doi:10.6041/j.issn.1000-1298.2023.01.025

基于改进 YOLO v5n 的猪只盘点算法

杨秋妹[1,2]　陈淼彬[1,2]　黄一桂[1,2]　肖德琴[1,2]　刘又夫[1,2]　周家鑫[1,2]
(1.华南农业大学数学与信息学院,广州 510642;2.农业农村部华南热带智慧农业技术重点实验室,广州 510642)

摘要:猪只盘点是规模化养殖中的重要环节,为生猪精准饲喂和资产管理提供了依据。人工盘点不仅耗时低效,而且容易出错。当前已有基于深度学习的生猪智能盘点算法,但在遮挡重叠、光照等复杂场景下盘点精度较低。为提高复杂场景下生猪盘点的精度,提出了一种基于改进 YOLO v5n 的猪只盘点算法。该算法从提升猪只目标检测性能出发,构建了一个多场景的生猪数据集;其次,在主干网络中引入 SE-Net 通道注意力模块,引导模型更加关注遮挡条件下猪只目标信息的通道特征。同时,增加了检测层进行多尺度特征融合处理,使模型更容易学习收敛并预测不同尺度的猪只对象,提升模型遮挡场景的检测性能;最后,对边界框损失函数以及非极大值抑制处理进行了改进,使模型对遮挡的目标有更好的识别效果。实验结果表明,与原 YOLO v5n 算法相比,改进算法的平均绝对误差(MAE)、均方根误差(RMSE)以及漏检率分别降低 0.509、0.708 以及 3.02 个百分点,平均精度(AP)提高 1.62 个百分点,达到 99.39%,在复杂遮挡重叠场景下具有较优的精确度和鲁棒性。算法的 MAE 为 0.173,与猪只盘点算法 CClusnet、CCNN 和 PCN 相比,分别降低 0.257、1.497 和 1.567。在时间性能上,单幅图像的平均识别时间仅为 0.056 s,符合实际猪场生产的实时性要求。

关键词:猪只计数;目标检测;注意力机制;多尺度感知
中图分类号:TP391　　**文献标识码**:A　　**文章编号**:1000-1298(2023)01-0251-12　　**OSID**:

Pig Counting Algorithm Based on Improved YOLO v5n

YANG Qiumei[1,2]　CHEN Miaobin[1,2]　HUANG Yigui[1,2]　XIAO Deqin[1,2]　LIU Youfu[1,2]　ZHOU Jiaxin[1,2]
(1. *College of Mathematics and Informatics*, *South China Agricultural University*, *Guangzhou* 510642, *China*
2. *Key Laboratory of Smart Agricultural Technology in Tropical South China*, *Ministry of Agriculture and Rural Affairs*, *Guangzhou* 510642, *China*)

Abstract: Pig counting is an important part in large-scale farming, which provides the basis for precise pig feeding and asset management. Manual counting is both time-consuming and inefficient, more than error-prone. In recent years, as the performance of deep learning systems far outperforms traditional machine learning systems, deep learning-based methods have demonstrated state-of-the-art performance in tasks such as image classification, segmentation, and object detection. Although there are currently existing intelligent pig counting algorithms based on deep learning, the counting accuracy is low in complex scenes such as occlusion and different illumination. So as to increase the accuracy of pig counting in complex scenarios, a pig counting algorithm was proposed based on improved YOLO v5n. Starting from improving the performance of pig target detection, the algorithm constructed a multi-scene pig dataset. In the field of target detection, each target was surrounded by the surrounding background, and the environment around the target object had rich contextual information. However, in the deep convolutional neural network, although the convolutional layer can capture the features of the image from the global receptive field to describe the image, it essentially only modeled the spatial information of the image without modeling the information between channels. By introducing the SE-Net channel attention module into the Backbone network, the model was guided to place greater emphasis on the channel features of the pig target information under occlusion conditions, so that it can better locate the features to be detected and enhance the network performance. At the same time, there may be pig targets of various

收稿日期:2022-09-30　修回日期:2022-11-08
基金项目:国家重点研发计划项目(2021YFD2000802)
作者简介:杨秋妹(1983—),女,讲师,博士,主要从事农业图像视频处理研究,E-mail:yqmbegonia@163.com
通信作者:肖德琴(1970—),女,教授,博士生导师,主要从事物联网和农业图像视频处理研究,E-mail:deqinx@scau.edu.cn

scales in an actual picture of a dense scene of a pig farm. In order to deal with the complex and densely occluded actual production pig farm scene and obtain more abundant and comprehensive feature information, a detection layer was added based on the original three detection layers of different scales for multi-scale feature detection, so as to better learn the multi-level features of the occlusion target and improve the detection performance of model complex occlusion scenes. Finally, the loss function of the boundary box and the non-maximum suppression processing were improved to make the model have better recognition effect on the occluded targets. According to experimental results, in contrast with the original YOLO v5n algorithm, the mean absolute error (MAE), root mean square error (RMSE) and missed detection rate of the improved algorithm were reduced by 0.509, 0.708 and 3.02 percentage points, respectively, and the average precision (AP) was improved by 1.62 percentage points to 99.39%. The improved algorithm had high accuracy and good robustness in complex occlusion overlapping scenarios. Compared with other pig inventory algorithms: CClusnet, CCNN and PCN, the MAE of this algorithm was 0.173, which was 0.257, 1.497 and 1.567 lower than that of the other three algorithms. In terms of time performance, it only took 0.056 s to recognize a single image on average, satisfying the real-time requirements of actual pig farm production.

Key words: pig counting; object detection; attention mechanism; multi-scale perception

0 引言

猪只盘点是猪场集约化资产管理中的一项重要任务。依据准确的猪只数量可以制定更加合理的养殖计划,帮助养殖户降低成本,减少不必要的损失,提高养殖场收益。传统的人工盘点方法需要高劳动力成本,且容易受到猪群重叠、相互遮挡、摄像头视角或者光照等影响,容易出错。另外,人工盘点会引起猪只的应激反应,出现猪只受惊或乱跑,增加盘点的难度,无法保证盘点的准确性。

随着规模化、信息化智慧养猪建设的快速推进,计算机视觉算法因其具有无侵扰、不间断、成本低等优势,成为智能猪只盘点方法的主要发展趋势[1]。早期的视觉猪只盘点方法利用形态学算法和区域生长法相结合的方法将背景与前景分割,从而实现猪只数量的统计[2-3]。深度学习方法在处理复杂遮挡计数问题上有很大的优势[4-8]。目前,有不少的研究人员已经将深度学习应用于猪只盘点。基于神经网络的密度图回归的计数方法受到不少研究人员的重视[9-11]。此类方法利用像素将图像转换为密度图回归,并对密度图进行积分得到计数。文献[12]提出了一个自底向上的猪只身体关键点检测方法,然后再基于 STRF(空间敏感时序响应)方法进行猪只计数。文献[13]用改进的 Mask R-CNN 网络减少了因光照和生猪遮挡致使猪只目标漏检的问题,其网络在 12~22 头猪只的单栏的计数准确率达到 98%。

综上,基于深度学习的猪只盘点算法虽然优于人工计数方法,但由于猪是一种群居的动物,通常挨靠在一起致使猪只互相遮挡重叠。而且同一猪舍的猪只颜色相近,当猪只体型发生变化的时候,更容易导致猪只个体难以区分,使得识别难度增大。因此,针对猪只相互遮挡、体型变化或者光线暗的情况下导致漏检或者多检的问题,本文结合注意力机制和多尺度特征检测,在 YOLO v5n 网络的基础上进行改进,拟实现一种改进 YOLO v5n 猪只计数算法。

