

YOLObin: Non-Decomposable Garbage Identification and Classification Based on YOLOv7

Md. Rakib Javed, Md. Shamsuzzaman

Department of Computer Science and Engineering, Begum Rokey University, Rangpur, Bangladesh
Email: rakibjaved58@gmail.com, szaman@brur.ac.bd

How to cite this paper: Javed, Md.R. and Shamsuzzaman, Md. (2022) YOLObin: Non-Decomposable Garbage Identification and Classification Based on YOLOv7. *Journal of Computer and Communications*, 10, 104-121.

<https://doi.org/10.4236/jcc.2022.1010008>

Received: September 18, 2022

Accepted: October 25, 2022

Published: October 28, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In order to safeguard the biological framework on which human creatures depend for presence and make society feasible improvement, it is becoming increasingly important to classify garbage. In any case, individuals are not recognizable with the classification strategy, so it is troublesome for individuals to accurately get the classification of each kind of rubbish. A proper waste management system is a primary task in building a smart and healthy city. In arrange to direct individuals to classify garbage accurately, this paper proposes a strategy of rubbish classification and acknowledgment based on YOLOv7 which is a cutting edge real time object detector. Performance of this model is compared along with two other object detectors where Mask-RCNN achieved f-measurement of 85%, YOLOv5 achieved f-measurement of 95.1% and YOLOv7 achieved f-measurement 95.9%. We have used non-decomposable multiclass garbage images which entails messy backgrounds with unwanted images as well. Four classes of non-decomposable garbage data namely chips packet, plastic bottle, polythene and image with multiclass garbage with 1000 images are prepared for our dataset. Our experimental models performed well in classifying garbage images with cluttered backgrounds. We compared our test results to previous studies in which the majority of the models were tested and trained using laboratory images. The test comes about illustrates that the classification framework features a sensible degree of accuracy and the segmentation recognition impact is way better within the case of point-by-point picture, which can proficiently and helpfully total the rubbish classification errand.

Keywords

Mask-RCNN, YOLOv5, YOLOv7, Garbage Classification, Machine Learning

1. Introduction

Waste materials that are thrown by people under various names-garbage, trash, junk or refuse usually reflect a perception of their lack of utility [1]. An astounding 2.12 billion tons of garbage are dumped worldwide each year. This enormous amount of garbage has caused severe environmental pollution and resource waste. The “classification” and “recycling” of solid waste is one of fundamental strategy for solving the garbage issue. A new type of cyclical economy that promotes sustainable growth and improves environmental quality has inspired an increasing number of countries to investigate recycling solutions in recent years [2]. Decomposable and non-decomposable wastes are the two main categories of waste. Non-decomposable garbage merely keeps piling up in landfills, which leads to hazardous situations. All the landfills around the world are reaching capacity [3]. Moreover, people discard used plastic bags, chips packets, plastic bottles, and other debris rather than placing them in the trash after use. This garbage travels through sewers to rivers and the ocean from there. Trees can't naturally grow there because of garbage that becomes buried in the soil [4].

The focus of this paper is on the detection and classification of garbage that has been thrown away eventually and non-decomposable. Garbage can be divided into multiple categories using a sophisticated algorithm for each disposal method. The labor load of municipal workers can be reduced and health concerns can be avoided with the help of a smart garbage classification system. In our surrounding environment, 5.8% of garbage is typically made up of metals, 3.5% of glass, 1.6% of plastic, 12.9% of papers, 1.8% of textiles, and 53.7% of it is biodegradable, which indicates that only 20.7% of the garbage should actually be dumped in landfills [4] [5]. Plastic bottles are most harmful of these wastes, because they are used to package liquids which people dumps a lot. A significant field of research involves the separation or classification of plastic bottles. A smart system can be developed to classify the garbage into different categories for separate disposal technique. Smart garbage classification system is very beneficial to lessen the work load of the municipal workers and preventing them from health risks.

The existing system for classifying garbage primarily uses human sorting after bin-level sorting, which is risky, time-consuming, and ineffective. As a result, there is a great need for a garbage categorization system to help individuals sort out the garbage. Deep learning's recent advancements have made it possible for numerous valuable computer vision applications, like object detection and classification, to be used. However, due to a number of challenges, it is still a challenging task. For example, gathering and annotating sufficient data for training may be rather expensive. The performance of the standard visual recognition model is lowered by the noise-filled garbage images, which make human labeling challenging [6]. The number of garbage categories or new classes are being produced each year. So, the publicly accessible garbage datasets size is also increasing day by day, as new products with different appearances come into existence

continuously.

This paper proposes a significant garbage classification and detection system based on a transfer learning approach and pre-trained models to achieve the best result of challenging classification using Mask-RCNN, YOLOv5, and YOLOv7 algorithms. Models are trained and tested with our self-prepared garbage dataset of captured images of on campus garbage data in Begum Rokeya University, Rangpur, Bangladesh. YOLOv7 (released on 06th July, 2022) is a state of the art model in detecting real time object. Our paper examines its performance in garbage detection and classification alongside other YOLO models.

The rest of this paper is organized as follows: A detailed analysis and discussion on the previous research works on garbage classification is made in Section 2 of this paper. Research methodology along with a recap of Network models, dataset preparation, augmentation, ground truth generation of the data adaptive with our training models is presented in Section 3. In Section 4, the experiment is done by training the network models and hyperparameter tuning accomplished with validation process. The overall performance of the algorithms is evaluated based on the results of the training and testing processes, as seen in the Result and discussion Section 5. Section 6 concludes the research work, with future remarks on better garbage classification models toward a healthy and clean residential environment.

2. Literature Review

Tremendous research is carrying on for the purpose of building a smart waste management system. Thus, a proper garbage classification system using Machine Learning is a trending research work in search of a sophisticated garbage detection model or framework.

Garbage production rate in urban area is high in comparison with rural area. An intelligent garbage detection approach was proposed in [7] with the dataset of urban area garbage. Faster-RCNN and Res-Net are two models that were used in their train and detection process. They fusion non-garbage urban images like roads, buildings and other images along with garbage data considering the diversity of urban areas to better recognition of garbage. The model can only detect an image whether garbage is present or not but can't classify garbage.

A mobile trash collector was developed in [8] with CNN, deep network architecture which can only collect trash in bin with its robotic arms. With their self-prepared garbage dataset, the trained model has a better prediction which was 0.96 for two classes (trash or not trash) of dataset.

