Supplementary Material

In this supplementary material, we add additional experimental results, which are not included in the paper due to space limitations. We report the results of our method on the BDD100K [1] dataset. In addition, we added more visualization results and provided these image files.

A   Results on GTA5 to BDD100K

BDD100K [1] contains 10,000 real-world dashcam video frames with accurate pixel-wise annotations. It has a compatible label space with Cityscapes. It is divided into 7000 images as the training set and 2000 images as the validation set. We use this training set as the target domain data and evaluate our model on this validation set.

Table 1. Results on GTA5 to BDD100K

Method	Arch.	Road	SW	Build	Wall	Fence	Pole	TL	TS	Veg.	Terrain	Sky	PR	Rider	Car	Truck	Bus	Train	Motor	Bike	mIoU
Source		67.4	29.7	51.6	6.2	30.9	27.9	33.6	22.1	69.5	25.8	58.3	53.7	13.5	68.7	17.6	11.0	0.0	34.9	23.1	33.9
Warm-up	A&S	85.5	35.7	74.7	12.2	33.3	**32.7**	37.1	25.3	**70.1**	**26.1**	86.0	53.5	26.2	81.3	20.2	15.2	0.0	33.3	17.4	40.4
IAST		**90.0**	**53.7**	**74.7**	**13.4**	**33.4**	30.8	**39.4**	**25.8**	69.3	19.6	**90.0**	**55.3**	**30.9**	**82.2**	**29.2**	**38.2**	0.0	**40.0**	**39.1**	**45.0**

B   Additional visualization results

We show some pseudo-label results 2 and prediction results 1. More images are in the 'predictions' and 'pseudo_label' folders.

References

1. Yu, F., Xian, W., Chen, Y., Liu, F., Liao, M., Madhavan, V., Darrell, T.: Bdd100k: A diverse driving video database with scalable annotation tooling. arXiv preprint arXiv:1805.04687 **2**(5), 6 (2018) (document), A

Fig. 1. Visualization results (GTA5 to Cityscapes). Columns correspond to ground truth, source only, warm-up and our IAST

Fig. 2. Pseudo-label results. Columns correspond to original images with ground truth label, constant threshold method, class-balanced threshold method and our instance adaptive method