

基于 SNSS – YOLO v7 的肉牛行为识别方法

段青玲^{1,2} 赵芷青^{1,2} 蒋涛³ 桂小飞³ 张宇航^{1,2}

(1. 中国农业大学信息与电气工程学院, 北京 100083; 2. 农业农村部智慧养殖技术重点实验室, 北京 100083;
3. 北京福通互联科技集团有限公司, 北京 101300)

摘要: 肉牛活动过程中所表现出的行为是肉牛健康状况的综合体现,实现肉牛行为的快速准确识别,对肉牛疾病防控、自身发育评估和发情监测等具有重要作用。基于机器视觉的行为识别技术因其无损、快速的特点,已应用在畜禽养殖行为识别中,但现有的基于机器视觉的肉牛行为识别方法通常针对单只牛或单独某个行为开展研究,且存在计算量大等问题。针对上述问题,本文提出了一种基于 SNSS – YOLO v7 (Slim – Neck & Separated and enhancement attention module & Simplified spatial pyramid pooling-fast – YOLO v7)的肉牛行为识别方法。首先在复杂环境下采集肉牛的爬跨、躺卧、探究、站立、运动、舔舐和互斗 7 种常见行为图像,构建肉牛行为数据集;其次在 YOLO v7 颈部采用 Slim – Neck 结构,以减小模型计算量与参数量;然后在头部引入分离和增强注意力模块 (Separated and enhancement attention module, SEAM) 增强 Neck 层输出后的检测效果;最后使用 SimSPPF (Simplified spatial pyramid pooling-fast) 模块替换原 YOLO v7 的 SPPCSPC (Spatial pyramid pooling cross stage partial conv) 模块,在增大感受野的同时进一步减少参数量。在自建数据集上测试,本文提出的肉牛行为识别方法的平均精度均值 (mAP_{@0.5}) 为 95.2%,模型内存占用量为 39 MB,参数量为 1.926×10^7 。与 YOLO v7、YOLO v6m、YOLO v5m、YOLOX-S、TPH – YOLO v5、Faster R – CNN 相比,模型内存占用量分别减小 47.9%、45.4%、7.6%、43.1%、57.8% 和 92.5%,平均精度均值 (mAP_{@0.5}) 分别提高 1.4、2.2、3.1、13.7、1.9、4.5 个百分点,试验结果表明,本文方法能够实现肉牛行为的准确识别,可以部署在计算资源有限的设备上,为实现畜禽养殖智能化提供支持。

关键词: 肉牛; 行为识别; 深度学习; YOLO v7; 多目标识别

中图分类号: TP391.41; S823 文献标识码: A 文章编号: 1000-1298(2023)10-0266-09 OSID: 

Behavior Recognition Method of Beef Cattle Based on SNSS – YOLO v7

DUAN Qingling^{1,2} ZHAO Zhiqing^{1,2} JIANG Tao³ GUI Xiaofei³ ZHANG Yuhang^{1,2}

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
2. Key Laboratory of Smart Breeding Technology, Ministry of Agriculture and Rural Affairs, Beijing 100083, China
3. Beijing Futong Internet Technology Group Co., Ltd., Beijing 101300, China)

Abstract: The behavior of beef cattle in the process of activity is the comprehensive embodiment of the health status of beef cattle. The rapid and accurate recognition of beef cattle behavior plays an important role in the prevention and control of beef cattle diseases, their own development assessment and estrus monitoring. Behavior recognition technology based on machine vision has been applied to behavior recognition of livestock and poultry breeding because of its lossless and fast characteristics. However, the existing behavior recognition methods of beef cattle based on machine vision were usually studied for a single cow or a single behavior, and there were problems such as large amount of calculation. In view of the above problems, a method based on Slim – Neck & Separated and enhancement attention module & Simplified spatial pyramid pooling-fast – YOLO v7 (SNSS – YOLO v7) was proposed. Firstly, seven common behavior images of beef cattle, such as mounting, lying, searching, standing, walking, licking and fighting, were collected in the complex environment to construct a beef cattle behavior dataset. Secondly, the Slim – Neck structure was used in the neck of YOLO v7 to reduce the amount of calculation and parameters of the model. Then, separated and enhancement attention module (SEAM) was introduced into the head to enhance the detection effect after the output of the Neck layer. Finally, the

simplified spatial pyramid pooling-fast (SimSPPF) module was used to replace the spatial pyramid pooling cross stage partial conv (SPPCSPC) module of the original YOLO v7, which further reduced the number of parameters while increased the receptive field. Tested on the self-built dataset, the mean average precision ($mAP_{@0.5}$) of the beef cattle behavior recognition method proposed was 95.2%, the model size was 39 MB, and the number of parameters was 1.926×10^7 . Compared with YOLO v7, YOLO v6m, YOLO v5m, YOLOX-S, TPH-YOLO v5 and Faster R-CNN, the model size was reduced by 47.9%, 45.4%, 7.6%, 43.1%, 57.8% and 92.5%, respectively. The mean average precision ($mAP_{@0.5}$) was improved by 1.4 percentage points, 2.2 percentage points, 3.1 percentage points, 13.7 percentage points, 1.9 percentage points, and 4.5 percentage points, respectively. The experimental results showed that the proposed method can achieve accurate recognition of beef cattle behavior, and can be deployed on devices with limited computing resources to provide support for intelligent livestock breeding.

