

基于多尺度融合与无锚点 YOLO v3 的鱼群计数方法

张璐^{1,2} 黄琳³ 李备备^{1,2} 陈鑫^{1,2} 段青玲^{1,2}

(1. 中国农业大学信息与电气工程学院, 北京 100083; 2. 中国农业大学国家数字渔业创新中心, 北京 100083;
3. 宁波市海洋与渔业研究院, 宁波 315000)

摘要: 准确实现鱼群计数对于水产养殖中的生物量估算、存活率评估、养殖密度控制和运输销售管理等有着重要的指导作用。针对目前鱼群计数方法难以处理复杂背景、多尺度鱼群图像的问题,提出了一种基于多尺度融合与无锚点 YOLO v3(Multi-scale fusion and no anchor YOLO v3, MSF-NA-YOLO v3)的鱼群计数方法。首先采集多源鱼群图像,构建鱼群计数数据集,其次采用基于多尺度融合的方法提取鱼群图像特征,最后基于 CenterNet 目标检测网络识别出鱼群图像中的鱼体目标,实现鱼群计数。在真实的鱼群数据集上进行测试,计数准确率为 96.26%,召回率为 90.65%,F1 值为 93.37%,平均精度均值为 90.20%。与基于 YOLO v3、YOLO v4 和 ResNet + CenterNet 的鱼群计数方法相比,召回率分别提高了 5.80%、1.84% 和 3.48%,F1 值分别提高了 2.26%、0.33% 和 1.68%,平均精度均值分别提高了 5.96%、1.97% 和 3.67%,表明基于本研究方法的计数结果与实际计数结果相差较小,综合性能更好。

关键词: 鱼群; 水产养殖; 深度学习; 计数; YOLO v3; CenterNet

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1000-1298(2021)S0-0237-08

Fish School Counting Method Based on Multi-scale Fusion and No Anchor YOLO v3

ZHANG Lu^{1,2} HUANG Lin³ LI Beibei^{1,2} CHEN Xin^{1,2} DUAN Qingling^{1,2}

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
2. National Innovation Center for Digital Fishery, China Agricultural University, Beijing 100083, China
3. Ningbo Institute of Oceanography and Fisheries, Ningbo 315000, China)

Abstract: Accurately obtaining the number of fish is a fundamental process for biomass estimation in fish culture. It not only helps farmers calculate the reproduction rate and estimate the production potential accurately but also serves as a guide for survival rate assessment, breeding density control, and transportation sales management. It can be said that fish counting runs through multiple links such as breeding, transportation, and sales. Among these links, fish live in different environments and their body size is also various, bringing certain difficulties to fish counting. Aiming at the above problems, a fish counting method based on multi-scale fusion and no anchor YOLO v3 (MSF-NA-YOLO v3) was proposed. Firstly, multi-source fish images were collected to construct a fish counting dataset with a total of 1 858 images. Secondly, the feature extraction network of YOLO v3 was improved, and a feature extraction method based on multi-scale fusion was proposed to enhance the feature expression of fish images. Finally, the CenterNet was used as the detection network of YOLO v3, and then a fish target detection network based on no anchor was proposed to identify fish targets in images and realize fish counting. The collected fish counting dataset was randomly divided into a training set, validation set and test set. The training set and validation set accounted for 90% of the dataset, with a total of 1 672 images, and the test set accounted for 10% of the dataset, with a total of 186 images. The ratio of the

收稿日期: 2021-07-20 修回日期: 2021-09-10

基金项目: 国家重点研发计划项目(2017YFE0122100)、山东省重大科技创新工程项目(2019JZZY010703)和宁波市公益性科技项目(202002N3034)

作者简介: 张璐(1993—),女,博士生,主要从事计算机视觉在水产养殖中的应用研究,E-mail: lumanman@cau.edu.cn

通信作者: 段青玲(1967—),女,教授,博士生导师,主要从事智能信息处理研究,E-mail: dqling@cau.edu.cn

training set to the validation set was 9:1, containing 1 505 and 167 images, respectively. The MSF-NA-YOLO v3 fish counting model was trained and validated by using the transfer learning method. When the training loss and validation loss became stable, the training stopped and the best fish counting model was obtained. Based on this model, the fish images of the test set were counted and a precision of 96.26%, recall of 90.65%, F1 value of 93.37%, and average precision of 90.20% were achieved. Compared with the fish counting model based on the original YOLO v3 feature extraction method and the single scale fusion feature extraction method, the precision of the fish counting model based on the feature extraction method proposed was increased by 0.51% and 0.72%, respectively, recall was increased by 0.44% and 1.72%, respectively, F1 value was increased by 0.47% and 1.24%, respectively, and mean average precision was increased by 0.45% and 1.87%, respectively, indicating that the proposed feature extraction method had better performance. Compared with the fish counting method based on YOLO v3, YOLO v4, and ResNet + CenterNet, the recall was increased by 5.80%, 1.84%, and 3.48%, respectively, F1 value was increased by 2.26%, 0.33%, and 1.68%, respectively, and mean average precision was increased by 5.96%, 1.97%, and 3.67%, respectively. Thus, the proposed method had a good overall performance and can provide support for the realization of fishery automation and intelligence.

