

水产养殖中水质与鱼类行为双向映射模型研究

魏天娇 胡祝华 范习禹

(海南大学信息与通信工程学院, 海口 570228)

摘要: 在水产养殖中,水质参数与鱼类活动之间有着密不可分的相互映射关系。过去的监测更多偏向于单向映射,一般都是通过鱼类的行为表明水质的情况。针对仅仅通过鱼类行为反映水质情况会产生误判和滞后的问题,本文构建一种基于随机森林的鱼类行为与水质情况双向映射模型。双向映射模型不仅可以提供更多的信息从而提高预测的准确性,而且也可以通过相互验证提高模型的可靠性。首先,通过引入可变形卷积模块对 YOLO v7 进行改进,利用改进模型检测出视频中鱼类的位置再通过前后帧的坐标量化出鱼的游动参数。随后,将采集到的鱼类游动参数及对应的水质参数作为输入,使用随机森林模型进行分类、回归,分别完成鱼类游动参数和水质参数具体数值的预测以及指标异常级别的预测,从而得到双向映射关系。为了表明模型的泛化能力,分别在黎安港和新村港渔场 2 个数据集下进行实验。实验结果表明:提出的方法可以较好地实现鱼类行为与水质关系的双向映射,其中,分类实验平均准确率可以达到 90.947%,回归实验决定系数 R^2 的平均值可以达到 0.801。

关键词: 智慧养殖; 鱼类行为; 水质; 深度学习; 随机森林; 双向映射模型

中图分类号: S927

文献标识码: A

文章编号: 1000-1298(2024)03-0290-10

OSID:



Construction of Bidirectional Mapping Model between Water Quality and Fish Behavior in Aquaculture

WEI Tianjiao HU Zhuhua FAN Xiyu

(School of Information and Communication Engineering, Hainan University, Haikou 570228, China)

Abstract: In aquaculture, there is an inseparable mutual mapping relationship between water quality and fish behaviors. In the past, monitoring was more biased towards one-way mapping, which generally indicated the water quality through fish behaviors. In order to solve the problem of misjudgment and lag only by fish behaviors, a bidirectional mapping model between fish behaviors and water quality was constructed based on random forest. The bidirectional mapping model can not only provide more information to improve the accuracy of prediction, but also improve the reliability of the model through mutual verification. Firstly, YOLO v7 was improved by introducing a deformable convolution module, and the position of fish in the video was detected by using the improved model, and then the swimming parameters of fish were quantified by the coordinates of the front and back frames. Then, the collected fish swimming parameters and the corresponding water quality parameters were taken as inputs, the random forest model was used for classification and regression, and the specific numerical values of fish swimming parameters and water quality parameters and the abnormal level of indicators were predicted respectively, so as to obtain a bidirectional mapping relationship. In order to show the generalization ability of the model, experiments were carried out under two data sets: Li'an Port and Xincun Port Fishing Ground. The experimental results showed that the proposed method can realize the bidirectional mapping between fish behaviors and water quality. Among them, the average accuracy of classification experiment can reach 90.947%, and the average value of regression experiment determination coefficient R^2 can reach 0.801.

Key words: intelligent farming; fish behavior; water quality; deep learning; random forest; bidirectional mapping model

收稿日期: 2023-10-28 修回日期: 2024-01-23

基金项目: 国家自然科学基金项目(61963012)、国家重点研发计划项目(2022YFD2400504)和海南省种业实验室项目(B23H10004)

作者简介: 魏天娇(2001—),女,硕士生,主要从事人工智能研究,E-mail: weitianjiao@hainanu.edu.cn

通信作者: 胡祝华(1979—),男,教授,博士生导师,主要从事人工智能和精准测控研究,E-mail: eagler_hu@hainanu.edu.cn

0 引言

鱼类的运动情况与水质参数密不可分,水质参数的变化会带来鱼类行为的变化。同时,水质参数的变化趋势也可一定程度地预测鱼类的行为。因此,鱼类行为与水质参数可互相参照。目前的监测更多偏向于单向映射,一般都是通过鱼类行为表明水质情况,然而仅仅通过鱼类行为去反映水质情况难免会产生滞后性。因此需要通过构建双向映射模型来深入了解鱼类行为与水质参数之间的复杂关系。

对于鱼类行为识别,VILLON 等^[1]通过 GoogLe - Net(Google inception net)架构提取鱼类特征,采用 Softmax 分类的方法实现对鱼的检测,并且还采用一个决策规则来提升检测的精度。RAUF 等^[2]提出了一个基于卷积神经网络 (Convolutional neural network, CNN)的鱼种识别框架,通过在网络中的每个级别的训练中即时添加 4 个卷积层来提高分类性能。王迎^[3]使用 CenterNet 检测网络,并利用无迹卡尔曼滤波算法实现鱼类的检测,最后使用长短期记忆网络 (Long short-term memory networks, LSTM)预测结果识别出水质异常。芦玥同等^[4]提出了通过电磁信号、视频信号以及其他方法追踪鱼类应激行为实现了水质在线监测预警。此外,对于水质预测的研究方法可分为 2 种:机理性水质预测方法和非机理性水质预测方法^[5-6]。由于机理性预测方法在实现过程中有一定难度,所以目前的研究中多使用非机理性水质预测方法。传统的非机理性水质预测方法有时间序列预测法、马夫科尔法、回归预测法和灰色系统预测法,这些方法主要应用污水处理领域。虽然这些方法容易实现,但在实验时考虑的因素较少会导致实验精度较低。随着机器学习的发展,很多与之相关的预测方法得以发展,如支持向量机和深度神经网络。这 2 种方法在水产养殖水质预测研究中取得了良好的应用效果^[7]。后来学者又提出了一些水产溶解氧非线性预测方法^[8],但这些方法并没考虑到其余水质参数对实验结果的影响。人工神经网络 (Artificial neural network, ANN)的发展恰恰可以解决这一问题。ANN 是一种模仿生物神经网络的结构和功能的数学模型,这个网络中由神经元与神经元连接形成,通过改动不同神经元之间的连接方式来实现对信息的处理,具有拟合复杂非线性关系的特点。CHEN 等^[9]利用简单循环神经网络模型 (Simple recurrent unit, SRU)和优化后的双向堆叠简单循环神经网络模型 (Bidirectional stack simple recurrent unit, Bi - SRU)分别对水质展开了

