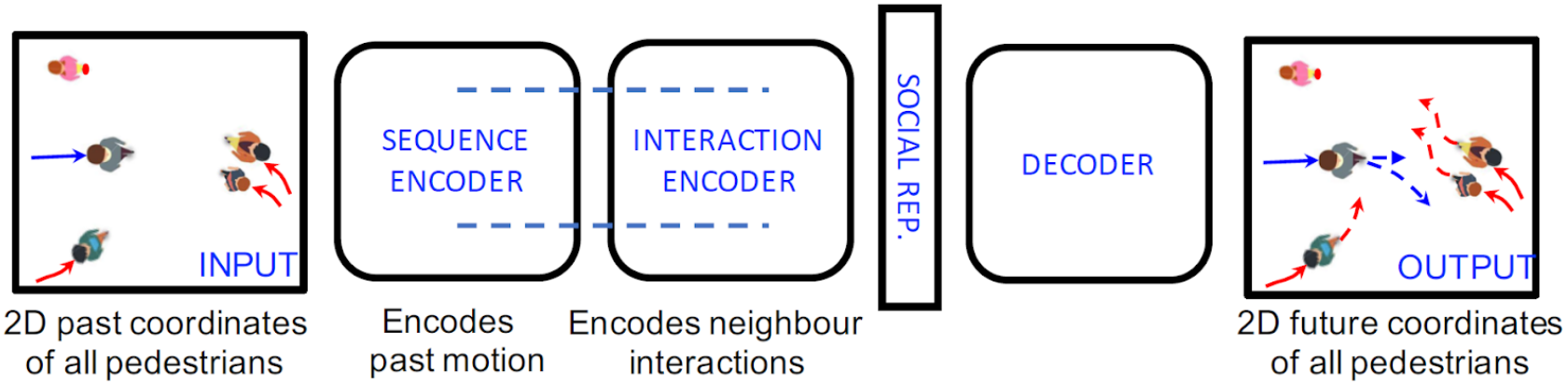
Asymmetrical Bi-RNNs, 3rd place solution at the ICCV Trajnet++ Challenge

Raphaël Rozenberg[1,2], Joseph Gesnouin[2,3] & Fabien Moutarde[2]

[1] *École Normale Supérieure, Université PSL*
[2] *MINES ParisTech, Université PSL, Centre de Robotique*
[3] *Institut VEDECOM*

October 17, 2021

Typical pipeline



- Modern approaches specifically focused on the presence of social interactions.
- We focus on the encoding part of the trajectories of individual people.

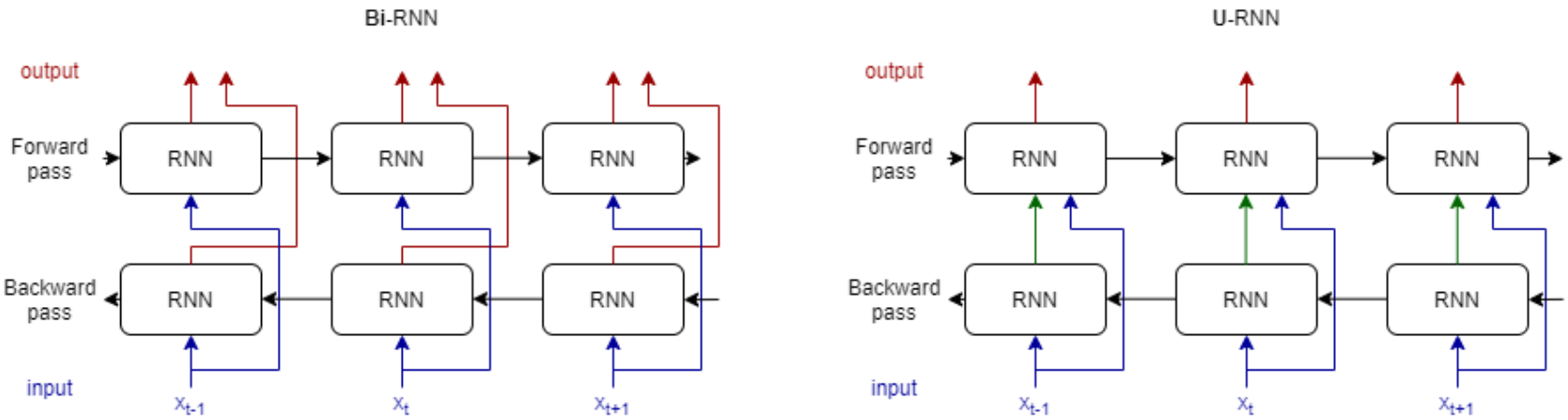
---

[1] [Kothari et al., 2021], "Human trajectory forecasting in crowds: A deep learning perspective"

[2] [Alahi et al., 2016] "Social lstm: Human trajectory prediction in crowded spaces"

A better encoder? From RNNs & Bi-RNNs to U-RNNs

The data has a **preferred direction** in time: the forward direction.



