

引用格式: 牛睿智, 刘志亮, 周雪, 等. 基于双模型决策级融合的鱼道分布外目标检测方法 [J]. 电子科技大学学报, 2025, 54(4): 554-565.  
NIU R Z, LIU Z L, ZHOU X, et al. Out-of-distribution fish detection in fishways based on decision-level fusion of two models[J]. Journal of University of Electronic Science and Technology of China, 2025, 54(4): 554-565.

# 基于双模型决策级融合的鱼道分布外目标检测方法



牛睿智<sup>1</sup>, 刘志亮<sup>1\*</sup>, 周雪<sup>2</sup>, 梅燕<sup>3</sup>

(1. 电子科技大学 机械与电气工程学院, 成都 611731; 2. 电子科技大学(深圳)高等研究院, 深圳 518110;  
3. 成都众柴科技有限公司, 成都 610052)

**摘要:** 针对鱼道过鱼目标检测鱼类特征模糊及数据集先验信息不足而导致的漏检、误检问题, 研究了一种基于粗-细粒度模型决策级融合的目标检测方法。该方法对粗粒度 YOLO 模型进行改进: 在主干层嵌入坐标注意力模块, 在特征融合部分嵌入自适应特征融合模块对不同尺度的特征层进行融合, 以提高粗粒度模型对任意鱼类的检测能力; 将改进后的粗粒度 YOLO 模型与细粒度 YOLO 模型的检测结果按照置信度筛选出需要融合的检测框, 并按照置信度的值对检测框进行加权融合, 以此降低在未知鱼类、模糊鱼类场景下进行分布外检测时的漏检率和误检率。应用此方法于真实环境下的鱼道过鱼数据集进行测试, 对未知鱼类的准确率达到 98.59%, 召回率达到 94.19%, 相比基于置信度的分布外检测方法分别提高了 9.25% 和 11.21%, 相比基于能量的分布外检测方法分别提高了 6.42% 和 3.69%。对模糊目标的识别准确率达到 95.45%, 召回率达到 91.8%, 相比基于置信度的分布外检测方法分别提高了 16.63% 和 18.58%, 相比基于能量的分布外检测方法分别提高了 11.27% 和 1.74%。研究成果对鱼道过鱼的目标检测有良好的借鉴意义。

**关键词:** 鱼道; 目标检测; YOLO; 注意力机制; 自适应特征融合; 决策级融合; 分布外检测  
**中图分类号:** TP39 **文献标志码:** A **DOI:** 10.12178/1001-0548.2024139

## Out-of-distribution fish detection in fishways based on decision-level fusion of two models

NIU Ruizhi<sup>1</sup>, LIU Zhiliang<sup>1\*</sup>, ZHOU Xue<sup>2</sup>, and MEI Yan<sup>3</sup>

(1. School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China;  
2. Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China, Shenzhen 518110, China;  
3. Chengdu Zoom Check Science & Technology Company Limited, Chengdu 610052, China)

**Abstract:** To address the issues of missed detection and false alarms resulting from ambiguous fish features and insufficient prior information in the dataset for fish detection in fishways, this paper proposes an object detection algorithm founded on decision fusion between coarse-grained and fine-grained YOLO models. The method enhances the coarse-grained YOLO model by embedding a coordinate attention module into the backbone layers and an adaptive spatial feature fusion module into the feature aggregation part to fuse features from different scale layers, thereby improving the model's detection capability for arbitrary fish species. The detection outcomes from the refined coarse-grained YOLO model and fine-grained YOLO model are filtered by confidence scores to identify boxes that need fusion. These boxes are then weighted fused based on their confidence values. This allows the proposed approach to attain a lower miss rate and false alarm rate for out-of-distribution detection on unknown and blurry fish. Evaluated on real-world fishway fish datasets, The proposed method achieves 98.59% accuracy and 94.19% recall for unknown fish, which are 9.25% and 11.21% higher than the confidence-based out-of-distribution detection method, and 6.42% and 3.69% higher than the energy-based out-of-distribution detection method,

收稿日期: 2024-06-12

基金项目: 国家自然科学基金(62372082); 中央高校基本科研业务费(ZYGX2024Z017); 深圳市自然科学基金(JCYJ20240813114206010); 四川省科技计划(2024JDHJ0057)

作者简介: 牛睿智, 主要从事复杂系统智能运维方面的研究。

\*通信作者 E-mail: Zhiliang\_Liu@uestc.edu.cn

respectively. The recognition accuracy of fuzzy targets reaches 95.45% and the recall rate reaches 91.8%, which is 16.63% and 18.58% higher than the confidence-based out-of-distribution detection method, and 11.27% and 1.74% higher than the energy-based out-of-distribution detection method, respectively. The research findings have valuable implications for fish detection in fishways.

**Key words:** fishways; object detection; YOLO; attention mechanism; adaptive feature fusion; decision fusion; out-of-distribution detection

修建大坝会产生大坝阻隔效应,切断洄游性鱼类的洄游通道,影响洄游性鱼类的种群繁殖<sup>[1]</sup>。鱼道作为减轻大坝阻隔对鱼类洄游影响的重要工程设施,保障鱼道工程的有效运行具有重要意义<sup>[2]</sup>。为评估鱼道过鱼效果,文献[3]提出相应的改进方法与措施以加强过鱼的有效性,需要在鱼道内开展过鱼监测。传统的监测技术包括张网法、堵截法等直接法和水声学法、遥测法等间接法<sup>[4]</sup>。直接法是直接捕获鱼类从而进行定量分析的方法,依赖人工操作,操作成本高,且对鱼类有伤害,无法长期监测;间接法中水声学法易受噪声干扰,而遥测法需要捕捉鱼类进行标记,单次标记仅能实现单次监测,无法满足过鱼监测的效率要求。

随着计算机视觉中深度学习技术的持续进步,学术界逐渐将其应用于鱼类目标检测相关研究中,并取得初步成果。相比传统的鱼类检测方法,其具有人工依赖度低、能够持续监测、对鱼类行为干扰小等优点<sup>[4]</sup>,能够有效提高鱼道过鱼目标检测的精度与效率。近年来, faster region-based CNN (faster R-CNN)<sup>[5]</sup>、single shot multibox detector (SSD)<sup>[6]</sup>、you only look once (YOLO)<sup>[7]</sup>等基于深度学习的方法开始被应用在鱼类监测领域中。

文献[8]将多尺度视网膜增强方法与 Faster R-CNN 相结合,提高模型对水下低清晰度图像的检测能力,在小规模水下鱼类数据集上达到 98.12% 的查准率。文献[9]将光流和高斯混合模型与 YOLO 相结合,以此提高 YOLO 在图像质量不佳时对鱼类的检测精度。文献[10]针对夜晚水下环境的鱼类检测,提出了基于 Cascade R-CNN 与多尺度 Retinex 色彩恢复算法相结合的方法,在夜晚水下鱼类图像检测中达到了 95.81% 的查准率。文献[11]使用多种图像增强算法对水下图像进行增强得到多个输入,使用 YOLOv4 模型进行检测后,将多个输出进行混合处理,提高了模型在模糊水下图像上的检测能力。文献[12]将 YOLOv4 模型与多尺度 Retinex 色彩恢复算法相结合,并采用生成对抗网络进行颜色校正和色彩恢复,在自行采集的 3 种鱼类数据集

上达到了 89.59% 的平均准确率。文献[13]通过添加注意力机制、更换模型损失函数的方式改进 YOLOv5s,在鱼类增殖站、水产研究所以及水电站实际采集的复杂水体数据集上达到了 91.9% 的平均精度,并利用 TensorRT 进行部署,实现了鱼道过鱼场景下的过鱼监测。

尽管深度学习技术的引入显著提升了鱼类检测的精度,但在通过分析鱼道监测图像可发现,当前过鱼图像检测仍存在以下难点:1)当水下环境存在光照不足或水体浑浊等问题时,鱼类在图像中的特征表现极为模糊,导致分类困难,并容易出现较高的漏检与误检;2)由于训练数据集中缺乏全面的先验信息,难以覆盖所有实际过鱼物种,一旦图像中出现未见过的鱼类,现有检测模型往往难以正确识别,从而进一步影响检测的准确性和稳定性。

