

Article

Not peer-reviewed version

Towards Better Predictive Models: The Role of Density in Pedestrian Trajectory Predictions

[Raphael Korbmacher](#)^{*} and [Antoine Tordeux](#)

Posted Date: 28 February 2024

doi: 10.20944/preprints202402.1646.v1

Keywords: Pedestrian trajectory prediction, deep learning, pedestrian trajectory dataset, density-based classification, collision avoidance



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Towards Better Predictive Models: The Role of Density in Pedestrian Trajectory Predictions

Raphael Korbmacher * and Antoine Tordeux

Department for Traffic Safety and Reliability, University of Wuppertal, 42119 Wuppertal, Germany

* Correspondence: korbmacher@uni-wuppertal.de

Abstract: Predicting human trajectories poses a significant challenge due to the complex interplay of pedestrian behavior, which is influenced by environmental layout and interpersonal dynamics. This complexity is further compounded by variations in scene density. To address this, we introduce a novel dataset from the Festival of Lights in Lyon 2022, characterized by a wide range of densities (0.2–2.2 *ped/m*²). Our analysis demonstrates that density-based classification of data can significantly enhance the accuracy of predictive algorithms. We propose an innovative two-stage processing approach, surpassing current state-of-the-art methods in performance. Additionally, we utilize a collision-based error metric to better account for collisions in trajectory predictions. Our findings indicate that the effectiveness of this error metric is density-dependent, offering prediction insights. This study not only advances our understanding of human trajectory prediction in dense environments but also presents a methodological framework for integrating density considerations into predictive modeling, thereby improving algorithmic performance and collision avoidance.

Keywords: pedestrian trajectory prediction; deep learning; pedestrian trajectory dataset; density-based classification; collision avoidance

1. Introduction

The challenge of predicting pedestrian trajectories has emerged as a pivotal challenge in recent years. This surge in interest is largely attributed to the profound implications it holds for autonomous vehicle navigation [1], service robot deployment [2], and the strategic planning of infrastructure and mass gatherings [3]. In addressing these intricate challenges, researchers have traditionally employed physics-based (PB) models to simulate and understand pedestrian behavior. These models have been instrumental in dissecting collective phenomena and enhancing our understanding of pedestrian dynamics, particularly in high-density contexts relevant to crowd management and evacuation strategies [3]. However, the landscape of pedestrian trajectory prediction has witnessed a paradigm shift over the last decade with the advent and integration of deep learning (DL) algorithms [4]. Despite the opaqueness of these models in terms of interpretability, their superiority in mirroring observed trajectories has been markedly pronounced, especially when juxtaposed with their PB counterparts [5]. Nonetheless, it is important to acknowledge that the domains of applicability for PB and DL models do not entirely overlap. While PB models excel in the realm of high-density simulations, providing insights into collective behavior, DL algorithms predominantly thrive in low-density environments where individual pedestrian movements are characterized by a greater degree of freedom and intricate long-range interactions [4]. This paper introduces a novel, real-world pedestrian trajectory dataset, gathered during the Festival of Lights in Lyon. Field pedestrian trajectory datasets are typically gathered from low-density situations. In contrast, this dataset captures the nuanced dynamics of pedestrian movements across a large spectrum of density levels. It ranges from sparse crowds observed during show moments to the densely packed throngs seen after the event. Utilizing this dataset, we train DL algorithms, including Long Short-Term Memory (LSTM) networks and Generative Adversarial Networks (GAN). Our methodology is underscored by a novel approach: we harness situational classification predicated on crowd density to refine our models' learning process. This two-stage process, which initially classifies the scene based on density and then predicts the

trajectories, not only bolsters the efficiency of our models but also substantially elevates the precision of trajectory predictions. This improvement is demonstrated by comparative analyses with traditional DL algorithms and PB models.

Another challenge of trajectory prediction that we face in this paper is overlapping and colliding of predicted trajectories [6]. To tackle that problem, we integrate a time-to-collision (TTC) [7] term into the loss function of the algorithms. A parameter, λ , is utilized to modulate the TTC's influence on the training. Our empirical research uncovers a significant relationship between the optimal λ values and the density levels, highlighting the intricacies of pedestrian behavior across different densities.

The remainder of this paper is organized as follows: first, we review related studies in Section 2. Then, our novel dataset is presented in Section 3 and the methodology for the empirical work is proposed in Section 4. The results are shown in Section 5. The last Section 6 includes a discussion of the results and an outlook on future works.

2. Related work

The domain of pedestrian trajectory prediction is multifaceted, drawing insights from various disciplines and methodologies. The two main stream are the physics-based models and data-based algorithms [4]. PB models have been the cornerstone of understanding pedestrian dynamics, especially in high-density scenarios. The Social Force model (SF), introduced by Helbing and Molnar [8], exemplifies this approach, simulating pedestrian movement by balancing attractive and repulsive forces. Other famous PB models are the Optimal Reciprocal Collision Avoidance (ORCA) from Van den Berg et al. [9] or the cellular automata model from Burstedde et al. [10]. However, these models are not without their challenges, particularly when it comes to encapsulating the full range of crowd behavior [4]. For more PB models see reviews like [11–13].

In pursuit of addressing these limitations, the research frontier has gradually shifted towards data-driven methodologies. Notably, the past decade has witnessed a burgeoning interest in DL approaches. Pioneering works like the *Social LSTM* from Alahi et al. [5] introduce the use of Recurrent Neural Networks (RNN), specifically LSTM networks, in conjunction with a novel concept known as *Social Pooling*. This innovative approach incorporates neighbouring information, thereby enriching the model's contextual understanding. This social concept was further enhanced by Gupta et al. [14] through *Social GAN*, where the generative adversarial framework allowed for the generation of multiple plausible future paths, addressing the inherent uncertainty in human movement. Alternative methods employed for social predictions include attention mechanisms [15], graph-based approaches [16], and the utilization of relative coordinates [17]. Additionally, deep learning architectures such as Convolutional Neural Networks [18,19] and Transformers [20] have been applied in trajectory prediction tasks.

