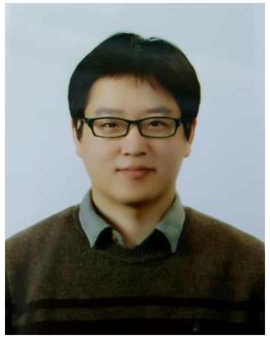
SAMSUNG ADVANCED INSTITUTE OF TECHNOLOGY

NEURAL INFORMATION PROCESSING SYSTEMS

DaDA: Distortion-aware Domain Adaptation for Unsupervised Semantic Segmentation

Sujin Jang



Joohan Na

Dokwan Oh

Samsung Advanced Institute of Technology (SAIT)

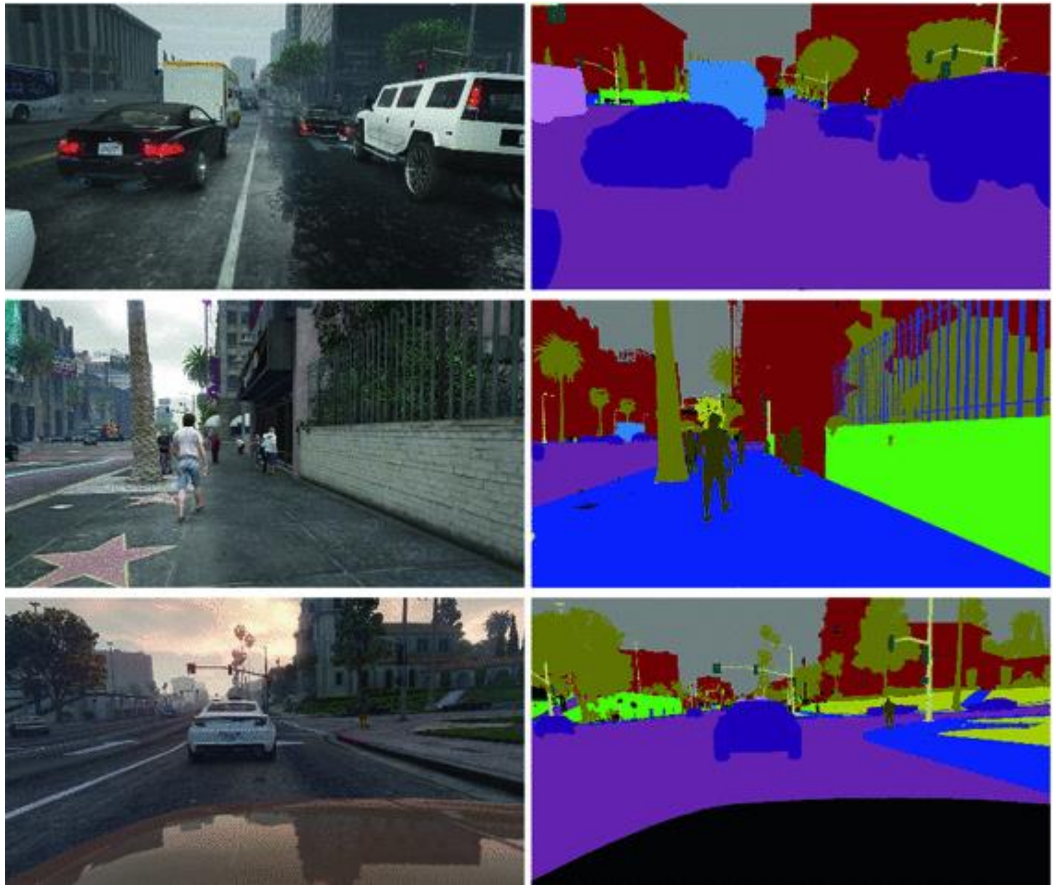
UDA for Semantic Segmentation

Target Domain

Cityscapes w/o Labels

Domain Adaptation

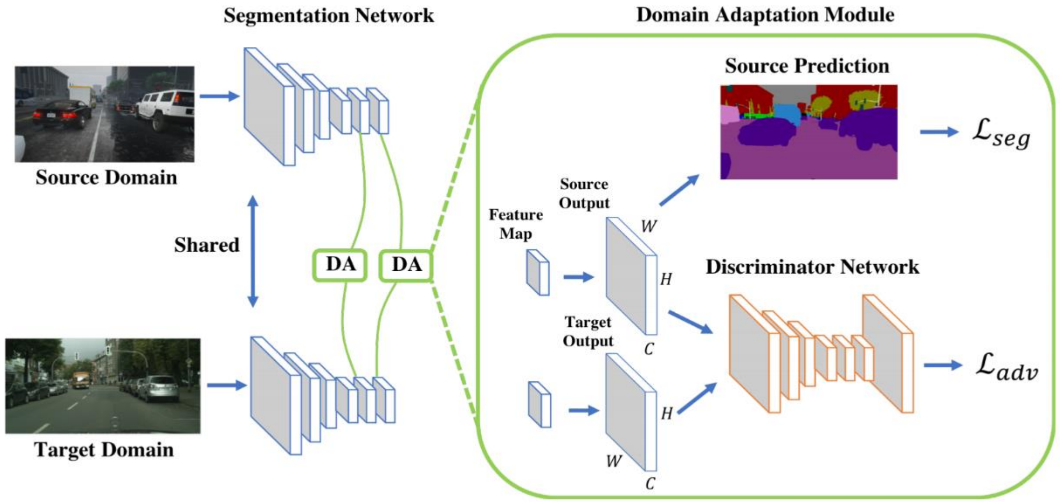
Source Domain



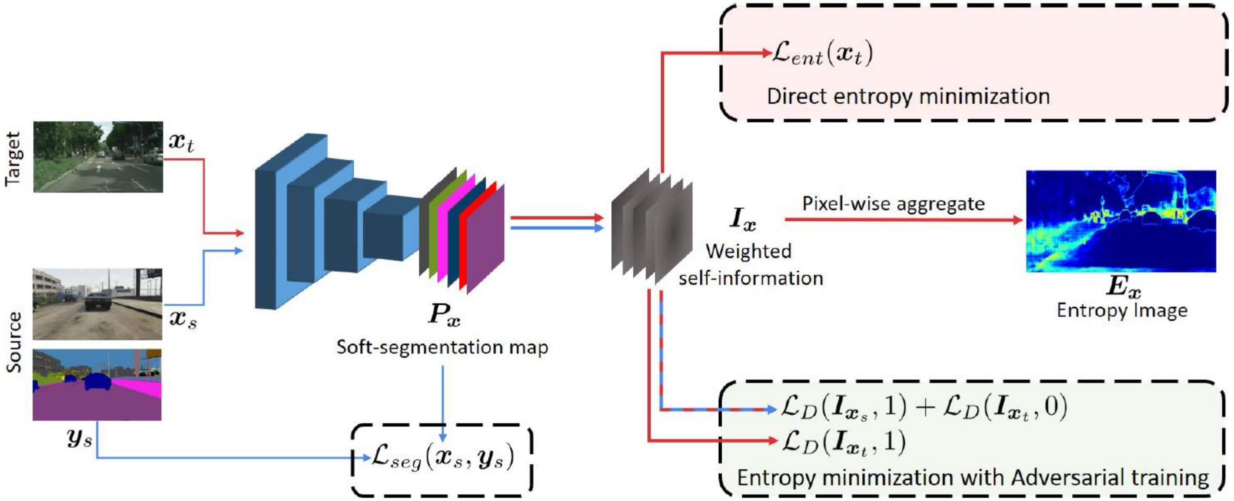
GTAV w/ Labels

*UDA: Unsupervised Domain Adaptation

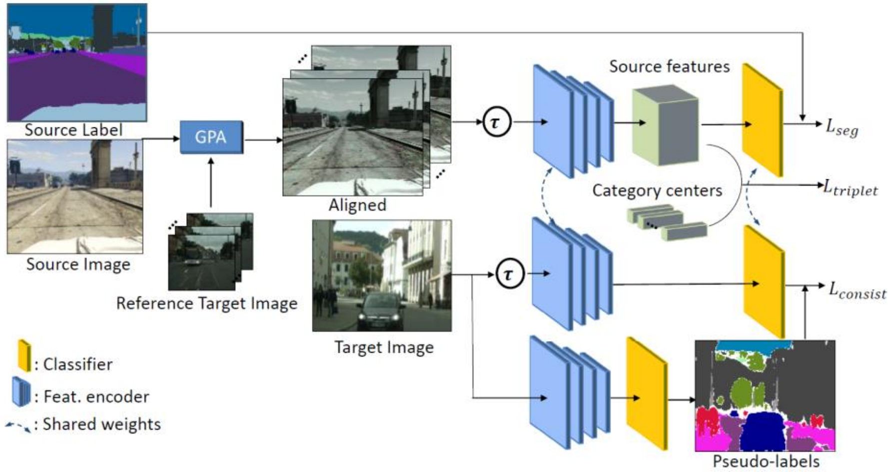
UDA for Semantic Segmentation



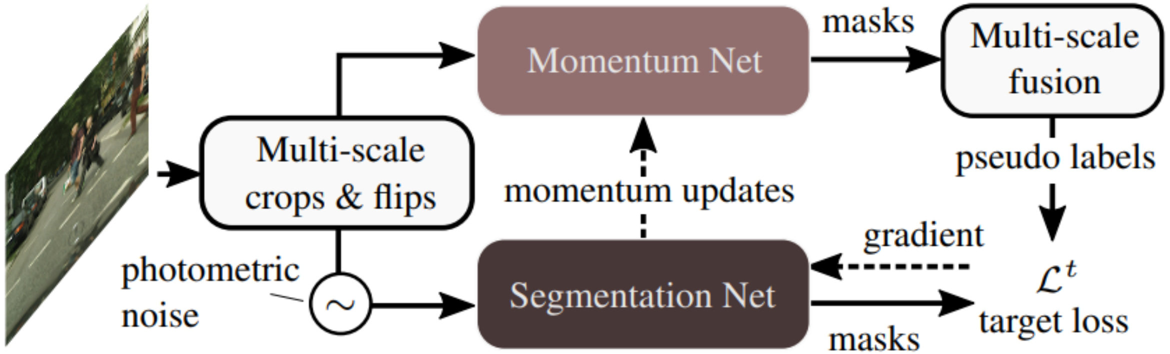
[Tsai et al. 2018]



