Rensselaer
why not change the world?®

Rensselaer
Radiation Measurement & Dosimetry Group

Multi-source Domain Adaptation for Semantic Segmentation

Huihua Yang | 02/05/2020

Contents

Multi-source Domain Adaptation for Semantic Segmentation:

1. Introduction

1. **Semantic segmentation** is Widely used in many applications
2. **Limitations on several semantic segmentation methods:** data needed are expensive and time-consuming to obtain
3. **Domain shift** or dataset bias

Domain adaptation (DA) or knowledge transfer techniques: mitigate the gap between different domains but do not consider a more practical scenario

unsupervised domain adaptation (UDA): single source
multi-source DA (MDA) :only focus on feature-level alignment

The author's method: MDA methods, feature level and pixel-level.

2. Problem setup

M sources domains: $\boldsymbol{S_1}, \boldsymbol{S_2}, \cdots, \boldsymbol{S_M}$:

labeled;

$X_i = \{x_i^j\}_{j=1}^{N_i}$: The observed data from S_i;

$Y_i = \{y_i^j\}_{j=1}^{N_i}$: The corresponding labels data drawn from source distribution $P_i(x, y)$, N_i is the number of samples in S_i;

One target domain T :

unlabeled;

$X_T = \{x_T^j\}_{j=1}^{N_T}$: the target data drawn from the target distribution $P_T(x, y)$ without label observation, N_T is the number of target samples

Two assumption:

1. homogeneity: The data from different domains are observed in the same image space but with different distributions;

(2) closed set: All the domains share the same space of classes.

3. Multi-source Adversarial Domain Aggregation Network: MADAN

Dynamic Adversarial Image Generation (DAIG): keep semantic information, close to target domain.

Adversarial Domain Aggregation (ADA): More aggregated unified domain

Feature-aligned Semantic Segmentation (FSS): domain invariant in feature-level.