为应对鱼道过鱼图像检测中存在的挑战,本文设计了一种融合粗粒度与细粒度模型、在决策层实现集成判断的分布外目标识别方法。

## 1 方法结构

在常见的目标检测模型中, YOLOv5 具有轻量、快速的特点,在各种目标检测任务中应用广泛。因此本文所采用的目标检测模型为 YOLOv5 系列。在计算机视觉的研究中,图像分类任务主要有粗粒度图像分类和细粒度图像分类两种<sup>[14]</sup>。其中,细粒度图像分类的对象通常属于同一个大类,进而对该大类下的子类别进行细分,如对同属于鱼类的多种鱼进行更细致的划分;而粗粒度图像分类的对象属性差异较大,如人、鸟、鱼等。这一概念在目标检测领域同样适用。针对鱼道环境中鱼类图像特征不清晰及鱼种未知等问题,本文在 YOLOv5 模型的基础上进行了两方面的改进:1)通过对输入图片进行灰度化预处理以及对粗粒度 YOLO 模型的改进,实现具有较强的泛化性的粗粒度鱼类检测方法;2)对粗粒度模型与细粒度模型的结果进行决策级融合。改进后的方法流程如图 1 所示。

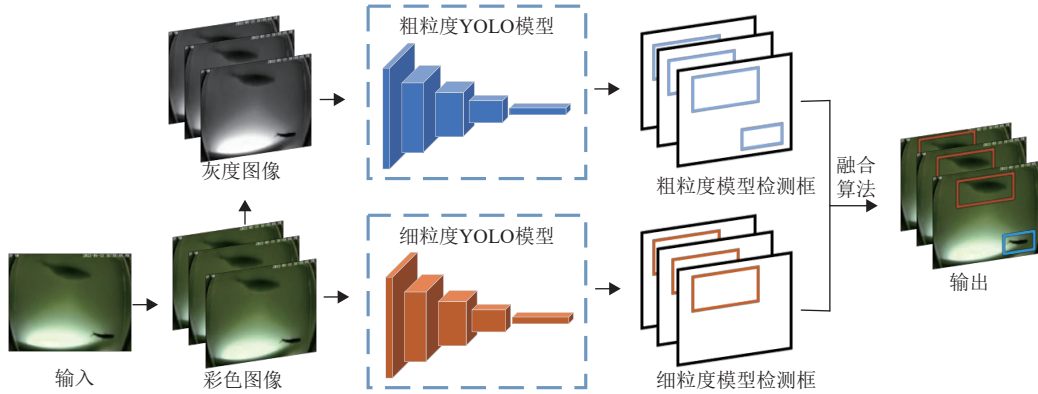


图 1 本文方法流程图

### 1.1 YOLOv5 模型改进

YOLOv5 系列包含 5 个不同参数规模的模型，即 YOLOv5x、YOLOv5l、YOLOv5m、YOLOv5s 和 YOLOv5n。其中 YOLOv5n 的参数量最小、计算速度最快。考虑到实际工程应用中，不同水电站的计算资源存在差异，为了适应多种实时检测需求，本文选用了 YOLOv5 系列中计算效率较高的 YOLOv5n 与 YOLOv5s，分别用于过鱼场景下的粗粒度与细粒度目标检测任务。相较于 YOLOv5n，YOLOv5s 的参数量和计算复杂度更大，对应卷积层的输出通道数更多，可构建更深层的网络来提

取目标的语义信息。YOLOv5s 虽然拥有对应的 YOLOv5s6 版本，但考虑到 YOLOv5s 本身已提供了足够的网络深度和宽度来捕捉细致的特征，因此对于细粒度目标检测任务，本文采用 YOLOv5s 作为检测模型。基于 YOLOv5n 架构，YOLOv5 官方后续推出了更深层的模型 YOLOv5n6。与 YOLOv5n 相比，YOLOv5n6 在特征提取网络层数和检测头数量上进行了扩展，从而实现了更强的检测性能。因此，本文选择 YOLOv5n6 作为后续粗粒度模型改进的基准模型。YOLOv5n6 与 YOLOv5s 的网络结构图如图 2 所示。

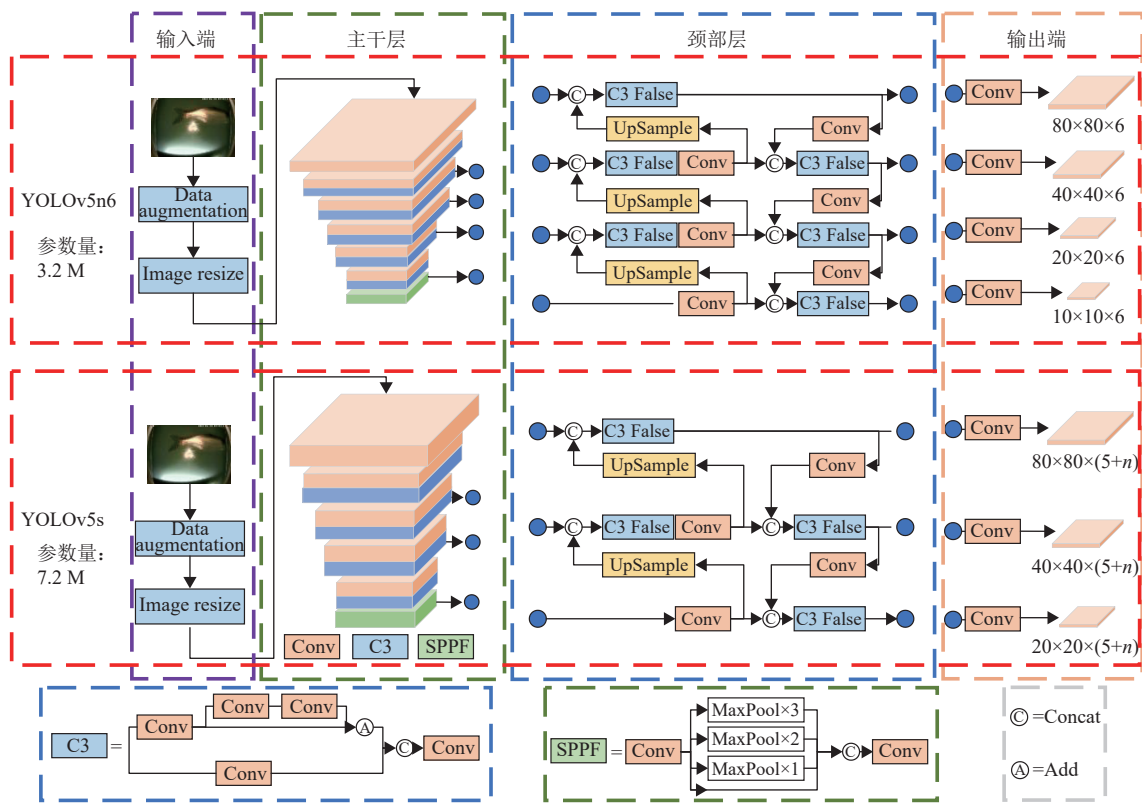


图 2 YOLOv5n6 与 YOLOv5s 的网络结构图

本文所采用的 YOLOv5 目标检测模型由输入模块、主干模块、颈部模块和头部模块组成。输入模块对原始图像进行一系列数据增强操作, 包括随机缩放、水平翻转、颜色空间增强、Mosaic 等, 输出经增强的图像。主干模块基于 CSPDarknet<sup>[15]</sup> 作为特征提取主干网络, 包含多个残差块, 内部通过残差连接实现特征重用, 并通过采样获取不同尺度特征, 构成特征金字塔。颈部模块采用基于 PAN<sup>[16]</sup> 的 CSP-PAN 结构, 高层特征上采样连接底层特征, 实现多尺度特征融合, 提升模型对不同大小目标的检测性能。头部模块在不同尺度的特征图上, 采用卷积层预测目标的框坐标、类别置信度与框置信度。由于本文将 YOLOv5n6 模型作为粗粒度模型, 输出类别仅有一类, 因此其输出端特征图的通道数为 6, 其中包含框坐标、框置信度以及一个类别置信度。而 YOLOv5s 作为细粒度模型, 根据预测类别数目  $n$  的不同, 会得到不同的特征图通道数。YOLOv5 的损失由边界框损失、目标损失和分类损失 3 部分组成。其中, 分类损失和目标损失采用二元交叉熵函数<sup>[17]</sup> 作为损失函数, 定位损失采用 CIOU<sup>[18]</sup> 作为损失函数。将 3 部分损失按照 0.05、1.00、0.50 的权重系数进行加权相加, 即为 YOLOv5 的总损失。

