HAL open science

Deep Learning for Predicting Pedestrian Trajectories in Crowds

Raphael Korbmacher, Antoine Tordeux

To cite this version:

Raphael Korbmacher, Antoine Tordeux. Deep Learning for Predicting Pedestrian Trajectories in Crowds. Intelligent Systems and Applications. IntelliSys 2023, Sep 2023, Amsterdam, Netherlands. pp.720-725, 10.1007/978-3-031-47718-8_46 . hal-04481327

HAL Id: hal-04481327

https://hal.science/hal-04481327v1

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.

Deep Learning for predicting pedestrian trajectories in crowds

Raphael Korbmacher[1] and Antoine Tordeux[1]

Bergische Universität Wuppertal, Wuppertal, Germany,
korbmacher@uni-wuppertal.de

Abstract. In the last decade, deep learning algorithms have gained a lot of attention in many fields. Even scientists who study pedestrian behavior and usually create physics-based models to simulate realistic behavior are increasingly using these algorithms. So far, however, this has only been tested in low-density environments. In this contribution, we want to investigate the question of whether the methods are also suitable for crowds. Therefore, we prepared high-density datasets and tried to predict the behaviors of the pedestrian with physics-based models as well as deep learning algorithms. We compare the prediction accuracy in terms of two error metrics: a distance-based metric and a metric that counts the number of collisions between pedestrians. The results show that the algorithms beat the models regarding the distance metric, but perform worst in terms of the collision metric. To be able to predict more realistic crowd behavior, the deep learning algorithm needs to improve collision avoidance. In order to make more accurate predictions, the deep learning algorithms need to be improve collision avoidance.

Keywords: deep learning, pedestrian predictions, crowds, physics-based models

1 Motivation

The number of people living in urban areas has been increasing rapidly for decades. Particularly in developing countries, the urban population is rising so rapidly in some cases that local infrastructures are overloaded. As a result, there are increasingly situations where there are so many people in one place that one can speak of crowds or high-density situations (2-8 pedestrians/m^2). These situations pose challenges for researchers studying pedestrian behavior when it comes to infrastructure, events, and evacuations. To prevent people from getting injured or even dying in high pedestrian density situations, tools are needed to safely plan and organise such situations. Typical tools used by scientists are physics-based models that define pedestrian behavior and can be used in simulation platforms [1]. In the last decade, Deep Learning has become an important topic in this field [2]. Many papers have shown that Deep Learning algorithms outperform the models in terms of trajectory prediction accuracy [3,4]. So far, however, they have only been used in low-density situations where there are

few pedestrians. We want to investigate whether these algorithms are useful for high-density situations. If this is the case, the algorithms can replace the models and be used for the safe planning of infrastructure, events, and much more.

2 Related work

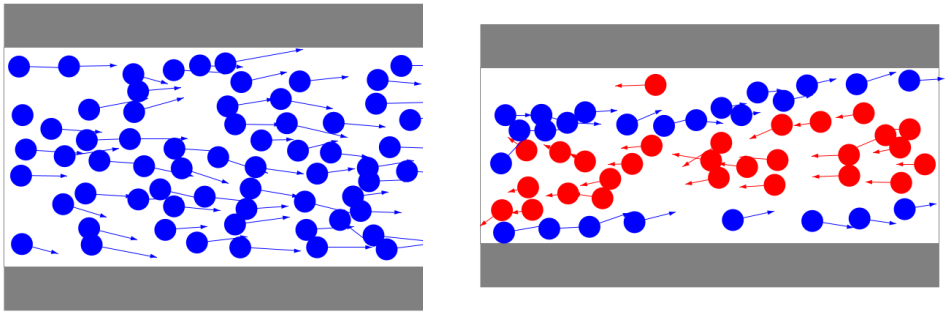
In the field of pedestrian dynamics, researchers develop physics-based models that help them understand and simulate certain situations. These models are defined by basic rules or generic functions and consider physical as well as social or psychological factors of the pedestrians. They can be specified by a few parameters that generally have physical interpretations. Prominent microscopic KB models are the social force model (SFM) [5] or the optimal reciprocal collision avoidance (ORCA) [6]. In general, the goal of these models is to improve the understanding of pedestrian dynamics and to identify microscopic mechanisms and parameters for the emergent patterns and organisation in given scenes.

Recently, deep learning algorithms have gained a lot of attention in this field. The reason for that is that predicted behavior is more accurate with algorithms than with models. This makes them relevant for systems like autonomous vehicles or social robots that have to interact with pedestrians. In contrast to the models, no functions or physical rules are needed, just data. Recurrent Neural Networks (RNN) are often used for trajectory predictions, because of their capability to handle sequence data. The first RNN published for this task was the Social-LSTM from Alahi et al. [3] that contains a social pooling layer. In this layer, both the past trajectories and those of the neighbors are processed to be able to model interactions. Generative adversarial networks (GAN) are another type of algorithm that has been shown to be useful for trajectory predictions. The most well-known is the Social-GAN from Gupta et al. [7].

