

轻量级多场景群养猪只行为识别模型研究

漆海霞 冯发生 尹选春 杨泽康 周子森 梁广升

(华南农业大学工程学院, 广州 510642)

摘要: 针对现有猪只行为识别模型体积大、识别场景单一、部署应用硬件要求高等问题,本文提出轻量级多场景群养猪只行为识别模型 YOLO v5n - PBR (YOLO v5n for pig behavior recognition)。首先通过拍摄和收集不同养殖场景、不同猪只数量及不同角度的群养猪只行为数据构建多场景群养猪只行为数据集,并根据该数据集中猪只行为目标的特点引入迁移学习方法和 OTA (Optimal transport assignment) 标签分配方法对 YOLO v5n 模型进行训练,加快模型收敛速度并提升模型精度,构建高精度多场景群养猪只行为识别模型;然后利用 L1 - norm 剪枝算法筛选并删减模型中不重要的通道,去除冗余参数;最后通过微调训练和中间特征知识蒸馏去除剪枝带来的性能劣化,从而得到轻量级多场景群养猪只行为识别模型 YOLO v5n - PBR 并进行嵌入式设备部署。实验结果表明, YOLO v5n - PBR 模型平均精度均值 (mean average precision, mAP) 为 96.9%, 参数量、计算量和内存占用量分别为 4.700×10^5 、 1.20×10^9 和 1.2 MB, 在两种不同系统和不同硬件配置的嵌入式设备上的部署实时识别帧率分别为 12.2 帧/s 和 66.3 帧/s, 与原始模型 YOLO v5n 相比, mAP 提高 1.1 个百分点, 参数量、计算量和内存占用量分别减少 73.3%、70.7% 和 68.4%, 部署实时识别帧率分别提高 74.3% 和 83.1%。此外, 基于多场景群养猪只行为数据集训练得到的 YOLO v5n - PBR 模型在 4 个单场景或双场景的群养猪只行为数据集上的 mAP 均能达到 98.1%, 对 2 种不同养殖场景的 6 段猪只行为视频的嵌入式设备部署识别统计结果与人工统计结果相近, 平均精确率和平均召回率均为 95.3%, 以较少的参数达到较强的泛化性。本文提出的 YOLO v5n - PBR 模型具有精度高、体积小、速度快、泛化性强等优点, 可满足嵌入式设备部署要求, 为猪只行为的实时、准确监测及猪只行为识别模型的部署应用提供技术基础。

关键词: 猪只行为识别; 模型轻量化; 通道剪枝; 知识蒸馏; YOLO v5n; 嵌入式设备

中图分类号: TP391.4; S828

文献标识码: A

文章编号: 1000-1298(2024)10-0306-12

OSID:



Research of Lightweight Multi-scene Group Pig Behavior Recognition Model

QI Haixia FENG Fasheng YIN Xuanchun YANG Zekang ZHOU Zisen LIANG Guangsheng
(College of Engineering, South China Agricultural University, Guangzhou 510642, China)

Abstract: In order to solve the problems of large size, single recognition scene and high hardware requirements for deploying application of existing pig behavior recognition models, a lightweight multi-scene group pig behavior recognition model YOLO v5n for pig behavior recognition (YOLO v5n - PBR) was proposed. Firstly, a multi-scene group pig behavior dataset was constructed by shooting and collecting group pig behavior data from different breeding scenes, different pig numbers and different angles, and based on the characteristics of pig behavior objectives in the dataset, the transfer learning method and the optimal transport assignment label assignment method were introduced to train the YOLO v5n model, which accelerated the model convergence speed and improved the model accuracy, and a high-precision multi-scene group pig behavior recognition model was constructed. Then the L1 - norm pruning algorithm was used to screen and delete the unimportant channels in the model to remove the redundant parameters. Finally, the performance degradation caused by pruning was removed by fine-tuning training and intermediate feature knowledge distillation, so that the lightweight multi-scene group pig behavior recognition model YOLO v5n - PBR was obtained and deployed as embedded devices.

收稿日期: 2023-12-10 修回日期: 2024-01-22

基金项目: 广州市农村科技特派员项目(20212100026)和特定高校学科建设项目(2023B10564002)

作者简介: 漆海霞(1969—),女,副教授,博士,主要从事农业信息监测与识别相关研究, E-mail: qihaxia_scau@126.com

Experimental results showed that the mean average precision (mAP) of the YOLO v5n – PBR model was 96.9% , with parameters, amount of computation, and memory footprint being 4.700×10^5 , 1.20×10^9 , and 1.2 MB, respectively. The deploy real-time recognition frame rates on embedded devices with different systems and hardware configurations were 12.2 frames/s and 66.3 frames/s. Compared with that of the original YOLO v5n model, the mAP was improved by 1.1 percentage points, and parameters, amount of computation, and memory footprint were decreased by 73.3% , 70.7% , and 68.4% , respectively. The deploy real-time recognition frame rates were increased by 74.3% and 83.1% . In addition, the YOLO v5n – PBR model trained based on the multi-scene group pig behavior dataset can reach 98.1% of mAP on four single-scene or dual-scene group pig behavior datasets, and the statistical results of embedded device deployment recognition of six pig behavior videos in two different breeding scenes were similar to those of manual statistics, with an average accuracy and average recall rate of 95.3% , which achieved strong generalization with fewer parameters. The YOLO v5n – PBR model proposed had the advantages of high accuracy, small size, fast speed, and strong generalization, which can meet the deployment requirements of embedded devices and provide a technical basis for real-time and accurate monitoring of pig behavior and the deploying application of pig behavior recognition model.

Key words: pig behavior recognition; model lightweight; channel pruning; knowledge distillation; YOLO v5n; embedded device

0 引言

猪只行为是猪只健康状态的重要表征和猪舍环境舒适度的重要体现。多数猪只疾病的临床或亚临床体征表现前常伴随着异常行为的出现,而当猪只受到养殖环境变化的刺激时也会通过调节行为来缓解外界环境带来的不适^[1]。因此,实时感知猪只行为并对其行为信息进行分析 and 挖掘,来评估猪只的健康状况和养殖环境舒适度等,实现疾病预警和环境预警^[2-3],对猪只的福利化养殖有着重要意义。

早期研究主要利用电子传感器^[4-5]和传统图像识别^[6-8]方法对猪只行为进行实时感知,但存在使用复杂、易导致猪只受伤和出现应激反应、识别速度慢等缺点,而将深度学习模型应用于猪只行为识别,可实现低应激、非接触式猪只行为识别^[9-13]。高云等^[14]基于 3D CONV 构建 3DconvNet 来识别群养猪只侵略性行为,识别准确率为 95.7% ,识别速度为 2 帧/s;薛月菊等^[15]通过改进 Faster R – CNN 对哺乳母猪站立、坐立、俯卧、腹卧和侧卧 5 种行为进行识别,平均识别准确率约为 93.3% ,识别速度约为 17 帧/s;ZHANG 等^[16]基于 SSD 和 MobileNet 构建 SBDA – DL 模型,识别猪只饮水、排尿和攀爬行为,平均识别准确率为 93.4% ,识别速度为 7 帧/s;这些模型虽能实现低应激、非接触式的猪只行为识别,但也存在以下问题:模型训练所用的数据集为单一养殖场景,泛化能力较差,较难应用到其他养殖场景;模型体积大,识别速度不高,难以满足视频流识别的实时性要求,同时在部署应用时对设备硬件要求高,不易进行部署应用;部分模型在训练平台完成训练及测试后,未开发相应平台进行模型部署应用,缺乏

