Review of Pedestrian Trajectory Prediction Methods: Comparing Deep Learning and Knowledge-Based Approaches

Raphael Korbmacher and Antoine Tordeux

Abstract— In crowd scenarios, predicting trajectories of pedestrians is a complex and challenging task depending on many external factors. The topology of the scene and the interactions between the pedestrians are just some of them. Due to advancements in data-science and data collection technologies deep learning methods have recently become a research hotspot in numerous domains. Therefore, it is not surprising that more and more researchers apply these methods to predict trajectories of pedestrians. This paper compares these relatively new deep learning algorithms with classical knowledge-based models that are widely used to simulate pedestrian dynamics. It provides a comprehensive literature review of both approaches, explores technical and application oriented differences, and addresses open questions as well as future development directions. Our investigations point out that the pertinence of knowledge-based models to predict local trajectories is nowadays questionable because of the high accuracy of the deep learning algorithms. Nevertheless, the ability of deep-learning algorithms for large-scale simulation and the description of collective dynamics remains to be demonstrated. Furthermore, the comparison shows that the combination of both approaches (the hybrid approach) seems to be promising to overcome disadvantages like the missing explainability of the deep learning approach.

Index Terms— Pedestrians, trajectory prediction, deep learning, knowledge-based models.

I. INTRODUCTION

THE prediction of future trajectories of pedestrians is a valuable but challenging task. Traditionally, researchers try to model and understand human behavior with simple rules and mechanisms that can be used to simulate realistic behavior and predict future trajectories [1], [2]. In the last decade, deep learning methods, that have the interesting property of being able to learn complex features from data alone, have caught a lot of attention in a variety of domains [3], [4]. They have led to a large number of practical applications, especially in machine perception, and have shown to outperform the prediction accuracy of the traditional models in many scientific
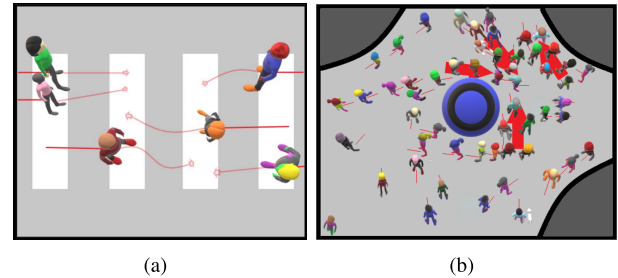
Fig. 1. For the prediction of pedestrian trajectories, it is useful to distinguish between scenes with few pedestrians (see (a)) and crowds of pedestrians (see (b)). Deep learning algorithms turn out to be accurate prediction tools for the scene (a), whereas knowledge-based approaches allow to describe collective phenomena at higher scales such as those described in the scene (b).

disciplines. Prominent examples occur in theoretical biology and active matter [5], [6], medicine [7], [8], or materials and chemical science [9], [10], [11], [12].

In the discipline of pedestrian dynamics, there is a high interest in making accurate predictions of pedestrian trajectories due to numerous real-world applications like facility, infrastructure, and building design [13], notably in the case of evacuation [14], [15], [16], autonomous driving cars [17], human-robot interactions [18], assistive technologies in industrial scenarios [19], and entertainment (e.g., augmented and virtual reality) [20].

A trajectory is defined as the time-profile of the pedestrian's position. Changes of the position in a given time step can be interpreted as the velocity. Predicting trajectories means to assess the future motion of pedestrians in a given scene. Besides the prediction of the simple physical motion of a single human, based on current or past observations, an important issue is to take into account their possible interactions as well as arising collective dynamics. Collective dynamics can be considered by predicting trajectories of many pedestrians in a scene simultaneously. In Fig. 1(a) few pedestrians are present. The challenge consists in predicting individual trajectories in local interactions with the neighbors over relatively small horizon times. In Fig. 1(b) a pedestrian crowd is shown. Predicting multiple trajectories and resulting collective crowd behavior at higher scale is necessary.

One approach to tackle the challenge of predicting trajectories is to use the knowledge-based (KB) approach, sometimes referred to by other names, such as physics-based [21], reasoning-based [22], expert-based [23] or traditional approach [24]. The KB approach contains models that are defined by basic rules or generic functions and considers
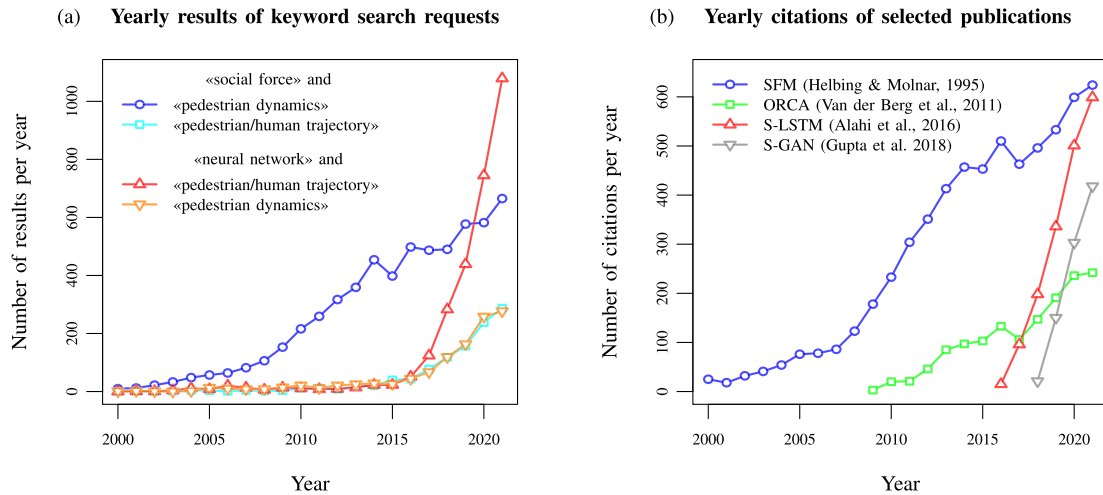
Fig. 2. Number of annual citations estimated by performing an online search with the engine Google Scholar [40]. Panel (a): Results of the tags «neural network» and «social force» in combination with either «pedestrian/human trajectory» or «pedestrian dynamics». Panel (b): Yearly citations for the articles by Helbing and Molnár [1] (SFM), Van den Berg *et al.* [25] (ORCA), Alahi *et al.* [34] (S-LSTM), and Gupta *et al.* [41] (S-GAN). SFM and ORCA are knowledge-based, whereas S-LSTM and S-GAN are deep learning approaches.

physical as well as social or psychological factors of the pedestrians. They can be specified by a few parameters which generally have physical *interpretations* and allow to adjust the model. Prominent microscopic KB models are the social force model (SFM) [1] or the optimal reciprocal collision avoidance (ORCA) [25]. In general, the goal of these models is to improve the understanding of pedestrian dynamics and to identify microscopic mechanisms and parameters for the emerge of patterns and organisation in given scenes [14], [26], [27], [28], [29], [30], [31]. These models can be successfully applied for trajectory predictions in both scenes shown in Fig. 1.

The other promising possibility to predict pedestrian trajectory consists of the use of *supervised deep learning* (DL) methods, especially long short-term memory networks (LSTM), convolutional neural networks (CNN), and generative adversarial networks (GAN). In the literature, this approach is often referred to as learning-based [22], pattern-based [32], or data-based approach. Deep learning refers to methods for training neural networks with more than two hidden layers (deep neural networks) [33] and supervised to methods that learn from data. Improving trajectory predictions for autonomous vehicles, service robots, and urban video surveillance are main goals [32]. In opposite to the KB approach, there are no interpretable parameters and rules necessary, but large amounts of data and flexible algorithms.

The KB approach is scientifically relevant since several decades, see for illustration in Fig. 2 the citation numbers of one of the most famous KB models, the social force model by Helbing and Molnár [1]. On the other hand, the DL approach is a youthful methodology that started to become highly relevant for the prediction of pedestrian trajectories after the publication of the social-LSTM from Alahi *et al.* [34]. After this date, the social force model started to get associated with pedestrian trajectory as well, see the publication number with the keywords combining social force and pedestrian trajectory in Fig. 2. Yet, the applications of KB and DL approaches are not identical. KB models are mainly designed to describe

and understand collective phenomena implying high numbers of agents in high density scenarios like in Fig. 1(b). They are based on few parameters that can be interpreted and calibrated. This results in high flexibility of the prediction scenarios, and notably the possibility to change pedestrian motion preferences (e.g., higher desired time-gap during pandemics). DL methods are designed to predict pedestrian trajectories at very local scales in space (few meters) and time (few seconds). If their predictions are accurate, they do not allow to control the motion and to reproduce different kinds of behaviors. In general, they are used for low density scenarios at local scale like in scene (a) presented Fig. 1. The question of whether DL approaches could also be suitable for large-scale simulation or successfully used to initiate collective dynamics is still open. On the other hand, the pertinence of KB models to predict local trajectories is, regarding the high accuracy of DL algorithms, questionable.

The combination of both KB and DL approaches seems to be especially promising. Recently, some authors try to implement components of the KB models in the DL algorithms to overcome crucial limitations of the DL approach, like the lack of interpretability or generality [35], [36], [37]. Furthermore, another possibility is to use DL methods to improve the accuracy of the KB models. This can be done by estimating the parameters based on the results of an DL algorithm or to implement a neural network in the KB simulations [38], [39]. These combinations are called the hybrid approach and they benefit from the strengths of both approaches and avoid their shortcomings. Other promising algorithms close to hybrid approaches rely on reinforcement learning (RL) and inverse reinforcement learning (IRL). The agents learn from their own experiences in RL methods while IRL partly rely on data.

In this article, we address a comprehensible bibliographical review of KB and DL approaches for the modeling and prediction of pedestrian trajectories. We critically compare the two modelling approaches from their technical aspects as well as their application fields. We highlight similarities and differences between the two methodologies and draw

future development perspectives. The manuscript is organised as follows. A thorough literature review of microscopic KB pedestrian models is presented in the next section. In Sec. III, we give a literature overview about the DL approach distinguishing between long short-term memory networks, convolutional neural networks, and generative adversarial neural networks. In Sec. IV we show a comprehensive comparison of both approaches focusing on their methodologies, phenomena of interest, and application scales. Finally, in the last section, we discuss future directions and perspectives of common developments of the KB and DL modelling approaches.

II. THE KNOWLEDGE-BASED APPROACH

In the beginnings of the discipline, pedestrian dynamics researchers mostly used direct observations, photographs and time-lapse films to improve knowledge about the behavior of pedestrians [42]. This knowledge was applied to develop level-of-service concepts, design elements of pedestrian facilities and planning guidelines [27], [43], [44], [45]. These concepts and guidelines are useful to understand and control pedestrian dynamics, but are not suited for predictions of pedestrian flows or trajectories. Therefore, in the next step researchers started to create simulation models like force-based microscopic models [46], queuing models [47], the transition matrix model [48] or the models of Henderson [49], [50] which conjectured that the behavior of pedestrian crowds is similar to that of gases or fluids. These last models describe aggregated quantities and not individual pedestrian performances. They are called macroscopic models. Currently, KB pedestrian models range from macroscopic, mesoscopic, and microscopic models among others modeling scale characteristics. Macroscopic and mesosopic approaches are borrowed from continuous fluid dynamics or gas-kinetic models describing the dynamics at an aggregated level, while microscopic approaches model individual pedestrian motions. In the literature, many reviews focus on the modelling scales of pedestrian dynamics and the passages from a modelling scale to another [28], [51], [52], [53], [54], [55]. Some other reviews highlight pedestrian collective dynamics [14], [56], [57], [58] or applications in layout design [59]. In the following, we propose a review of KB pedestrian models focusing on microscopic approaches and their applications for prediction of pedestrian trajectories.

A. Microscopic Pedestrian Models

In the three last decades, many researchers focused on individual motion of pedestrians using different microscopic models. One of the advantages of microscopic approaches compared to macroscopic ones is their natural ability to reproduce heterogeneous behaviors. Indeed, the pedestrian being individually considered, it is straightforward to attribute specific characteristics to each agent and to take into account behavioral heterogeneity, among other heterogeneous aspects. On the other hand, microscopic models can be computationally expensive and their use is limited in case of large-scale simulation.

Microscopic pedestrian models consider the individual behaviors and interactions between individuals. The crowd dynamic phenomena result from mutual influences at a macroscopic level [59]. Microscopic models are mainly designed to reproduce characteristics macroscopic features such as fundamental diagrams or collective organisations like band formation [60], [61]. These kinds of models, describing individually pedestrian dynamics, can be used to predict trajectories of pedestrians at any scale. The individual pedestrian behavior is described according to certain KB rules ground on physical, social or psychological factors [55]. These rules are formulated in hand-crafted dynamic equations based on Newton's laws of motion. Given the input information about the initial status of the pedestrians like position, velocity, and acceleration a forward simulation of these rules can be used to predict the future trajectories.

Depending on the inputs and outputs of the model, the motion of the pedestrians to a new position can be determined in different ways. If the output of the model is the new velocity or acceleration, which then allows calculating the new position, the model is classified as a velocity- or acceleration-based model, respectively. If the position is directly determined by certain rules and is not based on differential equations, the class of models is called decision-based.

1) Acceleration-Based Models: Acceleration-based models, typically *force-based* models, describe a class of microscopic models where the movement of pedestrians is defined by a superposition of exterior forces. One of the first force-based models dates backs to the 1975s and the work by Hirai and Tairu [46]. Nowadays, most of the acceleration-based models are force-based consisting of a relaxation (or anisotropic) term to the desired direction and an interaction term. This last term is generally the sum of repulsion (social force) with the neighbours and obstacles [62], [63]. The interaction force is the gradient of a potential on the distances and eventually the speed differences with the neighbors. This gradient may be exponential as in the social force model [1], algebraic as for the centrifugal and generalised centrifugal force models [64], [65], or partly linear as in the two-dimensional optimal velocity model [66]. The interaction force is weighted by a vision field concept, attributing more importance to the obstacles in front. Acceleration-based models, being of the second order, require relatively fine discretization scheme and may be subjected to numerical pitfalls [67]. Many currently developed acceleration-based pedestrian models are extensions of the social force model (see the review [68] and references therein). Yet, not all acceleration-based models rely on force concepts. See for instance the model by Moussaid *et al.* [69] based on the concept of desired time gap, the model by Karamouzas *et al.* [70] relying on time-to-collision variables, or a recent model by Lu *et al.* [71] based on anticipation mechanisms.

