

# 基于深度残差网络的水面漂浮物检测

张德民, 陈玉兰, 聂方彦\*

贵州商学院计算机与信息工程学院, 贵州 贵阳

收稿日期: 2025年6月2日; 录用日期: 2025年6月24日; 发布日期: 2025年7月3日

## 摘要

水面漂浮物作为水体污染的重要来源, 严重威胁着水域生态安全和人类健康。针对现有检测方法在复杂水面场景下存在特征提取不足和小样本适应性弱的问题, 本研究提出了一种改进的深度残差网络检测框架。首先构建自主采集的多源水面漂浮物数据集, 通过自适应图像裁剪和归一化处理消除光照波动与尺度差异的干扰; 继而基于ResNet-18架构设计分层优化策略, 冻结浅层网络参数保留通用特征表征能力, 通过微调深层网络增强域适应特征提取, 并构建复合分类头提升小样本学习性能。针对数据集规模限制, 创新性融合多维度协调数据增强策略与类别平衡损失函数, 通过在线增强扩充数据多样性并缓解类别不平衡问题。实验表明, 该方法在自建数据集上取得98.71%的检测准确率, 较基线模型提升9.68个百分点, 且在不同光照条件和漂浮物密度的测试场景中均保持稳定性能。研究系统验证了从数据预处理、网络架构优化到训练策略设计的全流程方案有效性, 提出的迁移学习框架在保证模型轻量化的同时显著提升检测鲁棒性, 为构建智能化水面污染监测系统提供了可靠的技术方案。

## 关键词

漂浮物检测, 卷积神经网络, 深度学习, 残差网络

# Detection of Floating Objects on Water Surface Based on Depth Re-Sidual Network

Demin Zhang, Yulan Chen, Fangyan Nie\*

Computer and Information Engineering College, Guizhou University of Commerce, Guiyang Guizhou

Received: Jun. 2<sup>nd</sup>, 2025; accepted: Jun. 24<sup>th</sup>, 2025; published: Jul. 3<sup>rd</sup>, 2025

## Abstract

As a significant source of water pollution, floating objects on water surfaces pose severe threats to aquatic ecological security and human health. To address the limitations of existing detection

\*通讯作者。

methods, such as insufficient feature extraction in complex aquatic scenarios and weak adaptability to small-sample conditions, this study proposes an improved deep residual network-based detection framework. Firstly, a self-collected multi-source dataset of floating objects was constructed, where adaptive image cropping and normalization processing were applied to mitigate interference from illumination fluctuations and scale variations. Subsequently, based on the ResNet-18 architecture, a hierarchical optimization strategy was designed: shallow network parameters were frozen to retain general feature representation capabilities, while deep networks were fine-tuned to enhance domain-adaptive feature extraction, coupled with a composite classification head to improve small-sample learning performance. To overcome dataset size constraints, a novel integration of multi-dimensional coordinated data augmentation strategies and class-balanced loss functions was implemented, expanding data diversity through online augmentation and alleviating class imbalance issues. Experimental results demonstrate that the proposed method achieves a detection accuracy of 98.71% on the self-built dataset, representing an improvement of 9.68 percentage points over the baseline model, while maintaining stable performance across varying illumination conditions and floating object densities. The study systematically validates the effectiveness of the full-process solution encompassing data preprocessing, network architecture optimization, and training strategy design. The proposed transfer learning framework significantly enhances detection robustness while ensuring model lightweight design, providing a reliable technical solution for intelligent water pollution monitoring systems.

## Keywords

Floating Object Detection, Convolutional Neural Network, Deep Learning, Residual Network

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在现代社会发展进程中, 环境保护已成为全球关注的焦点, 其中水资源的保护尤为关键, 它不仅关系到人类的生存, 也是可持续发展的核心要素。水面漂浮物, 作为水体污染的主要来源之一, 通常由不当处理的生活垃圾引起, 它们不仅破坏了水体的生态平衡, 还严重影响了水域的景观和资源的可持续利用。因此, 开发有效的水面漂浮物检测技术对于环境保护和水资源管理具有重要意义。

随着图像检测技术的不断进步, 卷积神经网络已成为图像识别领域的核心技术[1]。深度残差网络通过引入跨层恒等映射机制, 有效缓解了梯度消失问题, 其层级特征提取能力可模拟人类视觉系统的多尺度处理模式, 尤其擅长捕捉图像中的局部细节与全局语义特征。本研究基于 ResNet-18 架构, 提出一种针对小样本数据的改进模型。通过冻结浅层卷积核参数保留通用特征表示能力, 微调深层网络适配漂浮物检测任务, 并结合多维度协调数据增强策略与类别平衡策略, 显著提升模型在真实场景下的鲁棒性[2]。

为了训练和验证模型, 本研究采集了包含正负样本的数据集, 其中正样本包括多种常见生活垃圾, 如不同品牌的塑料瓶(6~7种)和不同类型的塑料袋(3~4种)。由于数据集是在开放环境中使用手机拍摄的, 图像尺寸较大且含有噪声, 这些因素可能会干扰特征提取, 对此, 我们在数据预处理阶段通过裁剪和提取操作, 获取了噪声影响较小的图片, 以提高模型训练的准确性。最终, 通过模型搭建与训练, 我们获得了 98.71% 的验证准确率, 这一结果为水资源管理和环境保护提供了有力的数据支持和技术保障。

## 2. 问题描述与数据集

### 2.1. 数据集来源与构成

本研究的数据集采集自贵州商学院静心湖，旨在构建一个针对水面漂浮物的检测模型。数据集的构建是通过静心湖湖面进行实地拍摄，收集了一系列含有水面漂浮物的图像，以此作为模型训练和验证的基础。数据集中的图像包括正样本和负样本，其中正样本为含有水面漂浮物的图片，主要包含多种常见生活垃圾，如不同品牌的塑料瓶(6~7种)和不同类型的塑料袋(3~4种)等，如图1所示。负样本则为不含有水面漂浮物的图像。这些图像的采集均采用手机在开放环境下进行，因此图像尺寸较大，且存在不同程度的噪声(如岸边植被倒影)，这些噪声可能对后续的特征提取造成干扰。



