Real time domain adaptation in semantic segmentation

Bernaudo J., Scarciglia L., Tasca M.

Introduction

- Real time semantic segmentation
- Domain adaptation

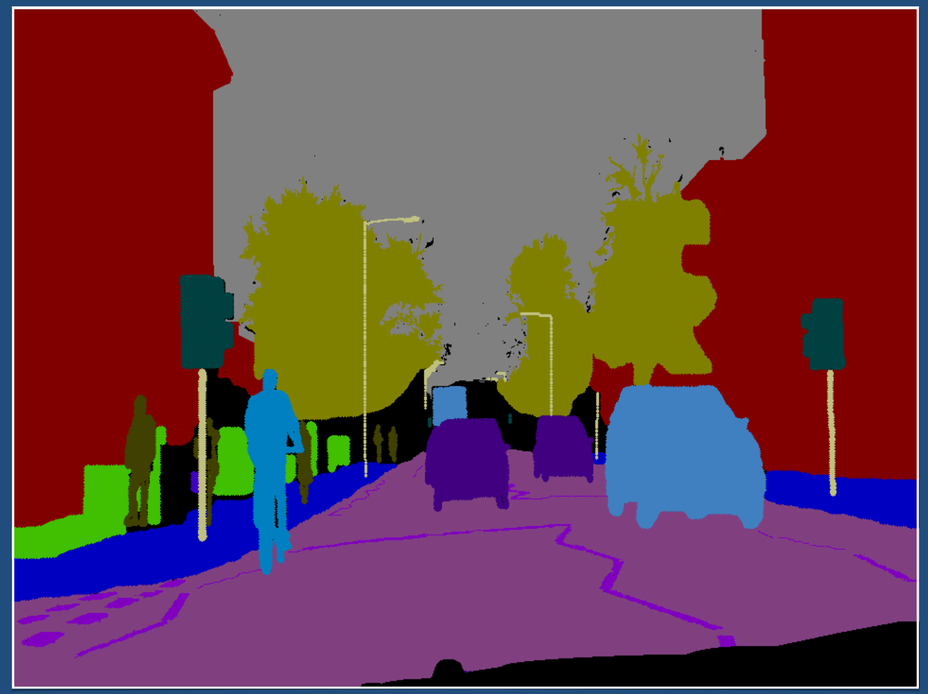


Image borrowed by CamVid dataset

CamVid

- Real world video dataset
- Over 700 images with ground truth labels
- 32 semantic classes

IDDA

- Synthetic dataset
- Over 1M FHD images with pixel wise semantic annotations
- 24 classes
- Large amount of scenarios

Project overview

- Baseline model: BiSeNet
- Domain adaptation: BiSeNet + Adversarial Learning
- Domain adaptation improvements: BiSeNet + Adversarial Learning + FDA

Bilateral Segmentation Network

BiSeNet

- Spatial path
- Context path
    - Backbone (lightweight model)
    - Global Average Pooling
    - Attention Refinement Module
- Feature fusion module
- Data augmentation

Experiment	Accuracy (%)	mIoU (%)	Training Time (avg per-epoch) (minutes)
BiSeNet(50 Epochs + ResNet-18)	0,859	0,602	04:58
BiSeNet(50 Epochs + ResNet-101)	0,872	0,650	04:34
BiSeNet(100 Epochs + ResNet-18)	0,867	0,620	04:57
BiSeNet(100 Epochs + ResNet-101)	**0,878**	**0,666**	04:56
BiSeNet(100 Epochs + ResNet-101) + Data augmentation	0,878	0,664	04:42
BiSeNet(100 Epochs + ResNet-101) + Adversarial learning	**0,673**	**0,308**	05:05
BiSeNet(100 Epochs + ResNet-101) + Adversarial learning + FDA	0,643	0,292	06:24