2) Velocity-Based Models: Acceleration and force-based models allow describing inertial effects, as well as delays in the dynamics. Such mechanisms are questionable for pedestrian dynamics. Indeed, in contrast to a driver in a vehicle, inertial effects are minor for a pedestrian and almost no latency takes place in the motion process. These assertions, upon other, carry the current of *velocity-based* modeling approaches. Velocity-based models have been developed since

the 2000s in the literature, later than the acceleration-based models [55]. They are partly inspired by robotic. Technically, velocity-based models rely on first order-differential equations, whereas acceleration-based models are based on second order equations. Velocity-based models are speed functions depending on the position differences with neighbors and obstacles. As for acceleration-based models, the speed of the neighbors (or speed difference) may be taken into account as well (making the system of speed equations implicit). A large class of velocity-based models are based on collision cones and collision avoidance techniques [25], [72], [73], [74], [75], [76], [77], [78]. The models are formulated as optimisation problems on the ensemble of feasible trajectories devoid of collisions [79], [80]. The presence of collisions is generally determined by assuming that the velocity of the neighbors is constant. The velocity ensemble leading to collisions describes then cones in space [72]. To avoid unrealistic oscillation effect (ping-pong effects), the models are extended to reciprocal velocity obstacle model (RVO) [73] or optimal reciprocal collision avoidance (ORCA) [25] for which avoidance techniques are determined in coordination between the pedestrians. ORCA models and their extensions are frequently used in computer graphics to reproduce crowd behaviors (see, e.g., [81], [82]). Other velocity-based models derive from concepts of bearing angle [83], gradient navigation [84], or, inspired from vehicular dynamics, time gap variable [85], [86].

3) Decision-Based Models and Cellular Automata: In *decision-based* or *rule-based* models the pedestrian behavior is not modeled based on differential equations, but on rules or decisions determining the new agent positions, velocities, etc. [55]. The time is considered to be discrete for this class of models. In synchronous approaches, the pedestrians make decisions at time $t + \Delta t$ knowing the state of the system at time t. The time step Δt, playing the role of reaction time, has a direct physical meaning and can be used for the calibration of the model. Cellular automata (CA) are typical decision-based models. Not only the time is discrete in CA models, the space and state (i.e., velocity) of the pedestrians are discrete as well. The pedestrians evolve on a lattice, that is generally squared or hexagonal. A cell can be occupied by a single pedestrian only (exclusion rule). The size of a cell corresponds to the size of a pedestrian, generally 40 cm × 40 cm on a squared lattice, i.e., a maximal density of 6.25 ped/m^2 [87]. The first pedestrian CA models date back to the end of the 1990s [88], [89], [90], [91]. In floor field CA [2], [92], [93], [94], the rules and transition probabilities to the neighboring cells result from static and dynamic floor fields. The static floor field describe the desired velocity of the pedestrian. The dynamic floor field models the interactions with the neighbors. These interactions are inspired from the process of chemotaxis [95] used by some insects, typically pheromones with ants. An important modeling part of the CA approaches consists in solving conflict cases when two pedestrians covet the same cell at the same time. A priority rule may be defined, which can be random [93]. In [96], friction probabilities are introduced for which no pedestrian reaches the desired cell in case of conflict. Such a mechanism allows notably to explain clogging effects at bottlenecks. Recent decision-based models are based on cognitive effect [97], [98] or learning process [38].

B. Trends During the Past Decades

The modeling and experimentation of pedestrian dynamics is a quite young research field. First investigations and models date back to the 1960s and the 1970s [42], [46], [48], [49]. However, the topic has mainly been the subject of significant research over the past three decades. Experimental studies on pedestrian dynamics in laboratory conditions have been intensively carried out during the 2010s. Pedestrian experiments rely on uni-directional flow, counter-flow, bottleneck, intersecting flow, etc. An open access data archive can be found in Germany; see [99] and references therein. At the same time, authors developed different types of KB pedestrian models, ranging from microscopic to macroscopic modeling scales; see, e.g., the reviews [28], [51], [52], [55], [56]. Most important in the literature of the KB models is undoubtedly the microscopic social force model by Helbing and Molnár, and more generally the force-based microscopic modeling approach (see Table I). Traditional KB approaches by cellular automata, queuing processes, or in analogy to fluid or gas dynamics seen currently to reach a plateau, even if the trends are still light increasing (see Fig. 3). The microscopic force-based models and approaches based on collision avoidance techniques remain relevant with an expanding number of citations. This is because they commonly serve as benchmark references to evaluate the quality of the predictions with deep learning methods (see, e.g., [34], [100]).

C. Knowledge-Based Models for Understanding and Predicting

The aim of knowledge-based models is mainly to identify mechanisms and fundamental parameters operating collectively in the pedestrian dynamics. First of all, body exclusion effects are responsible for jamming and clogging effects, and notions of maximal density. The parameters of KB models rely on the concept of *fundamental diagram*, a phenomenological uni-modal relationship between the flow and the density. The fundamental diagram relationship has been pointed out as early as the 1960s. Yet, investigations on its shape and scattering are still actively ongoing [105], [106], [107], [108], [109], [110], [111]. Identified parameters are, upon others, the desired speed, the agent size, or the reaction time at microscopic scales. They are the maximal density or the capacity at the macroscopic level. The quantitative estimations of the parameters, as well as their numbers and nature, are controversial and subject to diverse influences, type of the flows (e.g., uni-directional, bi-directional), context and motivation, ages, or cultural effects (see [14] and references therein). Simple microscopic rules allows explaining macroscopic shapes of the fundamental diagram [69], [112], [113]. Here, as for traffic flow, temporal parameters such as the reaction time or the time gap with the next pedestrian ahead turn out to be highly relevant [69], [85].

One of the main highlights of KB models is the identification of self-organisation phenomena and the emergence

TABLE I

SELECTION OF IMPORTANT ARTICLES IN THE LITERATURE OF KNOWLEDGE-BASED PEDESTRIAN MODELS WITH FOCUS ON MICROSCOPIC APPROACHES. (NUMBER OF CITATIONS BASED ON SEARCH FROM 31/12/2021)

| First author, year	Article's title and reference	Family	Citations
Burstedde, 2001	Simulation of pedestrian dynamics using a two-dimensional cellular automaton [2]	Cellular automata	1931
Kirchner, 2002	Simulation of evacuation processes using a bionics-inspired cellular automaton model for pedestrian dynamics [93]	—	1119
Van den Berg, 2008	Reciprocal velocity obstacles for real-time multi-agent navigation [73]	Collision avoidance	1363
Pellegrini, 2009	You'll never walk alone: Modeling social behavior for multi-target tracking [74]	—	1206
Van den Berg, 2011	Reciprocal n-Body Collision Avoidance [25]	—	1440
Helbing, 1995	Social force model for pedestrian dynamics [1]	Force-based	6245
Helbing, 2000	Simulating dynamical features of escape panic [101]	—	5382
Chraibi, 2010	Generalized centrifugal-force model for pedestrian dynamics [65]	—	335
Moussaid, 2011	How simple rules determine pedestrian behavior and crowd disasters [69]	—	1009
Karamouzas, 2014	A universal power law governing pedestrian interactions [70]	—	247
Treuille, 2006	Continuum crowds [102]	Queuing	1176
Henderson, 1971	The statistics of crowd fluids [49]	Gas-kinetic	778
Hughes, 2002	A continuum theory for the flow of pedestrians [103]	Fluid dynamics	1174
Chowdhury, 2000	Statistical physics of vehicular traffic and some related systems [28]	Review	2862
Helbing, 2001	Traffic and related self-driven many-particle systems [51]	—	3979
Castellano, 2009	Statistical physics of social dynamics [52]	—	3963
Schadschneider, 2009	Evacuation dynamics: Empirical results, modeling and applications [56]	—	766
Bellomo, 2011	On the modeling of traffic and crowds: A survey of models, speculations, and perspectives [53]	—	464
Bechinger, 2016	Active particles in complex and crowded environments [104]	—	1507
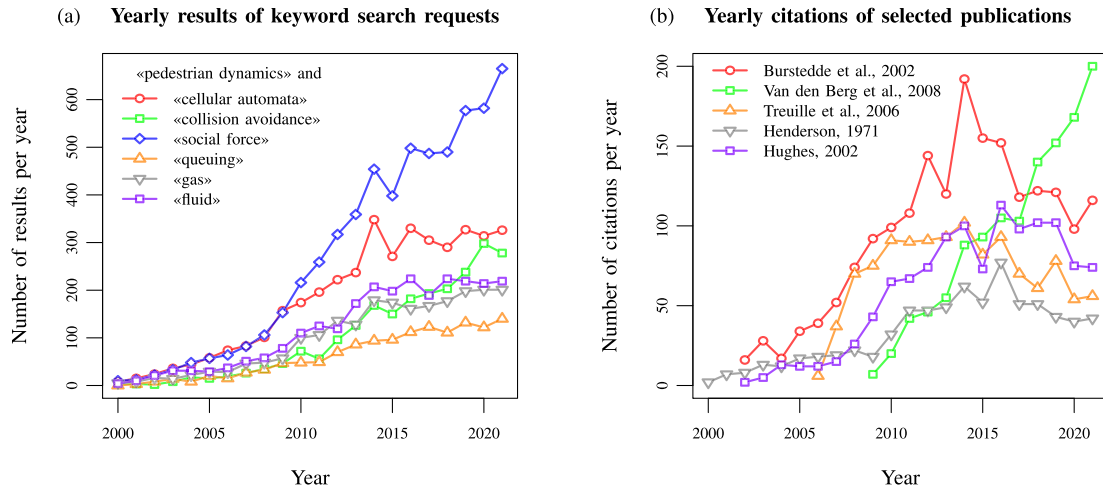
Fig. 3. Number of annual citations estimated by performing an online search with the engine Google Scholar [40]. Panel (a): Results for the tags «pedestrian dynamics» and keywords related to different model classes. Panel (b): Yearly citations for the articles by Burstedde *et al.* [92], Van den Berg *et al.* [73], Treuille *et al.* [102], Henderson [49], and Hughes [103].

of coordinated dynamics, patterns, structures, and orders at macroscopic scales. Multi-scale approaches allow understanding how microscopic individual behaviors initiate the emergence of macroscopic collective dynamics [60], [61]. Prominent examples of collective dynamics are lane formation [114], [115], stop-and-go waves [116], [117], freezing-by-heating effect [118], [119], herding effect [93], [118] or oscillations, intermittent flow, and pattern formation at bottlenecks and in intersections [27], [120], [121], [122]; see the review [14], [28], [29] and references therein. Comparable self-organisation phenomena arise in social systems and social networks, notably for opinion formation [123], [124], [125].

This includes a large class of non-equilibrium systems of self-driven or active particles often called *active matter* in the literature of statistical physics [52], [104], [126], [127], [128], [129], [130]. Understanding complex non-linear dynamics operating at different modeling scales remains challenging and currently attracts much attention, notably through data-based approaches [3], [6], [31], [131], [132].

The aims of microscopic KB models are mainly to provide a better understanding of large-scale dynamics from individual walking behaviors. Predictions of trajectories for a given dataset is no direct goal of KB models, but an indirect one. In the literature of pedestrian dynamics, some KB models have

been proven to be useful for predicting pedestrian trajectories. They are implemented in multi-agent simulation tools [133], [134], [135], [136] to analyse pedestrian dynamics in different types of infrastructures or for specific outdoor events. The KB model, typically the social force model, is the technical kernel of the dynamics in complement to furthers mechanisms and controls describing different types of behaviors or motivation levels, agent characteristics, and further context effects.

III. The Deep Learning Approach

The foregoing approach heavily relies on a theoretical modeling framework. Pre-identified key mechanisms were formulated in equations with a few meaningful parameters that have to be calibrated and validated. After that, the model can be used to simulate the scene, which can generate useful information for many applications. The predictive capacity for pedestrian trajectories almost comes as a by-product in the KB approach. In the DL approach the prediction of trajectories is not a by-product, but the main focus. As in many other domains, the DL approach captured a lot of attention over the last decade due to an increase of real and experimental databases, improvements in the computational capacity of computers [55], and the requirement of accurate pedestrian predictions for applications like autonomous vehicles or service robots [32].

There exist different possibilities in the literature to classify the DL methods. Rudenko *et al.* [32] classify the DL approach into sequential methods and non-sequential methods, based on the type of function approximation they are used for. Bighashdel and Dubbelman [24] classify the methods according to their main focus: The interaction-based approach, where the interactions between the pedestrians are addressed, the path-planning approach, where the trajectories are highly affected by the destinations, and the intention-based approach, where the intention is estimated. In this article, we differentiate between three classes of supervised DL algorithms: LSTM, CNN, and GAN. Currently, these three classes of methods seem to be most relevant in research, although there are also promising publications on transformer networks [137], [138] or variational autoencoder [139].

Before researchers start to use these DL algorithms, statistical models were applied to make predictions based on data. These statistical models learn pedestrian behavior by fitting different function estimates to data. One possibility is to use linear models with Kalman filters or extended Kalman filters [34], [140], [141]. Kalman filters can be used for predictions by propagating the current state with a dynamical model without the inclusion of new measurements. These simple models are not able to account for interactions between humans and thus do not fit for predictions of crowded scenes. The first statistical models that could learn interactions are those based on Gaussian process (GP) like the IGP from Trautman *et al.* [142]. They proposed an interacting GP where the trajectory of a pedestrian is represented as a Gaussian process. The interaction potential combines multiple trajectories, so that multi-modal distributions can be represented with relatively few parameters. It has been demonstrated, that they perform well with noisy observations and have closed-form predictive uncertainty. Also

in [143], [144], [145], [146], and [147] GP based models where proposed to model future pedestrian behavior. Other common statistical approaches that are not based on deep learning are approaches based on Markov property [148]. These approaches include hidden Markov models, in which the hidden state is the pedestrians intent [143], [147]. Whereas the GP uses the entire observed trajectory for the prediction of future trajectories, the predictions with Markov models only depend on the current state [148].

Besides these statistical models, reinforcement learning (RL) is an important method for modelling crowd behavior. Most of all, RL techniques are relevant for robotics to anticipate surrounding pedestrian behavior and to plan a collision-free paths [149], [150]. RL is an unsupervised machine learning method, where an objective is learned via trial and error associated to a *reward function* that rewards or penalizes agent behaviors. Therefore, no data is required. It is assumed that the reward function contains the necessary information [151]. An exception to this is *inverse* RL (IRL) where the design of the reward function is based on data. Kretzschmar *et al.* [152] introduce an IRL algorithm that uses a maximum entropy probability distribution for a joint set of continuous state-space for mobile robot navigation in crowds. Kitani *et al.* [153] propose an well-established IRL algorithm for plausible path predictions. To extend the flexibility of the IRL algorithms recent works use deep IRL algorithms that can estimate nonlinear, continuous reward functions. These deep IRL show promising results especially in robot path planning and vehicle driving behavior [154]. If it is the aim to predict pedestrian trajectories, these deep IRL algorithms are often combined with LSTM networks, which are supervised learning algorithms that will be discussed in the following. For more literature about RL in pedestrian dynamics refer to, e.g., [155], [156], [157]. For IRL methods used for trajectory predictions see [153], [158], for deep IRL methods see [159], [160], [161], or to get a wider perspective, refer to the surveys [54], [162].

