

基于ResNet50_SIMAM的水下目标检测模型研究

柏填晟, 张亚婷, 金珊, 李晓璇, 刘朝霞*

大连外国语学院软件学院, 辽宁 大连

收稿日期: 2024年2月16日; 录用日期: 2024年3月18日; 发布日期: 2024年3月25日

摘要

水下目标检测是海洋探索和监测领域的一个关键技术挑战, 具有广泛的应用。由于水下环境复杂以及视觉清晰度有限, 现有检测方法效果不佳。针对这一问题, 我们对现有的ResNet50模型进行改进, 通过引入SIMAM注意力机制来提高检测精确度。通过对数据集的预处理和增强, 模型成功适应了水下图像的特点。实验结果表明, 该模型在水下目标检测任务上表现卓越, Map值由原来的64.6上升到68.35, 验证了改进后的模型ResNet50_SIMAM在处理复杂水下视觉任务中的巨大潜力。

关键词

水下目标检测, 深度学习, ResNet50, TensorFlow, SimAM

Research on Underwater Target Detection Model Based on ResNet50_SIMAM

Tiancheng Bai, Yating Zhang, Shan Jin, Xiaoxuan Li, Zhaoxia Liu*

School of Software, Dalian University of Foreign Chinese, Dalian Liaoning

Received: Feb. 16th, 2024; accepted: Mar. 18th, 2024; published: Mar. 25th, 2024

Abstract

Underwater target detection is a key technological challenge in the field of ocean exploration and monitoring and has a wide range of applications. Due to the complexity of the underwater environment as well as the limited visual clarity, the existing detection methods are ineffective. To address this problem, we improve the existing ResNet50 model to enhance the detection accuracy

*通讯作者。

文章引用: 柏填晟, 张亚婷, 金珊, 李晓璇, 刘朝霞. 基于 ResNet50_SIMAM 的水下目标检测模型研究[J]. 计算机科学与应用, 2024, 14(3): 58-68. DOI: 10.12677/csa.2024.143058

by introducing the SIMAM attention mechanism. By preprocessing and enhancing the dataset, the model is successfully adapted to the characteristics of underwater images. The experimental results show that the model performs excellently on the underwater target detection task, and the Map value increases from the original 64.6 to 68.35, which verifies the great potential of the improved model ResNet50_SIMAM in dealing with the complex underwater vision tasks.

Keywords

Underwater Object Detection, Deep Learning, ResNet50, TensorFlow, SimAM

Copyright © 2024 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

水下目标检测在海洋探索、军事防御、环境监测以及水下资源开发等领域具有至关重要的作用，然而，由于水下环境的复杂性和特殊性，如光线衰减、水体散射和浑浊度等因素，传统的水下检测方法如声纳和基于光学的成像技术往往面临性能受限的问题。这些挑战使得水下目标检测成为一个亟待解决的技术难题，随着这些领域需求的增长，对于高效、准确的水下目标检测技术的迫切需求也随之增加[1]。

近年来，深度学习技术在图像识别和目标检测领域取得了显著的进展，为水下目标检测提供了新的解决方案[2]。深度学习技术，特别是卷积神经网络(CNN)，已经在处理复杂图像数据方面显示出卓越的性能，这为提高水下目标检测的准确性和效率提供了可能性[3]。

本研究旨在解决上述问题，进一步探索深度学习技术在水下目标检测中的应用，并重点解决如何适应水下图像的问题。为此，本文选择了已经广泛使用的深度卷积神经网络架构 ResNet50 作为基础[4]，并在 TensorFlow 框架下搭建了一个定制化的模型。通过对自制数据集的预处理和增强，本文致力于提高模型对水下目标的识别能力。本文将详细介绍数据集的准备、模型的构建和训练过程，并展示实验设置和结果，分析模型的性能，并讨论其在实际应用中的潜力。

