This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TITS.2025.3572254, IEEE Transactions on Intelligent Transportation Systems

Social Entropy Informer: A Multi-scale Model-data Dual-driven Approach for Pedestrian Trajectory Prediction

Zihan Jiang, *Student Member, IEEE*, Chengxuan Qin, Rui Yang, *Senior Member, IEEE*, Bingyu Shi, *Student Member, IEEE*, Fuad E. Alsaadi, and Zidong Wang, *Fellow, IEEE*

***Abstract*—Pedestrian trajectory prediction is fundamental in various applications, such as autonomous driving, intelligent surveillance, and traffic management. Existing methods generally fall into two categories: model-driven approaches and data-driven approaches. However, both approaches have inherent limitations when applied to real-world scenarios, particularly in capturing the complex interactions between pedestrians and modeling the stochastic nature of human motion. Notably, there is a lack of research on integrating the strengths of model-driven and data-driven paradigms, which can better address these challenges. This paper aims to fill these limitations by proposing a novel model-data dual-driven approach, called Social Entropy Informer (SEI), for pedestrian trajectory prediction. SEI simultaneously models local and global pedestrian interactions while incorporating information entropy to capture human motion's inherent randomness and uncertainty quantitatively, which provides a robust framework for predicting pedestrian trajectories. Furthermore, we propose a new loss function derived from information theory, which accounts for the stochasticity of pedestrian movement and enhances the model's ability to generalize across diverse scenarios. The SEI framework integrates feature extraction, entropy-based stochastic modeling, and the new loss function, improving prediction accuracy and model interpretability. Experimental results demonstrate that SEI outperforms other benchmark methods in prediction accuracy.**

***Index Terms*—Pedestrian trajectory prediction, information entropy, model-data dual-driven, stochasticity modeling, social interaction modeling, informer model**

I. Introduction

With the advancement of deep learning techniques [2], [6], [8] and neural networks [29], [68], [99], [100], pedestrian trajectory prediction has become a critical area of research, demonstrating significant value in applications such as autonomous driving [25], [76] and ensuring the safety of intelligent transportation systems [22]. The core of this technology lies in accurately predicting pedestrian movements in complex environments, which is crucial for reducing the risk of pedestrian-related traffic accidents. The rapid progress in data science and deep learning has further propelled pedestrian trajectory prediction into a significant research focus [12], [35], [78], [85]. However, this task remains highly challenging due to several key factors, including the complexity of scene topology and dynamic interactions between pedestrians [37], [64], [103]. Additionally, pedestrian trajectory prediction is vital for understanding complex crowd dynamics and social behaviors, providing new insights and tools for various fields such as autonomous driving [14], [46], [52], video surveillance [28], [30], [83], and logistics [23], [44], [47]. As a result, pedestrian trajectory prediction has attracted increasing attention from researchers across these domains.

Currently, pedestrian trajectory prediction relies on two main approaches: model-driven and data-driven methods [36], [60], [92]. Each approach offers unique strengths and is suited to different application scenarios. Model-driven methods are grounded in physical laws and behavioral principles, providing a solid theoretical understanding of pedestrian dynamics. However, model-driven methods often struggle with the complexities and non-deterministic nature of human behavior [55], [63], [74]. On the other hand, data-driven methods excel at capturing intricate, nonlinear patterns in the data but are highly dependent on the quality and quantity of the available data [20], [40], [70], [93]. The dichotomy between model-driven and data-driven methods highlights a critical gap in the current research landscape [87], [96], [101]. While model-driven approaches are often unable to adequately account for the unpredictable and complex nature of human behavior, data-driven methods excel in deciphering intricate and nonlinear patterns in pedestrian movements, albeit with a heavy reliance on the volume and quality of data [32], [33], [102].

This project was funded by the Jiangsu Provincial Qinglan Project, the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (23KJB520038), and the Research Enhancement Fund of XJTLU (REF-23-01-008). This project was also funded by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Saudi Arabia, under grant no. (GPIP: 108-135-2024). The authors, therefore, acknowledge with thanks DSR for technical and financial support.

Z. Jiang is with the Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, Shanghai 200092, China, and also with School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, 215123, China (e-mail: Zihan.Jiang22@alumni.xjtlu.edu.cn);

C. Qin is with School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, 215123, China, and also with School of Electrical Engineering, Electronics and Computer Science, University of Liverpool, Liverpool, L69 3BX, United Kingdom (e-mail: C.Qin8@liverpool.ac.uk);

R. Yang is with School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, 215123, China (e-mail: R.Yang@xjtlu.edu.cn);

B. Shi is with College of Computer and Control Engineering, Northeast Forestry University, Harbin, 150040, China (e-mail: bingyushi0904@outlook.com);

Fuad E. Alsaadi is with the Communication Systems and Networks Research Group, Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia (e-mail: falsaadi@kau.edu.sa; tel: +966 6952183);

Z. Wang is with the Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, United Kingdom (e-mail: Zidong.Wang@brunel.ac.uk).

Obviously, a significant gap in current research is the lack of a practical dual-drive approach to combine the advantages of model-driven and data-driven approaches, especially in modeling the complex interactions of pedestrian behavior and simulating the stochastic nature of pedestrians. Such a hybrid methodology can enhance the accuracy and reliability of trajectory predictions and bridge the gap between theoretical understanding and practical application [9], [34], [69]. However, further exploration and implementation of dual-driven approaches still need to be explored [10], [65], [98].

In the field of pedestrian trajectory prediction, researchers face two significant challenges that impede research progress and hinder practical applications [36], [92]. The first challenge is how to reasonably model the interactivity between pedestrians [90]. A central challenge in research on modeling pedestrian interactivity is accurately understanding and modeling the indirect interactions between pedestrians. Traditional models are usually based on a priori assumption that other pedestrians primarily influence pedestrian behavior within their direct perception range [1], [20], [40], [98]. However, this approach overlooks the complexity and dynamics of pedestrian behavior, especially considering the butterfly effect, where pedestrians outside the perception range can indirectly affect the walking trajectories of others through a series of mediating effects. This phenomenon suggests that pedestrian interactivity is not limited to physical contact or visual perception within perception, but also includes the influence of complex social and environmental factors. As a result, this paper extends the original social interaction hypothesis to propose a new approach for modeling pedestrian interactivity that pedestrians' social interactions extend beyond their immediate perceptions. This approach is crucial for accurately predicting pedestrian behavior, which can simulate both direct and indirect interaction between pedestrians. By incorporating this extended assumption into a dual-drive framework, the model can combine theoretical and data-driven paradigms to capture a complete understanding of pedestrian interactions.

The second challenge lies in effectively modeling the stochasticity of pedestrian behavior. Human movements are inherently uncertain, making it crucial for models to handle stochasticity effectively [59], [63], [74]. This has been a long-standing issue in predicting pedestrian trajectories. Past researchers have invested significant effort into modeling stochasticity [39], [48], [56], [59], but their studies may be limited by low robustness and inadequate simulation rationality [20], [36], [92]. Furthermore, it is challenging to account for the stochasticity of pedestrians in different states with varying pedestrian walking patterns. The simulations in these studies are primarily qualitative and lack quantitative analysis, posing difficulties in accurately modeling pedestrian stochasticity [4], [36], [48], [92]. Therefore, a novel approach is needed to effectively model the stochastic nature of pedestrian behavior. To address this challenge, this paper proposes an information entropy-based modeling approach. As a tool to measure uncertainty and randomness, information entropy can provide a more quantitative framework for modeling pedestrian behavior. Under the framework of the dual-driver approach, information entropy can provide more accurate guidance for the driving part of the model by measuring the stochasticity of pedestrian trajectories, thus enhancing the prediction ability of pedestrian behavior in complex dynamic environments.

Based on the above analysis and motivations, this paper proposes a novel model called the Social Entropy Informer (SEI) within the informer framework [94]. This approach introduces a new method for modeling pedestrian social interactions, incorporating both direct and indirect pedestrian interactions. Furthermore, the paper incorporates information entropy into pedestrian trajectory prediction to capture the randomness of pedestrian behavior. By utilizing information entropy, the model is able to effectively measure the stochasticity and uncertainty of pedestrian trajectories, resulting in an accurate representation of pedestrian behavior. Ultimately, this paper proposes a new model-data dual-driven approach that combines model-driven social interaction modeling and data-driven information entropy analysis to quantify and explain the regularity and stochasticity of pedestrian movement through information entropy, thus effectively integrating both advantages. This approach can address the complexity of pedestrian trajectory prediction on a theoretical basis while utilizing a data-driven approach to enhance the accuracy and robustness of the prediction further. Through this dual-driven framework, the model in this paper can overcome the limitations of a single method and achieve excellent pedestrian trajectory prediction in practical applications. In summary, the main contributions of this paper include:

1) This paper presents a novel approach for constructing a stochastic model of pedestrian behavior by incorporating information entropy, allowing for quantitative analysis of uncertainty in pedestrian trajectory;
2) Building upon a new and expanded hypothesis, which suggests that pedestrian social interactions are influenced by others both within and beyond their perception range, this study proposes a method to model both direct and indirect interactions among pedestrians;
3) A new model-data dual-driven approach, Social Entropy Informer, integrates the strengths of model-driven and data-driven approaches, and superior performance is demonstrated using public datasets.

The rest of the paper is presented as follows. The definition of the pedestrian trajectory prediction task and the a priori assumptions are described in Section II. Section III describes related work and progress in pedestrian trajectory prediction. The specific methodology of the proposed SEI model is outlined in Section IV. Section V details the datasets, experimental environment, and analysis of the experimental results. Finally, Section VI presents the conclusions and discusses relevant future work.

II. Problem formulation

According to previous studies [39], [45], [48], [98], this paper considers pedestrian trajectory prediction as the generation of future sequences based on past sequences. The sequence of pedestrian trajectories can be represented as $P = \{p_1, p_2, \ldots, p_n\}$, where each position $p_i = (x_i, y_i)$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TITS.2025.3572254, IEEE Transactions on Intelligent Transportation Systems

represents a two-dimensional spatial coordinate at time t_i. The goal of pedestrian trajectory prediction is to forecast future paths using historical trajectories. The mathematical expression of this task is shown in (1) below:

$$\hat{P}_{\text{future}} = F\left(P_{\text{past}}, \Theta, I\right) \tag{1}$$

where $P_{\text{past}} = \{p_1, p_2, \ldots, p_m\}$ and $\hat{P}_{\text{future}} = \{\hat{p}_{m+1}, \hat{p}_{m+2}, \ldots, \hat{p}_n\}$ represent the historical and predicted future trajectories respectively, F is a mapping function depending both historical trajectory and other environmental factors (e.g.: a machine learning model such as a deep neural network), Θ represents the model parameters (e.g.: behavioral characteristics, intentions, and other factors related to pedestrians), and I denotes environmental variables (e.g.: trajectories of other pedestrians, and layout of surrounding environment).