[Vu et al. 2019]

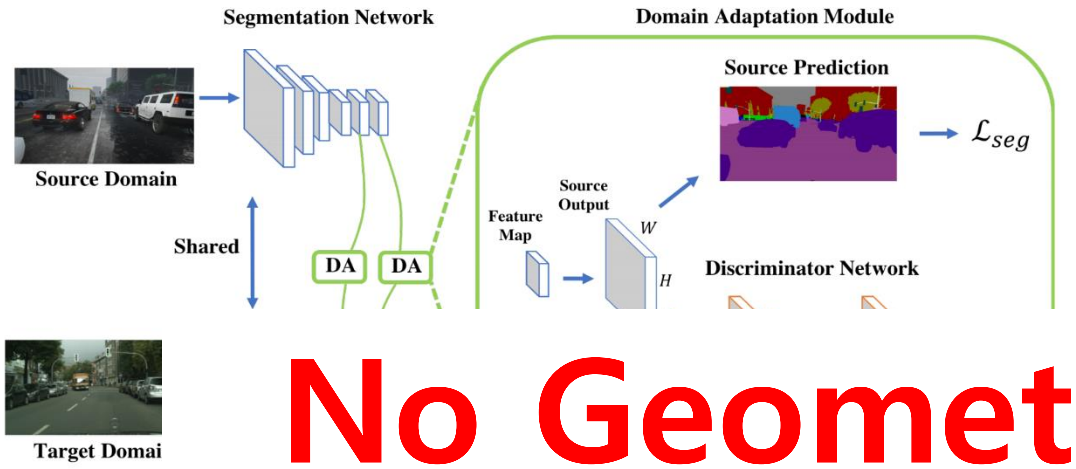


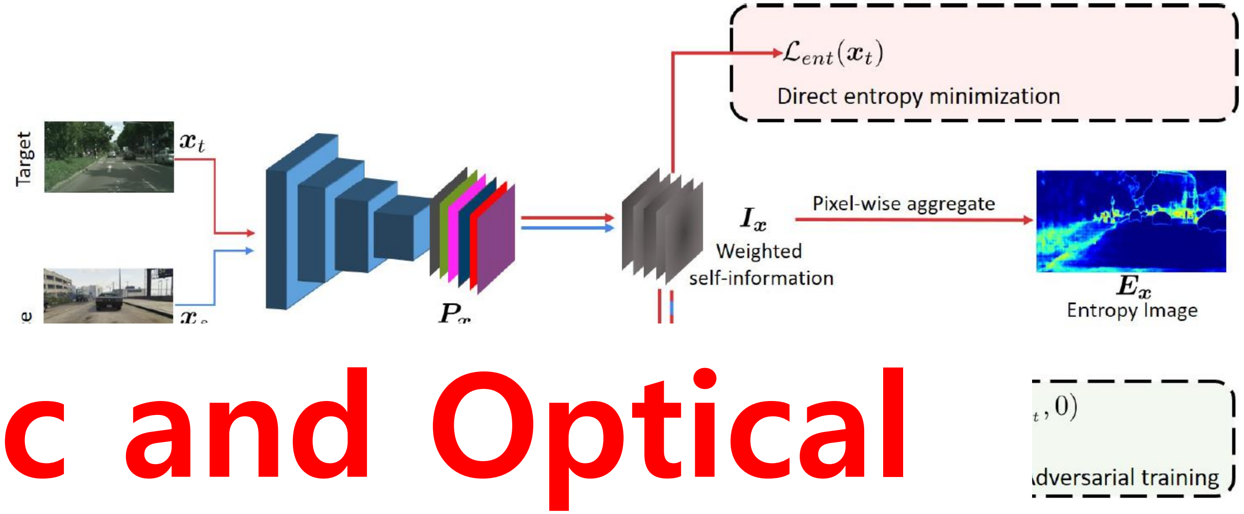
[Ma et al. 2021]



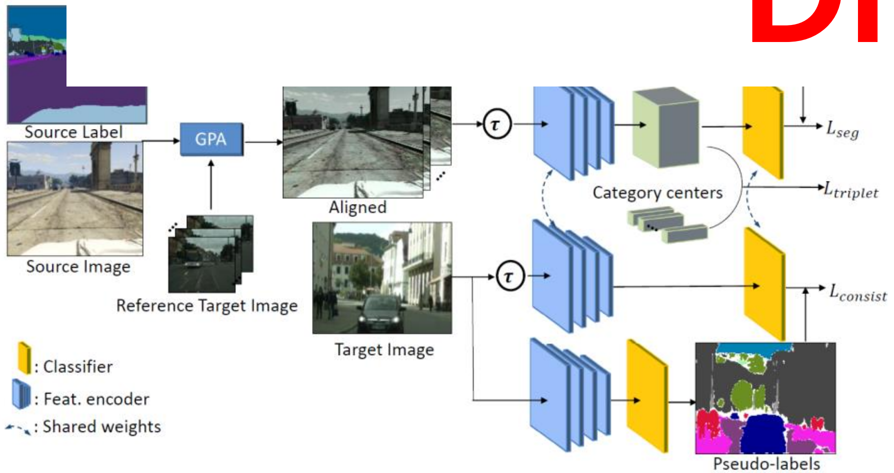
[Araslanov et al. 2021]

