D^{3}CTTA: Domain-Dependent Decorrelation for Continual Test-Time Adaption of 3D LiDAR Segmentation

Supplementary Material

A. Quantitative results

A.1. Quantitative Comparisons on Light Corruption Data

Tabs. A.1 and A.2 present the CTTA results for Synth4D to SemanticKITTI-C and nuScenes-C under light corruption levels, respectively. Our method achieves SOTA performance on both benchmarks. Specifically, for Synth4D to SemanticKITTI-C, our method achieves SOTA results across 5 domains, with an average mIoU improvement of 26.46% compared to the source model and 9.42% compared to the previous SOTA method. Similarly, for Synth4D to nuScenes-C, our method achieves SOTA results across 4 domains, with an average mIoU improvement of 7.91% over the source model and 2.45% over the previous SOTA method. These results further demonstrate the superiority of our approach.

A.2. Results on the inverse order

The results under the reversed perturbation order remain nearly identical to our original setting, with an average mIoU of 30.88 compared to 30.82 in our standard setup. This indicates that the overall performance is largely unaffected by the order of domain shifts. Among all domains, only Motion Blur and Crosstalk exhibit mIoU differences greater than 0.2, suggesting that while minor variations exist, the method maintains consistent adaptation across different perturbation sequences. This highlights the robustness of our approach to domain order, ensuring stable performance regardless of the sequence in which the corrupted domains appear.

A.3. Results on repeated domains

Tab. A.4 presents the performance over 13 rounds of repeated domains. Our method shows continuous improvement as the number of rounds increases, benefiting from

Table A.1. Quantitative comparison of continual test-time adaptation from Synth4D to SemanticKITTI-C under light corruption levels. We report the mIoU for each domain as well as the average mIoU across all methods. BP refers to backpropagation.

Method	DA Type	BP	Domain							mIoU
			Beam	Sensor	Crosstalk	Fog	Echo	Motion	Snow	
PSLabel	TTA	✓	**33.00**	**35.54**	29.20	26.09	27.61	25.33	24.72	28.78
TENT[5]	TTA	✓	24.94	18.22	15.97	10.57	7.62	12.68	12.57	14.65
GIPSO[4]	3D TTA	✓	30.23	25.10	28.97	29.19	31.82	24.18	23.71	27.60
EATA[3]	CTTA	✓	28.48	33.21	31.69	25.16	33.85	30.61	22.14	29.31
ViDA[2]	CTTA	✓	30.40	25.24	26.45	27.10	28.43	19.70	23.67	25.86
Source	-	✗	30.24	24.40	25.73	27.02	28.56	18.18	23.37	25.36
T3A[1]	TTA	✗	31.26	23.83	25.07	26.89	28.14	17.73	23.93	25.26
Ours	3D CTTA	✗	29.54	33.80	**33.20**	**31.54**	**36.80**	**32.75**	**26.87**	**32.07**

Table A.2. Quantitative comparison of continual test-time adaptation from Synth4D to nuScenes-C under light corruption levels. We report the mIoU for each domain as well as the average mIoU across all methods. BP refers to backpropagation.

Method	DA Type	BP	Domain							mIoU
			Beam	Sensor	Crosstalk	Fog	Echo	Motion	Snow	
PSLabel	TTA	✓	28.00	25.89	26.14	22.01	22.39	21.45	20.22	23.73
TENT	TTA	✓	13.27	5.71	3.95	6.89	5.88	7.68	10.63	7.72
GIPSO	3D TTA	✓	28.43	27.44	24.20	23.37	32.78	25.90	27.65	27.11
EATA	CTTA	✓	27.96	25.91	26.70	23.78	30.70	28.70	27.35	27.30
ViDA	CTTA	✓	**30.83**	**28.33**	20.74	23.21	**34.14**	19.48	27.35	26.30
Source	-	✗	30.56	28.29	19.77	22.63	34.08	18.94	27.14	25.92
T3A	TTA	✗	29.17	26.74	18.11	21.93	32.79	17.71	26.36	24.69
Ours	3D CTTA	✗	27.89	26.50	**27.91**	**24.45**	31.77	**29.14**	**28.10**	**27.97**