1 数据集构建与计数原理

1.1 数据来源

本文的实验数据来源于 5 种不同的猪场数据集,分别是 3 个自建实验猪场数据集:广州萝岗实验猪场、云浮范金树实验猪场、隔离舍实验猪场,以及两个公开数据集:文献[14-15]公开的数据集。广州萝岗实验猪场选取长白后备母猪为实验对象,猪栏尺寸为 4 m×5 m,单栏内饲养 4 头猪,猪只数量较少;云浮范金树实验猪场选取育肥期的长白母猪为实验对象,群养猪栏尺寸为 5 m×8 m,单栏内平均饲养 40 头猪,猪只数量较多,但局限于摄像头拍照角度肉眼可见的猪只数量平均为 26 头;隔离舍实验猪场选取育肥期的长白母猪为实验对象,猪栏尺寸为 2 m×4 m,单栏内饲养 6 头猪;文献[14]的数据集描述了 17 个不同的猪圈位置,包括 1.5~5.5 月龄猪只,单栏内饲养的猪只数量有 7、8、10、11、14、15 头,既包括日间拍摄的图像也包括晚上拍摄的图像;文献[15]实验猪场的单个猪圈尺寸为5.8 m×1.9 m,有 8 头猪,体质量约为 30 kg。各实验猪场信息如表 1 所示。

广州萝岗猪场、云浮范金树猪场、隔离舍猪场均采用海康威视 DS-2CD3345-I 型红外网络摄像头。广州萝岗猪场每天 24 h 共持续 32 d 对猪栏进行监控,但只选取 08:00—18:00 猪只比较活跃的时段。云浮范金树猪场与隔离舍猪场每天 24 h 不间

断拍摄监控，同时选取白天活跃时段和夜晚睡眠阶段拍摄。

表 1　各猪场信息

Tab. 1　Information of pig farms

猪场	生长阶段	猪舍尺寸/(m×m)	单栏饲养数量/只	选取时间
广州萝岗猪场	育肥期	4×5	4	08:00—18:00
云浮范金树猪场	育肥期	5×8	40	08:00—18:00
隔离舍猪场	育肥期	2×4	6	23:00—03:00
文献[14]	保育期、育肥期		7~15	白天、夜晚
文献[15]	育肥期	5.8×1.9	8	07:00—19:00

1.2　数据预处理与数据集构建

对采集到的视频每隔 35 帧截取一幅图像，并去除相似度过高、清晰度较差的图像。为更准确地计算单栏内猪只数量，排除其他相邻栏位猪只对实验计数结果的影响，本文对包含多个猪栏的图像进行了掩膜处理，在原始图像上以盘点猪栏为界限，添加一掩膜层，遮住其他栏位的猪只，掩膜的处理效果如图 1 所示。

(a) 处理前　　(b) 处理后

图 1　掩膜处理图像

Fig. 1　Mask processing pictures

最终数据集共有 3 000 幅图像，并用 LabelImg 标注软件进行标注。整个猪只盘点数据集按 6∶2∶2 的比例划分训练集、验证集和测试集。为保证数据集的多样性，利用 Python 脚本程序分别从各猪场随机选取 600 幅图像，其中训练集 360 幅图像，验证集和测试集均为 120 幅图像。依据上述方法本文构建的猪只盘点数据集如表 2 所示。

表 2　猪只盘点数据集

Tab. 2　Pig counting dataset

数据集	图像数量/幅	猪总数/头
训练集	1 800	19 205
验证集	600	6 481
测试集	600	6 316
合计	3 000	32 002

1.3　研究方法

本文提出了一种改进 YOLO v5n 猪只计数方法。首先，构建多场景的猪只盘点数据集（包括不同猪舍尺寸、不同群养猪只数量、不同遮挡程度、不同拍摄角度、不同光线强度等）。该方法在处理猪只遮挡重叠的问题，通过利用通道注意力机制在筛选特征时，能够提高感兴趣的区域信息权重，弱化与当前任务无关的特征信息以提升模型效果[16]，同时增加多尺度物体检测层，获得更丰富全面的特征信息，增强模型在复杂场景下多尺度学习的能力；并对边界框的损失函数以及加权的非极大值抑制（Non-maximum suppression，NMS）进行改进，提高模型在遮挡场景下的检测精度和速度，降低漏检率。最后经过非极大值抑制（DIoU - NMS）处理输出最终预测结果，并计算得到最终的猪只数量。本文方法流程如图 2 所示。

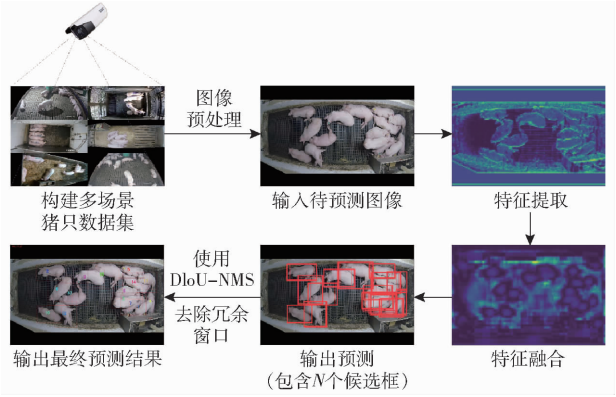


图 2　本文方法流程图

Fig. 2　Flow chart of proposed method

2　猪只计数网络实现

2.1　YOLO v5n 网络

YOLO v5[17]是基于 YOLO v4 算法基础上进行改进，同属于单阶段目标检测算法，其速度与精度较之前版本都得到了大幅提升。YOLO v5 支持 5 个不同规模的模型，根据网络的深度和宽度可分为 YOLO v5n、YOLO v5s、YOLO v5m、YOLO v5l、YOLO v5x 版本，其相对应模型的检测精度和模型参数量依次提高。为了从上述 5 个不同规模的模型中选择更适合猪只盘点的模型，使用本文的猪只盘点数据集重新训练上述 5 个模型，并分别用训练好的模型对测试数据集中的图像进行测试。其实验结果如表 3 所示，平衡模型的精度、参数量以及在猪只盘点实际应用

表 3　YOLO v5 不同规模模型的准确度比较

Tab. 3　Accuracy comparison of different scale models of YOLO v5

模型	准确率/%	召回率/%	AP/%	参数量	单幅图像平均识别时间/s
YOLO v5n	97.71	95.94	97.77	1.90×10^6	0.017 4
YOLO v5s	98.12	96.41	98.05	7.20×10^6	0.018 9
YOLO v5m	98.90	97.17	98.53	2.12×10^7	0.027 9
YOLO v5l	98.87	98.21	99.09	4.65×10^7	0.038 9
YOLO v5x	98.83	98.78	99.26	8.67×10^7	0.065 8

中实时 5 性要求 3 个指标，在精度相近的情况下，YOLO v5n 模型的参数量只有 1.9×10^6，其单幅图像平均识别时间仅为 0.017 4 s，相比较于参数量最多的 YOLO v5x，平均精确度（Average precision，AP）相差 1.49 个百分点，但单幅图像平均识别时间快了 48.4 ms。其次，与 YOLO v5s 模型相比，虽然识别速度相近，但两者的参数量相差近 4 倍。由于在猪场的盘点应用中，常采用边缘设备进行猪只盘点，此类设备对算法的处理速度要求较高。因此，本文最终选用参数量最少、识别速度最快的 YOLO v5n 作为本次实验的基础模型以便能够实时获取猪场猪只总数。

