# Single Use Plastic Bottle Recognition and Classification Using Yolo V5 and V8 Architectures

Venkata Durgarao Matta(✉), K. A. Venkata Ramana Raju Mudunuri, B. Ch. S. N. L. S. Sai Baba, Kompella Bhargava Kiran, C. H. Lakshmi Veenadhari, and B. V. Prasanthi

Computer Science and Engineering Department, Vishnu Institute of Technology, Bhimavaram, Andhra Pradesh, India

durgaraomatta@gmail.com

**Abstract.** Improper disposal of single use plastic bottles leads to many problems including danger to marine life and land pollution. Burning of plastic in turn releases dioxins and polychloride biphenyls. They are very harmful if inhaled and are threat to vegetation too. Manual sorting of plastic bottles and safe disposal is not an easy task. A lot of recycling initiatives use manual sorting for plastic recycling, which depends on plant staff visually identifying and selecting plastic bottles as they move along the conveyor belt. Automatic sorting of plastic bottles has advantage of non-intrusive sorting, speed, consistency, cost effectiveness in long run and even prevents health hazards to workers working in recycling environment. As a result, it is imperative to replace human sorting systems with intelligent automated systems. In this study, convolutional neural network architectures such as YOLOv5 and YOLOv8 were utilized to detect plastic bottles in images. Despite YOLOv8 having more parameters and requiring more computation time, it was found that YOLOv8 outperformed YOLOv5 in accurately identifying plastic bottles in the images.

**Keywords:** CNN · Computer Vision · Plastic bottles · Yolo V5 and Yolo V8

## 1 Introduction

It is estimated that 2.5 million tons of $CO_2$ is being released annually by improper water bottle disposals. Approximately, 1.1 million ocean creatures die because of plastic pollution caused by plastic water bottles. Beside this, many toxic gases like polychlorinated biphenyls are released by burning plastic bottles [1]. Determining the amount of plastic bottle waste is a big issue. Continuous usage of plastic bottles and no proper disposal might lead to many disasters and disturbs the harmony of nature. Determining the plastic bottles waste in public areas is an important area of research. In future, the cities may be ranked based on plastic management. One of the solutions to find the solution is to use computer vision models which are built using convolutional neural network (CNN) to

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2024
Published by Springer Nature Switzerland AG 2024. All Rights Reserved
P. Pareek et al. (Eds.): IC4S 2023, LNICST 537, pp. 99–106, 2024.
https://doi.org/10.1007/978-3-031-48891-7_8

find out the plastic bottle usage in an area. These models can be built on techniques like Histogram of Oriented Gradients (HOG), Region-based Convolutional Neural Networks (R-CNN), Faster R-CNN. The models used above gave less accuracy and performance is not up to the mark in many cases.

The basic concept of this paper is to use some good object detection techniques for uncovering plastic bottles in an given area. In this paper a we have used YOLO V5, V8 as they are Single shot detection techniques to find the targe image with a better accuracy. The output by using these techniques have bounding boxes along with probability of classification. The reason for selecting these models is for its speed. They are one short detection techniques. Means, the image is passed only once [3]. These features make these models methodical and are better algorithms. On the other hand, region-based CNN find the approximate region and recognizes the bounding box in a separate stage but are more accurate. The YOLO has gone through five levels of major iterations. Minimal requirement of data, simple architecture and easy implementation make these architectures unique.

The paper is organized as follows. The related work is first reviewed. The architectures are then described. The experimental outcomes are then emphasized. Finally, the paper is concluded.

