Development of Drifting Debris Detection System using Deep Learning on Coastal Cleanup

Shintaro Ogawa
Department of Intelligent and Control Systems, Kyushu Institute of Technology, 680-4, Kawazu Iizuka-city, Fukuoka, 820-8502, Japan

Takumi Tomokawa[1], Sylvain Geiser[1], Tan Chi Jie[1], Sakmongkon Chumkamon[1], Ayumu Tominaga[2], Eiji Hayashi[1]
[1]*Department of Mechanical Information Science and Technology, Kyushu Institute of Technology 680-4, Kawazu, Iizuka-city, Fukuoka, 820-8502, Japan*
[2]*Department of Creative Engineering Robotics and Mechatronics Course, National Institute of Technology Kitakyushu Colllege, 5-20-1 Shii, Kokuraminamiku, Kitakyushu-city, Fukuoka, 802-0985, Japan*

E-mail: ogawa.shintaro553@mail.kyutech.jp, m-san@mmcs.mse.kyutech.ac.jp tominaga@kct.ac.jp, haya@mse.kyutech.ac.jp, http://www.kyutech.ac.jp/

Abstract

In this paper, we developed a litter detector using deep learning to efficiently survey litter on beaches. The litter detector was developed using an HTC network. The HTC network and the mask R-CNN network were compared to evaluate the detector. The results showed that the HTC network was affected by small objects in the images.

Keywords: DNN, HTC, Mask R-CNN, Field Robot, Object detection, TACO

1. Introduction

Beach litter, a type of marine debris, has become a problem. The survey of litter is often conducted manually, which is a time-consuming task. In this study, a litter detection system was developed to reduce the amount of labor required for litter surveys. The creation of a litter detector is important for efficient and accurate litter surveys. The litter can range from small items such as microplastics to large items such as fishing equipment, and there are many beaches that are difficult for humans to access. In this study, we developed a detector that performs instance segmentation of beach images using deep learning. This detector detects trash pixel by pixel and classifies which type of trash it is. The accuracy of the detector was verified using a dataset containing images collected at a beach where cleanup activities were actually conducted.

2. HTC-based debris detection system

In this study, a deep learning model called HTC (Hybrid Task Cascade) was used for litter detection. HTC is an instance segmentation model that combines Cascade R-CNN and Mask R-CNN. Cascade R-CNN is improved from Faster R-CNN and expresses the threshold change of IoU (Intersection over Union, Eq. (1), which indicates the degree of overlap of two regions in a neural network. Mask R-CNN is a model for object detection, Faster R-CNN, with added segmentation functionality. Thus, HTC is created by combining two existing models.

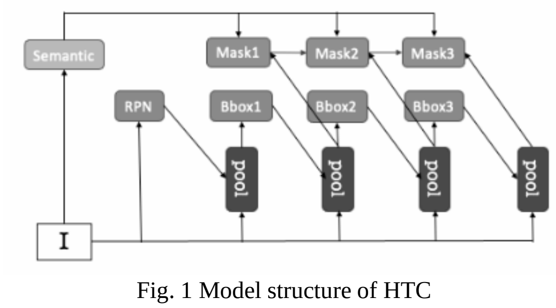

Fig. 1 Model structure of HTC

2.1. *Obtaining the dataset*

The dataset was created by combining a public dataset of general trash called TACO[1] and images taken at actual cleanup sites such as Hokuto Mizukumi Park. TACO is a dataset of overseas trash and includes many images taken at sites other than beaches, with a total of 1,500 images, and the latter includes 186 images taken at actual cleanup sites. The data set consists of 1686 images, 1636 of which are training data and 50 of which are evaluation data, for a total of 60 classes. In this study, in order to compare the detection accuracy between the different datasets, we prepared and trained two datasets, one using only TACO and the other using TACO plus images taken at the actual cleaning sites.

2.2. *Creating of Teacher Data*

In supervised learning, it is necessary to create correct answer data, called teacher data. In supervised learning, it is necessary to create correct data called teacher data. In supervised learning, the process of mapping teacher labels to target objects in images is called annotation. In this study, the annotation tool coco-annotator[2] was used to annotate each class of litter, and the teacher data was created. The image shown in Figure 2 shows the left image before annotation and the right image after annotation, in which a plastic bottle is annotated by dividing it into a bottle part and a cap part.

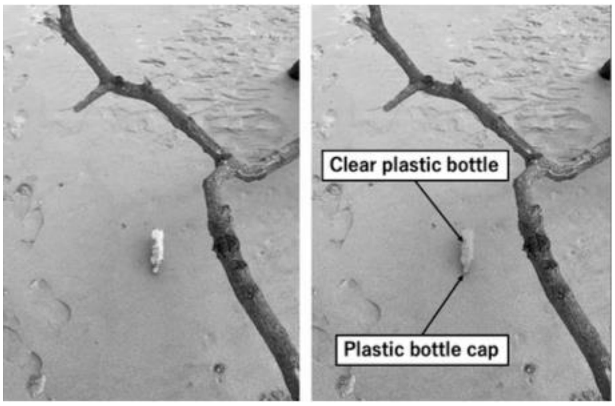

Fig. 2. Example Dataset

2.3. *Identifier creation and evaluation methods*

The discriminator was trained 2000 times with a learning rate of 0.002. Multi-class cross-entropy was used as the loss function. In addition to HTC, Mask R-CNN was used for training to compare real-time performance. The discriminators were evaluated using the unlearned discriminants. For evaluation of the discriminator, 50 untrained images were prepared and evaluated with the two training models. The accuracy was compared by calculating the goodness of fit, recall, and F value in the range of IoU values above 0.5. IoU is a measure of the accuracy of the expected region in object detection, and F value is the harmonic mean of the goodness of fit and recall. Equations (1) to (4) below show the equations for IoU, fit ratio, recall ratio, and F value.

$$IoU = \frac{Area\ of\ Intersection}{Area\ of\ Union} \tag{1}$$

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

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{3}$$

3. Experimental results

Below are the identification results of the evaluation data for each model and dataset as shown in Figure 3.1 and Figure 3.2.