BiSeNet + Adversarial Learning

WHY Domain Adaptation

WHAT is Adversarial Learning

HOW we implemented it

Expected results

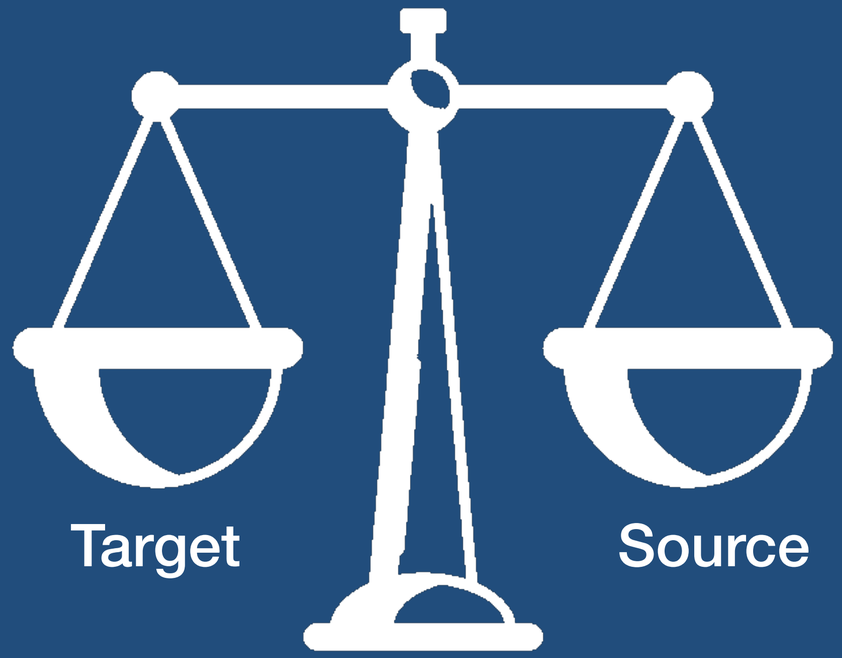
WHY Domain Adaptation

- Labeling: the crux of semantic segmentation
- Solution: Domain Adaptation
- Feature space VS Segmentation space

