INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Check for updates

OPEN ACCESS

Received: 26-06-2023

Accepted: 03-10-2023

Published: 13-11-2023

Citation: Srilatha J, Subashini TS, Vaidehi K (2023) Solid Waste Detection and Recognition using Faster RCNN. Indian Journal of Science and Technology 16(42): 3778-3785. https://doi.org/10.17485/IJST/v16i42.2005

* **Corresponding author**.

srilathajanakiraman@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Solid Waste Detection and Recognition using Faster RCNN

J Srilatha[1]*, T S Subashini[2], K Vaidehi[3]

1 Research Scholar, Department of CSE, Annamalai University, Chidambaram, 608002, Tamil Nadu, India
2 Professor, Department of CSE, Annamalai University, Chidambaram, 608002, Tamil Nadu, India
3 Associate Professor, Department of CSE, Stanley College of Engineering and Technology for Women, Hyderabad, 500001, Telangana, India

Abstract

Objective: To develop a two-stage object detection method based on convolutional neural networks (CNNs) to identify and classify solid waste, contributing to the creation of intelligent systems for society. **Methods:** The study utilizes a base network, ResNet 101, to generate convolution feature maps. In the first stage, a Region Proposal Network (RPN) is created on top of these convolution features, producing 256-dimensional feature vectors, objectness scores, and bounding rectangles for different anchor boxes. In the next stage, the region proposals are used to train a softmax layer and regressor, enabling the classification and localization of five types of solid waste, namely cardboard, glass, metal, paper and plastic. **Findings:** The proposed Faster RCNN demonstrates nearly real-time object detection rates. Experimental results reveal that the Faster RCNN with ResNet 101 and RPN achieves an accuracy of 96.7%, outperforming the Faster RCNN with a simple CNN, which achieves an accuracy of 86.7%. **Novelty:** Unlike traditional R-CNN, which relies on computationally inefficient selective search, the proposed Faster RCNN employs RPN, a small neural network sliding on the last convolution layer's feature map, predicting object presence and bounding boxes. This approach significantly improves efficiency compared to the exhaustive examination in R-CNN's selective search.

Keywords: Object Detection; RCNN; Fast RCNN; Faster RCNN; RPN; ROI pooling

1 Introduction

Despite the availability of various object detection methodologies in the literature, deep neural networks (DNNs) [1–3] have recently gained widespread usage. Numerous publications in the literature explore the utilization of CNN and its variations for solid waste object detection and classification. Some research papers focus on autonomous solid waste segregation using deep learning techniques. For instance, authors in [4] proposed a method based on SVM and CNN for image classification, successfully delineating photos into six classes of objects. In another study [5], a 50-layer ResNet-50

CNN model and SVM were employed to classify various types of waste, including glass, metal, paper, and plastic achieving effective categorization. Notably, CNNs offer maximum learning efficiency, require fewer training parameters, and exhibit superior accuracy compared to regular networks[6]. Moreover, in[7], authors applied Scale Invariant Feature Transform-Principal Component Analysis (SIFT-PCA) for feature extraction and SVM for classification, obtaining a 62% accuracy for automatic waste segregation. In[8], a pre-trained CNN, specifically MobileNet, was utilized as a feature extractor along with SVM as a classifier to achieve an accuracy of 94% for plastic and non-plastic detection. Overall, the literature underscores the potential of deep learning techniques in achieving high accuracy in object detection and classification from images. These approaches hold significant implications for the development of automated waste management systems and sustainable practices, leading to a cleaner and healthier environment for present and future generations. Nevertheless, further research is essential to enhance the robustness and effectiveness of these methods in real-world scenarios.

