Proposal of an Object Recognition Method Capable of Detecting and Clustering Unknown Objects

Yu Horiuchi, Shuichi Akizuki, Manabu Hashimoto, Takuma Mori and Kenji Nishimiya

***Abstract*— The proposed object recognition method can handle both known and unknown objects in images. Experimental results show that the proposed method can classify unknown objects while maintaining the accuracy of known object detection.**

***Keywords* — Object Detection; Out-of-distribution; Unsupervised Clustering; Unmanned Ground Vehicle;**

I. Introduction

Unmanned ground vehicle (UGV) is a self-driving work vehicle. UGV needs to detect and recognize people, vehicles, and other obstacles around itself. Depending on the situation, the UGV must stop or move around these obstacles. Generally, many unique objects are present at construction sites and factories where UGV operates. The UGV needs the ability to detect and recognize site-specific objects.

The goal of this study is unknown object detection and classification. While retaining the ability of the proposed method to classify known objects, we add to it the ability to detect and group unknown classes. It would handle features from known and unknown objects in the same space and optimize the distribution of features between known and unknown classes.

II. Proposed Method

The proposed method consists of two stages. In the first stage, we will aim to collect data on unknown objects. To do this, we will use a method called virtual outlier synthesis (VOS) [1] that detects unknown objects using only labeled data. First, we will train VOS offline using a common dataset. Next, we will apply VOS to scenes captured by a camera on a UGV. In the second stage of the proposed method, we will conduct a learning process. First, we will create teacher labels for the unknown object data obtained in the first stage. Unknown object data cannot be used to compute the loss of Faster R-CNN [2] because there are no teacher labels. Therefore, we will use deep clustering [3]. With deep clustering, we will cluster the unknown object data without teacher labels. In particular, we will perform principal component analysis and k-means clustering on the extracted image features, which will generate pseudo labels. We will extract features with ResNet and propose regions with region proposal network (RPN). Calculate the class likelihood of each object instances using the feature maps and the proposed regions. Next, the loss between this class likelihood and the teacher labels of the known objects or the pseudo labels of the unknown objects are computed.

III. Experiment And Discussion

The purpose of the experiment has following two parts: first, to show that the proposed method can classify unknown objects; second, to check the impact on the classification accuracy of known objects when considering unknown objects during learning.

The comparison method is Faster R-CNN.

The evaluation metric is mean average precision (mAP) that is an indicator of accuracy of detection of location and classification of the objects.

For training, we set 15 classes of the VOC dataset as known objects and 5 classes as unknown objects. For inference, we will use the COCO dataset (only using the same classes as the VOC dataset).

Fig.1 shows the classification results of the proposed method, and Table 1 shows the detection accuracy of the proposed method and the comparison method. From these results confirm that the proposed method can detect unknown classes. The classification results of the known objects in Table 1 show no difference in classification accuracy. This indicates that the proposed method does not adversely affect the classification accuracy of the known classes. Therefore, the proposed method is effective in classifying unknown objects.

Figure1. Clustering results of unknown objects by the proposed method

TABLE I. Object Detection Results For MS COCO

	ID (mAP↑)	OOD (mAP↑)	ALL (mAP↑)
Faster R-CNN	46.8	-	46.8
Ours	47.1	23.6	41.2

References

[1] Xuefeng Du, et al, "VOS:LEARNING WHAT YOU DON'T KNOW BY VIRTUAL OUTLIER SYNTHESIS", ICLR, (2022).
[2] S. Ren, et al, "Faster R-CNN: Towards RealTime Object Detection with Region Proposal Networks," IEEE Trans. PAMI, (2017).
[3] M. Caron, et al, "Deep Clustering for Unsupervised Learning of Visual Features", ECCV, (2018)

Yu Horiuchi is with the Graduate School of Engineering, Chukyo University, Nagoya, Aichi, Japan (corresponding author to provide e-mail: horiuchi@isl.sist.chukyo-u.ac.jp).

Shuichi Akizuki is with the Graduate School of Engineering, Chukyo University, Nagoya, Aichi, Japan (corresponding author to provide e-mail: s-akizuki@sist.chukyo-u.ac.jp).

Manabu Hashimoto is with the Graduate School of Engineering, Chukyo University, Nagoya, Aichi, Japan (corresponding author to provide e-mail: mana@isl.sist.chukyo-u.ac.jp).

Takuma Mori and Kenji Nishimiya are with Honda R&D Co., Ltd., Wako, Saitama, Japan.