HeightMapNet: Explicit Height Modeling for End-to-End HD Map Learning

Wenzhao Qiu, Shanmin Pang*, Hao Zhang, Jianwu Fang, Jianru Xue
Xi'an Jiaotong University
{a2801551754@stu., pangsm@, belief@stu., fangjianwu@, jrxue@}xjtu.edu.cn

A. Extended Ablation Studies

Table 1. Ablation study on the nuScenes `val` set, comparing our model with one where the foreground and height maps in HeightMapNet are replaced by a depth map. The depth loss is computed according to the MapTR V2 framework.

Method	Depth Map	+ Depth Loss	Height Map
mAP	50.1	55.2	56.4

Here, we further conduct experiments on the visual transformation module, and present the outcome of replacing the foreground map and height map with a depth map. As evident from Table 1, direct replacement leads to a substantial decline in performance. When the supervised loss supplied by lidar is introduced, the accuracy of the depth-based method is somewhat improved. However, it still lags behind that of the height-based method. In the actual implementation steps, this requires additional expense as lidar is needed as a sensor to provide depth signals. A plausible explanation is that compared to height, the depth-based approach is more intricate and therefore requires a stronger supervisory signal.

B. Additional Qualitative Results

B.1. Visualization Results under Different Environmental Conditions

In Figures 1, 2 and 3, we present more qualitative enhancements of our model compared to the baseline model on the nuScenes `val` set. The predictions on different environmental conditions are presented. From these examples, one can see that, our method can describe map elements accurately, even at the challenging scenarios, such as night and complex intersections.

B.2. Failure Cases

In Figure 4, we show the limitations of HeightMapNet* in certain challenging scenarios. Specifically, in rainy traffic intersections, although HeightMapNet* effectively identifies basic road elements, it falls short in accurately delineating distant boundaries. The detected features show noticeable discrepancies compared to actual conditions, failing to precisely demarcate the left side boundary of the road at a distance. Additionally, in nighttime traffic intersection scenarios, the model's predictions of curves at the rear of the vehicle also exhibit deviations from reality. This inaccuracy is likely due to low-light conditions and the complex interplay of environmental factors, leading to missed detections of certain dividers and boundaries. These findings highlight potential areas for improvement in the model's ability to perceive under adverse weather and low-light conditions.

*Corresponding author

C. Impact Statements

Our research focuses on developing the HeightMapNet algorithm, which significantly enhances the navigation accuracy and environmental adaptability of autonomous driving systems. As a result, it can provide more reliable map information, which in turn improves road safety and transportation efficiency. However, it is crucial to recognize the potential risks associated with this method. First, if the accuracy of the maps and their real-time updates are insufficient, it could mislead autonomous driving systems and increase the risk of traffic accidents. Additionally, reliance on this method could diminish drivers' skills over time, potentially affecting their ability to handle emergency situations.

Second, our method could be exploited by malicious actors, for example, by manipulating map data to mislead or control autonomous vehicles, posing a threat to public safety. Therefore, we advocate for stringent data verification and security measures in the development and deployment of such technologies to ensure the robustness and resilience against attacks.

Last, as autonomous driving technologies evolve, the associated regulatory and ethical issues also need thorough discussion and resolution to protect users' privacy and data security and ensure the fair and transparent use of technology. We encourage ongoing public and professional dialogues to assess and mitigate these potential societal and ethical concerns.