本文的结构如下：首先，本研究将回顾水下目标检测的相关工作，特别是深度学习在此领域的应用。其次，将详细介绍本研究的方法论，包括数据集的准备、模型的构建和训练过程。再次，展示本文的实验设置和结果，分析模型的性能，并讨论其在实际应用中的潜力。最后，总结本研究的主要发现，并提出未来研究的方向。

2. 相关工作

水下目标检测技术的发展历程反映了从传统方法到现代深度学习方法的转变，本节将回顾水下目标检测的传统方法、深度学习方法在图像识别和目标检测中的应用，以及 ResNet50 架构在此领域的相关研究。

2.1. 水下目标检测的传统方法

传统的水下目标检测技术主要依赖于声纳系统和光学成像[5]。声纳系统通过发射声波并分析其反射信号来检测和定位水下目标，但这种方法在分辨率和距离感知方面存在局限性。光学成像可以提供更高

的分辨率[6],但其性能受限于水下环境的光照条件和水质清晰度。这些方法在处理复杂水下环境时面临诸多挑战,如光线衰减、水体散射和浑浊度等。

2.2. 深度学习在图像识别和目标检测中的应用

深度学习技术,尤其是卷积神经网络(CNN),已经在图像识别和目标检测领域取得了显著成就[7]。CNN能够从大量图像数据中自动学习图像特征,并有效处理复杂和变化的图像场景。在水下目标检测领域,深度学习方法已被证明可以有效克服传统方法的局限性,提高检测的准确性和鲁棒性[8]。

2.3. ResNet50 架构的相关研究

ResNet(残差网络)是一种深度学习架构,因其独特的“跳跃连接”设计而闻名,该设计允许网络学习输入和输出之间的残差,从而解决了更深网络的训练难题[9]。ResNet50是ResNet系列中的一个变体,其包含50层网络,目前已在多个图像识别任务中表现出色。在水下目标检测领域,ResNet50的深层网络结构使其能够捕捉复杂的水下图像特征,提高检测的准确性。

综上所述,尽管传统方法在水下目标检测中有其应用价值,但深度学习,特别是ResNet50架构,为该任务提供了一种更为先进和有效的解决方案。本研究旨在进一步探索和验证ResNet50在水下目标检测中的应用潜力[10]。

3. Resnet50 基本结构

ResNet50是ResNet网络中的经典网络模型,其中的“50”表示网络的深度,即由50层组成[11],表1详细说明了ResNet50的网络架构,ResNet50作为深度卷积神经网络的主要模型,在图像识别任务中取得了显著的成功,其深度结构和残差块的设计为解决深度网络训练难题提供了有效的方案,成为深度学习领域的经典模型之一[12]。

Table 1. Basic architecture configuration of ResNet50
表 1. ResNet50 的基本结构配置

Conv layer	输出大小	50-layers	数目
Stage1: Conv_1	112 × 112	7 × 7, 64, stride2	
		3 × 3 max_pool, stride2	
Stage2: Conv_2	56 × 56	$\begin{bmatrix} 1 \times 1, (64) \\ 3 \times 3, (64) \\ 1 \times 1 (256) \end{bmatrix}$	
Stage3: Conv_3	28 × 28	$\begin{bmatrix} 1 \times 1, (128) \\ 3 \times 3, (128) \\ 1 \times 1 (512) \end{bmatrix}$	3
Stage4: Conv_4	14 × 14	$\begin{bmatrix} 1 \times 1, (256) \\ 3 \times 3, (256) \\ 1 \times 1 (1024) \end{bmatrix}$	4
Stage 5: Conv_5	7 × 7	$\begin{bmatrix} 1 \times 1, (512) \\ 3 \times 3, (512) \\ 1 \times 1 (2048) \end{bmatrix}$	6

