Real-time Domain Adaptation in Semantic Segmentation

Ianniello Luca, Martone Raffaele, Sirica Antonio

January 25, 2025

Abstract

This study addresses the problem of real-time domain adaptation in semantic segmentation, focusing on bridging performance gaps between source and target domains. Leveraging the LoveDA dataset, which encompasses rural and urban domains, we evaluate the performance of state-of-the-art models, including DeepLabv2 and PIDNet, under domain adaptation settings. We explore a range of adaptation techniques, such as data augmentation, adversarial learning, domain adaptation via cross-domain mixed sampling (DACS), Prototype-based Efficient MaskFormer (PEM), and unpaired image-to-image translation with CycleGAN. Our findings reveal the strengths and limitations of these methods in improving segmentation performance across domains while preserving computational efficiency.

1 Introduction

Semantic segmentation is a fundamental task in computer vision that involves partitioning an image into semantically meaningful regions, typically corresponding to different object classes. This task is crucial for various applications, including autonomous driving, medical imaging, and remote sensing. Recent advancements in deep learning have significantly improved the performance of semantic segmentation models [4].

While deep learning models such as DeepLab [2] and PIDNet [3] have achieved remarkable performance on benchmark datasets, their generalization to unseen domains remains a significant obstacle. Domain shifts, arising from differences in image characteristics, environments, and data distributions, often result in substantial performance degradation.

Domain adaptation aims to bridge the performance gap between source and target domains, enabling models trained on one domain to generalize effectively to another. The LoveDA dataset [7], with its distinct rural and urban domain settings, provides a robust benchmark for evaluating domain adaptation techniques in semantic segmentation.

This project investigates real-time domain adaptation techniques for semantic segmentation, focusing on the LoveDA dataset. We evaluate state-of-the-art models, such as DeepLabv2 and PIDNet, under various domain adaptation scenarios. Specifically, we analyze the challenges of adapting models from rural to urban domains and vice versa, and we explore solutions including data augmentation, adversarial training [6], image-to-image translation approaches (DACS) [5], real-time networks (PEM) [1],

and style transfer preprocessing models, like Cycle-GAN [8]. By experimenting with these approaches, we aim to identify effective methods for improving domain adaptation performance while maintaining real-time capabilities.

2 Related Work

3 Methodology

4 Experiments and results

5 Conclusion

References

[1] Niccolo Cavagnero, Gabriele Rosi, Claudia Cuttano, Francesca Pistilli, Marco Ciccone, Giuseppe Averta, and Fabio Cermelli. Pem: Prototype-based efficient maskformer for image segmentation. *arXiv preprint arXiv:2402.19422v3*, 2024. 1

[2] Liang-Chieh Chen, George Papandreou, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017. 1

[3] Hao Feng et al. Pidnet: A real-time semantic segmentation network inspired by pid controllers. *arXiv preprint arXiv:2103.12370*, 2021. 1

[4] Shijie Hao, Yuan Zhou, and Yanrong Guo. A brief survey on semantic segmentation with deep learning. *arXiv preprint arXiv:2007.04063*, 2020. 1

[5] Wilhelm Tranheden, Viktor Olsson, Juliano Pinto, and Lennart Svensson. Dacs: Domain adaptation via cross-domain mixed sampling. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2021. 1

[6] Yi-Hsuan Tsai, Wei-Chih Hung, Samuel Schulter, Kihyuk Sohn, Ming-Hsuan Yang, and Manmohan Chandraker. Learning to adapt structured output space for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018. 1

[7] Yucheng Wang et al. Loveda: A remote sensing land-cover dataset for domain adaptive semantic segmentation. *arXiv preprint arXiv:2110.08733*, 2021. 1

[8] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. *arXiv preprint arXiv:1703.10593v7*, 2020. 1