



DOI:10.12404/j.issn.1671-1815.2401029

引用格式:关昆仑,朱思文,张仰森,等.基于深度学习的生物资产检测模型 YOLOSC[J].科学技术与工程,2025,25(2):674-682.

Guan Kunlun, Zhu Siwen, Zhang Yangsen, et al. The biological asset detection model YOLOSC based on deep learning[J]. Science Technology and Engineering, 2025, 25(2): 674-682.

基于深度学习的生物资产检测模型 YOLOSC

关昆仑,朱思文,张仰森*,成琪昊,张学凯

(北京信息科技大学智能信息处理研究所,北京 100192)

摘要 为提高生物资产监盘审计过程中盘点准确性和盘点效率,提出了一种融入注意力机制和损失函数优化的生物资产检测模型 YOLOSC。首先,将压缩-激励网络(squeeze-and-excitation networks, SENet)注意力机制引入 YOLOv5s 模型的主干网络中,以增强对生物资产图片中关键特征的提取能力;其次,采用完全交并比(complete intersection over union, CIoU)作为检测框回归的损失函数,以提升训练过程中检测框的回归速度与定位精度;最后,构建了一个生物资产数据集对所提模型进行针对性训练,以提升模型检测效果。实验结果表明,相较于 YOLOv5 模型, YOLOSC 的精确率、召回率、 F_1 和 AP 分别提升了 2.3%、2.1%、2.7% 和 1.6%,证明了所提出的生物资产检测模型 YOLOSC 的有效性。

关键词 目标检测模型; YOLOv5; 注意力机制; 损失函数; 生物资产审计

中图分类号 TP391.4; **文献标志码** A

The Biological Asset Detection Model YOLOSC Based on Deep Learning

GUAN Kun-lun, ZHU Si-wen, ZHANG Yang-sen*, CHENG Qi-hao, ZHANG Xue-kai

(Institute of Intelligent Information Processing, Beijing Information Science and Technology University, Beijing 100192, China)

[Abstract] In order to improve the accuracy and efficiency of inventory counting in the process of monitoring and auditing biological assets, a biological asset detection model YOLOSC incorporating the attention mechanism and loss function optimization was proposed. Firstly, the SENet attention mechanism was introduced into the backbone network of the YOLOv5s model to enhance the ability of extracting the key features in the pictures of the biological assets. Secondly, the CIoU was adopted as the regression of the detection frames with the loss function to enhance the regression speed and localization accuracy of the detection frame during the training process. Finally, a biological asset datasets was constructed for targeted training of the proposed model to enhance the model detection effect. The experimental results show that compared with the YOLOv5 model, the precision, recall, F_1 value and AP of YOLOSC are improved by 2.3%, 2.1%, 2.7% and 1.6%, respectively, which proves the effectiveness of the proposed biological asset detection model YOLOSC.

[Keywords] target detection model; YOLOv5; attention mechanism; loss function; audit of biological assets

生物资产是一些企业资产的重要组成部分,如养殖企业中就有大量的生物资产。生物资产具有多样性、隐蔽性和价值动态性等特点,同时其生存的环境则具有一定的恶劣性、复杂性和危险性,例如,在陡峭山坡上吃草的羊群或在灯光黑暗的羊圈里的羊群,生物资产的这些特点使得在对企业的生物资产进行审计时存在着计量与确认不准确的风险、被盗与丢失的风险、人身安全风险、数据被篡改的风险,从而在对相关企业进行生物资产存货审计时给审计人员带来了一些困难和挑战^[1]。生物资产的多样性要求审计人员了解和掌握不同种类生

物资产的特点;其隐蔽性使得审计人员难以准确计量和确认生物资产;生物资产的价值波动性增加了审计人员在存货计量和确认过程中的不确定性^[2]。

为了降低生物资产存货审计过程中的各种风险并提高审计效果,引入生物目标检测技术作为解决方案显得尤为重要。生物目标检测技术是指利用计算机视觉和机器学习方法,自动检测和识别存货中的生物资产,减少人工操作的主观性和错误性。通过目标检测技术,审计人员能够实时、准确地进行存货监盘和资产审计,尤其是那些处于危险环境、恶劣环境或复杂环境中的生物资产审计,审

收稿日期:2024-02-07; 修订日期:2024-10-23

基金项目:北京市社会科学基金重点项目(21GLA007)

第一作者:关昆仑(1996—),男,满族,河北承德人,硕士研究生。研究方向:计算机视觉、信息安全。E-mail:guankunlun@outlook.com。

*通信作者:张仰森(1962—),男,汉族,山西临猗人,博士,教授。研究方向:人工智能、自然语言处理。E-mail:Zhangyangsen@163.com。

计人员不用很接近被审计对象,就可以对生物资产进行监盘审计,提高计量和确认的准确性。此外,目标检测技术还可以降低被盗和损失风险,提升存货的安全性。

针对生物资产审计的特点和风险,现提出一种基于YOLOv5的生物资产检测模型YOLOSC。该模型借鉴YOLOv5的架构,并将压缩-激励网络(squeeze-and-excitation networks, SENet)注意力机制融入主干网络,以增强对生物资产图片中关键特征的提取能力。同时,采用完全交并比(complete intersection over union, CIoU)作为检测框回归的损失函数,以提升训练过程中的检测框回归速度和定位精度。构建一套面向生物资产审计的训练数据集,以更好地适应生物资产审计需求。

