SocialVAE: Human Trajectory Prediction using Timewise Latents

Pei Xu [1,3], Jean-Bernard Hayet [2], Ioannis Karamouzas [1] peix@clemson.edu, jbhayet@cimat.mx, ioannis@clemson.edu

[1] Clemson University, [2] CIMAT, [3] Roblox

CLEMSON UNIVERSITY

CIMAT

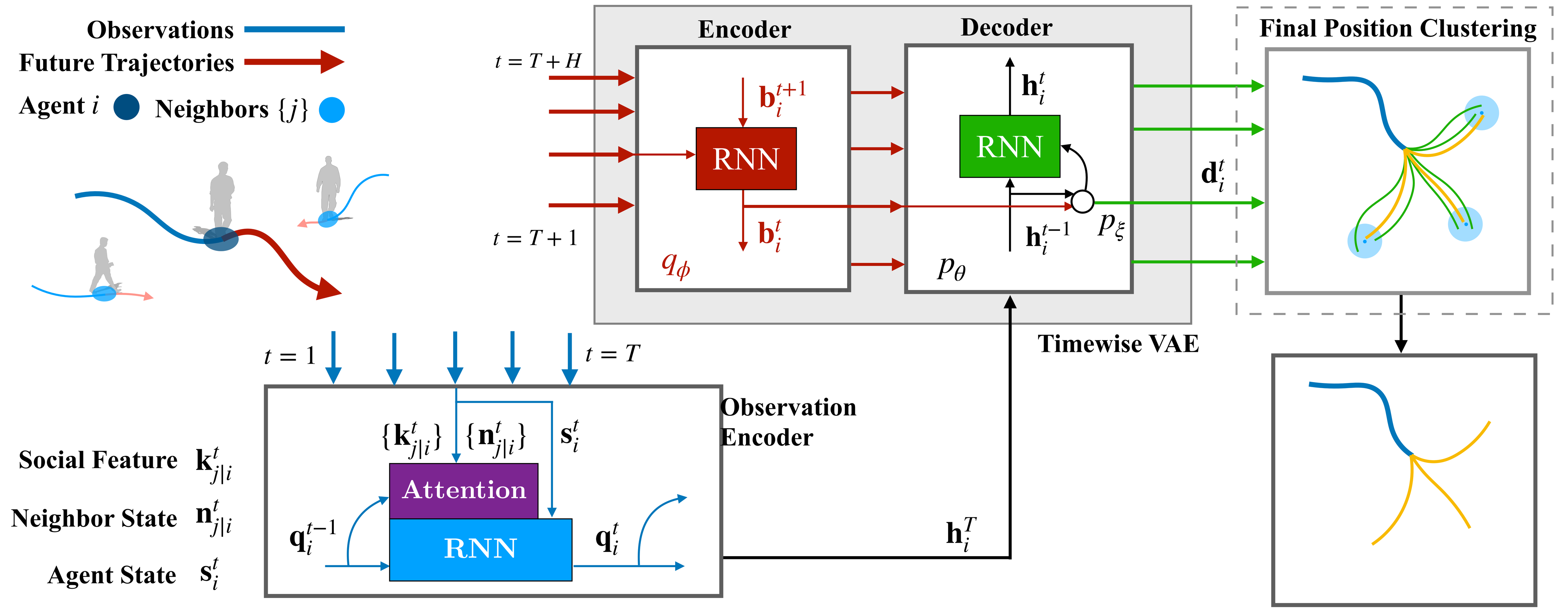
Introduction

Contribution: We propose SocialVAE, a novel approach to predict human trajectory distributions conditioned on short-term historical observations.

SocialVAE employs a ***timewise*** VAE architecture with a ***conditional prior*** and a ***backward posterior*** approximated bidirectionally from the whole trajectories, and uses a ***social attention*** mechanism to capture the influence from the neighboring agents.

We show that SocialVAE achieves state-of-the-art performance on the ETH/UCY and SDD benchmarks, and more complex scenes involving NBA players' movement and interactions.

Model Architecture



For each agent i, at each time step $t \in [T+1, T+H]$, our model is optimized to minimize the ELBO:

$$\sum_{t=T+1}^{T+H} \mathbb{E}_{\mathbf{z}_i^t \sim q_\phi(\cdot|\mathbf{b}_i^t, \mathbf{h}_i^{t-1})} \left[\log p_\xi(\mathbf{d}_i^t \,|\, \mathbf{z}_i^t, \mathbf{h}_i^{t-1})\right] - D_{KL}\left[q_\phi(\mathbf{z}_i^t \,|\, \mathbf{b}_i^t, \mathbf{h}_i^{t-1}) \,||\, p_\theta(\mathbf{z}_i^t \,|\, \mathbf{h}_i^{t-1})\right]$$