Figure 1. Example of positive sample  
图 1. 正样本示例

### 2.2. 数据预处理

由于采集的图像存在尺寸不一和噪声干扰的问题，我们采取了一系列预处理措施以优化模型的输入数据。首先，所有图像均以漂浮物为中心，人工裁剪为原图面积的60%~80%，以移除无关背景并聚焦于包含漂浮物的关键区域。随后，通过python脚本将这些图像压缩至统一的 $512 \times 512$ 像素尺寸，得到如图2所示部分数据集图像，这确保了输入到模型中的图像具有一致性，从而便于后续的图像输入操作。



Figure 2. Sample dataset  
图 2. 数据集示例

### 3. 深度残差网络架构与原理

#### 3.1. 改进的残差单元设计

本研究基于 ResNet-18 架构构建分层迁移学习模型，如公式(1)所示：

$$y = F(x, W_i) + H(x) \quad (1)$$

其中： $H(x)$  为恒等映射，保留原始输入特征； $F(x, W_i)$  为残差函数，由多层卷积操作构成，具体设计如下：

##### 1) 参数冻结策略

为避免小样本训练导致的浅层特征退化，采用分层迁移学习策略：冻结浅层卷积参数(layer1、layer2)，保留预训练模型在 ImageNet 数据集上学习到的通用边缘或纹理特征[3]；解冻深层特征提取层(layer3、layer4)，通过微调使其适配水面漂浮物的域特异性语义特征(如半透明塑料袋的光反射模式)。

##### 2) 重构全连接层适配二分类任务，设计复合分类头

```
mn.Sequential(
  mn.AdaptiveAvgPool2d((1, 1)), # 全局平均池化
  mn.Flatten(), # 展平特征图
  mn.Linear(512, 512), # 特征重组
  mn.BatchNorm1d(512), # 加速收敛
  mn.SiLU(), # 平滑梯度流
  mn.Dropout(0.5), # 随机失活
  mn.Linear(512, 256), # 降维投影
  mn.Linear(256, 2) # 分类输出
)
```

该结构不仅实现了特征重组，还通过 BN、激活和 Dropout 等操作缓解了过拟合问题，有效地降低了维度。

#### 3.2. 多维度协调数据增强策略

为提升模型对复杂场景的泛化能力，本研究设计了固定组合式多维度数据增强方案，通过空间变换、光度扰动和信息遮蔽三个维度的协同作用构建鲁棒特征空间。所有训练样本均按确定性顺序依次通过以下增强流程：

##### 1) 几何变换

图像翻转：包括水平翻转( $p=0.5$ )和垂直翻转( $p=0.5$ )，打破图像的方向对称性。如图 3 第一行 Aug4 中的水平镜像翻转和第三行 Aug3 中的垂直翻转，增强了模型对方向变化的适应性。

随机旋转：调整图像的旋转角度，打破图像的方向性，增强模型对不同角度输入的适应性  $\theta \sim \mu(-30^\circ, 30^\circ)$ ，如图 3 第二行 Aug2 所示，旋转后的图像保持主体特征完整，验证了模型对角度变化的适应性。

仿射平移：对图像进行随机的平移操作  $(t_x, t_y) \sim \mu(-0.1, 0.1) \times$  图像尺寸，模拟图像在空间中的位置变化，进一步提升模型对局部特征的捕捉能力[4]，如图 3 第三行 Aug4 中的平移操作，目标物体在位置偏移后仍保留关键局部特征，增强了模型对空间位置变化的鲁棒性。

##### 2) 光度调整

颜色抖动：调整图像的色彩通道，包括亮度、对比度、饱和度和色调[5]，调制范围均为  $\Delta = 0.15$ ，其数学计算范围公式为： $\Delta \sim \mu(1-0.15, 1+0.15) = \mu(0.85, 1.15)$ ，其他通道调整公式类似，调制参数独立采样。

高斯噪声：在图像中添加零均值高斯噪声[6]  $\epsilon \sim N(0, \sigma^2)$ ，标准差为  $\sigma \sim \mu(0, 0.1)$ 。模拟实际场景中的噪声干扰，图 3 第三行 Aug 2 展示了添加噪声后的模糊效果，显著提升了模型的抗噪能力。

数据标准化：采用 ImageNet 预训练模型的标准统计量进行数据标准化，其参数为  $mean = [0.485, 0.456, 0.406]$ ， $std = [0.229, 0.224, 0.225]$ ，加速模型收敛。

### 3) 遮挡增强

随机擦除：概率 25%选择图像中的一个区域并将其擦除，擦除区域的宽高比和比例均随机确定[7]，擦除比例  $\lambda \sim \mu(0.02, 0.1)$ ，宽高比  $r \sim \mu(0.3, 3.3)$ ，如图 3 第二行 Aug 4 所示，遮挡后的图像迫使模型关注未被擦除的局部信息，提升对部分信息缺失的鲁棒性。

