A YOLOv7 BASED VISUAL DETECTION OF WASTE

Zhuda Yang

Belarusian State University, 4, Nezavisimosti av., Minsk, 220050, Belarus,
judayeung@gmail.com

With the popularization of environmental awareness, how to correctly and efficiently classify the increasing domestic waste is attracting attention. In this study, the recently released YOLOv7 was used to detect and realize the classification of waste. Use the YOLOv7 model to train the waste data set collected in reality. As the experimental result, the experiment proves that the YOLOv7 model is better than other object detection models.

Keyword: Object Detection; Image-based waste detection; Waste classification; YOLOv7.

Introduce

With the rapid development of the economy and the increase of the population, the domestic garbage in the city is increasing day by day, and garbage classification is the best solution to deal with the pollution of garbage [1]. At present, the way of waste disposal in waste treatment plants is still mainly manual sorting, which has problems such as low efficiency and harsh environment.

In recent years, many researchers have used traditional methods for garbage classification and identification. A hierarchical classification method for the identification of low-density (LDPE) and high-density polyethylene (HDPE) in mixed plastic waste based on short-wave infrared (SWIR) hyperspectral imaging was proposed by Bonifazi et al [2]. Gundupalli et al. proposed a thermal imaging-based recyclable material identification method based on garbage, which successfully classified iron, plastic, paper, aluminum, stainless steel, and wood, among others [3].

With the rapid development of artificial intelligence technology, more and more scholars conduct research in the field of pattern recognition. There are many object detection methods developed. These detection algorithms can be roughly divided into two types. The first is the Two-Stage object detection algorithm. First, the selection network is used to select and suggest the part of the image that may be the target object, and then the detection algorithm is used. The network identifies and classifies objects. Such as R-CNN [4] model, Faster R-CNN [5] model, Mask R-CNN [6] model. The second is the One-stage object detection algorithm, which has a simple structure, strong scalability and wider application range. Such as YOLO [7] model, SSD [8] model.

So far, the most commonly studied and used models by researchers are the YOLO series models, such as YOLOv3, YOLOv4, and YOLOv5. YOLOv7 [9] is the latest work of the YOLO series, in this study, the purpose is to achieve waste classification by using image features of various household waste, so we study the YOLOv7 model and compare the results with experiments with different models.

1. YOLOv7

YOLOv7 further improves detection speed and accuracy on the basis of previous versions, while balancing speed and accuracy. The main optimization directions of current target detection: faster and stronger network architecture; more effective feature integration method; more accurate detection method; more accurate loss function; more effective label assignment method; more effective training method.

YOLOv7 is similar to YOLOv5 as a whole, mainly in the replacement of internal components of the network structure, auxiliary training heads, and label assignment ideas. First, YOLOv7 extends the efficient long-range attention network, called Extended-ELAN (E-ELAN for short). In large-scale ELANs, the network can reach a steady state regardless of the gradient path length and the number of blocks. E-ELAN expands, shuffles, and merges cardinality to the cardinality, which can improve the learning ability of the network without destroying the original gradient path. In terms of structure, as shown in the figure 1, E-ELAN only changes the architecture in the computing block, and does not change the architecture of the transition layer.

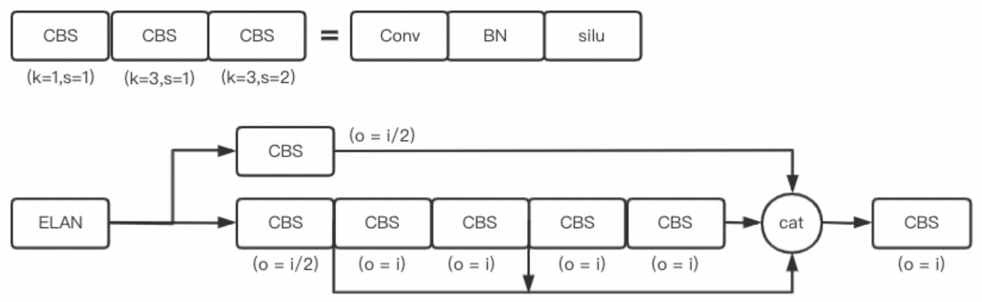

Figure 1. ELAN structure

In the past, in the training of deep networks, label assignments usually refer directly to. the ground truth and generate hard labels (without softmax) according to given rules. However, in recent years, taking object detection as an example, researchers often use the quality distribution predicted by the network to combine ground truth, and use some computational and optimization methods to generate reliable soft labels. For example, YOLO uses bounding box predictions and ground truth IoU as soft labels.

