

基于改进 YOLO v7 的笼养鸡/蛋自动识别与计数方法

赵春江^{1,2} 梁雪文^{1,2} 于合龙¹ 王海峰² 樊世杰³ 李 斌²

(1. 吉林农业大学信息技术学院, 长春 130118; 2. 北京市农林科学院智能装备技术研究中心, 北京 100097;
3. 北京市华都峪口禽业有限责任公司, 北京 101206)

摘要: 笼养模式下鸡/蛋自动识别与计数在低产能鸡判别及鸡舍智能化管理方面具有重要作用, 针对鸡舍内光线不均、鸡只与笼之间遮挡及鸡蛋粘连等因素导致自动计数困难的问题, 本研究以笼养鸡只与鸡蛋为研究对象, 基于 YOLO v7-tiny 提出一种轻量型网络 YOLO v7-tiny-DO 用于鸡只与鸡蛋识别, 并设计自动化分笼计数方法。首先, 采用 JRWT1412 型无畸变相机与巡检设备搭建自动化数据采集平台, 获取 2 146 幅笼养鸡只图像用于构建数据集。然后, 在 YOLO v7-tiny 网络基础上应用指数线性单元 (Exponential linear unit, ELU) 激活函数减少模型训练时间; 将高效层聚合网络 (Efficient layer aggregation network, ELAN) 中的常规卷积替换为深度卷积减少模型参数量, 并在其基础上添加深度过参数化组件 (深度卷积) 构建深度过参数化深度卷积层 (Depthwise over-parameterized depthwise convolutional layer, DO-DCov), 以提取目标深层特征; 同时在特征融合模块引入坐标注意力机制 (Coordinate attention mechanism, CoordAtt), 提升模型对目标空间位置信息的感知能力。试验结果表明, YOLO v7-tiny-DO 识别鸡只和鸡蛋的平均精确率 (Average precision, AP) 分别为 96.9% 与 99.3%, 与 YOLO v7-tiny 相比, 鸡只与鸡蛋的 AP 分别提高 3.2、1.4 个百分点; 改进后模型内存占用量为 5.6 MB, 比原模型减小 6.1 MB, 适合部署于算力相对有限的巡检机器人; YOLO v7-tiny-DO 在局部遮挡、运动模糊和鸡蛋粘连情况下均能实现高精度识别与定位, 在光线昏暗情况下识别结果优于其他模型, 具有较强的鲁棒性。最后, 将本文算法部署到 NVIDIA Jetson AGX Xavier 边缘计算设备, 在实际场景下选取 30 个鸡笼开展计数测试, 持续 3 d。结果表明, 3 个测试批次鸡只与鸡蛋的计数平均准确率均值分别为 96.7% 和 96.3%, 每笼平均绝对误差均值分别为 0.13 只鸡和 0.09 枚鸡蛋, 可为规模化养殖场智能化管理提供参考。

关键词: 笼养鸡/蛋; YOLO v7-tiny; 深度过参数化深度卷积层; 计数; 边缘计算

中图分类号: TP391.41 文献标识码: A 文章编号: 1000-1298(2023)07-0300-13

OSID:



Automatic Identification and Counting Method of Caged Hens and Eggs Based on Improved YOLO v7

ZHAO Chunjiang^{1,2} LIANG Xuewen^{1,2} YU Helong¹ WANG Haifeng² FAN Shijie³ LI Bin²

(1. College of Information and Technology, Jilin Agricultural University, Changchun 130118, China
2. Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China
3. Beijing Huadu Yukou Poultry Co., Ltd., Beijing 101206, China)

Abstract: In the cage mode, the elimination and death of laying hens will lead to changes in the number of hens and eggs production in the cage, so it is necessary to update the number of laying hens in the cage in a timely manner. Traditional machine vision methods recognized poultry by morphology or color, but their detection accuracy was low for complex scenarios such as uneven lighting in the cages, hens obscured by cages and the eggs adhesion. Therefore, based on deep learning and image processing, a lightweight network YOLO v7-tiny-DO was proposed for hens and eggs detection based on YOLO v7-tiny, and an automated counting method was designed. Firstly, the JRWT1412 distortion-free camera and the inspection equipment were used to build an automated data acquisition platform, and a total of 2 146

收稿日期: 2022-12-01 修回日期: 2023-02-10

基金项目: 北京市平谷区博士农场项目、国家科技创新 2030-“新一代人工智能”重大项目 (2021ZD0113804)、北京市农林科学院改革发展专项、北京市农林科学院科研创新平台建设项目 (PT2022-34) 和北京市博士后基金项目 (2022-ZZ-18)

作者简介: 赵春江 (1964—), 男, 研究员, 中国工程院院士, 主要从事智慧农业研究, E-mail: zhaocj@nercita.org.cn

通信作者: 李斌 (1983—), 男, 研究员, 主要从事畜禽智能装备研究, E-mail: lib@nercita.org.cn

images of caged hens and eggs were acquired as data sources. Then the exponential linear unit (ELU) was applied to the YOLO v7 - tiny network to reduce the training time of the model; the regular convolution was replaced in efficient layer aggregation network (ELAN) with depthwise convolution to reduce the number of model parameters, and on this basis, a depthwise over-parameterized depthwise convolutional layer (DO - DConv) was constructed by adding a depthwise over-parametric component (depthwise convolution) to extract the deep features of hens and eggs. At the same time, coordinate attention mechanism (CoordAtt) was embedded into the feature fusion module to improve the model's perception of the spatial location information of hens and eggs. The results showed that the average precision (AP) of YOLO v7 - tiny - DO was 96.9% and 99.3% for hens and eggs respectively, and compared with that of YOLO v7 - tiny, the AP of hens and eggs was increased by 3.2 percentage points and 1.4 percentage points, respectively. The model size of YOLO v7 - tiny - DO was 5.6 MB, which was 6.1 MB less than the original model, and it was suitable to be deployed in the inspection robot which lacked computing power. YOLO v7 - tiny - DO could achieve high-precision detection and localization under partial occlusion, motion blur and eggs adhesion, and outperformed other models in dim environment, with strong robustness. YOLO v7 - tiny - DO recognized that the F1 score of hens and eggs were 97.0% and 98.4% respectively. Compared with the mainstream object detection networks such as Faster R - CNN, SSD, YOLO v4 - tiny and YOLO v5n, the F1 score of hens were increased by 21.0 percentage points, 4.0 percentage points, 8.0 percentage points and 1.5 percentage points, respectively, and the F1 scores of eggs were increased by 31.4 percentage points, 25.4 percentage points, 6.4 percentage points and 4.4 percentage points, respectively. And frame rates were increased by 95.2 f/s, 34.8 f/s, 18.4 f/s and 8.4 f/s, respectively. Finally, the algorithm was deployed to the NVIDIA Jetson AGX Xavier edge computing device and 30 cages were selected for counting tests in a real-world scenario for 3 d. The results showed that the average precision of counting hens and eggs for the three test batches were 96.7% and 96.3%, respectively, and the mean absolute error were 0.13 hens and 0.09 eggs per cage, respectively, which can provide a reference for digital management of large-scale farms.

Key words: caged hens and eggs; YOLO v7 - tiny; depthwise over-parameterized depthwise convolutional layer; counting; edge computing

0 引言

作为重要的家禽种类之一,蛋鸡为人们提供了丰富的动物蛋白营养食源^[1]。近年来,随着消费需求日益增加,规模化养殖趋势加速,立体笼养成为主要饲养模式^[2-4]。死亡/低产蛋鸡的高效巡检与动态清零成为当前蛋鸡养殖产业的重要需求。一般来讲,笼内蛋鸡存活和产能状况通过笼内鸡/蛋数量精确盘点实现,而立体笼养模式下,养殖密度大、鸡只间遮挡和光线分布不均等因素使鸡只和鸡蛋准确识别与计数工作难度增加。

当前,人工观测方法清点鸡只与鸡蛋数目效率低下且易使蛋鸡产生应激反应^[5]。传统机器视觉方法主要应用特征信息提取和阈值分割技术开展鸡/蛋目标检测。毕敏娜等^[6]采用 SVM 算法检测散养病鸡,应用数码相机采集 RGB 图像,将鸡头鸡冠纹理特征和鸡眼瞳孔几何特征作为病鸡与健康鸡分类的特征向量,结果表明病鸡识别正确率为 92.5%。INDRA 等^[7]基于大津阈值法设计鸡蛋检测模型,以获取受精卵和不育卵的信息。MUVVA 等^[8]提出一种基于热度图和可见光图像的死鸡检

测方法,使用阈值分割技术提取热度图中活鸡像素和可见光图像中所有肉鸡(活鸡和死鸡)的像素,根据两幅图像像素差值,输出死鸡坐标信息,该方法在常规饲养密度和稀疏饲养密度下肉鸡(≤ 5 周龄)的死鸡识别正确率分别为 90.7% 与 95.0%。LI 等^[9]开发了基于图像分析和体质量计数的母鸡传感器计数系统,图像分析计数是将原始彩色图像转换为灰度图,然后根据最终生成的二值图像中白色像素数确定母鸡数量,计数准确率为 71.23%;体质量计数是指首先求得母鸡平均体质量,再根据传感器获得的笼内母鸡体质量计算母鸡数量,准确率达到 99.70%,但是这种方法每次仅能检测 8 只母鸡,效率低。上述基于各自研究场景下的检测研究虽然准确率较高,但无法解决笼养鸡舍内光线不均、鸡与笼之间遮挡及鸡蛋粘连等复杂场景下的鸡只与鸡蛋自动化识别与计数问题。