ResNet 网络模型中的残差块模型如图 1 所示。

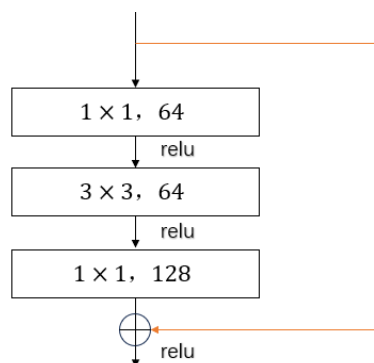


Figure 1. The graph residual block model

图 1. 图残差块模型

图 1 展示了深度网络中提出的一个块，称为“bottleneck”块，其主要目的是降维，首先使用 1×1 卷积将 256 维信道减少到 64 维信道，接着，使用 256 个信道的 1×1 卷积来恢复。

ResNet50，作为深度卷积神经网络的一个突出代表，是解决复杂图像识别任务，如水下生物检测的理想选择[12]，它的核心特性是残差学习，通过引入跳跃连接来解决深层网络训练中的退化问题。这种结构允许网络学习输入和输出之间的残差，从而有效地优化训练过程并构建更深的网络。ResNet50 的基本结构包括一个初始的卷积层和池化层，用于提取图像的基本特征，接着是 16 个残差块，每个块由 3 个卷积层组成，采用瓶颈设计以减少计算复杂度。跳跃连接在这些块中起到关键作用，有助于缓解梯度消失问题，提高学习能力。网络的末端是一个全局平均池化层，替代了传统的全连接层，进一步减少了模型的参数数量。在水下生物检测应用中，ResNet50 的这些特性使其能够有效地处理和识别在复杂水下环境中捕获的图像，从而提高检测的准确性和效率[13]。

4. 模型改进

4.1. 改进方法说明

为提高模型性能，本研究引入了 SIMAM 注意力机制，其原理是模仿大脑中激活神经元抑制周围神经元的方式，以强调图像中更重要的区域，其通过计算特征图的局部差异性来分配注意力权重，使用 Sigmoid 函数进行归一化，从而增强模型对关键特征的敏感度。加入 SIMAM 的设计不增加额外的模型参数，可以轻松集成到现有的神经网络中，在提高分类等任务的性能的同时不会增加计算负担。

4.2. 对比实验结果

对比实验结果如下表 2 所示。

Table 2. Comparison of experimental results

表 2. 实验结果对比

模型	Map	Imgsz	Param
ResNet50 (baseline)	64.6	256 × 256	91.44
ResNet50_SIMAM	68.35	256 × 256	91.44
MobileNetv2	68.15	256 × 256	9.6
ShuffleNet	65.14	256 × 256	5.68

