡量模型在所有 BCS 类别上的检测精度。

模型内存占用量、参数量和计算量是深度学习训练中衡量模型复杂度的指标。模型内存占用量是衡量模型所占用存储空间的参数。三者数值越小，越容易移植模型到低算力设备中。

4　结果与分析

4.1　轻量化处理对比

为验证改进模型轻量化处理后的性能，在奶牛 BCS 数据集基础上，将改进模型 YOLO v5s - MobileNetV3 - CA - EIoU(YOLO - MCE)与 YOLO v5s、YOLO v5s - MobileNetV3 和 YOLO v5s - MobileNetV3 - CA 模型进行了对比试验分析，其中对比模型采用 CIoU Loss。由表 3(表中 mAP@0.5 表示 IoU 为 0.5 的平均精度均值)得知，在主干网络中引入 MobileNetV3 网络后，YOLO v5s - MobileNetV3 模型相较于原 YOLO v5s 模型，其复杂度大幅度下降，模型内存占用量为 2.95 MB，计算量为 2.3×10^9，但其检测精度、召回率和平均精度均值均有所降低，其原因在于轻量级网络参数较少且模型深度较低，导致对奶牛尾部特征细节提取能力较低，即损失部分检测精度换取了模型的轻量化。故为了进一步提高模型检测性能，将引入坐标注意力机制的 M3CA 网络融入 YOLO v5s 主干网络，YOLO v5s - MobileNetV3 - CA 模型内存占用量为 2.28 MB，计算量为 2.0×10^9，相较 YOLO v5 - MobileNetV3，该模型的复杂度得到进一步下降，其检测精度和 mAP@0.5 均有所改善。预测层采用 EIOU Loss 后，YOLO - MCE 模型与 YOLO v5s - MobileNetV3 - CA 模型的复杂度一致，YOLO - MCE 模型的检测精度为 93.4%，召回率为 85.5%，mAP@0.5 为91.4%，较原 YOLO v5s 模型，其检测精度降低 0.2 个百分点，召回率提升 3.7 个百分点，平均精度均值提升 2.1 个百分点，模型内存占用量压缩 83.4%。由此得知，YOLO - MCE 模型实现了模型的轻量化，并在保持较高检测精度的同时大幅降低了模型的复杂度。

表 3　消融试验

Tab. 3　Ablation experiment

模型	精度/%	召回率/%	mAP@0.5/%	模型内存占用量/MB	参数量	计算量
YOLO v5s	93.6	81.8	89.3	13.72	7.02×10^6	1.58×10^{10}
YOLO v5s - MobileNetV3	91.8	81.6	88.6	2.95	1.38×10^6	2.30×10^9
YOLO v5s - MobileNetV3 - CA	94.0	83.2	90.3	2.28	1.02×10^6	2.00×10^9
YOLO - MCE	93.4	85.5	91.4	2.28	1.02×10^6	2.00×10^9

4.2　不同模型下奶牛体况评分对比

本研究在奶牛 BCS 数据集上进一步将 YOLO - MCE 与 Faster R - CNN、SSD、YOLO v3 - SPP 和 YOLO v5s 模型进行对比分析。由表 4 可以得知，YOLO -MCE 模型 mAP@0.5 为 91.4%，模型内存占用量为 2.28 MB，相比 Faster R - CNN、SSD、YOLO v3 -SPP 和 YOLO v5s 模型，其 mAP@0.5 分别提升 22.4、9.8、15.6、2.1 个百分点；其模型内存占用量分别缩小 1/137.5、1/45.6、1/141.7 和 1/6.0。由此可知，YOLO - MCE 模型相比 Faster R - CNN、SSD、YOLO v3 - SPP 和 YOLO v5s 模型，其在奶牛 BCS 数据集上的整体性能最优，且模型复杂度最低，有利于改进的奶牛体况评分模型部署于低算力设备。