Key words: fish school; aquaculture; deep learning; counting; YOLO v3; CenterNet

0 引言

计数是指准确估计出目标区域内的物体数量,便于及时掌握目标区域内的信息,进行调节控制,在多个领域都有广泛应用^[1-3]。鱼群计数对于鱼类的生物量估算、存活率评估、养殖密度控制、运输销售管理等具有很好的指导作用^[4-6]。水产养殖、运输、销售等多个环节中的鱼类所处环境复杂、鱼体尺寸规格多样,这些因素给鱼群计数带来了一定的困难。因此,需要一种方法,能够处理背景复杂多样、鱼体尺度不一、质量参差不齐的鱼群图像,满足水产行业中不同环节鱼群自动、准确、高效计数的需求。

计算机视觉是一种非接触、非侵入性的技术,因其具有效率高、准确度高、所需设备简单等优点,成为一种研究鱼群计数的重要手段^[7-9]。文献[10]采用二值化、膨胀、腐蚀等方法对鱼苗图像进行预处理,在此基础上分别采用了连通区域标记法和图像细化法实现了不同背景下的鱼苗计数。试验表明当图像背景复杂时,计数精度不高。文献[11]首先采用图像增强、边缘检测和阈值处理操作分割出目标鱼体,其次提取连通区域周长特征,通过设定特征阈值的方式实现了罗非鱼的计数。文献[12]首先采用图像处理技术对鱼群图像进行预处理,包括图像二值化、形态学操作等,其次采用链码算法和傅里叶描述子提取出鱼体边界特征,最后采用基于人工神经网络的特征回归方法实现了鱼群计数。该类方法通过人为设定的特征描述子提取鱼群图像中的有用信息进行目标识别,导致模型无法适用于可变、复杂的鱼群环境,因而计数结果不稳定。

以卷积神经网络为代表的深度学习技术具有检测精度高、适应性强等特点,在机器视觉领域逐渐得到广泛应用,并取得良好效果^[13-15]。在水产养殖中,深度学习技术已被应用于水产动物识别、行为分析、品质分级等研究中^[16-17]。在计数方面,文献[18]采用基于 mask region-based CNN 的方法对鱼群视频帧进行像素级的背景检测和目标分割,以识别出图像中的鱼群目标,在此基础上采用时间阈值法和距离阈值法确定了鱼的数量。文献[19]采用多列卷积神经网络提取不同感受野的鱼群图像特征信息,调整卷积核的大小来适应鱼类运动引起的角度、形状和大小变化,以更宽更深的扩张卷积神经网络检测鱼类目标,在此基础上实现了鱼群计数。文献[20]首先训练了一个 faster region-based CNN 水下目标检测器,获得目标的边界和类别,其次通过该检测器训练了一个核化相关滤波跟踪器,实现了海产品的识别和计数。

YOLO v3 是一种基于回归的目标检测算法,具有较好的速度与精度。该算法根据数据集中目标的尺寸设置锚点框尺寸,再基于锚点框识别出图像中的目标^[21]。但是,锚点框尺寸的设置不仅增加了网络的参数量,而且面对尺度多变的鱼群图像,也难以设置出最适锚点框尺寸,影响检测效果。针对上述问题,本文提出一种基于多尺度融合与无锚点 YOLO v3 的鱼群计数方法。首先采集多源鱼群图像,构建鱼群计数数据集,其次采用基于多尺度融合的方法提取鱼群图像特征,最后基于 CenterNet 目标检测网络识别出鱼群图像中的鱼体目标,实现鱼群计数。

1 材料与方法

1.1 试验数据

以复杂背景、多尺度鱼群图像作为研究对象进行计数试验,因此所采用的图像来源于多个渠道。①在实验室搭建鱼群图像采集平台,采用摄像头录

制鱼群俯视视频,通过截取视频帧获得鱼群计数图像(图 1a、1d)。②在水产养殖基地,通过手机录制运输箱中的鱼群视频,截取视频帧获得鱼群计数图像(图 1b、1e)。③在水产销售市场,通过手机拍摄售卖池中的鱼群照片,获得鱼群计数图像(图 1c、1f)。

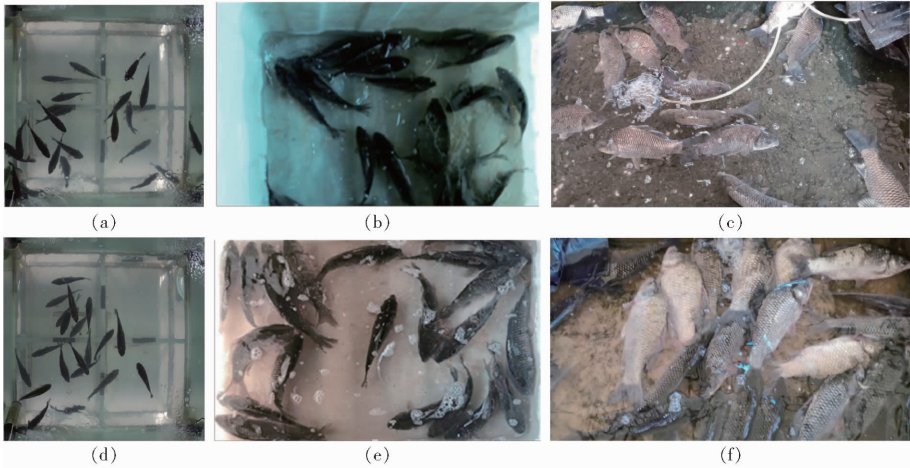


图 1 鱼群计数图像

Fig. 1 Fish counting images

从上述 3 个渠道分别获得鱼群计数图像 500、861、497 幅,采用 Labelimg 工具人工标注出每幅图

像中的鱼体目标,构成本研究的数据集,共计 1 858 幅图像。数据集详细描述如表 1 所示。

表 1 鱼群计数数据集描述

Tab. 1 Description of fish school counting dataset

序号	采集地点	数量/幅	尺寸/(像素×像素)	图像说明
1	实验室平台	500	450×450	玻璃鱼缸、增氧机背景,鱼体之间存在粘连
2	水产养殖基地	861	3 480×2 160	白色运输箱、泡沫背景,图像光照复杂,鱼体之间存在粘连
3	水产销售市场	497	3 480×2 160	水泥售卖池、增氧机、泡沫、黑色塑料袋等背景,图像光照复杂,鱼体之间存在粘连