实际应用价值。

针对以上问题,本文提出轻量级多场景群养猪只行为识别模型 YOLO v5n – PBR 并进行嵌入式设备部署应用,以期多场景群养猪只行为的实时准确监测及猪只行为识别模型的部署应用提供理论依据与技术基础。

1 多场景群养猪只行为数据集构建

1.1 数据来源

不同的养殖场景、猪只数量、猪只目标大小及拍摄角度都会对猪只行为识别模型的识别效果造成干扰。为提升模型的泛化性和鲁棒性,增强抗干扰能力,本文使用 3 个数据源构建多场景群养猪只行为数据集进行模型训练,分别是:河源市连平县江波养殖场、公开数据集 1^[17]及公开数据集 2^[18],其中在河源市连平县江波养殖场选取饲养 4 头育肥猪的猪栏和饲养 10 头育肥猪的猪栏作为主拍摄场景和次拍摄场景,获取的数据作为数据集 A 和数据集 B,拍摄周期为 2023 年 2 月 1—15 日;公开数据集 1 中获取的数据作为数据集 C;公开数据集 2 中获取的数据作为数据集 D。各数据集的详细信息如表 1 所示,各数据集部分养殖场景示例如图 1 所示,由于拍摄角度原因,数据集中会有部分猪只未出现在拍摄画面中。

1.2 数据集构建

人工筛选裁剪掉猪只长时间睡眠的部分后得到猪栏视频 411 段,其中数据集 A 视频 108 段,每段时长 5 ~ 32 min 不等;数据集 B 视频 45 段,每段时长 1 ~ 1.5 min 不等;数据集 C 视频 258 段,每段时长 1 min。以上视频帧率均为 30 帧/s,利用视频分帧技

表 1 各数据集详细信息
Tab.1 Details of each dataset

数据集	数据集来源	主要拍摄位置和角度	场景数量	猪只目标大小
A	河源市连平县江波养殖场	猪栏右方、斜对角及后方高约 2.3 m, 以向下倾斜约 45°的角度拍摄	1	大目标
B		猪栏前方和后方高约 1.8 m, 以向下倾斜 25°的角度拍摄	2	大目标、中目标
C	公开数据集 1	猪栏侧前方高 2.5 m, 以近乎俯视的角度拍摄	1	中目标
D	公开数据集 2	猪栏侧边高 3 m, 以俯视的角度拍摄	2	中目标、小目标

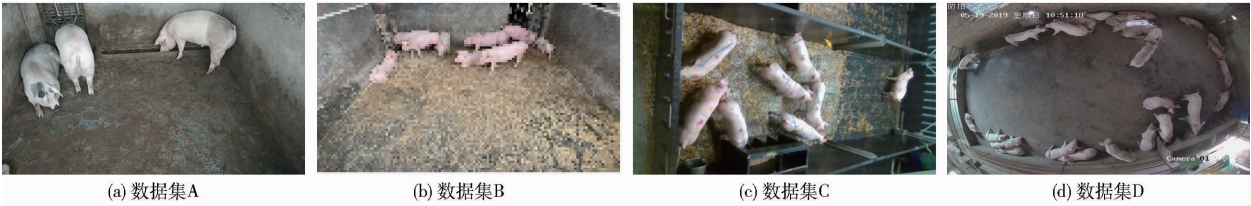


图 1 各数据集养殖场景示例
Fig.1 Examples of breeding scenarios in each dataset

术按每 30 帧取 1 帧对上述视频进行图像抽取,去除冗余图像及猪只目标特征模糊不清的糅杂图像后,筛选出 3 512 幅猪只行为图像,其中数据集 A 图像 1 621 幅,数据集 B 图像 270 幅,数据集 C 图像 1 621 幅,另外从公开数据集 2 中挑选 356 幅图像作为数据集 D,最终获得 3 868 幅猪只行为图像。为减少模型训练时间,将猪只行为图像的分辨率调整为 1 280 像素 × 720 像素。

数据集样本的多样性与模型的泛化性和鲁棒性

存在一定正比关系,会直接影响模型训练效果,因此本文采用垂直翻转、180°翻转、水平翻转、亮度调整及添加高斯噪声共 5 种方法对数据集 B 的全部图像和数据集 D 的 253 幅图像进行数据增强,以数据集 B 为例,数据增强示例如图 2 所示。上述数据增强方法可使训练后的猪只行为为识别模型适应不同的视角、光线及噪声影响,提升模型训练后的泛化性和鲁棒性,同时避免训练过程中因部分数据集样本过少而出现过拟合现象。



图 2 图像增强示例
Fig.2 Examples of image enhancement

数据集样本经数据增强后增加到 6 483 幅,采用 LabelImg 软件对数据集图像进行标注,标注框为

猪只行为目标的最小外接矩形,各行为定义如表 2 所示,以数据集 A 为例,猪只行为示例如图 3 所示。

表 2 猪只行为定义
Tab.2 Pig behavior definition

行为类别	行为定义	标签名称
进食	四腿与身体垂直,背朝上,腹部朝下,头部伸入食槽内	Feeding
俯卧	四腿收起,背朝上,腹部朝下并接触地面,身体不倾斜或微微倾斜	Lying prone
侧卧	四腿伸出在左侧或右侧,侧躺于地下,露出腹部,身体较大倾斜	Side prone
站立	四腿与身体垂直,背朝上,腹部朝下	Standing



图 3 猪只行为示例

Fig. 3 Examples of pig behavior

将数据集按照比例 7:2:1 随机划分为训练集 4 538 幅、验证集 1 296 幅和测试集 649 幅。至此,多场景群养猪只行为数据集构建完毕。

2 轻量级多场景群养猪只行为识别模型构建

本文技术路线如图 4 所示。首先根据多场景群养猪只行为数据集中猪只行为目标的特点,在 YOLO v5 模型的训练中引入迁移学习方法和 OTA 标签分配方法,使模型更快收敛及获得高精度,实现多场景下群养猪只行为的高精度识别,然后利用 L1 - norm 剪枝算法修剪模型中的冗余通道,对模型进行轻量化处理,最后利用微调训练及中间特征知识蒸馏一定程度上去除剪枝带来的性能劣化,从而获得精度高、体积小、速度快、泛化性强的轻量级多场景群养猪只行为识别模型,并通过开发猪只行为识别系统将模型进行嵌入式设备部署应用。

2.1 高精度模型构建

YOLO v5 模型具有 Mosaic 数据增强、自适应锚框计算、CSP 结构和 FPN + PAN 结构等改进,使其性能得到较大提升^[19]。在 YOLO v5 系列模型中,

YOLO v5n 模型在保证一定识别精度的同时有着最少的参数量、计算量和内存占用量,因此本文选择以 YOLO v5n 模型为基础模型来构建高精度多场景群养猪只行为识别模型,YOLO v5n 模型结构如图 4 所示。

2.1.1 迁移学习方法引入

迁移学习是一种将模型在源域学习到的知识迁移应用到目标域中的训练方法^[20]。模型在源域中学习到的特征表达,有助于其在训练中更快收敛及获得更好的初始性能。

在进行多场景群养猪只行为识别模型训练时,本文将 YOLO v5n 模型在 COCO 数据集上训练得到的权重作为预训练权重进行知识迁移,能有效提高 YOLO v5n 模型对多场景群养猪只行为的识别能力,对多场景群养猪只行为识别任务起积极作用。

2.1.2 OTA 标签分配方法引入

YOLO v5n 模型采用基于交并比 (Intersection over union, IoU) 的静态标签分配方法,该方法首先通过计算预测框与真实框 (Ground truth, GT) 之间的 IoU,并将其与设定的 IoU 阈值进行比较,随后