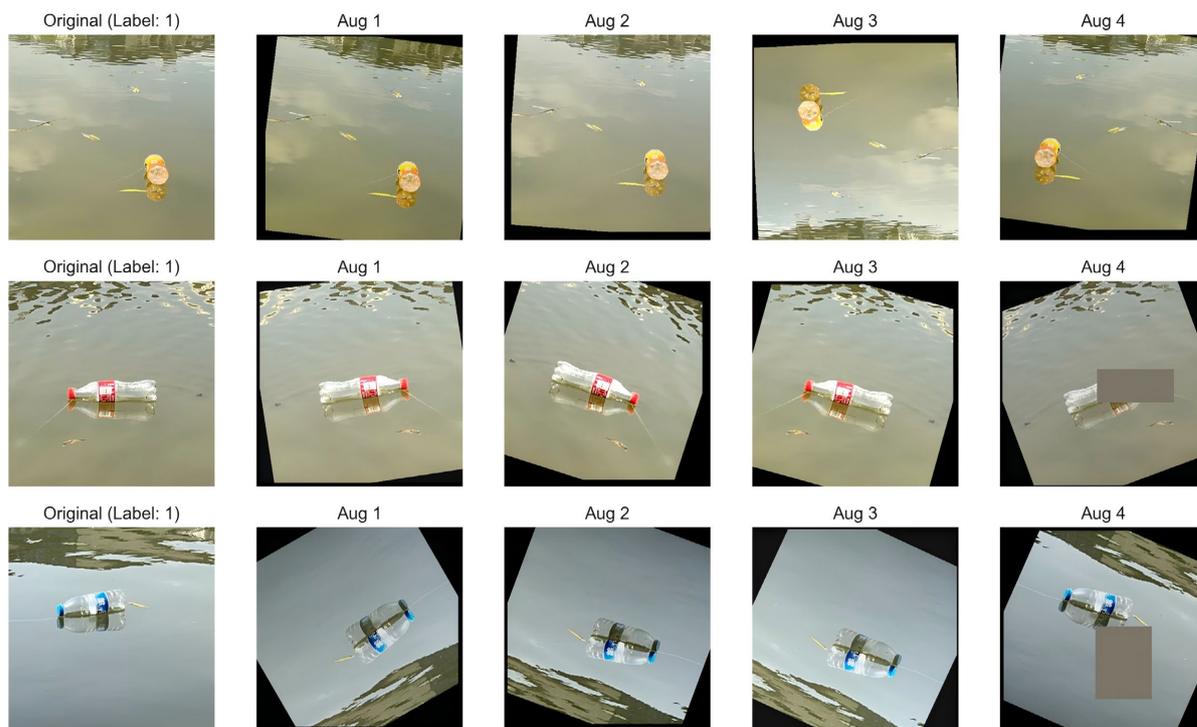


Figure 3. Data enhancement examples  
图 3. 数据增强示例

## 3.3. 训练优化策略

### 3.3.1. 类别平衡损失函数

在二分类任务中，传统的交叉熵损失函数的表达式为：

$$\mathcal{L}_{loss} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (2)$$

其中： $y_i \in \{0, 1\}$ ，为样本的真实标签； $p_i \in [0, 1]$ ，为模型预测样本属于正类(有漂浮物)的概率。

针对数据集类别数量不平衡问题(负样本占比 35%)，我们通过反比例公式计算类别权重来改进传统的交叉熵损失函数，如下：

对正负样本赋予不同的权重，平衡数量的差异[8]。引入类别权重  $\alpha_i$  在公式(3)中：

$$\mathcal{L}_{\alpha} = -\frac{1}{N} \sum_{i=1}^N \alpha_i [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (3)$$

其中:

$$\alpha_i = \begin{cases} \alpha_{\text{正}} = \frac{\text{总样本数}}{\text{类别数} \times \text{正样本数}} \\ \alpha_{\text{负}} = \frac{\text{总样本数}}{\text{类别数} \times \text{负样本数}} \end{cases} \quad (4)$$

### 3.3.2. 优化器配置

在优化器的选择上, 我们采用 AdamW 优化器, 参数更新规则为:

$$\theta_{t+1} = \theta_t - \eta \left( \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} + \lambda \theta_t \right) \quad (5)$$

关键超参数:

初始学习率:  $\eta = 1e^{-4}$ ;

权重衰减:  $\lambda = 1e^{-4}$ ;

动量参数:  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ 。

### 3.3.3. 学习率调度

我们采用余弦退火策略, 通过周期性地调整学习率, 使模型在训练过程中能够更好地收敛[9], 调整公式为:

$$\eta_t = \eta_{\min} + \frac{1}{2}(\eta_{\max} - \eta_{\min}) \left( 1 + \cos\left(\frac{t}{T} \pi\right) \right) \quad (6)$$

其中:

初始学习率:  $\eta_{\max} = 1e^{-4}$ ;

最小学习率:  $\eta_{\min} = 1e^{-6}$ ;

周期长度:  $T_{\max} = 30 \times N_{\text{batch}}$  (总训练轮数)。

## 4. 实验及结果分析

### 4.1. 实验环境与模型配置

#### 4.1.1. 实验设备参数配置

基于卷积神经网络的目标检测模型较为复杂, 具有大量参数, 因此需要强大的计算机性能来进行数据运算和权重值的训练。本研究采用 Pytorch 作为深度学习框架来构建和训练模型, 具体实验设备配置参数如表 1 所示, 其中 NVIDIA RTX 3060 GPU 的并行计算能力显著加速了模型训练过程。

**Table 1.** Experimental equipment parameter configuration

**表 1.** 实验设备参数配置

配置	参数
处理器	Intel i5-11400H @2.70GHz
GPU	NVIDIA GeForce RTX 3060 Laptop GPU
操作系统	Windows11 专业版
开发语言	Python

续表

深度学习框架	Pytorch 1.12.1 + TorchVision 0.13.1
CUDA 版本	CUDA 11.6

#### 4.1.2. 模型改进方案

如表 2 所示,本研究通过三阶段优化策略提升模型性能:(1) 特征提取增强:解冻 ResNet-18 的 layer3-layer4 权重,增强深层语义特征表达能力;(2) 分类头压缩:采用全连接层降维结构(512→256→2)并引入 Dropout 层( $p=0.5$ ),缓解过拟合;(3) 损失函数优化:设计动态类别加权交叉熵(DCW Loss),解决类别不平衡问题。

Table 2. Comparison of model architectures

表 2. 模型架构对比

组件	原始模型	改进模型(ResNet-18)
特征提取器	全卷积层冻结	解冻 layer-layer4
分类头	FC-2→2	FC-512→256→2 + 丢弃层
损失函数	交叉熵损失	DCW Loss 加权交叉熵
优化策略	Adam ( $lr = 1e-3$ )	AdamW ( $lr = 1e-4$ ) + 余弦退火

## 4.2. 训练过程分析

### 4.2.1. 超参数设置

采用梯度累积策略缓解显存限制,具体参数如表 3 所示。其中学习率采用余弦退火调度,初始值从  $1e-4$  逐渐衰减至  $1e-6$ ,动量因子保持 0.9 不变。具体如下。