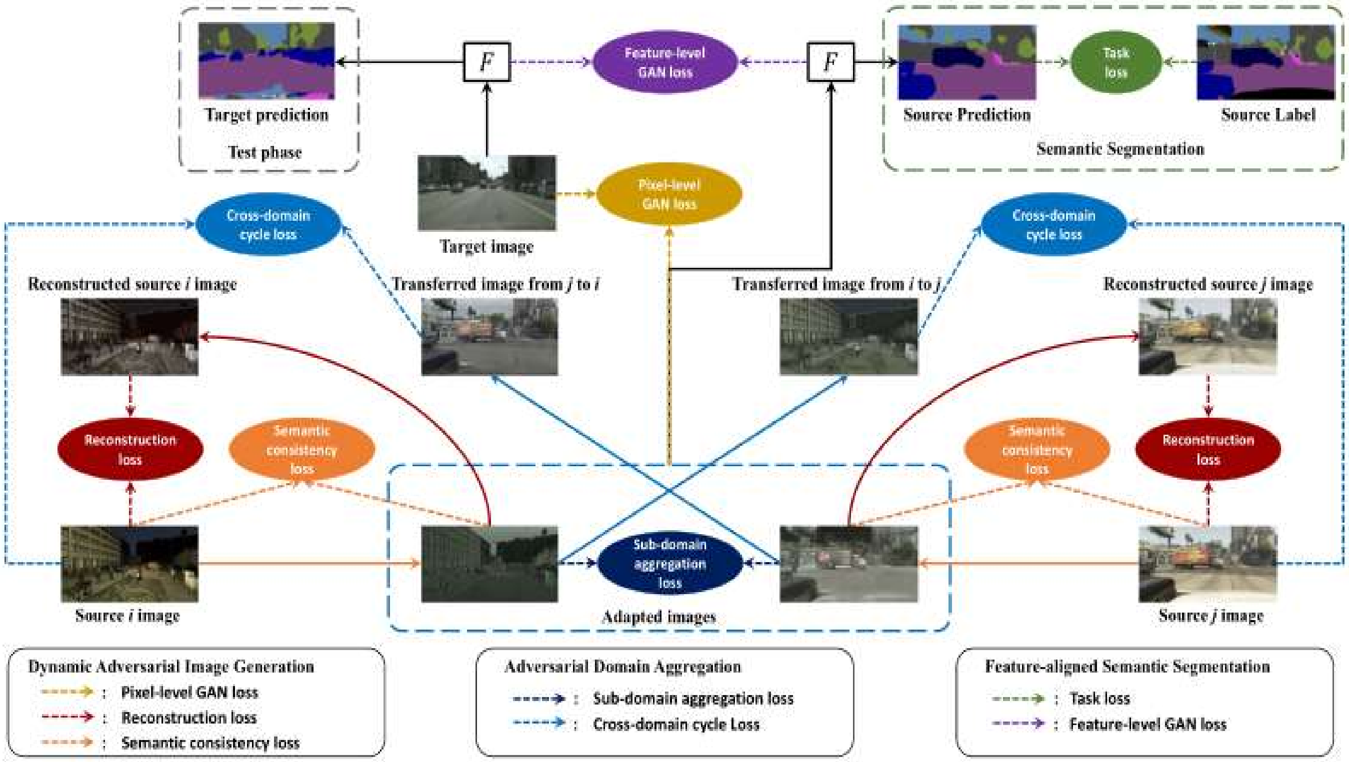

Figure 1: The framework of the proposed Multi-source Adversarial Domain Aggregation Network (MADAN). The colored solid arrows represent generators, while the black solid arrows indicate the segmentation network F. The dashed arrows correspond to different losses.

3. Multi-source Adversarial Domain Aggregation Network : MADAN

Dynamic Adversarial Image Generation (DAIG): keep semantic information, close to target domain.

GAN:

Generator $\boldsymbol{G_{S_i \to T}}$: generate adapted images that fool D_T from target T;
Discriminator $\boldsymbol{D_T}$: trained simultaneously with each $G_{S_i \to T}$ to classify real target images X_T from adapted images $G_{S_i \to T}$.

$$\mathcal{L}_{GAN}^{S_i \to T}(G_{S_i \to T}, D_T, X_i, X_T) = \mathbb{E}_{\mathbf{x}_i \sim X_i} \log D_T(G_{S_i \to T}(\mathbf{x}_i)) + \mathbb{E}_{\mathbf{x}_T \sim X_T} \log[1 - D_T(\mathbf{x}_T)]. \quad (1)$$

GAN:

Generator $\boldsymbol{G_{T \to S_i}}$: Inverse mapping of $\boldsymbol{G_{S_i \to T}}$;
Discriminator $\boldsymbol{D_i}$: classify X_i from $G_{T \to S_i}$.

$$\mathcal{L}_{GAN}^{T \to S_i}(G_{T \to S_i}, D_i, X_T, X_i) = \mathbb{E}_{\mathbf{x}_i \sim X_i} \log[1 - D_i(\mathbf{x}_i)] + \mathbb{E}_{\mathbf{x}_t \sim X_T} \log D_i(G_{T \to S_i}(\mathbf{x}_t)). \quad (2)$$

3. Multi-source Adversarial Domain Aggregation Network : MADAN

Dynamic Adversarial Image Generation (DAIG): keep semantic information, close to target domain.

The cycle-consistency loss:

$$\begin{aligned}\mathcal{L}_{cyc}^{S_i\leftrightarrow T}(G_{S_i\rightarrow T}, G_{T\rightarrow S_i}, X_i, X_T) =& \mathbb{E}_{\mathbf{x}_i\sim X_i} \parallel G_{T\rightarrow S_i}(G_{S_i\rightarrow T}(\mathbf{x}_i)) - \mathbf{x}_i \parallel_1 + \\ & \mathbb{E}_{\mathbf{x}_T\sim X_T} \parallel G_{S_i\rightarrow T}(G_{T\rightarrow S_i}(\mathbf{x}_t)) - \mathbf{x}_t \parallel_1 .\end{aligned} \tag{3}$$

Dynamic semantic consistency (DSC) loss

$$\mathcal{L}_{sem}^{S_i}(G_{S_i\rightarrow T}, X_i, F_i, F_A) = \mathbb{E}_{\mathbf{x}_i\sim X_i} KL(F_A(G_{S_i\rightarrow T}(\mathbf{x}_i))||F_i(\mathbf{x}_i)), \tag{4}$$

where $KL(\cdot||\cdot)$ is the KL divergence between two distributions.