长短期的预测实验。FU 等^[10]提出了一种基于时间卷积网络 (Temporal convolutional network, TCN)的水质预测方法,利用 Pearson 相关系数法分析水质参数之间的相关性,利用先验信息和预处理后的水质数据构建基于时域卷积网络的水质参数长期预测模型。

传统的基于滤波、分割和统计的方法可以建立简单场景的鱼体行为参数获取模型^[11],然而,构建复杂场景下的鱼体行为参数获取方法面临困难。此外,通过挖掘时间相关性和结构相关性来构建行为参数与对应的水质参数之间的互作用关系也是一个关键性的科学问题。因此,本文将通过引入可变形卷积模块得到 YOLO v7 改进模型 (Deformable Convolutional networks - you only look once version 7, DCN - YOLO v7)来检测视频中鱼类的位置,利用前后帧的坐标量化出鱼的游动参数,随后将采集到的游动参数及对应的水质参数作为输入,使用随机森林模型进行分类、回归,分别完成鱼类游动参数和水质参数具体数值的预测以及指标异常级别的预测,从而得到双向映射关系,以便养殖者可以更容易掌握养殖情况,对异常情况及时采取防范措施。

1 数据获取

1.1 目标检测方法

针对水下环境复杂、鱼体数量繁多、传统检测鱼体形态困难等问题,选择 DCN v3 - YOLO v7 模型作为检测鱼体形态的方法,目的是获得鱼体中心坐标,通过鱼体中心坐标差异计算鱼体速度、加速度和游动高度的实验指标,为后续随机森林的预测提供丰富的输入数据^[12-15]。

YOLO v7 模型是一种代表性的目标检测算法,其架构基于全卷积神经网络,利用深度学习技术对图像中的对象进行高效的识别和定位。该算法核心组成部分包括主干网络 (Backbone)、头部网络 (Head)、特征融合网络 (Neck)。

YOLO v7 模型参数和存储复杂度与采样点总数呈线性关系,模型效率低,检测时间较慢。因此,引入动态可变形卷积 (Deformable convolutional networks version 3, DCN v3)到 YOLO v7 模型中骨干网络的 P5 层,取代原来 P5 层所有卷积核 3×3 的卷积层,使得在骨干网络中的卷积神经元之间共享权重。其主要思想是将原始的卷积权重 w_k 分离为深度卷积和点卷积 2 部分,将空间聚集过程分成 g 组,每组都有单独的采样偏移 p_{gk} 和调制尺度 M_{gk} ,最后将卷积采样点间的调制标量归一化,使得在传输鱼体不同尺度下的模型训练过程更加稳定^[16]。具体

计算式为

$$y(p_0) = \sum_{g=1}^G \sum_{k=1}^K w_g m_{gk} x_g (p_0 + p_k + \Delta p_{gk}) \quad (1)$$

式中 G ——聚合组总数

p_0 ——当前像素 K ——维度

p_k ——预定义网格采样的第 k 个位置

w_g ——第 g 组的位置无关的投影权重

m_{gk} ——第 g 组中第 k 个采样点的调制标量, 由 Softmax 函数沿维度 K 归一化

x_g ——YOLO v7 模型的 P4 特征层经过采样后输出的特征图
 Δp_{gk} ——对应于第 g 组中栅格采样位置 p_k 的偏移量

为了减少过拟合,输入图像经过数据增强操作,如 Mosaic 和随机裁剪等,随后输入主干网络中,Backbone 用于提取图像的基本特征,Neck 则负责构建特征金字塔。最终,Head 部分生成最终的检测输出^[17]。YOLO v7 模型网络架构如图 1 所示。

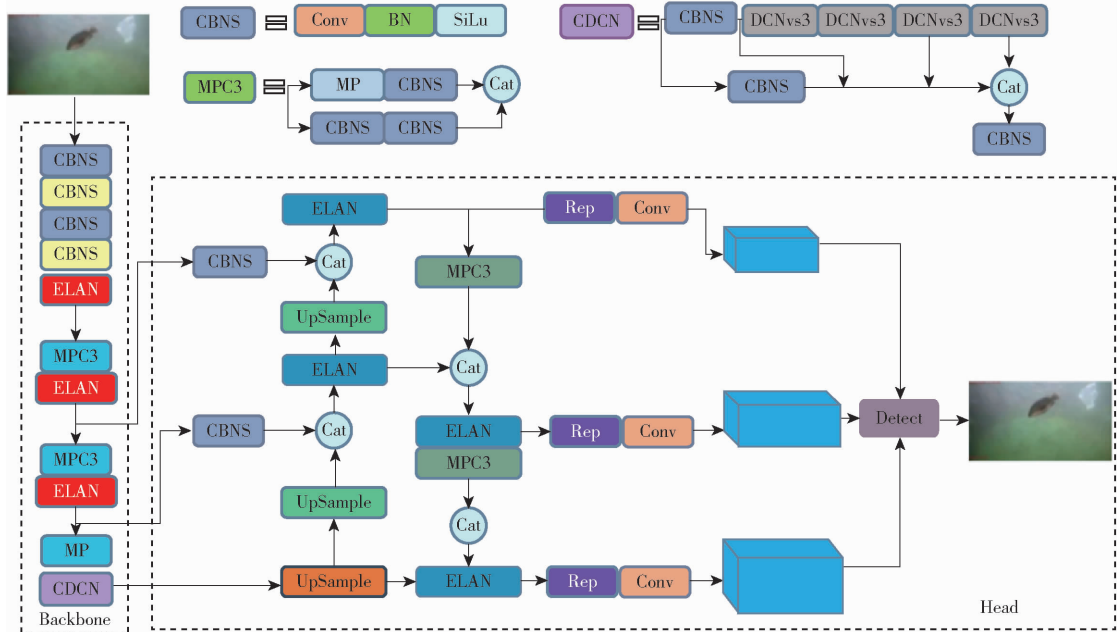


图 1 DCN - YOLO v7 检测模型结构以及检测流程
Fig. 1 Detection model structure and process of DCN - YOLO v7

Conv + Batch normalization + SiLU (CBNS) 模块由卷积、批量归一化层 (Batch normalization) 和激活函数 (Sigmoid-weighted linear unit, SiLU) 组成。高效层聚合网络 (Efficient layer aggregation networks, ELAN) 模块能够调整梯度路径的长度,使得深层网络能够高效学习和收敛,由多个 CBNS 模块组成。ELAN 模块的特点是,它不改变输入特征图的尺寸,仅更改输出通道数。Conv + Deformable convolutional networks (CDCN) 模块将骨干网络中所有 3×3 卷积替换成可变形卷积以适应特征图中鱼体的不规则形状,提高 bounding box 回归的精准性,从而得到更加精确的鱼体中心坐标,以计算鱼体的速度、加速度等参数。Maxpooling + CBNS (MPC3) 模块分为上下 2 部分,通过 Maxpooling 操作实现图像尺寸减半,然后通过 CBNS 层进行通道数减半。最后,这 2 部分特征通过 Concat 操作融合。

