End-to-End Trajectory Distribution Prediction Based on Occupancy Grid Maps (Supplementary Material)

1. Occupancy Grid Maps

In this section, we show examples of occupancy grid maps (OGMs) prediction results on the two datasets: Stanford Drone Dataset (SDD) and Intersection Drone Dataset (inD). Our temporal OGMs represent each future position distribution with an OGM, so we show the predicted OGMs from our model at 6 time steps. Fig. 1 and Fig. 2 show that our temporal OGMs can capture the scene compliant future position distribution while reflecting the multimodality.

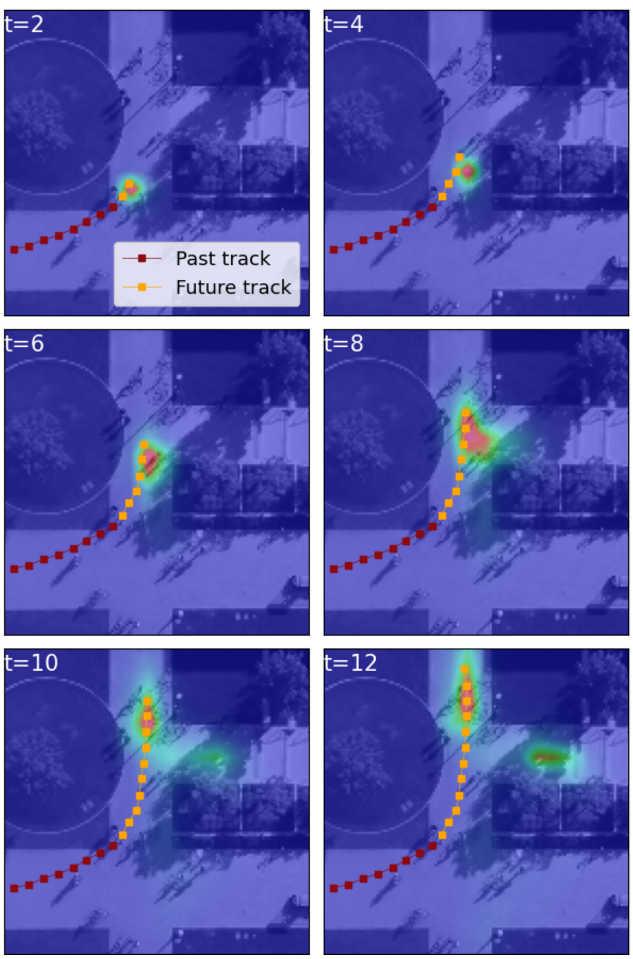

Figure 1. OGMs prediction results on the SDD. Warmer color indicates higher occupancy probability, while colder color represents lower occupancy probability.

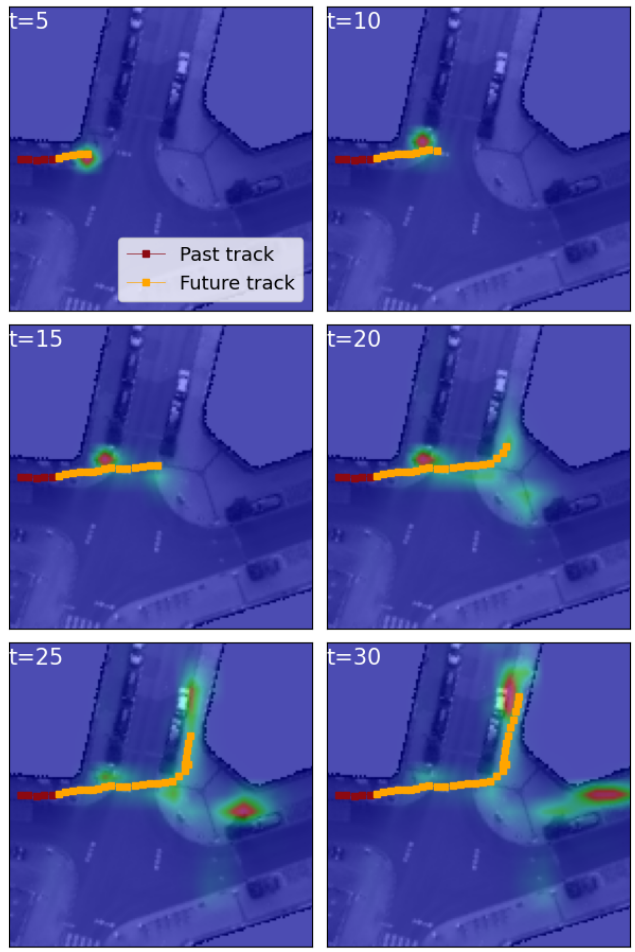

Figure 2. OGMs prediction results on the inD.