In the label assignment mechanism of YOLOv7, the network prediction results and the benchmark are considered at the same time, and then the soft labels (the labels after comprehensive consideration and optimization) are assigned to the "label assigner" mechanism. This new label assignment method generates coarse-to-fine hierarchical labels based on lead head prediction, which are used for the learning of lead head and auxiliary head, respectively. The specific method is shown in the following figure. In figure 2, by letting the shallow auxiliary head directly learn the information that the guide head has learned, the guide head can focus more on the residual information that has not yet been learned. Figure 3 The Coarse-to-fine guide head uses its own prediction and ground truth to generate soft labels and guide labels for assignment.

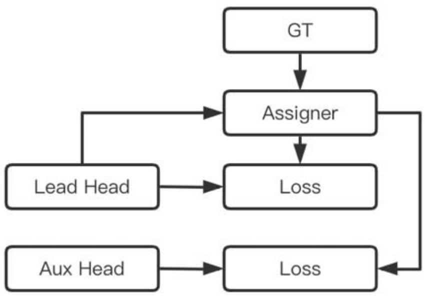

Figure 2. Lead guided assigner

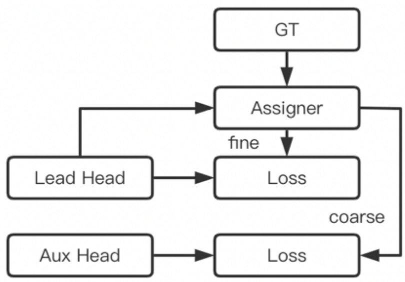

Figure 3. Coarse-to-fine lead guided assign

2. Experimental results

In this study, the YOLOv7 model was used to detect and classify domestic waste. To this end, this study takes common household waste as an example to conduct algorithm research, and collects common waste data, including peels, batteries, bottles, cigarettes, glass, tiles, etc. The dataset consists of 2400 pieces of data with different shooting angles, different scenes, different degrees of decay, etc., of which 70% are used for training and 30% are used for testing. The test output plot of the YOLOv7 model is shown in Figure 4.

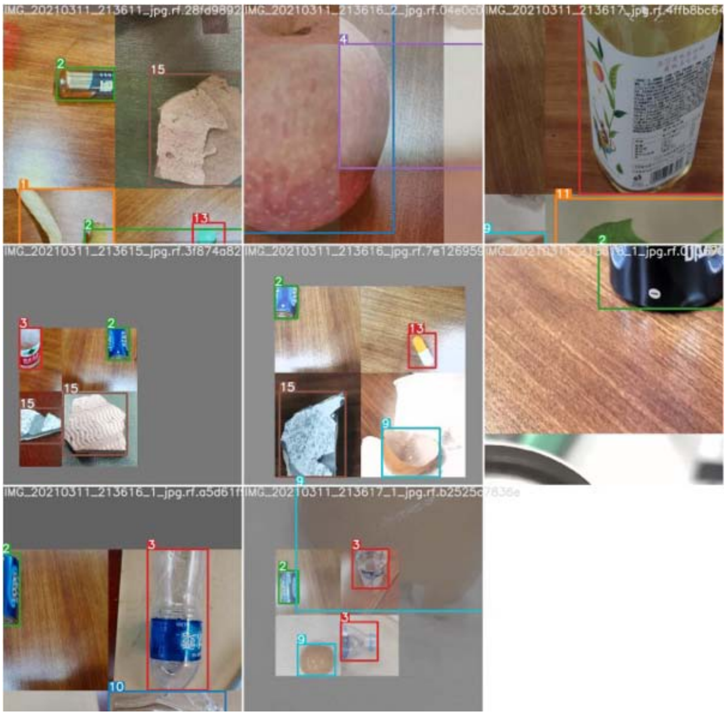

Figure 4. Test result of YOLOv7 on waste dataset

Figure 5 gives the mean precision (mAP), precision, and recall values of the YOLOv7. model.

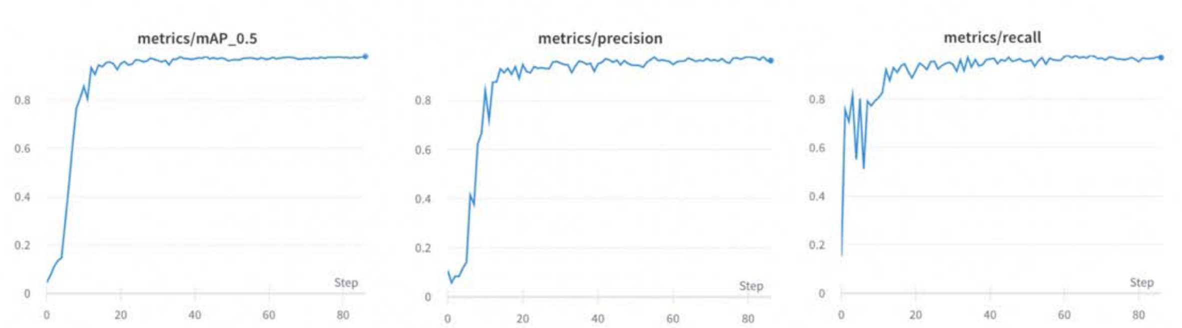

Figure 5. mAP, Precision, Recall values of the YOLOv7 model

In the literature, YOLOv7 based methods have not been used for waste detection. The mAPs of YOLOv7 and other models in this experiment are shown in Figure 6.

