# EC-YOLOX: A Deep-Learning Algorithm for Floating Objects Detection in Ground Images of Complex Water Environments

Jiaxin He, Yong Cheng, Wei Wang, Yakang Gu, Yixuan Wang, Wenjie Zhang, Achyut Shankar, Shitharth Selvarajan, and Sathish A. P. Kumar, *Senior Member, IEEE*

***Abstract*—Correct detection of floating objects in complex water environments is a challenge because of the problems of obscuration and dense floating objects. In view of the above issues, this article proposed a network called EC-YOLOX by introducing the coordinate attention (CA) and efficient channel attention (ECA) mechanism and improving the loss function to further the multifeature extraction and detection accuracy of floating objects. In this article, ablation experiments and comparison experiments were conducted on the river floating objects dataset. The ablation experiments showed that the ECA and CA mechanism played a great role in EC-YOLOX, which can reduce the missed detection rate by 5.86% and increase the mean average precision (mAP) by 5.53% compared with YOLOX. The EC-YOLOX was also applicable to different types of floating objects; the mAP of the ball, plastic garbage, plastic bag, leaf, milk box, grass, and branches were, respectively, improved by 4%, 4%, 4%, 6%, 4%, 18%, and 5%. The mAP of the comparison experiments was improved by 15.13%, 9.30%, and 8.03% compared to faster R-CNN, YOLOv5, and YOLOv3, respectively. This method facilitates the precise extraction of floating objects from images, which holds paramount importance for monitoring and safeguarding water environments. It offers significant contributions to water environment monitoring and protection.**

***Index Terms*—Attention mechanism, floating objects, loss function, missed detection rate, YOLOX.**

Manuscript received 31 August 2023; revised 30 October 2023 and 29 November 2023; accepted 14 February 2024. Date of publication 27 February 2024; date of current version 3 April 2024. This work was supported by the National Natural Science Foundation of China under Grant 42201053, Grant 41975183, and Grant 41875184. *(Corresponding author: Wenjie Zhang.)*

Jiaxin He and Wei Wang are with the School of Automation, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail: jiaxhe@nuist.edu.cn; wangwei087@nuist.edu.cn).

Yong Cheng, Yakang Gu, and Yixuan Wang are with the School of Software, Nanjing University of Information Science & Technology, Nanjing 210044, China (e-mail: yongcheng@nuist.edu.cn; holddestiny@nuist.edu.cn; kiken99@nuist.edu.cn).

Wenjie Zhang is with the School of Geographical Sciences, Nanjing University of Information Science & Technology, Nanjing 210044, China (e-mail: zhangwenjie@nuist.edu.cn).

Achyut Shankar is with the Department of Cyber Systems Engineering, WMG, University of Warwick, CV74AL Coventry, U.K., also with the Center of Research Impact and Outcome, Chitkara University, Rajpura 140401, India, and also with the School of Computer Science Engineering, Lovely Professional University, Phagwara 144411, India (e-mail: ashankar2711@gmail.com).

Shitharth Selvarajan is with the School of Built Environment, Engineering and Computing, Leeds Beckett University, LS1 3HE Leeds, U.K. (e-mail: s.selvarajan@leedsbeckett.ac.uk).

Sathish A. P. Kumar is with the Department of Computer Science, Cleveland State University, Cleveland, OH 44115 USA (e-mail: s.kumar13@csuohio.edu).

Digital Object Identifier 10.1109/JSTARS.2024.3367713

## I. Introduction

FLOATING objects not only pollute water environments but also cause water quality degradation [1], [2], [3], [4]. Floating objects can be classified into degradable and nondegradable types, with the latter being the majority, e.g., plastic bottles and garbage bags. The quantity and type of floating objects can reflect water quality, which is relatively easy to identify and acquire compared to the chemical composition. Automatic detection of floating objects through video or image processing is more convenient and efficient than manual recognition [38]. Consequently, image-processing-based methods for detecting floating objects on water surfaces have emerged as a prominent research area within the field of small object detection [5], [39].

Traditional image processing typically consists of three fundamental steps: preprocessing, feature extraction, and classification recognition. During the preprocessing stage, operations like filtering, enhancement, and denoising are performed on the image with the goal of improving image quality and reducing noise and errors in subsequent processing. Techniques, such as average filtering [8] and histogram equalization [9], are examples of this. Kataoka and Nihei [10] employed histogram equalization as a method to investigate the transportation of macrodebris in rivers. However, this technique can be susceptible to adverse effects caused by factors, such as aquatic plants, sunlight reflection, and water conditions. As a result, it may not be suitable for accurately detecting floating objects in complex water areas. In the feature extraction step, local features, morphology, color, and other information from the image are analyzed to extract features relevant to the target recognition. This prepares the groundwork for subsequent classification and recognition. However, feature extraction is typically performed manually and lacks adaptability and generalization [37]. Finally, in the classification recognition stage, traditional image processing employs support vector machines or clustering methods to classify and recognize the extracted features, thereby accomplishing tasks, such as detection, recognition, and tracking of the target [6], [7], [40]. However, traditional classifiers are typically based on artificial features and shallow models, making them less capable of handling complex image features and scenarios, and they often yield unsatisfactory results when dealing with nonlinear problems [41].

With the rapid development of computer vision, deep learning has showcased remarkable advantages in image processing, assuming a pivotal role in various aspects including image processing, object detection, crowd sensing, semantic segmentation, and behavior recognition [11], [12], [13], [14], [15], [16], [17], [42], [43]. Object detection algorithms can be categorized into two main groups: two-stage detectors, such as mask R-CNN [18] and faster R-CNN [19], and one-stage detectors, such as the YOLO series [20], [21], [22], [23], [24]. Among the two-stage detectors, Liu et al. [25] introduced a water surface floating object detection model by improving mask R-CNN, which effectively enhances the recognition capability of irregular floating objects. Li et al. [26] incorporated the Class Activation network into Faster R-CNN to tackle the problem about low recognition and localization accuracy of floating objects. Although two-stage detection algorithms improve the recognition of floating objects to a certain extent, they may extract some regions that are irrelevant to the target during the region proposal process, leading to false alarms [20].

Single-stage detection algorithms have gained popularity over two-stage detection algorithms due to their ability to directly generate prediction results from input images and solve the detection task through regression, thereby enhancing the detection accuracy of floating debris. Lin et al. [27] improved YOLOv5 by introducing an enhanced algorithm that extends the dataset and incorporates attention mechanisms, enabling rapid detection of multiple categories of floating objects. Li et al. [28] utilized an improved k-means clustering algorithm to obtain prior boxes and integrated category activation mapping into the YOLOv3 network, resulting in improved detection accuracy of floating debris and reduced localization errors. To tackle the challenges posed by small target sizes, complex scenes, and various noise sources, Chen et al. [29] presented a method for detecting and tracking floating debris. Their approach leverages temporal–spatial information fusion and utilizes the SSD network, specifically designed to address small floating debris detection tasks, while simultaneously estimating the position and size of the detected debris. Although these single-stage detection network frameworks perform well in floating debris detection tasks, in complex environmental backgrounds, considering the different scales of floating debris, they may not accurately extract each type of floating debris, resulting in missed detections and decreased detection accuracy [30], [31].