2. Reward Maps

Our model learn non-stationary rewards dependent on the history trajectories and neighboring scene for producing non-stationary policy. Our non-stationary rewards $\mathbf{r}^n : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ are dependent on the action and state, where the action set includes 4 adjacent movements $up, down, left, right$ and an end action leading to the absorbing state. For simplicity, we show the reward maps of taking four moving actions at the fixed MDP steps $n = 10$ on the SDD (see Fig. 3) while the end action at four different MDP steps $n = 5, 10, 15, 20$ on the inD (see Fig. 4). As shown in Fig. 3, the regions with the highest rewards (*i.e.*, in red color) are generally consistent with the action direction. Fig. 4 illustrates that the roadside or crosswalk regions have higher end rewards than the car lane and impassable regions, which is reasonable since the used training data of inD are pedestrian trajectories instead of vehicle tracks.

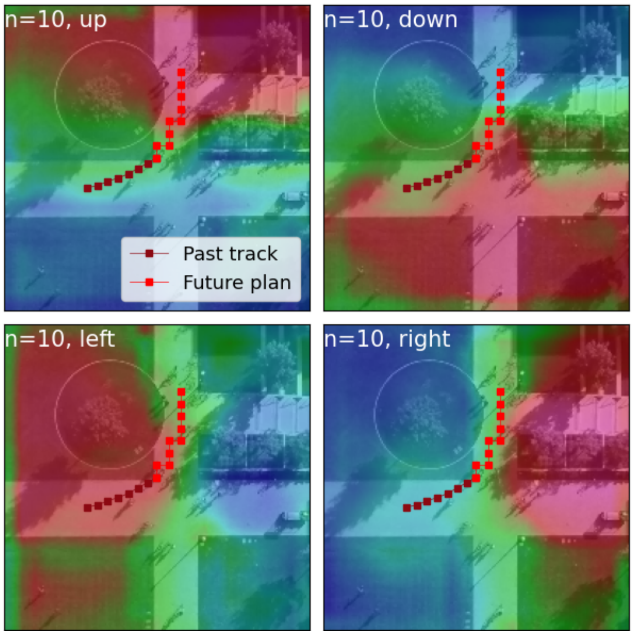

Figure 3. Reward maps of taking the 4 adjacent movements $up, down, left, right$ at MDP step $n = 10$ on the SDD. Warmer color represents higher reward value, while colder color implies lower reward value.

3. Policy Maps

We generate a non-stationary policy using an approximate value iteration network based on the non-stationary rewards. Similarly, we show the policy maps of taking four moving actions at the fixed MDP steps $n = 10$ on the SDD (see Fig. 5) while the end action at four different MDP steps $n = 5, 10, 15, 20$ on the inD (see Fig. 6).

Fig. 5 shows that our model learns to take an action to move on the road consistent with the action direction and avoid the terrains. Fig. 6 demonstrates that it learns to end at the roadsides in the last few MDP steps, or the impassable regions, where the pedestrian trajectories seen in the training data usually end in 30 seconds future.

3.1. Plan-conditioned Trajectory

The grid-based plans are sampled from the non-stationary policy by Gumbel-Softmax trick and bilinearly interpolation. Then, we recursively generate a trajectory using an RNN with multi-head attention on the sampled plan.

