VisDA Challenge 3rd place

# Adversarial Domain Adaptation for Semantic Segmentation

Wei-Chih Hung[1], Yi-Hsuan Tsai[2], Ming-Hsuan Yang[1]

[1]UC Merced, [2]NEC Labs America

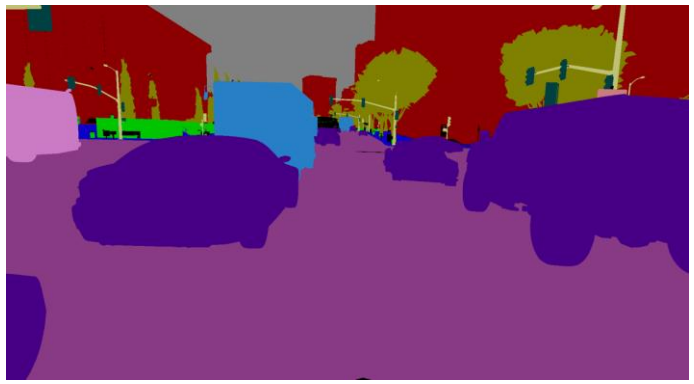
UNIVERSITY OF CALIFORNIA UCMERCED

NEC Laboratories America

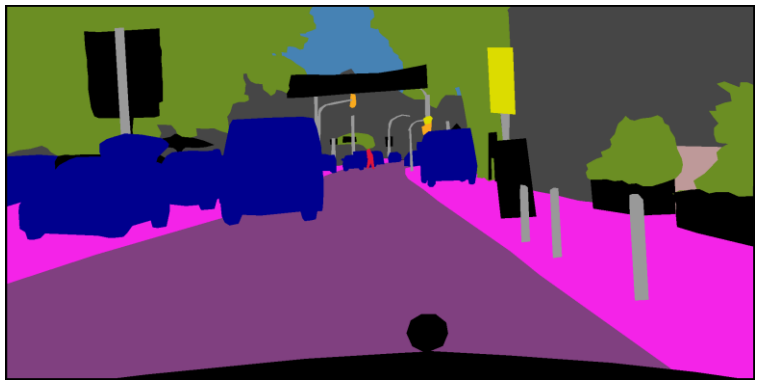
*Relentless* passion for innovation

# VisDA 2017 Challenge

Source Dataset: GTA5



Target Dataset (Validation): Cityscapes



Target Dataset (Challenge)