针对鱼道环境中鱼类检测所面临的问题, 本文选用 YOLOv5n6 模型为基础, 对其进行粗粒度方向的结构优化。在细粒度 YOLO 模型上, 由于 YOLOv5s 原模型可以很好地完成对分布内类别的检测, 因此细粒度 YOLO 模型直接使用 YOLOv5s 原模型。由于在鱼类特征极端模糊的情况下, 使用图像增强的方法难以恢复鱼类的种类特征, 同时由于水下光照复杂多变, 原始过鱼图像颜色变化较大。因此, 有必要对输入图像进行预处理, 而对彩色图像进行灰度化的处理可以有效减少图像颜色变化带来的问题。粗粒度 YOLO 模型的改进有 3 点: 1) 在数据预处理部分采用灰度算法将彩色图像转为灰度图, 减少因水下图像色彩变化对检测效果的影响; 2) 在 YOLOv5 模型主干中加入坐标注意力 (coordinate attention, CA) <sup>[19]</sup> 模块; 3) 在 YOLOv5 模型颈部加入自适应特征融合 (adaptively spatial feature fusion, ASFF) <sup>[20]</sup> 模块。改进后的方法流程如图 3 所示。

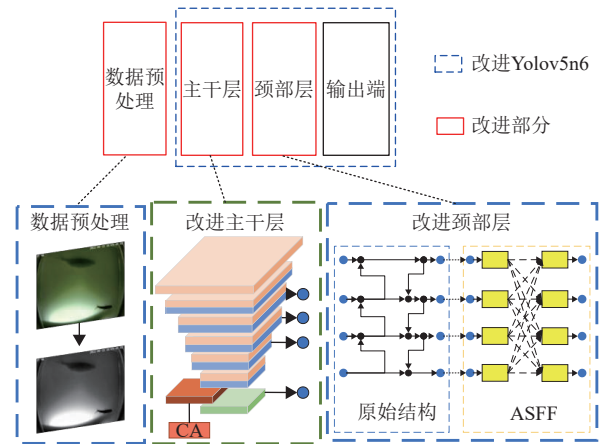


图3 YOLOv5n6 改进示意图

### 1.1.1 CA 模块

本文在 YOLOv5 的主干网络中引入 CA 注意力机制, 以增强模型的特征提取能力。该机制由文献 [19] 提出, 具备轻量、高效的特点, 能够兼顾通道间依赖和远距离位置信息的建模。其结构示意图如图 4 所示。

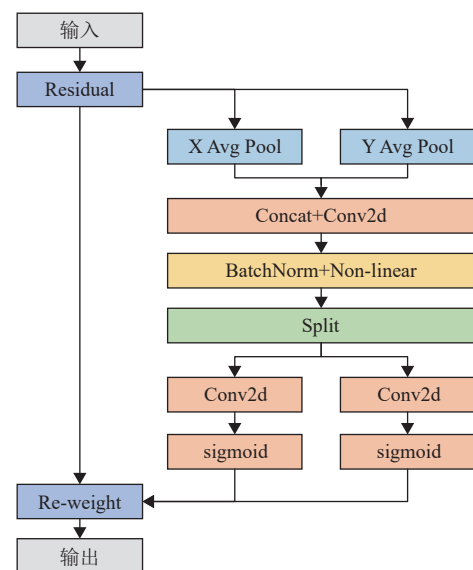


图4 CA 模块结构图

CA 模块首先通过 2 个并行的池化核  $(H,1)$  和  $(1,W)$  分别将垂直和水平方向的输入特征进行编码, 输出 1 对方向可知且位置敏感的特征图  $Z_c^h(h)$  和  $Z_c^w(w)$ :

$$Z_c^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} x_c(h, i) \quad (1)$$

$$Z_c^w(w) = \frac{1}{H} \sum_{0 \leq j \leq H} x_c(j, w) \quad (2)$$

式中,  $c$  表示第  $c$  维通道;  $w$  和  $h$  分别表示宽度和高

度； $Z_c^h(h)$ 即为高度为 $h$ 的第 $c$ 维通道的输出。利用 $Z_c^h(h)$ 和 $Z_c^w(w)$ 得到过程特征图 $f$ ：

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

式中， $F_1$ 代表 $1 \times 1$ 卷积； $\delta$ 代表线性整流（rectified linear unit, ReLU）激活函数；“ $[ ]$ ”代表张量拼接操作，用来聚合 $Z_c^h(h)$ 和 $Z_c^w(w)$ 的特征信息。

将 $f$ 沿宽度和高度 2 个维度拆分成 2 个特征向量 $f_h$ 和 $f_w$ ，得到 $g^h$ 和 $g^w$ ：

$$g^h = \sigma(F_1(f_h)) \quad (4)$$

$$g^w = \sigma(F_1(f_w)) \quad (5)$$

式中， $g^h$ 和 $g^w$ 为 2 个空间方向上的注意力权重图； $\sigma$ 代表 Sigmoid 激活函数。每个注意力权重图都带有特定空间方向的长程依赖。最终，将输入特征图与 2 个权重图相乘，增强特征图的特征表达：

$$y_{ij} = x_{ij} \times g_i^h \times g_j^w \quad (6)$$

式中， $x_{ij}$ 为输入特征图 $(i, j)$ 处的特征； $g_i^h$ 为权重图 $g^h$ 在 $(i, 1)$ 处的特征； $g_j^w$ 为权重图 $g^w$ 在 $(1, j)$ 处的特征； $y_{ij}$ 为输出特征图 $(i, j)$ 处的特征。

### 1.1.2 ASFF 模块

YOLOv5n6 在其颈部结构中采用 CSP-PAN 作为特征融合路径，通过双向融合，有效增强了特征图的语义表达能力。为进一步提升鱼道场景中特征融合的表现，本文在 CSP-PAN 模块之后引入 ASFF 机制。该模块为不同尺度的特征图分配动态权重，实现自适应的信息聚合，从而加强模型对关键鱼类语义的关注能力。

ASFF 最早由文献 [20] 提出，是一种具有可学习性且高效的特征融合策略。该模块的核心思想是在特征融合过程中动态分配权重，使模型能够聚焦于更具判别力的层级信息。

ASFF 结构如图 5 所示，其中  $X1$ 、 $X2$ 、 $X3$ 、 $X4$  表示 YOLOv5 中 CSP-PAN 结构的 4 个输出特征层。为实现多尺度特征融合，需要对这些特征层进行尺寸对齐，并统一其通道数。主要手段包括：通过卷积调整通道数量与降低空间分辨率，或采用插值上采样的方式提高分辨率，以确保融合前特征在空间尺寸与通道维度上的一致性。以 Level 1 为例，若需将  $X1$  与  $X2$  进行融合，需对  $X1$  施加一次步长为 2 的  $3 \times 3$  卷积，以降低分辨率并提升通道数，生成与  $X2$  尺寸一致的中间特征层，记为

$X1 \rightarrow 2$ 。类似地，若需将  $X1$  与  $X3$  融合，需先对  $X1$  进行一次步长为 2 的池化操作，然后再通过  $3 \times 3$  卷积进一步压缩分辨率与调整通道数，最终得到可与  $X3$  对齐的特征  $X1 \rightarrow 3$ 。

