Automated Debris Detection System Based on Computer Vision

Nur Athirah Zailan[1], Mohamad Haniff Junos[2], Khairunnisa Hasikin[3*], Anis Salwa Mohd Khairuddin[1], Uswah Khairuddin[4]

[1,2]Department of Electrical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Malaysia
[2]School of Aerospace Engineering, Universiti Sains Malaysia, Engineering Campus, Penang, Malaysia
[3]Department of Biomedical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Malaysia.
[4]Malaysia Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
[*]Corresponding Author
doi: https://doi.org/10.21467/proceedings.141.4

ABSTRACT

Marine litter has been one of the major challenges and a well-known issue across the globe for decades. 6.4 million tonnes of marine debris per year is estimated to enter water environments, with 8 million items entering each day. These statistics are so worrying, and mitigation steps need to be taken for the sake of a sustainable community. The major contributor to marine litter is none other than riverine litter. However, when there is not enough data about the amount of litter being transported, making quantitative data for monitoring impossible. Nowadays, most countries still use visual counting, which limits the feasibility of scaling to long-term monitoring at multiple locations. Therefore, an object detector using one of the deep learning algorithms, You Only Look Once version 4 (YOLOv4), is developed for floating debris of riverine monitoring system to mitigate the problem mentioned earlier. The proposed automated detection method has the capability to detect and categorize riverine litter, which can be improved in terms of detection speed and accuracy using YOLOv4. The detector is trained on five object classes such as styrofoam, plastic bags, plastic bottle, aluminium can and plastic container. Image augmentation technique is implemented into the previous datasets to increase training and validation datasets, which results in the increase of accuracy of the training. Some YOLOv4 and YOLOv4-tiny parameters have also been studied and manipulated to see their effects on the training.

Keywords: computer vision, deep learning, image processing, object detection, smart city

1 Introduction

Debris pollution in aquatic ecosystems has been a major environmental issue across the globe for decades now. This issue is not only affecting the ecology and endangers aquatic species, but it also may lead to bigger destruction in economics. Riverine ecosystems are known to affect marine environments directly because they are the heart of the oceans. Rivers transport land-based plastic waste from every part of the world and will finally get transported across the oceans. 6.4 million tonnes of marine debris per year are estimated to enter the rivers and the oceans, with about 8 million items entering each day [1]. In order to identify the sources of ocean plastic pollution, riverine litter monitoring is very crucial. This includes identifying the transport pathways and understanding the negative impacts they have towards our community before mitigation steps are taken to address the problem.

Riverine monitoring has become an important first step for most countries and an automated system has been demanded to support their efforts. This includes computer vision based for marine debris detection instead of using the conventional way such as manual counting, which can be labor intensive and lack harmonization across survey sites. In this way, information from riverine ecosystems can be recorded and gathered more effectively. Video cameras are used to record the water surface, making the monitoring of

Proceedings DOI: 10.21467/proceedings.141; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-957605-4-1

floating debris in turbid rivers possible. These areas are estimated to make up most of the total riverine debris transport [2].

2 Research Background

Object detection is a technology associated with computer vision that is used to detect instances of basically anything from different classes such as humans, trees, trains and many more, either in images or videos [3]. Previous works that adopted object detection models in various applications are tabulated in Table 1 below.

Table 1: *Previous Studies on YOLO-based network*

Author(s)	Type of YOLO-based network used	Object Classes	Mean Average Precision (mAP)
[4]	Modified YOLOv3 (YOLOv3-2SMA)	Water surface garbage (Plastic bag, plastic bottle, styrofoam)	91.43%
[5]	YOLOv3	Marine debris (bags, bottlecaps, bottles, buoys, containers)	52.38%
[6]	YOLOv4 based on deconvolutional single shot detector (DSSD) network	Citrus fruits in orchard environment	96.04%
[7]	YOLOv3	Water surface objects	78.60%
[8]	YOLOv3	Underwater sea-life (mixed-size fishes, crab) Marine debris (plastic bag, plastic bottles, driftwood)	69.60% 77.20%
[9]	YOLOv4 with channel pruning algorithm	Apple flowers (Fuji, Red Love, Gala)	97.31%
[10]	YOLOv3	Underwater fish (various species)	55.75%
[11]	Modified YOLOv3	Underwater images	87.42%
[12]	Fusion-YOLO (F-YOLO)	Tea Chrysanthemum	89.53%
[13]	YOLOv4 Improved YOLOv4-tiny	Random images from MS COCO Dataset	64.9 % 38.0%
[14]	YOLOv4 Improved YOLOv4-tiny	Vehicle and pedestrian	65% 41.4%
[15]	YOLOv4-tiny	Marine plastic debris	84.0%

