

基于改进 YOLOv5 网络的牧区牛粪检测方法

那日苏, 翟立浩, 邢海鹏, 张 泉

(内蒙古工业大学机械工程学院, 呼和浩特 010010)

摘要: 中国草原面积广阔, 牛粪作为草原畜牧业的副产品, 既对牧草生长有负面影响, 又是重要的资源。牛粪分布零散, 传统捡拾方式效率低, 为了解决上述问题, 该文提出一种改进牛粪图像检测模型, 用于检测牧区牛粪, 以提高牛粪捡拾的效率和智能化水平。1) 替换注意力机制轻量化主干网络 EfficientFormerV2, 提高对边界敏感性并减少模型复杂度。2) 改进颈部网络 BiFPN, 增强特征融合能力。3) 改进损失函数 Inner-IoU, 提高边界框回归的定位精度。4) 替换瓶颈层改进 FasterNet block, 以减少计算复杂度并加快推理速度。5) 添加捡拾判断机制, 利用预测阶段检测框对区域牛粪做出分类。改进后的网络模型在准确率为 92.6%, 召回率为 87.7%, 平均精度均值为 87.4% 的情况下, 参数量减少到 4.02M, 浮点运算量减少到 8.1G, 检测速度提升到 34.7 帧/s。试验结果表明, 改进模型在保持模型轻量化的同时, 具有较高的平均精度均值, 能够有效地完成牧区牛粪的识别与定位任务, 适合在资源受限的移动设备和牧区环境中应用, 为牛粪智能化捡拾车的研究提供技术支持。

关键词: 图像处理; 粪污; 卷积神经网络; 牛粪捡拾

doi: 10.11975/j.issn.1002-6819.202405003

中图分类号: S126

文献标志码: A

文章编号: 1002-6819(2024)-22-0145-09

那日苏, 翟立浩, 邢海鹏, 等. 基于改进 YOLOv5 网络的牧区牛粪检测方法[J]. 农业工程学报, 2024, 40(22): 145-153.

doi: 10.11975/j.issn.1002-6819.202405003 <http://www.tcsae.org>

NA Risu, ZHAI Lihao, XING Haipeng, et al. Detecting cattle manure in pastoral areas using the YOLO network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(22): 145-153. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202405003 <http://www.tcsae.org>

0 引言

中国的草原面积达 3.9 亿 hm^2 , 占国土面积的 40.9%^[1]。草原畜牧业是草原地区畜牧业的主要存在形式, 草原畜牧业是指以天然草原为放牧地、以草原植被为主要饲料源的家畜放牧饲养业, 草原畜牧业以牛羊活畜及其产品为主要产出。牛粪作为草原畜牧业的副产品, 分布于广袤的草原上。一方面由于草原地区气候干旱, 牛粪不易降解, 干燥的牛粪覆于草地上, 阻断了阳光与空气, 若不及时清理影响牧草生长。大量排放的畜禽粪便造成的环境污染问题日益加剧^[2]。另一方面, 牛粪也可以在微生物的作用下将有机物转化为稳定的腐殖质, 适合作为有机肥^[3], 同时牛粪也是重要的有潜力的燃料资源。

牛粪在天然草场的分布主要表现为大区域零散, 小区域聚集。牧区牛粪捡拾收集方式主要为手工捡拾, 草场中存在的大量生长中牧草决定了掘土类的收获机械无法在牧区使用。但随着计算机视觉技术的成熟, 可以将计算机视觉技术运用于牛粪的捡拾收集工作中。牛粪捡拾收集设备的工作原理为, 牛粪捡拾机通过与拖拉机悬挂连接提供动力, 牛粪捡拾机上的视觉相机捕捉草场画

面, 控制中心对相机画面进行处理, 当检测到大块牛粪或连片牛粪出现时, 捡拾爪启动下放收集牛粪。智能化捡拾有助于提升工作效率, 助力牧业现代化。

深度学习在目标检测领域表现出色, 目标检测方法可以分为两类: 单阶段目标检测和两阶段目标检测^[4]。一类是基于候选窗口的目标检测框架 (Two Stage), 代表有 R-CNN^[5]、Fast R-CNN^[6]、Faster R-CNN^[7] 等。在预测阶段首先选出包含目标概率较大的区域, 之后对所选区域进行预测。这类框架检测的精度比较高, 但检测速度比较慢^[8]。另一类是基于回归的目标检测框架 (One Stage), 代表有 YOLO^[9]、SSD^[10] 等。将检测问题看成回归问题。不需要选取候选框, 预测一步完成。这类框架检测速度比较快, 但在精度方面一般弱于基于候选窗口的目标检测框架^[11]。针对牧区牛粪识别问题所要求的实时监测以及移动设备应用, 检测速度更快的 One Stage 框架能更好的完成任务^[12]。YOLO 算法在 2015 年由 Joseph Redmon 提出, YOLO 的核心思想是把目标检测转变成一个回归问题, 利用整张图作为网络的输入, 再经过一个神经网络之后, 得到边界框的位置及其所属的类别^[13]。YOLOv5 为 YOLO 网络的第五代版本, 其主要由输入端、主干网络、颈部网络、检测头四部分组成。在输入端, 对输入图片进行 Mosaic 数据增强^[14]、计算锚框和统一图片尺寸。主干网络采用 New CSP-Darknet53^[15], 主要作用为提取图片信息。颈部网络采用 PANET^[16] 网络结构, 将图像高层的强语义特征与图像低层的强定位

收稿日期: 2024-05-13 修订日期: 2024-10-29

基金项目: 内蒙古自治区科技计划项目 (2022YFXZ0018); 内蒙古自治区重点研发和成果转化计划 (2023YFDZ0061)

作者简介: 那日苏, 博士, 副教授, 硕士生导师, 研究方向为机械系统动力学。Email: nrs3000@imut.edu.cn

特征融合, 进一步提高了神经网络的特征学习能力。检测头^[17]输出 3 种不同尺寸的特征图, 由分类损失, 目标损失和定位损失组成 YOLOv5 的总损失。

化肥和粪便造成的营养污染是全球关注的问题, 增加有关粪肥管理实践的数据是解决全球营养污染问题的重要一步。BRYAN 等^[18]使用电磁感应测量法收集饲料场表面的表观电导率数据, 验证了电磁感应数据与土壤中的氯、氮、磷和挥发性固体含量之间存在相关性, 利用拟合模型精确地识别粪便在饲料场表面的积累区域。KELLY 等^[19]使用随机森林模型, 利用雷达和光学卫星图像进行机器学习, 提供了一种在大范围内跟踪液体肥料管理模式的方法, 模型成功区分了玉米、大豆、草地和其他作物的饱和与非饱和土壤。目前中国养殖产业正处于智慧畜牧业初级阶段, 智能化技术正在逐步运用于畜禽养殖产业^[20]。胡振楠等^[21]采集包括温度、湿度、二氧化碳、氨气、硫化氢等环境参数, 建立清粪次数、清粪时间和环境参数关系模型, 可实现实时预测猪舍内粪污量并及时进行清理。

