CityWalks: An Extended Dataset for Attribute-aware Semantic Segmentation

Mahmud Dwi Sulistiyo[1,2], Yasutomo Kawanishi[1], Daisuke Deguchi[1], Ichiro Ide[1], Takatsugu Hirayama[1], and Hiroshi Murase[1]
[1]Nagoya University, Japan, [2]Telkom University, Indonesia

1. Introduction

A computer vision task to classify each and every pixel of the input image is called as semantic segmentation; It has been widely utilized for various purposes, including in ITS applications. Recently, the task is extended to attribute-aware semantic segmentation for enriching the output information by providing attribute values of a particular object, and thus give a better scene understanding [1]. Numerous datasets for the semantic segmentation task are publicly available, but none of them is sufficient for the attribute-aware semantic segmentation task. Therefore, we construct a novel dataset named CityWalks as an extension to the existing dataset with additional labels corresponding to pedestrian's body orientations as the attributes. This talk presents the construction of the CityWalks dataset and describes it from its qualitative and quantitative sides.

2. Related Work

For the semantic segmentation task, the Cityscapes is one of the most popular datasets which has mainly been built for traffic scene understanding [2]; It represents the complexity of real-world urban scenes featured with large amount of images, annotation richness, and scene diversity. The Cityscapes dataset, which has 19 classes for public challenge, provides 2,975 training images, 500 validation images, and 1,525 testing images. The training and validation sets are available with their fine annotated ground truths. However, the Cityscapes annotates only the object classes and hence it does not comply with the attribute-aware semantic segmentation task.

3. CityWalks Dataset

We construct the CityWalks dataset as an extension to the Cityscapes dataset for simultaneous semantic segmentation and attribute recognition tasks. The class *person* is extended into four classes corresponding to the pedestrians' orientations including *back*, *right*, *front*, and *left*. We keep the class label of *person* to refer to persons whose orientations are difficult to annotate (labeled as *unknown*), which results in 23 objects and attributes classes. Information of a pedestrian's orientation is important for ITS purposes such as pedestrian's movement prediction and collision avoidance. We manually re-annotated the Cityscapes' ground truth labels in pixel-level details using a conventional image editor; The process of re-annotation took around 2 to 10 minutes per image for 'easy' to 'difficult' cases.

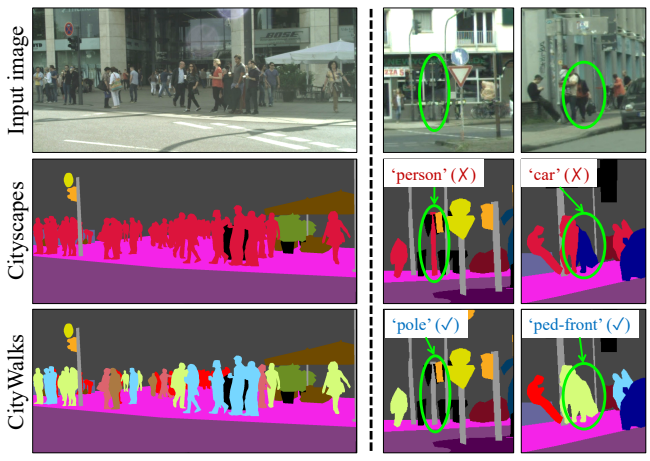

Fig. 1. Comparing the Cityscapes and the CityWalks annotations.

Table 1 Statistics of the CityWalks dataset

	Training set	Validation set
#image in the dataset	2,975	500
#image with *person*	2,345	402
#image with 4-oriented pedestrian	2,083	371
%*person* pixels in the dataset	1.08 %	1.15 %
%orientations pixels in the dataset	1.03 %	1.09 %

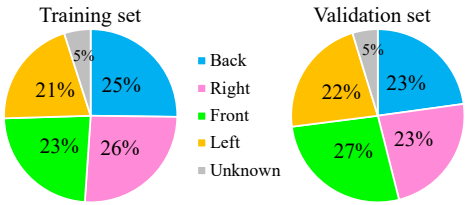

Fig. 2. Distributions of the pedestrians' orientation labels

Fig. 1 shows examples of the CityWalks ground truth labels compared to the original Cityscapes ones as well as some incorrect annotations we found in the original Cityscapes dataset. We summarize the statistics of the CityWalks dataset in Table 1. The *person* indicates all labels annotated with four orientation classes and pedestrian with *unknown* orientation. Fig. 2 shows distributions of the re-annotation results, especially for the four pedestrian orientation labels, covering all images in training and validation sets.

4. Conclusion

We introduced the CityWalks which is an extended traffic scene dataset based on the Cityscapes dataset for the task of attribute-aware semantic segmentation by considering the pedestrian's orientations as an additional attribute information. The CityWalks dataset is expected to improve the capability of the trained model to perform semantic segmentation and attribute recognition simultaneously; The dataset will be publicly available soon.

Acknowledgment: This research is supported by Indonesia's BUDI-LN Scholarship and Japan's MEXT Grants-in-Aid for Scientific Research.

Reference

(1) M. D. Sulistiyo, et al.: Proc. 21st IEEE ITSC, pp. 2698−2703, 2018.
(2) M. Cordts, et al.: Proc. 2016 IEEE CVPR, pp.3213−3223, 2016.