1.2 鱼群计数

1.2.1 鱼群计数模型架构

通过目标检测算法构建检测模型,识别出图像中的鱼体目标,实现鱼群图像计数。YOLO v3 将目标检测问题转化为回归问题,通过单个卷积神经网络处理图像可直接识别出目标及其位置,具有较好的速度与精度^[21]。该算法采用 Darknet53 网络提取出 3 个尺度的图像特征,并在这 3 个尺度上设置多个锚点框检测出目标。

由于实际生产中的鱼群图像具有背景复杂多样、鱼体尺度不一、质量参差不齐等特点,为了提高计数精度,满足水产行业中不同环节的计数需求,对 YOLO v3 网络进行了改进,提出了一个基于多尺度融合与无锚点 YOLO v3 的鱼群计数方法(如图 2 所示)。在特征提取部分,在保留原网络提取的 3 个尺度特征的同时,添加了 1 个多尺度融合特征,在目标检测部分,采用无锚点的 CenterNet 代替原 YOLO

v3 基于锚点的检测网络。

1.2.2 基于多尺度融合的鱼群图像特征提取

鱼群图像特征提取是指获取图像中具有代表性的信息来表示整个图像信息,是鱼群目标检测的基础。YOLO v3 采用 Darknet53 网络进行特征提取,并通过上采样和同维拼接,最终提取出 13 像素×13 像素、26 像素×26 像素和 52 像素×52 像素 3 个尺度的特征。为了进一步加强鱼群图像特征表达,提出了一种基于多尺度融合的鱼群图像特征提取方法。在 YOLO v3 原来 3 个尺度特征的基础上,增加了 1 个 104 像素×104 像素的多尺度融合特征。特征提取网络结构如图 3 所示。图中 DBL 表示卷积、批量归一化和激活函数操作。

Darknet53 网络使用了大量的 1×1 卷积层、3×3 卷积层和残差模块。1×1 卷积层用于特征压缩,3×3 卷积层用于特征提取和增加特征通道数,残差模块的作用是解决网络层数过多造成的梯度消失或

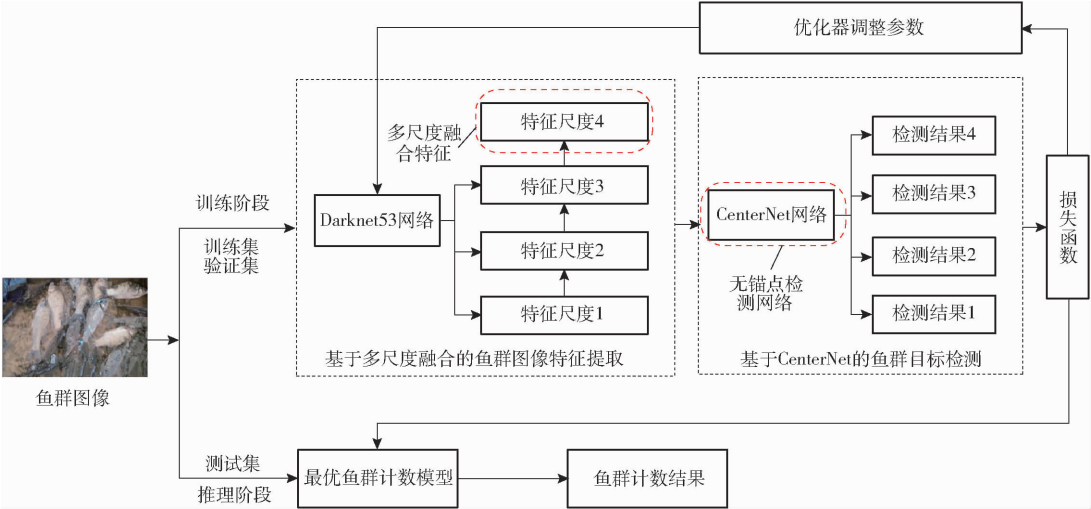


图 2 基于多尺度融合与无锚点 YOLO v3 的鱼群计数模型架构

Fig. 2 Fish counting model architecture based on multi-scale fusion and no anchor YOLO v3