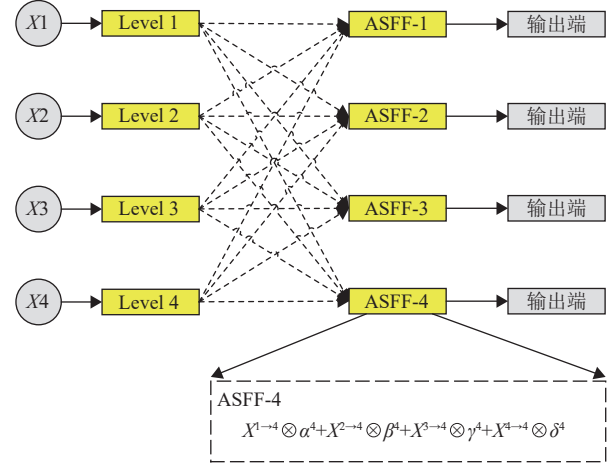


图 5 ASFF 模块结构图

特征缩放后，需要对特征进行融合。融合的公式如下：

$$Y_{ij}^l = \alpha_{ij}^l \times X_{ij}^{1 \rightarrow l} + \beta_{ij}^l \times X_{ij}^{2 \rightarrow l} + \gamma_{ij}^l \times X_{ij}^{3 \rightarrow l} + \delta_{ij}^l \times X_{ij}^{4 \rightarrow l} \quad (7)$$

式中， $Y_{ij}^l$ 表示特征融合后 $l$ 层特征层在 $(i, j)$ 处的特征； $\alpha_{ij}^l$ 、 $\beta_{ij}^l$ 、 $\gamma_{ij}^l$ 、 $\delta_{ij}^l$ 表示权重系数矩阵。其中， $\alpha_{ij}^l + \beta_{ij}^l + \gamma_{ij}^l + \delta_{ij}^l = 1$ 且 $\alpha_{ij}^l, \beta_{ij}^l, \gamma_{ij}^l, \delta_{ij}^l \in [0, 1]$ 。 $\alpha_{ij}^l$ 由 Softmax 归一化函数计算得到：

$$\alpha_{ij}^l = \frac{e^{\lambda_{\alpha_{ij}}^l}}{e^{\lambda_{\alpha_{ij}}^l} + e^{\lambda_{\beta_{ij}}^l} + e^{\lambda_{\gamma_{ij}}^l} + e^{\lambda_{\delta_{ij}}^l}} \quad (8)$$

式中， $\lambda_{\alpha_{ij}}^l$ 、 $\lambda_{\beta_{ij}}^l$ 、 $\lambda_{\gamma_{ij}}^l$ 、 $\lambda_{\delta_{ij}}^l$ 为 Softmax 函数的系数，均是通过特征缩放后的 3 个特征层进行  $1 \times 1$  卷积得到。与  $\alpha_{ij}^l$  同理，可得  $\beta_{ij}^l$ 、 $\gamma_{ij}^l$ 、 $\delta_{ij}^l$ 。权重系数矩阵通过反向传播进行训练，最终达到较优的特征融合效果。

## 1.2 决策级融合方法

粗粒度模型和细粒度模型检测后生成的结果中包含一系列置信度值与预测框坐标信息，为实现粗粒度信息与细粒度信息的融合，本文提出一种决策级融合方法，如图 6 所示。

在分布外目标检测任务中，基于置信度分数的分布外检测方法<sup>[21]</sup>较为常用。基于置信度分数的方法通过模型对测试样本预测得到的置信度来判断样本是否属于训练数据中的类别。如图 6 中对细粒

度模型结果的分割方法所示, 基于置信度分数的方法同样使用两个阈值来完成对检测结果的分割, 置信度值超过粗粒度阈值但小于细粒度阈值的预测框将被认为是分布外类别, 而置信度超过细粒度阈值的结果将被认定为分布内类别。

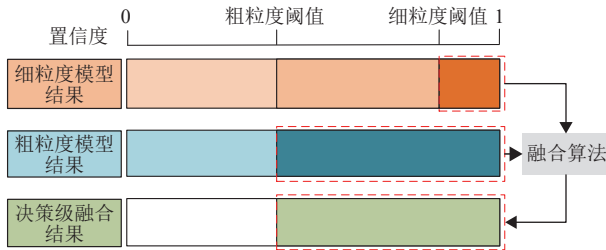


图6 决策级融合方法示意图

本文提出的决策级融合方法采用了上述基于置信度分数进行分布外检测的思想, 对于细粒度模型的检测结果, 通过设置粗粒度阈值与细粒度阈值完成对检测结果的筛选。置信度值超过粗粒度阈值但小于细粒度阈值的预测框将被认为识别到了某种鱼类, 类别未知; 而置信度超过细粒度阈值的预测框将被认为识别到了已知类别的鱼类。对于细粒度模型结果中超过细粒度阈值的部分, 将其与粗粒度模型结果中大于粗粒度阈值的预测框进行融合, 采用非极大值抑制的思想进行预测框融合后得到决策级融合结果。在后续实验中, 本文按照鱼道过鱼现场的实际经验, 将粗粒度阈值设定为0.4, 细粒度阈值设定为0.9。

对于参与融合的每一个细粒度预测框  $B_i^F$ , 首先使用计算交并比 (intersection over union, IoU) 的方式挑选出需要与其融合的粗粒度预测框。交并比的计算公式为:

$$\text{IoU}_F^C = \frac{S_B}{S_F + S_C - S_B} \quad (9)$$

式中,  $\text{IoU}_F^C$  为细粒度预测框  $B_F$  与粗粒度边界框  $B_C$  的交并比;  $S_F$  为细粒度预测框  $B_F$  的面积;  $S_C$  为粗粒度预测框  $B_C$  的面积;  $S_B$  为  $B_F$  与  $B_C$  交集部分的面积。计算完成后, 将所有交并比大于给定阈值的粗粒度预测框  $w_{\text{match}}^C$  与  $B_i^F$  进行加权融合, 计算公式为:

$$B_i^{\text{merge}} = \frac{w_i^F B_i^F + \sum_{0 \leq n \leq W} w_{\text{match},n}^C B_{\text{match},n}^C}{w_i^F + \sum_{0 \leq n \leq W} w_{\text{match},n}^C} \quad (10)$$

式中,  $B_i^{\text{merge}}$  为第  $i$  个细粒度预测框融合后的结果;  $w_i^F$  为第  $i$  个细粒度预测框的置信度;  $B_i^F$  为第  $i$  个细粒度预测框;  $w_{\text{match},n}^C$  为第  $n$  个与  $B_i^F$  交并比大于阈值的

粗粒度预测框置信度;  $B_{\text{match},n}^C$  为第  $n$  个与  $B_i^F$  交并比大于阈值的粗粒度预测框。

在进行检测框融合的过程中, 由于细粒度模型结果中参与融合的部分具有分布内类别的标签, 而粗粒度模型结果的标签均为分布外类别, 所以在融合的过程中不做分标签融合, 而是将筛选出的所有检测框放在一起进行融合。融合方法实际效果如图7所示。

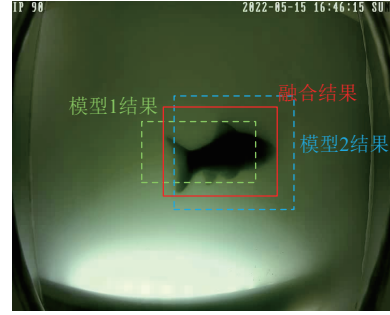


图7 预测框融合方法示意图

## 2 实验及分析

### 2.1 数据集构建

#### 2.1.1 数据采集

本文构建了常见鱼类的数据集 F9FD (fishway 9-species fish dataset) 和复杂环境条件下的数据集 FBFD (fishway blurred fish dataset)。原始视频的分辨率为  $704 \times 576$ , 帧率为 25 帧每秒。采集装置如图8所示。

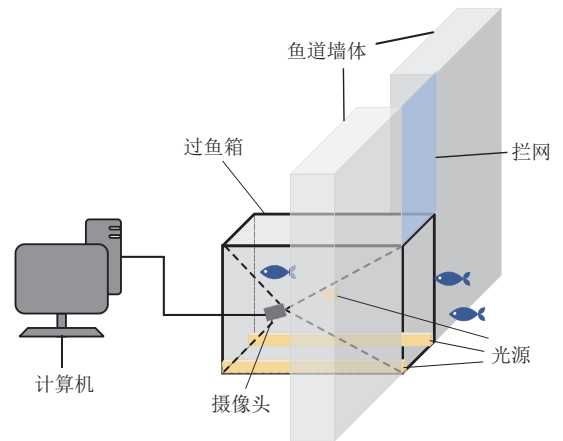


图8 数据采集装置原理图

#### 2.1.2 数据集 F9FD 介绍

F9FD 数据集中部分采集自福建某水电站的实际过鱼通道, 其余则采集自实验室条件下构建的过鱼箱。通过筛选与帧提取, 从中选取了 3 965 张分辨率为  $704 \times 576$  的图像, 并借助 LabelImg 对图像中鱼类目标进行标注。F9FD 中包含了 9 种鱼类数