近年来,深度学习技术发展迅速,目标检测网络已在畜禽个体识别研究领域表现出优越性能^[10-12]。GEFFEN 等^[13]利用 Faster R - CNN 网络实现笼养模式下母鸡的检测和计数工作,准确率为 89.6%,平均绝对误差为每笼 2.5 只母鸡。LI 等^[14]应用改进

后网络 QueryPNet 检测和分割大鹅,大鹅检测与分割的准确率均为 96.3%。JIANG 等^[15]通过改进的 YOLO v7 网络实现实时监控麻鸭数量,平均精度均值(Mean average precision, mAP)为 97.57%。上述模型检测性能良好,但模型参数量较大,难以部署到 GPU 资源匮乏的边缘计算设备进行实时检测。面对轻量化设备的应用需求,模型小、检测速度快的轻量级目标检测网络优势较为突出^[16-18]。其中应用较为广泛的轻量级网络有 SSD 与 YOLO 微型系列网络。刘慧等^[19]应用改进的 SSD 网络实时检测田间行人,当交并比阈值为 0.4 时,模型 mAP 与检测速度分别为 97.46% 和 62.5 f/s。易诗等^[20]提出一种基于 YOLO v3-tiny 的适宜于嵌入式系统部署的野鸡识别方法,模型 mAP 和检测速度分别为 86.5% 与 62 f/s。ZHENG 等^[21]使用改进的 YOLO v5 模型快速检测麻鸭,并对比了基于不同标注范围(鸭头和鸭身)的模型检测性能,结果表明,基于鸭头标注的模型检测性能更好,准确率与检测速度分别为 99.29% 和 269.68 f/s。综上,轻量化网络的高实时性、少参数量和易部署等特点为实际场景下的应用研发提供了可能。

基于此,针对立体笼养模式下巡检过程中存在的光线不均、鸡只易被遮挡和鸡蛋粘连等问题,本研究提出基于 YOLO v7-tiny-DO 网络的笼养鸡/蛋自动化识别与分笼计数方法(Counting in different cages, CDC)。首先,在高效聚合层网络内构建深度过参数化深度卷积层提取目标深层特征,在特征融合模块引入坐标注意力机制提高模型捕获鸡只和鸡蛋空间位置信息的能力;然后,设计和发展自动化分笼计数方法;最后,将其部署到边缘计算设备进行实地测试验证,以期为规模化蛋鸡养殖智能化管理提供思路。

1 材料与方法

1.1 试验平台搭建

为实现自动化数据采集和减少人工参与以保证数据的客观性,设计并搭建了鸡只视频数据采集平台,示意图如图 1 所示。该平台由视频采集单元、边缘计算单元、数据展示单元与电力供应单元组成。其中视频采集单元选用低成本 JRWT1412 型无畸变摄像头,内置 RGB 传感器(分辨率为 640 像素 × 480 像素,帧率为 30 f/s);边缘计算单元采用 NVIDIA Jetson AGX Xavier 开发者套件,内置 512 个 NVIDIA CUDA Cores、64 个 Tensor Cores、8 块 Carmel ARM CPU、2 个 DL Accelerator 和 1 个 Vision Accelerator 深度学习加速器引擎,以支持视频数据

采集、模型部署与算法推理等功能的实现;数据展示单元采用 11.6 寸 IPS 高清触摸屏,用以展示摄像头捕获的视频图像;电力供应单元采用便携式移动电源支撑巡检机器人运转,输出电压为 48 V。

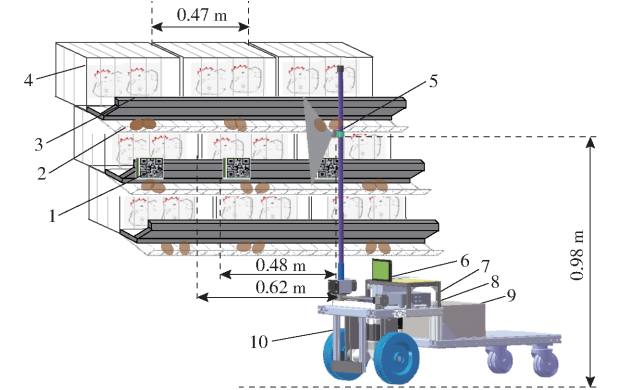


图 1 视频数据采集平台

Fig. 1 Collection platform of video data

1. 二维码 2. 蛋网 3. 食槽 4. 鸡笼 5. 视频采集单元 6. 数据展示单元 7. 边缘计算单元 8. USB 接口 9. 电力供应单元 10. 巡检机器人

1.2 数据获取与预处理

本研究所有数据均采集于北京市华都峪口禽业有限责任公司第 3 场阶梯式鸡舍。采集对象为中层 30 笼鸡,品种为京粉 1 号蛋鸡。采集时间为 2021 年 10 月 9—13 日与 2022 年 3 月 28 日—4 月 14 日 06:00—18:00,期间每隔 2 h 采集 1 批次视频,单个视频时长约 2.5 min,从中选取鸡头未被食槽完全遮挡的视频进行试验。

数据采集流程如下:首先,在食槽特定位置张贴二维码作为鸡笼编号;其次,在相机镜头距鸡笼水平距离 0.62 m 处运行巡检机器人,以速度 0.1 m/s 沿直线前进,行走过程中镜头距二维码与鸡笼的水平距离基本保持不变(误差小于 0.005 m);最后,获取的鸡只视频数据(分辨率为 640 像素 × 480 像素)通过 USB 接口传输至边缘计算单元,在数据展示单元显示并保存到本地磁盘。共拍摄视频 378 min,格式为 MP4,帧率为 30 f/s。采用 OpenCV 图像处理库每隔 10 帧截取一幅图像,从中筛选出 2 146 幅图像(分辨率为 640 像素 × 480 像素)作为数据源,格式为 JPG。由于笼养鸡舍内采用暖光灯照明,且鸡只灵活性较高,因此本文采集图像大多具有以下 1 种或多种特点:①鸡只被铁丝或食槽遮挡(图 2a 红色框)。②鸡只快速移动产生运动模糊(图 2b 蓝色框)。③多枚鸡蛋粘连在一起(图 2c 黄色框)。④两个暖光灯中间位置光线昏暗(图 2d 绿色框)。

由于笼养模式下养殖密度大,鸡笼边界不易区分,本研究通过检测鸡头识别鸡只个体,检测二维码

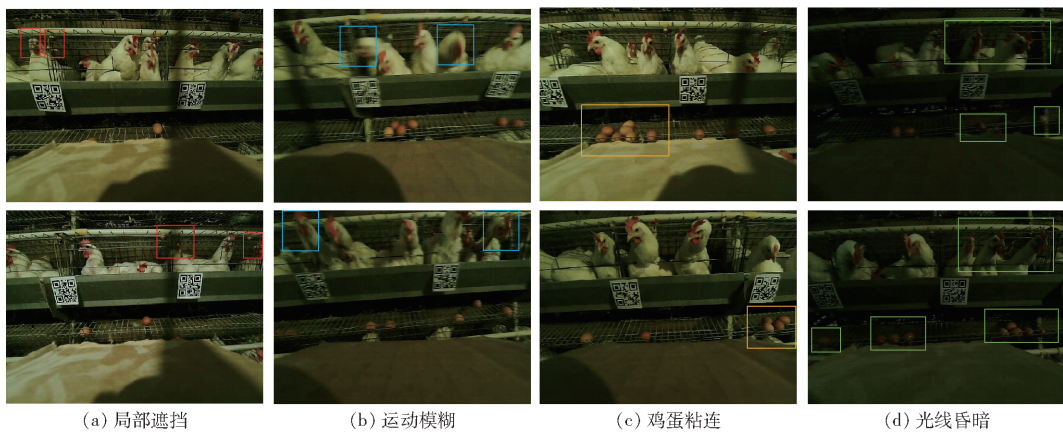


图 2 不同场景下图像示例

Fig. 2 Examples of images in different scenes

区分不同鸡笼。在人工可识别情况下,采用 Labellmg 标注鸡只头部(head)、鸡蛋(egg)与二维码(qrcode),并生成包含目标类别、坐标及图像像素信息的.xml文件,最终得到2146幅图像和对应标注文件用于训练与测试网络模型性能。为增强模型泛

化性和鲁棒性,在训练过程中应用在线数据增强,主要方法包括色调变换、仿射变换、Mosaic增强与Mixup增强,如图3所示。数据集按照比例7:2:1随机划分为训练集、验证集和测试集,对应图像分别为1501、430、215幅。



图 3 在线数据增强

Fig. 3 Online data augmentation

1.3 笼养鸡/蛋识别模型

1.3.1 YOLO v7-tiny 网络模型结构

YOLO v7-tiny^[22]是YOLO系列的轻量化模型之一,该网络在公开数据集 Microsoft COCO Dataset上取得较高的精度与较快的速度。YOLO v7-tiny整个网络结构可分为Input、Backbone与Head共3部分,网络结构如图4所示。Input主要对输入数据进行预处理,对于输入的任意尺寸图像被自适应放缩为模型适用分辨率(640像素×480像素),采用动态标签分配策略确定正负样本,并对数据进行Mosaic与Mixup增强。

