ISSN (ELECTRONIC): 2945-3135

Kumar, Abhishek (2024)*. Dynamic trajectory forecasting in crowded spaces: AnIn-Depth review of online and adaptive prediction methods. International Journal of Multidisciplinary Research & Reviews. 3(04), 87-102.*

INTERNATIONAL JOURNAL OF
MULTIDISCIPLINARY RESEARCH & REVIEWS

journal homepage: www.ijmrr.online/index.php/home

DYNAMIC TRAJECTORY FORECASTING IN CROWDED SPACES: AN IN-DEPTH REVIEW OF ONLINE AND ADAPTIVE PREDICTION METHODS

Abhishek Kumar

Assistant Professor Maulana Mazharul Haque A & P University, Patna, Bihar, India.

How to Cite the Article: ***Kumar, Abhishek (2024)****. Dynamic trajectory forecasting in crowded spaces: AnIn-Depth review of online and adaptive prediction methods. International Journal of Multidisciplinary Research & Reviews. 3(04), 87-102.*

DOI Information: **https://doi.org/10.56815/IJMRR.V314.2024/87-102**

Keywords

Pedestrian trajectory prediction, Automatic driving, Deep learning, Prediction method, Neural-network

Abstract

As they navigate city streets, autonomous cars will unavoidably collide with people. Research on pedestrian trajectory prediction is crucial for preventing route conflicts with pedestrians. This study summarizes the current public dataset for pedestrian trajectoryprediction and compares the performance of various methods. It also examines the pros and cons of depth learning-based trajectory prediction methods. In conclusion, we look forward to the current stateof pedestrian trajectory prediction and the difficulties and trends in its growth.

1. INTRODUCTION

The worldwide automobile business is booming along with the fast-paced urbanization and expanding global economy. By 2050, experts predict that 2.5 billion automobiles will circulate worldwide. One reason cars are so standard is their ease in people's daily lives. The concurrent rise in traffic congestion and the frequency of traffic accidents poses profound often overlooked threats to motorists' well-being on the road. As members of the most defenseless categories involved in road traffic, pedestrians and cyclists are particularly at risk when they do not have access to appropriate safety gear. The National Bureau of Statistics in China reports that road-vulnerable categories comprise 26% of all accident fatalities in the country each year. Hence, we need to find a solution to the issue if we want to make roads and pedestrian areas safer in the future and cut down on accidents and property damage. One critical approach to reducing traffic accidents is to improve the trajectory prediction of people crossing the street by autonomous cars. Theability to depict, perceive, and analyse human motion is fundamental for intelligent systems and people to live and interact. In the study of human mobility, prediction is crucial. "Over time, the model may anticipate scenarios involving several agents and actively incorporate this scene data, i.e., to improve the outcomes of active perception, predictive planning, model predictive control, or human-computer interaction." Consequently, several domains, including autonomous cars, intelligent transportation, command cities, service robots, etc., have devoted considerable resources to studying pedestrian trajectory prediction in the last few years. Prior to autonomous cars being widely used, traffic safety must be ensured [1]. Avoiding collisions and ensuring pedestrians are safe is possible if the autonomous vehicle can correctly identify where people are walking in the traffic scene. Predicting the paths of pedestrians is currently plagued by three major issues. (1) Predicting pedestrians' future trajectories is challenging because pedestrians are a more malleable player in the traffic scene. However, more sophisticated algorithms can approximately forecast pedestrian trajectories by looking at the paths taken by its past moments. In contrast to automobiles and other traffic players, pedestrians have more freedom of movement in real-world circumstances, including the ability to turn, stop, and proceed at will. Creating a kinematic model that works for pedestrians is challenging; drivers often need help to foretell walkers' paths. (2) A pedestrian's future movement in a real-world traffic scene is impacted by their free will, the traffic environment, and other pedestrians in the area, including factors like strolling in groups and interactions. Research by Moussaid M et al. [2] showsthat 70% of pedestrians travel in groups and that these groups often engage in simultaneous interactions. Simulating this interaction between pedestrians in the algorithm is challenging since it is highly abstract. Third, discovering a function that translates input to output is the implementation of the traditional pedestrian trajectory prediction algorithm concept. This is analogous to the mapping between various sequences in the trajectory prediction model. A compromised state (a forecast result that tends to anticipate a compromise trajectory) may be easily achieved using the traditional model or training procedure. “The traditional training model cannot predict the pedestrian's trajectory; the interaction between pedestrians and their surroundings becomes complicated and abstract due to the effect of both the pedestrians' subjective intentions and the objective environment.” Due to its limited contextual adaptability and inability to handle interactions in responsible situations, the classic pedestrian trajectory prediction model has limited prediction performance. Advancements in deep learning have allowed neural networks to achieve remarkable success in areas such as tracking, image recognition, and classification. When applied to studying pedestrian movements, deep learning meets all the requirements thanks to its comprehensive theoretical framework and robust network models. “Specifically, the most popular networks for pedestrian trajectory prediction modelling are generative adversarial networks (GANs), graph convolutional networks (GCNs), and recurrent neural networks (RNNs) for sequence learning.”

2. LITERATURE REVIEW

Guo et al. [3] proposed a novel method using a directional pooling grid for each agent of interest. There was a relationship between the relative velocities of the agent of interest and each cell in the grid. This technique detailed the scene and their interactions by simulating the individuals' relative locations and velocities. Two maps focused on the ego-vehicle were also used in a comparable portrayal by Kamenev et al. [4]. A bird's- eye view of the sky revealed other vehicles as rectangles on an occupancy map, each with its velocity vector. A lack of explicit modelling of interactions was observed. You may represent interactions using various approaches that do not rely on graphs. By combining the relative locations of the agent of interest with other agents in the scenario, Gupta et al. [5] created social pooling. This was sent via a max pooling layer and an MLP for processing. This approach enabled the network to transmit interaction data among all scene agents. Instead of using explicit characteristics, Mangalam et al. [6] used the encoded representations of trajectories and waypoints for each agent in their social pooling application. They developed a non-local attention mechanism to merge the agents' representations to teach the network how to pool resources. Several

ISSN (ELECTRONIC): 2945-3135
Kumar, Abhishek (2024). *Dynamic trajectory forecasting in crowded spaces: AnIn-Depth review of online and adaptive prediction methods. International Journal of Multidisciplinary Research & Reviews. 3(04), 87-102.*

ways use attention processes, particularly transformer layers, to encode the interactions between agents. “These methods include those by Sun and Sun [8], Xie et al. et al. ng et al. et al., and Ngiam et al. et al. Chen et al. [12] successfully modelled agent interactions by performing conditional predictions based on cross-attention between the encoded agent trajectory and the intents of surrounding agents. “Additionally, conditional predictions were used by Sun et al. [13] and Li et al. [14] in order to include interactions. To condition the forecast of the influenced agent on the influencer's marginal prediction, Sun et al. [13] first categorized agents as either influencers or impacted agents. In their goal-conditioned prediction technique,Li et al. [14] allowed the trajectories of involved agents to be affected by the objectives of other agents inthe interaction”.