2 Methodologies

The methodology is presented in Figure 1, utilizing two different backbone networks, namely simple CNN and Resnet101, to generate feature maps fed to the RPN for region proposal generation. For the experiments, a custom solid waste dataset was created, comprising images collected from Garbage classification dataset on kaggle, categorized into Cardboard, glass, metal, paper and plastic. These images were resized to 224 x 224 and annotated for training. The study employed a total of 2000 images, with 400 images per class. Among these, 320 images per class were used for training and 80 for testing, resulting in 1600 training images and 400 testing images. The work was implemented using TensorFlow in Python on Google Colab with a 4 GB NVIDIA GPU.

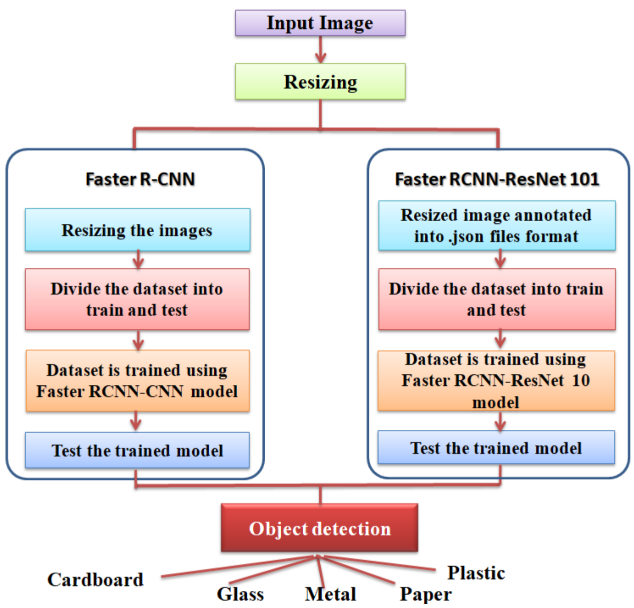

Fig 1. The flow chart of proposed solid waste object detection system

In this work, the most recent version of Region-CNN (R-CNN), namely faster RCNN, is proposed for classifying and locating five types of solid waste, including cardboard, glass, metal, paper and plastic objects. Faster RCNN operates as a two-stage detector, generating region proposals in the initial stage, which are then processed for object categorization and bounding-box regression in the subsequent stage. The study explores two variations of faster RCNN, one employing a simple CNN as a base network for feature extraction and the other utilizing ResNet 101. The region proposal network is created on top of the last layer feature map of the base network, and the region proposals obtained are used to train a softmax layer and regressor for solid waste classification.

2.1 Faster RCNN

Faster R-CNN[9] is an advanced object detection model that enhances Fast R-CNN by incorporating a region proposal network (RPN)[10] with the CNN model. The RPN efficiently shares full-image convolutional features with the detection network, resulting in nearly cost-free region proposals. Operating as a fully convolutional network, the RPN predicts object bounds and objectness scores simultaneously at each position. Through an end-to-end training process, the RPN generates high-quality region proposals, which are subsequently utilized by Fast R-CNN for detection. RPN and Fast R-CNN are seamlessly integrated into a single network by sharing their convolutional features. The Faster RCNN object detection network comprises a feature extraction network, which can be a simple CNN or a pre-trained CNN. Following this, two additional trainable sub-networks are introduced. The first network, known as the Region Proposal Network (RPN), predicts an object's actual class, while the second network, consisting of a classifier and regressor, predicts its category. For a visual representation, you can refer to the architectural diagram of Faster RCNN shown in Figure 2.

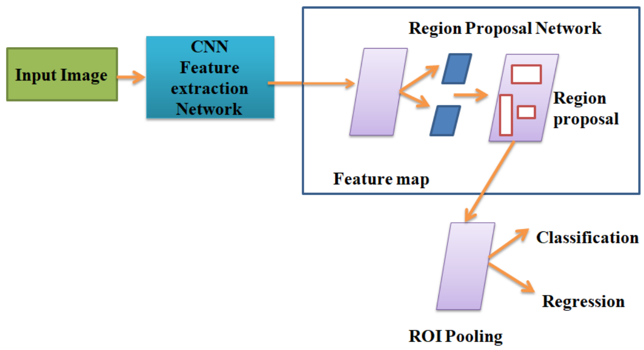

Fig 2. Architecture of Faster RCNN

2.2 Feature Extraction Network

In this study, two distinct networks, namely i) a simple CNN and ii) the ResNet 101 network, are utilized for generating feature maps.

