MadeAI 2025 - Modelling, Data Analytics and AI in Engineering
7-11 July 2025, Porto, Portugal

# A Dynamic Distance Social LSTM for Predicting Pedestrian Trajectories in Crowded Environments

**Ahmed Alia**[1,*], **Mohcine Chraibi**[1], **Armin Seyfried**[1,2]

[1]Institute for Advanced Simulation, Forschungszentrum Jülich, Jülich, 52428, Germany
[2] School of Architecture and Civil Engineering, University of Wuppertal, Wuppertal, Germany
E-mail: a.alia@fz-juelich.de. m.chraibi@fz-juelich.de, a.seyfried@fz-juelich.de

**Abstract:** This work introduces dynamic distance Social Long Short-Term Memory, a deep learning approach for pedestrian trajectory prediction in crowded environments. The approach integrates a new dynamic distance-based loss function into Social Long Short-Term Memory, enhancing collision avoidance without compromising displacement accuracy. The method is trained and evaluated on a heterogeneous density dataset and four homogeneous density datasets, covering various crowd-density levels. Experimental results show that the proposed approach outperforms baseline methods in reducing collision rates without decreasing displacement accuracy and, in most cases, even improving it.

**Keywords:** Deep Learning, Social LSTM, Pedestrian Trajectory Prediction, Collision Avoidance

## 1 Introduction

Predicting pedestrian trajectories in crowded environments is crucial for engineering applications such as intelligent transportation, robot navigation, urban planning, and event management. While deep learning has advanced trajectory forecasting, many models still struggle to prevent overlaps and collisions in high-density scenarios. A recent deep learning-based method attempted to address this issue [1], but it remains ineffective in high-density crowds, where it forecasts overlapping trajectories. This highlights the need for a model that can prevent collisions without increasing displacement error across varying crowd densities.

## 2 Proposed Approach

To address the above gap, this study proposes an approach that integrates Social Long Short-Term Memory with a novel dynamic distance-based loss function. The new loss function plays a critical role in enhancing Social Long Short-Term Memory's ability to learn collision avoidance without compromising distance-based performance in real-world crowded scenarios. This loss function ($\text{loss} = \text{ADE} + \lambda \times \text{DDC}$) combines the Average Displacement Error (ADE) with a Dynamic Distance-based Collision penalty (DDC). The Average Displacement Error helps maintain or improve the performance of distance-based metrics, while the Dynamic Distance-based Collision penalty prevents close proximity between individuals to reduce collision occurrences. Here, $\lambda$ is a predefined weight that controls the influence of the collision penalty on the overall loss function. The Dynamic Distance-based Collision penalty quantifies the total overlap distances between individuals at each time step. It dynamically adjusts a circular personal space around each pedestrian based on their ground-truth future trajectories. The radius of this space is adaptively modified according to the actual crowd density, ensuring that the model penalizes potential collisions while preserving natural movement behavior.

* Presenting author.

## 3 Preliminary Evaluation and Results

The proposed approach was evaluated against two Social Long Short-Term Memory-based methods: one optimizing distance-based accuracy [2], serving as the baseline, and another integrating Time to Collision alongside distance-based performance [1]. All models were implemented using the TrajNet++ framework [3] with hyperparameters from Ref [1]. The training and evaluation were conducted on five pedestrian trajectory datasets collected from the Festival of Lights in Lyon [1], including a heterogeneous density dataset and four homogeneous density datasets categorized by density based on the classification in Ref [1]. Performance was assessed using Average Displacement Error and Final Displacement Error (FDE) for accuracy, while the prediction Collision Rate metric (CR) [3] measured collision occurrences. As shown in Table 1, the proposed approach reduced collisions in heterogeneous, high, and very high density datasets while improving distance-based performance. In contrast, Time to Collision-based Social Long Short-Term Memory failed to reduce collisions without increasing displacement error. Both models successfully reduced collisions in low density and medium density datasets; however, our approach outperformed the Time to Collision-based model, achieving a 14.6% greater reduction in low density and 8.3% in medium density datasets.

**Table 1**: erformance comparison across a heterogeneous dataset and four homogeneous datasets (Low Density, Medium Density, High Density, and Very High Density). 'Diff.' represents the difference relative to the baseline model. 'Heterog.' refers to the heterogeneous dataset.

| **Model** | **Heterogen.** | | **Low Density** | | **Medium Density** | | **High Density** | | **Very High Density** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Value** | **Diff.** | **Value** | **Diff.** | **Value** | **Diff.** | **Value** | **Diff.** | **Value** | **Diff.** |
| **The baseline [2]** | | | | | | | | | | |
| ADE (m) | 0.257 | – | 0.499 | – | 0.345 | – | 0.241 | – | 0.259 | – |
| FDE (m) | 0.473 | – | 0.949 | – | 0.671 | – | 0.418 | – | 0.456 | – |
| CR (%) | 39.3 | – | 40.62 | – | 29.21 | – | 33.8 | – | 51.29 | – |
| **Time to Collision [1]** | | | | | | | | | | |
| Optimal $\lambda$ | 0.001 | – | 0.001 | – | 0.002 | – | 0.001 | – | 0.002 | – |
| ADE (m) | 0.288 | 0.031 ↑ | 0.469 | −0.03 ↓ | 0.307 | −0.038 ↓ | 0.251 | 0.01 ↑ | 0.319 | 0.06 ↑ |
| FDE (m) | 0.532 | 0.06 ↑ | 0.904 | −0.045 ↓ | 0.549 | −0.122 ↓ | 0.435 | 0.017 ↑ | 0.577 | 0.121 ↑ |
| CR (%) | 26.1 | −13.2 ↓ | 37.5 | −3.12 ↓ | 20.6 | −8.61 ↓ | 19.3 | −14.5 ↓ | 38.7 | −12.59 ↓ |
| **The proposed approach** | | | | | | | | | | |
| Optimal $\lambda$ | 0.003 | – | 0.01 | – | 0.01 | – | 0.001 | – | 0.002 | – |
| ADE (m) | 0.248 | −0.01 ↓ | 0.463 | −0.036 ↓ | 0.323 | −0.022 ↓ | 0.239 | −0.002 ↓ | 0.238 | −0.021 ↓ |
| FDE (m) | 0.445 | −0.028 ↓ | 0.876 | −0.073 ↓ | 0.621 | −0.05 ↓ | 0.413 | −0.005 ↓ | 0.42 | −0.036 ↓ |
| CR (%) | 29.9 | −9.4 ↓ | 22.9 | −17.72 ↓ | 12.3 | −16.91 ↓ | 25.5 | −8.3 ↓ | 47.4 | −3.89 ↓ |

## References

[1] Korbmacher, Raphael, and Antoine Tordeux. "Toward better pedestrian trajectory predictions: the role of density and time-to-collision in hybrid deep-learning algorithms." Sensors 24, no. 7 (2024): 2356.

[2] Alahi et al., 'Social LSTM: Human trajectory prediction in crowded spaces,' Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 961-971, 2016.

[3] Kothari, Parth, Sven Kreiss, and Alexandre Alahi. "Human trajectory forecasting in crowds: A deep learning perspective." IEEE Transactions on Intelligent Transportation Systems 23, no. 7 (2021): 7386-7400.