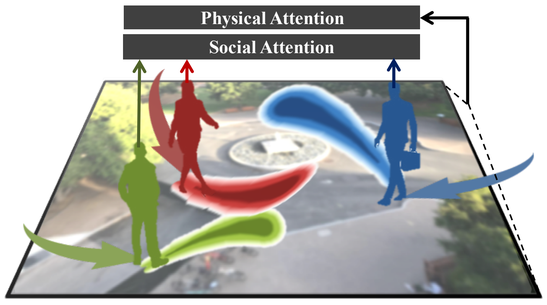
SoPhie: An Attentive GAN for Predicting Paths Compliant to Social and Physical Constraints

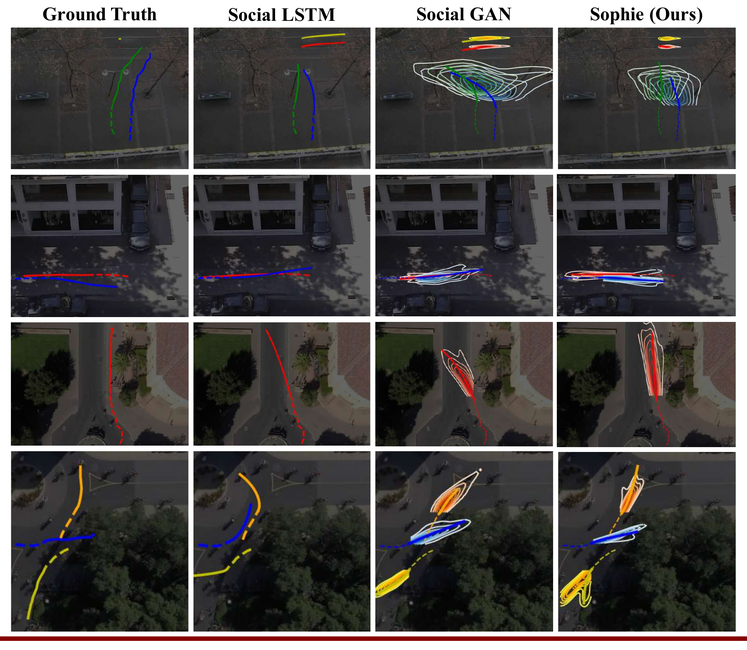
Amir Sadeghian*, Vineet Kosaraju*, Ali Sadeghian, Noriaki Hirose, S. Hamid Rezatofighi, Silvio Savarese
Stanford Vision and Learning Lab (SVL)

Motivation

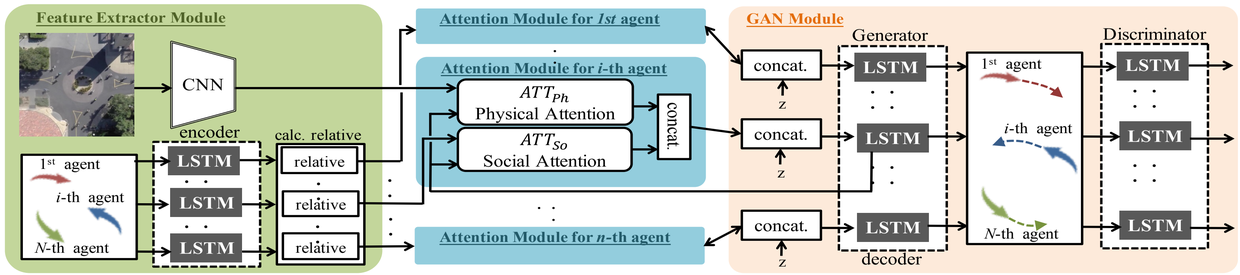
- Trajectory prediction can be formally stated as the problem of estimating the state of all agents in future, given the scene information and their past states.
- It applies to a wide range of domains from autonomous driving vehicles and social robot navigation, to abnormal behavior detection in surveillance
- Our method can **predict the trajectories** of all agent by:
  1. Obeying **physical constraints** of the environment
  2. Anticipating movements and **social behavior** of other people.
  3. Finding **distribution of** feasible paths
- Sophie allows us to **visualize and interpret** the fine-grained semantic elements of navigation scenes that influence the prediction of trajectories.