- **Simple CNN**

Convolutional neural networks (CNNs) offer an optimal architecture for identifying and comprehending key features within an image. With tens or even hundreds of layers, each specialized in detecting distinct image attributes. CNNs apply filters to training images at various resolutions, using the outputs of convolved images as inputs for subsequent layers. These filters begin by recognizing simple features like brightness and edges, and progressively advance to more complex features that uniquely define objects. The general CNN architecture is depicted in Figure 3, and for a more detailed understanding of CNN, refer to [11,12].

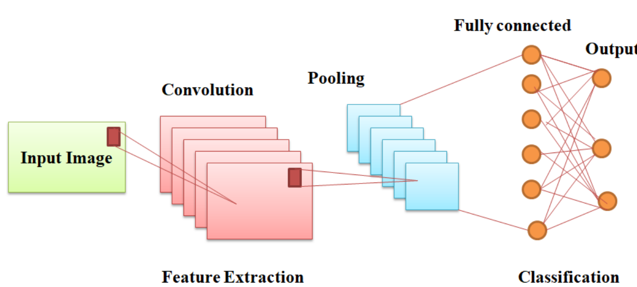

Fig 3. Architecture of Simple CNN

- **Resnet 101**

Deep neural networks face a bottleneck where their performance declines as they become deeper. This happens because the gradients used to calculate the loss function diminish to near-zero values after multiple applications of the chain rule. Consequently, the network's weights do not update, leading to a lack of learning. ResNets address this issue by allowing gradients to flow directly through identity/skip connections, back from later layers to earlier filters. The ResNet architecture comprises one convolution and pooling step, followed by four layers with similar behavior. Each layer involves 3x3 convolutions with a fixed feature map dimension (F) and bypasses the input every two convolutions. Moreover, the width (W) and height (H) dimensions remain constant throughout each layer. Before entering the common layer behavior, ResNet employs a 'block,' denoted as Conv1, which consists of convolution, batch normalization, and max pooling operations. Every ResNet layer is composed of several such blocks. When ResNets deepen, they typically achieve it by increasing the number of operations within a block, while maintaining a total of four layers. An operation in this context refers to a convolution, batch normalization, and ReLU activation applied to an input, with the last operation of a block omitting the ReLU. For a comprehensive explanation of ResNet, refer to [(13)].

In this study, a variant of ResNet, specifically ResNet-101, serves as the backbone network, and it comes with fixed weights preloaded in the PyTorch repository. ResNet-101 is a convolutional neural network with an impressive depth of 101 layers. The initial convolution layer in the ResNet-101 architecture employs 7x7 filters, and down-sampling between the stages is achieved using a stride of 2 instead of max pooling layers. The final layer is an overall average pooling layer connected with a softmax layer. To enhance the network's performance and stability during training, batch normalization is applied after each convolute on layer. The architecture of the 101-layer ResNet can be visualized in Figure 4.