3 Materials and Methods

In this work, five major object classes that are commonly found floating in the river are identified. The classes are styrofoam, plastic bag, plastic bottle, plastic container, and aluminium can, which are some of the most common ones according to [2]. A total of 300 original images are augmented by adjusting the brightness level to imitate various environmental conditions. Hence, a total of 900 images are used as training dataset. On the other hand, 30 test images for each class are used to test the effectiveness of the proposed detection model. The model is trained on the Google Colabratory platform, which is a free cloud service based on Jupyter Notebooks. Free GPU is provided, which makes deep learning applications possible by using well-known libraries such as PyTorch, TensorFlow, Keras and OpenCV.

YOLOv4-tiny is basically the simpler version of YOLOv4. It is established based on YOLOv4 that is compressed in order to produce an undemanding structure and minimized parameters. YOLOv4-tiny has become very practical in creating on mobile and embedded devices. YOLOv4-tiny is well-known for its faster training time and detection speed because it contains only two YOLO heads. It has been trained from 29-pretrained convolutional layers. This architecture is a bit different compared to YOLOv4, which has three heads and has been trained from 137 pre-trained convolutional layers.

YOLOv4-tiny also has Frames Per Second (FPS), which is roughly around eight times more than the regular YOLOv4. However, its accuracy on MS COCO dataset is one-third lesser than that from YOLOv4. The model reaches 443 FPS with 22.0% Average Precision (AP) of 42% for AP50 [13]. This statistic is recorded on RTX 2080Ti. It also reached 1774 FPS using TensorRT. Batch size is initialized at 4 and at FP16-precision.

For real-time object detection, YOLOv4-tiny is the better option when compared with YOLOv4 as faster inference time is more important than precision or accuracy when working with a real-time object detection environment. The YOLOv4-tiny implemented the backbone of the CSPDarknet53-tiny network, which has CSPBlock module in cross stage partial network.

Aside from accuracy, computation is also improved by 10-20% using this module. The calculations of bottleneck with bigger values are eliminated to decrease the number of calculations. The sustained or minimized calculation increases accuracy for YOLOv4-tiny. The Mish activation function is eliminated in CSPDarknet53-tiny for extra simplification in the computation process. Instead, it uses the LeakyReLu function.

4 Theory and Calculation

Confusion matrix is constructed for each of the object classes and the model's performance is evaluated in terms of accuracy, precision, recall and F1-score using the formula listed below:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{1}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{2}$$

$$F1\ score = 2 \times \frac{Precision * Recall}{Precision + \ Recall} \tag{3}$$

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Negative + True\ Negative + False\ Positive} \tag{4}$$

5 Results and Discussion

Several challenging images that are blurry can still be detected, as can be seen from Figure 1. This shows that the proposed detector is feasible during rainy or sunny day. The testing is conducted using different IoU threshold values of 0.3, 0.5, 0.7 and 0.9 for each of the classes. IoU threshold value simply limits the detection to the model's confidence to detect the object.

Figure 1: *Detected object from aluminium Can class*

Figure 2: *Detected object from Styrofoam class*

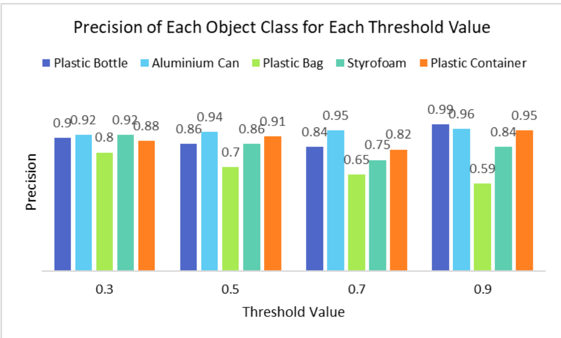

Figure 3: T*he precision versus threshold value for each object class*

In Figure 3, aluminium can maintain as the class with the highest precision value among all classes while plastic bag has the lowest precision value. However, the pattern for precision value is irregular for some classes as the threshold value increases.