据,分别为花鲈、白甲鱼、鳊鲴、草鱼、鲫鱼、罗非鱼、翘嘴、鲢鳙、鲤鱼。其中,草鱼、鲫鱼、罗非鱼在同一时刻采集,该部分图片中同时包含上述 3 种鱼类。F9FD 中的鱼类图片覆盖了多种光照条件,且鱼类特征较为清晰。该数据集为多分类数据集,用于目标检测模型的训练工作以及本文方法在过鱼种类未知情况下检测能力的测试工作。该数据集图片示例如图 9 所示。

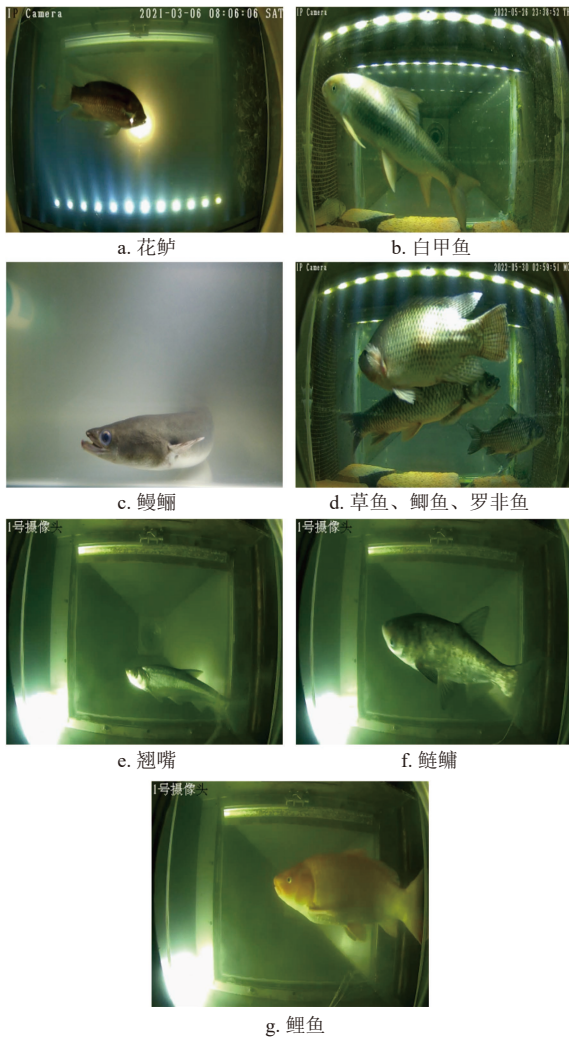


图 9 F9FD 数据集图片示例

### 2.1.3 数据集 FBFD 介绍

FBFD 中的数据采集自四川某水电站现场过鱼箱。通过筛选与帧提取,最终选取了 176 张分辨率为  $704 \times 576$  的鱼类图像。受环境中光照条件与水体清洁度变化的影响,图像在清晰度与亮度方面存在一定差异,部分鱼类目标边缘模糊、特征不明显,给后续的目标识别带来一定挑战。该数据集为单分类数据集,用于验证本文方法在鱼类特征模糊情况下的检测能力。该数据集图片示例如图 10 所示。

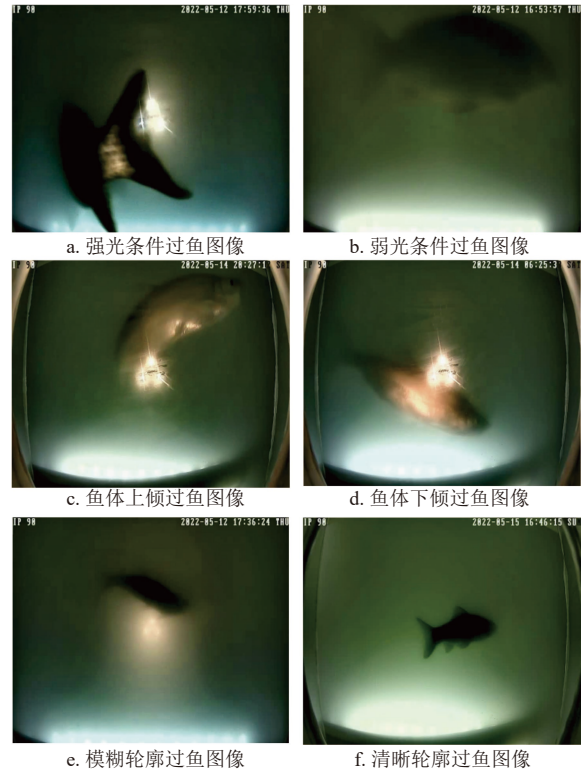


图 10 FBFD 数据集图片示例

## 2.2 实验配置

本实验在 Ubuntu 20.04.4 LTS 操作系统上进行,基于 Pytorch 1.7.0 深度学习框架搭建。硬件环境包括 NVIDIA GeForce RTX 2080 GPU、Intel i9-9900K 处理器及 32 GB 内存。训练过程中,超参数设置为:迭代 150 次、批量大小 16、初始学习率设为 0.001、输入图像分辨率为  $640 \times 640$ 。训练时采用在 COCO 数据集上预训练的权重进行迁移学习。

## 2.3 评价标准

为评估本文方法在鱼类模糊、过鱼种类未知条件下解决漏检和误检问题的有效性,本文选用精确率和召回率作为衡量指标。精确率衡量了模型标记的正类别中有多少是真正的正类别,直接反映了模型的检测准确性。召回率则衡量了模型能够检测到的正确正类别的比例,是漏检率的直接度量。相较于所有类别的平均精度均值 (mean average precision, mAP),本研究将重心放在精确率和召回率上更具实际意义,因为本文主要目标在于减少鱼类模糊、过鱼种类未知条件下的漏检和误检,而精确率和召回率直接反应了这两方面的性能。精确率与召回率的计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (11)$$

$$R = \frac{TP}{TP + FN} \quad (12)$$

式中,  $P$  (positive) 为精确率;  $R$  (recall) 为召回率; TP (true positive) 为正确识别的目标数量; FP (false positive) 为误检的目标数量; FN (false negative) 为漏检的目标数量。若预测结果与真实值的 IoU 大于给定阈值, 则该预测结果属于 TP, 否则属于 FP。

## 2.4 粗粒度模型改进实验结果分析

为了验证本文所提出的粗粒度 YOLO 模型各项改进的效果, 设计了一系列消融实验。具体如表 1 所示, 在 YOLOv5n6 模型的基础上, 分别引入图像灰度化处理、CA 注意力模块、ASFF 模块, 以及同时加入 CA 和 ASFF 模块的多种改进方案。所有模型均在 F9FD 数据集上进行训练, 并在 FBFD 数据集上比较其精确率和召回率表现。

由表 1 可知, 将彩色图像转换为灰度图后, 模型性能显著提升, 各项指标均有明显改善。单独引入 CA 模块后, YOLOv5n6 模型在所有指标上均获得较大提升, 表明该注意力机制有效增强了模型的检测能力。单独添加 ASFF 模块也带来了不同程度的性能提升。值得注意的是, 同时结合 CA 模块和 ASFF 模块的模型, 在精确率和召回率上均优于仅添加单一模块的版本, 进一步验证了多项改进的协同作用。采用灰度图进行训练和测试的改进模型, 相较于使用彩色图像的原始 YOLOv5n6 模型, 精确率提升了 12.9%, 召回率提升了 14.6%, 充分证明了所提改进方案对模型性能的显著优化。

表 1 粗粒度模型改进消融实验 %

方法	精确率	召回率
YOLOv5n6+彩色图	80.6	74.7
YOLOv5n6+灰度图	88.4	87.1
YOLOv5n6+CA+灰度图	90.2	88.2
YOLOv5n6+ASFF+灰度图	89.1	87.6
YOLOv5n6+CA+ASFF+灰度图	93.5	89.3

