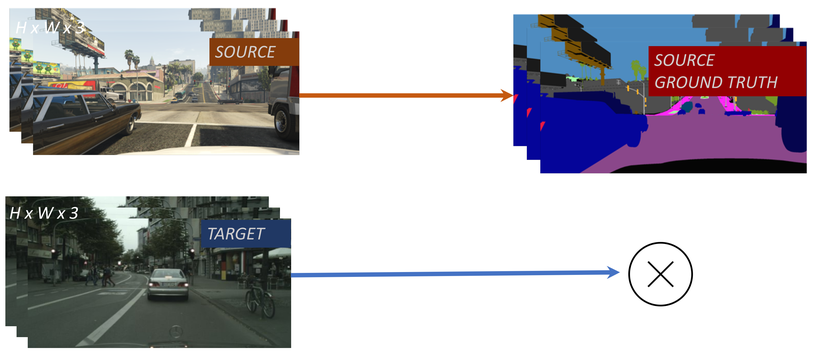
JOINT GLOBAL-LOCAL ALIGNMENT FOR DOMAIN ADAPTIVE SEMANTIC SEGMENTATION

Sudhir Yarram[1], Ming Yang[2], Junsong Yuan[1], Chuming Qiao[1]

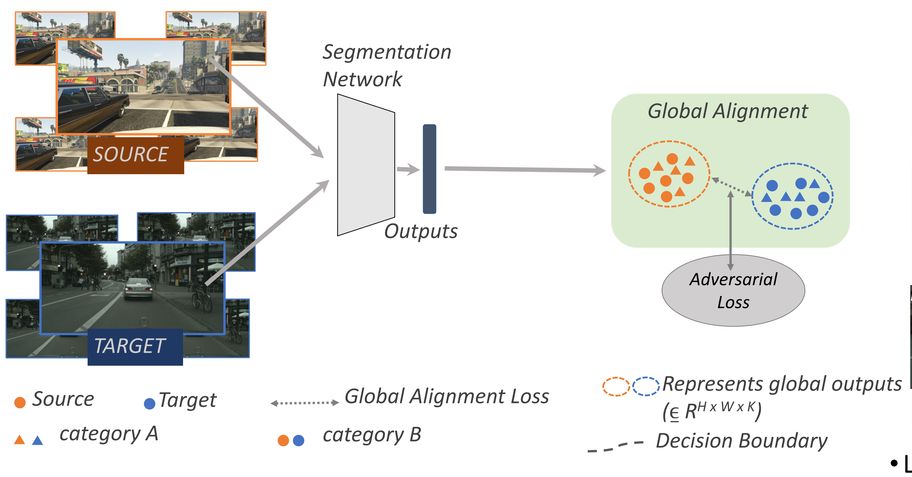
[1]University at Buffalo [2]Horizon Robotics

1 Unsupervised Domain Adaptation for Semantic Segmentation



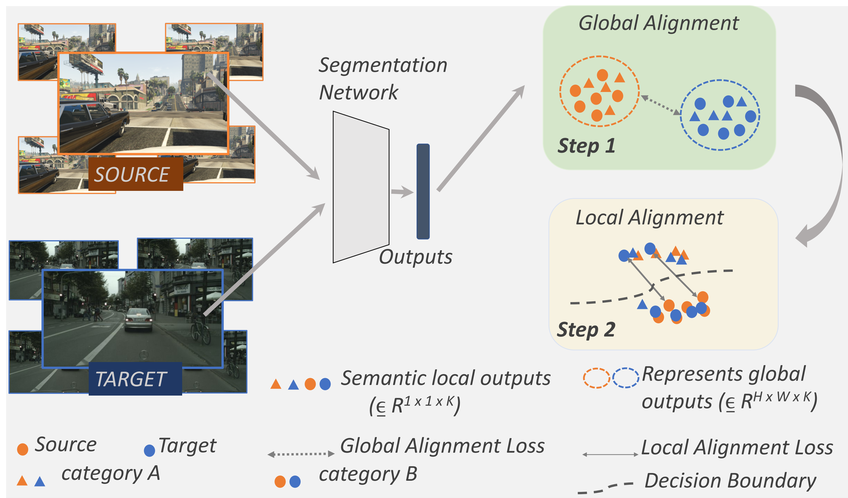
- Unsupervised domain adaptation deals with leveraging the large labeled data available in source domain to achieve good performance in the target domain.