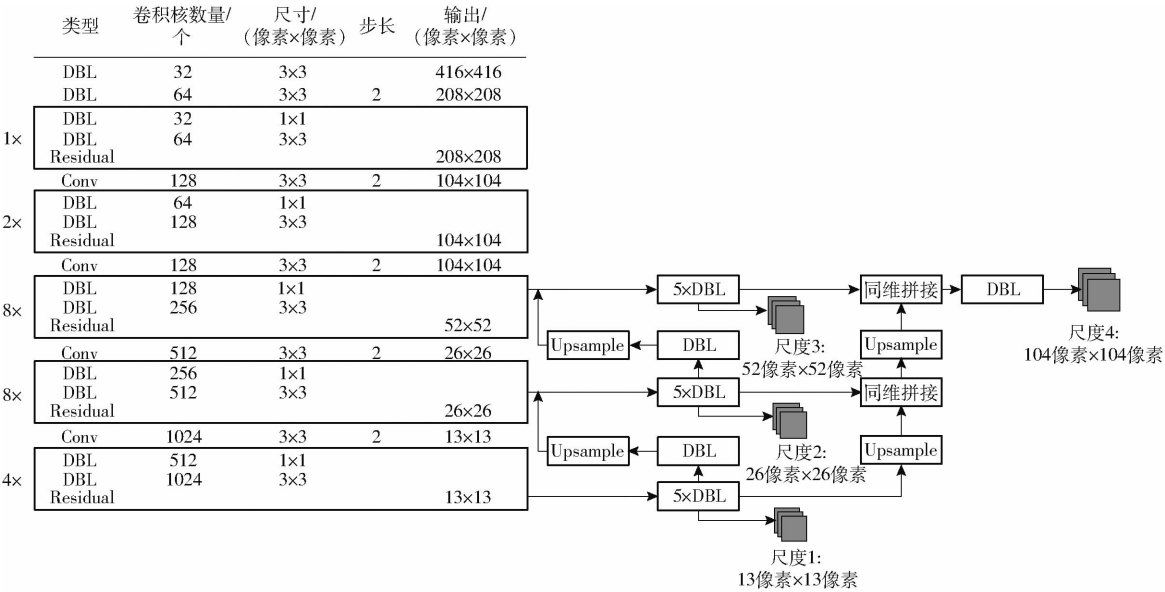


图 3 基于多尺度融合的鱼群图像特征提取网络结构

Fig. 3 Network structure diagram of fish image feature extraction based on multi-scale fusion

者梯度爆炸问题,降低深层网络训练难度。

首先将鱼群图像尺寸缩放至 416 像素 × 416 像素输入网络,其次对输入图像进行 2 次 3 × 3 卷积,步长分别为 1 和 2,然后进行多次 1 × 1 卷积、3 × 3 卷积和残差操作,分别初步获得 52 像素 × 52 像素、26 像素 × 26 像素和 13 像素 × 13 像素 3 个尺度的特征。对于 13 × 13 特征层,进行卷积操作后直接获得最终的 13 像素 × 13 像素尺度的特征;对于 26 × 26 特征层,对 13 × 13 特征层进行卷积和上采样,与 26 × 26 特征层进行同维拼接后,再进行卷积操作获得最终的 26 像素 × 26 像素尺度的特征;对于 52 × 52 特征层,对 26 × 26 特征层进行卷积和上采样,与 52 × 52 特征层进行同维拼接后,再进行卷积操作获得最终的 52 像素 × 52 像素尺度的特征。

在此基础上,将 13 × 13 特征层进行 3 次上采

样、批量归一化和激活操作,将 26 × 26 特征层进行 2 次上采样、批量归一化和激活操作,将 52 × 52 特征层进行 1 次上采样、批量归一化和激活操作,分别获得 104 × 104 像素尺度特征,对 3 个 104 像素 × 104 像素尺度特征进行同维拼接和卷积,最终获得多尺度融合特征。

经过基于多尺度融合的特征提取网络之后,提取出了 13 像素 × 13 像素、26 像素 × 26 像素、52 像素 × 52 像素和 104 像素 × 104 像素 4 个尺度的鱼群图像特征,分别在 4 个尺度特征层上进行鱼体目标检测。

1.2.3 基于 CenterNet 的鱼群目标检测

基于提取的鱼群图像特征,构建目标检测模型,识别出图像中的鱼体目标,实现鱼群图像计数。原 YOLO v3 的检测网络需要根据鱼体尺寸设置锚点框

尺寸,再基于锚点框识别出图像中的鱼体目标。实际生产中的鱼体尺度不一,采用原 YOLO v3 的检测网络难以设置出最适锚点框尺寸,影响模型检测效果。因此,针对上述问题,采用 CenterNet 作为 YOLO v3 的检测网络,提出了一种基于无锚点的鱼群目标检测方法。

CenterNet 目标检测算法将目标边框的获取转换为目标中心点的预测分类问题,通过关键点估计找到中心点,利用中心点的图像信息得到目标的尺寸和位置。无需预先设置锚点框尺寸,减少了网络参数数量和计算量,避免了参数设置不准确影响模型检测效果的问题^[22]。基于 CenterNet 的鱼群目标检测网络包括解码网络和检测子网络,如图 4 所示。

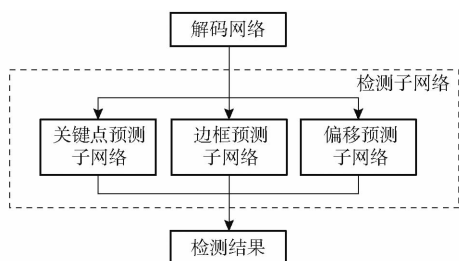


图 4 基于 CenterNet 的鱼群目标检测网络结构

Fig. 4 Network structure diagram of fish target detection based on CenterNet

假设输入的鱼群计数图像为 $I \in R^{W \times H \times 3}$, 其中 R 为图像缩放比, W 、 H 分别为鱼群计数图像的宽和高。在解码网络中,对特征提取阶段获得的 4 个尺度的特征分别进行上采样操作,获得 $W/R \times H/R$ 像素的特征。检测子网络包括关键点预测子网络、边框预测子网络和偏移预测子网络 3 部分,分别输出关键点对应目标的热力图、尺寸和偏移量,将 3 个输出结合后获得最终的检测结果。

关键点的热力图、尺寸和偏移量分别表示为

$$Y \in [0, 1]^{W/R \times H/R \times C} \quad (1)$$

$$(\hat{w}_i, \hat{h}_i) = \hat{S}_{\hat{x}_i, \hat{y}_i} \in R^{W/R \times H/R \times 2} \quad (2)$$

$$(\delta \hat{x}_i, \delta \hat{y}_i) = \hat{O}_{\hat{x}_i, \hat{y}_i} \in R^{W/R \times H/R \times 2} \quad (3)$$

