International Journal of Multidisciplinary Research and Publications
ISSN (Online): 2581-6187

Research and Design of a Multi-Target Garbage Classification Model Based on Improved YOLOv5 Algorithm in Complex Environments

Wuzheng Xu[1], Zhang Qing[2], Haoyu Hu[3], Wu Ling[4*]
[1]Anhui University of Finance and Economics
[2]Zhejiang Gongshang University Hangzhou College of Commerce
[3]Anhui University of Finance and Economics
[4]Anhui University of Finance and Economics
No. 962 Caoshan Road, Bengbu City, Anhui Province, China 233030
*Corresponding author: Wu Ling

***Abstract*—*With the acceleration of urbanization and the enhancement of environmental awareness, efficient and accurate garbage classification has become an urgent need for smart city construction. Deep learning-based object detection techniques offer a feasible solution for automated garbage sorting. However, in real-world scenarios, complex environmental factors such as varying lighting conditions, object occlusions, small-scale targets, and cluttered backgrounds significantly degrade the recognition accuracy and robustness of existing models. To address these challenges, this paper proposes an improved YOLOv5-based multi-target garbage classification model designed to enhance detection performance in complex environments. Building upon the YOLOv5s baseline model, we implement three key improvements: Firstly, the Convolutional Block Attention Module (CBAM) is integrated into the backbone network to enable the model to focus on critical features of garbage objects, effectively suppressing background interference. Secondly, the original Path Aggregation Network (PANet) in the neck is replaced with a Bidirectional Feature Pyramid Network (BiFPN) with weighted feature fusion, which strengthens the model's capability in extracting features from garbage items of various scales. Lastly, the SIoU loss function, which considers angle cost, is adopted to replace the original CIoU loss, improving bounding box regression accuracy and convergence speed. Experimental results on a comprehensive garbage classification dataset demonstrate that the mean Average Precision (mAP@0.5) of the proposed model has been significantly improved, and it also outperforms other mainstream models in key metrics. Visualization results confirm that the model maintains high detection stability in challenging scenarios, providing an effective technical approach for real-time and accurate garbage classification in complex environments*.

***Keywords*—** *Garbage Classification; Object Detection; YOLOv5; Complex Environments; Attention Mechanism.*

I. INTRODUCTION

With the acceleration of global urbanization and continuous improvement in residents' consumption levels, the "garbage siege" phenomenon has become a severe challenge to the sustainable development of modern cities. Implementing efficient garbage classification serves as a fundamental approach to achieving waste reduction, resource recovery, and harmless treatment, holding crucial strategic significance for building "zero-waste cities," promoting circular economy development, and implementing national environmental policies [1]. However, traditional manual sorting methods not only suffer from low efficiency and high costs but also pose potential health risks to workers, making it difficult to meet the demands of large-scale, routine garbage classification.

In recent years, artificial intelligence technologies, particularly computer vision based on deep learning, have achieved breakthrough progress, providing novel technical solutions for automated and intelligent garbage classification [2]. Among these, the single-stage object detection algorithm YOLO has been widely adopted in real-time detection scenarios due to its excellent balance between speed and accuracy [3]. Compared to early traditional image recognition methods based on handcrafted features and image classification models based on deep learning, object detection algorithms can directly locate and identify multiple different categories of garbage in images, better aligning with practical application requirements.

Despite the outstanding performance of existing object detection algorithms on datasets under ideal conditions, their recognition accuracy and robustness often significantly decline when deployed in real-world complex scenarios such as communities, streets, or transfer stations. These "complex environments" are specifically characterized by: varying lighting conditions (significant appearance differences of the same object under daylight and nighttime, shadows and strong light interference), severe occlusion and overlap between targets (items stacked within garbage bags, garbage blocking each other in bins leading to incomplete target features), difficulties in small target detection (small-sized garbage like batteries, bottle caps, cigarette butts occupying small proportions in images with weak feature information, easily missed), complex background interference (ground textures, vegetation, and other irrelevant clutter with features similar to garbage targets causing false detections), and large scale variations among targets (large cartons and small packaging bags possibly coexisting in the same frame, challenging the model's multi-scale feature extraction capability).