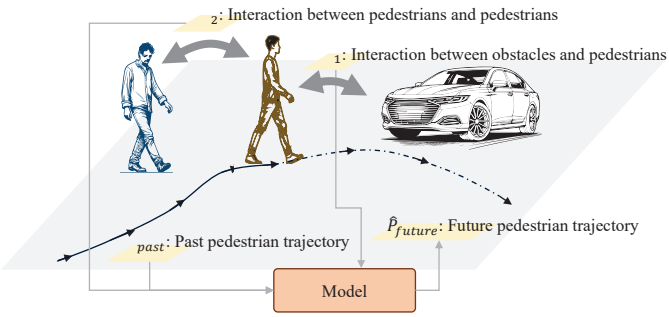


Fig. 1. Illustration of pedestrian trajectory prediction task.

Fig. 1 illustrates the pedestrian trajectory prediction task, highlighting two main components: past trajectories and predicted future trajectories. The model combines these elements to forecast future paths, accounting for interactions between pedestrians and their environment. The depicted framework captures the complexity of human interaction, accounting for individual trajectory histories and the influence of obstacles. Past methods for predicting pedestrian trajectories often assumed that the interactions experienced by pedestrians would only be limited to their perceptual range, as shown in (2) below:

$$I = g\left(I_{in}\left(\Omega_i\right), t\right) \tag{2}$$

where I represents the impact of other external factors on pedestrians, Ω_i represents the perceived range of the pedestrian, $I_{in}\left(\Omega_i\right)$ represents all inputs within the perceived range, and t represents time. This paper extends the hypothesis by considering the actual situation: the interactive influence on pedestrians is not limited to their perceptual range and will also be influenced by factors outside of this range, as shown in (3) below:

$$I = g\left(I_{in}\left(\Omega_i\right), I_{out}\left(\Xi_i\right), t\right) \tag{3}$$

where Ξ_i represents the range outside of pedestrian perception, and $I_{out}\left(\Xi_i\right)$ represents all inputs outside of the perception range. Fig. 2 compares two different approaches for modeling social interactions: (a) the original approach assumes that pedestrians interact only within each other's perceptual range, thereby limiting the potential for interactions; (b) our approach allows for interaction both within and beyond each other's perceptual range. This is reflective of real life, where a change in direction or speed by one pedestrian can cause a chain reaction that affects others even at a distance. Therefore, approach (b) can capture a broad range and complexity of interaction behaviors.

III. Related Work

A. Current research status of model-driven methods

Model-driven methods occupy an essential place in pedestrian trajectory prediction, and these methods are primarily based on theoretical knowledge and empirical judgments from multidisciplinary fields. They apply sociology [38], [41], [49], psychology [3], [15], [54], and physics theories [7], [26], [61], [71] to understand and predict pedestrian behavior. From a sociological and psychological perspective, pedestrian behavior is viewed as a product of social interactions and individual psychological states. Sociological theories [24], [57], [75], help us understand behavior patterns when people walk in groups. Moreover, psychological theories [18], [31], [80], especially those on attention, perception, and decision-making, provide insights into how individuals process information about their environment and make walking decisions. Physics and kinematics models provide an alternative perspective on predicting pedestrian motion [72], [89], [97]. These models are usually based on physical laws, such as Newton's laws of motion, to describe the trajectory of pedestrians, and social force models [26], which treat pedestrians as particles subject to forces, predict the trajectory of pedestrians by modeling gravitational and repulsive forces. These methods provided valuable insights to researchers, but when dealing with complex environments, the limitations of these traditionally based methods began to show.

As a result, researchers have increasingly turned to various machine learning methods to address the challenges of pedestrian trajectory prediction. Techniques such as Kalman filtering [21], Gaussian process dynamical models (GPDM) [51], and XGBoost [107], among others [66], [91], [105], have been widely explored. These methods aim to identify patterns in the data by constructing models that, while not explicitly relying on physical laws as traditional physical modeling does, are still grounded in some form of modeling assumptions. For instance, Kalman filtering is effective for linear systems with Gaussian noise, while GPDM captures temporal dependencies in trajectory data through probabilistic modeling. XGBoost, on the other hand, leverages ensemble learning to handle structured data with high accuracy. However, despite the successes in certain scenarios, these methods often struggle when applied to high-dimensional, complex trajectory data. One major limitation is their reliance on manual feature engineering, which requires domain expertise and may be time-consuming. Additionally, these methods often make strong assumptions about the data, such as linearity, Gaussian noise, or specific functional forms, which may not hold in real-world pedestrian dynamics.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TITS.2025.3572254, IEEE Transactions on Intelligent Transportation Systems

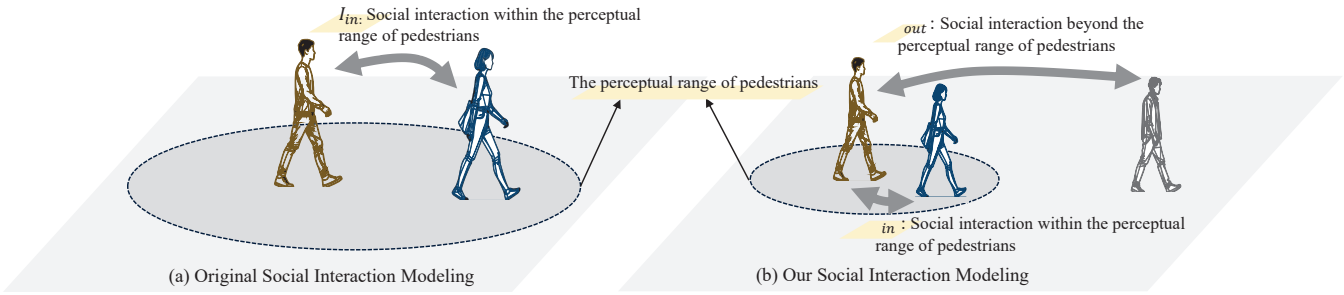


Fig. 2. Variations between different pedestrian interaction modeling.

B. Current research status of data-driven methods

In contrast to model-driven methods, data-driven approaches have gained significant traction in recent years due to their ability to learn complex patterns directly from data without relying heavily on predefined theoretical frameworks. These methods leverage large-scale datasets and advanced computational techniques to flexibly and adaptively model pedestrian behavior. Among the most prominent data-driven techniques are deep learning models, which have demonstrated remarkable success in capturing the intricate dynamics of pedestrian trajectories. Recurrent neural networks (RNN) and their variants [36], such as long short-term memory (LSTM) networks [1], [84], [98] and gated recurrent units (GRU) [81], have been widely adopted for sequence modeling tasks, including trajectory prediction. These models excel at capturing temporal dependencies in pedestrian motion by processing sequential data over time. For instance, LSTM has been used to model the influence of past trajectories on future movements, enabling more accurate predictions in dynamic environments. However, traditional RNN-based approaches often struggle to account for interactions between multiple pedestrians, which are crucial in crowded scenarios.

To address this limitation, researchers have turned to graph-based models that explicitly model pedestrian relationships as a graph, such as graph neural networks (GNN) [59], [65], [86]. GNN can effectively capture social and spatial dependencies, such as collision avoidance and group behavior, by representing pedestrians as nodes and their interactions as edges. For example, social BiGAT [39] and social STGCNN [48] are notable frameworks that integrate graph structures with deep learning to improve trajectory prediction accuracy in crowded settings. Another significant advancement in data-driven methods is the use of attention mechanisms [11], [42], [104], which allow models to focus on the most relevant parts of the input data. Transformers [67], initially developed for natural language processing, have been adapted for trajectory prediction tasks due to their ability to handle long-range dependencies and parallelize computations. These models have shown promise in capturing complex interactions and global context, making them suitable for large-scale and heterogeneous environments.

Despite their strengths, data-driven methods also face several challenges. One major issue is the reliance on large amounts of labeled data, which can be difficult and expensive. Additionally, the interpretability of these models remains a concern, as their "black-box" nature makes it challenging to understand the underlying decision-making process. Furthermore, while data-driven methods excel in capturing complex patterns, they may lack the generalizability of model-driven approaches when applied to unseen scenarios or environments with limited data. Hybrid approaches have emerged as a promising direction to bridge the gap between model-driven and data-driven methods. These methods combine the interpretability and theoretical grounding of model-driven techniques with the flexibility and scalability of data-driven models. For example, some studies have integrated physical constraints, such as social force models, into deep learning frameworks to ensure that predictions adhere to fundamental physical principles [64], [77], [98], [103].

IV. Social Entropy Informer

Based on the structure diagram shown in Fig. 3, this section provides a comprehensive explanation of the proposed SEI model, including three essential components: (1) social interaction module; (2) information entropy-based pedestrian stochasticity modeling; and (3) variety information entropy loss function. In (a). data processing, to incorporate environmental information into the pedestrian trajectory prediction model, this paper designs an environmental information encoding module responsible for extracting compelling features from the input RGB scene images and mapping them into a high-dimensional space for fusion with the pedestrian trajectory data. The module gradually extracts low-level and high-level features in the image through multiple convolutional layers and further integrates these features through fully connected layers.

A. Social interaction module

Based on the proposed extended social interaction hypothesis, this paper presents a new approach for modeling pedestrian social interaction, as shown in Fig. 4. The social interaction method consists of two main components: the global social interaction modeling module modelling pedestrian-perceived in-range and out-range interactions, and the local social interaction modeling module modelling pedestrian-perceived in-range interactions.

The global social interaction modeling module utilizes multi-head attention, add & norm, and feed forward techniques

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TITS.2025.3572254, IEEE Transactions on Intelligent Transportation Systems

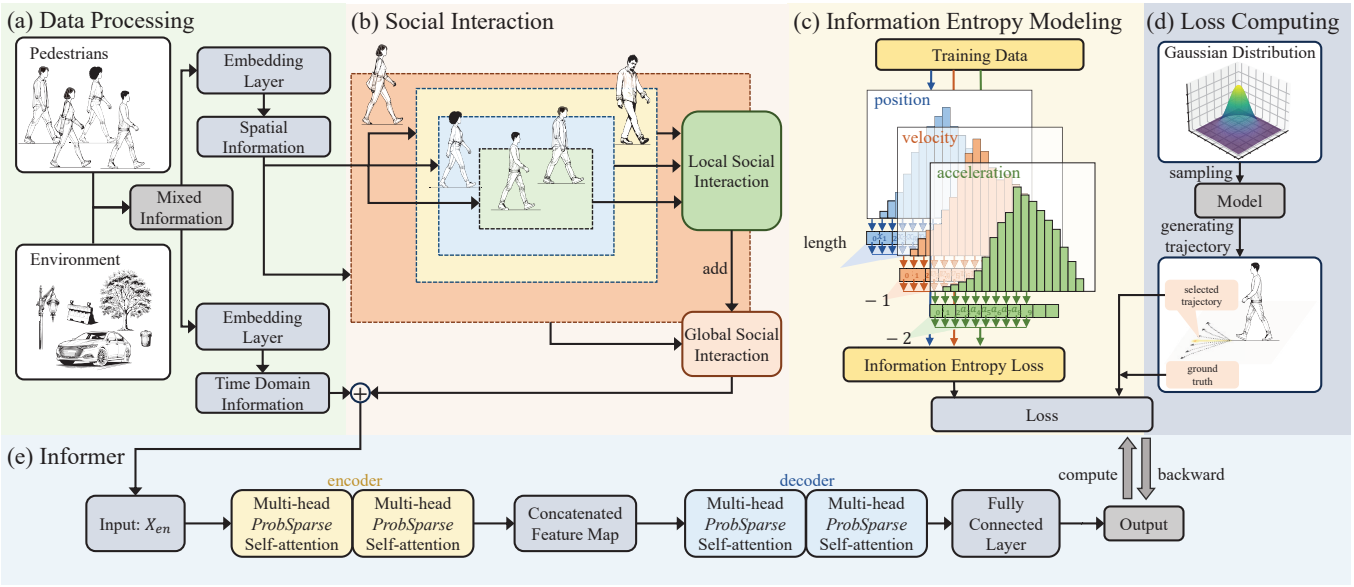


Fig. 3. SEI model structure diagram: An overview of data processing, social interaction, information entropy modeling, and loss computing.