Backbone由卷积模块、高效层聚合网络(Efficient layer aggregation network,ELAN)与MP模块构成。卷积模块包含卷积层、数据归一化层与激活函数3部分,激活函数采用Sigmoid加权线性单元(Sigmoid-weighted linear unit,SiLU)^[23],其具有无上界、有下界、平滑与非单调等特点。ELAN通过控制最短与最长的梯度路径,使网络在不破坏原有梯度路径的情况下不断增强网络的特征学习能力,提高模型的鲁棒性。MP模块采用两个不同分支分别进行下采样操作,二者采用残差连接,以增强模型的特

征提取能力。

Head主要包含SPPCSPC、ELAN与Detect3个模块。SPPCSPC由金字塔池化结构(Spatial pyramid pooling,SPP)^[24]与跨阶段区域网络(Cross stage partial,CSP)^[25]组成,CSP将特征拆分为两部分,特征1经过SPP池化操作后与特征2融合,获取4个不同尺寸(1、5、9、13)的感受野,以便Detect层检测不同大小的目标。Detect层输出目标边界框坐标、置信度和类别信息。

原始YOLO v7-tiny对粘连的鸡蛋和光线较暗处目标检测性能较差,为提高模型整体检测精度,满足笼养鸡舍复杂环境的应用需求,本研究在YOLO v7-tiny网络的基础上,首先将SiLU激活函数替换为指数线性单元(Exponential linear unit,ELU),减少模型训练时间和提高目标检测网络的整体性能。其次在ELAN结构中采用深度过参数化深度卷积层(Depthwise over-parameterized depthwise convolutional layer,DO-DCnv)^[26]提取目标深层特征,以提升光线较暗条件下目标的识别能力。最后在Head的第2、4层引入坐标注意力机制(Coordinate attention mechanism,CoordAtt)^[27],提取特征图像中目标空间

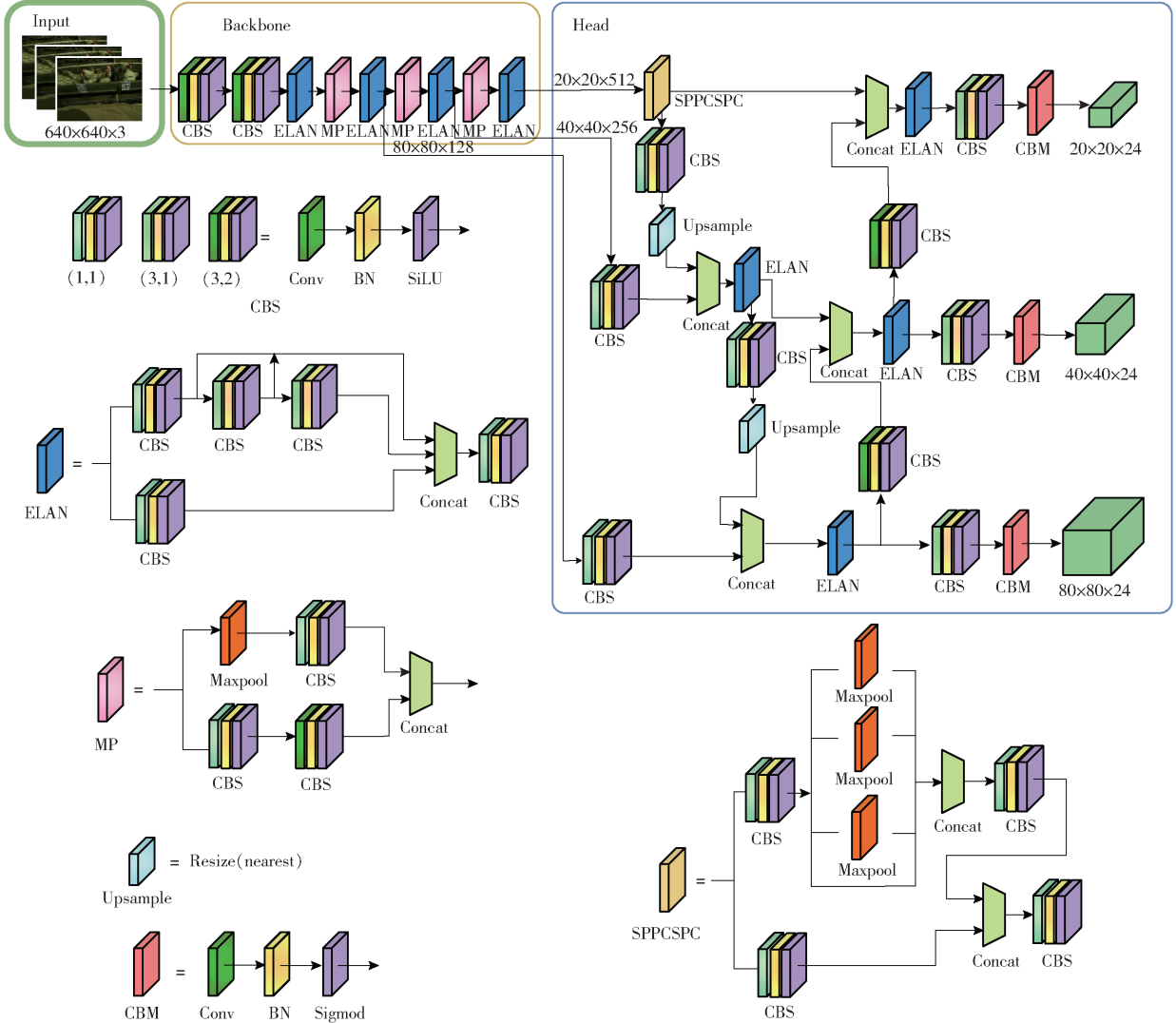


图4 YOLO v7 - tiny 网络结构

Fig.4 Network structure of YOLO v7 - tiny

位置信息。

1.3.2 激活函数

深度学习中激活函数是将神经网络非线性化, 激活隐藏结点, 此类函数通常具有连续且可微等特点。SiLU 激活函数虽然满足以上要求, 但仅在深度神经网络的隐藏层中可用^[28]。因此, 本研究引入 ELU 激活函数。图 5 为 ELU 与 SiLU 激活函数对比曲线图, 可以看出, ELU 函数在定义域内均连续且可微, 满足激活函数的基本要求。当输入特征向量 x 大于 0 时, 输出为 x 本身, 减轻了梯度弥散问题 (导数为 1); 当 x 小于 0 时, 采用指数函数连接, 避免陷入局部最优。ELU 函数输出均值为 0, 与 SiLU 函数相比, 降低了模型计算复杂度, 且单侧饱和, 使模型收敛效果更好。

1.3.3 改进的高效层聚合网络

目标检测模型中特征提取是后续特征融合的基础数据来源。YOLO v7 - tiny 网络中 ELAN 采用常规卷积提取目标特征, 但实际鸡舍环境较为复杂, 常

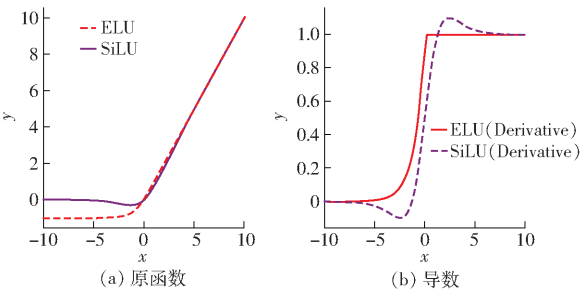


图5 激活函数 ELU 与 SiLU 的对比图

Fig.5 Comparison diagrams of activation functions between ELU and SiLU

规卷积操作难以挖掘目标深层次特征。因此, 本研究将 ELAN 模块内的常规卷积层替换为深度卷积层 (Depthwise convolutional layer, DConv) 以减少模型参数量, 并在其基础上添加过参数化组件 (深度卷积) 构建深度过参数化深度卷积层, 改进后的 ELAN 结构如图 6 所示。与常规卷积不同, DO - DConv 模块不仅能够通过增加网络深度提高网络的表达能力, 而且采用多层复合线性操作, 使其在推理阶段折

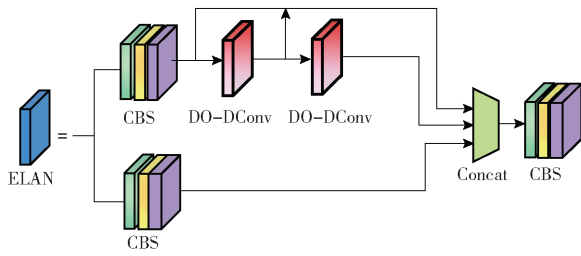


图 6 改进后的 ELAN 结构

Fig. 6 Structure of improved efficient layer aggregation network

叠为单层形式,从而减少计算量。

图 7 为深度过参数化深度卷积层结构,可以看出,DO-DConv 模块是由深度卷积内核 $D \in \mathbf{R}^{(M \times N) \times D_{mul} \times C_{in}}$ 与 $W \in \mathbf{R}^{D_{mul} \times D_{mul}^w \times C_{in}}$ 构成 ($D_{mul} \geq M \times N$)。其计算流程如下:首先将深度卷积算子应用于深度卷积内核 D 和输入特征 $P (P \in \mathbf{R}^{(M \times N) \times C_{in}})$,经过深度卷积计算生成特征 $P' = D \circ P$;然后再将 P' 应用于深度卷积内核 W 与 P' 生成特征 $O = W \circ P'$ 。最后可将深度卷积过参数化深度卷积的输出表示为 $O = W \circ (D \circ P)$ 。 M, N 为输入特征的空间维度, C_{in} 、 C_{out} 分别为输入和输出特征的通道数, D_{mul} 为深度卷积的深度乘数。