Although researchers have conducted numerous studies applying deep learning to garbage classification, most current work exhibits certain limitations: idealized environmental

Wuzheng Xu, Zhang Qing, Haoyu Hu, and Wu Ling, "Research and Design of a Multi-Target Garbage Classification Model Based on Improved YOLOv5 Algorithm in Complex Environments," *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 8, Issue 4, pp. 153-158, 2025.

assumptions (most research is validated in relatively simple, well-lit experimental environments with clear targets, lacking consideration for complex, dynamic real-world environments), insufficient model specificity (studies mostly focus on directly applying general detection models, with relatively few works specifically optimizing models for unique challenges in garbage classification scenarios), and inadequate integration of advanced technologies (techniques proven effective in general object detection such as attention mechanisms, more efficient feature fusion architectures, and next-generation loss functions have not been systematically explored and integrated into garbage classification models for complex environments).

To address the aforementioned challenges, this paper takes YOLOv5s as the baseline model and deeply researches and designs an improved multi-target garbage classification model suitable for complex environments. The main contributions of this paper can be summarized as follows:

Proposed a feature enhancement method incorporating attention mechanism.

Designed a feature fusion network based on weighted bidirectional feature pyramid.

Introduced a bounding box regression loss function considering directional matching degree.

Constructed a garbage classification dataset covering multiple complex scenarios and conducted comprehensive experimental validation.

II. Related Work

2.1 Research Progress in Object Detection Algorithms

Object detection is one of the core tasks in computer vision, aiming to locate specific target positions (typically represented by bounding boxes) and identify their categories in given images. In recent years, deep learning-based object detection methods have developed rapidly, primarily divided into two main schools: two-stage detectors and single-stage detectors.

Two-stage detectors, represented by the R-CNN series, can be summarized as a two-stage process: "candidate region generation" and "region classification and regression." R-CNN [8] proposed by Girshick et al. innovatively introduced convolutional neural networks into object detection, generating candidate regions through selective search and then performing feature extraction and classification for each region. However, it suffered from computational redundancy and slow speed. Subsequently, Fast R-CNN [9] significantly improved efficiency by sharing convolutional feature maps and performing feature extraction only once for the entire image. Faster R-CNN [10] further introduced the Region Proposal Network (RPN), achieving end-to-end training and making candidate box generation highly synergistic with the object detection task itself. This reached new heights in accuracy and established a classic framework for two-stage methods.

Single-stage detectors abandon the candidate region generation step, directly dividing the image into grids and predicting bounding boxes and category probabilities on grid cells, thus possessing inherent advantages in speed. YOLO [11] is the pioneering work in this field, treating object detection as a regression problem and achieving extremely fast detection speeds. Since then, the YOLO series algorithms have been continuously updated. For example, YOLOv3 [12] introduced multi-scale prediction and an improved backbone network, Darknet-53. YOLOv4 [13] built upon YOLOv3, integrating various "bag of freebies" techniques including Mosaic data augmentation, CmBN, and SAT self-adversarial training, significantly improving accuracy while maintaining speed. YOLOv5, as an efficient engineering implementation based on the PyTorch framework, quickly became a popular solution widely adopted in both industry and academia due to its simple code structure, flexible data augmentation strategies, and excellent performance. Meanwhile, other single-stage detectors like SSD [14] and RetinaNet [15] also made important contributions to the balance of accuracy and speed through technologies such as multi-scale feature map prediction and the focal loss function. Overall, single-stage detectors, with their higher computational efficiency, are more suitable for scenarios requiring real-time or quasi-real-time processing, such as garbage classification.

2.2 Research Status of Garbage Classification Based on Deep Learning

With the widespread adoption of deep learning technology, its application in the field of garbage classification has become increasingly in-depth. Early research mainly focused on image classification tasks, i.e., determining which category of garbage the entire image belongs to. For example, Mittal et al. [4] used pre-trained CNN models to classify garbage images, proving the effectiveness of deep features in this task. However, classification methods cannot handle situations where images contain multiple different categories of garbage.

To address this, researchers began to apply object detection algorithms to garbage classification. Yang et al. [5] used Faster R-CNN to detect recyclable garbage, achieving high recognition accuracy, but the inference speed of the two-stage model was slow. To meet real-time requirements, subsequent work increasingly adopted single-stage detectors. Liu et al. [6] applied the YOLOv3 model to domestic waste detection and verified its feasibility on a self-built dataset. Awe et al. [7] compared the performance of SSD and YOLOv4 in garbage classification, indicating that YOLOv4 has better comprehensive performance. In recent years, the latest models such as YOLOv5 and YOLOv7 have also been attempted for garbage recognition tasks, showing promising potential.