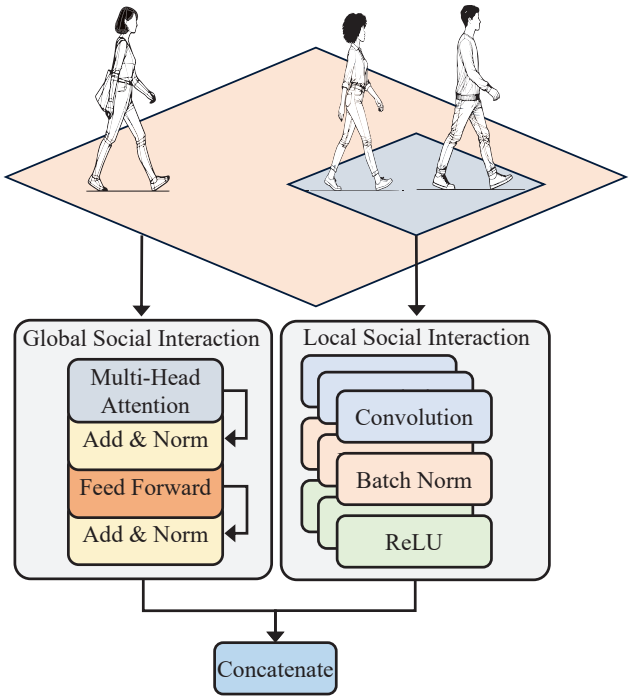


Fig. 4. Multi-scale pedestrian social interaction method schematic diagram.

to comprehensively model pedestrian social interactions on a global scale. The multi-attention mechanism is one of the critical components of this module. By simultaneously focusing on different attention subspaces, it effectively captures pedestrian trajectory features at multiple levels and aspects. Each head focuses on a specific feature subspace, which in the pedestrian trajectory prediction task are expressed as velocity, acceleration, and direction of motion. In addition, it considers the social interactions between pedestrians based on their location and movement. By considering interactions between heads, the model captures correlations between pedestrian trajectories and contextual information about individual pedestrian trajectories. In addition, it performs feature selection and weighting to emphasize features that are relevant to the task while reducing the impact of redundant features on the results. This mechanism is shown in (4) below:

$$\text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Concat}(\text{head}_1, \ldots, \text{head}_h) \cdot \mathbf{W}^O \quad (4)$$

where $\mathbf{Q}$, $\mathbf{K}$, and $\mathbf{V}$ denote the input matrices of the query, key, and value, respectively, head_i denotes the output of the ith attention header, h denotes the number of headers, and $\mathbf{W}^O$ is the weight matrix of the output. The individual attention mechanisms are computed as shown in (5) below:

$$\text{Attention } (\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^{\mathbf{T}}}{\sqrt{\mathbf{d_k}}}\right)\mathbf{V} \quad (5)$$

where the matrices $\mathbf{Q}$, $\mathbf{K}$, and $\mathbf{V}$ represent the Query, Key, and Value, respectively, and the term $\mathbf{d_k}$ denotes the dimension of the key matrices.

The local social interaction modeling module is designed to capture the interactions within a pedestrian's perceptual range by employing multiple CNNs at varying scales. Building upon the methodologies outlined in previous research [13], [17], [43], this study introduces three distinct scales: 3×3, 9×9, and 27×27. The 3×3 scale focuses on capturing fine-grained, close-proximity interactions that are crucial for immediate understanding. The 9×9 scale broadens this perspective to include mid-range interactions, balancing detail and context. Finally, the 27 × 27 scale is responsible for capturing long-distance interactions, expanding the model's perceptual range, which facilitates the model's understanding of social dynamics of the pedestrian environment.

Each CNN employs convolutional operations in the local social interaction modeling module to extract spatial features

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TITS.2025.3572254, IEEE Transactions on Intelligent Transportation Systems

in the input trajectory data. Pooling layers are used to summarize these features, and capture relevant information at each scale (3×3, 9×9, and 27×27). The outputs of each CNN are then connected to form a composite feature of the local social interactions at different scales. With the composite feature, the module can capture different levels of social interactions ranging from individual behavior to broad group dynamics.

B. Information Entropy-based Pedestrian Stochasticity Modeling

Pedestrian trajectories typically exhibit complexity and stochasticity because individual purposes, environmental layout, and social interactions often influence them. These trajectories show many different dynamic patterns, reflecting the uncertainty of pedestrian behavior. To understand and quantify this complexity, a statistical measure of randomness and uncertainty within a system is needed, and this paper addresses the challenge by introducing information entropy. The level of uncertainty and randomness in pedestrian behavior patterns can be effectively assessed by computing the information entropy of pedestrian trajectories, as shown in (6) below:

$$H = -\sum_{i=1}^{n} p\left(Y_i\right) \log_2 p\left(Y_i\right) \tag{6}$$

where $p\left(x_i\right)$ represents the current pedestrian trajectory probability. Therefore, it is evident that this paper aims to model the probability of pedestrian trajectories. This paper takes a statistical approach to probabilistic modeling of pedestrian trajectories, focusing on three main directions: position, velocity, and acceleration. The position of the pedestrian is first modeled with information entropy, as detailed in the specific modeling process below:

1) The distance traveled by the pedestrian trajectory in each frame is computed and divided into intervals $A = \{A_1, A_2, ..., A_j\}$, where each interval size is $[0, a], [a, 2a], [2a, 3a], ..., [ja, (j+1)a]$.
2) When modeling the higher-order case, additional computations must be performed before modeling, as shown in (7) below:

$$d_i^{(k)} = \frac{d_{i+1}^{(k-1)} - d_i^{(k-1)}}{\Delta t} \tag{7}$$

where Δt represents the time interval, $d_i^{(0)}$ represents the sequence representing the pedestrian's location, k represents the k-order derivative, and i represents the ith value in the sequence.
3) All the computed results are summarized, and the probability of each interval is computed through (8) below:

$$p\left(Y_i\right) = \frac{m_i}{\sum_{j=0}^{M} m_{[ja,(j+1)a]}} \tag{8}$$

where m represents the number counted in each interval, M represents the total number of intervals, and i represents the interval of the current trajectory. This process is specifically shown in Algorithm 1.

Algorithm 1: Computation of Distribution Probability of Pedestrian Trajectory

Input: Pedestrian trajectory dataset $D = \{d_1, d_2, ..., d_N\}$, where d_i represents the trajectory of the i-th pedestrian.
Output: List of probabilities for each distance, velocity, and acceleration interval $P = [p(A_1), p(A_2), ..., p(A_M)]$, $P' = [p'(A_1), p'(A_2), ..., p'(A_M)]$ and $P'' = [p''(A_1), p''(A_2), ..., p''(A_M)]$.

```
Divide the distances, velocity, and acceleration between adjacent frames into several intervals A = {A_1, A_2, ..., A_M}, A' = {A'_1, A'_2, ..., A'_M} and A'' = {A''_1, A''_2, ..., A''_M}.
for d_i in D do
    for t = 1 to T (Total frames) do
        Compute the Euclidean distance, velocity, and acceleration between adjacent frames t and t + 1.
    end for
end for
for j = 1 to M do
    Compute the probability p(A_j), p'(A_j) and p''(A_j) of interval A_j, A'_j and A''_j using (8).
end for
return the list of interval probabilities P, P', and P''.
```

C. Variety Information Entropy Loss Function

Traditional pedestrian trajectory prediction methods often employ L2 loss to measure the discrepancy between the actual and predicted trajectories, as demonstrated in (9) below:

$$L_2 = \sum_{i=t_{obs}}^{t_{\text{obs}}+t_{\text{pre}}} \left(Y_i - \hat{Y}_i\right)^2 \tag{9}$$

where Y_i represents the actual trajectory, and $\hat{Y}_i$ represents the predicted trajectory. However, this prediction usually takes into account the average of all possible trajectories and cannot effectively model the stochastic nature of pedestrian movement. To address this issue, the variety loss approach is introduced, as shown in (10) below, incorporating the stochasticity of pedestrian movement by utilizing a Gaussian distribution for multiple sampling:

$$L_{\text{variety}} = \min_k \left\| Y_i - \hat{Y}_i^k \right\|_2 \tag{10}$$

where k is a hyperparameter representing the number of samples in this study. The variety loss models pedestrian stochasticity by generating multiple trajectories. However, this approach also presents a new issue as it assumes that pedestrians adhere to a predetermined probability distribution throughout their walk. Moreover, the variety loss promotes the generation of multiple trajectories without a defined limit.

To address the inherent stochasticity of pedestrian trajectories, this paper introduces an innovative approach by incorporating information entropy into the modeling process. This approach is based on one premise: the information entropy of pedestrian trajectories remains relatively stable.

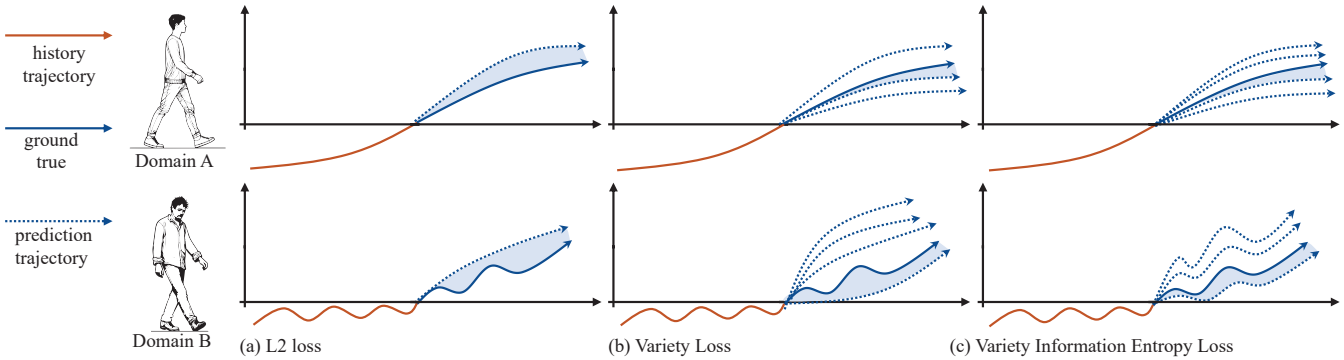


Fig. 5. Differences between different loss functions in the pedestrian trajectory prediction.

This paper aims to accurately capture the uncertainty in pedestrian movement patterns by integrating information entropy. The theoretical rationale for this approach is that pedestrian movement trajectories, although seemingly random, have underlying patterns and constraints, which can be quantified using information entropy. This quantification can significantly provide further insight into pedestrian behavior when various environmental factors influence complex pedestrian patterns.