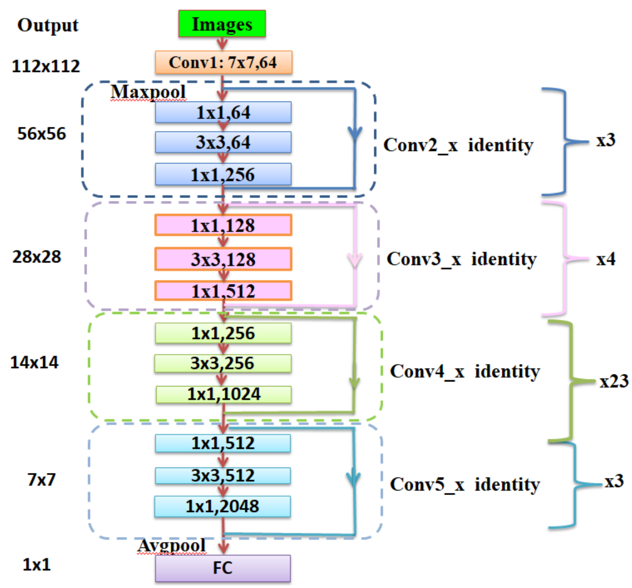

Fig 4. Architecture of ResNet-101

2.3 Region Proposal Network

The Region Proposal Network (RPN) is a fully convolutional network responsible for generating proposals with different scales and aspect ratios. It takes the convolutional feature map generated by the backbone layer as input and uses sliding window convolution to produce the anchors. The RPN's primary task is to detect "good" anchor boxes, which are then forwarded to the next layer. It consists of a classifier and a regressor. The classifier uses Intersection over Union (IoU) metric to predict if an anchor box contains an object or parts of the background based on certain IoU thresholds. On the other hand, the regressor predicts offsets to tightly fit the anchor boxes containing objects to the ground truth labels. For a visual representation, refer to Figure 5, which illustrates the architecture of RPN [(14)].

- **NMS and RoI Pooling**

In addition to IOU thresholding for filtering, Non-maximum suppression (NMS) is applied as an extra technique to remove overlapping bounding boxes [(15)]. Subsequently, the RoI pooling layer is utilized to convert the variable-sized generated proposals

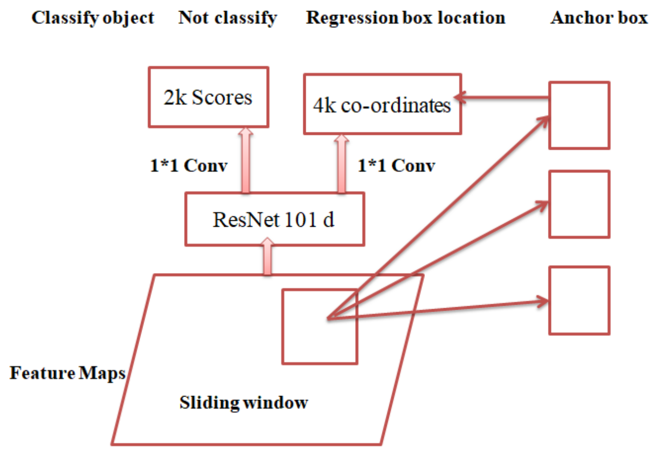

Fig 5. Architecture of RPN feature extractor

into a fixed size, enabling the execution of a classifier and bounding box regression on top of them.

2.4 Proposed enhanced-Faster RCNN with ResNet 101

The Faster RCNN model is an advanced two-stage network for detecting objects, inheriting the robust qualities of the R-CNN family while exhibiting highly precise detection capabilities. To enhance the speed of detection, a regional proposal network (RPN) is introduced to replace the less efficient selective search algorithm, which doesn't fully utilize GPU resources. The RPN, functioning as a fully convolutional network, generates features for both box regression and box classification using predefined anchors as reference points for regression. After RPN generates proposals, they, along with backbone features, are fed into the classifier. The anchor design already accounts for various scale and ratio patterns, enabling the architecture to be effectively trained on single-scale images. For rapid adaptation to specific applications, a transfer learning approach is adopted, utilizing a pretrained ResNet-101 [15] backbone with an input image size of 224x224 for training the Faster RCNN model. Configuration parameter adjustments include the number of classes, training steps, batch size, and maximum number of detection boxes. The architectural design of Faster RCNN ResNet-101 [16] [17] can be observed in Figure 6.