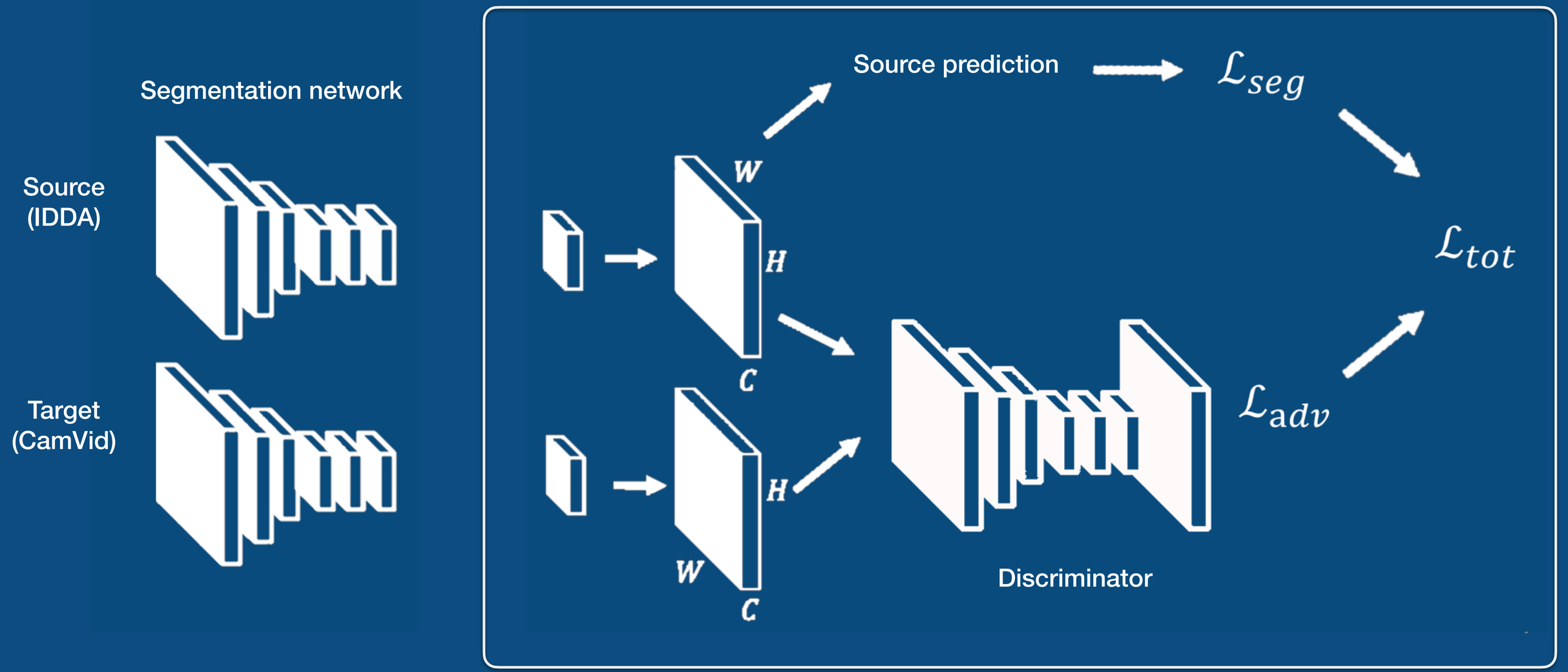
WHAT is Adversarial Learning

GAN approach

- Generator and Discriminator
- Aim of the Discriminator



HOW we implemented it