上述国内外研究虽然利用深度学习在粪污管理方面进行了大量研究, 但部分关注于大尺度遥感影像, 部分关注于养殖场粪污管理, 不适用于牧区牛粪检测任务中准确快速的识别, 分类与定位的要求。针对上述问题, 本文提出一种基于改进 YOLOv5 网络的牧区牛粪检测方法, 通过替换注意力机制轻量化主干网络, 改进特征融合网络, 改进损失函数, 替换瓶颈层, 添加捡拾判断机制, 以期改进后的模型大小和运算量减少, 检测速度增加, 提升准确率、召回率及平均精度均值, 能够实现牧区牛粪检测任务, 为牛粪智能化捡拾车的研究提供技术支持。

1 试验材料

1.1 数据采集与预处理

本文图像数据采集于内蒙古自治区锡林郭勒盟阿巴嘎旗草原牧场, 分别为不同光照、影子遮盖、牧草遮挡、不同干湿程度和不同尺寸形状随机采集的牛粪图像。拍摄形式为手持拍摄, 拍摄位置为垂直拍摄, 设备与牛粪距离为 80~100 cm 以模拟牛粪捡拾车拍摄效果。为了提高数据多样性, 增强模型的泛化能力与鲁棒性, 对采集的数据通过翻转、缩放平移、运动模糊、随机遮挡、更改亮度 5 种方法随机组合的方式进行数据增广, 最终得到有效图像 1 320 张, 处理前后图像对比如图 1 所示。以 7:2:2 的比例将其划分为训练集 840 张、验证集 240 张、测试集 240 张, 并使用 LabelImg 标签工具对图像进行标注。

1.2 模型训练与评价标准

本文试验运行平台采用 Windows10 操作系统, 试验平台处理器 12th Gen Intel(R) Core(TM) i5-12490F 3.00 GHz, 16.0GB 运行内存, 显卡型号 NVIDIA GeForce RTX 3060 Ti, 8.0GB 显存, 软件环境为 Python 3.11, 深度学习框架为 PyTorch 2.0.1。

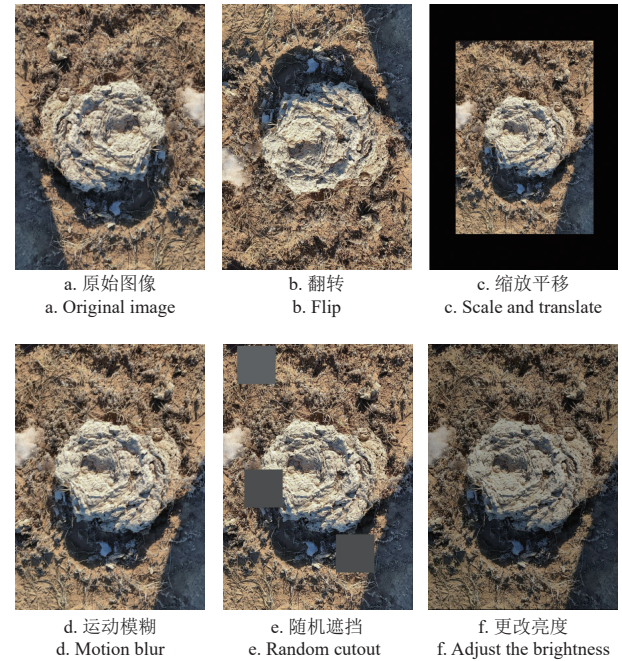


图 1 牛粪图像数据增广

Fig.1 Augmentation of cow dung image data

训练网络所使用的数据集为牧区牛粪图像数据集, 输入图像分辨率为 4 096×3 072 像素, 批量大小 (batchsize) 设置为 8, 训练次数 (epoch) 设置为 500, 参数更新使用随机梯度下降 (SGD) 方法, 学习率设置为 0.01。

为评价改进网络的性能, 采用精确率 (precision)、召回率 (recall)、平均精度均值 (mean average precision)、参数量 (params)、浮点运算量 (floating point operations)、检测速度 (frames per second) 共 6 个指标对模型进行评价。

2 改进 YOLOv5 目标检测算法

2.1 替换主干网络

针对牧区牛粪检测任务, 其存在有牛粪边缘与土地边界不明显且牧草遮盖及在移动端部署使用的特点, 既需要网络对边界处敏感又需要网络轻量化。注意力机制能够以高权重聚焦重要信息, 以低权重忽略不相关的信息, 将注意力集中在图片中特定的部分, 可以有效解决牛粪与土地边界模糊的问题。然而注意力机制由于大量的参数和其模型设计, 使得其速度通常比轻量级网络慢几倍, EfficientFormerV2^[22-23]以纯 Transformer 为设计范例, 以延迟为目标进行优化设计, 在保证准确率的同时, 增加了推理速度。故在本文中用 EfficientFormerV2 替换 New CSP-Darknet53 作为主干网络提取特征。

EfficientFormerV2 的网络结构如图 2 所示, 采用四阶段分层式设计, 四阶段分别输出 1/4, 1/8, 1/16, 1/32 分辨率的尺寸图。输入图像首先经过 CONV stem 由两个 3×3 步长为 2 的卷积核构成, 计算如式 (1) 所示,

$$X_{ij=1, j=1}^{B,C,j=1, \frac{H}{4}, \frac{W}{4}} = \text{stem}(X_0^{B,3,H,W}) \quad (1)$$

式中 X_0 表示输入图像, B 表示批量大小 (batch size), C 表示通道维数, H 表示特征图高度, W 表示特征图宽

度, i 表示网络层数, j 表示输入变量。stem 表示网络的第一个模块, 负责对输入数据进行初步的特征提取和降维处理。前两个阶段使用统一的前馈神经网络^[24] (feed forward neural network, FFN) 以高分辨率捕捉特征图信息, 该 FFN 使用深度可分离卷积^[25] (DWCONV) 代替显式的残留连接的局部令牌混合器 (token mixers), 且每个 FFN 之间都用残差连接, 计算如式 (2) 所示,

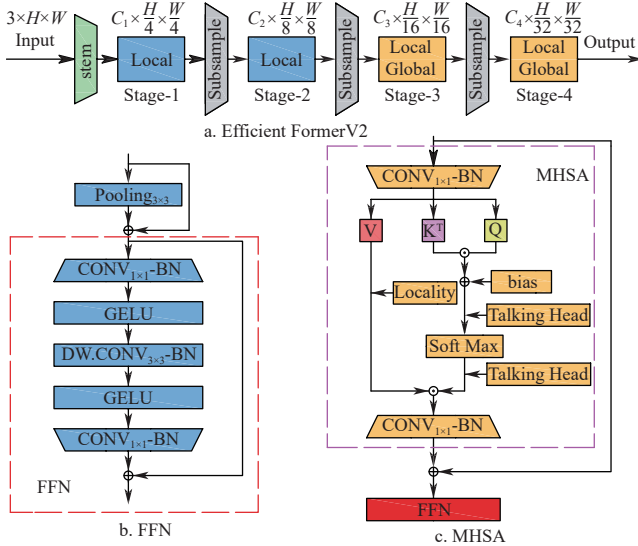
$$X_{i+1,j}^{B,C_j, \frac{H}{2^{j+1}}, \frac{W}{2^{j+1}}} = S_{i,j} \cdot \text{FFN}^{C_j, E_{i,j}}(X_{i,j}) + X_{i,j} \quad (2)$$