式中 Y ——关键点对应目标

$(\hat{w}_i, \hat{h}_i)$ ——预测出的关键点尺寸

$(\delta \hat{x}_i, \delta \hat{y}_i)$ ——预测出的关键点偏移量

C ——检测目标的类别数目,取 1

在结合输出之前,首先对网络输出的关键点进行筛选,获得中心点集合。将热力图上的所有关键点与其连接的 8 个临近点进行比较,如果该点的响应值大于或等于其 8 个临近点则保留,最终保留所有满足条件的前 n 个峰值点,构成中心点集合 $\hat{P}_c$,表示为

$$\hat{P}_c = \{(\hat{x}_i, \hat{y}_i)\}_{i=1}^n \quad (4)$$

式中 $(\hat{x}_i, \hat{y}_i)$ ——第 i 个中心点的坐标

每一个关键点的位置表示为 (x_i, y_i) ,之后,CenterNet 使用关键点值作为其检测强度的度量,并在位置上产生一个边界框,表示为

$$B = \begin{bmatrix} \hat{x}_i + \delta \hat{x}_i - \frac{\hat{w}_i}{2} & \hat{y}_i + \delta \hat{y}_i - \frac{\hat{h}_i}{2} \\ \hat{x}_i + \delta \hat{x}_i + \frac{\hat{w}_i}{2} & \hat{y}_i + \delta \hat{y}_i + \frac{\hat{h}_i}{2} \end{bmatrix} \quad (5)$$

使用基于 CenterNet 的鱼群目标检测网络,首先根据热力图中关键点分布回归计算确定鱼体目标的中心点,再根据中心点位置回归计算确定鱼体目标的尺寸和偏移量,最终在图像中绘制出目标边界框,识别出鱼体目标,实现鱼群计数。

1.3 模型训练与评价指标

1.3.1 模型训练

鱼群计数数据集共 1 858 幅图像,划分为训练集、验证集和测试集,其中训练集和验证集占 90%,共 1 672 幅图像,测试集占 10%,共 186 幅图像。在鱼群计数模型训练的过程中,训练集用于模型训练,验证集用于模型验证以调节模型参数获得最优鱼群计数模型。训练集和验证集的比例为 9:1,分别包含 1 505 幅和 167 幅图像。

采用迁移学习方法^[23]进行鱼群计数模型的训练。在粗略训练阶段,冻结全部卷积层,只对全连接层进行训练,迭代 50 次,批尺寸设置为 16,初始学习率设置为 0.001。在精调训练阶段,解冻全部卷积层,对所有层进行训练,迭代 100 次,批尺寸设置为 8,初始学习率设置为 0.000 1。为了防止模型过拟合,采用早停法进行训练,即当模型在验证集上的误差比上一次训练结果差的时候停止训练,使用上一次迭代结果中的参数作为模型的最终参数。动量参数设置为 0.9,采用 Adam 优化算法。

鱼群计数模型训练时采用的损失函数 L_{det} 由中心点定位损失函数 L_k 、目标尺寸损失函数 L_{size} 和偏移量损失函数 L_{off} 组成,计算公式^[22]为

$$L_{det} = L_k + \lambda_{size} L_{size} + \lambda_{off} L_{off} \quad (6)$$

式中 λ_{size} —— L_{size} 的权重调节参数,设置为 1

λ_{off} —— L_{off} 的权重调节参数,设置为 0.1

L_k 、 L_{size} 和 L_{off} 的计算公式分别为

$$L_k = \begin{cases} \frac{-1}{N} (1 - \hat{Y}_{xyc})^\alpha \lg \hat{Y}_{xyc} & (Y_{xyc} = 1) \\ \frac{-1}{N} (1 - Y_{xyc})^\beta \hat{Y}_{xyc}^\alpha \lg (1 - \hat{Y}_{xyc}) & (Y_{xyc} \neq 1) \end{cases} \quad (7)$$

$$L_{size} = \frac{1}{N} \sum_{k=1}^N |\hat{S}_{pk} - s_k| \quad (8)$$

$$L_{off} = \frac{1}{N} \sum_p \left| \hat{O}_{\tilde{p}} - \left(\frac{p}{R} - \tilde{p} \right) \right| \quad (9)$$

- 式中
- α, β ——权重调节参数,取 2 和 4
- N ——鱼群计数图像 I 的关键点数目
- Y_{xyc} ——热力图中中心点分布函数的真实值
- $\hat{Y}_{xyc}$ ——热力图中中心点分布函数的预测值
- s_k ——目标鱼体尺寸的真实值
- $\hat{s}_{pk}$ ——目标鱼体尺寸的预测值
- $\hat{O}_{\tilde{p}}$ ——中心点偏移量的预测值
- p ——中心点的真实值
- $\tilde{p}$ ——中心点的预测值

试验的硬件环境:Ubuntu 18.04 操作系统,中央处理器 (Central processing unit, CPU) 为 Intel (R) Xeon (R) Silver 4114, 主频为 2.2 GHz, 内存为 125 GB, 图形处理器 (Graphics processing unit, GPU) 为 GeForce RTX 2080Ti; 软件配置: 运算平台为 Cuda Toolkit 10.0, 深度神经网络的 GPU 加速库为 Cudnn 7.6.1, 编程语言为 Python, 深度学习库为 Keras 2.1.5 和 Tensorflow-gpu 1.13.2。