如图 3 所示，YOLO v5n 网络结构分为输入端、主干特征提取网络 Backbone、颈部 Neck 网络和输出端共 4 部分。

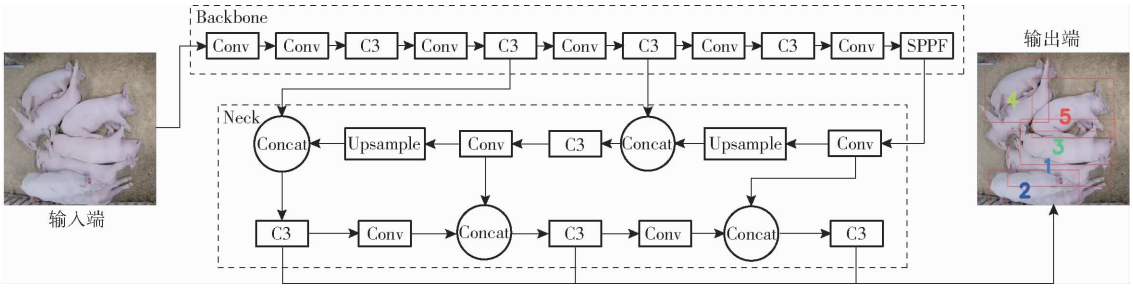


图 3 YOLO v5n 网络结构图

Fig. 3 YOLO v5n network structure diagram

YOLO v5n 的输入端对输入图像预处理，即将任意尺寸的输入图像缩放为网络的输入尺寸，并进行 Mosaic 数据增强丰富图像检测目标的背景，而且能够提高小目标的检测效果，并且在归一化计算时一次性处理多幅图像。其次，YOLO 系列检测算法中，针对不同的数据集往往需要设定特定长宽的锚框。在网络训练阶段，模型在初始锚框的基础上输出边界框，从而跟真实框对比，计算两者间的差距，再反向更新，迭代网络参数。在 YOLO v3 和 YOLO v4 检测算法中，在网络训练前单独采用 K－Means 聚类方法以获取不同数据集的最佳锚框值。但在 YOLO v5 检测算法中通过内嵌自适应锚框计算方法，每次训练前对数据集标注信息进行计算，得到该数据集标准信息针对默认锚框的最佳召回率，当最佳召回率大于或等于 98% 时，则不需要更新锚框；否则，重新计算符合该数据集的锚框，从而能够自适应地计算出不同训练集中最佳锚框值。

主干特征提取网络 Backbone 使用 CSPDarknet53 结构，且加入 Conv 卷积，CSP 及 SPPF[18] 结构。通过使用一个卷积核尺寸为 6×6，步长为 2 的卷积结构代替 Focus 结构，既能达到 2 倍下采样特征图的效果，又可避免多次采用切片操作，提高计算和推理速度[19]。其次，不同于 YOLO v4，YOLO v5n 设计了两种 CSP 结构，CSP1_X 结构应用于主干特征提取网络 Backbone 中，CSP2_X 结构应用于 Neck 网络中，其结构图如图 4 所示。SPPF 结构（图 5）与 SPP 结构一样，均采用 1×1、5×5、9×9 和 13×13 的最大池化方式，进行多尺度特征融合，但 SPPF 结构可以降低每秒浮点运算次数，运行的更快。

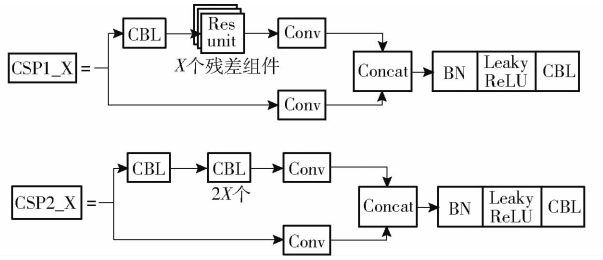


图 4 两种 CSP 结构

Fig. 4 Two CSP structures

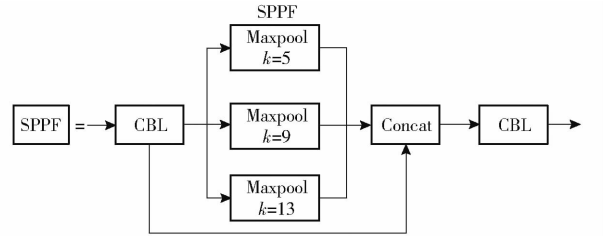


图 5 SPPF 结构

Fig. 5 SPPF structure

Neck 部分将特征金字塔的不同层级的图像组合在一起来完成多尺度混合，并将图像特征传递给预测层[20]。该部分仍采用 FPN 和 PANet 结合的双金字塔结构。特征金字塔网络（Feature pyramid networks，FPN）自顶向下，将高层的语义特征往下传递，得到预测的特征图[21]。路径聚合网络（Path aggregation network，PANet）是针对 FPN 只增强了语义信息，但并没有对定位信息进行传递的缺点[22]。通过在特征金字塔 FPN 的后面增加一个自底向上的路径，用底层精确的定位信息来增强整个特征层次结构，从而缩短底层特征与高层特征之间的信息路径。

输出端采用 GIoU Loss 作为边界框损失函数，其解决了当边界框与真实框没有交集时，交并比

(Intersection over union, IoU)等于 0,其损失函数导数也为 0,从而无法优化两框重叠的问题。在目标检测的最后处理过程中,YOLO v5n 使用加权的 NMS 方式筛选 N 个目标框。

2.2　SE－Net 注意力模块

注意力机制是通过参考人的视觉感知能力,即人在处理视觉信息初期会集中专注于当前情景下重点区域,而其他区域将相应降低,这为更高层级的视觉感知和逻辑推理以及更加复杂的计算机视觉处理任务提供更易于处理且更相关的信息[23]。

在深度卷积神经网络中,通过构建一系列的卷积层、非线性层和下采样层使得网络能够从全局感受野上提取图像特征来描述图像,但归根结底只是建模了图像的空间特征信息而没有建模通道之间的特征信息,整个特征图的各区域均被平等对待。然而,在猪只遮挡重叠、光线不足、皮毛颜色相似等实际猪场生产场景下,原始 YOLO v5n 算法难以检测到被遮挡的猪只,或者容易漏检一些边缘区域的猪只。在目标检测领域中每一个目标都被周围背景所包围,目标物体周围的环境有着丰富的上下文信息,这可以作为网络判断目标类别和定位的重要辅助信息[24]。通常在实际猪场生产图中,通过对目标区域的特征进行加权能使其更有效地定位到待检测猪只,提高网络性能[25]。SE－Net 通道注意力模块基于特征图上通道间的相关关系,网络通过损失函数去学习权重,强化有效特征,弱化低效或无效特征,提高了特征图的表现能力,从而提高了模型识别精度,能够在光线不足、密集遮挡的情况下有效检测出目标物体[26－27]。

针对 YOLO v5n 目标检测算法存在边界框定位不够准确导致难以区分重叠物体、鲁棒性差等问题,本文在主干特征提取网络中加入 SE－Net(Squeeze-and-Excitation networks)通道注意力模块来改进算法,在特征图的通道间建立特征映射,使网络自动学习全局特征信息并突出重要的特征信息,同时弱化对当前任务无关的特征信息,使模型更专注于训练遮挡对象的目的。

SE－Net 结构单元图如图 6 所示,SE－Net 模块具体含括 3 个操作:压缩(Squeeze)、激励(Excitation)和缩放(Scale)。首先对原始输入进行全局平均池化操作(即压缩过程)将原始特征层维度 $H\times W\times C$ 压缩为 $1\times1\times C$,使得感受野更大,得到每个通道的全局特征。然后用两个全连接层(第一个全连接层用于降低维度,另一个还原为原始的维度)融合各特征通道的特征图信息。再使用 Sigmoid 函数对权值进行归一化。最后是缩放操作,将激励操作后输出的权重映射为一组特征通道的权重,然后与原始特征图的特征相乘加权,实现对原始特征在通道维度上的特征重标定。网络通过加强训练 SE－Net 模块学习后的特征,提高了猪只遮挡的通道权重,指导模型更侧重学习猪只间遮挡的特征信息,进而提升模型在猪只遮挡场景下的检测性能[28]。

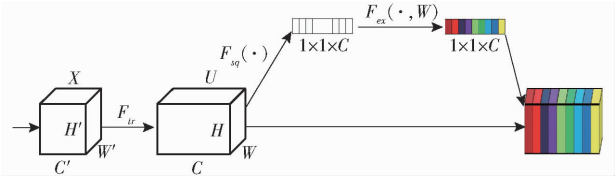


图 6　SE－Net 结构单元图

Fig. 6　SE－Net block

如图 7 所示,本文将 SE－Net 通道注意力模块分别加入到原始 YOLO v5n 主干特征提取网络 Backbone 的 C3 卷积模块后(图 7a)和颈部网络 Neck 的 C3 卷积模块后(图 7b)。两种添加方法的检测结果对比如图 8 所示:图 8a 为原始 YOLO v5n 检测效果图,漏检的情况比较严重;图 8b 为在主干特征提取网络中加入 SE－Net 模块,图 8c 为在颈部网络中加入 SE－Net 模块,加入 SE－Net 注意力模块后漏检猪只数量相对较小,但在主干特征提取网络中引入相比于在颈部网络中能够更好地检测到遮挡猪只。因此本文采用在 YOLO v5n 主干特征提取网络引入 SE－Net 通道注意力模块。