式中 $S_{i,j}$ 是可学习层的尺度, 每个阶段的通道维数 C_j 和每个 FFN 的扩展比 $E_{i,j}$ 是 FFN 的两个性质。在最后两个阶段中, 同时使用局部 FFN 和全局 MHSA block (多头注意力), 计算如式 (3)、式 (4) 所示,

$$X_{i+1,j}^{B,C_j, \frac{H}{2^{j+1}}, \frac{W}{2^{j+1}}} = S_{i,j} \cdot \text{MHSA}(\text{Proj}(X_{i,j})) + X_{i,j} \quad (3)$$

$$\text{MHSA}(Q, K, V) = \text{Softmax}(Q \cdot K^T + ab) \cdot V \quad (4)$$

式中 MHSA 表示多头自注意力机制, Softmax 是一种将数值向量归一化为概率分布向量的激活函数, Q, K, V 表示通过线性层 $\text{Proj}(X_{i,j})$ 由输入特征投影学习的查询值、键值和权重值, ab 表示位置编码的可学习注意力偏差, 为在不增加模型大小和延迟的情况下提高注意模块性能的技术, 首先通过添加 3×3 的卷积将局部信息融入到 V 矩阵中, 其次在 Head 维度之间添加全连接层 (即 talking head), 增强不同头之间的信息交互。



注: C 代表特征图的通道数; H 代表特征图的高度; W 代表特征图的宽度; Pooling 表示池化层; CONV 表示卷积层; BN 表示批量归一化操作; GELU 表示高斯误差线性单元激活函数; DW.CONV 表示深度可分离卷积。
Note: C represents the number of channels in the feature map; H represents the height of the feature map; W represents the width of the feature map; Pooling represents the pooling layer; CONV stands for convolutional layer; BN stands for batch normalization operation; GELU represents the Gaussian Error Linear Unit activation function; DW. CONV stands for deep separable convolution.

图 2 EfficientFormerV2 及其关键模块网络模型结构

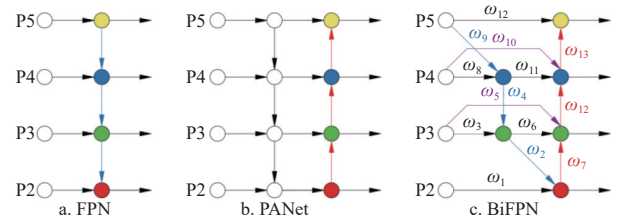
Fig.2 EfficientFormerV2 and its key module network model structure

2.2 改进特征融合网络

YOLOv5 的颈部采用的是 PANet 网络结构, 它是在

自顶而下横向连接将高低层特征融合的 FPN 网络基础上, 增加了自底而上的 PAN 网络。PANet 引入了自底向上的路径, 利用准确的低层定位信号增强整个特征层次, 从而缩短了低层与顶层特征之间的信息路径, 有效提高了网络的特征融合能力。但直接加在 FPN 网络之后的 PAN 网络输入全部来自于 FPN 网络处理之后的特征信息, 会使得原始特征信息存在部分丢失, 影响检测准确性。并且 PANet 网络存在的单边输入节点增加较多模型的参数却只获取了较少的信息。

BiFPN^[26] 全称为双向特征金字塔网络 (bidirectional feature pyramid network), 网络原理如图 3 所示, 是一种改进版的 FPN 网络结构, BiFPN 在 PANet 网络的基础上删除了单边输入节点, 并添加了一同层由原始输入直接特征融合的通路, 这使得各节点之间同时具备自顶向下、自底而上、跨跃式连接的特点。针对于牧区牛粪检测任务, BiFPN 网络可以使不同尺度特征信息更好地融合利用, 解决了原始网络提取不同尺度的特征时存在信息丢失的问题。



注: P2, P3, P4, P5 表示 1/4, 1/8, 1/16, 1/32 分辨率的四阶段尺寸图; 箭头部分表示传递各层级信息的通路; ω 表示不同阶段学习到的用于区分特征重要程度的权重参数。

Note: P2, P3, P4, P5 represent four-stage dimensional drawings with 1/4, 1/8, 1/16, 1/32 resolutions; The arrow represents the pathway for transmitting information at each level; ω represents the weight parameters learned at different stages to distinguish the importance of features.

图 3 特征融合网络模型原理

Fig.3 Feature fusion network model principle

传统的特征融合是将不同分辨率的特征图以相同权重进行叠加, 但不同分辨率的特征图对输出特征的贡献是不平等的。为平衡每一个输入特征, BiFPN 结构采用快速归一化融合 (fast normalized fusion) 的加权融合方式为输入添加了一个额外的权重, 具体的融合过程如式 (5) ~ 式 (10) 所示。

$$P_2^{\text{out}} = \text{Conv} \left(\frac{\omega_1 \cdot P_2^{\text{in}} + \omega_2 \cdot \text{Resize}(P_3^{\text{in}})}{\omega_1 + \omega_2 + \varepsilon} \right) \quad (5)$$

$$P_3^{\text{td}} = \text{Conv} \left(\frac{\omega_3 \cdot P_3^{\text{in}} + \omega_4 \cdot \text{Resize}(P_4^{\text{td}})}{\omega_3 + \omega_4 + \varepsilon} \right) \quad (6)$$

$$P_3^{\text{out}} = \text{Conv} \left(\frac{\omega_5 \cdot P_3^{\text{in}} + \omega_6 \cdot P_3^{\text{td}} + \omega_7 \cdot \text{Resize}(P_2^{\text{td}})}{\omega_5 + \omega_6 + \omega_7 + \varepsilon} \right) \quad (7)$$

$$P_4^{\text{td}} = \text{Conv} \left(\frac{\omega_8 \cdot P_4^{\text{in}} + \omega_9 \cdot \text{Resize}(P_5^{\text{in}})}{\omega_8 + \omega_9 + \varepsilon} \right) \quad (8)$$

$$P_4^{\text{out}} = \text{Conv} \left(\frac{\omega_{10} \cdot P_4^{\text{in}} + \omega_{11} \cdot P_4^{\text{td}} + \omega_{12} \cdot \text{Resize}(P_3^{\text{out}})}{\omega_{10} + \omega_{11} + \omega_{12} + \varepsilon} \right) \quad (9)$$

$$P_5^{\text{out}} = \text{Conv} \left(\frac{\omega_{12} \cdot P_5^{\text{in}} + \omega_{13} \cdot \text{Resize}(P_4^{\text{out}})}{\omega_{12} + \omega_{13} + \varepsilon} \right) \quad (10)$$