Hence, surmounting the challenges posed by complex water environment interferences and enhancing feature map scales are crucial factors in improving the detection of floating objects on water surfaces. The coordinate attention (CA) mechanism has proven to be effective in addressing the issue of complex background interference [32]. The efficient channel attention (ECA) module can better handle feature information of different scales by adaptively weighting between multiscale features, effectively addressing the problem of objects with varying sizes [33]. The varifocal loss function can dynamically match positive samples, solving the problem of sample imbalance [34]. To further improve detection accuracy, this article introduces the CA and ECA attention mechanisms, as well as the zoom loss function module based on the YOLOX network. This approach ensures effective detection of targets in dense and complex backgrounds while also enhancing the detection ability of multiscale targets. This article makes the following primary contributions.

1) We introduce a floating object detection method, EC-YOLOX, designed for complex water environments. Ablation experiments confirm substantial performance enhancements attributable to the CA module, the ECA module, and the varifocal loss function.
2) Comparative experiments conducted on a river floating object dataset demonstrate that EC-YOLOX exhibits superior performance, thereby substantiating its effectiveness.
3) Compared to baseline models, EC-YOLOX achieves the lowest overall missed detection rate for various categories of floating objects. This improvement underscores its accuracy and precision in detection.

## II. Methodology

### A. The Improved Network

In 2021, Megvii Technology introduced a novel network called YOLOX, exhibiting superior property and advantages over YOLOv3 and YOLOv5. YOLOX adopts anchor-free methods and dynamically matches positive samples of different target sizes, improving both detection speed and accuracy through data augmentation and decoupling heads. First, the network takes in $640 \times 640$ images, which are further augmented using Mixup and Mosaic. Additionally, the output is subjected to processing through the Darknet53 backbone network and the feature pyramid network (FPN), thereby enhancing the accuracy of detection. Finally, three decoupled heads are introduced at the output end to predict the feature map output by the neck layer.

YOLOX has shown good detection performance in object detection while also having advantages, such as low parameter count and being conducive to detecting small objects. However, when detecting floating objects, YOLOX may encounter difficulties in accurately identifying each class of object due to the diverse range of sizes that such objects can exhibit. The ECA and CA attention mechanisms can effectively address this issue. Therefore, this article chooses to improve YOLOX by introducing the ECA and CA attention mechanisms, as shown in Fig. 1.

As shown in Fig. 1, first, Mixup and Mosaic techniques are used for data augmentation, and the augmented data are introduced into the CSPdarknet53 module to extract feature information. Subsequently, the obtained feature maps are passed through the ECA attention module, which effectively captures interdependencies among different channels, enhancing the discriminative ability of the characteristics. Moreover, to further enhance the algorithm's performance, the CA mechanism is integrated into features at different scales, namely P3, P4, and P5. Last, the decoupling head generates predictions for three distinct multiscale features, capturing information from various scales.

### B. Coordinate Attention Mechanism

In the context of detecting floating objects on water surfaces, several challenges arise, including complex water environments

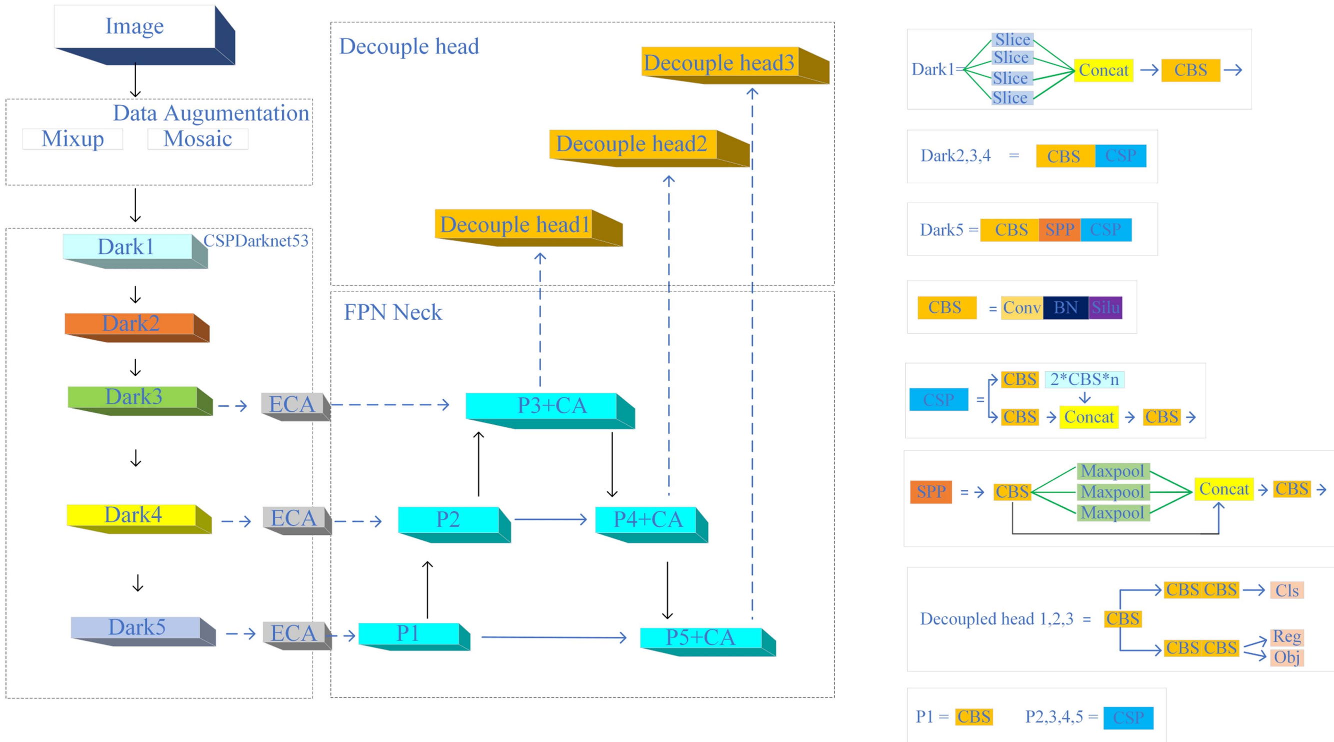


Fig. 1. Improved network structure: EC-YOLOX.

and the scattered distribution of floating objects. Existing networks cannot ignore the interference of background information when facing these problems, which leads to missed detection and other issues, as the models cannot focus on each type of floating object. This article presents the CA attention module, which addresses the challenge of background interference by selectively attending to more relevant parts of the target. By focusing on these specific regions, the CA attention module reduces the influence of the background environment and extracts more crucial features. The introduction of the CA attention module contributes to improved detection accuracy in detecting floating objects.