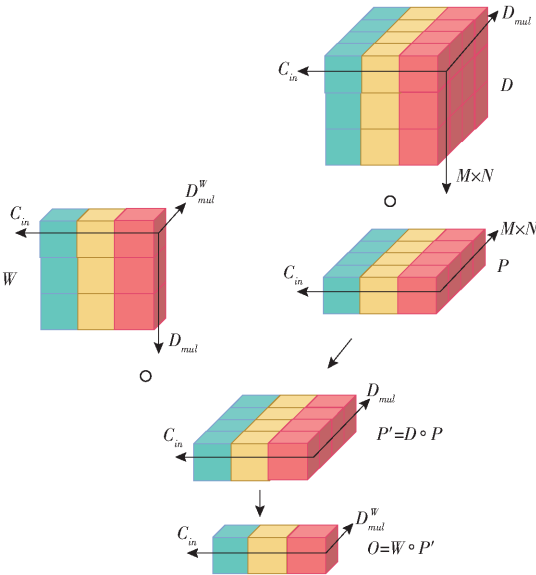


图 7 深度过参数化深度卷积层结构

Fig. 7 Structure of depthwise over-parameterized depthwise convolutional layer

1.3.4 坐标注意力机制

为了进一步提升模型对鸡只和鸡蛋空间位置信息的捕获能力,本研究在网络 Head 端引入坐标注意力机制。CoordAtt 通过精确的位置信息对通道关系和长期依赖性进行编码,形成一个对方向和位置敏感的特征图用于增强目标的特征表示。图 8 为 CoordAtt 编码过程,总体可分为坐标信息嵌入与坐标注意力生成两部分。

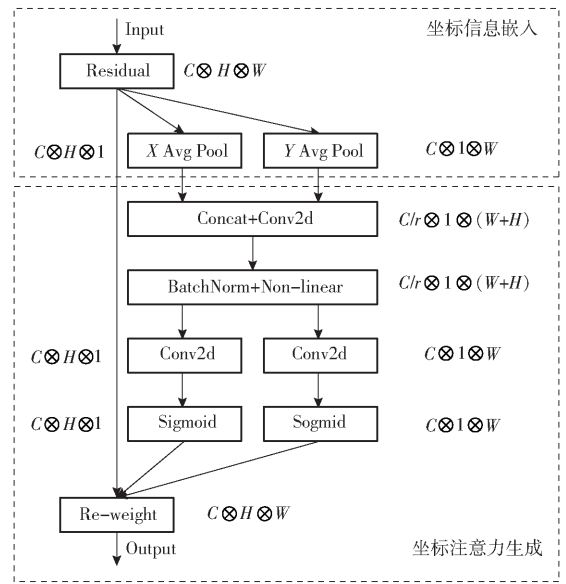


图 8 坐标注意力机制

Fig. 8 Coordinate attention mechanism

由图 8 坐标信息嵌入模块可知,对于输入特征 $X \in \mathbf{R}^{C \times H \times W}$,采用尺寸为 $(H, 1)$ 和 $(1, W)$ 的池化卷积核分别沿水平坐标与垂直坐标对每个通道进行编码,计算平均池化后,第 c 个通道上高度 h 和宽度 w 的编码可表示为 $z_c^h(h)$ 与 $z_c^w(w)$,计算公式为

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} X_c(h, i) \quad (1)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} X_c(j, w) \quad (2)$$

式中 $X_c(h, i)$ ——高度 h 在第 c 个通道的分量

$X_c(j, w)$ ——宽度 w 在第 c 个通道的分量

这种转换使注意力机制模块可以捕捉空间方向的长期依赖关系,并保存沿着另一空间方向特征的精确位置信息。经过信息嵌入模块的变换后,将两个方向的坐标信息进行拼接,然后输入 1×1 卷积函数 F_1 中进行变换操作,计算公式为

$$f = \delta(F_1([z^h, z^w])) \quad (3)$$

式中 $[,]$ ——沿空间维度的串联操作

δ ——非线性激活函数

f ——对空间信息在水平与垂直方向进行编码的中间特征

将 f 沿水平与垂直方向分解为张量 $f^h \in \mathbf{R}^{\frac{C}{r} \times H}$ 和 $f^w \in \mathbf{R}^{\frac{C}{r} \times W}$,然后分别采用 1×1 卷积 F_h 和 F_w 将 f^h 与 f^w 变换为同输入 X 具有相同通道数的张量,变换过程为

$$g^h = \sigma(F_h(f^h)) \quad (4)$$

$$g^w = \sigma(F_w(f^w)) \quad (5)$$

式中 σ ——Sigmoid 激活函数

g^h ——沿水平方向特征张量

g^w ——沿垂直方向特征张量

由于 σ 计算耗时较大,为了加快运算速度,采用下采样比 r 减少 f 通道数。综上,对于输入 X 的第 c 个通道上高度 i 和宽度 j 的特征 $X_c(i,j)$,经过坐标注意力模块的输出可表示为

$$y_c(i,j) = X_c(i,j) \times g_c^h(i) \times g_c^w(j) \quad (6)$$

式中 $y_c(i,j)$ ——嵌入坐标注意力模块后的输出结果

$g_c^h(i)$ ——第 c 个通道上水平方向的注意力权重

$g_c^w(j)$ ——第 c 个通道上垂直方向的注意力权重

1.4 分笼计数法

ALPATOV 等^[29]提出车辆检测与道路标线检测算法,并根据道路标线算法划分车道从而实现多通道车辆计数,结果表明,5 个通道平均计数准确率为 98.9%。基于此,本研究通过检测二维码获取检测框坐标和存储信息区分不同鸡笼,并设计分笼计数法(Counting in different cages, CDC)实现鸡只与鸡蛋自动化计数。该算法可分为计数感兴趣区域(Counting in region of interest, CRoI)设置与自动化分笼计数两个阶段。CRoI 设置流程如下:机器运动过程中,采用 YOLO v7-tiny-DO 检测图像中目标,当识别到 2 个二维码时,根据其检测框坐标分别计算待测图像中鸡只与鸡蛋计数区域的横、纵坐标,并以此为边界划分目标 CRoI,如图 9a(红色框表示鸡只 CRoI,绿色框为鸡蛋 CRoI),计算过程为

$$\begin{cases} x_{\min} = \min(x_q, x_{q+1}) \\ x_{\max} = \max(x_q, x_{q+1}) \\ y_{\min}^{\text{head}} = 0 \\ y_{\max}^{\text{head}} = (y_q + y_{q+1})/2 + 40 \\ y_{\min}^{\text{egg}} = y_q \\ y_{\max}^{\text{egg}} = (y_q + y_{q+1})/2 + 190 \end{cases} \quad (q = 0, 1, \dots, 29) \quad (7)$$

式中 q ——二维码存储的鸡笼编号

x_q ——二维码检测框左侧上顶点横坐标

y_q ——二维码检测框左侧下顶点纵坐标

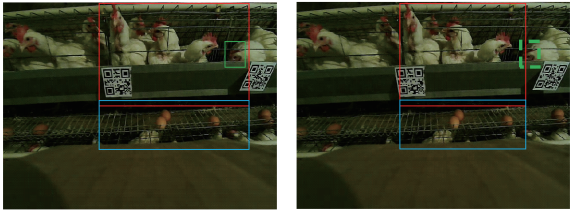
$[x_{\min}, x_{\max}]$ ——计数区域横坐标取值范围

$[y_{\min}^{\text{head}}, y_{\max}^{\text{head}}]$ ——鸡只计数区域的纵坐标取值范围

$[y_{\min}^{\text{egg}}, y_{\max}^{\text{egg}}]$ ——鸡蛋计数区域的纵坐标取值范围

值得注意的是,若未识别到 2 个二维码,则不进行 CRoI 设置操作,跳至下一帧。

选定 CRoI 后进行自动化分笼计数。过程如下:首先采用 pyzbar 函数识别二维码信息,选择较小数



(a) 修正前 (b) 修正后

图 9 计数感兴趣区域

Fig. 9 Counting in regions of interest

值作为待测鸡笼编号 q ;然后获取当前帧 $F(k,q)$ 所有鸡只与鸡蛋检测框的中心点坐标 (x_i, y_i) ,计算其在 CRoI 范围内的目标个数观测值 $F(k,q)_{\text{num}}$,计算过程为

$$F(k,q)_{\text{num}} = \sum_{i=1}^n C(k,q)_{(x_i,y_i)} \quad (8)$$

其中 $C(k,q)_{(x_i,y_i)} = \begin{cases} 1 & ((x_i, y_i) \in \text{CRoI}) \\ 0 & ((x_i, y_i) \notin \text{CRoI}) \end{cases}$