式中 P_2^{in} , P_2^{d} , P_2^{out} 表示 1/4 分辨率尺寸图的输入特征, 中间特征和输出特征, P_3^{in} , P_3^{d} , P_3^{out} 表示 1/8 分辨率尺寸图的输入特征, 中间特征和输出特征, P_4^{in} , P_4^{d} , P_4^{out} 表示 1/16 分辨率尺寸图的输入特征, 中间特征和输出特征, P_5^{in} , P_5^{out} 表示 1/32 分辨率尺寸图的输入特征和输出特征。Conv 表示用于特征处理的卷积操作, ω_1 、 ω_3 、 ω_6 、 ω_8 、 ω_{11} 、 ω_{12} 表示横向连接阶段学习到的用于区分特征重要程度的权重参数, ω_2 、 ω_4 、 ω_9 表示自顶向下连接阶段学习到的用于区分特征重要程度的权重参数, ω_7 、 ω_{12} 、 ω_{13} 表示自底向上连接阶段学习到的用于区分特征重要程度的权重参数, ω_5 、 ω_{10} 表示跨跃式连接阶段学习到的用于区分特征重要程度的权重参数, Resize 表示调整输入数据尺寸以适应模型输入要求的操作, $\varepsilon=0.0001$ 是保证数值稳定的小值。与 PANet 网络相比, BiFPN 作为特征融合网络实现了更高的精度和更全面的特征图信息融合。

2.3 改进损失函数

针对牧区牛粪检测任务, 边界框回归 (bounding box regression) 是决定牛粪定位精度的关键步骤。IoU (intersection over union) 是目前主流边界框回归函数的主要组成部分, 它表示为预测边界框和真实边界框相交区域面积与合并区域面积的比值。YOLOv5 模型采用 CIoU Loss^[27] 作为边界框回归损失函数, CIoU Loss 在 IoU Loss 的基础上增加了距离损失项及形状损失项, 计算过程如式 (11)、式 (12) 所示,

$$\text{IoU} = \frac{|B \cap B^{\text{gt}}|}{|B \cup B^{\text{gt}}|} \quad (11)$$

$$L_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{\text{gt}})}{c^2} + \alpha v \quad (12)$$

式中 B 和 B^{gt} 分别表示预测框和真实框, b 和 b^{gt} 分别表示预测框和真实框的中心点位置, ρ 表示预测边界框中心点和真实边界框中心点距离, c 表示其最小方框的对角线距离, α 表示正权衡参数, v 表示修正因子, L_{CIoU} 表示 CIoU Loss 损失函数。

为提升 IoU Loss 在不同检测任务中的普遍性和增加收敛速度, ZHANG 等^[28] 分析了边界框回归过程与模式, 基于边界框回归问题固有特征, 发现了在模型训练过程中使用较小的辅助边界框计算损失对高 IoU 样本的回归有增益效果, 低 IoU 样本则与之相反, 提出了 Inner-IoU Loss, 针对不同的数据集, 引入尺度因子 ratio 控制辅助边界框的尺度大小用于计算损失。Inner-IoU Loss 的定义如式 (13) ~ 式 (20) 所示,

$$b_l^{\text{gt}} = x_c^{\text{gt}} - \frac{w^{\text{gt}} \times r}{2}, b_r^{\text{gt}} = x_c^{\text{gt}} + \frac{w^{\text{gt}} \times r}{2} \quad (13)$$

$$b_t^{\text{gt}} = y_c^{\text{gt}} - \frac{h^{\text{gt}} \times r}{2}, b_b^{\text{gt}} = y_c^{\text{gt}} + \frac{h^{\text{gt}} \times r}{2} \quad (14)$$

$$b_l = x_c - \frac{w \times r}{2}, b_r = x_c + \frac{w \times r}{2} \quad (15)$$

$$b_t = y_c - \frac{h \times r}{2}, b_b = y_c + \frac{h \times r}{2} \quad (16)$$

$$i = (\min(b_l^{\text{gt}}, b_r) - \max(b_l^{\text{gt}}, b_l)) \times (\min(b_b^{\text{gt}}, b_b) - \max(b_t^{\text{gt}}, b_t)) \quad (17)$$

$$u = (w^{\text{gt}} \cdot h^{\text{gt}}) \cdot (r)^2 + (w \cdot h) \cdot (r)^2 - i \quad (18)$$

$$\text{IoU}^i = \frac{i}{u} \quad (19)$$

$$L_{\text{inner-CIoU}} = L_{\text{CIoU}} + \text{IoU} - \text{IoU}^i \quad (20)$$

式中 $(x_c^{\text{gt}}, y_c^{\text{gt}})$ 表示真实边界框内部中心点, (x_c, y_c) 表示锚框内部中心点, w^{gt} 和 h^{gt} 表示真实边界框的宽度和高度, w 和 h 表示锚框的宽度和高度。变量 r 表示比例因子, 通常取值范围为 [0.5, 1.5], 针对牧区牛粪检测任务, r 的取值为 1.10。 b_l^{gt} 、 b_r^{gt} 、 b_t^{gt} 、 b_b^{gt} 分别表示为引入尺度因子后真实框的顶点坐标, b_l 、 b_r 、 b_t 、 b_b 分别表示为引入尺度因子后预测框的顶点坐标。 i 定义为预测框与真实框交集的面积, u 定义为预测框与真实框并集的面积。 $L_{\text{inner-CIoU}}$ 表示改进后的损失函数。改进损失函数 CIoU 为 Inner-IoU, 提升了检测效果, 增加了定位精度, 减少了误检和漏检。

2.4 替换 C3 模块中的 Bottleneck

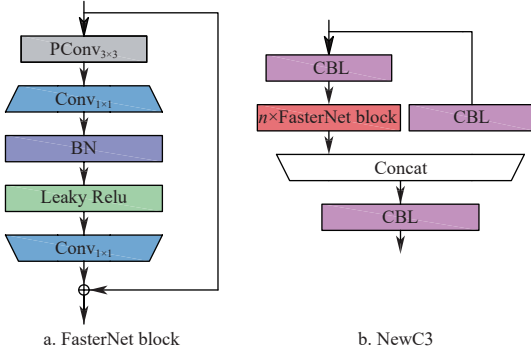
YOLOv5 的 C3 模块采用 CSP 架构, 包含了 3 个标准卷积层以及 3 个 Bottleneck 模块用于增加网络的深度和感受野, 提高目标检测的性能。其中, 每个 Bottleneck 由一系列卷积层和残差连接组成, 用于提取特征并减少参数量。C3^[29] 模块拥有高效的特征提取以及多尺度感受野的优点, 但由于 C3 模块采用了多个卷积层的堆叠, 增加了网络的计算复杂度, 导致了模型的推理速度较慢, 难以适用低算力的边缘化设备。