$\boldsymbol{F_i}$: A segmentation model pretrained on ($\boldsymbol{X_i}$, $\boldsymbol{Y_i}$) ;
$\boldsymbol{F_A}$: The task segmentation model trained on the adapted domain

3. Multi-source Adversarial Domain Aggregation Network : MADAN

Adversarial Domain Aggregation (ADA):

More aggregated unified domain

The sub-domain aggregation discriminator (SAD)

$$\mathcal{L}_{SAD}^{S_i}(G_{S_1\to T},\ldots G_{S_i\to T},\ldots,G_{S_M\to T},D_A^i) = \mathbb{E}_{\mathbf{x}_i\sim X_i}\log D_A^i(G_{S_i\to T}(\mathbf{x}_i))+ \frac{1}{M-1}\sum\nolimits_{j\neq i}\mathbb{E}_{\mathbf{x}_j\sim X_j}\log[1-D_A^i(G_{S_j\to T}(\mathbf{x}_j))]. \tag{5}$$

The cross-domain cycle discriminator (CCD)

$$\mathcal{L}_{CCD}^{S_i}(G_{T\to S_1},\ldots G_{T\to S_{i-1}},G_{T\to S_{i+1}},\ldots,G_{T\to S_M},G_{S_i\to T},D_i) = \mathbb{E}_{\mathbf{x}_i\sim X_i}\log D_i(\mathbf{x}_i)+ \frac{1}{M-1}\sum\nolimits_{j\neq i}\mathbb{E}_{\mathbf{x}_j\sim X_j}\log[1-D_i(G_{T\to S_i}((G_{S_j\to T}(\mathbf{x}_j)))]. \tag{6}$$

3. Multi-source Adversarial Domain Aggregation Network : MADAN

Feature-aligned Semantic Segmentation (FSS): domain invariant in feature-level.

The cross-entropy loss:

$$\mathcal{L}_{task}(F, X', Y) = -\mathbb{E}_{(\mathbf{x}', \mathbf{y}) \sim (X', Y)} \sum_{l=1}^{L} \sum_{h=1}^{H} \sum_{w=1}^{W} \mathbb{1}_{[l=\mathbf{y}_{h,w}]} \log(\sigma(F_{l,h,w}(\mathbf{x}'))), \quad (7)$$

where L is the number of classes, H, W are the height and width of the adapted images, σ is the softmax function, $\mathbb{1}$ is an indicator function, and $F_{l,h,w}(\mathbf{x}')$ is the value of $F(\mathbf{x}')$ at index (l, h, w).

$$X' = \bigcup_{i=1}^{M} X'_i$$

$$Y = \bigcup_{i=1}^{M} Y_i.$$

The feature-level GAN loss

$$\mathcal{L}_{feat}(F_f, D_{F_f}, X', X_T) = \mathbb{E}_{\mathbf{x}' \sim X'} \log D_{F_f}(F_f(\mathbf{x}')) + \mathbb{E}_{\mathbf{x}_T \sim X_T} \log[1 - D_{F_f}(F_f(\mathbf{x}_T))], \quad (8)$$

where $F_f(\cdot)$ is the output of the last convolution layer (*i.e.* a feature map) of the encoder in F.

3. Multi-source Adversarial Domain Aggregation Network: MADAN

Dynamic Adversarial Image Generation (DAIG): keep semantic information, close to target domain.

Adversarial Domain Aggregation (ADA): More aggregated unified domain

Feature-aligned Semantic Segmentation (FSS): domain invariant in feature-level.

$$\begin{aligned}
&\mathcal{L}_{MADAN}(G_{S_1\to T}\cdots G_{S_M\to T}, G_{T\to S_1}\cdots G_{T\to S_M}, D_1\cdots D_M, D_A^1\cdots D_A^M, D_{F_f}, F)\\
&= \sum_i \Big[\mathcal{L}_{GAN}^{S_i\to T}(G_{S_i\to T}, D_T, X_i, X_T) + \mathcal{L}_{GAN}^{T\to S_i}(G_{T\to S_i}, D_i, X_T, X_i)\\
&+ \mathcal{L}_{cyc}^{S_i\leftrightarrow T}(G_{S_i\to T}, G_{T\to S_i}, X_i, X_T) + \mathcal{L}_{sem}^{S_i}(G_{S_i\to T}, X_i, F_i, F)\\
&+ \mathcal{L}_{SAD}^{S_i}(G_{S_1\to T}, \ldots G_{S_i\to T}, \ldots, G_{S_M\to T}, D_A^i)\\
&+ \mathcal{L}_{CCD}^{S_i}(G_{T\to S_1}, \ldots G_{T\to S_{i-1}}, G_{T\to S_{i+1}}, \ldots, G_{T\to S_M}, G_{S_i\to T}, D_i)\Big]\\
&+ \mathcal{L}_{task}(F, X', Y) + \mathcal{L}_{feat}(F_f, D_{F_f}, X', X_T).
\end{aligned} \tag{9}$$

