Multimodal Trajectory Predictions for Autonomous Driving without a Detailed Prior Map (*Supplementary Material*)

In this supplementary material, we provide an additional ablation study on the performance gap between different input maps (an OGM and a detaild prior map).

1. Ablation studies on different input maps

We investigate how the proposed method performs differently compared to those with a detailed prior map input. Our motivation is to use only perceptual data (OGM), but the proposed method can also use the prior map as input. Therefore, we compared the prediction performance when changing the input map and the presence / absence of our discriminator $\mathcal{D}$ as shown in Table 5. For example, method 3 used the prior map as input and was trained using $\mathcal{D}$. Method 6 used a map that concatenated the OGM and the prior map in channel and was trained without $\mathcal{D}$.

	OGM	Prior map	$\mathcal{D}$	mFDE [m] ↓	mADE [m] ↓	PoP [%] ↑
Method 1 (our setting)	✓		✓	2.85	1.49	77.2
Method 2	✓			2.90	1.51	74.3
Method 3		✓	✓	2.67	1.34	80.3
Method 4		✓		2.67	1.33	80.5
Method 5	✓	✓	✓	**2.64**	**1.32**	**81.6**
Method 6	✓	✓		2.65	**1.32**	81.1

Table 5. An ablation study on different input maps. We prepare three configure parameters (OGM, prior map, and $\mathcal{D}$) and compare six methods, which are the combination of those parameters. The tick of OGM indicates that an OGM is used for an input image. The tick of prior map indicates that a detailed prior map including lane information as shown in Fig. 6 is used for an input image. The tick of $\mathcal{D}$ indicates that the model is trained using our proposed discriminator $\mathcal{D}$. As with table 1, we measured three metrics (mFDE, mADE, and PoP) on our dataset with $K = 5$.

Table 5 shows that the performance of method 5, which use the prior map and the OGM as input and are trained using $\mathcal{D}$, is the highest in those of other settings. However, we can see that the performance gaps between the methods using the prior map (method 3, 4, 5, and 6) are small. Prior maps have sufficient information to understand road rules, and we think that other factors such as social interaction are necessary to improve the prediction performance further. Comparing the methods using only OGM (method 1 and 2), we can see that the introduction of $\mathcal{D}$ improves PoP by 3 %. From these results, it can be inferred that our method works more effectively in the absence of prior maps.