UDA for Semantic Segmentation

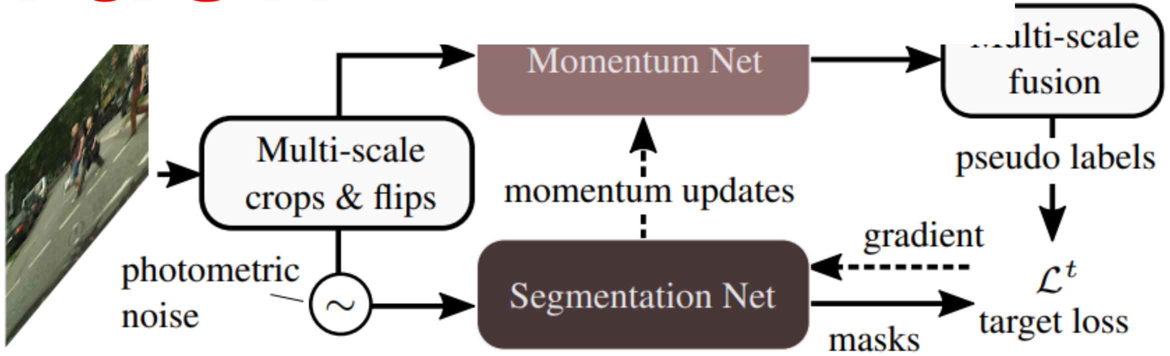




No Geometric and Optical Distortion

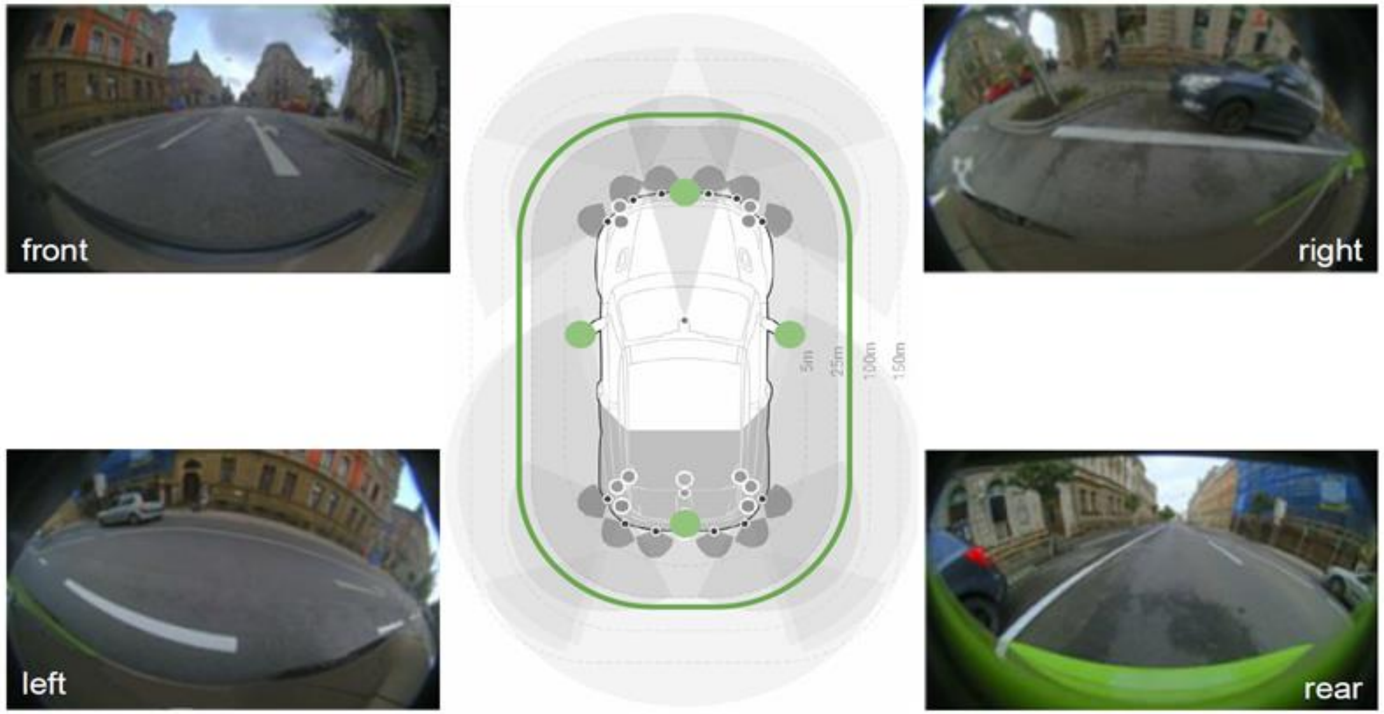


[Ma et al. 2021]

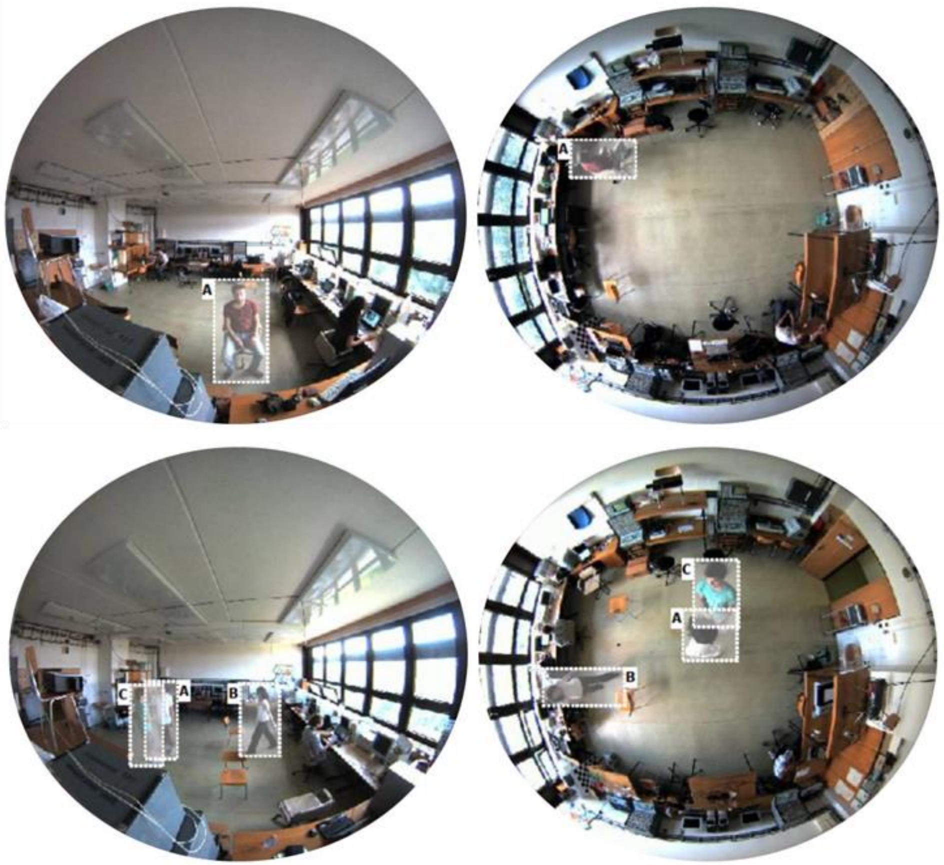


[Araslanov et al. 2021]

Wide-Angle Cameras



Woodscape Dataset for
Autonomous Driving

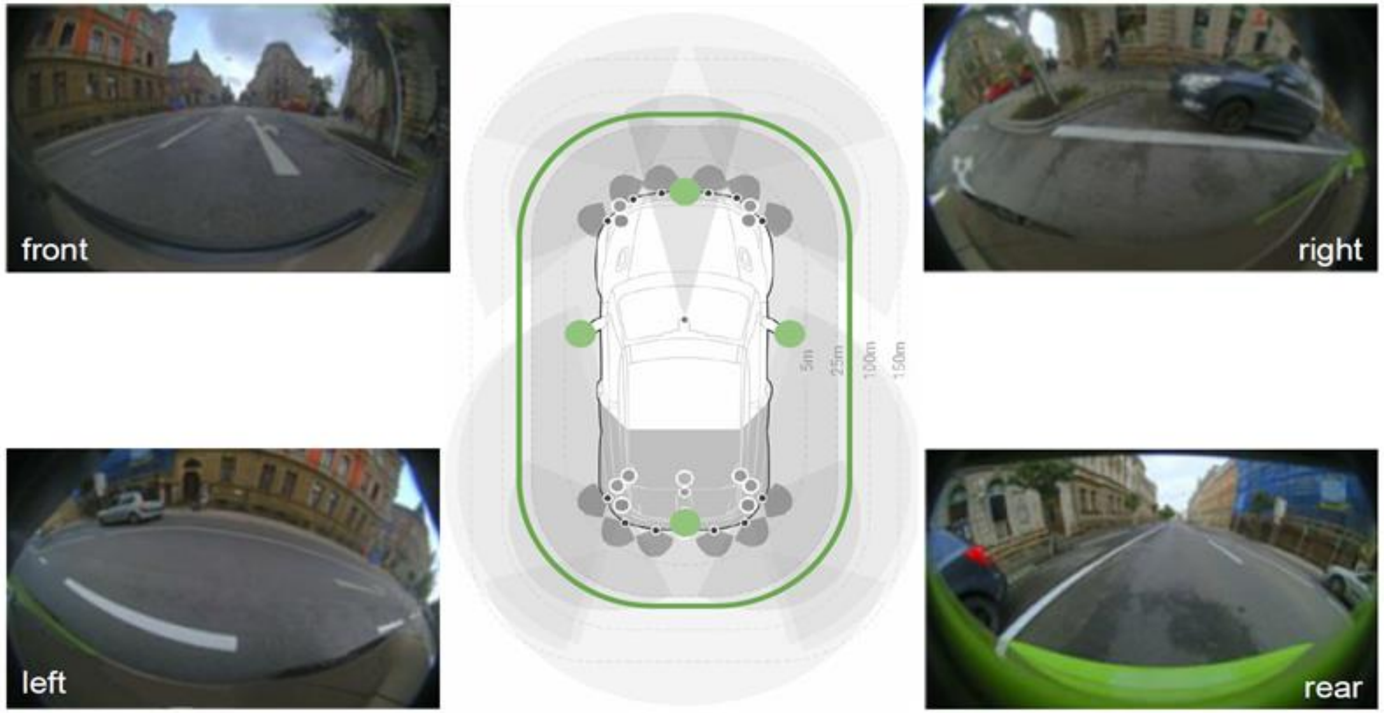


Bomni-DB Dataset for
Surveillance Applications

SAMSUNG
SAMSUNG ADVANCED
INSTITUTE OF TECHNOLOGY

NEURAL INFORMATION
PROCESSING SYSTEMS

Wide-Angle Cameras



Woodscape Dataset for
Autonomous Driving

No Labeled Data

Distortion-aware Unsupervised Domain Adaptation

Target Domain

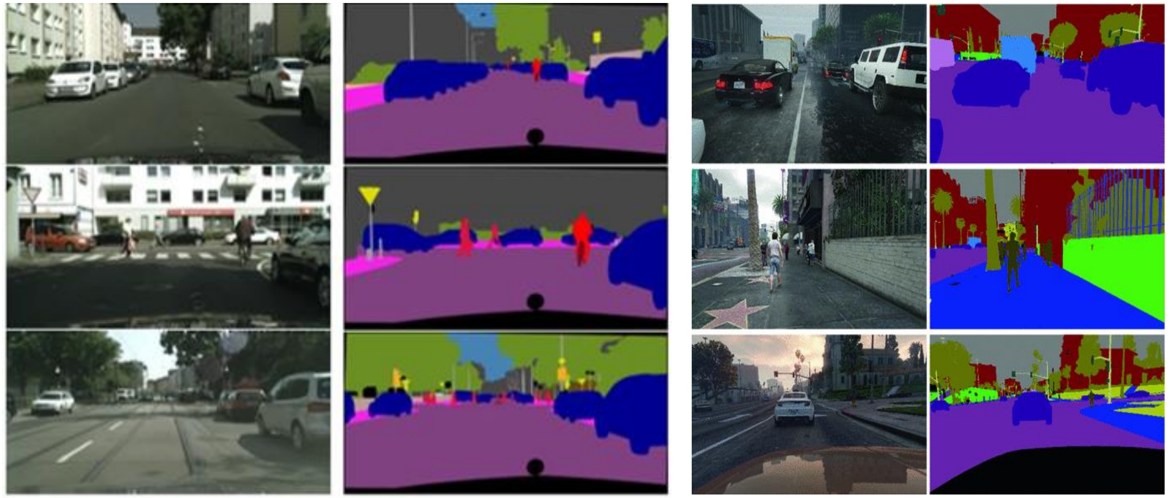
Woodscape   Fisheye Driving Dataset(FDD)*