Key words: beef cattle; behavior recognition; deep learning; YOLO v7; multi-object recognition

0 引言

肉牛是一种以生产牛肉为主的牛,是我国重要的畜禽养殖品种之一。牛肉因富含蛋白质和维生素等营养物质,对人体健康有益而深受消费者喜欢^[1-2]。在肉牛养殖中,肉牛活动过程中所表现出的行为是其健康状况的综合体现,对疾病防控、自身发育评估和发情监测有重要意义。快速准确地识别肉牛爬跨、躺卧、探究、站立、运动、舔舐和互斗等常见行为能够帮助养殖者及时评估肉牛的生理健康状况,实现肉牛的精准化管理。

国内外研究学者开展了肉牛行为识别技术研究,主要分为基于传感器的接触式肉牛行为识别^[3-4]和基于计算机视觉的非接触式肉牛行为识别^[5-6]两种方式。基于传感器的接触式肉牛行为识别方法是在牛只不同部位佩戴传感器,根据采集的肉牛活动量、温度和声音等数据,判断肉牛的不同行为^[7-8]。该方法容易造成牛只的应激反应,易损坏传感器,从而影响牛只行为识别的准确率。

基于计算机视觉的非接触式肉牛行为识别方法通过采集视频图像,利用机器学习技术对图像进行处理,实现肉牛行为识别。该方法成本较低,可以避免接触式识别方法带来的肉牛应激反应问题。GUO等^[9]提出了基于图像区域几何和光流特征分析的爬跨行为识别方法,首先使用融合颜色和纹理的背景减法获得奶牛目标,然后使用帧间差分法提取奶牛的几何和光流特征,最后使用支持向量机(Support vector machine, SVM)算法进行特征分类,实现奶牛爬跨行为的识别,识别准确率为 90.9%。JIANG等^[10]提出了基于噪声和双向长短期记忆模型的跛行检测方法,首先采用 FLYOLO v3 (Fliter layer YOLO v3)算法提取奶牛背部位置坐标,并获得奶牛目标,然后提取奶牛背部曲率,采用噪声和双向长短期记忆模型,实现奶牛跛行的识别,识别

准确率为 96.1%。刘月峰等^[11]提出了基于改进 YOLO v3 的奶牛进食行为识别方法,使用幅值迭代剪枝算法对 YOLO v3 进行优化,实现奶牛进食行为的识别,识别准确率为 79.9%。王少华等^[12]提出了基于改进 YOLO v3 的奶牛发情行为识别方法,使用锚点框聚类、DenseBlock 和边框损失函数对 YOLO v3 进行优化,实现奶牛发情行为的识别。上述方法仅实现了牛只单一行为的识别,如果要识别多种行为,需要对多个模型进行集成,无法满足实际生产的需求。

在多行为识别方面, QIAO 等^[13]提出了基于 C3D (Convolutional 3D) 网络和 ConvLSTM (Convolutional LSTM network) 的牛只行为识别方法,实现了小牛和奶牛的采食、探索、舔舐、行走和站立 5 种行为的分类,小牛和奶牛分类准确率分别为 90.32% 和 86.67%。WU 等^[14]提出了基于卷积神经网络和长短期记忆网络的奶牛行为识别方法,实现了单只奶牛的躺卧、站立、行走、饮水和反刍 5 种行为的识别,识别准确率为 97.6%。YIN 等^[15]提出了基于 EfficientNet-LSTM 的奶牛行为识别方法,使用双向特征金字塔网络 (Bi-directional feature pyramid network, BiFPN) 优化 EfficientNet 实现特征聚合,并使用注意力机制优化双向长短期记忆网络实现时间维度的串联,最后实现了单只奶牛的行走、站立、躺卧、哺乳和饮水 5 种行为的识别,识别准确率为 95.2%。MA 等^[16]提出了基于 ReXNet 3D (Rank expansion networks 3D) 的奶牛行为识别方法,在 ReXNet (Rank expansion networks) 中添加时间信息,实现了单只奶牛的躺卧、站立和行走 3 种行为的识别,识别准确率为 95%。上述方法仅实现了单只牛的多种行为识别,无法直接应用在规模化养殖场多只牛的场景中。

综上所述,基于计算机视觉的非接触式牛只行为识别是目前研究热点,已开展了大量的相关研究,

但仍存在以下问题:在实际养殖环境中存在多只肉牛且行为多种多样,现有模型无法适用于养殖场的实际情况;现有模型存在参数多和模型大的特点,不利于在算力有限的边缘端设备上部署应用。

针对上述问题,本文提出一种基于 SNSS - YOLO v7 的肉牛行为识别方法,首先在实际养殖环境下采集肉牛行为数据,构建肉牛行为数据集;其次在 YOLO v7 颈部采用 Slim - Neck 结构,以减小模型计算量与参数量;然后在头部引入 SEAM 模块增强 Neck 层输出后的检测效果;最后使用 SimSPPF 模块替换原 YOLO v7 的 SPPCSPC 模块,在增大感受野的同时进一步减少参数量。

1 材料与方 法

1.1 试验数据

本文试验数据通过肉牛养殖基地实地采集和网络抓取两种渠道获取。以内蒙古锡林郭勒盟正蓝旗鑫源牧场为试验场地,通过视频摄像设备(海康威视 DS - 2CD3T86FWDV2 - I3S)采集养殖场的肉牛行为数据。通过上述两种渠道获取肉牛多天的日常活动视频,人工筛选出具有爬跨、躺卧、探究、站立、运动、舔舐和互斗 7 种行为的视频片段 75 条,视频片段长度不等。按照比例 8 : 2 随机选择 60 条视频片段做训练集,15 条视频片段做测试集。为避免图像相似度过高,对视频片段采取视频抽帧的方式,每间隔 25 帧抽取一幅图像,共获得训练集图像 1 959 幅,测试集图像 501 幅,构成本文的行为识别数据集。表 1 给出了肉牛 7 种行为判定标准和数据集信息。

表 1 肉牛行为判定标准和数据集信息
Tab.1 Beef cattle behavior recognition criteria and dataset description

行为序号	行为类别	行为描述	行为数量	标签
①	爬跨	一只牛身体的 1/3 以上部分爬跨在另一只牛身上	470	mounting
②	躺卧	腿弯曲与地面接触	5 197	lying
③	探究	头靠近地面或与地面接触	503	searching
④	站立	4 条腿支撑身体站立,头未与地面进行接触	2 779	standing
⑤	运动	身体与 4 条腿的位置发生移动,头未与地面进行接触	492	walking
⑥	舔舐	头转向腹部,用舌头舔食身体	449	licking
⑦	互斗	2 头肉牛头部互相顶撞	214	fighting

为了模拟实际生产环境,保证算法的实用性,构建肉牛行为识别数据集时综合考虑以下因素:①姿态变化。采集包含肉牛多种姿态的图像,以增加样

本的多样性。②密集程度。采集不同密集程度下的肉牛图像,每幅图像中包含多头肉牛和多种行为,以增加识别算法对不同场景的适应性。③光照变化。采集不同自然光照条件下的肉牛图像,以增加识别算法对不同光照条件的适应性。