The training process corresponds to solving for a target model F according to the optimization:

$$F^* = \arg\min_F \min_D \max_G \mathcal{L}_{MADAN}(G, D, F), \tag{10}$$

4. Experiments & Results: The Visualization results

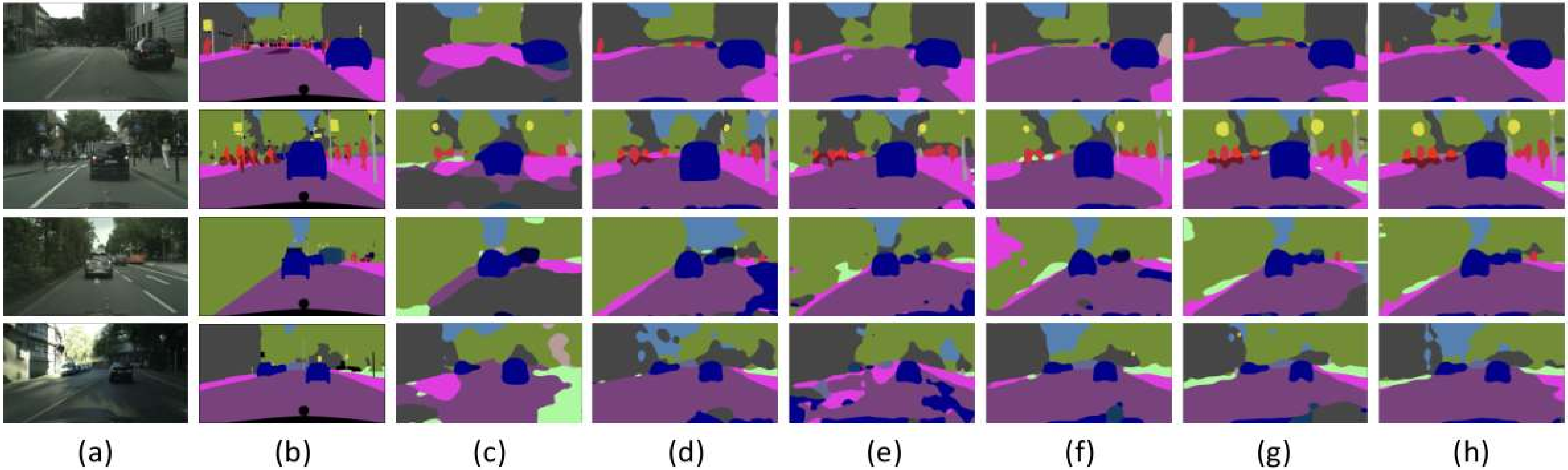

Figure 2: Qualitative semantic segmentation result from GTA and SYNTHIA to Cityscapes. From left to right are: (a) original image, (b) ground truth annotation, (c) source only from GTA, (d) CycleGANs on GTA and SYNTHIA, (e) +CCD+DSC, (f) +SAD+DSC, (g) +CCD+SAD+DSC, and (h) +CCD+SAD+DSC+Feat (MADAN).

