Appendix

Contents

The following items are included in the supplementary material:

- Exploration of the source domain shift in Section A.
- Class-wise performance of semantic segmentation in Section B.
- Ablation study on the scalar weight λ and the number of iterations for distribution optimization training in Section C.
- More visualization results in Section D.

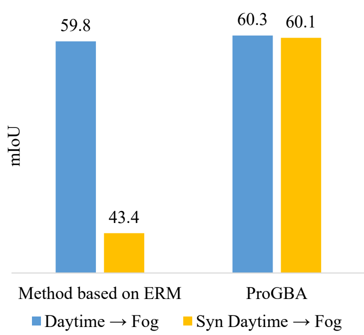

Fig. A: Impact of the source domain shifts. Performance of the adaptation parameters that are trained in Daytime → Fog and evaluated in Synthetic Daytime → Fog is illustrated.

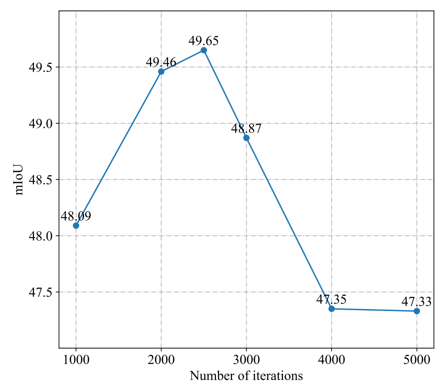

Fig. B: Impact of the number of distribution optimization iterations. The averaged mIoU of ProGBA adaptation from Cityscapes to all target domains in ACDC is provided.

A Exploration of the source domain shift

To further investigate the generalizability of distributions learned by ProGBA, we initially train a augmentation distribution tailored for Daytime → Fog. This distribution is then utilized to assess performance in Synthetic Daytime → Fog scenario. Similar operations are conducted on an ERM-based approach for comparison. We evaluate the model's ability to capture crucial characteristics by examining if the adaptation distributions developed for Daytime → Fog enhance performance in the Synthetic Daytime → Fog scenario. Results shown in Fig. A indicate that the adaptive augmentation learned by the ERM-based method underperforms in Synthetic Daytime → Fog, suggesting its ineffectiveness in discerning the fundamental disparities between daytime and foggy day. Conversely, despite the shift from actual to the synthetic daytime scenario, the

Table A: Zero-shot domain adaptation in semantic segmentation. Performance of ProGBA compared against PØDA [2] and source-only baseline. This table provides details of the main results in Table 1. All models are trained on the Cityscapes [1]. The best performance in each column is highlighted in **bold**.