Expected results

Increment of:

- Accuracy
- mIoU = mean Intersection over Union

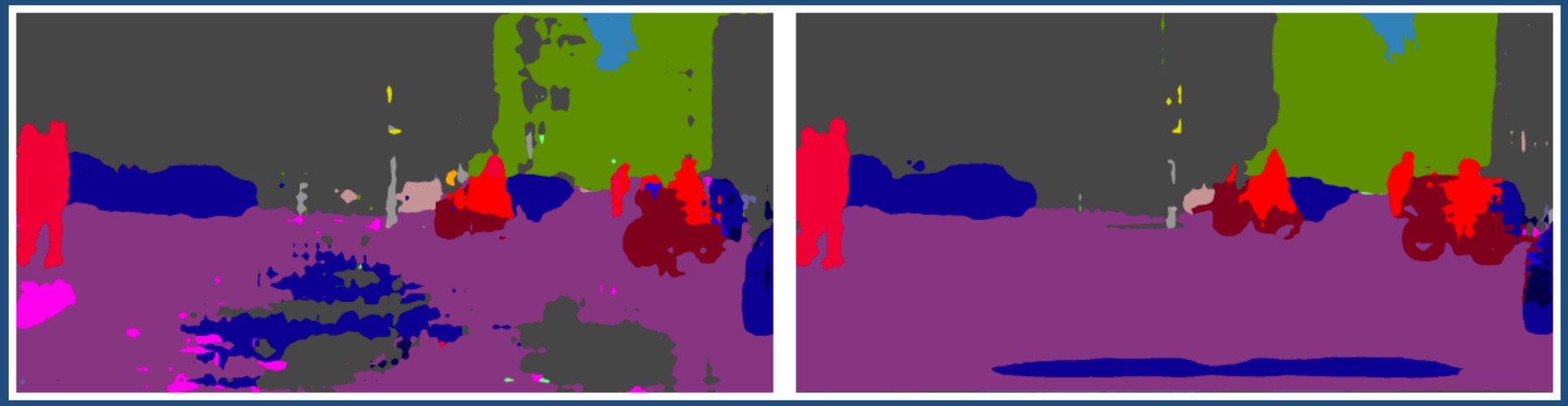
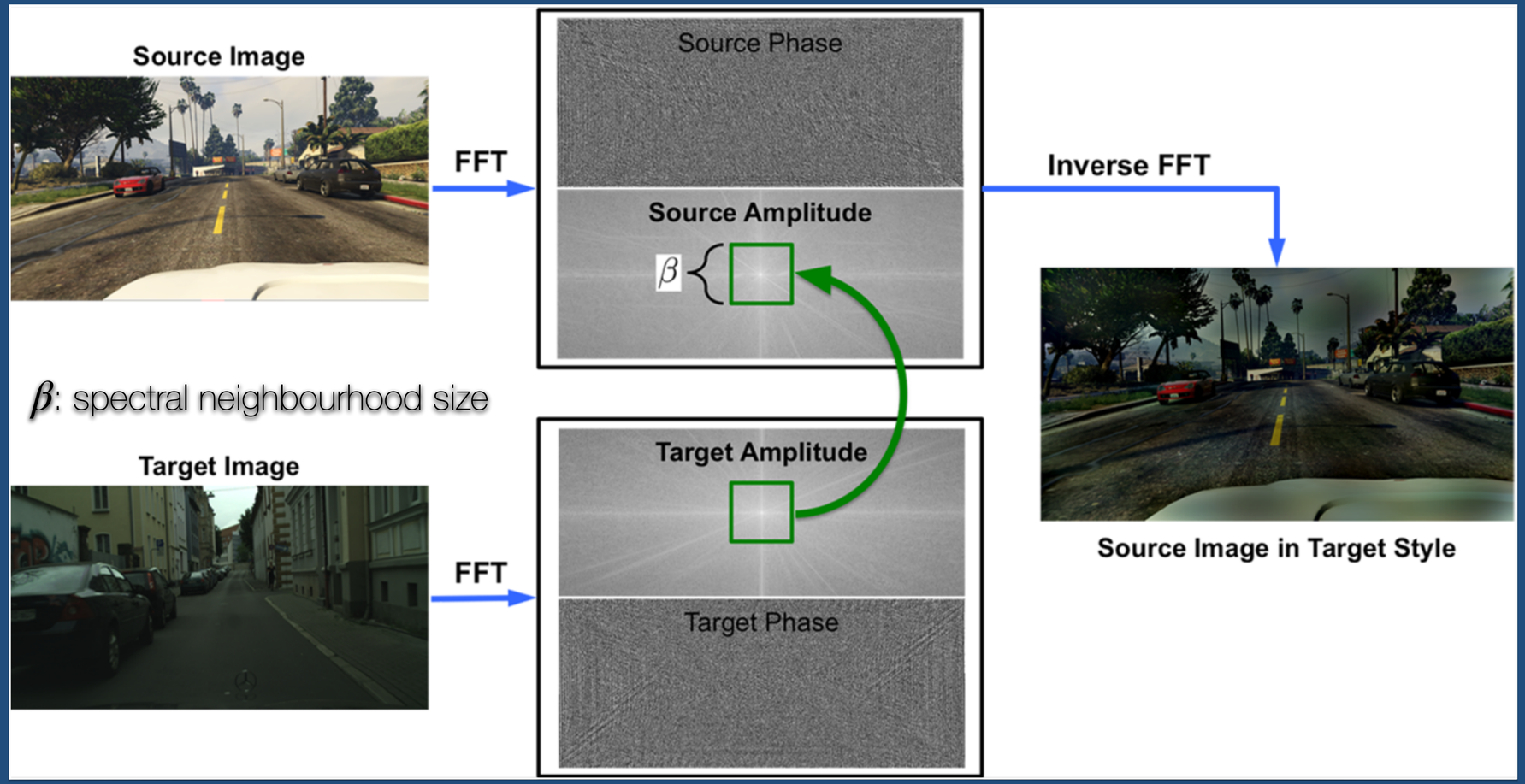