Similarly, Tsao et al. [15] presented an approach whereby a Transformer encoder was trained to model interactions using two pretext tasks: first, determining the kind of interaction, and second, determining if the two agents were socially or non-socially near. The grid that was built around the agent of interest was a n × m grid, and it was represented by a sparse tensor that included the encoded previous trajectories of other agents (Wang et al. [16]). Using a convolutional auto encoder, they proceeded to simulate the interactions onthis grid. Potential fields are an additional, intriguing way to model interactions [17]. The authors constructed a potential field to encircle the agent of interest. This field indicated the agent's cost, which was more significant in areas with other agents and infrastructural barriers. This technique taught the trajectory predictor to steer clear of impediments and other actors.

Mangalam et al. [18] used waypoints connecting the most recent observed locations to the predicted target location as an extra measure. In this case, the model forecasted a distribution of probabilities for the endpoint and the intermediate checkpoints. In contrast to Dendorfer et al. [19], who solely used goal distribution estimation to model the task's multimodality and claimed that it made the trajectory decoder's job easier, Mangalam et al. [18] used predicted distributions to sample goals and waypoints and they generated a conditioned probability distribution for each timestep of the trajectory. Using this method, Chiara et al. [20] sampled from distributions of goals and introduced random noise into the moduleresponsible for generating trajectories. Chiara et al. [20] just projected non-parametric probability distributions for potential objectives in image space, in contrast to Mangalam et al. [18] who conducted all their calculations in image space. A recurrent network that dealt with the 2D locations of the agent ofinterest formed the basis of the trajectory creation module. “Mangalam et al. [19] supplied an encoded segmentation map of the environment to both their goal and waypoint module and their trajectory prediction module, in contrast to Dendorfer et al. [19] and Chiara et al. [20], who solely used images to include scene information in their goal estimation modules”. The authors Gu et al. [21] avoided explicitly estimating the goal probability distribution in favor of intensively sampling goals based on the likely future lanes of the street network, which they scored first. “After that, the attention mechanism and a softmax function were used to determine the likelihood of each objective.” The authors used a trainable goal set predictor to choosethe most probable objectives before employing this set. So far, every technique has used a position-based probability distribution model to choose a subset of potential conditioning objective locations.

The first use of GANs for human trajectory prediction was SocialGAN by Gupta et al. [5]. “To create new trajectories, SocialGAN employs a basic LSTM-based generator to encode previous ones, then runs the output through a social pooling module (see Section 3.2 for more information on this), and finally decodes the result.” Concatenating the social pooling module's output with a random noise

Kumar, Abhishek (2024). Dynamic trajectory forecasting in crowded spaces: AnIn-Depth review of online and adaptive prediction methods. International Journal of Multidisciplinary Research & Reviews. 3(04), 87-102.

vector from the standard distribution introduces diversity and uncertainty, making this technique unimodal. When training an adversarial network, the encoder module acts as a discriminator. Using scene and interaction modeling, the approaches SocialBiGAT and Sophie developed by Kosaraju et al. [22] and Sadeghian et al. [23], respectively, improved upon SocialGAN. SoPhie used SocialGAN's framework but swapped out its social pooling module with one that relies on attention-based interaction. Agent's distance from its neighbors is used as a metric for feature fusion. The input of the interaction module was the encoded features, decoder module hidden states, and scene characteristics retrieved from a scene picture using a convolutional neural network (CNN). In addition, Kosaraju et al. [22] revised their approach to model interactions. Using the scene agents' LSTM-encoded attributes, SocialBiGAT simulates the interaction with a GAT [24]. During adversarial training, two people act as monitors. In SocialGAN, the local discriminator remains unchanged. The global discriminator, on the other hand, uses a convolutional neural network (CNN) to distinguish between different types of produced trajectories according to physical scene restrictions. Incorporating multimodality, the authors trained a latent encoder to map the produced trajectorie back to the latent space, drawing inspiration from CycleGAN [25]. Using the SocialGAN architecture, which is focused on

pedestrians, Wang et al. [26] extended it to car highway situations. This required swapping out the interaction module with a social module based on autoencoders (for more on this, see Section 3.2), but otherwise, the framework remained the same. Even though they used a pyramid network to extract characteristics from trajectories at varying temporal resolutions, Li et al. [27] also used adversarial trainingto solve the pedestrian trajectory prediction issue. The trajectories were decrypted after the fused features were applied using a CNN. The model's generator module introduced unimodal noise, resulting in various paths. A state refinement module was also added to the original SocialGAN by Zhou et al. [28]. A GAT- based interaction module was substituted for SocialGAN's social pooling module in this state refinement module. “This module organized the agents into groups and then used the interactions between groups and agents inside them to create a scene graph. This enabled finer-grained interaction modeling compared to the suggested social pooling module.”

A. Trajectory prediction/imputation

This class of application cases includes imputation of trajectories [32], trajectory prediction in artificial [29], urban [30], and marine [31] settings. “As a result of its interchange ability with movement prediction [31] and future 3location prediction (FLP) [22], the next location/destination prediction use case is not always clear when discussing trajectory prediction.” However, this use case is different from the one we will talk about later—predicting the following location or destination—because it is about the future path of a mover or the route the mover will take to its next place. Training using high-resolution data to forecast trajectory routes with great spatial information is essential. Consequently, these models' training data is composed of resampled, generalized, or discretized trajectories of individual movers, which are relatively dense. For example, Capobianco et al. [33] build 12-position, fixed-length trajectories with a defined duration of three hours by resampling each trajectory to a fifteen-minute fixed sampling frequency. Before feeding them into an LSTM, they are fed three vertices at a time. Mehri et al. [31] use context-aware piece-wise linear segmentation to generalize AIS trajectories. Their model can now provide very accurate, short-term trajectory predictions.