4. Experiments & Results : Visualization

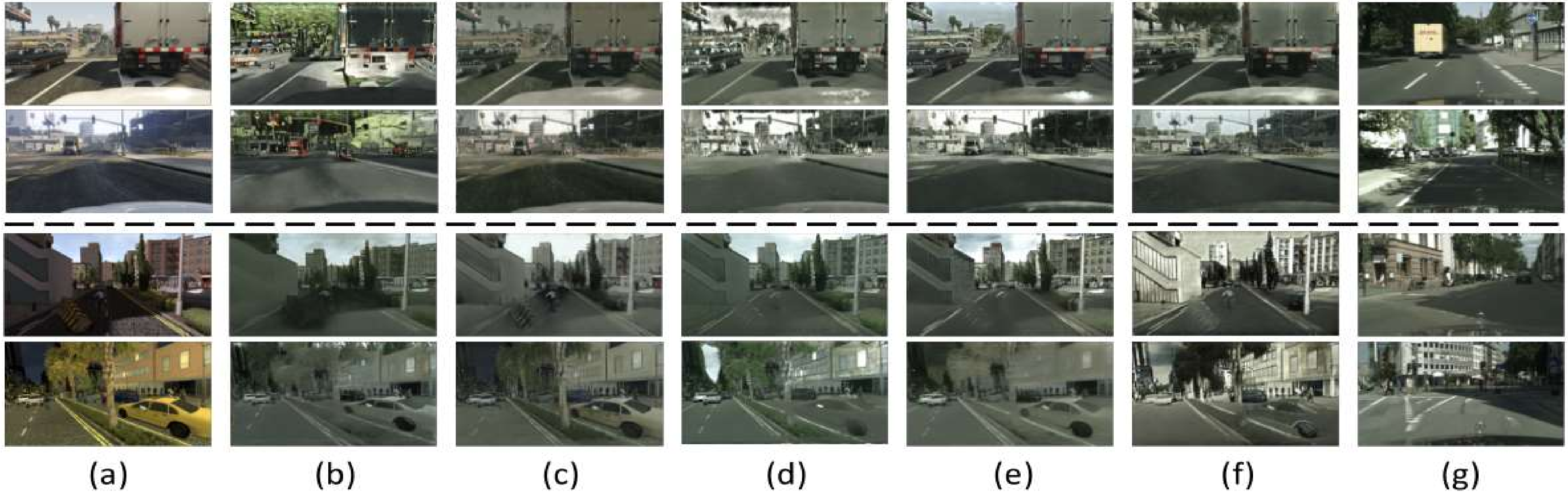

Figure 3: Visualization of image translation. From left to right are: (a) original source image, (b) CycleGAN, (c) CycleGAN+DSC, (d) CycleGAN+CCD+DSC, (e) CycleGAN+SAD+DSC, (f) CycleGAN+CCD+SAD+DSC, and (g) target Cityscapes image. The top two rows and bottom rows are GTA $\rightarrow$ Cityscapes and SYNTHIA $\rightarrow$ Cityscapes, respectively.

4. Experiments & Results : Comparison with State-of-the-art

The source-only method obtains the worst performance.

All adaptation methods perform better

MADAN achieves the highest mIoU score

Table 2: Comparison with the state-of-the-art DA methods for semantic segmentation from GTA and SYNTHIA to Cityscapes. The best class-wise IoU and mIoU trained on the source domains are emphasized in bold (similar below).