Table A.3. Quantitative results under reverse order

domain	Beam	Sensor	Crosstalk	Fog	Echo	Motion	Snow	Avg
			t $\longrightarrow$					
forward	29.47	34.56	32.36	26.45	35.34	32.67	24.92	30.82
			t $\longleftarrow$					
inverse	29.66	34.67	32.16	26.30	35.51	32.99	24.98	30.88

Table A.4. Results on repeated domains with 13 rounds

Round	1	4	7	10	13
Ours(w domain)	31.14	31.30	31.37	31.40	31.42
Ours(w/o domain)	31.04	31.00	31.00	30.99	30.99
T3A	24.98	24.98	24.98	24.98	24.98
EATA	29.35	27.68	8.77	3.27	2.54

Table A.5. Results on Synth4D-SemanticKITTI

method	source	T3A	ViDA	EATA	Ours
mIoU	34.21	34.23	34.23	39.67	40.66

Table A.6. Ablation of K in KNN

K	5	10	20	50	100
mIoU	30.70	30.73	30.82	30.67	30.65

the domain detection module, whereas other methods either remain unchanged or experience significant performance degradation.

A.4. Results on Synth4D-SemanticKITTI

We conduct the TTA setting using Synth4D-SemanticKITTI, and our method still achieves state-of-the-art performance, with a 18.85% improvement over the source model and a 2.50% improvement over the previous SOTA method, EATA, as shown in Tab. A.5.

B. Ablations

B.1. Ablation of K in KNN

As shown in Tab. A.6, selecting a small K increases false positives in pseudo labels, while a large K results in the omission of more true positives. The optimal performance is achieved at K=20.

B.2. Ablation of domain similarity thresholds τ_c

We test the domain detection accuracy under different domain thresholds τ_c, as shown in the Tab. A.7. Smaller thresholds fail to effectively distinguish different domains, while larger thresholds tend to split point clouds within the same domain into too many separate domains.

Table A.7. Ablation of domain thresholds.

threshold	0.75	0.8	0.85	0.9	0.95
accuracy	58.91	76.13	84.28	78.01	48.05

C. Qualitative Comparisons

Fig. A.1, A.2, and A.3 present visualization results from the SemanticKITTI-C dataset. We compared the source model, ground truth, the SOTA BP-free method (T3A), the SOTA BP method (EATA), and our method. The regions highlighted in red illustrate the differences in predictions. It is evident that both EATA and our method significantly improve the performance of manmade structures and vegetation (highlighted in yellow and gray) compared to the source model and T3A. Additionally, there is a noticeable improvement in the accuracy of vehicle predictions. Compared to EATA, our method often provides more comprehensive predictions for vehicles and manmade structures.

D. Visualizations of Pseudo Labels

Fig. A.4 shows the visualization results of pseudo labels filtered using entropy, probability, and KNN. We ensure that the number of retained points is consistent across the three filtering methods. It can be observed that pseudo labels filtered through KNN exhibit better spatial smoothness, ensuring prediction consistency for neighboring points, thereby achieving improved performance. However, this approach may inadvertently remove boundary points where different objects are adjacent. This case represents a limitation of the current method.

