Supplemental Material: Where are you heading? Dynamic Trajectory Prediction with Expert Goal Examples

He Zhao
York University
zhufl@eecs.yorku.ca

Richard P. Wildes
York University
wildes@cse.yorku.ca

Abstract

This document provides additional material that is supplemental to our main submission. In particular, we provide the experimental details for the feasibility evaluation on the synthetic Hyang4-SDD scene, shown in Section 3.4 of the main submission.

1. Hyang4 scene synthetic dataset

The Hyang4 scene synthetic dataset (Hyang4 dataset) is proposed by the authors of Goal-GAN [1] to specifically evaluate the multi-modality and feasibility of their trajectory predictions. Two essential steps are followed to achieve this end: First, the authors manually segment the road boundaries, where the walking lanes and stairs are precisely denoted as accessible areas; second, they synthesize a set of new trajectories automatically using the social force model [2], such that they all reside within the feasible road areas, *i.e.* white region in Figure 1. In total, there are 1300 simulated trajectories and we follow the data split proposed by the authors [1], namely 800 trajectories for training, 200 for validation and 300 for testing.

To evaluate the feasibility, two extra metrics are suggested: mode-coverage (MC) that measures the portion of goal predictions that have distance to ground-truth goals up to 2 meters in world coordinates; the feasibility metric, F, that denotes the ratio of trajectories lying inside the feasible area. The transformation matrix between pixel and world coordinates for this scene is provided by [3]. Formally, the MC evaluation metric is defined according to

$$\mathbf{C}(\cdot) = \begin{cases} 1, & ||\mathrm{W}_{\text{pixel-to-world}}(\hat{\mathbf{Y}}_i^{t_{end}} - \mathbf{Y}_i^{t_{end}})||_2 \leqslant 2 \\ 0, & \text{else} \end{cases} \tag{1}$$

and

$$\mathrm{MC} = \frac{\sum_{i=1}^{N} \mathbf{C}(\hat{\mathbf{Y}}_i^{t_{end}}, \mathbf{Y}_i^{t_{end}})}{N}, \tag{2}$$

where $\hat{\mathbf{Y}}_i^{t_{end}}$ and $\mathbf{Y}_i^{t_{end}}$ are estimated and groundtruth trajectory end points, resp., $\mathrm{W}_{\text{pixel-to-world}}$ denotes the image-to-world coordinate transformation matrix, $\mathbf{C}(\cdot)$ is the mode-coverage calculating function and MC is the final score considering all N testing data. The feasibility score, $\mathbf{F}$, is computed via

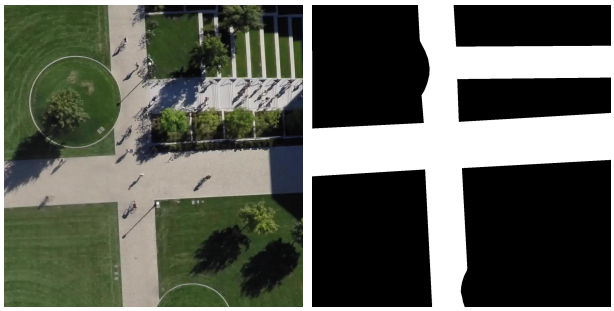

Figure 1: Hyang4 scene image and its road segmenation provided by [1]. The white region is accessible to pedestrians; the dark region is prohibited.

$$\mathbf{H}(\cdot) = \begin{cases} 1, & \hat{\mathbf{Y}}_i \notin \mathbf{S}_{\text{boundary}} \\ 0, & \text{else} \end{cases} \tag{3}$$

and

$$\mathrm{F} = \frac{\sum_{i=1}^{N} \mathbf{H}(\hat{\mathbf{Y}}_i, \mathbf{S}_{\text{boundary}})}{N}, \tag{4}$$

where $\mathbf{S}_{\text{boundary}}$ is the road contour and $\mathbf{H}(\cdot)$ is the function that decides if predicted trajectories traverse boundaries.

Results of our proposed approach to trajectory prediction on this evaluation are in Sec. 3.4 of the main submission.

References

[1] Patrick Dendorfer, Aljosa Osep, and Laura Leal-Taixé. Goal-GAN: Multimodal trajectory prediction based on goal position estimation. In *Proc. ACCV*, 2020. 1

[2] Dirk Helbing and Peter Molnar. Social force model for pedestrian dynamics. *Physical review E*, 51(5):4282, 1995. 1

[3] Amir Sadeghian, Vineet Kosaraju, Ali Sadeghian, Noriaki Hirose, Hamid Rezatofighi, and Silvio Savarese. Sophie: An attentive GAN for predicting paths compliant to social and physical constraints. In *Proc. CVPR*, 2019. 1