为了进一步增强图像的复杂性,本文采用马赛克增强的方法进行数据增强。采用 Labellmg 人工标注每幅图像中肉牛的爬跨、躺卧等 7 种行为。数据集集中的部分图像如图 1 所示。其中,图 1a、1d 是网络抓取的图像,图 1b、1c、1e、1f 是在肉牛养殖基地获取的图像。图中标注的目标①为爬跨肉牛、目标②为躺卧肉牛、目标③为探究肉牛、目标④为站立肉牛、目标⑤为运动肉牛、目标⑥为舔舐肉牛、目标⑦为互斗肉牛。



图 1 肉牛不同行为示意图

Fig.1 Diagrams of different behaviors of beef cattle

1.2 肉牛行为识别模型

1.2.1 肉牛行为识别模型架构

YOLO v7 算法是优秀的目标检测算法,其网络模型主要由 CBS 模块、ELAN 模块、MP 模块、SPPCSPC 模块以及 Concat 模块构成,包含大量的普通卷积,内存消耗较大,难以部署在边缘端设备上。本文提出基于 SNSS - YOLO v7 的肉牛行为识别模型,首先在网络颈部采用 Slim - Neck 结构轻量化网络;其次头部引入 SEAM 模块增强 Neck 层输出后的检测效果;最后使用 SimSPPF 模块替换原 YOLO v7 的 SPPCSPC 模块。在保持精度的同时有效减少模型内存占用量,从而实现肉牛行为的准确识别。

肉牛行为识别模型架构如图 2 所示。当输入图像为 640 像素 × 640 像素时,主干网络从输入图像中提取 80 像素 × 80 像素、40 像素 × 40 像素和 20 像素 × 20 像素共 3 种特征图;颈部采用 SimSPPF 模块增大感受野,使用 Slim - Neck 结构对主干网络

提取的 3 种不同大小的特征图进行特征融合,且降低网络复杂度;头部引入 SEAM 模块增强 Neck 层输出后的检测效果。通过损失函数计算训练损失,使用优化器调整参数,获得肉牛行为识别模型,并对测试集进行结果预测。

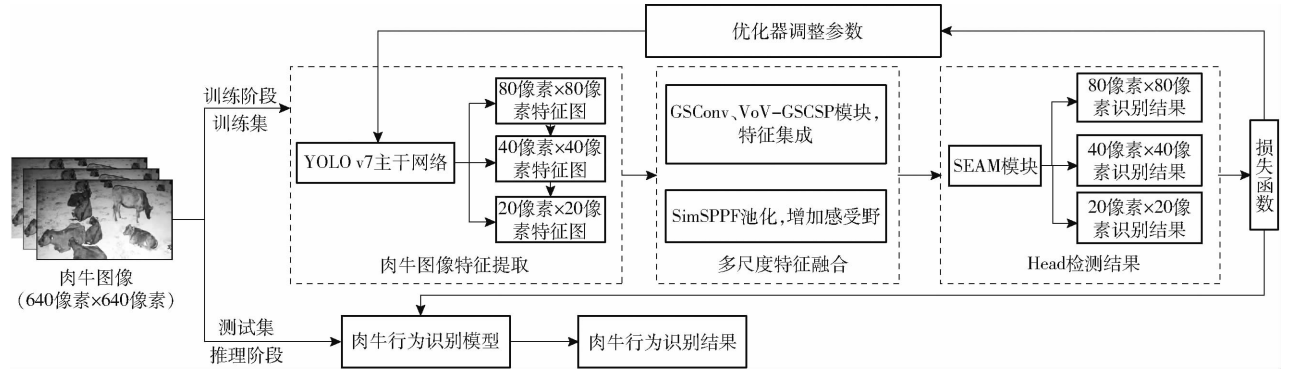


图 2 肉牛行为识别模型架构

Fig. 2 Beef cattle behavior recognition model architecture

1. 2. 2 SNSS - YOLO v7 模型网络结构

基于 SNSS - YOLO v7 的轻量级肉牛行为识别模型网络结构如图 3 所示。

采用 Slim - Neck 结构轻量化颈部网络,减轻模型的复杂度同时保持精度;引入 SEAM 模块增强 Neck 层输出后的检测效果;使用 SimSPPF 模块替换