A. Long Short-Term Memory Networks

A recurrent neural network (RNN) is an important class of machine learning methods, that uses feedback connections to store representations of recent input events in form of activations [169]. These feedback connections can exist between different time steps providing a temporal memory to the network [170]. Because of this capability, they are especially suited for sequence modeling tasks such as time series prediction and sequence labeling tasks [171]. Most successful are the LSTM architectures of RNNs, that use purpose-built memory cells to store information. They have achieved impressive results in many sequential prediction tasks like speech recognition [172], machine translation [173], handwriting recognition and generation [174], and image captioning [175]. LSTM networks can be trained for sequence generation by processing real data sequences one step at a time and predicting what comes next so that temporal coherence can be ensured. The training of a LSTM network can be lengthy and difficult, but the forward-propagation of new data (testing) can happen quite fast. That is why they are a popular choice for real-time tracking and predicting.

TABLE II

SELECTION OF IMPORTANT ARTICLES OF DEEP LEARNING ALGORITHMS PREDICTING PEDESTRIAN TRAJECTORIES.
(NUMBER OF CITATION BASED ON SEARCH FROM 31/12/2021)

| First author, year	Article's title and reference	Family	Citations
Trautman, 2010	Unfreezing the robot: Navigation in dense, interacting crowds [144]	Gaussian	459
Kitani, 2012	Activity forecasting [153]	IRL	727
Alahi, 2016	Social-LSTM: Human trajectory prediction in crowded spaces [34]	LSTM	1770
Kretzschmar, 2016	Socially compliant mobile robot navigation via IRL[152] [153]	IRL	326
Yi, 2016	Pedestrian behavior understanding and prediction with deep neural networks [163]	CNN	124
Lee, 2017	DESIRE: Distant future prediction in dynamic scenes with interacting agents [159]	LSTM	604
Fernando, 2018	Soft+ hardwired attention: A LSTM framework for human trajectory prediction and abnormal event detection [164]	LSTM	219
Gupta, 2018	Social GAN: Socially acceptable trajectories with generative adversarial networks [41]	GAN	910
Nikhil, 2018	Convolutional neural network for trajectory prediction [165]	CNN	83
Vemula, 2018	Social attention: Modeling attention in human crowds [166]	LSTM	357
Xue, 2018	SS-LSTM: A hierarchical LSTM model for pedestrian trajectory prediction [167]	LSTM	217
Sadeghian, 2019	SoPhie: An attentive GAN for predicting paths compliant to social and physical constraints [168]	GAN	457
Rudenko, 2020	Human motion trajectory prediction: A survey [32]	Review	282
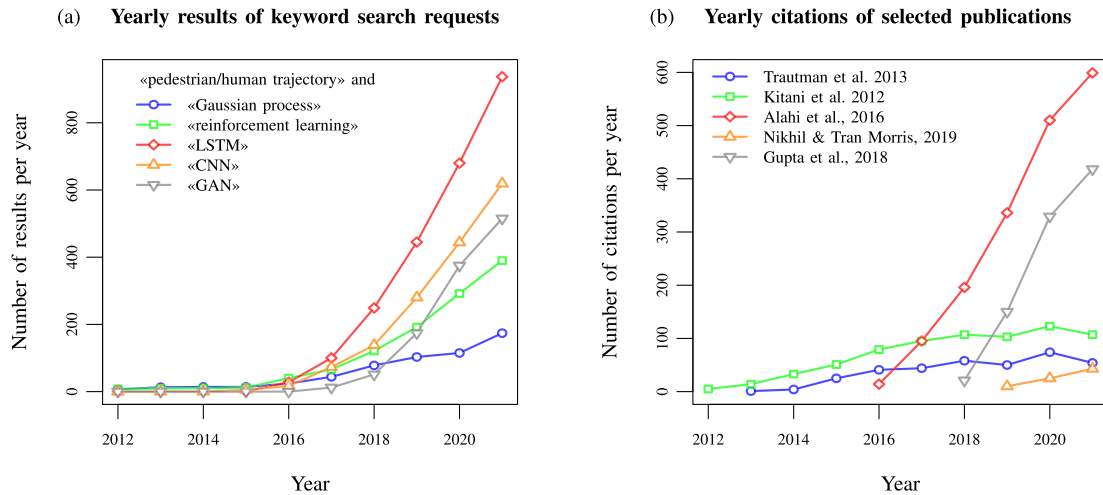
Fig. 4. Number of annual citations estimated by performing an online search with the engine Google Scholar [40]. Panel (a): Results for the tags «pedestrian trajectory» and keywords related to the class of algorithm. Panel (b): Yearly citations for the articles by Trautman *et al.* [142], Kitani *et al.* [153], Alahi *et al.* [34], Yi *et al.* [163], and Gupta *et al.* [41].

Recently, many researchers use these LSTM networks to predict trajectories of pedestrians. Most famous is the social-LSTM proposed by Alahi *et al.* [34]. It is the first major DL application that takes the interactions between pedestrians in crowded scenarios into account. The main contribution of the social-LSTM is a novel pooling layer called social pooling that gathers the hidden states of nearby pedestrians. With that technique, the influence of neighboring pedestrians on the movement of the ego pedestrian can be included as an input for the following prediction step. Inspired by the success of the social-LSTM many researchers try to use LSTM networks with different settings. The scene-LSTM incorporates additional scene information [176]. The social-scene LSTM [167] considers the scene influence, the social information of the person, and scene scale information in three different LSTM networks. Pfeiffer *et al.* [177] incorporate static obstacles that have to be avoided. Other approaches that process scene information are [159], [178]. Besides the capturing of scene

information, some LSTM algorithms integrate an attention mechanism in the interaction module to capture the relative importance of each person in the scene. Other LSTM networks focus on attention mechanisms in the interaction module to capture the relative importance of each person in the scene. In some of these works, the attention weights are learned by data or by added handcrafted based on domain knowledge [164], [179], [180], [181]. Another relevant work is STGAT by Huang *et al.* [180] that uses graph neural networks (GNN) instead of a pooling module. To share information across the pedestrians, each agent is treated as a node of a graph. Derived from this work DAG-Net [182] presents a double attentive GNN combined with a RNN and STR-GGRNN [183] introduces an online framework that automatically infers the social interactions by completing the graph edges. An overview of DL pedestrian trajectory prediction algorithms relying on LSTM networks is provided in Table III. The works are sorted by the publication year to give

TABLE III
OVERVIEW OF DL PEDESTRIAN TRAJECTORY PREDICTION ALGORITHMS RELYING ON LSTM NETWORKS

First author, year	Name	Main characteristics
Alahi, 2016 [34]	Social-LSTM	Social pooling layer to handle interactions
Lee, 2017 [159]	DESIRE	Ranks and refines the generated trajectories
Fernando, 2018 [164]	Soft + Hardwired Attention	Utilises "soft attention" as well as "hard-wired" attention
Xue, 2018 [167]	SS-LSTM	Three different LSTMs to capture person, social and scene scale information
Manh, 2018 [176]	Scene-LSTM	Incorporates scene information
Hasan, 2018 [184]	MX-LSTM	Takes head pose and vision range into account
Vemula, 2018 [166]	Social-Attention	Based on social-LSTM, but captures the relative importance of each person when navigating in crowds
Bisagno, 2018 [185]	Group- LSTM	Coherent filtering algorithm to segment groups
Cheng, 2018 [186]	Social-Grid LSTM	Combines social pooling and grid-LSTM methods
Bartoli, 2018 [187]	Context-aware-Social-LSTM	Interactions with static elements and dynamic agents
Pfeiffer, 2018 [177]	Static-LSTM	Angular pedestrian grid combined with CNN
Ivanovic, 2019 [188]	The Trajectron	LSTMs combined with CVAEs and dynamic spatiotemporal grahical strucutres
Lisotto, 2019 [178]	SNS-LSTM	Social, navigation and semantic pooling mechanism
Huang, 2019 [180]	STGAT	Spatial-Temporal Graph neural network combined with LSTM
Haddad, 2019 [179]	Situation-Aware-LSTM	Spatio- temporal graph that operates on the local and global contexts
Syed, 2019 [189]	SSeg-LSTM	Semantic segmentation to incorporate scene information
Zhao, 2019 [190]	MATF	Encodes past trajectory and scene context into a Multi-Agent Tensor
Monti, 2020 [182]	DAG-Net	Double attentive GNN that deals with past interactions and future goals

the reader an impression about the overall progress in this field.

B. Convolutional Neural Network

Convolutional neural network (CNN) are machine learning methods mostly used for computer vision tasks like image classification [191], object detection, object tracking and image segmentation [192]. For studies of pedestrian dynamic these methods are highly important because of their impressive performance in classifying and tracking of objects like pedestrians or vehicles [193], [194], [195]. Another advantage is the reduced computational load outperforming regular neural networks [165], [196]. However, CNN are not widely used to predict pedestrian trajectories, because these are non-sequential methods, which makes it difficult to design the network input and output [163]. They are mostly used for trajectory predictions of road vehicles [197], the prediction of pedestrian behaviors for autonomous vehicles [198], [199], or pose/action recognition [200], [201]. The first CNN designed to model and predict pedestrian trajectories is the "Behavior-CNN" from Yi *et al.* [163]. It is a 3-stage deep CNN that encodes the pedestrian behavior into sparse displacement volumes which can be directly used as network input at the first stage. The CNN from Nikhil and Tran Morris [165] utilizes a highly parallelizable convolutional layer to handle temporal dependencies. Trajectory histories are embedded by means of a fully connected layer. Stacked convolutional layers are used to learn temporal dependencies in a consistent manner. The features from the final convolutional layer are passed through a fully-connected layer which generates all predicted positions simultaneously.

An extension of the CNNs are the graph CNNs which were first introduced by [202]. Mohammed *et al.* [196] propose the Social-STGCNN that uses graph CNN by modeling the interactions as a spatio-temporal graph, whose edges model the social interactions between the pedestrians. Dan [203] proposed a graph CNN combined with a LSTM network. The graph CNN extract the feature from the pedestrians and the scene for which every pedestrian is regarded as a node and the relationship between each node and its neighbors is obtained by graph embedding. The LSTM encodes the relationship so that the model can predict nodes trajectories. Other approaches combing CNNs with LSTM networks are [204], [205], [206]. An overview of DL pedestrian trajectory prediction algorithms relying on convolutional neural networks is given in Table IV.

C. Generative Adversarial Networks

A problem that is inherent with predictions of trajectories is the multimodal nature of future pedestrian trajectories. GANs have a high potential to cope with this problem, because of their capabilities to generate multimodal samples [212]. With the use of GANs, it is possible to predict a distribution of potential future trajectories and not just the best single trajectory [213]. The architecture of a GAN consists of two part: a generator and a discriminator. In the DL approach, a neural network is trained to match the desired data distribution and achieve a low error rate. In a non-intuitive way, the generator of a GAN is trained so that the error rate of the discriminator increases at first. The generator takes the training data as input and creates different output samples. In turn, the discriminator takes the generated samples and the training data and tries to distinguish whether a given sample belongs to the true data distribution or is generated by the generator. Both components are engaged in a competition similar to a two-player min-max game where each one tries

TABLE IV
OVERVIEW OF DL PEDESTRIAN TRAJECTORY PREDICTION ALGORITHMS RELYING ON CONVOLUTIONAL NEURAL NETWORKS

First author, year	Name	Main characteristics
Yi, 2016 [163]	Behavior CNN	3-stage deep CNN that creates sparse displacement volumes
Varshneya, 2017 [207]	SSCN	Static spatial context modeled with CNN
Rehder, 2018 [198]	RMDN	CNN for infering destination from images and position. LSTM for prediction
Nikhil, 2018 [165]	CNN	Highly parallelizable CNN to handle temporal dependencies
Mohamed, 2020 [196]	Social-STGCNN	Models the interactions as a graph using social spatio-temporal graph CNN
Yu, 2020 [208]	TGConv	Transformer based graph convolu- tion mechanism
Dan, 2020 [203]	Spatial-Temporal Block	Spatial Temporal Graph CNN combined with LSTM
Ridel, 2020 [206]	COVLSTM	2-D grid combined with CNN and LSTM
Jain, 2020 [205]	DRF-Net	Discrete residual flow network
Zhang, 2021 [209]	Social-IWSTCNN	CNN with spatial and temporal features
Zhao, 2021 [210]	STUGCN	CNN with spatio-temporal graph architecture
Zamboni, 2021 [211]	Conv2D	2D Convolutional models with different network achitectures

to outsmart the other one [41], [197]. From this process, the generator learns to generate data that resemble the true data distribution. Although the results of the GAN-based methods are promising, there are two main difficulties. First, GANs can be hard to train and second, GAN learning often suffers from mode collapse [214], [215]. Applied for trajectory prediction, most GANs are combined with LSTM networks. In the social GAN from Gupta *et al.* [41] the generator is composed of an LSTM-based encoder, a context pooling module, and an LSTM-based decoder. The discriminator uses LSTMs as well. The pooling module adopted in the social GAN uses a uniform weight for all surrounding pedestrians. Therefore, it can not distinguish the different effects exerted on a target pedestrian by pedestrians at different distances and traveling at different speeds. For that reason, many authors added attention mechanisms. Sadegehian *et al.* [168] add an attention module to assign different soft attention distribution weights to the surrounding pedestrians and the static environment. Using an attention module with a physical and social component and a feature extractor module composed of a CNN and several LSTM network encoder, the model SoPhie learns the interaction information of different agents and extracts the most important information from the neighbors. Social Ways [216] applies info-GAN [217], which introduces latent code to enhance the multi-modality of prediction and computes the attentive social features to generate a more convincing result. It uses discrimination loss for the discriminator, adversarial loss for the generator and information loss for both. Social-BiGAT [218] relies on BiGAN architecture to help reduce the variance of the predicted trajectory distributions and allow for better generalization. In Lv *et al.* [219] the authors propose a GAN combined with transformer networks which generates trajectory distributions to capture the uncertainty of the predictions. However, these methods focus on the trajectory prediction in homogeneous environment without considering the types of road users. Lai *et al.* [220] use an attention module, containing two components, in order to alleviate the issues given by the complexity of a scene with many heterogeneous interacting agents [220]. An overview of DL pedestrian trajectory prediction algorithms relying on generative adversarial neural networks is proposed in Table V.

IV. COMPARING THE APPROACHES

In this Section, we present a comparison of the KB and the supervised DL approaches. First, we focus on technical differences (Sec. IV-A) and second, we describe the differences regarding the applications for predicting pedestrian trajectories (Sec. IV-B).