1.2 实验数据集

该实验中鱼类的数据集在海南省陵水县新村港和黎安港渔场采集得到,采集装置如图 2 所示,采集

到地鱼类信息以视频方式呈现,视频图像分辨率为 $3\,840 \text{ 像素} \times 2\,160 \text{ 像素} \times 3 \text{ 通道}$ ^[18]。

将采集到的视频数据,按照每帧单独分离成分辨率为 $3\,840 \text{ 像素} \times 2\,160 \text{ 像素}$ 的图像,并且使用软件 Labelme 对图像中的鱼体进行标注。实验环境是 ubuntu 20.04 操作系统,GPU 为 RTX 3080Ti,Pytorch 平台,CUDA 11.3 以及 Python 3.8。实验中采用 Labelme 对鱼类视频数据集进行标注。在模型训练中,为了保证训练的充分性,关键参数设置为:训练周期均为 300 轮,batch_size 调整为 8,优化函数采用 SGD 函数,动量参数为 0.937,权重衰减参数为 0.000 5。设置初始学习率为 0.01,并且在第 3 轮时以 0.1 进行衰减^[19]。将训练好的最优模型在测试集下测试,检测结果表明精确率 P 为 99.8%;召回率 R 为 100%;当 IOU 的阈值为 0.5 时,精度为 0.998;当 IOU 阈值以步长 0.05 从 0.5 增长到 0.95 时,精度为 0.960。通过 P 可以看出,DCN v3 - YOLO v7 模型对于鱼体可以实现精准检测,通过 R 看出模型没有出现漏检误检的情况,因此可以精准



Fig. 2 Underwater acquisition equipment of fishing ground

定位鱼体的中心坐标。

随后,将检测出的最优模型放在 detect.py 中运行,并且 save-txt 参数为开启状态,选取连续时间、平面鱼体视图较好的 1 000 幅图像进行检测,检测得到图像中鱼体锚框左上角和右下角的坐标,从而计算鱼体的中心坐标。再利用每一帧之间的时间差,计算出鱼体的速度、加速度、游动高度这 3 种数据,为后续随机森林的水质与鱼类行为双向作用模型提供丰富的输入数据^[20-22]。

(1) 鱼体速度

在正常水质中鱼类会正常游动,活动量不大,但是当水质处于异常情况时,鱼类可能就会出现加速游动的行为,此时活动量加大。当鱼类生存在经过严重污染的水质中便会出现游动速度降低的情况。因此,速度在一定程度上可以反映出鱼类受异常水质影响的情况。通过量化鱼体的速度可作为水质情况判断的依据。鱼类处于不同帧的位置能够反映其游动的情况,在实验过程中以视频左下角为原点建立坐标系。

鱼体不同帧游动距离计算式为

$$d = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \quad (2)$$

式中 (x_i, y_i) ——视频中某一帧鱼体中心坐标
 (x_{i+1}, y_{i+1}) ——视频中下一帧鱼体中心坐标
假设鱼体在两帧间做匀速直线运动,若视频帧率为 γ ,时间间隔为 Δf ,则鱼体运动时间 t 的计算式为

$$t = \frac{\Delta f}{\gamma} \quad (3)$$

则鱼体此时运动速度 v 计算式为

$$v = \frac{d}{t} \quad (4)$$

(2) 鱼体加速度

水质对鱼类加速度有着直接或间接的影响。良好的水质条件能够提供适宜的生存环境,促进鱼类正常的生理功能。相反,溶解氧含量、温度和 pH 值等参数偏离适宜范围时,鱼体加速度也会随之受到影响。因此,鱼体加速度也在一定程度上反映水质情况。加速度也可作为衡量水质的参数。

若鱼类在不同帧间的速度分别为 v_1 和 v_2 ,则鱼体加速度计算式为

$$\alpha = \frac{|v_2 - v_1|}{t_2 - t_1} \quad (5)$$

式中 t_1 ——某一帧图像所处的时间

t_2 ——下一帧图像所处的时间

(3) 鱼体游动高度

鱼体游动高度可间接地反映鱼类对水质的敏感程度,因此,将鱼体游动高度也作为一个量化参数。通过鱼类检测跟踪方法得出目标鱼体在某一时刻的坐标 (x, y) ,假定视频图像的高度为 H ,则此时鱼体游动高度用 h 表示,该游动高度是针对整个视频图像而言,具体计算式为

$$h = \frac{y}{H} \quad (6)$$

1.3 数据预处理

信息采集后需要对数据进行预处理,其中主要包括填补数据和异常数据纠错。滑动平均法经常作为数据完整性处理方法,可以有效地消除数据中的随机起伏,同时其计算过程也较为容易、耗时较短。因此本文利用滑动平均法来实现缺失数据填补、异常值纠错的操作^[23]。

(1) 填补缺失数据

数据在传输过程中难免会有一些数据出现空值的情况,模型中通过调用 ismissing 函数生成逻辑矩阵,再通过循环不断地更新矩阵,当检测到空白值时将窗口长度设置为 30 进行预测填补直至所有数据检测完毕。详细计算式为

$$A = \frac{y_1 + y_2 + \cdots + y_n}{n} \quad (7)$$

式中 $y_1, y_2, \cdots, y_n$ ——滑动窗口内的邻近观测值
 n ——滑动窗口的长度
 A ——填补值
(2) 异常值纠错
数据除了缺失,还会存在异常值,因此纠正异常

数据是必要步骤。使用滑动平均法对异常值进行纠错,纠正异常值是将其左右两边临近值的平均数作为修正值。计算式为

$$y_j = \begin{cases} \frac{y_{j-1} + y_{j+1}}{2} & (|y_{j+1} - y_j| > \alpha \text{ 或 } |y_j - y_{j-1}| > \beta) \\ y_j & (\text{其他}) \end{cases}$$

(8)

式中 y_j ——待检查的异常值
 y_{j+1} —— y_j 之后相差一个单元的数据
 y_{j-1} —— y_j 之前相差一个单元的数据
 α, β ——前后数据误差的阈值