This paper introduces a new constraint in the trajectory generation process to achieve this purpose, as shown in (11). This constraint effectively regulates the trajectories generated by the variety loss mechanism to ensure that the generated trajectories conform to the natural range of variability in the natural behavior of pedestrians. The information entropy loss prevents the model from generating trajectories that do not conform to the pedestrian pattern, thus keeping the trajectories reasonable for pedestrians.

$$L_{\text{entropy}} = \min_{k} \left\| Y_i - \hat{Y}_i^k \right\|_2 + |H_{pre} - H_{obs}| \tag{11}$$

where H_{pre} represents the information entropy of the predicted trajectory, and H_{obs} represents the information entropy of the actual trajectory. The H_{pre} and H_{obs} computations are shown in (12) and (13) below:

$$H_{pre} = \left| H_{pre}^{p} + H_{pre}^{v} + H_{pre}^{a} \right| \tag{12}$$

$$H_{obs} = \left| H_{obs}^{p} + H_{obs}^{v} + H_{obs}^{a} \right| \tag{13}$$

where p represents position, v represents velocity, and a represents acceleration. L2 loss, variety loss, and variety information entropy loss in the pedestrian trajectory prediction task are schematically shown in Fig. 5.

The difference between different losses in pedestrian trajectory prediction is demonstrated in Fig. 5, where the orange solid line indicates the historical trajectory, the blue solid line shows the actual future trajectory, and the blue dashed line represents the predicted trajectory.

V. Experimental Results and Analysis

A. Dataset Introduction

The ETH [16], UCY [5], and SDD [58] datasets have emerged as pivotal benchmarks in pedestrian trajectory prediction. However, each dataset's distinct characteristics and scenarios highlight their unique applications and strengths. The ETH dataset, collected by ETH Zurich, contains video sequences captured from two different environments: a schoolyard and a street. The videos are recorded at a resolution of 640x480 pixels and a frame rate of 14 frames per second (fps), providing moderate detail and smoothness in capturing pedestrian movements. The dataset contains 1550 pedestrian trajectories that reflect pedestrian behavior and interactions in these environments.

The UCY dataset from the University of Cyprus provides a wide variety of recording scenarios. Specifically, these scenarios are university campuses, streets, and public squares. These different scenarios have a large number of pedestrian trajectories and social interactions. In addition, the video resolution of the dataset varies from 720x576 to 1080x1920 pixels, with frame rates ranging from 25 to 50 fps. The UCY dataset provides 7724 pedestrian trajectories that reflect different pedestrian dynamics in various environments.

The SDD dataset was collected by researchers at Stanford University and is based on real outdoor environments such as campuses and streets, which highly reproduces the complex behaviors of pedestrians and traffic conditions in the real world. The dataset contains many high-definition video frames captured by UAV-mounted cameras, which are of high resolution and can present the appearance, posture, and trajectory of pedestrians and other detailed information. At the same time, the pedestrians in the video are thoroughly and accurately labeled with various attributes such as location, speed, direction of movement, and behavioral categories.

B. Experiment Environment

The experimental environment used in this paper consists of an i7-11800H processor, NVIDIA GeForce RTX 3080 graphics card, Windows 11 operating system, and Python 3.8. The hyperparameter settings employed in the experiments are detailed as follows: the learning rate was set to 0.0015, the number of epochs was fixed at 300, and the batch size was set to 4 to optimize the training process. For the temporal aspects of the proposed model, an observation time of 3.2 seconds (equivalent to 8 frames), and a prediction time of 4.8 seconds (equivalent to 12 frames) were used, with the sampling number set to 20.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TITS.2025.3572254, IEEE Transactions on Intelligent Transportation Systems

TABLE I
COMPARISON OF EXPERIMENTAL RESULTS

Method (ADE/FDE)	ETH	HOTEL	UNIV	ZARA1	ZARA2	Average	SDD
Vanilla LSTM [1]	1.09/2.41	0.86/1.91	0.61/1.31	0.41/0.88	0.52/1.11	0.70/1.52	×
Social LSTM [1]	1.09/2.35	0.79/1.76	0.67/1.40	0.47/1.00	0.56/1.17	0.71/1.53	31.19/56.97
SGAN [20]	0.87/1.62	0.67/1.37	0.60/1.26	0.34/0.69	0.42/0.84	0.58/1.12	27.23/41.44
SoPhie [59]	0.70/1.43	0.76/1.67	0.54/1.24	0.30/0.63	0.38/0.78	0.54/1.15	16.27/29.38
GAT [39]	0.68/1.29	0.68/1.40	0.57/1.29	0.29/0.60	0.37/0.75	0.52/1.07	×
Social BiGAT [39]	0.69/1.29	0.49/1.01	0.55/1.32	0.30/0.62	0.36/0.75	0.48/1.00	×
BR-GAN [56]	0.73/1.37	0.55/1.13	0.53/1.07	0.35/0.71	0.35/0.72	0.50/1.00	×
Social STGCNN [48]	0.64/1.11	0.49/0.85	0.44/0.79	0.34/0.53	0.30/0.48	0.44/0.75	×
EvoSTGAT [65]	0.64/1.19	0.35/0.51	0.44/0.82	0.31/0.50	0.28/0.47	0.41/0.70	×
AST-GNN [95]	0.66/1.02	0.37/0.61	0.46/0.83	0.32/0.52	0.28/0.45	0.42/0.69	×
LSTM-NLL+DE [53]	0.60/1.11	0.40/0.67	0.53/0.97	0.56/1.12	0.25/0.50	0.47/0.87	
VIKT-D [104]	0.62/1.26	0.26/0.52	0.53/1.22	0.42/0.95	0.31/0.69	0.43/0.93	12.59/23.51
Social SAGAN [86]	0.65/1.19	0.36/0.70	0.54/1.14	0.33/0.66	0.29/0.61	0.43/0.86	×
SEEM [77]	0.48/0.86	0.52/1.18	0.35/0.65	0.28/0.47	0.24/0.50	0.38/0.71	×
TPPO [88]	0.75/1.27	0.36/0.70	0.39/0.74	**0.22/0.37**	0.23/0.45	0.39/0.71	×
Atten-GAN [19]	0.64/1.12	0.36/0.72	0.51/1.13	0.35/0.61	0.24/0.65	0.42/0.84	×
IA-LSTM [84]	0.43/0.77	0.50/0.80	0.48/0.73	0.44/0.46	0.36/0.55	0.44/0.66	×
E-SR-LSTM [98]	0.44/0.79	**0.19/0.31**	0.32/0.64	0.27/0.54	0.50/1.05	0.34/0.67	×
Social NSTransformers [42]	0.40/0.71	0.29/0.47	0.39/0.73	0.34/0.62	0.31/0.57	0.35/0.62	10.92/18.01
RCPNet [106]	0.48/0.86	0.38/0.68	0.31/**0.58**	0.25/0.44	0.23/**0.35**	0.33/0.58	**8.18**/13.83
Social Entropy Informer	**0.34/0.64**	**0.19**/0.33	**0.29**/0.61	0.24/0.52	**0.22**/0.46	**0.26/0.51**	8.72/**13.35**

C. Evaluation Indicators

In this study, the evaluation metrics used are Average Displacement Error (ADE) and Final Displacement Error (FDE), hich are commonly used in similar studies [1], [20], [56], [59]. The formulations for these indicators are shown in (14) and (15), respectively:

$$ADE = \frac{\sum_{n=1}^{N}\sum_{t=t_{\text{obs}}+1}^{t_{\text{pre}}}\left\|Y_n^t - \hat{Y}_n^t\right\|_2}{N * (t_{\text{pre}} - t_{\text{obs}} - 1)} \tag{14}$$

$$FDE = \frac{\sum_{n=1}^{N}\left\|Y_n^{t_{\text{pre}}} - \hat{Y}_n^{t_{\text{pre}}}\right\|_2}{N} \tag{15}$$

where Y_n^t and $\hat{Y}_n^t$ represent the actual and predicted trajectories of pedestrian n at time t, and N represents the current total number of pedestrians.

D. Comparative Experimental Results and Analysis

The experiments in this paper are performed on ETH, UCY, and SDD datasets. ETH and UCY have five sub-datasets: ETH, HOTEL, UNIV, ZARA01 and ZARA02. Table I demonstrates the experimental results for the above public dataset. In this table, bold denotes the best results, and underline denotes the sub-optimal results. The numbers before and after the slash for these results represent the ADE and FDE values, respectively. As seen from Table I, SEI performs very well on all datasets. Specifically, SEI exhibits low values on all evaluation metrics (ADE and FDE), indicating its high accuracy and robustness in the pedestrian trajectory prediction task. On the ETH dataset, SEI has an ADE of 0.34 and an FDE of 0.64, showing excellent prediction accuracy and trajectory stability. On the HOTEL dataset, SEI has an ADE of 0.19 and an FDE of 0.33, maintaining a low error and demonstrating high prediction accuracy in complex environments. The ZARA2 dataset results in an ADE of 0.22 and an FDE of 0.46, proving SEI's effectiveness in handling social interactions and dynamic environments. SEI performs consistently across multiple datasets, effectively reducing prediction errors and improving trajectory accuracy.

However, it is evident that the SEI model still does not achieve optimal results on some datasets. In the Hotel scene, the movement of pedestrians may be highly influenced by a few key pedestrians around them. E-SR-LSTM performs deeper mining of relationships between localized pedestrians by state refinement for nodes and spatial edges to the best performance on the FDE. In UNIV, ZARA1, and ZARA2, RCPNet estimates human's goal intention through Destination Variational Auto-Encoder (DVAE), which learns the future path distribution and outputs multiple predictions so that the predicted trajectories are more in line with the actual human intention. Conversely, SEI captures pedestrian behavioral characteristics mainly through the social interaction module and information entropy. It is less targeted than RCPNet in intent modeling, making RCPNet more advantageous on some datasets. In the ZARA1 dataset, pedestrians have diverse behavioral patterns. TPPO is a model based on GAN, which generates multimodal predictions by introducing potential variables that can be changed to deal with future uncertainty. This multimodal output capability of TPPO allows it to cover various possible trajectories better and more accurately predict pedestrians' future locations. SEI is relatively weak in dealing with multimodality and cannot fully consider the diversity of pedestrian trajectories as TPPO.