## 2 Related Work

Walden et al. [4] proposed a paper in which the plastic bottles are converted into hue, saturation and (HSV). Filters are then applied along with binary thresholding. These images are then converted into gray pixels. Blob analysis is then utilised to determine the number of plastic bottles. Here, no CNN based approach is used and there is no way to increase the accuracy of models. Dhokley et al. [5] proposed a similar method to detect plastic waste and uses YOLO V3. But the proposed approach uses much advanced models. Christopher et al. [6] proposed a paper "PET-Bottle-Recognizer" to detect Polyethylene-Terephthalate Based- Bottles. This accuracy of the model is 85.70%. They have sued mean average pooling to calculate the accuracy. Jungiu et al. [7] have come with an approach based on improved yolo V3. The system extracts the features by using shufferNet network. The screening accuracy is around 91.3% provided the detection rate is set at 26 frames per second. An approach proposed by keqiong et al. uses stochastic configuration network and yolo v5 to detect the plastic bottles[8]. They have dataset is generated using Hikvision MV CE050-30GM camera. Gilroy et al. [9] have performed the detection of plastic bottles in river by using yolo V5 algorithm which is focused on custom dataset. The model has an accuracy of 84%, a precision rate of 79.14%, and a recall rate of 57.37% when deployed on raspberry pie.

Most of the researchers have used models like VGG 16, YOLO V3, YOLO V5 and YOLOV5s and the accuracy of the models were around 85%. This paper uses the latest version YOLO V5 and YOLO V8. Our model is tested on Plastic Bottles in the wild Image Dataset from Kaggle.

## 3 Architectures Used

This dataset has 8000 images out of which 70% is used for training, and 20% & 10% for testing and validation respectively. The rationale behind this choice is to balance the trade-off between having enough data for training, assessing model performance on unseen data, and tuning hyperparameters. [10]. The model built is a binary classifier, where we check for plastic bottle in an image. In this paper, we are using YOLO for object detection. Earlier, Object detection was performed using sliding window method. Later more faster versions like, Region based Convolutional neural network(R-CNN) [11], fast Region based convolutional neural network (Fast R-CNN) [12], faster region based convolutional neural network (Faster R-CNN) [13]. In 2016, YOLO (you only look once) were invented which outperformed all the previous pervious object detection algorithms. In case of image classification, we just look if the object is present in an image. But, in the case of object detection, we exactly look for object inside an image using bounding boxes. This is referred to as object localization. In terms of a bounding box, we have a vector of minimal elements [PC, BX, BY, BW, BH C1, C2]. Pc is the probability of a given class, Bx, By is the center coordinate, and Bw and Bh are the width and height of the bounding box, C1 and C2 are the class labels. If there is no object in an image, the value of Pc is 0 and the rest of the values in the vectors do not have any meaning. Here if we need to detect multiple objects, the vector size will be increased accordingly. For example, if 10 objects need to be detected, then vector size will be 70. In case of yolo algorithm, the image is divided into grids. There is no rule for dividing the data into specific number of grids. If an image is divided into a $4 \times 4$ grid, then each grid is individually searched for the object based on the coordinates of center. And, if each grid is represented by a vector of size 7, the image will have $4 \times 4 \times 7$ volume of information. So, the training attribute is images with grid and bounding boxes and training labels would be a three-dimensional vector. While predicting for objects in an image, the output would be 16 vectors in case of $4 \times 4$ grid.

Basic YOLO algorithm has some limitations, at first it can detect multiple bounding rectangles for a given object. One approach to solve this issue is to select the bounding box with highest probability. This approach works in case of single object detection. In case of multiple object detection, another approach called Intersection over union might work well. This method finds the rectangles with overlapping area

$$IOU = \frac{Intersect\ area}{Union\ area} \tag{1}$$

The Intersection over union (IOU) method is also known as Non max suppression. The larger the value of IOU, the better the accuracy. In some cases, an object can be inside another object, this scenario can be handled by concatenation of 2 vectors resulting in a vector of size 14.