1 目标检测相关研究

目标检测是人工智能领域中的重要研究内容,在安全监控、智能交通、车辆检测、大地遥感等领域都有广泛的应用。目标检测算法主要分为两阶段和单阶段目标检测算法。两阶段目标检测算法通常先对图像中的物体进行位置定位,再对定位到的目标进行分类;而单阶段目标检测算法直接对目标检测框进行回归和分类。两阶段目标检测算法主要有区域卷积神经网络(regions with CNN features, RCNN)、快速区域卷积神经网络(fast Rich features for object detection, Fast-RCNN)和掩码区域卷积神经网络(mask regional convolutional neural networks for instance-level segmentation, Mask R-CNN)等,这些方法具有较高的精确度,但它们的检测速度不能满足实时性的需求。单阶段目标检测算法主要以YOLO和单阶段检测器(single shot multiBox detector, SSD)为代表,具有检测速度快、能轻松地实现实时运行的特点^[3-4]。尤其是YOLO系列算法,其核心思想是把目标检测转换成一个回归问题。YOLOv1算法其结构的全连接层部分由两个全连接层组成,用来预测目标的位置和类别概率值,对YOLOv1的网络结构等进行改进并引入Anchors box机制,提出的YOLOv2算法检测速度和精度均有提高。YOLOv3由于特征提取部分的改进,性能进一步提高;YOLOv4加入了一些实用技巧,实现了检测速度与精度的最佳平衡。YOLOv5是在YOLOv4的基础上,对input、backbone、neck和head 4个模块都进行了一些新的改进,使其速度与精度都得到了极大的性能提升,它有4个版本,能够更好地权衡检测速度和精度,方便用户选择。YOLOv5采纳了交叉阶段部分暗网53(cross stage partial darknet53, CSPDark-

net53)作为其骨干网络架构,并集成了空间金字塔池化层(spatial pyramid pooling, SPP)层以增强特征提取能力,同时采用了路径聚合网络(path aggregation network, PAN)作为其特征融合的纽带^[5]。此外,它配备了先进的检测头^[6],专门针对实时检测任务进行了优化,以确保快速响应,而不是单纯追求最高的检测精度。YOLOv5与YOLOv8相比,两个算法在骨干网络架构、Anchor Boxes、非极大值抑制(non-maximum suppression, NMS)、后处理、优化器、激活函数等方面都比较相似,但在处理速度(frame per second, FPS)方面,YOLOv5在CPU上的处理速度高于YOLOv8,即相比之下,YOLOv5在不安装GPU的普通计算机上具有更快的训练速度和检测速度,且已在中外的多个领域得到广泛的应用和推广。

由于YOLO系列算法发展较快,近年来,中外学者对各版本的算法都有不少的研究和应用。周奥等^[7]提出了一种基于改进YOLOv3模型的虹膜快速定位方法。针对眼周图像中虹膜内、外圆尺寸变化不大,将YOLOv3网络的多尺度结构改进为双尺度检测。陈亚晨等^[8]提出一种基于改进YOLOv3的眼机交互模型,提高了网络像素特征提取细粒度并加快了检测速度。覃紫馨等^[9]对YOLOv4算法进行了改进,使其更适用于探地雷达异常目标的自动检测。柏乐^[10]在YOLOv5的主干网络和颈部网络中分别引入轻量级的幽灵网络(ghost network, ghost-Net)和自适应感受野的选择性卷积(selective kernel convolution, SK),以提升对海洋底栖生物检测的速度与精度。陈宇梁等^[11]在YOLOv5的基础上,提出了快速空间金字塔池化模块和基于旋转窗口的Swin-Transformer模块,提升了模型在弱光下的检测能力。葛云飞等^[12]利用改进的YOLOv5构建了一个轻量级口罩人脸检测模型。他们将注意力机制嵌入到YOLOv5的主干网络中,并改进了边框回归损失函数,以提高检测速度和精度,用于检测人脸口罩。魏鑫宇等^[13]在SSD算法基础上引入同步特征提取结构,同时提出相邻三层融合机制,以增强对中低层特征图中的语义信息提取,提升高速路面资产检测效率。王美华等^[14]利用改进的YOLOv5s中对病理学图像中猪只小肠绒毛进行检测。李跃华等^[15]利用改进的YOLOv5s对着装不规范检测算法进行研究,他们改进特征金字塔结构(feature pyramid networks, FPN),采用跨层级联的方式融合更多的特征,一定程度上防止了浅层语义信息的丢失。

尽管研究人员在YOLO算法的应用上开展了许多研究,但在生物审计领域应用研究还较少。在生

物资的实际存货监盘审计中,系统应具备实时快速响应的特点,且系统设备不要求安装 GPU 卡,因此,最新提出的 YOLOv8 算法由于需要 GPU 环境而不符合实际工作的要求。

由于 YOLOv5 在 CPU 环境下的性能表现很好,因此,现提出一种基于 YOLOv5 的生物资产检测模型 YOLOSC,旨在实现审计过程的智能化,提高审计工作效率的,同时确保审计人员的安全。该模型基于 YOLOv5s 的架构进行改进,将 SENet 注意力机制融入主干网络中,增强模型对黑暗、模糊或光线不足等复杂场景下生物资产的检测能力。同时,将 CIoU 损失函数应用于边界框回归,增强模型在复杂遮挡场景下目标的检测效果,降低漏检概率。这一研究为解决生物资产审计问题提供了新的技术参考和智能化方案。

2 网络结构的选择

前面已经指出,在对目标进行检测,尤其是对小物体目标进行检测时, YOLOv5 和 YOLOv8 都是有效的选择。但是,如果要在普通的 CPU 设备上部署解决方案, YOLOv5 可能更合适。它具有推理速度快、检测精度高的优势,同时还具有权重比较小的优势。这使得 YOLOv5 适合于部署到嵌入式设备进行实时检测,能够满足生物资产审计过程中对检测系统准确性、实时性和轻量级的要求,具备在生物资产审计过程中应用的潜力。因此,选择 YOLOv5 作为基础网络架构,开展监盘存货审计辅

助检测系统的研究。该网络架构如图 1 所示。

YOLOv5 的架构包括 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 4 种变体^[16]。这 4 种架构的模型大小和参数逐渐增加,从而提升了特征提取能力,但同时会降低检测速度。在生物资产审计过程中,要求快速高效,因此实时性和轻量级特性是两个最重要的考虑因素。综合考虑检测准确性、效率和模型大小等多个方面因素,本文选择 YOLOv5 中参数规模最小的 YOLOv5s 网络架构作为基础^[17]。