式中 k ——帧序号

最后,当二维码信息发生变化时,统计鸡笼编号 q 所有视频帧的观测值,计算其均值作为目标数量预测值 $P(k,q)$,计算过程为

$$P(k,q) = \lceil \text{mean}_{a \in [0, L]} (F(k_a, q)_{\text{num}}) \rceil \quad (9)$$

式中 $\lceil \rceil$ ——小数向上取整计算符号

mean——所有帧预测值的均值

实际计数过程中,巡检机器人移动时获取的计数感兴趣区域与实际鸡笼区域相比误差较大,导致分笼计数错误(图 9a 绿色虚线框)。为减小运动过程中计数区域的误差,本研究采用单目相机标定模型^[30]对 CRoI 进行校正,实现 CRoI 在像素坐标与世界坐标之间的相互转换,坐标变换过程为

$$\begin{cases} W_x = \frac{(i_x - c_x)z_1}{f_x} \\ W_y = \frac{(i_y - c_y)z_1}{f_y} \\ W_z = z_1 \end{cases} \quad (10)$$

$$\begin{cases} I_x = \frac{W_x f_x}{z_2} + c_x \\ I_y = \frac{W_y f_y}{z_2} + c_y \end{cases} \quad (11)$$

式中 (c_x, c_y) ——图像中心点坐标

(i_x, i_y) ——鸡笼边界坐标

z_1 ——镜头到二维码的水平距离

z_2 ——镜头到鸡笼的水平距离

(W_x, W_y, W_z) —— (i_x, i_y) 转换后的世界坐标

(I_x, I_y) —— (i_x, i_y) 纠正后坐标

f_x, f_y ——横、纵像素焦距

1.5 模型训练

1.5.1 训练环境

模型训练硬件环境主要包括 Intel(R) Core(TM) i9-10900X 处理器,主频 3.70 GHz,运行内存 24 GB,GPU 型号为 Titan RTX,操作系统为 Ubuntu 18.04。软件环境为 Python 3.7 语言版本、PyTorch 1.7 框架、OpenCV 4.3.5 图像处理库与 pyzbar 0.1.8 二维码识别库。

1.5.2 参数设置

网络训练参数设置如下:输入图像尺寸为 640 像素×480 像素,批量大小为 32,初始学习率设置为 0.001。采用随机梯度下降(Stochastic gradient descent,SGD)方法优化模型,动量因子为 0.937,权重衰减因子为 0.000 5。模型共训练 300 轮,迭代 1 个轮次保存一次模型权重,最终选择识别精度最高的模型。

1.5.3 评价指标

采用精确率(Precision, P)、召回率(Recall, R)、平均精确率(Average precision,AP)、F1 值、模型内存占用量(Model size)和帧率(Frames per seconds,

FPS)来评价模型性能,采用平均绝对误差(Mean absolute error,MAE)和平均准确率(Mean accuracy rate,MAR)评估分笼计数算法。

2 结果与分析

2.1 不同检测模型性能对比

为验证 YOLO v7-tiny-DO 模型对鸡只与鸡蛋的识别能力,本研究选择主流目标检测网络 Faster R-CNN、SSD、YOLO v4-tiny、YOLO v5n、YOLO v7-tiny 与 YOLO v7-tiny-DO 进行对比试验。训练过程中,所有网络使用相同训练集,并应用 215 幅测试集图像对模型性能进行综合评估。经完整测试后,不同模型检测结果如表 1 所示,表中 $AP_{0.5}$ 表示交并比为 0.5 时的平均精确率。从表 1 可以看出,与 Faster R-CNN、SSD、YOLO v4-tiny、YOLO v5n 和 YOLO v7-tiny 相比,YOLO v7-tiny-DO 的 F1 值优势较为突出,鸡只的 F1 值为 97.0%,分别高于以上模型 21.0、4.0、8.0、1.5、1.7 个百分点;鸡蛋的 F1 值为 98.4%,比其他模型分别高 31.4、25.4、6.4、4.4、0.7 个百分点。

表 1 不同检测网络对比
Tab.1 Comparison of different detection networks

模型	$P/\%$		$R/\%$		$AP_{0.5}/\%$		F1 值/ $\%$		模型内存占 用量/MB	帧率/ ($f\cdot s^{-1}$)
	鸡只	鸡蛋	鸡只	鸡蛋	鸡只	鸡蛋	鸡只	鸡蛋		
Faster R-CNN	63.2	50.9	95.7	96.3	93.7	90.9	76.0	67.0	108.0	11.2
SSD	96.7	91.0	88.9	61.4	92.7	67.6	93.0	73.0	91.6	71.6
YOLO v4-tiny	89.8	92.5	87.6	91.0	93.5	95.8	89.0	92.0	22.4	88.0
YOLO v5n	96.5	94.6	94.5	93.4	97.0	96.1	95.5	94.0	3.7	98.0
YOLO v7-tiny	97.4	97.9	93.2	97.5	93.7	97.9	95.3	97.7	11.7	109.9
YOLO v7-tiny-DO	97.4	98.0	96.5	98.8	96.9	99.3	97.0	98.4	5.6	106.4

对比 YOLO v7-tiny-DO 与 YOLO v7-tiny,二者帧率均高于 100 f/s,能够满足实时检测任务的需要;YOLO v7-tiny-DO 模型检测鸡只和鸡蛋的 AP 分别为 96.9%与 99.3%,相比 YOLO v7-tiny 分别提高 3.2、1.4 个百分点;YOLO v7-tiny-DO 模型内存占用量为 5.6 MB,比原模型减小 6.1 MB,适用于算力匮乏的巡检机器人。综上,YOLO v7-tiny-DO 网络在提升模型精度的情况下降低了模型参数量,证明了改进后模型的有效性。

2.2 不同场景下检测结果分析

由于实际鸡舍环境较为复杂,采集图像中通常存在局部遮挡、运动模糊、鸡蛋粘连与光线昏暗等问题,导致鸡只与鸡蛋自动化识别与计数困难。因此,针对以上 4 种场景分别进行讨论。

2.2.1 局部遮挡情况下识别结果对比

笼养鸡舍环境中铁丝或食槽遮挡易造成目标

部分特征丢失,为验证改进后模型在局部遮挡情况下检测能力,将图 2a 的 2 幅图像输入 6 个模型进行测试,检测结果如图 10 所示,图中绿色、橙色、蓝色和紫色椭圆分别表示鸡只漏检、鸡只错检、鸡蛋漏检与鸡蛋错检。表 2 为对应检测结果,可以看出,YOLO v7-tiny-DO、YOLO v7-tiny 和 YOLO v5n 能够识别图像中所有目标;Faster R-CNN 虽然能够识别被遮挡目标,但误将红色饮水器识别为鸡只或鸡蛋(图 10a),与其模型精确率低、召回率高相对应,发生错检次数较多;与 Faster R-CNN 不同,SSD 和 YOLO v4-tiny 召回率低、精确率高,在遮挡较严重情况下,鸡只或鸡蛋特征缺失,易发生漏检现象(图 10b、10c)。经过对比分析,YOLO v7-tiny-DO、YOLO v7-tiny 和 YOLO v5n 模型对于遮挡目标检测结果较好,未发生漏检或错检情况。

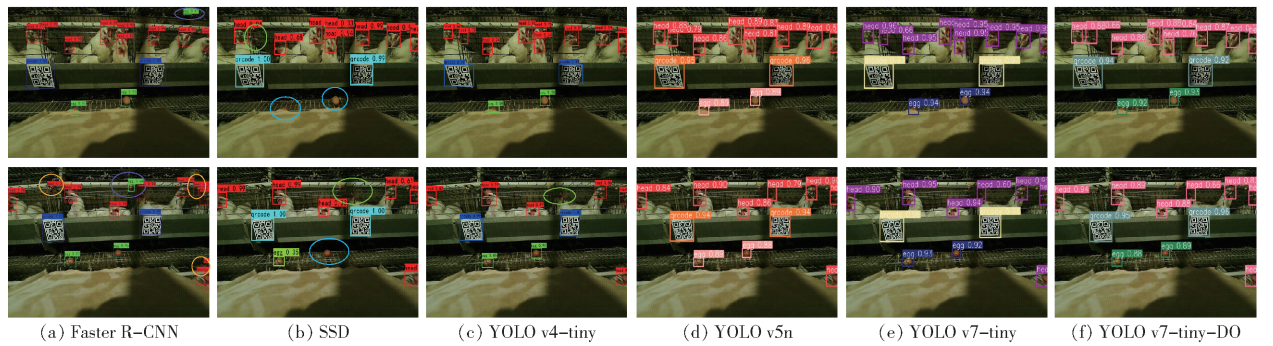


图 10 局部遮挡情况图像检测结果对比

Fig. 10 Comparison of images detection results under partial occlusion

表 2 局部遮挡情况下检测结果对比

模型	Tab.2 Comparison of detection results under partial occlusion							
	实际		正确预测		漏检		错检	
	目标数		目标数		目标数		目标数	
	鸡只	鸡蛋	鸡只	鸡蛋	鸡只	鸡蛋	鸡只	鸡蛋
Faster R-CNN	16	4	16	4	0	0	3	2
SSD	16	4	14	3	2	3	0	0
YOLO v4-tiny	16	4	15	4	1	0	0	0
YOLO v5n	16	4	16	4	0	0	0	0
YOLO v7-tiny	16	4	16	4	0	0	0	0
YOLO v7-tiny-DO	16	4	16	4	0	0	0	0