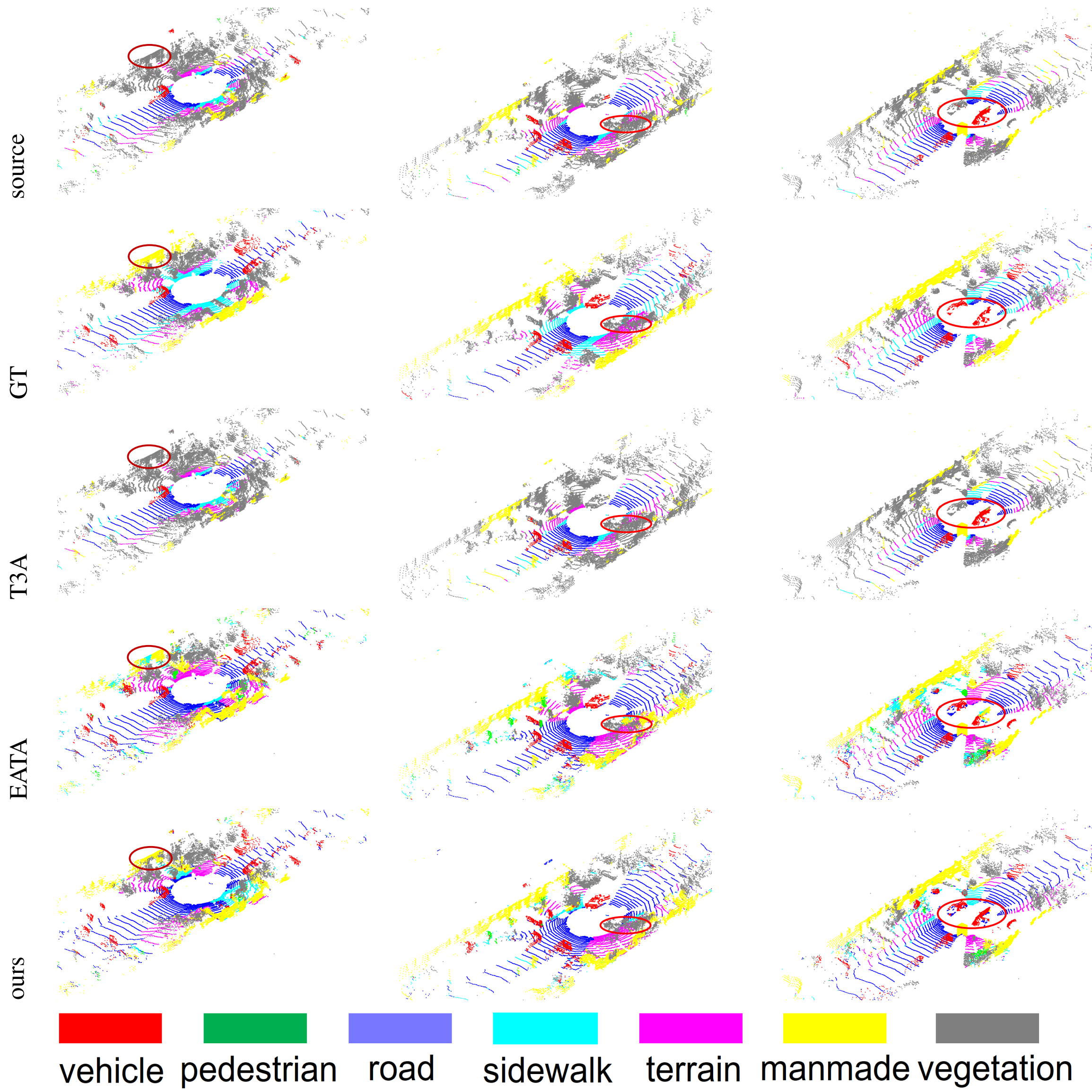

Figure A.1. Visualization results of the source model, ground truth, T3A, EATA and our method. Regions highlighted in red illustrate the differences in predictions.