表 4　各模型检测性能对比

Tab. 4　Comparison of detection performance of different models

模型	mAP@0.5/%	模型内存占用量/MB	参数量
Faster R - CNN	69.0	313.47	1.37×10^8
SSD	81.6	104.04	2.62×10^7
YOLO v3 - SPP	75.8	323.12	6.30×10^7
YOLO v5s	89.3	13.72	7.02×10^6
YOLO - MCE	91.4	2.28	1.02×10^6

4.3　不同 BCS 类别下检测结果

本研究选取置信度阈值为 0.5，IoU 阈值为 0.5。YOLO v5s 和 YOLO - MCE 模型的精度-召回率曲线如图 9 所示。可见在 BCS 为 3.75、4.00、4.25 的类别下 YOLO - MCE(88.2%、92.2%、

95.5%)的分类效果优于 YOLO v5s(86.4%、89.1%、88.5%);BCS 为 3.25、3.50 的类别下,YOLO-MCE(92.1%、89.0%)的结果略低于 YOLO v5s(92.5%、90.2%)。改进后的模型在类别识别分类的总体性能得到了改善。

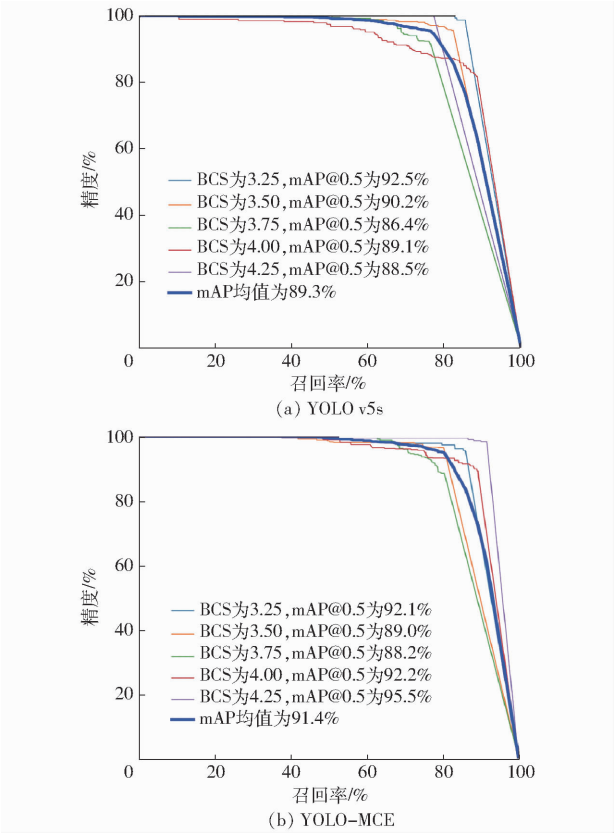


图 9 YOLO v5s 与 YOLO-MCE 的精度-召回率曲线

Fig. 9 P-R curves of YOLO v5s and YOLO-MCE

为验证模糊数据集下模型的检测性能,分别对比分析了 BCS 不同类别下的 YOLO v5s 和 YOLO-MCE 的检测效果。各类别模糊样本下的 BCS 检测结果如图 10 所示。由图 10 得知,引入坐标注意力机制和 EIoU Loss 的 YOLO-MCE 模型,更好地结合了牛尾区域的位置信息,从而具有良好的特征提取、特征融合以及高质量锚框生成能力。降低了漏检率、提高了运动模糊状况下的检测性能以及目标边界框定位能力,在牛尾区域模糊的识别分类精度上有所改善。针对 YOLO v5s 在模糊样本下出现的漏检(图 10a、10c、10g)和误检(图 10e、10i)情况,改进模型均有所改善,如图 10b、10d、10f、10h、10j 所示。试验结果表明,改进模型相较原模型在模糊样本下具有更好的检测和泛化能力。

为了更直观体现模型的注意区域,本研究借助梯度加权类激活映射方法(Gradient-weighted class activation map, Grad-CAM)[30]对牛尾重要特征区域进行颜色可视化。在模型预测层中的最后一个卷积层后引入 Grad-CAM 方法,直观展示了模型在牛