YOLOv5 的核心思想是将整个图像输入到类似于 RCNN 的更快网络中,并直接回归到边界框的位置和类别输出层^[18]。该模型由 4 个部分组成:输入端、主干网络、颈部网络和输出端。输入端包括 Mosaic 数据增强、自适应锚定框和自适应图片缩放。主干网络包括 Focus 层和 C3 模块。颈部网络采用 FPN + PAN 结构。输出端包括边界框的损失函数和 NMS 非最大抑制。

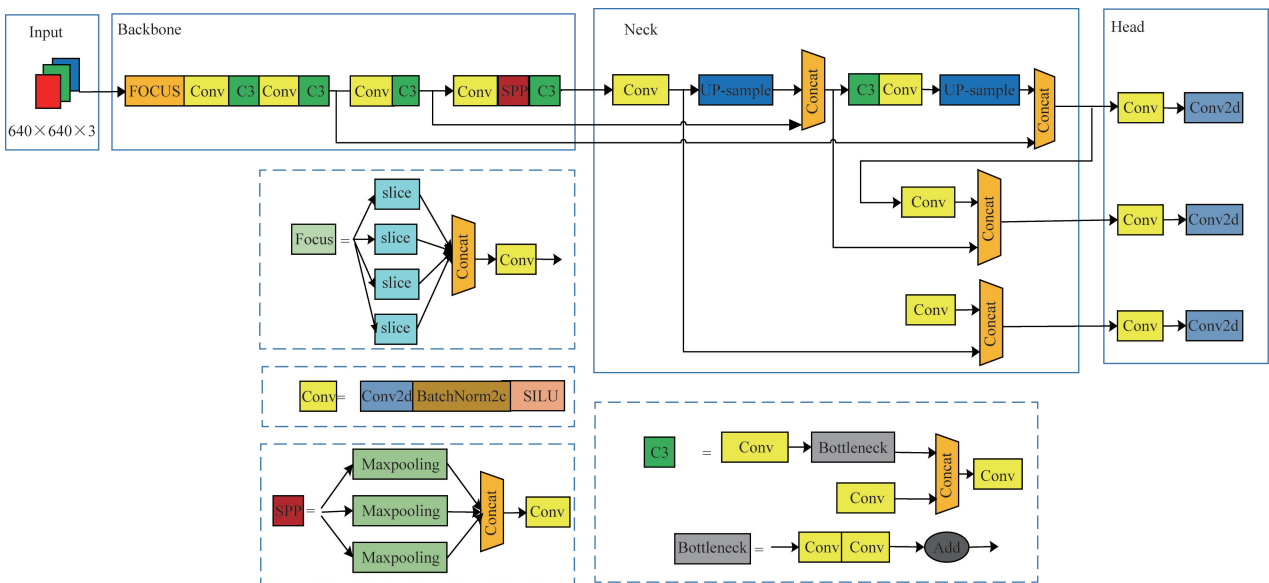
2.1 输入端

2.1.1 Mosaic 数据增强

Mosaic 数据增强是一种针对包含 4 个输入图片的数据集的技术,它通过在管道内进行随机缩放、随机裁剪和随机排列等操作,旨在提高模型的鲁棒性和准确性,同时减少了整个检测过程对 CPU 的负荷。

2.1.2 自适应锚定框

YOLOv5 可针对各种目标的不同长宽比输出预测框。然后,将预测框与标注的真实框进行比较,计算它们之间的差异,并通过反向传播更新网络参



Input 为输入端;Backbone 为主干网络;Neck 为颈部结构;Head 为头部结构;Focus 为焦点层;Conv 为卷积层;C3 为网络模块;UP-sample 为上采样层;Concat 为拼接操作;Conv2d 为 2D 卷积层;SIUD 为数据归一化层;Bottleneck 为瓶颈结构;Add 为加法操作

图 1 YOLOv5s 网络结构图

Fig. 1 YOLOv5s network architecture diagram

数,以适应缺陷区域的形状。

2.1.3 自适应图片缩放

在该网络中,针对具有不同长宽比的数据集,采用了统一的图像缩放策略,将输入图片调整到相同的尺寸。为了实现这一目标,其使用了最少量的黑色背景填充。这样做旨在确保模型对于不同尺寸的输入数据具有一致的处理能力,以提高模型的泛化性能。

2.2 主干层

2.2.1 Focus 层

首先,通过对输入图像按照间隔像素进行切片,获得4幅大小减半的图像。在进入骨干网络之前,在Focus层中进行了额外的处理,将宽度和高度空间的信息集中在通道空间中。随后,进行进一步的卷积操作,生成更高维度的特征图。在YOLOv5架构中,YOLOv5s包含了32个卷积核,是4种架构中卷积核数量最多的一种。这就使得YOLOv5s在计算资源有限的设备上表现最佳、检测速度最快,也是本文研究中选择它作为网络架构的主要考虑^[19]。

2.2.2 C3 模块

YOLOv5在设计中汲取了跨阶段局部网络(cross-stage partial network, CSPNet)的跨阶段合并策略,将其融入C3模块的构建,C3模块以并行处理方式处理输入特征:一方面通过多个Bottleneck层和标准卷积(convolutional layer-batch normalization layer - LeakyReLU, CBL)堆叠处理特征,而另一方面仅通过一个基本卷积模块处理,最终通过Concat操作将这两部分特征合并,实现在保持精度的同时降低计算次数。具备残差结构的C3模块有效地增强了梯度信息在反向传播中的传递,缓解了Backbone网络的深度加深可能导致的梯度消失和梯度重复问题。因此,C3网络作为YOLOv5主干网络的结构具有增强网络学习能力的优势,从而降低了计算瓶颈和内存成本^[20]。

