Supplementary Materials

AH-OCDA: Amplitude-based Curriculum Learning and Hopfield Segmentation Model for Open Compound Domain Adaptation

A. Datasets

Labeled synthetic source domain GTA5 [26] consists of 24,966 synthesized images of road scenes taken from a video game. Each image sized to 1280×720, and random cropped to 1024×512 for training. Following previous works, we selected 19 classes (total 33 classes provided by the original game engine). SYNTHIA [27] consists of 2224 synthetic images of virtual urban road scenes. The images are randomly cropped to 1024×512 during training. Following previous works, we selected 11 classes (road, sidewalk, building, wall, fence, pole, light, vegetation, sky, person, and car).

Unlabeled compound and open domain. BDD100K [39] comprises 14,697 unlabeled real images of mixed domains and an additional 1430 images with labels for validation. BDD100K includes 'rainy', 'snowy', and 'cloudy' as compound domains, and 'overcast' as as unseen domain. For the scope of our research, other domains in the original BDD100K dataset were not employed.

Extended unlabeled open domain. Cityscapes [3] provides 2,975 real urban road scenes from various European cities for semantic segmentation task. We utilize its validation set with 500 images for extended open evaluation. KITTI [5] is real road scene images collected from Karlsruhe, Germany. We utilize its validation set with 200 images for extended open evaluation.

B. Intuitive Insights into the Hopfield Segmentation Model

Figure 1 offers high-level insights into why the Hopfield segmentation model in AH-OCDA achieves strong performance and highlights the effectiveness of its approach in addressing the OCDA problem. We compare the Hopfield segmentation model with the prototype-based DA method in the OCDA setting especially focusing on how would each handle compound domains and obtain representations.

Prototype-based models use explicit feature vectors as class representations by updating the prototype through the training of compound domains (Fig.1(a)) thus it may incorrectly predict when C1B is close to the prototype of class A (Fig.1(b)). A similar result occurs when UA is close to the prototype of class B (Fig.1(c)). In contrast, our Hopfield segmentation model implicitly stores both intra-class and inter-class information in its parameters, enabling accurate predictions even for distorted features from compound or unseen domains (Fig.1(d)). Furthermore, this clearly distinguishes from segmentation models that utilize the conventional Hopefield network as an auxiliary optimization module. In addition, it supports the findings of the ablation study presented in Section 4.5.3, which demonstrates that freezing the Hopfield segmentation model during training on the compound domain yields better performance.

C. Qualitative Result

The qualitative results of GTA5 $\rightarrow$ C-Driving are presented in Fig. 2 $\sim$ Fig. 6. In addition to the AH-OCDA, we present the qualitative results of each component of AH-OCDA. 'w/o Curr' and 'w/o Hopf' denote AH-OCDA without curriculum learning and AH-OCDA without the Hopfield network, respectively. While AH-OCDA shows phenomenal segmentation results, there are some consistent tendencies corresponding to each component of AH-OCDA.

Specifically, it is possible to infer the significance of amplitude-based curriculum learning through the comparative analysis of 'w/o Hopf' and 'Source only'. As discussed in the main manuscript, amplitude-based curriculum learning gradually trains the segmentation model to achieve proficient performance even on the compound domain far from the source domain. Upon comparing 'w/o Hopf' with 'Source only', it is evident that amplitude-based curriculum learning substantially improved segmentation results. Especially for the 'Night' domain, which is far from the source domain, 'Source Only' shows poor performance where it is difficult to distinguish the car shape. In contrast, 'w/o Hopf' shows improved segmentation results, enabling the differentiation of the car's pixels from the road. These results support that amplitude-based curriculum learning led to the successful adaptation of the segmentation model to