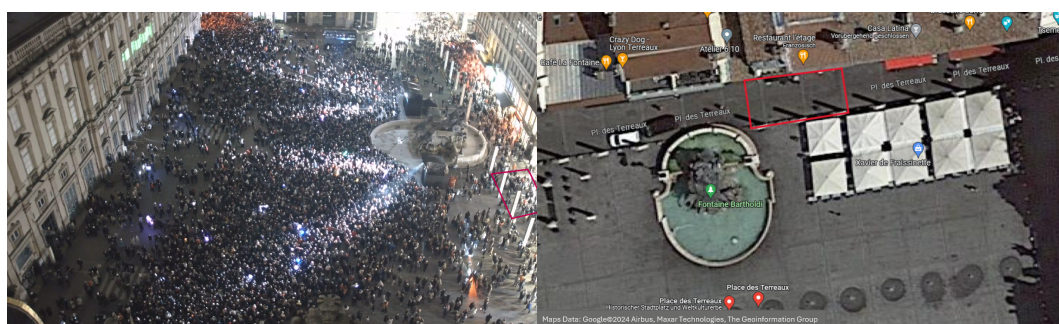
A pivotal aspect of this paper is the innovative classification of trajectory scenes based on crowd density prior to prediction. To the best of our knowledge, this approach is a novel paradigm, potentially owing to the scarcity of high-density, real-world pedestrian trajectory datasets. Xue et al. [21] predict pedestrian destinations using bidirectional LSTM classification. This involves an additional classification stage to distinguish between possible destinations of pedestrians. They classify the route manually into four distinct categories. In another paper from the authors [22] the classification is done based on a clustering algorithm. Kothari et al. [23] categorize pedestrian trajectories based on the nature of interactions observed, identifying behaviors such as collision avoidance, leader-follower dynamics, and grouping behavior. An alternative methodology involves classifying trajectories based on individual pedestrian characteristics. Papathanasopoulou et al. [24] concentrate on attributes such as age, gender, height, and speed to inform their classification. A second cornerstone of this work is the seamless integration of a PB concept, TTC, into the loss function of DL algorithms. This synthesis of PB principles and DL models is not an isolated endeavor. Alahi et al. [5] and Khadka et al. [25] utilized simulated data from PB models for training DL algorithms. Antonucci et al. [26] embedded a PB model

directly into the DL architecture. Furthermore, the works of Silvestri et al. [27] and Kothari et al. [28] stand out for their use of PB principles within the loss function to eliminate unrealistic predictions.

3. The Dataset

With a growing interest in data-based methods the significance of pedestrian trajectory data has been elevated in recent research. This area has seen a proliferation of datasets published by researchers, which can be categorized into field data and experimental data obtained in laboratory conditions. In the field studies, real-world settings are employed where individuals, unaware of their participation in a study, navigate through various scenarios. Famous field datasets are the ETH [29] and UCY [30] datasets, which are widely used in the machine learning community. Originating from surveillance videos, these datasets capture pedestrians scene of low density ($0.1\text{--}0.5 \text{ ped}/m^2$). Other field datasets are the Stanford Drone Dataset [31], the Grand Central Station Dataset [32], and the Edinburgh Informatics Forum Dataset [33]. None of these have densities above $0.2 \text{ ped}/m^2$. In the following we will present a field dataset with pedestrian densities between $0.2\text{--}2.2 \text{ ped}/m^2$.

The data was collected at Lyon's Festival of Lights. The event running for four days from 7 pm to 11 pm attracts millions (2 million in 2022) of visitors each year. Key attractions are light shows at Place des Terreaux and Place Saint-Jean. We have installed cameras at the Place des Terreaux to film the area which is represented by the red rectangle in Figure 1.



(a) View over Place des Terreaux (Lyon).

(b) View over the tracking area (red box).

Figure 1. Area for the trajectory tracking at Lyon Festival of Lights 2022.

In Figure 1a, the entirety of Place des Terreaux is depicted. The red box on the right-hand side delineates our designated tracking region. Figure 1b offers an aerial perspective of this same area. This designated zone measures 9 meters in length and 6.5 meters in width. On average, we concurrently tracked 55 pedestrians, resulting in a mean density of $0.95 \text{ ped}/m^2$. The distribution of pedestrian density exhibited significant variability. During the light show, the majority of pedestrians congregated in the central area of the square, remaining largely stationary. Consequently, the pedestrian density within our tracking zone was relatively low. However, when the show concluded—which consistently lasts approximately 9 minutes—the crowd dynamics shifted dramatically as most individuals sought to exit towards another event. In this transition phase, the density within our tracking corridor surged, often exceeding 120 pedestrians moving simultaneously. For video calibration, we meticulously established nine calibration points, ensuring the precise tracking of pedestrian trajectories using the PeTrack software [34]. Throughout our study, we recorded 5195 individual trajectories, which averaged a duration of 12.38 seconds and a mean velocity of 0.62 m/s . For the training and testing of the algorithms we need trajectories of a minimal length of 7 seconds (see 4.1). Because many trajectories are more than 14 seconds long, they can be used more than once. All in all we get 7450 trajectories for training and testing.

4. Methodology

4.1. Overview

For predicting pedestrian trajectories we have i^{th} pedestrian in a scene represented by image coordinates (x_t^i, y_t^i) for each time instant $t = k \cdot dt$, with $k \in \mathbb{N}$ und $dt = 1/3$ s the time step. The observed positions from $t = T_1$ to $t = T_{obs}$ is taken as input and the aim is to predict future trajectories from $t = T_{obs} + dt$ to $t = T_{pred}$. Every scene involves a primary pedestrian and his neighbours over the timespan T_1 to T_{pred} . A neighbours is a pedestrian whose position at T_1 is closer to the position of the primary pedestrian than a radius R . Our dataset has a framerate of three observation for each second. We choose input trajectories of 9 observations (3 sec.) and want to predict 12 timesteps (4 sec.).