An application of image recognition technology in garbage classification was developed for education platform. For garbage classification the deep learning image recognition algorithm inception v3 was used for this project. Based on the cultivation of children's awareness of garbage classification they design an education platform for children using image recognition technology [9].

An algorithm name PublicGarbageNet was developed by using CNN archi-

ture. This algorithm is a multitask classification algorithm in which one task identifies four major categories of garbage and the other task achieves recognition of ten subclass garbage. The accuracy of this model reaches 96.35% [10].

A garbage recognition and classification system was developed based on convolutional neural network VGG16. This project classifies the domestic garbage in four categories, recyclable garbage, hazardous garbage, kitchen waste and other garbage. After actual tests, the correct classification rate is 75.6% [11].

A garbage object recognition and classification system was developed based on Mask Scoring RCNN. Using Mask-Scoring-RCNN they classify four types of garbage, hazardous garbage, kitchen garbage, recyclable garbage and other garbage. The average accuracy of mask scoring RCNN in garbage classification is 65.8% [12].

To reduce the pollution and the burden of people sorting garbage, a method of garbage recognition and classification was proposed based on transfer learning, which migrates the existing inception v3 model recognition task on the ImageNet dataset to garbage identification. The test accuracy of garbage classification was 93.2% [13].

A novel intelligent garbage classification system based on deep learning and an embedded linux system is proposed for garbage classification. A Gnet model and improved mobilenetv3 is used for classification based on transfer learning. The accuracy of the proposed system was 92.62% [14].

An automatic garbage detection and collection system was proposed for detect and collect garbage. This project was developed by using CNN. They propose a fully automatic system which detects and collects the garbage. The detection accuracy was 90% [15].

To collect outdoor garbage an autonomous trash collecting robot was developed by using YOLOv4-Tiny. For garbage detection the experimented algorithm was Mask-RCNN, YOLOv4 and YOLOv4-Tiny. The calculation rate of Average Precision was 83% for Mask-RCNN, 97.1% for YOLOv4 and 95.2% for YOLOv4-tiny [16].

In this research, we are going to use the very recently released YOLOv7 training model with our self-prepared dataset containing 1000 images of multiclass garbage images, captured in a real environment with messy background. The recently released YOLO model's performance is observed with our real garbage image datasets. YOLOv7 model is not tested yet with garbage dataset. Hence, most of the previously research work done with garbage image dataset which were collected in laboratory environment. But, in real world garbage image contains blurry, noisy image with messy background with complex multiclass garbage. Moreover, in the literature we reviewed, most of the papers basically did binary classification of whether the class is trash or not trash. Thus, in our research we have used image datasets that were collected from real world garbage in our University campus area and used those for both training and testing. An automated garbage classifier will always capture the real world garbage data with

messy, complex backgrounds, not just laboratory processed images with static solid color backgrounds. Hence, in our research work we have captured the garbage image from real messy garbage dumped areas and have chosen the most sophisticated object detector for the classification task and achieved better performance compared with the prior works, as shown in **Table 1**.

Table 1. An Overview of the prior research works.

Network Model	dataset	Accuracy (%)	Limitations
[7] Faster RCNN, 2018	Urban garbage from sanitation dpt.	89	Garbage or Not; binary classifications
[8] CNN, 2019	Self-prepared dataset	96	Trash or not trash; can't classify the garbage
[9] Inception V3, 2019	Three classes of image from internet	-	Raspberry Pi application but did not clear about their classification task
[10] Public Garbage Net, 2020	Self-prepared data	96.35	Only single class or object in an image; captured data in laboratory with static gray background
[11] VGG16, 2020	Web crawled	75.6	Web crawled laboratory images with white clear background
[12] Mask-RCNN, 2020	Beijing Municipal Garbage data	65.8	Segmentation recognition is worse with complex or similar background color as image color
[13] Inception V3, 2020	Crawled from web	93.2	Crawled web data of some garbage-like object but seems not real garbage data. Poor recognition effect on residual waste and hazardous waste
[14] Gnet, 2021; Improved MobilNetv3	Huawei Garbage Classification Challenge Cup dataset	92.62	40 class of garbage data but it seems a typical object detection; images are not like typical garbage as their background says
[15] CNN,	-	90	Raspberry model for valuables and garbage with robotic arm. Can't classify to specific garbage
[16] Mask-RCNN, YOLOv4, YOLOv4-tiny, 2021	TACO open dataset	83, 97.1, 95.2	Arduino based trash collection but can't categories or classify the collected trash

3. Materials and Methodology

3.1. Transfer Learning

Deep neural networks require massive amounts of labeled data, which are typically unavailable for certain applications, such as garbage classification. The most popular approach is to use transfer learning, which trains the CNN model using open datasets and then fine-tunes the classifier using particular data. Pre-training is a crucial component of transfer learning, and it can be accomplished by using standard supervised learning, weakly-supervised learning, and unsupervised learning [17] to help the model develop a strong feature extractor. It should be noted that pretraining for the target task should be done in a wider domain just in case “negative transfer” does not occur. The domain adaptation approach, which addresses the issue of insufficient data for visual recognition and ubiquitous computing, assists in aligning the marginal distribution among domains if there is a domain shift during the transfer process [18]. In our research, we used pre-trained Mask-RCNN, YOLOv5, and YOLOv7 models on our self-prepared garbage dataset.

3.2. Network Revisited

3.2.1. MASK-RCNN

Mask R-CNN or regional convolutional neural network is a deep neural network which was developed as a solution to the instance segmentation problem [19]. To put it another way, Mask R-CNN can distinguish between various objects in an image or video. If the input of Mask RCNN model is an image, it returns the object’s bounding boxes, classes, and masks.

The color RGB images from the garbage datasets were processed by the pipeline illustrated in **Figure 1**. The feature extraction process from the captured image was carried out using the convolution layers. The feature map acquired from the convolutional layer was then transferred to the Regional Proposal Network. Region of Interests or 5 ROIs was then generated from the RPN. The features extracted from each ROI were chosen by the Roi Align layers. The Roi align layers passed the selected features to the fully connected layers. In this

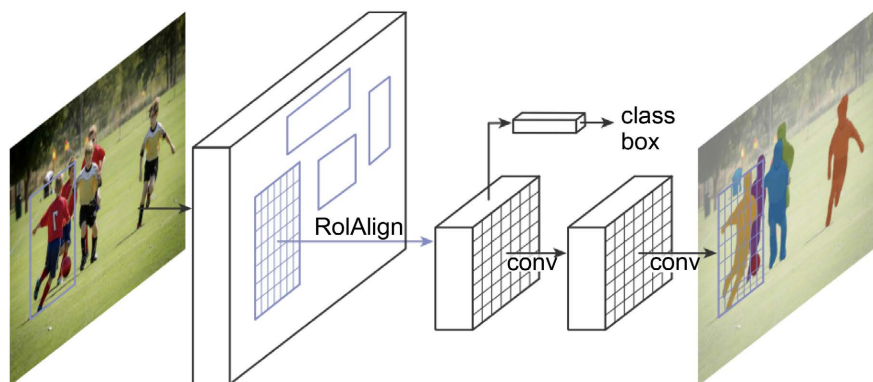


Figure 1. MASK-RCNN architecture.

layer, classification, mask and bounding-box prediction occurred. Image annotation for the training, validation and test sets were performed manually.