The CA attention mechanism is a position encoding-based attention mechanism that utilizes the coordinate information of each position in the input image to calculate attention weights and obtain the feature representation of the entire image [32]. This mechanism partitions the input image into multiple grids, where each grid comprises position information, such as the center or upper-left coordinates. Then, it uses the position information of each grid as input to calculate its attention weight through an attention model, which can be a simple multilayer perceptron (MLP) or a complex neural network. Finally, CA uses the calculated attention weights to enhance the feature representation and obtain the feature information from all targets in the entire image. Through the utilization of position encoding, the CA attention mechanism is able to direct attention to various regions within the image and effectively capture spatial information. The CA attention module is illustrated in Fig. 2, providing a visual representation of its structure and functionality.


Fig. 2. Coordinate attention mechanism.

As per Fig. 2, during the recognition of floating objects, the input feature map, denoted as "Input" with dimensions of $C \times H \times W$, undergoes pooling operations. This process generates two separate feature graphs: one with dimensions of $C \times H \times 1$, obtained from pooling along the $X$ direction, and another with dimensions of $C \times 1 \times W$, obtained from pooling along the $Y$ direction. The pooling operations in both directions are performed to capture and extract relevant information along the horizontal and vertical dimensions of the feature graph. Subsequently, the generated feature maps undergo a transformation process and are fused together to generate a new feature graph, called $f_1$. Next, $f_1$ undergoes further processing by applying a $1 \times 1$ convolutional kernel, resulting in a reduction and activation of the feature map. This operation generates a new feature map, denoted as $f_2$. Last, the feature map $f_2$ is split along the spatial dimension, resulting in the creation of two new feature maps: $f_3$ and $f_4$. Subsequently, the feature maps $f_3$ and $f_4$ are upsampled using $1 \times 1$ convolutional kernels. This upsampling operation increases the spatial resolution of the feature graphs, expressing more localized information. The upsampled feature maps are then imported into the sigmoid activation function, resulting in generating final attention vectors $g_3$ and $g_4$.

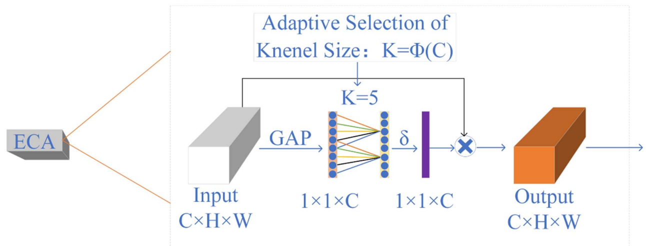


Fig. 3. Efficient channel attention mechanism.

### C. Efficient Channel Attention Mechanism

Floating objects on water surfaces exhibit a diverse range of types, and their sizes can vary significantly. However, the varying sizes of these objects pose challenges for the network in extracting accurate and discriminative features across different types of floating objects. Current networks cannot extract feature information of floating objects at different scales, leading to missed detection as the models cannot accurately detect each type of floating object. This article introduces the ECA module, which adaptively weights multiscale features for better handling of feature information. This enables the network to improve its feature representation ability for floating objects at different scales, thereby enhancing detection performance. The ECA module is depicted in Fig. 3, providing a visual representation of its structure and functionality.

The process of the ECA is depicted in Fig. 3, and it can be divided into several steps for its operation. First, the input feature map undergoes global average pooling, resulting in a channel vector. Second, the weights of each channel are computed by a small MLP. The channel vector serves as input to the MLP, and a scalar weight is generated as output. Third, each channel's feature representation is multiplied by its corresponding weight, thereby adjusting the importance or relevance of different channels. Last, the adjusted feature representation is passed as input to the next layer for further processing.

### D. Improved Loss Function

The loss function of the YOLOX algorithm mainly includes intersection over union (IoU) loss ($\mathrm{Loss_{IoU}}$), classification loss ($\mathrm{Loss_{cls}}$), and confidence loss ($\mathrm{Loss_{obj}}$), where obj represents the predicted result of whether the feature points and positive/negative samples contain the target object. However, when using the BCELoss function to calculate $\mathrm{Loss_{obj}}$, the imbalance of samples can produce error. Therefore, we use the varifocal loss function to improve the detection ability for targets in dense environment [34]. This loss function divides the target samples based on their number and then selects positive samples and rejects negative samples accordingly. This approach results in a more precise detection ranking of dense targets. Formulas (1) and (2) show the calculation method of the varifocal loss.

From formulas (1) and (2), it can be seen that $p$ represents the predicted IoU-aware classification score and $q$ represents the target score. For foreground points, the true value of $q$ is the IoU; otherwise, it is 0. On the other hand, for background points or regions that are not part of any object, the true value of $q$ is 0, indicating no overlap or relevance to any specific object.

In YOLOX, $\mathrm{Loss_{IoU}}$ represents the IoU loss, which serves as a metric to measure the detection performance. Formula (3) reflects the power function calculation of $\mathrm{Loss_{IoU}}$. $\mathrm{Loss_{cls}}$ in YOLOX represents the classification loss associated with the recognition of floating objects. This loss term is computed using the binary cross entropy (BCELoss) function, which is commonly employed for binary classification tasks. By utilizing BCELoss, the network model aims to optimize the classification accuracy of floating objects, enhancing the stability and performance of the overall network.

To tackle the challenges posed by the mixed distribution and varying scales of floating objects on water surfaces, we optimize the loss function, specifically $\mathrm{Loss_{obj}}$, by employing varifocal loss. This approach aims to enhance the accuracy of the model's detection ranking and improve overall performance. By utilizing varifocal loss, the model can effectively handle the inherent variations in object sizes and distributions, allowing for more precise and reliable detection results. Equation (4) is the final expression

$$\mathrm{VFL} = -q(q\log(p) + (1-q)\log(1-p)(q>0) \tag{1}$$

$$\mathrm{VFL} = -\alpha p^{\gamma}\log(1-p)(q=0) \tag{2}$$

$$\mathrm{Loss_{IoU}} = 1 - \mathrm{IoU}^2 \tag{3}$$

$$\mathrm{Loss} = \mathrm{Loss_{IoU}} + \mathrm{Loss_{obj}} + \mathrm{Loss_{cls}}. \tag{4}$$

## III. Experiment

### A. Experiment Environment

The experimental setup involves using the Windows 10 operating system, an NVIDIA GeForce RTX 3070Ti GPU, and 8 GB of memory. The deep-learning framework utilized is PyTorch 1.7.1 with CUDA 11.6. For this experiment, pretrained weights from the VOC dataset are employed, and the training process is divided into two phases: the freezing phase and the thawing phase. Among them, the number of training times in the freezing phase is set to 50 epochs, and the batch-size is set to 4. The number of training times after the unfreezing phase is set to 300 epochs, and the batch-size is set to 4. The model's optimizer is stochastic gradient descent, and the learning rate is adjusted using the cosine annealing method, which helps optimize the training process over iterations. To enhance the available data, the Mixup and Mosaic methods are employed, which provide data augmentation techniques to improve the model's generalization and robustness by combining and manipulating training samples. These techniques collectively contribute to the performance improvement and accuracy enhancement of the model in object detection tasks.

