Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 8
AICT 2024

Computer Vision Applications in Garbage Management: A Survey

Yiyi Li[1, a]

[1]School of Information Science and Technology, Northwest University, Xi'an, China

[a]2021117462@stumail.nwu.edu.cn

Abstract. In recent years, computer vision (CV) technique has made great breakthroughs, and it has been applicated in the fields related to garbage management. Due to quick expansion of consumerism and urbanization, the amount of garbage generated by humans has increased rapidly, which has caused harmful results for both the environment and people's lives. Identifying garbage automatically by means of CV technique plays an important role in urban road garbage detection, marine litter detection and garbage classification. This paper, gathering research from various sources, show how computer vision techniques can help garbage management to become smarter and more accurate by giving a view of the literature on different CV applications in the fields related to garbage management. In addition, the paper also aims to find and analyze the common problem when practicing the theories in practice and several promising directions for future research.

Keywords: Computer Vision, Garbage Management, Deep Learning, Smart garbage bin, Garbage detection.

1. Introduction

Since the world's population increasement and economic development, the amount of garbage generated keeps growing. In general, different kind of garbage come from 4 various source: (i) industry solid waste, (ii) household waste, (iii) construction waste, (iv) agricultural solid waste. Without control, improper garbage would be harmful to both human health and environment: (i) air pollution, (ii) water pollution, (iii) soil pollution, (iv) threats to food safety, (v) epidemic disease,

(vi) bad impact on city appearance and management. In order to reduce detriments caused by improper garbage, both developed and developing countries are taking action on garbage management.

For the past decades, Deep Learning (DL) has evolved rapidly, which also promote the development of CV techniques. As the algorithms related to Convolutional Neutral Network (CNN), such as Deep Convolutional Neutral Network (DCNN), Dilated Convolutions and Prompt Learning (NLP) continue to refine, CV tasks, or derivation of meaningful information from digital images, videos, and other visual inputs, could achieve more specific and accurate results. CV techniques drive tasks related to garbage management such as waste detection and waste classification more intelligent. This survey, focusing on three fields related to garbage management: (i) urban road garbage detection, (ii) marine litter detection, (iii) garbage classification, introduces and assesses different developed algorithms in literature published in the last three years of CV that applicated in the fields listed above.

This survey is structured as follows. In Section 2, different CV techniques applicated in garbage management are introduced. In Section 3, CV-technique applications in the three fields related to garbage management: urban road garbage detection, marine litter detection and garbage classification are described. In Section 4, current achievements approached as well as deficiency of CV applications in garbage management are concluded. In Section 5, the difficulties of CV applications in garbage management are discussed. In Section 6, the prospects and future directions for future research are provided.

2. Computer Vision Studies in the Field of Garbage Management

2.1. Evolution of Computer Vision Studies

2.1.1 Geometric Model Solution

Early researchers used simple geometric models to understand objects in images. Different handcrafted geometric features and traditional image-processing methods are applied at this beginning stage of CV development [1]. This phase is represented by the work of Moravec at Bell Labs in 1966, who conducted experiments in visual navigation for robots. The main emphasized application fields of this stage are image processing.

2.1.2 Machine Learning

Machine learning algorithms such as Support Vector Machines (SVMs), decision tree and Adaboost were used in CV in 1990s. Handcrafted feature extraction and traditional machine learning algorithms, such as Support Vector Machine (SVM), K Nearest Neighbors (KNN) and so on, perform well on some tasks, but have limited effectiveness for complex visual tasks.

2.1.3 Deep Learning

Convolutional Neural Network (CNN) approach was used to implement a handwritten digit recognition system in 1986, which laid the foundation for the developed application in CV. Deep learning has dominated computer vision as computational power increases and the availability of large-scale datasets increases.