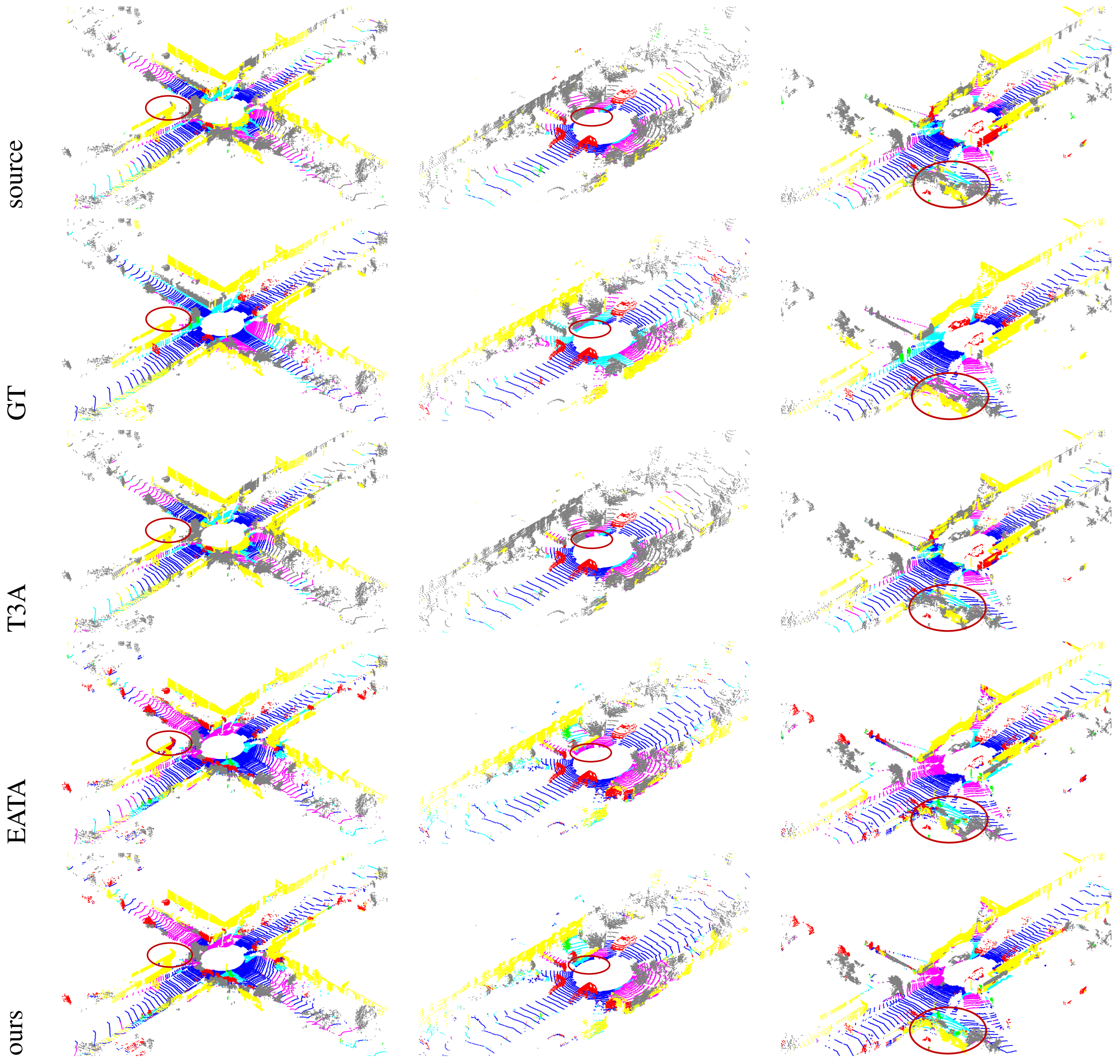

Figure A.2. Visualization results of the source model, ground truth, T3A, EATA and our method. Regions highlighted in red illustrate the differences in predictions.

