HAL open science

Differences in pedestrian trajectory predictions for high- and low-density situations

Raphael Korbmacher, Huu-Tu Dang, Antoine Tordeux, Benoit Gaudou, Nicolas Verstaevel

To cite this version:

Raphael Korbmacher, Huu-Tu Dang, Antoine Tordeux, Benoit Gaudou, Nicolas Verstaevel. Differences in pedestrian trajectory predictions for high- and low-density situations. 14th International Conference on Traffic and Granular Flow (TGF 2022), Oct 2022, Dehli, India. hal-03793751

HAL Id: hal-03793751

https://hal.science/hal-03793751v1

Submitted on 28 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Empirical comparison of different pedestrian trajectory prediction methods at high densities

Raphael Korbmacher, Huu-Tu Dang, Antoine Tordeux, Benoit Gaudou, Nicolas Verstaevel

Abstract Predicting human trajectories is a challenging task due to the complexity of pedestrian behavior, which is influenced by external factors such as the scene's topology and interactions with other pedestrians. A special challenge arises from the dependence of the behaviour on the density of the scene. In the literature, deep learning algorithms show the best performance in predicting pedestrian trajectories, but so far just for situations with low-densities. In this study, we aim to investigate the suitability of these algorithms for high-density scenarios by evaluating them using two error metrics and comparing their accuracy to that of knowledge-based models. The first metric is distance-based, while the second counts the number of collisions between pedestrians. Our findings reveal that deep learning algorithms provide improved trajectory accuracy in the distance metric, but knowledge-based models perform better in avoiding collisions.

1 Introduction

The rapid increase in urban populations in recent decades, particularly in developing countries, has led to a strain on local infrastructure and an increase in high-density situations, where the number of people in one area can reach 2-8 pedestrians per square meter. This poses a challenge for researchers studying pedestrian behavior and how it relates to infrastructure, events, and evacuations. To ensure safety in these high-density situations, tools are needed to plan and organize them effectively. Scientists typically use knowledge-based models (KB) to understand and simulate

Raphael Korbmacher, Antoine Tordeux
School of Mechanical Engineering and Safety Engineering, University of Wuppertal, Germany e-mail: korbmacher@uni-wuppertal.de; tordeux@uni-wuppertal.de

Huu-Tu Dang, Benoit Gaudou, Nicolas Verstaevel
UMR 5505 IRIT, University Toulouse 1 Capitole, France e-mail: huu-tu.dang@ut-capitole.fr; benoit.gaudou@ut-capitole.fr; nicolas.verstaevel@ut-capitole.fr

pedestrian dynamics [1]. These models are based on basic rules or generic functions, and can be divided into two main categories: microscopic and macroscopic. Microscopic models, such as the acceleration-based and velocity-based models, focus on simulating the behavior of individual pedestrians and their local interactions [2]. On the other hand, macroscopic models, such as the Hughes model [3], treat pedestrian flow as fluid dynamics, focusing on the aggregation of crowd information.

In recent years, deep learning (DL) algorithms have gained popularity as a method for predicting pedestrian trajectories [4]. Notable examples include Social-LSTM [5], which incorporates information from nearby pedestrians using a social pooling layer, and Social-GAN [6], which utilizes a generator and discriminator to generate and evaluate predictions. These DL algorithms have been shown to be more accurate than traditional KB models, but have only been tested in low-density situations.

This study aims to investigate the performance of DL algorithms in high-density situations, and to compare their accuracy to that of KB models. To evaluate the predictions, we use two metrics: a distance-error metric and a collision metric. These metrics allow us to assess the accuracy of the predictions, as well as their realism, in both low-density and high-density situations.

2 Methodology

2.1 Datasets

For the evaluation of predictions on low-density datasets, two mostly-used public datasets are selected: ETH [7] and UCY [8]. There are a total of 750 trajectories in the ETH dataset, which have been subdivided into two sets: ETH and Hotel. In the UCY dataset, three sets have been chosen: ZARA01, ZARA02, and UNIV which contain 786 trajectories in total. Both data sets are collected in outdoor settings and include numerous pedestrian traffic patterns: unidirectional, bidirectional, and multidirectional.