3.2.2. YOLOv5

Glenn Jocher released YOLOv5 on June 26, 2020 [20]. YOLO is a real time object detection model series. YOLOv5 is a single stage object detector. It has three important part like other object detectors. They are Model Backbone, Model Neck and Model Head. With the help of 58 open-source contributors, YOLOv5 has set a new standard for object detection models. It already outperforms the Efficient Det and previous YOLO versions. YOLOv5 used coco dataset for training. Cocoval 2017 dataset was used for performance evaluation over various inference sizes from 256 to 1536. GPU speed is measured using an AWS p3.2xlarge V100 instance with a batch size of 32 with the COCO val 2017 dataset to determine the average inference time for each image [21].

3.2.3. YOLOv7

YOLOv7 was released in July 7, 2022 by Wong Kin Yiu. It is state-of-the-art and the most recent version of the YOLO (you only look once). The paper claims that it is the current fastest and most accurate real-time object detector. There are three primary parts to the YOLO framework: Backbone, Head and Neck. The main function of the backbone is to extract important information from an image and send it through the neck to the head. The neck compiles feature maps that the backbone extracted and creates feature pyramids. The head's output layers with final detections make up the last part of the structure. YOLOv7 is not restricted to a single head. The lead head is in charge of the final output, while the auxiliary head is in charge of assisting training in the middle layers. In addition, to improve deep network training, a label assigner mechanism was introduced, which takes into account network prediction results as well as ground truth and then assigns soft labels as shown in **Figure 2**. In contrast to traditional label assignment, which relies solely on the ground truth to generate hard labels based on given rules, reliable soft labels employ calculation and optimization methods that take into account the quality and distribution of prediction output in addition to the ground truth. The YOLOv4, scaled YOLOv4 and YOLO-R are the ancestor of YOLOv7. Among all real-time object detectors with 30 FPS or more on GPU V100 in the range of 5 to 160 FPS, Yolov7 has the highest accuracy (56.8% AP), according to Paper. YOLOv7 was trained entirely from scratch using the MS COCO dataset without the use of any pre-trained weights [22].

3.3. Dataset Preparation

3.3.1. Data Collection

Multiclass garbage data was captured and collected from campus area in Begum Rokeya University, Rangpur, Bangladesh. Mostly three categories of non decomposable garbage data: Chips Packet, Polythene and Plastic Bottle found while we were conducting a survey in the university campus. Initially a total number of 164 noise free, fine images were selected among our captured images. Next, data

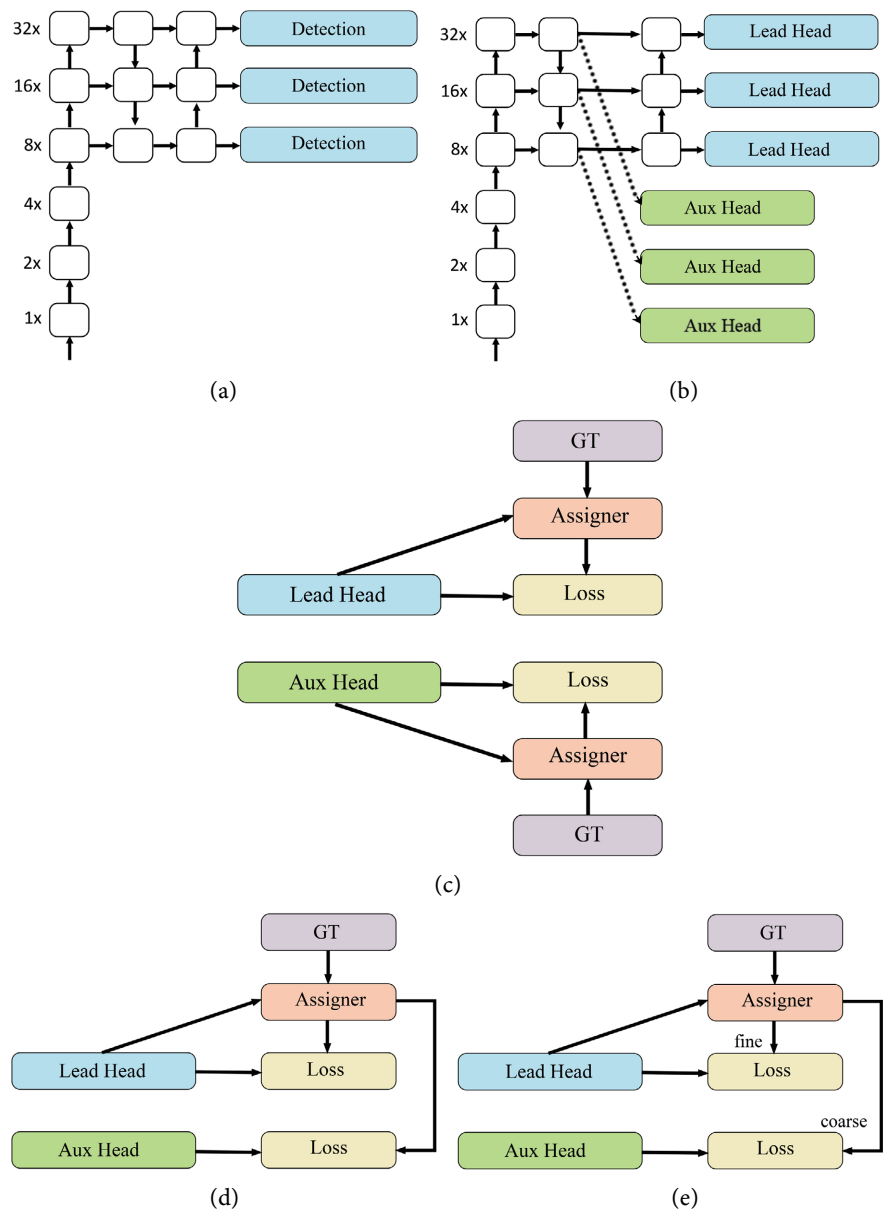


Figure 2. YOLOv7 architecture's auxiliary head and lead head guided label assigner. (a) Normal model; (b) Model with auxiliary head; (c) Independent Assigner; (d) Lead guided assigner; (e) Coarse-to-fine lead guided assigner.

augmentation was done to increase the dataset sizes with larger number of images. The device used for data collection was Redmi Note 10 Pro with 64 MP camera resolution and captured image size was 4640×3472 . Sample data of different classes are given in **Figure 3**.