B. Arrival time prediction

Examples of use cases that fall under this category are rail network arrival time prediction [35] and street network arrival time prediction [36]. The acronym ETA stands for "estimated time of arrival," another name for this assignment. Routing algorithms that combine journey time information with data from street networks are often used to calculate ETA in traditional intelligent transportation systems (ITS). This trip time data may be derived from more sophisticated ML/DL prediction models or historical averages for a particular season, day of the week, and time of day. Recurrent mechanisms often forecast trip times because they align with past travel times (e.g., at certain times and on specific days) [3, 36]. One study that uses aggregated trajectories to train GNNs for Google Maps journey time forecasts is Derrow-Pinion et al. [36]. Segment and super segment embedding vectors make up the GNN graph. Nodes keep track of information about individual street segments, such as average current and past segment travel times and speeds, segment length, and road type.

In contrast, edges track information about whole super segments, such as real-time travel times for those segments. Alternatively, instead of using aggregated trajectories for training, Wang et al. [37] provide the GEO-convolutional network layer (GEO-Conv), which is also used by Buijse et al. [35]). Using dense trajectories as input, the proposed GEO-Conv layer first applies a non-linear mapping to each trajectory point, then uses a GEO-Conv step with multiple kernels.

C. Anomaly detection

This class of use cases is concerned with finding and dealing with anomalies, outliers, or otherwise odd data observations and patterns. Many people use the two words interchangeably. "Deviate so much from other observations as to arouse suspicions that a different mechanism generated it" is one way to describe an outlier finding. Similarly, out-of-the-ordinary patterns are known as anomalies [38]. We say there have been anomalous movements when "the trajectories of individual movers are unremarkable, but their combined spatio-temporal pattern is unusual" [39]. “Other movement anomalies include anomalous recordings, which have unusual spatiotemporal or thematic properties, and anomalous (sub) trajectories."Data with ground truth labels is uncommon since anomalies are typically defined differently depending on the context. Anomaly detection methods look for significant deviations from observed trajectories using geographical, spatiotemporal, or other metrics to find unusual trajectories. For instance, GeoTrackNet [40] is a model for detecting anomalies in maritime trajectories. It incorporates a probabilistic representation of AIS tracks based on Recurrent Neural Networks (RNNs) and a contrario detection [41]. After GeoTrackNet identified any anomalies, AIS specialists assessed them.

Similarly, Singh et al. [42] provide a method for detecting anomalies such as unexpected turns, on-off switching, and trajectories using RNN regression models. Once again, the absence of ground truth data makes a quantitative accuracy study impractical. Some studies use synthetically created outliers to compensate for the absence of ground truth [43]. To automatically identify movement abnormalities, such as synthetic anomalies in taxi trajectories, Liatsikou et al. [43] built an LSTM-based network. “Since the autoencoder only accepts inputs of a certain defined length, all trajectories are trimmed to nine points, and shorter ones are rejected, limiting the duration of the anomalies that may be discovered.”

D. Next location/destination prediction

Predicting where a journey will end up is the purview of this use case category [44]. In most cases, it comes down to selecting an ultimate destination or following a site from a limited pool of possibilities. Along with GPS traces, social media check-ins (such as those from Foursquare) are popular data sources in this area. Next, we need to figure out where we will be checking in. Attention methods are pretty standard for predicting one's upcoming location. For instance, Gao et al. [45]

train VANext, a convolutional network for semi- supervised network trajectories using check-in data. Using a causal embedding approach analogous to a high- order Markov process, they transform each user's itinerary into sequence embeddings. This includes the check-in and POI sequences. Their GRU learns the trajectory patterns from the resultant embeddings.

Additionally, they focus on the embeddings to forecast the user's next point of interest. In order to create separate latent vectors from sparse and significant trajectories, Feng et al. [46] modify two attention methods. After that, a historical attention module and its DeepMove GRUs get these embeddings. The learned attention weights may easily explain predictions based on the user's past movement behavior. The STSAN is a new network that Li et al. [47] present. “The trajectory embeddings are created by merging the time (activity sequence), space (distance matrices of locations), and location attentions (location sequence and their categories).” Based on these embeddings, they employ a softmax layer to forecast the user's subsequent POI. To address the issue of heterogeneity, they use a federated learning framework. “Liao and colleagues [48] design embedding using both location sequences and graph embedding's-where and what one does their MCARNN multi-task context-aware recurrent neural network can handle both activity and location prediction tasks using graphs representing locations.”

E. Traffic volume prediction

Traffic volume forecast on roadway segments [49], human activity prediction at particular POIs [50] and metropolitan regions [51], and animal movement dynamics prediction [52] are all examples of use cases that fall under the umbrella of crowd or traffic flow prediction. Aggregated trajectory data is often used to train models for this use case. To illustrate the point, the training material for the Traffic4Cast 2021 competition was traffic movies. The objective was to forecast traffic in the face of changes in both space and time, such as the COVID-19 epidemic and the transfer to other cities. After using CNN (U-Net) and multi-tasklearning, Lu [51] was able to win this challenge. Using a multi-task learning strategy, they train the U-Net model to simultaneously anticipate future traffic conditions for several cities by randomly sampling from all accessible cities. Similar to Wang et al. [53], the traffic movie method is also used. Their approach involves combining individual-level trajectories into a grid. Inflow is the sum of all the incoming traffic from other areas during a specific time frame, while outflow is the sum of all the traffic leaving the region. Zhang et al. [54] also construct time-series grids with average traffic speeds and cab arrivals per cell. Another popular method for this application is graph-based modeling. To illustrate their GCN, Li et al. [55] combine CDR data into a graph, where nodes indicate geographical statistical units and edges show their relationships (physical distance, physical movement, phone calls, etc.). Like this, Lippert et al. [56] construct temporal networks using bird migration data, with nodes representing radar sites and edges representing flows between the Voronoi tessellation cells of those locations. Alternatively, one might follow the instructions given by Buroni et al. [57], who construct and train a Direct LSTM encoder-decoder model using vehicle counts obtained from GPS tracks.

The goal of training the model was to anticipate the number of cars per edge for the Belgian highway network at each time step. “In a similar vein, Gao et al [50] train a GCN+GRU model to anticipate these visit counts by measuring the number of cars per POI each time step (hourly) using their GPS tracks.” Also, to leverage (and fine-tune) pre-trained natural language processing models like Bert, Roberta, GPT-2, and XLNet to forecast daily POI visit counts, Xue et al. [58] suggest a translator they call mobility prompting.