A. Technically Oriented Comparison

There are considerable differences between the two approaches, which makes it difficult to create a common framework that both can share for a fair comparison. Kothari *et al.* [222] define trajectory predictions as "given the past trajectories of all humans in a scene, forecast the future trajectories which conform to the social norms". This definition fits for the DL approach, but not the KB approach where the past trajectories are not used to predict (simulate) the future ones. In KB methods, inputs describe the current state of the system through, e.g., instantaneous relative position [1], relative velocity [65], time gap [69], or collision-related indicators such as bearing angle [83], collision cone [73] or time-to-collision [70]. Moreover, the approaches have different criteria in terms of quality of the fit. A "good" KB model has few interpretable parameters, can simulate realistic pedestrian trajectories, and improves understanding of the phenomena. To evaluate the performance of a KB method, the parameters have to be calibrated and the model needs to be validated. The calibration phase is done using knowledge on the values of the parameters (i.e., 1 or 1.5 m/s for the pedestrian desired speed, see [87]) or is formulated using empirical data as an optimisation problem (by least squares or by maximum likelihood, see, e.g., [223], [224], [225], [226]). The validation of KB methods is generally done using three typical ways: by using fundamental diagrams [112], [227], using data [228], [229], [230] and by comparing the resulting phenomena with real-world self-organization phenomena. The goal of the DL approach is to predict trajectories that are as close as possible

TABLE V
Overview of DL Pedestrian Trajectory Prediction Algorithms Relying on Generative Adversarial Neural Networks

First author, year	Name	Main characteristics
Gupta, 2018 [41]	Social-GAN	GAN composed of LSTM encoder, context-pooling module and LSTM decoder
Fernando, 2018 [221]	GD-GAN	GAN for pedestrian trajectory predicting and group detection
Sadeghian, 2019 [168]	SoPhie	GAN that uses path history and scence information
Amirian, 2019 [216]	Social Ways	Uses discrimination-, adversarial- and information loss
Kosaraju, 2019 [218]	Social-BiGAT	GAN for multimodal trajectory predictions
Lai, 2020 [220]	AEE-GAN	Info-GAN architecture with recurrent feedback
Huang, 2021 [212]	STI-GAN	Combination of graph attention network and GAN

to the real trajectories. Common key figures for the performance of the DL approaches are the average displacement error (ADE) [74] and the final displacement error (FDE) [34]. 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. When the two approaches are compared in the literature the ADE and FDE are used as reference (see, e.g., [34], [41], [100]).

The algorithms of the DL approach have a large number of parameters that have no direct physical meaning and therefore can be referred to as coefficients. These coefficients have to be trained by one part of the data set (training phase). In general, this is done by using an training algorithm like the back-propagation algorithm [231]. The aim of this training algorithm is to adjust the network settings in a way, that minimizes the given cost function [232]. Common cost functions are the mean square error or the cross entropy. After the training is complete, the algorithms are feed with the rest of the data set to evaluate predictions (testing phase). This cross-validation method allows detecting overfitting phenomena, when the algorithm presents low training error but poor performance for the prediction of new data. The training is usually computationally expensive and is made offline, whereas the predictions of the trajectories (i.e., the computation), once the training is completed, are quite fast and can be made online (in real-time) [35]. We summarize our findings regarding the technical differences of both approaches in Table VI.

B. Application Oriented Comparison

Besides the technical orientated differences, there are also great differences regarding the applications of both approaches. KB approaches are mostly used for crowded situations. Collective dynamics are described at macroscopic scales. Applications mainly rely on large-scale simulations to analyze, e.g., infrastructure design or evacuation situations. In contrast to that, the DL approach focuses more on single pedestrians and their interactions with other pedestrians or the environment locally and in low density situations. Applications are mainly for automatised mobile systems like autonomous vehicles or industrial robots that have to anticipate the future behavior/trajectory of pedestrians to avoid collisions. The DL approach surpasses the KB approach in complexity (i.e., number of inputs as well as number of parameters/coefficients).

This provides a high flexibility and enables the DL algorithms to learn complex interactions and motion patterns when the amount of data is sufficient. This is especially useful in situations where complexity prohibits the explicit programming of a system's exact physical nature.

1) The Importance of Data: There are great differences in the role of data for the two KB and DL approaches. The performance of the DL algorithms highly depends on the quality of the data. The majority of researchers in this discipline use datasets containing low density situations with few interacting pedestrians to train the algorithms. A collection of datasets that are widely used to train and evaluate the algorithms are TrajNet++ [222] or OpenTraj [233]. The KB models do not need data to learn pedestrian behavior. Data is just needed for calibrating the parameters. In practice, this is often done by using experimental data including high density situations, see, e.g., [99]. The data used for the KB approach contains the trajectories of each pedestrian as well as the structure of the environment and locations of obstacles. For the DL algorithms, the data can be more varied because the algorithms are more flexible in terms of inputs. In the graph neural networks, the pedestrians and their interactions are described through a graph [166]. Other authors use spatial information represented as points of interest [187] or as occupancy maps [177]. By using CNNs it is even possible to use images or videos as inputs for the predictions [168], [218]. Without sufficient amounts of data, the DL algorithms are not capable to learn pedestrian behavior and predict future trajectories successfully. In addition to the amount of data, pre-processing is an important step to archive good results and efficiently train the algorithm. One pre-processing technique is data normalization where the coordinates of the trajectories are normalized to coordinates with origin in the first observation, coordinates with origin in the last observation, or relative coordinates [234]. Another important pre-processing technique is data augmentation. Using GANs [235] is one form of data augmentation, but there are also more basic forms like rotating the input trajectories, mirror the trajectories, or applying Gaussian filter [234]. Schöller *et al.* [236] show that data augmentation helps to prevent the DL algorithm from learning environmental priors instead of pedestrian behavior.

2) Numerical Comparison: Many studies report on numerical comparisons of KB and DL approaches for the prediction of trajectories. Authors generally used average and final displacement errors (ADE and FDE) to quantify the prediction

TABLE VI

TECHNICALLY ORIENTED COMPARISON OF THE KB AND THE DL APPROACH FOR PREDICTION OF PEDESTRIAN DYNAMICS

	Knowledge-based	Deep learning
Semantics	Model	Algorithm
	Parameter	Coefficient
	Calibration	Training
	Validation	Testing
Methodology	Differential equation systems or cellular automata	Neural network, mostly RNN, LSTM, CNN, GAN
Inputs	System actual state (e.g., pedestrian relative positions, velocities, etc. at time t)	Past trajectories discretised over the time interval $[t-T, t]$, $T \approx 2$–4 s
Outputs	Future pedestrian positions at time $t+\delta t$, δt ranging from 0.01 s (discretised differential system) to 0.5–1 s (cellular automata)	Future pedestrian positions discretised over time interval $[t, t + T^{\star}]$, $T^{\star} \approx 3$–5 s
Modeling concepts	Fundamental diagram	Learning human-interaction
	Body exclusion, maximal density	Social pooling
	Desired velocity and social force	Attention mechanisms
	Static and dynamic floor field	Group dynamics
	Collision avoidance techniques, spatial and temporal anticipation mechanisms	Human-space interactions
	Vision angle, bearing angle	
	Temporal interaction indicators (time gap, time-to-collision, time-to-interaction)	
Performance evaluation	Flow-density relationship	Average displacement error (ADE)
	Comparison to experimental data	Final displacement error (FDE)
	Description of collective behaviors and self-organised phenomena	Modified Hausdorff distance (MHD)

TABLE VII

QUANTITATIVE COMPARISON OF ADE AND FDE METRICS FOR ARTICLES USING SOCIAL-LSTM AS BENCHMARK WITH DIFFERENT DATASETS

First author, year	Average	ETH	HOTEL	ZARA1	ZARA2	UCY
Syed, 2019 [189]	0.08 / 0.14	0.15 / 0.295	0.05 / 0.08	0.05 / 0.08	0.07 / 0.1	0.1 / 0.16
Xue, 2018 [167]	0.12 / 0.17	0.2 / 0.37	0.08 / 0.13	0.08 / 0.11	0.07 / 0.12	0.2 / 0.24
Manh, 2018 [176]	0.25 / 0.22	0.18 / 0.34	0.25 / 0.29	0.37 / 0.33	0.19 / 0.1	0.25 / 0.03
Alahi, 2016 [34]	0.27 / 0.61	0.5 / 1.07	0.11 / 0.23	0.22 / 0.48	0.25 / 0.5	0.27 / 0.77
Vemula, 2018 [166]	0.37 / 3.32	0.46 / 4.56	0.42 / 3.57	0.21 / 0.65	0.41 / 3.39	0.36 / 4.45
Hossain, 2022 [239]	0.44 / 0.98	0.60 / 1.31	0.15 / 0.33	0.43 / 0.93	0.51 / 1.09	0.52 / 1.25
Zhu, 2019 [240]	0.45 / 0.91	0.73 / 1.48	0.49 / 1.01	0.27 / 0.56	0.33 / 0.7	0.41 / 0.84
Hasan, 2018 [184]	0.64 / 1.45	—	—	0.68 / 1.53	0.63 / 1.43	0.62 / 1.40
Gupta, 2018 [41]	0.72 / 1.54	1.09 / 2.35	0.79 / 1.76	0.47 / 1.00	0.56 / 1.17	0.67 / 1.40

accuracy. Several articles include the social-LSTM [34] and the social force model [1] as references for, respectively, DL and KB approach. In Table VII, the ADE and FDE metrics from selected studies using the social-LSTM as a benchmark are shown. All these articles use the same algorithm, the same datasets, the same error metrics, the same length of observed (3.2s) and predicted trajectories (4.8s), and the same cross-validation methods. Nevertheless, the results vary significantly from one analysis to another. This is not just a problem of the social-LSTM, but of the evaluation of the DL approach in general. Reasons for this high variation could be the randomness of the dataset splitting, the randomness of the starting coefficients, differences in the implementation of the algorithms, or different hyperparameter settings for the training, among others. An attempt to solve this problem

is the trajectory forecasting challenge Trajnet++ [222]. The challenge provides a uniform sampling and evaluation system allowing to compare rigorously different algorithms in the same framework. Table VIII reports on ADE and FDE metrics for selected articles using the social force model and social-LSTM algorithm. We can observe that the error estimates also vary significantly with the social force model. Yet, in contrast to Table VII, the studies do not systematically use the same dataset and prediction length making direct comparisons potentially biased. The objective is to compare the KB and DL approaches using relative errors. Assuming that the setting for KB and DL approaches are identical for a given study, the comparison of relative errors is fair. The variations from one study to another are again very significant. However, except in the work of Cheng *et al.* [237], the social-LSTM

TABLE VIII
QUANTITATIVE COMPARISON OF THE SOCIAL FORCE MODEL AND THE SOCIAL-LSTM (AVERAGE ADE/AVERAGE FDE)

First author Approach	Alahi [34]	Fernando [164]	Cheng [237]	Fernando [221]	Hasan [184]	Song [238]
Social force model	0.39 / 0.60	3.24 / 4.86	0.37 / 1.27	1.5 / 2.46	4.28 / 7.63	0.61 / 0.96
Social-LSTM	0.27 / 0.61	1.76 / 3.51	0.67 / 3.1	1.1 / 1.89	0.97 / 2.08	0.25 / 0.22
Relative error [%]	44 / -1.6	84 / 38	-80 / -145	36 / 24	341 / 267	144 / 336

systematically outperforms the social force model in terms of prediction accuracy. The mean ADE and FDE relative errors are, respectively, 105 and 80.4%. In the work from Hasan *et al.* [184] or from Song *et al.* [238] the differences are especially huge, the social-LSTM being up to 341% more accurate. Such results are statistically not surprising since the DL approach is based on much more free parameters (coefficients) than the KB approach.

a) Advantages and disadvantages of KB and DL approaches: The last paragraph shows that the DL algorithms outperform the KB models in terms of prediction accuracy of pedestrian trajectories in low density situations and that the results of the error metrics have a high fluctuation. This fluctuation is mainly a problem of the DL algorithms because the evaluation of these algorithms is based on these error metrics. In addition to this lack of reproducibility, there is also the problem of missing explainability of the DL approach. This means that the coefficients can not be physically understood and interpreted. Therefore it is not clear why the predicted trajectories have the given shapes. In applications where autonomous systems make decisions, it is important to understand and communicate why specific choices have been made [241]. But these in large numbers existing coefficients are also an advantage of the DL approach because they make it possible to learn complex behaviors by fitting a set of values such that the predicted behavior fits the observed behavior. These algorithms do not need a priori knowledge about the system, just observations.

The comparison of the KB models with the DL algorithms is mostly done numerically, disregarding the missing reproducibility and high fluctuation. Furthermore, important advantages of the KB models are not taken to account in the numerical comparison. First of all, the KB models have the advantage of simple forms and interpretable parameters, which makes it easy to reproduce the results and to understand the predictions. The interpretability of the parameters makes the models flexible and easy to cope with environmental and behavioral changes. For example, if the preferences of pedestrians change because of a pandemic, which entails social distance and desired time-gap to rise, the KB methods can change the parameters to model these new situations, without needing new data. In the DL approach, new data is necessary to learn behavioral changes. But the KB models also have disadvantages, like the need of domain knowledge. Furthermore, only average pedestrian behavior is simulated and it is difficult to capture the complete crowd behavior range with a single model [69]. The advantages and disadvantages, as well as other differences regarding the applications of both KB and DL approaches, are summarized in Table IX.

V. FUTURE DIRECTIONS

In the last part of this work we want to look into the future and describe which trends we can identify.

A. The Hybrid Approach

In Table VIII it is shown that many points that can be criticized in the predictions with the DL algorithms are strengths of the KB models. KB models have few parameters with physical interpretations requiring few data for calibration, while the coefficients of the DL algorithms are generally not interpretable and much data is needed in the training phase. For this reason, the combination of both approaches seems to have potential as some pioneer studies point out [36], [242]. This combination is called the hybrid approach and even though the idea of combining both approaches has picked up momentum just in the last few years, there is already a vast amount of work in diverse disciplines. Examples of the hybrid approach are given in applications like the discovering of novel climate patterns [243], [244], the finding of novel compounds in material science [245], [246], the designing density functionals in quantum chemistry [247], or the improving imaging technologies in bio-medical science [248], [249]. Recently, the first applications for pedestrian trajectory predictions can be found in [38], [35], and [37].

In general, there are three ways of combining both KB and DL approaches. First, KB models can be used to improve the DL algorithms. One possibility to do so is *data generation*. Simulations based on KB models are carried out to obtain a synthetic data set, that is used to train and test the neural network. A popular application of data generation can be found in the training of autonomous vehicles for the augmentation of data for scenarios that are not sufficiently represented in the available data set [250], [251]. In [252], [38], and [34], the authors apply the social force model to generate a synthetic data-set for the training of neural networks and setting of hyper-parameters. *Knowledge-guided design of architecture* is another possibility to improve the DL algorithms with knowledge. The modular and flexible nature of the networks enables the use of knowledge to specify node connections that capture dependencies among variables. In this way, Antonucci *et al.* [35] successfully embed the social force model in the architecture of a neural network to generate predictions of human motion. A common technique to improve the output of DL algorithms is *knowledge-guided loss function*. It makes the output consistent with physical laws so that unrealistic prediction can be ruled out [253], [254]. Because the training of DL algorithms is an iterative process, they require an initial choice of coefficients as a first step to

TABLE IX
APPLICATION ORIENTED COMPARISON OF THE KB AND THE DL APPROACH FOR PREDICTION OF PEDESTRIAN DYNAMICS

	Knowledge-based	Deep learning
Applications	Simulation tool	Socially-aware mobile robots
	Infrastructure design	Design of intelligent tracking systems
	Evacuation situations	Pedestrian trajectory prediction for autonomous vehicles
	Intelligent transport systems	
Application scales	Large-scale simulation	Local prediction scale (in time and space)
	Large infrastructures (train station, stadium, commercial mall), urban centers, outdoor events	Few interacting pedestrians
Crowd density	Low density (long range interaction, e.g., collision avoidance models)	Low density situations
		Long range pedestrian interaction
	High density (short range interaction, e.g., force-based models)	
Advantages	Interpretable parameter	Accurate predictions
	Explainable predictions	Learn complex interaction
	Reproducible	No domain knowledge necessary
	Few data needed	No modelling-bias
		Can process different types of data
Disadvantages	Low use of data	Not interpretable
	Averaged behavior only	Not reproducible in praxis
	Not suitable to complex interaction	Lack of generalisation
	Complete crowd behavior range with a single model difficult to capture	Require large amount of data
		Necessary complexity of the network unknown

commence the learning process. *Knowledge-guided initialization of the network* can be used to guide the network at an early stage to archive generalizable and physically consistent results [255].