Standards	Method	road	sidewalk	building	wall	fence	pole	t-light	t-sign	vegettion	sky	person	rider	car	bus	m-bike	bicycle	mIoU
Source-only	GTA	54.1	19.6	47.4	3.3	5.2	3.3	0.5	3.0	69.2	43.0	31.3	0.1	59.3	8.3	0.2	0.0	21.7
	SYNTHIA	3.9	14.5	45.0	0.7	0.0	14.6	0.7	2.6	68.2	68.4	31.5	4.6	31.5	7.4	0.3	1.4	18.5
	GTA+SYNTHIA	44.0	19.0	60.1	11.1	13.7	10.1	5.0	4.7	74.7	65.3	40.8	2.3	43.0	15.9	1.3	1.4	25.8
GTA-only DA	FCN Wld [47]	70.4	32.4	62.1	14.9	5.4	10.9	14.2	2.7	79.2	64.6	44.1	4.2	70.4	7.3	3.5	0.0	27.1
	CDA [48]	74.8	22.0	71.7	6.0	11.9	8.4	16.3	11.1	75.7	66.5	38.0	9.3	55.2	18.9	16.8	14.6	28.9
	ROAD [50]	85.4	31.2	78.6	**27.9**	**22.2**	21.9	23.7	11.4	80.7	68.9	48.5	14.1	78.0	23.8	8.3	0.0	39.0
	AdaptSeg [71]	**87.3**	29.8	78.6	21.1	18.2	22.5	21.5	11.0	79.7	71.3	46.8	6.5	**80.1**	26.9	10.6	0.3	38.3
	CyCADA [32]	85.2	37.2	76.5	21.8	15.0	23.8	22.9	21.5	80.5	60.7	50.5	9.0	76.9	28.2	4.5	0.0	38.7
	DCAN [55]	82.3	26.7	77.4	23.7	20.5	20.4	**30.3**	15.9	**80.9**	69.5	**52.6**	11.1	79.6	21.2	17.0	6.7	39.8
SYNTHIA-only DA	FCN Wld [47]	11.5	19.6	30.8	4.4	0.0	20.3	0.1	11.7	42.3	68.7	51.2	3.8	54.0	3.2	0.2	0.6	20.2
	CDA [48]	65.2	26.1	74.9	0.1	0.5	10.7	3.7	3.0	76.1	70.6	47.1	8.2	43.2	20.7	0.7	13.1	29.0
	ROAD [50]	77.7	30.0	77.5	9.6	0.3	**25.8**	10.3	15.6	77.6	**79.8**	44.5	16.6	67.8	14.5	7.0	23.8	36.2
	CyCADA [32]	66.2	29.6	65.3	0.5	0.2	15.1	4.5	6.9	67.1	68.2	42.8	14.1	51.2	12.6	2.4	20.7	29.2
	DCAN [55]	79.9	30.4	70.8	1.6	0.6	22.3	6.7	**23.0**	76.9	73.9	41.9	16.7	61.7	11.5	10.3	**38.6**	35.4
Source-combined DA	CyCADA [32]	82.8	35.8	78.2	17.5	15.1	10.8	6.1	19.4	78.6	77.2	44.5	15.3	74.9	17.0	10.3	12.9	37.3
Multi-source DA	MDAN [69]	64.2	19.7	63.8	13.1	19.4	5.5	5.2	6.8	71.6	61.1	42.0	12.0	62.7	2.9	12.3	8.1	29.4
	MADAN (Ours)	86.2	**37.7**	**79.1**	20.1	17.8	15.5	14.5	21.4	78.5	73.4	49.7	**16.8**	77.8	**28.3**	**17.7**	27.5	**41.4**
Oracle-Train on Tgt	FCN [5]	96.4	74.5	87.1	35.3	37.8	36.4	46.9	60.1	89.0	89.8	65.6	35.9	76.9	64.1	40.5	65.1	62.6

Table 3: Comparison with the state-of-the-art DA methods for semantic segmentation from GTA and SYNTHIA to BDDS. The best class-wise IoU and mIoU are emphasized in bold.

Standards	Method	road	sidewalk	building	wall	fence	pole	t-light	t-sign	vegettion	sky	person	rider	car	bus	m-bike	bicycle	mIoU
Source-only	GTA	50.2	18.0	55.1	3.1	7.8	7.0	0.0	3.5	61.0	50.4	19.2	0.0	58.1	3.2	**19.8**	0.0	22.3
	SYNTHIA	7.0	6.0	50.5	0.0	0.0	15.1	0.2	2.4	60.3	**85.6**	16.5	0.5	36.7	3.3	0.0	3.5	17.1
	GTA+SYNTHIA	54.5	19.6	64.0	3.2	3.6	5.2	0.0	0.0	61.3	82.2	13.9	0.0	55.5	16.7	13.4	0.0	24.6
GTA-only DA	CyCADA [32]	**77.9**	26.8	68.8	13.0	**19.7**	13.5	**18.2**	**22.3**	64.2	84.2	39.0	**22.6**	**72.0**	11.5	15.9	2.0	35.7
SYNTHIA-only DA	CyCADA [32]	55	13.8	45.2	0.1	0.0	13.2	0.5	10.6	63.3	67.4	22.0	6.9	52.5	10.5	10.4	13.3	24.0
Source-combined DA	CyCADA [32]	61.5	27.6	**72.1**	6.5	2.8	15.7	10.8	18.1	78.3	73.8	44.9	16.3	41.5	21.1	21.8	**25.9**	33.7
Multi-source DA	MDAN [69]	35.9	15.8	56.9	5.8	16.3	9.5	8.6	6.2	59.1	80.1	24.5	9.9	53.8	11.8	2.9	1.6	25.0
	MADAN (Ours)	60.2	**29.5**	66.6	**16.9**	10.0	**16.6**	10.9	16.4	**78.8**	75.1	**47.5**	17.3	48.0	**24.0**	13.2	17.3	**36.3**
Oracle-Train on Tgt	FCN [5]	91.7	54.7	79.5	25.9	42.0	23.6	30.9	34.6	81.2	91.6	49.6	23.5	85.4	64.2	28.4	41.1	53.0

4. Experiments & Results : Ablation Study

For all simulation to real adaptations, DSC achieves better results.