As can be seen from the figure, the YOLOv7 model gives better results compared to. other models. The effect of YOLOv7 model on waste recognition is shown in the Figure 7.

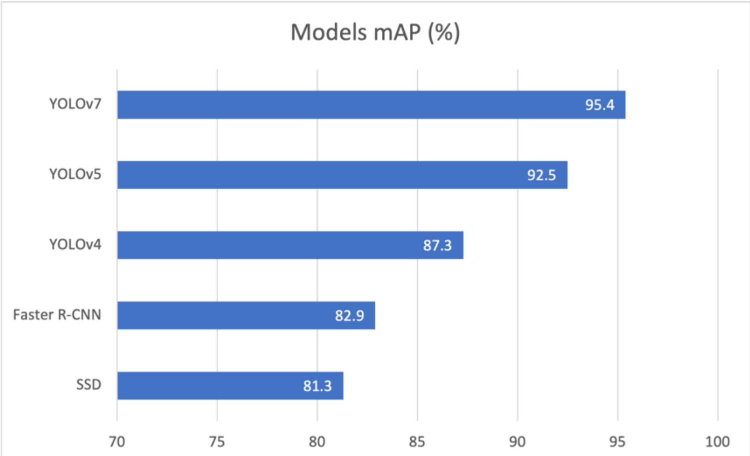

Figure 6. mAP of Models

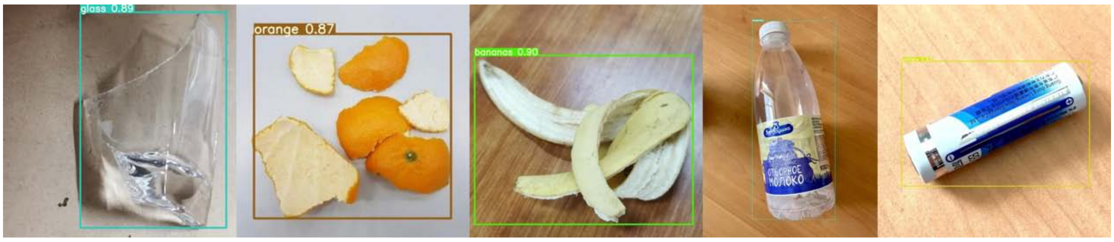

Figure 7. Model recognition effect

Conclusion

This paper studies a visual detection system based on the YOLOv7 model for domestic waste classification, predicts the input image, and completes the classification of domestic waste at high speed and accurately. The experimental results show that the system can basically meet the actual needs.

References

1. Wei Xiaoxiao, Wang Xiaoming, Li Lei, et al. Temporal and spatial characteristics of municipal solid waste generation and treatment in China from 1979 to 2016[J] // China Environmental Science. 2018. № 38(10). P. 3833–3843.
2. Bonifazi G., Capobianco G., Serranti S. A hierarchical classification approach for recognition of low-density (LDPE) and high-density polyethylene (HDPE) in mixed plastic waste based on short-wave infrared (SWIR) hyperspeRWCMal imaging[J]. SpeRWCMochimica Acta Part A Molecular & Biomolecular SpeRWCMoscopy, 2018: S1386142518301975.
3. A S P G , B S H , A A T . Multi-material classification of dry recyclables from municipal solid waste based on thermal imaging[J] // Waste Management. 2017. № 70. P. 13–21.
4. Girshick, R., Donahue, J., Darrell, T., & Malik, J. Rich feature hierarchies for accurate

object detection and semantic segmentation // Proceedings of the IEEE conference on computer vision and pattern recognition. 2014. P. 580–587.

5. Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J] // IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017. № 39(6). P. 1137–1149.
6. He, K., Gkioxari, G., Dollár, P., & Girshick, R. Mask r-cnn // In Proceedings of the IEEE international conference on computer vision. 2017. P. 2961–2969.
7. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. You only look once: Unified, real-time object detection // In Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. P. 779–788.
8. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). Ssd: Single shot multibox detector. In European conference on computer vision (p. 21–37). Springer, Cham.
9. Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. ”YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real- time object detectors”. 2022, arXiv preprint arXiv:2207.02696.