2.2.2 运动模糊情况下识别结果对比

实际巡检过程中,鸡只快速移动易导致图像中鸡只头部发生模糊现象,图 11 为 6 个模型检测图 2b 中模糊鸡头情况的结果,对应试验结果如表 3 所示。由表 3 可以看出,YOLO v7-tiny-DO 能够识别图像中所有模糊鸡头,Faster R-CNN 漏检 1 只鸡(图 11a),SSD 漏检 2 只鸡(图 11b),YOLO v4-tiny 漏检 2 只鸡(图 11c),YOLO v5n 漏检 1 只鸡(图 11d),YOLO v7-tiny 漏检 2 只鸡(图 11e)。当图像中鸡头发生运动模糊时,目标特征较为发散,不易识别。经过对比,在运动模糊情况下本研究提出

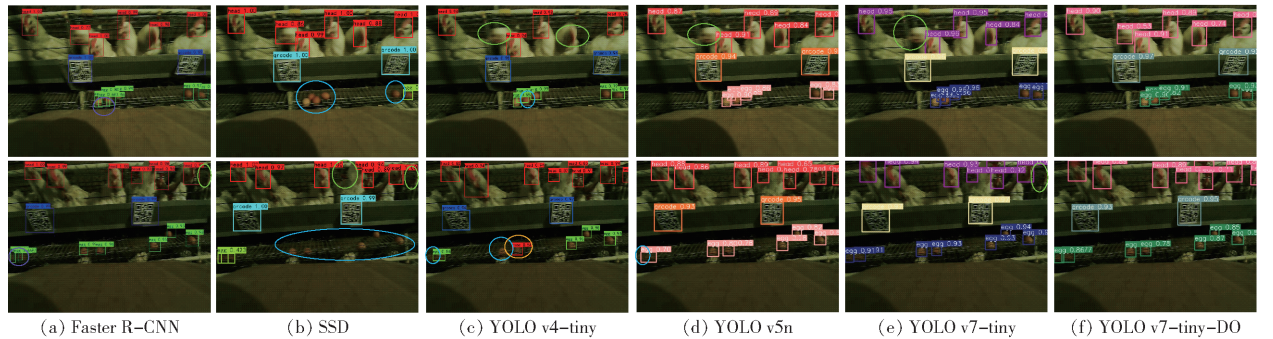


图 11 运动模糊情况图像检测结果对比

Fig. 11 Comparison of images detection results under motion blur

表 3 运动模糊情况下检测结果对比

模型	Tab.3 Comparison of detection results under motion blur							
	实际		正确预测		漏检		错检	
	目标数		目标数		目标数		目标数	
	鸡只	鸡蛋	鸡只	鸡蛋	鸡只	鸡蛋	鸡只	鸡蛋
Faster R-CNN	14	13	13	13	1	0	0	2
SSD	14	13	12	3	2	10	0	0
YOLO v4-tiny	14	13	12	9	2	3	1	0
YOLO v5n	14	13	13	12	1	1	0	0
YOLO v7-tiny	14	13	12	13	2	0	0	0
YOLO v7-tiny-DO	14	13	14	13	0	0	0	0

连现象较为严重,图 12 为 6 个模型在图 2c 中粘连鸡蛋情况下的识别结果,对应试验结果如表 4 所示。由表 4 可知,YOLO v7-tiny-DO 准确识别图像中所有粘连鸡蛋,其他模型均存在不同程度的鸡蛋漏检情况(图 12a~12e),其中 SSD 识别结果最差(图 12b),仅能检测到多个粘连鸡蛋中的一枚鸡蛋。经过对比,YOLO v7-tiny-DO 在鸡蛋粘连情况下的鸡蛋计数性能最优。

2.2.4 光线昏暗情况下识别结果对比

笼养鸡舍内昏暗的环境为蛋鸡和鸡蛋识别与计数工作带来巨大挑战。图 13 为 6 个模型在图 2d 光线昏暗环境下的检测结果,对应试验结果如表 5 所示。由表 5 可以看出,YOLO v7-tiny-DO 漏检了 1 只鸡(图 13f);Faster R-CNN 错检 3 只鸡和 2 个

的 YOLO v7-tiny-DO 检测性能优于其他模型。

2.2.3 鸡蛋粘连情况下识别结果对比

笼养模式下蛋鸡饲养密度较大,蛋网上鸡蛋粘

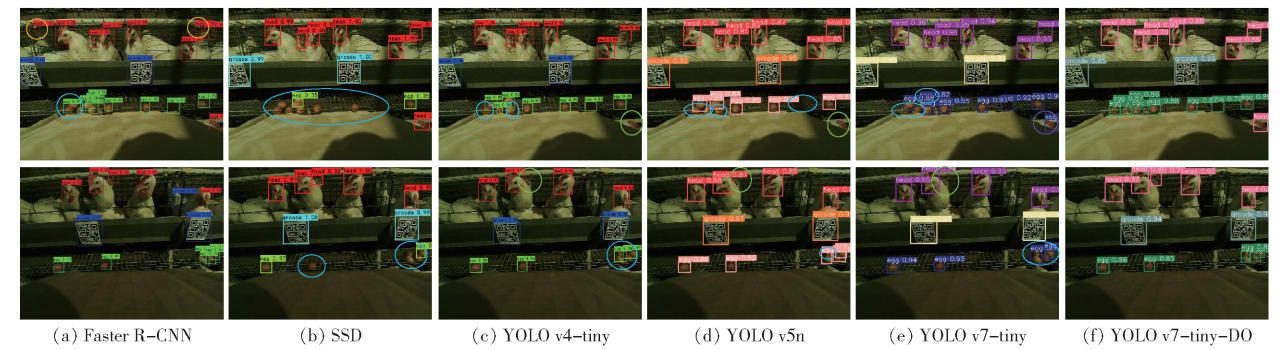


图 12 鸡蛋粘连情况图像检测结果对比

Fig. 12 Comparison of images detection results under eggs adhesion

表 4 鸡蛋粘连情况下检测结果对比

模型	Tab.4 Comparison of detection results under eggs adhesion							
	实际		正确预测		漏检		错检	
	目标数		目标数		目标数		目标数	
	鸡只	鸡蛋	鸡只	鸡蛋	鸡只	鸡蛋	鸡只	鸡蛋
Faster R-CNN	12	16	12	15	0	1	2	0
SSD	12	16	12	4	0	12	0	0
YOLO v4-tiny	12	16	10	12	2	4	0	0
YOLO v5n	12	16	11	12	1	4	0	0
YOLO v7-tiny	12	16	11	11	1	5	0	1
YOLO v7-tiny-DO	12	16	12	16	0	0	0	0

鸡蛋(图 13a); SSD 漏检 2 只鸡和 12 枚鸡蛋(图 13b); YOLO v4-tiny 漏检 2 只鸡和 2 个鸡蛋, 错检 2 只鸡(图 13c); YOLO v5n 漏检 2 只鸡和 2 枚鸡蛋(图 13d); YOLO v7-tiny 漏检 2 只鸡和 2 枚鸡蛋(图 13e)。综上, 图中鸡只与鸡蛋的形态特征较为模糊, 所有模型均存在漏检或错检情况, 但相比其他模型, YOLO v7-tiny-DO 漏检与错检数目最少, 说明本研究提出的模型性能更好。

2.3 计数试验结果与分析

将 YOLO v7-tiny-DO 部署到 NVIDIA Jetson AGX Xavier 边缘计算平台, 经深度学习引擎

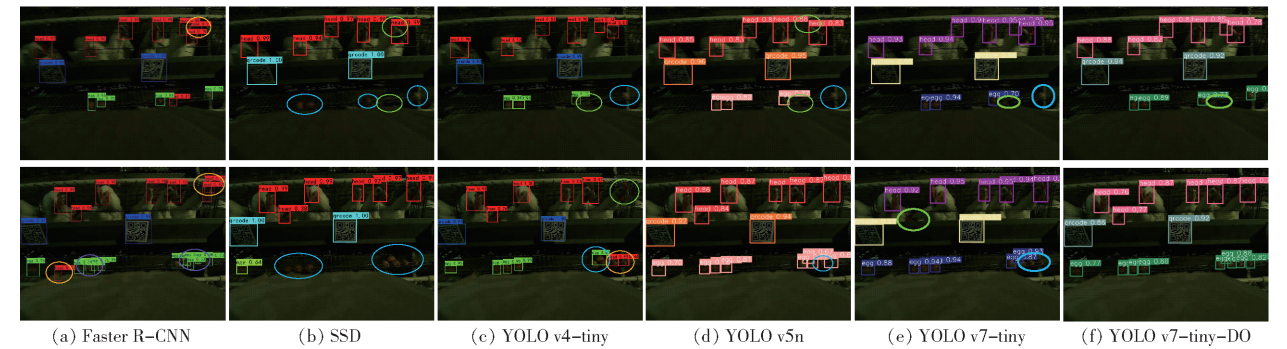


图 13 光线昏暗环境图像检测结果对比

Fig. 13 Comparison of images detection results in dim environment

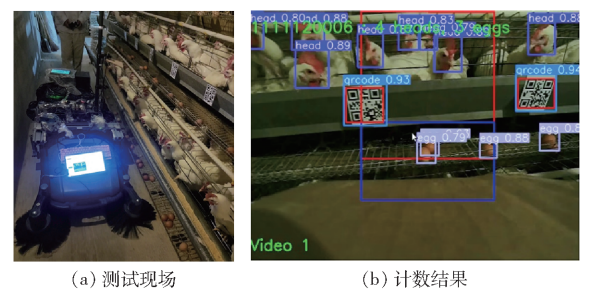
表 5 光线昏暗环境下检测结果对比

模型	Tab.5 Comparison of detection results in dim environment							
	实际		正确预测		漏检		错检	
	目标数		目标数		目标数		目标数	
	鸡只	鸡蛋	鸡只	鸡蛋	鸡只	鸡蛋	鸡只	鸡蛋
Faster R-CNN	12	13	12	13	0	0	3	3
SSD	12	13	10	1	2	12	0	0
YOLO v4-tiny	12	13	10	8	2	3	2	0
YOLO v5n	12	13	10	11	2	2	0	0
YOLO v7-tiny	12	13	10	9	2	4	0	0
YOLO v7-tiny-DO	12	13	11	13	1	0	0	0

