Asymmetrical Bi-RNN for pedestrian trajectory encoding

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

[1] Ecole Normale Superieure - Université PSL
[2] Centre de Robotique, MINES ParisTech - Universite PSL
[3] Institut Vedecom

Objectives

- Provide insight into designing improved motion encoders.
- Introduce a new Bi-RNN that respects the preferred direction of the data in time.
- Evaluate the proposed Bi-RNN approach by working out some word problems: trajectory forecasting.

Human Trajectory Forecasting

Pedestrian motion behavior involves a **combination of individual goals and social interactions** with other agents.

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

From Bi-RNNs to U-RNNs

Some movements are made because pedestrians **anticipate a potential obstacle**, however, using Bi-RNNs leads to two drawbacks:

- **The architecture is symmetrical in both time directions**.
- The output is simply the **concatenation of two naive readings of the input in both directions**.

The idea behind U-RNN is to use during the forward pass information about the future that was acquired during the backward pass. We accumulate information while knowing which part will be useful in the future with respect to the preferred direction of the data.

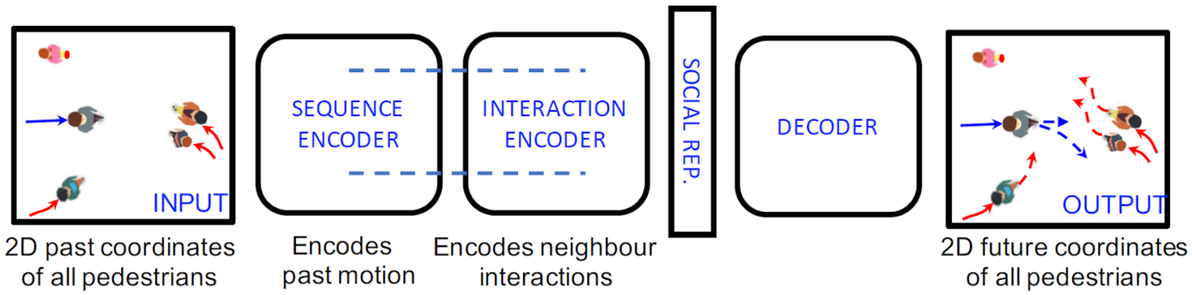

Figure 1: Typical pipeline for human trajectory forecasting.

The Asymmetrical Bi-RNN (U-RNN) architecture

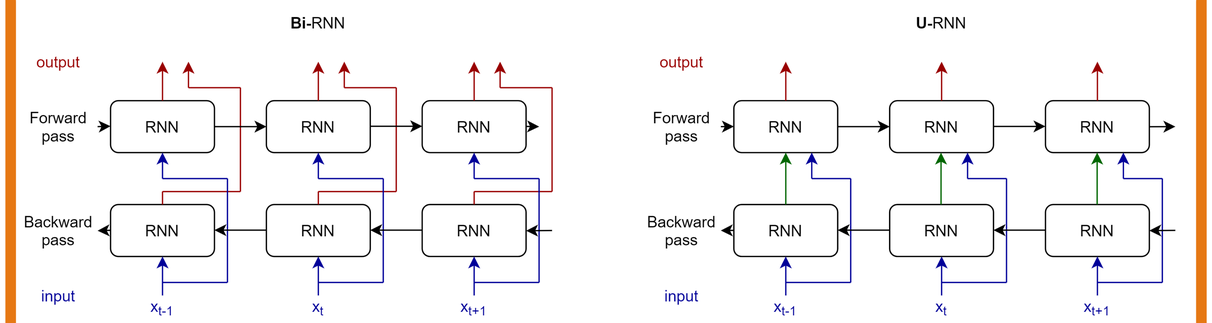

Figure 2: Comparison between Bi-RNN and U-RNN architectures (blue: inputs - red: outputs - black: hidden states - green: intermediate output).

Comparison of motion-encoding designs with respect to various interactions modules architectures on TrajNet++.

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

Results Analysis

- U-LSTM encoder helped get significantly better ADE, FDE and Col-I for **a variety of interaction modules**, suggesting that there was indeed **unused information in past trajectories**.
- The better performance of U-LSTM compared to U-GRU strongly indicates that the additional in formation extracted by the U-RNN architecture came from **long-term dependencies**.
- The hypothesis that our non-symmetrical architecture should better leverage information by using the preferred direction of the data is supported by the **absence of performance improvement when using a reversed U-LSTM encoder**.

Conclusion

- A new sequence encoder.
- Interactions are **NOT** the only aspect on which pedestrian trajectory can progress.
- Could be used to significantly **improve current pedestrian trajectory prediction algorithms**.

Ressources

- Raphaël Rozenberg, Joseph Gesnouin, Fabien Moutarde. Asymmetrical Bi-RNNs (U-RNNs), 2nd place solution at the Trajnet++ Challenge for pedestrian trajectory forecasting. Workshop on Long-term Human Motion Prediction, ICRA 2021.
- Rozenberg, Raphaël, Joseph Gesnouin, and Fabien Moutarde. "Asymmetrical Bi-RNN for pedestrian trajectory encoding." arXiv preprint arXiv:2106.04419 (2021).