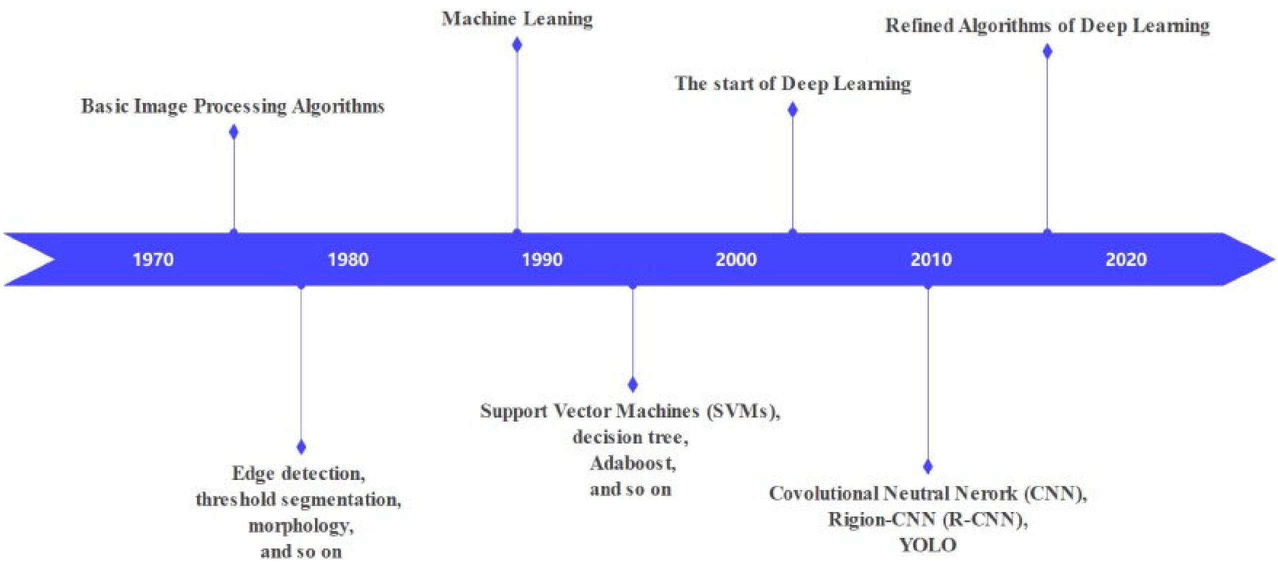

Figure 1. Evolution of CV Studies

2.2. Deep Learning Algorithms Applied in Garbage Management

2.2.1 Convolutional Neural Network (CNN)

CNN is a deep learning model that is primarily used to process data with a grid structure, such as images and videos. In the context of garbage management, CNN could be applied with various purposes: (i) image recognition, (ii) visual recognition, (iii) image characterization, (iv) anomaly detection. [2]

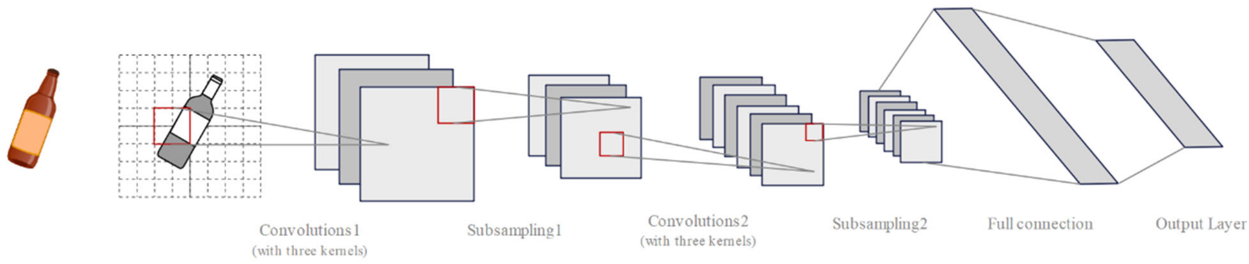

Figure 2. The Process of CNN

The input of image is the input layer, and it might consist of multiple channel matrix such as "RGB" channel. Mostly, the image will be altered to grey level image to reduce the calculation cost. Then, the input image will be sent into convolutional layer to calculate inner product with multiple kernels that indicate different features, and the result matrices are regarded as "Feature Maps". After that, in order to decrease the dimension of data and avoid overfitting, there would be a pooling layer after each convolution operation. After several operations of convolution and pooling, the feature maps will be input to fully connected layer and the output layer usually is the vector with elements indicate each kind of stuff that should be detected.