即,当 y_j 与 y_{j+1} 差值的绝对值大于 α 或 y_j 与 y_{j-1} 差值的绝对值大于 β 时则应该被纠正为二者的平均值。

2 基于随机森林模型的双向映射模型构建方案

2.1 双向映射模型概述

双向映射模型既可以通过水质情况反映鱼类的行为也可以通过鱼类的行为反映水质的情况,若发现异常,可以尽早地进行控制,避免鱼类大量死亡造成巨大损失的情况^[24]。因此,执行双向映射方案十分必要,该模型采用基于随机森林算法的预测方案,从而达到双向预测的目的。

将采集到的鱼类行为与其对应水质参数的数据集打乱,以保证模型的有效性。以特定的比例进行抽取,将原始数据集划分为训练集与测试集。随后对数据集进行转置处理,将转置后的训练集归一化至 0~1 之间,并且将输入输出值再进行转置以适应随机森林模型,即保证每一行是一个样本,每一列是一个特征。建立随机森林回归模型,首先设置超参数即决策树数目为 100,每棵决策树最小叶子数为 3,再利用 Treebagger 函数建立模型(Treebagger 函数会返回一个随机森林模型对象 B,可以利用该对象进行预测)。模型建立后需对其进行测试,将训练集的输入与测试集的输入都放至模型中从而得到训练集与测试集的预测值,随后将这些数据进行反归一化得到需要的预测值。最后将决定系数(Coefficient of determination, R^2)、平均绝对误差(Mean absolute error, MAE)、平均偏差误差(Mean bias error, MBE)、平均绝对百分比误差(Mean absolute percentage error, MAPE)和均方根误差(Root mean square error, RMSE)作为评估指标来衡量模型的性能^[25]。随机森林模型结构如图 3 所示。

该模型的核心思想是首先在不同水质情况下监测鱼类的行为参数从而形成鱼类行为与水质参数相

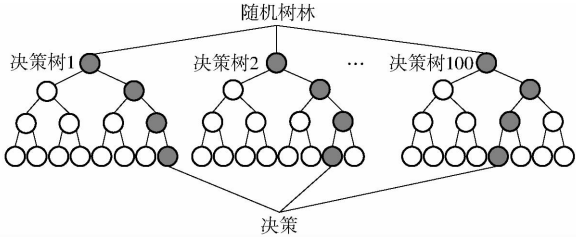


图 3 随机森林模型结构图

Fig. 3 Structural diagram of random forest

对应的数据集,接下来将水质情况与鱼类行为参数分别作为模型的输入进行训练以此达到预测双向映射的效果,即当鱼类行为参数作为输入时可以反映出水质的情况,同时当水质参数作为输入时也可以反映出鱼类的行为参数^[26]。最后按特定比例选取的测试集评估整个模型的性能。选取的鱼类特征分别为速度、加速度、游动高度;选取的水质参数分别为 pH 值、温度、溶解氧含量和盐度。

2.2 双向映射模型构建

随机森林由多棵独立训练的决策树构成,这样的结构可以提高模型的准确性和稳定性。在随机森林当中,每棵决策树都是基于随机选择的样本和特征进行训练的。随机森林会从训练数据集当中随机选择一定数量的样本(有放回的抽样),再从所有特征中随机选择几个特征进行训练。在处理回归问题时,随机森林对每棵独立决策树预测的结果求平均,作为回归预测结果;处理分类问题时,每棵决策树都会得出各自分类的结果,随机森林对不同预测结果采取投票的方式得出最终结果作为随机森林的输出值,即分类预测结果^[27]。模型预测流程为:

(1) 导入数据

```
res = xlsread('数据集.xlsx') % 导入数据集
```

(2) 数据归一化

```
[train, p-input] = mapminmax(P-train, 0, 1) % 将数据集放入 mapminmax 函数中归一化  
P-test = mapminmax('apply', p-test, p-input) % 将测试集按照归一化准则进行归一化
```

(3) 训练模型

```
Trees = 100 % 决策树数目为 100  
Leaf = 3 % 每棵决策树最小叶子数为 3  
OOBPredictorImportance = 'on' % 打开误差图
```

```
Method = 'regression' % 判断是分类实验还是回归实验
```

(4) 测试

```
sim1 = predict(net, p-train) % 得到训练集
```

预测值

```
sim2 = predict( net, p-test) % 得到测试集预测值
```

(5) 数据反归一化

```
Sim1 = mapminmax( 'reverse', sim1, p-output)
```

% 对步骤(4)得出的训练集预测值进行反归一化

```
Sim2 = mapminmax( 'reverse', sim2, p-output)
```

% 对步骤(4)得出的测试集预测值进行反归一化

(6) 绘制结果图。

在本研究也添加了基于随机森林算法的时间序列预测,时间序列分析考虑数据点之间的时间依赖性,时间序列模型可以捕捉到数据中的自相关性和趋势,从而提供更准确的预测结果。这意味着可以更好地预测未来的水质情况和鱼类行为,有助于养殖户提前采取相应措施来保护水质和鱼类的健康。基于随机森林模型的时间序列预测流程为:

(1) 导入数据

```
result = xlsread( '数据集.xlsx') % 导入数据集

(2) 数据分析



```
num_samples = length(result) % 样本个数
```


```

```
kim = 10 % 自变量个数
zim = 1 % 预测数
```

(3) 构建数据集。

(4) 数据归一化

```
[ p_train, p_input ] = mapminmax( P_train, 0, 1) % 将数据集放入 mapminmax 函数中归一化
P-test = mapminmax( 'apply', p-test, p-input)
```

% 将测试集按照归一化准则进行归一化

(5) 训练模型。

(6) 数据反归一化。

(7) 绘制实验结果。

随机森林的优点在于其随机性和其并行化,随机性主要体现在样本随机性和特征随机性。样本随机性是通过有放回抽样来实现的,而特征随机性则是通过随机选择特征来实现的。这种随机性可以减小过拟合的风险,提高模型的泛化能力^[28]。并行化体现在随机森林的决策树可以独立地训练,并行化处理可以提高模型的训练速度^[29-30]。