### B. Dataset

The dataset used in this experiment is taken from river scenes and annotated in Pascal VOC format, containing eight object classes, including bottle, grass, branch, milk box, plastic bag, plastic garbage, ball, and leaf. The dataset consists of 2400

Fig. 4. River floating objects dataset.

images and is divided into training, test, and validation sets according to the 8:1:1 ratio [27]. Fig. 4 displays a selection of sample images extracted from the dataset.

### C. Evaluation Metrics

To assess the effectiveness of EC-YOLOX in detecting floating objects on water surfaces, two evaluation metrics, namely recall and frames per second (FPS), are employed. Equations (5) and (8) are used to express recall and precision. The average precision (AP) and mean average precision (mAP) are represented by (6) and (7), respectively.

In the evaluation metrics, TP represents the quantity of affirmative examples in which a class is correctly predicted, FP denotes the quantity of affirmative examples in which a class is misclassified, FN represents the quantity of unfavorable examples in which a class is misclassified, $P(r)$ represents the precision, and $K$ represents the number of categories

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{5}$$

$$\mathrm{mAP} = \frac{\sum_{i=1}^{k} \mathrm{AP}_i}{K} \tag{6}$$

$$\mathrm{AP} = \int_0^1 P(r)dr \tag{7}$$

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}. \tag{8}$$

### D. Ablation Experiment

To explore the effectiveness of improvement strategies, ablation experiments were conducted on the YOLOX baseline model on the floating objects dataset. Ablation experiments involve adding, replacing, or removing specific modules from the YOLOX baseline model and are designed to explore the impact of different improvement strategies on model performance. The settings and parameters were kept consistent across all experiments conducted, ensuring a fair and reliable comparison. The results obtained from these experiments are presented in Tables I and II.

According to Table I, the introduction of the CA attention module improves the AP (mAP) of the improved network by 3.31% compared to the baseline network, while the recall rate is improved by 3.29%. As depicted in Fig. 5(a) and (b), the incorporation of the CA module has resulted in notable enhancements in object detection. Specifically, the CA attention module has enabled the network to accurately detect and classify various objects, such as plastic bags, plastic garbage, leaves, branches, grass, and balls. The introduction of the CA attention module has significantly improved the detection accuracy of plastic bags, plastic garbage, and leaves. Moreover, it has also yielded a notable enhancement in the detection accuracy of branches, grass, and balls, with an increase of 14%, 10%, and 16%, respectively. These findings demonstrate the CA module that can improve the detection ability of the network across various categories of floating objects in complex water environments. The successful integration of the CA attention module has resulted in improved accuracy and robustness. According to Table I and Fig. 5(a) and (c), introducing the ECA attention module at the FPN structure results in a 3.42% increase in mAP and a 3.44% increase in recall rate. The enhancement of detection capabilities for objects, such as bottles, branches, plastic bags, and grass, was observed, with a significant improvement in detection accuracy for grass, branches, and balls by 17%, 14%, and 26%. This underscores that the ECA module can enhance the representational capacity of floating objects across diverse scales. As indicated by Table I and Fig. 5(a) and (d), the implementation of the improved loss function has resulted in notable

TABLE I
ABLATION EXPERIMENT OF THE IMPROVED MODULE

| CA | ECA | Vfl | Precision (%) | Recall (%) | mAP (%) | FPS (f/s) |
|---|---|---|---|---|---|---|
| - | - | - | 72.37 | 73.51 | 76.27 | 106.88 |
| √ | - | - | 69.15 | 76.80 | 79.58 | 78.51 |
| - | √ | - | 71.92 | 76.95 | 79.69 | 88.22 |
| - | - | √ | 66.56 | 73.96 | 78.35 | 111.34 |
| √ | √ | √ | 68.54 | 79.37 | 81.80 | 80.59 |

TABLE II
ABLATION EXPERIMENT RESULTS OF THE IMPROVED STRATEGY IN TERMS OF AP VALUES ARE SHOWN, WHERE B1–B8 REPRESENT BOTTLE, MILK BOX, BALL, PLASTIC BAG, PLASTIC-GARBAGE, BRANCH, GRASS, AND LEAF, RESPECTIVELY

| CA | ECA | Varifocalloss | B1 (%) | B2 (%) | B3 (%) | B4 (%) | B5 (%) | B6 (%) | B7 (%) | B8 (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| – | – | – | 93.77 | **79.29** | 90.70 | 82.27 | 87.01 | 59.61 | 43.68 | 73.86 |
| √ | – | – | 93.29 | 78.92 | 88.78 | **86.36** | 90.52 | 63.53 | 55.16 | 80.09 |
| – | √ | – | 92.29 | 77.30 | 87.28 | 83.46 | 90.22 | 65.24 | 59.50 | 82.24 |
| – | – | √ | **93.92** | 78.65 | 93.41 | 85.89 | 87.72 | 64.51 | 54.92 | 76.75 |
| √ | √ | √ | 92.92 | 78.24 | **94.81** | 85.59 | **91.03** | **65.49** | **61.65** | **83.36** |

The bold values indicate the highest detection accuracy for each type of floater under the various improvement strategies.

TABLE III
COMPARISON EXPERIMENT RESULTS

| Detecting algorithm | mAP (%) | FPS (f/s) |
|---|---|---|
| detr | 64.86 | 41.27 |
| Faster R-CNN | 66.67 | 37.15 |
| Retinanet | 69.19 | 33.47 |
| SSD (mobilenet) | 70.37 | 103.16 |
| Efficientdet | 72.50 | 27.50 |
| YOLOv5 | 72.50 | 124.75 |
| YOLOv3 | 73.77 | 76.21 |
| SSD(vgg) | 75.71 | 51.64 |
| YOLOX | 76.27 | 106.88 |
| EC-YOLOX (ours) | 81.80 | 80.59 |

improvements in the detection performance. Specifically, the improved loss function has enhanced the ability to screen and eliminate unfavorable examples, leading to a 2.08% increase in mAP and a 0.45% increase in recall rate. The improved loss function achieves this by more effectively penalizing false negatives and false positives, thus reducing their impact on the overall model performance. The enhanced screening and elimination capabilities enable the model to more accurately and reliably classify floating objects. The detection accuracy of bottles, balls, branches, plastic bags, and grass is improved, with the detection accuracy of branches, balls, grass, and plastic bags increasing by 2%, 20%, 17%, and 11%, respectively. In summary, the incorporation of the CA and ECA attention modules, along with the varifocal loss, has significantly enhanced the detection ability of EC-YOLOX. These improvements have resulted in a 5.53% gain in mAP and a 5.86% gain in recall rate, demonstrating that EC-YOLOX can achieve better detection property compared to YOLOX. Overall, the CA and ECA attention modules, as well as the varifocal loss, enhanced the detection performance of the model, resulting in a 5.53% mAP gain and a 5.86% recall rate gain. The obtained results provide strong evidence of the effectiveness of EC-YOLOX in achieving significantly better detection accuracy than the baseline network.