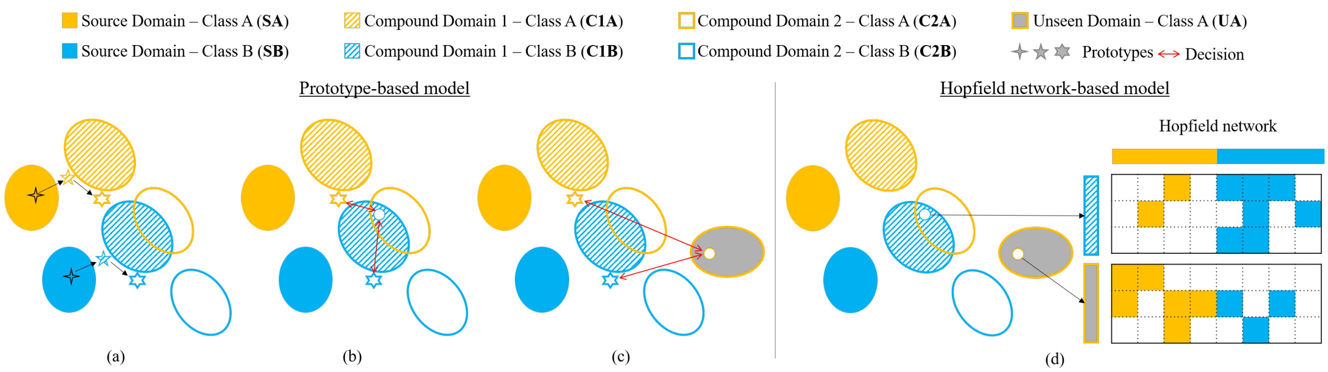

Figure 1. The high-level illustration of prototype-based model and Hopfield network-based model on OCDA setting.

compound domains, including a domain that is close to the source domain to the farthest domain. Moreover, the Hopfield segmentation model maps the feature distribution of the target image to the source domain. The 'w/o Curr' results show stable segmentation across all domains, compared to 'Source Only' and 'w/o Hopf'. Specifically, segmentation results for 'Source Only' and 'w/o Hopf' exhibit unstable predictions often showing a mosaic pattern. The different patterns in segmentation results support that the Hopfield segmentation model retrieves the pattern of features in the source domain for the consistent prediction of one object. AH-OCDA produces precise and stable segmentation results in both compound and open domains, as the amplitude-based curriculum learning and the Hopfield segmentation model complement each other, leveraging their respective strengths.

Lastly, we acknowledge the efforts of those who contributed to the BDD100K dataset and inevitable errors in labeling even with meticulous efforts. Through examples of the qualitative results, we draw attention to the presence of mislabeled samples within the BDD100K dataset to inform future researchers in the field. As of what we have found, in Fig. 4, the road is labeled as the sky, and in Fig. 5, most of the pixels except for the car are labeled as the road when it's not. Nonetheless, we utilized the original labels provided from the dataset for a fair comparison with prior works.

D. Limitations

As demonstrated in Table 5 of the main manuscript, AH-OCDA exhibits robustness to variations in the number of K when compared to the previous method. However, when there is an imbalance in image distribution within a compound domain—for example, when images are predominantly distributed at the beginning of the curriculum— AH-OCDA could experience performance degradation due to dividing the sorted compound domain images by K splits. In such a situation, we may alternatively define a threshold value rather than the number of K. When the distance between two adjacent images in the sorted compound domain images exceeds the defined threshold, we can consider that they have distant domain information. This approach may even be beneficial when there is a huge domain gap between the domains in the compound domain. Nevertheless, we opted not to implement thresholding in AH-OCDA as thresholding the distance value requires more precise consideration and is considerably more sensitive compared to the use of the number of K. In our future research, we aim to remove the hyperparameter K to enhance generalizability.