2.3　多尺度特征检测

原始 YOLO v5n 网络使用 3 种类型的输出特征图来检测不同尺寸的物体,当输入图像尺寸为 1 280 像素×1 280 像素时,其对应的 3 种检测尺度分别为 40 像素×40 像素、80 像素×80 像素、160 像素×160 像素,用来检测大、中、小目标,其感受野大小依次递减。在现有的 3 个尺度特征图上进行猪只检测,对于某些尺度超大的猪只,感受野最大的特征图可能仍然提取不到有效的高级语义信息,相反,对于某些尺度超小的猪只,感受野最小的特征图也会丢失其轮廓和位置信息[29]。在实际猪场应用场景中,猪只姿态各式各样,猪只间相互拥挤遮挡,而且,由于摄像头安装位置不同,其拍摄的猪只体积、体形也有较大的差异。例如,受限于猪场的高度,范金树实验猪场的摄像头拍摄出猪只形状既存在超大尺度的猪只,也存在超小尺度的猪只,所以映射到图像中的猪只也有多种不同尺度。为了应对复杂密集遮挡的实际生产猪场场景,得到更加丰富全面的特征信息,本文在原有 3 个不同尺度的检测层的基础上,增加一个超小尺度的检测层[30]。选择此

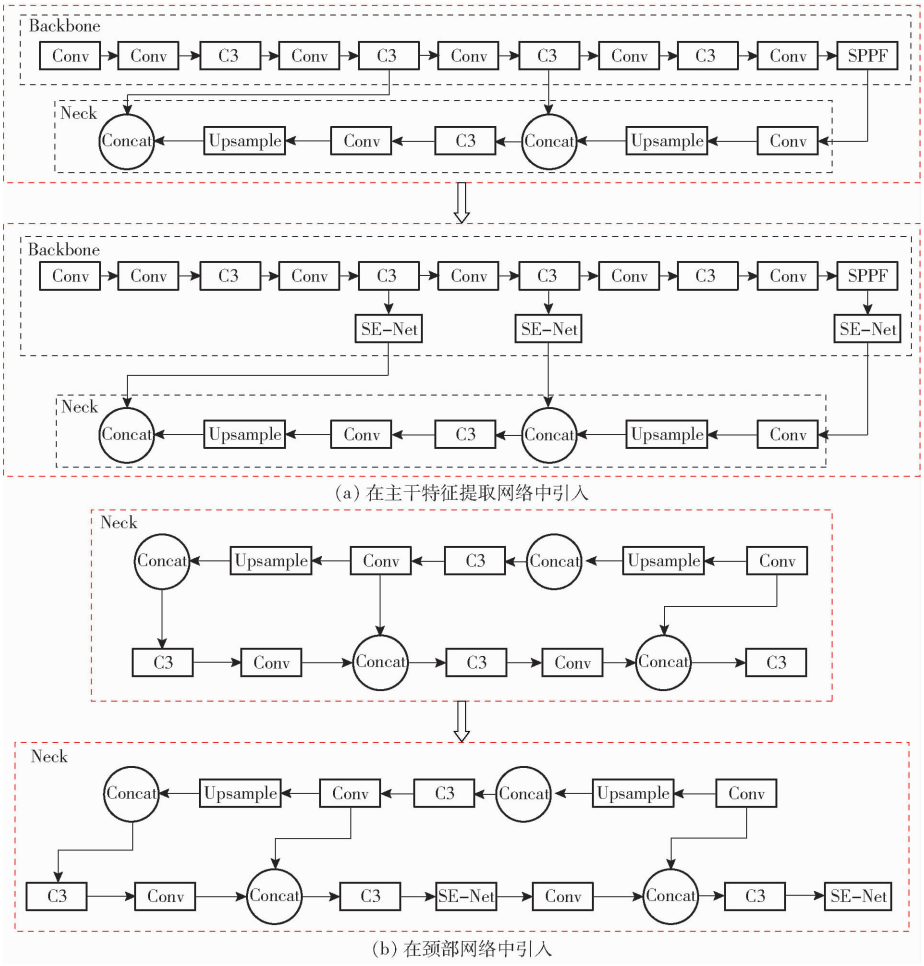


图 7 引入 SE－Net 通道注意力模块

Fig. 7 Introducing SE－Net channel attention module

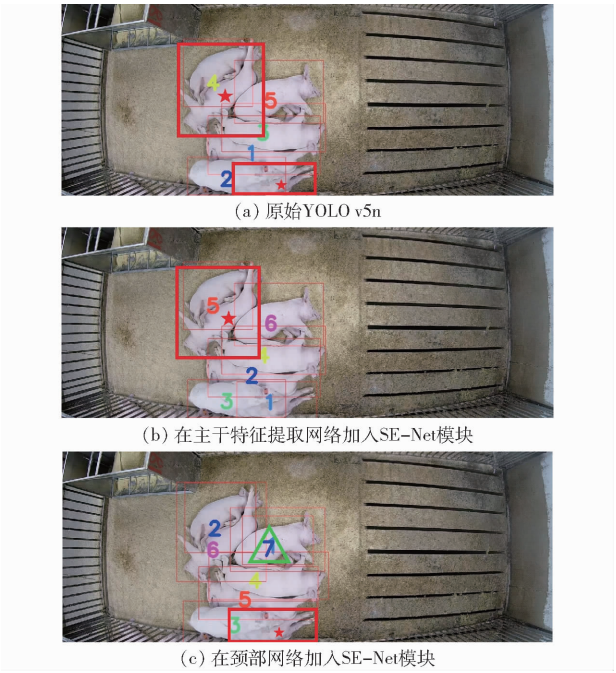


图 8 增加 SE－Net 模块前后检测效果对比

Fig. 8 Comparison of results before and after adding SE－Net module

尺度的检测层对于超小的猪只来说，能够保留下有效语义信息的可能性更大；另外，对尺寸超大的猪只来说，能够提取到更高层次的语义信息，同时也包含了更多的轮廓信息和位置信息，将此类特征图融合之后会使得融合之后的特征信息更加丰富全面，从而有利于提高猪只定位的准确率[29]。

本文在原始 YOLO v5n 的主干网络中加入 SE－Net 模块的基础上，在 32 倍下采样后继续增加一次下采样得到 20×20 的特征图，同时颈部 Neck 网络相应增加一次上采样，与主干网络的第 8 层拼接（原始颈部网络只经过两次上采样，分别与主干网络的第 6 层和第 4 层拼接）。改进后的 YOLO v5n 颈部网络经过 3 次上采样后再进行下采样操作。即主干网络在 32 倍下采样后继续一次 2 倍下采样得到 20×20 的特征尺寸图。同时，在颈部网络增加一次 2 倍上采样，然后经过 8、16、32 倍下采样分别得到 160×160、80×80、40×40 尺度的检测层特征图。再进行 64 倍的下采样得到 20×20 尺度的检测层特征图。

经过改进后的 YOLO v5n 网络（图 9）存在 4 个检测层，加深了网络的深度，进一步提高模型的表达能力，使得模型能够提取到更高层次更丰富的语义信息，帮助模型在复杂猪舍场景下提取多尺

度特征信息，将复杂的目标区分开来，提升模型的检测性能[31]。在猪只遮挡密集场景下 YOLO v5n 网络改进前后的检测效果如图 10 所示，左边为原始 YOLO v5n 检测结果，右边为增加一个检测层后的检测结果。由图 10 可以看出，改进后的模型在猪只挤靠相互遮挡的场景下，仍能更好地检测识别出被大面积遮挡的猪只，且也可检测出图 10d 中边缘的猪只，而原始 YOLO v5n 模型并没有检测到被大面积遮挡的猪只。由此可见添加一个检测层后的模型在复杂遮挡场景下的检测效果更佳。