### E. Comparison Experiments Results

To assess the viability of EC-YOLOX, comparative experiments were conducted utilizing detr, faster R-CNN, Retinanet, SSD, Efficientdet, YOLOv3, YOLOv5, and YOLOX algorithms. As indicated in Table III, EC-YOLOX demonstrated superior detection accuracy compared to detr, faster R-CNN, Retinanet, SSD (Mobilenet), Efficientdet, YOLOv5, YOLOv3, SSD (vgg), and YOLOX, with mAP improvements of 16.94%, 15.13%, 12.61%, 11.43%, 9.30%, 9.30%, 8.03%, 6.09%, and 5.53%, respectively. Regarding prediction speed, while EC-YOLOX was outperformed by SSD (Mobilenet), YOLOv5, and YOLOX, it exhibited superior speed compared to all other algorithms at 80.59 FPS. These findings demonstrate that EC-YOLOX achieves both high detection accuracy and rapid prediction speed, affirming its feasibility and efficiency for the water surface detection of floating objects task.


(a) (b) (c) (d)

Fig. 5. Results of the improved strategies are presented as follows: (a) result for YOLOX, (b) results about adding CA for YOLOX, (c) results about adding ECA for YOLOX, and (d) results about adding varifocal loss for YOLOX.

As shown in Fig. 6, YOLOv3, YOLOv5, and YOLOX all exhibit missed detection issues when confronted with complex water environments and dense distributions of floating objects. However, as demonstrated in Fig. 6(a) and (d), EC-YOLOX did not exhibit missed detections for any category of floating object and even showed improved accuracy in detecting leaves compared to YOLOv3. Moreover, as shown in Fig. 6(b) and (d), when compared to YOLOv5, EC-YOLOX did not exhibit missed detections for balls, plastic garbage, and branches in dense distributions of floating objects on water surfaces, and did not miss detecting grass or plastic bags in complex backgrounds, such as areas densely populated with water plants. Furthermore, EC-YOLOX showed significantly improved accuracy in detecting branches and leaves, with an 11% and 17% increase in accuracy, respectively. When comparing Fig. 6(c) and (d), it is evident that EC-YOLOX did not exhibit missed detections in any water environment compared to YOLOX, and exhibited significantly improved accuracy in detecting bottles, branches, and grass. Overall, EC-YOLOX demonstrated superior detection performance compared to YOLOX. The AP and mAP values of YOLOX and EC-YOLOX are shown in Fig. 7. As shown in the figure, EC-YOLOX exhibited comparable accuracy to YOLOX in detecting bottles, while showing an increased accuracy of 4%, 4%, 4%, 6%, 4%, 5%, and 18% in detecting balls, plastic garbage, plastic bags, leaves, milk boxes, branches, and grass, respectively. These results demonstrate that EC-YOLOX can achieve further improvements in detecting floating objects on water surfaces, and validate its effectiveness in detecting different types of floating objects in complex and densely distributed water environments.


(a) (b) (c) (d)

Fig. 6. Detection results for the dataset are presented as follows: (a) for YOLOv3, (b) for YOLOv5, (c) for YOLOX, and (d) for EC-YOLOX.

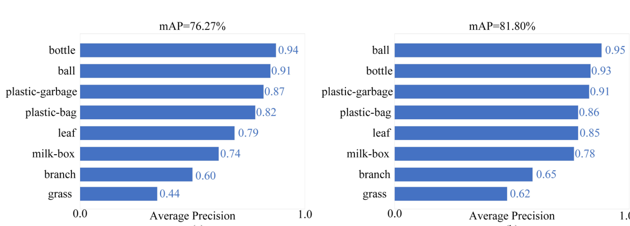


Fig. 7. Dataset comparison experiment AP and mAP values: (a) for YOLOX and (b) for EC-YOLOX.

## IV. Discussion

According to the ablation experiments, the EC-YOLOX network can improve the detection rate of floating objects on water surfaces, with a 5.53% increase in mAP compared to the YOLOX algorithm. The results presented in Tables I and II, along with Fig. 5, provide compelling evidence that the incorporation of the CA mechanism, as opposed to the baseline network, effectively captures crucial feature information of floating objects, overcoming the challenges posed by complex water environments. This improvement leads to a significant 3.31% increase in mAP. The inclusion of the ECA module effectively enhances the capability to represent floating objects of varying scales. This improvement leads to a notable 3.42% increase in mAP. The utilization of the varifocal loss function further enhances the detection ability of EC-YOLOX. In the dataset comprising floating objects on water surfaces, our method was compared against well-known algorithms, including faster R-CNN [19], YOLOv3 [22], and YOLOv5. As shown in Table III, EC-YOLOX achieves a higher mAP, with improvements of 15.13%, 8.03%, and 9.30% compared to faster R-CNN, YOLOv3, and YOLOv5, respectively, successfully completing the task of detecting floating objects. Overall, our method exhibits higher accuracy compared to other mainstream object detection algorithms.

TABLE IV
COMPARATIVE EXPERIMENT RESULTS PRESENT AP VALUES FOR DIFFERENT CATEGORIES, AS DEPICTED IN TABLES III AND IV

| Models | mAP (%) | FPS(f/s) | B1 (%) | B2 (%) | B3 (%) | B4 (%) | B5 (%) | B6 (%) | B7 (%) | B8 (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| **detr** | 64.86 | 41.27 | 92.40 | 73.18 | 61.31 | 81.30 | 60.06 | 59.10 | 30.23 | 61.33 |
| **Faster R-CNN** | 66.67 | 37.15 | 83.25 | 71.86 | 73.41 | 71.19 | 76.04 | 50.53 | 33.91 | 73.14 |
| **Retinanet** | 69.19 | 33.47 | 93.07 | 66.82 | 81.92 | 82.82 | 54.45 | 57.99 | 45.36 | 71.12 |
| **SSD (mobilenet)** | 70.37 | 103.16 | 90.23 | 72.90 | 89.32 | 64.89 | 78.67 | 56.11 | 47.27 | 63.56 |
| **Efficientdet** | 72.50 | 27.50 | 93.68 | 65.31 | 88.68 | 83.25 | 72.65 | 52.95 | 47.13 | 76.34 |
| **YOLOv5** | 72.50 | 124.75 | 91.41 | 67.56 | 91.37 | 80.89 | 79.00 | 56.65 | 41.44 | 71.72 |
| **YOLOv3** | 73.77 | 76.21 | 91.41 | 69.33 | 92.76 | 82.06 | 79.15 | 54.15 | 39.11 | 82.19 |
| **SSD (vgg)** | 75.71 | 51.64 | 92.12 | 76.56 | 91.37 | 73.26 | 84.34 | 53.92 | 56.62 | 77.49 |
| **YOLOX** | 76.27 | 106.88 | **93.77** | **79.29** | 90.70 | 82.27 | 87.01 | 59.61 | 43.68 | 73.86 |
| **Ours** | **81.80** | 80.59 | 92.92 | 78.24 | **94.81** | **85.59** | **91.03** | **65.49** | **61.65** | **84.64** |