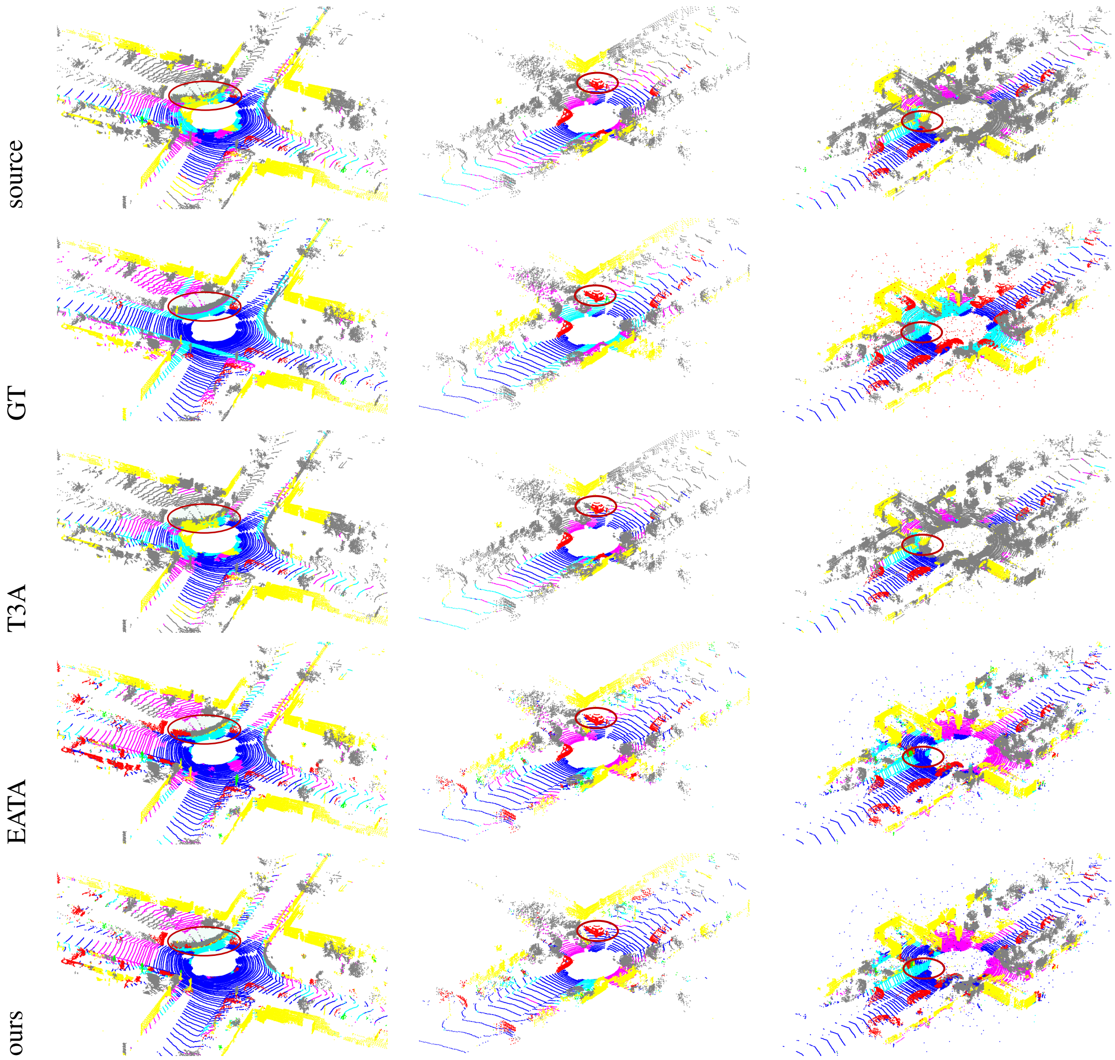

Figure A.3. Visualization results of the source model, ground truth, T3A, EATA and our method. Regions highlighted in red illustrate the differences in predictions.

