Semantic Segmentation and Adversarial Domain Adaptation

Student: Yifan Bao Rui Feng

Professor: Min Chi

Introduction

Image Segmentation: Deep learning-based image segmentation models, like **deeplab family**, often achieving the highest accuracy, have received a lot of attention.

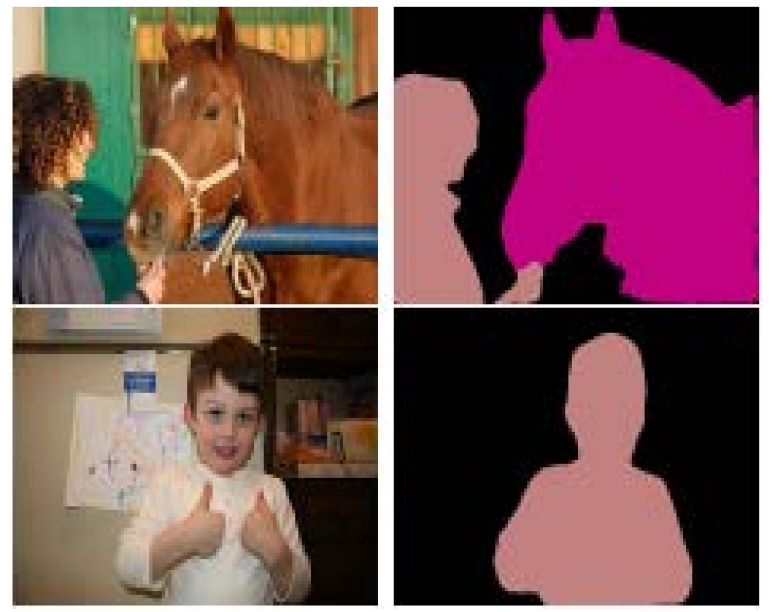

Figure1: Examples of Image Segmentation.

Domain Adaptation: To reduce the labeling cost, unsupervised domain adaptation (UDA) approaches are proposed to transfer knowledge from labeled synthesized datasets to unlabeled real-world datasets.

Datasets

Cityscapes: It includes 5,000 annotated images with 2048 × 1024 resolution, grouped into 8 categories.

SYNTHIA: We use the SYNTHIA-RANDCITYSCAPES subset consisting of 9,400 1280×760 synthetic images.

Metrics

Intersection over Union (IoU) is defined as

$$IoU = \frac{Intersection}{Union} = \frac{TP}{TP + FP + FN}$$

where TP (true positive) represents a pixel that is correctly predicted to belong to the given class. Similar as FP and FN.

Mean-IoU (mIoU) is defined as the average IoU over all classes.

$$mIoU == \frac{\sum_n IoU_i}{n}$$

where n is the number of classes and IoU_i is the IoU metric for the i-th class

Model Framework

AdaptSegNet: Figure2 shows that AdaptSegNet consists of two parts: one is the **segmentation network**, the other is the **domain adaptation module.** We use DeepLabV2 and DeepLabV3+ as the segmentation network separately