5. 实验与分析

5.1. 数据获取

本文使用数据集[11]分为 20 个类, 共有 5443 张图片, 图 2 为部分数据集展示, 其中图左为灰鲨, 图右为肥胖三叶虫。

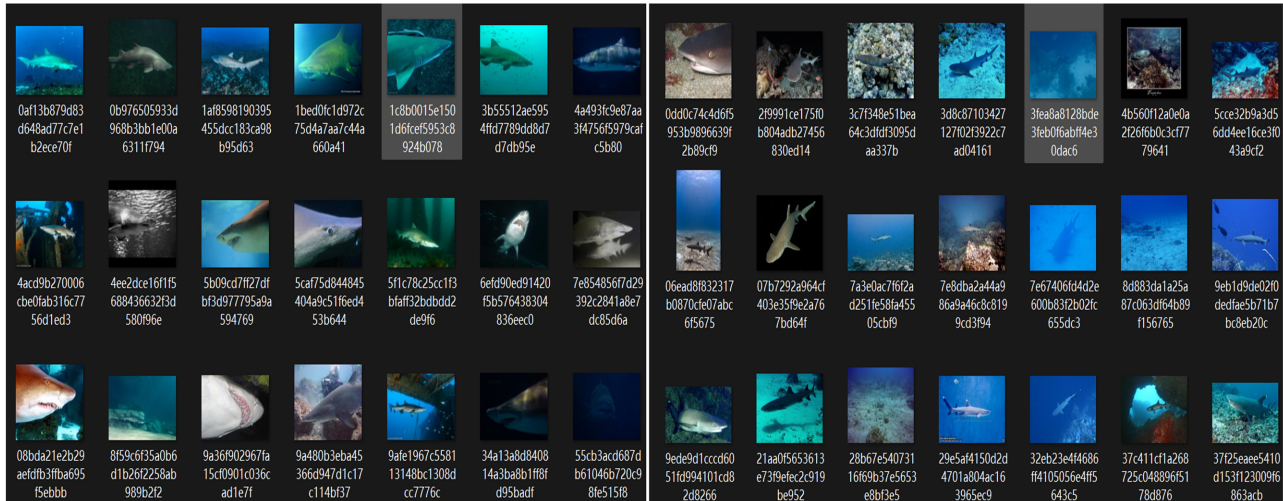


Figure 2. Part of the dataset images

图 2. 部分数据集图片展示

5.2. 数据读取

本文定义了一个函数 `read_data(data_dir)`, 用于从指定的目录中读取图像文件, 这一步骤是数据预处理的基础, 确保了模型能够接收到正确的输入数据。在读取图像时同时从文件名中提取标签信息, 这些标签对应于图像中的水下生物类别。为了标准化输入数据并减少模型训练时的计算负担, 本文对图像数据进行了归一化处理, 即将像素值除以 255, 从而将数据范围限定在 0 到 1 之间。此外, 所有图像数据被转换为 `numpy` 数组, 以便于后续的处理和模型训练。

5.3. 数据预处理

1) 数据打乱: 本研究使用了 `numpy` 的随机种子以及 `shuffle` 函数对图像数据以及标签进行了打乱, 有助于防止过拟合的现象发生, 同时可以使模型更加具有泛化性以及鲁棒性[14]。

2) 数据增强: 采用了 `ImageDataGenerator` 对训练数据进行增强。这包括了一系列的图像变换操作, 如随机旋转、宽度偏移、高度偏移和缩放等。这些操作能够人为地增加数据集的多样性, 模拟水下环境中可能出现的各种变化, 如光照变化、视角差异和物体的部分遮挡等。通过这种方式, 能够显著提高模型对新、未见过数据的泛化能力, 从而提高其在实际应用中的准确性和可靠性。

5.4. 数据训练

在本研究中, 采用了基于 `TensorFlow` 框架的 `ResNet50` 模型进行水下生物检测的数据训练。`ResNet50`, 作为一种深度卷积神经网络, 被配置为包含一个全局平均池化层和两个密集连接层, 最终层采用 `softmax` 激活函数以实现多类别分类。模型编译时, 选择了 `Adam` 优化器和 `Sparse Categorical Crossentropy` 损失函数, 以及 `sparse_categorical_accuracy` 作为评估指标[15]。

为了优化训练过程，实施了多个回调函数，包括模型检查点以保存训练过程中的最佳模型，早停机制以防止过拟合，以及学习率调整以提高模型训练的精度。训练过程中，本文使用了 `model.fit` 方法，设置批量大小为 4，并进行了 200 个训练周期，同时将 15% 的数据用作验证集，以监控和评估模型性能。

训练完成后，模型被保存为 H5 文件格式，以便于未来的应用和复用。此外，本研究利用 `matplotlib` 库绘制了训练和验证过程中的准确率和损失曲线，以直观评估模型的性能和学习过程。这一全面而系统的训练方法为水下生物检测任务提供了一个坚实且有效的深度学习模型基础。在进行数据增强的消融实验时，目的是观察去除或修改其中某些增强技术对模型性能的影响。因此需要对每个配置重新训练模型，然后记录和比较其性能，数据增强配置的消融实验如表 3 所示。