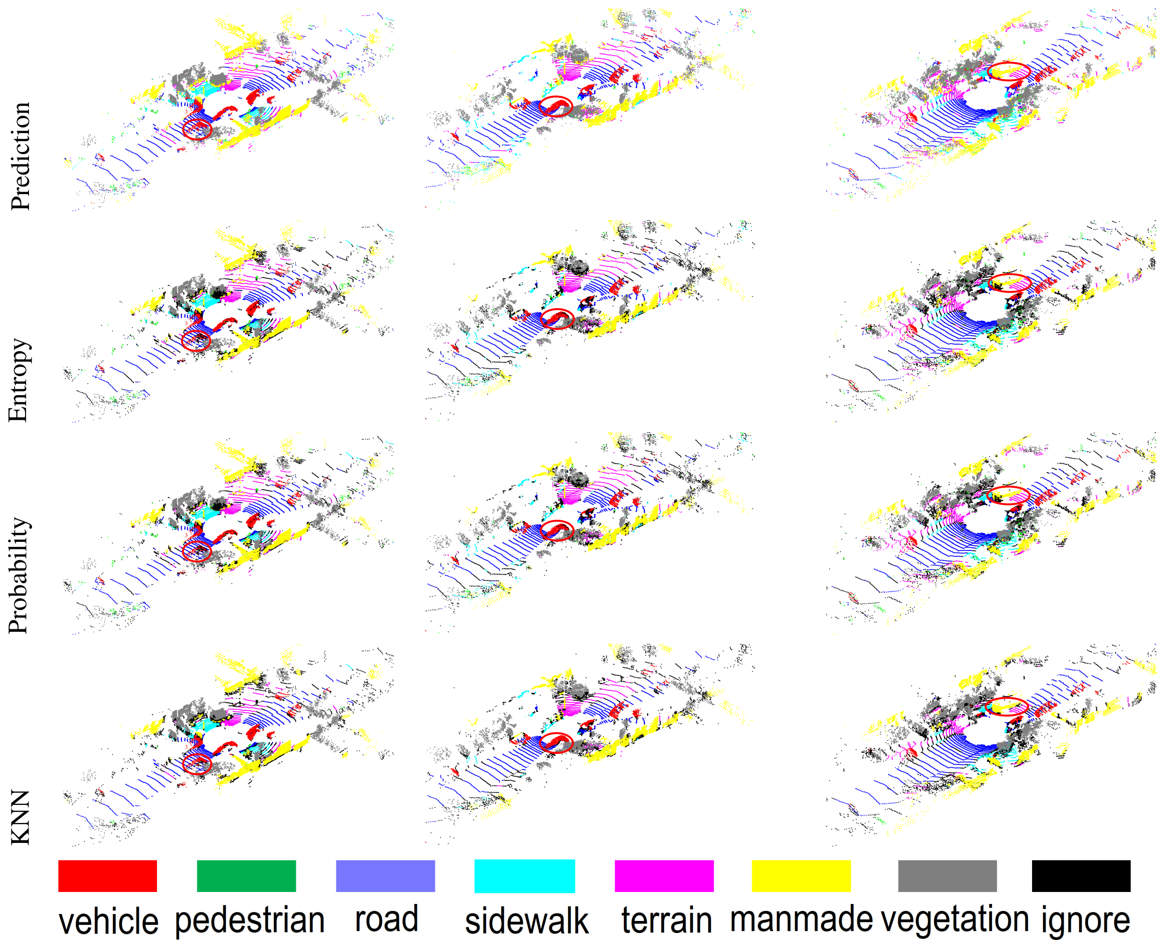

Figure A.4. Visualization of source model predictions and pseudo labels filtered by entropy, probability, and KNN. Regions highlighted in red indicate differences in predictions.

References

[1] Yusuke Iwasawa and Yutaka Matsuo. Test-time classifier adjustment module for model-agnostic domain generalization. *Advances in Neural Information Processing Systems*, 34: 2427–2440, 2021. 1

[2] Jiaming Liu, Senqiao Yang, Peidong Jia, Renrui Zhang, Ming Lu, Yandong Guo, Wei Xue, and Shanghang Zhang. ViDA: Homeostatic visual domain adapter for continual test time adaptation. In *The Twelfth International Conference on Learning Representations*, 2024. 1

[3] Shuaicheng Niu, Jiaxiang Wu, Yifan Zhang, Yaofo Chen, Shijian Zheng, Peilin Zhao, and Mingkui Tan. Efficient test-time model adaptation without forgetting. In *International conference on machine learning*, pages 16888–16905. PMLR, 2022. 1

[4] Cristiano Saltori, Evgeny Krivosheev, Stéphane Lathuiliére, Nicu Sebe, Fabio Galasso, Giuseppe Fiameni, Elisa Ricci, and Fabio Poiesi. Gipso: Geometrically informed propagation for online adaptation in 3d lidar segmentation. In *European Conference on Computer Vision*, pages 567–585. Springer, 2022. 1

[5] Dequan Wang, Evan Shelhamer, Shaoteng Liu, Bruno Olshausen, and Trevor Darrell. Tent: Fully test-time adaptation by entropy minimization. In *International Conference on Learning Representations*, 2021. 1