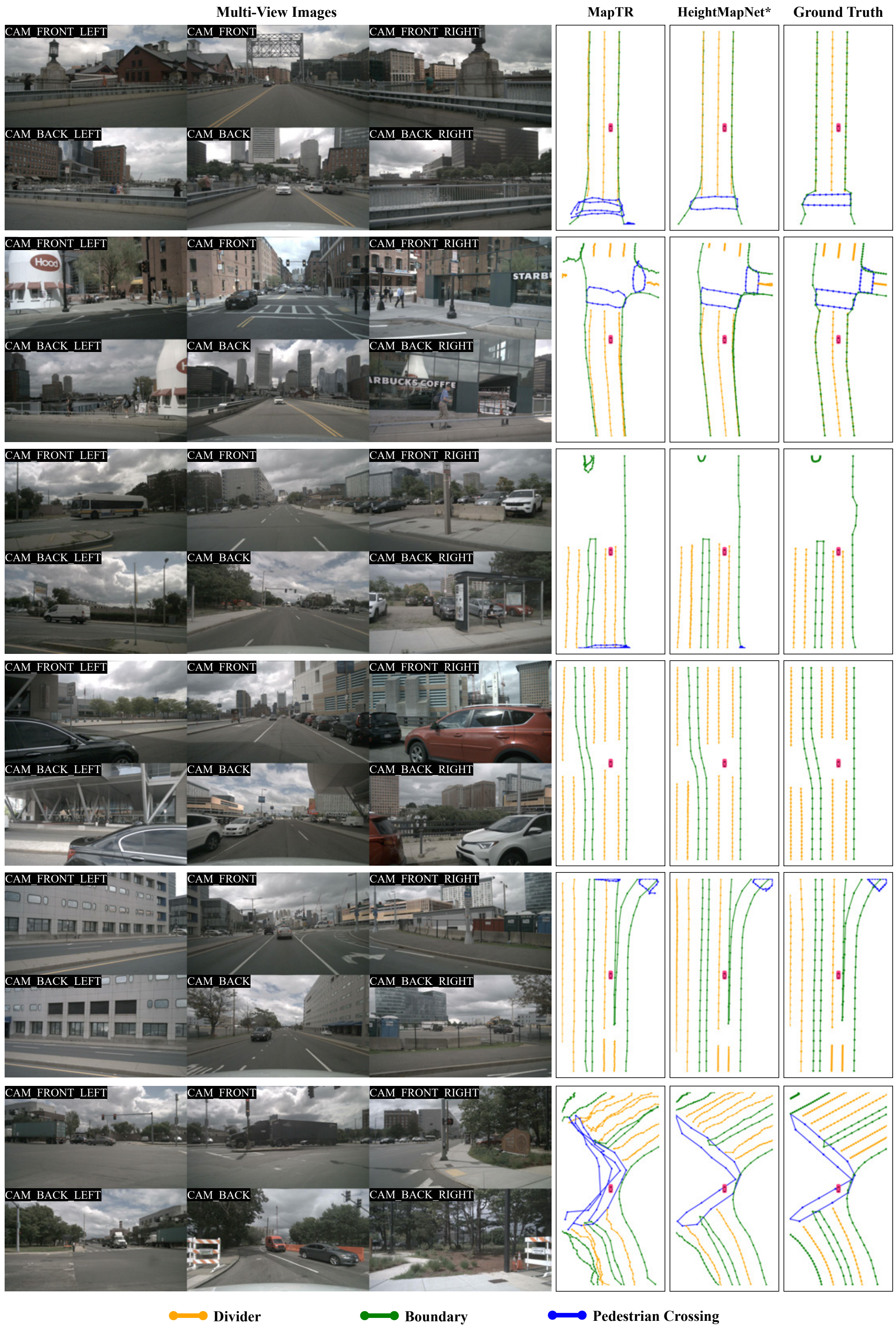

Figure 1. The visualization results of predicted map elements on the nuScenes `val` set under normal weather. Our method has better map shapes than the baseline MapTR, and is robust to various driving scenes. Even at complex intersections, the map near the vehicle closely matches the ground truth.