The detectors' performance from YOLOv4 and YOLOv4-tiny are compared based on their overall performances, as can be seen in Figure 4.

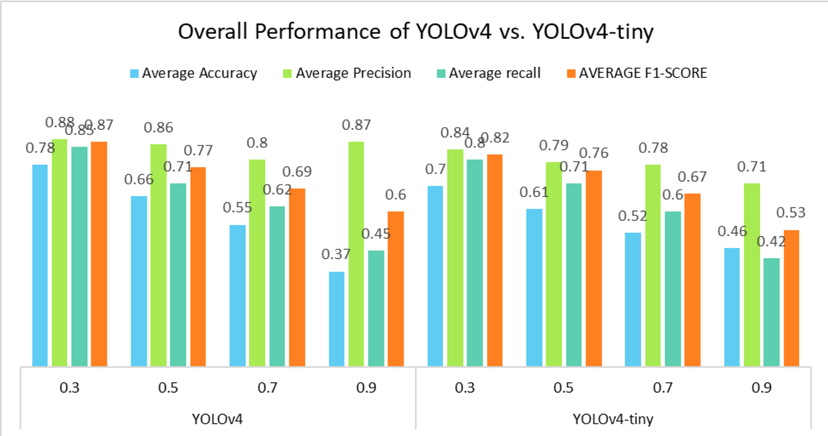

Figure 4: *The overall performance of YOLOv4 vs YOLOv4-tiny*

As can be seen in Figure 4, the YOLOv4 detector has a better overall performance in terms of accuracy, precision, recall and F1-score as compared to YOLOv4-tiny. However, the only trade-off between these two models is the time taken to train YOLOv4-tiny is so much shorter, which is approximately only 20% than that of YOLOv4. With YOLOv4, the average accuracy, average precision (AP), average recall and average F1-score are obtained at 66%, 86%, 71% and 77%, respectively for 0.5 threshold value. When comparing it to [5] the results are better than the YOLOv3 used for marine debris detection with a mAP of 52.38%. Meanwhile, [8] and [10] obtained a result of 77.20% and 55.75% mAP, respectively, also using YOLOv3.

For YOLOv4-tiny, the average accuracy, average precision (AP), average recall and average F1-score are obtained at 61%, 79%, 71% and 76 %, respectively for 0.5 threshold value. For marine debris detection, Tata et al. (2020) has the best results using YOLOv4-tiny with a mAP of 84%. However, for [14] and [15], they only could get 41.1% and 38% mAPs, respectively, which are worse than their YOLOv4 models.

6 Conclusions

In conclusion, an object detection module to find and detect plastic debris in the river has been successfully developed. A YOLO-based object detection framework has been developed for both YOLOv4 and YOLOv4-tiny. Furthermore, fine-tuning has been done for both models to find the best parameters to be used to obtain an effective object detection model. It is found that the best YOLOv4 model is the one that is trained using CSPDarknet53 backbone, uses width and height of 416*416, subdivisions of 8 and batch of 64. As for YOLOv4-tiny, the best model is obtained with CSPDarkenet53-tiny backbone, the size parameter of 2, width and height of 288*288, subdivisions of 8 and batch of 64.

Other than that, the effectiveness of the proposed models has been tested during testing process. Both models can detect images with a lot of variations such as blurry, noisy, dark, and bright images as well as objects from different perspectives or angles. In other words, these models are reliable and can be used under different weather conditions. It is observed that YOLOv4 yields the best results in terms of accuracy, precision, recall and F1-score when comparing it to the YOLOv4-tiny model. However, there is a trade-off between these two when it comes to the time taken to train the model because YOLOv4-tiny takes a lot shorter time to train than that of YOLOv4.

In the future, more image databases can be included to make the detector more general and reliable, especially because the detection for riverine debris can be really wide. This is due to various types of debris that may exist in the water environments. More classes can also be included to cover for other floating debris such as microplastics, glass and fishing equipment that can possibly be found in the river or the ocean.