Target eval	Method	road	sidewalk	building	wall	fence	pole	traffic light	traffic sign	vegetation	terrain	sky	person	rider	car	truck	bus	train	motorcycle	bicycle	mIoU
ACDC Night	Source-only	78.84	26.73	58.84	20.88	18.39	26.36	**22.16**	16.39	45.94	8.86	**4.53**	14.27	0.30	44.14	0.71	0.00	56.06	7.12	23.93	26.36
	PØDA [2]	**81.06**	33.11	**60.46**	28.11	25.66	27.42	21.16	18.78	43.92	10.58	0.34	**24.78**	0.01	40.65	**1.07**	0.00	55.02	**10.17**	26.14	28.24
	ProGBA	80.87	**34.03**	56.05	**28.49**	**27.26**	**28.97**	17.92	**22.66**	**47.32**	**12.74**	1.00	23.06	**2.59**	**47.69**	0.00	0.00	**65.38**	8.93	**38.33**	**30.18**
ACDC Snow	Source-only	75.95	34.35	76.73	39.33	**33.54**	**30.10**	47.72	**48.14**	79.28	**9.99**	94.41	38.82	**0.86**	73.54	70.36	**28.07**	65.29	**29.94**	26.48	47.47
	PØDA [2]	77.73	41.26	**78.71**	41.82	31.12	28.12	**57.83**	43.75	79.85	7.60	94.87	42.52	0.64	72.04	71.73	26.11	66.88	29.38	**31.70**	48.61
	ProGBA	**81.54**	**46.43**	77.70	**44.01**	31.59	26.66	55.68	43.59	**79.92**	8.91	**94.90**	**43.34**	0.04	**73.83**	**73.42**	26.89	**70.80**	28.07	22.53	**48.93**
ACDC Rain	Source-only	77.82	31.98	80.84	21.57	24.78	31.28	49.69	**45.64**	**87.19**	**53.41**	96.34	45.34	0.55	**77.86**	36.91	79.81	38.76	**16.23**	9.24	47.64
	PØDA [2]	78.81	33.80	80.74	21.68	25.20	**32.41**	**52.19**	42.65	86.32	52.74	96.42	**46.78**	0.18	77.23	36.14	**85.25**	31.70	10.10	**19.63**	47.89
	ProGBA	**80.62**	**44.83**	**82.84**	**26.40**	**30.72**	28.17	51.74	42.78	87.07	51.49	**96.44**	45.24	**0.59**	77.18	**49.85**	75.29	**42.03**	15.73	14.28	**49.65**
ACDC Fog	Source-only	92.58	70.48	**82.29**	47.47	33.30	37.87	57.79	**49.06**	**84.52**	51.53	95.86	27.80	**35.47**	**78.89**	**76.84**	82.57	**89.90**	19.11	23.50	59.83
	PØDA [2]	**92.91**	71.05	81.27	42.18	31.00	**38.09**	**56.98**	44.36	83.10	25.10	**96.10**	28.10	16.38	77.42	71.86	84.55	87.86	**50.53**	**31.20**	59.84
	ProGBA	92.81	**71.47**	81.84	**49.85**	**34.08**	30.11	56.94	46.52	84.45	**53.59**	95.57	**31.61**	34.68	77.52	72.73	**90.78**	87.84	28.67	25.46	**60.34**
GTA5	Source-only	76.69	27.19	78.67	40.03	14.74	27.47	30.02	**12.52**	**68.10**	41.34	**89.00**	61.44	39.68	78.71	62.58	**63.05**	0.00	44.84	23.08	46.27
	PØDA [2]	77.44	29.20	78.69	**40.53**	13.83	27.21	30.75	11.16	67.69	**43.13**	88.66	63.61	41.77	78.48	**64.34**	62.75	0.00	40.82	22.77	46.46
	ProGBA	**79.27**	**37.27**	**79.59**	39.75	**16.44**	**27.73**	**31.11**	11.39	67.95	42.55	88.60	**65.21**	**43.05**	**78.69**	61.93	57.90	0.00	**53.29**	**24.43**	**47.69**

Table B: Zero-shot domain adaptation in semantic segmentation. This table provides details of the main results in Table 1. All models are trained on the GTA5 [4]. The best performance in each column is highlighted in **bold**.

Target eval	Method	road	sidewalk	building	wall	fence	pole	traffic light	traffic sign	vegetation	terrain	sky	person	rider	car	truck	bus	train	motorcycle	bicycle	mIoU
Cityscapes	Source-only	**85.77**	23.11	83.02	38.59	31.64	20.83	**32.90**	16.83	84.60	**45.58**	**87.45**	54.64	11.22	78.89	25.80	48.38	**0.15**	**18.73**	31.37	43.13
	PØDA [2]	78.47	**31.17**	**83.66**	43.12	**32.97**	**23.23**	31.59	**20.44**	**85.00**	29.66	87.01	**55.60**	13.92	78.08	28.63	43.18	0.00	18.56	**43.20**	43.55
	ProGBA	84.46	25.98	82.75	**44.65**	28.76	19.96	32.16	16.69	83.93	43.95	86.95	53.62	**16.26**	**79.70**	**30.13**	**49.31**	0.00	17.79	32.76	**43.67**

augmentation distribution learned by ProGBA successfully adjusts to foggy day conditions, maintaining the same level of performance as in Daytime → Fog. This highlights ProGBA's capacity to discern more profound distinctions between source and target domains, thereby affirming its efficiency and robustness.