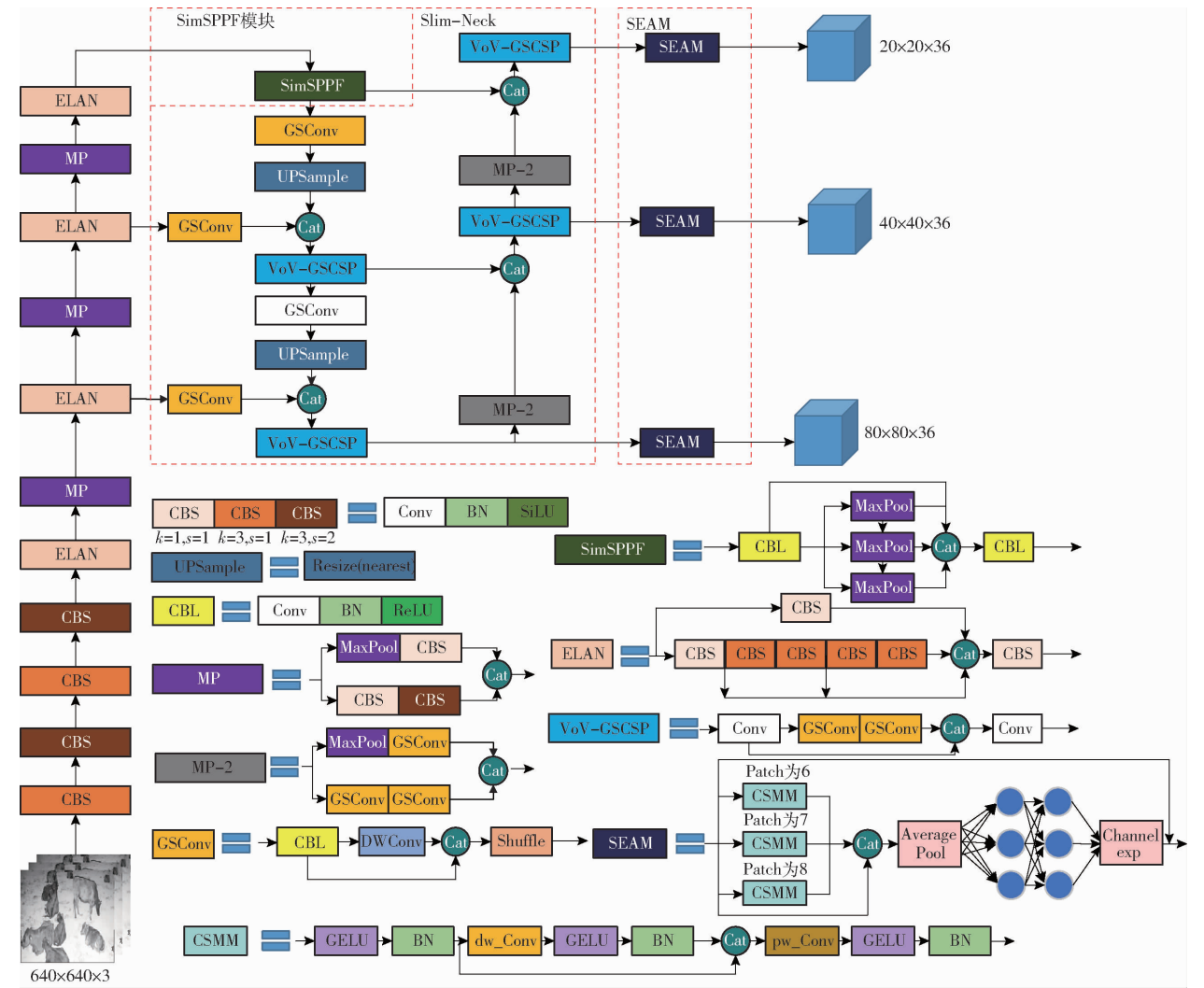


图 3 SNSS - YOLO v7 网络结构

Fig. 3 SNSS - YOLO v7 network structure

原网络的 SPPCSPC 模块,增加感受野的同时进一步减小参数量。

(1) Slim - Neck 结构轻量化颈部

YOLO v7 网络使用大量的标准卷积 (Standard convolution, SC) 提取图像的特征信息,但是标准卷积会耗费大量的时间,轻量化网络 Xception^[17]、MobileNet^[18] 和 ShuffleNet^[19] 等使用的深度可分离卷积 (Depthwise separable convolution, DSC) 可以有效地解决标准卷积的耗时问题,但模型的检测精度也随之降低。GSConv^[20] 采用 SC、DSC 以及 Shuffle 相结合的方式,其计算成本仅是标准卷积的 60% ~ 70%,但效果与 SC 不相上下。因此,本文使用 GSConv 代替原网络中 Neck 层中的标准卷积,并在 GSConv 的基础上引入 VoV - GSCSP 模块^[20]。通过 GSConv 和 VoV - GSCSP 模块所构成的 Slim - Neck 结构对主干网络提取的特征进行增强,在充分提取深层特征和浅层特征信息的同时,既减轻了模型的复杂度,又保持了模型的精度。GSConv 卷积操作和 VoV - GSCSP 网络结构分别如图 4 和图 5 所示。

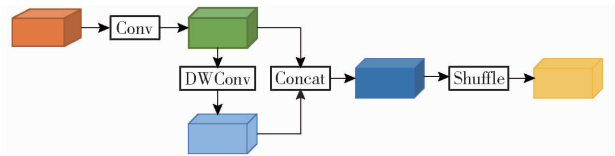


图 4 GSConv 卷积操作
Fig. 4 GSConv convolution operation

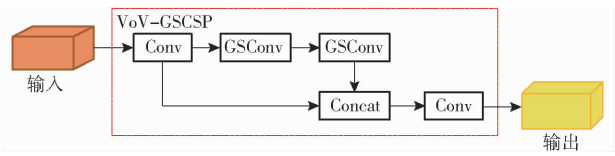


图 5 VoV - GSCSP 网络结构
Fig. 5 VoV - GSCSP structure

(2) SEAM 模块

为进一步降低模型复杂度,本文使用 SEAM 模块替换原 YOLO v7 的 Rep 结构来增强 Neck 层输出部分的响应能力。SEAM 模块^[21] 中使用带有残差连接的深度可分离卷积,采用一个两层的全连接网络来融合每个通道的信息,以加强所有通道之间的联系。通过上一步学习到的被遮挡行为与无遮挡行为的关系来弥补遮挡场景下的损失,可以有效解决行为之间的遮挡问题。

(3) SimSPPF

空间金字塔池化 (Spatial pyramid pooling, SPP)^[22] 是一种将局部特征映射到不同维度空间并将其融合的池化方法,输入不同大小的图像可以产生固定大小的特征向量,进而有效提取出不同尺度的特征信息。

YOLO v7 中的空间金字塔池化使用的是 SPPCSPC 模块(图 6a),将输入分成 2 个不同的分支,其中一个进行常规处理,另一个进行 SPP 结构处理,最后将两部分通过 Concat 操作合并在一起。其中 SPP 结构处理将输入分成 4 部分,其中 3 部分分别进行尺寸为 5×5 、 9×9 和 13×13 的最大池化操作,第 4 部分输入与其他 3 个经过池化操作的输入通过 Concat 连接,生成固定尺寸的特征。SPPCSPC 模块虽然有较好的表现效果,但参数量与计算量过大。