FasterNet^[30] 模型采用轻量化设计, 大体分为 4 个阶段, 每个阶段的主体都由 FasterNet block 堆叠组成, 实现精度保持的同时压缩了网络结构和运算量。FasterNet block 模块采用倒残差结构设计, 先进行 3×3 的 PConv 层升维, 然后 1×1 的 Conv 层, 最后再通过 1×1 的 Conv 层降维, 利用 Shortcut 将模块初始层与模块终止层张量相加, 重复了输入特征, 加快了梯度传播, 避免了梯度发散的问题。为实现高性能的同时保持特征多样性并实现更低的延迟, FasterNet block 模块仅在中部 Conv 层后进行归一化和激活处理, 在激活层中, 为缓解激活值过小而导致神经元参数无法更新的问题, 本文使用 Leaky Relu^[31] 替代 Relu 作为 FasterNet block 模块的激活函数。其中, PConv 是只对部分输入通道应用常规 Conv 进行空间特征提取的部分卷积, 为了充分有效地利用所有通道的信息, 一般在 PConv 后附加逐点卷积, 组成具有类似 T 形感受野的卷积层, 该卷积具有更低的浮点运算数和较低的内存访问量。

使用改进 FasterNet block 替换 C3 模块中的 Bottleneck, 可以有效降低模型参数量和计算量, 减少冗余计算增加运行速度。改进后的 FasterNet block 模块结构及改进后的 C3 模块结构如图 4 所示, 改进后的 YOLOv5 网络模型如图 5 所示。

2.5 添加牛粪捡拾判断机制

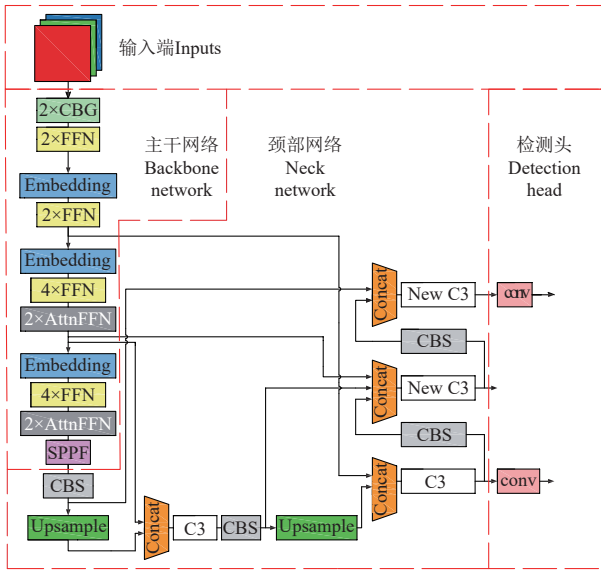
针对牧区牛粪捡拾任务，牛粪块的大小及牛粪群的疏密程度是判断牛粪是否值得捡拾的两个重要标准。小块且分布稀疏的牛粪一方面易被自然界降解，另一方面利用便利性也不如大块牛粪，故在牧区牛粪捡拾任务中大块集中牛粪比起小块稀疏牛粪更具捡拾优先性。可以利用预测阶段 YOLOv5 检测框划分阈值对牛粪块的大小及牛粪群的疏密程度做出判断。



注：PConv 表示部分卷积；Leaky Relu 表示一种修正线性单元激活函数；CBL 表示由卷积层、批归一化层和激活函数层组成的模块。
Note: PConv stands for partial convolution; Leaky Relu represents a modified linear element activation function; CBL represents a module consisting of a convolutional layer, a batch normalization layer, and an activation function layer.

图 4 FasterNet block 模块结构及新 C3 模块结构

Fig.4 IFasterNet block module structure and new C3 module structure



注：CBG 表示由卷积层、批归一化层和激活函数层组成的模块；SPPF 表示快速空间金字塔池化层；CBS 表示由卷积层、批归一化层和激活函数层组成的模块；New C3 表示改进后的具有 3 个卷积层的 CSP 瓶颈结构。
Note: CBG represents a module consisting of a convolutional layer, a batch normalization layer, and an activation function layer. SPPF stands for Spatial Pyramid Pooling - Fast; CBS represents a module consisting of a convolutional layer, a batch normalization layer, and an activation function layer; New C3 represents an improved CSP bottleneck structure with three convolutional layers.

图 5 改进 YOLOv5 网络模型

Fig.5 Improved YOLOv5 network model

YOLOv5 使用维度集群作为锚定框来预测边界框^[32]，每个边界框预测 4 个坐标信息 X_1 、 X_2 、 Y_1 、 Y_2 ，其中，

X_1 、 X_2 表示矩形边界框的横坐标信息， Y_1 、 Y_2 表示矩形边界框的纵坐标信息，矩形边界框的 4 个顶点坐标分别为 (X_1, Y_1) 、 (X_1, Y_2) 、 (X_2, Y_1) 和 (X_2, Y_2) 。边界框宽和高的定义如式 (21) ~ 式 (23) 所示，

$$w = X_2 - X_1, h = Y_2 - Y_1 \quad (21)$$

$$\sum_{i=1}^n w_i = w_1 + w_2 + \dots + w_n \quad (22)$$

$$\sum_{i=1}^n h_i = h_1 + h_2 + \dots + h_n \quad (23)$$

式中 w 和 h 表示边界框的宽度和高度， n 表示一张检测图像上所有边界框的数量， $\sum_{i=1}^n w_i$ 和 $\sum_{i=1}^n h_i$ 表示一张检测图像上所有边界框的宽度高度之和。具体到牧区牛粪捡拾任务中，单个边界框的 w 和 h 可映射单个牛粪块的大小， n 可映射牛粪群的疏密程度， $\sum_{i=1}^n w_i$ 和 $\sum_{i=1}^n h_i$ 可映射牛粪群的大小。牛粪块的大小及牛粪群的疏密程度判断定义如式 (24) ~ 式 (26) 所示，

$$w \geq w_0 \cup h \geq h_0 \quad (24)$$

$$n \geq n_0 \quad (25)$$

$$\sum_{i=1}^n w_i \geq w_0 \cup \sum_{i=1}^n h_i \geq h_0 \quad (26)$$

式中 w_0 和 h_0 表示判断牛粪块尺寸大小的阈值， n_0 表示判断牛粪群疏密程度的阈值。 w_0 、 h_0 和 n_0 的取值为根据牧区环境和牛群数量质量综合考虑的常数，该常数可灵活取定以适应不同环境不同要求的捡拾任务。当边界框宽度大于等于 w_0 或边界框高度大于等于 h_0 时，判断牛粪块尺寸为大，否则为小，当一张检测图像上所有边界框的数量大于等于 n_0 时，判断牛粪群的疏密程度为密，否则为疏。添加牛粪捡拾判断机制在判定牛粪存在的前提下利用预测阶段 YOLOv5 检测框据尺寸大小和疏密程度对某一地区牛粪做出分类，为智能化条件化捡拾大块集中牛粪提供依据。