整个实验系统架构如图4所示。

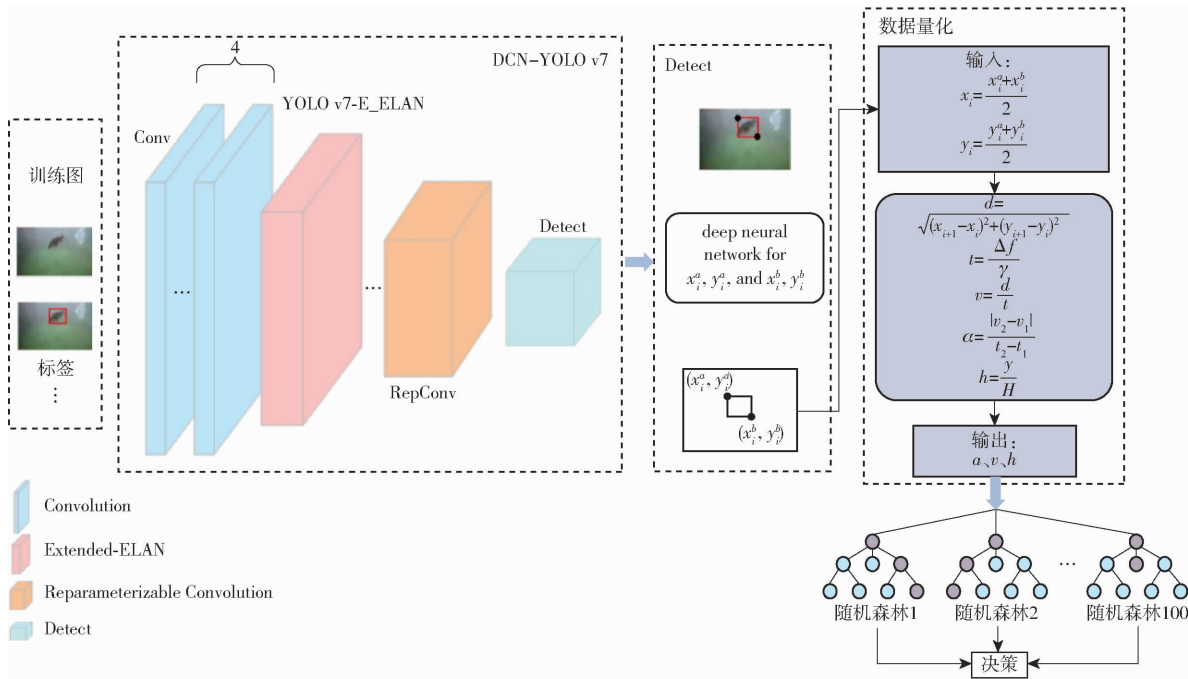


图4 系统架构图

Fig.4 System architecture diagram

3 实验结果与分析

首先按照比例4:1将数据集分为训练集和测试集。在训练过程中,将鱼类行为和水质参数作为输入,并将其与相应的标签进行配对进而使模型能够学习到鱼类行为与水质之间的关系。

2个数据集通过鱼体速度、加速度、游动高度预测溶解氧含量回归模型测试集测试100次的RMSE

预测结果可视化示意图如图5所示。

2个数据集通过温度、溶解氧含量、pH值和盐度来预测鱼类游动高度回归模型测试集测试100次的RMSE预测结果可视化示意图如图6所示。

上述2个数据集预测溶解氧含量回归实验测试集评估结果如表1所示。

上述2个数据集预测游动高度回归实验测试集评估结果如表2所示。

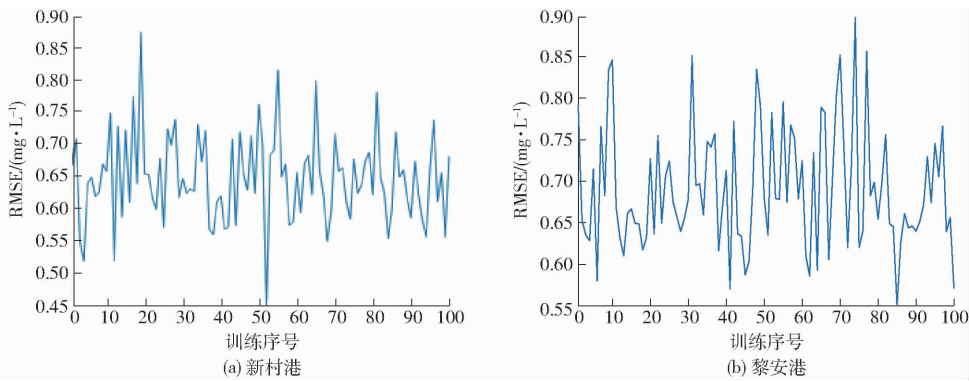


图 5 溶解氧含量测试集 RMSE 预测结果可视化示意图

Fig. 5 Visualization of RMSE prediction results of dissolved oxygen content test set

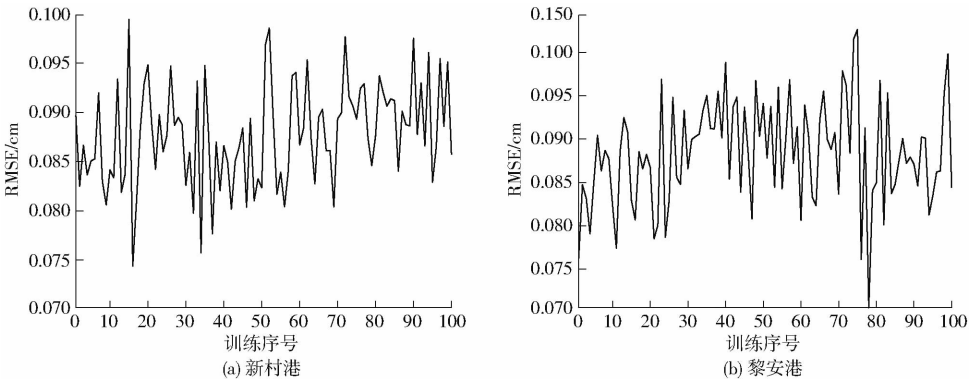


图 6 鱼类游动高度测试集 RMSE 预测结果可视化示意图

Fig. 6 Visualization of RMSE prediction results of fish swimming height test set

表 1 预测溶解氧含量回归实验测试集评估结果
Tab.1 Evaluation of dissolved oxygen content regression experimental test set evaluation results

数据集	RMSE/ (mg·L ⁻¹)	MAE/ (mg·L ⁻¹)	MBE/ (mg·L ⁻¹)	R ²
新村港	0. 652	0. 364	- 0. 020	0. 862
黎安港	0. 691	0. 565	- 0. 106	0. 819

表 2 预测鱼类游动高度回归实验测试集评估结果
Tab.2 Evaluation results of regression experimental test set for predicting fish swimming height

数据集	RMSE/cm	MAE/cm	MBE/cm	R ²
新村港	0. 088	0. 061	0. 010	0. 764
黎安港	0. 088	0. 059	- 0. 016	0. 755