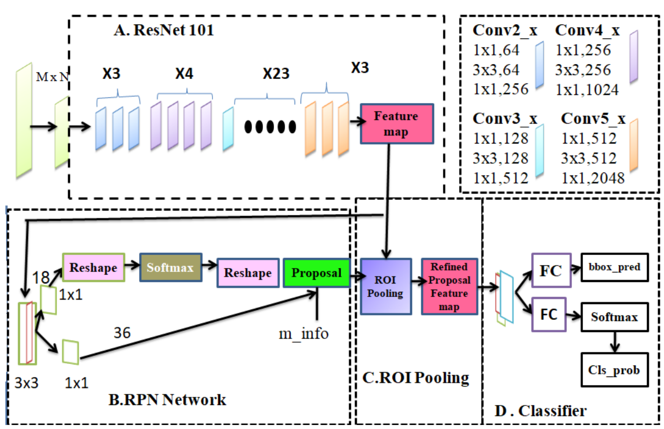

Fig 6. Architecture of enhanced-Faster RCNN with ResNet 101

The novelty of using Faster RCNN with ResNet-101 in the context of the Garbage detection is to identify waste objects of different scales and aspect ratios. This novel method combines cutting-edge object detection with a strong image recognition foundation, enabling very precise and effective identification of different categories of cardboard, glass, metal, paper and plastic waste object in the form of images. This combination makes use of deep learning's capacity to understand subtle features and correlations within waste images, in contrast to conventional solutions, which frequently rely on rudimentary image processing techniques. The proposed work unleashes the potential to not only classify solid waste items but also precisely localize and outline them inside images by combining enhanced Faster RCNN's area proposal network with ResNet-101 excellent feature extraction capabilities. This Proposed technique has the potential to have a big impact on Solid waste management, and environmental conservation initiatives, opening the door for more sensible and sustainable waste handling and reduction techniques.

- **Dataset**

The dataset has five categories of image data, namely, cardboard, glass, metal, paper and plastic, with a total of 2000 image data samples. The images are of various sizes and have been labeled with the corresponding category. It was created by Yousefi in 2021 and the dataset is available in Kaggle repository. The dataset is split into training and testing groups.

3 Result

The classification outcomes of faster-RCNN using a simple CNN as the backbone network are illustrated in Figure 7, displaying a sensitivity=86.8%, specificity=97%, f1-score=86.8% overall accuracy of 86.7%.

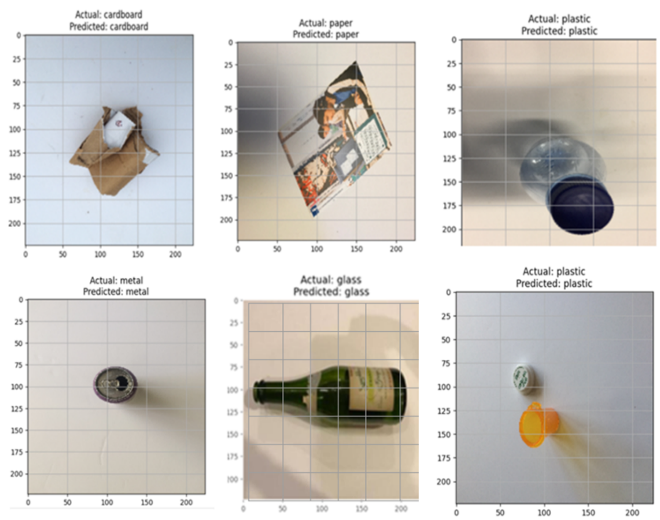

Fig 7. Classification Results of Faster RCNN (simple CNN as a backbone network)

Fig 8. Classification Results of Faster-RCNN with ResNet 101 as a backbone network

The classification results of faster-RCNN using ResNet 101 as the backbone network are presented in Figure 8, showcasing a sensitivity=96.8%, specificity=99%, f1-score=96.9% and impressive overall accuracy of 96.7%.

The accuracy comparison graph of the proposed waste detection systems shown in Figure 9 clearly demonstrates that Faster RCNN with ResNet 101 as its backbone is highly efficient in accurately detecting and localizing solid waste when compared to Faster RCNN employing simple CNN as backbone.