2 Previous Methods



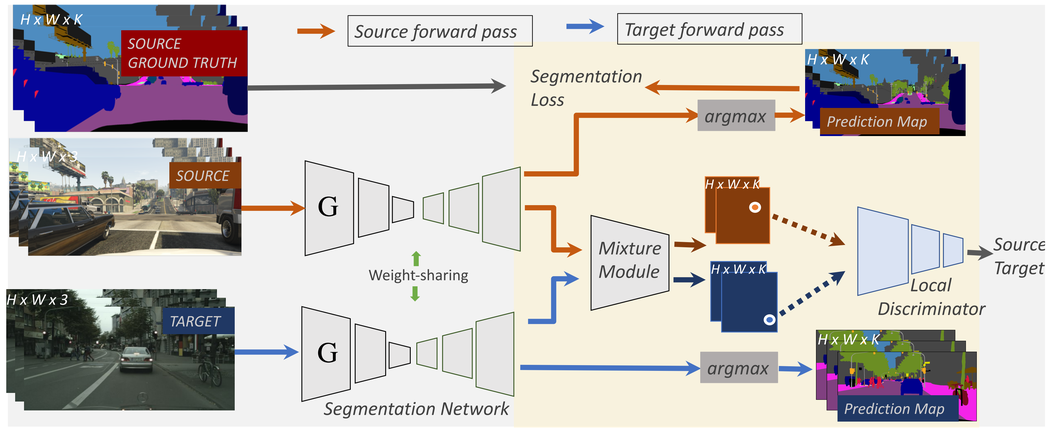
- Drawbacks
  - Inherent high dimensionality of global features involved in global alignment does not ensure alignment of various local semantic outputs.

3 Our Approach



- We propose a two step approach. In the first step, we conduct global alignment using popular adversarial loss.
- Further, we conduct local alignment by aligning the local pixel level features which can further reduce the domain gap by aligning semantically encoded local outputs.

4 Our Framework



- Local adversarial alignment.
- It consists of a Segmentation Model G and a local discriminator D along with a mixture module.
- For an input source image, the output map is used to calculate adversarial loss as well as the segmentation loss.
- For an input target image, the output map is used to calculate the adversarial loss.

5 Experiments: Comparison with state-of-the-arts

Method	Arch.	GTA2City mIoU(%)	SYN2City mIoU(%)
AdapSegNet [3]	V	35.0	37.6
Our AdapSegNet	V	**36.8** ± 0.2 **(+1.8)**	**38.9** ± 0.3 **(+1.3)**
BDL [21]	V	41.3	46.1
Our BDL	V	**43.5** ± 0.2 **(+2.2)**	**47.7** ± 0.2 **(+1.6)**
Source only	-	36.6	38.6
AdapSegNet [3]	R	41.4	45.9
Our AdapSegNet	R	**45.2** ± 0.1 **(+3.8)**	**46.9** ± 0.3 **(+1.0)**
ADVENT [11]	R	43.8	47.6
Our ADVENT	R	**46.4** ± 0.3 **(+2.6)**	**48.3** ± 0.4 **(+0.7)**
BDL [21] $M_2^{(2)}(F^{(2)})$	R	48.5	51.4
Our BDL	R	**49.4** ± 0.2 **(+0.9)**	**52.5** ± 0.3 **(+1.1)**

- Semantic segmentation performance in mIoU(%) on GTA5 to Cityscapes (GTA2City) and Synthia to Cityscapes (SYN2City) Adaptation task.
- 'R' means the ResNet-101 and 'V' means the VGG-16 backbone. Our global-local alignment approach shows consistent improvement over baselines.

References:

[3] Learning to Adapt Structured Output Space for Semantic Segmentation, CVPR'18

[11] ADVENT: Adversarial Entropy Minimization for Domain Adaptation in Semantic Segmentation, CVPR'19

[21] Bidirectional Learning for Domain Adaptation of Semantic Segmentation, CVPR'19

Github: github.com/skrya/globallocal