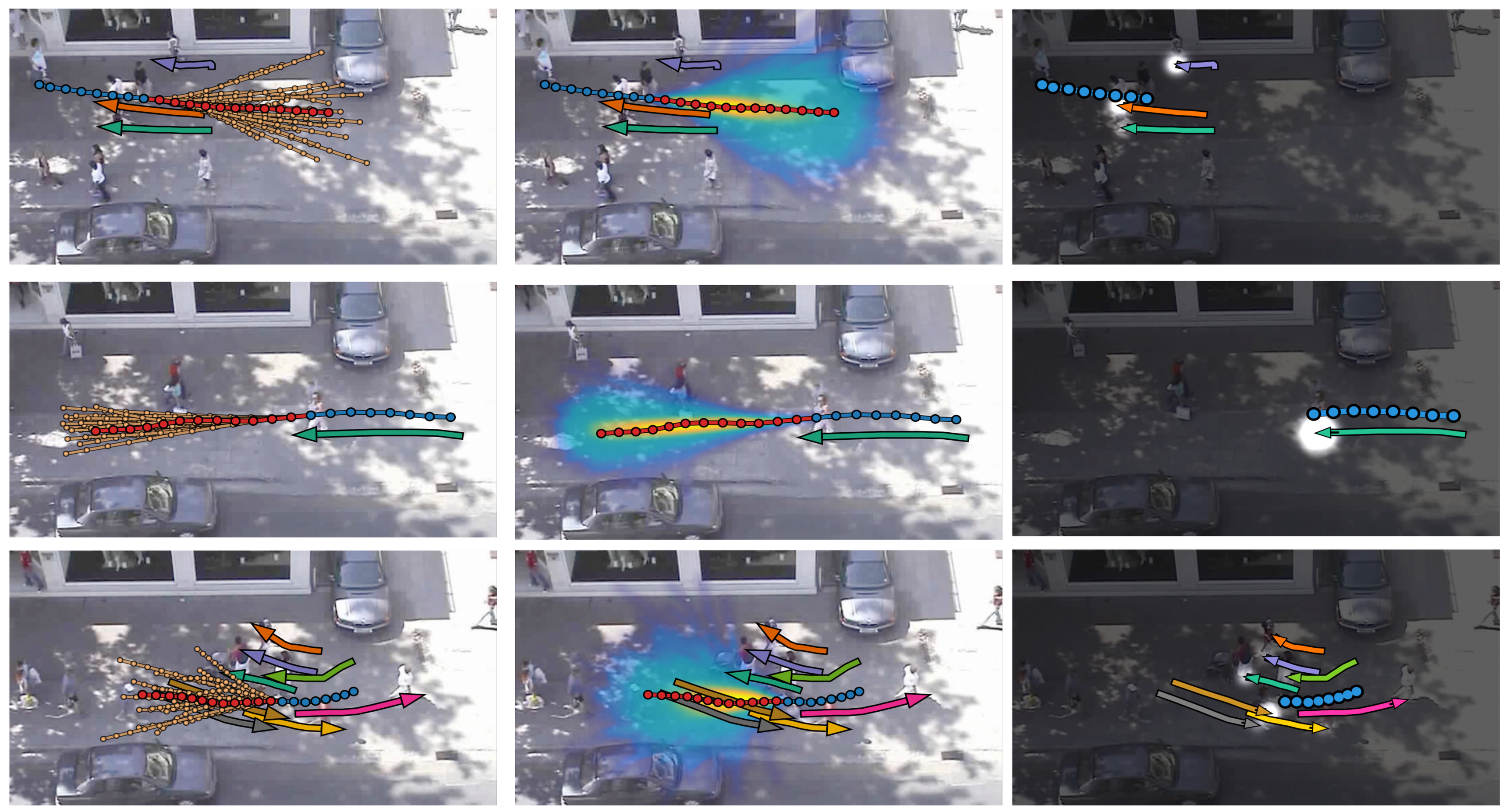
with timewise latent $\mathbf{z}_i^t$, a conditional prior $p_\theta(\mathbf{z}_i^t \,|\, \mathbf{h}_i^{t-1})$, and a backward posterior $q_\phi(\mathbf{z}_i^t \,|\, \mathbf{b}_i^t, \mathbf{h}_i^{t-1})$ where ξ, ϕ and θ are parameters of the distributions modeled by networks.

We also propose an optional post-processing approach *Final Position Clustering* (FPC) to improve the diversity of predictions.