B Class-wise performance

Table C: Performance evaluation of different λ values. ProGBA adaptation results(mIoU) from Cityscapes to ACDC are reported. The best performance in each column is highlighted in **bold**.

λ	ACDC Fog	ACDC Night	ACDC Rain	ACDC Snow	Mean
0.001	**60.69**	30.00	48.93	46.89	47.59
0.01	59.57	29.55	**49.88**	47.69	**48.45**
0.1	58.82	**30.53**	46.72	**48.83**	48.18
1.0	60.10	27.85	48.69	47.29	47.44

We present class-wise IoUs for semantic segmentation in Table A and Table B. In detail, Table A showcases outcomes from models trained on the Cityscapes dataset [1] and evaluated on various domain validation sets from ACDC [5]. Meanwhile, Table B details results from models trained on the GTA5 [4] dataset and assessed on the Cityscapes [1] validation set, which represents the CS $\rightarrow$ Synthesis scenario. Specifically, the lower IoU results for certain classes may be due to their complexity and distinct characteristics. Traffic signs are smaller and less frequent, while vegetation has high variability. These factors challenge retaining consistent semantic information during domain adaptation.

C Additional ablation study

C.1 Scalar weights

Table C showcases a performance comparison across various target domains within the ACDC dataset [5], evaluating the effect of different λ parameter values. With an increase in λ, there's a noticeable uptick in the mean IoU, peaking at $\lambda = 0.01$. Beyond this point, performance begins to decline with further increases in λ. Therefore, 0.01 as the optimal value for λ is selected.

C.2 Number of iterations

In all experiments conducted, we carry out 2500 iterations of distribution optimization. Fig. B illustrates the impact of varying the total number of iterations, where we identify a critical inflection point at 2500 iterations. Iterations below this threshold prove inadequate for achieving style alignment, whereas exceeding 2500 iterations leads to a decline in performance. This phenomenon may be caused by *over-stylization* [3].

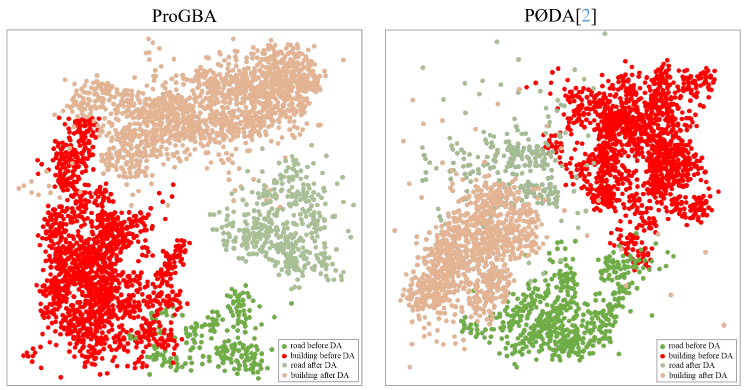

Fig. C: The t-SNE visualization of features. Different categories and the use of domain adaptation are indicated by colors.

D Visualization

D.1 Feature visualization

Fig. C shows the t-SNE visualizations of PØDA and ProGBA. PØDA's features are more dispersed due to randomly sampling stored (μ_i, σ_i), leading to less clear decision boundaries. In contrast, ProGBA produces clearer feature boundaries because augmented features are sampled from the same distribution for the same target domain, ensuring less deviation.

D.2 More qualitative examples

We offer additional visualization results for semantic segmentation of Cityscapes $\rightarrow$ ACDC in Fig. D and Cityscapes $\rightarrow$ GTA5(Synthesis) in Fig. E. As for Cityscapes $\rightarrow$ ACDC, our predictions with ProGBA show better performance in dark light or with heavy occlusion, demonstrating superior object segmentation visual results compared to both the source-only approach and the model utilizing PØDA [2]. In the context of Cityscapes $\rightarrow$ GTA5(Synthesis), the ProGBA-enhanced model excels in handling large objects, yielding precise segmentation outcomes.