3 结果与分析

3.1 多模块消融结果

为进一步验证改进方法的合理性和有效性，本文以牧区牛粪图像为数据集进行一系列消融试验，试验性能对比如表 1 所示。首先，替换主干网络 New CSP-Darknet53 为 EfficientFormerV2，与基准模型相比，召回率小幅降低的同时，参数量与浮点运算量分别下降了 41% 和 48%，计算复杂度较原模型大大降低，优势明显。其次，改进特征融合网络 PANET 为 BiFPN，召回率提升了 0.4 个百分点，平均精度均值提升了 0.3 个百分点，检测精度有所提升。然后，改进损失函数 CIoU 为 Inner-IoU，当 r 值大于 1 时，通过生成更大的辅助边界框达到对样本加速收敛的效果，经 r 不同取值对比试验， r 的取值大于 1 时，检测精度均有提升，但当 r 的取值为 1.10 时，效果最佳，准确率提升了 0.9 个百分点，平均精度均值提升了 0.7 个百分点，实现了在不增长运算量

的情况下精度的提升。最后，使用改进 FasterNet block 替换 C3 模块中的 Bottleneck，与原模块中 Relu 作为激活函数相比，使用 Leaky Relu 替代 Relu 作为 FasterNet block 模块的激活函数，准确率、召回率和平均精度均值都有较大幅度的提升，与基准模型相比，使用了 New C3 的模型准确率提升了 0.3 个百分点，召回率提升了

0.8 个百分点，平均精度均值提升了 0.2 个百分点，且参数量与浮点运算量都下降明显。相较于单独替换主干网络 EfficientFormerV2，改进特征融合网络 BiFPN 有助于提升检测精度，降低误检几率。同时使用 BiFPN、Inner-IoU 和 New C3 三种改进策略不仅减低参数量与运算量，也提升了检测精度。

表 1 消融试验性能对比
Table 1 Comparison of performance of ablation experiments

EfficientFormerV2	BiFPN	Inner-IoU	New C3	准确率 Precision/%	召回率 Recall/%	平均精度均值 mean Average Precision@0.5 (mAP @0.5)/%	参数量 Params/M	浮点运算量 Floating point operations FLOPs/G
×	×	×	×	92.3	89.2	87.8	7.02	15.9
√	×	×	×	90.1	87.9	86.8	4.12	8.2
×	√	×	×	92.2	89.6	88.1	7.08	16.0
×	×	1.05	×	92.4	89.2	87.7	7.01	15.9
×	×	1.10	×	93.2	89.4	88.5	7.01	15.9
×	×	1.15	×	92.6	89.6	88.3	7.01	15.9
×	×	×	Relu	91.9	88.8	87.1	6.46	14.6
×	×	×	Leaky Relu	92.6	90.0	88.0	6.47	14.8
√	√	×	×	91.8	87.4	87.0	4.13	8.2
×	√	1.10	Leaky Relu	92.9	89.5	88.2	6.64	15.6
√	√	1.10	Leaky Relu	92.6	87.7	87.4	4.02	8.1

注：×表示移除改进模块；√表示增添改进模块。
Note: × represents the removal of improved modules; √ represents the addition of improved modules.

综合以上改进方法，改进模型参数量下降了 43%，浮点运算量下降了 49%，准确率略有提升，召回率与平均精度均值略有下降。为对比模型改进前后对牧区牛粪特征提取能力强弱的变化，利用 Grad-CAM^[33] 通过热力图的方式实现网络训练过程的可视化。图 6 为模型改进前后的 Grad-CAM 可视化结果，在 Grad-CAM 图中以特殊高亮的形式显示更加关注的区域。相对于原始模型只关注牛粪内部纹理特征，改进模型不仅关注牛粪内部纹理特征，也关注牛粪边界处特征，具体体现在热度区域覆盖了更多的牛粪部分，在背景区域收到的干扰也较小，改进后模型的特征提取能力明显优于初始模型。改进模型在保证检测精度和有效性的前提下明显减少了参数量和运算量，显著提升了模型的性能，适合在资源受限的移动设备与牧区环境中应用。

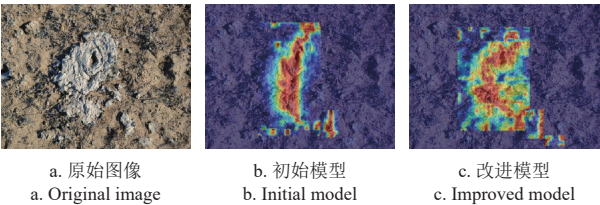


图 6 模型改进前后 Grad-CAM 可视化结果
Fig.6 Grad-CAM visualization before and after model improvement

3.2 多网络对比结果

为进一步评价改进 YOLOv5 网络对牧区牛粪的检测效果，本文以相同数据集分别采用 YOLOv9-s、YOLOv5s、YOLOv7-tiny、EfficientFormerV2 和 Fasternet网络以相同的训练条件进行一系列对比试验，试验性能对比如表 2 所示。其中，YOLOv9-s 的平均精度均值最高，同时其参数量及浮点运算量也是最高，模型复杂程度最高。EfficientFormerV2 的检测速度最快，同时其平均精度均值也较低。Fasternet 的参数量、计算量和平均精度均值最低。YOLOv5-EBIN 与其他主流检测模型相比，拥有最高的准确率及较高的召回率和平均精度均值。以 YOLOv5s 为例，YOLOv5-EBIN 略低于 YOLOv5s 的召回率，但 YOLOv5s 的参数量是改进模型的 1.75 倍，浮点运算量是改进模型的 1.93 倍，检测速度却只有 YOLOv5-EBIN 的 2/3。改进后网络模型的准确率为 92.6%，召回率为 87.7%，平均精度均值为 87.4%，参数量为 4.02M，浮点运算量为 8.1G。试验证明，与其他模型相比，改进模型 YOLOv5-EBIN 拥有不逊色于 YOLOv5s 的平均精度均值和逼近 Fasternet 的参数量及计算量和逼近 EfficientFormerV2 的检测速度，在保证模型轻量化的同时，具有相对较高的准确率、平均精度均值和检测速度，能够很好地完成牧区牛粪的检测工作。

表 2 不同网络试验性能对比
Table 2 Comparison of experimental performance of different networks

模型 Models	准确率 Precision/%	召回率 Recall/%	mAP @0.5/%	参数量 Params/M	浮点运算量 FLOPs/G	检测速度 Frames per second/(帧·s ⁻¹))
YOLOv9-s	92.4	88.2	88.6	9.60	38.7	31.1
YOLOv5s	92.3	89.2	87.8	7.02	15.9	23.9
YOLOv7-tiny	92.1	81.8	85.7	6.01	13.2	30.8
EfficientFormerV2	88.7	82.3	82.5	4.21	7.5	38.6
Fasternet	87.4	80.9	81.7	3.49	6.6	28.8
YOLOv5-EBIN	92.6	87.7	87.4	4.02	8.1	34.7