The first version of YOLO was released in year 2016 [14]. The concept of YOLO was related to regression. It was able to predict images at 45 frames per second. The similar version of yolo known as lighter YOLO could predict at a speed of 144 frames per second but with lesser layers. The next version of YOLO as known as YOLO 2 was able to input of different sizes and was able to balance between speed and accuracy [15]. In 2018, YOLO V3 was released which was based on Darknet-53 architecture

[16]. The YOLO versions up to three were proposed by Joe Redmon. Later version V4 was having features like Weighted-Residual-Connections (WRC) Cross Stage Partial connections (CSP), Cross Mini-Batch Normalization (CmBN), Self-adversarial training (SAT), Mish activation, Mosaic data augmentation, DropBlock regularization, CIoU loss. Later YOLO V5 was released and was the first version from yolo which was developed using Pytorch and removed the drawbacks of Darknet framework [17]. In year 2023, YOLO v7 was released. The Extended Efficient Layer Aggregation Network (E-ELAN) stands for the computing block in the YOLOv7 backbone [18]. By employing "expand, shuffle, merge cardinality" to accomplish the capacity to constantly increase the learning ability of the network without breaking the original gradient route, the YOLOv7 E-ELAN architecture helps the model learn better. Both YOLO v5 and V8 architectures use CSPDarknet53 architecture and the detection accuracy is improved by using anchor boxes.

The main advantage of YOLO v5 and YOLO v8 is their simplicity, single forward pass, lightweight and opensource. In order to reduce detection of same object multiple times, we use non-Maximum suppression technique. They use Adam and Mish as their optimizer and activation function respectively. [19]. Yolo architectures suffer with accurately localizing small objects. These are overcome using good resolution input images, more data argumentation and adjusting the anchor boxes is done. Using these approaches help to reduce false positive and false negative.

## 4 Experimental Results

We have ploted F1- Confidence curve, precision Recall curve, precision confidence curve and recal confindence curve to check the performance of the model. The relationship between a binary classifier's precision and recall as the decision threshold changes is depicted graphically by the F1 confidence curve. The F1 score, which is the harmonic mean of precision and recall, is plotted versus the confidence threshold. The trade-off between recall and precision is depicted by the curve, which displays how the f1 score changes as the confidence threshold is altered. The precision typically rises while the recall falls as the threshold rises, and vice versa. For a certain classification task, the f1 confidence curve can assist in determining the best threshold by balancing precision and recall. A high F1 score means that the classifier is successfully striking a balance between recall and precision. Using the f1 confidence curve,

An illustration of a binary classifier's performance at various classification thresholds is the precision-recall curve. The accuracy and recall values for various threshold values are plotted, Precision is the proportion of true positives among all projected positives, while recall is the proportion of true positives among all real positives. The precision-recall curve can be used to visualize the trade-off between precision and recall at different decision thresholds. A perfect classifier would provide a point in the top-right corner of the curve with precision and recall both equal to 1.0. On the curve, a random classifier would generate a straight line from (0,0) to (1,1). A decent classifier's curve ought to be as near the top-right as possible.A binary classifier's recall as a function of the degree of certainty or confidence in its predictions is shown graphically as the recall-confidence curve. It displays a visualization of the recall values at various classifier confidence levels.

The recall-confidence curve can also be used to assess the effects of several feature sets or hyperparameters on the performance of a single classifier or to compare the effectiveness of various classifiers. In general, better performance of the classifier is indicated by a higher recall value for a particular confidence level. It should be emphasized that the recall-confidence curve cannot give a thorough assessment of a classifier's effectiveness because it only examines recall and ignores accuracy or other metrics. Additional evaluation metrics, such as the precision-recall curve, F1 score, or area under the ROC curve, must be considered in addition to the recall-confidence curve for a more full assessment of the classifier's performance. The curve makes determining the confidence level at which a classifier performs well and the confidence level over which that performance begins to decline easy. It enables the selection of the confidence level that optimizes remembrance, which might be useful when the goal is to achieve high recall at the expense of lower precision.