The most important operation of the CNN model is to calculate the inner product of kernel matrix with the relevant image's parts. With the slides of the kernel, the feature map related to that kernel can be achieved. In addition, in order to avoid the risk of edge feature disappearance, there might be a padding operation. After the convolution operation, the feature maps should be downsampled with the pooling operation.

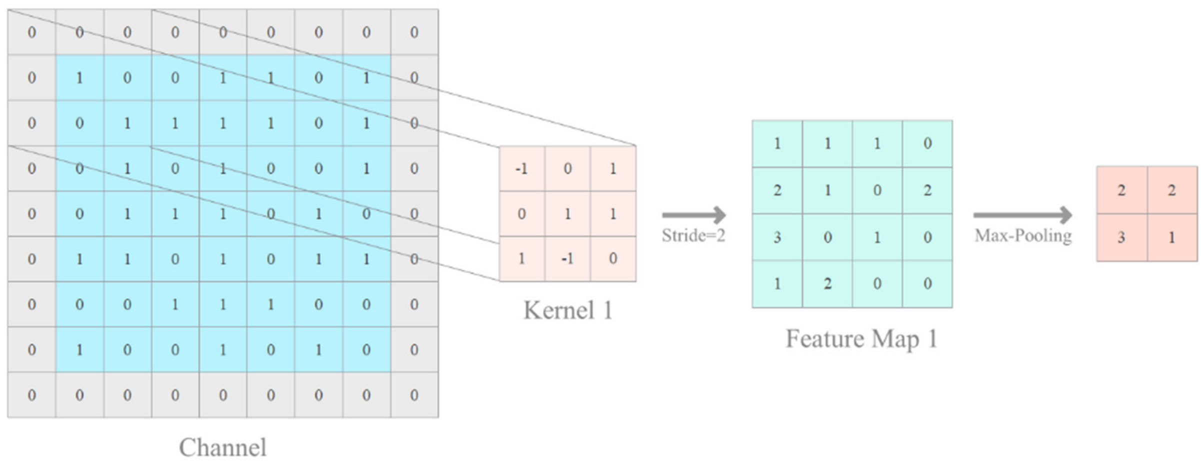

Figure 3. The Introduction of Convolution Layers

A majority of algorithms to deal with the problems that encountered when implementing the smart garbage management system are developed by CNN. The ResNet model can solve the gradient disappearance and explosion problem when the number of layers increases. Originally, the model is to learn H(x), but it became more and more difficult to learn H(x) with the deepening of the hierarchy, so it changed to learn F(x)=H(x)-x, and H(x) can be obtained through simple operations. With such operation, the channel to H(x) is not only decided by the output of the layers but also can be reached through the result achieved by the previous layer's result. The simple process schematic diagram of ResNet is shown in Figure 4.

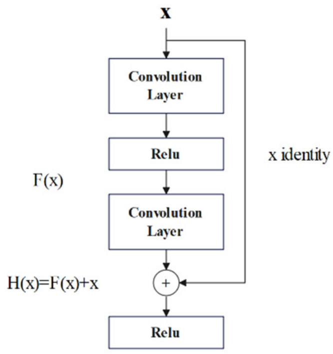

Figure 4. The ResNet

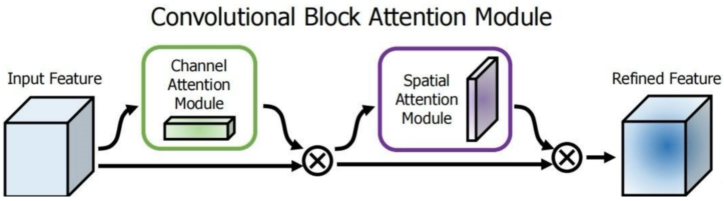

Figure 5. The CBAM Process Image in [3]

A novel method is proposed by combining two CNN technologies ResNeXt and ResNet50 to create an effective waste management solution in [2]. The modified ResNeXt mentioned in the paper is a model that combine the ResNet with Convolutional Block Attention Module (CBAM). With the channel attention module and spatial attention model, not only the computing power can be saved, but the module also can be integrated into the existing network architecture as a plug-and-play module.[3] The modified model can achieve accuracy to be 98.9%.