$\longrightarrow$ We accumulate information while knowing which part of the information will be useful in the future.

[1] [Xue et al., 2017], "Bi-prediction: pedestrian trajectory prediction based on bidirectional LSTM classification."

Results on Trajnet++ real world dataset

Model (Encoder - Decoder)	**Interaction**	**ADE (m)** ± 0.01 m	**FDE (m)** ± 0.01 m	**Col-I (%)** ± 0.5%	**Col-II (%)** ± 1%
Constant velocity	None	0.68	1.42	14.3	15.2
None - GRU	Dir.	0.63	1.33	6.9	12.1
LSTM - LSTM	Occ.	0.58	1.23	11.5	13.9
U-LSTM - LSTM	Occ.	**0.57**	**1.22**	**10.2**	14.9
GRU - GRU	Dir.	0.58	1.24	6.5	12.4
Bi-GRU - GRU	Dir.	0.59	1.26	6.7	11.7
U-GRU - GRU	Dir.	0.58	1.25	6.5	11.7
reversed U-GRU - GRU	Dir.	0.58	1.25	6.5	**11.0**
LSTM - LSTM	Dir.	0.58	1.25	6.4	11.4
Bi-LSTM - LSTM	Dir.	0.59	1.28	6.2	11.9
U-LSTM - LSTM	Dir.	**0.56**	**1.22**	**5.2**	11.9
reversed U-LSTM - LSTM	Dir.	0.58	1.26	6.6	**11.1**
LSTM - LSTM	Soc.	0.55	1.18	6.9	12.7
U-LSTM - LSTM	Soc.	**0.53**	**1.15**	**6.5**	11.5
Social NCE	Soc. + contr.	**0.53**	**1.14**	**5.3**	11.3

[1] [Liu et al., 2021], "Social NCE: Contrastive Learning of Socially-aware Motion Representations"

Conclusion

- A new sequence encoder.
- Interactions are **NOT** the only aspect on which pedestrian trajectory prediction can progress.
- Could be used to significantly **improve current trajectory prediction algorithms**:
  - TrouSPI-Net [Gesnouin et al., 2021]
  - (GO)-Home [Gilles et al., 2021a, Gilles et al., 2021b]

$\longrightarrow$ [Rozenberg et al., 2021], "Asymmetrical Bi-RNN for pedestrian trajectory encoding", *https://arxiv.org/abs/2106.04419*

Bibliography

Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., and Savarese, S. (2016).

Social lstm: Human trajectory prediction in crowded spaces.
In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Gesnouin, J., Pechberti, S., Stanciulescu, B., and Moutarde, F. (2021).

TrouSPI-Net: Spatio-temporal attention on parallel atrous convolutions and U-GRUs for skeletal pedestrian crossing prediction.
arXiv e-prints, page arXiv:2109.00953.

Gilles, T., Sabatini, S., Tsishkou, D., Stanciulescu, B., and Moutarde, F. (2021a).

Gohome: Graph-oriented heatmap output forfuture motion estimation.
arXiv preprint arXiv:2109.01827.

Gilles, T., Sabatini, S., Tsishkou, D., Stanciulescu, B., and Moutarde, F. (2021b).

Home: Heatmap output for future motion estimation.
arXiv preprint arXiv:2105.10968.

Kothari, P., Kreiss, S., and Alahi, A. (2021).

Human trajectory forecasting in crowds: A deep learning perspective.
IEEE Transactions on Intelligent Transportation Systems.

Liu, Y., Yan, Q., and Alahi, A. (2021).

Social nce: Contrastive learning of socially-aware motion representations.
In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 15118–15129.

Rozenberg, R., Gesnouin, J., and Moutarde, F. (2021).

Asymmetrical bi-rnn for pedestrian trajectory encoding.
arXiv preprint arXiv:2106.04419.

Xue, H., Huynh, D. Q., and Reynolds, M. (2017).

Bi-prediction: pedestrian trajectory prediction based on bidirectional lstm classification.
In *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, pages 1–8. IEEE.