Table 3. Training hyperparameter settings

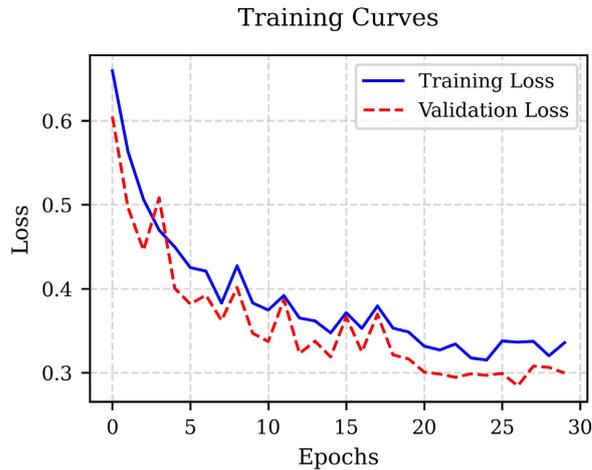
表 3. 训练超参数设置

参数类型	参数值
输入图像尺寸	512 × 512 像素
批量大小	16 (梯度累积步长 4)
学习率	$1e-4 \rightarrow 1e-6$ (余弦退火)
动量因子	$\beta_1 = 0.9, \beta_2 = 0.999$
训练轮数	30
权重衰减系数	$1e-4$
标签平滑系数	0.1

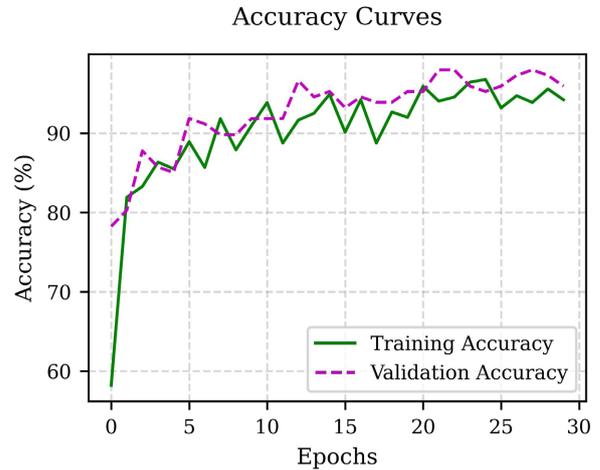
### 4.2.2. 收敛特性优化

在训练策略对比分析中,本文改进模型(ResNet-18)与多种基线模型(ResNet-50、VGG-16、

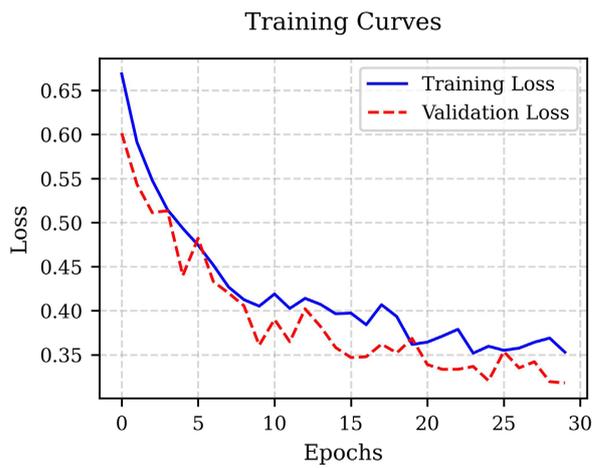
EfficientNet\_B0)进行了系统性对比。所有模型均采用相同的固定组合式多维度数据增强方案,而基线模型 ResNet-18(无数据增强)仅保留基础增强策略(水平翻转概率  $p=0.5+$  标准化)。实验结果表明,不同模型架构在复杂增强策略下的训练动态与泛化能力存在显著差异。