The predicted trajectory of all primary pedestrians are evaluated on two commonly utilized Euclidean distance metric and a collision metric. In the first distance-based metrics, called average displacement error (ADE) [29], the distance between the predicted trajectory and the ground truth trajectory is measured at any time step t

$$\text{ADE} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \|\hat{x}_i(t) - x_i(t)\|. \quad (1)$$

$x_i(t)$ is the actual position of the i^{th} pedestrian at time t while $\hat{x}_i(t)$ is the predicted position. The Euclidean distance is denoted as $\|\cdot\|$. The second distance-based metric, called final displacement error (FDE) [30] displays the distance between the final point $t = T_{pred}$ of the predicted trajectory and the ground truth trajectory

$$\text{FDE} = \frac{1}{N} \sum_{i=1}^N \|\hat{x}_i(T) - x_i(T)\|. \quad (2)$$

These distance-based metrics are widely used in pedestrian trajectory predictions for their effectiveness in quantifying the goodness-of-fit. However, repulsive forces, which are pivotal in shaping interactions between pedestrians, are not taken into account [35]. Consequently, these metrics do not account for potential overlaps or collisions between pedestrians. Therefore the collision metric is used to enhance the evaluating process

$$\text{COL} = \frac{1}{|S|} \sum_{\hat{Y} \in S} \text{COL}(\hat{Y}), \quad (3)$$

with

$$\text{COL}(\hat{Y}) = \min \left(1, \sum_{t=1}^T \sum_{i=1}^N \sum_{j>i}^N [\|\hat{x}_i(t) - \hat{x}_j(t)\| \leq 2R] \right). \quad (4)$$

S includes all scenes in the test set, $\hat{Y}$ represents a scene prediction containing N agents, and $\hat{y}_i$ is the prediction of agent i over the prediction time of T , while $[\cdot]$ is the Iverson bracket

$$[P] = \begin{cases} 1 & \text{if } P \text{ is true,} \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

This metric counts a prediction as a collision when a predicted pedestrian trajectory intersects with neighboring trajectories, thus indicating the proportion of predictions where collisions occur. A vital factor in this calculation is the chosen pedestrian size R . An increase in R will likewise increase the number of collisions.

4.2. Prediction approaches

In the subsequent trajectory predictions, various trajectory prediction approaches ranging from traditional PB models to modern DL algorithms are chooses as benchmarks for comparison

with our two-stage approach. We present the results of the Constant Velocity model (CV) and SF model [8] as well as the results of a Vanilla LSTM, the Social LSTM (SLSTM) [5] and the Social GAN [14]. The position and velocity of the i^{th} pedestrian are denoted as $x_i \in \mathbb{R}^2$ and $v_i \in \mathbb{R}^2$, respectively. For a system of N pedestrians, the position and velocity vectors, $x = (x_1, \dots, x_N)$ and $v = (v_1, \dots, v_N)$, have dimensions of $2N$. All variables, including $x(t)$ and $x_i(t)$, are functions of time t .

4.2.1. Constant Velocity Model

The CV model assumes pedestrian velocities remain unchanged over time. It serves as a baseline for more complex models. The future position of a pedestrian is predicted as:

$$x_i(t + t_p) = x_i(t) + t_p v_i(t), \quad \forall t_p \in [0, T_p]. \quad (6)$$

4.2.2. Social Force Model

Introduced by Helbing and Molnar [8], the SF model treats pedestrians as particles influenced by forces. Within this framework, the model calculates acceleration based on the cumulative effect of three distinct forces, as delineated in equation 7

$$m_i \frac{dv_i}{dt} = m_i \frac{v_i^0 - v_i}{\tau} + \sum_{j \neq i} \nabla U(x_j - x_i) + \sum_W \nabla V(x_W - x_i) \quad (7)$$

Here, m_i , v_i , and v_i^0 signify the mass, current velocity, and desired velocity of pedestrian i . The term $\nabla U(x_j - x_i)$ represents the repulsive force from other pedestrians, while $\nabla V(x_W - x_i)$ indicates the repulsive force from obstacles. The potential functions $U(d)$ and $V(d)$ are given by:

$$U(d) = A B e^{-|d|/B}, \quad A, B > 0 \quad \text{and} \quad V(d) = A' B' e^{-|d|/B'}, \quad A', B' > 0 \quad (8)$$

The first term of Equation 7 signifies the driving force experienced by the i^{th} pedestrian. This force propels the individual towards their desired speed and direction within a relaxation time $\tau > 0$. The second term encapsulates the summation of social forces, originating from the repulsive effects as pedestrians endeavour to maintain a comfortable distance from one another. The third term accounts for the aggregate interaction forces between pedestrian i and various obstacles.

Whereas the CV model has no parameter, the SF model has three parameters, preferred velocity, interaction potential, and reaction time, that can be optimized to get accurate predictions.

4.2.3. Vanilla LSTM

LSTM networks, a class of RNN designed to learn long-term dependencies, have proven effective in handling sequential data, particularly for time series prediction tasks. Introduced by Hochreiter and Schmidhuber [36], LSTMs address the vanishing and exploding gradient problems common in traditional RNNs, making them suitable for complex sequence modeling tasks such as trajectory prediction. The vanilla LSTM model considers historical trajectories to predict future positions.

$$x_i(t + t_p) = x_i(t) + \text{LSTM}(t_p, (x_i(t - t_o), t_o \in [0, T_o])), \quad \forall t_p \in [0, T_p]. \quad (9)$$

Here, $x_i(t + t_p)$ predicts the future trajectory of a pedestrian i at time $t + t_p$, based on its past positions $x_i(t - t_o)$, over an observation window $t_o \in [0, T_o]$.

4.2.4. Social LSTM

LSTM networks have demonstrated effective performance in sequence learning tasks. One such task, the prediction of pedestrian trajectories, presents the challenges, that the trajectory of a pedestrian

can be significantly influenced by the trajectories of surrounding pedestrians. The number of these neighboring influences can fluctuate widely, especially in densely crowded environments [37].

Enhancing the LSTM framework, the SLSTM by Alahi et al. [5] incorporates a social pooling layer, enabling the model to consider the influence of neighboring pedestrians explicitly. This is a key distinction from the Vanilla LSTM, reflecting the model's capacity to capture social interactions:

$$x_i(t + t_p) = x_i(t) + \text{SLSTM}(i, t_p, (x(t - t_o), t_o \in [0, T_o])), \quad \forall t_p \in [0, T_p]. \quad (10)$$

In this formulation, the inclusion of the index i and the collective pedestrian state x emphasizes the model's attention to the surrounding pedestrians' trajectories, making it adept at handling complex social behaviors in dense scenarios.