Unlabeled Fisheye Images

Visual
+
Distortion

Domain Adaptation

Source Domain



Cityscapes   GTAV

Labeled Rectilinear Images

SAMSUNG ADVANCED INSTITUTE OF TECHNOLOGY

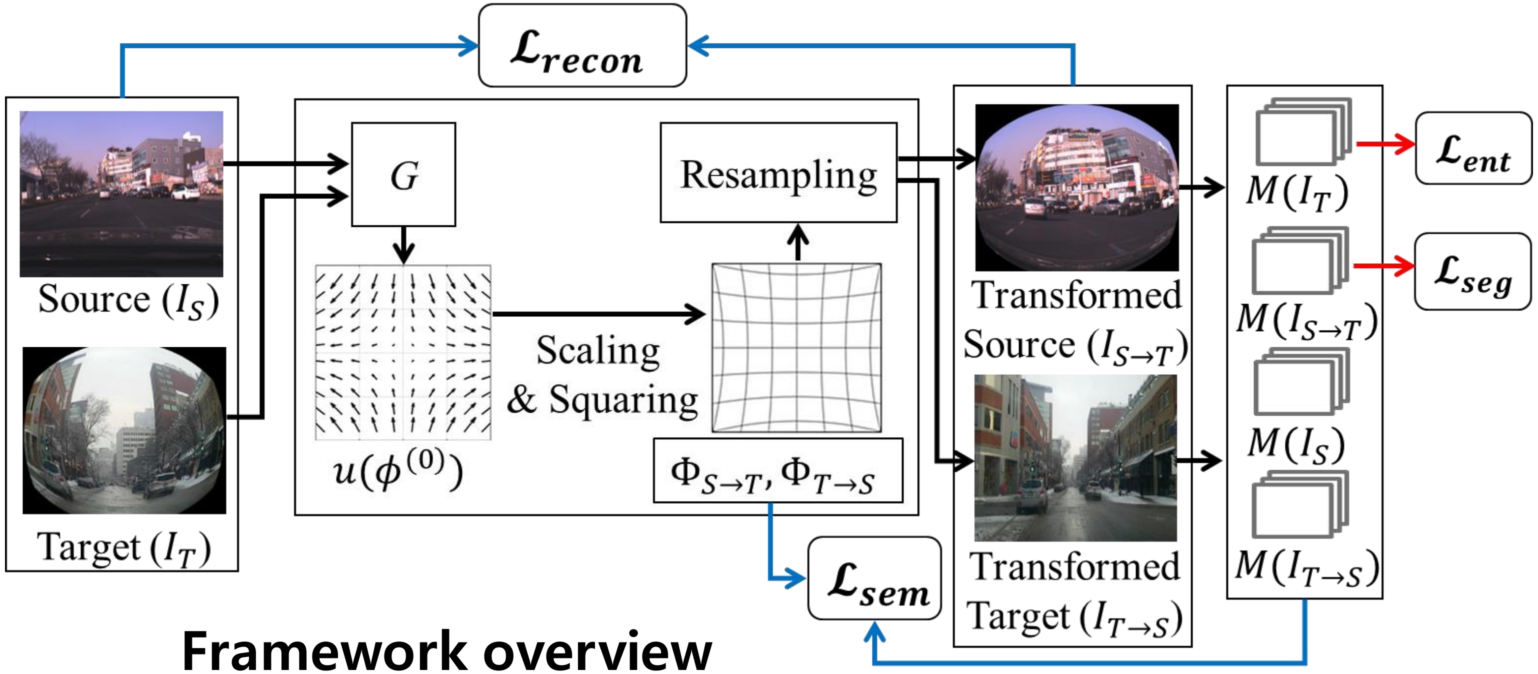
*Fisheye Driving Dataset (FDD): Our fisheye dataset available at https://sait-fdd.github.io/

NEURAL INFORMATION PROCESSING SYSTEMS

DaDA: Distortion-aware Domain Adaptation

Our Contributions:

- New UDA benchmarks introducing geometric optical distortion;
- DaDA framework to solve such challenging but practically important tasks;
- Extensive experimental results to validate our approach.

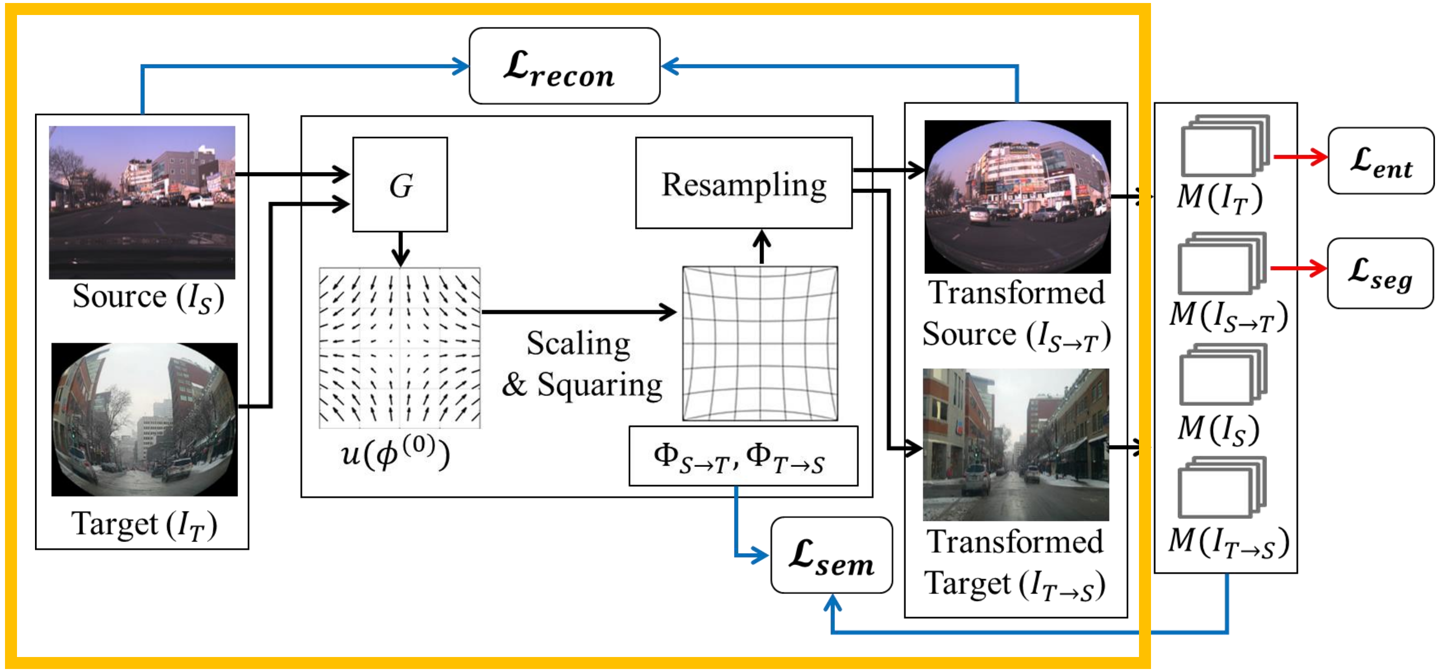


Framework overview

SAMSUNG
SAMSUNG ADVANCED INSTITUTE OF TECHNOLOGY

NEURAL INFORMATION PROCESSING SYSTEMS

DaDA: Distortion-aware Domain Adaptation

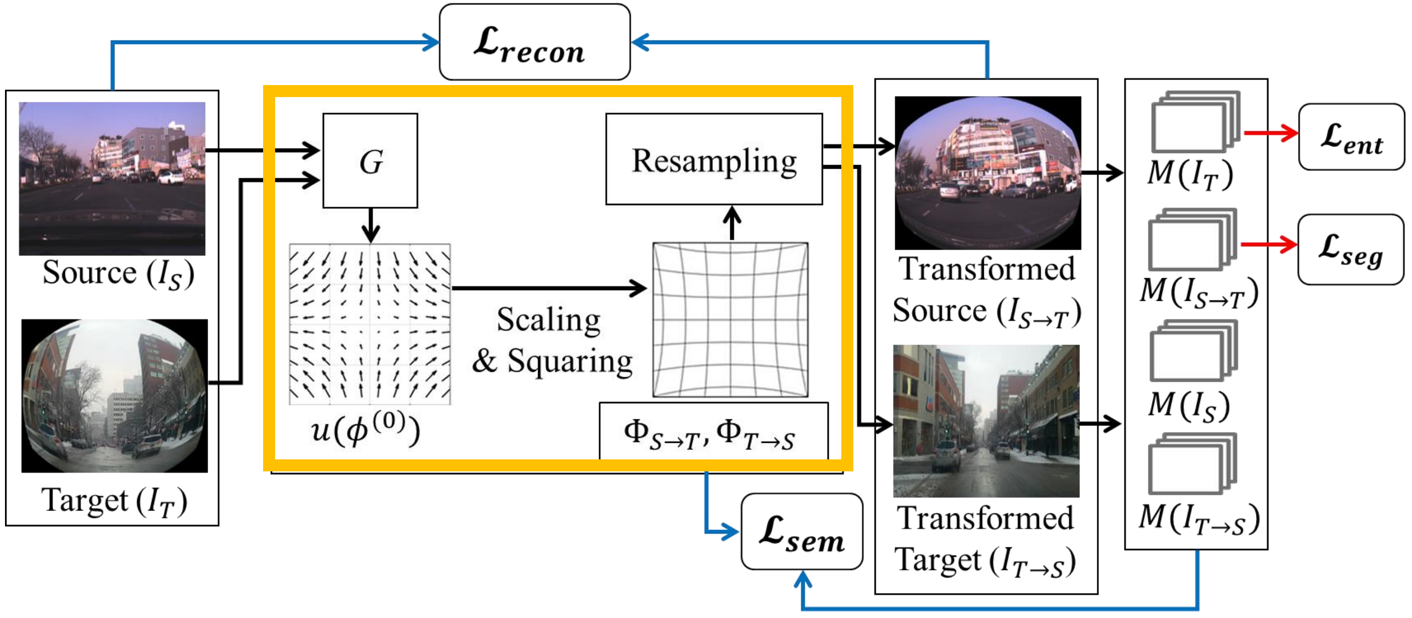


Relative Distortion Learnig (RDL)