2.3 Research Status Review and Positioning of This Paper

Through a systematic review of existing research, it can be found that although deep learning-based object detection algorithms have achieved significant results in garbage classification tasks, most studies still have certain limitations:

a. Idealized Environmental Assumptions: Most work is validated in relatively simple experimental environments with good lighting and clear targets, lacking sufficient consideration for complex and dynamic real-world environments.

Wuzheng Xu, Zhang Qing, Haoyu Hu, and Wu Ling, “Research and Design of a Multi-Target Garbage Classification Model Based on Improved YOLOv5 Algorithm in Complex Environments,” *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 8, Issue 4, pp. 153-158, 2025.

b. Insufficient Model Specificity: Research mostly focuses on directly applying general detection models, with relatively few works specifically optimizing models for the unique challenges in garbage classification scenarios (such as large target scale differences, appearance deformation, dense small targets, etc.).
c. Inadequate Integration of Advanced Technologies: Techniques proven effective in general object detection, such as attention mechanisms, more efficient feature fusion architectures (e.g., BiFPN), and next-generation loss functions (e.g., SIoU), have not been systematically explored and integrated into garbage classification models designed for complex environments.

Therefore, the research positioning of this paper is to directly address the challenges of garbage classification in real complex environments. Based on the high-performance and high-efficiency YOLOv5 model, through a series of targeted innovations such as introducing attention mechanisms, optimizing feature fusion networks, and improving loss functions, we aim to build a robust and high-precision multi-target garbage classification model to compensate for the deficiencies of existing research in dealing with complex environments.

III. Improved YOLOv5 Garbage Classification Model

3.1 YOLOv5 Algorithm Foundation

YOLOv5 is a series of efficient single-stage object detection models, divided into four versions (s, m, l, x) based on their depth and width. This paper adopts the lighter and faster YOLOv5s as the baseline model for improvement, aiming to enhance accuracy while ensuring real-time performance. Its network structure primarily consists of four parts:

Input: The input end utilizes technologies such as Mosaic data augmentation, adaptive anchor box calculation, and adaptive image scaling, which enrich the dataset and improve model training efficiency and generalization ability.

Backbone Network: The backbone network employs CSPDarknet53. Through its Cross-Stage Partial network structure, it enhances feature extraction capability while reducing computational load and alleviates the vanishing gradient problem.

Neck Network: The neck structure uses the Path Aggregation Network (PANet). It integrates semantic information from deep layers with localization information from shallow layers through a bottom-up and top-down feature pyramid structure, enabling the detection of targets at different scales.

Detection Head: The detection head uses a decoupled structure to predict the bounding box, objectness score, and classification separately. It ultimately outputs feature maps at three different scales for detecting large, medium, and small targets.

Although YOLOv5s is a powerful baseline model, its native design still has room for improvement in terms of feature focus, multi-scale fusion efficiency, and localization accuracy when facing garbage classification tasks in complex environments.

3.2 Overall Design of the Improved Model

To address the aforementioned challenges, this paper introduces three targeted improvements to the baseline model:

Embed the CBAM attention module at the end of the backbone network to enhance the model's ability to extract key features of garbage and suppress complex backgrounds.

Replace the PANet in the neck with the BiFPN structure, achieving more efficient multi-scale feature fusion through weighted bidirectional cross-scale connections.

Replace the bounding box regression loss function in the detection head from CIoU Loss to SIoU Loss, introducing directional consideration to improve localization accuracy and convergence speed.

3.3 Integration of the CBAM Attention Mechanism

3.3.1 Motivation

In complex environments, background information (e.g., ground texture, vegetation) often interferes with the garbage targets themselves, causing the model to produce false detections or missed detections. Convolutional Neural Networks process features from all regions equally, lacking the ability to focus on key information. The attention mechanism simulates the human visual system, guiding the model to allocate limited computational resources to more informative regions.

3.3.2 Principle

We selected the Convolutional Block Attention Module (CBAM), a lightweight and general-purpose module that can be seamlessly integrated into CNN architectures. CBAM sequentially infers attention maps along two independent dimensions, channel and space, and then multiplies these maps with the input feature map for adaptive feature optimization.