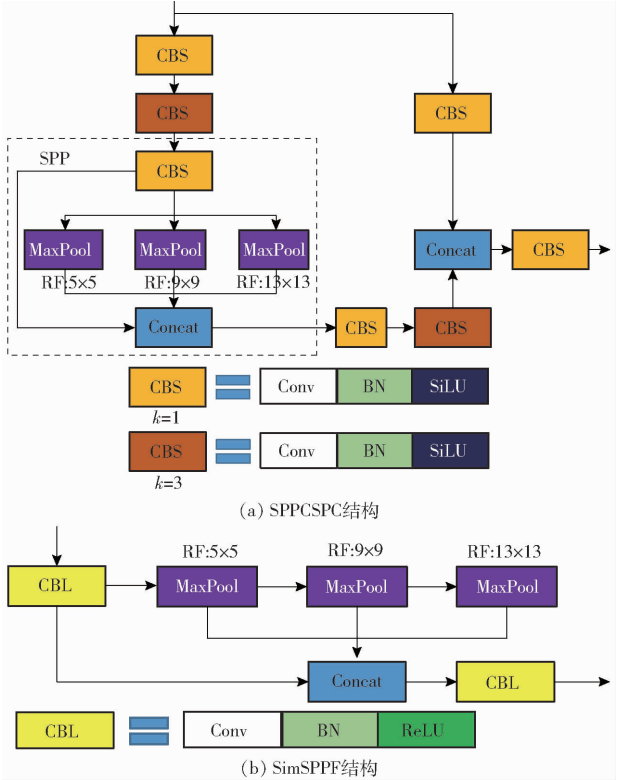


图 6 SPPCSPC 结构和 SimSPPF 结构
Fig. 6 SPPCSPC structure and SimSPPF structure

本文使用 SimSPPF 模块^[23] (图 6b) 将输入分为 2 部分,一部分串联 3 个尺寸为 5×5 的最大池化层 (Max Pooling)。其中,串联 2 个 5×5 的最大池化层和 1 个 9×9 的最大池化层计算结果相同,串联 3 个 5×5 的最大池化层和 1 个 13×13 的最大池化层计算结果相同。通过串联最大池化层对输入进行处理,极大地减少了计算量,提高了速度和效率。另一部分输入直接与经过 3 个池化操作的输出通过 Concat 连接,生成固定尺寸的特征,在很大程度上增加感受野,起到特征增强的作用。在 SimSPPF 模块中使用 ReLU 激活函数,而 SPPCSPC 模块使用 SiLU 激活函数,单个 CBL 速度比 CBS 快。

1.3 试验设计及评价

为验证本文所提方法的有效性,将本文提出的肉牛行为识别方法与 YOLO v6m、YOLO v5m、

YOLOX - S、TPH - YOLO v5 和 Faster R - CNN 方法进行对比;将本文提出的肉牛行为识别原网络与各模块改进后的网络进行对比,验证本文对原算法改进的有效性。

采用精确率 (Precision)、召回率 (Recall)、平均精度均值 ($mAP_{@0.5}$)、模型内存占用量、参数量和浮点运算量 (GFLOPs) 评估肉牛行为识别模型的性能,平均精度 (AP) 是肉牛各个行为识别的平均精度, mAP 是 AP 的平均值,模型内存占用量使用模型所占存储空间来衡量,GFLOPs 表示每秒 10 亿次的浮点运算数。

试验服务器配置如下:操作系统为 Ubuntu 18.04, GPU 为 GeForce RTX 2080Ti, CPU 为 Intel (R) Xeon (R) Silver 4114, 主频 2.2 GHz, 内存 125 GB,深度学习库为 PyTorch 1.7.1, 运算平台为 CUDA 10.0、CUDNN 7.6.1, 编程语言为 Python。

模型训练过程中参数设置有:输入图像尺寸为 640 像素 $\times$ 640 像素、Epoch 为 300、批量大小为 8、初始学习率为 0.01、动量参数为 0.937、权重衰减系数为 0.000 5。

2 试验结果与分析

2.1 肉牛行为识别结果

为验证本文提出的 SNSS - YOLO v7 肉牛行为识别方法对肉牛爬跨、躺卧、探究、站立、运动、舔舐和互斗 7 种行为识别效果的优越性,将本文方法与原 YOLO v7 算法进行比较,对比结果如表 2 所示。

表 2 不同目标平均精度均值对比结果
Tab.2 Comparison result of $mAP_{@0.5}$ of different targets

行为类别	行为数量	$mAP_{@0.5}/\%$	
		本文方法	YOLO v7
爬跨	94	99.6	99.6
躺卧	1 000	99.5	99.6
探究	94	92.6	91.0
站立	581	97.8	96.8
运动	105	88.5	84.1
舔舐	98	94.7	94.1
互斗	49	93.4	91.6
全部行为	2 021	95.2	93.8

由表 2 可以看出,本文方法对爬跨行为识别的平均精度均值 ($mAP_{@0.5}$) 为 99.6%,与 YOLO v7 相同。本文方法对探究、舔舐和互斗行为的识别平均精度均值 ($mAP_{@0.5}$) 分别为 92.6%、94.7% 和 93.4%,较 YOLO v7 分别提高 1.6、0.6、1.8 个百分点。

YOLO v7 对 7 种行为的识别中,运动行为识别的平均精度均值 ($mAP_{@0.5}$) 最低,仅有 84.1%,分析其

原因是肉牛的站立行为和运动行为表现相似,模型容易产生混淆,而误将运动行为识别成站立行为。本文方法对运动行为的识别平均精度均值 ($mAP_{@0.5}$) 为 88.5%,较 YOLO v7 提高 4.4 个百分点,对肉牛站立行为识别的平均精度均值 ($mAP_{@0.5}$) 为 97.8%,较 YOLO v7 提高 1 个百分点。说明本文方法对肉牛运动和站立相似行为识别能力有所提高。

本文提出的肉牛行为识别方法在测试集上的识别结果如图 7 所示。



图 7 肉牛行为识别结果

Fig.7 Results of behavior recognition of beef cattle

图 7a、7d 中肉牛行为数量较少,图 7a 中的躺卧、探究和站立行为均被正确识别,图 7d 中的爬跨行为被正确识别。图 7b、7e 为自然光照条件下存在多只肉牛且存在遮挡行为的图像,图 7b 中由于站立与运动行为高度相似,导致其中一个运动行为被错误识别成站立行为,其余站立行为与运动行为均被正确识别,图 7e 中的躺卧、站立和探究行为均被正确识别。图 7c、7f 为在夜晚存在多只肉牛且存在遮挡行为的图像,夜晚肉牛行为多表现为躺卧行为,图 7c 中的躺卧和探究行为均全部被正确识别,图 7f 中的躺卧行为均被正确识别。总体来说,本文方法对肉牛的多种行为均取得较好的识别效果。

2.2 消融试验结果

为验证提出的基于 SNSS - YOLO v7 的肉牛行为识别方法对原算法改进的有效性,设计了消融试验。将原网络与进行各模块改进后的网络在测试集上进行测试,对比结果如表 3 所示。