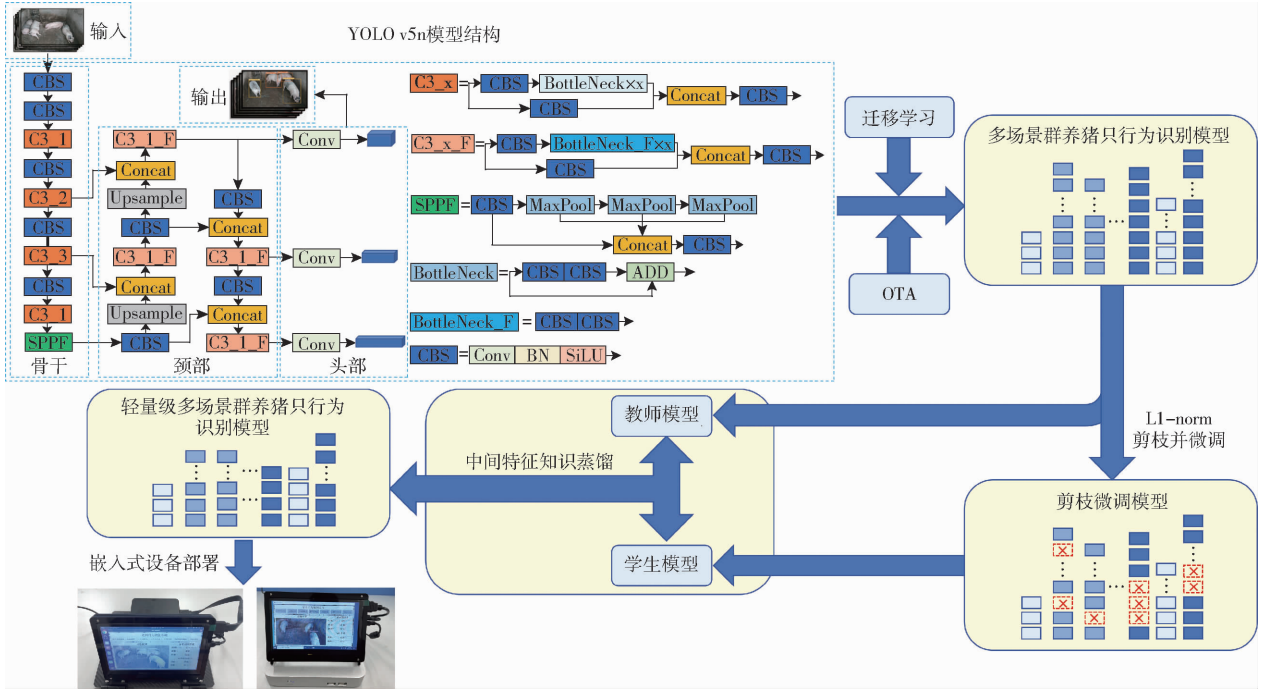


图 4 技术路线

Fig. 4 Technical route

根据比较结果将预测框与对应的 GT 相匹配,为预测框分配类别标签和回归目标。IoU 阈值的设置会非常影响模型的识别结果,将直接决定模型识别精度。

在本文的多场景群养猪只行为识别任务中,摄像机的拍摄角度、拍摄高度和猪只数量会使数据集中出现大量不同遮挡、不同大小和不同角度的猪只行为,而基于 IoU 的标签分配方法属于静态分配,在模型训练过程中始终根据设置的 IoU 阈值进行标签分配,对于不同遮挡、不同大小和不同角度的猪只行为的正负样本的判定条件均相同,将会影响模型性能。为解决该问题,本文在 YOLO v5n 模型的训练中引入 OTA 标签分配方法。OTA 是一种考虑全局最优的动态标签分配方法,将标签分配问题转化为最优传输问题,通过计算 GT 与所有预测框间的正样本运输成本和背景与所有预测框间的负样本运输成本,寻找合适的映射关系,使得整体运输成本最低,让模型自己学习最优的标签分配方式^[21],有效提高模型对多场景群养猪只行为识别精度。

在 OTA 标签分配方法中,真实框 GT_i 、背景与预测框 P_j 间的正、负标签运输成本计算公式为

$$c_{ij}^{fg} = L_{cls}(P_j^{cls}(\theta), G_i^{cls}) + \alpha L_{reg}(P_j^{box}(\theta), G_i^{box}) \quad (1)$$
$$c_j^{bg} = L_{cls}(P_j^{cls}(\theta), \emptyset) \quad (2)$$

- 式中
- c_{ij}^{fg}

——正标签运输成本
- L_{cls}

——分类损失
- P_j^{cls}

——预测框分类分数
- θ

——模型参数
- G_i^{cls}

——真实框类别
- α

——平衡系数
- L_{reg}

——回归损失
- P_j^{box}

——预测框边界框
- G_i^{box}

——真实框边界框
- c_j^{bg}

——负标签运输成本
- $\emptyset$

——背景类

将正、负标签运输成本拼接后得到成本矩阵,进

一步结合供给矩阵和需求矩阵便可通过 Sinkhorn - Knopp^[22]解决最优传输问题。

2.2 轻量级模型构建

2.2.1 基于 L1 - norm 剪枝的模型轻量化

本文通过引入迁移学习方法和 OTA 标签分配方法对 YOLO v5n 模型进行训练,构建了多场景群养猪只行为识别模型,但其结构中存在冗余。为减小模型内存占用量并加快识别速度,使其更易在巡检机器人、猪舍实时监测系统应用的嵌入式设备中部署,有必要使用剪枝算法对模型进行修剪。本文使用 L1 - norm 剪枝算法对模型进行剪枝,该剪枝算法通过采用 L1 正则化对通道批量归一化 (Batch normalization, BN) 层的缩放因子进行稀疏训练,以此来判断 BN 层缩放因子接近于 0 的通道,即对模型识别性能贡献较小的通道,然后剪掉这些通道,获得剪枝模型,最后对剪枝模型进行微调训练,一定程度上恢复剪枝带来的模型性能下降^[23],通道剪枝过程如图 5 所示。

2.2.1.1 稀疏化训练

模型在训练时需要不断学习新的数据分布情况,这导致模型收敛慢,对此,通过添加 BN 层对数据进行批归一化处理^[24],公式为

$$Z_{out} = \gamma \hat{Z} + \beta \quad (3)$$

其中

$$\hat{Z} = \frac{Z_{in} - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} \quad (4)$$