2.3 颈部网络

YOLOv5的颈部采用了FPN+PAN结构,其中FPN层传递了来自特征金字塔顶部的强语义特征,而PAN层则传递了来自底部的强定位特征。通过在不同检测层上进行特征聚合,提升了骨干网络的特征提取能力。这种设计使得模型能够更好地学习密集目标的多级特征,从而提高其在复杂密集场景中的应对能力。

2.4 输出端

2.4.1 边界框的损失函数

YOLOv5模型采用广义交叉比损失(generalized

intersection over union loss, GIoU Loss)作为边界框的损失函数,代替了原本的交叉比损失(intersection over union loss, IoU_Loss)损失函数。这一改进解决了当真实框与预测框没有交集时,导致的交并比为0、损失函数导数为0,无法解决真实框与预测框重叠的问题。

2.4.2 NMS 非极大值抑制

在检测并分类缺陷区域后,采用NMS对检测框进行筛选,以解决重叠检测框的问题。该方法首先选择一个概率最大的框,并将其添加到最终结果中,然后计算其他框与该最大框的交并比IoU。如果某个框的IoU值较大,则将其排除。重复上述步骤,直到每个对象只保留一个相应的检测框。在YOLOv5中,引入了距离交叉比(distance intersection over union, DIoU)来量化重叠部分,并使用DIoU-NMS来筛选检测框,从而提高重叠检测缺陷区域的成功率。

3 融入注意力机制和损失函数优化的模型 YOLOSC

3.1 在主干层融入 SENet 注意力机制

注意力机制源于人类视觉研究,人们在处理信息时会有选择性地关注重要信息,而忽略其他可见信息,以应对信息处理的限制。注意力机制的基本原理是允许神经网络通过选择性地关注输入的特定部分来提取特征子集。注意力机制在资源有限的情况下解决信息过载问题,将计算资源分配给更重要的任务,从而提高模型的计算效率。注意力机制可以应用于不同类型的特征输入中。在目标检测领域的各种实验研究中,不同的注意力机制被引入到模型的不同位置,以提高模型的识别和检测性能。

在生物资产审计中,由于生物资产的形状、颜色和背景各不相同,不可避免地存在遮挡重叠、目标模糊和环境黑暗等问题,这导致了YOLOv5在生物资产检测中的目标定位精度下降。为了提高审计过程中存货监盘的目标检测精度,本文研究将SENet网络结构引入存货检测审计模型的主干网络中,在通道维度上引入注意力机制,以提高模型的精确度和泛化能力,更好地提取各种生物资产目标的特征。

SENet模块具有计算量小且能提高模型学习能力的特点,它通过特征重标定方案自学习每个特征通道的重要性权重,并利用这些权重增强重要特征并抑制不重要特征。SENet模块由3个关键操作组成:压缩(squeeze)、激励(excitation)和缩放(scale)

操作,如图 2 所示。首先,给定一个尺寸为 $H' \times W' \times C'$ 的输入特征 X ,将 X 输入主干网络进行卷积操作,得到一个尺寸为 $H \times W \times C$ 的特征 U 。然后,利用 $F_{sq}(\cdot)$ 函数进行压缩操作,具体是通过全连接层搭配 ReLU 激活函数的方式将特征图 U 聚合压缩成维度为 $1 \times 1 \times C$ 的特征图,实现全局感受野,得到每个通道的全局特征。接下来,利用 $F_{ex}(\cdot, W)$ 函数进行激励操作,具体是通过全连接层通过非线性操作拟合通道间的相关性,并使用 Sigmoid 函数得到每个通道的权重值。最后,通过缩放操作(重标定操作),即使用 $F_{scale}(\cdot, \cdot)$ 函数根据通道权重值逐通道加到输入特征图 U 上,最终得到通道特征加权的特征图 \tilde{X} 。这样的设计和引入 SENet 模块将有助于提升模型的检测精度,改善 YOLOv5s 的检测性能。

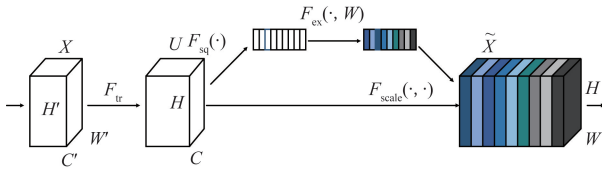


图 2 SENet 结构图

Fig. 2 SENet structure diagram

在主干网络的每个 C3 卷积模块之后嵌入了 SENet 通道注意力模块。嵌入 SENet 的 YOLOV5 局部网络结构如图 3 所示。

3.2 YOLOv5s 边界框定位损失函数优化

IoU 损失函数是目标检测领域常用的评估模型预测值和实际值之间差距的方法。在 YOLOv5 网络结构中,使用了 GIoU 损失定位函数来解决边界框不重叠的问题。然而,在目标检测过程中,当预测框和真实框固定不变且预测框位于真实框内部时,GIoU 无法准确反映两个框的包含关系,并且缺乏进一步的优化操作。为解决该问题,提出了使用 CIoU 损失函数代替 GIoU 损失函数。CIoU 损失函数不仅考虑了重叠