2.2.2 YOLO Algorithm

YOLO's basic algorithm is a neural network based on deep learning techniques. It is a real- time object detection algorithm which is capable of identifying and locating multiple objects in images or videos and assigning the objects with corresponding category labels.

Algorithm models such as Region-based CNN (R-CNN) and Faster R-CNN are two-stage method that consists of region proposal and classifier like SVM, while algorithm models such as SSD and YOLO are one stage method that apply the CNN network to extract features and train the network in full-connection layer, which means the goal of object detection is transformed to a regression problem. The core idea of YOLO is using the whole image as the input of the network and return the location of bounding box and the category of bounding box.

The input image would be firstly put into a set of convolution layers. Then the feature maps are fed into full connection layer with relevant weights and bias. The output tensor of the last full connection layer can be reshaped as a three-dimensional tensor with the grids' dimension and bounding box's parameters. Each grid of the image input will generate n bounding boxes (n=2 in the example figure), and the bounding box has the parameters: (i) confidence level, (ii) center coordinate, (iii) anchor box's size, (iv) the classification of the stuff in the bounding box. Therefore, when calculating the loss function of YOLO network, there are actually 4 kind of loss element: position loss, confidence level loss with object detected, confidence level loss with object undetected and classification loss.

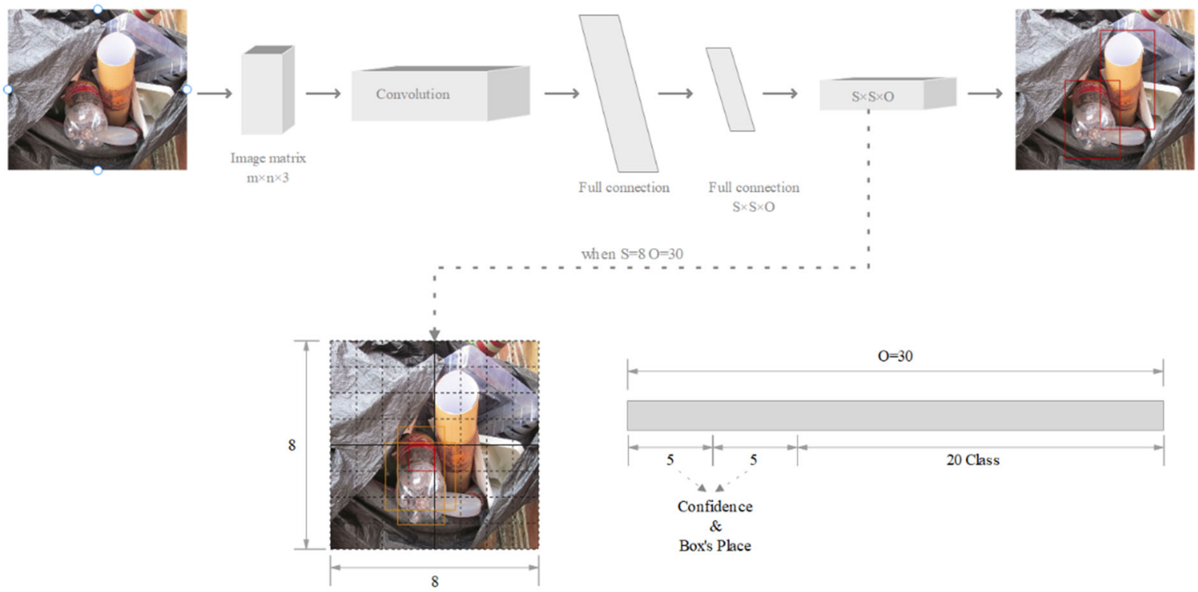

Figure 6. The Process of YOLO Network

YOLO can be well used in the field of waste classification with powerful object categorization function. YOLO has been modified and refined continuously, the first version YOLOv1 can process 45 frames of pictures per second, which solve the most criticized real-time problem of RCNN. However, it has flaws in recall and positioning accuracy, which means it can't detect objects close to each other and small groups. Therefore, YOLO network has been refined continuously, from YOLOv1 to YOLOv8.[4]