For high-density datasets, to the best of our knowledge, the only ones available come from laboratory experiments at Forschungszentrum Jülich [9]. There are two types of pedestrian flows in the Jülich datasets: unidirectional (UNI) and bidirectional (BIDI). There are some scenarios in which pedestrians are asked to perform particular behaviors; for the purpose of keeping natural behavior, these scenarios will be removed. Finally, the UNI dataset includes 3638 trajectories with 5 data sets, whereas the BIDI dataset contains 3096 trajectories in total with 6 data sets.

Given each set in the dataset, the average density of this set is calculated by averaging the instantaneous densities over some different time steps. The average density of the data sets in the low-density datasets are $0.13 - 0.38\ ped/m^2$. In the high-density datasets, the UNI dataset has average densities of $0.38 - 2.33\ ped/m^2$, whereas the BIDI dataset average denisties of $0.38 - 3.0\ ped/m^2$.

2.2 Models and algorithms

Two state-of-the-art KB approaches are considered: SFM [10] and ORCA [11]. These modeling techniques take information from neighbors into account when computing the exact velocity for the next step. Furthermore, the constant-velocity (CV) model is chosen to check how good these prediction methods are compared to just using instantaneous velocity as predictions.

For the DL approach, we use the two most commonly used algorithms in the pedestrian trajectory prediction domain. These are the DL architectures: LSTM and GAN. For both architectures, we used two types of algorithms. One of them is quite simple and just takes the historical trajectory, of the pedestrian whose trajectory is predicted, into account. It is called the vanilla LSTM or vanilla GAN. These algorithms are Grid-based which means that the input is a local grid constructed around the pedestrian. As input, they receive single positions and linearly embed them in a 32-dimensional vector, which is appropriate to the spatial pooling size. Three LSTM layers with 128 hidden dimensions and a linear output layer follow. The other forms of algorithms additionally have social pooling mechanism implemented in their architectures to aggregate neighbor information in the grid. These two algorithms are the Social-LSTM [5] and the Social-GAN [6]. These social pooling layers make it possible to not just use the history of the pedestrian as input, but also the history trajectories of the surrounding pedestrians. This allows interactions to be taken into account in the predictions. Table 1 presents different kinds of information needed for each model or algorithm to make predictions.

Table 1: Kind of information feed to the algorithms/models.

	Past Trajectory	Neighbor Information
CV	-	-
SF	-	✓
ORCA	-	✓
Vanilla-LSTM	✓	-
Social-LSTM	✓	✓
Social-GAN	✓	✓

2.3 Calibration/Training

For KB models, the parameters of the ORCA and SF models are calibrated according to the given data set, and the CV model does not need any training or calibration. Relevant information about some specific frames is extracted from real datasets and used as inputs to simulate pedestrians in the following steps. The simulated trajectories are then compared to the real trajectories.

For DL algorithms, in contrast to the configuration of related contributions [4], we do not use a prediction length of $4.8s$, but of $1.6s$. The observation length is $1.2s$. The reason for choosing smaller time intervals is that at high densities, the interaction scales are much shorter. If higher time scales are chosen, the number of neighbours that have to be taken into account at those densities would be too high. The frame rate for the high-density datasets is 0.125, which means 9 observations are taken to make 12 predictions. For the low-density datasets, the frame rate is 0.4. Therefore, 3 observations are taken to make 4 predictions. The DL algorithms are trained with a learning rate of 0.0015, and a RMS-prop is used as the ADAM optimizer. The batch size is 8, and we train for 15 epochs. As a loss function, the Mean Squared Error is used. For the validation and testing, we use a hold-out validation strategy. 15% of the data is used for validation, 15% for testing and the rest for training. The calculations are done using PyTorch [1].