SAMSUNG ADVANCED INSTITUTE OF TECHNOLOGY

NEURAL INFORMATION PROCESSING SYSTEMS

DaDA: Distortion-aware Domain Adaptation



Diffeomorphic Transformation

- Globally one-to-one mapping
- Continuous and smooth
- Differentiable and invertible

$$G(I_S, I_T) = u(\phi^{(0)})$$

$$\frac{\partial \phi^{(t)}}{\partial t} = u(\phi^{(t)}), \quad u \in \mathbb{R}^{2\times w\times h}$$

Squaring-and-Scaling Integration

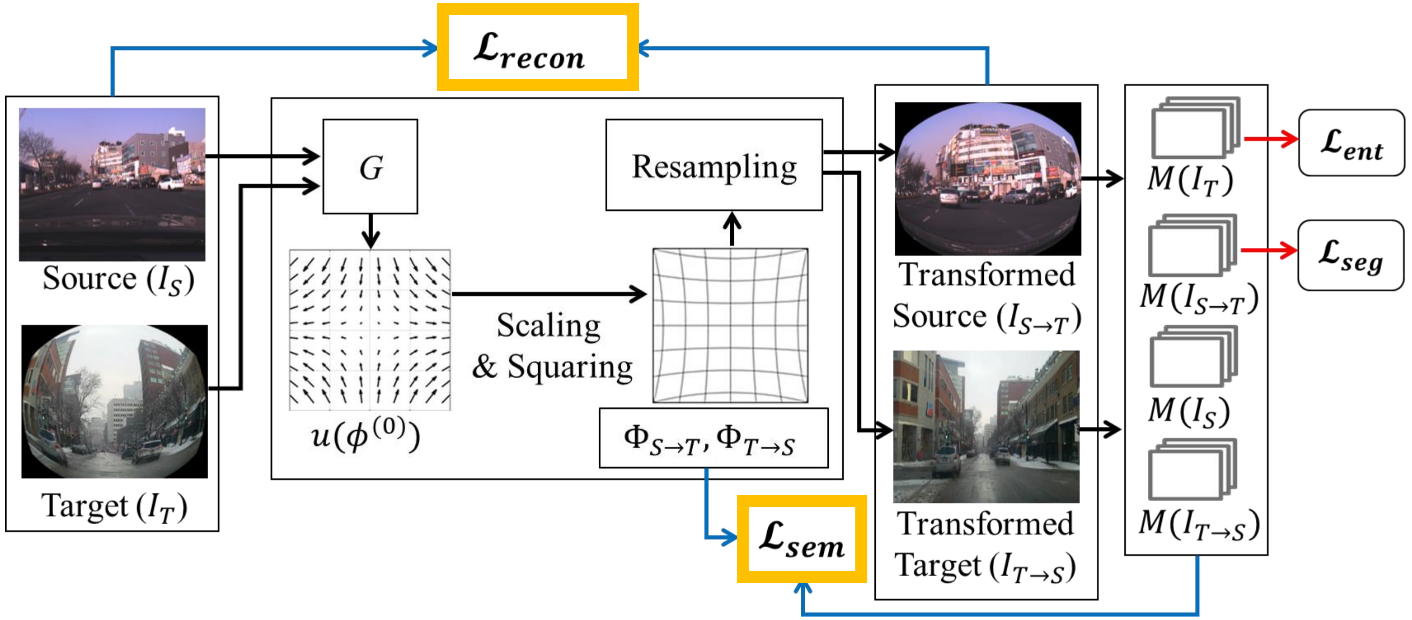
$$\phi^{(1/2^T)} = \phi^{(0)} + u/2^T$$

$$\phi^{(1/2^{t-1})} = \phi^{(1/2^t)} \circ \phi^{(1/2^t)}$$

Forward field: $\Phi_{S\to T}$

Backward field: $\Phi_{T\to S}$

DaDA: Distortion-aware Domain Adaptation



Distortion-aware Losses

$$\mathcal{L}_{recon} = \|I_S - I'_S\|_1 + \|I_T - I'_T\|_1,$$
$$\text{where } I'_S = I_{S\to T} \circ \Phi_{T\to S},\ I'_T = I_{T\to S} \circ \Phi_{S\to T}$$

$$\mathcal{L}_{sem} = \|M(I_S) \circ \Phi_{S\to T} - M(I_{S\to T})\|_1 + \|M(I_T) \circ \Phi_{T\to S} - M(I_{T\to S})\|_1.$$

Distortion-aware Discriminator and Adversarial Loss

$$\mathcal{L}_{D_G} = \mathbb{E}_{I_S\sim\mathcal{S}, I_T\sim\mathcal{T}}[1 - D_G(I_S \circ \Phi_{S\to T}, \nabla(I_S \circ \Phi_{S\to T}))] + \mathbb{E}_{I_T\sim\mathcal{T}}[D_G(I_T, \nabla I_T)].$$

$$\mathcal{L}_{adv\_G} = \mathbb{E}_{I_S\sim\mathcal{S}, I_T\sim\mathcal{T}}[D_G(I_S \circ \Phi_{S\to T}, \nabla(I_S \circ \Phi_{S\to T}))],$$

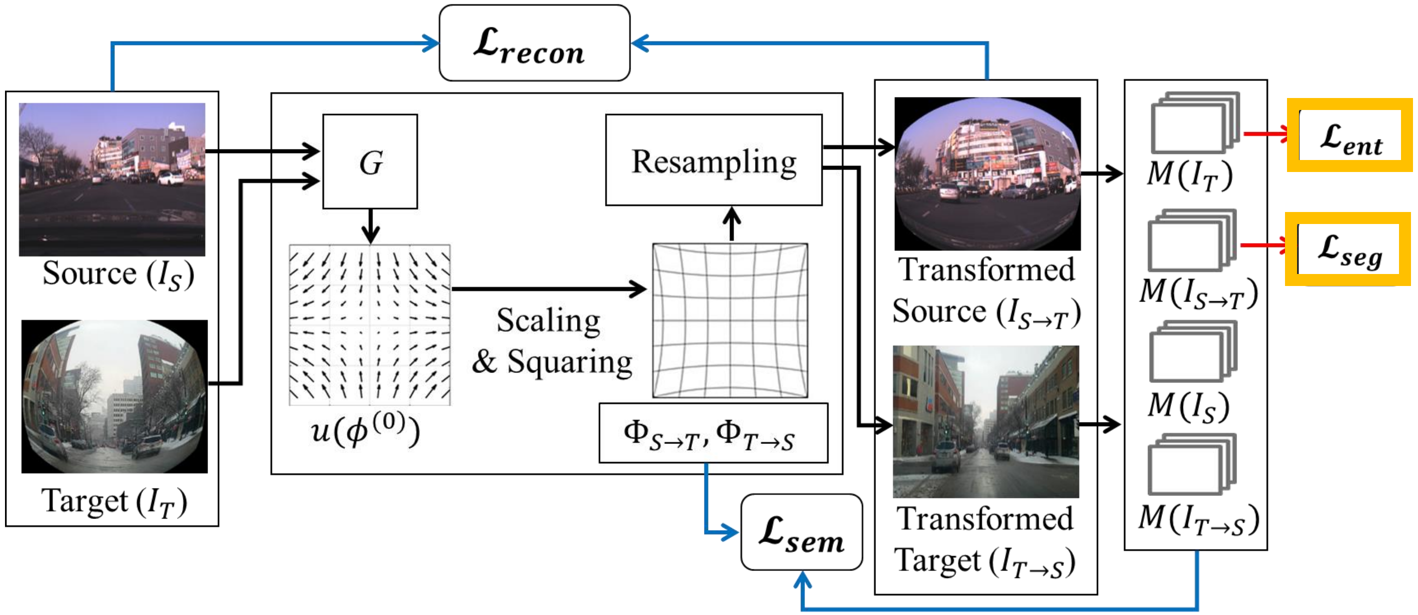
Total Loss for RDL

$$\mathcal{L}_{rdl} = \beta_1\mathcal{L}_{recon} + \beta_2\mathcal{L}_{sem} + \beta_3\mathcal{L}_{adv\_G}$$