3.3.2. Data Preprocessing

The captured image size was extremely large. So, the image needed to resize for avoid high computation of memory. The garbage data was resized by using adobe photoshop script. Using adobe photoshop the main captured image was resized by 1024×766 . **Figure 4** shows the resized image.



Figure 3. Sample garbage data.



Figure 4. (a) Captured and (b) Resized Image. (a) 4640×3472 , (b) 1027×766 .

3.3.3. Data Augmentation

Data augmentation is an excellent approach to enlarge the dataset and make the dataset as diversified as feasible. By utilizing data augmentation, it is possible to significantly lessen network overfitting and increase the generalizability of the training model [23]. The generalization of the image classification model was improved with the use of data augmentation. The images are often either randomly translated by a few pixels or turned horizontally during the data augmentation phase. Total augmented images of different classes are shown in **Table 2**.

3.3.4. Ground Truth Generation

Data annotation is the process of labeling the data to show the outcome for ma-

chine learning model to predict. For image annotation, VIA (VGG image annotator) tool was used which saved the annotations in a JSON file [24]. During annotation for Mask R-CNN, a polygon bounding box was used for object detection. To annotate photos for YOLO, we used Piotr Skalski's Make-Sense method. A square bounding box was used for YOLO series [25] as shown in **Figure 5**.

4. Experiment

The whole training process was carried out in Google Collaboratory Pro with a NVIDIA P100 25 Gb graphics processing unit. To train our models, 800 images were used for training and 200 images were used for validation. The ratio of training and validation dataset were 0.8 and 0.2.

Training and Hyperparameter Tuning

Transfer learning was used as a training and validation technique. It allows for the customization of a model that has already been trained on a sizable image classification dataset, such as ImageNet or the MS COCO (Microsoft Common Objects in Context) dataset, to carry out a specific task. The MS COCO dataset comprises 330 k images [26], whereas the ImageNet collection has a massive 14.1 million plus images, estimated [27].

In this study, weights of pre-trained models on the MS COCO dataset were used for Mask-RCNN. Pretrained YOLOv5 medium, and YOLOv7 weights were

Table 2. Augmented data.

Class	Number of Images	Total Image
Chips packet	250	1000
Polythene	250	
Plastic bottle	250	
Multi-class garbage	250	

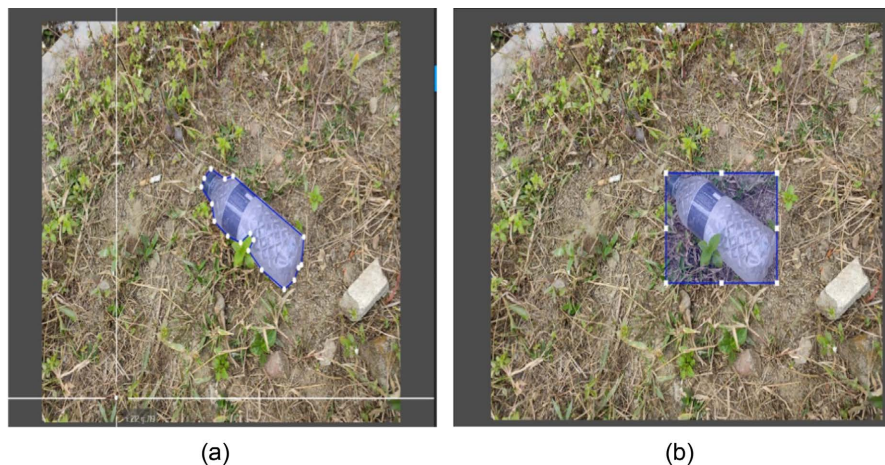


Figure 5. Image annotation for (a) Mask-RCNN and (b) YOLO models.

used for train YOLOv5, and YOLOv7. The top few layers, or the head, of the pretrained model's convolutional neural network were trained on the "garbage" dataset with the intention of achieving the specified classification, while the base of the network was frozen to be utilized as a feature extractor. This new training dataset for object recognition had a low learning rate at 0.001 for Mask-RCNN, 0.01 for YOLOv5 and 0.01 for the YOLOv7 algorithm initially. This was done to minimize undesirable divergent behavior in the loss function by making modest adjustments to the weights during the training. Since the model was not trained entirely from scratch here, but rather made use of certain pretrained weights, transfer learning significantly boosted the model's efficacy by resulting in far more accurate outputs. This method also helped to get over the problem caused by the tiny dataset. The average precision (AP), which for all three models was 0.5% or 50%, was determined using the intersection over union (IoU) threshold. 200 training epochs were used to train the Mask-RCNN model, with 10 training steps each epoch. YOLOv5 model was trained for 200 epochs with a batch size of 16 and input image size for YOLOv5 was 416×416 . The YOLOv7 model was trained for 200 epoch and the batch size was set by 16 and input image size was 640×640 . The training and validation batch is shown in **Figure 6** and **Figure 7**. Those we observed in our YOLO algorithms training and validation stage. We