Photo taken from “Learning to adapt structured output space for semantic segmentation”, Y. Tsai, W. Hung, S. Schulter, K. Sohn, M. Yang, and M. Chandraker

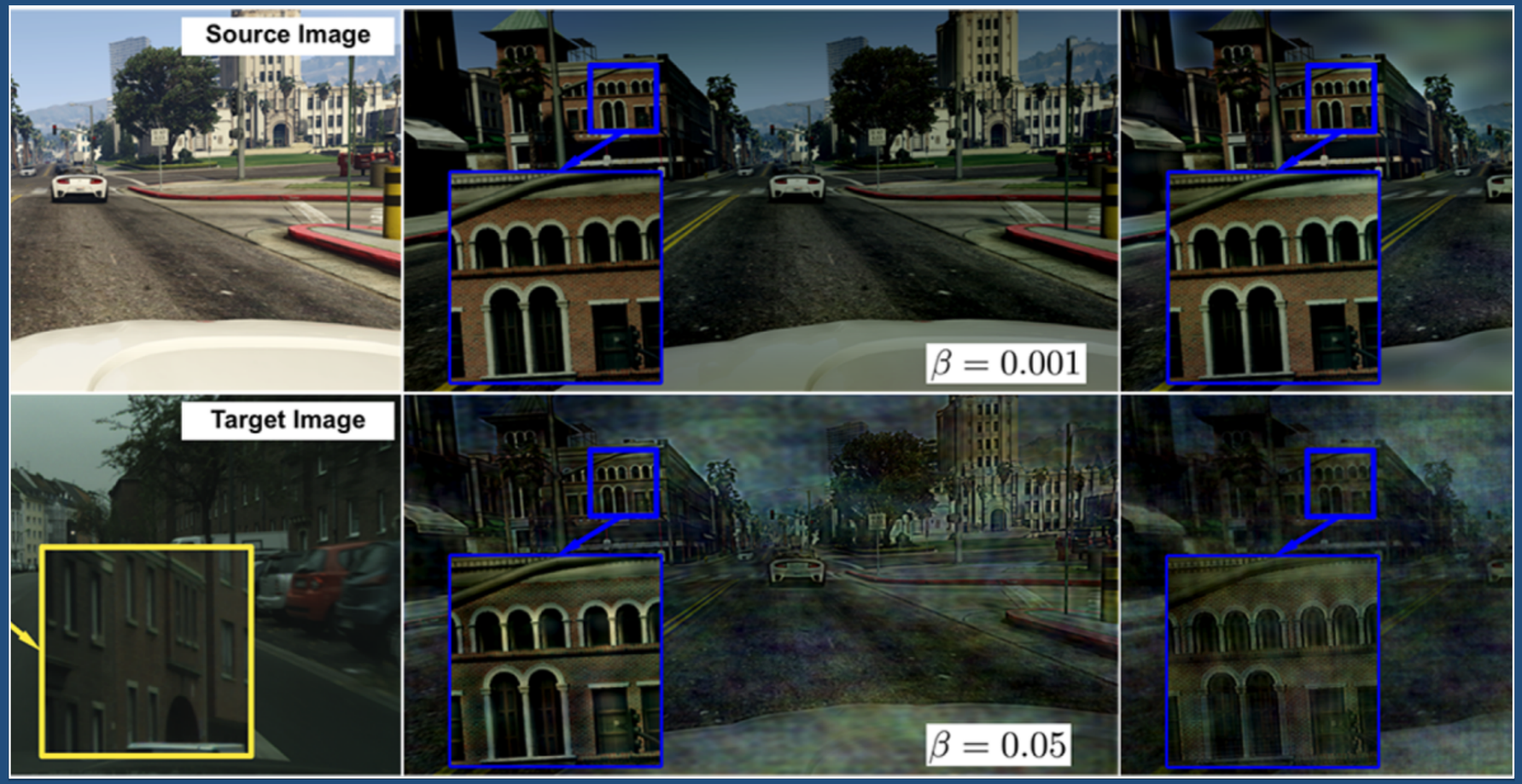
FDA: Fourier Domain Adaptation

- Unsupervised domain adaptation
- Reduce discrepancy between source and target domain
- Swap low frequency spectrums
- No training required



Spectral transfer: mapping a source image to a target "style" without altering semantic content

Yang and S. Soatto, “FDA: fourier domain adaptation for semantic segmentation,” CoRR, vol. abs/2004.05498, 2020.