- 式中
- Z_{out}

——BN 层的输出
- γ

——BN 层的缩放因子
- β

——BN 层的偏置
- $\hat{Z}$

——归一化处理后的数据
- Z_{in}

——BN 层的输入
- μ_B

——每个批次的均值
- σ_B

——每个批次的标准差
- ε

——避免除数为 0 而使用的微小正数

一般情况下,模型每个通道会有一个对应的 γ , γ 表示通道重要性, γ 越大意味着通道对模型性能贡献

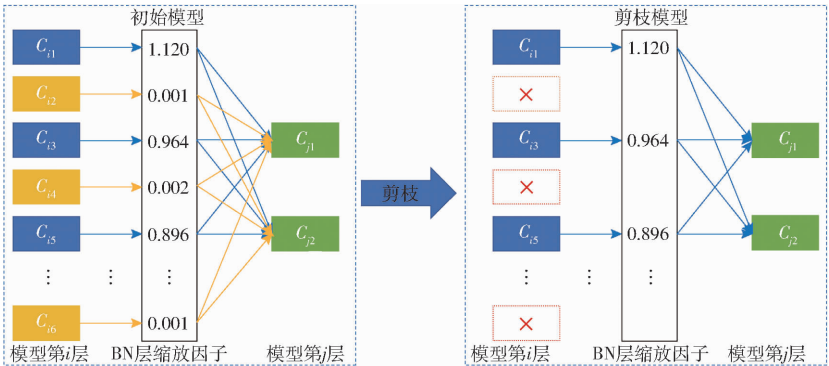


图 5 通道剪枝过程

Fig. 5 Channel pruning process

越大, γ 越小则意味着通道越不重要, 可以剪掉。

为判断识别出 γ 较小的通道, 需要对模型进行稀疏化训练。在模型稀疏训练过程中, 通过在损失函数上添加对 BN 层 γ 的正则化损失, 使其趋于零, 以此来判断 γ 接近于 0 的通道^[23], 模型稀疏训练的损失函数公式为

$$L = l + \lambda \sum_{\gamma \in \Gamma} g(\gamma) \quad (5)$$

式中 L ——模型稀疏训练损失值

l ——模型正常训练损失值

λ ——平衡因子

$g(\gamma)$ ——缩放因子 γ 惩罚函数

对于 $g(\gamma)$, 本文选择广泛用于实现参数稀疏化的 L1 范数, 即 $g(\gamma) = |\gamma|$ 。

2.2.1.2 剪枝与微调

稀疏化训练完成后, γ 接近于 0 的通道对模型识别性能贡献较低, 因此可以将这些通道修剪掉, 以达到压缩模型结构、提高识别速度的目的。剪枝后的模型由于参数大量减少会不可避免地出现精度损失, 需要进行微调训练, 让模型更好地适应新结构, 一定程度上恢复模型丢失的性能。

2.2.2 基于中间特征知识蒸馏的性能恢复

尽管剪枝后的微调训练能帮助模型恢复部分剪枝丢失的性能, 但效果有限。对此, 本文通过知识蒸馏来进一步去除这种性能劣化, 从一个性能较好但参数量和计算量较大的教师模型中提取知识, 将其传递给一个性能较差但参数量和计算量较小的学生模型, 以此来指导学生模型达到更好的性能^[25]。知识蒸馏有标签知识蒸馏和特征知识蒸馏两大类, 其中标签知识蒸馏是将教师模型输出的标签作为学生模型直接学习的知识, 使学生模型的输出标签与教师模型相拟合, 而特征知识蒸馏是将教师模型某些层的特征表达作为学生模型直接学习的知识, 直接匹配中间的特征或学习特征之间的转换关系, 使学生模型某些层的特征表达拟合教师模型对应层的特征表达^[26]。

本文选择多场景群养猪只行为数据集作为模型知识蒸馏的输入, 剪枝前的多场景群养猪只行为识别模型作为教师模型, 剪枝并微调训练后的剪枝微调模型作为学生模型。多场景群养猪只行为数据集中包含不同角度、不同大小、不同遮挡及不同数量的猪只行为目标, 这需要模型有更强的特征学习能力去学习多个场景下不同猪只行为的特征表达。此外, 经过剪枝后, 教师模型和学生模型的对应中间层之间存在一定的通道差异, 这也导致它们的特征表达能力不同。

如果采用标签知识蒸馏, 学生模型仅学习教师模型输出的标签, 而缺乏特征层的指导, 其所能学习到的猪只行为信息有限, 不利于自身性能恢复, 所以学生模型需要学习隐含在教师模型中的特征知识, 用于解决由于通道差异导致的特征表达能力不同的问题^[27], 因此本文采用如图 6 所示的中间特征知识蒸馏进行模型知识蒸馏。

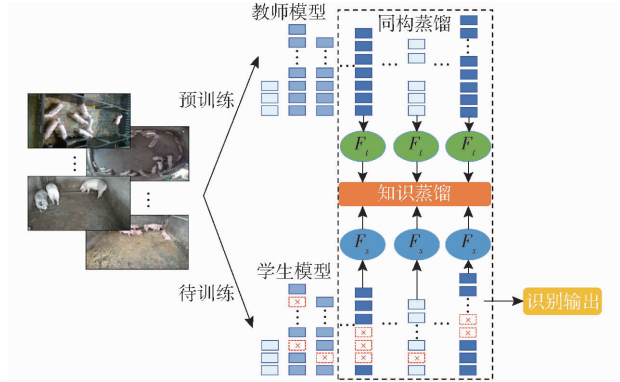


图 6 中间特征知识蒸馏

Fig. 6 Intermediate feature knowledge distillation

由于剪枝并不改变模型的整体架构, 教师模型和学生模型具有相同的架构, 层与层之间一一对应, 因此本文采用如图 6 所示的同构蒸馏, 直接利用教师模型中间层的特征表达指导学生模型对应层进行学习。本文选择教师模型和学生模型的 17、20 和 23 层作为中间特征知识蒸馏时的指导层和被指导层, 同时采用 mimic 方法^[28]计算中间特征知识蒸馏损失, 公式为

$$L_{kd} = \sum_i^n \frac{1}{N_i} \| F_t^{(i)} - r(F_s^{(i)}) \|_2^2 \quad (6)$$

式中 L_{kd} ——中间特征知识蒸馏损失

N_i ——中间层特征维度

$F_t^{(i)}$ ——教师模型中间层特征

$F_s^{(i)}$ ——学生模型中间层特征

r ——特征维度适配器

n ——中间特征知识蒸馏层数量

本文采用 1×1 卷积作为特征维度适配器 r , 以此来补齐学生模型中间层的特征维度, 使其和教师模型间对应中间层的特征维度一致。经特征维度对齐和 mimic 学习后, 学生模型中间层特征表达能尽可能与教师模型对应中间层的特征表达相拟合, 从而提升中间特征知识蒸馏的效果。

3 实验与结果分析

3.1 实验训练平台、嵌入式设备部署及评价指标

3.1.1 训练平台环境配置及训练参数设定

本文的模型实验在 Windows 11 操作系统下进行, 处理器型号为 Intel(R) Core(TM) i5-12400F

2. 50 GHz,显卡型号为 Nvidia GeForce RTX 3060,深度学习框架为 PyTorch 1.12,编程平台为 PyCharm,编程语言为 Python 3.8。

实验设定猪只行为识别模型训练时图像输入尺寸为 640 像素×640 像素,初始学习率为 0.01,权重衰减系数为 0.000 5,优化器为 SGD,动量为 0.937,训练批次大小为 32,训练轮次为 200。

3.1.2 嵌入式设备部署

猪只行为识别模型构建完成后,将其进行部署应用,具体步骤如下:

(1)开发猪只行为识别系统。本文基于 PyQt5 开发可用于 Ubuntu 和 Windows 两种系统的猪只行为识别系统,系统界面如图 7 所示。该系统可对图像、视频进行识别或调用摄像头进行猪只行为实时识别,并统计各行为的整体持续时间。



图 7 猪只行为识别系统界面
Fig. 7 Pig behavior recognition system

(2)猪只行为识别系统部署。猪只行为识别系统开发完成后使用 PyInstaller 将其打包为 .exe 应用程序,无需安装依赖环境便可在其他设备进行部署。本文将系统分别部署于新创云 4200U 和 Intel NUC 两种不同操作系统、不同硬件配置的嵌入式设备上,两种设备的具体配置如表 3 所示。这两种设备小巧轻便,放置空间小,能轻松集成于猪舍巡检机器人。