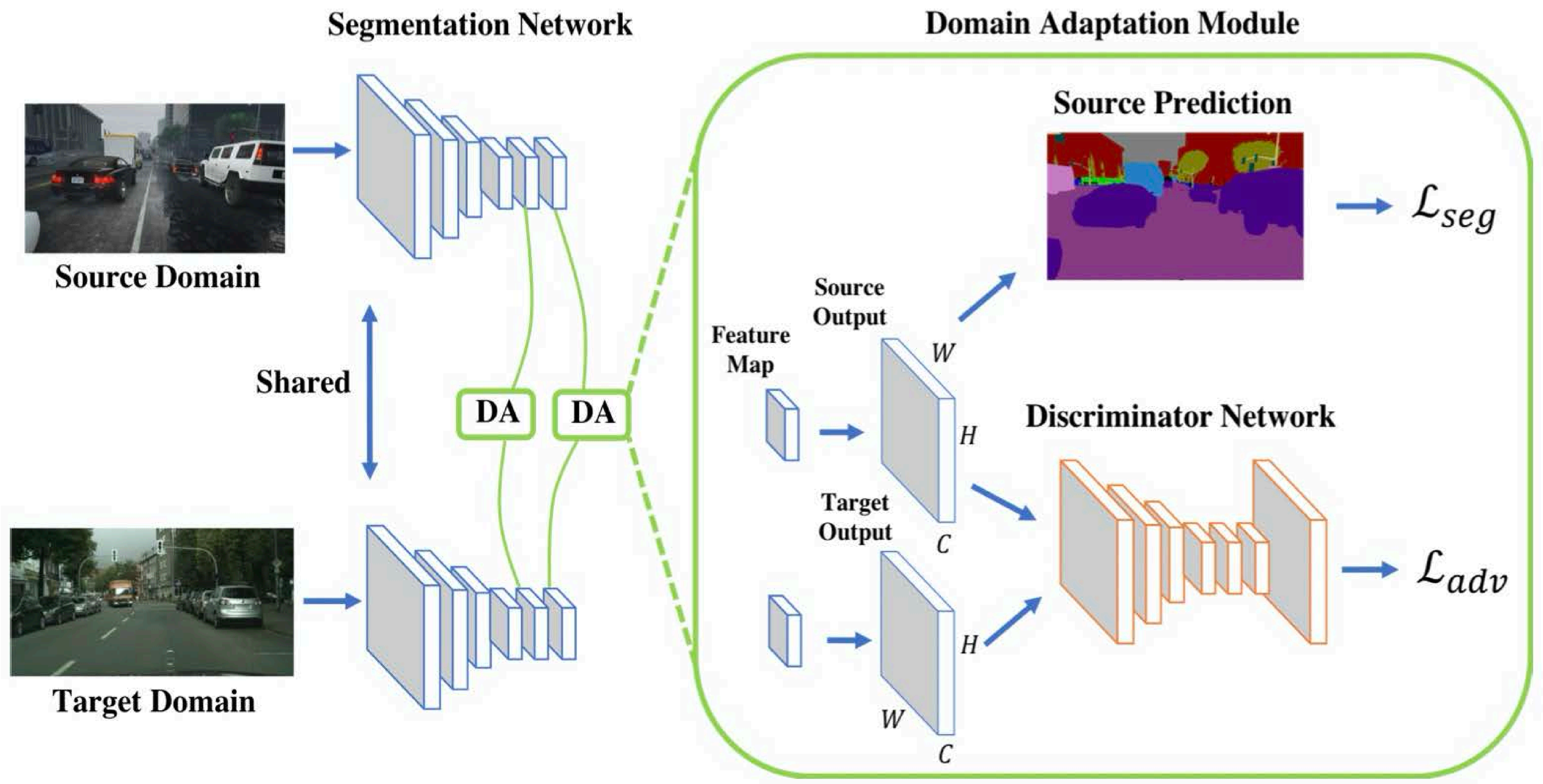

Figure2: AdaptSegnet Architecture .

Qualitativie and Quantitative Results

Figure3 shows the training results of DeeplabV2 model on SYNTHIA dataset.

