学位報告 4

報告番号	※甲　　第　　号

主　論　文　の　要　旨

論文題目　Attribute-Aware Semantic Segmentation from an In-vehicle Camera（車載カメラ映像からの属性付きセマンティックセグメンテーション）

氏　　名　Mahmud Dwi Sulistiyo

論 文 内 容 の 要 旨

Today, the need for the existence of an autonomous driving is increasing, while various methods in the fields of Computer Vision and Machine Learning have been rapidly developed. In this regard, semantic segmentation has become one of the most important techniques in performing environmental perception that supports the intelligent vehicle. However, the conventional Semantic Segmentation task is not sufficient to produce a comprehensive information because in the implementation, it is only able to inform what objects are captured by the camera, not yet explaining the states of these objects. Therefore, there is a demand to simultaneously combine the Semantic Segmentation and the Attribute Recognition tasks to improve the ability of understanding the traffic scenes. This thesis is presented to answer this challenge by introducing the Attribute-Aware Semantic Segmentation task for images captured by an in-vehicle camera as well as contributing some proposed approaches with sufficient experiments and comparative analysis.

The main objective of this thesis is to enhance the semantic segmentation result and improve the performance both quantitatively and qualitatively. This research focuses on pedestrian and its body orientation as the target object and attribute, in addition to performing semantic segmentation for other object categories. This thesis reveals three important factors, which corresponding to the three research topics, to actualize the attribute-aware semantic segmentation, including the lack of sufficient dataset, the general perspective in developing the method (Domain-Free approach), and the special perspective in improving the method’s performance (Domain-Specific approach). Each research topic is discussed in a separate chapter.

The first aspect to address is the construction of a sufficient dataset. As we know, there are numerous datasets publicly available for the conventional Semantic Segmentation task. However, none of them is sufficient for the task of Attribute-Aware Semantic Segmentation task since it is considered a newly introduced task. This thesis introduces a novel dataset named CityWalks as an extension to the Cityscapes dataset with additional attribute labels corresponding to a pedestrian's four body orientations. The CityWalks dataset aims to provide high quality ground-truth annotations including training and validation sets, in developing deep learning-based models for attribute-aware semantic segmentation.

The next aspect of concern is how to develop a method for the Attribute-Aware Semantic Segmentation task that is independent to the data domain, generally applicable, and powerful in performance. In this regard, a Domain-Free approach is proposed to solve the problem of attribute-aware semantic segmentation regardless of the image data domain. With this approach, an attribute-aware loss function is introduced as the Domain-Free method. It proposes a unified formula that is able to treat both object and attribute classes in the same manner and hence is applicable to an arbitrary base model. Experiments with various settings are conducted and the results show that the proposed method successfully outperforms the baseline methods.

Despite the advantages demonstrated by the Domain-Free approach, the method has difficulties to handle ambiguous attributes such as pedestrian's body orientation in object-level. To tackle this problem, a Domain-Specific approach, which focuses on the target object and attribute as well as the characteristics of the input data domain, is proposed through a Multi-Task Learning-based custom model named ColAtt-Net. The model performs two functions separately, including the pixel-wise classification on object categories and the column-wise prediction on pedestrians' body orientations. The improvements shown by the proposed Domain-Specific method in several experiments are quite significant thanks to its ability to reduce the ambiguity in predicting a pedestrian's orientation.

The research presented in this thesis has answered a part of the global problems that may exist when solving the Attribute-Aware Semantic Segmentation task especially from an in-vehicle camera. Since the topic is initially introduced in this thesis, there will be several ideas for further development as well as other challenges.