## 2.5 分布外检测实验结果分析

### 2.5.1 过鱼种类未知实验

为验证本文方法在过鱼种类不在训练集的情况下, 对未知鱼类进行分布外检测的能力, 本文依次从数据集 F9FD 中提取一类鱼的数据作为测试集, 剩余 8 类按照 8:2 的比例分割为训练集和验证集, 用于训练粗粒度模型和细粒度模型。其中, 粗粒度模型在转换为灰度图像的数据上进行训练。由于 F9FD 数据集中草鱼、鲫鱼、罗非鱼在同一时间采集, 在图片数据中同时出现, 因此草鱼、鲫鱼、罗非鱼将同时作为未知鱼类, 其余 6 种作为训练与

验证集。本文将按照图 6 中对细粒度模型结果的划分方法, 将高于粗粒度阈值但小于细粒度阈值的结果的类别信息视为未知类别, 最终在测试集上对比基于能量的分布外检测方法<sup>[22]</sup>、基于置信度的方法和本文方法。

图 11 和图 12 分别展示了 7 组粗粒度模型和 7 组细粒度模型在训练过程中的训练集损失曲线。从图中观察可知, 损失值在前 20 次迭代内迅速下降, 随后随着迭代次数的增加, 下降速度逐渐减缓, 最终趋于稳定。最终, 粗粒度和细粒度模型在验证集上的 mAP 均在 95% 以上。

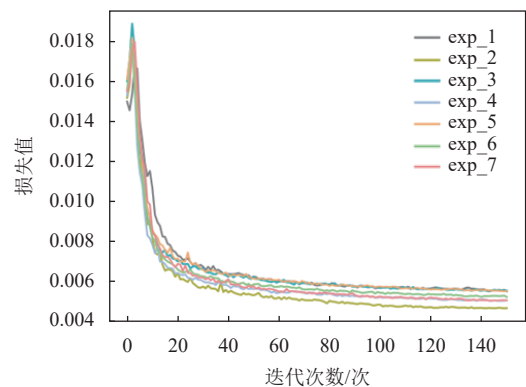


图 11 粗粒度模型训练集损失曲线

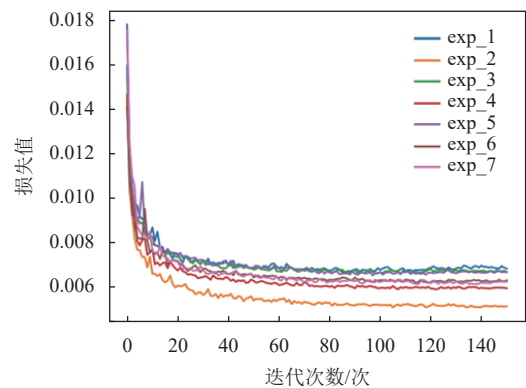


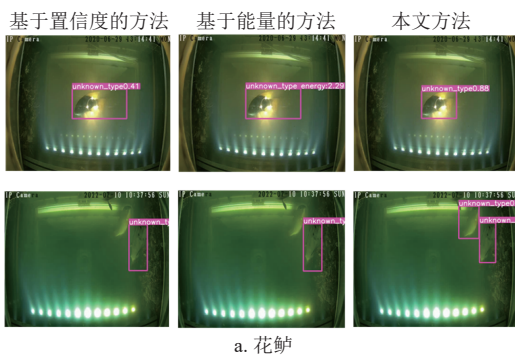
图 12 细粒度模型训练集损失曲线

为验证本文方法在过鱼种类未知条件下的性能, 本文分别选取 7 组实验的过鱼图片进行测试, 同时将本文方法与基于置信度的分布外检测方法的检测结果进行对比。

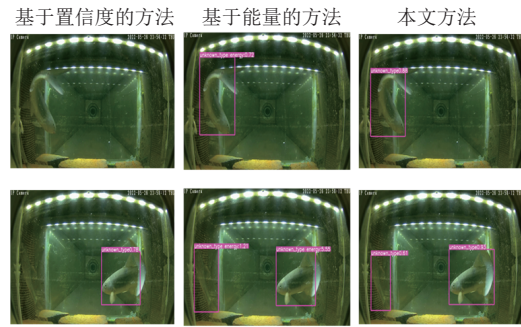
如图 13a 所示, 第 1 张图片中基于置信度的分布外检测方法和基于能量的方法对未经训练的花鲈类别进行检测时预测框位置不准确, 本文方法对鱼类的框选更为精准。第 2 张图片中基于置信度的分布外检测方法和基于能量的分布外检测方法出现了漏检情况, 而本文方法准确检测到了 2 条未知鱼类。如图 13b 所示, 对白甲鱼类别进行

检测时，2 张图片中基于置信度的分布外检测方法均出现了漏检情况，基于能量的分布外检测方法和本文方法在这 2 张过鱼图片上均未出现漏检。同时，本文方法对鱼类的框选更为精准。如图 13c 所示，对草鱼、鲫鱼、罗非鱼进行检测时基于置信度的分布外检测方法出现了严重的漏检情况，基于能量的分布外检测方法漏检数量较少，但在第 1 张图片上依然有漏检的情况出现，而本文方法则能准确检出鱼类。如图 13d 所示，对鳊鲴进行检测时，由于鳊鲴的特征与其他鱼类的特征差异较大，基于置信度的分布外检测方法几乎无法准确检出鱼类，且由于鳊鲴的数据采集自实验室，图片左上角出现了增氧泵，基于置信度的分布外检测方法和基于能量的分布外检测方法均对其出现了误检。而在这种条件下本文方法依然准确检测到了鱼类。如图 13e 所示，对鲤鱼进行检测时，基于置信度的分布外检测方法在第 1 张图片上出现了误检，第 2 张图片中出现了漏检情况。基于能量的分布外检测方法在第 1 张图片上出现了误检，且将未知鱼类错误分类为罗非鱼。而本文方法检测结果相比于其他方法更为准确。如图 13f 所示，3 种方法在对翘嘴进行检测时均出现了漏检，本文方法漏检数量少于其他 2 种方法，表现较优。基于置信度的分布外检测方法结果最差，第 2 张图片中基于置信度的方法对图像中的 2 条鱼均未能检出，基于能量的分布外检测方法检测出了 1 条鱼类，本文方法则准确检测出 2 条鱼类。

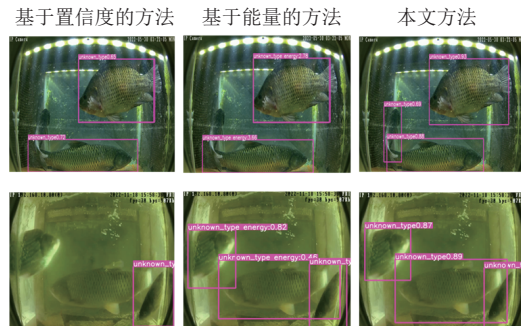
如图 13g 所示，在对鲢鳙类别进行检测时，本文方法对鱼类的框选更为精准，且漏检情况更少。总体而言，基于置信度的分布外检测方法和基于能量的分布外检测方法在过鱼种类未知的情况下出现了较多的漏检和误检，本文方法则在该情况下漏检、误检较少。