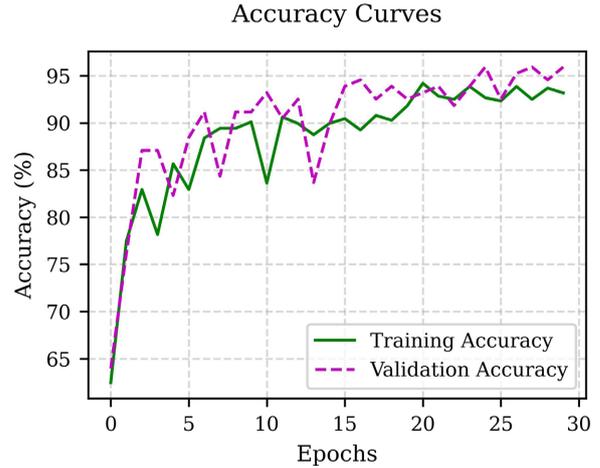
(a) ResNet-18 (增强)损失值曲线



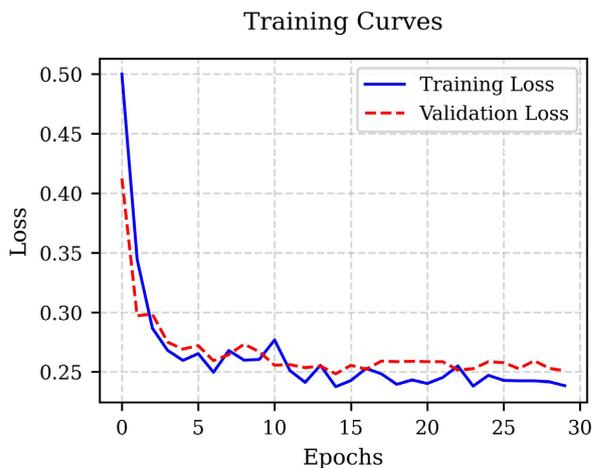
(b) ResNet-18 (无增强)准确率百分比



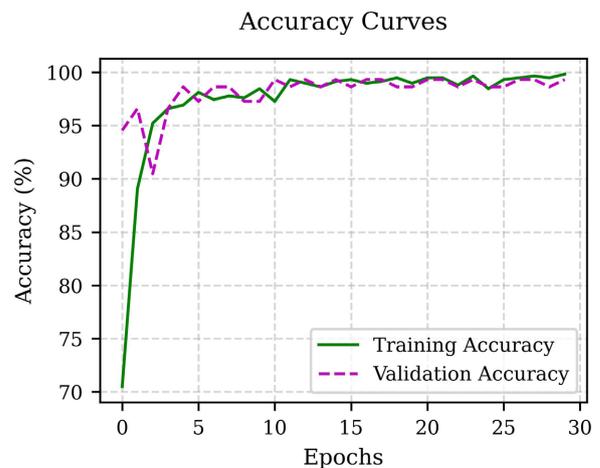
(c) ResNet-18 (基础增强)损失值曲线



(d) ResNet-18 (无基础增强)准确率百分比



(e) 改进模型(ResNet-18)损失值曲线



(f) 改进模型(ResNet-18)准确率百分比

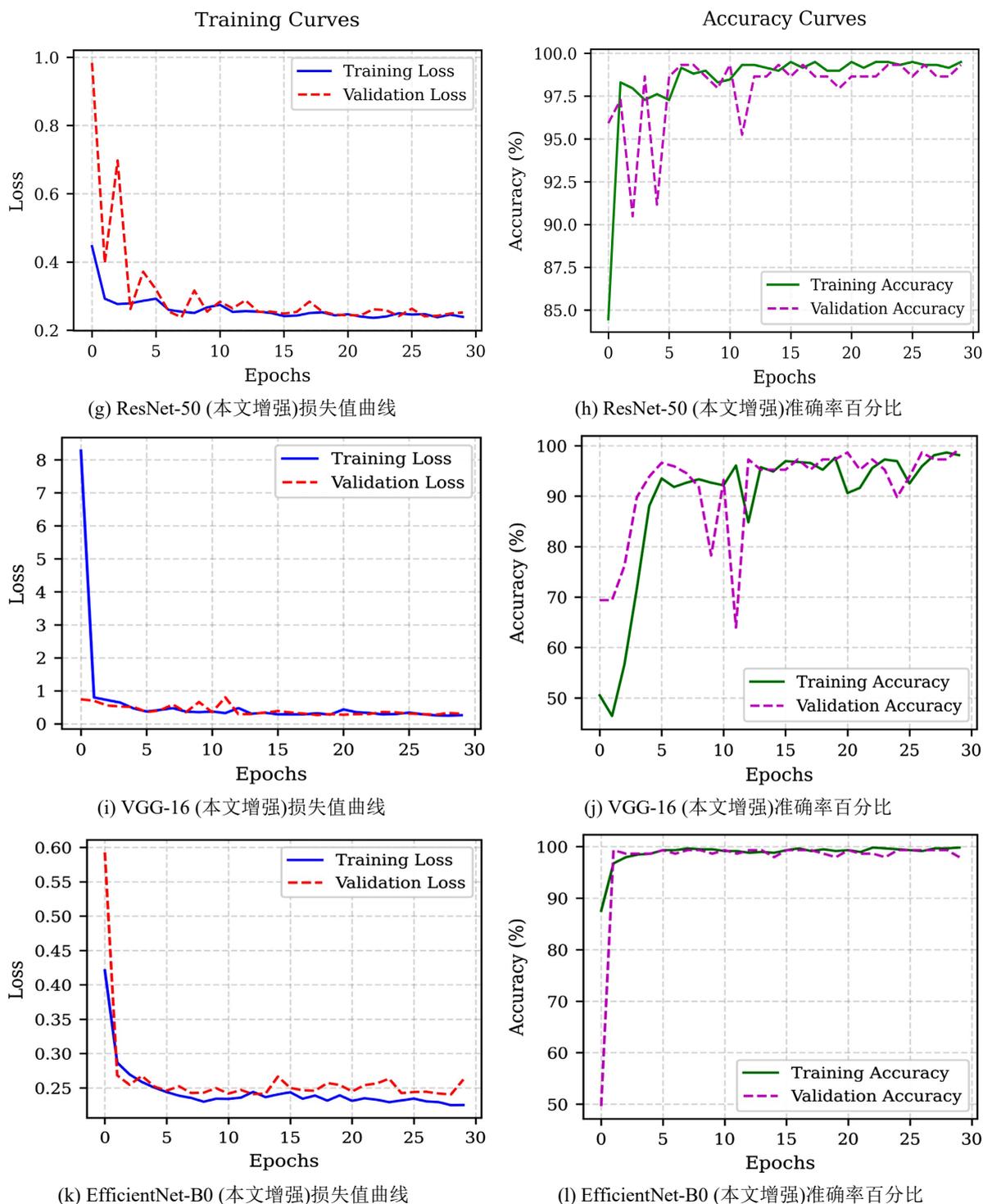


Figure 4. Loss value and accuracy percentage curve

图 4. 损失值和准确率百分比曲线

## 1) ResNet-18 (本文数据增强) 的局限性

当数据增强策略的噪声分布与真实场景复杂度匹配时, ResNet-18 (本文数据增强) 表现出显著的不稳定性。其验证损失曲线在 0.25~0.55 区间剧烈震荡(见图 4(c)), 训练准确率波动幅度达 8.7%(见图 4(d))。

这种现象可能源于两方面：其一，复杂增强策略(如随机擦除和仿射变换)引入的噪声模式超出原始模型的特征能力；其二，确定性增强序列导致的特征空间突变，使模型在局部最优解之间频繁震荡。

### 2) 改进模型(ResNet-18)的优化效果

改进模型(ResNet-18)通过解冻卷积层赋予特征空间动态重构能力，深层分类头提供足够的非线性表达能力，而 DCW 损失与 AdamW 优化则确保训练过程在复杂噪声干扰下保持收敛稳定性，其展现出更优的收敛特性。损失从 0.50 降至约 0.15，验证损失从 0.45 降至约 0.20(见图 4(e))，训练和验证准确率均提升至 97%以上(见图 4(f))，显示出改进策略在提升模型收敛速度和准确率方面的有效性，尤其是在验证集上的泛化能力得到了显著增强。

### 3) 扩展模型对比分析

为进一步验证方法的鲁棒性，引入 ResNet-50、VGG-16 和 EfficientNet\_B0 进行对比。ResNet-50：深层网络(50 层)凭借更强的特征提取能力，损失值收敛速度最快(见图 4(g))，但因模型复杂度高，验证准确率出现周期性震荡(见图 4(h))，表明过拟合风险加剧。VGG-16：受限于全连接层的冗余参数，其损失值收敛速度较快且波动平缓(见图 4(i))，但验证准确率仍存在震荡(见图 4(j))，推测因浅层特征表达能力不足导致噪声敏感性升高。EfficientNet\_B0：通过复合缩放策略平衡深度、宽度与分辨率，其损失值与准确率曲线均优于原始 ResNet-18(见图 4(k)、图 4(l))，但在测试集当中，改进模型(ResNet-18)在识别上相较于该模型较好。