The second way of combining the approaches is obtained by using DL algorithms to improve the prediction accuracy of the KB models. One of the oldest and most common way for addressing the imperfections of the KB approach is *residual modeling*. A DL algorithm learns to correct the errors made by the KB model by predictions the model residuals [253]. Bahari *et al.* [256] proposed a so-called "realistic residual block" to improve vehicle trajectory predictions. Another way to improve the KB predictions consists in calibrating the parameters of the KB models by using DL algorithms. Göttlich and Knapp [257] use neural networks to estimate the parameters of the social force model. Hossain *et al.* [239] extent the social force model [1] using group forces, based on neural networks, to consider interactions with static obstacles, other pedestrian, and pedestrian groups. Kreiss [258] uses deep neural networks to estimate different interaction potentials for the social force model.

Given enough data, DL algorithms are capable of predicting pedestrian trajectories of a given scene with relative high accuracy. However, pedestrians experience different interactions in different situations. Whether they are able to make accurate predictions for the hole range of possible scenes and interactions is a open question. The hybrid approach is useful for this problem because it can compensate scarcity of data with available knowledge. The hybrid approaches are promising to improve the predictions of pedestrian trajectories because they benefit from the strengths of both approaches and reduce shortcomings like the missing explainability of the DL approach.

B. Other Directions

As it has been shown in Fig. 2, the DL approach has gained attention just a few years ago. At the current state-of-the-art, DL is mostly used for learning human behavior and predicting single pedestrian trajectories. In the future, the DL approach could be used in more applications in the discipline of pedestrian dynamics and notably for large scale simulation and the simulation of collective dynamics. Nowadays, the behavior and interactions of agents in simulation platforms currently available rely on KB approaches [133], [134], [135]. Following the success of the DL algorithms for predictions of pedestrians in low density situations, there is a high potential to use these methods for accurate predictions of crowd dynamics like in evacuation situations.

In addition to the KB models, RL algorithms are often used for simulating crowd dynamics, but they have some disadvantages, like the need of a reward function, a priori given goal/destination, or the difficulty to incorporate interactions. Such difficulties are overcome by the supervised DL algorithms. The combination of RL and DL algorithms seems to be promising to solve these disadvantages as Everett *et al.* [161] have recently shown. Other works have shown successful applications of deep RL for crowd simulations [160] or even evacuation dynamics [259].

There are still open questions to be tackled, before supervised DL algorithms could be successfully used for large-scale crowd simulations, like:

- Which kind of neural networks, which complexity and which training data should be used for large-scale simulation including different types of geometries?
- Should the type and complexity of appropriate networks as well as the data used for the training depend on the scenarios of the simulation?

Besides these considerations, one may expect to develop deep networks trained on large amounts of data that could predict accurate trajectories for any density levels and any type of facility. Such networks could be used in agent-based simulation platforms whose objective is to simulate any type of scenario. Yet, the question whether such universal networks, as supervised approaches, require training on datasets representing the full diversity of pedestrian dynamics remains open. It may be possible, as in unsupervised approaches, that training on few data is sufficient to obtain accurate predictions even for scenarios the networks are not trained for. Such question raises more generally on the robustness of the predictions against new data, i.e., new situations, scenarios, density levels, or types of facilities. Some preliminary results obtained in a corridor and a bottleneck have shown that neural networks are quite robust to new types of facilities and may even overcome KB models in terms of prediction robustness [21], [260].

Another possible development direction of DL approaches for prediction pedestrian trajectory concerns the nature and type of the inputs. The inputs of the algorithms are mostly the trajectories of the pedestrian over finite past horizon. Yet, studies of pedestrian dynamics with KB models identified different relevant variables like, besides the relative position [1], the relative velocity [65], the time gap [69], or indicators related to possibilities of collision such as bearing angle [83], collision cone [73] or time-to-collision [70]. The use of these variables as inputs, even if they can theoretically be deduced from the trajectories, could allow obtaining prediction improvement, especially in the case of low amount of data for the training phase. We can already observe a trend to use more inputs and notably the speed difference [21], [222], [261] and other hidden variables estimated by training [41], [216]. One may expect to include further variables as inputs. The time-to-collision and the time gap, whose roles are fundamental in KB pedestrian dynamic models [69], [70], are promising candidates. Such variables contain information about the relevance of the interactions and may even substitute the interaction or attention modules of DL approaches.

REFERENCES

[1] D. Helbing and P. Molnar, "Social force model for pedestrian dynamics," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 51, no. 5, pp. 4282–4286, 1995.

[2] C. Burstedde, K. Klauck, A. Schadschneider, and J. Zittartz, "Simulation of pedestrian dynamics using a two-dimensional cellular automaton," *Phys. A, Stat. Mech. Appl.*, vol. 295, no. 3, pp. 507–525, 2001.

[3] A. Ourmazd, "Science in the age of machine learning," *Nat. Rev. Phys.*, vol. 2, no. 7, pp. 342–343, 2020.

[4] G. Saporta, "Models for understanding versus models for prediction," in *Proc. COMPSTAT*. Heidelberg, Germany: Springer, 2008, pp. 315–322.

[5] A. W. Senior *et al.*, "Improved protein structure prediction using potentials from deep learning," *Nature*, vol. 577, no. 7792, pp. 706–710, 2020.

[6] F. Cichos, K. Gustavsson, B. Mehlig, and G. Volpe, "Machine learning for active matter," *Nature Mach. Intell.*, vol. 2, no. 2, pp. 94–103, Feb. 2020.

[7] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, pp. 115–118, Jan. 2017.

[8] V. Gulshan *et al.*, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *J. Amer. Med. Assoc.*, vol. 316, no. 22, pp. 2402–2410, 2016.

[9] K. T. Schütt, F. Arbabzadah, S. Chmiela, K. R. Müller, and A. Tkatchenko, "Quantum-chemical insights from deep tensor neural networks," *Nature Commun.*, vol. 8, no. 1, pp. 1–8, Apr. 2017.

[10] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural message passing for quantum chemistry," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 1263–1272.

[11] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, "Machine learning for molecular and materials science," *Nature*, vol. 559, pp. 547–555, Jul. 2018.

[12] K. G. Reyes and B. Maruyama, "The machine learning revolution in materials?" *MRS Bull.*, vol. 44, no. 7, pp. 530–537, 2019.

[13] S. Bitgood, "An analysis of visitor circulation: Movement patterns and the general value principle," *Curator, The Museum J.*, vol. 49, no. 4, pp. 463–475, 2006.

[14] M. Boltes, J. Zhang, A. Tordeux, A. Schadschneider, and A. Seyfried, *Empirical Results of Pedestrian and Evacuation Dynamics*. Berlin, Germany: Springer, 2018, pp. 1–29.

[15] D. Helbing, I. Farkas, P. Molnár, and T. Vicsek, "Simulation of pedestrian crowds in normal and evacuation situations," in *Pedestrian and Evacuation Dynamics*, M. Schreckenberg and S. D. Sharma, Eds. Berlin Heidelberg: Springer, 2002, pp. 21–58.

[16] X. Zheng, T. Zhong, and M. Liu, "Modeling crowd evacuation of a building based on seven methodological approaches," *Building Environ.*, vol. 44, no. 3, pp. 437–445, 2009.

[17] A. Poibrenski, M. Klusch, I. Vozniak, and C. Müller, "M2P3: Multimodal multi-pedestrian path prediction by self-driving cars with egocentric vision," in *Proc. 35th Annu. ACM Symp. Appl. Comput.*, Mar. 2020, pp. 190–197.

[18] S. Scheggi, M. Aggravi, F. Morbidi, and D. Prattichizzo, "Cooperative human-robot haptic navigation," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2014, pp. 2693–2698.

[19] M. Leo, G. Medioni, M. Trivedi, T. Kanade, and G. M. Farinella, "Computer vision for assistive technologies," *Comput. Vis. Image Understand.*, vol. 154, pp. 1–15, Jan. 2017.

[20] I. Rodin, A. Furnari, D. Mavroedis, and G. M. Farinella, "Predicting the future from first person (egocentric) vision: A survey," *Comput. Vis. Image Underst.*, vol. 211, Oct. 2021, Art. no. 103252.

[21] A. Tordeux, M. Chraibi, A. Seyfried, and A. Schadschneider, "Prediction of pedestrian dynamics in complex architectures with artificial neural networks," *J. Intell. Transp. Syst.*, vol. 24, no. 6, pp. 556–568, Nov. 2020.

[22] T. Kruse, A. K. Pandey, R. Alami, and A. Kirsch, "Human-aware robot navigation: A survey," *Robot. Auto. Syst.*, vol. 61, no. 12, pp. 1726–1743, Dec. 2013.

[23] H. Cheng, F. T. Johora, M. Sester, and J. P. Müller, "Trajectory modelling in shared spaces: Expert-based vs. deep learning approach?" in *Proc. Int. Workshop Multi-Agent Syst. Agent-Based Simul.* Cham, Switzerland: Springer, 2020, pp. 13–27.

[24] A. Bighashdel and G. Dubbelman, "A survey on path prediction techniques for vulnerable road users: From traditional to deep-learning approaches," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, Oct. 2019, pp. 1039–1046.

[25] J. van den Berg, S. Guy, M. Lin, and D. Manocha, "Reciprocal *n*-body collision avoidance," in *Robotics Research*. Berlin, Germany: Springer, 2011, pp. 3–19.

[26] A. Portz and A. Seyfried, "Analyzing stop-and-go waves by experiment and modeling," in *Pedestrian and Evacuation Dynamics*, R. D. Peacock, E. D. Kuligowski, and J. D. Averill, Eds. Boston, MA, USA: Springer, 2011, pp. 577–586.

[27] D. Helbing, L. Buzna, A. Johansson, and T. Werner, "Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions," *Transp. Sci.*, vol. 39, no. 1, pp. 1–24, 2005.

[28] D. Chowdhury, L. Santen, and A. Schadschneider, "Statistical physics of vehicular traffic and some related systems," *Phys. Rep.*, vol. 329, nos. 4–9, pp. 199–329, May 2000.

[29] N. Bellomo, B. Piccoli, and A. Tosin, "Modeling crowd dynamics from a complex system viewpoint," *Math. Models Methods Appl. Sci.*, vol. 22, Aug. 2012, Art. no. 1230004.

[30] M. Moussaïd *et al.*, "Traffic instabilities in self-organized pedestrian crowds," *PLOS Comput. Biol.*, vol. 8, no. 3, pp. 1–10, 2012.

[31] Z. Shahhoseini and M. Sarvi, "Collective movements of pedestrians: How we can learn from simple experiments with non-human (ant) crowds," *PLoS ONE*, vol. 12, pp. 1–20, 2017.

[32] A. Rudenko, L. Palmieri, M. Herman, K. M. Kitani, D. M. Gavrila, and K. O. Arras, "Human motion trajectory prediction: A survey," *Int. J. Robot. Res.*, vol. 39, no. 8, pp. 895–935, 2020.

[33] J. Frochte, *Maschinelles Lernen: Grundlagen und Algorithmen in Python*. Munich, Germany: Carl Hanser Verlag GmbH Co KG, 2019.

[34] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social LSTM: Human trajectory prediction in crowded spaces," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 961–971.

[35] A. Antonucci, G. P. R. Papini, L. Palopoli, and D. Fontanelli, "Generating reliable and efficient predictions of human motion: A promising encounter between physics and neural networks," 2020, *arXiv:2006.08429*.

[36] Z. Hu, X. Ma, Z. Liu, E. Hovy, and E. Xing, "Harnessing deep neural networks with logic rules," in *Proc. 54th Annu. Meeting Assoc. Comput. Linguistics*, Berlin, Germany, 2016, pp. 2410–2420.

[37] P. Kothari, B. Sifringer, and A. Alahi, "Interpretable social anchors for human trajectory forecasting in crowds," 2021, *arXiv:2105.03136*.

[38] Z. Zhang and L. Jia, "Direction-decision learning based pedestrian flow behavior investigation," *IEEE Access*, vol. 8, pp. 15027–15038, 2020.

[39] P. Kielar and A. Borrmann, "An artificial neural network framework for pedestrian walking behavior modeling and simulation," *Collective Dyn.*, vol. 5, pp. 290–298, Mar. 2020.

[40] *Google Scholar*. Accessed: Dec. 31, 2021. [Online]. Available: https://scholar.google.com

[41] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese, and A. Alahi, "Social GAN: Socially acceptable trajectories with generative adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2255–2264.

[42] D. Oeding, "Verkehrsbelastung und dimensionierung von gehwegen und anderen anlagen des fußangerverkehrs," Technische Hochschule Braunschweig, Forschungsbericht, Braunschweig, Germany, Tech. Rep. 22, 1963.

[43] J. J. Fruin, *Designing for Pedestrians a Level Service Concept*. New York, NY, USA: Polytechnic Univ., 1970.

[44] A. Polus, J. Schofer, and A. Ushpiz, "Pedestrian flow and level of service," *J. Transp. Eng.*, vol. 109, no. 1, pp. 46–56, 1983.

[45] S. Holl, M. Boltes, and A. Seyfried, "Level-of-safety-konzept für den fußverkehr bei Großveranstaltungen," in *Veranstaltungskommunikation*. Wiesbaden, Germany: Springer, 2019, pp. 253–277.

[46] K. Hirai and K. Tarui, "A simulation of the behavior of a crowd in panic," in *Proc. Int. Conf. Cybern. Soc.*, San Francisco, CA, USA, 1975, pp. 409–411.

[47] G. G. Løvås, "Modeling and simulation of pedestrian traffic flow," *Transp. Res. Part B Meth.*, vol. 28, no. 6, pp. 429–443, 1994.

[48] D. Garbrecht, "Describing pedestrian and car trips by transition matrices," *Traffic Quart.*, vol. 27, no. 1, pp. 89–110, 1973.

[49] L. F. Henderson, "The statistics of crowd fluids," *Nature*, vol. 229, pp. 381–383, Mar. 1971.

[50] L. F. Henderson, "On the fluid mechanics of human crowd motion," *Transp. Res.*, vol. 8, no. 6, pp. 509–515, 1974.