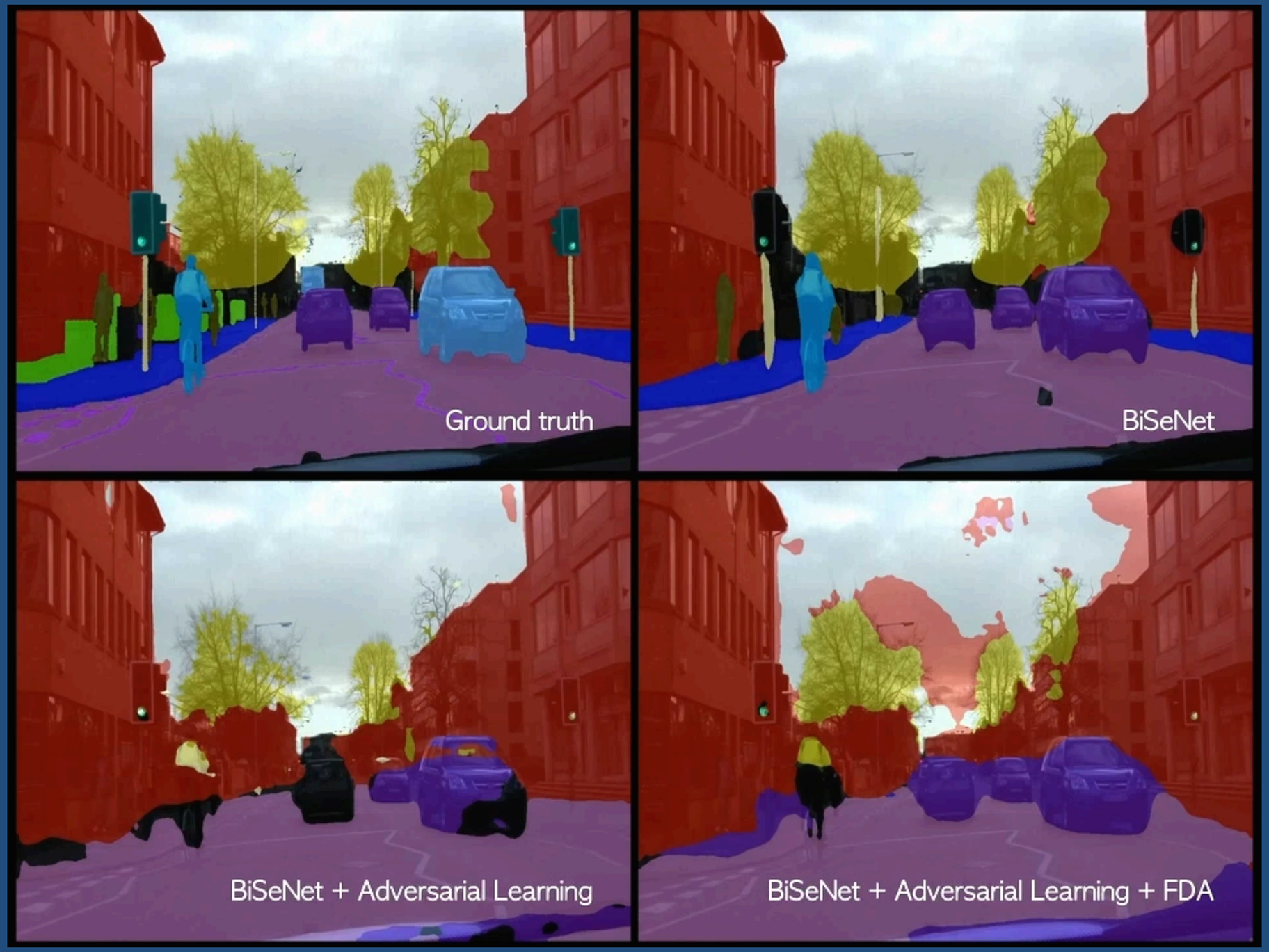


Effect of the size of the domain β swapping the spectrum

Yang and S. Soatto, “FDA: fourier domain adaptation for semantic segmentation,” CoRR, vol. abs/2004.05498, 2020.

Results

Experiment	Accuracy (%)	mIoU (%)	Training Time (avg per-epoch) (minutes)
BiSeNet(50 Epochs + ResNet-18)	0,859	0,602	04:58
BiSeNet(50 Epochs + ResNet-101)	0,872	0,650	04:34
BiSeNet(100 Epochs + ResNet-18)	0,867	0,620	04:57
BiSeNet(100 Epochs + ResNet-101)	**0,878**	**0,666**	04:56
BiSeNet(100 Epochs + ResNet-101) + Data augmentation	0,878	0,664	04:42
BiSeNet(100 Epochs + ResNet-101) + Adversarial learning	**0,673**	**0,308**	05:05
BiSeNet(100 Epochs + ResNet-101) + Adversarial learning + FDA	0,643	0,292	06:24


Ground truth
BiSeNet
BiSeNet + Adversarial Learning
BiSeNet + Adversarial Learning + FDA

**“The journey of a thousand miles begins with a single step…
batch after batch, with SGD. Usually”**

-Lao Tzu et al.