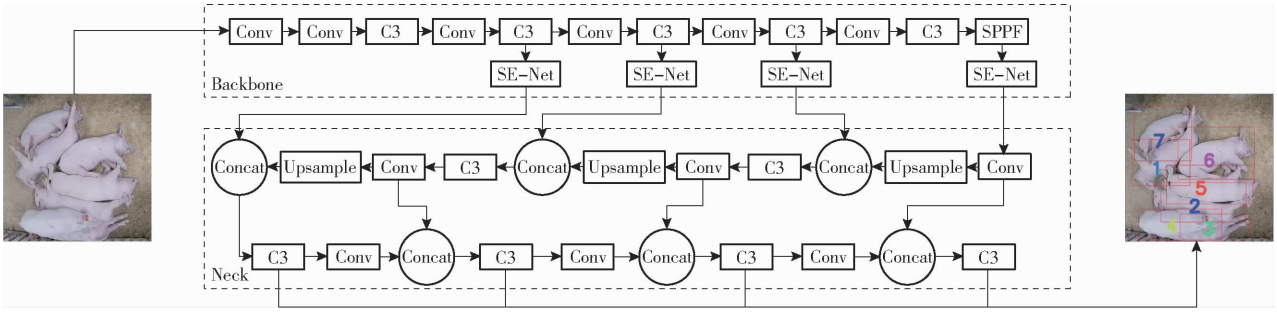
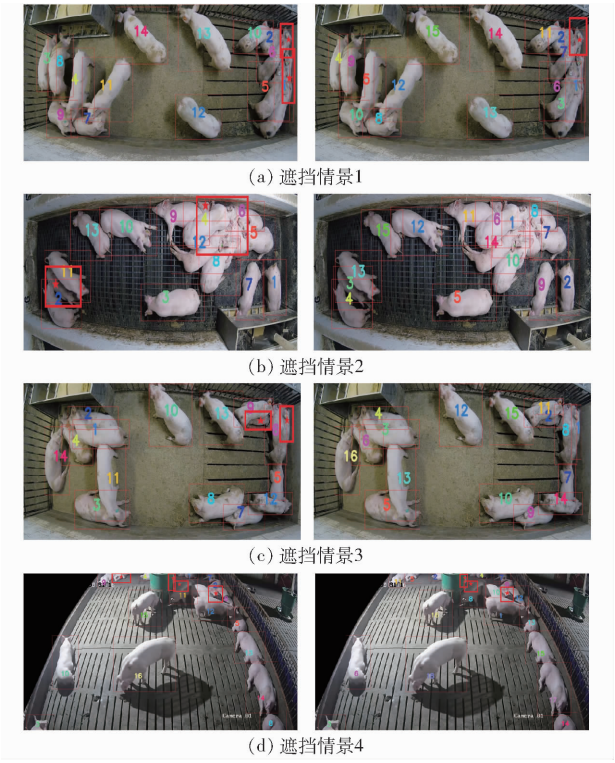


图 9　添加一个检测层后的 YOLO v5n 网络结构

Fig. 9　YOLO v5n network structure after adding a detection layer



(a) 遮挡情景1

(b) 遮挡情景2

(c) 遮挡情景3

(d) 遮挡情景4

图 10　增加检测层前后效果对比

Fig. 10　Comparison of effects before and after adding a detection layer

2.4　改进边界框损失函数

根据 2.1 节，YOLO v5n 的边界框（Bounding box）采用 GIoU Loss 作为损失函数，在 IoU 的基础上，解决了边界框不重合的问题，但当边界框与真实框有包含情况时（如图 11 所示，绿色和红色分别表示目标框和预测框），GIoU 退变为 IoU，导致难以分辨两框的相对位置关系，此时 GIoU 的优势消失[32]。因此，GIoU 不能准确反映两个边界框之间的重叠关系，也不能给出一个边界框被另一个边界框包围时的优化方向[33]。文献[34]提出的 CIoU 是在 DIoU

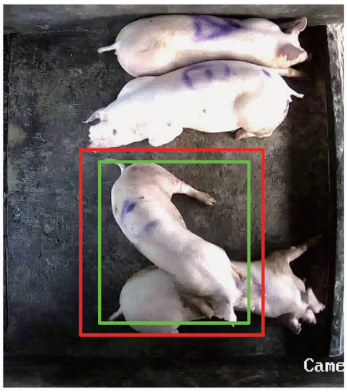


图 11　边界框包含检测目标的示例

Fig. 11　Schematic of a bounding box containing the detection target

的基础上，不仅考虑边界框的重叠区域和中心点的距离，还权衡了边界框的纵横比，弥补了 GIoU 损失函数的不足。因此，本文采用更优的 CIoU Loss 测量方法，包括重合面积、中心距离和宽高比 3 个测量维度，具有更快的收敛速度和更佳的性能。CIoU Loss 函数定义为

$$L_{CIoU} = 1 - I_{oU} + \frac{\rho^2(b, b^{(gt)})}{c^2} + \alpha\nu \tag{1}$$

其中

$$\alpha = \frac{\nu}{1 - I_{oU} + \nu} \tag{2}$$

$$\nu = \frac{4}{\pi^2}\left(\arctan\frac{w^{(gt)}}{h^{(gt)}} - \arctan\frac{w}{h}\right)^2 \tag{3}$$

式中　b——边界框中心点坐标

$b^{(gt)}$——真实框中心点坐标

ρ——两个中心点间的欧氏距离

I_{oU}——交并比

c——边界框和真实框最小外接矩形的对角线距离

w、h——边界框的宽、高

$w^{(gt)}$、$h^{(gt)}$——真实框的宽、高

α——权重系数

ν——宽高比度量函数

2.5 DIoU - NMS 代替加权 NMS

在经典的 NMS 中，对同一类别中所有边界框的得分排序，选择分数最高的边界框，然后计算出其与剩余的边界框的 IoU 值，并将大于阈值的边界框全部过滤掉[35]。但是在实际猪场中，猪群的重叠相互遮挡，当几头猪相互靠近时，IoU 值变大，若采用加权 NMS，那么其筛选结果可能只留下一个边界框，从而可能致使多头猪只漏检。采用 DIoU - NMS 算法[36]代替原始的加权 NMS，同时考虑了重叠区域和两个边界框的中心点距离，即如果出现两框间的 IoU 值较大，且中心距离也较大时，DIoU - NMS 将此类情况视为是两个目标的边界框，不会被筛选，从而有效提高检测精度，降低漏检率[32]，DIoU - NMS 的更新公式为

$$s_i = \begin{cases} s_i & (I_{oU} - R_{DIoU}(M, B_i) < \varepsilon) \\ 0 & (I_{oU} - R_{DIoU}(M, B_i) \geqslant \varepsilon) \end{cases} \tag{4}$$

其中

$$R_{DIoU} = \frac{\rho^2(b, b^{(gt)})}{c^2} \tag{5}$$

式中 s_i——分类得分　　ε——NMS 的阈值

M——最高分类得分的检测框

B_i——其余初始检测框

如图 12 所示，当两头猪只互相粘连时，对应的两个边界框重叠区域较大，但它们的中心点距离较远，若采用 DIoU - NMS，因为其不仅考虑到两框的重合部分，还综合权衡了两框中心点的欧氏距离，将两个中心点距离较远的框可以看成是不同猪只目标产生的，所以，在筛选边界框时保留这类边界框，不会将其删除，从而降低了猪只漏检率。因此，本文算法选用 DIoU - NMS 替换加权的 NMS 是切实有效的。

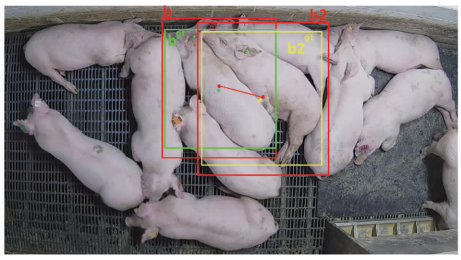


图 12 DIoU - NMS 示意图

Fig. 12 Schematic of DIoU - NMS

3 实验结果与分析

3.1 实验设置

本次实验用 Pytorch 框架搭建，操作平台为 64 位 Windows 10，GPU 模式（GPU 型号为 Nvidia RTX3070），显存 8GB。本文模型训练的超参数设置如表 4 所示。