图 10 YOLO v5s 和 YOLO-MCE 在模糊样本下各 BCS 的检测效果

Fig. 10 Detection results of YOLO v5s and YOLO-MCE in each BCS value under motion blur sample

尾检测中注意力关注区域的分布情况,如图 11 所示。奶牛尾部沟壑区域的凹陷程度对奶牛尾部识别分类做出了关键决策。而在 BCS 为 3.25、3.50 和 3.75 时,原模型对牛尾特征感兴趣区域较为集中,BCS 为 4.00 和 4.25 时,感兴趣区域较为分散,未能正确学习到牛尾关键特征。而改进模型的感兴趣区域覆盖范围更广,针对牛尾关键特征区域的定位更加精准,使模型对牛尾特征信息的学习更加充分,从而提升了实际检测效果。

4.4 应用案例

本模型已部署于盱眙卫岗牧业有限公司奶牛养殖场,其工控机配置为 CPU:Intel12thGenIntel(R) Core(TM) i5-12400F;GPU:NVIDIA GeForce RTX 3060, Microsoft Basic Render Driver;内存:16 GB;Windows 11 系统。工控机现场测试如图 12 所示。监视界面中可实时监测奶牛 BCS 数值。此外,牧场电子耳标记录奶牛 ID 信息,可进一步将 ID 信息与个体 BCS 数值相关联。