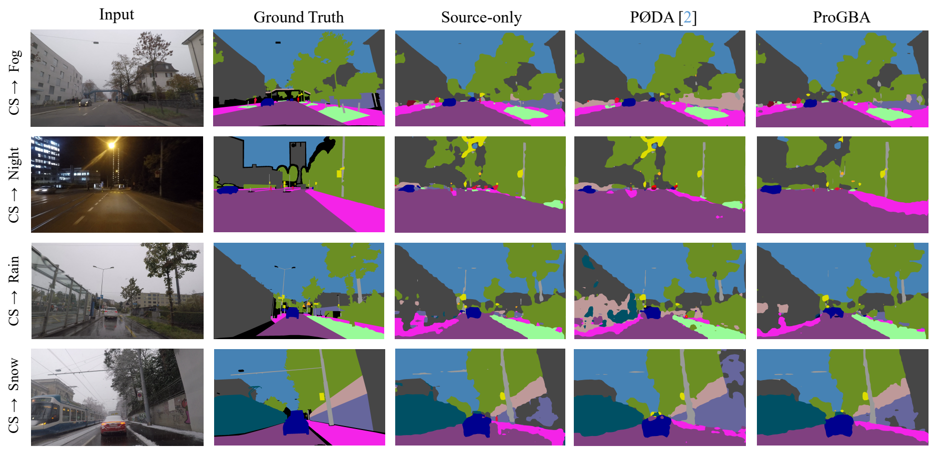

Fig. D: Qualitative examples on Cityscapes → ACDC validation set.

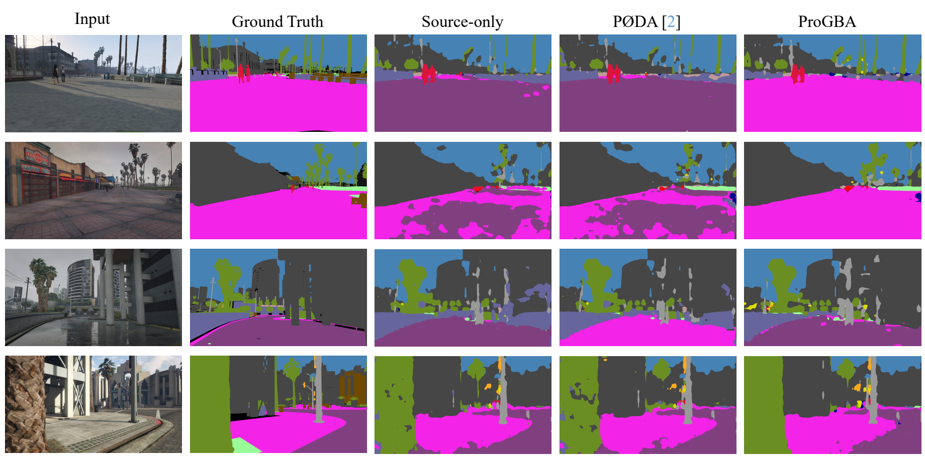

Fig. E: Qualitative examples on Cityscapes → GTA5(Synthesis) validation set.

References

1. Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B.: The cityscapes dataset for semantic urban scene understanding. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3213–3223 (2016)
2. Fahes, M., Vu, T.H., Bursuc, A., Pérez, P., De Charette, R.: Poda: Prompt-driven zero-shot domain adaptation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 18623–18633 (2023)
3. Kwon, G., Ye, J.C.: Clipstyler: Image style transfer with a single text condition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 18062–18071 (2022)
4. Richter, S.R., Vineet, V., Roth, S., Koltun, V.: Playing for data: Ground truth from computer games. In: Computer Vision–ECCV 2016: 14th European Conference,

Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14. pp. 102–118. Springer (2016)

5. Sakaridis, C., Dai, D., Van Gool, L.: Acdc: The adverse conditions dataset with correspondences for semantic driving scene understanding. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 10765–10775 (2021)