(DL accelerator 与 Vision accelerator)加速后,快速检测视频帧中目标,实现笼养模式下鸡只与鸡蛋自动化分笼计数。为验证算法有效性,本研究在北京市

华都峪口禽业有限公司第 3 场 8 舍选取 30 个鸡笼开展计数试验,巡检机器人运动过程中,基于内置取帧算法每天采集 15 min 视频数据用于测试,持续 3 d,期间第 21 号鸡笼有 1 只死鸡被淘汰。图 14 为测试现场及计数结果。

表 6 为不同测试批次鸡只与鸡蛋的自动化分笼计数结果。由表 6 可知,测试批次 1 中鸡只和鸡蛋分笼计数 MAR 分别为 96.7% 和 97.2%,MAE 分别为 0.13 只和 0.07 枚;测试批次 2 中鸡/蛋分笼计数 MAR 分别为 97.5% 和 94.5%,MAE 分别为 0.10 只和 0.13 枚;测试批次 3 中鸡只和鸡蛋分笼计数 MAR 分别为 96% 和 97.2%,MAE 分别为 0.17 只和 0.07 枚。经过分析可知,3 次测试鸡只与鸡蛋分笼计数平均准确率均值分别为 96.7% 和 96.3%,每笼



(a) 测试现场 (b) 计数结果

图 14 测试现场与计数结果

Fig. 14 Testing on spot and results of counting

表 6 鸡/蛋自动化分笼计数结果

Tab. 6 Results of automatic counting hens and eggs in different cages						
测试批次	目标类别	鸡笼数量/个	人工观测值/个	模型预测值/个	MAR/%	MAE
1	鸡只	30	120	116	96.7	0.13 只
	鸡蛋	30	55	53	97.2	0.07 枚
2	鸡只	30	119	116	97.5	0.10 只
	鸡蛋	30	66	62	94.5	0.13 枚
3	鸡只	30	119	114	96.0	0.17 只
	鸡蛋	30	63	61	97.2	0.07 枚

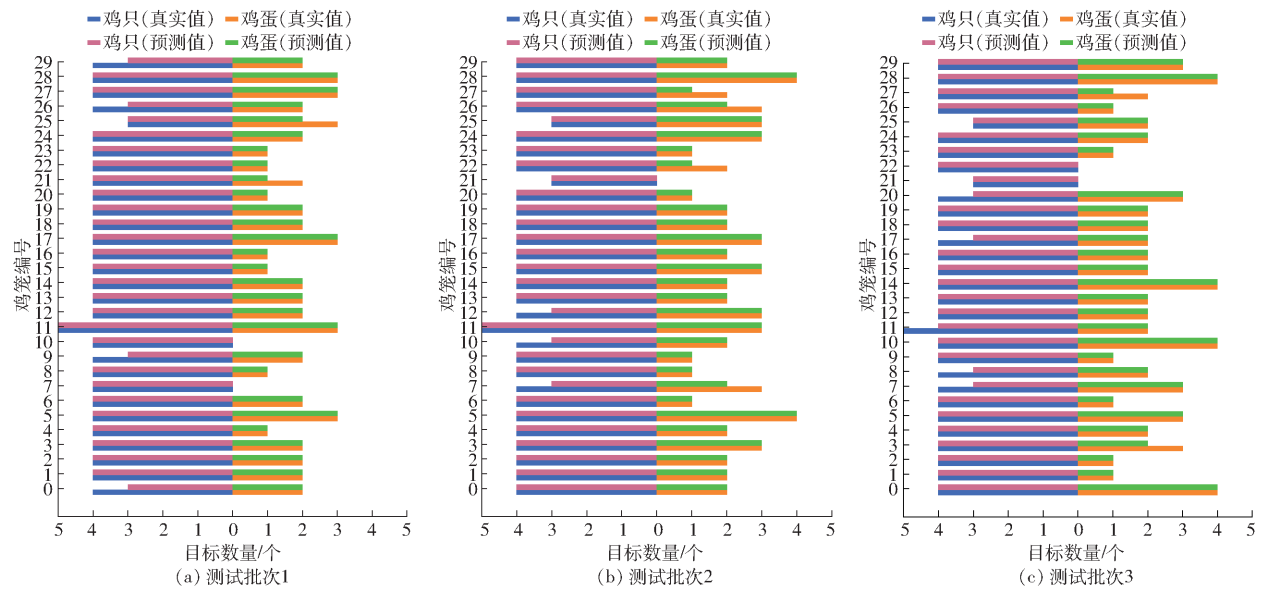


图 15 鸡/蛋自动化分笼计数可视化结果

Fig. 15 Visualization of results of automatic counting hens and eggs in different cages

为 106.4 f/s,模型内存占用量为 5.6 MB,与原模型相比,识别鸡只与鸡蛋的 AP 分别提高了 3.2、1.4 个百分点,模型内存占用量减少了 6.1 MB,适用于部署到算力匮乏的巡检机器人上。

(2)与主流目标检测网络 Faster R-CNN、SSD、YOLO v4-tiny 和 YOLO v5n 相比,YOLO v7-tiny-DO 的 F1 值更优,分别高 21.0、4.0、8.0、1.5 个百分点,鸡蛋的 F1 值分别高 31.4、25.4、6.4、4.4 个百分点。对比以上模型在复杂场景下的检测性能,结果

平均绝对误差均值分别为 0.13 只和 0.09 枚,说明通过本文算法获取笼内鸡只与鸡蛋个数是可行的。

图 15 为不同测试批次鸡只与鸡蛋自动化分笼计数结果统计图。从图中可以看出,测试批次 1 中 0、9、26、29 号鸡笼均漏检 1 只鸡,21、25 号鸡笼均漏检 1 枚鸡蛋;测试批次 2 中 7、10、12 号鸡笼存在鸡只计数错误,7、22、26、27 号鸡笼存在鸡蛋计数错误,平均绝对误差均为 1 枚;测试批次 3 中 7、8、11、17、20 号鸡笼均漏检 1 只鸡,3、27 号鸡笼均漏检 1 枚鸡蛋。对于漏检目标,可能是因为鸡舍内光线较为昏暗导致。

3 结论

(1)选择 YOLO v7-tiny 为基础网络,首先将模型激活函数替换为 ELU,然后在 ELAN 内构建深度过参数化深度卷积层,最后在 Head 端引入坐标注意力机制。试验结果表明,YOLO v7-tiny-DO 识别鸡只与鸡蛋的 AP 分别为 96.9% 与 99.3%,帧率

表明,YOLO v7-tiny-DO 在局部遮挡、动态模糊和鸡蛋粘连情况下均能实现高精度识别与定位,在光线昏暗情况下识别结果优于其他模型。

(3)将 YOLO v7-tiny-DO 模型部署到搭载 NVIDIA Jetson AGX Xavier 的巡检机器人上,开展为期 3 d 的计数测试。结果表明,3 个测试批次鸡只与鸡蛋的分笼计数平均准确率均值分别为 96.7% 和 96.3%,每笼平均绝对误差均值分别为 0.13 只鸡和 0.09 枚鸡蛋。

参 考 文 献

[1] JOAT N, VAN T T H, STANLEY D, et al. Temporal dynamics of gut microbiota in caged laying hens: a field observation from hatching to end of lay[J]. *Applied Microbiology and Biotechnology*, 2021,105(11):4719–4730.

[2] 李保明, 王阳, 郑炜超, 等. 中国规模化养鸡环境控制关键技术与设施设备研究进展[J]. *农业工程学报*, 2020, 36(16):212–221.

LI Baoming, WANG Yang, ZHENG Weichao, et al. Research progress in environmental control key technologies, facilities and equipment for laying hen production in China[J]. *Transactions of the CSAE*, 2020, 36(16): 212–221. (in Chinese)

[3] 梁超, 尹欢欢, 李保明, 等. 蛋鸡舍热湿环境参数全年逐时动态预测模型[J]. *农业工程学报*, 2021,37(8):229–235.