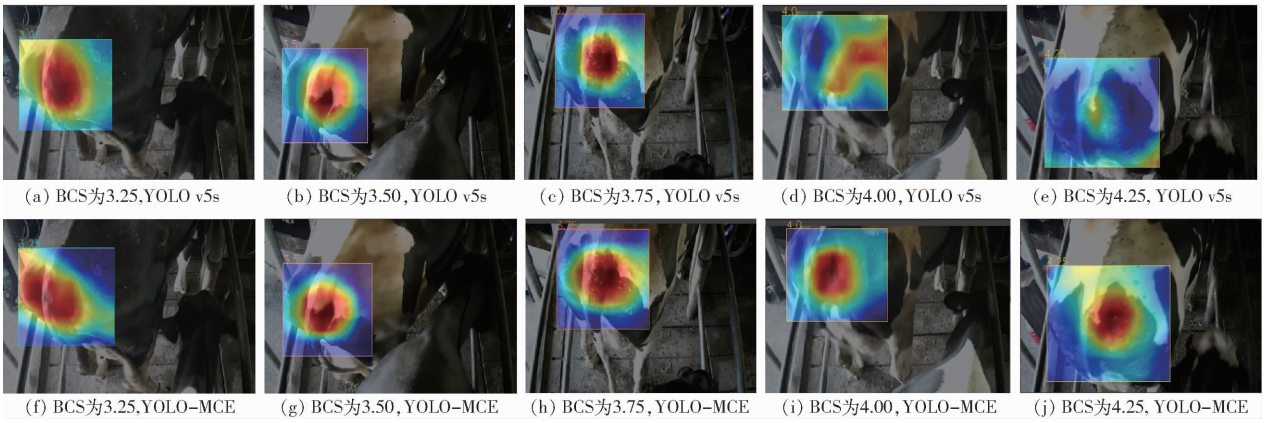


图 11　YOLO v5s 和 YOLO－MCE 的 Grad－CAM 热力图对比

Fig. 11　Comparison of Grad－CAM heat maps of YOLO v5s and YOLO－MCE

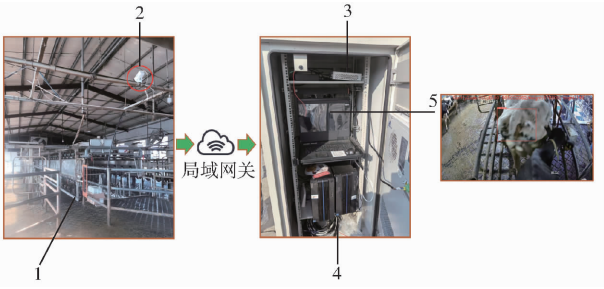


图 12　工控机现场测试图

Fig. 12　Field test diagram of industrial computer

1. 挤奶通道　2. 2D 摄像机　3. 硬盘录像机　4. 工控机　5. 监控界面

5　结论

(1)提出了一种轻量化奶牛体况评价模型。将融入坐标注意力机制的 MobileNetV3 网络作为 YOLO v5 的主干网络,优化了目标检测模型的处理速度,并加强了模型对奶牛尾部区域空间和位置信息的提取能力。预测层采用 EIoU Loss 来融合更多的边界框信息,优化了收敛速度和边界框回归精度。整体提高了模型在奶牛体况评分领域的检测性能和实用性。

(2)改进模型的目标检测精度为 93.4%,召回率为 85.5%,mAP@0.5 为 91.4%,模型内存占用量为 2.28 MB,参数量为 1.02×10^6,计算量为 2.0×10^9。改进的检测模型具有较低的模型复杂度,同时具备较高的牛尾区域识别与体况评估性能。

(3)所提方法在检测精度与实时性较高的情况下,可极大程度地降低模型对硬件设备算力的依赖。通过在实际场景中的测试,本文方法可实现对奶牛体况的高效评分。

参　考　文　献

[1] 高芬,李秋波,周杰,等.我国肉牛规模化养殖变动趋势及影响因素研究——基于面板分位数回归的实证分析[J/OL].中国农业资源与区划:1－10[2023－02－13].
GAO Fen, LI Qiubo, ZHOU Jie, et al. Study on the changing trend and influencing factors of large-scale beef cattle breeding in China—empirical analysis based on panl quantile regression[J/OL]. Chinese Journal of Agricultural Resources and Regional Planning:1－10[2023－02－13]. (in Chinese)

[2] 徐哲,寇秋雯,贾亚雄,等.科技创新驱动奶业可持续健康发展[J].农业科技管理,2021,40(4):25－27,37.
XU Zhe, KOU Qiuwen, JIA Yaxiong, et al. Promoting sustainable and healthy development of dairy industry by scientific and technological innovation[J]. Management of Agricultural Science and Technology, 2021,40(4):25－27,37. (in Chinese)

[3] PETROVSKA S, JONKUS D. Relationship between body condition score, milk productivity and live weight of dairy cows[J]. Res. Rural. Dev., 2014, 1(1): 100－106.

[4] 李吉楠, 孙鹏, 覃春富, 等. 体况评分在奶牛饲养管理上应用的研究进展[J]. 中国畜牧兽医, 2013, 40(10): 115－119.
LI Ji'nan, SUN Peng, QIN Chunfu, et al. Research advance of body condition score in the management of feeding the dairy cows[J]. China Animal Husbandry & Veterinary Medicine, 2013, 40(10): 115－119. (in Chinese)

[5] ROCHE J R, MACDONALD K A, SCHÜTZ K E, et al. Calving body condition score affects indicators of health in grazing dairy cows[J]. Journal of Dairy Science, 2013, 96(9): 5811－5825.

[6] 赵一广, 杨亮, 郑姗姗, 等. 家畜智能养殖设备和饲喂技术应用研究现状与发展趋势[J]. 智慧农业, 2019, 1(1): 20－31.
ZHAO Yiguang, YANG Liang, ZHENG Shanshan, et al. Advances in the development and applications of intelligent equipment and feeding technology for livestock production[J]. Smart Agriculture, 2019, 1(1): 20－31. (in Chinese)