7 Declarations

7.1 Competing Interests

There is no conflict of interest.

7.2 Publisher's Note

AIJR remains neutral with regard to jurisdiction claims in published maps and institutional affiliations.

References

[1] A. McIlgorm, H.F. Campbell, M.J Rule, "Understanding the economic benefits and costs of controlling marine debris in the APEC region (Report No. MRC 02/2007)," *National Marine Science Centre,* 2008. doi: 10.13140/2.1.4323.9042

[2] T. van Emmerik, M. Loozen, K. van Oeveren, F. Buschman, G. Prinsen, "Riverine plastic emission from Jakarta into the ocean," *Environmental Research Letters,* vol. 14, no.8, 2019. Access online on 24 March 2021 at https://iopscience.iop.org/article/10.1088/1748-9326/ab30e8

[3] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, M. Pietikäinen, "Deep learning for generic object detection: A survey," *International journal of computer vision*, vol. 128, no. 2, pp. 261-318, 2020. doi: 10.48550/arXiv.1809.02165

[4] X. Li, M. Tian, S. Wu, L., J. Yu, "A modified YOLOv3 detection method for vision-based water surface garbage capture robot," *International Journal of Advanced Robotic Systems*, vol. 14, no. 8, 2020. https://doi.org/10.1177%2F1729881420932715

[5] L. Sherwood, M. Tian, S. Kong, L. Wu, J. Yu, "Applying object detection to monitoring marine debris," *Tropical Conservation Biology and Environmental Science TCBES Theses*, vol. 14, no. 8, 2020. https://doi.org/10.1177%2F1729881420932715

[6] C. Fu, W. Liu, A. Ranga, A. Tyagi and A. Berg, "DSSD : Deconvolutional Single Shot Detector", 2019. Access online on 26 March 2022 at https://arxiv.org/abs/1701.06659.

[7] L. Zhang, Y. Zhang, Z. Zhang, J, Shen, H. Wang, "Real-time water surface object detection based on improved faster R-CNN," *Sensors*,vol. 19, no. 16, 2019. doi: 10.3390/s19163523

[8] J. Watanabe, I. Shao, Y., N. Miura, "). Underwater and airborne monitoring of marine ecosystems and debris," *Journal of Applied Remote Sensing,* vol.13, no.4, 2019. doi: 10.1117/1.JRS.13.044509

[9] D. Wu, S. Ly, M. Jiang, H. Song. "Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments," *Computers and Electronics in Agriculture,* vol. 178, 2020. https://doi.org/10.1016/j.compag.2020.105742

[10] W. Xu, S. Matzner, "Underwater fish detection using deep learning for waterpower applications," *International Conference on Computational Science and Computational Intelligence (CSCI),* 2018. Access online on 24 March 2021 at https://arxiv.org/abs/1811.01494v1

[11] F. Han, J. Yao, H. Zhu, C. Wang, "Underwater Image Processing and Object Detection Based on Deep CNN Method," *Journal of Sensors,* 2020. doi:10.1155/2020/6707328

[12] C. Qi, I. Nyalala, K. Chen, "Detecting the Early Flowering Stage of Tea Chrysanthemum Using the F-YOLO Model," *2020 IEEE International Conference on Consumer Electronics - Asia (ICCE-Asia,* 2020. https://doi.org/10.3390/ agronomy11050834

[13] Z. Jiang, L. Zhao, L. Li, S., Y. Jia, "Real-time object detection method based on improved YOLOv4-tiny," *Computer Vision and Pattern Recognition*, 2020. Access online on 25 March 2021 at https://arxiv.org/abs/2011.04244

[14] L. Ma, Y. Chen, J. Zhang, "Vehicle and Pedestrian Detection Based on Improved YOLOv4-tiny Model," *Journal of Physics: Conference Series,* 2021. Access online on 24 March 2021 at https://iopscience.iop.org/article/10.1088/1742-6596/1920/1/012034

[15] G. Tata, S.J. Royer, O. Poirion, J. Lowe, "DeepPlastic: A Novel Approach to Detecting Epipelagic Bound Plastic Using Deep Visual Models," *Environmental Science,* 2021. Access online on 23 March 2021 at https://arxiv.org/pdf/2105.01882.pdf