[51] D. Helbing, "Traffic and related self-driven many-particle systems," *Rev. Modern Phys.*, vol. 73, no. 4, pp. 1067–1141, 2001.

[52] C. Castellano, S. Fortunato, and V. Loreto, "Statistical physics of social dynamics," *Rev. Mod. Phys.*, vol. 81, no. 2, pp. 591–646, Apr./Jun. 2009.

[53] N. Bellomo and C. Dogbe, "On the modeling of traffic and crowds: A survey of models, speculations, and perspectives," *SIAM Rev.*, vol. 53, no. 3, pp. 409–463, 2011.

[54] F. Martinez-Gil, M. Lozano, I. García-Fernández, and F. Fernández, "Modeling, evaluation, and scale on artificial pedestrians: A literature review," *ACM Comput. Surv.*, vol. 50, no. 5, pp. 1–35, 2017.

[55] M. Chraibi, A. Tordeux, A. Schadschneider, and A. Seyfried, *Modelling of Pedestrian and Evacuation Dynamics*. Berlin, Germany: Springer, 2018, pp. 1–22.

[56] A. Schadschneider, W. Klingsch, H. Klüpfel, T. Kretz, C. Rogsch, and A. Seyfried, "Evacuation dynamics: Empirical results, modeling and applications," in *Encyclopedia of Complexity and Systems Science*. New York, NY, USA: Springer, 2009, pp. 3142–3176.

[57] D. C. Duives, W. Daamen, and S. P. Hoogendoorn, "State-of-the-art crowd motion simulation models," *Transp. Res. C, Emerg. Technol.*, vol. 37, pp. 193–209, Dec. 2013.

[58] A. Schadschneider, M. Chraibi, A. Seyfried, A. Tordeux, and J. Zhan, *Pedestrian Dynamics—From Empirical Results to Modeling*. Cham, Switzerland: Birkhäuser, 2018.

[59] H. Dong, M. Zhou, Q. Wang, X. Yang, and F.-Y. Wang, "State-of-the-art pedestrian and evacuation dynamics," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 5, pp. 1849–1866, May 2019.

[60] E. Cristiani, B. Piccoli, and A. Tosin, *Multiscale Modeling of Pedestrian Dynamics*, vol. 12. Switzerland: Springer, 2014.

[61] S. P. Hoogendoorn, F. L. M. van Wageningen-Kessels, W. Daamen, and D. C. Duives, "Continuum modelling of pedestrian flows: From microscopic principles to self-organised macroscopic phenomena," *Phys. A, Stat. Mech. Appl.*, vol. 416, pp. 684–694, Dec. 2014.

[62] M. Chraibi, U. Kemloh, A. Schadschneider, and A. Seyfried, "Force-based models of pedestrian dynamics," *Netw. Heterogeneous Media*, vol. 6, no. 3, p. 425, 2011.

[63] C. Totzeck, "An anisotropic interaction model with collision avoidance," *Kinetic Rel. Models*, vol. 13, no. 6, pp. 1219–1242, 2020.

[64] W. Yu, R. Chen, L. Dong, and S. Dai, "Centrifugal force model for pedestrian dynamics," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 72, no. 2, Aug. 2005, Art. no. 026112.

[65] M. Chraibi, A. Seyfried, and A. Schadschneider, "Generalized centrifugal-force model for pedestrian dynamics," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 82, no. 4, Oct. 2010, Art. no. 046111.

[66] A. Nakayama, K. Hasebe, and Y. Sugiyama, "Instability of pedestrian flow and phase structure in a two-dimensional optimal velocity model," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 71, no. 3, Mar. 2005, Art. no. 036121.

[67] G. Köster, F. Treml, and M. Gödel, "Avoiding numerical pitfalls in social force models," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 87, no. 6, Jun. 2013, Art. no. 063305.

[68] X. Chen, M. Treiber, V. Kanagaraj, and H. Li, "Social force models for pedestrian traffic–state of the art," *Transp. Rev.*, vol. 38, no. 5, pp. 625–653, Sep. 2018.

[69] M. Moussaïd, D. Helbing, and G. Theraulaz, "How simple rules determine pedestrian behavior and crowd disasters," *Proc. Nat. Acad. Sci. USA*, vol. 108, no. 17, pp. 6884–6888, 2011.

[70] I. Karamouzas, B. Skinner, and S. J. Guy, "A universal power law governing pedestrian interactions," *Phys. Rev. Lett.*, vol. 113, no. 5, Dec. 2014, Art. no. 238701.

[71] Y.-X. Lü, Z.-X. Wu, and J.-Y. Guan, "Pedestrian dynamics with mechanisms of anticipation and attraction," *Phys. Rev. Res.*, vol. 2, no. 4, Nov. 2020, Art. no. 043250.

[72] S. Paris, J. Pettré, and S. Donikian, "Pedestrian reactive navigation for crowd simulation: A predictive approach," *Comput. Graph. Forum*, vol. 26, no. 3, pp. 665–674, 2007.

[73] J. Van den Berg, M. Lin, and D. Manocha, "Reciprocal velocity obstacles for real-time multi-agent navigation," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2008, pp. 1928–1935.

[74] S. Pellegrini, A. Ess, K. Schindler, and L. Van Gool, "You'll never walk alone: Modeling social behavior for multi-target tracking," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 261–268.

[75] S. Guy *et al.*, "ClearPath: Highly parallel collision avoidance for multi-agent simulation," in *Proc. ACM SIGGRAPH/Eurographics Symp. Comput. Animation*, New Orleans, LA, USA, 2009, pp. 177–187.

[76] S. Guy, M. Lin, and D. Manocha, "Modeling collision avoidance behavior for virtual humans," in *Proc. 9th Int. Conf. Auto. Agents Multiagent Syst.*, Toronto, ON, Canada, vol. 2, 2010, pp. 575–582.

[77] S. Kim *et al.*, "BRVO: Predicting pedestrian trajectories using velocity-space reasoning," *Int. J. Robot. Res.*, vol. 34, no. 2, pp. 201–217, Feb. 2015.

[78] K. Guo, D. Wang, T. Fan, and J. Pan, "VR-ORCA: Variable responsibility optimal reciprocal collision avoidance," *IEEE Robot. Autom. Lett.*, vol. 6, no. 3, pp. 4520–4527, Jul. 2021.

[79] B. Maury, J. Venel, A.-H. Olivier, and S. Donikian, "A mathematical framework for a crowd motion model," *Comp. Rendus Math.*, vol. 346, pp. 1245–1250, 2008.

[80] B. Maury and J. Venel, "A discrete contact model for crowd motion," *ESAIM: Math. Model. Numer. Anal.*, vol. 45, no. 1, pp. 145–168, Jan. 2011.

[81] Z. Yin, J. Liu, and L. Wang, "Less-effort collision avoidance in virtual pedestrian simulation," in *Proc. Int. Conf. Artif. Intell. Comput. Sci.*, 2019, pp. 488–493.

[82] J. Charlton, L. R. M. Gonzalez, S. Maddock, and P. Richmond, "Fast simulation of crowd collision avoidance," in *Proc. Comput. Graph. Int. Conf.* Cham, Switzerland: Springer, 2019, pp. 266–277.

[83] J. Ondřej and J. Pettré, "A synthetic-vision based steering approach for crowd simulation," *ACM Trans. Graph.*, vol. 29, p. 123, Jul. 2010.

[84] F. Dietrich and G. Köster, "Gradient navigation model for pedestrian dynamics," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 89, no. 6, Jun. 2014, Art. no. 062801.

[85] A. Tordeux, M. Chraibi, and A. Seyfried, "Collision-free speed model for pedestrian dynamics," in *Traffic and Granular Flow*. Cham, Switzerland: Springer, 2016, pp. 225–232.

[86] Q. Xu, M. Chraibi, A. Tordeux, and J. Zhang, "Generalized collision-free velocity model for pedestrian dynamics," *Phys. A, Stat. Mech. Appl.*, vol. 535, Dec. 2019, Art. no. 122521.

[87] U. Weidmann, "Transporttechnik der fußgänger—Transporttechnische eigenschaften des fußgängerverkehrs (literaturauswertung)," ETH Zürich, Schriftenreihe des IVT, Zürich, Switzerland, Tech. Rep. 90, 1993.

[88] M. Fukui and Y. Ishibashi, "Self-organized phase transitions in cellular automaton models for pedestrians," *J. Phys. Soc. Jpn.*, vol. 68, p. 2861, Aug. 1999.

[89] M. Muramatsu, T. Irie, and T. Nagatani, "Jamming transition in pedestrian counter flow," *Phys. A, Stat. Mech. Appl.*, vol. 267, nos. 3–4, pp. 487–498, 1999.

[90] V. J. Blue and J. L. Adler, "Cellular automata microsimulation of bidirectional pedestrian flows," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1678, no. 1, pp. 135–141, Jan. 1999.

[91] H. Klüpfel, T. Meyer-König, J. Wahle, and M. Schreckenberg, "Microscopic simulation of evacuation processes on passenger ships," in *Theory and Practical Issues on Cellular Automata*, S. Bandini and T. Worsch, Eds. Berlin, Germany: Springer, 2000.

[92] C. Burstedde, A. Kirchner, K. Klauck, A. Schadschneider, and J. Zittartz, "Cellular automaton approach to pedestrian dynamics—Applications," in *Pedestrian and Evacuation Dynamics*, M. Schreckenberg and S. D. Sharma, Eds. Berlin, Germany: Springer, 2002, pp. 87–98.

[93] A. Kirchner and A. Schadschneider, "Simulation of evacuation processes using a bionics-inspired cellular automaton model for pedestrian dynamics," *Phys. A, Stat. Mech. Appl.*, vol. 312, pp. 260–276, Sep. 2002.

[94] A. Schadschneider, "Cellular automaton approach to pedestrian dynamics—Theory," in *Pedestrian and Evacuation Dynamics*, M. Schreckenberg and S. D. Sharma, Eds. Berlin, Germany: Springer, 2002, pp. 75–86.

[95] E. Ben-Jacob, "From snowflake formation to growth of bacterial colonies. Part II. Cooperative formation of complex colonial patterns," *Contemp. Phys.*, vol. 38, pp. 205–241, May 1997.

[96] A. Kirchner, K. Nishinari, and A. Schadschneider, "Friction effects and clogging in a cellular automaton model for pedestrian dynamics," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 67, May 2003, Art. no. 056122.

[97] C. von Krüchten and A. Schadschneider, "A cognitive, decision-based model for pedestrian dynamics," in *Traffic and Granular Flow*. Cham, Switzerland: Springer, 2020, pp. 141–147.

[98] C. Von Krüchten and A. Schadschneider, "Concept of a decision-based pedestrian model," *Collective Dyn.*, vol. 5, pp. 316–323, Mar. 2020.

[99] *Forschungszentrum Jülich, Jülich Supercomputing Centre. Data Archive of Experiments on Pedestrian Dynamic.* Accessed: Aug. 1, 2021. [Online]. Available: http://ped.fz-juelich.de/dataarchive

[100] I. Hasan *et al.*, "Forecasting people trajectories and head poses by jointly reasoning on tracklets and vislets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 4, pp. 1267–1278, Apr. 2019.

[101] D. Helbing, I. Farkas, and T. Vicsek, "Simulating dynamical features of escape panic," *Nature*, vol. 407, no. 6803, pp. 487–490, 2000.

[102] A. Treuille, S. Cooper, and Z. Popović, "Continuum crowds," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 1160–1168, Jul. 2006.

[103] R. L. Hughes, "A continuum theory for the flow of pedestrians," *Transp. Res. B, Methodol.*, vol. 36, no. 6, pp. 507–535, 2002.

[104] C. Bechinger, R. Di Leonardo, H. Löwen, C. Reichhardt, G. Volpe, and G. Volpe, "Active particles in complex and crowded environments," *Rev. Modern Phys.*, vol. 88, no. 4, Nov. 2016, Art. no. 045006.

[105] B. D. Hankin and R. A. Wright, "Passenger flow in subways," *J. Oper. Res. Soc.*, vol. 9, no. 2, pp. 81–88, 1958.

[106] S. Older, "Movement of pedestrians on footways in shopping streets," *Traffic Eng. Control.*, vol. 10, no. 4, pp. 160–163, 1968.

[107] F. P. Navin and R. J. Wheeler, "Pedestrian flow characteristics," *Traffic Eng., Inst Traffic Engr*, vol. 39, pp. 31–36, Jul. 1969.

[108] A. Seyfried, B. Steffen, W. Klingsch, and M. Boltes, "The fundamental diagram of pedestrian movement revisited," *J. Stat. Mech., Theory Exp.*, vol. 2005, no. 10, Oct. 2005, Art. no. P10002.

[109] U. Chattaraj, A. Seyfried, and P. Chakroborty, "Comparison of pedestrian fundamental diagram across cultures," *Adv. Complex Syst.*, vol. 12, no. 3, pp. 393–405, 2009.

[110] J. Zhang, W. Klingsch, A. Schadschneider, and A. Seyfried, "Ordering in bidirectional pedestrian flows and its influence on the fundamental diagram," *J. Stat. Mech., Theory Exp.*, vol. 2012, no. 2, Feb. 2012, Art. no. P02002.

[111] R. Subaih, M. Maree, M. Chraibi, S. Awad, and T. Zanoon, "Gender-based insights into the fundamental diagram of pedestrian dynamics," in *Proc. Int. Conf. Comput. Collective Intell.* Cham, Switzerland: Springer, 2019, pp. 613–624.

[112] A. Schadschneider and A. Seyfried, "Validation of CA models of pedestrian dynamics with fundamental diagrams," *Cybern. Syst.*, vol. 40, no. 5, pp. 367–389, 2009.

[113] S. N. Gomes, A. M. Stuart, and M.-T. Wolfram, "Parameter estimation for macroscopic pedestrian dynamics models from microscopic data," *SIAM J. Appl. Math.*, vol. 79, no. 4, pp. 1475–1500, Jan. 2019.

[114] J. Cristín, V. Méndez, and D. Campos, "General scaling in bidirectional flows of self-avoiding agents," *Sci. Rep.*, vol. 9, p. 18488, Dec. 2019.

[115] G. H. Goldsztein, "Self-organization when pedestrians move in opposite Directions. Multi-lane circular track model," *Appl. Sci.*, vol. 10, no. 2, p. 563, Jan. 2020.

[116] N. Bain and D. Bartolo, "Dynamic response and hydrodynamics of polarized crowds," *Science*, vol. 363, no. 6422, pp. 46–49, Jan. 2019.

[117] M. Friesen, H. Gottschalk, B. Rüdiger, and A. Tordeux, "Spontaneous wave formation in stochastic self-driven particle systems," *SIAM J. Appl. Math.*, vol. 81, no. 3, pp. 853–870, Jan. 2021.

[118] D. Helbing, I. J. Farkas, and T. Vicsek, "Freezing by heating in a driven mesoscopic system," *Phys. Rev. Lett.*, vol. 84, no. 6, pp. 1240–1243, 2000.

[119] H. Stanley, "Non-equilibrium physics: Freezing by heating," *Nature*, vol. 404, p. 718, Apr. 2000.