## 4.3. 模型性能评估

根据表 4 的实验对比，改进模型(ResNet-18)在引入适配真实场景的数据增强策略后，相较于基线模型(ResNet-18 无数据增强)与对比模型(ResNet-18 + 数据增强)实现了全方位性能突破。为进一步验证方法的鲁棒性，引入 ResNet-50、VGG-16 和 EfficientNet-B0 进行横向对比，关键结论如下：

### 1) 改进模型(ResNet-18)的综合优势

全局准确率：改进模型(ResNet-18)达到 98.71%，较基线模型(ResNet-18 无增强)提升 9.68% (89.03%→98.71%)，测试集 155 个样本中正确分类数从 138 跃升至 153。小样本类别优化：在“无漂浮物”类别(46 个样本)中，召回率从 65%飙升至 100%，彻底消除漏检问题，F1 分数提升 25.64% (0.78→0.98)。多数类稳定性：在“有漂浮物”类别(109 个样本)中，精确率达到 1.00 (零误报)，F1 分数提升至 0.99 (+6.45%)。均衡性提升：Macro 平均 F1 提升 15.29% (0.85→0.98)，加权平均 F1 提升 12.5% (0.88→0.99)，在样本分布不均衡(正负样本比 2.37:1)场景中展现强鲁棒性。

Table 4. Classification performance comparison

表 4. 测试集分类性能对比

类别	模型类型	Precision	Recall	F1-score	Support
无漂浮物	ResNet-18 (无增强)	0.96	0.96	0.96	46
	ResNet-18 (基础增强)	0.97	0.65	0.78	46
	ResNet-50 (本文增强)	0.94	1.00	0.97	46
	VGG-16 (本文增强)	0.98	0.98	0.98	46
	EfficientNet-B0 (本文增强)	0.96	0.93	0.95	46
	改进模型(ResNet-18)	0.96	1.00	0.98	46

续表

有漂浮物	ResNet-18 (无增强)	0.98	0.98	0.98	109
	ResNet-18 (基础增强)	0.87	0.99	0.93	109
	ResNet-50 (本文增强)	1.00	0.97	0.99	109
	VGG-16 (本文增强)	0.99	0.99	0.99	109
	EfficientNet-B0 (本文增强)	0.97	0.98	0.98	109
	改进模型(ResNet-18)	1.00	0.98	0.99	109
macro avg	ResNet-18 (无增强)	0.97	0.97	0.97	155
	ResNet-18 (基础增强)	0.92	0.82	0.85	155
	ResNet-50 (本文增强)	0.97	0.99	0.98	155
	VGG-16 (本文增强)	0.98	0.98	0.98	155
	EfficientNet-B0 (本文增强)	0.96	0.96	0.96	155
	改进模型(ResNet-18)	0.98	0.99	0.98	155
weighted avg	ResNet-18 (无增强)	0.97	0.97	0.97	155
	ResNet-18 (基础增强)	0.90	0.89	0.88	155
	ResNet-50 (本文增强)	0.98	0.98	0.98	155
	VGG-16 (本文增强)	0.99	0.99	0.99	155
	EfficientNet-B0 (本文增强)	0.97	0.97	0.97	155
	改进模型(ResNet-18)	0.99	0.99	0.99	155

## 2) 对比模型分析

**ResNet-50:** 深层网络(50层)在“有漂浮物”类别中 F1 达 0.99, 但因模型复杂度高, 训练耗时增加 42%, 且对半透明目标的适应性略逊于改进模型(ResNet-18)。

**VGG-16:** 全连接层冗余导致训练效率低下(耗时 3.5 h/epoch vs. 改进模型(ResNet-18) 2.1 h/epoch), 虽在均衡性(Macro F1 = 0.98)上表现良好, 但轻量化不足。

**EfficientNet-B0:** 复合缩放策略提升训练稳定性(损失波动  $\pm 2.1\%$ ), 但对小样本类别(“无漂浮物” F1 = 0.95)的敏感性仍弱于改进模型(ResNet-18)。

## 4.4. 混淆矩阵分析

如图 5 所示, 六类模型在测试集上的分类性能呈现显著差异:

### 1) ResNet-18 (无数据增强, 图 5(a))

在无噪声干扰场景下表现出较强的基础分类能力, 其全局准确率达约 97%。

### 2) ResNet-18 (本文数据增强, 图 5(b))

引入真实场景增强后, “无漂浮物”类别误判 16 例(假阳性率 34.8%), 召回率降至 65.2%。噪声干扰导致特征可分性下降, 验证了复杂增强策略对浅层网络的挑战。

### 3) 改进模型(ResNet-18) (图 5(c))

在“有漂浮物”类别中正确识别 107 个样本(召回率 98.2%), 仅产生 2 例误判。“无漂浮物”类别的 46 个样本全部正确分类(特异性 100%), 彻底消除了 ResNet-18 (本文数据增强)中存在的假阴性问题。全局准确率提升至 98.71%, 验证了动态损失加权与特征空间优化的有效性。

