Efficient Semantic Segmentation for Autonomous Vehicles with Limited Computational Resources

Felix Blanco, *Member, UTEC*

Abstract—The abstract goes here.

Index Terms—Computer Society, Computer Vision and Pattern Recognition

1 Introduction

Semantic segmentation, which involves classifying each pixel in an image into a set of predefined categories, is a fundamental task in computer vision, and has a crucial role in autonomous driving systems to enable the vehicle to understand and interpret its surroundings. Dividing the image into multiple semantic regions is a key step in the perception pipeline of autonomous vehicles, as it provides a rich representation of the environment identifying the instances of objects and their relations.

Autonomous vehicles operate in a wide range of environments, and thus, the real-time performance of the segmentation algorithm is critical as it's a key component required for safe and reliable decision-making. However, the computational resources available in autonomous vehicles are often limited due to physical constraints including memory capacity due to existence of large models and processing speed required for real-time responses. These limitations pose significant challenges for implementing efficient and accurate semantic segmentation algorithms on autonomous vehicles, as traditional methods typically rely on computationally expensive deep neural networks with many parameters.

In addition, complex scenes, very often contain many instances and abstract levels, and high-resolution images to capture fine-grained details, pose additional challenges for semantic segmentation algorithms. **Complex scenes** require the model to classify numerous objects accurately, even more if the model request some hierarchical information. However, achieving high accuracy in semantic segmentation is a challenging task due to presence of occlusions, object overlaps and lighting variations, which can lead to inaccurate classification. **High-resolution** of the input images increase the computational demands of semantic segmentation models, which can lead to a significant drop in performance, making it difficult to achieve real-time processing required by autonomous vehicles.

The objective is to enable real-time semantic segmentation on autonomous vehicles without compromising the model accuracy while keeping the computational requirements low. Also, should be able to handle different levels of scene detail and adapt to different driving scenarios, due to dynamic and changing environments.

To evaluate the proposed approach, made use of benchmark datasets(*i.e.* Cityscapes[1], Mapillary Vistas [2]) for semantic segmentation in autonomous driving scenarios. This paper findings have important implications for the development of *practical* semantic segmentation algorithms for autonomous vehicles, deploying embedded systems with constrained computational resources, in order to contribute to the advancement of safe and reliable autonomous driving systems.

2 Theorical Foundations

2.1 Semantic Segmentation

A

2.2 Hierarchical Semantic Segmentation

B

2.3 Transformers

C

3 Related Work

3.1 Deep Hierarchical Semantic Segmentation[3]

B

3.2 REF-Net: robust, efficient, and fast network for semantic segmentation applications using devices with limited computational resources[4]

A

References

[1] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 3213–3223.

[2] G. Neuhold, T. Ollmann, S. Rota Bulo, and P. Kontschieder, "The mapillary vistas dataset for semantic understanding of street scenes," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 4990–4999.

[3] L. Li, T. Zhou, W. Wang, J. Li, and Y. Yang, "Deep hierarchical semantic segmentation," *arXiv preprint arXiv:2203.14335*, 2022.

[4] B. Olimov, J. Kim, and A. Paul, "Ref-net: robust, efficient, and fast network for semantic segmentation applications using devices with limited computational resources," *IEEE Access*, vol. 9, pp. 15 084–15 098, 2021.