Qualitative results



Overall model



Our model consists of three key components including:

1. A **feature extractor** module
2. An **attention** module, and
3. An LSTM based **GAN** module

Quantitative results on benchmark datasets

	Baselines					SoPhie (Ours)				
Dataset	**Lin**	**LSTM**	**S-LSTM**	**S-GAN**	**S-GAN-P**	$\mathbf{T_A}$	$\mathbf{T_O + I_O}$	$\mathbf{T_O + I_A}$	$\mathbf{T_A + I_O}$	$\mathbf{T_A + I_A}$
ETH	1.33 / 2.94	1.09 / 2.41	1.09 / 2.35	0.81 / 1.52	0.87 / 1.62	0.90 / 1.60	0.86 / 1.65	0.71 / 1.47	0.76 / 1.54	**0.70 / 1.43**
HOTEL	**0.39 / 0.72**	0.86 / 1.91	0.79 / 1.76	0.72 / 1.61	0.67 / 1.37	0.87 / 1.82	0.84 / 1.80	0.80 / 1.78	0.83 / 1.79	0.76 / 1.67
UNIV	0.82 / 1.59	0.61 / 1.31	0.67 / 1.40	0.60 / 1.26	0.76 / 1.52	**0.49 / 1.19**	0.58 / 1.27	0.55 / 1.23	0.55 / 1.25	0.54 / 1.24
ZARA1	0.62 / 1.21	0.41 / 0.88	0.47 / 1.00	0.34 / 0.69	0.35 / 0.68	0.38 / 0.72	0.34 / 0.68	0.35 / 0.67	0.32 / 0.64	**0.30 / 0.63**
ZARA2	0.77 / 1.48	0.52 / 1.11	0.56 / 1.17	0.42 / 0.84	0.42 / 0.84	0.38 / 0.79	0.40 / 0.82	0.43 / 0.87	0.41 / 0.80	**0.38 / 0.78**
AVG	0.79 / 1.59	0.70 / 1.52	0.72 / 1.54	0.58 / 1.18	0.61 / 1.21	0.61 / 1.22	0.61 / 1.24	0.57 / 1.20	0.58 / 1.20	**0.54 / 1.15**

	Baselines						SoPhie (Ours)				
Dataset	**Lin**	**SF**	**S-LSTM**	**S-GAN**	**CAR-Net**	**DESIRE**	$\mathbf{T_A}$	$\mathbf{T_O + I_O}$	$\mathbf{T_O + I_A}$	$\mathbf{T_A + I}$	$\mathbf{T_A + I_A}$
SDD	37.11 / 63.51	36.48 / 58.14	31.19 / 56.97	27.246 / 41.440	25.72 / 51.8	19.25 / 34.05	17.76 / 32.14	18.40 / 33.78	16.52 / 29.64	17.57 / 33.31	**16.27 / 29.38**

Analysis

Model	Complex	Simple
LSTM	31.31	30.48
CAR-Net	24.32	30.92
S-GAN	29.29	**22.24**
SoPhie	**15.61**	**21.08**

Impact of physical constraints: Performance of baselines on SDD, split into physically simple and complex scenes

	GT	LIN	S-GAN	SoPhie
ETH	0.000	3.137	2.509	**1.757**
HOTEL	0.092	**1.568**	1.752	1.936
UNIV	0.124	1.242	**0.559**	0.621
ZARA1	0.000	3.776	1.749	**1.027**
ZARA2	0.732	3.631	2.020	**1.464**
Avg	0.189	2.670	1.717	**1.361**

Impact of social constraints: Average % of colliding pedestrians per frame for each of the scenes

Nexus 6 Little 1 Huang 1

Using Sophie to generate highly accurate traversability maps.