4.2.5. Social GAN

Another approach we take into account is the Social GAN (SGAN) introduced by Gupta et al. [14]. This model extends traditional approaches by incorporating GANs to predict future trajectories. GANs, conceptualized by Goodfellow et al. [38], consist of two competing networks: a Generator, which generates data samples, and a Discriminator, which evaluates the authenticity of the samples against real data. SGAN leverages this architecture to generate plausible future trajectories of pedestrians, addressing the complex dynamics of pedestrian movement in crowded spaces. A key feature of the SGAN model is its pooling mechanism, which processes the relative positions of pedestrians to each other. This mechanism is crucial for understanding the social interactions and dependencies among individuals in crowded environments

$$x_i(t + t_p) = x_i(t) + \text{SGAN}(i, t_p, (x(t - t_o), t_o \in [0, T_o])), \quad \forall t_p \in [0, T_p]. \quad (11)$$

4.3. Two-stage process

The foundation of our innovative classification framework lies in its capacity to predict trajectories across varying density levels, marking a departure from traditional models that typically utilize a single algorithm to process a wide array of scenarios within a dataset. Our strategy entails segmenting the dataset according to the density of each scene, thereby generating distinct subsets. At the inception of our methodology, we establish well-defined criteria for classification. This process is underpinned by two distinct methodologies: a statistical analysis and a review of existing literature. The results of this clustering process are depicted in Figure 2. These figures visually represent each dataset item as a point measurement each second, with the left side of Figures 2 illustrating points based on their average density

$$\rho(t) = \frac{N(t)}{A}, \quad (12)$$

and average velocity

$$\bar{v}(t) = \frac{1}{N(t)} \sum_{i \in S(t)} v_i(t), \quad (13)$$

and the right side based on their average density and flow, the flow being the product of the density by the average velocity

$$J(t) = \rho(t) \bar{v}(t). \quad (14)$$

In these illustrations, the use of varied colors signifies the identification of distinct clusters. Additionally, to corroborate these findings, a hierarchical cluster analysis was undertaken, which yielded analogous outcomes.

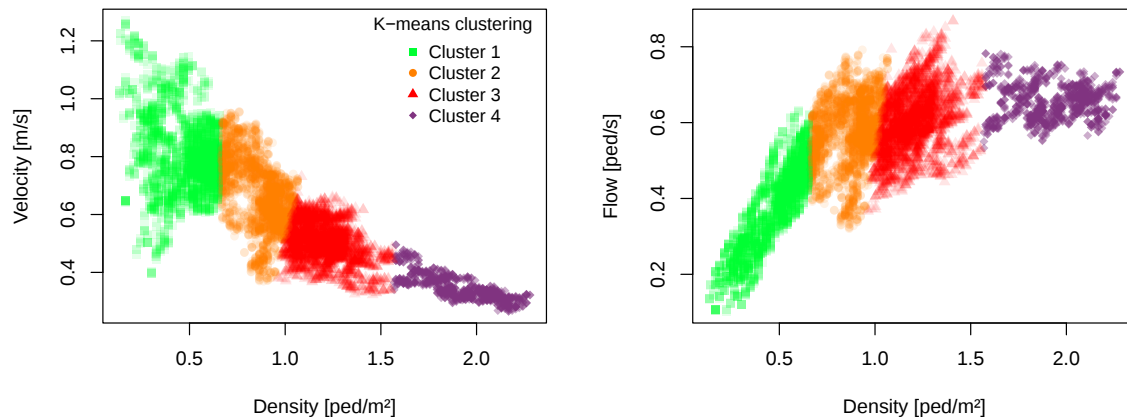


Figure 2. Results of the K-Means clustering. Trajectory scenes are clustered as shown by the different colors of the points.

We can see clear vertical colour switch's of the points at densities around 0.7 , 1.1 , and 1.6 ped/m^2 . Remarkably, without presetting the number of clusters or explicitly focusing on density levels, the K-Means algorithm autonomously reveals density-dependent clustering. The delineation of clusters and their boundary values align closely with those identified in the literature. Stefan Holl [39] delineates critical density thresholds for various infrastructures, signifying points at which pedestrian behavior undergoes significant changes. According to Holl, densities below 0.7 ped/m^2 indicate a free flow state, densities below 1.3 ped/m^2 represent a bound flow, and values above 1.3 ped/m^2 are indicative of congested flow. In our model, we refine these categories by slightly narrowing the bound flow range and subdividing the congested flow category into two distinct segments.

Figure 12 illustrates the procedural steps undertaken to evaluate our proposed methodology and within the sizes of the clusters, which are taken from Figure 2.

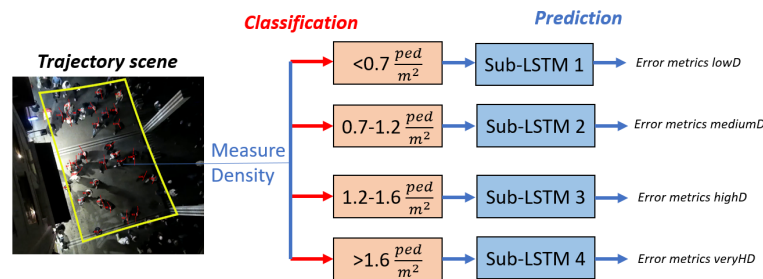


Figure 3. Schemata of our two-stage prediction approach

New trajectory scene are given to our framework, where the initial step involves calculating the scene's density using (12) in individuals per square meter (ped/m^2), N represents the total number of pedestrians observed within the scene, and A is the scene's total area in square meters.

The density categorization is as follows: scenes with a density below 0.7 ped/m^2 are labelled as lowD; densities ranging from 0.7 to 1.2 ped/m^2 are classified as mediumD; densities between 1.2 and 1.6 ped/m^2 are designated as highD; and densities exceeding 1.6 ped/m^2 are identified as veryHD. Following classification, the scene is bifurcated into two segments: the initial segment spans 9 timesteps and serves as input for one of the four specialized Sub-LSTMs, while the subsequent segment, encompassing 12 timesteps, is utilized to appraise the LSTMs' performance through the

computation of error metrics ADE, FDE and COL.

Figures 4a to 4d provide illustrative examples of each density level encountered in our dataset.