表 3 消融试验结果

Tab.3 Results of ablation experiments

模型	精确率/%	mAP _{@0.5} /%	参数量	浮点运算量	模型内存占用量/MB
YOLO v7	91.1	93.8	3.651×10^7	103.3	74.90
YOLO v7 + Slim - Neck	91.5	94.1	3.112×10^7	89.2	64.20
YOLO v7 + Slim - Neck + SEAM	91.3	94.2	2.529×10^7	78.5	51.10
本文方法	91.7	95.2	1.926×10^7	73.7	39.00

由表 3 可以看出,本文方法所做的第 1 个改进为使用 Slim - Neck 轻量化 Neck 结构,使精确率提高 0.4 个百分点,平均精度均值($mAP_{@0.5}$)提高 0.3 个百分点,模型内存占用量减小 14.3%,参数量减少 14.8%,浮点运算量减少 13.6%;本文方法所做的第 2 个改进为使用 SEAM 模块替换 Rep 结构,虽然精确率并未提升,但平均精度均值($mAP_{@0.5}$)提高 0.1 个百分点,模型内存占用量减小 20.4%,参数量减少 18.7%,浮点运算量减少 12%;本文方法所做的第 3 个改进为使用 SimSPPF 模块替换 SPPCSPC 模块,精确率提高 0.4 个百分点,平均精度均值($mAP_{@0.5}$)提高 1 个百分点,模型内存占用量减小 23.7%,参数量减少 23.8%,浮点运算量减少 6.1%。

2.3 不同模型对比试验结果

为验证本文提出的基于 SNSS - YOLO v7 的肉牛行为识别方法的有效性,将本文方法与 YOLO v6m^[23]、YOLO v5m、YOLOX - S^[24]、TPH - YOLO v5^[25]和 Faster R - CNN^[26]肉牛行为识别方法进行对比。其中,YOLO v6m、YOLO v5m、YOLOX - S 方法是当前主流的目标检测算法,TPH - YOLO v5 和 Faster R - CNN 是在其他文献中性能较好的算法。为了验证本文方法的优越性,所有模型训练过程中参数设置均保持一致。对比结果如图 8 所示。

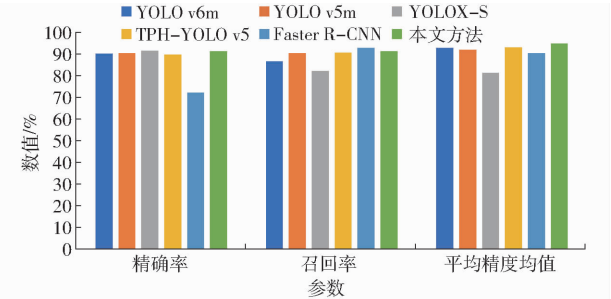


图 8 不同模型识别效果比较

Fig.8 Comparison of recognition effect of different models

由图 8 可以看出,本文方法与 YOLO v6m、YOLO v5m 和 TPH - YOLO v5 方法相比,精确率分别提高 1.1、0.8、1.7 个百分点,召回率分别提高 4.3、1.1、0.4 个百分点,平均精度均值($mAP_{@0.5}$)分别提高 2.2、3.1、1.9 个百分点。与 YOLOX - S 方法相比,精确率持平,召回率提高 9 个百分点,平均精

度均值($mAP_{@0.5}$)提高 13.7 个百分点。与 Faster R - CNN 方法相比,召回率虽有所下降,但精确率提高 19.3 个百分点,平均精度均值($mAP_{@0.5}$)提高 4.5 个百分点。综上,本文方法在识别效果上优于其他方法,验证了本文方法的优越性。

为进一步验证本文方法是否适合部署在边缘端设备上,将本文方法的参数量和模型内存占用量与 YOLO v6m、YOLO v5m、YOLOX - S、TPH - YOLO v5、Faster R - CNN 模型进行对比。对比结果如表 4 所示。

表 4 不同模型参数量和模型内存占用量比较

Tab.4 Comparison of different model parameters and model sizes

模型	参数量	模型内存占用量/MB
YOLO v6m	3.424×10^7	71.40
YOLO v5m	2.088×10^7	42.20
YOLOX - S	8.94×10^6	68.54
TPH - YOLO v5	4.541×10^7	92.50
Faster R - CNN	1.3683×10^8	521.90
本文方法	1.926×10^7	39.00

由表 4 可知,本文方法与 YOLO v6m 方法相比模型内存占用量减少 45.4%,参数量减少 43.8%;与 YOLO v5m 方法相比,模型内存占用量减少 7.6%,参数量减少 7.8%;与 YOLOX - S 方法相比,参数量虽然略有提升,但模型内存占用量减少 43.1%;与 TPH - YOLO v5 方法相比,模型内存占用量减少 57.8%,参数量减少 57.6%;与 Faster R - CNN 方法相比,模型内存占用量减少 92.5%,参数量减少了 85.9%。综上,本文方法在识别性能方面优于对比方法,验证了本文方法的有效性。

3 讨论

3.1 牛只不同密集程度下行为识别效果

为了验证本文提出的基于 SNSS - YOLO v7 的肉牛行为识别方法在肉牛不同密集程度下的识别效果,分别对轻微、中度和重度 3 种密集程度进行了测试,测试结果如表 5 所示。轻微密集程度的肉牛数量少于 5 头,且肉牛行为之间几乎无遮挡;中度密集程度的肉牛数量在 5 ~ 10 头之间,且部分肉牛行为

之间存在遮挡;重度密集程度下的肉牛数量大于 10 头,且肉牛行为之间互相遮挡。

表 5 不同密集程度下肉牛行为识别效果对比

Tab.5 Comparison of behavior recognition effect under different intensities

密集程度	精确率	mAP _{@0.5}
轻微密集	98.3	97.0
中度密集	87.4	94.3
重度密集	83.6	92.7

由表 5 可知,轻微密集和中度密集程度下,本文方法可以较准确地识别出肉牛行为;在重度密集情况下,肉牛行为的识别效果受到了较大的影响,因肉牛行为之间互相遮挡严重,导致精确率和 mAP_{@0.5}有所下降。

图 9 为不同密集程度的肉牛行为识别结果。图 9a 为轻微密集环境,肉牛站立行为和运动行为均被正确识别。图 9b 为中度密集环境,肉牛站立、舔舐和躺卧行为均被准确识别。图 9c 为重度密集环境,两只肉牛因遮挡严重导致躺卧行为被漏检,其它探究、躺卧、站立、舔舐行为均被正确识别。