Table 4: Comparison between the proposed dynamic semantic consistency (DSC) loss in MADAN and the original SC loss in [32] on Cityscapes. The better mIoU for each pair is emphasized in bold.

Source	Method	road	sidewalk	building	wall	fence	pole	t-light	t-sign	vegettion	sky	person	rider	car	bus	m-bike	bicycle	mIoU
GTA	CycleGAN+SC	85.6	30.7	74.7	14.4	13.0	17.6	13.7	5.8	74.6	69.9	38.2	3.5	72.3	5.0	3.6	0.0	32.7
	CycleGAN+DSC	76.6	26.0	76.3	17.3	18.8	13.6	13.2	17.9	78.8	63.9	47.4	14.8	72.2	24.1	19.8	10.8	**38.1**
	CyCADA w/ SC	85.2	37.2	76.5	21.8	15.0	23.8	21.5	22.9	80.5	60.7	50.5	9.0	76.9	28.2	9.8	0.0	38.7
	CyCADA w/ DSC	84.1	27.3	78.3	21.6	18.0	13.8	14.1	16.7	78.1	66.9	47.8	15.4	78.7	23.4	22.3	14.4	**40.0**
SYNTHIA	CycleGAN+SC	64.0	29.4	61.7	0.3	0.1	15.3	3.4	5.0	63.4	68.4	39.4	11.5	46.6	10.4	2.0	16.4	27.3
	CycleGAN + DSC	68.4	29.0	65.2	0.6	0.0	15.0	0.1	4.0	75.1	70.6	45.0	11.0	54.9	18.2	3.9	26.7	**30.5**
	CyCADA w/ SC	66.2	29.6	65.3	0.5	0.2	15.1	4.5	6.9	67.1	68.2	42.8	14.1	51.2	12.6	2.4	20.7	29.2
	CyCADA w/ DSC	69.8	27.2	68.5	5.8	0.0	11.6	0.0	2.8	75.7	58.3	44.3	10.5	68.1	22.1	11.8	32.7	**31.8**

Table 5: Comparison between the proposed dynamic semantic consistency (DSC) loss in MADAN and the original SC loss in [32] on BDDS. The better mIoU for each pair is emphasized in bold.

Source	Method	road	sidewalk	building	wall	fence	pole	t-light	t-sign	vegettion	sky	person	rider	car	bus	m-bike	bicycle	mIoU
GTA	CycleGAN+SC	62.1	20.9	59.2	6.0	23.5	12.8	9.2	22.4	65.9	78.4	34.7	11.4	64.4	14.2	10.9	1.9	31.1
	CycleGAN+DSC	74.4	23.7	65.0	8.6	17.2	10.7	14.2	19.7	59.0	82.8	36.3	19.6	69.7	4.3	17.6	4.2	**32.9**
	CyCADA w/ SC	68.8	23.7	67.0	7.5	16.2	9.4	11.3	22.2	60.5	82.1	36.1	20.6	63.2	15.2	16.6	3.4	32.0
	CyCADA w/ DSC	70.5	32.4	68.2	10.5	17.3	18.4	16.6	21.8	65.6	82.2	38.1	16.1	73.3	20.8	12.6	3.7	**35.5**
SYNTHIA	CycleGAN+SC	50.6	13.6	50.5	0.2	0.0	7.9	0.0	0.0	63.8	58.3	21.6	7.8	50.2	1.8	2.2	19.9	21.8
	CycleGAN + DSC	57.3	13.4	56.1	2.7	14.1	9.8	7.7	17.1	65.5	53.1	11.4	1.4	51.4	13.9	3.9	8.7	**22.5**
	CyCADA w/ SC	49.5	11.1	46.6	0.7	0.0	10.0	0.4	7.0	61.0	74.6	17.5	7.2	50.9	5.8	13.1	4.3	23.4
	CyCADA w/ DSC	55	13.8	45.2	0.1	0.0	13.2	0.5	10.6	63.3	67.4	22.0	6.9	52.5	10.5	10.4	13.3	**24.0**

4. Experiments & Results : Ablation Study

Both domain aggregation methods can obtain better performance while SAD outperforms CCD

Add DSC loss could further improve the mIoU score;

Feature-level alignments also contribute to the adaptation task

Adding each one of them does not introduce performance degradation.