ISSN (ELECTRONIC): 2945-3135

Kumar, Abhishek (2024). *Dynamic trajectory forecasting in crowded spaces: AnIn-Depth review of online and adaptive prediction methods. International Journal of Multidisciplinary Research & Reviews. 3(04), 87-102.*

3. INTRODUCTION OF PEDESTRIAN TRAJECTORY PREDICTION METHOD

It is possible to classify pedestrian trajectory prediction algorithms into two broad categories: those that rely on deep learning and those that rely on conventional probability prediction. The pedestrian prediction is reformulated as a probability prediction issue using the usual probability prediction approach. The goal of developing the pedestrian kinematics model was to predict how the pedestrian's motion state would evolve in order to derive the pedestrian's trajectory. When the situation is relatively certain—that is, when pedestrians are prepared to cross the road—Nicholas Schneider et al. [59] utilized a dynamic Bayesian network to ascertain pedestrian trajectories based on pedestrian head posture, the proximity of pedestrians and vehicles to the point of collision, and the distance between pedestrians and roadside barriers. “QM The balanced Gaussian process dynamics model (B-GPDMs) was used by Raul et al. [60] to extract three- dimensional time-related information from the pedestrian skeleton's key points into a low-dimensional space.” They then obtained multiple models of the four behaviors of walking, stopping, starting, and standing while using the model. They picked the most similar model to predict the pedestrians' future path, posture, and intention. Nevertheless, these approaches are noise-sensitive and unable to represent rich ambient features, complicated and dynamic motion characteristics, and real-world people. There are significant practical issues with the limited time span in which they may receive the forecast findings. Researchers started using deep learning approaches to pedestrian trajectory prediction challenges as deep learning advanced. RNN-based, GAN-based, and GCN-based pedestrian trajectory prediction are the most popular deep learning-based approaches. Using RNN for pedestrian trajectory prediction. The network that learns from past data, which predicts the path of a pedestrian, was first attempted using RNN. By feeding data into the historical network, the outcome is calculated. Using this function, RNN can extrapolate values from past sequences. With its recursive organizational structure and focus on sequence modeling, RNN has exceptional modeling skills in time analysis and sequence learning. However, RNN's limitations become more apparent with longer time series. Due to RNN's inability to acquire long-term memory of the data state, the network layer stops learning. When training big networks, RNNs will cause gradients to either vanish or explode since they keep all past information in the network. When it comes to operational efficiency, RNN models are not very good. “They have slow training and reasoning speeds since their present states depend on the hidden states of the past moments, which makes parallel processing impossible.” The accuracy of pedestrian trajectory prediction relies on training the network using massivedata sets and many network nodes. Consequently, pedestrian trajectory prediction is one area where conventional RNNs fall short. In order to address the issue of transmitting information in lengthy sequences, researchers have created similar variations based on RNN networks, such as GRUs and long short-term memory networks (LSTM). These variants use gating techniques to control the flow and selection of information. Using RNN research as a focal point, the variant above techniques have accomplished outstanding results in pedestrian trajectory prediction [61]. Pedestrian trajectory prediction based on GCN. A convolutional neural network based on graphs GCN is a way to train and learn from graph data, which allows for deep learning on graph data. This data comes from the graph's edges and nodes. Despite RNN's strong sequence modeling capabilities, it lacks a comprehensible high-level spatio-temporal framework. The area of pedestrian trajectory prediction deals with unpredictable pedestrian numbers and interactions. One obvious way to depict pedestrian interaction is using a graph structure. It outperforms the aggregation-based RNN technique in terms of effectiveness and ease of use. The processing of graph data in non-Euclidean space is significantly impacted by GCN

[62]. The basic premise is to use deep learning techniques to convert the graph's nodes and edges into vector space, after which grouping and classification may occur. Graph structure is now the foundational component of several approaches [63]. These approaches often employ a combination of deep sequence models, including long short-term memory networks, to simulate pedestrians, with their interactions serving as connections and pedestrians themselves represented as nodes.

When GCN incorporates spatio-temporal data into pedestrian trajectory prediction, it better understands pedestrian behavior and can simulate social interaction more quickly. While GCN's potential uses in trajectory prediction are vast, the network's shallowness, instability, and lack of adaptive capacity remain issues. Predicting the path of pedestrians using the GAN technique. The network that generates hostile content By playing a game between the generation model and the discriminant model, the GAN network, an unsupervised deep learning model, can overcome the challenge of computing the generation probability. Incorporating a GAN network into pedestrian trajectory prediction helps address the limitation of being able to predict only one "optimal" route. The network is capable of predicting several realistic paths and, using game theory, may further maximize the accuracy of its predictions. “Simultaneously, the trajectory sampler

utilizes the GAN input random vector in conjunction with the hidden representation of other pedestrian trajectories to handle the interaction between all observed pedestrians, building on the successful experience of GAN networks in super-resolution, image conversion, and image synthesis.” Mode collapse and sluggish convergence are two issues that might arise in a neural network that is based on a GAN model.

F. Trajectory prediction method based on shallow learning

An early approach to trajectory prediction used a combination of the fundamental kinematic model, the Bayesian filter, and its extensions to project the present state into the future. Methods based on multi-model interaction and methods based on a single kinematic model were compared by Schneider et al. [66]. Findings demonstrate that a 30-centimeter reduction in lateral position estimate error is achievable using the multi-model interaction-based approach while navigating. “But, the model's assumption of constant speed makes it hard for the Bayesian filter-based model to capture pedestrians' switching dynamics, and there are not enough data sets or motion types tested to support the development and prediction of more complicated motion models To depict nonlinear and time-varying dynamics, Pavlovic et al. [67] used the switched linear dynamic system (SLDS) model [68]”. In order to anticipate nonlinear motion in the real world, the model— which relies on a Markov chain for probability transfer—switches between several linear kinematic models. Although this model is helpful for simpler motion models, the information about motion features it uses is not always enough to facilitate state switching. It does not work as well for more complicated models. For complicated motion model testing, bigger motion capture data sets are required for accuracy. Kooij et al.

[69] developed a context-based dynamic Bayesian network (DBN) model to predict pedestrian paths. “This model controls the switching state of the SLDS model by combining context information (pedestrian head direction, emergency degree, and environmental space layout) as a potential state and adding it to the top of the model. This model can provide more precise forecasts compared to the SLDS model.” While SLDS model-based prediction methods and DBN model-based prediction methods consume a lot of computing power during model prediction reasoning and mathematical model building, they do an excellent job of reflecting basic motion types (like turning) and additional scenes (like traffic lights and crosswalks) based on the SLDS model. A model of appealing and

repulsive pedestrian motion, the social force model, was suggested by Helbing et al. [70]. Robotics and activity recognition makes heavy use of this approach [71]. While Alahi et al. [72] suggested displaying social affinity characteristics by analyzing pedestrian paths in a crowd to determine their relative positions, Yi et al. [73] suggested using human traits to enhance crowd forecasts. In order to keep up with the ever-increasing advancements in machine learning, kinematics-based approaches make use of tracking algorithms derived from machine learning, such as the Kalman filter (KF), Markov model (MM), and Gaussian process [74] (GP). When solving the pedestrian trajectory prediction issue, the KF model excels because it quickly achieves high prediction accuracy while efficiently processing data from noise-free sites.