3 Methodology

3.1 Data sets

We use two different environments of data to train the algorithms and calibrate the models. Because there are no real-world trajectory datasets of high-density situations available we use data sets from pedestrian experiments [8]. The first environment contains six different data sets with different experiment settings. The density is between 0.25 and 2 ped/m^2. In this experiment, there is a 5 m wide and 12 m long corridor with pedestrians just coming from one direction. It is called the unidirectional (Uni) environment. This setting is shown in Fig. 1 (a). In the other environment, the corridor is 4 m wide and 12 m long, and the pedestrians are coming from both directions, as shown in Fig. 1(b). It is called the bidirectional (Bidi) environment. The density of these data sets is between 0.4 and 3 ped/m^2.



(a) Uni. (b) Bidi.

Fig. 1: Experiments reconstruction with knowledge-based approaches.

3.2 Models and algorithms

For making the predictions, we use three deep learning algorithms and three physics-based models. The deep learning algorithms are the two Social-LSTM [3] and the Social-GAN[7] as well as the V-LSTM. In contrast to the first two algorithms, the V-LSTM does not have a social layer and just uses the past trajectory of a single pedestrian to predict the future trajectory. For the physics-based approach, we use the Social-Force model (SF) [5] and the optimal reciprocal collision avoidance model (ORCA) [6]. The third physics-based model is the constant velocity model. In this simple model, it is assumed that the direction and speed do not change.

3.3 Calibration/ Training

As an input, we use past trajectories with a length of 1.2 sec. The prediction length is 1.6 sec. The frame rate of the data sets is 0.125, which means 9 observations are taken to make 12 predictions. The learning rate is 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 of the distance between the predicted trajectory and real trajectory is used. For 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. The parameters of the ORCA and SF are calibrated according to the given data set. Important parameters are the desired velocity and the desired distance from other pedestrians.

3.4 Evaluation/ Testing

The most commonly used metrics for pedestrian trajectory predictions are the distance-based metrics ADE [9] and FDE [10]. The first one averages the Euclidean distance between points of the predicted trajectory and the ground truth

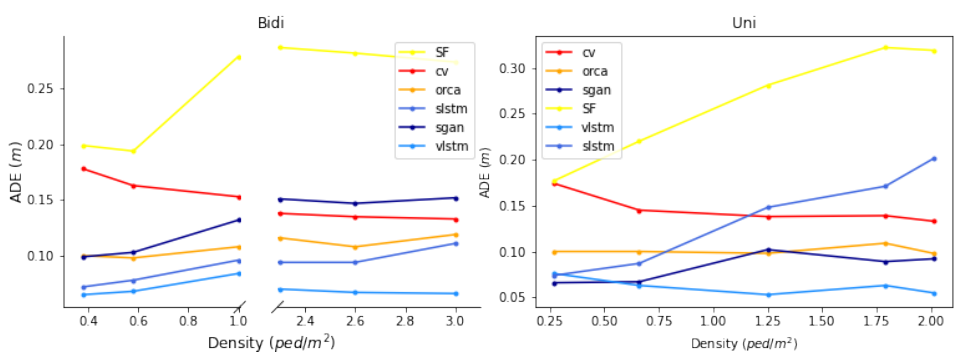
that have the same temporal distance, and the second one measures the distance between the final predicted position and the ground truth position. These two metrics have been the only metrics for evaluating pedestrian trajectory predictions for some time. Recently, more and more works emphasize the importance of metrics that takes the number of collisions into account [11]. Pedestrians have an intrinsic motivation not to collide with each other. Especially in high-density situations, these metrics provide more insights into how realistic the predicted behavior is. These metrics are important for high-density situations because avoiding collisions with other pedestrians is the most important motivation when navigating through crowds. For calculating the collision metric, we use the term in Eq. 1 which was developed by Kothari et al. [11].

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

If the predicted trajectory of a pedestrian intersects with neighbor trajectories the prediction is counted as a prediction with collision. The metrics indicate the percentage of predictions in which collisions occurred.

4 Results

First, we will have a look at the distance error metric. In Fig. 2 two diagrams are shown. On the left side, the prediction accuracy for the Bidi environment is presented and on the right side for the Uni environment. These data environments contain data sets from different experimentation settings which have different densities. The average densities of each of the data sets are plotted on the x-axis and the distance error metirc on the y-axis.