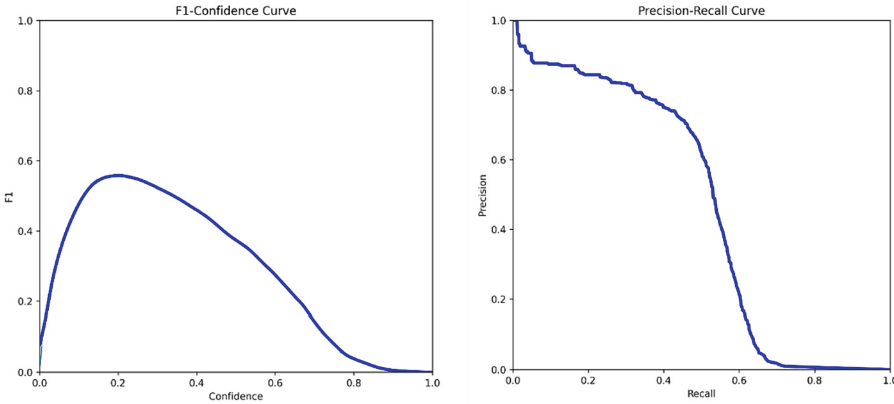


**Fig. 1.** F1 confidence curve for YOLO V8 (left). Precision Recall curve for YOLO V8 (Right)

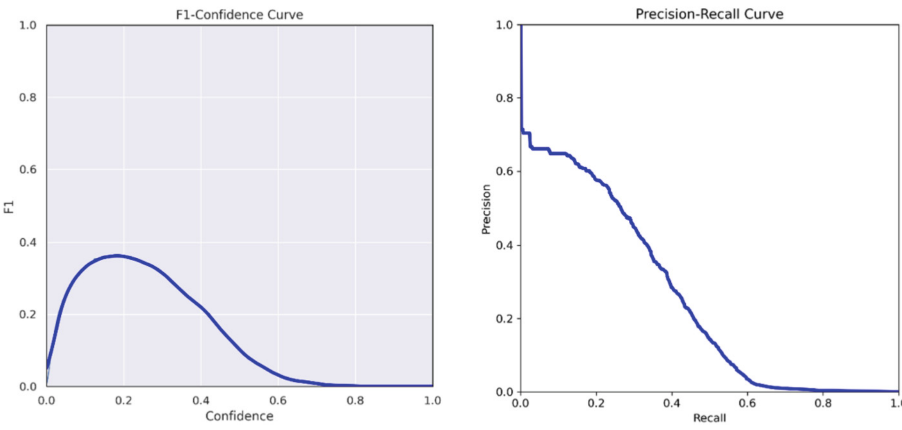


**Fig. 2.** F1 confidence curve for YOLO V5 (left). Precision Recall curve for Yolo V5 (Right)

Single-use plastic bottles come in various shapes, sizes, and materials. Some bottles may have unique designs or labels that differ significantly from the original training data. This problem can be reduced to an extent by using multiple data argumentation techniques.

As we can see above Fig. 1 and 2, the mAp value of models build using YOLO V5 and V8 stands at 0.252 and 0.46 respectively. Similarly, F1 score of both models stands 0.56 and 0.36. The same is displayed in the below table. Yolo V5 was run for 40 epochs and Yolo V8 was run for 25 epochs. Overfitting was observed after increasing the number of epochs. The images were resized to 416 × 416 and the batch size was 16 in both environments.

The experiment was performed on machine with NVIDIA QuADro GV100 having 5120 CUDA cores and 640 Tensor cores (Table 1).

**Table 1.** Mean average precision and F1 score of Yolo V5 and YOLO V8

| | mAP@ 0.5 | F1 Score |
|---|---|---|
| Yolo V5 | 0.252 | 0.36 |
| Yolo V8 | 0.46 | 0.56 |

The above performance comes with a price. YOLO V5 uses 10.6 Million parameters and YOLO V8 uses 61.3 M parameters. Figure 3 shows the sample identification of plastic bottles in different settings. Bounding boxes can be seen around the plastic bottles and displays the probability of the object to be a single use plastic bottle. The accuracy lies around 0.3 to 0.6 most of the time.