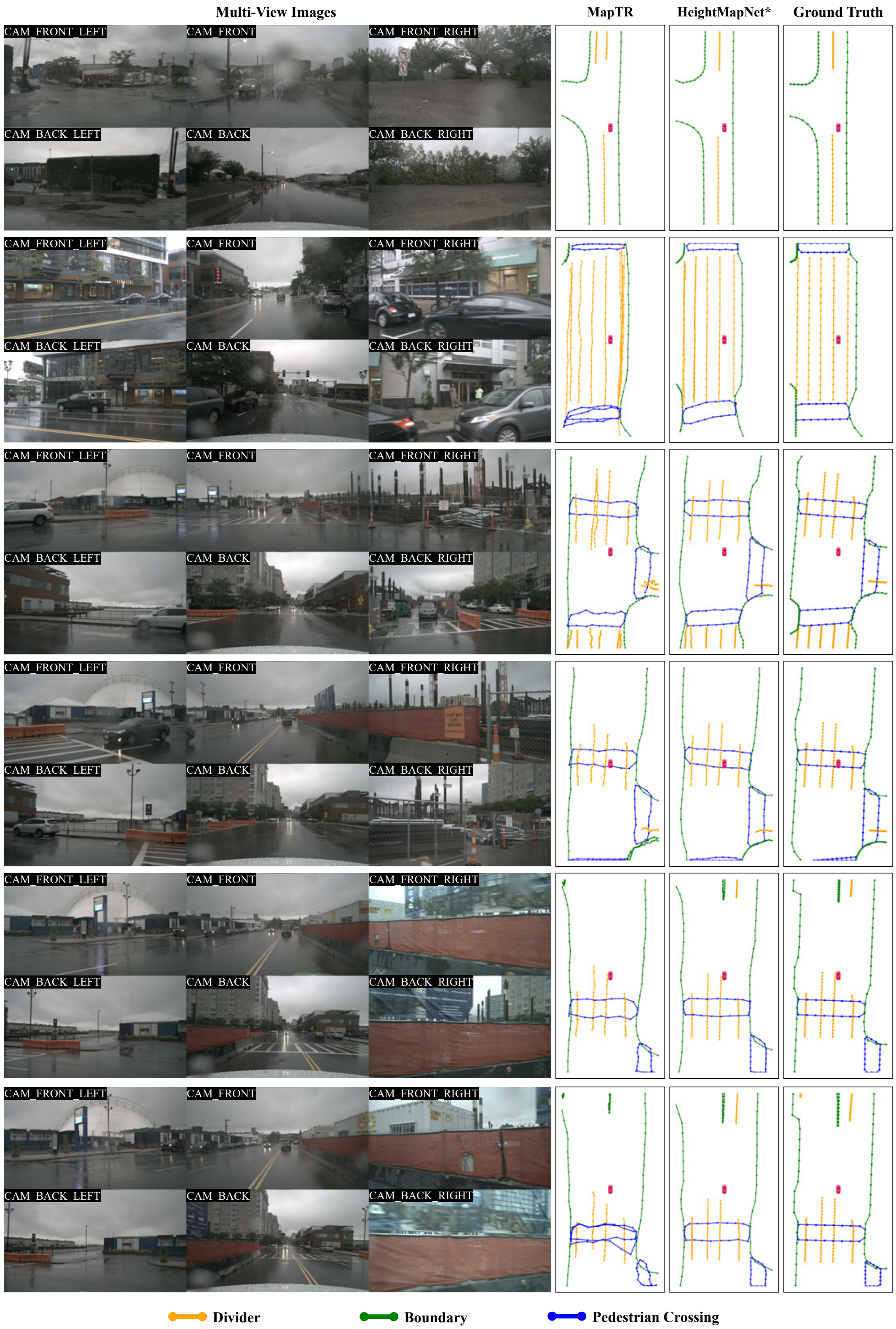

Figure 2. The visualization results of predicted map elements on the nuScenes `val` set under rainy weather.