In Figure 4a, the scene exhibits very low density, with pedestrian movement primarily from two directions, leading to numerous interactions and avoidance behaviors. This is characteristic of our lowD data. Conversely, Figure 4b showcases a moderately higher density, yet still affords space for interactions, avoidance, and bidirectional pedestrian flow. In Figure 4c, representing highD data, the dynamics of pedestrian movement markedly differ from those observed in lowD and mediumD scenes, with movement predominantly unidirectional from the top, indicating a tendency to follow the pedestrian ahead. This pattern is even more pronounced in Figure 4d, where the flow from the top is so dense that passage from the bottom becomes challenging, leading pedestrians to follow the leader with limited freedom of movement and space. These observed behavioral differences underpin our classification rationale.

4.4. Collision weight

Predictions of pedestrian trajectories presents the challenge to predict trajectory paths that do not collide with neighbours. Accurately measuring these collisions is challenging due to the shapes of pedestrians, which can vary from person to person. Traditionally, collisions are defined by the overlap of the radii of two pedestrians, as delineated in Equations 3 to 5. However, this method proves sub optimal for inclusion as a penalty function within the loss function of DL algorithms. Analysis of pedestrian trajectory data frequently reveals instances where collisions are not genuine but rather instances of grouping behavior, with individuals walking closely, sometimes shoulder-to-shoulder. It is not these interactions we aim to deter, but rather scenarios in which individuals move directly towards one another without any attempt to avoid collision—behaviors that are unrealistic and undesirable. To address this, we adopt the TTC concept, a widely recognized principle in the study of pedestrian dynamics [7]. Implementing this variable in the loss function of an DL algorithm would reduce predicted situations, where pedestrians walk straight towards each other without avoidance mechanism. Integrating TTC into a DL algorithm's loss function significantly mitigates predictions where pedestrians are on a direct collision course without any avoidance mechanisms. The TTC term calculates the time until two pedestrians would collide if they continue moving at their current velocities, a concept validated by Karamouzas et al. [7]. The relative position and velocity between the pedestrian i and j can be denoted by $p_{ij} = (x_i - x_j, y_i - y_j)$ and $v_{ij} = (v_{x_i} - v_{x_j}, v_{y_i} - v_{y_j})$, respectively. A collision between pedestrian i and pedestrian j occurs if a ray, originating from (x_i, y_i) and extending in the direction of v_{ij} , intersects the circle centred at (x_j, y_j) with a radius of $R_i + R_j$ at some time τ_{ij} in the future. This condition can be mathematically represented as $\|p_{ij} + v_{ij} \cdot t\|^2 < (R_i + R_j)^2$ where $\|\cdot\|$ denotes Euclidean norm. Solving this quadratic inequality for t yields τ_{ij} as the smallest positive root:

$$\tau_{ij} = \frac{-p_{ij} \cdot v_{ij} - \sqrt{(p_{ij} \cdot v_{ij})^2 - \|v_{ij}\|^2(\|p_{ij}\|^2 - (r_i + r_j)^2)}}{\|v_{ij}\|^2} \quad (15)$$

A collision is imminent when $\tau_{ij} = 0$, whereas a large positive value for τ_{ij} indicates no collision risk. To implement τ_{ij} into the loss function we have to use an sigmoid function f that has high values, if τ_{ij} is low and vice versa:

$$f(\tau) = \frac{1}{1 + e^{s(\tau - \delta)}}, \quad (16)$$

where $s = 10$ and $\delta = 0.4$ are slope and threshold parameters, respectively. This function is then integrated into the loss function, traditionally focused solely on minimizing the ADE. The revised loss function combines ADE with TTC loss, optimized through minimization:

$$\mathbf{L}_i = \sum_{t=1}^T \|x_i(t) - \hat{x}_i(t)\| + \lambda \sum_{t=1}^T f(\min_{j \neq i} \{\tau_{ij}\}), \quad (17)$$

where $\lambda > 0$ modulates the influence of the TTC component in their loss function. The calculation of τ_{ij} considers all nearby pedestrians to the primary pedestrian, employing the minimum τ_{ij} to identify and mitigate the most critical potential collision scenario in the model.

4.5. Implementation details

The algorithms are implemented in the commonly accepted configurations of related contributions [23]. All computations are performed using the PyTorch framework. The learning rate is set to 0.001 and an ADAM optimizer is utilized. The batch size is set to 8 and training is carried out for 15 epochs, if not the early stop mechanism interrupt. This is the case, when the validations error starts to rise for three epochs. For validation and testing, a hold-out validation strategy is adopted by allocating 15% of the dataset for each validation and testing, while the remaining data serves as

the training set. For capturing pedestrian interactions, we choose a circles with a radius of $R = 4.5$ m surrounding the primary pedestrian.

5. Results

We will unveil the predictive outcomes of our dataset using two distinct yet synergistic methods. Initially, we will showcase the performance of our two-stage prediction framework, comparing it with contemporary state-of-the-art methodologies. Subsequently, we will demonstrate the seamless integration of our two-stage process with the incorporation of the TTC term into the loss function, illustrating its efficacy in mitigating collision instances.

5.1. Two-Stage Predictions

The results of our predictions will be presented in Table 1. As described in Section 4.3 we evaluated the predictions on the different density levels lowD, mediumD, highD, and veryHD. For every approach we measure ADE, FDE and COL metrics.

Table 1. Quantitative comparison of ADE and FDE metrics for articles using Social LSTM as benchmark with different datasets.