区域和中心点的距离,还考虑了边界框的长宽比,弥补了 GIoU 的不足之处。通常,CIoU 损失函数为

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

$$\alpha = \frac{v}{(1 - \text{IoU}) + v} \quad (2)$$

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

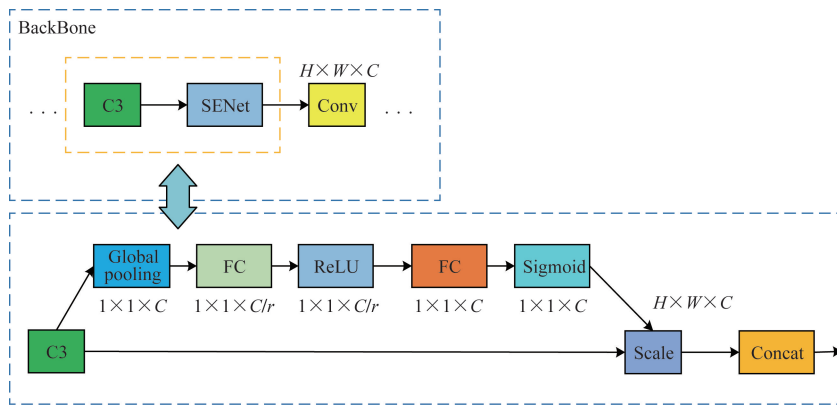
式中: $\text{Loss}_{\text{CIoU}}$ 为 CIoU 损失函数的损失值; $\rho^2(b, b^{gt})/c^2$ 为预测框中心与真实框之间的欧氏距离; b 和 b^{gt} 分别为预测框和真实框的中心点; c 为最小外接框对角线的距离; α 为平衡长宽比例对损失贡献度的系数; IoU 为损失的平衡因子; v 为衡量两框的长宽比; h, w 为预测框的高和宽; h^{gt} 和 w^{gt} 为真实框的高和宽。

根据式(1),当预测框与真实框存在重叠时,CIoU 的值为 IoU 除以一个调整系数,该调整系数综合考虑了中心点距离和长宽比的影响,因此能更好地匹配真实框。而当预测框与真实框没有重叠时,CIoU 的值为 -1,能更迅速地将预测框拉回。此外,通过引入平衡长宽比例对损失贡献度的系数 α ,CIoU 还能防止过大的预测框出现,进一步改进了 YOLOv5s 检测精度。

4 实验部分

4.1 实验环境与相关参数

模型训练使用的硬件环境: Intel i5-7 代/4 核 CPU、12 GB 内存和 NVIDIA GTX 1080 Ti GPU。软件环境: python 3.8 开发语言、Windows10 操作系统、CUDA10.1 和 PyTorch 深度学习框架。在训练之前,使用的预训练权重文件进行模型初始化,预训练权重是在 COCO 数据上训练得到的。具体训练参数设置如表 1 所示。



C 为输入特征图的通道数; r 为缩小比例

图 3 YOLOv5 中嵌入 SENet 结构图

Fig. 3 Embedded SENet structure diagram in YOLOv5

表 1 训练参数设置

Table 1 Training parameter settings

参数	数值
输入图像规格	640 × 640
训练轮次	300
初始学习率	0.01
动量参数	0.921
权重衰减系数	0.00038
优化器	SSD

4.2 数据采集与处理

为了进行实验,构建了一个包含常见生物资产的审计数据集,并采用了现实拍摄和网络爬虫相结合的方法进行数据收集。采集到的部分数据集如图 4 所示。

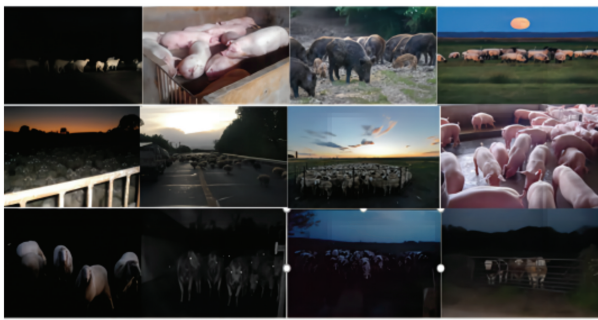


图 4 采集数据样例

Fig. 4 Sample data collection

现实拍摄部分选择了生物资产企业实际场景进行拍摄,在不同地点连续一段时间进行监测,以获取包含复杂背景信息(如光照条件、颜色背景、季节背景)的图片和视频。对于视频数据,按照 2 s 的时间间隔进行采样,然后筛选并保留有价值的图片,去除相似度高和分辨率低的图片。另外,采用网络爬虫技术从生物资产企业的官方网站爬取各种类型的图片和视频,并对视频进行截取,以完善数据集的收集。

鉴于数据集包含多种类别,选择了生物资产企业中常见的 36 种生物资产作为检测对象,其中包括牛、羊和猪等。同时使用图像标注软件 LabelImg,手动框选图像中的目标,并标注目标的类别。对应每个目标,保存了包含目标框位置和长宽信息的 txt 文件。数据集的标注效果如图 5 所示。

最终获得了包含 36 种生物资产的图片数据集,共计 7 356 张图片。按照 7:2:1 的比例,将数据集划分为训练集、验证集和测试集,其中训练集包含 5 150 张图片,验证集包含 1 471 张图片,测试集包含 735 张图片。

4.3 评价指标

使用 4 个指标来评估模型性能:精确率

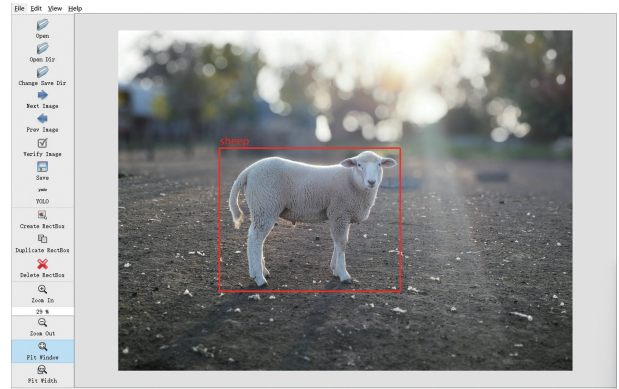


图 5 数据集标注

Fig. 5 Labeling of data sets

(P_{recision})、召回率(R_{ecall})、平均精度(average precision, AP)和 F_1 。精确率表示正确预测为正的样本数量占所有预测为正的样本数量的比例,召回率表示正确预测为正的样本数量占所有实际为正的样本数量的比例。平均精度是为了综合考虑精确率和召回率而计算的指标,它代表了模型的综合性能。平均精度是针对多类别检测任务的精度评估指标,通过计算所有类别的检测精度的平均值得到。这些评价指标的公式为

$$P_{\text{recision}} = \frac{TP}{TP + FP} \quad (4)$$

$$R_{\text{ecall}} = \frac{TP}{TP + FN} \quad (5)$$

$$F_1 = \frac{2TP}{2TP + FP + FN} \quad (6)$$

$$AP = \int_0^1 P_{\text{recision}}(R_{\text{ecall}}) dR_{\text{ecall}} \quad (7)$$