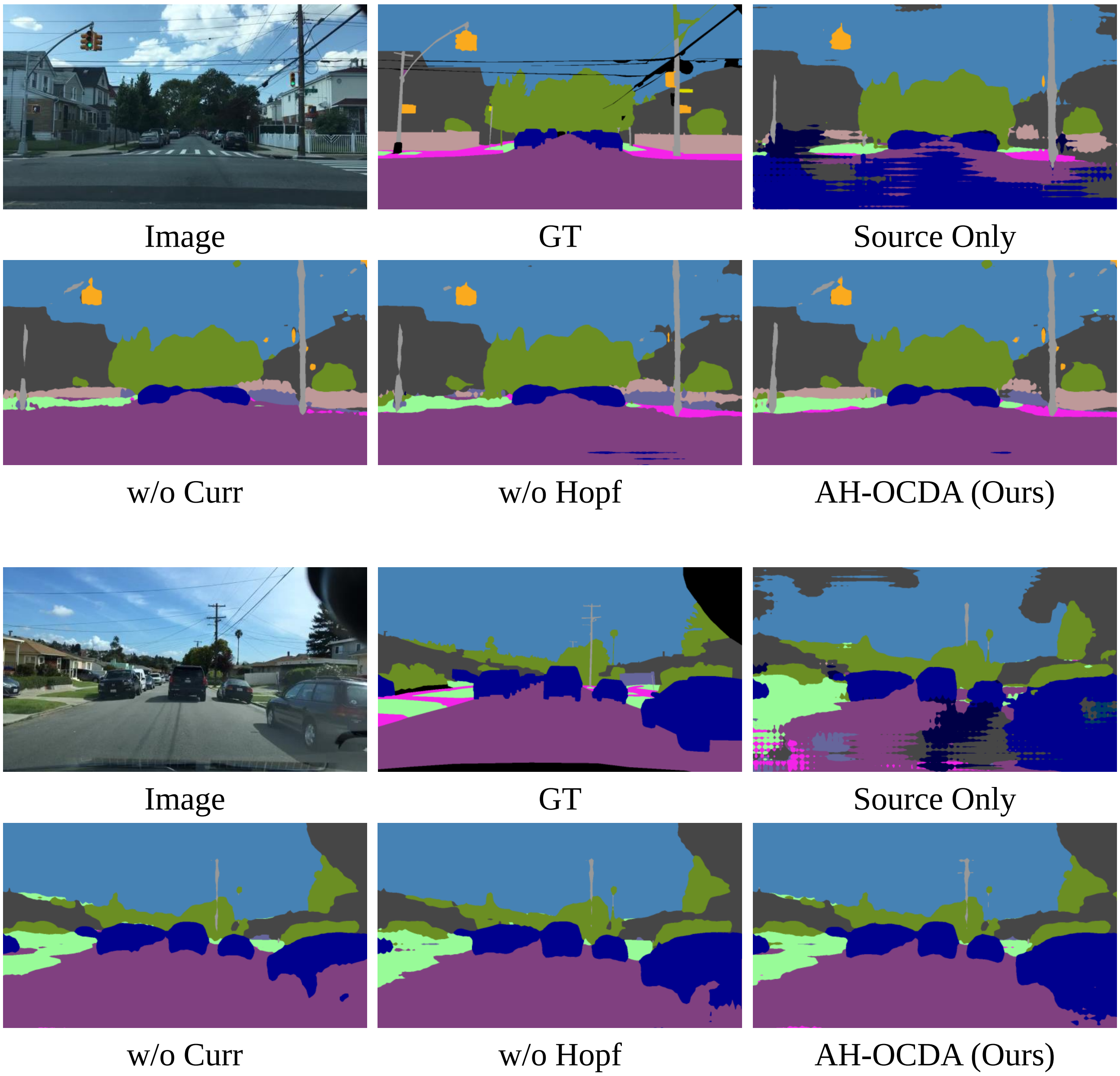

Figure 2. Qualitative analysis on ‘Cloudy’ domain of GTA5 $\rightarrow$ C-Driving

