# Report

January 6, 2025

# 1 Testi Luca

## 1.1 Semantic Segmentation

Semantic segmentation is a fundamental task in computer vision that involves classifying each pixel in an image into a predefined category, thereby enabling detailed scene understanding. Recent advancements in deep learning have significantly improved the accuracy and efficiency of semantic segmentation methods. For instance, DeepLab employs atrous convolutions and fully connected Conditional Random Fields (CRFs) to capture multi-scale context and refine segmentation boundaries [?]. BiSeNet introduces a bilateral network architecture to balance spatial detail and semantic context, achieving real-time performance [?]. Meanwhile, PIDNet draws inspiration from PID controllers to develop a lightweight network for high-quality real-time segmentation [?]. Domain adaptation techniques further enhance semantic segmentation by addressing challenges posed by domain shifts, as seen in methods like DACS, which utilizes cross-domain mixed sampling [?], and LoveDA, a dataset specifically designed for domain adaptation in remote sensing [?]. These advancements highlight the diverse strategies employed to tackle the complexities of semantic segmentation across different applications and domains.

## 1.2 PIDNet

PIDNet is a real-time semantic segmentation network inspired by the principles of Proportional-Integral-Derivative (PID) controllers, which are widely used in control systems to achieve precise and stable performance [?]. By integrating ideas from PID control theory, PIDNet introduces a unique architecture that balances low-latency processing with high-quality segmentation results. The network comprises three branches—P (proportional), I (integral), and D (derivative)—that are designed to capture complementary information: the P-branch focuses on spatial detail, the I-branch accumulates global context, and the D-branch enhances boundary precision. This innovative design enables PIDNet to achieve state-of-the-art performance in real-time segmentation tasks, particularly in scenarios that demand both accuracy and efficiency, such as autonomous driving and robotics. Furthermore, its lightweight structure ensures applicability in resource-constrained environments without significant performance degradation, demonstrating its practical utility in a wide range of applications.

## 1.3 ADAM

ADAM (Adaptive Moment Estimation) is a popular optimization algorithm that combines the benefits of adaptive learning rates and momentum-based updates to accelerate convergence and improve generalization in deep learning models. The algorithm maintains two moving averages of gradients: the first moment (mean) and the second moment (uncentered variance). These estimates are used to update the model parameters by adjusting the learning rate based on the gradient magnitudes and the historical gradients. ADAM's adaptive learning rate mechanism allows it to automatically adjust the step size for each parameter, enabling faster convergence and better generalization performance compared to traditional optimization methods like Stochastic Gradient Descent (SGD). By incorporating momentum, ADAM also benefits from faster convergence and improved stability, making it a popular choice for training deep neural networks across various tasks and domains.

## 1.4 DACS

Domain adaptation addresses the challenge of generalizing a model trained on one domain to perform well on a different but related domain, a critical issue in semantic segmentation when labeled data from the target domain is scarce. Traditional methods often struggle with the significant domain gap caused by differences in visual appearance, texture, or lighting conditions. DACS (Domain Adaptation via Cross-domain Mixed Sampling) introduces an innovative approach to bridge this gap by leveraging cross-domain mixed sampling [?]. This technique combines image regions from both the source and target domains to create hybrid training samples, effectively enhancing the model's ability to adapt to target domain features. Additionally, DACS uses a self-supervised learning objective to further refine its predictions on the target domain, achieving state-of-the-art performance on several benchmarks. The simplicity and effectiveness of DACS make it a robust solution for domain adaptation challenges in semantic segmentation, particularly in applications like autonomous driving and remote sensing where domain shifts are prevalent.

@articlechen2017deeplab, title=DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs, author=Chen, Liang-Chieh and Papandreou, George and Murphy, Kevin and Yuille, Alan L, journal=IEEE Transactions on Pattern Analysis and Machine Intelligence, year=2017

@inproceedingsyu2018bisenet, title=BiSeNet: Bilateral Segmentation Network for Real-time Semantic Segmentation, author=Yu, Changqian and Wang, Jingbo and Peng, Chao and Gao, Changxin and Yu, Gang and Sang, Nong, booktitle=Proceedings of the European Conference on Computer Vision (ECCV), year=2018

@articlefeng2021pidnet, title=PIDNet: A Real-time Semantic Segmentation Network Inspired by PID Controllers, author=Feng, Hao and others, journal=arXiv preprint arXiv:2103.12370, year=2021

@inproceedingstranheden2021dacs, title=DACS: Domain Adaptation via Cross-domain Mixed Sampling, author=Tranheden, Wilhelm and Olsson, Viktor and Pinto, Juliano and Svensson, Lennart, booktitle=Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), year=2021

@articlewang2021loveda, title=LoveDA: A Remote Sensing Land-Cover Dataset for Domain Adaptive Semantic Segmentation, author=Wang, Yucheng and others, journal=arXiv preprint arXiv:2110.08733, year=2021