图 7 展示了不同网络模型牛粪图像检测结果，除 EfficientFormerV2 发生漏检情况外，其余模型均能很好的检测并定位 3 处牛粪位置，改进后的 YOLOv5 模型在

推理阶段可依据牛粪尺寸大小的不同对牛粪做出分类并使用不同颜色的边界框做出区分。与不同网络检测结果对比，该模型均表现出了较好的目标定位和识别能力。

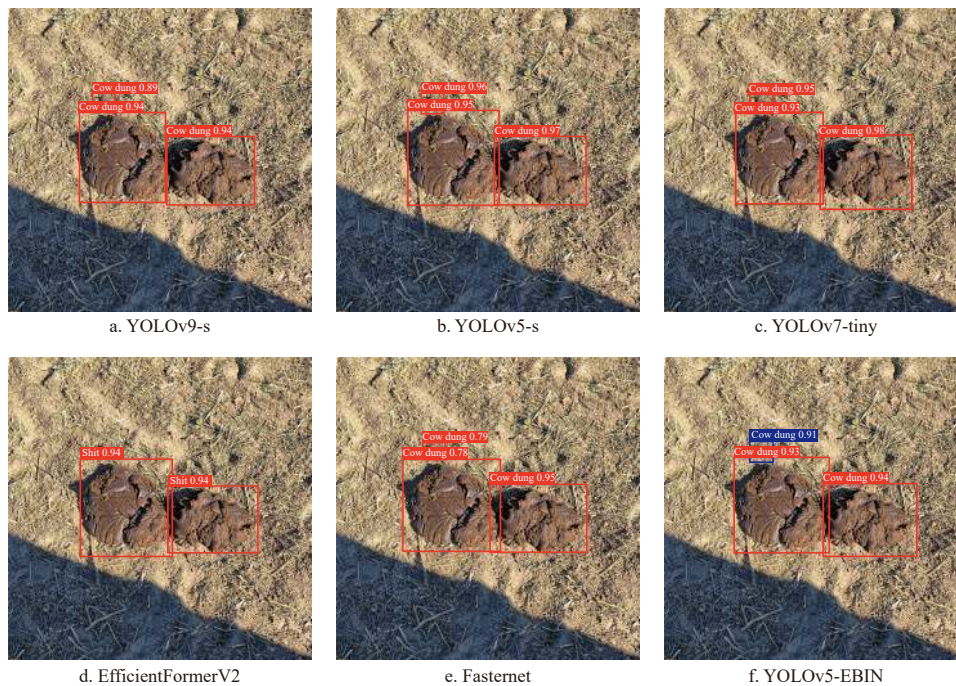


图 7 不同网络模型牛粪图像检测结果

Fig.7 Detection results of cow fecal images from different network models

4 结 论

1) 本研究在 YOLOv5 目标检测模型的基础上结合 EfficientFormerV2, BiFPN, Inner-IoU 及改进 FasterNet block, 提出了一种改进 YOLOv5 牛粪图像检测模型。改进后网络模型的准确率为 92.6%，召回率为 87.7%，平均精度均值为 87.4%，参数量为 4.02M，浮点运算量为 8.1G，改进模型在保证检测精度的同时明显减少了参数量和运算量，显著提升了模型的性能，可以完成牧区牛粪检测任务。

2) 本研究在目标预测阶段添加了捡拾判断机制，利用 YOLOv5 检测框对牛粪块的尺寸大小及牛粪群的疏密程度做出判断，并据此对某一地区牛粪做出分类，为牛粪的智能化条件化捡拾提供依据。

3) 本研究建立一套牧区牛粪数据集，为提高数据多样性对该数据集进行了数据增广。在自建数据集上对改进模型进行了多网络对比试验，试验结果表明，改进模型优于相同量级参数量的其他模型。

【参 考 文 献】

- [1] 丁勇, 春亮, 孙娟娟, 等. 中国, 40% 国土是草原[N]. 中国绿色时报, 2021-03-11(8).
- [2] 邵浩宸, 衣淑娟, 沈玉君, 等. 不同工艺对牛粪堆沤肥过程气体排放的影响[J]. 农业工程学报, 2024, 40(2): 323-331.

- Engineering (Transactions of the CSAE), 2024, 40(2): 323-331. (in Chinese with English abstract)
- [3] 周海宾, 刘娟, 丁京涛, 等. 采用 EEM-FRI 方法研究黑曲霉对牛粪堆肥腐熟及纤维素降解影响[J]. 农业工程学报, 2022, 38(1): 276-286.
- ZHOU Haibin, LIU Juan, DING Jingtao, et al. Effects of *Aspergillus niger* on maturity and cellulose degradation of cow manure composting using EEM-FRI method[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(1): 276-286. (in Chinese with English abstract)
- [4] 张胜茂, 李佳康, 唐峰华, 等. 基于深度学习的鱼类养殖监测研究进展[J]. 农业工程学报, 2024, 40(5): 1-13.
- ZHANG Shengmao, LI Jiakang, TANG Fenghua, et al. Research progress on fish farming monitoring based on deep learning technology[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(5): 1-13. (in Chinese with English abstract)
- [5] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE, 2014: 580-587.
- [6] GIRSHICK R. Fast R-CNN[C]// Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 1440-1448.
- [7] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-