表 3 两种设备的具体配置

Tab. 3 Specific configurations for two devices			
设备	操作系统	CPU	内存/GB
新创云 4200U	Windows 10	i5-4200U	4
Intel NUC	Ubuntu 20.04	i7-1165G7	32

(3)猪只行为识别模型部署。将猪只行为识别模型权重转化为 ONNX (Open neural network

exchange) 权重,并在两种设备上部署的猪只行为识别系统中调用 ONNX 权重对猪只行为进行识别,统计相应的猪只行为识别持续时间,嵌入式设备部署如图 8 所示。

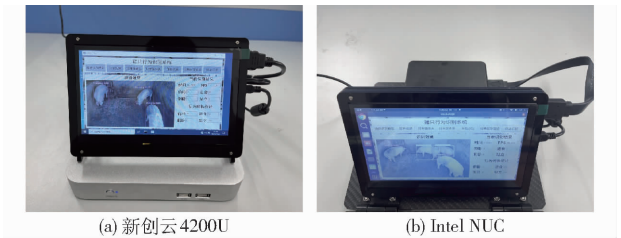


图 8 嵌入式设备部署
Fig. 8 Embedded device deployment

3.1.3 实验评价指标

为综合衡量本文改进对模型性能及嵌入式设备部署的影响,选取精确率 (Precision, P)、召回率 (Recall, R)、平均精度均值 (mean average precision, mAP)、参数量、计算量及内存占用量作为基本评价指标,同时考虑到模型在猪舍监测系统、猪舍巡检机器人等应用的部署要求,增加模型在嵌入式设备上的部署实时识别帧率作为评价指标,用于评价模型部署在嵌入式设备上的性能。

3.2 模型训练结果

3.2.1 测试集上的模型性能

本文基于多场景群养猪只行为数据集,通过模型改进、模型剪枝以及知识蒸馏构建了轻量级多场景群养猪只行为识别模型 YOLO v5n-PBR。该模型在测试集上所有类别的 mAP 为 96.9%,参数量为 4.7×10^5 ,计算量为 1.2×10^9 ,内存占用量为 1.2 MB。图 9 为 YOLO v5n-PBR 模型对多养殖场场景下猪只的俯卧、侧卧、进食及站立 4 种行为的识别效果,可见该模型整体上能很好地识别多场景下的群养猪只行为。

3.2.2 模型泛化性能

为验证 YOLO v5n-PBR 模型的泛化性能,本文将在多场景群养猪只行为数据集上训练得到的 YOLO v5n-PBR 模型和 YOLO v5n 模型分别用于 4 个单场景或双场景的群养猪只行为数据集进行泛化性实验,实验结果如表 4 所示。

由表 4 可知,在仅有少量参数的情况下,基于多场景群养猪只行为数据集训练得到的 YOLO v5n-



图 9 YOLO v5n-PBR 模型识别效果
Fig. 9 Recognition effects of YOLO v5n-PBR model

表 4 泛化性实验结果

Tab. 4 Generalization experimental results %				
数据集	模型	P	R	mAP
数据集 A	YOLO v5n	95.3	95.0	98.0
	YOLO v5n-PBR	95.1	95.9	98.2
数据集 B	YOLO v5n	97.0	95.3	98.7
	YOLO v5n-PBR	97.4	95.1	98.6
数据集 C	YOLO v5n	96.2	95.2	98.0
	YOLO v5n-PBR	95.6	95.7	98.1
数据集 D	YOLO v5n	95.4	92.7	97.6
	YOLO v5n-PBR	95.2	94.8	98.1

PBR 模型在 4 个单场景或双场景的群养猪只行为数据集上的 mAP 均能达到 98.1%，与原始模型 YOLO v5n 相近。实验结果表明,YOLO v5n-PBR 模型能较好地识别单场景或双场景下的群养猪只行为,具有较强的泛化性,同时基于多场景数据集训练的方式也能为模型后续迁移应用于其他养殖场景提供基础。

3.3 消融实验

为验证各方法的有效性,以 YOLO v5n 模型为基准模型进行消融实验,实验结果如表 5 所示。由表 5 可知,在 YOLO v5n 模型的训练过程中,是否采用迁移学习方法对模型识别精度有显著影响:采用迁移学习方法后,YOLO v5n 模型 mAP 为 97.0%,比未采用迁移学习方法时提高 1.2 个百分点。实验结果表明,迁移学习方法能使 YOLO v5n 模型在掌握 COCO 数据集的通用特征后,更好地学习多场景下的群养猪只行为特征。

表 5 消融实验结果

Tab.5 Results of ablation experiment					%
基准模型	迁移学习	OTA	<i>P</i>	<i>R</i>	mAP
YOLO v5n			93.1	91.1	95.8
	√		94.6	94.9	97.0
	√	√	95.7	95.4	97.6

注:“√”表示训练过程中使用该方法。下同。

为进一步验证 OTA 标签分配方法的效果,本文进行了 OTA 标签分配方法在模型剪枝及知识蒸馏中的消融实验,实验结果如表 6 所示。

表 6 OTA 消融实验结果

Tab. 6 Results of OTA ablation experiment						%
基准模型	剪枝	知识蒸馏	OTA	<i>P</i>	<i>R</i>	mAP
多场景群养猪只	✓			91.1	89.3	94.8
	✓		✓	91.8	91.8	96.0
行为识别模型	✓	✓		92.6	91.8	96.0
	✓	✓	✓	94.2	92.9	96.9

结合表 5 和表 6 可知,引入 OTA 标签分配方法,能有效解决模型对多场景群养猪只行为数据集

中大量不同程度遮挡、不同大小目标和不同角度猪只行为标签的分配问题,实现动态最优分配,对模型初始训练、剪枝和知识蒸馏的精确率、召回率和 mAP 都有一定的提升。

引入迁移学习方法和 OTA 标签分配方法前后的 YOLO v5n 模型的 mAP 变化曲线如图 10 所示,通过图 10 可以直观地看到:引入迁移学习方法和 OTA 标签分配方法在一定程度上加快了模型收敛速度,提升了模型性能。

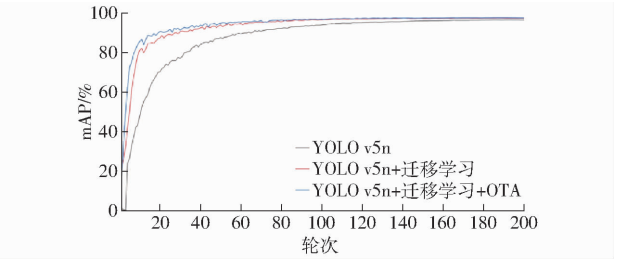


图 10 mAP 变化曲线
Fig. 10 Change curves of mAP

在 YOLO v5n 模型的训练中引入迁移学习方法和 OTA 标签分配方法后,得到多场景群养猪只行为识别模型。该模型 mAP 为 97.6%,比原始模型 YOLO v5n 提高 1.8 个百分点,在多场景群养猪只行为数据集上可准确识别猪只行为,有效完成猪只行为识别任务。

3.4 模型剪枝效果

3.4.1 不同稀疏率对模型稀疏训练的影响

为确定最佳稀疏率,以稀疏率为 0.000 1、0.001、0.002 及 0.003 分别对多场景群养猪只行为识别模型进行稀疏训练。稀疏训练过程中,不同稀疏率的 γ 变化趋势如图 11 所示。