式中:TP 为正确预测为正的样本数量;FP 为错误预测为正的样本数量;FN 为错误预测为负的样本数量; n 为样本类型的数量; $P_{\text{recision}}(R_{\text{ecall}})$ 为在不同召回率下的精确率。

4.4 YOLOSC 模型与其他目标检测模型对比

为评估 YOLOSC 模型在生物资产检测中的性能优势,本文模型与 YOLOv4、Faster R-CNN 和 SSD 3 个热门目标检测模型进行了对比实验。实验使用 4.3 节中的评价指标对模型性能进行评价,并引入了帧率(FPS)来评估各个模型的检测速度。实验结果如表 2 所示。

从表 2 可得,YOLOSC 模型在精确度、召回率、 F_1 和平均精度方面均优于对比模型。尽管在检测速度方面稍低于 SSD 模型,但仅相差 3 帧/s 的 FPS 值。综合考虑所有性能指标,可以看出,本文提出的 YOLOSC 模型在审计生物资产检测任务中具有明显优势,优于上述对比模型。

表2 YOLOSC 与其他模型实验对比结果

Table 2 Experimental comparison results of YOLOSC with other models

评价指标	YOLOv4	Faster R-CNN	SSD	YOLOSC
$P_{recision}$	0.896	0.911	0.905	0.924
R_{ecall}	0.819	0.829	0.821	0.846
F_1	0.856	0.868	0.861	0.883
AP	0.895	0.901	0.899	0.913
FPS	71.2	63.5	49.3	52.3

表3 消融实验结果

Table 3 Results of ablation experiments

评价指标	YOLOv5	YOLOv5 + SENet	YOLOv5 + CIoU	YOLOSC
$P_{recision}$	0.901	0.910	0.908	0.924
R_{ecall}	0.825	0.836	0.833	0.846
F_1	0.856	0.871	0.869	0.883
AP	0.897	0.905	0.906	0.913

4.5 实验结果与分析

通过实验将 YOLOSC 模型与原始 YOLOv5 模型进行对比,对其进行全面综合性能评估。训练过程中各项指标的变化曲线如图 6 所示。

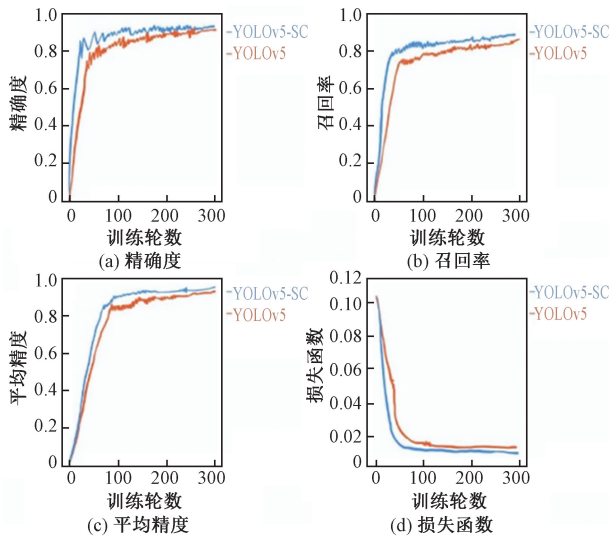


图6 指标变化曲线

Fig. 6 Indicator change curves

随着训练轮数的增加, YOLOSC 模型的指标值逐渐提升。当训练至约 300 轮时,其综合性能指标值已趋于稳定。YOLOSC 相较于 YOLOv5,具有更快的收敛速度和更低的损失值。在 300 轮时, YOLOSC 的损失值达到 0.01,而 YOLOv5 的损失值达到 0.012,之后均未发生变化。精确率、召回率和平均精度相较于 YOLOv5 也有提升,精确度达到 0.924,召回率达到 0.846,平均精度达到 0.913, F_1 达到 0.883。

上述结果表明,通过 SENet 注意力机制和 CIoU 损失函数的引入, YOLOSC 模型在模型性能上取得了显著改进。

4.6 消融实验

为探究改进的模块对模型带来的提升效果,进行了一组消融实验。实验结果如表 3 所示。

通过上述消融实验结果可得以下结论:首先,对比 YOLOv5 模型, YOLOv5 + SENet 在所有评价指标上都取得了一定的提升。这表明引入 SENet 通道注意力机制有助于提升模型的精确性和召回率,并在综合性能上有所改善。

其次, YOLOv5 + CIoU 模型在精确率、召回率、 F_1 和平均精度等指标上都相对于 YOLOv5 有所提高。这显示出 CIoU 损失函数的应用对于定位框的准确性有积极影响。

最后, YOLOSC 模型在所有评价指标上都达到了最佳结果。消融实验结果表明,引入 SENet 通道注意力机制和采用 CIoU 损失函数对 YOLOv5 模型的性能有明显的正向影响。YOLOSC 模型在审计生物资产检测任务中表现出较大优势,相较于其他模型具有更高的精确性和召回率以及更好的综合性能。

4.7 实验结果实例

为直观地观察 YOLOSC 模型相较于原始 YOLOv5 模型在真实环境下的检测效果,在测试集中随机选取了一些存在遮挡或光线不佳情况的生物资产图片进行检测,并将两种场景下的检测结果如图 7 和图 8 所示。