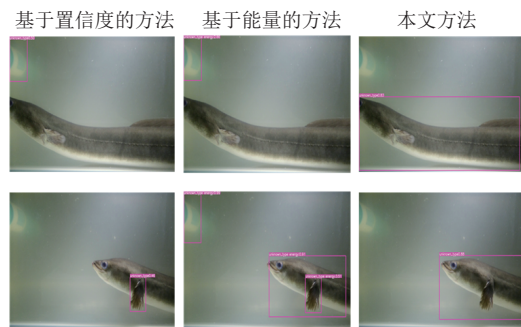
a. 花鲈



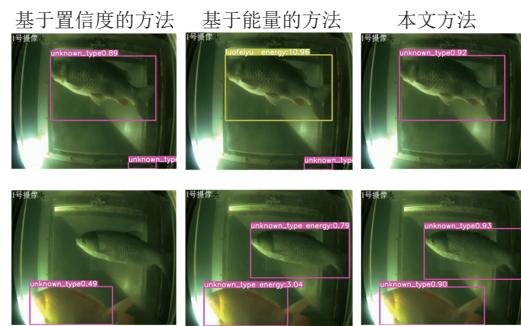
b. 白甲鱼



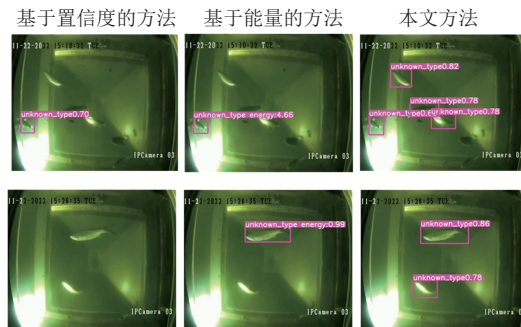
c. 草鱼、鲫鱼、罗非鱼



d. 鳊鲴



e. 鲤鱼



f. 翘嘴

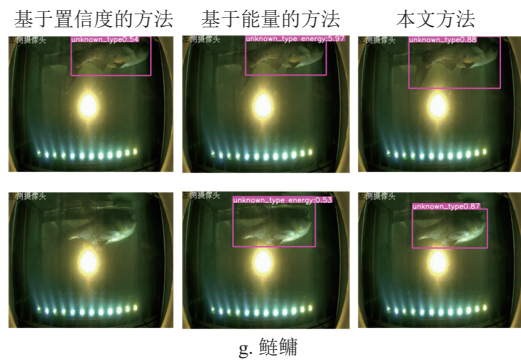


图 13 过鱼种类未知实验过鱼检测结果对比

为进一步验证本文方法在过鱼种类未知条件下的性能, 本文选取 7 组实验中基于置信度的分布外检测方法和基于能量的分布外检测方法与本文方法的召回率与精确率进行对比, 如图 14 所示。在 7 组实验中, 本文方法的召回率和精确率均高于其他 2 种对比方法。除翘嘴类别外, 本文方法的召回率均高于 90%, 而在翘嘴类别的实验中本文方法的召回率比基于置信度的分布外检测方法高出 13.46%, 相比基于能量的分布外检测方法高出 5.77%。

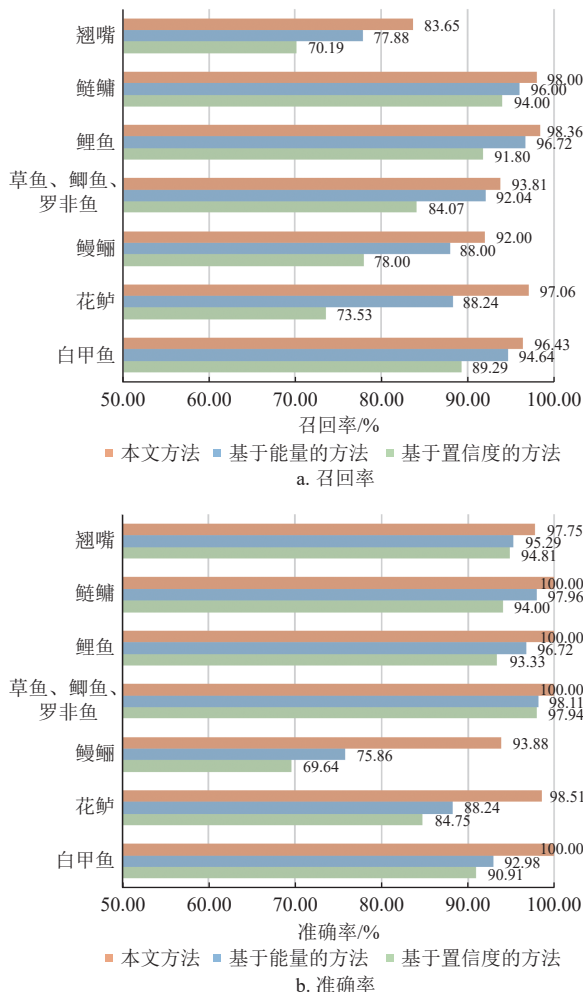


图 14 过鱼种类未知实验召回率与准确率对比图

同时, 本文对基于置信度的分布外检测方法与本文方法 7 组实验结果的 TP、FP、FN 取均值, 并进行对比。如表 2 所示, 本文方法在漏检数量和误检数量上远低于基于置信度的分布外检测方法和基于能量的分布外检测方法, 大大降低了漏检率和误检率, 表明了本文方法在过鱼种类未知场景下的有效性。

表 2 过鱼种类未知实验检测结果数量对比

方法	TP	FP	FN
基于置信度的分布外检测方法	585	62	131
基于能量的分布外检测方法	641	50	75
本文方法	668	8	48

### 2.5.2 过鱼特征模糊实验

为验证本文方法在鱼类特征模糊的情况下, 对模糊鱼类进行识别和分类的能力, 本文使用数据集 FBFD 作为测试集, 将数据集 F9FD 按照 8:2 的比例分割作为训练集和验证集, 用于训练粗粒度模型和细粒度模型。粗粒度模型在转换为灰度图像的数据上进行训练。最终, 在测试集上将基于置信度的分布外检测方法和基于能量的分布外检测方法与本文方法进行对比。

图 15 展示了粗粒度模型和细粒度模型在训练过程中的训练集损失变化情况。可以看到, 损失在前 20 次迭代时快速降低, 随后随着迭代的继续, 下降速度逐渐放缓并趋于平稳。最终, 粗粒度模型在验证集上的 mAP 为 98.9%, 细粒度模型在验证集上的 mAP 为 98.1%。

为验证本文方法在鱼类特征模糊条件下的性能, 将本文与基于置信度的分布外检测方法和基于能量的分布外检测方法进行对比, 结果如图 16 所示。基于置信度的分布外检测方法和基于能量的分布外检测方法在应对鱼类特征模糊情况下的检测中出现了较为严重的漏检和误检。第 1 行图片中, 基于置信度的分布外检测方法和基于能量的分布外检测方法均将鱼道下方补光灯弯弧识别为了鱼类, 第 2~第 4 行图片中基于置信度的分布外检测方法均出现了漏检, 基于能量的分布外检测方法结果稍好, 但在第 3 行图片中也出现了漏检。而本文方法在鱼类特征模糊的情况下能够更为准确地识别到鱼类。总体而言, 在过鱼特征模糊的情况下本文方法相比于基于置信度的分布外检测方法漏检、误检更少, 表现更优。

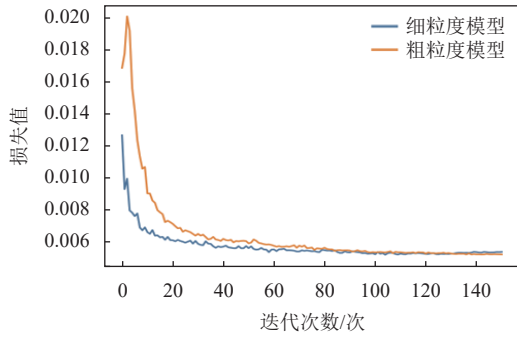
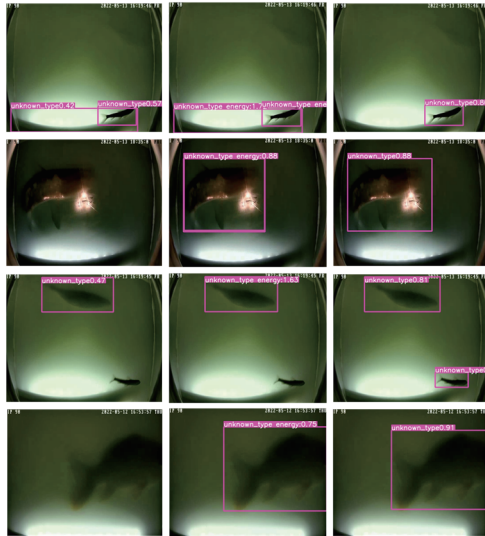