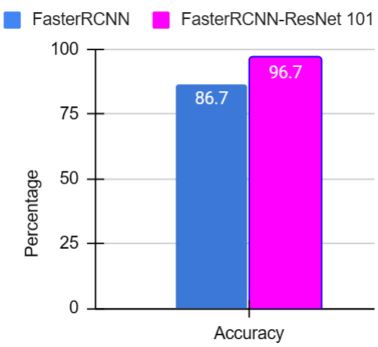

Fig 9. Accuracy comparison of Faster RCNN with simple CNN and ResNet 101 as a backbone network

4 Discussion

Table 1. Accuracy comparision of the proposed methods with existing methods

Existing Methods	Methods	Accuracy
(18)	Faster RCNN with ResNet 50	93.3 %
Proposed Method	Faster R-CNN(with simple CNN)	86.7%
	Faster RCNN-Resnet 101	96.7%

Table 1 shows comparison between existing methods and proposed method. In proposed work exhibits the best accuracy of 96.7% using Faster RCNN with ResNet 101. Therefore the contributions of this paper may be used as a novel object detection model that automatically detects garbage with high detection rates.

5 Conclusion

This study focused on investigating the performance of the CNN-based two-stage object detector, Faster RCNN, for solid waste detection, classification, and localization. Two backbone networks, Simple CNN and ResNet 101, were utilized for feature extraction. The proposed Faster RCNN employs RPN, a small neural network sliding on the last convolution layer's feature map, predicting object presence and bounding boxes. The comparison of their performance revealed that Faster RCNN with ResNet 101 as the feature extractor showed remarkable efficiency in accurately segregating solid waste. There is no benchmark data set exists, which is a major drawback for the researchers comparing their model performance with benchmark results. So, firstly annotated benchmark data set may be constructed for each waste category to facilitate future research.

References

1) Yang Z, Bao Y, Liu Y, Zhao Q, Zheng H, Bao Y. Research on deep learning garbage classification system based on fusion of image classification and object detection classification. *Mathematical Biosciences and Engineering*. 2023;20(3):4741–4759. Available from: https://www.aimspress.com/article/doi/10.3934/mbe.2023219.
2) Adedeji O, Wang Z. Intelligent waste classification system using deep learning conv olutional neural network. *Procedia Manufacturing*. 2019;35:607–612. Available from: https://doi.org/10.1016/j.promfg.2019.05.086.
3) Atikuzzaman M, Hossain MP, Islam MZ, Kabir SA. A Comparative Analysis of Convolutional Neural Networks for Trash Classification. *GUB Journal of Science and Engineering*. 2021;8(1):17–23. Available from: https://doi.org/10.3329/gubjse.v8i1.62327.
4) Habib G, Qureshi S. Optimization and acceleration of convolutional neural networks: A survey. *Journal of King Saud University - Computer and Information Sciences*. 2022;34(7):4244–4268. Available from: https://doi.org/10.1016/j.jksuci.2020.10.004.
5) Ma X, Li Z, Zhang L. An Improved ResNet-50 for Garbage Image Classification. *Tehnicki vjesnik - Technical Gazette*. 2022;29(5):1552–1559. Available from: https://hrcak.srce.hr/file/408378#:~:text=Adedeji%20and%20Wang%20%5B16%5D%20proposed,accuracy%20of%2087%25%20was%20achieved.
6) Taye MM. Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions. *Computation*. 2023;11(3):1–23. Available from: https://doi.org/10.3390/computation11030052.