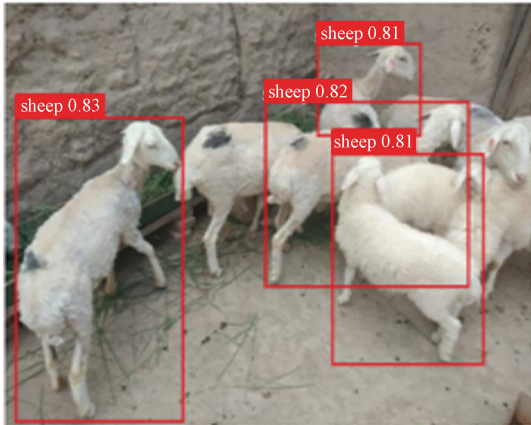
图 7 展示了在遮挡场景下的检测效果对比。从图 7 可以明显看出,改进后的 YOLOSC 模型能够准确地对图片中生物资产进行检测,而原始的 YOLOv5 模型在遮挡情况下则出现了对生物资产目标检测不全面的情况。

图 8 展示了在模糊场景下的检测效果对比。当环境较为黑暗或光照不佳时,一些目标可能会出现模糊的情况。然而改进后的 YOLOSC 模型仍然能够有效地进行检测,而 YOLOv5 模型则可能会遗漏一些生物资产目标。

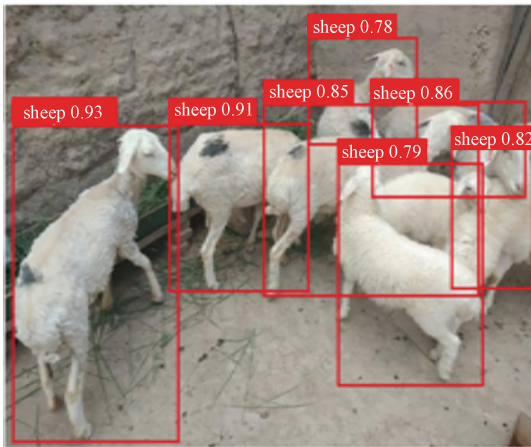
通过对比图 7 和图 8,可以明显看出 YOLOSC 模型相对于 YOLOv5 模型在遮挡和模糊场景下的检测效果有显著的改进,进一步验证了 YOLOSC 模型在审计生物资产检测任务中的优越性和鲁棒性。



(a) 输入图片



(b) YOLOv5检测结果



(c) YOLOv5-SC检测结果

图7 遮挡场景下实验效果对比图

Fig. 7 Comparison chart of experimental effects in occluded scenes

5 结论

为了推动审计行业向智能化、自动化的方向发展,特别是在存货监盘过程中,对生物资产进行快速而准确的盘点是一项关键任务。传统的盘点方法往往依赖于人工操作,不仅耗时耗力,而且容易受到主观因素的影响,导致盘点结果的准确性和一致性难以保证。因此,提出了一种基于深度学习的



(a) 输入图片 (b) YOLOv5检测结果 (c) YOLOv5-SC检测结果

图8 模糊场景下实验效果对比图

Fig. 8 Comparison chart of experimental effects in blurred scenes

目标检测模型——YOLOSC。它是北京市社会科学基金重点项目“审计过程数字化路径与方法研究”的成果,被北京市审计局大兴分局应用于某些具有生物资产的企业监盘审计中。

YOLOSC 模型是在 OLOv5 的基础上进行了改进。主要的改进点如下。

(1)引入了 SENet 注意力机制。这种机制能够使模型更加关注图片中的关键特征,从而在复杂的视觉环境中,如遮挡和模糊情况下,更好地识别和定位生物资产。

(2)采用了 CIoU 损失函数。这是一种用于衡量目标检测模型预测边界框与真实边界框之间重叠程度的损失函数,它能够提供更精确的定位性能,从而提高模型的整体检测能力。

为了验证 YOLOSC 模型的有效性,在自制的生物资产数据集上进行一系列实验,将 YOLOSC 模型与其他几种主流的目标检测模型进行了对比。实验结果显示,在精确率、召回率、 F_1 以及平均精度等性能指标上,YOLOSC 模型都展现出了明显的优势。这些结果表明,通过引入 SENet 注意力机制和采用 CIoU 损失函数,提升了 YOLOv5 模型在处理遮挡、模糊等复杂场景下的目标检测性能。也就是说,YOLOSC 模型在面向生物资产审计时,对审计人员较高的实用价值,有助于推动审计行业的技术进步和智能化发展。

尽管本文模型具有较好的目标检测性能,但仍存在潜在的改进空间。例如,可以进一步优化 SENet 注意力机制的设计,探索更有效的特征融合策略,进一步提升模型性能。也可考虑引入其他先进的目标检测技术和网络架构,以提高检测精度和速度。同时,扩充生物资产数据集,并考虑增加数据增强操作,以进一步提升模型在生物资产检测中的精度,为生物资产审计提供更好的智能化工具。