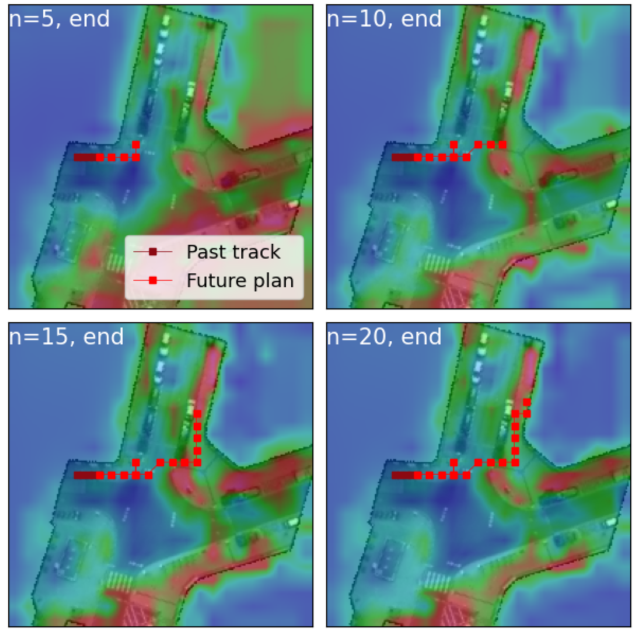

Figure 4. Reward maps of taking the *end* action on the inD.

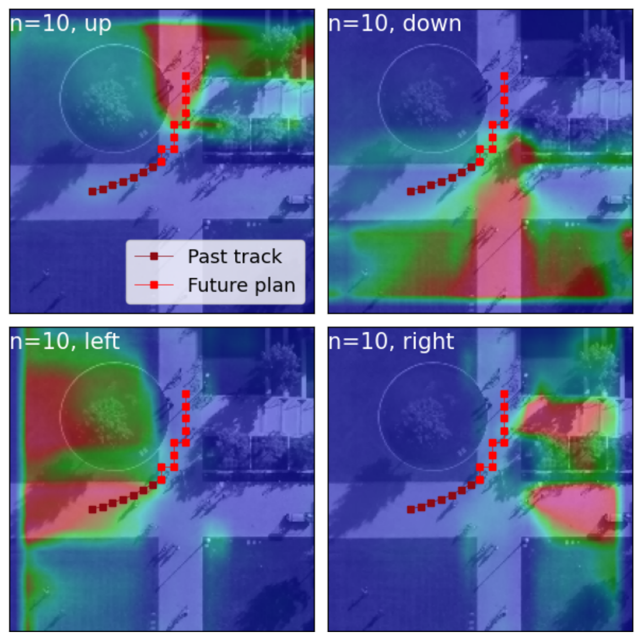

Figure 5. Policy maps of taking four movements $up, down, left, right$ at MDP step $n = 10$ on the SDD. Warmer color means higher action probability, while colder color means lower action probability.

We show four examples of an sampled plan and its corresponding trajectory on SDD in Fig. 7 and inD in Fig. 8. We can observe that the sampled plans are scene compliant and the trajectory follows its corresponding plan while keeping smooth.

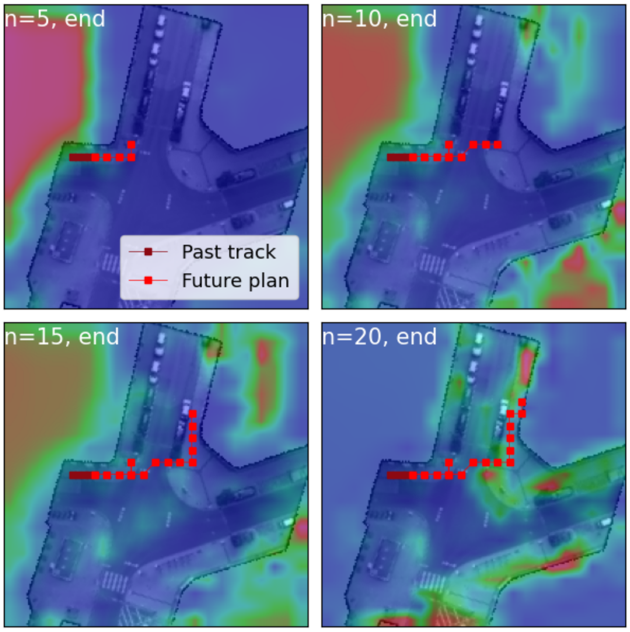

Figure 6. Policy maps of taking the *end* action on the inD.