On the contrary, the accuracy of predictions is severely impacted by the high model complexity and the substantial prediction error over extended periods. As noise levels rise, the KF model becomes more sensitive, leading to a linear decline in prediction accuracy. While the MM performs well for predicting the current stage of a pedestrian's movement process, it is not well-suited for predicting the trajectory of a pedestrian over the medium to long term due to its sensitivity to fluctuations in the trajectory. "In contrast to high-order MM, which considerably increases the computational complexity of the model, first-order MM simply takes into account the impact of the present pedestrian trajectory point on the future trajectory point, limiting the use of data information from the historical trajectory point." Assuming the latent variables follow the Gaussian distribution, the GP model offers a non-parametric approach to probability prediction. Historical trajectory data is used to learn the GP model's anticipated trajectory [75], and an appropriatekernel function is specified in the GP model to clarify the uncertainty associated with line prediction using trajectory modeling. As a result of its ability to anticipate data with noise points, GP models can circumvent the fact that trajectory data is not discrete and accurately portray the statistical features of the distribution of pedestrian trajectories.

To summarize, the shallow learning-based trajectory prediction technique has made some early progress andcontributed to the development of pedestrian trajectory prediction. "Unfortunately, there is a significant disparity between the real situation and the final prediction impact because of the limits of kinematics-based approaches, inadequate extraction of motion feature information, lack of unique scene information, and the complexity of model creation." Predicting more complicated pedestrian motion models and scenarios using conventional approaches is challenging.

G. Trajectory Prediction Method Based on RNN

In order to produce an output that is sequential in time, the RNN first encodes the incoming sequence data into a hidden representation of a defined size. A second RNN is then used to decode the hidden representation. Long Short-Term Memory (LSTM) networks can learn and repeat lengthy sequences but are not good at picking up on relationships between different related sequences. For this reason, the social pooling layer of S-LSTM [76] was suggested, which would heuristically compile data from nearby pedestrians. One of the earliest recurrent neural network applications, S-LSTM has become standard for pedestrian trajectory prediction by simulating human-to-human interaction in congested environments. You can see the results of several RNN-based pedestrian trajectory prediction algorithms in Table 1. One approach to categorizing prediction models is by the information they draw from their neighbors; these may be models that are based on past states [77] or models that are based on current outcomes (speed, position, etc.) [78]. Specific RNN-based pedestrian trajectory prediction systems use the attention mechanism to provide several approaches to neighbor weight calculation, allowing for the differentiation of neighbor significance. To illustrate, a combination of "soft attention" and "hard attention" in an attention model allows for the calculation of the pairwise speed correlation [79] and the soft attention score based on the hidden state. “Two temporal attention mechanisms are developed to compute the hidden state vectors of the position and velocity LSTM layers, based on the LSTM's position-velocity-time-attention model [80].” Pedestrian interaction and trajectory prediction credibility are both improved by attention mechanism deployment.

4. CONCLUSION

Finally, dynamic trajectory forecasting in congested areas involves many moving parts, and dealing with the complexity of people and vehicle movements calls for innovative and flexible approaches. Substantial improvements have been made in the examined online and adaptive prediction algorithms' capacity to model and forecast trajectories in real-time. Machine learning, statistical modeling, and heuristic methods are just a few of the numerous methodologies that these approaches use. Each of these methods has its own set of strengths and weaknesses. Applications in dynamic settings, including autonomous driving and crowd management, are perfect for online prediction techniques because they thrive in circumstances that need quick reactions and continual updates. Conversely, adaptive approaches improve resilience and dependability over time by enhancing prediction accuracy by incorporating new data and adjusting models to emerging trends. These technologies may be integrated to provide a holistic solution by combining the advantages of real-time responsiveness with adaptive learning capabilities. More work must be done to make these approaches more accurate and efficient, especially in dense and dynamic situations. Deploying these technologies in realistic, real-world applications also requires tackling computing problems and assuring scalability. Ultimately, as dynamic trajectory forecasting algorithms continue to advance, we should be able to anticipate better and control complicated movement patterns, leading to increased efficiency, safety, and system performance in congested areas.

5. AUTHORS CONTRIBUTION

The authors agreed to have no connections or engagements with any group or body that provides financial and non-financial assistance for the topics and resources covered in the article.

6. ACKNOWLEDGEMENT

Individuals / resources participated in the work are acknowledged properly.

7. SOURCES OF FUNDING

The authors received no financial aid to support the study.

8. PLAGIARISM POLICY

The authors declare that any kind of violation of plagiarism, copyright, and ethical matters will be handled by all authors. Journalists and editors are not liable for the aforesaid matters.

9. CONFLICT OF INTEREST

The authors declared that no potential conflicts of interest concerning the research, authorship, and/or publication of this article.

REFERENCES

[1] Wang, Jinxiang, Dai, et al. Output-feedback robust control for vehicle path tracking considering different human drivers' characteristics[J]. Mechatronics: The Science of Intelligent Machines, 2018, 50:402-412.

[2] MOUSSAID M, PEROZO N, GARNIER S, et al. The walking behaviour of pedestrian social groups and its impact on crowd dynamics[J]. PLoS ONE, 2010, 5(4): e10047.

[3] Guo, K.; Liu, W.; Pan, J. End-to-End Trajectory Distribution Prediction Based on Occupancy Grid Maps. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 2242–2251

[4] Kamenev, A.; Wang, L.; Bohan, O.B.; Kulkarni, I.; Kartal, B.; Molchanov, A.; Birchfield, S.; Nister, D.; Smolyanskiy, N. PredictionNet: Real-Time Joint Probabilistic Traffic Prediction for Planning, Control, and Simulation. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Philadelphia, PA, USA, 23–27 May 2022; pp. 8936–8942.

[5] Gupta, A.; Johnson, J.; Fei-Fei, L.; Savarese, S.; Alahi, A. Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; pp. 2255–2264

[6] Mangalam, K.; Girase, H.; Agarwal, S.; Lee, K.H.; Adeli, E.; Malik, J.; Gaidon, A. It Is Not the Journey But the Destination Endpoint Conditioned Trajectory Prediction. In Proceedings of the European Conference on Computer Vision (ECCV), Glasgow, UK, 23–28 August 2020; pp. 759–776.