Channel Attention Module: This module focuses on "what features are meaningful". It first performs global average pooling and global max pooling on the input feature map respectively. These two outputs are then fed into a shared Multi-Layer Perceptron (MLP). After element-wise addition of the feature vectors output by the MLP, the channel attention weight Mc is generated via the Sigmoid activation function. This process enables the model to learn the importance of different feature channels.

Spatial Attention Module: This module focuses on "where the informative part is in the feature map". It first performs average pooling and max pooling along the channel dimension on the channel-attention-weighted feature map and concatenates the two results into a single feature map. Then, a standard convolutional layer is applied, followed by the Sigmoid activation function to generate the spatial attention weight Ms. This process enables the model to focus on the key spatial locations of the target.

Finally, the input feature map F is multiplied sequentially by Mc and Ms to obtain the refined feature map F'.

3.3.3 Integration Implementation

This paper embeds the CBAM module after the last C3 module in the CSPDarknet53 backbone network. The feature map at this location contains rich deep semantic information while retaining certain spatial detail information. Introducing CBAM here allows the model to pre-strengthen channel and

Wuzheng Xu, Zhang Qing, Haoyu Hu, and Wu Ling, "Research and Design of a Multi-Target Garbage Classification Model Based on Improved YOLOv5 Algorithm in Complex Environments," *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 8, Issue 4, pp. 153-158, 2025.

spatial features related to garbage targets and suppress irrelevant background noise before entering the feature fusion stage, providing higher-quality feature representations for the subsequent neck and detection head.

3.4 Feature Fusion Network Optimization Based on BiFPN

3.4.1 Motivation

The native PANet used in YOLOv5s aggregates features at different scales through bidirectional paths, but its feature fusion process is "indiscriminate", meaning it treats different input features equally. However, in garbage classification tasks, input feature maps at different scales contribute differently to the final output. Treating all input features equally may lead to insufficient information fusion, especially when dealing with garbage targets with significant scale variations.

3.4.2 Principle

This paper uses the Bidirectional Feature Pyramid Network (BiFPN) with weighted feature fusion to replace the original PANet. BiFPN optimizes feature fusion through the following two core concepts

Efficient Cross-Scale Connections: BiFPN removes nodes in PANet that have only one input edge and adds an extra edge between input and output nodes at the same level, thereby achieving more feature fusion without significantly increasing computational cost.

Weighted Feature Fusion: Traditional feature fusion mostly uses direct addition or concatenation, whereas BiFPN introduces a learnable weight parameter for each input feature (which can be a scalar, vector per channel, or pixel-wise weight). This allows the network to learn the importance of features at different resolutions during training. This paper adopts the fast normalized fusion strategy.

$$O = \sum_i \frac{\omega_i}{\epsilon + \sum_j \omega_j} \cdot I_i$$

This method allows the model to adaptively adjust the contribution of different feature maps, achieving more flexible and efficient feature fusion

3.4.3 Integration Implementation

We completely replace the original PANet structure in YOLOv5s with the BiFPN structure. BiFPN takes three feature maps (P3, P4, P5) of different scales from the CSPDarknet53 backbone network as input. Through its unique weighted bidirectional fusion mechanism, it generates output feature maps of the same three scales but with higher fusion quality, which are then sent to the detection head for prediction. This improvement significantly enhances the model's detection capability for garbage targets of varying scales, from large cartons to small batteries.

3.5 SIoU Loss Function

3.5.1 Motivation

The loss function for object detection typically consists of classification loss and bounding box regression loss. The native bounding box regression loss in YOLOv5s is CIoU Loss, which considers the overlap area, center point distance, and aspect ratio. However, CIoU does not consider the direction between bounding boxes, which may cause the prediction box to "wander" during training and slow down the convergence speed.

3.5.2 Principle

This paper uses SIoU Loss to replace CIoU Loss. SIoU redefines the penalty metric by introducing a directionality term. It consists of the following four cost functions:

Angle Cost: First, calculate the angle α between the line connecting the centers of the ground truth box and the prediction box and the coordinate axis. The model will prioritize minimizing this angle, encouraging the prediction box to move towards the ground truth box along the X or Y axis. This defines a new distance calculation method.

Distance Cost: Calculate the distance penalty between the centers of the prediction box and the ground truth box based on the angle cost.