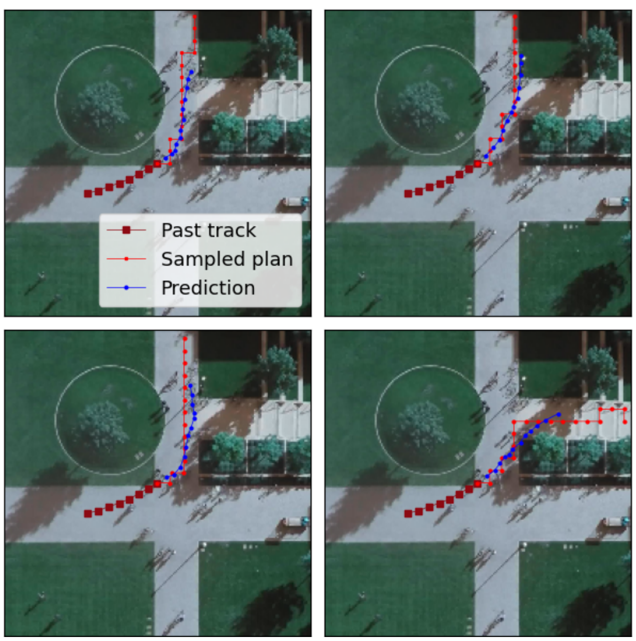

Figure 7. Four grid-based plans sampled from the non-stationary policy and the generated trajectories conditioned on the plans on the SDD.

4. Trajectory Distribution

By minimizing the approximate symmetric cross-entropy loss, we learn a trajectory distribution close to the ground-truth. Examples of the predictive trajectory distribution from SDD and inD are shown in Fig. 9 and Fig. 10. We can see that our predicted trajectory distributions are diverse and feasible.

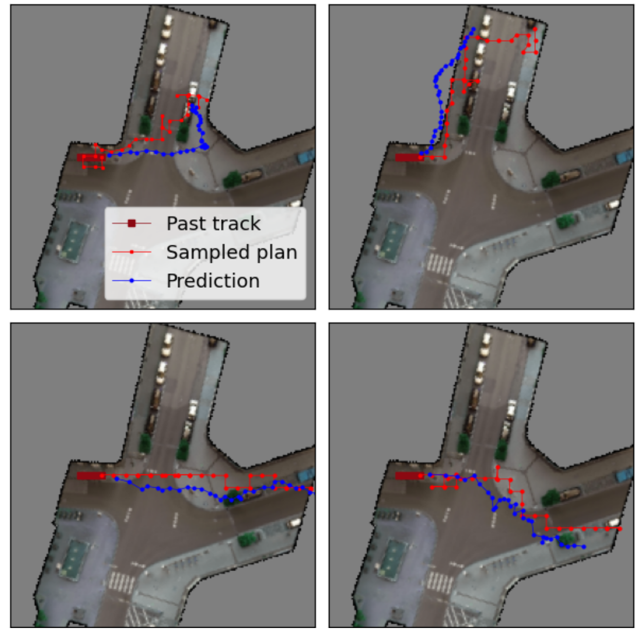

Figure 8. Four grid-based plans and corresponding trajectories on the inD.

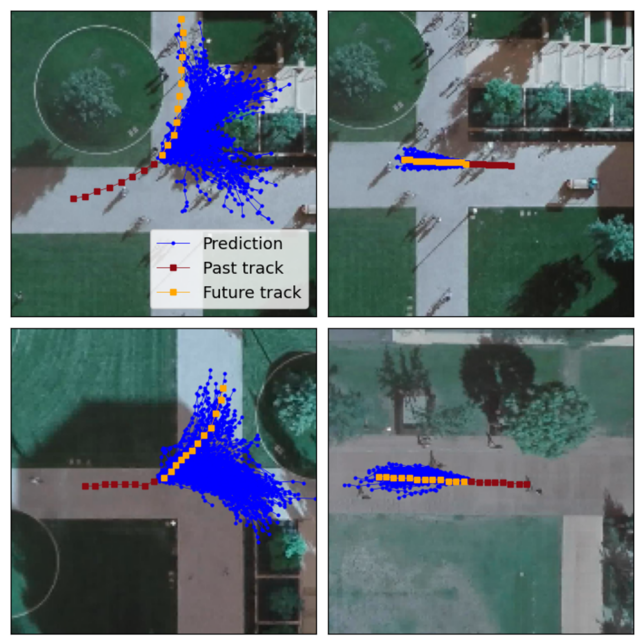

Figure 9. Predicted trajectory distributions on the SDD.

5. Representative Trajectories

We design a Transformer-based refinement network to generate a small set of representative trajectories based on a large number of trajectories sampled from the trajectory distribution. Fig. 11 and Fig. 12 show qualitative examples from SDD and inD. We note that the representative trajectories are diverse and capable to cover the ground-truth future trajectory due to the adopted variety loss.