表 4 训练参数设置

Tab. 4 Parameters setting for training procedure

参数	数值	含意
batch - size	16	每批处理的图像数量
epoch	200	训练迭代次数
img - size	1 280 × 1 280	输入图像尺寸
lr	0.001	学习率
adam	N/Y	是否使用 adam 优化器
momentum	0.843	动量
weight_decay	0.000 36	权重衰减项

3.2 评估指标

采用准确率、召回率、平均精度（AP）作为衡量目标检测算法性能的指标。

针对深度学习的计数研究常使用平均绝对误差（Mean absolute error，MAE）和均方根误差（Root mean square error，RMSE）。一般情况下，MAE 越小，预测的准确率就越高，RMSE 越小，鲁棒性也越好。本文选取上述两个指标作为猪只盘点模型的评估指标。

另外，加入漏检率作为模型性能偏差和变化的辅助指标。

3.3 消融实验

为了更直观地评估改进技术对模型性能的影响，本文进行了消融实验。如表 5 所示，以未做任何改进的原始 YOLO v5n 为基准，+ 表示模块混合改进，分别训练划分好的 6 组实验。如表 5 所示，在本文的测试集中原始 YOLO v5n 的 MAE、RMSE、AP 和漏检率分别为 0.682、1.21 和 97.71% 和 4.06%。以此为基线，每一处改进后各评估指标都有所提升。由表 5 可知，在主干特征提取网络中引入 SE - Net 模块，其 MAE、RMSE 以及漏检率均比添加在颈部网络中低，分别低 0.095、0.38 以及 1.34 个百分点，AP 也提高 0.67 个百分点。无论是原始 YOLO v5n，还是增加一个检测层后，在主干特征提取网络加入注意力机制模块比在颈部网络加入注意力机制模块，MAE、RMSE、AP 和漏检率均比未增加相应模块有所提升。而增加检测层和在主干特征提取网络中引入 SE - Net 模块，MAE、RMSE 以及漏检率也比基线分别降低 0.509、0.708 以及 3.02 个百分点，AP 达到 99.39%，较基线提高 1.62 个百分点。其次，结合表 3 可知，改进后的模型的参数量是 YOLO v5x 模型的 1/46，而且两者的 AP 很相近。综合而言，增加一个检测层与在主干特征提取网络加入注意力机

表 5　不同模型的消融实验

Tab. 5　Ablation experiment of different modules

模型	MAE	RMSE	准确率/%	召回率/%	AP/%	漏检率/%	参数量
YOLO v5n	0.682	1.210	97.71	95.94	97.77	4.06	1.900×10^6
+ SE－Net(Backbone)	0.488	1.020	98.98	97.85	98.91	2.15	1.940×10^6
+ SE－Net(Neck)	0.583	1.400	98.84	96.51	98.24	3.49	1.910×10^6
+ Adding Detect Layer	0.327	0.742	99.06	98.69	99.12	1.31	3.087×10^6
+ Adding Detect Layer + SE－Net(Backbone)	0.173	0.502	99.14	98.96	99.39	1.04	3.115×10^6
+ Adding Detect Layer + SE－Net(Neck)	0.340	0.753	98.92	98.43	98.95	1.57	3.102×10^6

制模块的组合改进效果最佳，虽然模型参数量相对比较大，但其在复杂遮挡场景下的检测效果比其他组合更好。

3.4　不同盘点算法对比

为更好地分析验证本文算法的性能，将本文改进后 YOLO v5n 与 YOLO v4[37]、YOLO X[38]、Faster R－CNN[39]、专门修改用于猪计数的 CCNN[9]、基于多尺度感知的猪计数网络 PCN[10]、基于关键点检测和 STRF 的猪计数网络[12]和中心集群计数网络 CClusnet[40]的猪只计数网络进行比较。其中，YOLO v4、YOLO X 和 Faster R－CNN 网络均采用本文的猪只盘点数据集进行训练和测试；而基于多尺度感知的猪计数网络 PCN、专门修改用于猪计数的 CCNN、中心集群计数网络 CClusnet 以及基于关键点检测和 STRF 的猪计数网络则引用文献中的相关实验结果。实验结果如表 6 所示。

表 6　不同猪只盘点算法性能比较

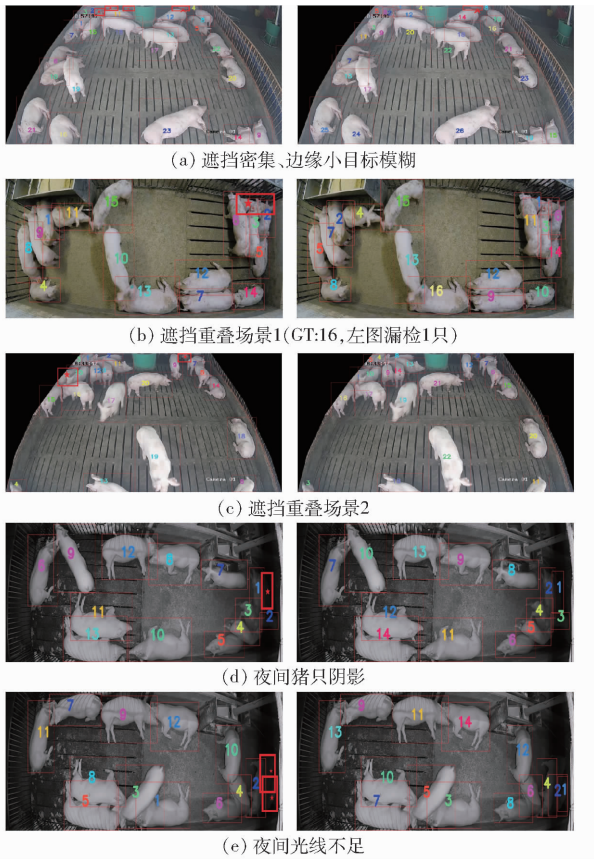
Tab. 6　Performance comparison of different pig counting algorithms

网络	MAE	RMSE	单幅图像平均识别时间/s
CCNN[9]	1.67	2.13	0.042
PCN[10]	1.74	2.28	0.108
Keypoint Tracking + STRF[12]	3.32		
YOLO v4[37]	0.193	0.574	0.045
YOLO X[38]	0.242	0.647	0.066
Faster R－CNN[39]	0.88	1.65	0.071
CClusnet[40]	0.43		
改进后的 YOLO v5n(本文)	0.173	0.502	0.056

由表 6 可知，CCNN、CClusnet、YOLO v4、YOLO X 和 Faster R－CNN 的 MAE 分别为 1.67、0.43、0.193、0.242 和 0.88，说明这 5 种计数方法在一定程度上能够处理遮挡问题。虽然 CCNN 计数方法的单幅图像平均识别时间最低(0.042 s)，本文改进后的 YOLO v5n 网络模型与其只差 0.014 s，但本文的计数网络准确性和鲁棒性均优于其他计数网络。这是因为本文的计数网络在处理不同场景时，如光照变化、遮挡和重叠，通过增加注意力机制和多尺度特征检测，减少了猪只漏检多检情况，且不需要控制特定的环境，速度和精度均能达到实际猪场的猪只计数要求。

3.5　实验结果分析

猪只遮挡是造成猪只盘点精度低的主要原因之一。当猪只严重拥挤相互遮挡时，猪只目标轮廓信息缺失。图 13 为 YOLO v5n 改进前后的检测结果，左侧为原始 YOLO v5n 检测图，右侧为改进后的 YOLO v5n 检测效果图。图 13a 可知，改进后的 YOLO v5n 对于边缘小目标的检测效果有一定的提升，图 13a 中改进后的模型能够检测出原始模型漏检的目标；图 13b、13c 为不同场景下遮挡重叠的图像。通过观察可知，原始模型对于猪只相互遮挡密集或者边缘小目标模糊的图像漏检相对严重，而