The performance of different behavioral prediction models on multiple datasets is further illustrated using visualization techniques, as shown in Fig. 6. The visualization results reveal that our method exhibits higher adaptability and accuracy in processing complex intersections than BRGAN and social STGCNN models. Compared with E-SR-LSTM, SEI is less affected by the disturbance of the surrounding environment and can effectively predict the walking paths of pedestrians in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TITS.2025.3572254, IEEE Transactions on Intelligent Transportation Systems

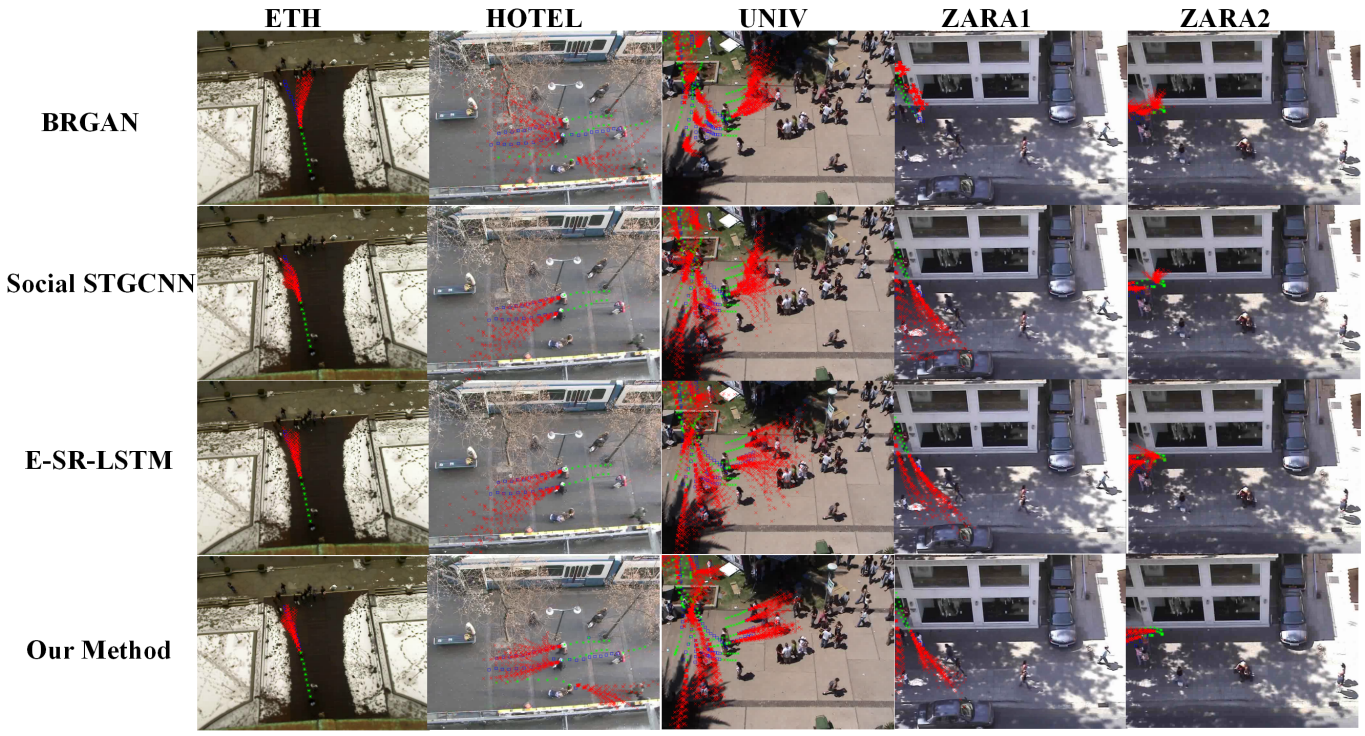


Fig. 6. The comparison of experimental visualization results. Green represents the historical trajectory, blue represents the future labeled trajectory, and red represents the predicted trajectory.

TABLE II
RESULTS OF ABLATION EXPERIMENTS

Method (ADE/FDE)	S	V	I	ETH	HOTEL	UNIV	ZARA01	ZARA02	Average
Informer	×	×	×	0.64/1.22	0.24/0.56	0.38/0.77	0.31/0.68	0.26/0.56	0.37/0.76
Informer	✓	×	×	0.41/0.79	0.23/0.55	0.35/0.75	0.27/0.63	0.23/0.53	0.30/0.65
Informer	×	✓	×	0.44/0.77	0.23/0.49	0.38/0.75	0.26/0.56	0.22/0.51	0.31/0.62
Informer	×	✓	✓	0.39/0.70	0.21/0.43	0.36/0.69	0.27/0.57	0.24/0.51	0.30/0.58
Informer	✓	✓	×	0.35/0.68	0.20/0.37	0.31/0.67	0.25/0.53	**0.22**/0.48	0.27/0.55
Informer	✓	✓	✓	**0.34/0.64**	**0.19/0.33**	**0.29/0.61**	**0.24/0.52**	**0.22/0.46**	**0.26/0.51**

dense crowds. In the SDD dataset experiments, × represents that no experimental comparison of the method was conducted. Overall, Social Entropy Informer achieves excellent results with an ADE of 8.72 and an FDE of 13.35, indicating that the method has significantly less error than other methods in predicting pedestrians' future and final locations. This advantage is not only reflected in the comparison with the relatively weaker methods but also the comparison with the better RCPNet (ADE: 8.18, FDE: 13.83), which is very close to the optimal values in ADE and FDE, which fully demonstrates the stability and efficiency of the method on SDD dataset. The visualization of SEI on the SDD dataset is shown in Figure 7.

E. Ablation Experimental Results and Analysis

In Table II, S stands for social interaction module, V stands for variety loss, and I stands for information entropy. The experimental results in the table show that the performance of the informer method in several scenarios is significantly improved with the introduction of the three key factors, namely, social interaction module (S), variety loss (V), and information entropy (I), in turn. The overall average ADE/FDE value is 0.37/0.76 in the initial stage when no factor is introduced. After introducing the S factor, the average value drops to 0.30/0.65, indicating that the social interaction module plays an important role in improving the model performance. This significant change indicates that the social interaction module is important in improving the model performance. Next, when the V factor alone is introduced, the average index also decreases to 0.31/0.62, which is strong evidence of the positive effect of variety loss on the model performance.

When the V and I factors are introduced, the average index decreases even further to 0.30/0.58. In the specific scenarios where S and V are introduced but not I, the model already

TABLE III
MATCHED SAMPLES T-TEST RESULTS

Index	Mean diff.	Std. deviation diff.	P-value
ADE	-0.027	0.018288	0.001171
FDE	-0.041	0.020248	0.000125

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TITS.2025.3572254, IEEE Transactions on Intelligent Transportation Systems

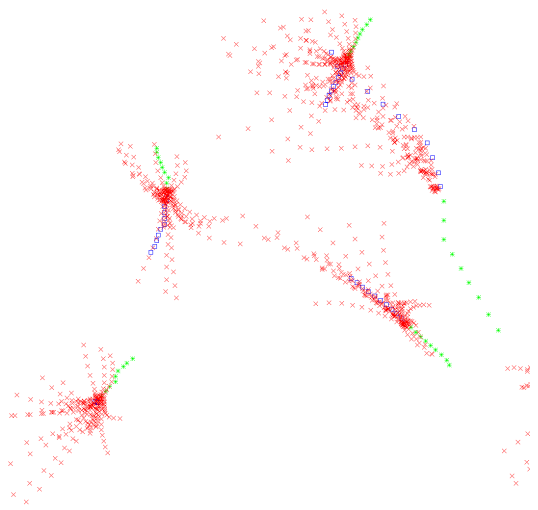


Fig. 7. SEI visualization results on the SDD dataset. Green represents the historical trajectory, blue represents the future labeled trajectory, and red represents the predicted trajectory

achieves good results, with some indicators reaching the underlined superior levels. It is worth mentioning that when all three factors are introduced, the ADE and FDE metrics in all scenarios have reached the optimal state, which are presented in bold in the table. It can be seen that the synergistic effect of the three factors, namely social interaction module, variety loss, and information entropy, is crucial for improving the performance of the informer model, and their joint efforts can effectively reduce the prediction error of the model and significantly improve the accuracy and generalization ability of the model.

Hence, it is imperative to test this statistically in the present study. Besides, integrating information entropy into variety loss transforms the underlying probability distribution, initially represented by a standard Gaussian distribution. This integration process involves adjusting the distribution based on the difference between the entropies of predicted and observed trajectories. The difference in entropies is computed by subtracting the entropy of the predicted probabilities from the entropy of the ground true probabilities. The logarithmic function's inherent nonlinearity introduces a deviation from the standard Gaussian distribution in the original probability distribution. Therefore, it is necessary to statistically test this in the present study, and a matched samples T-test is utilized to verify the validity of information entropy. The hypothesis is that model A's pedestrian trajectory prediction accuracy (with information entropy) is higher than model B's (without information entropy). The two models were tested ten times each, with the results of the matched samples t-test shown in Table III.

The results in Table III demonstrate that the mean difference between model A and model B is -0.027 for ADE and -0.041 for FDE. These negative differences suggest that incorporating information entropy is effective in reducing prediction errors. The p-values are notably low (0.001171 for ADE and 0.000125 for FDE), providing statistically robust evidence that these differences are not coincidental but significantly enhance the model's predictive power by information entropy. While there is some fluctuation in performance, the overall standard deviations observed in the tests indicate that model A significantly outperforms model B in both ADE and FDE metrics.

F. Detailed analysis of the diversity and naturalness of generated trajectories

The ADE and FDE often represent only the accuracy of the trajectories, but this paper involves trajectory sampling, so analyzing the accuracy alone is insufficient. In order to comprehensively analyze the diversity and naturalness of the generated trajectories, this paper analyzes the generated trajectories in terms of both the Kolmogorov-Smirnov (KS) test and the cosine similarity of the trajectories, and the specific results are shown in the Table IV.

In Table IV, several trajectories were randomly selected for this paper's KS test and cosine similarity calculation. From the results, it can be seen that almost all the trajectories generated in this paper can pass the KS test, which can be regarded as a small probability of observing a current difference or a more extreme difference between these two samples when the original hypothesis (the two samples come from the same distribution) holds. This means there is strong evidence that these two samples are not from the same distribution, i.e., the generated trajectories are diverse. Besides, it is then clear from the similarity calculations that the generated trajectories are highly similar to the predicted trajectories, and the generated trajectories are natural. The results of the related qualitative analysis are shown in Figure 8.

In Figure 8, A represents the SEI realized based on variety loss, and B represents the SEI realized based on variety information entropy loss. Green represents the historical trajectory, blue represents the future labeled trajectory, and red represents the predicted trajectory. In B, the predicted trajectories are more closely clustered around the accurate future trajectories regarding spatial distribution than A. The predicted trajectories are more closely clustered around the accurate future trajectories regarding spatial distribution. This indicates that the distribution of the prediction results of the SEI realized based on the entropy loss of diversity information in predicting the future trajectories of pedestrians or objects is closer to the actual future trajectories, which suggests that the prediction model of B performs better in capturing the characteristics and trends of the future trajectories. The accuracy and reliability of the predictions are relatively higher.