Table 3. Ablation experiments for data augmentation configurations

表 3. 数据增强配置的消融实验

实验编号	重缩放 (rescale)	旋转范围 (rotation_range)	宽度偏移 (width_shift_range)	高度偏移 (height_shift_range)	缩放范围 (zoom_range)	训练集 准确率	验证集 准确率
基线	1./255	0	0	0	0	99.81%	91.17%
实验 1	1./255	45	0.15	0.15	0.5	99.87%	92.23%
实验 2	1./255	0	0.15	0.15	0.5	99.75%	91.52%
实验 3	1./255	45	0	0	0.5	99.75%	91.87%
实验 4	1./255	45	0.15	0.15	0	99.87%	91.17%
实验 5	1./255	45	0.15	0.15	0.5	99.82%	91.17%

由表 3 可见，实验 1 在经过四种数据增强后，取得了最高的检测精准度。其中最优训练过程的 loss 和 accuracy 曲线如图 3 所示。

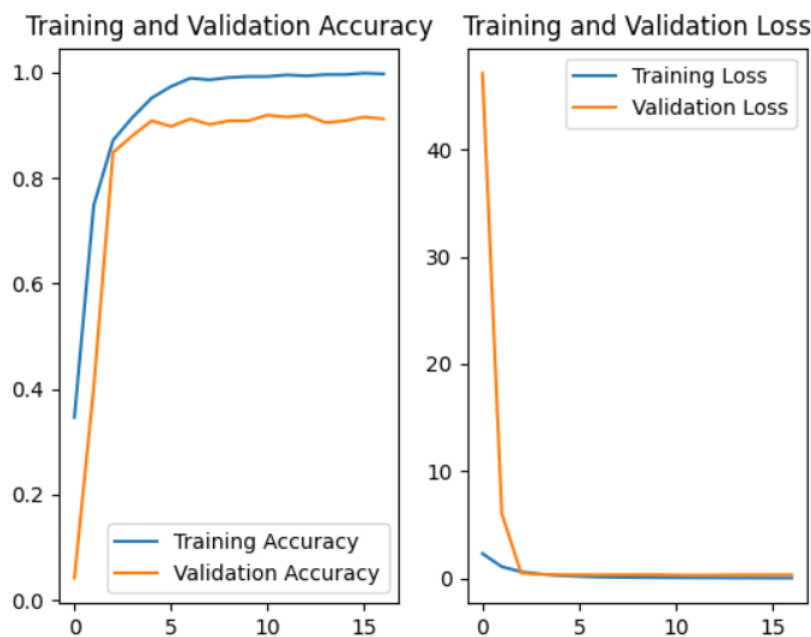


Figure 3. Loss and Accuracy curves

图 3. Loss 和 Accuracy 曲线

5.5. 实验结果

本文使用 Python 编写了 Resnet50 模型，并在一台 GPU 服务器上进行了训练。训练过程中，使用了数据增强技术，包括随机裁剪、旋转、缩放等，以增加数据集的多样性。最终训练的模型在测试集上取得了较好的效果，能够将图片的不同种类的鱼类分类出来，并且有较高的精准度，具有实际应用价值。

实验结果具有创新性与现有研究成果对比：

在这项实验中，本文使用了基于 Resnet (Residual Network)的方法来进行水下生物检测。与现有的研究成果相比，本研究在以下几个方面具有创新性：

一是数据预处理和增强策略：在数据预处理阶段，采用了先进的数据增强技术，如随机旋转、缩放和平移等，以模拟水下环境中可能遇到的各种情况。这些策略显著提高了模型对新场景的适应能力和泛化性。

二是复杂水下环境的适应性：模型特别考虑了水下环境的复杂性，如光照变化、水体散射和浑浊度等因素。通过对模型进行细致的调整和训练，本研究提高了模型在这些具有挑战性条件下的识别准确率和鲁棒性。

三是实际应用场景的考虑：在设计时充分考虑了实际应用场景，确保模型不仅在实验室条件下有效，而且适用于真实的水下环境，提高了本研究的实用价值和应用前景。

综上所述，本研究通过结合先进的深度学习技术和针对性的数据处理策略，为水下生物检测领域带来了新的视角和方法，展示了深度学习在解决复杂环境下生物识别问题中的巨大潜力。