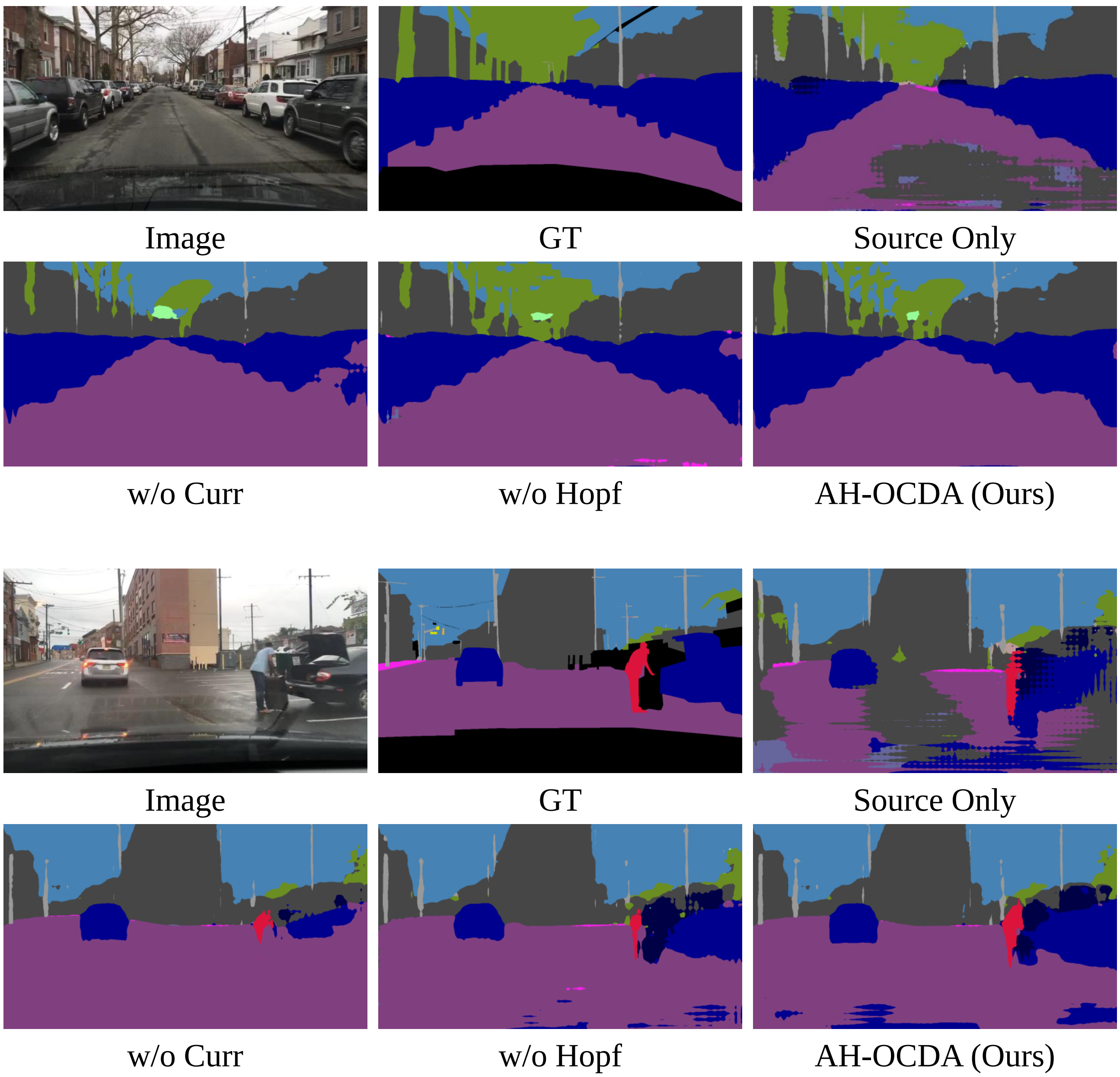

Figure 3. Qualitative analysis on ‘Rainy’ domain of GTA5 → C-Driving