2.4 Evaluation/Testing

The most commonly used metric for pedestrian trajectory predictions is the distance-based metric Average Displacement Error (ADE) [12]. The ADE averages the Euclidean distance between points of the predicted trajectory and the ground truth over all predicted steps. Recently, more and more works emphasize the importance of a metrics that takes the number of collisions into account [12, 13]. Collision metrics provide more insights about how realistic the predicted behavior is, especially in high density situations, since avoiding collisions with other pedestrians is the most important motivation when navigating through crowds. For calculating the collision metric, we use the equation that was developed by Kothari et al. [12].

$$\mathbf{Col} = \frac{\sum_{\hat{Y} \in S} Col(\hat{Y})}{|S|}, \; Col(\hat{Y}) = min(1, \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j>i}^{N} [||\hat{y}_i^{(t)} - \hat{y}_j^{(t)}||_2^2 \leq \delta]) \quad (1)$$

where S includes all scenes in the test set, $\hat{Y}$ is a scene prediction which contains N agents, $\hat{y}_i$ is the prediction of agent i over the prediction length of T.

If the predicted trajectory of a pedestrian intersects with neighbour trajectories, the prediction is counted as a prediction with collision. The metrics indicates the percentage of predictions in which collisions occurred. Important for that is the chosen size of the pedestrians, which is given by the variable δ in equation 1. If δ is increased, the number of collisions also increases. Adapted from the heuristic of Moussaïd et al. [14] for estimating the size of pedestrians, we choose a pedestrian radius of $0.2m$, hense $\delta = 0.4m$.

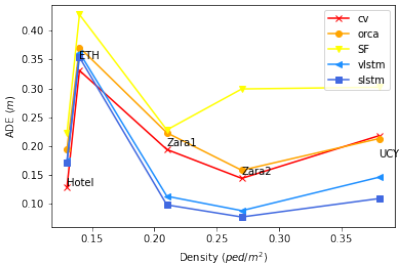
[1] http://pytorch.org

3 Results

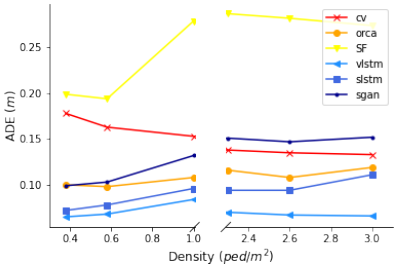
In this section, we provide the evaluation results of trajectory predictions for both low-density and high-density datasets. The ADE and collision metrics have been applied and the results are compared to provide insights on the performance of the prediction methods.

3.1 Distance metric

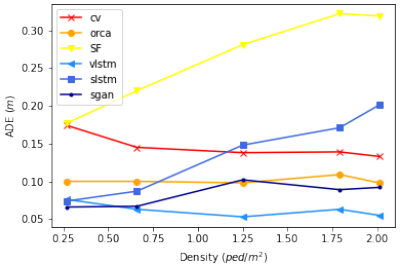
Fig. 1a shows the ADE results of prediction for the low-density datasets. The average density of each set of the low-density datasets are displayed on the x-axis, and the result of the ADE metric is displayed on the y-axis. It can be seen that the DL algorithms outperform the KB models for almost all densities. For the Zara1, Zara2, and UCY data sets, the results of DL algorithms are much more accurate. The Social-LSTM performed slightly better than the Vanilla-LSTM, which was expected. The SF model performs worst, in particular for the two data sets Zara2 and UCY which have higher densities.



(a) Low-density datasets.



(b) BIDI dataset.



(c) UNI dataset.

Fig. 1: Results of distance error-metric (ADE).