A plenty of high-precise garbage detection systems are implemented by YOLO Algorithm, such as YOLOv5. The YOLOv5-OCDS model proposed in [5] uses ODConv to replace a part of ordinary convolution in the neck and introduces adaptive adjustment to perceptual field and shape in convolution operation, which makes the convolution operation more flexible and adaptable and increases adaptability to small target objects and objects with large shape changes. The deformable convolution and C3 are fused to form a new structure, the C3DCN structure, which replaces a part of the C3 structure in the neck.

2.2.3 Lightweight Neural Networks

Lightweight neural networks are neural network models that need fewer parameters and computational effort. These models are typically designed to be deployed in resource-constrained or computationally limited environments, such as mobile devices, embedded systems, or IoT devices. Therefore, the models of lightweight neural networks can be applied widely, and in some economic backward areas, which have considerable application prospects.

In CNN model, the convolution layers with strides larger than 1 and the pooling layers can subsample the input images, which means the image can be compressed, and the weights needed in full connection layers will be decreased. In the model SqueezeNet, squeeze layer is proposed, which is actually a convolution layer with a 1×1 kernel to reduce feature maps' dimension. In addition, the MobileNet proposed a depth-wise separable convolution instead of the traditional convolution, which includes two operations: (i) depth-wise convolution, (ii) point-wise convolution. Instead of down-sampling by pooling, the step size is set to 2 when using depth-wise convolution to achieve the purpose of down-sampling.

Table 1. The introduction of YOLO series

Version	Network Structure	Improvements	Features
YOLOv1	●Input image: 448*448*3 (colored image). ●Intermediate Layer: several convolution layers and maximum pooling layers. ●Full-connection Layers with 7*7*30 predicted outcome.	-	●Less than 25ms delay, real-time video processing. ●There are problems of loss function and positioning.
YOLOv2	●Feature extraction network: Darknet-19 with 19 convolution layers and 5 pooling layers. ● Can accept input images of different sizes.	●Lead-in anchor mechanisms with K-Means. ●Add batch normalization to convolution layer ●Each cell can predict nine Anchor box.	●When the speed is 67fps, the accuracy can reach 76.8 mAP; and when the speed is 40fps, the accuracy can reach 78.6 mAP. ●It can be a good trade-off between speed and accuracy.
YOLOv3	●Feature extraction network: Darknet-53 with several 1×1 and 3×3 convolution layers, and every convolution layer is followed by a batch normalization and a Leaky Relu.	●Substitute pooling layers with convolution layers with stride=2. ●The idea of residual network introduced in order to make the network extract deeper features.	● It has faster speed and repetitive gradient information is optimized in network.
YOLOv5	●Backbone Network: CBAM and CSPDarknet-53	●By CBAM attention structure, input feature images are compressed. ●Utilize gray level filling method to unify input size. ●Loss function: utilize BEC loss function to optimize classification, while CLoU function is utilized to optimize bounding boxes.	●The detection accuracy increases while the detection speed decreases. ●According to the requirements of the project, the model can be chosen to achieve the best balance between accuracy and speed.
YOLOv8	●CSPDarknet-53 introduces C2f rich gradient flow.	●The C3 module in YOLOv5 was replaced by C2f module. ●Anchor-Free.	●Further lightweight is realized.

In [6], a garbage detection system is implemented by the methods of lightweight neural networks with light workload and has considerable efficiency. It proposed YOLOG which is an improved model an improved model based on YOLOv4 with optimized data augmentation, CSPResNet, overall structure, and a new activation function. It ensures real-time and accurate results and simplifies the hardware requirements, reduces computational costs, and meets the needs of practical applications.