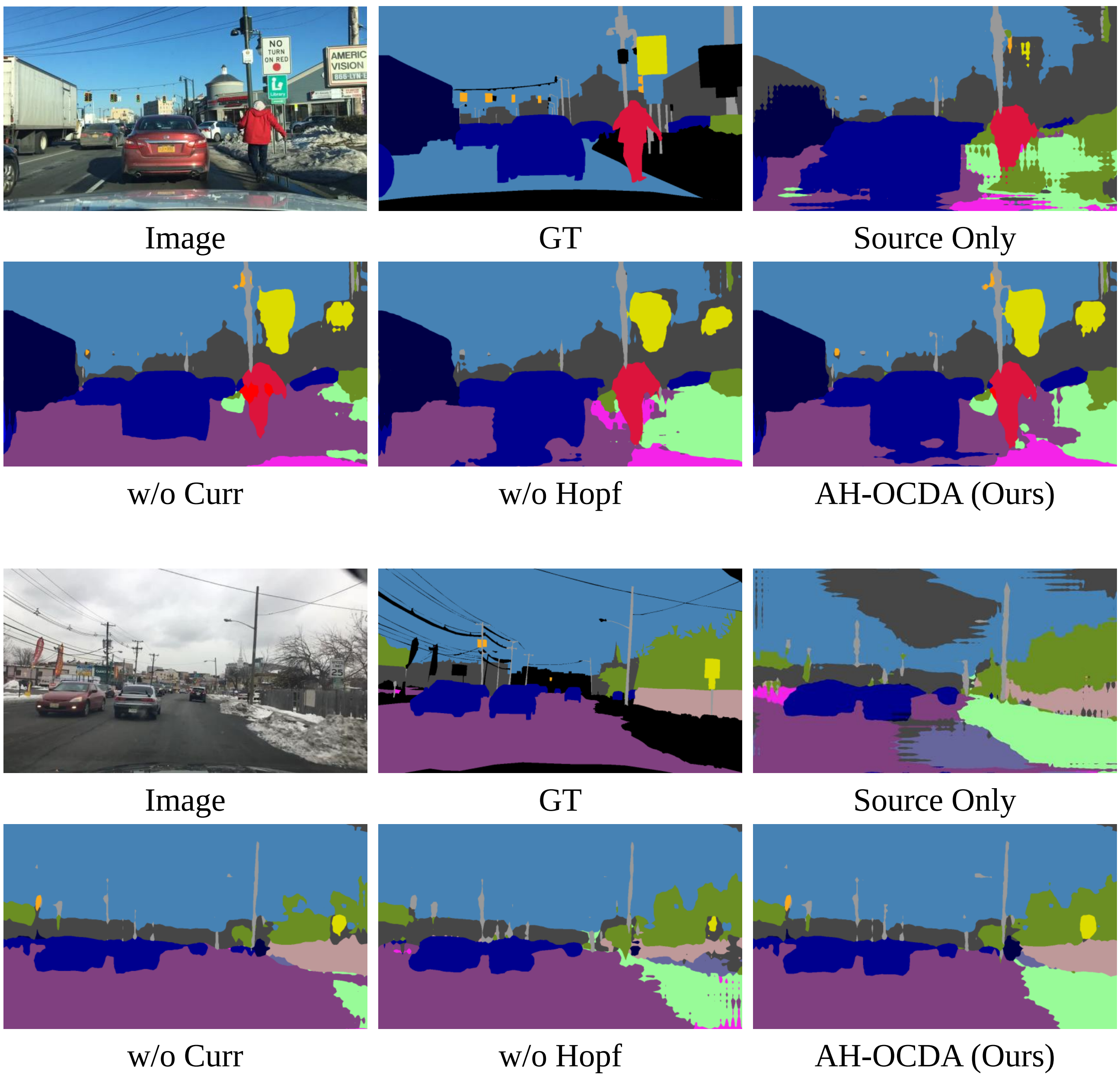

Figure 4. Qualitative analysis on ‘Snowy’ domain of GTA5 $\rightarrow$ C-Driving