Shape Cost: Calculate the difference in width and height between the prediction box and the ground truth box

IoU Cost: This is the traditional Intersection over Union

The total cost for SIoU Loss is

$$\mathcal{L}_{SIoU} = 1 - IoU + \frac{\Delta + \Omega}{2}$$

By introducing the angle cost, SIoU provides a better optimization direction for bounding box regression, thereby accelerating model convergence and achieving higher localization accuracy.

3.5.3 Integration Implementation

During the model training phase, we directly replace the loss function used for calculating bounding box regression in the YOLOv5s detection head from CIoU Loss to SIoU Loss. This change does not require adjusting the network structure but only modifying the loss calculation part of the code. It can effectively improve the localization accuracy of the bounding boxes, especially when dealing with partially occluded or densely arranged garbage targets.

IV. Experiments and Results Analysis

4.1 Experimental Setup

4.1.1 Dataset and Data Augmentation

This study constructed a garbage classification dataset comprising 6,845 high-resolution images, covering 12 sub-categories under four major classes: recyclables, kitchen waste, hazardous waste, and other garbage. The dataset specifically considers real-world complexities, with approximately 40% of the images containing one or multiple complex scenarios, including: uneven lighting (overexposure/underexposure), target occlusion, small target aggregation (e.g., multiple bottle caps, cigarette butts), and cluttered backgrounds (e.g., garbage scattered on grass or streets). The dataset was randomly divided into training, validation, and test sets in an 8:1:1 ratio.

To enhance the model's generalization capability and robustness, this study adopted the following data preprocessing and augmentation strategies:

Input images were uniformly resized to 640×640 pixels

Mosaic augmentation was applied, randomly stitching four images to simulate small target aggregation and complex backgrounds

Color space perturbations were implemented, randomly

Wuzheng Xu, Zhang Qing, Haoyu Hu, and Wu Ling, "Research and Design of a Multi-Target Garbage Classification Model Based on Improved YOLOv5 Algorithm in Complex Environments," *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 8, Issue 4, pp. 153-158, 2025.

adjusting image brightness, contrast, saturation, and hue

Geometric transformations were performed, including random rotation, scaling, cropping, and horizontal flipping

Random erasing was used to simulate occlusion, enhancing the model's ability to recognize partially occluded targets

4.1.2 Evaluation Metrics

To quantitatively assess model performance, this study employed evaluation metrics commonly used in object detection:

Precision (P): P = TP/(TP+FP)

Recall (R): R = TP/(TP+FN)

Average Precsion (AP): Calculated for a single category

mean Average Precision (mAP@0.5): The average of AP across all categories at an IoU threshold of 0.5, serving as the core metric for overall model performance

mAP@0.5:0.95: The average mAP calculated over IoU thresholds ranging from 0.5 to 0.95 (step size 0.05), representing a more stringent comprehensive metric.

V. Conclusion and Future Work

5.1 Conclusion

Addressing the challenges of multi-target garbage classification in complex environments, this paper conducted an in-depth study based on YOLOv5s and proposed an improved solution. Through systematic analysis of the baseline model, we identified its shortcomings in feature focus, multi-scale fusion, and localization accuracy, and accordingly developed a comprehensive improvement strategy.

The main contributions of this paper include proposing a YOLOv5 model integrated with the CBAM attention mechanism. By embedding the CBAM module at the end of the backbone network, the model can adaptively calibrate feature responses in both channel and spatial dimensions, effectively enhancing its ability to extract key garbage features while significantly suppressing interference from complex backgrounds. The introduction of the SIoU loss function optimizes bounding box regression. By considering the directionality between bounding boxes, SIoU Loss provides better gradient directions for model convergence, thereby improving localization accuracy while accelerating training speed.

In summary, the improved YOLOv5 multi-target garbage classification model researched and designed in this paper effectively overcomes the challenges posed by complex environments. It provides a high-precision, high-efficiency, and high-robustness technical solution for the practical deployment of intelligent garbage classification systems, possessing certain theoretical significance and practical application value.

5.2 Future Work

Although the model proposed in this paper has achieved good performance, limited by research time and conditions, some aspects warrant further investigation and improvement in future work:

Model Lightweighting and Efficiency Optimization: This study primarily focused on improving detection accuracy, resulting in increased model parameters and computational complexity compared to the baseline. Future work will explore lightweight technologies such as model pruning, knowledge distillation, or neural architecture search to significantly reduce model size and computational costs while maintaining accuracy as much as possible, facilitating better deployment on edge devices with limited computing resources (e.g., embedded systems, mobile terminals).