[7] Huang, Z.; Mo, X.; Lv, C. Multi-modal Motion Prediction with Transformer-based Neural Network for Autonomous Driving. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Philadelphia, PA, USA, 23–27 May 2022; pp. 2605–2611

[8] Sun, H.; Sun, F. Social-Transformer: Pedestrian Trajectory Prediction in Autonomous Driving Scenes. In Proceedings of the International Conference on Cognitive Systems and Signal Processing (ICCSIP), Suzhou, China, 20–21 November 2021; Springer Science and Business Media Deutschland GmbH: Berlin/Heidelberg, Germany, 2021; pp. 177–190

[9] Xie, C.; Li, Y.; Liang, R.; Dong, L.; Li, X. Synchronous Bi-Directional Pedestrian Trajectory Prediction with Error Compensation. In Proceedings of the Asian Conference on Computer Vision (ACCV), Macau, China, 4–8 December 2022; pp. 2796–2812.

[10] Wang, Z.; Guo, J.; Hu, Z.; Zhang, H.; Zhang, J.; Pu, J. Lane Transformer: A High-Efficiency Trajectory Prediction Model. IEEE Open J. Intell. Transp. Syst. 2023, 4, 2–13.

[11] Ngiam, J.; Caine, B.; Vasudevan, V.; Zhang, Z.; Chiang, H.T.L.; Ling, J.; Roelofs, R.; Bewley, A.; Liu, C.; Venugopal, A.; et al. Scene transformer: A unified architecture for predicting future trajectories of multiple agents. In Proceedings of the International Conference on Learning Representations (ICLR), Virtual, 25–29 April 2022

[12] Chen, H.; Liu, Y.; Hu, C.; Zhang, X. Vulnerable Road User Trajectory Prediction for Autonomous Driving Using a Data-Driven Integrated Approach. IEEE Trans. Intell. Transp. Syst. 2023, 24, 7306–7317.

[13] Sun, Q.; Huang, X.; Gu, J.; Williams, B.C.; Zhao, H. M2I: From Factored Marginal Trajectory Prediction to Interactive Prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 6543–6552

[14] Li, D.; Zhang, Q.; Lu, S.; Pan, Y.; Zhao, D. Conditional Goal-oriented Trajectory Prediction for Interacting Vehicles with Vectorized Representation. IEEE Trans. Neural Netw. Learn. Syst. 2021, 1–13.

[15] Tsao, L.W.; Wang, Y.K.; Lin, H.S.; Shuai, H.H.; Wong, L.K.; Cheng, W.H. Social-SSL: Self-supervised Cross-Sequence Representation Learning Based on Transformers for Multi-agent

Trajectory Prediction. In Proceedings of the European Conference on Computer Vision (ECCV), Tel Aviv, Israel, 23–27 October 2022; pp. 234–250.

[16] Wang, Y.; Zhao, S.; Zhang, R.; Cheng, X.; Yang, L. Multi-Vehicle Collaborative Learning for Trajectory Prediction with SpatioTemporal Tensor Fusion. IEEE Trans. Intell. Transp. Syst. 2022, 23, 236–248

[17] Bhatt, N.P.; Khajepour, A.; Hashemi, E. MPC-PF: Socially and Spatially Aware Object Trajectory Prediction for Autonomous Driving Systems Using Potential Fields. IEEE Trans. Intell. Transp. Syst. 2023, 24, 5351–6361

[18] Mangalam, K.; An, Y.; Girase, H.; Malik, J. From Goals, Waypoints and Paths to Long Term Human Trajectory Forecasting. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Virtual, 11–17 October 2021; pp. 15233–15242

[19] Dendorfer, P.; Osep, A.; Leal-Taixe, L. Goal-GAN: Multimodal Trajectory Prediction Based on Goal Position Estimation. In Proceedings of the Asian Conference on Computer Vision (ACCV), Kyoto, Japan, 30 November–4 December 2020.

[20] Chiara, L.F.; Coscia, P.; Das, S.; Calderara, S.; Cucchiara, R.; Ballan, L. Goal-Driven Self-Attentive Recurrent Networks for Trajectory Prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, New Orleans, LA, USA, 18–24 June 2022; pp. 2518–2527

[21] Gu, J.; Sun, C.; Zhao, H. Densetnt: End-to-end trajectory prediction from dense goal sets. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Virtual, 11–17 October 2021; pp. 15303–15312

[22] Kosaraju, V.; Sadeghian, A.; Martín-Martín, R.; Reid, I.; Rezatofighi, S.H.; Savarese, S. Social-bigat: Multimodal trajectory forecasting using bicycle-gan and graph attention networks. In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), Vancouver, BC, Canada, 8–14 December 2019

[23] Sadeghian, A.; Kosaraju, V.; Sadeghian, A.; Hirose, N.; Rezatofighi, S.H.; Savarese, S. SoPhie: An Attentive GAN for Predicting Paths Compliant to Social and Physical Constraints. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 1349–1358.

[24] Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; Bengio, Y. Graph attention networks. In Proceedings of the International Conference on Learning Representations (ICLR), New Orleans, LA, USA, 6–9 May 2019.

[25] Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired Image-To-Image Translation Using Cycle-Consistent Adversarial Networks. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2223–2232

[26] Wang, Y.; Zhao, S.; Zhang, R.; Cheng, X.; Yang, L. Multi-Vehicle Collaborative Learning for Trajectory Prediction with SpatioTemporal Tensor Fusion. IEEE Trans. Intell. Transp. Syst. 2022, 23, 236–248.

[27] Li, Y.; Liang, R.; Wei, W.; Wang, W.; Zhou, J.; Li, X. Temporal pyramid network with spatial-temporal attention for pedestrian trajectory prediction. IEEE Trans. Netw. Sci. Eng. 2022, 9, 1006–1019

[28] Zhou, L.; Zhao, Y.; Yang, D.; Liu, J. GCHGAT: Pedestrian trajectory prediction using group constrained hierarchical graph attention networks. Appl. Intell. 2022, 52, 11434–11447

[29] Carroll M, Lin J, Paradise O, Georgescu R, Sun M, Bignell D, Milani S, Hofmann K, Hausknecht M, Dragan A, Devlin S (2022) Towards Flexible Inference in Sequential Decision Problems via Bidirectional Transformers. arXiv:2204.13326

[30] Fan Z, Yang X, Yuan W, Jiang R, Chen Q, Song X, Shibasaki R (2022) Online trajectory prediction for metropolitan scale mobility digital twin. In: Proceedings of the 30th International Conference on Advances in Geographic Information Systems, pp 1–12. ACM, Seattle Washington. https://doi.org/10. 1145/3557915.3561040. https://dl.acm.org/doi/10.1145/3557915.3561040

[31] Mehri S, Alesheikh AA, Basiri A (2021) A Contextual Hybrid Model for Vessel Movement Prediction. IEEE Access 9:45600–45613. https://doi.org/10.1109/ACCESS.2021.3066463. https://ieeexplore.ieee. org/document/9380635/