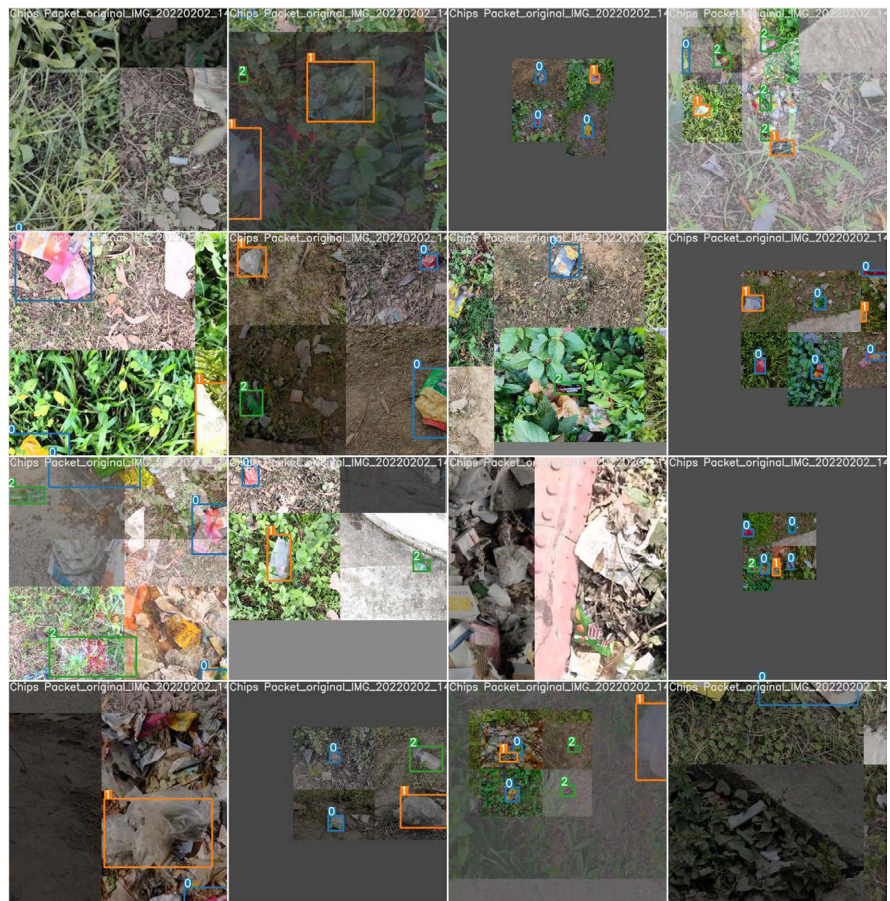


Figure 6. Training batch for Yolo algorithm.

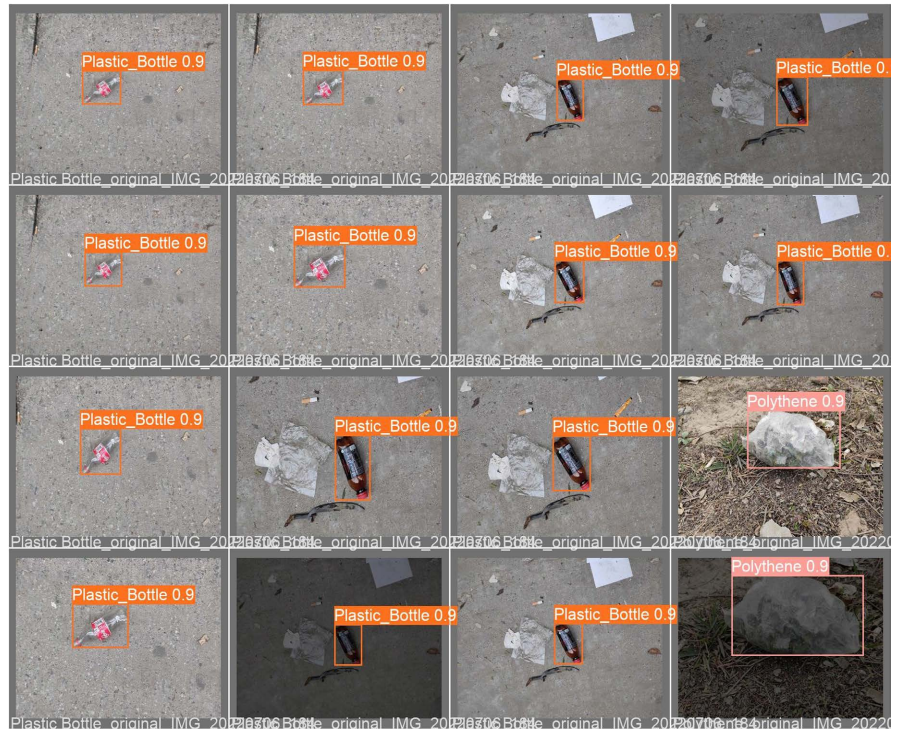


Figure 7. Validation batch for Yolo algorithm.

can see there are images with messy backgrounds and in some cases specifically for polythene the object color is the same as the background but, we were able to successfully identify and classify that garbage. On the other hand, multiple objects other than the garbage is also present in the images. Those also identified and also with multiclass garbage in a single image were perfectly identified by the detector.

5. Result and Discussion

5.1. Performance Evaluation

Precision, Recall and F1 score are performance evaluation metrics used to identify the overall accuracy of training models for some datasets.

Precision measures the number of positive predictions that are correct (TP) as shown in Equation (1). Recall (also called sensitivity) measures the number of positive cases that a model is able to predict correctly among all positive cases in dataset which can be calculated using Equation (2). F1 score tells us the harmonic mean on precision and recall in Equation (3). F1 score has some advantages over precision and recall. In case of very small values of precision and recall, it balances the metrics.

$$\text{Precision} = \frac{Tp}{Tp + Fp} \quad (1)$$

$$\text{Recall} = \frac{Tp}{Tp + FN} \quad (2)$$

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

Here, true positive rate is represented by TP. True positive or TP is the total number of garbage detected correctly. False positive rate is represented by FP. False positive or FP is the false predicted garbage value from total garbage value. False negative rate is represented by FN. False negative or FN is the garbage predicted value was falsely predicted. The harmonic mean of recall and precision is used to calculate the F measure, giving each the same weight. It enables evaluation of a model taking into consideration both precision and recall using a single score, which is useful for explaining the performance of the model and when comparing models.

5.2. Result Analysis

The loss function of Mask-RCNN during the train shown in **Figure 8**. The final loss obtained at the end of training for 200 epoch was 0.392, Final mrcnn bounding box loss was 0.1001, mrcnn mask loss was 0.1474 and final class loss was 0.0177. **Table 3** shows a clear outline of precision, recall and f measurement of Mask-RCNN. In Mask-RCNN we got average precision 0.796, average recall 0.913 and the f-measurement was 0.850.

Figure 9 shows the loss function of YOLOv5 model as the training progresses. The box loss after 200 epoch was 0.01402, the object loss was 0.006536 and the class loss was 0.000156. **Figure 10** and **Table 3** show a clear outline of Precision, Recall, F measurement, mAP at IOU 0.5 and IOU 0.95. The precision we get for YOLOv5 is 0.942, recall is 0.961, mAP at 0.5 is 0.972, mAP at 0.95 is 0.688 and the f measurement is 0.951.