1.3.2 评价指标

采用准确率、召回率、F1 值和平均精度均值评估鱼群计数模型的性能。

2 结果与分析

2.1 模型训练结果

采用迁移学习训练方法对所提出的基于多尺度融合与无锚点 YOLO v3 鱼群计数模型进行训练。模型在训练集和验证集上的损失值曲线如图 5 所示。

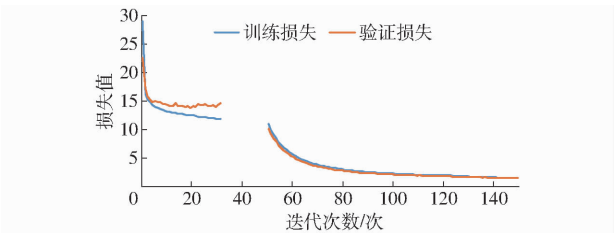


图 5 训练集和验证集上的损失值
Fig.5 Loss value on training set and validation set

在粗略训练阶段,模型刚开始训练时,训练集和验证集上的损失值都非常大,随着训练迭代次数的增加,两者的损失值均在下降,最终迭代 32 次时,基于早停法模型停止训练,最终粗略训练阶段获得训练集上的最小损失值为 11.973 7,验证集上的最小损失值为 13.882 1。进入精调训练阶段,训练集和验证集上的损失值均急剧下降,当迭代 150 次时,损失值趋于平稳,模型停止训练。最终训练集上的最小损失值为 1.613 2,验证集上的最小损失值为 1.589 3,

此时获得最优鱼群计数模型。

2.2 鱼群计数结果

采用训练得到的最优 MSF-NA-YOLO v3 鱼群计数模型对测试集的 186 幅图像进行测试,通过多个指标对模型性能进行了评价。在测试集 186 幅图像中共有 2 758 条鱼,正确检测出的鱼体数目为 2 500,错误检测出的鱼体数目为 97,计数准确率为 96.26%,召回率为 90.65%,F1 值为 93.37%,平均精度均值为 90.20%。计数准确率达到了 96% 以上,召回率、F1 值和平均精度均值也均高于 90%,验证了本研究方法的有效性。

计数模型在测试集样例上的检测结果如图 6 所示。图中绿色框表示正确检测出的鱼体目标,红色框表示错误检测出的鱼体目标,紫色框表示未检测出的鱼体目标。

图 6a、6d、6g 为在实验室平台采集的鱼群图像,图像背景比较简单,分别有 23、18、17 条鱼,均被全部正确检测出来,计数准确率为 100%。图 6b、6e、6h 为在水产养殖基地运输箱中采集的鱼群图像,鱼体粘连程度较大且表面含有一些泡沫。图 6b、6e 分别有 15、21 条鱼,均被全部正确检测出来,图 6h 中有 21 条鱼,19 条被正确检测出来,2 条被漏检。图 6c、6f、6i 为在水产销售市场售卖池中采集的鱼群图像,图像背景很复杂、鱼体粘连程度严重且表面含有较多泡沫。图 6c、6f 中分别有 17、14 条鱼,均被正确检测出来,图 6i 中有 5 条鱼,3 条被正确检测出来,1 条被漏检,1 条被误检,还有 1 条(位于图像左上角)由于检测绘制的边界框与实际标注的边界框存在较大差距,既属于被误检,也属于被漏检。总体来说,采用所提出的 MSF-NA-YOLO v3 鱼群计数模型对于不同生产环节中复杂背景下的多尺度鱼群图像均取得较好的计数结果。

2.3 对比试验结果

2.3.1 特征提取方式对比试验结果

为了验证所提出的基于多尺度融合的鱼群图像特征提取方法的有效性,将其与原 YOLO v3 的特征提取方式(3 个尺度特征层)和单一融合尺度特征提取方式(3 个尺度特征层融合成 1 个尺度特征层)进行了对比。3 种特征提取方式均在基于 CenterNet 的鱼群目标检测网络上进行计数。基于不同特征提取方式的鱼群计数模型的对比结果如图 7 所示。图中 YOLO v3-CenterNet-3 表示基于原 YOLO v3 特征提取方式的鱼群计数模型,YOLO v3-CenterNet-1 表示基于单一融合尺度特征提取方式的鱼群计数模型。