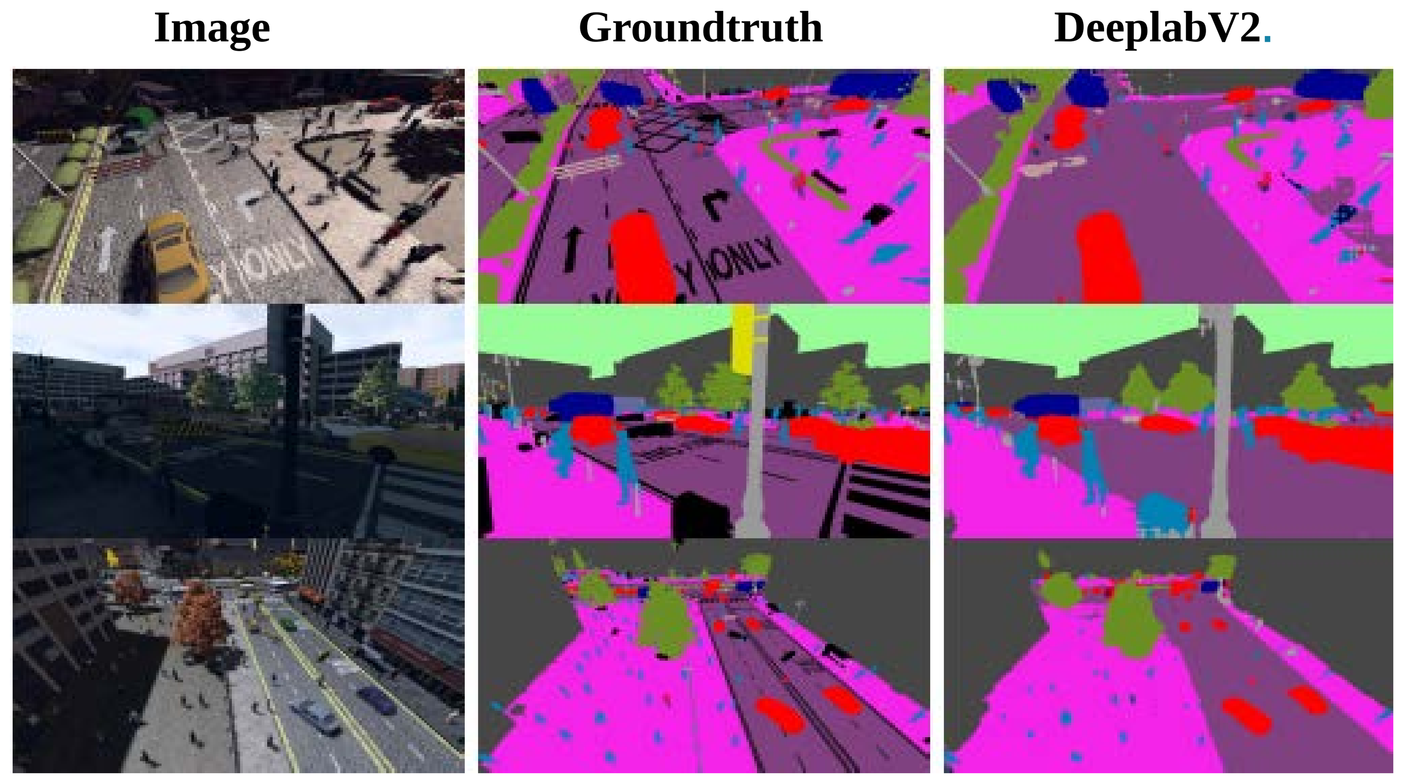

Figure3: Segmentation Result of DeeplabV2 on SYNTHIA.

Qualitativie and Quantitative Results (Cond)

Figure4 compares the segmentation results with and without domain adaptation. We can see that adversarial training along with DeeplabV3+ achieves the best visual results on target domain.