LIANG Chao, YIN Huanhuan, LI Baoming, et al. Hourly model for predicting year-round temperature and relative humidity of the environment in laying hen houses[J]. *Transactions of the CSAE*, 2021, 37(8): 229–235. (in Chinese)

[4] 宗超, 曹晏飞, 曹孟冰, 等. 笼养和栖架养殖模式下蛋鸡的发声特征[J]. *农业工程学报*, 2021,37(6):135–141.

ZONG Chao, CAO Yanfei, CAO Mengbing, et al. Vocalization characteristics of laying hens under cage and perch systems[J]. *Transactions of the CSAE*, 2021, 37(6): 135–141. (in Chinese)

[5] 方鹏, 郝宏运, 李腾飞, 等. 基于注意力机制和可变形卷积的鸡只图像实例分割提取[J]. *农业机械学报*, 2021,52(4):257–265.

FANG Peng, HAO Hongyun, LI Tengfei, et al. Instance segmentation of broiler image based on attention mechanism and deformable convolution[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(4): 257–265. (in Chinese)

[6] 毕敏娜, 张铁民, 庄晓霖, 等. 基于鸡头特征的病鸡识别方法研究[J]. *农业机械学报*, 2018, 49(1): 51–57.

BI Minna, ZHANG Tiemin, ZHUANG Xiaolin, et al. Recognition method of sick yellow feather chicken based on head features [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2018, 49(1): 51–57. (in Chinese)

[7] INDRA D, HASANUDDIN T, SATRA R, et al. Eggs detection using otsu thresholding method[C]//2018 2nd East Indonesia Conference on Computer and Information Technology (EIConCIT). IEEE, 2018: 10–13.

[8] MUVVA V V, ZHAO Y, PARAJULI P, et al. Automatic identification of broiler mortality using image processing technology [C] // 10th International Livestock Environment Symposium (ILES X). American Society of Agricultural and Biological Engineers, 2018.

[9] LI Guoming, LI Baoming, SHI Zhengxiang, et al. Design and evaluation of a lighting preference test system for laying hens[J]. *Computers and Electronics in Agriculture*, 2018, 147: 118–125.

[10] WANG Shuo, PU Ziyuan, LI Qianmu, et al. Estimating crowd density with edge intelligence based on lightweight convolutional neural networks[J]. *Expert Systems with Applications*, 2022, 206: 117823.

[11] LI Pei, LU Huishan, WANG Fujie, et al. Detection of sick laying hens by infrared thermal imaging and deep learning[J]. *Journal of Physics: Conference Series*, 2021, 2025(1): 12008.

[12] YANG X, CHAI L L, BIST R B, et al. A deep learning model for detecting cage-free hens on the litter floor[J]. *Animals*, 2022, 12(15): 1983.

[13] GEFFEN O, YITZHAKY Y, BARCHILON N, et al. A machine vision system to detect and count laying hens in battery cages [J]. *Animal*, 2020, 14(12): 2628–2634.

[14] LI Jiao, SU Houcheng, ZHENG Xingze, et al. Study of a querypnnet model for accurate detection and segmentation of goose body edge contours[J]. *Animals*, 2022, 12(19): 2653.

[15] JIANG Kailin, XIE Tianyu, YAN Rui, et al. An attention mechanism-improved YOLO v7 object detection algorithm for hemp duck count estimation[J]. *Agriculture*, 2022, 12(10): 1659.

[16] 景亮, 王瑞, 刘慧, 等. 基于双目相机与改进 YOLO v3 算法的果园行人检测与定位[J]. *农业机械学报*, 2020,51(9): 34–39.

JING Liang, WANG Rui, LIU Hui, et al. Orchard pedestrian detection and location based on binocular camera and improved YOLO v3 algorithm[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2020, 51(9): 34–39. (in Chinese)

[17] LI Huipeng, LI Changyong, LI Guibin, et al. A real-time table grape detection method based on improved YOLO v4-tiny network in complex background[J]. *Biosystems Engineering*, 2021, 212: 347–359.

[18] PRASETYO E, SUCIATI N, FATICHAH C. YOLO v4-tiny with wing convolution layer for detecting fish body part[J]. *Computers and Electronics in Agriculture*, 2022, 198: 107023.

[19] 刘慧, 张礼帅, 沈跃, 等. 基于改进 SSD 的果园行人实时检测方法[J]. *农业机械学报*, 2019, 50(4): 29–35.

LIU Hui, ZHANG Lishuai, SHEN Yue, et al. Real-time pedestrian detection in orchard based on improved SSD [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 50(4): 29–35. (in Chinese)

[20] 易诗, 沈练, 周思尧, 等. 基于增强型 Tiny – YOLO V3 模型的野鸡识别方法[J]. *农业工程学报*, 2020, 36(13): 141–147.

YI Shi, SHEN Lian, ZHOU Siyao, et al. Recognition method of pheasant using enhanced Tiny – YOLO V3 model [J]. *Transactions of the CSAE*, 2020, 36(13): 141–147. (in Chinese)

- [21] ZHENG Xingze, LI Feiyi, LIN Bin, et al. A two-stage method to detect the sex ratio of hemp ducks based on object detection and classification networks[J]. *Animals*, 2022, 12(9): 1177.
- [22] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLO v7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[J]. *arXiv preprint arXiv:2207.02696*, 2022.
- [23] ELFWING S, UCHIBE E, DOYA K. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning[J]. *Neural Networks*, 2018, 107: 3 – 11.
- [24] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904 – 1916.
- [25] LIU W, HASAN I, LIAO S C. Center and scale prediction: anchor-free approach for pedestrian and face detection[J]. *Pattern Recognition*, 2023, 135: 109071.
- [26] CAO Jinming, LI Yangyan, SUN Mingchao, et al. DO – Conv: depthwise over-parameterized convolutional layer[J]. *IEEE Transactions on Image Processing*, 2022, 31: 3726 – 3736.
- [27] HOU Qibin, ZHOU Daquan, FENG Jiashi. Coordinate attention for efficient mobile network design[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021: 13713 – 13722.
- [28] JUNG H K, CHOI G S. Improved YOLO v5: efficient object detection using drone images under various conditions[J]. *Applied Sciences*, 2022, 12(14): 7255.
- [29] ALPATOV B A, BABAYAN P V, ERSHOV M D. Vehicle detection and counting system for real-time traffic surveillance[C] // *2018 7th Mediterranean Conference on Embedded Computing (MECO)*. IEEE, 2018: 1 – 4.
- [30] 曾宏伟, 雷军波, 陶建峰, 等. 基于单目视觉的谷物联合收获机产量测量方法[J]. *农业机械学报*, 2021, 52(12): 281 – 289.
- ZENG Hongwei, LEI Junbo, TAO Jianfeng, et al. Yield monitoring for grain combine harvester based on monocular vision[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(12): 281 – 289. (in Chinese)
- ~~~~~

(上接第 251 页)

- [19] 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. Research on high-density pig counting network based on multi-scale perception[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(9): 172 – 178. (in Chinese)
- [20] 孙俊, 杨锴锋, 罗元秋, 等. 基于无人机图像的多尺度感知麦穗计数方法[J]. *农业工程学报*, 2021, 37(23): 136 – 144.
- SUN Jun, YANG Kaifeng, LUO Yuanqiu, et al. Multi-scale sensory ear counting method based on UAV images[J]. *Transactions of the CSAE*, 2021, 37(23): 136 – 144. (in Chinese)
- [21] 鲍文霞, 张鑫, 胡根生, 等. 基于深度卷积神经网络的田间麦穗密度估计及计数[J]. *农业工程学报*, 2020, 36(21): 186 – 193, 323.
- BAO Wenxia, ZHANG Xin, HU Gensheng, et al. Estimation and counting of wheat ear density in field based on deep convolutional neural network[J]. *Transactions of the CSAE*, 2020, 36(21): 186 – 193, 323. (in Chinese)
- [22] LI Y, ZHANG X, CHEN D. CSRNet: dilated convolutional neural networks for understanding the highly congested scenes[C] // *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, 2018: 1091 – 1100.
- [23] LIU W, SALZMANN M, FUA P. Context-aware crowd counting[C] // *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, 2019: 5099 – 5108.
- [24] GAO J, WANG Q, YUAN Y. SCAR: spatial-/channel-wise attention regression networks for crowd counting[J]. *Neurocomputing*, 2019, 363: 1 – 8.
- [25] MA Z, WEI X, HONG X, et al. Bayesian loss for crowd count estimation with point supervision[C] // *IEEE/CVF Conference on Computer Vision*, IEEE, 2019: 6142 – 6151.
- [26] WANG B, LIU H, SAMARAS D, et al. Distribution matching for crowd counting[C] // *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2020, 33: 1595 – 1607.
- [27] SONG Q, WANG C, JIANG Z, et al. Rethinking counting and localization in crowds: a purely point-based framework[C] // *IEEE/CVF Conference on Computer Vision*, IEEE, 2021: 3365 – 3374.
- [28] WANG Q, GAO J, LIN W, et al. NWPU-crowd: a large-scale benchmark for crowd counting and localization[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 43(6): 2141 – 2149.
- [29] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. *arXiv preprint arXiv:1409.1556*, 2014.
- [30] SHI Z, METTES P, SNOEK C G. Counting with focus for free[C] // *IEEE/CVF Conference on Computer Vision*, IEEE, 2019: 4200 – 4209.
- [31] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C] // *IEEE/CVF Conference on Computer Vision*, IEEE, 2017: 2980 – 2988.