(a) 遮挡密集、边缘小目标模糊

(b) 遮挡重叠场景1(GT:16，左图漏检1只)

(c) 遮挡重叠场景2

(d) 夜间猪只阴影

(e) 夜间光线不足

图 13　检测结果对比

Fig. 13　Comparison of test results

改进后的模型依旧可以准确检测到上述复杂情景下的猪只目标。

另外,光照也是影响猪只盘点精度的原因之一。在实际生产应用中,猪场可能会突然遇到短暂光照骤变等噪声致使出现猪只阴影[41],或者夜间光线不足导致猪只目标轮廓特别不清晰。尽管 YOLO v5n 算法能在一定程度上尽量减少光照影响,但是猪只阴影、目标轮廓不清晰还是会造成算法的误检或者漏检。如图 13d 所示,右侧围栏处的两头猪只间形成了阴影,加之光线暗等原因导致原始 YOLO v5n 算法将中间的猪只阴影识别为一头猪只,而改进后的算法能够准确检测出阴影处的所有猪只。如图 13e 所示,图中绿框的猪只因光线不足,且被其他猪只夹在中间,猪只目标轮廓不清晰致使原始 YOLO v5n 算法漏检该猪只,但改进后的模型能够检测出原始模型因阴影漏检的猪只。

综上所述,改进的 YOLO v5n 算法对复杂场景中的猪只图像盘点鲁棒性更强,这表明改进后的网络在不同遮挡重叠目标场景下泛化能力有所提升。但目前本文算法在应对特别严重遮挡目标导致目标轮廓的特征信息不明显的情况时会存在一定程度的漏检。

4 结束语

为了改进现有的猪只盘点算法在遮挡、体型变化或者光线较暗等场景下的不足,提出并实现了一种改进的 YOLO v5n 猪只盘点算法。该算法针对复杂场景下的猪只盘点问题提出了在主干特征提取网络中加入 SE - Net 注意力模块,从而提升了模型对于复杂密集遮挡、光照不足等场景下的鲁棒性;增加了一个尺度的检测层,加深网络的深度,进一步提取多尺度特征信息将复杂的目标区分开来,提升模型的检测性能;改进边界框损失函数和 NMS,以此提高模型在复杂遮挡场景下的泛化能力。消融实验结果表明,改进的 YOLO v5n 模型在本文测试集中的检测精度较原 YOLO v5n 模型得到改善,其 MAE、RMSE 以及漏检率分别比原 YOLO v5n 降低 0.509、0.708 以及 3.02 个百分点,平均精度提高 1.62 个百分点,其 AP 为 99.39%。对比实验结果表明,与 Faster R - CNN、CCNN 和 PCN 猪只盘点算法相比,本文算法的 MAE 和 RMSE 分别为 0.173 和 0.502,MAE 比其他 3 种算法分别减少 0.707、1.497 和 1.567,RMSE 比其他 3 种算法分别降低 1.148、1.628 和 1.778。在遮挡拥挤、光照不良等复杂场景下,本文提出的 YOLO v5n 算法减少了猪只漏检的情况,能得到理想的计数结果,具备较好的准确性和鲁棒性;模型的单幅图像识别时间仅为 0.056 s,比大多数猪只盘点算法快,能够满足实际应用场景的实时性要求,为后续应用到边缘设备(如猪只盘点仪等)提供了算法支撑。

参 考 文 献

[1] KAMILARIS A, PRENAFETA-BOLDU F X. Deep learning in agriculture: a survey[J]. Computers and Electronics in Agriculture, 2018, 147(1): 70 - 90.

[2] 张天昊,梁炎森,何志毅. 图像识别计数在储备生猪统计的应用[J]. 计算机应用与软件,2016,33(12):173 - 178.
ZHANG Tianhao, LIANG Yansen, HE Zhiyi. Applying image recognition and counting to reserved live pigs statistics[J]. Computer Applications and Software, 2016, 33(12): 173 - 178. (in Chinese)

[3] 梁炎森,张天昊,何志毅. 畜牧养殖场图像远程采集与目标计数系统[J]. 桂林电子科技大学学报,2017,37(6):437 - 441.
LIANG Yansen, ZHANG Tianhao, HE Zhiyi. A remote image acquisition and target counting system for livestock farm[J]. Journal of Guilin University of Electronic Technology, 2017, 37(6): 437 - 441. (in Chinese)

[4] 张宏鸣,付振宇,韩文霆,等. 基于改进 YOLO 的玉米幼苗株数获取方法[J]. 农业机械学报,2021,52(4):221 - 229.
ZHANG Hongming, FU Zhenyu, HAN Wenting, et al. Detection method of maize seedling number based on improved YOLO [J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(4): 221 - 229. (in Chinese)

[5] ONORO-RUBIO D, LOPEZ-SASTRE R J. Towards perspective-free object counting with deep learning[C] // European Conference on Computer Vision, 2016: 615 - 629.

[6] XU M, GE Z, JIANG X, et al. Depth information guided crowd counting for complex crowd scenes[J]. Pattern Recognition Letters, 2019, 125: 563 - 569.

[7] 赵睿,刘辉,刘沛霖,等. 基于改进 YOLOv5s 的安全帽检测算法[J/OL]. 北京航空航天大学学报, https://doi.org/10.13700/j.bh.1001 - 5965.2021.0595.

[8] CHEN S W, SHIVAKUMAR S S, DCUNHA S, et al. Counting apples and oranges with deep learning: a data-driven approach [J]. IEEE Robotics and Automation Letters, 2017, 2(2): 781 - 788.

[9] TIAN M, GUO H, CHEN H, et al. Automated pig counting using deep learning[J]. Computers and Electronics in Agriculture, 2019, 163: 104840.

[10] 高云,李静,余梅,等. 基于多尺度感知的高密度猪只计数网络研究[J]. 农业机械学报,2021,52(9):172 - 178.
GAO Yun, LI Jing, YU Mei, et al. High-density pig counting net based on multi-scale aware[J]. Transactions of the Chinese