3. CV-Technique Applications in Fields of Garbage Management

There are several image processing projects needed by garbage management fields. The most significant process works are: (i) object detection, (ii) object classification, (iii) image enhancement, anomaly detection. Table 2 outlines problems in the field of garbage management can be solved by CV techniques.

Table 2. CV Techniques Used in Garbage Management Areas

CV Function	Model Examples	Application Areas
Object Detection	●Modified YOLOv4 in [7]: deformable convolution is applied to adapt to perishable garbage anomalie. Then, Soft pool is changed in the spatial pyramid pool. Finally, the PIoU loss function is replaced by the CIoU loss function.	●Garbage Detection ●Anomaly Detection
Object Classification	●YOLOv5-OCDS model has been proposed in [5]. ●YOLOG updated YOLOv4 with CSPResNet proposed in [6].	●Waste Classification
Image Enhancement	●In [8], image is enhanced with the application of framework of probabilistic UIE, and reduces the loss of feature information by adding eSE attention. ●In [9], a multi-input network AMI-CycleGAN combined with adaptive spatial feature fusion is proposed.	●Removal of dirty on images obtained from camera sensors ●Bringing low-resolution objects up to high resolution

3.1 Urban Road Garbage Detection

With real-time litter detection system (by installing cameras on roads or using equipment such as drones), computer vision technology can monitor litter on roads in real time. It can recognize and detect various kinds of garbage on the road, such as paper, plastic bottles or cigarette butts, thus helping the city administration to clean up in time. With different CNN algorithm models of Deep Learning, the object detection function can be implemented. [10]

3.2 Marine Litter Detection

With the application of computer vision technology, a marine litter monitoring system can be established to monitor the distribution of litter in the ocean and provide early warning information. This will help to protect the marine ecosystem and prevent litter from harming and polluting marine life. Combining autonomous robots such as unmanned ships or submersibles and monitors or cameras, CV can be applied to recognize different types of trash, such as plastic, metal or glass. The main challenges for marine litter detection are the limits caused by environment and poor devices.[11]

3.3 Garbage Classification

Intelligent garbage classification system can be implemented by Deep Learning method, such as YOLO or Arithmetic Optimization Algorithm (AOA). Computer vision can be used to automatically identify and sort various types of waste, such as recyclables, hazardous materials and wet waste. The automatic garbage identification system would improve the efficiency of waste disposal. A smart bin is implemented by Arithmetic Optimization Algorithm (AOA) with Improved RefneDet (IRD) and EfcientNet-B0 model in [12].

4. Discussions

4.1 Current achievements

The application of CV techniques in the field of garbage management has already achieved significant results. Firstly, Important advances have been made in techniques for detecting real-time waste contamination using edge computing and video analytics. The refined algorithms of different Deep Learning model are achieved constantly, which improves the accuracy and efficiency of the process. In addition, computer vision-based platforms for waste sorting and recycling have also been developed with a rapid speed.

4.2 Deficiency

Although the algorithms of Deep Learning applied in the fields of garbage management are being refined, the actually utilization factor of the intelligent garbage management system is still low, which means a lot of algorithms are not really working in practice. Due to the complexity of the external environment, accuracy and stability are still a challenge to all sorts of algorithms. Recognition algorithms may be affected by the factors such as lighting conditions, occlusion, and changes in trash shape.

4.3 Challenges in the Future

(i) Deploying a computer vision system requires a huge cost in hardware equipment, software development, and maintenance, which can be a significant burden, especially for some smaller waste disposal regions. Therefore, reduction of algorithm implementation costs and improving algorithmic intelligence are still two main subjects of intelligent garbage management.

(ii) Real-time monitoring and processing of garbage needs quick response from computer vision systems in order to recognize and process garbage in time. However, some complex recognition algorithms would lead to long processing time and affect the real-time performance of the system. However, if the accuracy of algorithms is reduced, the algorithms then would not adapt the large number of waste types and diverse environments' complexity. Therefore, researchers need to find a balance between the accuracy and efficiency of the algorithm models when updating and applying them to the fields of garbage management.