SAMSUNG ADVANCED INSTITUTE OF TECHNOLOGY

NEURAL INFORMATION PROCESSING SYSTEMS

DaDA: Distortion-aware Domain Adaptation



Distortion-aware Adversarial Adaptation

$$\mathcal{L}_{seg} = -\sum_{h,w}\sum_{c\in C} Y_{S\to T}^{(h,w,c)} \log(M(I_{S\to T})^{(h,w,c)})$$

$$\mathcal{L}_{ent} = \frac{-1}{\log(C)} \sum_{h,w}\sum_{c\in C} M(I_T)^{(h,w,c)} \log M(I_T)^{(h,w,c)}$$

Distortion-aware Discriminator and Adversarial Loss

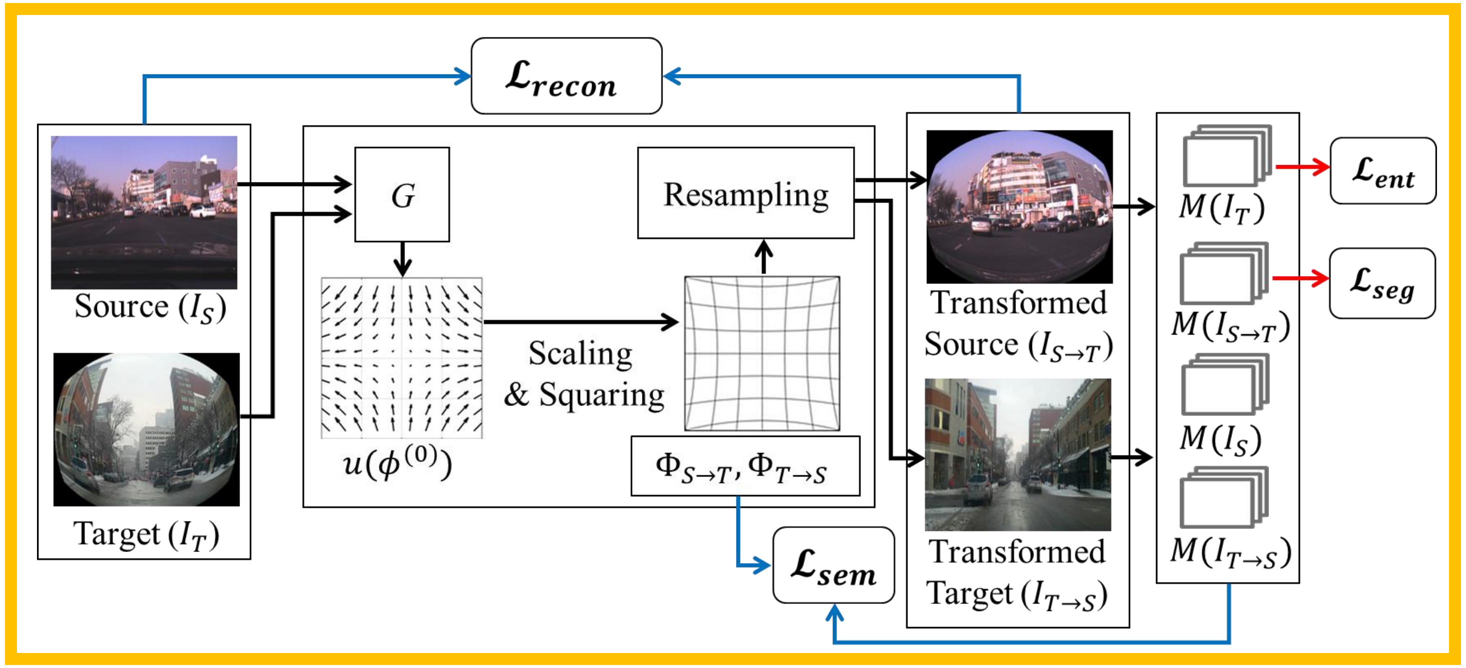
$$\mathcal{L}_{D_M} = \mathbb{E}_{I_S\sim\mathcal{S}, I_T\sim\mathcal{T}}[1 - D_M(M(I_S \circ \Phi_{S\to T}))] + \mathbb{E}_{I_T\sim\mathcal{T}}[D_M(M(I_T))]$$

$$\mathcal{L}_{adv\_M} = \mathbb{E}_{I_T\sim\mathcal{T}}[D_M(1 - M(I_T))]$$

SAMSUNG ADVANCED INSTITUTE OF TECHNOLOGY

NEURAL INFORMATION PROCESSING SYSTEMS

DaDA: Distortion-aware Domain Adaptation



Distortion-aware Adversarial Adaptation

$$\mathcal{L}_{seg} = -\sum_{h,w}\sum_{c\in C} Y_{S\to T}^{(h,w,c)} \log(M(I_{S\to T})^{(h,w,c)})$$

$$\mathcal{L}_{ent} = \frac{-1}{\log(C)}\sum_{h,w}\sum_{c\in C} M(I_T)^{(h,w,c)} \log M(I_T)^{(h,w,c)}$$

Distortion-aware Discriminator and Adversarial Loss

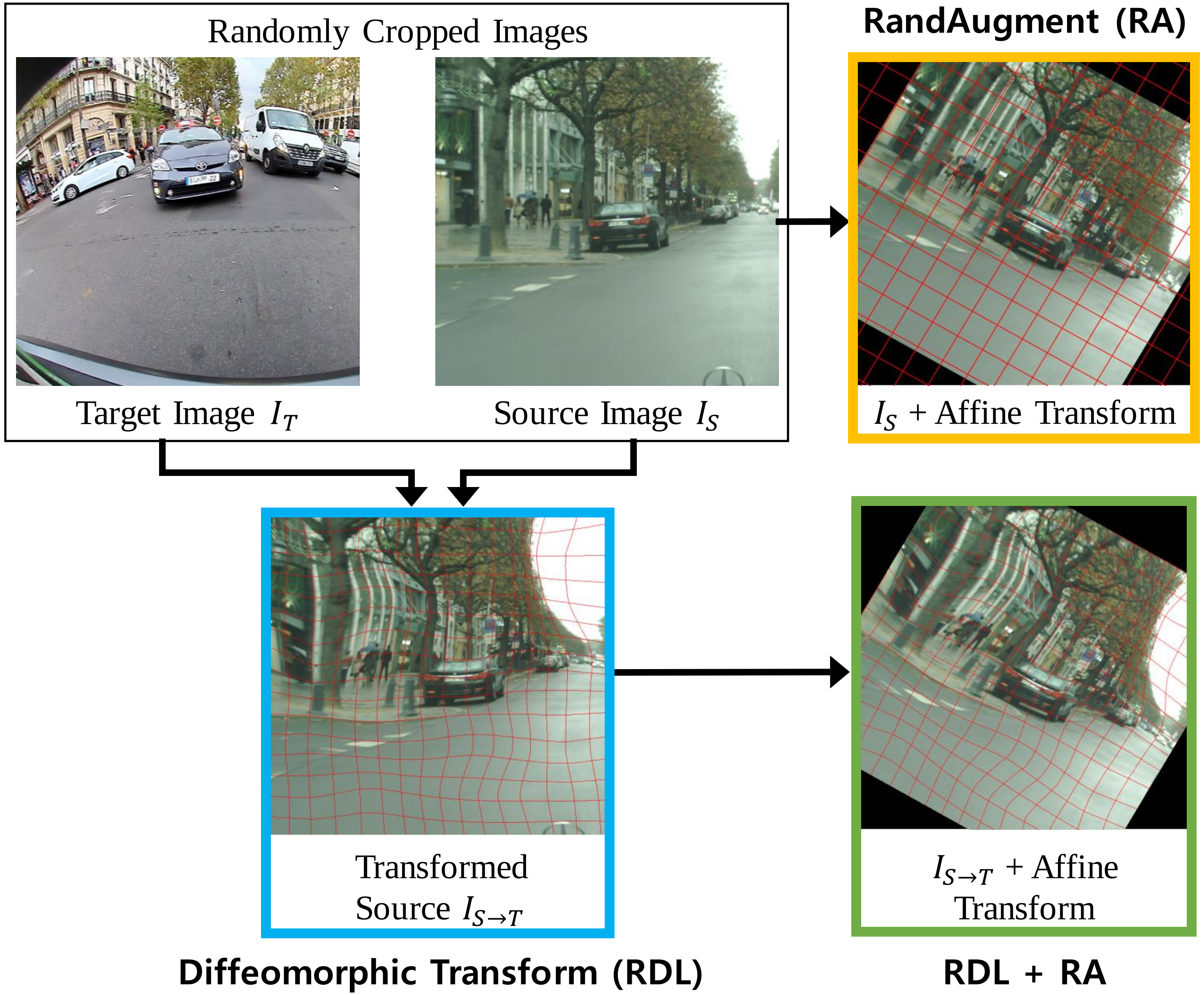
$$\mathcal{L}_{D_M} = \mathbb{E}_{I_S\sim\mathcal{S}, I_T\sim\mathcal{T}}[1 - D_M(M(I_S \circ \Phi_{S\to T}))] + \mathbb{E}_{I_T\sim\mathcal{T}}[D_M(M(I_T))]$$

$$\mathcal{L}_{adv\_M} = \mathbb{E}_{I_T\sim\mathcal{T}}[D_M(1 - M(I_T))]$$

Total Loss for Segmentation Adaptation

$$\mathcal{L}_{all} = \mathcal{L}_{rdl} + \mathcal{L}_{adv\_M} + \mathcal{L}_{seg} + \gamma\mathcal{L}_{ent}$$

Experiments – Diffeomorphic and Affine Transformation



Experiments – Distortion Style Translation

Target Image (I_T) | Source Image (I_S) | CycleGAN [44] | RDL ($I_{S\to T}$) | $\Phi_{T\to S}$ | $I'_S = I_{S\to T} \cdot \Phi_{T\to S}$

SAMSUNG ADVANCED INSTITUTE OF TECHNOLOGY

NEURAL INFORMATION PROCESSING SYSTEMS

Experiments – Quantitative Results

Comparisons with the baseline adaptation methods.