Model	LowD		MediumD		HighD		VeryHD	
	ADE/FDE	COL	ADE/FDE	COL	ADE/FDE	COL	ADE/FDE	COL
CV	0.71/0.97	54.76	0.85/0.98	45.73	0.53/0.8	62.35	0.44/0.67	81.74
Social Force [8]	0.78/1.33	24.4	0.55/0.89	31.16	0.5/0.82	36.43	0.36/0.63	54.78
Vanilla LSTM	0.5/0.99	31.55	0.33/0.63	37.69	0.29/0.52	36.43	0.24/0.41	63.8
Social LSTM [5]	0.53/1.02	57.74	0.37/0.73	59.3	0.41/0.78	64.26	0.35/0.66	75.37
Social GAN [14]	0.53/0.99	31.36	0.39/0.72	32.16	0.36/0.61	32.33	0.25/0.41	55.94
Our 2stg. SLSTM	0.48/0.93	30.95	0.3/0.63	36.18	0.26/0.4	42.02	0.24/0.41	52.23
Our 2stg. SGAN	0.44/0.83	32.74	0.27/0.52	40.2	0.28/0.5	35.33	0.26/0.43	58.6
Our 2stg. TTC-SLSTM	0.39/0.73	29.17	0.3/0.62	22.61	0.23/0.36	36.29	0.24/0.41	52.23

The initial insight gleaned from Table 1 reveals a clear trend: as density increases, the COL metric rises while the ADE/FDE diminish. This pattern emerges because higher densities naturally lead to reduced distances between individuals, consequently resulting in increased overlaps among agents. Additionally, it’s observed that velocities decrease as density intensifies, leading to trajectories that are shorter in spatial extent. This reduction in travel distance directly contributes to the observed decrease in both ADE and FDE metrics at higher densities. Furthermore, it is clear that the DB algorithm outperform the traditional models CV and SF in terms of ADE/FDE. In terms of COL metric SF performs very well.

In the last three rows of Table 1, we present the effectiveness of our two-stage approach and its combination with the TTC term. The results clearly show a significant improvement in the algorithm’s precision, attributed to the strategy of classification before prediction. Our enhanced two-stage SLSTM model consistently outperforms the traditional SLSTM across all evaluated datasets, demonstrating superior performance in terms of ADE, FDE, and COL metrics. Similarly, our adapted SGAN model shows marked improvements over the standard SGAN in three out of four datasets with respect to ADE. Integrating the TTC term further enhances the SLSTM results, notably in reducing collisions. A more detailed discussion on this enhancement is provided in the subsequent Section 5.2.

5.2. Collision weight

In this study, we propose to integrate TTC in the loss function with the two-stage approach outlined in Section 4.3. As described in Equation 17 the collision part in the loss function can be adjusted by a parameter λ [35]. If λ is high, the impact of the TTC term is high compared to the impact

of ADE and vice versa. In the following diagram, the impact of different values of λ on the prediction accuracy of the SLSTM algorithm is displayed. In Figure 5a for lowD data in Figure 5b for mediumD data, in Figure 5c for highD data, and in Figure 5d for veryHD data. The first observation in each Figures is for $\lambda = 0$, which means, that it is equivalent to the value of our two-stage SLSTM in Table 1.

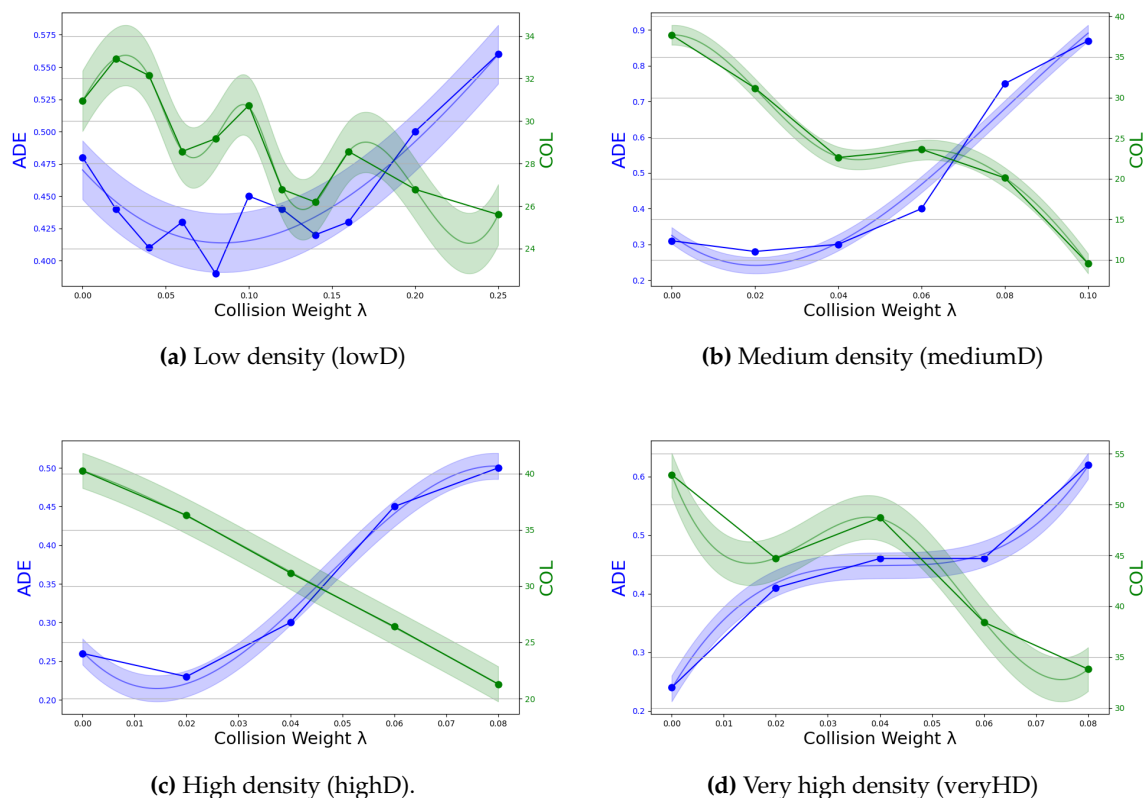


Figure 5. ADE and COL metrics for the two-stage SLSTM algorithm according to the collision weight λ for the four density levels.

Across all figures, a consistent pattern emerges: increasing the collision weight λ generally results in fewer collisions. Optimal predictions occur at a specific λ value, where ADE and FDE are equivalent to or lower than those at $\lambda = 0$. Each dataset exhibits a maximum effective λ value beyond which ADE sharply increases. In the lowD dataset (Figure 5a), the ideal λ is 0.08, reducing ADE by 19% and collisions by 6%. Values slightly higher than 0.08 are still beneficial, yielding fewer collisions and enhanced avoidance behavior, but λ values exceeding 0.16 lead to a significant increase in ADE. In the mediumD dataset (Figure 5b), the optimal λ is 0.04, reducing ADE and collisions by 3% and 40%, respectively. Here, a λ value above 0.06 results in increased ADE, although a λ of 0.1 reduces the collision metric by 75%. In the highD dataset, improvements in ADE are marginal too, only notable at a λ of 0.02. However, the collision metric significantly decreases, by up to 48%, at a λ of 0.08. Conversely, in the very high-density (veryHD) dataset, increasing collision weight initially results in a rise in ADE, with no subsequent decrease. While the collision metric decreases by 37% at $\lambda = 0.08$ values, there's no improvements for ADE. These empirical observations lead to the insight, that pedestrian behavior at different densities is very different and need different parameter configurations. At lower densities our TTC term can improve overall accuracy (ADE and COL) in higher densities we can only reduce COL, by taking higher ADE into account.