图 15 模型训练过程损失曲线



a. 基于置信度的方法 b. 基于能量的方法 c. 本文方法

图 16 过鱼特征模糊实验过鱼检测结果对比

为进一步验证本文方法在鱼类特征模糊条件下的性能，将基于置信度的分布外检测方法和基于能量的分布外检测方法与本文方法检测结果的准确率和召回率进行对比，如图 17 所示。本文方法的召回率与准确率均高于基于置信度的分布外检测方法，本文方法的准确率达到 95.45%，相比基于置信度的分布外检测方法提升了 16.63%，相比基于能量的分布外检测方法提升了 11.27%；本文方法的召回率达到 91.8%，相比基于置信度的分布外检测方法提升了 18.58%，相比基于能量的分布外检测方法提升了 1.74%。

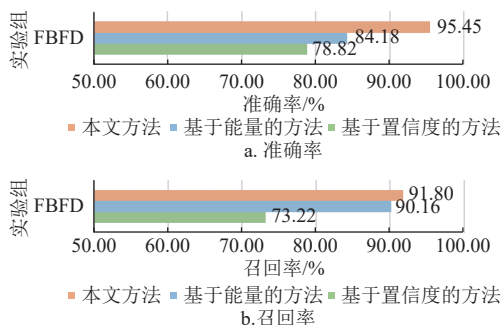


图 17 过鱼特征模糊实验准确率、召回率对比图

同时，将基于置信度的分布外检测方法和基于能量的分布外检测方法与本文方法检测结果的 TP、FP、FN 数量进行对比。

如表 3 所示，本文方法在漏检数量和误检数量上远低于其余 2 种对比方法，证明了本文方法在过鱼种类未知场景下的有效性和先进性。

表 3 过鱼特征模糊实验检测结果数量对比

方法	TP	FP	FN
基于置信度的分布外检测方法	134	36	49
基于能量的分布外检测方法	165	31	18
本文方法	168	8	15

### 3 结束语

本文针对鱼道过鱼场景下因鱼类特征模糊及数据集先验信息不足导致的误检漏检问题，使用非极大值抑制的思想对粗粒度模型和细粒度模型检测结果进行决策级融合，通过结合粗粒度模型对鱼类检测的泛化能力以及细粒度模型对已知鱼类的检测能力，在保证对已知鱼类的正常分类的前提下减少对未知鱼类和模糊鱼类的漏检和误检。在目标检测模型的改进上，本文在粗粒度模型的数据预处理部分采用灰度算法对输入图像进行处理，将彩色图像转为灰度图像，减少因图像色彩变化对检测结果的干扰。在主干层添加 CA 模块、颈部层添加 ASFF 模块，提高模型的检测性能。通过上述改进提高粗粒度模型对鱼类检测的泛化能力，实现对未知鱼类、模糊鱼类的准确检测能力。实验结果显示，本文所提方法在真实环境中的过鱼测试集中，对未知鱼类的检测精度达到了 98.59%，召回率为 94.19%，相比基于置信度的分布外检测方法分别提高了 9.25% 和 11.21%，相比基于能量的分布外检测方法分别提高了 6.42% 和 3.69%。本文方法对模糊目标的识别精确率达到了 95.45%，召回率达到了 91.8%，相比基于置信度的分布外检测方法分别提高了 16.63% 和 18.58%，相比基于能量的分布外检测方法分别提高了 11.27% 和 1.74%，证明了本文方法的有效性和先进性，在鱼道过鱼目标检测方面有良好的应用前景。

### 参考文献

[1] 朱瑶. 大坝对鱼类栖息地的影响及评价方法述评[J]. 中国水利水电科学研究院学报, 2005, 3(2): 100-103.  
ZHU Y. Commentary on dam influence on fish habitat and evaluation on assessment method[J]. Journal of China

- Institute of Water, 2005, 3(2): 100-103.
- [2] 王兴勇, 郭军. 国内外鱼道研究与建设[J]. 中国水利水电科学研究院学报, 2005, 3(3): 222-228.  
WANG X Y, GUO J. Brief review on research and construction of fish-ways at home and abroad[J]. Journal of China Institute of Water, 2005, 3(3): 222-228.
- [3] 温静雅, 陈昂, 曹娜, 等. 国内外过鱼设施运行效果评估与监测技术研究综述[J]. 水利水电科技进展, 2019, 39(5): 49-55.  
WEN J Y, CHEN A, CAO N, et al. A review of effectiveness assessment and monitoring technologies for fish passage facilities[J]. Advances in Science and Technology of Water Resources, 2019, 39(5): 49-55.
- [4] 王义川, 王煜, 林晨宇, 等. 鱼道过鱼效果监测方法述评[J]. 生态学杂志, 2019, 38(2): 586-593.  
WANG Y C, WANG Y, LIN C Y, et al. A review on monitoring methods for the effectiveness of fishway[J]. Chinese Journal of Ecology, 2019, 38(2): 586-593.
- [5] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[EB/OL]. [2024-06-01]. [https://link.springer.com/chapter/10.1007/978-3-319-46448-0\\_2](https://link.springer.com/chapter/10.1007/978-3-319-46448-0_2).
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016: 779-788.
- [8] 袁红春, 张硕. 基于 Faster R-CNN 和图像增强的水下鱼类目标检测方法[J]. 大连海洋大学学报, 2020, 35(4): 612-619.  
YUAN H C, ZHANG S. Detection of underwater fish based on faster R-CNN and image enhancement[J]. Journal of Dalian Ocean University, 2020, 35(4): 612-619.
- [9] JALAL A, SALMAN A, MIAN A, et al. Fish detection and species classification in underwater environments using deep learning with temporal information[J]. Ecological Informatics, 2020, 57: 101088.
- [10] 张明华, 龙腾, 宋巍, 等. 基于改进 Cascade R-CNN 和图像增强的夜晚鱼类检测[J]. 农业机械学报, 2021, 52(9): 179-185.  
ZHANG M H, LONG T, SONG W, et al. Object detection of underwater fish at night based on improved cascade R-CNN and image enhancement[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 179-185.
- [11] 覃学标, 黄冬梅, 宋巍, 等. 模糊水下图像多增强与输出混合的鱼类检测方法[J]. 农业机械学报, 2022, 53(7): 243-249.  
QIN X B, HUANG D M, SONG W, et al. Fish detection method of multiple enhanced and outputs blend for blurred underwater images[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 243-249.
- [12] 杨文静, 陈明, 冯国富. 基于图像增强的水下视频鱼类识别方法[J]. 激光与光电子学进展, 2021, 58(22): 2211002.  
YANG W J, CHEN M, FENG G F. Fish recognition method for underwater video based on image enhancement[J]. Laser & Optoelectronics Progress, 2021, 58(22): 2211002.
- [13] 李健源, 柳春娜, 卢晓春, 等. 基于改进 YOLOv5s 和 TensorRT 部署的鱼道过鱼监测[J]. 农业机械学报, 2022, 53(12): 314-322.  
LI J Y, LIU C N, LU X C, et al. Fish passage monitoring based on improved YOLOv5s and TensorRT deployment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(12): 314-322.
- [14] 冯语珊, 王子磊. 自上而下注意力分割的细粒度图像分类[J]. 中国图象图形学报, 2016, 21(9): 1147-1154.  
FENG Y S, WANG Z L. Fine-grained image categorization with segmentation based on top-down attention map[J]. Journal of Image and Graphics, 2016, 21(9): 1147-1154.
- [15] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[EB/OL]. [2024-06-01]. <https://doi.org/10.48550/arXiv.2004.10934>.
- [16] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 8759-8768.
- [17] DEBOER P T, KROESE D P, MANNOR S, et al. A tutorial on the cross-entropy method[J]. Annals of Operations Research, 2005, 134(1): 19-67.
- [18] REZATOFI G H, TSOI N, GWAK J, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2019: 658-666.
- [19] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2021: 13708-13717.
- [20] LIU S, HUANG D, WANG Y. Learning spatial fusion for single-shot object detection[EB/OL]. [2024-06-01]. <https://doi.org/10.48550/arXiv.1911.09516>.
- [21] BLEI Y, JOURDAN N, GÄHLERT N. Identifying out-of-distribution samples in real-time for safety-critical 2D object detection with margin entropy loss[EB/OL]. [2024-06-01]. <https://doi.org/10.48550/arXiv.2209.00364>.
- [22] WANG H, LIU W, BOCCHIERI A, et al. Can multi-label classification networks know what they don't know?[J]. Advances in Neural Information Processing Systems, 2021, 34: 29074-29087.