Fig. 3.1. Example of Identification Result
(Left: HTC, Right: Mask R-CNN, Dataset: TACO)

Fig. 3.2. Example of Identification Result
(Left: HTC, Right: Mask R-CNN, Dataset: TACO+Real env.)

The experimental results are evaluated in term of the fit rate, repeatability, and F value for the class as a whole, and the fit rate for PET bottles as shown in Table 1 while

Table 2 shows the individual precision performance for plastic bottle class.

Table 1. Precision, Recall, F-score

	Precision	Recall	F-score
HTC(TACO)	0.034	0.074	0.047
Mask R-CNN(TACO)	0.025	0.03	0.027
HTC(TACO+Real env.)	0.041	0.053	0.046
Mask R-CNN (TACO+Real env.)	0.038	0.053	0.044

Table 2. Precision in plastic bottle

	Precision
HTC(TACO)	0.164
Mask R-CNN(TACO)	0.127
HTC(TACO+Real env.)	0.141
Mask R-CNN (TACO+Real env.)	0.167

4. Consideration

Although Table 1 shows a relatively low precision score, the individual class of plastic bottle demonstrate a better precision result as shown in Table 2. Taking the identification of pet bottles as an example, the Mask R-CNN showed a higher conformance rate with the addition of real environment data, while the HTC showed a higher conformance rate without the addition of real environment data. The F-measure of the class as a whole does not increase much with the addition of the dataset for HTC. Why is it that HTC does not obtain good results even with the addition of a dataset? One possible reason is the compatibility between the added real-world dataset and HTC's model structure. As a characteristic of the data set, the real environment data set contained more litter than TACO, including many small litters. As shown in Figure 1, the model structure of HTC is such that the features of the Region Proposal Network (RPN) are optimized through the pooling layer. Therefore, it is thought that the detection of small debris was omitted in the feature optimization process due to the addition of the real environment data set. Classification of small objects is also extremely difficult, and misclassification is considered to be one of the reasons. For prospects, to solve these problems, the input resolution of the model should be increased, and the number of pixels occupied by small objects should be increased by raising the resolution of the input images. In addition, it may be possible to prevent misclassification by avoiding annotation of too small garbage and reducing the number of classes to a minimum in the creation of training data.

5. Summary

In this study, we attempted to detect debris at a beach cleaning site using deep learning. We attempted to detect debris at a beach cleanup site using deep learning. As a result, we found that the number of classes for the dataset. The results suggest that it is necessary to set an appropriate number of classes for the dataset and to devise a way to detect small objects and that it is necessary to devise a method for small objects. In the future, we will aim to further improve the detection accuracy. We will further improve the detection accuracy and build a system with the aim of increasing its usefulness in actual cleaning sites. We will construct a system with the aim of further improving the detection accuracy and increasing the usefulness of the system in actual cleaning sites.

References

1. "TACO Dataset ", http://tacodataset.org/
2. Justine Brooks, https://github.com/jsbroks/coco-annotator

Authors Introduction

Mr. Shintaro Ogawa

He received bachelor degree in Engineering in 2022 from intelligent and Control Systems, Kyushu Institute of Technology in Japan. He is currently a Master student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Dr. Sakmongkon

Dr. Sakmongkon Chumkamon received Doctor of Engineering degree from Kyushu Institute of Technology in 2017. He was a postdoctoral Guangdong Technology Presently he is a postdoctoral researcher in Kyushu Institute of Technology since 2019. His research interests include factory automation robots and social robots.

Projected Assist. Prof. Ayumu Tominaga

Projected Assist. Prof. Ayumu Tominaga is a professor in Department of Creative Engineering Robotics and Mechatronics Course at National Institute of Technology Kitakyushu College. He received the Ph.D. (Dr. Eng.) degree from Kyushu Institute of Technology in 2021. His research interests include Intelligent mechanics, Mechanical systems and Perceptual information processing

Mr. Takumi Tomokawa

He received bachelor degree in Engineering in 2021 from mechanical system engineering, Kyushu Institute of Technology in Japan. He is currently a Master student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Mr. Tan Chi Jie

He received his Bachelor of Engineering Electronics Majoring in Robotics and Automation from the Faculty of Engineering, Multimedia University, Malaysia in 2020. He is currently a Master student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Mr. Sylvain Geiser

He studied Engineering and Computer Science in Ecole des Mines de Nancy, France, between September 2018 and June 2020. He is currently a Master student at Kyushu Institute of Technology, Japan, and conducts research at Hayashi Laboratory.

Prof. Eiji Hayashi

Prof. Eiji Hayashi is a professor in the Department of Intelligent and Control Systems at Kyushu Institute of Technology. He received the Ph.D. (Dr. Eng.) degree from Waseda University in 1996. His research interests include Intelligent mechanics, Mechanical systems and Perceptual information processing. He is a member of The Institute of Electrical and Electronics Engineers (IEEE) and The Japan Society of Mechanical Engineers (JSME).