[32] Musleh M (2022) Towards a unified deep model for trajectory analysis. In: Proceedings of the 30th International Conference on Advances in Geographic Information Systems, pp 1–2. ACM, Seattle Washington. https://doi.org/10.1145/3557915.3565529. https://dl.acm.org/doi/10.1145/3557915.3565529

[33] Capobianco S, Forti N, Millefiori LM, Braca P, Willett P (2023) Recurrent Encoder-Decoder Networks for Vessel Trajectory Prediction With Uncertainty Estimation. IEEE Trans Aerosp Electron Syst 59(3):2554– 2565. https://doi.org/10.1109/TAES.2022.3216823. https://ieeexplore.ieee.org/document/9946391/

[34] Fan Z, Yang X, Yuan W, Jiang R, Chen Q, Song X, Shibasaki R (2022) Online trajectory prediction for metropolitan scale mobility digital twin. In: Proceedings of the 30th International Conference on Advances in Geographic Information Systems, pp 1–12. ACM, Seattle Washington. https://doi.org/10. 1145/3557915.3561040. https://dl.acm.org/doi/10.1145/3557915.3561040

[35] Buijse BJ, Reshadat V, Enzing OW (2021) A Deep Learning-Based Approach for Train Arrival Time Prediction. In: Yin H, Camacho D, Tino P, Allmendinger R, Tallón-Ballesteros AJ, Tang K, Cho SB, Novais P, Nascimento S (eds) Intelligent Data Engineering and Automated Learning - IDEAL 2021, vol 13113, pp 213–222. Springer International Publishing, Cham. series Title: Lecture Notes in Computer Science. https://doi.org/10.1007/978-3-030-91608-4_22. https://link.springer.com/10.1007/978-3-030- 91608-4_22

[36] Derrow-Pinion A, She J, Wong D, Lange O, Hester T, Perez L, Nunkesser M, Lee S, Guo X, Wiltshire B, Battaglia PW, Gupta V, Li A, Xu Z, Sanchez-Gonzalez A, Li Y, Velickovic P (2021) ETA Prediction with Graph Neural Networks in Google Maps. In: Proceedings of the 30th ACM International Conference on Information & Knowledge Management. pp 3767–3776. ACM, Virtual Event Queensland Australia. https://doi.org/10.1145/3459637.3481916. https://dl.acm.org/doi/10.1145/3459637.3481916

[37] Wang D, Zhang J, Cao W, Li J, Zheng Y (2018) When Will You Arrive? Estimating Travel Time Based on Deep Neural Networks. Proceedings of the AAAI Conference on Artificial Intelligence 32(1). https:// doi.org/10.1609/aaai.v32i1.11877. https://ojs.aaai.org/index.php/AAAI/article/view/11877

[38] Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: A survey. ACM Computing Surveys 41(3):1–58. https://doi.org/10.1145/1541880.1541882. http://portal.acm.org/citation.cfm? doid=1541880.1541882

[39] Hawkins DM (1980) Identification of outliers, vol 11. Springer

[40] Nguyen D, Vadaine R, Hajduch G, Garello R, Fablet R (2022) GeoTrackNet -A Maritime Anomaly Detector Using Probabilistic Neural Network Representation of AIS Tracks and A Contrario Detection. IEEE Trans Intell Transp Syst 23(6):5655–5667. https://doi.org/10.1109/TITS.2021.3055614. https:// ieeexplore.ieee.org/document/9353410

[41] Desolneux A, Moisan L, Morel JM (2007) From gestalt theory to image analysis: a probabilistic approach, vol 34. Springer Science & Business Media

[42] Singh SK, Fowdur JS, Gawlikowski J, Medina D (2022) Leveraging Graph and Deep Learning Uncertainties to Detect Anomalous Maritime Trajectories. IEEE Trans Intell Transp Syst 23(12):23488–23502. https://doi.org/10.1109/TITS.2022.3190834. https://ieeexplore.ieee.org/document/9839418/

[43] Liatsikou M, Papadopoulos S, Apostolidis L, Kompatsiaris I (2021) A Denoising Hybrid Model for Anomaly Detection in Trajectory Sequences. In: Workshop Proceedings of the EDBT/ICDT 2021 Joint Conference. Nicosia, Cyprus. https://ceur-ws.org/Vol-2841/BMDA_4.pdf

[44] Liao D, Liu W, Zhong Y, Li J, Wang G (2018) Predicting activity and location with multi-task context aware recurrent neural network. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence, pp 3435–3441

[45] Gao Q, Zhou F, Trajcevski G, Zhang K, Zhong T, Zhang F (2019) Predicting Human Mobility via Variational Attention. In: The World Wide Web Conference on - WWW'19, pp 2750–2756. ACM Press, San Francisco, CA, USA. https://doi.org/10.1145/3308558.3313610. http://dl.acm.org/citation. cfm?doid=3308558.3313610

[46] Feng J, Li Y, Zhang C, Sun F, Meng F, Guo A, Jin D (2018) DeepMove: Predicting Human Mobility with Attentional Recurrent Networks. In: Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18. pp 1459–1468. ACM Press, Lyon, France. https://doi.org/10.1145/3178876. 3186058. http://dl.acm.org/citation.cfm?doid=3178876.3186058

[47] . Li A, Wang S, Li W, Liu S, Zhang S (2020) Predicting Human Mobility with Federated Learning. In: Proceedings of the 28th International Conference on Advances in Geographic Information Systems, pp 441–444. ACM, Seattle WA USA. https://doi.org/10.1145/3397536.3422270. https://dl.acm.org/doi/10. 1145/3397536.3422270

[48] Liao D, Liu W, Zhong Y, Li J, Wang G (2018) Predicting activity and location with multi-task context aware recurrent neural network. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence, pp 3435–3441

[49] Kashyap AA, Raviraj S, Devarakonda A, Nayak KSR, K V S, Bhat SJ, (2022) Traffic flow prediction models - A review of deep learning techniques. Cogent Eng 9(1):2010510. https://doi.org/10.1080/23311916. 2021.2010510. https://www.tandfonline.com/doi/full/10.1080/23311916.2021.2010510

[50] Gao Y, Chiang Y, Zhang X, Zhang M (2022) Traffic volume prediction for scenic spots based on multisource and heterogeneous data. Transactions in GIS 26(6):2415–2439. https://doi.org/10.1111/tgis.12975. https://onlinelibrary.wiley.com/doi/10.1111/tgis.12975