Table 6: Ablation study on different components in MADAN on Cityscapes. Baseline denotes using piexl-level alignment with cycle-consistency, +SAD denotes using the sub-domain aggregation discriminator, +CCD denotes using the cross-domain cycle discriminator, +DSC denotes using the dynamic semantic consistency loss, and +Feat denotes using feature-level alignment.

Method	road	sidewalk	building	wall	fence	pole	t-light	t-sign	vegettion	sky	person	rider	car	bus	m-bike	bicycle	mIoU
Baseline	74.9	27.6	67.5	9.1	10.0	12.8	1.4	13.6	63.0	47.1	41.7	13.5	60.8	22.4	6.0	8.1	30.0
+SAD	79.7	33.2	75.9	11.8	3.6	15.9	8.6	15.0	74.7	**78.9**	44.2	17.1	68.2	24.9	16.7	14.0	36.4
+CCD	82.1	36.3	69.8	9.5	4.9	11.8	12.5	15.3	61.3	54.1	49.7	10.0	70.7	9.7	19.7	12.4	33.1
+SAD+CCD	82.7	35.3	76.5	15.4	**19.4**	14.1	7.2	13.9	75.3	74.2	**50.9**	**19.0**	66.5	26.6	16.3	6.7	37.5
+SAD+DSC	83.1	36.6	78.0	**23.3**	12.6	11.8	3.5	11.3	75.5	74.8	42.2	17.9	72.2	27.2	13.8	10.0	37.1
+CCD+DSC	**86.8**	36.9	78.6	16.2	8.1	**17.7**	8.9	13.7	75.0	74.8	42.2	18.2	74.6	22.5	**22.9**	12.7	38.1
+SAD+CCD+DSC	84.2	35.1	78.7	17.1	18.7	15.4	**15.7**	**24.1**	77.9	72.0	49.2	17.1	75.2	24.1	18.9	19.2	40.2
+SAD+CCD+DSC+Feat	86.2	**37.7**	**79.1**	20.1	17.8	15.5	14.5	21.4	**78.5**	73.4	49.7	16.8	**77.8**	**28.3**	17.7	**27.5**	**41.4**

Table 7: Ablation study on different components in MADAN on BDDS.

Method	road	sidewalk	building	wall	fence	pole	t-light	t-sign	vegettion	sky	person	rider	car	bus	m-bike	bicycle	mIoU
Baseline	31.3	17.4	55.4	2.6	12.9	12.4	6.5	18.0	63.2	79.9	21.2	5.6	44.1	14.2	6.1	11.7	24.6
+SAD	58.9	18.7	61.8	6.4	10.7	17.1	20.3	17.0	67.3	83.7	21.1	6.7	66.6	22.7	4.5	14.9	31.2
+CCD	52.7	13.6	63.0	6.6	11.2	17.8	**21.5**	**18.9**	67.4	**84.0**	9.2	2.2	63.0	21.6	2.0	14.0	29.3
+SAD+CCD	61.6	20.2	61.7	7.2	12.1	**18.5**	19.8	16.7	64.2	83.2	25.9	7.3	66.8	22.2	5.3	14.9	31.8
+SAD+DSC	60.2	29.5	66.6	16.9	10.0	16.6	10.9	16.4	**78.8**	75.1	**47.5**	17.3	48.0	**24.0**	13.2	**17.3**	34.3
+CCD+DSC	61.5	27.6	72.1	6.5	12.8	15.7	10.8	18.1	78.3	73.8	44.9	16.3	41.5	21.1	21.8	15.9	33.7
+SAD+CCD+DSC	64.6	**38.0**	75.8	17.8	13.0	9.8	5.9	4.6	74.8	76.9	41.8	**24.0**	**69.0**	20.4	23.7	11.3	35.3
+SAD+CCD+DSC+Feat	**69.1**	36.3	**77.9**	**21.5**	**17.4**	13.8	4.1	16.2	76.5	76.2	42.2	16.4	56.3	22.4	**24.5**	13.5	**36.3**

5. Conclusion

Their contribution:

- Might be the first work on multi-source structured domain adaptation
- Besides feature-level alignment, pixel-level alignment is further considered
- Conduct extensive experiments from synthetic data to real data

Further direction:

- Investigate multi-modal DA to better boost the adaptation performance
- Improve the computational efficiency of MADAN with techniques such as neural architecture search

Rensselaer

why not change the world?®