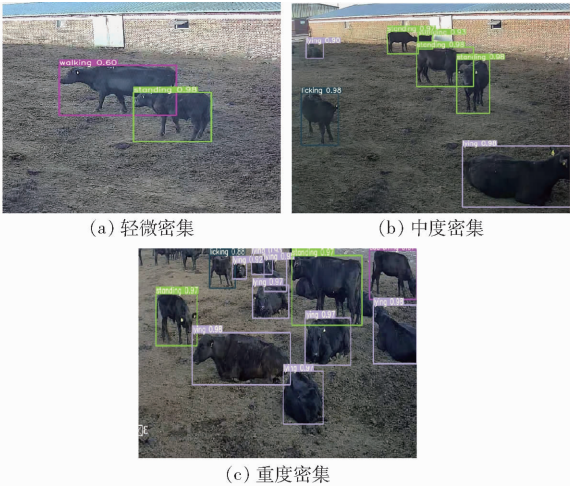


图 9 不同密集程度下的肉牛行为识别结果

Fig.9 Results of beef cattle behavior recognition under different intensity levels

3.2 不同光照情况下行为识别效果

为了验证本文提出的基于 SNSS - YOLO v7 的肉牛行为识别方法的实用性,对不同光照情况下肉牛行为的识别效果进行了测试,测试结果如图 10 所示。

图 10a、10b 光照较强,图像中的肉牛行为均被正确识别。图 10c 中光线较弱且肉牛重度密集,导致图像中的两个站立行为被漏检。图 10d 中光线较暗且肉牛中度密集,图像中的行为均被正确识别。

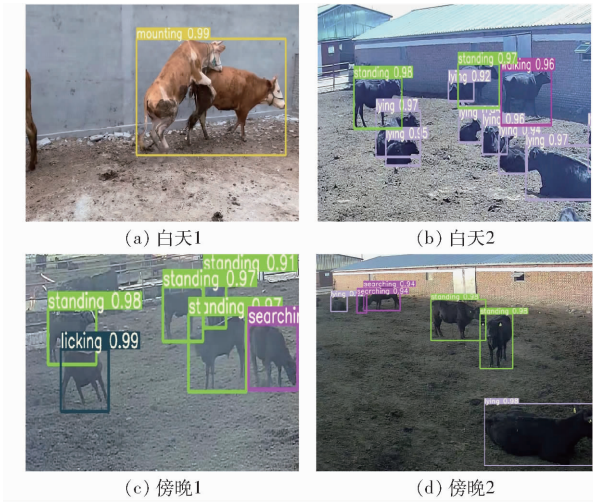


图 10 不同光照情况下肉牛行为识别结果

Fig.10 Recognition results of beef cattle behavior under different lighting conditions

3.3 站立、运动和探究行为分析

站立、运动和探究行为虽然相似,但可以反映肉牛的健康状况,比如,患病的肉牛运动行为减少,这些相似行为对识别精度有一定影响。如果将站立、运动和探究行为合并为一类,则肉牛行为可分为爬跨、躺卧、舔舐、互斗及其他 5 类,测试结果如表 6 所示。

表 6 不同行为类别数量识别效果比较

Tab.6 Comparison of recognition effect of different behavior categories

行为类别数量	精确率	mAP _{@0.5}
5 类	96.0	97.4
7 类	91.7	95.2

由表 6 可知,将站立、运动和探究 3 类相似行为归为一类,精确率大幅提高,但无法准确识别肉牛行为是站立、运动还是探究,导致无法及时掌握肉牛的健康状况。详细区分站立、运动和探究 3 类相似行为,精确率和 mAP_{@0.5}虽然略有下降,但仍然全部在 90% 以上,可以及时发现肉牛养殖过程中存在的异常情况,对于肉牛健康养殖具有重要作用。

4 结论

(1)提出了一种基于 SNSS - YOLO v7 的肉牛行为识别方法。首先在 YOLO v7 的 Neck 部分使用 Slim - Neck 结构,有效降低模型复杂度;其次在 YOLO v7 的 Head 引入 SEAM 模块,增强对 Neck 层输出的响应能力;最后使用 SimSPPF 模块替换原 YOLO v7 的 SPPCSPC 模块,在减小参数的同时进一步增大感受野。试验结果表明,本文方法能有效识别肉牛的爬跨、躺卧、探究、站立、运动、舔舐和互斗

7 种行为,平均精度均值为 95.2%,模型内存占用量为 39 MB,参数量为 1.926×10^7 ,识别效果较好。模型内存占用量分别减小 47.9%、45.4%、7.6%、43.1%、57.8% 和 92.5%,平均精度均值分别提高 1.4、2.2、3.1、13.7、1.9、4.5 个百分点。表明本文方法行为识别效果更好。

(2) 与 YOLO v7、YOLO v6m、YOLO v5m、YOLOX-S、TPH-YOLO v5 和 Faster R-CNN 相比,

参 考 文 献

[1] 许金奎. 肉牛养殖技术及饲养管理措施[J]. 中国畜牧业, 2022(24): 69-70.
XU Jinkui. Beef cattle breeding technology and feeding management measures[J]. China Animal Industry, 2022(24): 69-70. (in Chinese)

[2] 陈浩,王纯洁,斯木吉德,等. 牛肉品质及其影响因素研究进展[J]. 动物营养学报, 2021, 33(2): 669-678.
CHEN Hao, WANG Chunjie, SI Mujide, et al. Research progress of beef quality and its influencing factors[J]. Chinese Journal of Animal Nutrition, 2021, 33(2): 669-678. (in Chinese)

[3] QUANG T H, KHANH P, BUI N, et al. Cow behavior monitoring using a multidimensional acceleration sensor and multiclass SVM[J]. International Journal of Machine Learning and Networked Collaborative Engineering, 2018, 2: 110-118.

[4] WANG J, ZHANG H, JI J, et al. Development of a wireless measurement system for classifying cow behavior using accelerometer data and location data[J]. Applied Engineering in Agriculture, 2019, 35(2): 135-147.

[5] LI Z Y, SONG L, DUAN Y C, et al. Basic motion behaviour recognition of dairy cows based on skeleton and hybrid convolution algorithms[J]. Computers and Electronics in Agriculture, 2022, 196: 106889.