[51] Lu Y (2021) Learning to Transfer for Traffic Forecasting via Multi-task Learning. publisher: arXiv Version Number: 1. https://doi.org/10.48550/ARXIV.2111.15542. arXiv:2111.15542

[52] Lippert F, Kranstauber B, Forré PD, van Loon EE (2022) Learning to predict spatiotemporal movement dynamics from weather radar networks. Methods Ecol Evol pp 2041–210X.14007.

https://doi.org/10. 1111/2041-210X.14007. https://onlinelibrary.wiley.com/doi/10.1111/2041-210X.14007

[53] Wang C, Liang Y, Tan G (2022) Periodic residual learning for crowd flow forecasting. In: Proceedings of the 30th International Conference on Advances in Geographic Information Systems, pp. 1–10. ACM, Seattle Washington. https://doi.org/10.1145/3557915.3560947. https://dl.acm.org/doi/10.1145/3557915. 3560947

[54] Zhang Y, Li Y, Zhou X, Kong X, Luo J (2020) Curb-GAN: Conditional Urban Traffic Estimation through Spatio-Temporal Generative Adversarial Networks. In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 842–852. ACM, Virtual Event CA USA. https://doi.org/10.1145/3394486.3403127. https://dl.acm.org/doi/10.1145/3394486.3403127

[55] Li M, Gao S, Lu F, Liu K, Zhang H, Tu W (2021) Prediction of human activity intensity using the interactions in physical and social spaces through graph convolutional networks. Int J Geogr Inf Sci 35(12):2489–2516. https://doi.org/10.1080/13658816.2021.1912347. https://www.tandfonline.com/doi/ full/10.1080/13658816.2021.1912347

[56] Lippert F, Kranstauber B, Forré PD, van Loon EE (2022) Learning to predict spatiotemporal movement dynamics from weather radar networks. Methods Ecol Evol pp 2041–210X.14007. https://doi.org/10. 1111/2041-210X.14007. https://onlinelibrary.wiley.com/doi/10.1111/2041-210X.14007

[57] Buroni G, Bontempi G, Determe K (2021) A tutorial on network-wide multi-horizon traffic forecasting with deep learning. In: Workshop Proceedings of the EDBT/ICDT 2021 Joint Conference. Nicosia, Cyprus. https://ceur-ws.org/Vol-2841/BMDA_6.pdf

[58] Xue H, Voutharoja BP, Salim FD (2022) Leveraging language foundation models for human mobility forecasting. In: Proceedings of the 30th International Conference on Advances in Geographic Information Systems, pp. 1–9. ACM, Seattle Washington. https://doi.org/10.1145/3557915.3561026. https://dl.acm. org/doi/10.1145/3557915.3561026

[59] Nachiket Deo, Mohan M. Trivedi et al Learning and Predicting On-road Pedestrian Behavior Around Vehicles[J]. ITSC,2017

[60] Raúl Quintero Mínguez, Ignacio Parra Alonso et al. Pedestrian Path, Pose, and Intention Prediction Through Gaussian Process Dynamical Models and Pedestrian Activity Recognition.TITS,2018

[61] Liu Q, Wu S, Wang L, et al. Predicting the next location: A recurrent model with spatial and temporal contexts[C]. Proceedings of the AAAI Conference on Artificial Intelligence. Phoenix: AAAI, 2016: 194-200.

[62] Zhang P, Ouyang W L, Zhang P F, et al. SR-LSTM: State refinement for lstm towards pedestrian trajectory prediction[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE CS, 2019: 12077- 12086

[63] SCHLICHTKRULL M, KIPF T N, BLOEM P, et al. Modeling relational data with graph convolutional networks[C]//Proceedings of the European Semantic Web Conference. Berlin: Springer Press, 2018: 593-607.

[64] Ivanovic B, Pavone M. The trajectron: Probabilistic multiagent trajectory modeling with dynamic spatiotemporal graphs[C]. Proceedings of the IEEE International Conference on Computer Vision. Seoul: CV/IEEE, 2019: 2375-2384

[65] Gupta A, Johnson J, Li F F, et al. Social gan: Socially acceptable trajectories with generative adversarial networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE CS, 2018: 2255- 2264.

[66] SCHNEIDER N, GAVRILA D M. Pedestrian path prediction with recursive Bayesian filters: a comparative study[C]//Proceedings of the 35th German Conference on Pattern Recognition. Berlin: Springer Press, 2013: 174-183.

[67] PAVLOVIC V, REHG J M, MACCORMICK J. Learning switching linear models of human motion[C]//Proceeding of the Conference and Workshop on Neural Information Processing Systems. New York: Curran Associates Press, 2000: 981-987.

[68] FOX E B, SUDDERTH E B, JORDAN M, et al. Bayesian nonparametric inference of switching dynamic linear models[J]. IEEE Transactions on Signal Processing, 2011, 59(4): 1569- 1585.

[69] KOOIJ J F P, SCHNEIDER N, FLOHR F, et al. Context-based pedestrian path prediction[C]//Proceedings of the 13th European Conference on Computer Vision. Berlin: Springer Press, 2014: 618-633

[70] HELBING D, MOLNAR P. Social force model for pedestrian dynamics[J]. Physical Review E, 1995, 51(5): 4282-4286

[71] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[J]. Computer Science, 2014.

[72] ALAHI A, RAMANATHAN V, LI F F. Socially-aware largescale crowd forecasting[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2014: 2211-2218.

[73] YI S, LI H S, WANG X G. Understanding pedestrian behaviors from stationary crowd groups[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2015: 3488-3496.

[74] RASMUSSEN C E, WILLIAMS C K I. Gaussian processes for machine learning[M]. Cambridge: MIT Press, 2005.

[75] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780

[76] Alahi A, Goel K, Ramanathan V, et al. Social lstm: Human trajectory prediction in crowded spaces[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE CS, 2016: 961-971.

[77] Vemula A, Muelling K, Oh J. Social attention: Modeling attention in human crowds[C]. Proceedings of the IEEE International Conference on Robotics and Automation. Brisbane: IEEE, 2018: 4601-4607.

[78] Hasan I, Setti F, Tsesmelis T, et al. MX-LSTM: Mixing tracklets and vislets to jointly forecast trajectories and head poses[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE CS, 2018: 6067-6076.

[79] Fernando T, Denman S, Sridharan S, et al. Soft+hardwired attention: An lstm framework for human trajectory prediction and abnormal event detection[J]. Neural Networks, 2018, 108: 466-478.

[80] Xue H, Huynh D Q, Reynolds M. A locationvelocity-temporal attention lstm model for pedestrian trajectory prediction[J]. IEEE Access, 2020, 8: 44576-44589.