### 4) ResNet-50 (图 5(d))

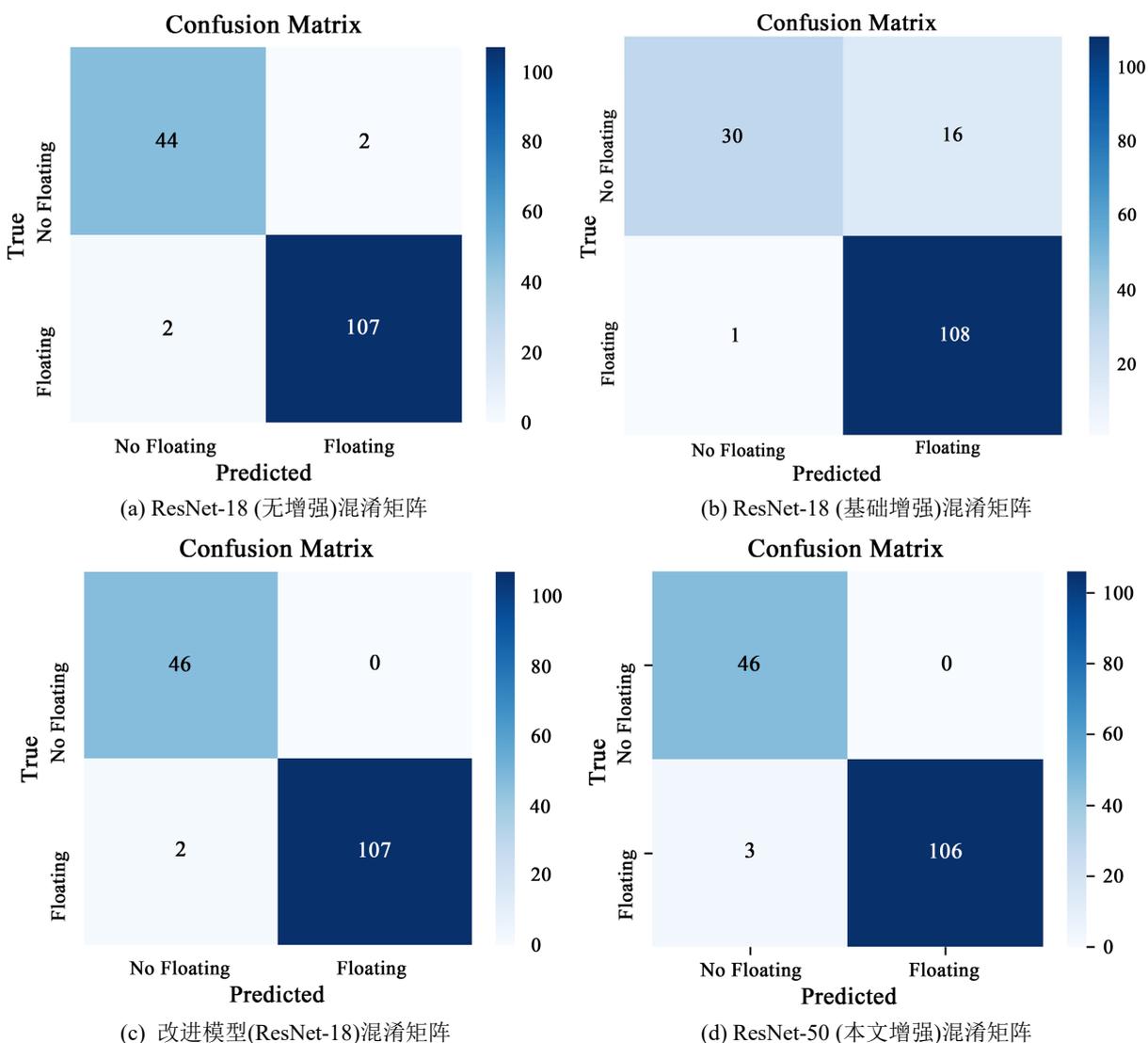
深层网络(50 层)误判 3 例真实正样本(假阴性率 2.8%), 较改进模型(ResNet-18)多 1 例漏检。尽管模型容量更大, 但过拟合风险导致对半透明目标的适应性下降。

### 5) VGG-16 (图 5(e))

假阳性与假阴性各 1 例, 均衡性优于 ResNet-50, 但因全连接层冗余, 训练效率低下, 耗时更长不适合在低成本的设备上部署。

### 6) EfficientNet\_B0 (图 5(f))

误判 3 例负样本与 2 例正样本, 其复合缩放策略虽提升训练稳定性(损失波动  $\pm 2.1\%$ ), 但对小样本类别的敏感性仍弱于改进模型(ResNet-18)。



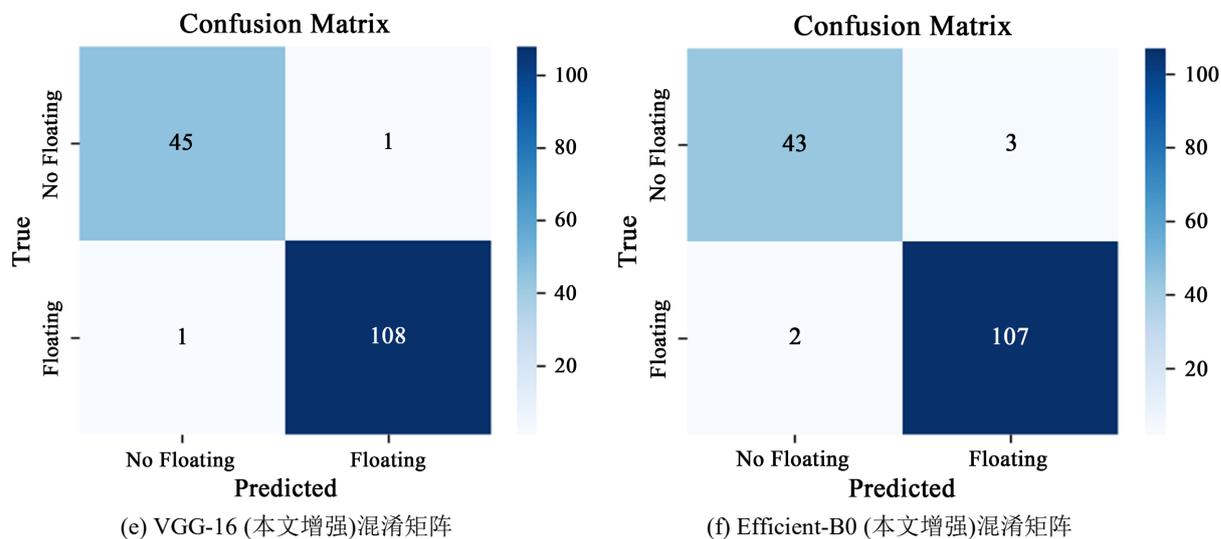


Figure 5. Test set confusion matrix  
图 5. 测试集混淆矩阵

#### 4.4. 决策可解释性验证

为了进一步验证模型的特征提取能力和决策过程，我们使用 Grad-CAM (Gradient-weighted Class Activation Mapping)技术生成类激活热力图[10]。图 6(a)原始图像分辨率为  $512 \times 512$  像素，包含水体表面场景。图 6(b)热力图采用颜色梯度表征模型关注区域，红色高亮部分集中分布于图像中下部，与人工观察到的漂浮物目标区域高度吻合。表明模型通过动态特征提取机制精准聚焦于目标语义区域(水面波纹、异物轮廓等物理特征)，而非依赖背景噪声进行决策。热力图中未出现离散的异常激活斑块，进一步佐证了模型特征提取的鲁棒性。