由图 7 可知,基于多尺度融合的鱼群图像特征

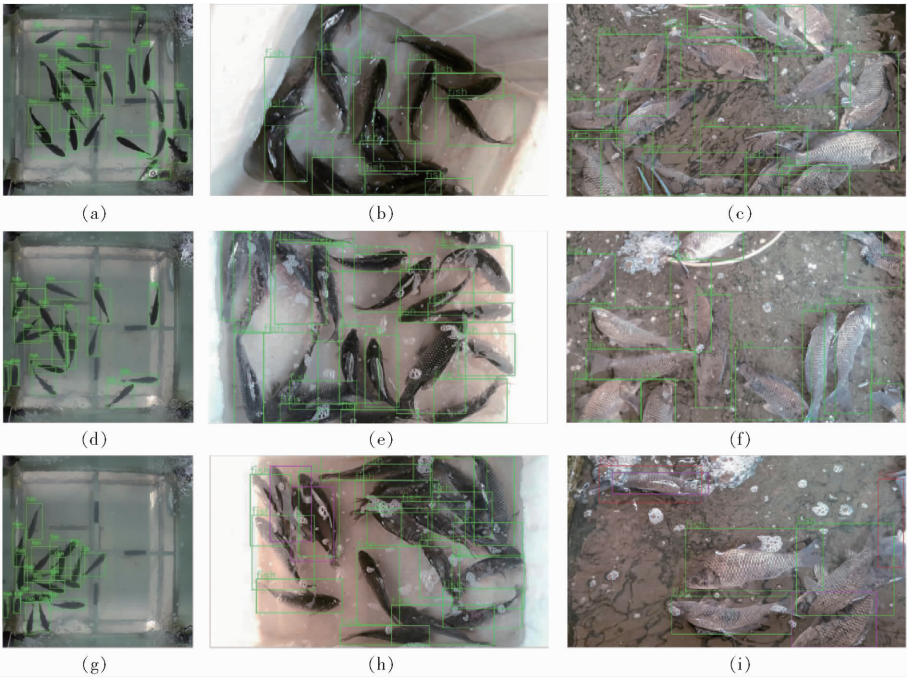


图 6 鱼群图像检测结果

Fig. 6 Fish school image detection results

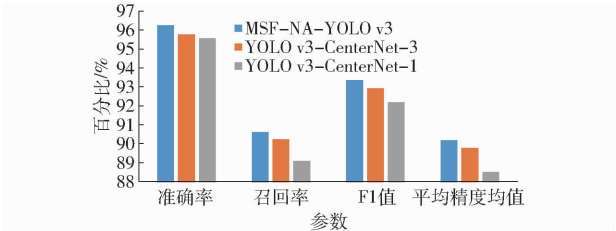


图 7 基于不同特征提取方式的鱼群计数模型对比结果

Fig. 7 Comparison results of fish counting models based on different feature extraction methods

提取方式的性能优于对比方法。在所提出的特征提取方式下、原 YOLO v3 特征提取方式下和单一融合尺度特征提取方式下,鱼群计数模型的准确率分别为 96.26%、95.77% 和 95.57%,召回率分别为 90.65%、90.25% 和 89.12%,F1 值分别为 93.37%、92.93% 和 92.23%,平均精度均值分别为 90.20%、89.80% 和 88.54%。与两种对比特征提取方式相比,本文方法的准确率分别提高了 0.51% 和 0.72%,召回率分别提高了 0.44% 和 1.72%,F1 值分别提高了 0.47% 和 1.24%,平均精度均值分别提高了 0.45% 和 1.87%。在所提出的鱼群特征提取方式下,鱼群计数模型的各项指标均优于对比方法,验证了本文方法的有效性与优越性。

2.3.2 鱼群计数模型对比试验结果

为了进一步验证所提出的鱼群计数方法的性能,将其与基于 YOLO v3、YOLO v4^[24] 和 ResNet + CenterNet^[22] 的鱼群计数方法进行了对比。ResNet + CenterNet 为无锚点目标检测算法。对比结果如表 2 所示。

表 2 鱼群计数模型对比结果

Tab.2 Comparison results of fish school counting models

方法	准确率	召回率	F1 值	平均精度均值
MSF - NA - YOLO v3	96.26	90.65	93.37	90.20
YOLO v3	97.73	85.68	91.31	85.13
YOLO v4	97.50	89.01	93.06	88.46
ResNet + CenterNet	96.49	87.60	91.83	87.01

由表 2 可知,所提出的基于 MSF - NA - YOLO v3 的鱼群计数方法的性能优于对比方法。与基于 YOLO v3 的鱼群计数方法相比,召回率提高了 5.80%,F1 值提高了 2.26%,平均精度均值提高了 5.96%,与基于 YOLO v4 的鱼群计数方法相比,召回率提高了 1.84%,F1 值提高了 0.33%,平均精度均值提高了 1.97%,与基于 ResNet + CenterNet 的鱼群计数方法相比,召回率提高了 3.48%,F1 值提高了 1.68%,平均精度均值提高了 3.67%。对比方法的召回率明显低于 MSF - NA - YOLO v3 方法,即鱼群计数图像中有很多鱼体目标未被检测出来。究其原因,YOLO v3 和 YOLO v4 方法均为基于锚点的目标检测算法,需根据数据集中待检测目标的尺寸设置对应的锚点框尺寸,对于实际生产中尺度复杂多样的鱼群计数图像并不适用。与此同时,虽然 ResNet + CenterNet 为无锚点目标检测算法,但由于其特征提取不佳,其召回率、F1 值和平均精度均值也明显低于 MSF - NA - YOLO v3 方法,也验证了所提出的鱼群计数方法的优越性。MSF - NA - YOLO v3 方法的准确率略低于对比方法,但其他评价指标均高于对

比方法,表明本研究方法的综合性能较好。

3 结论

(1)针对水产养殖中鱼群图像具有复杂背景、多尺度的特点,本文改进 YOLO v3 的特征提取网络和目标检测网络,提出了一种基于多尺度融合与无锚点 YOLO v3 的鱼群计数方法。试验结果表明,准确率为 96.26%,召回率为 90.65%,F1 值为 93.37%,平均精度均值为 90.20%,计数效果较好。

(2)与基于原 YOLO v3 特征提取方式和单一融合尺度特征提取方式的鱼群计数模型相比,基于所

提出的特征提取方式的鱼群计数模型准确率分别提高了 0.51% 和 0.72%,召回率分别提高了 0.44% 和 1.72%,F1 值分别提高了 0.47% 和 1.24%,平均精度均值分别提高了 0.45% 和 1.87%,表明本研究提出的特征提取方式具有更好的性能。

(3)与基于 YOLO v3、YOLO v4 和 ResNet + CenterNet 的鱼群计数方法相比,召回率分别提高了 5.80%、1.84% 和 3.48%,F1 值分别提高了 2.26%、0.33% 和 1.68%,平均精度均值分别提高了 5.96%、1.97% 和 3.67%,表明基于本文方法的计数结果与实际计数结果相差较小,综合性能更好。