针对回归实验,从上述实验结果可以看出,对于输入鱼类行为参数预测水质参数,新村港数据集预测模型具有较高的 R^2 , 达到 0. 862;黎安港数据集预测模型的 R^2 为 0. 819。这表明该模型能够较好地预测溶解氧含量。新村港数据集预测结果的 MAE 为 0. 364 mg/L;黎安港数据集预测结果的 MAE 为 0. 565 mg/L,这说明平均预测误差相对较小。而 2 个数据集的 MBE 分别为 - 0. 020 mg/L 和 - 0. 106 mg/L,同时 RMSE 分别为 0. 652 mg/L 和 0. 691 mg/L,这表示预测结果整体上与实际值接近。

同时,对于输入水质参数预测鱼类行为参数的实验在测试集上也有较低的 MAE、MBE 以及 RMSE, R^2 也比较接近于 1。这表明鱼类行为参数和水质参数可以作为可靠的预测指标,并且该模型在双向映射方面可以较好地挖掘二者之间的关系。

通过鱼体速度、加速度、游动高度预测溶解氧的分类模型预测结果示例图以及对分类模型测试集训练 100 次的预测准确率可视化示意图分别如图 7 和图 8 所示。

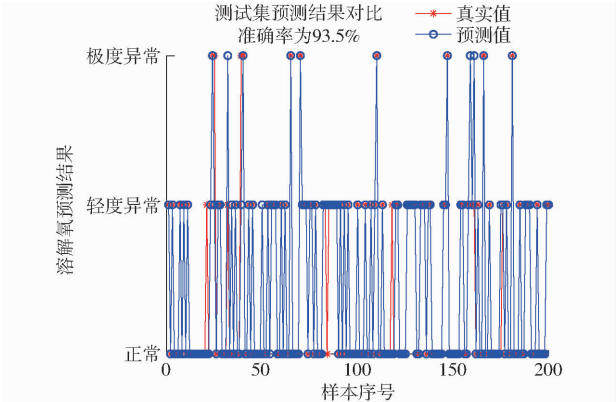


图 7 溶解氧含量测试集分类模型预测结果示例图
Fig. 7 Schematic diagram of prediction results of dissolved oxygen content test set classification model

通过温度、溶解氧含量、pH 值和盐度来预测鱼

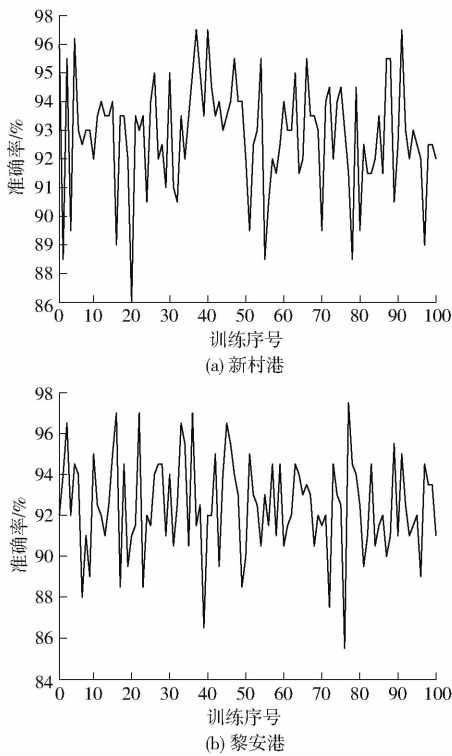


图 8 溶解氧含量测试集预测准确率可视化示意图
Fig.8 Schematics of accuracy of dissolved oxygen content test set

类游动速度的分类模型预测结果示例图以及对分类模型测试集训练 100 次的预测准确率可视化示意图分别如图 9 和图 10 所示。

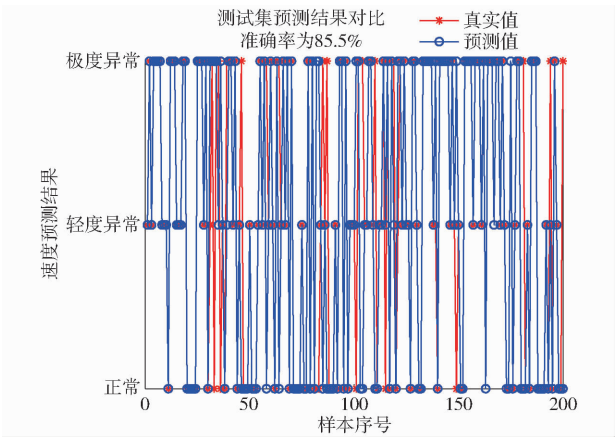


图 9 速度测试集分类模型预测结果示例图
Fig.9 Schematic diagram of prediction results of classification model of the swimming speed test set

针对分类实验,即输入水质参数反映鱼类行为异常级别和输入鱼类行为反映水质异常级别的实验结果显示,新村港数据集双向预测模型的平均准确率分别可达到 91.272% 和 92.848%;黎安港数据集双向预测模型的平均准确率分别可达到 92.461% 和 87.205%。这意味着该模型能够较为准确地预测异常级别,并且与实际情况具有较高的一致性。