In these tables, the categories are denoted by B1-B8, representing Bottle, Milk-Box, Ball, Plastic-Bag, Plastic-Garbage, Branch, Grass, and Leaf, respectively. The bold values indicate the highest detection accuracy for each type of floater under the various deep learning network.

The results from Tables I and II indicate that the CA attention module effectively reduces the missed detection rate and achieves the highest accuracy for detecting plastic bags. Additionally, the mAP for plastic garbage exceeds 90%, which is attributed to the comprehensive utilization of its features and its resilience to interference from complex aquatic environments. However, the overall model experiences a notable reduction in prediction speed. When compared to the other two enhancement strategies, the introduction of the ECA attention module detects a greater number of floating objects, resulting in a significant overall accuracy improvement. Nevertheless, it marginally decreases the prediction speed. With regard to various types of floating objects, the varifocal loss enhances the detection accuracy for all objects except for milk box, with bottles achieving the highest accuracy. Moreover, it slightly improves the prediction speed. The combined results of the three enhancement strategies demonstrate that the detection accuracy for bottles, balls, and plastic garbage exceeds 90%. This is primarily because these objects possess distinct features and exhibit high-contrast coloration compared to the aquatic environment. However, only the overall model achieves an mAP of over 60% for grass, underscoring the effectiveness of the improvement algorithm. Nonetheless, the detection accuracy for grass remains significantly lower than that of other floating object categories due to the complexity of its features and the challenges associated with feature extraction.

According to Fig. 5(a) and (b), the improved network with the introduced CA attention mechanism did not exhibit missed detections for plastic bags, plastic garbage, and leaves, but only missed detecting balls, while the detection accuracy for balls, bottles, branches, and grass all showed significant improvements. This observation suggests that although the CA mechanism improves the ability of detecting various floating objects within complex water environments, it still exhibits limitations in accurately detecting multiscale floating objects. In Fig. 5(a) and (c), introducing the ECA mechanism yielded a substantial enhancement in the detection accuracy of balls, branches, and grass. However, plastic garbage was not successfully detected. This indicates that while ECA can enhance the feature extraction ability for multiscale floating objects, it cannot achieve precise detection in densely distributed background environments. Based on the analysis of Table I and Fig. 5(a) and (d), the introduction of the varifocal loss function in the YOLOX network resulted in a comparatively smaller improvement in overall detection accuracy compared to the first two improvement strategies mentioned. However, significant enhancements were observed, specifically in the accuracy of detecting balls, bottles, plastic bags, and grass. This indicates that the varifocal loss can balance positive and negative samples within the dataset.

Based on comparative experiments detailed in Tables III and IV, EC-YOLOX achieves the highest overall detection accuracy. Furthermore, it exhibits a certain level of improvement in the detection accuracy of floating objects other than bottles and milk boxes. Particularly, for grass and leaf categories, EC-YOLOX demonstrates the most significant enhancement in detection performance. These findings indicate that in aquatic environments with high contrast, EC-YOLOX can fully leverage its advantages, especially enhancing the detection of brightly colored floating objects. However, for small-sized floating objects like bottles and milk boxes, EC-YOLOX may not fully extract the corresponding features, thereby impacting detection accuracy.

Additionally, according to Fig. 6, YOLOv3, YOLOv5, and YOLOX all exhibited missed detections of balls and plastic garbage, while the improved algorithm did not exhibit any missed detections. However, the detection accuracy for balls, plastic bags, and milk boxes did not surpass 60%. This suggests that although EC-YOLOX can detect various floating objects in complex water environments, the lack of distinct feature representations for different categories of floating objects in intricate and densely populated aquatic environments hinders the model's ability to extract complete corresponding features, thus affecting its detection performance. Moreover, other algorithms also exhibited missed detections of branches and plastic bags, indicating that compared to these algorithms, the detection of less distinctive floating objects, such as branches and plastic bags, in complex water environments is still challenging. The detection accuracy of leaves, grass, branches, and plastic garbage showed significant improvements, indicating the effectiveness of EC-YOLOX in detecting floating objects of different scales

(Fig. 6). In Fig. 7(b), we observed a significant improvement in the detection accuracy of leaves and grass with EC-YOLOX, as the feature information of these objects was fully utilized to enhance the detector's accuracy. However, according to Fig. 7(a) and (b), we found that the mAP of grass and branches was much lower than that of other targets, possibly due to occlusion by other floating objects or missed detections due to their similar color to the water environment. Additionally, linear discriminant analysis can cluster objects [35] and accurately describe the internal structure of samples, thereby improving accuracy, while marginal distribution can determine categories based on test instances [36], allowing for improved instance detection performance under multiclass supervision and achieving better generalization performance.

Additionally, it is worth noting that the FPS of the EC-YOLOX algorithm is measured to be 80.59, which is slightly lower compared to the baseline network. This reduction in FPS can be attributed to the introduction of the CA module, which increases the computation of EC-YOLOX and consequently slows down the speed of the detector. Therefore, in future article, it may be beneficial to consider lightweighting the model to further optimize its performance. First, from a hardware perspective, using servers with larger memory can improve the model's inference speed. Second, reducing the model's parameter count is an achievable optimization method, such as introducing depthwise separable convolutional networks to maintain efficiency and deeper feature information, enabling precise detection of floating objects.