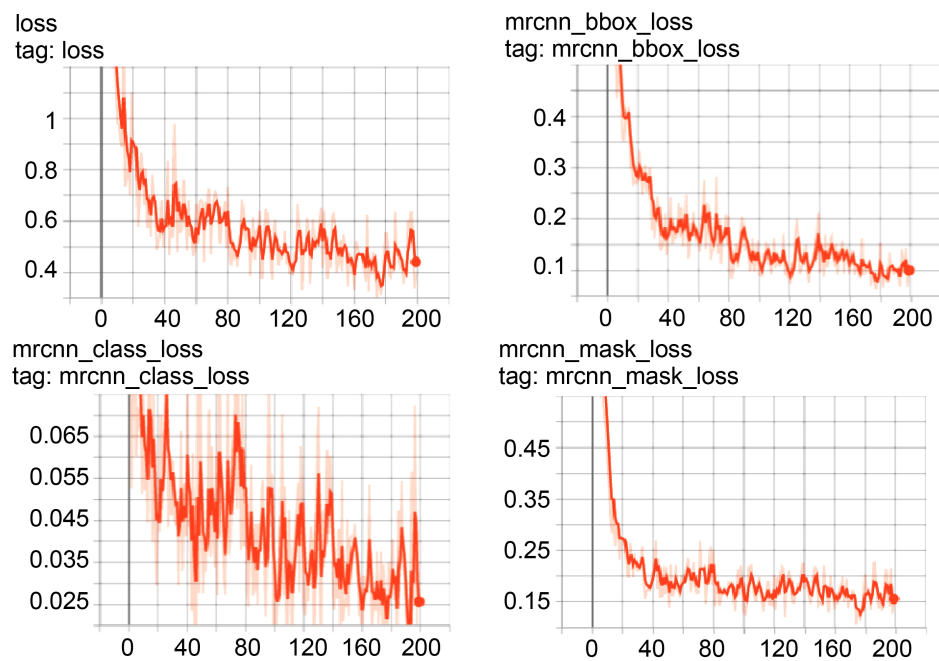


Figure 8. Mask-RCNN loss function in training.

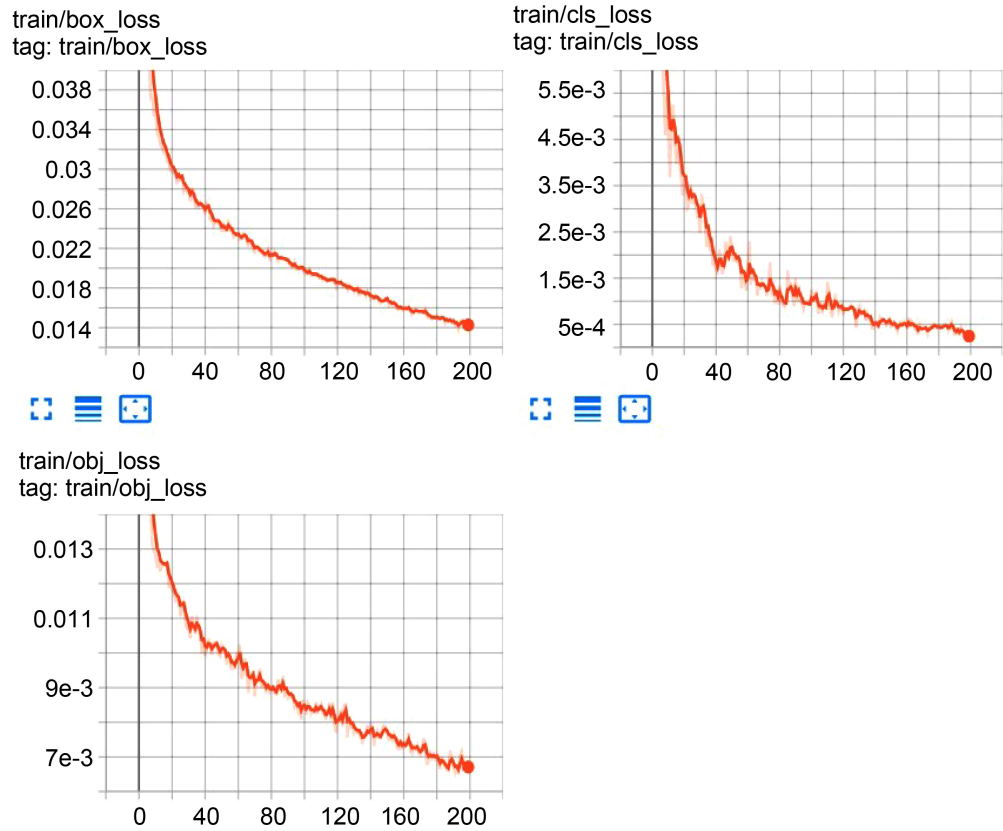


Figure 9. Loss function in training YOLOv5.

