Report of Social-Gan on the Interaction dataset and Argoverse dataset

Xiaogang Jia
Mechanical Systems Control Lab

August 7, 2020

1 Motivation

The original Social-Gan is proposed to complete the prediction for human motion and its key method, Pooling Module, computes relative positions between the ego one and all other people. However, this pattern may not work well in the motion prediction for autonomous driving, because for an ego vehicle, not all other vehicles contribute to its motion planning. Besides, the Pooling Module in Social-Gan cannot capture the spatial information which matters a lot for the planning of vehicles. In order to verify the influence of the Pooling Module, we have conducted several experiments on the Interaction dataset and Argoverse dataset.

2 Experiment

Interaction dataset holds two test sets: Regular track and Generalizability track. Data from the same situations of training set is used in the Regular track, while data from different situations is tested in the Generalizability track. Here we have trained three models of Social-Gan to better verify the pooling mechanism. The method of sgan-ori is the original Social-Gan with only changing the hyper parameters. To test the influence of data augmentation, we apply random rotation and sequence reverse to train a new model, sgan-aug. Finally, we simply remove the Pooling Module and only use the history trajectories of vehicles to encode. All the hyper parameters keep constant in three models and the results are reported on Table 1 and Table 2.

According to the results on the challenge dataset, data augmentation brings less improvement on the Regular track, but boosts the model a lot on the Generalizability track. For the Pooling Module, surprisingly, the model performs better without the pooling information. Therefore, the global social interaction may harm the performance of vehicle trajectory prediction.

Table 1: Results on the Regular Track

Method	ADE	FDE	MON	Note
sgan-ori	0.552	1.620	0.228	Original
sgan-aug	0.529	1.651	0.215	Data-aug
sgan-nopool	0.437	1.444	0.176	No pooling

Table 2: Results on the Generalizability Track

Method	ADE	FDE	MON	Note
sgan-ori	2.872	6.195	2.498	Original
sgan-aug	1.095	2.980	0.594	Data-aug
sgan-nopool	0.996	2.965	0.474	No pooling

In order to find an effective pooling method with social features, we test two additional ideas. First, considering that local features are more important for the ego vehicle, we define an effective area(in practice, size of 40×40) around the predicted vehicle and then utilizes the angles between the velocity vector and the distance vector to confine the field of view of the predicted vehicle. Cars which are out of the view of the predicted vehicle are removed in the pooling module. Besides, we also test the influence of the lane. We generate the surrounding lanes of vehicles and use a simple CNN to process the images. Results are reported on the Table 3.

From Table 3, the new pooling method sgan-new-pooling and the method that combines map of lanes sgan-map produce almost the same results of the sgan-nopool.

Table 3: Results on the Validation dataset

Method	ADE	FDE	Note
sgan-nopool	0.34	1.05	No pooling
sgan-new-pooling	0.34	1.03	New pooling
sgan-map	0.36	1.10	Combine map of lane

To further verify the Pooling Module, we test it on the Argoverse dataset. Due to the huge computational costs of feature extraction, we simply utilize the social features provided by the authority. For one vehicle, its social features contain three parts: minimum distance of the tracks in front, minimum distance of the tracks in back, number of the surrounding vehicles. Here, we first train two models: Social-Gan without Pooling Module(sgan-nopool) and Social-Gan pooled with official social features(sgan-social). To finish the training process in a short time, we only predict once(k=1) for each track during the training. Finally, we train another model using 10 predictions during the training process. The Argoverse leaderboard allows each model to predict 6 times at most. Results are reported on Table 4.

From Table 4, we find that pooling with social features does not improve the performance, which can further prove that this pooling mechanism contribute less to the trajectory prediction of vehicles.

Table 4: Results on the Argoverse Challenge

Method	ADE(k=1)	FDE(k=1)	ADE(k=6)	FDE(k=6)
sgan-nopool(k=10)	3.3765	8.2234	1.4896	3.1058
sgan-nopool(k=1)	2.1933	5.0622	2.1627	4.9897
sgan-social(k=1)	2.1956	5.0665	2.1671	4.9983

3 Summary

To verify the effectiveness of the pooling module in Social-Gan, we compare the results between no-pooling model and pooling model with social features on both Interaction dataset and Argoverse dataset. From the experimental results, we can draw the following conclusions:

- Pooling with social features like relative distance between vehicles, numbers of surrounding vehicles etc. brings less benefits for vehicle trajectory prediction. Since pooling with fully-connected layers can hardly capture the spatial information in the environment, it would be less surprised to have such results.
- Prediction with effective spatial information could be more beneficial. Whether combining the bird view pictures or utilizing the social convolution makes it more reasonable for the model to learn the spatial information, thereby improving the performance.