Integration of Emerging Technologies: In recent years, the Transformer architecture has demonstrated strong performance in computer vision. Future work could consider introducing Transformer modules into the model or exploring pure Transformer-based detectors to capture richer global contextual information, potentially further enhancing model performance.

Expansion and Deepening of Application Scenarios: Future plans include applying this model to more specific practical scenarios, such as integrating it into the vision system of garbage sorting robots or developing real-time monitoring applications for community smart garbage bins. Additionally, exploring garbage detection in video streams rather than static images could enable real-time analysis and recording of garbage disposal behaviors.

Dataset Expansion and Fine-Grained Annotation: Although the current dataset already includes multiple complex scenes, the data scale could be further expanded in the future to cover more diverse and extreme harsh conditions (e.g., rain/snow weather, severe contamination). Furthermore, introducing instance segmentation annotations and upgrading the detection task to a segmentation task could achieve finer-grained contour recognition of garbage targets, meeting the demands of higher-precision sorting.

Through continuous exploration in the above directions, we anticipate advancing garbage classification technology in complex environments to new heights, contributing greater strength to the realization of a fully automated and intelligent garbage treatment system.

Acknowledgment

This study was supported by the Innovation and Entrepreneurship Program of Anhui University of Finance and Economics(Project No. 202410378067).

References

[1] Wang Wei, Li Jing. Research on the model of urban do mestic waste classification collection and treatment[J]. Environmental Science and Management, 2021, 46(3): 45-49.

[2] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.

[3] Zhao Z Q, Zheng P, Xu S T, et al. Object detection with deep learning: A review[J]. IEEE transactions on neural networks and learning systems, 2019, 30(11): 3212-3232.

[4] Mittal G, Yagnik K B, Garg M, et al. SpotGarbage: smartphone app to detect garbage using deep learn ing[C]//Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing. 2016: 940-945. 9

[5] G Thung,Mingxiang Yang.Classification of trash for recyclability status[J]. CS229 project report, 2016, 2016: 3.

[6] Liu J, Wang X, Wang C, et al. Garbage detection and classification using a deep learning algorithm[C]//2019 14th International Conference on Computer Science & Education (ICCSE). IEEE, 2019: 800-804.

Wuzheng Xu, Zhang Qing, Haoyu Hu, and Wu Ling, "Research and Design of a Multi-Target Garbage Classification Model Based on Improved YOLOv5 Algorithm in Complex Environments," *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 8, Issue 4, pp. 153-158, 2025.

[7] Awe O, Mengistu R, Sreedhar V. Smart trash net: Application of deep learning in smart waste manage ment[C]//2020 International Conference on Computa tional Science and Computational Intelligence (CSCI). IEEE, 2020: 282-287.
[8] Girshick R, Donahue J, Darrell T, et al. Rich feature hi erarchies for accurate object detection and semantic seg mentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 580-587. [9] Girshick R. Fast r-cnn[C]//Proceedings of the IEEE in ternational conference on computer vision. 2015: 1440 1448.
[10] Ren S, He K, Girshick R, et al. Faster r-cnn: To wards real-time object detection with region proposal networks[J]. Advances in neural information processing systems, 2015, 28.
[11] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detec tion[C]//Proceedings of the IEEE conference on com puter vision and pattern recognition. 2016: 779-788.
[12] Redmon J, Farhadi A. Yolov3: An incremental improve ment[J]. arXiv preprint arXiv:1804.02767, 2018.
[13] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: Op timal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.
[14] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Nether lands, October 11–14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016: 21-37.
[15] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE interna tional conference on computer vision. 2017: 2980-2988.
[16] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the Euro pean conference on computer vision (ECCV). 2018: 3 19.
[17] Tan M, Pang R, Le Q V. Efficientdet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 10781-10790.

Wuzheng Xu, Zhang Qing, Haoyu Hu, and Wu Ling, “Research and Design of a Multi-Target Garbage Classification Model Based on Improved YOLOv5 Algorithm in Complex Environments,” *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 8, Issue 4, pp. 153-158, 2025.