[120] D. Helbing and A. Johansson, *Pedestrian, Crowd and Evacuation Dynamics*. New York, NY, USA: Springer, 2009, pp. 6476–6495.

[121] J. Cividini, C. Appert-Rolland, and H. J. Hilhorst, "Diagonal patterns and Chevron effect in intersecting traffic flows," *Europhys. Letters*, vol. 102, no. 2, p. 20002, Apr. 2013.

[122] A. Nicolas, S. Ibáñez, M. N. Kuperman, and S. Bouzat, "A counterintuitive way to speed up pedestrian and granular bottleneck flows prone to clogging: Can 'more' escape faster?" *J. Stat. Mech., Theory Exp.*, vol. 2018, no. 8, Aug. 2018, Art. no. 083403.

[123] G. Hermann and J. Touboul, "Heterogeneous connections induce oscillations in large-scale networks," *Phys. Rev. Lett.*, vol. 109, no. 1, Jul. 2012, Art. no. 018702.

[124] M. Moussaïd, J. E. Kämmer, P. P. Analytis, and H. Neth, "Social influence and the collective dynamics of opinion formation," *PLoS ONE*, vol. 8, no. 11, pp. 1–8, 2013.

[125] J. D. Touboul, "The hipster effect: When anti-conformists all look the same," *Discrete Continuous Dyn. Syst. B*, vol. 24, p. 4379, Oct. 2019.

[126] F. Schweitzer, *Brownian Agents and Active Particles: Collective Dynamics in the Natural and Social Sciences*. Berlin, Germany: Springer, 2003.

[127] S. Ramaswamy, "The mechanics and statistics of active matter," *Annu. Rev. Condens. Matter Phys.*, vol. 1, no. 1, pp. 323–345, Aug. 2010.

[128] T. Vicsek and A. Zafeiris, "Collective motion," *Phys. Rep.*, vol. 517, nos. 3–4, pp. 71–140, Aug. 2012.

[129] M. C. Marchetti *et al.*, "Hydrodynamics of soft active matter," *Rev. Modern Phys.*, vol. 85, no. 3, p. 1143, 2013.

[130] J. Elgeti, R. G. Winkler, and G. Gompper, "Physics of microswimmers—Single particle motion and collective behavior: A review," *Rep. Prog. Phys.*, vol. 78, no. 5, May 2015, Art. no. 056601.

[131] D. Needleman and Z. Dogic, "Active matter at the interface between materials science and cell biology," *Nature Rev. Mater.*, vol. 2, no. 9, pp. 1–14, Sep. 2017.

[132] A. R. Dulaney and J. F. Brady, "Machine learning for phase behavior in active matter systems," *Soft Matter*, vol. 17, no. 28, pp. 6808–6816, 2021.

[133] T. Korhonen, S. Hostikka, S. Heliövaara, and H. Ehtamo, "FDS+Evac: An agent based fire evacuation model," in *Pedestrian and Evacuation Dynamics*, W. W. F. Klingsch, C. Rogsch, A. Schadschneider, and M. Schreckenberg, Eds. Berlin, Germany: Springer, 2010, pp. 109–120.

[134] S. Curtis, A. Best, and D. Manocha, "Menge: A modular framework for simulating crowd movement," *Collective Dyn.*, vol. 1, pp. 1–40, Mar. 2016.

[135] P. Taillandier *et al.*, "Building, composing and experimenting complex spatial models with the GAMA platform," *GeoInformatica*, vol. 23, no. 2, pp. 299–322, Apr. 2019.

[136] B. Kleinmeier, B. Zönnchen, M. Gödel, and G. Köster, "Vadere: An open-source simulation framework to promote interdisciplinary understanding," *Collective Dyn.*, vol. 4, pp. 1–18, Sep. 2019.

[137] F. Giuliari, I. Hasan, M. Cristani, and F. Galasso, "Transformer networks for trajectory forecasting," in *Proc. 25th Int. Conf. Pattern Recognit. (ICPR)*, Jan. 2021, pp. 10335–10342.

[138] L. Achaji, J. Moreau, T. Fouqueray, F. Aioun, and F. Charpillet, "Is attention to bounding boxes all you need for pedestrian action prediction?" 2021, *arXiv:2107.08031*.

[139] B. Ivanovic, K. Leung, E. Schmerling, and M. Pavone, "Multimodal deep generative models for trajectory prediction: A conditional variational autoencoder approach," *IEEE Robot. Automat. Lett.*, vol. 6, no. 2, pp. 295–302, Apr. 2020.

[140] N. Schneider and D. M. Gavrila, "Pedestrian path prediction with recursive Bayesian filters: A comparative study," in *Proc. German Conf. Pattern Recognit.* Berlin, Germany: Springer, 2013, pp. 174–183.

[141] M. Meuter, U. Iurgel, S.-B. Park, and A. Kummert, "The unscented Kalman filter for pedestrian tracking from a moving host," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2008, pp. 37–42.

[142] P. Trautman, J. Ma, R. M. Murray, and A. Krause, "Robot navigation in dense human crowds: Statistical models and experimental studies of human–robot cooperation," *Int. J. Robot. Res.*, vol. 34, no. 3, pp. 335–356, 2015.

[143] M. Bennewitz, W. Burgard, G. Cielniak, and S. Thrun, "Learning motion patterns of people for compliant robot motion," *Int. J. Robot. Res.*, vol. 24, no. 1, pp. 31–48, 2005.

[144] P. Trautman and A. Krause, "Unfreezing the robot: Navigation in dense, interacting crowds," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2010, pp. 797–803.

[145] C. G. Keller and D. M. Gavrila, "Will the pedestrian cross? A study on pedestrian path prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 2, pp. 494–506, Apr. 2013.

[146] D. Ellis, E. Sommerlade, and I. Reid, "Modelling pedestrian trajectory patterns with Gaussian processes," in *Proc. IEEE 12th Int. Conf. Comput. Vis. Workshops, ICCV Workshops*, Sep. 2009, pp. 1229–1234.

[147] K. Kim, D. Lee, and I. Essa, "Gaussian process regression flow for analysis of motion trajectories," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 1164–1171.

[148] S. Ferguson, B. Luders, R. C. Grande, and J. P. How, "Real-time predictive modeling and robust avoidance of pedestrians with uncertain, changing intentions," in *Algorithmic Foundations of Robotics XI*. Cham, Switzerland: Springer, 2015, pp. 161–177.

[149] Y. F. Chen, M. Everett, M. Liu, and J. P. How, "Socially aware motion planning with deep reinforcement learning," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 1343–1350.

[150] Z. Wan, C. Jiang, M. Fahad, Z. Ni, Y. Guo, and H. He, "Robot-assisted pedestrian regulation based on deep reinforcement learning," *IEEE Trans. Cybern.*, vol. 50, no. 4, pp. 1669–1682, Nov. 2018.

[151] F. Martinez-Gil, M. Lozano, I. García-Fernández, P. Romero, D. Serra, and R. Sebastián, "Using inverse reinforcement learning with real trajectories to get more trustworthy pedestrian simulations," *Mathematics*, vol. 8, no. 9, p. 1479, Sep. 2020.

[152] H. Kretzschmar, M. Spies, C. Sprunk, and W. Burgard, "Socially compliant mobile robot navigation via inverse reinforcement learning," *Int. J. Robot. Res.*, vol. 35, no. 11, pp. 1289–1307, 2016.

[153] K. M. Kitani, B. D. Ziebart, J. A. Bagnell, and M. Hebert, "Activity forecasting," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2012, pp. 201–214.

[154] M. Wulfmeier, D. Z. Wang, and I. Posner, "Watch this: Scalable cost-function learning for path planning in urban environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 2089–2095.

[155] T. V. Le, S. Liu, and H. C. Lau, "A reinforcement learning framework for trajectory prediction under uncertainty and budget constraint," in *Proc. ECAI*, 2016, pp. 347–354.

[156] F. Martinez-Gil, M. Lozano, and F. Fernández, "MARL-Ped: A multi-agent reinforcement learning based framework to simulate pedestrian groups," *Simul. Model. Pract. Theory*, vol. 47, pp. 259–275, Sep. 2014.

[157] F. Martinez-Gil, M. Lozano, and F. Fernández, "Multi-agent reinforcement learning for simulating pedestrian navigation," in *Proc. Int. Workshop Adapt. Learn. Agents*. Berlin, Germany: Springer, 2011, pp. 54–69.

[158] N. Rhinehart and K. M. Kitani, "First-person activity forecasting from video with online inverse reinforcement learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 2, pp. 304–317, Feb. 2018.

[159] N. Lee, W. Choi, P. Vernaza, C. B. Choy, P. H. S. Torr, and M. Chandraker, "DESIRE: Distant future prediction in dynamic scenes with interacting agents," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 336–345.

[160] J. Lee, J. Won, and J. Lee, "Crowd simulation by deep reinforcement learning," in *Proc. 11th Annu. Int. Conf. Motion, Interact., Games*, 2018, pp. 1–7.

[161] M. Everett, Y. F. Chen, and J. P. How, "Collision avoidance in pedestrian-rich environments with deep reinforcement learning," *IEEE Access*, vol. 9, pp. 10357–10377, 2021.

[162] W. van Toll and J. Pettré, "Algorithms for microscopic crowd simulation: Advancements in the 2010s," *Comput. Graph. Forum*, vol. 40, no. 2, pp. 731–754, May 2021.

[163] S. Yi, H. Li, and X. Wang, "Pedestrian behavior understanding and prediction with deep neural networks," in *Computer Vision*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham, Switzerland: Springer, 2016, pp. 263–279.

[164] T. Fernando, S. Denman, S. Sridharan, and C. Fookes, "Soft+ hard-wired attention: An LSTM framework for human trajectory prediction and abnormal event detection," *Neural Netw.*, vol. 108, pp. 466–478, Dec. 2018.

[165] N. Nikhil and B. Tran Morris, "Convolutional neural network for trajectory prediction," in *Proc. Eur. Conf. Comput. Vis. (ECCV) Workshops*. Cham, Switzerland: Springer, 2018, pp. 186–196.

[166] A. Vemula, K. Muelling, and J. Oh, "Social attention: Modeling attention in human crowds," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2018, pp. 4601–4607.

[167] H. Xue, D. Q. Huynh, and M. Reynolds, "SS-LSTM: A hierarchical LSTM model for pedestrian trajectory prediction," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2018, pp. 1186–1194.

[168] A. Sadeghian, V. Kosaraju, A. Sadeghian, N. Hirose, H. Rezatofighi, and S. Savarese, "SoPhie: An attentive GAN for predicting paths compliant to social and physical constraints," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 1349–1358.

[169] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[170] S. Li, W. Li, C. Cook, C. Zhu, and Y. Gao, "Independently recurrent neural network (IndRNN): Building a longer and deeper RNN," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 5457–5466.

[171] A. Azzouni and G. Pujolle, "A long short-term memory recurrent neural network framework for network traffic matrix prediction," 2017, *arXiv:1705.05690*.

[172] A. Graves, A.-R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 6645–6649.

[173] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 3104–3112.

[174] A. Graves, "Generating sequences with recurrent neural networks," 2013, *arXiv:1308.0850*.

[175] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, "Show and tell: A neural image caption generator," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3156–3164.

[176] H. Manh and G. Alaghband, "Scene-LSTM: A model for human trajectory prediction," 2018, *arXiv:1808.04018*.

[177] M. Pfeiffer, G. Paolo, H. Sommer, J. Nieto, R. Siegwart, and C. Cadena, "A data-driven model for interaction-aware pedestrian motion prediction in object cluttered environments," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2018, pp. 5921–5928.

[178] M. Lisotto, P. Coscia, and L. Ballan, "Social and scene-aware trajectory prediction in crowded spaces," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 1–8.

[179] S. Haddad, M. Wu, H. Wei, and S. K. Lam, "Situation-aware pedestrian trajectory prediction with spatio-temporal attention model," in *Proc. 24th Comput. Vis. Winter Workshop*, 2019, pp. 4–16.

[180] Y. Huang, H. Bi, Z. Li, T. Mao, and Z. Wang, "STGAT: Modeling spatial–temporal interactions for human trajectory prediction," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 6272–6281.

[181] C. Tao, Q. Jiang, L. Duan, and P. Luo, "Dynamic and static context-aware LSTM for multi-agent motion prediction," in *Computer Vision*. Cham, Switzerland: Springer, 2020, pp. 547–563.

[182] A. Monti, A. Bertugli, S. Calderara, and R. Cucchiara, "DAG-Net: Double attentive graph neural network for trajectory forecasting," in *Proc. 25th Int. Conf. Pattern Recognit. (ICPR)*, Jan. 2021, pp. 2551–2558.

[183] S. Haddad and S.-K. Lam, "Self-growing spatial graph network for context-aware pedestrian trajectory prediction," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2021, pp. 1029–1033.

[184] I. Hasan, F. Setti, T. Tsesmelis, A. Del Bue, F. Galasso, and M. Cristani, "MX-LSTM: Mixing tracklets and vislets to jointly forecast trajectories and head poses," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6067–6076.

[185] N. Bisagno, B. Zhang, and N. Conci, "Group LSTM: Group trajectory prediction in crowded scenarios," in *Proc. Eur. Conf. Comput. Vis. (ECCV) Workshops*. Cham, Switzerland: Springer, 2018, pp. 213–225.

[186] B. Cheng, X. Xu, Y. Zeng, J. Ren, and S. Jung, "Pedestrian trajectory prediction via the social-grid LSTM model," *J. Eng.*, vol. 2018, no. 16, pp. 1468–1474, 2018.

[187] F. Bartoli, G. Lisanti, L. Ballan, and A. Del Bimbo, "Context-aware trajectory prediction," in *Proc. 24th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2018, pp. 1941–1946.

[188] B. Ivanovic and M. Pavone, "The trajectron: Probabilistic multi-agent trajectory modeling with dynamic spatiotemporal graphs," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 2375–2384.

[189] A. Syed and B. T. Morris, "SSeg-LSTM: Semantic scene segmentation for trajectory prediction," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 2504–2509.

[190] T. Zhao et al., "Multi-agent tensor fusion for contextual trajectory prediction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 12126–12134.

[191] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural. Inf. Process. Syst.*, vol. 25, 2012, pp. 1097–1105.

[192] Y. Bengio, *Learning Deep Architectures for AI*. Boston, MA, USA: Now Publishers Inc, 2009.

[193] M. Fabbri, F. Lanzi, S. Calderara, A. Palazzi, R. Vezzani, and R. Cucchiara, "Learning to detect and track visible and occluded body joints in a virtual world," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*. Cham, Switzerland: Springer, 2018, pp. 430–446.

[194] E. Ristani and C. Tomasi, "Features for multi-target multi-camera tracking and re-identification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6036–6046.

[195] Y.-C. Yoon, A. Boragule, Y.-M. Song, K. Yoon, and M. Jeon, "Online multi-object tracking with historical appearance matching and scene adaptive detection filtering," in *Proc. 15th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, Nov. 2018, pp. 1–6.