## V. Conclusion

To further enhance the ability of detecting floating objects in complex and densely populated water environments, we put forward EC-YOLOX for detecting floating object based on the YOLOX framework. This method incorporates three key components: the CA mechanism, the ECA mechanism, and a specialized loss function. These components collectively address the challenges posed by the varying scales of floating objects and the complexities of water environments, thereby improving the detection accuracy and mitigating missed detections. The CA module is effective in mitigating interference from complex aquatic environments and enhancing the ability to extract features from various types of floating objects, with a 3.31% improvement in mAP. The ECA attention module significantly improves the representational capacity for multiscale floating objects, achieving an mAP gain of 3.42%. The varifocal loss function, through detection ranking, reduces the impact of negative samples on detection results, resulting in a 2.08% improvement in mAP. Experimental results on the river floating objects dataset demonstrate that EC-YOLOX exhibits superior detection performance. According to the ablation experiments, all three improvement strategies have been shown to enhance the accuracy of floating object detection, resulting in an overall increase in mAP of 5.53%. Comparative experiments indicate that EC-YOLOX achieves an mAP of 81.80% and significantly outperforms other mainstream algorithms in terms of the mAP for detecting grass. Meanwhile, the detection accuracy of balls and bottles has reached 95% and 93%, respectively. This demonstrates the effectiveness and feasibility of EC-YOLOX. Despite the notable improvements in detecting grass and branches achieved by the EC-YOLOX algorithm, its detection accuracy still significantly lags behind that of other categories. Hence, future article endeavors should prioritize the enhancement of the detection capability for grass and branches, facilitating the adoption of improved algorithms in the detection of floating objects within river and ocean environments. This focus on improving grass and branch detection will provide robust support for the protection of ecological environments.

## References

[1] T. Huang and B. Chai, "Research progress of water pollution and eutrophication control technology in water source reservoirs," *Adv. Earth Sci.*, vol. 24, no. 6, 2009, Art. no. 588.

[2] F. Cui, "Urban water supply plant to deal with sudden water quality pollution technical measures to consider," *Water Resour. Manage.*, vol. 32, no. 7, pp. 7–9, 2006.

[3] Y. Cai, X. Tang, and W. Jiang, "Impact of river floating materials on engineering and current status of research," *J. Yangtze River Sci. Res. Inst.*, vol. 30, no. 8, 2013, Art. no. 84.

[4] Y. Cai, Z. Li, L. Li, and S. Liu, "Exploration of comprehensive management measures of river type reservoir floating materials," *J. Yangtze River Sci. Res. Inst.*, vol. 27, no. 12, 2010, Art. no. 31.

[5] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, May 2010.

[6] J. Platt, "Sequential minimal optimization: A fast algorithm for training support vector machines," *Microsoft Res. Tech. Rep.*, vol. 10, 1998, Art. no. 4376.

[7] C. Mu, Z. Yuan, J. Song, and Y. Chen, "A new approach to track moving target with improved mean shift algorithm and Kalman filter," in *Proc. 4th Int. Conf. Intell. Human-Mach. Syst. Cybern.*, 2012, pp. 359–362.

[8] A. Pepe, "A 3D space-time non-local mean filter (NLMF) for land changes retrieval with synthetic aperture radar images," *Remote Sens.*, vol. 14, no. 23, 2022, Art. no. 5933.

[9] H. Rahman and G. Paul, "Tripartite sub-image histogram equalization for slightly low contrast gray-tone image enhancement," *Pattern Recognit.*, vol. 134, 2023, Art. no. 109043.

[10] T. Kataoka and Y. Nihei, "Quantification of floating riverine macro-debris transport using an image processing approach," *Sci. Rep.*, vol. 10, no. 1, 2023, Art. no. 2198.

[11] B. Ni, Z. Liu, X. Cai, M. Nappi, and S. Wan, "Segmentation of ultrasound image sequences by combing a novel deep Siamese network with a deformable contour model," *Neural Comput. Appl.*, vol. 35, no. 20, pp. 14535–14549, 2023.

[12] H. Wang, D. Zhang, S. Ding, Z. Gao, J. Feng, and S. Wan, "Rib segmentation algorithm for X-ray image based on unpaired sample augmentation and multi-scale network," *Neural Comput. Appl.*, vol. 3, pp. 1–15, 2021.

[13] Y. Wu, H. Guo, C. Chakraborty, M. Khosravi, S. Berretti, and S. Wan, "Edge computing driven low-light image dynamic enhancement for object detection," *IEEE Trans. Netw. Sci. Eng.*, vol. 10, no. 5, pp. 3086–3098, Sep./Oct. 2023.

[14] J. Gou, L. Sun, B. Yu, S. Wan, and D. Tao, "Hierarchical multi-attention transfer for knowledge distillation," *ACM Trans. Multimedia Comput., Commun. Appl.*, vol. 51, pp. 1–20, 2022.

[15] Y. Zhang, F. Zhang, Y. Jin, Y. Cen, V. Voronin, and S. Wan, "Local correlation ensemble with GCN based on attention features for cross-domain person Re-ID," *ACM Trans. Multimedia Comput., Commun. Appl.*, vol. 19, no. 2, pp. 1–22, 2023.

[16] X. Gao et al., "Joint metric learning-based class-specific representation for image set classification," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published, doi: 10.1109/TNNLS.2022.3212703.

[17] Y. Cheng et al., "A multi-feature fusion and attention network for multi-scale object detection in remote sensing images," *Remote Sens.*, vol. 15, no. 8, 2023, Art. no. 2096.

[18] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 9, pp. 386–397, Sep. 2017.

[19] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2015.
[20] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 779–788.
[21] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 6517–6525.
[22] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.
[23] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," 2020, *arXiv:2004.10934*.
[24] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, "YOLOX: Exceeding YOLO series in 2021," 2021, *arXiv:2107.08430*.
[25] W. Liu, Y. Wang, S. Jiang, T. Ma, and W. Xiao, "Research on Mask R-CNN based floating object recognition method on water surface," *Yangtze River*, vol. 52, no. 11, pp. 226–233, 2021.
[26] Z. Yi, D. Yao, G. Li, J. Ai, and W. Xie, "Detection and localization for lake floating objects based on CA-faster R-CNN," *Multimedia Tools Appl.*, vol. 81, no. 12, pp. 17263–17281, 2022.
[27] F. Lin, T. Hou, Q. Jin, and A. You, "Improved YOLO based detection algorithm for floating debris in waterway," *Entropy*, vol. 23, no. 9, 2021, Art. no. 1111.
[28] G. Li, D. Yao, J. Ai, Z. Yi, S. Lei, and W. Wang, "Floating objects detection based on improved YOLOv3," *J. Guangxi Univ.*, vol. 46, pp. 1569–1578, 2021.
[29] R. Chen, J. Wu, Y. Peng, Z. Li, and H. Shang, "Detection and tracking of floating objects based on spatial-temporal information fusion," *Expert Syst. Appl.*, vol. 225, 2023, Art. no. 120185.
[30] S. Lei, J. Ai, J. Peng, and D. Yao, "Evaluation of deep neural networks for floating objects detection," *Environ. Develop.*, vol. 31, no. 6, pp. 117–120, 2019.
[31] J. Wang, "Research on detection of floating objects on water surface based on YOLOv3," M.S. thesis, Dept. Comput. Tech., Xinjiang Normal Univ., Xinjiang, China, 2022.
[32] Q. Hou, D. Zhou, and J. Feng, "Coordinate attention for efficient mobile network design," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 13708–13717.
[33] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 11534–11542.
[34] H. Zhang, Y. Wang, F. Dayoub, and N. Sunderhauf, "VarifocalNet: An IoU-aware dense object detector," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 8514–8523.
[35] F. Zhu, J. Gao, J. Yang, and N. Ye, "Neighborhood linear discriminant analysis," *Pattern Recognit.*, vol. 123, 2022, Art. no. 108422.
[36] F. Zhu, W. Zhang, X. Chen, X. Gao, and N. Ye, "Large margin distribution multi-class supervised novelty detection," *Expert Syst. Appl.*, vol. 224, 2023, Art. no. 19937.
[37] V. Kyrki and D. Kragic, "Computer and robot vision," *IEEE Robot. Automat. Mag.*, vol. 18, no. 2, pp. 121–122, Feb. 2011.
[38] S. Armitage, K. Awty-Carroll, D. Clewley, and V. Martinez-Vicente, "Detection and classification of floating plastic litter using a vessel-mounted video camera and deep learning," *Remote Sens.*, vol. 14, no. 14, 2022, Art. no. 3425.
[39] M. Skerjanec, K. Kregar, G. Stebe, and G. Rak, "Analysis of floating objects based on non-intrusive measuring methods and machine learning," *Geomorphology*, vol. 408, 2022, Art. no. 108254.
[40] B. Basu, S. Sannigrahi, A. S. Basu, and F. Pilla, "Development of novel classification algorithms for detection of floating plastic debris in coastal waterbodies using multispectral sentinel-2 remote sensing imagery," *Remote Sens.*, vol. 13, no. 8, 2021, Art. no. 1598.
[41] G. Qiao, M. Yang, and H. Wang, "A detection approach for floating debris using ground images based on deep learning," *Remote Sens.*, vol. 14, no. 17, 2021, Art. no. 4161.
[42] J. Zhang, Q. Yan, X. Zhu, and K. Yu, "Smart industrial IoT empowered crowd sensing for safety monitoring in coal mine," *Digit. Commun. Netw.*, vol. 9, no. 2, pp. 296–305, 2023.
[43] Y. Cheng et al., "Multi-scale feature fusion and transformer network for urban green space segmentation from high-resolution remote sensing images," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 124, 2023, Art. no. 103514.