实验结果如图 4 所示，其中左图为待检测图片，右图为图片检测结果。

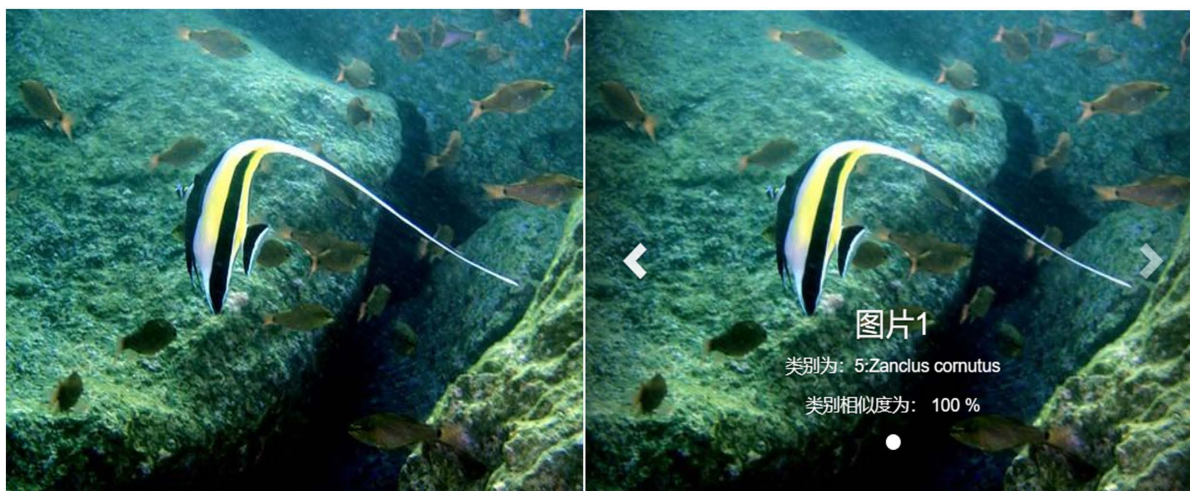


Figure 4. Example figure of detection results

图 4. 检测结果示例图

在复杂的水下环境下，本研究克服了诸多技术挑战，采用了创新的生物检测方法，最终成功地实现了高准确率的水下生物检测。这一成果不仅证明了技术的有效性，也为未来的水下生物研究提供了新的思路和方法。因此，本研究在复杂的水下环境下取得了很高的检测准确率，为相关领域的发展做出了重要贡献。

6. 结语

本研究提出并验证了一种基于深度学习的高效准确的水下目标检测方法。通过定制 ResNet50 模型并

引入 SIMAM 注意力机制，成功地提高了检测的精确度。实验证明该方法可以有效应对水下环境的复杂性和视觉清晰度的限制，具有状态良好的检测性能。这也为广泛的水下目标检测应用奠定了基础，为深度学习在复杂水下视觉任务中的应用提供了范例。本研究展示了深度学习模型通过定制化和创新机制的引入，可以克服水下目标检测面临的独特挑战，为未来建立适用于多种极端水下条件的检测系统提供了宝贵经验和启发。总体而言，本研究创新性地实现了针对水下目标检测任务的高效深度学习模型设计和优化，为相关领域的研究和应用提供了高效的工具。