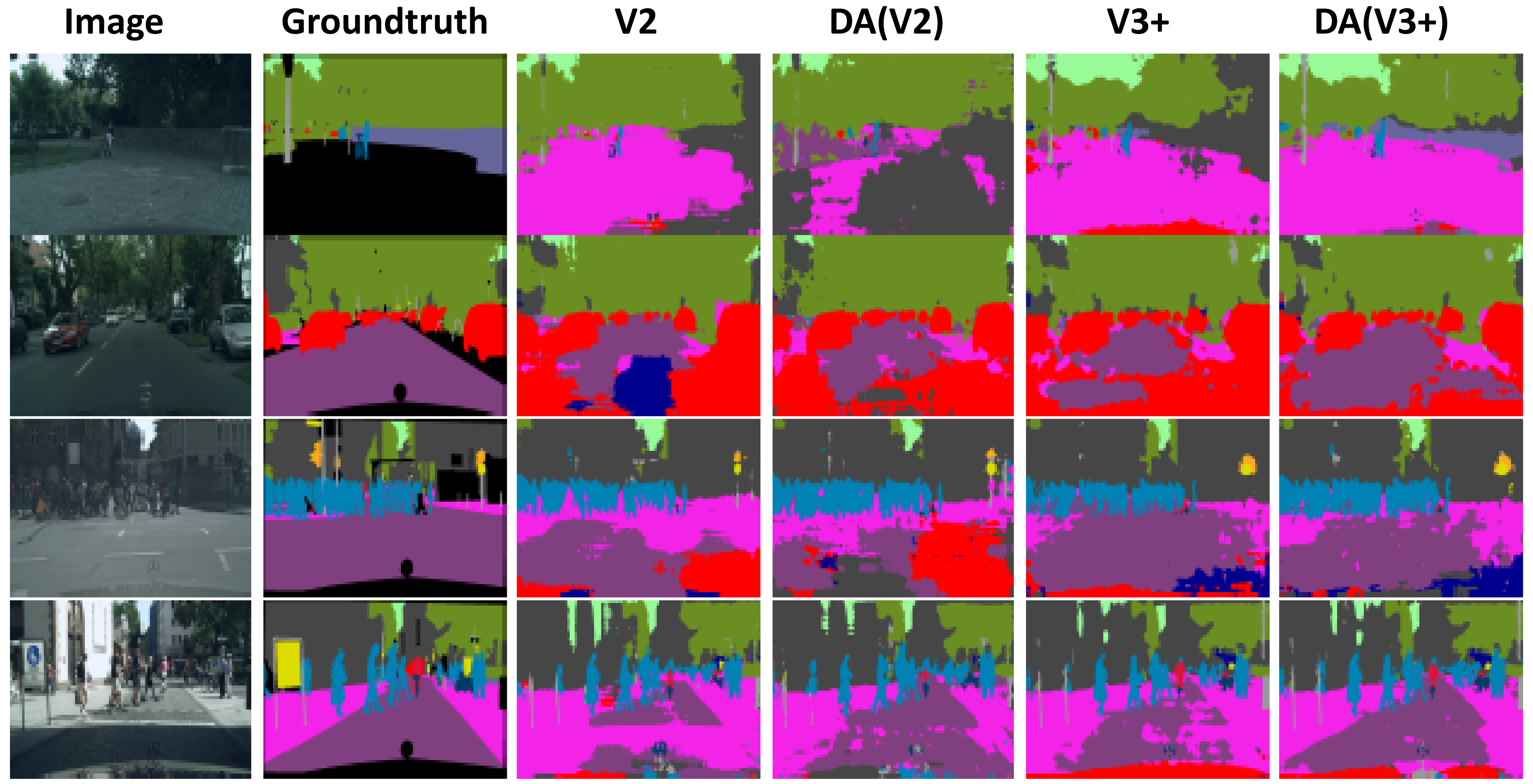

Figure4: Qualitative Image Segmentation Results on the Cityscapes Dataset.

Table1 is our quantitative results. As expected, the adaptation model along with DeeplabV3+ have the best mIoU(mean intersection over union) results.

Table1: Quantitative Results for SYNTHIA-to-Cityscapes experiments

Method	Seg Model	building	walls*	veg	sky	person	motor	bike	mIoU (%)	mIoU* (%)
Source only	DeeplabV2	75.8	6.4	**79.0**	**81.1**	**57.6**	**20.8**	28.2	34.1	39.4
Source only	DeeplabV3+	74.8	5.3	74.0	71.9	52.0	10.0	22.5	33.2	38.6
AdaptSegNet	DeeplabV2	64.9	4.7	76.9	76.3	52.8	16.0	**34.7**	35.3	41.1
AdaptSegNet	DeeplabV3+	**76.8**	**10.1**	75.7	73.4	53.1	19.1	29.9	**37.0**	**42.8**

Conclusion

- For segmentation before adaptation, DeeplabV2 performs better due to our limited time for fine-tuning DeeplabV3+ model.
- Results show that domain adaptation with DeeplabV3+ model performs favorably against DeeplabV2(which is used in the original paper) even it is less accurate before adaptation.
- The adversarial domain adaptation works well on big datasets with relatively big domain discrepancy.