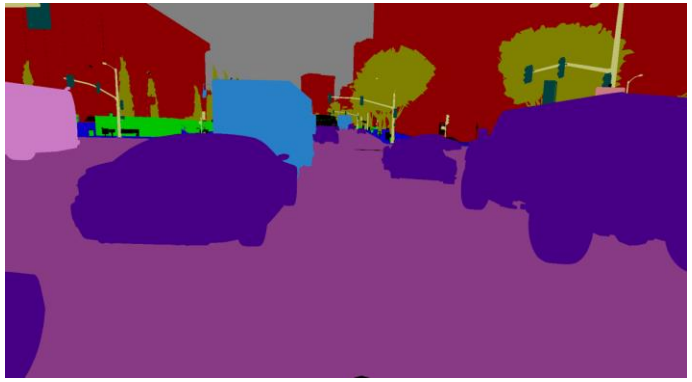
# Motivation

Source

Input images are different

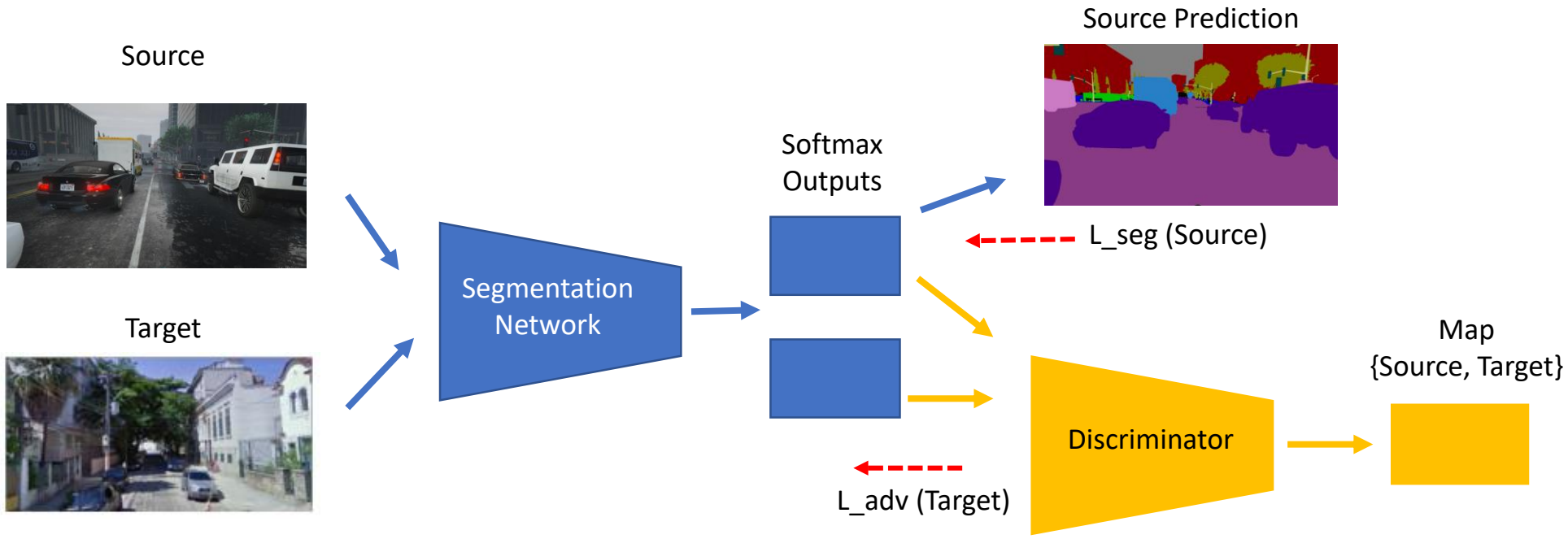
Target

**Idea: make the predictions on both datasets look similar**



Ground truth labels' layout distributions are similar

# Adversarial Learning for Domain Adaptation



# Implementation Details

- We use PyTorch
- Baseline model: DeepLab-v2 without multi-scale
  - ResNet-101
  - Pretrained only with ImageNet
  - ~65% mean IOU on Cityscapes
- It is essential to balance:
  - Segmentation network and discriminator
  - L_seg and L_adv

# Experimental Results

- GTA -> Cityscapes

| | Baseline | Adapt |
|---|---|---|
| Mean IOU | 32.33 | 42.44 |

- GTA -> Test Set

| | Baseline | Adapt |
|---|---|---|
| Mean IOU | 30.3 | 42.4 |

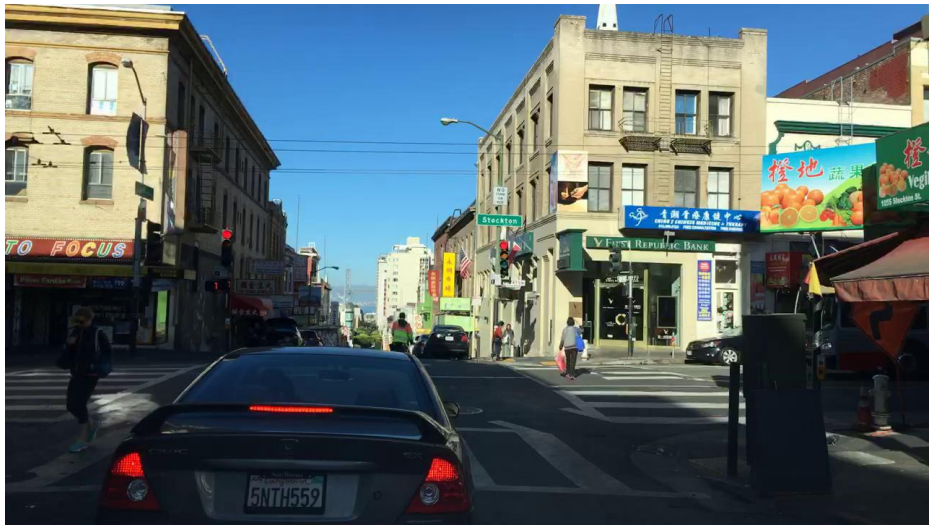
# Detailed Class Performance on Test Set

| # | User | Team Name | road | | sdwlk | | bldng | | wall | | fence | | pole | | light | | sign | | vgttn | | trrn | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | src | adapt | s | a | s | a | s | a | s | a | s | a | s | a | s | a | s | a | s | a |
| 3 | whung | VLLAB | 30.8 | 87.2 | 16.3 | 33.3 | 44.7 | 70.2 | 2.6 | 13.6 | 20.5 | 27.8 | 28.2 | 29.3 | 34.4 | 32.9 | 27.9 | 27.9 | 67.8 | 77.2 | 11.4 | 28.6 |

| # | User | Team Name | sky | | person | | rider | | car | | truck | | bus | | train | | mcycl | | bcycl | | MeanIoU | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | s | a | s | a | s | a | s | a | s | a | s | a | s | a | s | a | s | a | s | a |
| 3 | whung | VLLAB | 70.0 | 90.3 | 43.2 | 47.0 | 23.0 | 35.7 | 63.5 | 78.0 | 10.6 | 24.8 | 7.8 | 18.0 | 8.4 | 9.1 | 41.0 | 37.4 | 23.8 | 38.1 | 30.3 | 42.4 |

- Improve 17 classes
- 11 classes have improvement over 10%
- 2 classes (turn light, motorcycle) perform a bit worse

# Visualization




baseline

adapt

# Visualization

baseline

adapt

# Visualization

baseline

adapt

# Conclusions

- Adversarial learning can help domain adaptation without any hand-crafted criterions
- Our designed model is end-to-end, one-stage training, and can be adapted to other segmentation networks
- During inference, there is no extra computation