6. Conclusions

Pedestrian behavior is inherently complex, exhibiting a wide variety of patterns across different contexts. This paper introduced a novel pedestrian trajectory dataset, characterized by its diversity

in situational contexts, including varying densities and motivations. Our analysis of the data reveals variations in pedestrian behaviors correlating with the density of the scene. To address these variations, we propose a novel two-stage classification and prediction process. This approach first classifies scenes based on density and then applies the suitable model for predicting behavior within that specific density context. Implementing this framework enhanced the prediction accuracy of two famous DL algorithms, Social LSTM and Social GAN.

Further, we integrated a TTC based term into the loss function of the SLSTM to improve avoidance behaviors, consequently reducing potential collisions. Our empirical studies indicate that the effectiveness of the TTC-based term varies with density; it significantly benefits scenarios of low density by correlating higher TTC values with reduced collision incidents. However, the outcomes in high-density situations were more ambiguous, suggesting a nuanced impact of density on the efficacy of this approach. This observation could be attributed to the nuanced dynamics of pedestrian behavior across different densities. Specifically, in environments with lower densities, pedestrians tend to navigate more through avoidance and interactions, making TTC particularly relevant. Conversely, in higher density settings, pedestrian movement is more characterized by forced leader-follower dynamics, diminishing the prominence of TTC in explaining behavior. This study underscores the complexity of pedestrian behavior, which varies significantly under different environmental conditions. It highlights the necessity of adopting a flexible modeling approach to accurately predict pedestrian trajectories in diverse settings.

This research opens several avenues for future investigation in the field of pedestrian trajectory prediction, especially concerning heterogeneous datasets characterized by variable densities. Current methodologies typically rely on a one-size-fits-all model for behavior prediction across all conditions. We advocate for the development and application of multiple specialized models, each tailored to different scene characteristics, with scene density being a pivotal factor. While focusing on density has proven to be a successful strategy, exploring additional factors could yield further improvements. For instance, our methodology utilized an estimate of overall scene density. However, pedestrians do not take global densities for their decision-making into account, but rather local densities. Wirth et al. [40] demonstrate that pedestrian decisions are primarily influenced by their visual neighborhood. Future studies should investigate the impact of assessing local density variations within a scene, which could be particularly beneficial in environments exhibiting a wide range of density levels. This direction could unlock new dimensions of accuracy and reliability in trajectory prediction models.

Moreover, incorporating the TTC concept into the loss function has shown promise in enhancing prediction accuracy at lower density levels. Future research should explore alternative loss functions, particularly for high-density scenarios, where the traditional ADE based approaches may not suffice. Investigating other metrics that could more effectively capture the complexities of high-density pedestrian behavior is crucial for advancing the field.

Author Contributions: Conceptualization, R.K. and A.T.; methodology, R.K. and A.T.; software, R.K.; validation, R.K. and A.T.; formal analysis, R.K.; investigation, R.K.; resources, R.K.; data curation, R.K.; writing—original draft preparation, R.K.; writing—review and editing, R.K. and A.T.; visualization, R.K.; supervision, A.T.; project administration, A.T.; funding acquisition, A.T. All authors have read and agreed to the published version of the manuscript.

Funding: 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.

Data Availability Statement: The data is available on the following website: <https://madras-data-app.streamlit.app/>. For more information see: <https://www.madras-crowds.eu/>

Acknowledgments: We would like to express our gratitude to Professor Armin Seyfried for his engaging and insightful discussions, which have borne fruit in the form of valuable ideas.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DL	Deep learning
PB	Physics-based
LSTM	Long Short-Term Memory
SLSTM	Social Long Short-Term Memory
GAN	Generative Adversarial Network
SGAN	Social Generative Adversarial Network
TTC	Time-to-collision
SF	Social Force model
CV	Constant Velocity model
ORCA	Optimal Reciprocal Collision Avoidance
RNN	Recurrent Neural Network
ADE	Average displacement error
FDE	Final displacement error
COL	Collision metric
lowD	Low density
mediumD	Medium density
highD	High density
VeryHD	Very high density

References

1. Poibrenski, A.; Klusch, M.; Vozniak, I.; Müller, C. M2p3: multimodal multi-pedestrian path prediction by self-driving cars with egocentric vision. *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, 2020, pp. 190–197.
2. Scheggi, S.; Aggravi, M.; Morbidi, F.; Prattichizzo, D. Cooperative human-robot haptic navigation. 2014 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2014, pp. 2693–2698.
3. Boltes, M.; Zhang, J.; Tordeux, A.; Schadschneider, A.; Seyfried, A. Empirical results of pedestrian and evacuation dynamics. *Encyclopedia of complexity and systems science* **2018**, 16, 1–29.
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. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 961–971.
6. Dang, H.T.; Korbmacher, R.; Tordeux, A.; Gaudou, B.; Verstaevel, N. TTC-SLSTM: Human trajectory prediction using time-to-collision interaction energy. 2023 15th International Conference on Knowledge and Systems Engineering (KSE). IEEE, 2023, pp. 1–6.
7. Karamouzas, I.; Skinner, B.; Guy, S.J. Universal power law governing pedestrian interactions. *Physical review letters* **2014**, 113, 238701.
8. Helbing, D.; Molnar, P. Social force model for pedestrian dynamics. *Physical Review E* **1995**, 51, 4282.
9. Van Den Berg, J.; Guy, S.J.; Lin, M.; Manocha, D. Reciprocal n-body collision avoidance. *Robotics Research: The 14th International Symposium ISRR*. Springer, 2011, pp. 3–19.
10. Burstedde, C.; Klauck, K.; Schadschneider, A.; Zittartz, J. Simulation of pedestrian dynamics using a two-dimensional cellular automaton. *Physica A: Statistical Mechanics and its Applications* **2001**, 295, 507–525.
11. Bellomo, N.; Dogbe, C. On the modeling of traffic and crowds: A survey of models, speculations, and perspectives. *SIAM review* **2011**, 53, 409–463.
12. Chraïbi, M.; Tordeux, A.; Schadschneider, A.; Seyfried, A. Modelling of pedestrian and evacuation dynamics. *Encyclopedia of complexity and systems science* **2018**, pp. 1–22.
13. Duives, D.C.; Daamen, W.; Hoogendoorn, S.P. State-of-the-art crowd motion simulation models. *Transportation research part C: emerging technologies* **2013**, 37, 193–209.