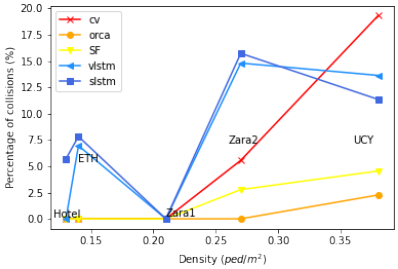
The ADE of the predictions for the high-density datasets is shown in Fig. 1b,c. The ADE results of the BIDI dataset are depitched in Fig. 1b, while Fig. 1c shows ADE results of the UNI dataset. There are a couple of similarities to the results in Fig. 1a. The SF model makes the worst predictions, especially at higher densities, for the UNI dataset as well as the BIDI dataset. Furthermore, the DL algorithms have better prediction accuracy than the KB models. In the BIDI dataset, the Vanilla-LSTM and the Social-LSTM perform best, whereas the Vanilla-LSTM and the SGAN work best in the UNI scenario. In both cases, the ADE rises for the Social-LSTM with increasing densities. Contrary to that, the ADE for the Vanilla-LSTM has the tendency to fall with increasing densities. Likewise, the Social-GAN and the SF show growing ADE with increasing densities, while the ADE of the CV seems to fall with increasing densities. It seems that approaches that take neighbour information into account, like the Social-LSTM, the SGAN, the SF, and the ORCA do, show a rising ADE, while approaches that do not use neighbour information's as an input have fallen ADE's, like the Vanilla-LSTM and CV.

3.2 Collision metrics

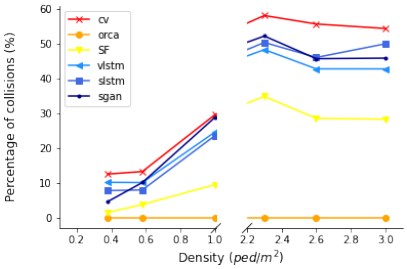
Fig. 2a shows the results of collision metric for the low-density datasets. The ORCA model has the fewest collisions, followed by the SF model, both of which have far less collisions than the DL algorithms. In the Zara2 and UCY data sets, some grouping behavior can be observed which results in a high local density. In these groups, the pedestrians walk very close together with almost touching shoulders. This is counted as a collision, although it is not what we would expect as a collision. Besides that, it is clearly shown that the percentage of collisions predicted by the DL algorithms is too high. There is only the Zara1 data set which is collision-free, but for the others the collision number is always higher than in the real trajectories and especially higher than in the predictions by the KB models.

Fig. 2b,c present the percentage of collisions in the high-density datasets. In Fig. 2b, the results for the BIDI dataset are presented, while in Fig. 2c shows the results for the UNI dataset. In general, the results of the collision metric of the high-density datasets are higher than those in the low-density datasets. In both high-density datasets, it could clearly be shown that there is a tendency for a rising percentage of collisions with increasing densities. This applies in particular to the bidirectional environment, where the percentage of collisions in situations with densities above 2 ped/m^2.

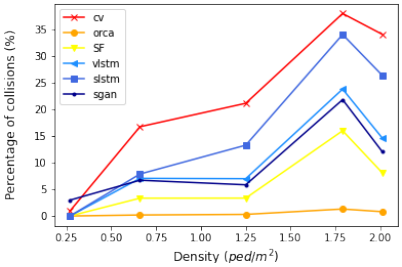
In summary, with the distance metric, the DL algorithms get better ADE results in the high-density datasets as there is a decline in the ADE results at higher density compared to those in the low-density datasets. In contrary to this, the collision metric indicates a completely opposite trend with an increase in the collision metric results from the low-density datasets to the high-density datasets. This demonstrated that the prediction of DL algorithms in high-density situations gains better distance accuracy, but the realism is worse as the predicted trajectories result in many collisions. KB



(a) Low-density datasets.



(b) BIDI dataset.



(c) UNI dataset.

Fig. 2: Results of collision metric.

models, on the other hand, handle collisions well despite producing predictions with a high ADE.