Results

ETH/UCY Scenarios

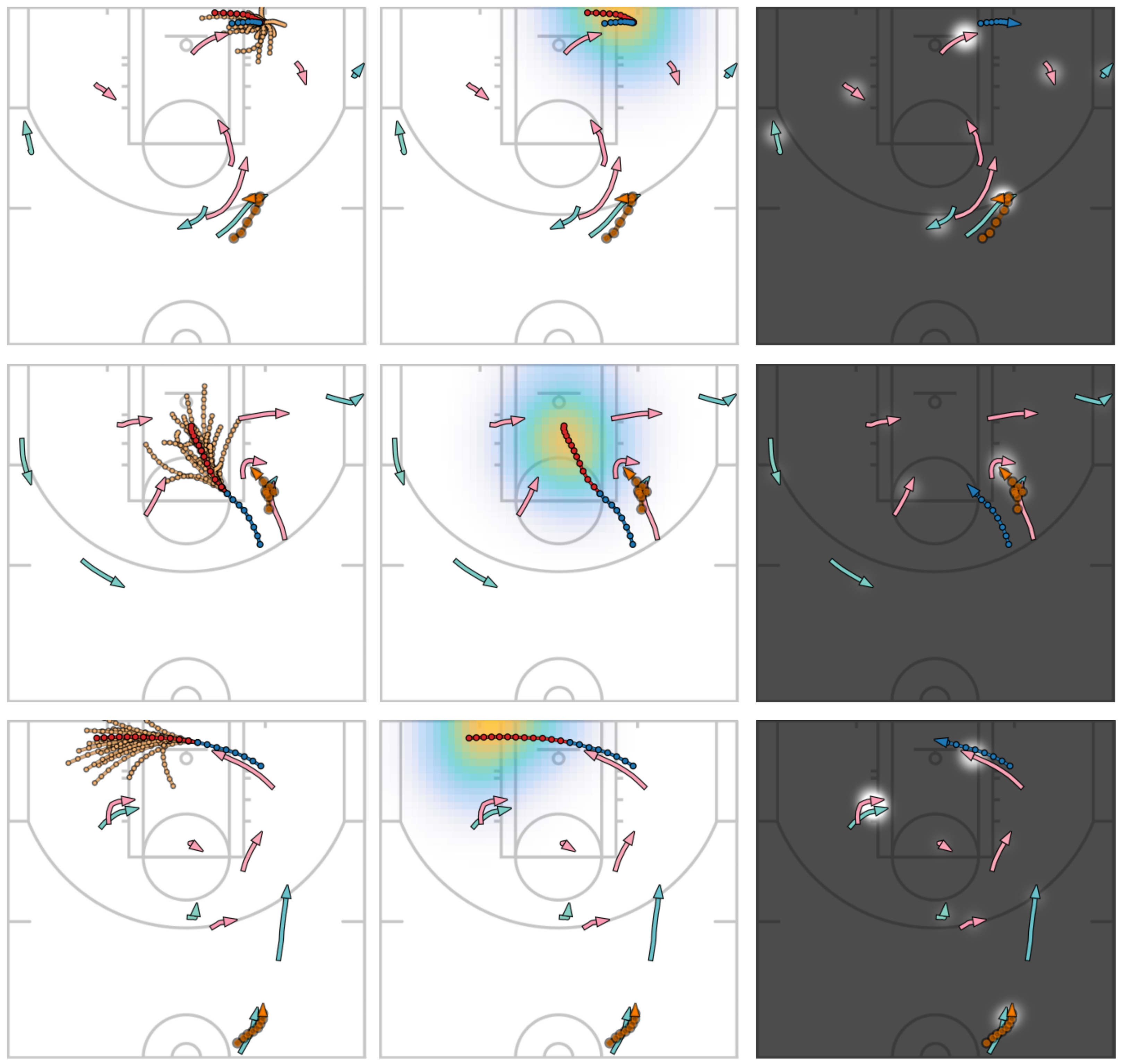
observation ground truth prediction neighbors ball



Metric: ADE/FDE*	ETH/Hotel	UCY	SDD
Constant Velocity	0.69/1.45	0.42/0.94	19.74/40.04
Trajectron++ [Salzmann et al. 2020]	0.35/0.61	0.21/0.43	10.00/17.15
BiTraP [Yao et al. 2021]	0.37/0.63	0.21/0.42	9.09/16.31
SGNet [Wang et al. 2022]	0.34/0.61	0.22/0.41	9.69/17.01
SocialVAE	0.31/0.49	0.20/0.37	8.88/14.81
SocialVAE + FPC	**0.27/0.39**	**0.17/0.29**	**8.10/11.72**

*ADE: Average Displacement Error; FDE: Final Displacement Error.

NBA Scenarios



From left to right: Predicted Trajectories, Prediction Heatmap, and Attention Map.

Metric: ADE/FDE*	Rebounding	Scoring
Constant Velocity	2.14/5.09	2.07/4.81
Trajectron++ [Salzmann et al. 2020]	0.98/1.93	0.73/1.46
BiTraP [Yao et al. 2021]	0.83/1.72	0.74/1.49
SGNet [Wang et al. 2022]	0.78/1.55	0.68/1.30
SocialVAE	0.72/1.37	0.64/1.17
SocialVAE + FPC	**0.66/1.10**	**0.58/0.95**

*ADE: Average Displacement Error; FDE: Final Displacement Error.

Conclusion

SocialVAE brings more than **10 % improvement** and in certain test cases more than **50 % improvement** over existing trajectory prediction methods.

Code Available at

https://motion-lab.github.io/SocialVAE

Acknowledgement

This work was supported by the National Science Foundation under Grant No. IIS-2047632.

ECCV TEL AVIV 2022