参 考 文 献

- [1] 赵春燕. 存货监盘审计程序存在的风险及对策[J]. 国际商务财会, 2021(16): 79-81.
Zhao Chunyan. Risks and countermeasures of inventory monitoring audit procedures [J]. International Business Accounting, 2021 (16): 79-81.
- [2] 刘俊彤. 生物资产审计风险及对策[J]. 合作经济与科技, 2022(12): 160-162.
Liu Juntong. Biological asset audit risks and countermeasures[J]. Cooperative Economy and Science and Technology, 2022 (12): 160-162.
- [3] Rahman S, Rony J H, Uddin J, et al. Real-time obstacle detection with YOLOv8 in a WSN using UAV aerial photography[J]. Journal of Imaging, 2023, 9(10): 216-233.
- [4] 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. New York: IEEE, 2023: 7464-7475.
- [5] Zhu X, Lü S, Wang X, et al. TPH-YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios[C]//Proceeding of the IEEE/CVF International Conference on Computer Vision. New York: IEEE, 2021: 2778-2788.
- [6] Bharat M, Mishra K K. Small-object detection based on YOLOv5 in autonomous driving systems [J]. Pattern Recognition Letters, 2023, 168: 115-122.
- [7] 周奥, 杨岗, 闫磊, 等. 基于改进 YOLO 的虹膜快速定位检测算法[J]. 科学技术与工程, 2022, 22(33): 14801-14808.
Zhou Ao, Yang Gang, Yan Lei, et al. Iris fast location detection algorithm based on improved YOLO[J]. Science Technology and Engineering, 2022, 22(33): 14801-14808.
- [8] 陈亚晨, 韩伟, 白雪剑, 等. 基于改进 YOLO-v3 的眼机交互模型研究及实现 [J]. 科学技术与工程, 2021, 21 (3): 1084-1090.
Chen Yachen, Han Wei, Bai Xuejian, et al. Research and implementation of eye computer interaction model based on improved YOLO-v3[J]. Science Technology and Engineering, 2021, 21 (3): 1084-1090.
- [9] 覃紫馨, 姜彦南, 徐立, 等. 基于 YOLO 算法的探地雷达道路图像异常自动检测 [J]. 科学技术与工程, 2023, 23 (27): 11505-11512.
Qin Zixin, Jiang Yannan, Xu Li, et al. Automatic detection of anomalies in GPR images based on YOLO algorithm [J]. Science Technology and Engineering, 2023, 23(27): 11505-11512.
- [10] 柏乐. 基于深度学习的海洋底栖生物检测[D]. 大连: 大连海事大学, 2022.
Bai Le. Deep learning-based detection of marine benthic organisms [D]. Dalian: Dalian Maritime University, 2022.
- [11] 陈宇梁, 董绍江, 孙世政, 等. 改进 YOLOv5s 的弱光水下生物目标检测算法[J]. 北京航空航天大学学报, 2024, 50(2): 499-507.
Chen Yuliang, Dong Shaojiang, Sun Shizheng, et al. Improvement of YOLOv5s for weak light underwater biological target detection algorithm[J]. Journal of Beihang University, 2024, 50 (2): 499-507.
- [12] 葛云飞, 祁云嵩, 孟祥宇. YOLOv5 改进的轻量级口罩人脸检测[J]. 计算机系统应用, 2023, 32(3): 195-201.
Ge Yunfei, Qi Yunsong, Meng Xiangyu. YOLOv5 improved lightweight mouthpiece face detection[J]. Computer Systems & Applications, 2023, 32(3): 195-201.
- [13] 魏鑫宇, 王池社. 基于改进 SSD 的高速路面资产检测算法研究[J]. 信息与电脑(理论版), 2022, 34(4): 101-104.
Wei Xinyu, Wang Chishe. Research on high speed pavement asset detection algorithm based on improved SSD [J]. Information & Computer, 2022, 34(4): 101-104.
- [14] 王美华, 王安邦, 肖德琴, 等. 改进 YOLOv5s 对病理学图像中猪只小肠绒毛的检测[J]. 农业工程学报, 2024, 40(5): 207-215.
Wang Meihua, Wang Anbang, Xiao Deqin, et al. Improving YOLOv5s for detecting small intestinal villi in pathological images of pigs[J]. Journal of Agricultural Engineering, 2024, 40(5): 207-215.
- [15] 李跃华, 仲新, 姚章燕, 等. 基于改进 YOLOv5s 的着装不规范检测算法研究[J]. 图学学报, 2024, 45(3): 433-445.
Li Yuehua, Zhong Xin, Yao Zhangyan, et al. Research on dress irregularity detection algorithm based on improved YOLOv5s [J]. Journal of Graphics, 2024, 45(3): 433-445.
- [16] 兰玉彬, 孙斌书, 张乐春, 等. 基于改进 YOLOv5s 的自然场景下生姜叶片病虫害识别[J]. 农业工程学报, 2024, 40(1): 210-216.
Lan Yubin, Sun Binshu, Zhang Lechun, et al. Identification of ginger leaf diseases and pests in natural scenes based on improved YOLOv5s [J]. Journal of Agricultural Engineering, 2024, 40 (1): 210-216.
- [17] 何月, 王丽颖, 包霞, 等. 基于改进 YOLOv5s 的指针式水表读数检测[J]. 科学技术与工程, 2024, 24(7): 2734-2741.
He Yue, Wang Liying, Bao Xia, et al. Reading detection of pointer water meter based on improved YOLOv5s [J]. Science Technology and Engineering, 2024, 24(7): 2734-2741.
- [18] 蒋博, 万毅, 谢显中. 改进 YOLOv5s 的轻量化钢材表面缺陷检测模型[J]. 计算机科学, 2023, 50(S2): 271-277.
Jiang Bo, Wan Yi, Xie Xianzhong. Improving the lightweight steel surface defect detection model for YOLOv5s [J]. Computer Science, 2023, 50 (S2): 271-277.
- [19] 董子平, 陈世国, 廖国清. 基于 YOLOv5s 的密集多人脸检测算法[J]. 计算机工程与科学, 2023, 45(10): 1838-1846.
Dong Ziping, Chen Shiguo, Liao Guoqing. A dense multi face detection algorithm based on YOLOv5s [J]. Computer Engineering and Science, 2023, 45 (10): 1838-1846.
- [20] 刘雅洁, 伊力哈木·亚尔买买提, 席凌飞, 等. 改进 YOLOv5s 的安全帽佩戴检测算法研究[J]. 计算机工程与应用, 2023, 59(20): 184-191.
Liu Yajie, Ilham Yarmaimaiti, Xi Lingfei, et al. Research on improving the safety helmet wearing detection algorithm of YOLOv5s [J]. Computer Engineering and Applications, 2023, 59 (20): 184-191.