致 谢

感谢大连外国语大学软件学院及创新创业学院对本次研究的支持。

基金项目

大连外国语大学 2023 年大学生创新创业项目。

参考文献

- [1] Chen, L., Yin, T., Zhou, S., *et al.* (2023) Underwater Target Detection Algorithm Based on Feature Fusion Enhancement. *Electronics*, **12**, Article 2756. <https://doi.org/10.3390/electronics12132756>
- [2] Wang, X., Xue, G., Huang, S., *et al.* (2023) Underwater Object Detection Algorithm Based on Adding Channel and Spatial Fusion Attention Mechanism. *Journal of Marine Science and Engineering*, **11**, Article 1116. <https://doi.org/10.3390/jmse11061116>
- [3] Chen, R. and Chen, Y. (2023) Improved Convolutional Neural Network YOLOv5 for Underwater Target Detection Based on Autonomous Underwater Helicopter. *Journal of Marine Science and Engineering*, **11**, Article 989. <https://doi.org/10.3390/jmse11050989>
- [4] Hua, X., Cui, X.P., Xu, X.H., *et al.* (2023) Underwater Object Detection Algorithm Based on Feature Enhancement and Progressive Dynamic Aggregation Strategy. *Pattern Recognition*, **139**, Article ID: 109511. <https://doi.org/10.1016/j.patcog.2023.109511>
- [5] Song, P.H., Li, P.T., Dai, L.H., *et al.* (2023) Boosting R-CNN: Reweighting R-CNN Samples by RPN's Error for Underwater Object Detection. *Neurocomputing*, **530**, 150-164.
- [6] Xu, S.B., Zhang, M.H., Song, W., *et al.* (2023) A Systematic Review and Analysis of Deep Learning-Based Underwater Object Detection. *Neurocomputing*, **527**, 204-232.
- [7] Chen, L., Yang, Y.Y., Wang, Z.H., *et al.* (2023) Lightweight Underwater Target Detection Algorithm Based on Dynamic Sampling Transformer and Knowledge-Distillation Optimization. *Journal of Marine Science and Engineering*, **11**, Article 426. <https://doi.org/10.3390/jmse11020426>
- [8] Chen, Y., Song, P.H., Liu, H., *et al.* (2023) Achieving Domain Generalization for Underwater Object Detection by Domain Mixup and Contrastive Learning. *Neurocomputing*, **528**, 20-34.
- [9] Fu, C.P., Liu, R.S., Fan, X., *et al.* (2023) Rethinking General Underwater Object Detection: Datasets, Challenges, and Solutions. *Neurocomputing*, **517**, 243-256.
- [10] Yuan, W.Q., Fu, C.P., Liu, R.S., *et al.* (2022) SSoB: Searching a Scene-Oriented Architecture for Underwater Object Detection. *The Visual Computer*, **39**, 5199-5208. <https://doi.org/10.1007/s00371-022-02654-4>
- [11] Cheng, P.Y., Ting, Y.Z., Dong, G.L., *et al.* (2023) Water Source Identification in Mines Combining LIF Technology and ResNet. *Journal of Mountain Science*, **20**, 3392-3401. <https://doi.org/10.1007/s11629-023-8189-0>
- [12] Xu, L.Y., Huang, J.J., Zhang, H., *et al.* (2022) Direction of Arrival Estimation of Acoustic Sources with Unmanned Underwater Vehicle Swarm via Matrix Completion. *Remote Sensing*, **14**, Article 3790. <https://doi.org/10.3390/rs14153790>
- [13] Fayaz, S., Parah, S.A., *et al.* (2022) Underwater Object Detection: Architectures and Algorithms – A Comprehensive Review. *Multimedia Tools and Applications*, **81**, 20871-20916. <https://doi.org/10.1007/s11042-022-12502-1>
- [14] Wang, X.H., Zhu, Y.G., Li, D.Y. *et al.* (2022) Underwater Target Detection Based on Reinforcement Learning and Ant Colony Optimization. *Journal of Ocean University of China*, **21**, 323-330. <https://doi.org/10.1007/s11802-022-4887-4>
- [15] Qi, J.H., Gong, Z.Q., Yao, A.H., *et al.* (2021) Bathymetric-Based Band Selection Method for Hyperspectral Underwater Target Detection. *Remote Sensing*, **13**, Article 3798. <https://doi.org/10.3390/rs13193798>