参 考 文 献

- [1] ARCIDIACONO C, PORTO S M C, MANCINO M, et al. A threshold-based algorithm for the development of inertial sensor-based systems to perform real-time cow step counting in free-stall barns[J]. *Biosystems Engineering*, 2017,153(1):99–109.
- [2] 张领先,陈运强,李云霞,等. 基于卷积神经网络的冬小麦麦穗检测计数系统[J]. *农业机械学报*, 2019,50(3):144–150.
- [3] ZHANG Lingxian, CHEN Yunqiang, LI Yunxia, et al. Detection and counting system for winter wheat ears based on convolutional neural network[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019,50(3):144–150. (in Chinese)
- [3] 姜海燕,徐灿,陈尧,等. 基于田间图像的局部遮挡小尺寸稻穗检测和计数方法[J]. *农业机械学报*, 2020,51(9):152–162.
- [4] JIANG Haiyan, XU Can, CHEN Yao, et al. Detecting and counting method for small-sized and occluded rice panicles based on in-field images[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2020,51(9):152–162. (in Chinese)
- [4] LI Daoliang, MIAO Zheng, PENG Fang, et al. Automatic counting methods in aquaculture: a review[J]. *Journal of the World Aquaculture Society*, 2020,52(2):1–15.
- [5] ZHANG Lu, LI Wensheng, LIU Chunhong, et al. Automatic fish counting method using image density grading and local regression[J]. *Computers and Electronics in Agriculture*, 2020,179(12):105844.
- [6] JING Danxiang, HAN Jun, WANG Xiaoding, et al. A method to estimate the abundance of fish based on dual-frequency identification sonar (DIDSON) imaging[J]. *Fisheries Science*, 2017,83(5):685–697.
- [7] PUIG-PONS V, MUNOZ-BENAVENT P, ESPINOSA V, et al. Automatic Bluefin Tuna (*Thunnus thynnus*) biomass estimation during transfers using acoustic and computer vision techniques[J]. *Aquacultural Engineering*, 2019,85(5):22–31.
- [8] HERNANDEZ-ONTIVEROS J M, INZUNZA-GONZALEZ E, GARCIA-GUERRERO E E, et al. Development and implementation of a fish counter by using an embedded system[J]. *Computers and Electronics in Agriculture*, 2018,145(2):53–62.
- [9] GARCIA V, SANT A D A, GARCIA Z V A, et al. A new image dataset for the evaluation of automatic fingerlings counting[J]. *Aquacultural Engineering*, 2020,89(5):102064.
- [10] ZHANG Jianfei, PANG Haitong, CAI Weiming, et al. Using image processing technology to create a novel fry counting algorithm[J/OL]. *Aquaculture and Fisheries*, 2021. <https://doi.org/10.1016/j.aaf.2020.11.004>.
- [11] AWALLUDIN E A., MUHAMMAD W N A W, ARSAD T N T, et al. Fish larvae counting system using image processing techniques[J]. *Journal of Physics: Conference Series*, 2020,1529(5):052040.
- [12] ALIYU I, GANA K J, MUSA A A, et al. Incorporating recognition in catfish counting algorithm using artificial neural network and geometry[J]. *KSII Transactions on Internet and Information Systems*, 2020,12(14):4866–4888.
- [13] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. *Nature*, 2015,521(7553):436–444.
- [14] LIU Weibo, WANG Zidong, LIU Xiaohui, et al. A survey of deep neural network architectures and their applications[J]. *Neurocomputing*, 2017,234(4):11–26.
- [15] KAMILARIS A, PRENAFETA-BOLDU F X. Deep learning in agriculture: a survey[J]. *Computers and Electronics in Agriculture*, 2018,147(4):70–90.
- [16] QIN Hongwei, LI Xiu, LIANG Jian, et al. DeepFish: accurate underwater live fish recognition with a deep architecture[J]. *Neurocomputing*, 2016,187(4):49–58.
- [17] TAHERI-GARAVAND A, NASIRI A, BANAN A, et al. Smart deep learning-based approach for non-destructive freshness diagnosis of common carp fish[J]. *Journal of Food Engineering*, 2020,278(8):109930.
- [18] TSENG C H, KUO Y F. Original article detecting and counting harvested fish and identifying fish types in electronic monitoring system videos using deep convolutional neural networks[J]. *ICES Journal of Marine Science*, 2020,77(4):1367–1378.
- [19] ZHANG Song, YANG Xinting, WANG Yizhong, et al. Automatic fish population counting by machine vision and a hybrid deep neural network model[J]. *Animals*, 2020,10(2):364.
- [20] XU Fengqiang, DING Xueyan, PENG Jinjia, et al. Real-time detecting method of marine small object with underwater robot vision[C]//OCEANS-MTS/IEEE Kobo Techno-Oceans Conference (OTO), Kobe: IEEE, 2018.
- [21] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. [2018–04–08]. <https://arxiv.org/abs/1804.02767>.
- [22] ZHOU X, WANG D, KRAHENBUHL P. Objects as points[EB/OL]. [2019–04–25]. <https://arxiv.org/abs/1904.07850v2>.
- [23] 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.
- [23] ZHUANG Fuzhen, LUO Ping, HE Qing, et al. Survey on transfer learning research[J]. *Journal of Software*, 2015,26(1):26–39. (in Chinese)
- [24] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. [2020–04–23]. <https://arxiv.org/abs/2004.10934>.