[7] 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)
[8] QIAO Y, KONG H, CLARK C, et al. Intelligent perception for cattle monitoring: a review for cattle identification, body condition score evaluation, and weight estimation[J]. Computers and Electronics in Agriculture, 2021, 185: 106143.
[9] GUO Y, ZHANG Z, HE D, et al. Detection of cow mounting behavior using region geometry and optical flow characteristics [J]. Computers and Electronics in Agriculture, 2019, 163: 104828.
[10] HUANG X, HU Z, QIAO Y, et al. Deep learning-based cow tail detection and tracking for precision livestock farming[J/OL]. IEEE/ASME Transactions on Mechatronics, 2022. DOI:10.1109/TMECH.2022.3175377.
[11] RODRÍGUEZ ALVAREZ J, ARROQUI M, MANGUDO P, et al. Estimating body condition score in dairy cows from depth images using convolutional neural networks, transfer learning and model ensembling techniques[J]. Agronomy, 2019, 9(2): 90.
[12] 赵凯旋,刘晓航,姬江涛. 基于 EfficientNet 与点云凸包特征的奶牛体况自动评分[J]. 农业机械学报,2021,52(5):192 - 201,73.
ZHAO Kaixuan, LIU Xiaohang, JI Jiangtao. Automatic body condition scoring method for dairy cows based on EfficientNet and convex hull feature of point cloud[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(5):192 - 201, 73. (in Chinese)
[13] ÇEVIK K K, MUSTAFA B. Body condition score (BCS) segmentation and classification in dairy cows using R - CNN deep learning architecture[J]. Avrupa Bilim ve Teknoloji Dergisi, 2019(17): 1248 - 1255.
[14] HUANG Xiaoping, LI Xinru, HU Zelin. Cow tail detection method for body condition score using Faster R - CNN[C]// IEEE International Conference on Unmanned Systems and Artificial Intelligence, ICUSAI, 2019:347 - 351.
[15] 宋怀波,尚钰莹,何东健. 果实目标深度学习识别技术研究进展[J]. 农业机械学报,2023,54(1):1 - 19.
SONG Huaibo, SHANG Yuying, HE Dongjian. Review on deep learning technology for fruit target recognition [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1):1 - 19. (in Chinese)
[16] 罗志聪,李鹏博,宋飞宇,等. 嵌入式设备的轻量化百香果检测模型[J]. 农业机械学报,2022,53(11):262 - 269,322.
LUO Zhicong, LI Pengbo, SONG Feiyu, et al. Lightweight passion fruit detection model based on embedded device[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(11):262 - 269,322. (in Chinese)
[17] 王卓,王健,王枭雄,等. 基于改进 YOLO v4 的自然环境苹果轻量级检测方法[J]. 农业机械学报,2022,53(8):294 - 302.
WANG Zhuo, WANG Jian, WANG Xiaoxiong, et al. Lightweight real-time apple detection method based on improved YOLO v4[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(8):294 - 302. (in Chinese)
[18] 武星,齐泽宇,王龙军,等. 基于轻量化 YOLOv3 卷积神经网络的苹果检测方法[J]. 农业机械学报,2020,51(8):17 - 25.
WU Xing, QI Zeyu, WANG Longjun, et al. Apple detection method based on Light - YOLOv3 convolutional neural network [J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8):17 - 25. (in Chinese)
[19] GUO Y, QIAO Y, SUKKARIEH S, et al. Bigru-attention based cow behavior classification using video data for precision livestock farming[J]. Transactions of the ASABE, 2021, 64(6): 1823 - 1833.
[20] WILDMAN E E, JONES G M, WAGNER P E, et al. A dairy cow body condition scoring system and its relationship to selected production characteristics[J]. Journal of Dairy Science, 1982, 65(3): 495 - 501.
[21] LI X, HU Z, HUANG X, et al. Cow body condition score estimation with convolutional neural networks[C]//2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC). IEEE, 2019: 433 - 437.
[22] 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.
[23] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 390 - 391.
[24] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904 - 1916.
[25] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759 - 8768.
[26] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713 - 13722.
[27] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1314 - 1324.
[28] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132 - 7141.
[29] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146 - 157.
[30] 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.