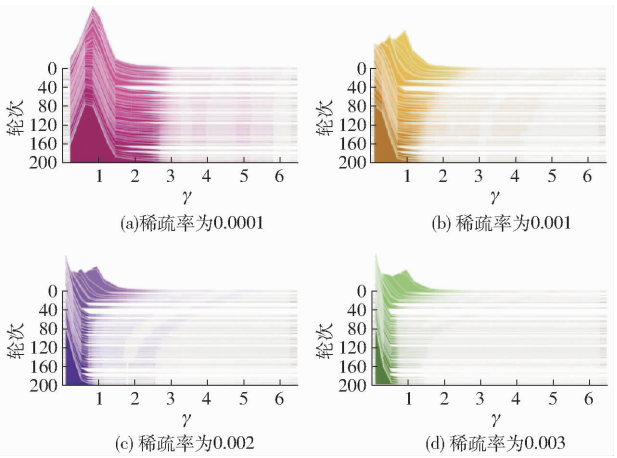


图 11 不同稀疏率的 γ 变化趋势
Fig. 11 The γ trends of different sparsity rates

由图 11 可知, γ 随着稀疏训练的进行逐渐趋近于 0,并且稀疏率越大, γ 趋近于 0 的速度越快。稀疏训练完成后,不同稀疏率的稀疏训练模型性能如

表 7 所示,结合图 7 可知:稀疏率过小时,稀疏训练的速度慢,辨别不出通道重要性;稀疏率过大时,模型性能下降较多。综合考虑稀疏训练速度及模型性能,本文稀疏率取 0.002。

表 7 不同稀疏率的模型性能

Tab.7 Performance of models with different sparsity rates				%
稀疏率	P	R	mAP	
0.000 1	96.2	95.5	97.5	
0.001	96.2	95.5	97.5	
0.002	95.9	94.1	97.5	
0.003	94.4	93.6	97.1	

3.4.2 不同剪枝率及微调对模型的影响

稀疏训练完成后,本文设置不同的剪枝率对稀疏率 0.002 稀疏训练后的多场景群养猪只行为识别模型进行剪枝,将缩放因子 γ 较小的通道剪掉,随后进行微调训练,剪枝微调后的模型性能如表 8 所示。

表 8 不同剪枝率的模型性能

Tab.8 Model performance with different pruning rates						
剪枝率	P/%	R/%	mAP/%	参数量	计算量	内存占用量/MB
0.40	95.7	95.4	97.6	9.6×10^5	2.3×10^9	2.2
0.50	95.1	95.7	97.6	8.7×10^5	2.0×10^9	2.0
0.60	94.0	94.0	97.0	6.8×10^5	1.6×10^9	1.6
0.70	91.8	91.8	96.0	4.7×10^5	1.2×10^9	1.2
0.75	90.9	91.1	95.3	4.2×10^5	1.0×10^9	1.1

由表 8 可知,模型经过不同的剪枝率剪枝并微调后,参数量、计算量及内存占用量均有不同程度的减小,而 mAP 也随着模型的压缩不可避免地出现了下降。当剪枝率小于 0.70 时,模型剪枝效果显著;当剪枝率大于 0.70 时,模型中冗余参数基本已被修剪完,模型剪枝效果较差,mAP 下降较多。因此,本文选择剪枝率为 0.70,可在保持较高精度的同时有效减少模型参数量、计算量及内存占用量。

3.4.3 剪枝后模型变化

以稀疏率 0.002、剪枝率 0.70 对多场景群养猪只行为识别模型进行剪枝后,模型各层通道数变化如图 12 所示。由图 12 可以看出,模型多数层的通道数量大幅降低,表明 L1-norm 算法可有效减少模型参数量、计算量及内存占用量。

3.5 不同知识蒸馏算法效果

为验证本文所采用的中间特征知识蒸馏算法的蒸馏效果,本文对中间特征知识蒸馏和标签知识蒸馏 2 种知识蒸馏算法进行知识蒸馏模型的性能对比实验,实验结果如表 9 所示。

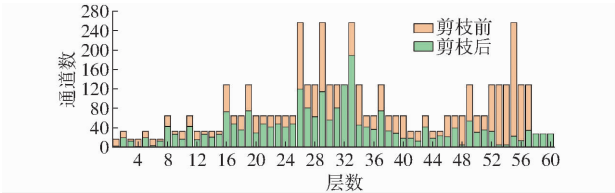


图 12 模型通道数变化

Fig.12 Change in number of channels in model

表 9 不同知识蒸馏模型性能

Tab.9 Performance of different knowledge distillation models				%
知识蒸馏方法	P	R	mAP	
无蒸馏	91.8	91.8	96.0	
标签知识蒸馏	93.4	93.0	96.3	
中间特征知识蒸馏	94.2	92.9	96.9	

由表 9 可知,经中间特征知识蒸馏后,模型的精确率、召回率和 mAP 分别提高 2.4、1.1、0.9 个百分点,其中 mAP 比采用标签知识蒸馏时高 0.6 个百分点。实验结果表明,采用中间特征知识蒸馏能将教师模型中间层更具表征能力的特征知识转移给学生模型,以此来提高知识蒸馏过程中传输的知识的表征能力和信息量,使得学生模型能学习到教师模型对于多场景群养猪只行为的特征表达,有效提升知识蒸馏效果。

3.6 与主流深度学习模型对比实验

为验证 YOLO v5n-PBR 模型对多场景群养猪只行为的识别效果,本文将其与猪只行为识别任务中常用的主流深度学习模型 SSD512、Faster R-CNN、YOLO v5n、YOLO v6n^[29]、YOLO v7-tiny^[30]及 YOLO v8n^[31]进行了性能对比实验,其中 YOLO v7-tiny 为 YOLO v7 的轻量级模型,实验结果如表 10 所示。

由表 10 可知,与原始模型 YOLO v5n 相比,YOLO v5n-PBR 模型在 mAP 提高 1.1 个百分点的情况下,参数量、计算量和内存占用量分别减少 73.3%、70.7% 和 68.4%。与其余 5 种模型相比,YOLO v5n-PBR 模型的 mAP 仅比 YOLO v8n 模型低 0.1 个百分点,在参数量、计算量和内存占用量上均为最佳,其中参数量分别比 SSD512、Faster R-CNN、YOLO v6n、YOLO v7-tiny 和 YOLO v8n 模型减少 2.354×10^7 、 1.363×10^8 、 3.760×10^6 、 5.540×10^6 和 2.540×10^6 ,计算量分别减少 8.66×10^{10} 、 2.00×10^{11} 、 1.06×10^{10} 、 1.18×10^{10} 和 6.90×10^9 ,内存占用量分别减少 90.9、520.7、7.5、11.1、5.0 MB。

3.7 模型嵌入式设备部署实验

为验证 YOLO v5n-PBR 模型在嵌入式设备部署时的实时识别速度,本文将 YOLO v5n、YOLO v5n-

表 10 不同模型的性能

Tab. 10 Performance of different models

模型	<i>P</i> /%	<i>R</i> /%	mAP/%	参数量	计算量	内存占用量/MB
SSD512	84. 1	79. 2	88. 1	2.401×10^7	8.78×10^{10}	92. 1
Faster R – CNN	78. 1	90. 3	89. 9	1.368×10^8	2.01×10^{11}	521. 9
YOLO v5n	93. 1	91. 1	95. 8	1.760×10^6	4.10×10^9	3. 8
YOLO v5n – PBR	94. 2	92. 9	96. 9	4.700×10^5	1.20×10^9	1. 2
YOLO v6n	92. 6	91. 3	95. 5	4.230×10^6	1.18×10^{10}	8. 7
YOLO v7 – tiny	92. 5	92. 0	96. 3	6.010×10^6	1.30×10^{10}	12. 3
YOLO v8n	93. 9	94. 0	97. 0	3.010×10^6	8.10×10^9	6. 2