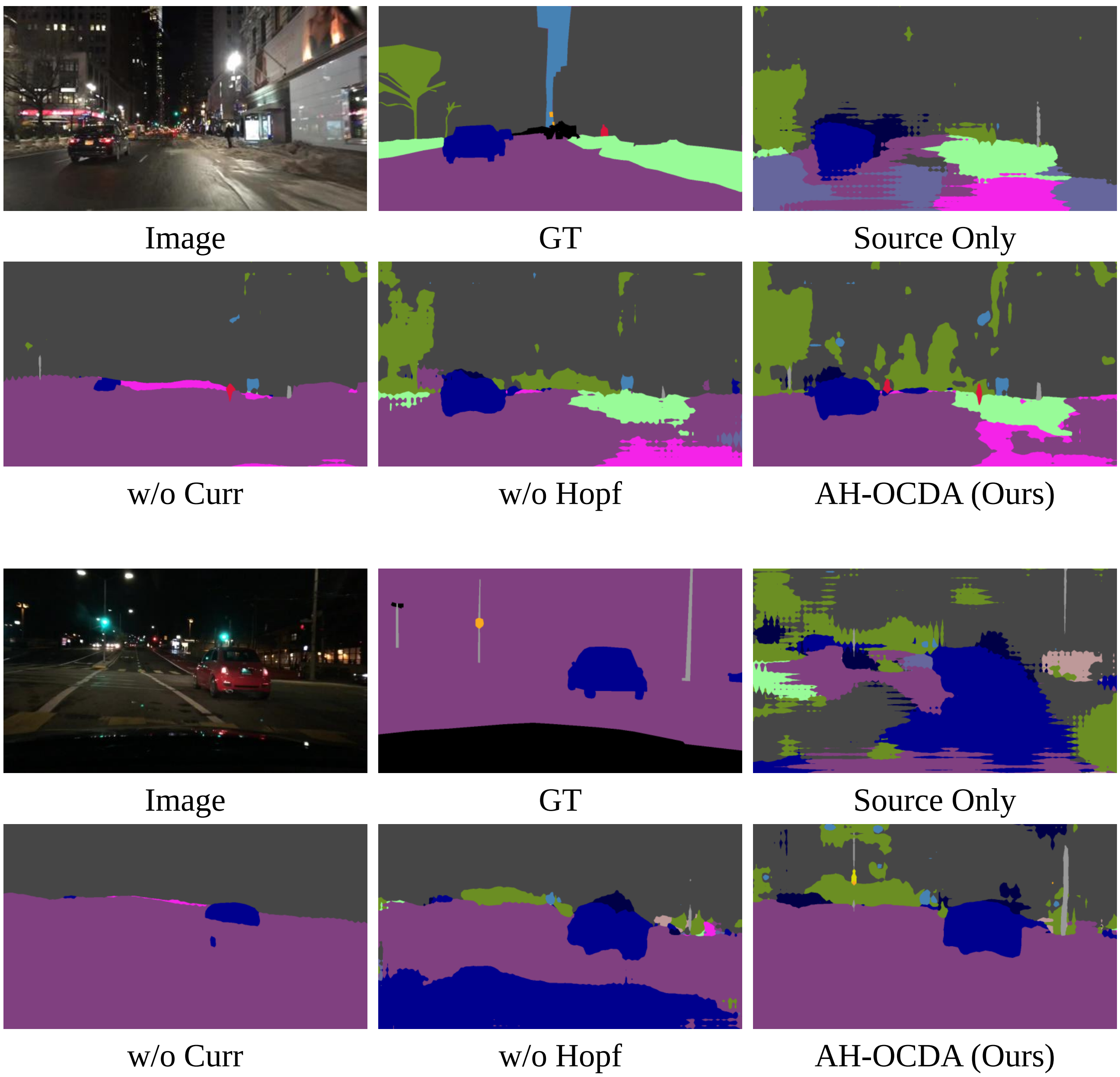

Figure 5. Qualitative analysis on 'Night' domain of GTA5 $\rightarrow$ C-Driving