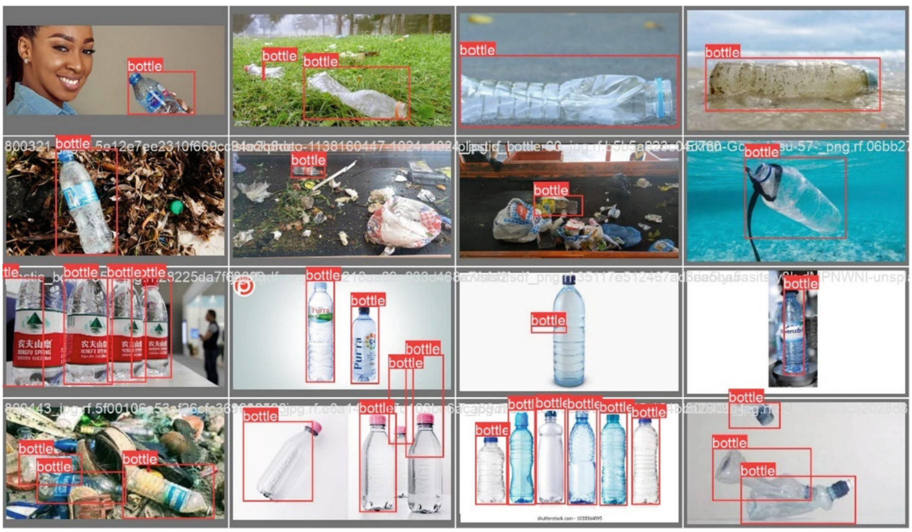

**Fig. 3.** Sample detection of plastic bottles by YOLO V5 and V8.

## 5 Conclusion

Given the abundance of plastic bottle waste in the environment, it is easy to come across scenes and photos of this trash. The main goal of this paper is to detect single-use plastic bottles in each image by employing cutting-edge convolutional neural networks. The dataset consists of 8000 images of pre-annotated single use plastic bottles collected from publicly available sources. This paper contrasts the performance of models build using YOLO v5 and v8 architectures to detect plastic bottles. It can be observed that efficiency of model created using YOLO V8 has outperformed the model build using YOLO V5. The models are validated using mean average precision and F1 score. Mean average precision of YOLO V8 is 0.56. Whereas the F1 Score is approximately 0.36.

## References


1. Yuan, Z., Nag, R., Cummins, E.: Human health concerns regarding microplastics in the aquatic environment - from marine to food systems. Sci. Total. Environ. **823**, 153730 (2022). https://doi.org/10.1016/j.scitotenv.2022.153730
2. Cao, Y., Wang, H.: Object detection: algorithms and prospects. In: 2022 International Conference on Data Analytics, Computing and Artificial Intelligence (ICDACAI), Zakopane, Poland, pp. 1–4 (2022). https://doi.org/10.1109/ICDACAI57211.2022.00031
3. Diwan, T., Anirudh, G., Tembhurne, J.V.: Object detection using YOLO: challenges, architectural successors, datasets and applications. Multimed Tools Appl **82**, 9243–9275 (2023). https://doi.org/10.1007/s11042-022-13644-y
4. Walden, K., Mehrubeoglu, M.: Quantifying plastic bottle debris in waterways using image processing (2020). In: Proceedings of the 2020 IEEE 7th International Conference on Engineering Technologies and Applied Sciences (ICETAS), Kuala Lumpur, Malaysia, pp. 1658–1663 (2020). https://doi.org/10.1109/CSCI51800.2020.00305
5. Dhokley, W., Khambati, H., Kapadia, K., Dubey, A.: Identification of products in plastic waste using object detection. Int J Creative Res Thoughts **8**, 6 (2020). ISSN 2320–2882
6. Cunanan, C.F.: PET-Bottle-recognizer: a machine vision recognition of polyethylene-terephthalate based-bottle for plastic waste classification and recycling. In: Proceedings of the 2020 IEEE 7th International Conference on Engineering Technologies and Applied Sciences (ICETAS), Kuala Lumpur, Malaysia, pp. 1–5 (2020). https://doi.org/10.1109/ICETAS51660.2020.9484217
7. Xiao, J., Tang, Y., Zhao, Y., Yan, Y.: Design of plastic bottle image recognition system based on improved YOLOv3. In: Proceedings of the 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), Harbin, China, pp. 2047–2050 (2020). https://doi.org/10.1109/ICMCCE51767.2020.00445
8. Chen, K., An, J., Fang, Y., Bu, T.: Research on solid waste plastic bottle cognitive based on YOLOv5s and deep stochastic configuration network. In: Proceedings of the 2022 7th International Conference on Automation, Control and Robotics Engineering (CACRE), Xi'an, China, pp. 275–280 (2022). https://doi.org/10.1109/CACRE54574.2022.9834213
9. Sio, G.A., Guantero, D., Villaverde, J.: Plastic waste detection on rivers using YOLOv5 algorithm. In: Proceedings of the 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, pp. 1–6 (2022). https://doi.org/10.1109/ICCCNT54827.2022.9984439
10. Shi, J., Zhou, Y., Zhang, W.X.Q.: Target detection based on improved mask rcnn in service robot. In: Proceedings of the 2019 Chinese Control Conference (CCC), Guangzhou, China, pp. 8519–8524 (2019). https://doi.org/10.23919/ChiCC.2019.8866278