	Cityscapes → Woodscape		GTAV → Woodscape		Cityscapes → FDD		GTAV → FDD	
Method	mIoU(%)	gain	mIoU(%)	gain	mIoU(%)	gain	mIoU(%)	gain
SourceOnly	32.39		29.32		34.76		32.13	
AdaptSeg [33]	46.33		35.94		39.07		36.90	
AdaptSeg+RA	50.44	+4.11	36.88	+0.94	39.42	+0.35	37.22	+0.32
AdaptSeg+RDL	50.88	+4.55	37.36	+1.42	**41.35**	**+2.28**	39.29	+2.39
AdaptSeg+RA+RDL	**52.59**	**+6.26**	**37.73**	**+1.78**	41.07	+2.00	**39.64**	**+2.74**
AdvEnt [34]	45.26		34.70		38.87		37.25	
AdvEnt+RA	50.60	+5.34	36.64	+1.94	41.58	+2.71	38.75	+1.50
AdvEnt+RDL	50.94	+5.68	36.39	+1.69	**42.43**	**+3.56**	39.93	+2.68
AdvEnt+RA+RDL	**52.64**	**+7.38**	**37.62**	**+2.92**	42.32	+3.45	**40.87**	**+3.62**

SAMSUNG
SAMSUNG ADVANCED INSTITUTE OF TECHNOLOGY

NEURAL INFORMATION PROCESSING SYSTEMS

Experiments – Quantitative Results

Effect of DaDA on Self-Supervised Learning (SSL)

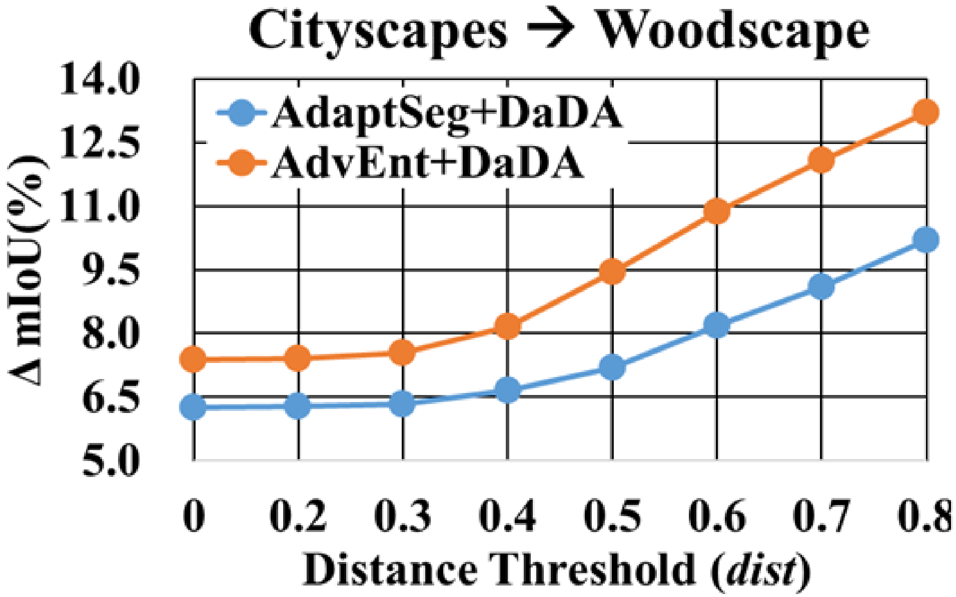
		Cityscapes → Woodscape		GTAV → Woodscape		Cityscapes → FDD		GTAV → FDD	
SSL Method	+DaDA	mIoU(%)	gain	mIoU(%)	gain	mIoU(%)	gain	mIoU(%)	gain
IAST [26]		47.00		38.83		39.60		37.47	
	✓	53.82	+6.82	**40.75**	**+1.92**	44.46	+4.86	40.06	+2.59
IntraDA [27]		48.92		36.10		40.36		38.61	
	✓	53.24	+4.32	39.85	+3.75	**45.28**	**+4.92**	**42.10**	**+3.49**
ProDA [40]		50.69		34.44		39.72		35.97	
	✓	**54.83**	**+4.14**	35.75	+1.31	42.14	+2.42	37.09	+1.12

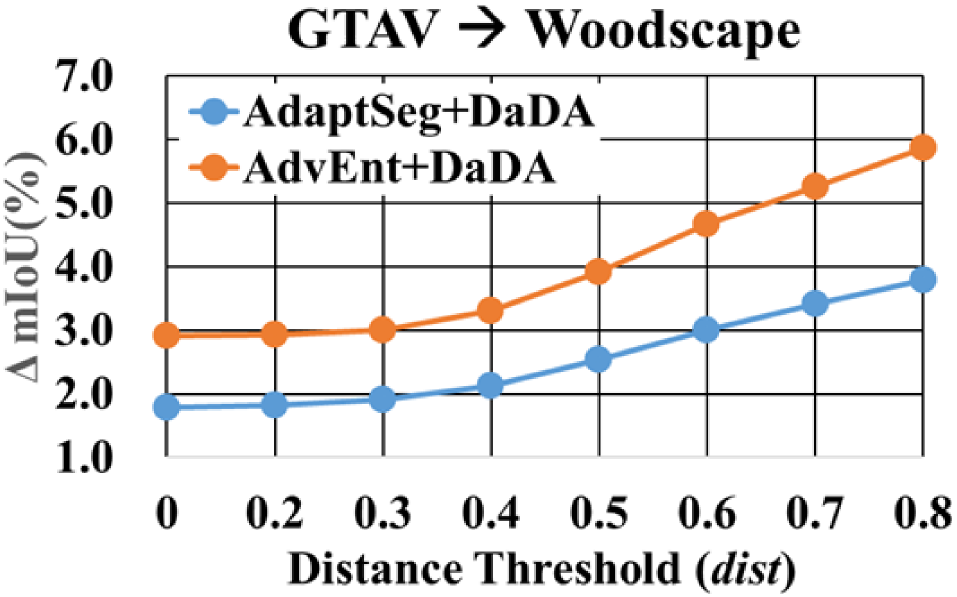
Experiments – Quantitative Results

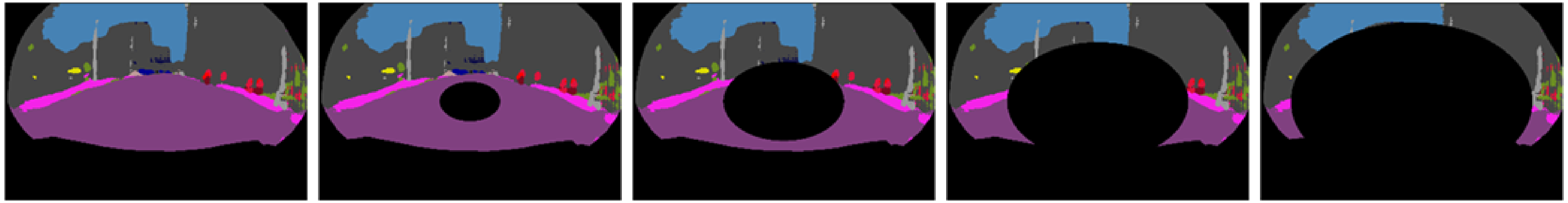
Ablation results on the distortion-aware losses.