PBR 和 YOLO v7 – tiny 模型部署于新创云 4200U 和 Intel NUC,分别在河源市连平县江波养殖场数据源和公开数据集 1 数据源中选取 3 段时长 1 min 的猪只活跃视频作为模型嵌入式设备部署测试数据,上述 6 段视频包含俯卧、侧卧、进食及站立 4 种行为,3 种模型的部署实时识别帧率如表 11 所示。

表 11 部署实时识别帧率

Tab. 11 Deploy real-time recognition of frame rates

模型	帧/s	
	新创云 4200U	Intel NUC
YOLO v5n	7. 0	36. 2
YOLO v5n – PBR	12. 2	66. 3
YOLO v7 – tiny	4. 6	21. 6

由表 11 可知,YOLO v5n – PBR 模型在两种设备上的部署实时识别帧率在 3 种模型中最高。在新创云 4200U 上的部署实时识别帧率为 12. 2 帧/s,分别比 YOLO v5n 模型和 YOLO v7 – tiny 模型高 74. 3% 和 165. 2%;在 Intel NUC 上的部署实时识别帧率为 66. 3 帧/s,分别比 YOLO v5n 模型和 YOLO v7 – tiny 模型高 83. 1% 和 206. 9%。

为进一步验证 YOLO v5n – PBR 模型在嵌入式设备部署时的识别精确率和召回率,本文用猪只行为视频每秒的最后一帧图像中的猪只行为代表猪只当前 1 s 的行为,分别由 YOLO v5n、YOLO v5n – PBR、YOLO v7 – tiny 和人工对 6 个视频段内群体猪只行为次数进行识别统计,3 种模型和人工识别统计结果如表 12 所示。

由表 12 可知,YOLO v5n – PBR 模型对猪只群体 4 类行为的整体识别统计结果与人工识别统计相近,并且整体识别统计效果优于 YOLO v5n 模型和 YOLO v7 – tiny 模型,表明该模型对猪只行为视频中的猪只行为的识别统计误差较小。

表 13 为 YOLO v5n – PBR 模型的详细识别统计结果。由表 13 可知,YOLO v5n – PBR 模型对于猪只群体 4 种行为漏识别和误识别次数均较小,平均精确率和平均召回率均为 95. 3%。

综合表 11 ~ 13 可知,本文开发的猪只行为识别

表 12 模型与人工识别统计结果

Tab. 12 Model and human recognition statistics

行为类别	人工统计	模型统计		
		YOLO v5n	YOLO v5n – PBR	YOLO v7 – tiny
俯卧	662	662	650	636
侧卧	334	323	331	336
进食	319	343	329	339
站立	879	841	876	887

表 13 详细识别统计结果

Tab. 13 Detailed recognition of statistical results

行为类别	漏识别数	误识别数	<i>P</i> /%	<i>R</i> /%
俯卧	44	32	95. 1	93. 4
侧卧	27	24	92. 7	91. 9
进食	7	17	94. 8	97. 8
站立	16	13	98. 5	98. 2

系统能在两种不同系统、不同配置的嵌入式设备中运行,并能调用相应的模型对猪只行为视频进行不同行为的实时识别统计。此外,虽然 YOLO v5n – PBR 模型经过轻量化后仅保留少量参数,但将其部署于两种嵌入式设备上对两种不同养殖场景的猪只行为视频进行实时识别时,部署实时识别帧率远优于 YOLO v5n 模型和 YOLO v7 – tiny 模型,同时有着和人工识别统计相近的识别统计结果、较高的平均精确率和平均召回率,表明 YOLO v5n – PBR 模型在嵌入式设备部署时依然有识别速度快、精度高、泛化性强等优点。

4 结论

(1)提出了一种轻量级多场景群养猪只行为识别模型 YOLO v5n – PBR。首先构建多场景群养猪只行为数据集,并根据该数据集中猪只行为目标的特点,在 YOLO v5n 模型的训练中引入迁移学习方法和 OTA 标签分配方法,得到多场景群养猪只行为识别模型,然后利用 L1 – norm 剪枝算法对模型进行轻量化,最后采用微调训练和中间特征知识蒸馏去除剪枝后的性能劣化。YOLO v5n – PBR 模型的 mAP 为 96. 9%,参数量为 4.700×10^5 ,计算量为

1.20 × 10⁹, 内存占用量为 1.2 MB。与原始模型 YOLO v5n 相比, YOLO v5n - PBR 模型在 mAP 提高 1.1 个百分点的情况下, 参数量、计算量和内存占用量分别减少 73.3%、70.7% 和 68.4%。

(2) 基于多场景群养猪只行为数据集上训练得到的 YOLO v5n - PBR 模型可很好地泛化应用于 4 个单场景或双场景的群养猪只行为数据集上进行猪只行为识别, 识别 mAP 均能达到 98.1%, 与原始模型 YOLO v5n 相近; 采用中间特征知识蒸馏能更好地帮助学生模型学习教师模型中关于多场景群养猪只行为的特征表达, 有效恢复剪枝丢失的性能, 知识蒸馏后的模型 mAP 比蒸馏前提高 0.9 个百分点; YOLO v5n - PBR 模型与猪只行为

识别常用的深度学习模型相比, mAP 仅比 YOLO v8n 模型低 0.1 个百分点, 在参数量、计算量等方面优于其他模型。

(3) 将 YOLO v5n - PBR 模型部署于两种嵌入式设备上, 并对两种不同养殖场景的 6 段猪只行为视频进行识别统计, 部署实时识别帧率分别为 12.2 帧/s 和 66.3 帧/s, 比原始模型 YOLO v5n 高 74.3% 和 83.1%, 识别统计结果与人工识别统计结果相近, 4 种行为的平均精确率和平均召回率均为 95.3%, 表明 YOLO v5n - PBR 模型能满足嵌入式设备部署要求, 且实时识别帧率和行为识别精度能满足猪只行为视频流实时、准确识别的要求。

参 考 文 献

[1] 叶建伟, 薛亚辉, 王明宇, 等. 行为自动检测技术在生猪健康和福利方面的研究进展[J]. 家畜生态学报, 2019, 40(10): 8 - 13, 27.
YE Jianwei, XUE Yahui, WANG Mingyu, et al. Research progress on automatic pig behavior detection technology in animal health and welfare[J]. Journal of Domestic Animal Ecology, 2019, 40(10): 8 - 13, 27. (in Chinese)

[2] 汪开英, 赵晓洋, 何勇. 畜禽行为及生理信息的无损监测技术研究进展[J]. 农业工程学报, 2017, 33(20): 197 - 209.
WANG Kaiying, ZHAO Xiaoyang, HE Yong. Review on noninvasive monitoring technology of poultry behavior and physiological information[J]. Transactions of the CSAE, 2017, 33(20): 197 - 209. (in Chinese)

[3] 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)

[4] 刘龙申, 沈明霞, 姚文, 等. 基于加速度传感器的母猪产前行为特征采集与分析[J]. 农业机械学报, 2013, 44(3): 192 - 196, 191.
LIU Longshen, SHEN Mingxia, YAO Wen, et al. Acquisition and analysis of sows' behavior before farrowing based on acceleration sensor[J]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(3): 192 - 196, 191. (in Chinese)

[5] RINGGENBERG N, BERGERON R, DEVILLERS N. Validation of accelerometers to automatically record sow postures and stepping behaviour[J]. Applied Animal Behaviour Science, 2010, 128(1 - 4): 37 - 44.