G. Experimental Results and Analysis of Local and Global Perceptual Ranges in the Social Interaction Module

The experimental results on different pedestrian perception ranges are shown in the Table V, involving the experimental data of SEI at different perception ranges and the corresponding average data. First, when only the local perceptual range is turned on, the model performs more prominently in the ETH scenario and achieves better results in some indicators. However, in terms of the overall average data, the average

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TITS.2025.3572254, IEEE Transactions on Intelligent Transportation Systems

TABLE IV
QUANTITATIVE RESULTS FOR DIVERSITY AND NATURALNESS OF THE GENERATED TRAJECTORIES

Evaluation indicators	1	2	3	4	5	6	7	Average
KS test (P vaule)	$2.72 * 10^{-11}$	$2.26 * 10^{-28}$	$9.35 * 10^{-22}$	$1.21 * 10^{-69}$	$2.98 * 10^{-49}$	$6.81 * 10^{-6}$	$1.04 * 10^{-21}$	$9.73 * 10^{-7}$
Cosine similarity	0.9956	0.9946	0.9964	0.9956	0.9969	0.9907	0.9923	0.9952

TABLE V
EXPERIMENTAL RESULTS OF DIFFERENT PEDESTRIAN PERCEPTION RANGE

Method	Local	Global	ETH	HOTEL	UNIV	ZARA01	ZARA02	Average
Social Entropy Informer	✓	×	0.41/0.69	0.32/0.85	0.36/0.69	0.27/0.58	0.25/0.51	0.32/0.66
Social Entropy Informer	×	✓	0.39/0.71	0.20/0.42	0.32/0.66	0.26/0.53	**0.22**/0.44	0.28/0.55
Social Entropy Informer	✓	✓	**0.34/0.64**	**0.19/0.33**	**0.29/0.61**	**0.24/0.52**	**0.22/0.46**	**0.26/0.51**

achievement is 0.32/0.66, which is in the middle level. This indicates that by relying only on the local sensing range, the model can capture a specific range of pedestrian interaction information and better adapt to some particular scenes. Secondly, when only the global perception range is turned on, the model significantly improves some of the metrics of multiple scenes. In terms of the average data, the average score improves to 0.28/0.55, indicating that the global perceptual range can help the model to acquire a broader range of information and thus achieve better results in multiple scenarios, which is a significant performance improvement compared to relying on the local perceptual range only.

Finally, when both local and global perceptual ranges are turned on, the model achieves optimal results in all scenarios. This demonstrates that combining local and global perceptual ranges at the same time allows the model to synthesize the information from different scales to capture both local pedestrian interaction details and global scene information. In summary, the experimental results clearly show that considering both local and global perceptual ranges in SEI can bring the optimal performance for the model, which can deal with the pedestrian interaction problem more efficiently and improve the accuracy and stability of the model in different scenarios compared with using local or global perceptual ranges alone.

H. Impact of Different Size of Convolution on Model Performance and Analysis

In order to investigate the effect of convolutional kernel size on model performance, this paper designs experiments with different combinations of multi-scale convolutional kernels of different sizes. The subsection selects several combinations of convolutional kernel sizes and applies them to different datasets to evaluate their impact on the model. The experimental results are shown in Table VI.

From the experimental results in Table VI, it can be seen that different combinations of convolutional kernels significantly

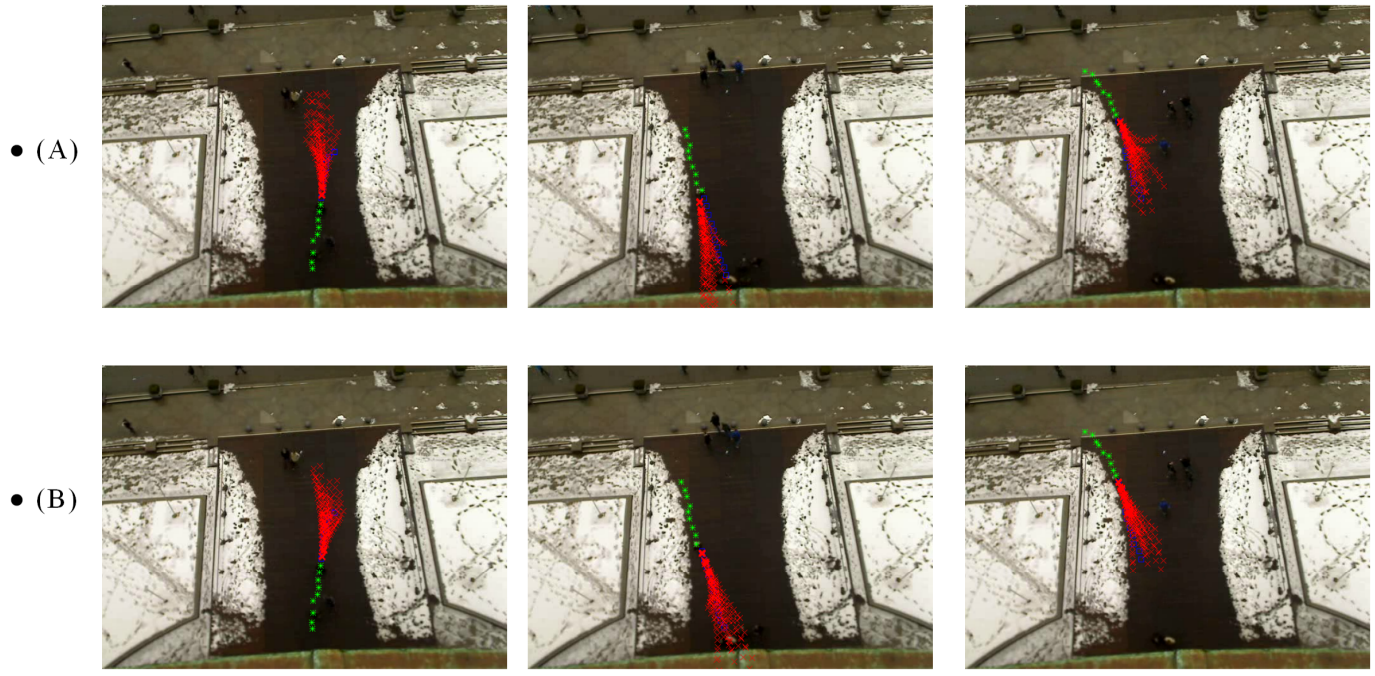


Fig. 8. Quantitative comparative analysis of the generated sampling trajectories results.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TITS.2025.3572254, IEEE Transactions on Intelligent Transportation Systems

TABLE VI
EXPERIMENTAL RESULTS FOR DIFFERENT CONVOLUTIONAL KERNEL SIZES

Method (ADE/FDE)	ETH	HOTEL	UNIV	ZARA01	ZARA02	Average
Multi-scale convolutional scales: 3 9 27	**0.34/0.64**	**0.19/0.33**	**0.29/0.61**	**0.24/0.52**	**0.22/0.46**	**0.26/0.51**
Multi-scale convolutional scales: 5 11 21	0.40/0.67	0.22/0.37	0.32/0.65	0.27/0.56	0.25/0.50	0.29/0.55
Multi-scale convolutional scales: 9 15 25	0.41/0.71	0.23/0.39	0.34/0.68	0.29/0.59	0.26/0.52	0.31/0.58
Multi-scale convolutional scales: 7 13 23	0.38/0.71	0.21/0.35	0.30/0.62	0.26/0.55	0.24/0.49	0.28/0.54

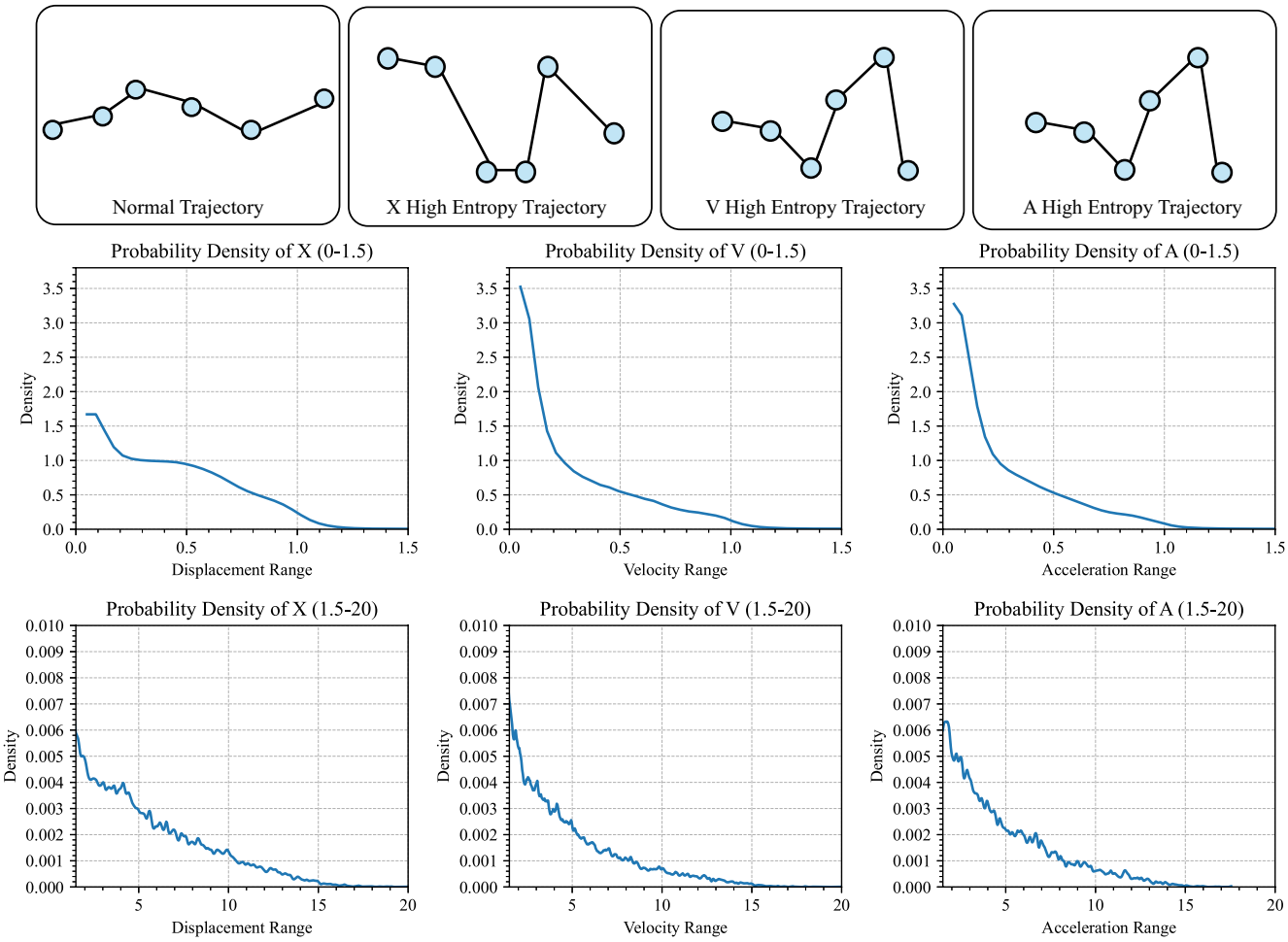


Fig. 9. Information entropy statistics results and trajectory schematics

affect the accuracy of the model. In this paper, we analyze the effect of different combinations of convolutional kernel sizes on the ADE and FDE of different datasets. The results show that the multi-scale convolutional kernel (3 9 27) performs the best on all the datasets, which indicates that this combination has a strong feature extraction capability and can effectively capture the spatio-temporal information of the trajectories. The performance of other combinations is also closer, with certain combinations performing better on specific datasets.

I. Detailed Analysis and Discussion of Information Entropy in SEI

To summarize the relationship between information entropy and pedestrian trajectory, the information entropy statistical process is visualized in this paper, as shown in Figure 9. In Figure 9, the statistical results of this paper are shown, where the probability densities of displacement, velocity, and acceleration all show a rapid decrease with increasing variables for a smaller range of values (0 - 1.5). This implies that smaller displacement, velocity, and acceleration values are more common in this range, with a relatively low probability of larger values occurring. In the more extensive range of values (1.5 - 20), the probability densities of all three are maintained at a low level overall. The changes are relatively smooth, indicating that larger displacement, velocity, and acceleration values occur less frequently, and their distribution is somewhat uniform.