11. Sah, S.: Plastic Bottles in the wild Image Dataset, Version 1. Retrieved 28 March 2023 (2023). https://www.kaggle.com/general/46091
12. Ullah, A., Xie, H., Farooq, M.O., Sun, Z.: Pedestrian detection in infrared images using fast RCNN. In: Proceedings of the 2018 Eighth International Conference on Image Processing Theory, Tools and Applications (IPTA), Xi'an, China, pp. 1–6 (2018). https://doi.org/10.1109/IPTA.2018.8608121
13. Wang, D., Wang, L., Peng, D., Qi, E.: Research on appearance defect detection of power equipment based on improved faster-RCNN. In: Proceedings of the 2021 6th International Conference on Power and Renewable Energy (ICPRE), Shanghai, China, pp. 290–295 (2021). https://doi.org/10.1109/ICPRE52634.2021.9635270
14. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 779–788 (2016). https://doi.org/10.1109/CVPR.2016.91
15. Bhuvaneshwary, N., Jayameenakshi, M., Lakshmi, S.A., Shrivalli, K.: People detection and identification using yolo. In: Proceedings of the 2021 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT), Mysuru, India, pp. 85–87 (2021). https://doi.org/10.1109/ICEECCOT52851.2021.9708016
16. Bi, F., Yang, J.: Target detection system design and FPGA implementation based on YOLO v2 algorithm. In: Proceedings of the 2019 3rd International Conference on Imaging, Signal Processing and Communication (ICISPC), Singapore, pp. 10–14 (2019). https://doi.org/10.1109/ICISPC.2019.8935783
17. Ting, L., Baijun, Z., Yongsheng, Z., Shun, Y.: Ship detection algorithm based on improved YOLO V5. In: Proceedings of the 2021 6th International Conference on Automation, Control and Robotics Engineering (CACRE), Dalian, China, pp. 483–487 (2021). https://doi.org/10.1109/CACRE52464.2021.9501331
18. Balaji, R., Prabaharan, G., Singh, A.R., Athisayamani, S., Sarveshwaran, V., Dani-ya, S.: Multi-scale features fusion with YOLOv3 for detecting small and fine tumors in MRI images. In: Proceedings of the 6th International Conference on Electronics, Communication, and Aerospace Technology (ICECA 2022), Coimbatore, India, pp. 1545–1549 (2022). https://doi.org/10.1109/ICECA55336.2022.10009122
19. Kim J.H., Kim N., Won C.S.: High-speed drone detection based on Yolo-V8. In: Proceedings of the ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, pp. 1–2 (2023). https://doi.org/10.1109/ICASSP49357.2023.10095516