14. 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*, 2018, pp. 2255–2264.
15. Vemula, A.; Muelling, K.; Oh, J. Social attention: Modeling attention in human crowds. 2018 IEEE international Conference on Robotics and Automation (ICRA). IEEE, 2018, pp. 4601–4607.
16. Monti, A.; Bertugli, A.; Calderara, S.; Cucchiara, R. Dag-net: Double attentive graph neural network for trajectory forecasting. 2020 25th International Conference on Pattern Recognition (ICPR). IEEE, 2021, pp. 2551–2558.
17. Shi, X.; Shao, X.; Guo, Z.; Wu, G.; Zhang, H.; Shibasaki, R. Pedestrian trajectory prediction in extremely crowded scenarios. *Sensors* **2019**, *19*, 1223.
18. Yi, S.; Li, H.; Wang, X. Pedestrian behavior understanding and prediction with deep neural networks. *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. Springer, 2016, pp. 263–279.
19. Li, Y.; Xu, H.; Bian, M.; Xiao, J. Attention based CNN-ConvLSTM for pedestrian attribute recognition. *Sensors* **2020**, *20*, 811.
20. Yao, H.Y.; Wan, W.G.; Li, X. End-to-end pedestrian trajectory forecasting with transformer network. *ISPRS International Journal of Geo-Information* **2022**, *11*, 44.
21. Xue, H.; Huynh, D.Q.; Reynolds, M. Bi-prediction: Pedestrian trajectory prediction based on bidirectional LSTM classification. 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA). IEEE, 2017, pp. 1–8.
22. Xue, H.; Huynh, D.Q.; Reynolds, M. PoPPL: Pedestrian trajectory prediction by LSTM with automatic route class clustering. *IEEE transactions on neural networks and learning systems* **2020**, *32*, 77–90.
23. Kothari, P.; Kreiss, S.; Alahi, A. Human trajectory forecasting in crowds: A deep learning perspective. *IEEE Transactions on Intelligent Transportation Systems* **2021**, *23*, 7386–7400.
24. Papathanasopoulou, V.; Spyropoulou, I.; Perakis, H.; Gikas, V.; Andrikopoulou, E. A data-driven model for pedestrian behavior classification and trajectory prediction. *IEEE Open Journal of Intelligent Transportation Systems* **2022**, *3*, 328–339.
25. Khadka, A.; Remagnino, P.; Argyriou, V. Synthetic crowd and pedestrian generator for deep learning problems. ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020, pp. 4052–4056.
26. Antonucci, A.; Papini, G.P.R.; Bevilacqua, P.; Palopoli, L.; Fontanelli, D. Efficient prediction of human motion for real-time robotics applications with physics-inspired neural networks. *IEEE Access* **2021**, *10*, 144–157.
27. Silvestri, M.; Lombardi, M.; Milano, M. Injecting domain knowledge in neural networks: a controlled experiment on a constrained problem. *Integration of Constraint Programming, Artificial Intelligence, and Operations Research: 18th International Conference, CPAIOR 2021, Vienna, Austria, July 5–8, 2021, Proceedings 18*. Springer, 2021, pp. 266–282.
28. Kothari, P.; Alahi, A. Safety-compliant generative adversarial networks for human trajectory forecasting. *IEEE Transactions on Intelligent Transportation Systems* **2023**, *24*, 4251–4261.
29. 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. IEEE, 2009, pp. 261–268.
30. Lerner, A.; Chrysanthou, Y.; Lischinski, D. Crowds by example. *Computer graphics forum*. Wiley Online Library, 2007, Vol. 26, pp. 655–664.
31. Robicquet, A.; Sadeghian, A.; Alahi, A.; Savarese, S. Learning social etiquette: Human trajectory understanding in crowded scenes. *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VIII 14*. Springer, 2016, pp. 549–565.
32. Zhou, B.; Wang, X.; Tang, X. Understanding collective crowd behaviors: Learning a mixture model of dynamic pedestrian-agents. 2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2012, pp. 2871–2878.
33. Majecka, B. Statistical models of pedestrian behaviour in the forum. *Master's thesis, School of Informatics, University of Edinburgh* **2009**.
34. Boltes, M.; Seyfried, A. Collecting pedestrian trajectories. *Neurocomputing* **2013**, *100*, 127–133. Special issue: Behaviours in video.

35. Korbmacher, R.; Dang, H.T.; Tordeux, A. Predicting pedestrian trajectories at different densities: A multi-criteria empirical analysis. *Physica A: Statistical Mechanics and its Applications* **2024**, *634*, 129440.
36. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural computation* **1997**, *9*, 1735–1780.
37. Korbmacher, R.; Dang-Huu, T.; Tordeux, A.; Verstaev, N.; Gaudou, B. Differences in pedestrian trajectory predictions for high-and low-density situations. 14th International Conference on Traffic and Granular Flow (TGF) 2022. Springer, 2022, pp. à-paraitre.
38. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Advances in neural information processing systems* **2014**, *27*.
39. Holl, S. *Methoden für die Bemessung der Leistungsfähigkeit multidirektional genutzter Fußverkehrsanlagen*; Number FZJ-2017-00069, Jülich Supercomputing Center, 2016.
40. Wirth, T.D.; Dachner, G.C.; Rio, K.W.; Warren, W.H. Is the neighborhood of interaction in human crowds metric, topological, or visual? *PNAS nexus* **2023**, *2*, pgad118.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.