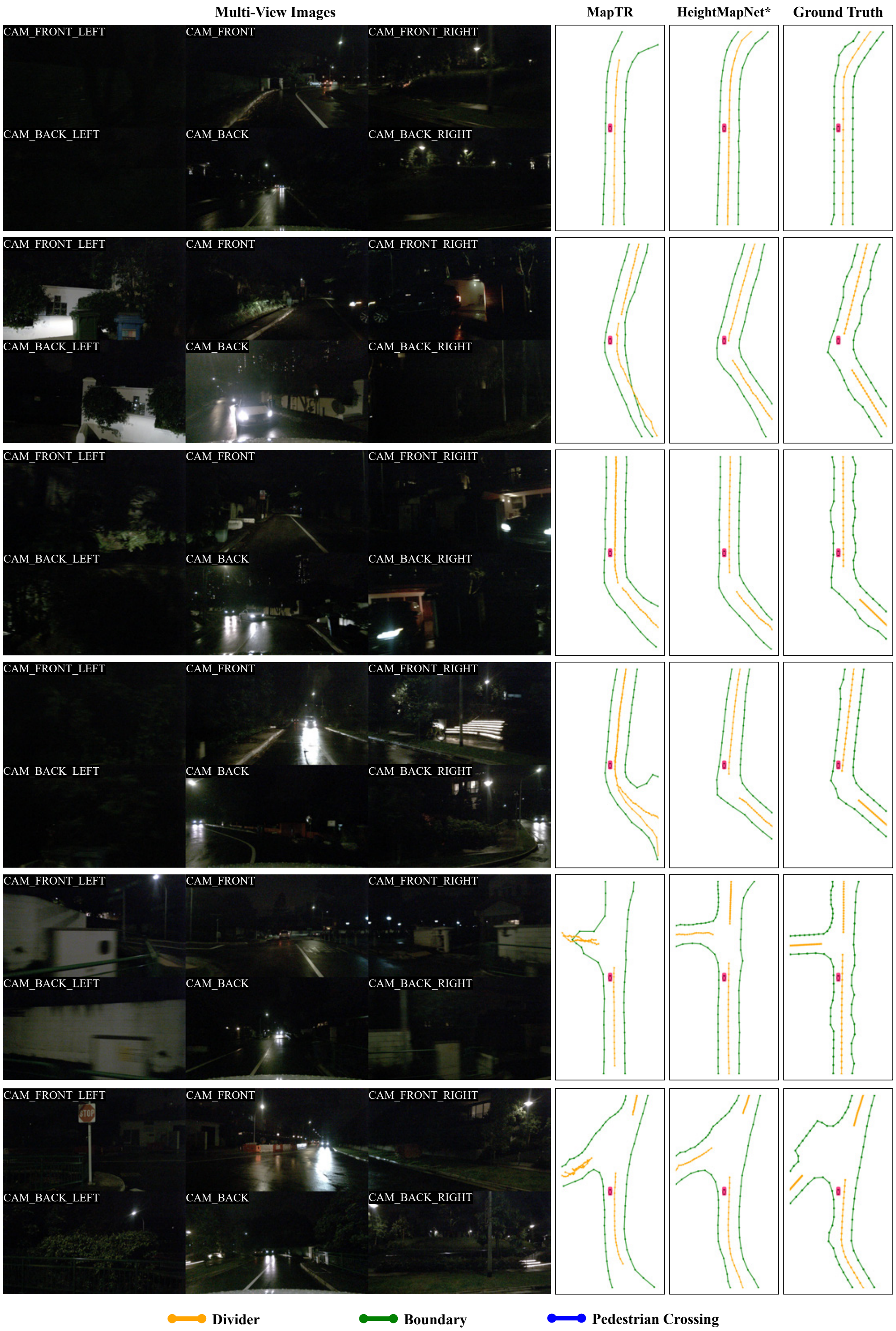

Figure 3. The visualization results of predicted map elements on the nuScenes `val` set at night.

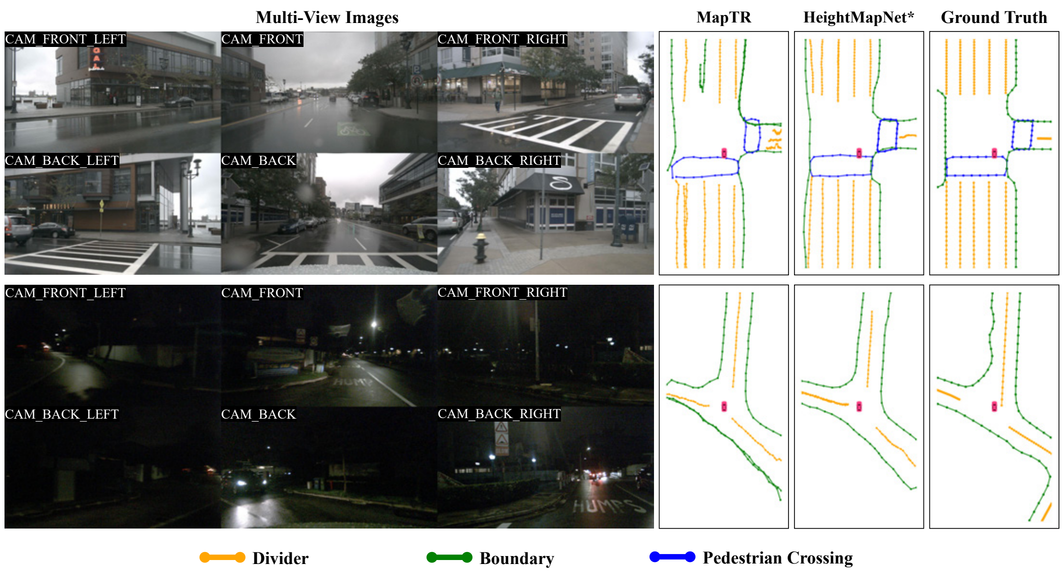

Figure 4. Visualization of '*Failure Cases*' on the nuScenes `val` set.