Furthermore, several different pedestrian trajectories are shown in the figure, with the typical trajectories exhibiting smooth and stable characteristics with a low degree of fluctu-

ation. In contrast, the high entropy trajectories all show significant fluctuations and irregularities. Combining the trajectory schematic and probability density analysis results, there is a clear difference between the normal motion state and the high entropy motion state. The normal state has more stable motion characteristics, while the high-entropy state shows more substantial randomness and unpredictability. In addition, the results of the probability density distribution provide a quantitative basis for understanding the motion behavior of the studied pedestrians and help to identify common and rare states during motion.

J. Experimental results and analysis of model runtime performance

TABLE VII
THE RESULTS OF MODELS RUNTIME PERFORMANCE

Method	Model parameters	Inference time
Social LSTM [1]	0.26 Mb	0.04 s
STGAT [27]	0.0047 Mb	0.01 s
STGCNN [48]	0.01 Mb	2.05 s
MemoNet [79]	5.18 Mb	3.51 s
STAR [82]	0.80 Mb	0.45 s
GE-Trans [50]	44.15 Mb	0.05 s
Social Entropy Informer	0.80 Mb	1.51 s

As can be seen from the data in the Table VII, there are significant differences between the different models regarding parameter size and inference time. Regarding model parameters, the GE-Trans model has 44.15 Mb parameters, which is the most parameterized among all models. In comparison, the STGAT model has only 0.0047 Mb parameters, which is the model with the fewest parameters, and it has a particular advantage in scenarios with limited resources. In terms of inference time, the STGAT model performs the best with only 0.01 s, which can give the inference results quickly; the inference time of STGCNN and MemoNet is longer with 2.05 s and 3.51 s, respectively, which may affect their use in real-time demanding application scenarios. The model parameter of SEI is 0.80 Mb, which indicates that it can be used in real-time scenarios. Medium level indicates that it is in a relatively balanced position regarding model complexity and resource requirements. In the future, the model structure or algorithm can be further optimized for this problem to improve its inference efficiency.

VI. CONCLUSIONS

This paper proposes the SEI model, an excellent pedestrian trajectory prediction framework that skillfully integrates the information entropy theory into the informer framework. The SEI model enhances the accuracy and generalization of pedestrian trajectory prediction by quantitatively analyzing trajectory uncertainty and utilizing an innovative information entropy-based loss function for optimizing model training. Extensive experiments on multiple standard datasets verify the SEI model's excellent performance in predicting pedestrian behavior in complex and dynamic environments.

Although SEI has shown some advantages in pedestrian trajectory prediction, there are still significant limitations. In complex indoor scenarios such as large shopping centers and airport terminals, the complex spatial structure and various pedestrian purposes make it difficult for SEI to capture the pedestrian decision-making process accurately. In extreme traffic conditions, such as high-density crowds and emergencies at intersections or subway stations during rush hours, as well as multi-modal interactions in comprehensive transportation hubs, SEI's ability to predict changes in pedestrian behavior is insufficient. For pedestrians with unique behaviors, including children, the elderly, and those carrying heavy or unique objects, SEI lacks an effective response mechanism due to their unique behavioral patterns, resulting in biased predictions, and subsequent research needs to focus on solving these problems in order to improve the effectiveness of the model.

REFERENCES

[1] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, F. Li and S. Savarese, "Social LSTM: Human Trajectory Prediction in Crowded Spaces," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 961-971.

[2] F. E. Alsaadi, Z. Wang, Y. Luo, N. S. Alharbi and F. W. Alsaade, "H∞ State Estimation for BAM Neural Networks With Binary Mode Switching and Distributed Leakage Delays Under Periodic Scheduling Protocol," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 9, pp. 4160-4172, 2022.

[3] A. Bera, T. Randhavane and R. Prinja and D. Manocha, "SocioSense: Robot navigation amongst pedestrians with social and psychological constraints," *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Vancouver, BC, Canada, 2017, pp. 7018-7025.

[4] D. E. Benrachou, S. Glaser, M. Elhenawy and A. Rakotonirainy, "Use of Social Interaction and Intention to Improve Motion Prediction Within Automated Vehicle Framework: A Review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 22807-22837, 2022.

[5] A. Lerner, Y. Chrysanthou, and D. Lischinski, "Crowds by Example." *Computer Graphics Forum*, vol. 26, no. 3, pp. 655-664, 2007.

[6] Z. Chen, R. Yang, M. Huang, F. Li, G. Lu, and Z. Wang, "EEGProgress: A Fast and Lightweight Progressive Convolution Architecture for EEG Classification", *Computers in Biology and Medicine*, vol. 169, pp. 107901, 2023.

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

[8] Z. Chen, R. Yang, M. Huang, Z. Wang and X. Liu, "Electrode Domain Adaptation Network: Minimizing the Difference Across Electrodes in Single-Source to Single-Target Motor Imagery Classification", *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 8, no. 2, pp. 1994-2008, 2023.

[9] W. Chen, H. Sang, J. Wang and Z. Zhao, "DSTIGCN: Deformable Spatial-Temporal Interaction Graph Convolution Network for Pedestrian Trajectory Prediction," *IEEE Transactions on Intelligent Transportation Systems*, early access.

[10] W. Chen, H. Sang, J. Wang and Z. Zhao, "DSTCNN: Deformable spatial-temporal convolutional neural network for pedestrian trajectory prediction," *Information Sciences*, vol. 666, pp. 120455, 2024.

[11] X. Chen, H. Zhang, F. Deng, J. Liang and J. Yang, "Stochastic Non-Autoregressive Transformer-Based Multi-Modal Pedestrian Trajectory Prediction for Intelligent Vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 5, pp. 3561-3574, 2024.

[12] Y. Chen, R. Yang, M. Huang, Z. Wang and X. Liu, "Single-Source to Single-Target Cross-Subject Motor Imagery Classification Based on Multisubdomain Adaptation Network," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 1992-2002, 2022.

[13] C. Zhou, M. Han, Q. Liang, Y. Hu and S. Kuai, "A Social Interaction Field Model Accurately Identifies Static and Dynamic Social Groupings", *Nature human behaviour*, vol.3, no.8, pp. 847–855, 2019.

[14] W. Chen, P. Li, and H. Zhao, "Automatic Model-Based Dataset Generation for High-Level Vision Tasks of Autonomous Driving in Haze Weather," *Neurocomputing*, vol. 494, pp. 23-32, 2022.

[15] Ruth Conroy Dalton, "The secret is to follow your nose: Route path selection and angularity," *Environment and Behavior*, vol. 35, no. 1, pp. 107-131, 2003.

[16] S. Pellegrini, A. Ess, and L. Van Gool, "Improving data association by joint modeling of pedestrian trajectories and groupings," *Computer Vision–ECCV 2010: 11th European Conference on Computer Vision (ECCV)*, Heraklion, Crete, Greece, 2010, pp.452-465.

[17] L. Feng and B. Bhanu, "Understanding Dynamic Social Grouping Behaviors of Pedestrians," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 2, pp. 317-329, 2015.

[18] Meiqing Fu, Rui Liu and Yu Zhang, "Do people follow neighbors? An immersive virtual reality experimental study of social influence on individual risky decisions during evacuations," *Automation in Construction*, vol. 126, pp. 103644, 2021.

[19] F. Fang, X. Wang, Z. Li, K. Qian and B. Zhou, "A Unified Framework for Pedestrian Trajectory Prediction and Social-Friendly Navigation," *IEEE Transactions on Industrial Electronics*, vol. 71, no. 9, pp. 11072-11082, 2024.

[20] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese and A. Alahi, "Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR)*, Salt Lake City, UT, USA, 2018, pp. 2255-2264.

[21] M. Goldhammer, S. Köhler, S. Zernetsch, K. Doll, B. Sick and K. Dietmayer, "Intentions of Vulnerable Road Users—Detection and Forecasting by Means of Machine Learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 7, pp. 3035-3045, 2020.

[22] J. Gan, S. Li, C. Wei, L. Deng and X. Tang, "Intelligent Learning Algorithm and Intelligent Transportation-Based Energy Management Strategies for Hybrid Electric Vehicles: A Review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 10, pp. 10345-10361, 2023.

[23] Y. Gao, L. T. Yang, J. Yang, D. Zheng and Y. Zhao, "Jointly Low-Rank Tensor Completion for Estimating Missing Spatiotemporal Values in Logistics Systems," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 2, pp. 1814-1822, 2023.

[24] Adrien Gregorj, Zeynep Yücel, Francesco Zanlungo, Claudio Feliciani and Takayuki Kanda, "Social aspects of collision avoidance: a detailed analysis of two-person groups and individual pedestrians," *Scientific reports*, vol. 13, no. 1, pp. 5756, 2023.

[25] C. Huang, H. Huang, J. Zhang, P. Hang, Z. Hu and C. Lv, "Human-Machine Cooperative Trajectory Planning and Tracking for Safe Automated Driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 12050-12063, 2022.

[26] Dirk Helbing and Péter Molnár, "Social force model for pedestrian dynamics," *Physical review E*, vol. 51, no. 5, pp. 4282, 1995.

[27] Y. Huang, H. Bi, Z. Li, T. Mao and Z. Wang, "STGAT: Modeling Spatial-Temporal Interactions for Human Trajectory Prediction," *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea (South), 2019, pp. 6271-6280.

[28] N. Li, J. Zhong, X. Shu, and H. Guo, "Weakly-supervised anomaly detection in video surveillance via graph convolutional label noise cleaning," *Neurocomputing*, vol. 481, pp. 154-167, 2022.

[29] J. Li, H. Dong, Z. Wang and X. Bu, "Partial-Neurons-Based Passivity-Guaranteed State Estimation for Neural Networks With Randomly Occurring Time Delays," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 9, pp. 3747-3753, 2020.

[30] J. Leng, H. Wang, X Gao, Y. Zhang, Y. Wang, and M. Mo, "Where to look: Multi-granularity occlusion aware for video person re-identification," *Neurocomputing*, vol. 536, pp. 137-151, 2023.

[31] Jing Lin, Lijun Cao and Nan Li, "How the completeness of spatial knowledge influences the evacuation behavior of passengers in metro stations: A VR-based experimental study," *Automation in Construction*, vol. 113, pp. 103136, 2020.

[32] J. Lian, F. Yu, L. Li and Y. Zhou, "Causal Temporal–Spatial Pedestrian Trajectory Prediction With Goal Point Estimation and Contextual Interaction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 24499-24509, 2022

[33] L. Li, B. Zhou, J. Lian, X. Wang and Y. Zhou, "Multi-PPTP: Multiple Probabilistic Pedestrian Trajectory Prediction in the Complex Junction Scene," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 13758-13768, 2022

[34] W. Li, Y. Zhang, L. Li, Y. Lv and M. Wang, "A Pedestrian Trajectory Prediction Model for Right-Turn Unsignalized Intersections Based on Game Theory," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 8, pp. 9643-9658, 2024.