Society for Agricultural Machinery,2021,52(9):172 - 178. (in Chinese)
[11] ZHANG Y, ZHOU D, CHEN S, et al. Single-image crowd counting via multi-column convolutional neural network[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2016:589 - 597.
[12] CHEN G, SHEN S, WEN L, et al. Efficient pig counting in crowds with keypoints tracking and spatial-aware temporal response filtering[C]//2020 IEEE International Conference on Robotics and Automation. IEEE,2020:10052 - 10058.
[13] 胡云鸽,苍岩,乔玉龙. 基于改进实例分割算法的智能猪只盘点系统设计[J]. 农业工程学报,2020,36(19):177 - 183.
HU Yunge, CANG Yan, QIAO Yulong. Design of intelligent pig counting system based on improved instance segmentation algorithm[J]. Transactions of the CSAE,2020,36(19):177 - 183. (in Chinese)
[14] PSOTA E T, MITTEK M, PEREZ L C, et al. Multi - pig part detection and association with a fully-convolutional network[J]. Sensors,2019,19(4):852 - 876.
[15] BERGAMINI L, PINI S, SIMONI A, et al. Extracting accurate long-term behavior changes from a large pig dataset[C]//16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 2021: 524 - 533.
[16] 胡志伟,杨华,娄甜田. 采用双重注意力特征金字塔网络检测群养生猪[J]. 农业工程学报,2021,37(5):166 - 174.
HU Zhiwei, YANG Hua, LOU Tiantian. Instance detection of group breeding pigs using a pyramid network with dual attention feature[J]. Transactions of the CSAE,2021,37(5):166 - 174. (in Chinese)
[17] JOCHER G. Ultralytics/yolov5: v6.1-tensorrt, tensorflow edge tpu and openvino export and inference[EB/OL]. 2022 - 02 - 22[2022 - 11 - 03]. https://github.com/ultralytics/yolov5.
[18] 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.
[19] LI J, LIU C, LU X, et al. CME - YOLOv5: an efficient object detection network for densely spaced fish and small targets[J]. Water, 2022, 14(15): 2412.
[20] YAO J, QI J, ZHANG J, et al. A real-time detection algorithm for kiwifruit defects based on YOLOv5[J]. Electronics, 2021, 10(14): 1711.
[21] 江磊, 崔艳荣. 基于 YOLOv5 的小目标检测[J]. 电脑知识与技术,2021,17(26):131 - 133.
JIANG Lei, CUI Yanrong. Small target detection based on YOLOv5[J]. Computer Knowledge and Technology,2021,17(26): 131 - 133. (in Chinese)
[22] 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.
[23] 王文冠, 沈建冰, 贾云得. 视觉注意力检测综述[J]. 软件学报, 2019, 30(2): 416 - 439.
WANG Wenguan, SHEN Jianbing, JIA Yunde. Review of visual attention detection[J]. Journal of Software,2019,30(2): 416 - 439. (in Chinese)
[24] ZOU Z, SHI Z, GUO Y, et al. Object detection in 20 years: a survey[J]. arXiv preprint arXiv:1905.05055, 2019.
[25] YIN D, TANG W, CHEN P, et al. Pig target detection from image based on improved YOLO v3[C]// International Conference on Artificial Intelligence and Security. Springer, Cham, 2021: 94 - 104.
[26] 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.
[27] 刘天真,滕桂法,苑迎春,等. 基于改进 YOLO v3 的自然场景下冬枣果实识别方法[J]. 农业机械学报,2021, 52(5):17 - 25.
LIU Tianzhen, TENG Guifa, YUAN Yingchun, et al. Winter jujube fruit recognition method based on improved YOLO v3 under natural scene[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(5):17 - 25. (in Chinese)
[28] ZHANG S, YANG J, SCHIELE B. Occluded pedestrian detection through guided attention in CNNS[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6995 - 7003.
[29] 赵婉月. 基于 YOLOv5 的目标检测算法研究[D]. 西安:西安电子科技大学,2021.
ZHAO Wanyue. Study of object detection algorithm based on YOLOv5[D]. Xi'an:Xidian University,2021. (in Chinese)
[30] LI S, LI Y, LI Y, et al. YOLO - FIRI: improved YOLOv5 for infrared image object detection[J]. IEEE Access, 2021, 9: 141861 - 141875.
[31] 曹帅,张晓伟,马健伟. 基于跨尺度特征聚合网络的多尺度行人检测[J]. 北京航空航天大学学报,2020,46(9):1786 - 1796.
CAO Shuai, ZHANG Xiaowei, MA Jianwei. Trans-scale feature aggregation network for multiscale pedestrian detection[J]. Journal of Beijing University of Aeronautics and Astronautics,2020,46(9):1786 - 1796. (in Chinese)
[32] 于硕,李慧,桂方俊,等. 复杂场景下基于 YOLOv5 的口罩佩戴实时检测算法研究[J]. 计算机测量与控制,2021, 29(12):188 - 194.
YU Shuo, LI Hui, GUI Fangjun, et al. Research on real-time mask-wearing detection algorithm based on YOLOv5 in complex scenes[J]. Computer Measurement and Control,2021,29(12):188 - 194. (in Chinese)
[33] GAO J, CHEN Y, WEI Y, et al. Detection of specific building in remote sensing images using a novel YOLO - S - CIOU model. Case: gas station identification[J]. Sensors, 2021, 21(4): 1375.

[34] ZHENG Z, WANG P, REN D, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. IEEE Transactions on Cybernetics, 2021,52(8):8574 - 8586.
[35] 张长伦,张翠文,王恒友,等. 基于注意力机制的 NMS 在目标检测中的研究[J]. 电子测量技术,2021,44(19):82 - 88.
ZHANG Changlun, ZHANG Cuiwen, WANG Hengyou, et al. Research on non-maximum suppression based on attention mechanism in object detection[J]. Electronic Measurement Technology,2021,44(19):82 - 88. (in Chinese)
[36] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993 - 13000.
[37] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLO v4: optimal speed and accuracy of object detection[J]. arXiv preprint arXiv, 2020: 2004.10934.
[38] GE Z, LIU S, WANG F, et al. YOLO X: exceeding YOLO series in 2021[J]. arXiv preprint arXiv:2107.08430, 2021.
[39] REN S, HE K, GIRSHICK R, et al. Faster R - CNN: towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing Systems, 2015, 28:91 - 99.
[40] HUANG E, MAO A, GAN H, et al. Center clustering network improves piglet counting under occlusion[J]. Computers and Electronics in Agriculture,2021,189:106417.
[41] 杨秋妹,肖德琴,张根兴. 猪只饮水行为机器视觉自动识别[J]. 农业机械学报,2018,49(6):232 - 238.
YANG Qiumei, XIAO Deqin, ZHANG Genxing. Automatic pig drinking behavior recognition with machine vision[J]. Transactions of the Chinese Society for Agricultural Machinery,2018,49(6):232 - 238. (in Chinese)

(上接第 250 页)

[15] 李振波,李萌,吴宇峰,等. 基于优化 SSD 算法的冰鲜鲳鱼新鲜度评估方法研究[J]. 农业机械学报, 2021, 52(增刊): 472 - 481.
LI Zhenbo, LI Meng, WU Yufeng, et al. Evaluation method of iced pomfret freshness based on improved SSD[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(Supp.): 472 - 481. (in Chinese)
[16] 成怡,张宇,李宝全. 改进 CenterNet 的交通标志检测算法[J]. 信号处理,2022,38(3):511 - 518.
CHENG Yi, ZHANG Yu, LI Baoquan. Traffic sign detection algorithm based on improved CenterNet[J]. Journal of Signal Processing, 2022,38(3):511 - 518. (in Chinese)
[17] 杨蜀秦,刘江川,徐可可,等. 基于改进 CenterNet 的玉米雄蕊无人机遥感图像识别[J]. 农业机械学报, 2021, 52(9): 206 - 212.
YANG Shuqin, LIU Jiangchuan, XU Keke, et al. Improved CenterNet based maize tassel recognition for UAV remote sensing image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9):206 - 212. (in Chinese)
[18] 闫建伟,赵源,张乐伟,等. 改进 Faster - RCNN 自然环境下识别刺梨果实[J]. 农业工程学报, 2019,35(18):143 - 150.
YAN Jianwei, ZHAO Yuan, ZHANG Lewei, et al. Recognition of *Rosa roxbunghii* in natural environment based on improved Faster RCNN[J]. Transactions of the CSAE, 2019,35(18):143 - 150. (in Chinese)
[19] 何东健,刘建敏,熊虹婷,等. 基于改进 YOLO v3 模型的挤奶奶牛个体识别方法[J]. 农业机械学报, 2020, 51(4): 250 - 260.
HE Dongjian,LIU Jianmin,XIONG Hongting,et al. Individual identification of dairy cows based on improved YOLO v3[J]. Transactions of the Chinese Society for Agricultural Machinery,2020,51(4):250 - 260. (in Chinese)
[20] XIAO J, LIU G, WANG K, et al. Cow identification in free-stall barns based on an improved Mask R - CNN and an SVM[J]. Computers and Electronics in Agriculture, 2022, 194: 106738.
[21] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: optimal speed and accuracy of object detection[J]. arXiv preprint arXiv: 2004.10934, 2020.
[22] ZHANG Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330 - 1334.
[23] JIANG X F, TAO C K. Advanced multi-scale Retinex algorithm for color image enhancement[C]// International Symposium on Photoelectronic Detection and Imaging 2007: Related Technologies and Applications, SPIE, 2008: 325 - 332.
[24] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.
[25] LIU S, HUANG D, WANG Y. Receptive field block net for accurate and fast object detection[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 385 - 400.
[26] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception - v4, Inception - ResNet and the impact of residual connections on learning[C]// AAAI Conference on Artificial Intelligence,2016.
[27] NEUBECK A, VAN G L. Efficient non-maximum suppression[C]// 18th International Conference on Pattern Recognifion. IEEE, 2006:850 - 855.