4.4 Prospects and Future Directions

The application of computer vision in the fields of garbage management has a broad prospect, which can make an important contribution to the improvement of environmental quality, resource utilization efficiency and the promotion of urban intelligent development. With the continuous progress of technology, it is believed that its role in the field of garbage management will be more and more prominent. Real-time garbage monitoring and management, waste analysis and resource recovery and intelligent waste separation system can be implemented with the application of appropriate CV methods, which will contribute to improving the efficiency of waste disposal and reducing negative impacts on the environment.

5. Conclusions

This survey discussed the basic and tradition networks of CV, which is CNNs and YOLO series. Accuracy and speed are the two main performance parameters. The researches in order to improve the algorithms performance in the area of garbage management are mainly based on the network such as YOLOv4, YOLOv5 and ResNet. And the main methods to improve the algorithms are (i) substituting the traditional convolution layer with new-means convolution layer, (ii) utilizing different kinds of activation function, (iii) adding attention modules. With the continuous development of CV algorithms, there's a greater possibility to overall intelligentize garbage management systems.

The application of computer vision in the fields of garbage management is very useful, helping improve the efficiency of waste disposal, reduce the negative impact on the environment, and advance the goals of sustainable development. Different kinds of algorithm model of CV have been applied in the fields of garbage management. However, the accuracy and stability of the algorithms and the cost to implement the algorithms are still challenges that need to be overcome in the future research. Researchers are required to develop more practical methods to apply the advanced CV techniques in the fields of garbage management.

References

[1] Patrikar, D.R.; Parate, M.R. Anomaly Detection Using Edge Computing in Video Surveillance System: Review. Int. J. Multimed. Inf. Retr. 2022, 11, 85–110. [CrossRef] [PubMed]

[2] Gunaseelan J ,Sundaram S ,Mariyappan B .A Design and Implementation Using an Innovative Deep-Learning Algorithm for Garbage Segregation[J].Sensors,2023,23(18):

[3] S. Woo, J. Park, J. Y. Lee and I. S. Kweon, "CBAM: convolutional block attention module[J]", arXiv:1807.06521, pp. 3-19, 2018.

[4] XU Yanwei, LI Jun, DONG Yuanfang, et al. Survey of the Development of YOLO Object Detection Algorithms[J/OL]. Journal of Frontiers of Computer Science and Technology, 1-19[2024-07-17].http://kns.cnki.net/kcms/detail/11.5602.TP.20240611.1631.006.html.

[5] Sun Q ,Zhang X ,Li Y , et al.YOLOv5-OCDS: An Improved Garbage Detection Model Based on YOLOv5[J].Electronics,2023,12(16):

[6] Cai Xinchen, Shuang Feng, Sun Xiangming, et al. Towards Lightweight Neural Networks for Garbage Object Detection. Sensors. 22. 7455. 10.3390/s22197455.

[7] Guo, Mengting. Research on the Algorithm for Detection of Abnormal Objects of Perishable Garbage and Identification of Littering Behavior[D].China Jiliang University, 2021.DOI:10.27819/d.cnki.gzgjl.2021.000437.

[8] Liao, Chenjin. Garbage Target Detection Based on Enhanced YOLOv7 in Marine Environment [J]. Application of Electronic Technology, 2024, 50 (06): 66-70. DOI:10.16157/j.issn.0258-7998.244869.

[9] Ren Xiaotian, Zuo Guanfang, Tao Xu, et al. Marine Garbage Image Enhancement Algorithm based on AMI-CycleGAN [J]. Foreign Electronic Measurement Technology, 2023, 42 (06): 33-42. DOI:10.19652/j.cnki.femt.2304823.

[10] Esma D ,Murat D .Computer Vision Applications in Intelligent Transportation Systems: A Survey[J].Sensors,2023,23(6):2938-2938.

[11] Fang R ,Duan ZY ,Liu ZZ, et al. Review on marine litter treatment technologies [J]. Integrated Smart Energy,2023,45(05):70-79.

[12] Yu S ,Xin H ,Xiwang T , et al.DEEPBIN: Deep Learning Based Garbage Classification for Households Using Sustainable Natural Technologies[J].Journal of Grid Computing,2023,22(1)