**Jiaxin He** is currently working toward the master's degree in control science and engineering with the Nanjing University of Information Science and Technology, Nanjing, China.

His research interests include deep learning and remote sensing image processing, mainly focusing on the study of urban building extraction from high-resolution remote sensing images based on deep learning.

Mr. He has participated in the National Key Research and Development Program, the National Natural Science Foundation of China Youth Fund Project,and other topics.

**Yong Cheng** received the Ph.D. degree in computer science and technology from the Wuhan University, Wuhan, China, in 2009.

He is currently a Full Professor with the School of Software, Nanjing University of Information Science and Technology, Nanjing, China. He has authored/coauthored some articles in international journals, including IEEE INTERNET OF THINGS JOURNAL, etc. His research interests include big data, Internet of Things, and deep learning.

**Wei Wang** is currently working toward the master's degree in electronic information with the Nanjing University of Information Science and Technology, Nanjing, China.

His research interests include deep learning and remote sensing image processing, and he mainly focuses on the study of urban green space classification of high-resolution remote sensing images based on deep learning, to further realize the extraction and mapping of urban green space.

Mr. Wang has participated in the National Key Research and Development Program, the National Natural Science Foundation of China Youth Fund Project, and other topics. He has participated in several academic conferences and made oral presentations.

**Yakang Gu** is currently working toward the master's degree in software engineering with the Nanjing University of Information Science and Technology, Nanjing, China.

His research interests include deep-learning-based downscaling studies of global climate models to further obtain high-resolution future projected temperatures under different scenarios.

**Yixuan Wang** is currently working toward the master's degree in software engineering with the Nanjing University of Information Science and Technology, Nanjing, China.

She has coauthored one SCI paper. Her research interests include deep learning and streetscape image processing, mainly focusing on the research of urban green landscape segmentation algorithms for streetscape images based on deep learning, and further realizing urban green landscape extraction and mapping.

**Wenjie Zhang** received the M.Sc. and Ph.D. degrees from the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China in 2014 and 2019 respectively.

He is currently an Associate Professor and Master Supervisor with the School of Geosciences, Nanjing University of Information Science and Technology, Nanjing, China. His research interests include the influence of urban vegetation change pattern on the vertical structure of urban heat island, the influence of microtopography and micrometeorology on the ice cover of power grid transmission lines, and feature recognition and classification based on deep learning and geographic knowledge.

**Achyut Shankar** received the bachelor's degree in computer science and engineering from the Dr. M.G.R. University, Chennai, India, the master's degree in computer science and engineering from the SRM University, Chennai, India, and the Ph.D. degree in computer science and engineering with a major in wireless sensor network from the VIT University, Vellore, India.

He has been working as an Assistant Professor with the Amity University Noida, Noida, India, since March 2018. He has authored/coauthored more than 20 research articles in reputed international conferences and journals. His research interests include wireless sensor networks, machine learning, the Internet of Things, blockchain, and cloud computing.

Dr. Shankar is a member of ACM. He was the recipient of the research award for excellence in research for the year 2016 and 2017. He had organized many special sessions with Scopus Indexed International Conferences worldwide, proceedings of which were published by Springer, IEEE, Elsevier, and so on. He is currently serving as a Reviewer for IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, IEEE SENSORS JOURNAL, and other prestigious conferences.

**Shitharth Selvarajan** received the Ph.D. degree from the Department of Computers Science and Engineering, Anna University, Chennai, India.

He completed his Postdoc at The University of Essex, Colchester, UK. He has worked in various institutions with a teaching experience of seven years. He is currently a Lecturer in cyber security with Leeds Beckett University, Leeds, U.K. He has authored or coauthored more than 100 international journals and 20 international and national conferences, and invented four patents in IPR. His research interests include cyber security, blockchain, critical infrastructure and systems, network security, and ethical hacking.

Dr. Selvarajan is also an active Member of IEEE Computer Society and five more professional bodies. He is also a Member of the International Blockchain organization. He is a certified Hyperledger Expert and certified Blockchain Developer. He is an active Researcher, Reviewer and Editor for many international journals.

**Sathish A. P. Kumar** (Senior Member, IEEE) received the Ph.D. degree in computer science and engineering from the University of Louisville, Louisville, KY, USA, in 2007.

He is currently an Associate Professor of computer science with the Department of Electrical Engineering and Computer Science, Cleveland State University, Cleveland, OH, USA. He is also directing the ISCAR Laboratory, Cleveland State University. He has authored/coauthored more than 75 technical papers in international journals and conference proceedings. His research interests include cybersecurity, machine learning, quantum computing, and secured distributed systems and their applications.

Dr. Kumar serves as an Associate Editor for IEEE ACCESS, *PLOS One*, and *Machine Learning With Applications* (Elsevier), and as an Editorial Board Member for Scientific Reports (Nature).