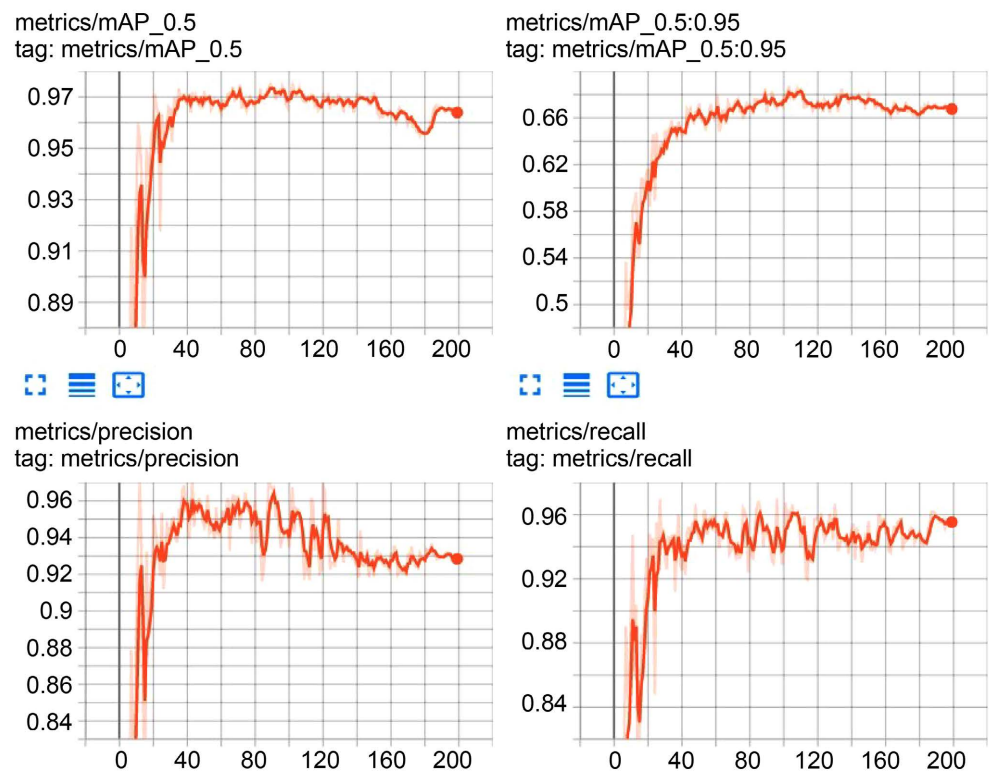


Figure 10. Precision, recall and mAP in YOLOv5.

Table 3. A garbage detection and classification quality metrics of Mask-RCNN, YOLOv5, and YOLOv7.

Models	Number of Images	Precision	Recall	F1 Score	mAP	mAP (0.95)
Mask-RCNN	200	0.796	0.913	0.850	-	-
YOLOv5	200	0.942	0.961	0.951	0.972	0.688
YOLOv7	200	0.972	0.947	0.959	0.963	0.677

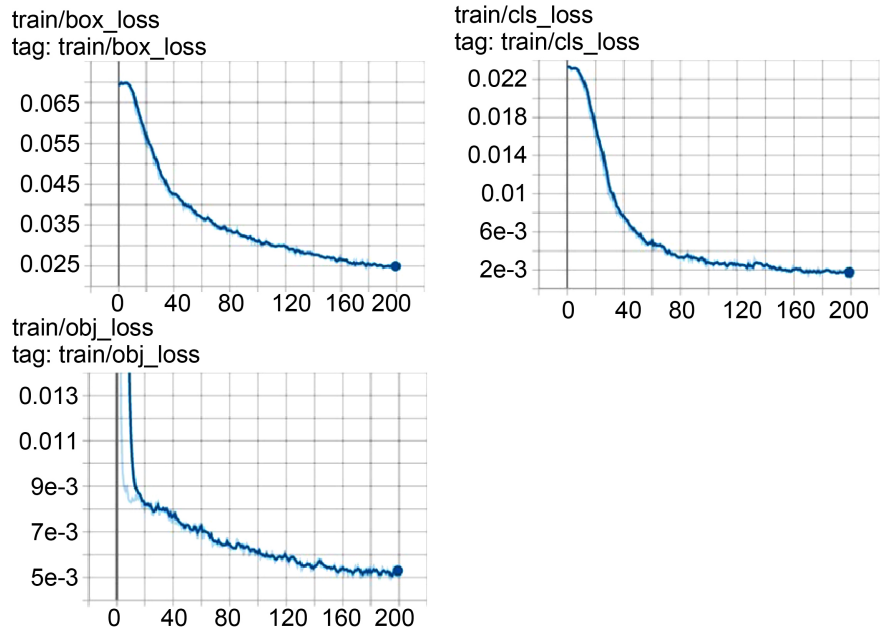


Figure 11. Loss function in training YOLOv7.

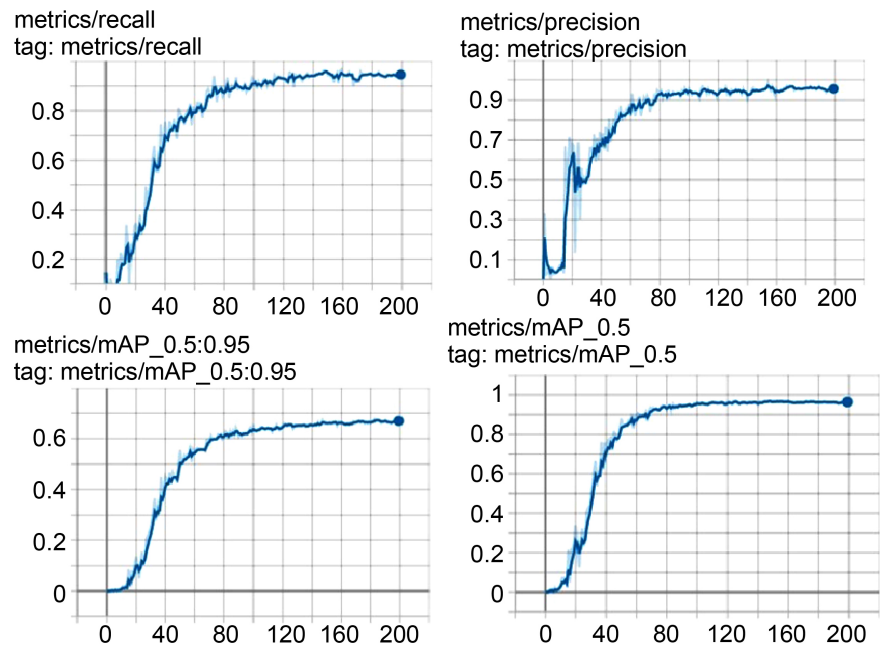


Figure 12. Precision, recall and mAP in YOLOv7.

Figure 11 shows the loss function of YOLOv7 as the training progresses. After 200 epoch the box loss is 0.02503, object loss is 0.005487, the class loss is 0.001726. **Table 3** and **Figure 12** show the clear outline of precision, recall, f measurement and mAP at IOU 0.5 and IOU 0.95. The precision we got 0.972, recall 0.947, f-measurement was 0.959, mAP at 0.5 was 0.963 and at 0.95 was 0.677. **Table 2** shows that YOLO model performs better than Mask-RCNN. Among those three models YOLOv5 gives the best mean average precision in our dataset but YOLOv7 gives the best f1 score which is 0.959. The table tells that the mean average precision of YOLOv5 is 0.972. F-measurement tells us the overall performance of a model. Here, f1 score of YOLOv7 is high.