Base Method	$+\mathcal{L}_{adv\_G}$	$+\mathcal{L}_{sem}$	$+\mathcal{L}_{recon}$	Cityscapes → Woodscape	GTAV → Woodscape
AdaptSeg [33]				46.33	35.94
	✓			49.61	36.45
	✓	✓		50.29	36.75
	✓		✓	49.97	37.17
	✓	✓	✓	**50.88**	**37.36**
AdvEnt [34]				45.26	34.70
	✓			47.77	35.36
	✓	✓		49.22	35.77
	✓		✓	50.32	36.11
	✓	✓	✓	**50.94**	**36.39**

Experiments – Distortion-aware mIoU(%)







dist=0.0 *dist*=0.2 *dist*=0.4 *dist*=0.6 *dist*=0.8

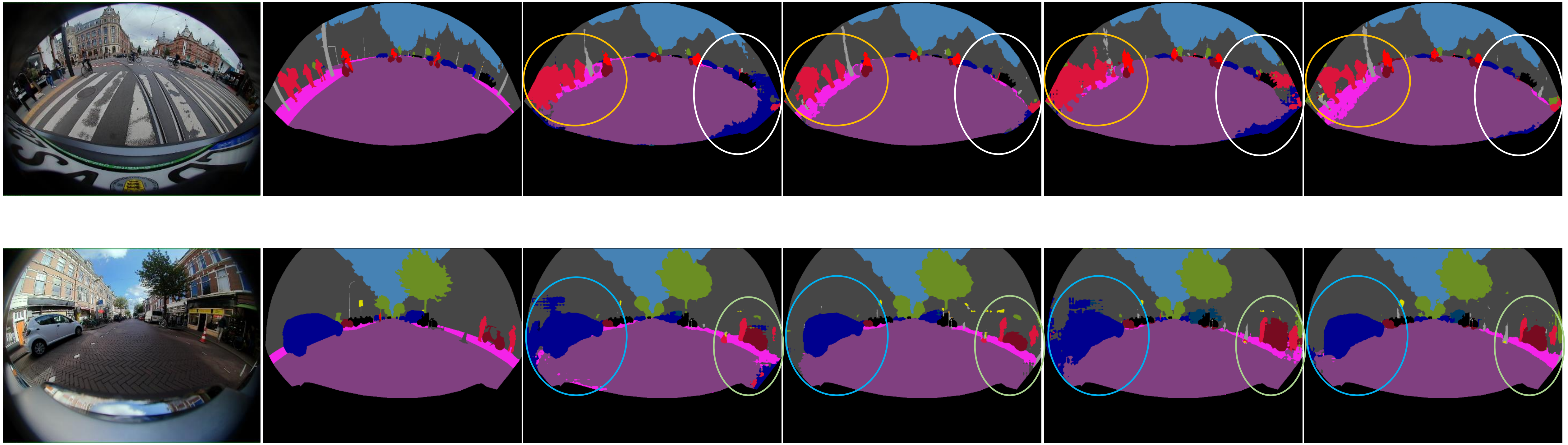
Experiments – Distortion-aware mIoU(%)

Table 8: **Performance gain achieved by adding DaDA increases as** $dist$ **increases.**

Cityscapes → Woodscape										
Method	dist=0.0	dist=0.2	dist=0.4	dist=0.6	dist=0.8	gain@0.0	gain@0.2	gain@0.4	gain@0.6	gain@0.8
AdaptSeg [10]	46.33	46.27	46.11	43.89	38.22					
AdaptSeg+DaDA	52.59	52.55	52.78	52.08	48.42	+6.26	+6.28	+6.67	+8.19	+10.20
AdvEnt [7]	45.26	45.19	44.97	42.54	37.44					
AdvEnt+DaDA	52.64	52.60	53.14	53.41	50.65	+7.38	+7.41	+8.17	+10.87	+13.21
GTAV → Woodscape										
Method	dist=0.0	dist=0.2	dist=0.4	dist=0.6	dist=0.8	gain@0.0	gain@0.2	gain@0.4	gain@0.6	gain@0.8
AdaptSeg [10]	35.94	35.92	35.68	33.95	30.46					
AdaptSeg+DaDA	37.73	37.74	37.80	36.95	34.25	+1.78	+1.82	+2.12	+3.00	+3.79
AdvEnt [7]	34.70	34.67	34.54	32.94	28.94					
AdvEnt+DaDA	37.62	37.61	37.85	37.60	34.81	+2.92	+2.94	+3.31	+4.66	+5.87

Experiments – Qualitative Result

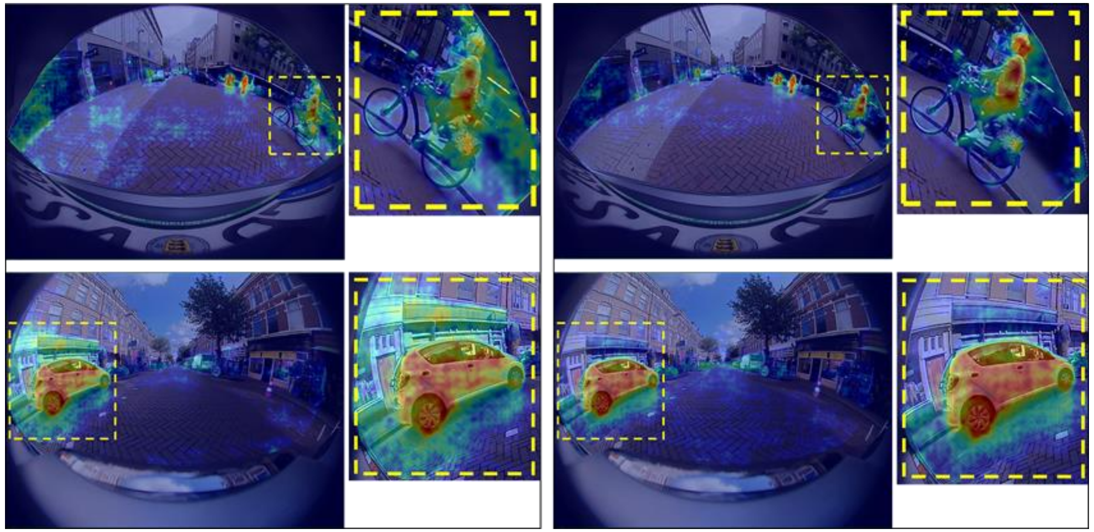
Qualitative Examples



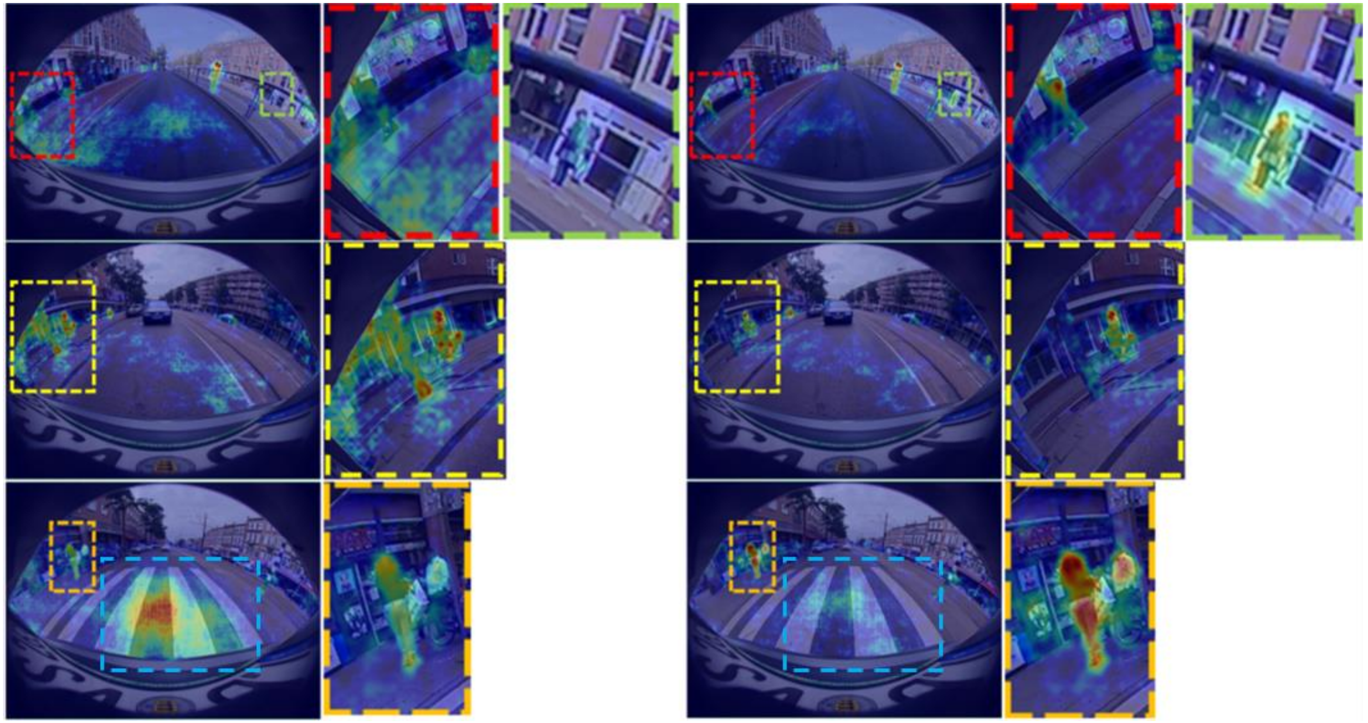
Target Image | Ground-Truth | AdaptSeg Only | AdaptSeg+DaDA | AdvEnt Only | AdvEnt+DaDA

Experiments – Qualitative Result

Class Activation Visualizations



w/o DaDA　　　　w/ DaDA (ours)



w/o DaDA　　　　w/ DaDA (ours)

Conclusion

- Practically Meaningful and New unsupervised domain adaptation benchmarks posing challenging tasks
  - Visual + **Distortion** domain gaps;
  - Fisheye Driving Dataset (FDD) available at https://sait-fdd.github.io/
- A novel distortion-aware domain adaptation (DaDA) framework;
  - Unsupervised and Unpaired Relative Distortion Learning;
  - Relative Deformation Field Generator based on diffeomorphism
- A solid baseline and new perspective on geometric distortion in unsupervised domain adaptation.

SAMSUNG ADVANCED INSTITUTE OF TECHNOLOGY

NEURAL INFORMATION PROCESSING SYSTEMS

Thanks for Watching!

FDD Dataset
&
More Information
https://sait-fdd.github.io/