- time object detection with region proposal net works[C]// Proceedings of International Conference on Neural Information Processing Systems. Montreal, Canada: NIPS, 2015: 91-99.
- [8] 赵永强, 饶元, 董世鹏, 等. 深度学习目标检测方法综述[J]. *中国图象图形学报*, 2020, 25(4): 629-654.
- ZHAO Yongqiang, RAO Yuan, DONG Shipeng, et al. Survey on deep learning object detection[J]. *Journal of Image and Graphics*, 2020, 25(4): 629-654. (in Chinese with English abstract)
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015: 779-788.
- [10] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]// Proceedings of European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016: 21-37.
- [11] 陈科圻, 朱志亮, 邓小明, 等. 多尺度目标检测的深度学习研究综述[J]. *软件学报*, 2021, 32(4): 1201-1227.
- CHEN Keqi, ZHU Zhiliang, DENG Xiaoming, et al. Deep learning for multi-scale object detection: A survey[J]. *Journal of Software*, 2021, 32(4): 1201-1227. (in Chinese with English abstract)
- [12] 黄健, 张钢. 深度卷积神经网络的目标检测算法综述[J]. *计算机工程与应用*, 2020, 56(17): 12-23.
- HUANG Jian, ZHANG Gang. Survey of object detection algorithms for deep convolutional neural networks[J]. *Computer Engineering and Applications*, 2020, 56(17): 12-23. (in Chinese with English abstract)
- [13] 吕璐, 程虎, 朱鸿泰, 等. 基于深度学习的目标检测研究与应用综述[J]. *电子与封装*, 2022, 22(1): 010307.
- LYU Lu, CHENG Hu, ZHU Hongtai, et al. Progress of research and application of object detection based on deep learning[J]. *Electronics & Packaging*, 2022, 22(1): 010307. (in Chinese with English abstract)
- [14] SANGDOO Yun, DONGYOON Han, SEONG Joon, et al. CutMix: Regularization strategy to train strong classifiers with localizable features[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE, 2019: 6023-6032.
- [15] 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 Conference on Computer Vision and Pattern Recognition. Virtual: IEEE, 2020: 390-391.
- [16] LIU S, QI L, QIN HH, et al. Path aggregation network for instance segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 8759-8768.
- [17] JOSEPH Redmon, ALI Farhadi. YOLOv3: An incremental improvement[EB/OL]. (2018-5-8) [2024-10-30]. <https://arxiv.org/pdf/1804.02767>.
- [18] BRYAN L W, SCOTT M L, ROGER A E, et al. Electromagnetic induction sensor data to identify areas of manure accumulation on a feedlot surface[J]. *Soil Science Society of America Journal*, 2009, 12: 2068-2077.
- [19] KELLY S, DANICA S, REBECCA L. Muenich. Using remote sensing to identify liquid manure applications in eastern North Carolina[J]. *Journal of Environmental Management*, 2022, 317(9): 115334.
- [20] 唐瑜嵘, 沈明霞, 薛鸿翔, 等. 人工智能技术在畜禽养殖业的发展现状与展望[J]. *智能化农业装备学报*, 2023, 4(1): 1-16.
- TANG Yurong, SHEN Mingxia, XUE Hongxiang, et al. Development status and prospect of artificial intelligence Technology in livestock and poultry breeding[J]. *Journal of Intelligent Agricultural Mechanization*, 2023, 4(1): 1-16. (in Chinese with English abstract)
- [21] 胡振楠, 孙红敏, 李晓明, 等. 猪舍自动清粪控制系统设计与实现[J]. *南方农机*, 2021, 52(1): 8-11.
- [22] LI YY, HU J, WEN Y, et al. Rethinking vision transformers for mobilenet size and speed[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE, 2023: 16843-16854.
- [23] LI YY, YUAN G, WEN Y, et al. EfficientFormer: Vision transformers at mobilenet speed[EB/OL]. (2022-6-2) [2024-10-30]. <https://arxiv.org/pdf/2206.01191>.
- [24] GONG C Y, WANG D L, LI M, et al. NASViT: Neural architecture search for efficient vision transformers with gradient conflict aware supernet training[C]// International Conference on Learning Representations. Virtual: OpenReview. net. 2022.
- [25] FRANÇOIS Chollet. Xception: Deep learning with depthwise separable convolutions[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016: 1800-1807.
- [26] TAN MX, PANG RM, QUOC V. Le. EfficientDet: scalable and efficient object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE, 2019: 10778-10787.
- [27] ZHENG Z, WANG P. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. *IEEE Transactions on Cybernetics*, 2020, 52: 8574-8586.
- [28] ZHANG H, XU C, ZHANG S. Inner-IoU: More effective intersection over union loss with auxiliary bounding box[EB/OL]. (2023-11-14) [2024-10-30]. <https://arxiv.org/pdf/2311.02877>.
- [29] HYJOIN P, YOUNGJOON Y, GEONSEOK S, et al. C3: Concentrated-comprehensive convolution and its application to semantic segmentation[EB/OL]. (2019-7-28) [2024-10-30]. <https://arxiv.org/pdf/1812.04920>.
- [30] CHEN J R, KAO S H, He H, et al. Run, Do not walk: Chasing higher FLOPS for faster neural networks[C]// Proceedings of

- the IEEE Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE, 2023: 12021-12031.
- [31] ANDREW L. Maas, Awni Y. Hannun, ANDREW Y. Ng. Rectifier nonlinearities improve neural network acoustic models[C]// Proceedings of the 30 th International Conference on Machine Learning. Atlanta, USA, 2013: 16489696.
- [32] JOSEPH R, ALI F. Yolo9000: Better, faster, stronger[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Hawaii, USA: IEEE, 2017: 6517-6525.
- [33] 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. Venice, Italy: IEEE, 2017: 618-626.

Detecting cattle manure in pastoral areas using the YOLO network

NA Risu , ZHAI Lihao , XING Haipeng , ZHANG Quan

(School of Mechanical Engineering, Inner Mongolia University of Technology, Hohhot 010010, China)

Abstract: Grassland animal husbandry is the main form of animal husbandry. Cow manure is the by-product of grassland animal husbandry on the vast grassland areas. Since cow manure can serve as an important source of energy, there is a negative impact on grass growth. The distribution of cow manure is also characterized by scattered and concentrated areas in natural grasslands. However, manual picking is still used to collect cow manure in pastoral areas at present, leading to the high labor intensity and cost. Alternatively, computer vision can be expected to apply to the collection of cow manure in recent years. In this article, an image detection model was proposed for cow manure using improved YOLOv5. Firstly, New CSP-Darknet53 was replaced with EfficientFormerV2, in order to improve the boundary sensitivity for low complexity of the model. Experimental results showed that the computational complexity was significantly reduced, compared with the original model. Secondly, PANET was replaced with BiFPN to enhance the feature fusion capability for the high accuracy of detection. There was an increase in the accuracy of detection. Furthermore, CIoU Loss was replaced with Inner IoU Loss, in order to improve the localization accuracy of bounding box regression. Once the ratio value was greater than 1, the larger auxiliary bounding boxes were generated to accelerate the convergence of the sample. A comparison with different ratio values showed that the best performance was achieved when the ratio value was 1.10. From then on, the Bottleneck was replaced in the C3 module with an improved FasterNet block. Leaky Relu was used instead of Relu as the activation function for the FasterNet block module, resulting in significant improvements in the accuracy, recall, and average precision. Additionally, the number of parameters and floating-point operations decreased significantly. The computational complexity was reduced to accelerate the inference speed. Finally, a picking judgment mechanism was added using the YOLOv5 detection box, in order to evaluate the size of the cow manure block and the density of the cow manure group. The cow manure was classified in a certain area for intelligent conditional picking of cow manure. The accuracy of the improved network model was 92.6%, the recall rate was 87.7%, the average accuracy was 87.4%, the parameter count was 4022847B, and the floating-point operation count was 8.1G. The improved model significantly reduced the parameter and operation for the high detection accuracy, in order to significantly improve the performance of the model. A dataset of cow manure was established in pastoral areas, in order to improve data diversity. The collected data was randomly combined using five operations: flipping, scaling and translation, motion blur, random occlusion, and brightness change, in order to enhance data augmentation. Comparison experiments were conducted on the multiple networks and the improved model on the dataset. The experimental results showed that the accuracy of the improved model was superior to other models with the same level of parameter and computational complexity. The improved YOLOv5 model classified the cow manure, according to its different sizes during the inference stage, and then distinguished it using different colored bounding boxes. The recognition and localization of cow manure were achieved in the pastoral areas. The finding can also provide technical support to the intelligent picking vehicles of cow manure.

Keywords: image processing; manures; convolutional neural networks; collecting cow manure