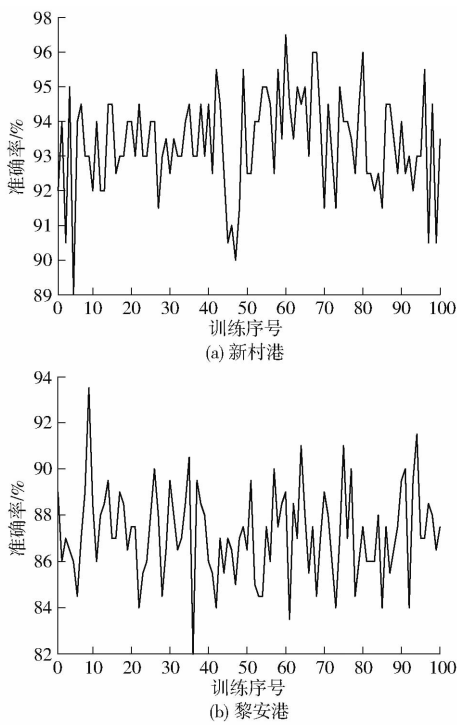


图 10 速度测试集预测准确率可视化示意图
Fig.10 Schematics of accuracy of swimming speed test set

2 个时间序列实验测试集训练 100 次结果的不同指标可视化示意图分别如图 11、12 所示。

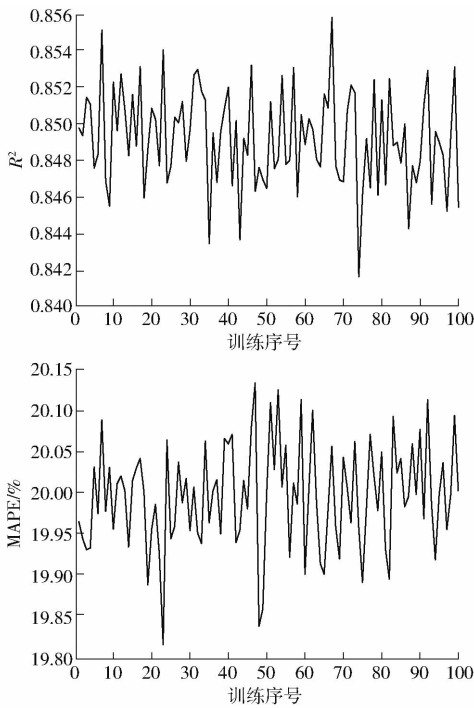


图 11 新村港时间序列实验测试集的可视化训练结果
Fig.11 Visual training results of time series experiment of Xincun Port

时间序列的预测实验可以通过时间序列分析对未来的水质变化和鱼类行为做出预测。经过 100 次训练后,2 个数据集的测试集 MAPE 平均值分别为 19.995% 和 17.636%, R^2 平均值分别为 0.849 和 0.871。实验结果显示,模型在预测水质参数方面表

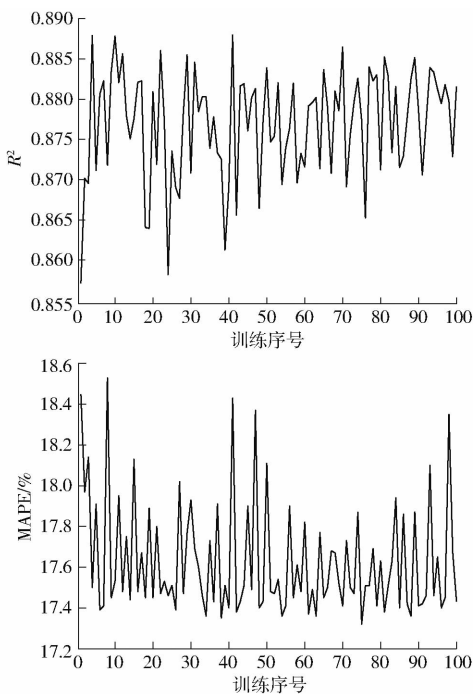


图 12 黎安港时间序列实验测试集的可视化训练结果

Fig.12 Visual training results of time series experiment of Li'an Port

现出了一定的准确性。时间序列分析的应用前景和优势明显。可以帮助养殖者更好地理解水质和鱼类行为之间的关系。

4 结束语

引入 DCN v3 在 YOLO v7 中骨干网络的 P5 层,取代原来 P5 层所有卷积核 3×3 的卷积层,使得在骨干网络中的卷积神经元之间共享权重,以适应特征图中鱼体的不规则形状,提高 bounding box 回归的精准性,使得在传输鱼体不同尺度下的模型训练过程更加稳定。

同时还通过引入随机性来构建多个不同的决策树从而减少了过拟合的风险。因此本实验将 DCN v3 - YOLO v7 与随机森林相结合,搭建了鱼类行为与水质参数的双向映射模型,它不仅可以提供更多的信息从而提高预测的准确性,而且还可以通过相互验证提高预测的可靠性。并且为了说明模型泛化能力,在 2 个数据集下进行了实验。分类实验平均准确率为 90.94%,回归实验决定系数 R^2 的平均值为 0.801。

参 考 文 献

[1] VILLON S, MOUILLOT D, CHAUMONT M, et al. A deep learning method for accurate and fast identification of coral reef fishes in underwater images[J]. Ecological Informatics,2018, 48:238 - 244.

[2] RAUF T H, LALI U I M, ZAHoor S, et al. Visual features based automated identification of fish species using deep convolutional neural networks[J]. Computers and Electronics in Agriculture,2019,167:105075.

[3] 王迎. 鱼体目标跟踪及运动分析在水质监测中的应用[D]. 秦皇岛:燕山大学,2022:31 - 61.

WANG Ying. Application of fish target tracking and motion analysis in water quality monitoring[D]. Qinhuangdao: Yanshan University, 2022:31 - 61. (in Chinese)

[4] 芦玥同,黄毅,刘金涛,等. 基于鱼类应激反应的水质在线监测预警技术研究进展[J]. 环境污染与防治,2022,44(11): 1526 - 1530.

LU Yuetong, HUANG Yi, LIU Jintao, et al. Research progress of water quality online monitoring and early warning technology based on fish stress response[J]. Environmental Pollution and Control, 2022,44(11):1526 - 1530. (in Chinese)

[5] 张鼎渊. 基于 EEMD - LSTM 水质预测方法的研究与应用[D]. 合肥:中国科学技术大学, 2021:18 - 25.

ZHANG Dingyuan. Research and application of water quality prediction method based on EEMD - LSTM[D]. Hefei: University of Science and Technology of China, 2021: 18 - 25. (in Chinese)

[6] LIU S, TAI H, DING Q, et al. A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction[J]. Mathematical and Computer Modelling, 2013, 58(3 - 4): 458 - 465.

[7] LIU S, XU L, LI D. Prediction of aquaculture water quality based on combining principal component analysis and least square support vector regression[J]. Sensor Letters, 2013, 11(6 - 7): 1305 - 1309.

[8] 杨文博. 面向水产养殖溶解氧浓度的组合预测模型研究[D]. 烟台:山东工商学院,2023: 5 - 6.

YANG Wenbo. Research on combination prediction model of dissolved oxygen concentration in aquaculture [D]. Yantai: Shandong Technology and Business University, 2023: 5 - 6. (in Chinese)

[9] CHEN Z, HU Z, XU L, et al. DA - Bi - SRU for water quality prediction in smart mariculture[J]. Computers and Electronics in Agriculture, 2022, 200: 107219.

[10] FU Yuexin, HU Zhuhua, ZHAO Yaochi, et al. A long-term water quality prediction method based on the temporal convolutional network in smart mariculture[J]. Water, 2021, 13(20): 2907.