[196] A. Mohamed, K. Qian, M. Elhoseiny, and C. Claudel, "Social-STGCNN: A social spatio-temporal graph convolutional neural network for human trajectory prediction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 14424–14432.

[197] F. Leon and M. Gavrilescu, "A review of tracking and trajectory prediction methods for autonomous driving," *Mathematics*, vol. 9, no. 6, p. 660, Mar. 2021.

[198] E. Rehder, F. Wirth, M. Lauer, and C. Stiller, "Pedestrian prediction by planning using deep neural networks," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 5903–5908.

[199] K. M. Abughalieh and S. G. Alawneh, "Predicting pedestrian intention to cross the road," *IEEE Access*, vol. 8, pp. 72558–72569, 2020.

[200] S. Yan, Y. Xiong, and D. Lin, "Spatial temporal graph convolutional networks for skeleton-based action recognition," in *Proc. 32nd AAAI Conf. Artif. Intell.*, 2018, pp. 1–10.

[201] G. Cheron, I. Laptev, and C. Schmid, "P-CNN: Pose-based CNN features for action recognition," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 3218–3226.

[202] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*.

[203] X. Dan, "Spatial–temporal block and LSTM network for pedestrian trajectories prediction," 2020, *arXiv:2009.10468*.

[204] Y. Xu, Z. Piao, and S. Gao, "Encoding crowd interaction with deep neural network for pedestrian trajectory prediction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 5275–5284.

[205] A. Jain et al., "Discrete residual flow for probabilistic pedestrian behavior prediction," in *Proc. Conf. Robot Learn.*, 2020, pp. 407–419.

[206] D. Ridel, N. Deo, D. Wolf, and M. Trivedi, "Scene compliant trajectory forecast with agent-centric spatio–temporal grids," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 2816–2823, Apr. 2020.

[207] D. Varshneya and G. Srinivasaraghavan, "Human trajectory prediction using spatially aware deep attention models," 2017, *arXiv:1705.09436*.

[208] C. Yu, X. Ma, J. Ren, H. Zhao, and S. Yi, "Spatio-temporal graph transformer networks for pedestrian trajectory prediction," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2020, pp. 507–523.

[209] C. Zhang, C. Berger, and M. Dozza, "Social-IWSTCNN: A social interaction-weighted spatio-temporal convolutional neural network for pedestrian trajectory prediction in urban traffic scenarios," 2021, *arXiv:2105.12436*.

[210] Z. Zhao and C. Liu, "STUGCN: A social spatio-temporal unifying graph convolutional network for trajectory prediction," in *Proc. 6th Int. Conf. Autom., Control Robot. Eng. (CACRE)*, Jul. 2021, pp. 546–550.

[211] S. Zamboni, Z. T. Kefato, S. Girdzijauskas, C. Norén, and L. Dal Col, "Pedestrian trajectory prediction with convolutional neural networks," *Pattern Recognit.*, vol. 121, Jan. 2022, Art. no. 108252.

[212] L. Huang, J. Zhuang, X. Cheng, R. Xu, and H. Ma, "STI-GAN: Multimodal pedestrian trajectory prediction using spatiotemporal interactions and a generative adversarial network," *IEEE Access*, vol. 9, pp. 50846–50856, 2021.

[213] T. Salzmann, B. Ivanovic, P. Chakravarty, and M. Pavone, "Trajectron++: Dynamically-feasible trajectory forecasting with heterogeneous data," in *Computer Vision*, A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, Eds. Cham, Switzerland: Springer, 2020, pp. 683–700.

[214] M. Arjovsky and L. Bottou, "Towards principled methods for training generative adversarial networks," 2017, *arXiv:1701.04862*.

[215] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training GANs," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 2234–2242.

[216] J. Amirian, J.-B. Hayet, and J. Pettré, "Social ways: Learning multimodal distributions of pedestrian trajectories with GANs," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2019, pp. 2964–2972.

[217] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever, and P. Abbeel, "InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets," in *Proc. 30th Int. Conf. Neural Inf. Process. Syst.*, 2016, pp. 2180–2188.

[218] V. Kosaraju, A. Sadeghian, R. Martín-Martín, I. Reid, S. H. Rezatofighi, and S. Savarese, "Social-BiGAT: Multimodal trajectory forecasting using bicycle-GAN and graph attention networks," 2019, *arXiv:1907.03395*.

[219] Z. Lv, X. Huang, and W. Cao, "An improved GAN with transformers for pedestrian trajectory prediction models," *Int. J. Intell. Syst.*, vol. 37, no. 8, pp. 4417–4436, Aug. 2022.

[220] W.-C. Lai et al., "Trajectory prediction in heterogeneous environment via attended ecology embedding," in *Proc. 28th ACM Int. Conf. Multimedia*, Oct. 2020, pp. 202–210.

[221] T. Fernando, S. Denman, S. Sridharan, and C. Fookes, "GD-GAN: Generative adversarial networks for trajectory prediction and group detection in crowds," in *Proc. Asian Conf. Comput. Vis.* Cham, Switzerland: Springer, 2018, pp. 314–330.

[222] P. Kothari, S. Kreiss, and A. Alahi, "Human trajectory forecasting in crowds: A deep learning perspective," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 1–15, Jul. 2021.

[223] A. Corbetta, A. Muntean, and K. Vafayi, "Parameter estimation of social forces in pedestrian dynamics models via a probabilistic method," *Math. Biosci. Eng.*, vol. 12, no. 2, pp. 337–356, 2015.

[224] R. Lovreglio, E. Ronchi, and D. Nilsson, "Calibrating floor field cellular automaton models for pedestrian dynamics by using likelihood function optimization," *Phys. A, Stat. Mech. Appl.*, vol. 438, pp. 308–320, Nov. 2015.

[225] A. Tordeux and A. Schadschneider, "White and relaxed noises in optimal velocity models for pedestrian flow with stop-and-go waves," *J. Phys. A, Math. Theor.*, vol. 49, no. 18, May 2016, Art. no. 185101.

[226] N. W. F. Bode and E. Ronchi, "Statistical model fitting and model selection in pedestrian dynamics research," *Collective Dyn.*, vol. 4, pp. 1–32, Apr. 2019.

[227] C. Feliciani and K. Nishinari, "An improved cellular automata model to simulate the behavior of high density crowd and validation by experimental data," *Phys. A, Stat. Mech. Appl.*, vol. 451, pp. 135–148, Jun. 2016.

[228] M. Li, Y. Zhao, L. He, W. Chen, and X. Xu, "The parameter calibration and optimization of social force model for the real-life 2013 Ya'an earthquake evacuation in China," *Saf. Sci.*, vol. 79, pp. 243–253, Nov. 2015.

[229] T. Robin, G. Antonini, M. Bierlaire, and J. Cruz, "Specification, estimation and validation of a pedestrian walking behavior model," *Transp. Res. B, Methodol.*, vol. 43, no. 1, pp. 36–56, 2009.

[230] M. Ko, T. Kim, and K. Sohn, "Calibrating a social-force-based pedestrian walking model based on maximum likelihood estimation," *Transportation*, vol. 40, no. 1, pp. 91–107, 2013.

[231] M. A. Nielsen, *Neural Networks and Deep Learning*, vol. 25. San Francisco, CA, USA: Determination Press, 2015.

[232] S. Göttlich and C. Totzeck, "Parameter calibration with stochastic gradient descent for interacting particle systems driven by neural networks," *Math. Control, Signals, Syst.*, vol. 34, pp. 1–30, Mar. 2021.

[233] J. Amirian, B. Zhang, F. V. Castro, J. J. Baldelomar, J.-B. Hayet, and J. Pettré, "OpenTraj: Assessing prediction complexity in human trajectories datasets," in *Proc. Asian Conf. Comput. Vis.*, Nov. 2020, pp. 1–17.

[234] S. Zamboni, Z. T. Kefato, S. Girdzijauskas, C. Norén, and L. Dal Col, "Pedestrian trajectory prediction with convolutional neural networks," *Pattern Recognit.*, vol. 121, Jan. 2022, Art. no. 108252.

[235] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, pp. 1–48, Dec. 2019.

[236] C. Schöller, V. Aravantinos, F. Lay, and A. Knoll, "What the constant velocity model can teach us about pedestrian motion prediction," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 1696–1703, Apr. 2020.

[237] H. Cheng, W. Liao, M. Y. Yang, B. Rosenhahn, and M. Sester, "AMENet: Attentive maps encoder network for trajectory prediction," *ISPRS J. Photogramm. Remote Sens.*, vol. 172, pp. 253–266, Feb. 2021.

[238] X. Song et al., "Pedestrian trajectory prediction based on deep convolutional LSTM network," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3285–3302, Jun. 2020.

[239] S. Hossain, F. T. Johora, J. P. Müller, S. Hartmann, and A. Reinhardt, "SFMGNet: A physics-based neural network to predict pedestrian trajectories," 2022, *arXiv:2202.02791*.

[240] Y. Zhu, D. Qian, D. Ren, and H. Xia, "StarNet: Pedestrian trajectory prediction using deep neural network in star topology," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Nov. 2019, pp. 8075–8080.

[241] F. Pasquale, "Toward a fourth law of robotics: Preserving attribution, responsibility, and explainability in an algorithmic society," *Ohio St. LJ*, vol. 78, p. 1243, Oct. 2017.

[242] D. Pedreschi, F. Giannotti, R. Guidotti, A. Monreale, S. Ruggieri, and F. Turini, "Meaningful explanations of black box AI decision systems," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, 2019, pp. 9780–9784.

[243] J. Kawale et al., "A graph-based approach to find teleconnections in climate data," *Stat. Anal. Data Mining, ASA Data Sci. J.*, vol. 6, no. 3, pp. 158–179, Jun. 2013.

[244] J. H. Faghmous, I. Frenger, Y. Yao, R. Warmka, A. Lindell, and V. Kumar, "A daily global mesoscale ocean eddy dataset from satellite altimetry," *Sci. Data*, vol. 2, no. 1, pp. 1–16, Dec. 2015.

[245] G. Hautier, C. C. Fischer, A. Jain, T. Mueller, and G. Ceder, "Finding nature's missing ternary oxide compounds using machine learning and density functional theory," *Chem. Mater.*, vol. 22, no. 12, pp. 3762–3767, 2010.

[246] C. C. Fischer, K. J. Tibbetts, D. Morgan, and G. Ceder, "Predicting crystal structure by merging data mining with quantum mechanics," *Nature Mater.*, vol. 5, no. 8, pp. 641–646, Aug. 2006.

[247] L. Li et al., "Understanding machine-learned density functionals," *Int. J. Quantum Chem.*, vol. 116, no. 11, pp. 819–833, Jun. 2016.

[248] K. C. Wong, L. Wang, and P. Shi, "Active model with orthotropic hyperelastic material for cardiac image analysis," in *Proc. Int. Conf. Funct. Imag. Model. Heart*. Springer, 2009, pp. 229–238.

[249] X. Xu, J. L. Sapp, A. R. Dehaghani, F. Gao, M. Horacek, and L. Wang, "Robust transmural electrophysiological imaging: Integrating sparse and dynamic physiological models into ECG-based inference," in *Proc. Int. Conf. Med. Image Comput.-Assist. Intervent*. Cham, Switzerland: Springer, 2015, pp. 519–527.

[250] K.-H. Lee, G. Ros, J. Li, and A. Gaidon, "SPIGAN: Privileged adversarial learning from simulation," 2018, *arXiv:1810.03756*.

[251] L. von Rueden, S. Mayer, R. Sifa, C. Bauckhage, and J. Garcke, "Combining machine learning and simulation to a hybrid modelling approach: Current and future directions," in *Proc. Int. Symp. Intell. Data Anal.* Cham, Switzerland: Springer, 2020, pp. 548–560.

[252] A. Khadka, P. Remagnino, and V. Argyriou, "Synthetic crowd and pedestrian generator for deep learning problems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Barcelona, Spain, May 2020, pp. 4052–4056.

[253] J. Willard, X. Jia, S. Xu, M. Steinbach, and V. Kumar, "Integrating scientific knowledge with machine learning for engineering and environmental systems," 2020, *arXiv:2003.04919*.

[254] M. Silvestri, M. Lombardi, and M. Milano, "Injecting domain knowledge in neural networks: A controlled experiment on a constrained problem," in *Proc. Int. Conf. Integr. Constraint Program., Artif. Intell., Oper. Res.* Springer, 2021, pp. 266–282.

[255] A. Karpatne et al., "Theory-guided data science: A new paradigm for scientific discovery from data," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 10, pp. 2318–2331, Jun. 2017.

[256] M. Bahari, I. Nejjar, and A. Alahi, "Injecting knowledge in data-driven vehicle trajectory predictors," *Transp. Res. C, Emerg. Technol.*, vol. 128, Jul. 2021, Art. no. 103010.

[257] S. Göttlich and S. Knapp, "Artificial neural networks for the estimation of pedestrian interaction forces," in *Crowd Dynamics*, vol. 2. Springer, 2020, pp. 11–32.

[258] S. Kreiss, "Deep social force," 2021, *arXiv:2109.12081*.

[259] Z. Zhang, D. Lu, J. Li, P. Liu, and G. Zhang, "Crowd evacuation simulation using hierarchical deep reinforcement learning," in *Proc. IEEE 24th Int. Conf. Comput. Supported Cooperat. Work Design (CSCWD)*, May 2021, pp. 563–568.

[260] A. Tordeux, M. Chraibi, A. Seyfried, and A. Schadschneider, "Prediction of pedestrian speed with artificial neural networks," in *Proc. Int. Conf. Traffic Granular Flow*. Springer, 2017, pp. 327–335.

[261] Y. Ma, E. W. M. Lee, and R. K. K. Yuen, "An artificial intelligence-based approach for simulating pedestrian movement," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 11, pp. 3159–3170, Nov. 2016.

[262] M. Schreckenberg and S. D. Sharma, *Pedestrian and Evacuation Dynamics*. Berlin, Germany: Springer, 2002.

Raphael Korbmacher received the B.Sc. degree in engineering and the M.Sc. degree in quality engineering from the University of Wuppertal, Germany, in 2018 and 2021, respectively, and the B.Sc. degree in economics from FernUniversity, Hagen, Germany. He is currently pursuing the Ph.D. degree with the University of Wuppertal, Germany, under the direction of Prof. Antoine Tordeux with focus on the use of deep learning methods in reliability engineering and traffic science.

Antoine Tordeux received the Diploma degree in mathematics from Gustave Eiffel University, France, in 2007, and the Ph.D. degree in mathematics from Gustave Eiffel University in 2010, under the direction of Prof. Michel Roussignol and Prof. Sylvain Lassarre. From 2013 to 2018, he worked at Forschungszentrum Jülich, Germany, with Prof. Armin Seyfried and heads currently the Chair for Traffic Safety and Reliability as a Junior Professor with the University of Wuppertal, Germany. His research topics concern the modeling, analysis and simulation of pedestrian and road traffic flow, traffic networks, and evacuation dynamics.