[6] NILSSON M, HERLIN H A, ARDÖ H, et al. Development of automatic surveillance of animal behaviour and welfare using image analysis and machine learned segmentation technique[J]. Animal, 2015, 9(11): 1859 - 1865.

[7] NASIRAHMADI A, RICHTER U, HENSEL O, et al. Using machine vision for investigation of changes in pig group lying patterns[J]. Computers and Electronics in Agriculture, 2015, 119: 184 - 190.

[8] 刘东, 何东健, 陈晨, 等. 基于 ALR - GMM 的群养猪攻击行为识别算法研究[J]. 农业机械学报, 2021, 52(1): 201 - 208.
LIU Dong, HE Dongjian, CHEN Chen, et al. Recognition of aggressive behaviour in group-housed pigs based on ALR - GMM [J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(1): 201 - 208. (in Chinese)

[9] 滕光辉, 冀横溢, 庄晏榕, 等. 深度学习在猪只饲养过程的应用研究进展[J]. 农业工程学报, 2022, 38(14): 235 - 249.
TENG Guanghui, JI Hengyi, ZHUANG Yanrong, et al. Research progress of deep learning in the process of pig feeding[J]. Transactions of the CSAE, 2022, 38(14): 235 - 249. (in Chinese)

[10] 邓铭辉, 龚俊杰, 郑飘逸, 等. 基于改进 YOLOX 的群养生猪轻量化目标检测方法[J]. 农业机械学报, 2023, 54(11): 277 - 285.
DENG Minghui, GONG Junjie, ZHENG Piaoyi, et al. Lightweight target detection method for group-raised pigs based on improved YOLOX [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(11): 277 - 285. (in Chinese)

[11] 涂淑琴, 汤寅杰, 李承桀, 等. 基于改进 ByteTrack 算法的群养生猪行为识别与跟踪技术[J]. 农业机械学报, 2022, 53(12): 264 - 272.
TU Shuqin, TANG Yinjie, LI Chengjie, et al. Behavior recognition and tracking of group-housed pigs based on improved ByteTrack algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(12): 264 - 272. (in Chinese)

[12] 李菊霞, 李艳文, 牛帆, 等. 基于 YOLOv4 的猪只饮食行为检测方法[J]. 农业机械学报, 2021, 52(3): 251 - 256.
LI Juxia, LI Yanwen, NIU Fan, et al. Pig diet behavior detection method based on YOLOv4[J]. Transactions of the Chinese

Society for Agricultural Machinery, 2021, 52(3): 251 – 256. (in Chinese)

[13] CHEN C, ZHU W X, NORTON T. Behaviour recognition of pigs and cattle: journey from computer vision to deep learning [J]. Computers and Electronics in Agriculture, 2021, 187: 106255.

[14] 高云,陈斌,廖慧敏,等. 群养猪侵略性行为的深度学习识别方法[J]. 农业工程学报, 2019, 35(23): 192 – 200.
GAO Yun, CHEN Bin, LIAO Huimin, et al. Recognition method for aggressive behavior of group pigs based on deep learning [J]. Transactions of the CSAE, 2019,35(23): 192 – 200. (in Chinese)

[15] 薛月菊,朱勋沐,郑婵,等. 基于改进 Faster R – CNN 识别深度视频图像哺乳母猪姿态[J]. 农业工程学报, 2018, 34(9):189 – 196.
XUE Yueju, ZHU Xunmu, ZHENG Chan, et al. Lactating sow postures recognition from depth image of videos based on improved Faster R – CNN[J]. Transactions of the CSAE, 2018, 34(9): 189 – 196. (in Chinese)

[16] ZHANG Y Q, CAI J H, XIAO D Q, et al. Real-time sow behavior detection based on deep learning[J]. Computers and Electronics in Agriculture, 2019, 163:104884.

[17] BERGAMINI L, PINI S, SIMONI A, et al. Extracting accurate long-term behavior changes from a large pig dataset[C] // Beijing;16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP), 2021: 524 – 533.

[18] 科大讯飞股份有限公司. 猪只盘点挑战赛[EB/OL]. (2021 – 06 – 21) [2023 – 12 – 08]. <https://challenge.xfyun.cn/topic/info?type=pig-check&ch=ds22-kol-gl02>.

[19] 薛鸿翔,沈明霞,刘龙申,等. 基于改进 YOLO v5s 的经产母猪发情检测方法研究[J]. 农业机械学报, 2023, 54(1): 263 – 270.
XUE Hongxiang, SHEN Mingxia, LIU Longshen, et al. Estrus detection method of parturient sows based on improved YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 263 – 270. (in Chinese)

[20] 庄福振,罗平,何清,等. 迁移学习研究进展[J]. 软件学报,2015, 26(1):26 – 39.
ZHUANG Fuzhen, LUO Ping, HE Qing, et al. Survey on transfer learning research[J]. Journal of Software, 2015, 26(1): 26 – 39. (in Chinese)

[21] GE Z, LIU S T, LI Z M, et al. OTA:optimal transport assignment for object detection[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 303 – 312.

[22] CUTURI M. Sinkhorn distances: lightspeed computation of optimal transport [C] // Proceedings of the 26th International Conference on Neural Information Processing Systems, 2013: 2292 – 2300.

[23] LIU Z, LI J, SHEN Z, et al. Learning efficient convolutional networks through network slimming[C] // Proceedings of the IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 2736 – 2744.

[24] BJORCK J, GOMES C, SELMAN B, et al. Understanding batch normalization[C] // Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montreal: MIT Press, 2018: 7705 – 7716.

[25] 孟宪法,刘方,李广,等. 卷积神经网络压缩中的知识蒸馏技术综述[J]. 计算机科学与探索, 2021, 15(10): 1812 – 1829.
MENG Xianfa, LIU Fang, LI Guang, et al. Review of knowledge distillation in convolutional neural network compression[J]. Journal of Frontiers of Computer Science and Technology, 2021, 15(10): 1812 – 1829. (in Chinese)

[26] 邵仁荣,刘宇昂,张伟,等. 深度学习中知识蒸馏研究综述[J]. 计算机学报,2022, 45(8):1638 – 1673.
SHAO Renrong, LIU Yuang, ZHANG Wei, et al. A survey of knowledge distillation in deep learning[J]. Chinese Journal of Computers, 2022,45(8): 1638 – 1673. (in Chinese)

[27] 黄震华,杨顺志,林威,等. 知识蒸馏研究综述[J]. 计算机学报, 2022, 45(3):624 – 653.
HUANG Zhenhua, YANG Shunzhi, LIN Wei, et al. Knowledge distillation: a survey[J]. Chinese Journal of Computers, 2022, 45(3): 624 – 653. (in Chinese)

[28] LI Q Q, JIN S Y, YAN J J. Mimicking very efficient network for object detection[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6356 – 6364.

[29] LI C, LI L, JIANG H, et al. YOLOv6: a single-stage object detection framework for industrial applications[EB/OL]. (2022 – 09 – 07) [2023 – 12 – 08]. <https://arxiv.org/abs/2209.02976>.

[30] WANG C Y,BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464 – 7475.

[31] 王金鹏,何萌,甄乾广,等. 基于改进 COF – YOLOv 8n 的油茶果静态与动态检测计数方法[J]. 农业机械学报, 2024, 55(4): 193 – 203.
WANG Jinpeng, HE Meng, ZHEN Qianguang, et al. *Camellia oleifera* fruit static and dynamic detection counting based on improved COF – YOLO v8n[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(4): 193 – 203. (in Chinese)