[11] HU Zhuhua, LI Ruqing, XIA Xin, et al. A method overview in smart aquaculture [J]. Environmental Monitoring and Assessment,2020,192(8): 1 - 25.

[12] HAN Bing, HU Zhuhua, SU Zhengwei, et al. Mask_LaC R - CNN for measuring morphological features of fish [J]. Measurement,2022,203: 111859.

[13] YU Chuang,HU Zhuhua,HAN Bing, et al. An intelligent measurement scheme for basic characters of fish in smart aquaculture [J]. Computers and Electronics in Agriculture,2023, 204:107506.

[14] 胡祝华,张逸然,赵瑶池,等. 权重约束 AdaBoost 鱼眼识别及改进 Hough 圆变换瞳孔智能测量[J]. 农业工程学报,2017,33(23):226-232.
HU Zhuhua, ZHANG Yiran, ZHAO Yaochi, et al. Fish eye recognition based on weighted constraint AdaBoost and pupil diameter automatic measurement with improved Hough circle transform[J]. Transactions of the CSAE, 2017,33(23):226-232. (in Chinese)

[15] 刘斌,王凯歌,李晓蒙,等. 基于语义部位分割的条纹斑竹鲨鱼体运动姿态解析[J]. 农业工程学报,2021,37(3):179-187.
LIU Bin, WANG Kaige, LI Xiaomeng, et al. Motion posture parsing of *Chiloscyllium plagiosum* fish body based on semantic part segmentation[J]. Transactions of the CSAE, 2021,37(3):179-187. (in Chinese)

[16] WANG W, DAI J, CHEN Z, et al. Internimage: exploring large-scale vision foundation models with deformable convolutions[J]. arXiv preprint, arXiv:2211.05778,2022.

[17] 孙月莹,陈俊霖,张胜茂,等. 基于改进 YOLOv7 的毛虾捕捞渔船作业目标检测与计数方法[J]. 农业工程学报,2023,39(10):151-162.
SUN Yueying, CHEN Junlin, ZHANG Shengmao, et al. Target detection and counting method for *Acetes chinensis* fishing vessels operation based on improved YOLOv7[J]. Transactions of the CSAE, 2023,39(10):151-162. (in Chinese)

[18] 巩龙勤. 基于深度学习的水下鱼体跟踪研究[D]. 海口:海南大学,2023:34.
GONG Longqin. Research on underwater fish tracking based on deep learning[D]. Haikou: Hainan University,2023:34. (in Chinese)

[19] ZHAO Shili,ZHANG Song,LIU Jincun, et al. Application of machine learning in intelligent fish aquaculture: a review[J]. Aquaculture,2021,540:736724.

[20] GONG L, HU Z, ZHOU X. A real time video object tracking method for fish[C]//Proceedings of the 2022 6th International Conference on Machine Learning and Soft Computing, 2022:78-84.

[21] 范伟康. 基于鱼群行为分析的水质监测系统研究[D]. 杭州:浙江工业大学,2015:8-11.
FAN Weikang. Research on water quality monitoring system based on the behavior of fish school[D]. Hangzhou: Zhejiang University of Technology,2015:8-11. (in Chinese)

[22] 彭红梅. 基于生物水质监测的鱼体运动状态检测系统研究[D]. 西安:西安邮电大学,2017:22-31.
PENG Hongmei. Research on fish movement state detection system based on biological water quality monitoring[D]. Xi'an: Xi'an University of Posts & Telecommunications,2017:22-31. (in Chinese)

[23] 袁自祥. 水利设施时间序列监测数据预测方法研究[D]. 绵阳:西南科技大学,2023:13-17.
YUAN Zixiang. Research on time-series prediction methods for monitoring data of water conservancy facilities[D]. Mianyang: Southwest University of Science and Technology, 2023:13-17. (in Chinese)

[24] 孙龙清,王泊宁,王嘉煜,等. 基于 G-RepVGG 和鱼类运动行为的水质监测方法[J]. 农业机械学报,2022,53(增刊2):210-218.
SUN Longqing, WANG Boning, WANG Jiayu, et al. Water quality monitoring based on fish movement behavior and G-RepVGG[J]. Transactions of the Chinese Society for Agricultural Machinery,2022,53(Supp.2):210-218. (in Chinese)

[25] 王涌,陆卫,左楚涵,等. 基于改进随机森林模型的水质 BOD 快速预测研究[J]. 传感技术学报,2021,34(11):1482-1488.
WANG Yong, LU Wei, ZUO Chuhan, et al. Research on water quality BOD prediction based on improved random forest model[J]. Chinese Journal of Sensors and Actuators,2021,34(11):1482-1488. (in Chinese)

[26] 李旭杰,史灵,花思洋,等. 基于随机森林的水质监测指标预测[J]. 江苏水利,2022(5):6-10.
LI Xujie, SHI Ling, HUA Siyang, et al. Water quality monitoring indicators prediction based on random forests[J]. Jiangsu Water Resources, 2022(5):6-10. (in Chinese)

[27] 叶丽珠,郑冬花,刘月红,等. 基于鲸群优化随机森林算法的非平衡数据分类[J]. 南京邮电大学学报(自然科学版),2022,42(6):99-105.
YE Lizhu, ZHNEG Donghua, LIU Yuehong, et al. Unbalanced data classification based on whale swarm optimization random forest algorithm[J]. Journal of Nanjing University of Posts and Telecommunications(Natural Science),2022,42(6):99-105. (in Chinese)

[28] 李道亮,姜国旗,杨建安,等. 基于计算机视觉的鱼类低氧胁迫行为检测与跟踪算法[J]. 农业机械学报,2023,54(10):399-406.
LI Daoliang, JIANG Guoqi, YANG Jian'an, et al. Detection and tracking algorithm of fish hypoxia stress behavior based on computer vision[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023,54(10):399-406. (in Chinese)

[29] 张霄. 基于随机森林的森林火灾风险等级预测的研究与应用[D]. 南昌:南昌大学,2022:22-25.
ZHANG Xiao. Research and application of forest fire risk grade prediction based on random forest[D]. Nanchang: Nanchang University,2022:22-25. (in Chinese)

[30] 项涛,董响红,郭超,等. 基于随机森林模型预测气候变化对黄颡鱼适生区的影响[J]. 水生生物学报,2019,43(6):1262-1270.
XIANG Tao,DONG Xianghong,GUO Chao, et al. The impacts of climate change on the distribution of *Pelteobagrus fulvidraco* by random forest model[J]. Acta Hydrobiologica Sinica,2019,43(6):1262-1270. (in Chinese)