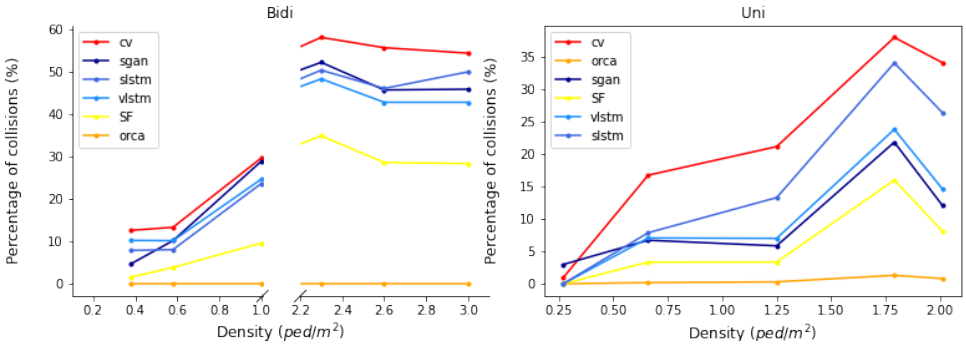


(a) Bidi data set (b) Uni data set

Fig. 2: Distance error-metric (ADE) for high-density data set.

It can be seen that the algorithms outperform the models for almost all densities. Especially the V-LSTM shows a good performance which is an indicator, that the neighbor information is not as important at high density as it is in low densities. The reason for that is, that the pedestrians do not have many degrees of freedom and their behavior is mostly defined by following the preceding pedestrian. Interaction do not play an important role like they do in low-density situations.

Next we want to present the results for the collision metric, which are shown in Fig. 3. In Fig. 3 (a) the results for the bidirectional flow are presented and in



(a) Bidi data set (b) Uni data set

Fig. 3: Collision error-metric for high-density data set.

Fig. 3 (b) the results for the unidirectional flow. In both environments, it could clearly be shown, that there is a tendency for a rising percentage of collisions with rising density. This applies in particular to the Bidi environment where the percentage of collisions in situations with densities above 2 ped/m^2 is very high. Furthermore, it can be seen that the models SF and ORCA are much better in terms of this metric. The predictions based on ORCA have almost no collisions. Both models outperform the algorithms for every data set and density level.

5 Conclusion

In Fig. 2 it could be clearly shown, that the deep learning algorithms outperform the physics-based models in terms of prediction accuracy. This is also the result that can be found in the literature for low density data sets. To the best of our knowledge, this is the first work to show that this is also true for predictions of trajectories in crowds. Nevertheless, the predictions produce too many collisions, as shown in Fig 3. Therefore, future work must focus on reducing the number of collisions. Promissing ways to do that is to combine the physics-based and deep learning approach (hybrid approach).

References

1. Schadschneider, A., Chraibi, M., Seyfried, A., Tordeux, A., Zhang, J. (2018). Pedestrian dynamics: From empirical results to modeling. In Crowd Dynamics, Volume 1 (pp. 63-102). Birkhäuser, Cham.
2. Korbmacher, R., Tordeux, A. (2022). Review of Pedestrian Trajectory Prediction Methods: Comparing Deep Learning and Knowledge-Based Approaches. IEEE Transactions on Intelligent Transportation Systems.
3. Smith, T.F., Waterman, M.S.: Identification of common molecular subsequences. J. Mol. Biol. 147, 195?197 (1981). doi:10.1016/0022-2836(81)90087-5
4. Lee, N., Choi, W., Vernaza, P., Choy, C. B., Torr, P. H., Chandraker, M. (2017). Desire: Distant future prediction in dynamic scenes with interacting agents. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 336-345).
5. Helbing, D., Molnar, P. (1995). Social force model for pedestrian dynamics. Physical review E, 51(5), 4282.
6. Berg, J. V. D., Guy, S. J., Lin, M., Manocha, D. (2011). Reciprocal n-body collision avoidance. In Robotics research (pp. 3-19). Springer, Berlin, Heidelberg.
7. Gupta, A., Johnson, J., Fei-Fei, L., Savarese, S., Alahi, A. (2018). Social gan: Socially acceptable trajectories with generative adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2255-2264).
8. Forschungszentrum Jülich, Jülich Supercomputing Centre. Data archive of experiments on pedestrian dynamic. Accessed: 2021-08-01.
9. Pellegrini, S., Ess, A., Schindler, K., Van Gool, L. (2009, September). You'll never walk alone: Modeling social behavior for multi-target tracking. In 2009 IEEE 12th international conference on computer vision (pp. 261-268). IEEE.
10. Lerner, A., Chrysanthou, Y., Lischinski, D. (2007, September). Crowds by example. In Computer graphics forum (Vol. 26, No. 3, pp. 655-664). Oxford, UK: Blackwell Publishing Ltd.
11. Kothari, P., Kreiss, S., Alahi, A. (2021). Human trajectory forecasting in crowds: A deep learning perspective. IEEE Transactions on Intelligent Transportation Systems.