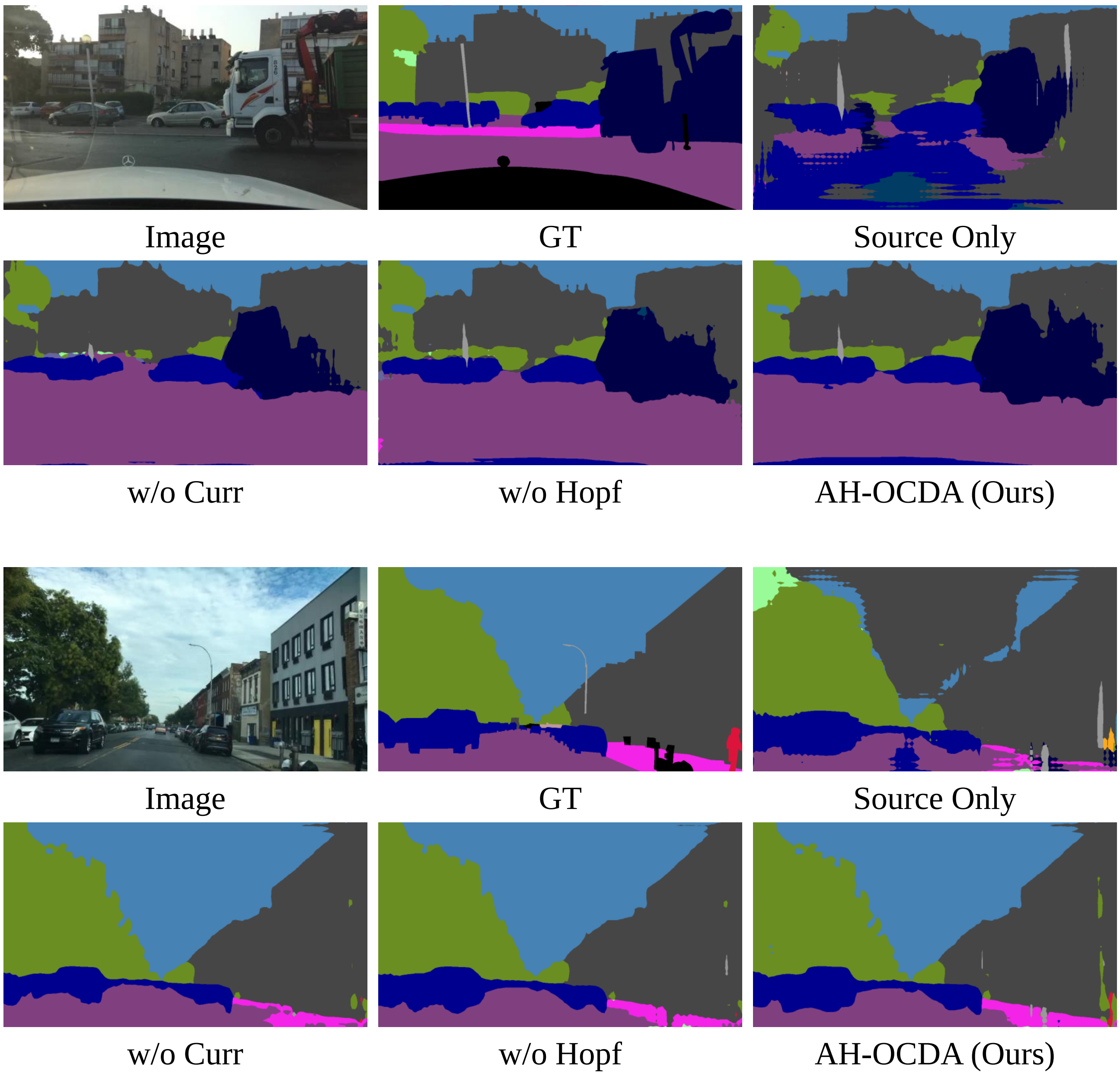

Figure 6. Qualitative analysis on ‘Overcast’ domain of GTA5 $\rightarrow$ C-Driving