7) Puspaningrum AP, Endah SN, Sasongko PS, Kusumaningrum R, Khadijah, Rismiyati, et al. Waste Classification Using Support Vector Machine with SIFT-PCA Feature Extraction. In: 2020 4th International Conference on Informatics and Computational Sciences (ICICoS), 10-11 November 2020, Semarang, Indonesia. IEEE. 2020;p. 1–6. Available from: https://ieeexplore.ieee.org/document/9298982/authors#authors.
8) Srilatha J, Subashini TS, Vaidehi K. Classification of Plastic and Non-Plastic Wastes Using Mobile net and SVM. *Mathematical Statistician and Engineering Applications*. 2022;71(4):7267–7278. Available from: https://doi.org/10.17762/msea.v71i4.1344.
9) Faisal M, Chaudhury S, Sankaran KS, Raghavendra SJ, Chitra RJ, Eswaran M, et al. Faster R-CNN Algorithm for Detection of Plastic Garbage in the Ocean: A Case for Turtle Preservation. *Mathematical Problems in Engineering*. 2022;2022:1–11. Available from: https://doi.org/10.1155/2022/3639222.
10) Sarosa M, Muna N, Rohadi E. Performance of Faster R-CNN to Detect Plastic Waste. *International Journal of Advanced Trends in Computer Science and Engineering*. 2020;9(5):7756–7762. Available from: https://www.warse.org/IJATCSE/static/pdf/file/ijatcse120952020.pdf.
11) Sharma AK, Jain A, Chaudhary D, Tiwari S, Mahdin H, Baharum Z, et al. An Approach to Automatic Garbage Detection Framework Design ing using CNN. *International Journal of Advanced Computer Science and Applications*. 2023;14(2):257–262. Available from: https://thesai.org/Downloads/Volume14No2/Paper_31-An_Approach_to_Automatic_Garbage_Detection_Framework.pdf.
12) Appe SRN, Arulselvi G, Balaji GN. Tomato Ripeness Detection and Classification using VGG based CNN Models. *International Journal of Intelligent Systems and Applications in Engineering*. 2023;11(1):296–302. Available from: https://www.ijisae.org/index.php/IJISAE/article/view/2538.
13) Patel D, Patel F, Patel S, Patel N, Shah D, Patel V. Garbage Detection using Advanced Object Detection Techniques. In: 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 25-27 March 2021, Coimbatore, India. IEEE. 2021. Available from: https://ieeexplore.ieee.org/abstract/document/9395916.
14) Chen Y, Sun J, Bi S, Meng C, Guo F. Multi-objective solid waste classification and identification model based on transfer learning method. *Journal of Material Cycles and Waste Management*. 2021;23(6):2179–2191. Available from: https://doi.org/10.1007/s10163-021-01283-8.
15) Ding J, Zhang J, Zhan Z, Tang X, Wang X. A Precision Efficient Method for Collapsed Building Detection in Post-Earthquake UAV Images Based on the Improved NMS Algorithm and Faster R-CNN. *Remote Sensing*. 2022;14(3):1–23. Available from: https://doi.org/10.3390/rs14030663.
16) Zhou QQQ, Wang J, Tang W, Hu ZCC, Xia ZYY, Li XSS, et al. Automatic Detection and Classification of Rib Fractures on Thoracic CT Using Convolutional Neural Network: Accuracy and Feasibility. *Korean Journal of Radiology*. 2020;21(7):869–879. Available from: https://doi.org/10.3348/kjr.2019.0651.
17) Guo Q, Liu L, Xu W, Gong Y, Zhang X, Jing W. An Improved Faster R-CNN for High-Speed Railway Dropper Detection. *IEEE Access*. 2020;8:105622–105633. Available from: https://ieeexplore.ieee.org/document/9110596.
18) Nie Z, Duan W, Li X. Domestic garbage recognition and detection based on Faster R-CNN. In: 2020 2nd International Conference on Electronics and Communication, Network and Computer Technology (ECNCT), 23-25 October 2020, Chengdu, China;vol. 1738 of Journal of Physics: Conference Series. IOP Publishing. 2021;p. 1–7. Available from: https://iopscience.iop.org/article/10.1088/1742-6596/1738/1/012089.