2018 21st International Conference on Intelligent Transportation Systems (ITSC)

Attribute-aware Semantic Segmentation of Road Scenes for Understanding Pedestrian Orientations

M. D. Sulistiyo[1,2], Y. Kawanishi[1], D. Deguchi[3], T. Hirayama[4], I. Ide[1], J. Y. Zheng[1,5], H. Murase[1]

***Abstract*— Semantic segmentation is an interesting task for many deep learning researchers for scene understanding. However, recognizing details about objects' attributes can be more informative and also helpful for a better scene understanding in intelligent vehicle use cases. This paper introduces a method for simultaneous semantic segmentation and pedestrian attributes recognition. A modified dataset built on top of the Cityscapes dataset is created by adding attribute classes corresponding to pedestrian orientation attributes. The proposed method extends the SegNet model and is trained by using both the original and the attribute-enriched datasets. Based on an experiment, the proposed attribute-aware semantic segmentation approach shows the ability to slightly improve the performance on the Cityscapes dataset, which is capable of expanding its classes in this case through additional data training.**

I. Introduction

In recent years, computer vision technologies by deep learning have become popular for various complex tasks. One of the most interesting topics is the semantic segmentation, which is a more complex task compared to *simple* image segmentation. Semantic segmentation is a pixel-wise labeling task to simultaneously classify pixels in the input image into predefined classes. Semantic segmentation with deep-learning approaches have been widely utilized for the purposes such as indoor and outdoor scene understanding as well as applications like autonomous vehicle sensing [1] and robotics navigation [2]. For instance, from a road scene, a model trained for semantic segmentation will inform an autonomous vehicle of surrounding objects before it acts.

Many studies on this task have been conducted so far. A Fully Convolutional Network was introduced for the semantic segmentation task as an end-to-end trained model [3]. Almost in the same period, SegNet with a different network architecture was also published [1]. It was then modified into Bayesian SegNet and obtained better performance by adding a probabilistic element [4]. Later on, the Pyramid Scene Parsing network (PSPNet) [5] outperformed other existing methods and achieved a state-of-the-art performance for PASCAL VOC 2012 [6] and Cityscapes datasets [7]. The Mask R-CNN presented a conceptual framework for instance-level segmentation, which combined object detection and semantic segmentation tasks. The approach outperformed some previous methods for several challenges [8]. Reference [9] provides a summary of popular deep neural networks that were adopted to solve the semantic segmentation.

However, the existing semantic segmentation methods only recognize an object's name as its class, e.g., road, building, car, person, and so on, while for autonomous vehicles, additional information describing a particular object in detail such as its attributes could help the better understanding of scenes. Therefore, this paper proposes an attribute-aware semantic segmentation method, which is a more difficult challenge. Yet we can obtain better evaluation results if the process is successfully performed. With simultaneous recognition and segmentation tasks, it can enhance environment perception tasks in Intelligent Transportation Systems that need precision such as parking, road following, pedestrian detection, action identification, and localization.

Here we chose pedestrian as the target class, out of other classes like vehicle and traffic signs, as it has many attributes to be explored. Since a pedestrian is one of the most important moving objects in a street scene, it should be useful for assisting an autonomous vehicle. In this paper, we particularly consider the body orientation of a pedestrian as the attribute and divide it into a predefined number of classes. In the autonomous vehicle system, being informed of pedestrian orientations can be helpful in collision avoidance. Although *only* the pedestrian body orientation was added as a new semantic information, it can give hint and insight on other object attributes when the method is extended.

The contributions of this paper are as follows:

1) We introduce a new concept of attribute-aware semantic segmentation that is general and possible to be applied to various tasks.
2) We design an attributes-dependent loss function that is modified from the original one for increasing the number of classes regarding the attribute-aware concept.
3) We modify the annotation of Cityscapes dataset for the attribute-aware semantic segmentation task.
4) By extending the Cityscapes to include additional attribute classes, we achieved better performance of semantic segmentation.

The rest of this paper will discuss related work in Sec. II, our proposed idea and its implementation in Sec. III, and our modified dataset annotation in Sec. IV. Experimental results and discussions are presented in Sec. V followed with a conclusion.

[1]Mahmud Dwi Sulistiyo, Yasutomo Kawanishi, Ichiro Ide, Hiroshi Murase are with Graduate School of Informatics, Nagoya University, Japan {kawanishi, ide, murase} @i.nagoya-u.ac.jp

[2]Mahmud Dwi Sulistiyo is also with School of Computing, Telkom University, Indonesia mahmuddwis@telkomuniversity.ac.id

[3]Daisuke Deguchi is with Information Strategy Office, Nagoya University, Japan ddeguchi@nagoya-u.jp

[4]Takatsugu Hirayama is with Institutes of Innovation for Future Society, Nagoya University, Japan takatsugu.hirayama@nagoya-u.jp

[5]Jiang Yu Zheng is with Dept. of Computer Science, Indiana University-Purdue University Indianapolis, USA jzheng@iupui.edu