4 Conclusion

In the last decade, DL algorithms have shown great promise as one of the state-of-the-art approaches for human trajectory prediction. However, the performance of such algorithms is mostly proven in low-density scenarios. This work demonstrated their poor performance in dense environments by revealing an increase in collision metric findings in comparison to KB models. Specifically, we showed that the Social-LSTM, which excels in low-density settings, fares worse than the Vanilla-LSTM in high-density situations. Hence, the need to reduce the number of collisions in the predicted trajectories in crowded situations is also important in terms of reality rather than just focusing on distance-error metrics. Different improvements can be used to tackle this problem. One is to combine distance-error and collision metrics for training the algorithms. Another method is to use a hybrid approach, where DL algorithms are combined with KB models. Furthermore, a more realistic elliptical pedestrian shape, grouping behavior, and new collision metrics should be taken into

account for better collision evaluation. Overcoming these problems by improving the collision metric would be an important step toward the use of DL algorithms for making realistic trajectory predictions at high densities.

Acknowledgements The authors acknowledge the Franco-German research project MADRAS funded in France by the Agence Nationale de la Recherche (ANR, French National Research Agency), grant number ANR-20-CE92-0033, and in Germany by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), grant number 446168800.

References

1. Schadschneider, A., Chraibi, M., Seyfried, A., Tordeux, A., & Zhang, J. Pedestrian dynamics: From empirical results to modeling. *Crowd Dynamics*, Volume 1 (pp. 63-102). Birkhäuser, Cham. (2018)
2. Chraibi, M., Tordeux, A., Schadschneider, A., & Seyfried, A. (2018). Modelling of pedestrian and evacuation dynamics. Encyclopedia of complexity and systems science, 1-22.
3. Hughes, R. A continuum theory for the flow of pedestrians. *Transportation Research Part B: Methodological*. **36**, 507-535 (2002)
4. Korbmacher, R. & Tordeux, A. Review of Pedestrian Trajectory Prediction Methods: Comparing Deep Learning and Knowledge-Based Approaches. *IEEE Transactions On Intelligent Transportation Systems*. (2022)
5. Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L. & Savarese, S. Social LSTM: Human Trajectory Prediction in Crowded Spaces. *2016 IEEE Conference On Computer Vision And Pattern Recognition (CVPR)*. pp. 961-971 (2016)
6. Gupta, A., Johnson, J., Fei-Fei, L., Savarese, S. & Alahi, A. Social gan: Socially acceptable trajectories with generative adversarial networks. *Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition*. pp. 2255-2264 (2018)
7. Pellegrini, S., Ess, A., Schindler, K. & Van Gool, L. You'll never walk alone: Modeling social behavior for multi-target tracking. *2009 IEEE 12th International Conference On Computer Vision*. pp. 261-268 (2009)
8. Lerner, A., Chrysanthou, Y. & Lischinski, D. Crowds by example. *Computer Graphics Forum*. **26** pp. 655-664 (2007)
9. Cao, S., Seyfried, A., Zhang, J., Holl, S. & Song, W. Fundamental diagrams for multidirectional pedestrian flows. *J. Of Statistical Mechanics: Theory And Experiment*. **2017**, 033404 (2017)
10. Helbing, D., Farkas, I. & Vicsek, T. Simulating dynamical features of escape panic. *Nature*. **407**, 487-490 (2000)
11. Berg, J., Guy, S., Lin, M. & Manocha, D. Reciprocal n-body collision avoidance. *Robotics Research*. pp. 3-19 (2011)
12. Kothari, P., Kreiss, S. & Alahi, A. Human trajectory forecasting in crowds: A deep learning perspective. *IEEE Transactions On Intelligent Transportation Systems*. (2021)
13. Mohamed, A., Qian, K., Elhoseiny, M. & Claudel, C. Social-stgcnn: A social spatio-temporal graph convolutional neural network for human trajectory prediction. *Proceedings Of The IEEE/CVF Conference On Computer Vision And Pattern Recognition*. pp. 14424-14432 (2020)
14. Moussaïd, M., Helbing, D. & Theraulaz, G. How simple rules determine pedestrian behavior and crowd disasters. *Proc. Of The National Academy Of Sciences*. **108**, 6884-6888 (2011)