[6] 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)

[7] BARKER Z E, VÁZQUEZ DIOSDADO J A, CODLING E A, et al. Use of novel sensors combining local positioning and acceleration to measure feeding behavior differences associated with lameness in dairy cattle[J]. Journal of Dairy Science, 2018, 101(7): 6310-6321.

[8] CHELOTTI J O, VANRELL S R, MARTINEZ R L S, et al. An online method for estimating grazing and rumination bouts using acoustic signals in grazing cattle[J]. Computers and Electronics in Agriculture, 2020, 173: 105443.

[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] JIANG B, SONG H, WANG H, et al. Dairy cow lameness detection using a back curvature feature[J]. Computers and Electronics in Agriculture, 2022, 194: 106729.

[11] 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)

[12] 王少华,何东健. 基于改进 YOLO v3 模型的奶牛发情行为识别研究[J]. 农业机械学报, 2021, 52(7): 141-150.
WANG Shaohua, HE Dongjian. Dysfunctional behavior recognition of dairy cows based on improved YOLO v3 model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 141-150. (in Chinese)

[13] QIAO Y, GUO Y, YU K, et al. C3D-ConvLSTM based cow behaviour classification using video data for precision livestock farming[J]. Computers and Electronics in Agriculture, 2022, 193: 106650.

[14] 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.

[15] YIN X, WU D, SHANG Y, et al. Using an EfficientNet-LSTM for the recognition of single cow's motion behaviours in a complicated environment[J]. Computers and Electronics in Agriculture, 2020, 177: 105707.

[16] MA S, ZHANG Q, LI T, et al. Basic motion behavior recognition of single dairy cow based on improved Rexnet 3D network[J]. Computers and Electronics in Agriculture, 2022, 194: 106772.

[17] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017.

[18] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [EB/OL]. [2017-04-17]. <https://arxiv.org/abs/1704.04861>.

[19] MA N, ZHANG X, ZHENG H, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[C]//15th European Conference on Computer Vision (ECCV), 2018.

[21] 汪攀峰, 丁启朔. 黏土单层击实试验与制样因子研究[J]. 岩土力学, 2010,31(6):1797–1802.
WANG Panfeng, DING Qishuo. Study of dynamic compaction of monolayer cohesive soil and factors for remolded soil[J]. Rock and Soil Mechanics, 2010, 31(6):1797–1802. (in Chinese)

[22] 李笑吟, 毕华兴, 张志, 等. 晋西黄土区土壤水分有效性分析的克立格法[J]. 土壤学报, 2006, 43(6):1004–1010.
LI Xiaoyin, BI Huaxing, ZHANG Zhi, et al. Kriging theory on the research of soil moisture validation in the Loess Plateau [J]. Acta Pedologica Sinica, 2006, 43(6):1004–1010. (in Chinese)

[23] XIONG P, ZHANG Z, WANG Y, et al. Variable responses of maize roots at the seedling stage to artificial biopores in noncompacted and compacted soil[J]. Journal of Soils and Sediments, 2022,22(4):1155–1164.

[24] STIRZAKER R J, PASSIOURA J B, WILMS Y. Soil structure and plant growth: impact of bulk density and biopores[J]. Plant and Soil, 1996,185:151–162.

[25] ATKINSON J A, HAWKESFORD M J, WHALLEY W R, et al. Soil strength influences wheat root interactions with soil macropores[J]. Plant Cell and Environment, 2020,43(1):235–245.

[26] ŠIMANSKY V. Soil structure as a significant indirect factor affecting crop yields[J]. Acta Fytotechnica et Zootechnica, 2021, 24:129–136.

[27] 葛双洋. 水稻土深松对土壤物理性质及小麦生长影响的研究[D]. 南京: 南京农业大学, 2017.
GE Shuangyang. Effect of subsoiling on paddy soil physical properties and growth of wheat[D]. Nanjing: Nanjing Agricultural University, 2017. (in Chinese)

[28] 曾健, 费良军, 陈琳, 等. 添加 γ -聚谷氨酸对土壤结构及持水特性的影响[J]. 水土保持学报, 2018,32(1):217–224.
ZENG Jian, FEI Liangjun, CHEN Lin, et al. Effect of γ -PGA on soil structure and water-holding characteristics[J]. Journal of Soil and Water Conservation, 2018,32(1):217–224. (in Chinese)

[29] 孙梅, 黄运湘, 孙楠, 等. 农田土壤孔隙及其影响因素研究进展[J]. 土壤通报, 2015,46(1):233–238.
SUN Mei, HUANG Yunxiang, SUN Nan, et al. Advance in soil pore and its influencing factors[J]. Chinese Journal of Soil Science, 2015,46(1):233–238. (in Chinese)

[30] 刘继龙, 李佳文, 周延, 等. 秸秆覆盖与耕作方式对土壤水分特性的影响[J]. 农业机械学报, 2019,50(7):333–339.
LIU Jilong, LI Jiawen, ZHOU Yan, et al. Effects of straw mulching and tillage on soil water characteristics[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019,50(7):333–339. (in Chinese)

(上接第 274 页)

[20] LI H, LI J, WEI H, et al. Slim-neck by GSConv: a better design paradigm of detector architectures for autonomous vehicles [EB/OL]. [2022–08–17]. <https://arxiv.org/abs/2206.02424>.

[21] YU Z, HUANG H, CHEN W, et al. YOLO–FaceV2: a scale and occlusion aware face detector[EB/OL]. [2022–08–03]. <https://arxiv.org/abs/2208.01944>.

[22] 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.

[23] LI C, LI L, JIANG H, et al. YOLOv6: a single-stage object detection framework for industrial applications[EB/OL]. [2022–09–07]. <https://arxiv.org/abs/2209.02976>.

[24] GE Z, LIU S, WANG F, et al. YOLOX: exceeding YOLO series in 2021[EB/OL]. [2021–08–06]. <https://arxiv.org/abs/2107.08430>.

[25] ZHU X, LYU S, WANG X, et al. TPH–YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios [C] // 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), 2021.

[26] FUENTES A, YOON S, PARK J, et al. Deep learning-based hierarchical cattle behavior recognition with spatio-temporal information[J]. Computers and Electronics in Agriculture, 2020, 177:105627.