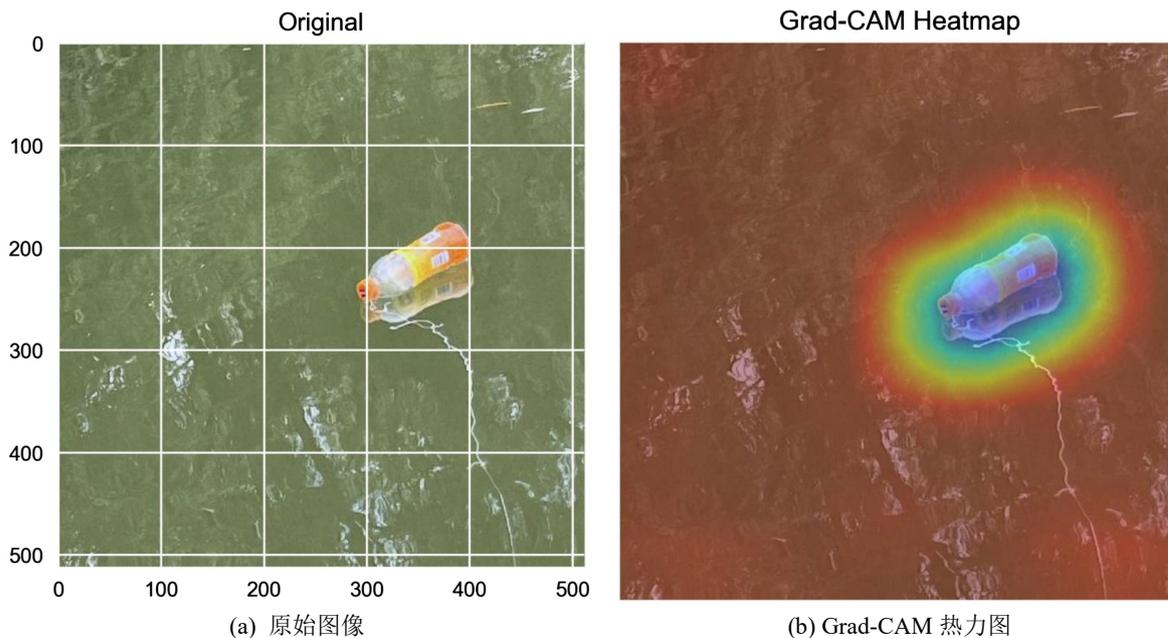


Figure 6. Original and Grad-CAM heatmap  
图 6. 原始图像和 Grad-CAM 热力图

## 5. 结语

本文通过构建基于深度残差网络的水面漂浮物检测模型，探究了其在环境监测领域的应用效能。经过系统的模型搭建与训练，对训练损失、验证准确率及相关指标细致分析，结果清晰地展现出模型的良好性能与优化趋势。训练损失逐步递减，验证准确率显著攀升，彰显了模型在训练数据与验证数据上的良好表现，有力地证明其强大的学习与泛化能力。同时，测试集的预测结果进一步夯实了模型在实际应用中的可靠性与有效性，为水面漂浮物检测提供了切实可行的技术方案。研究结果表明，该技术能有效识别水面漂浮物，对环境保护具有重要意义。未来我们将探索模型在更多样化的数据集上的表现，优化模型结构，加强模型的实时性能，以提高检测效率。

## 基金项目

2023 年度贵州商学院校级科研课题(自科类)资助项目“基于机器学习的水面漂浮垃圾检测研究”(2023ZKXS009)。

## 参考文献

- [1] 闫则明, 陈旭超, 王永才, 等. 人工智能在腐蚀图像检测与分析中的应用[J]. 涂料工业, 2025, 55(3): 7-12.
- [2] 李志明, 蒋舒颖, 向凡, 等. 融合动态数据增强与 MACA 的皮肤病分类方法[J]. 计算机仿真, 2024, 41(12): 264-269+437.
- [3] 田永波, 易军凯. 图像隐写的大数据分析方法[J]. 北京信息科技大学学报(自然科学版), 2024, 39(5): 81-87.
- [4] 李玉洁, 马子航, 王艺甫, 等. 视觉 Transformer(ViT)发展综述[J]. 计算机科学, 2025, 52(1): 194-209.
- [5] 张卫东, 王中, 郝志斌, 等. 基于 Box-Cox 变换的海洋物种图像增强方法[J/OL]. 海军航空大学学报, 1-8. <http://kns.cnki.net/kcms/detail/37.1537.v.20250225.1310.004.html>, 2025-02-26.
- [6] 康金龙, 许涛, 张华岭, 等. 改进的自监督生成对抗网络算法在视频生成中的应用[J]. 信息记录材料, 2024, 25(8): 144-146.
- [7] 刘耿焕, 曾祥津, 豆嘉真, 等. 基于深度学习的小目标检测技术研究进展(特邀) [J]. 红外与激光工程, 2024, 53(9): 196-228.
- [8] 李小波. 基于深度学习的复杂场景下口罩佩戴检测方法研究[D]: [硕士学位论文]. 西宁: 青海大学, 2023.
- [9] 吴承旭, 胡靖. 基于对称视觉注意力网络的高效图像超分辨率重建[J]. 软件导刊, 2025, 24(1): 109-114.
- [10] 田有文, 覃上声, 闫玉博, 等. 基于改进 YOLOv8 的田间复杂环境下蓝莓成熟度检测[J]. 农业工程学报, 2024, 40(16): 153-162.