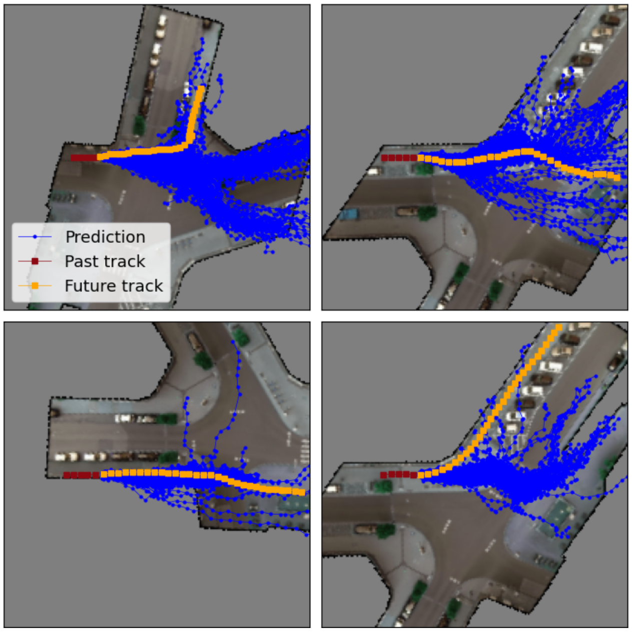

Figure 10. Predicted trajectory distributions on the inD.

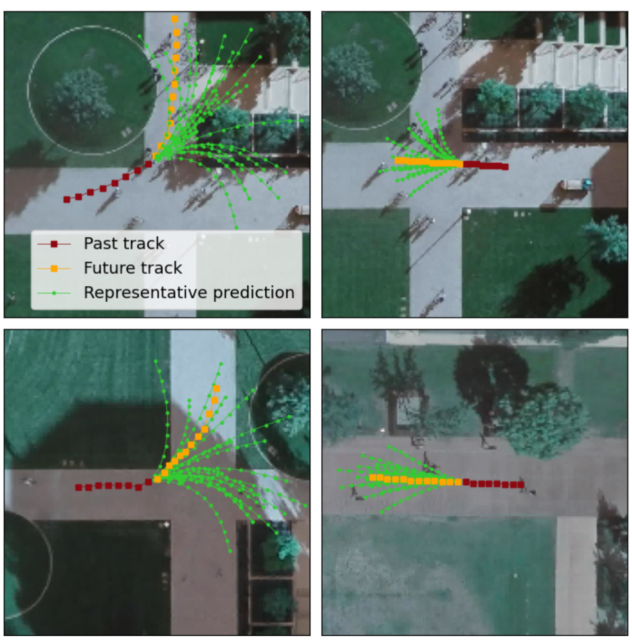

Figure 11. Representative trajectories predicted on the SDD.

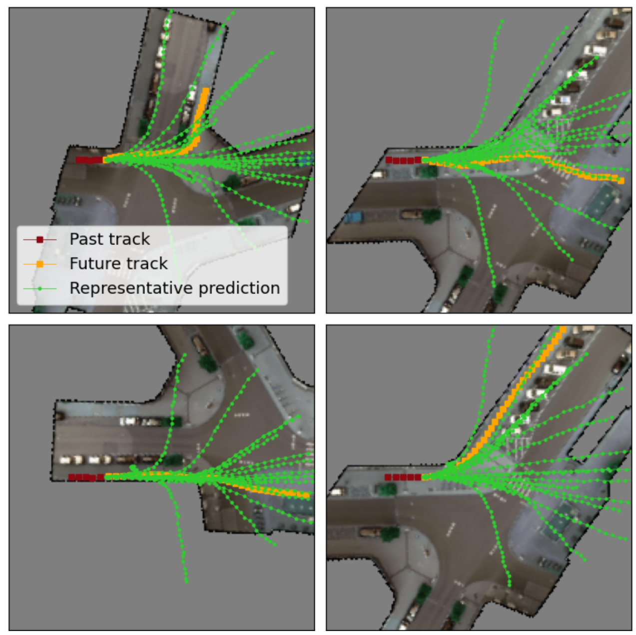

Figure 12. Representative trajectories predicted on the inD.