6. Conclusion

Sorting out garbage and its proper recycling process will lead to build green, clean smart living place. In our research, object identification and classification task are tested with real garbage dataset with messy background images with a state of the art object detection model, YOLOv7 where in most of the previous work used white background images as their dataset. Performance is compared with Mask-RCNN object identifier and other prior YOLO algorithms as well. Among these, YOLOv7 achieved higher F_1 score which is 0.959. As, new category or classes of garbage is booming every year, in this case YOLOv7 can be a better choice to deal with garbage sorting and classification task for a smart garbage management system. Because YOLO algorithm's performance is impressive in object detection especially for real time object. In our work, there were bit a less number of garbage class. In future, we intend to work with large number of classes as new categories of garbage is being produced and dumped each year and build an automated garbage classifier to deal with these ever growing garbage dataset with higher accuracy.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Garbage. <https://en.wikipedia.org/wiki/Garbage>
- [2] Hoornweg, D. and Perinaz, B.T. (2012) What a Waste: A Global Review of Solid Waste Management. World Bank, Washington DC. <http://hdl.handle.net/10986/17388>
- [3] Zhang, D., Keat, T.S. and Gersberg, R.M. (2010) A Comparison of Municipal Solid Waste Management in Berlin and Singapore. *Waste Management*, **30**, 921-923. <https://doi.org/10.1016/j.wasman.2009.11.017>
- [4] Rokade, R., Maurya, A., Khade, V. and Mali, P.J. (2018) Smart Garbage Separation Robot with Image Processing Technique. *International Journal of Engineering Research & Technology (IJERT)*, **6**, 1-4. <https://www.ijert.org/iciate-2018-volume-6-issue-12>

- [5] Saravana Kannan, G., Sasi Kumar, S., Ragavan, R. and Balakrishnan, M. (2016) Automatic Garbage Separation Robot Using Image Processing Technique. *International Journal of Scientific and Research Publications*, **6**, 326-328. <https://www.ijsrp.org/research-paper-0416/ijsrp-p5250.pdf>
- [6] Yang, J., Zeng, Z., Wang, K., Zou, H. and Xie, L. (2021) GarbageNet: A Unified Learning Framework for Robust Garbage Classification. *IEEE Transactions on Artificial Intelligence*, **2**, 372-380. <https://doi.org/10.1109/TAI.2021.3081055>
- [7] Wang, Y. and Zhang, X. (2018) Autonomous Garbage Detection for Intelligent Urban Management. *MATEC Web of Conferences*, **232**, Article No. 01056. <https://doi.org/10.1051/mateconf/201823201056>
- [8] Hossain, S., Debnath, B., Anika, A. and Al-Hossain, M.J. (2019) Autonomous Trash Collector Based on Object Detection Using Deep Neural Networ. *TENCON2019—2019 IEEE Region 10 Conference (TENCON)*, Kochi, 17-20 October 2019. <https://doi.org/10.1109/TENCON.2019.8929270>
- [9] Chen, G., Wang, H. and Zheng, J. (2019) Application of Image Recognition Technology in Garbage Classification Education. *2019 5th International Conference on Control, Automation and Robotics*, Beijing, 19-22 April 2019. <https://doi.org/10.1109/ICCAR.2019.8813481>
- [10] Zeng, M., Lu, X., Xu, W. and Liu, Y. (2020) PublicGarbageNet: A Deep Learning Framework for Public. *2020 39th Chinese Control Conference*, Shenyang, 27-29 July 2020. <https://doi.org/10.23919/CCC50068.2020.9189561>
- [11] Hao, W. (2020) Garbage Recognition and Classification System Based on Convolutional Neural Network VGG16. *2020 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)*, Shenzhen, 24-26 April 2020.
- [12] Li, S., Yan, M. and Xu, J. (2020) Garbage Object Recognition and Classification Based on Mask Scoring RCNN. *2020 International Conference on Culture-Oriented Science & Technology (ICCST)*, Beijing, 28-31 October 2020. <https://doi.org/10.1109/ICCST50977.2020.00016>
- [13] Cao, L. and Xiang, W. (2020) Application of Convolutional Neural Network Base on Transfer Learning for Garbage Classification. *2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, Chongqing, 12-14 June 2020. <https://doi.org/10.1109/ITOEC49072.2020.9141699>
- [14] Fu, B., Li, S., Wei, J., Li, Q., Wang, Q. and Tu, J. (2021) A Novel Intelligent Garbage Classification System Based On Deep Learning and an Embedded Linux System. *IEEE Access*, **9**, 131134-131146. <https://doi.org/10.1109/ACCESS.2021.3114496>
- [15] Bansal, S., Patel, S. and Shah, I. (2019) AGDC: Automatic Garbage Detection and Collection. ArXiv: 1908.05849.
- [16] Kulshreshtha, M., Chandra, S.S., Randhawa, P. and Tsaramiris, G. (2021) OATCR: Outdoor Autonomous Trash-Collecting Robot Design Using YOLOv4-Tiny. *Electronics*, **10**, Article No. 2292. <https://doi.org/10.3390/electronics10182292>
- [17] Pan, S.J. and Yang, Q. (2010) A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, **22**, 1345-1359. <https://doi.org/10.1109/TKDE.2009.191>
- [18] Yang, J., Zou, H., Zhou, Y., Zeng Z. and Xie, L. (2020) Mind the Discriminability: Asymmetric Adversarial Domain Adaptation. *Computer Vision-ECCV 2020: 16th European Conference*, Glasgow, 23-28 August 2020. https://doi.org/10.1007/978-3-030-58586-0_35
- [19] He, K., Gkioxari, G., Dollár, P. and Girshick, R. (2017) Mask R-CNN. 2017 *IEEE*

International Conference on Computer Vision (ICCV), Venice, 22-29 October 2017. <https://doi.org/10.1109/ICCV.2017.322>

- [20] Maithani, M. (2020) Guide to Yolov5 for Real-Time Object Detection. <https://analyticsindiamag.com/yolov5/>
- [21] YOLOv5. <https://github.com/ultralytics/yolov5>
- [22] Wang, C.Y., Bochkovskiy, A. and Liao, H.Y.M. (2022) YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. ArXiv: 2207.02696.
- [23] Dilmegani C. (2022) What Is Data Augmentation? Techniques & Examples & Benefits. AI Multiple. <https://research.aimultiple.com/data-augmentation/>
- [24] Dutta, A. and Zisserman, A. (2019) The VIA Annotation Software for Images, Audio and Video. *MM19: Proceedings of the 27th ACM International Conference on Multimedia*, New York, 21-25 October 2019, 2276-2279. <https://doi.org/10.1145/3343031.3350535>
- [25] Makesense.ai. <https://github.com/SkalskiP/make-sense>
- [26] COCO Dataset. <https://cocodataset.org/#home>
- [27] ImageNet. <https://www.image-net.org/>