[35] F. Deng, Y. Ming and B. Lyu, "CCE-Net: causal convolution embedding network for streaming automatic speech recognition," *International Journal of Network Dynamics and Intelligence*, vol. 3, no. 3, art. no. 100019, 2024.

[36] R. Korbmacher and A. Tordeux, "Review of Pedestrian Trajectory Prediction Methods: Comparing Deep Learning and Knowledge-Based Approaches," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 24126-24144, 2022.

[37] M. Kang, S. Wang, S. Zhou, K. Ye, J. Jiang and N. Zheng, "FFINet: Future Feedback Interaction Network for Motion Forecasting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 9, pp. 12285-12296, 2024.

[38] C. G. Keller and D. M. Gavrila, "Will the Pedestrian Cross? A Study on Pedestrian Path Prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 2, pp. 494-506, 2014.

[39] V. Kosaraju, A. Sadeghian, R. Martín-Martín, I. Reid, H. Rezatofighi, and S. Savarese, "Social-BiGAT: Multimodal Trajectory Forecasting using Bicycle-GAN and Graph Attention Networks," *2019 Advances in Neural Information Processing Systems (NIPS)*, Vancouver, Canada, 2019, vol.32.

[40] Z. Jiang, B. Jin and Y. Song, "A Novel Pet Trajectory Prediction Method for Intelligent Plant Cultivation Robot," *IEEE Sensors Letters*, vol. 7, no. 2, pp. 1-4, 2023.

[41] Jolyon J. Faria, Stefan Krause and Jens Krause, "Collective behavior in road crossing pedestrians: the role of social information," *Behavioral Ecology*, vol.21, no. 6, pp. 1236–1242, 2010.

[42] Z. Jiang et al., "Social NSTransformers: Low-Quality Pedestrian Trajectory Prediction," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 11, pp. 5575-5588, 2024.

[43] D. Field, A. Hayes and R. Hess, "Contour integration by the human visual system: evidence for a local "association field"," *Vision research*, vol. 33, no. 2, pp. 173–193, 1993.

[44] K. Leng and S. Li, "Distribution Path Optimization for Intelligent Logistics Vehicles of Urban Rail Transportation Using VRP Optimization Model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 2, pp. 1661-1669, 2022.

[45] Y. Li, X. -Y. Lu, J. Wang and K. Li, "Pedestrian Trajectory Prediction Combining Probabilistic Reasoning and Sequence Learning," *IEEE Transactions on Intelligent Vehicles*, vol. 5, no. 3, pp. 461-474, 2020.

[46] J. Li, F. Jiang, J. Yang, B. Kong, M. Gogate, K. Dashtipour and A. Hussain, "Lane-deeplab: Lane semantic segmentation in automatic driving scenarios for high-definition maps," *Neurocomputing*, vol. 465, pp. 15-25, 2021.

[47] K. Lyu, H. Chen and A. Che, "A Bid Generation Problem in Truckload Transportation Service Procurement Considering Multiple Periods and Uncertainty: Model and Benders Decomposition Approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 9157-9170, 2022.

[48] A. Mohamed, K. Qian, M. Elhoseiny and C. Claudel, "Social-STGCNN: A Social Spatio-Temporal Graph Convolutional Neural Network for Human Trajectory Prediction," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, 2020, pp. 14412-14420.

[49] Mehdi Moussaïd, Dirk Helbing and Guy Theraulaz, "How simple rules determine pedestrian behavior and crowd disasters," *Proceedings of the National Academy of Sciences*, vol. 108, no. 17, pp. 6884–6888, 2011.

[50] J. Ma, C. Yang, S. Mao, J. Zhang, S. C. Periaswamy and J. Patton, "Human Trajectory Completion with Transformers," *ICC 2022 - IEEE International Conference on Communications*, Seoul, Korea, Republic of, 2022, pp. 3346-3351.

[51] R. Quintero Mínguez, I. Parra Alonso, D. Fernández-Llorca and M. Á. Sotelo, "Pedestrian Path, Pose, and Intention Prediction Through Gaussian Process Dynamical Models and Pedestrian Activity Recognition," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 5, pp. 1803-1814, 2019.

[52] A. Mohammadi, K. Jamshidi, H. Shahbazi, and M. Rezaei, "Efficient deep steering control method for self-driving cars through feature density metric," *Neurocomputing*, vol. 515, pp. 107–120, 2023.

[53] A. Nayak, A. Eskandarian, Z. Doerzaph and P. Ghorai, "Pedestrian Trajectory Forecasting Using Deep Ensembles Under Sensing Uncertainty," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 9, pp. 11317-11329, 2024.

[54] Anna Sieben, Jette Schumann and Armin Seyfried, "Collective phenomena in crowds—Where pedestrian dynamics need social psychology," *PLoS one*, vol.12 no. 6, pp. e0177328, 2017.

[55] R. Quintero Mínguez, I. Parra Alonso, D. Fernández-Llorca, and M. Á. Sotelo, "Pedestrian Path, Pose, and Intention Prediction Through Gaussian Process Dynamical Models and Pedestrian Activity Recognition," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 5, pp. 1803-1814, 2019.

[56] S. Pang, J. Cao, M. Jian, J. Lai and Z. Yan, "BR-GAN: A Pedestrian Trajectory Prediction Model Combined With Behavior Recognition," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 24609-24620, 2022.

[57] Karthika P. Sobhana and Ashish Verma, "Walking in Social Groups: Role of Intra-Group Interactions," *Adaptive Behavior*, vol. 32, no. 1, pp. 33–46, 2024.

[58] A. Robicquet, A. Sadeghian, A. Alahi, and S. Savarese, "Learning Social Etiquette: Human Trajectory Understanding In Crowded Scenes," *Computer Vision–ECCV 2016: 14th European Conference (ECCV)*, Amsterdam, North Holland, The Netherlands, 2016, pp. 549–565.

[59] 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," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 2019, pp. 1349-1358.

[60] K Shen, R Quan, L Zhu, J Xiao and Y Yang, "Neural Interaction Energy for Multi-Agent Trajectory Prediction," *Proceedings of the 32nd ACM International Conference on Multimedia (ACM MM)*, Melbourne, VIC, Australia, 2024, pp. 1952–1960.

[61] Armin Seyfried, Bernhard Steffen and Thomas Lippert, "Basics of modelling the pedestrian flow," *Physica A: Statistical Mechanics and its Applications*, vol. 368, no. 1, pp. 232-238, 2006.

[62] M. Szankin and A. Kwasniewska, "Can AI see bias in X-ray images?," *International Journal of Network Dynamics and Intelligence*, vol. 1, no. 1, pp. 48-64, 2022.

[63] Z. Su, C. Wang, H. Cui, N. Djuric, C. Vallespi-Gonzalez and D. Bradley, "Temporally-Continuous Probabilistic Prediction using Polynomial Trajectory Parameterization," *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Prague, Czech Republic, 2021, pp. 3837-3843.

[64] L. Shi et al., "Representing Multimodal Behaviors With Mean Location for Pedestrian Trajectory Prediction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 9, pp. 11184-11202, 2023.

[65] H. Tang, P. Wei, J. Li, N. Zheng, "EvoSTGAT: Evolving Spatiotemporal Graph Attention Networks for Pedestrian Trajectory Prediction," *Neurocomputing*, vol. 491, page. 333-342, 2022.

[66] D. Vasquez, T. Fraichard and C. Laugier, "Incremental Learning of Statistical Motion Patterns With Growing Hidden Markov Models," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 3, pp. 403-416, 2009.

[67] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is All You Need," *Advances in Neural Information Processing Systems (NIPS)*, Long Beach, CA, USA, vol. 30, 2017.

[68] D. Wu, M. Shang, X. Luo and Z. Wang, "An L1-and-L2-Norm-Oriented Latent Factor Model for Recommender Systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 10, pp. 5775-5788, 2022.

[69] C. Wong, B. Xia, Z. Zou, Y. Wang and X. You, "SocialCircle: Learning the Angle-based Social Interaction Representation for Pedestrian Trajectory Prediction," *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, 2024, pp. 19005-19015.

[70] F. Wen, M. Li and R. Wang, "Social Transformer: A Pedestrian Trajectory Prediction Method based on Social Feature Processing Using Transformer," *2022 International Joint Conference on Neural Networks (IJCNN)*, Padua, Italy, 2022, pp. 1-7.

[71] W. Wu, M. Chen, J. Li, B. Liu and X. Zheng, "An Extended Social Force Model via Pedestrian Heterogeneity Affecting the Self-Driven Force," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 7974-7986, 2022.

[72] W. Wu, M. Chen, J. Li, B. Liu, X. Wang and X. Zheng, "Visual information based social force model for crowd evacuation," *Tsinghua Science and Technology*, vol. 27, no. 3, pp. 619-629, 2022.

[73] A. Dong, A. Starr and Y. Zhao, "Neural network-based parametric system identification: a review," *International Journal of Systems Science*, vol. 54, no. 13, pp. 2676–2688, 2023.

[74] Y. Wang, Z. Liu, Z. Zuo, Z. Li, L. Wang and X. Luo, "Trajectory Planning and Safety Assessment of Autonomous Vehicles Based on Motion Prediction and Model Predictive Control," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 8546-8556, 2019.

[75] W. Wu and X. Zheng, "A Systematic Analysis of Subgroup Research in Pedestrian and Evacuation Dynamics," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 2, pp. 1225-1246, 2024.

[76] Y. Wang, Y. Niu, W. Zhu, W. Chen, Q. Li and T. Wang, "Predicting Pedestrian Crossing Behavior at Unsignalized Mid-Block Crosswalks Using Maximum Entropy Deep Inverse Reinforcement Learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 5, pp. 3685-3698, 2024.

[77] D. Wang, H. Liu, N. Wang, Y. Wang, H. Wang and S. McLoone, "SEEM: A Sequence Entropy Energy-Based Model for Pedestrian Trajectory All-Then-One Prediction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 1, pp. 1070-1086, 2023.

[78] Y. Xue, R. Yang, X. Chen, Z. Tian and Z. Wang, "A Novel Local Binary Temporal Convolutional Neural Network for Bearing Fault Diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1-13, 2023.

[79] C. Xu, W. Mao, W. Zhang and S. Chen, "Remember Intentions: Retrospective-Memory-based Trajectory Prediction," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, 2022, pp. 6478-6487.

[80] W. Yi, W. Wu, X. Wang and X. Zheng, "Modeling the Mutual Anticipation in Human Crowds With Attention Distractions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 9, pp. 10108-10117, 2023.

[81] X. Yan, J. Yang, L. Song and Y. Liu, "PSA-GRU: Modeling Person-Social Twin-Attention Based on GRU for Pedestrian Trajectory Prediction," *2021 40th Chinese Control Conference (CCC)*, Shanghai, China, 2021, pp. 8151-8157.

[82] C. Yu, X. Ma, J. Ren, H. Zhao, and S. Yi, "Spatio-temporal graph transformer networks for pedestrian trajectory prediction," *Computer Vision–ECCV 2020: 16th European Conference*, Glasgow, UK, August 23–28, 2020, Proceedings, Part XII 16.

[83] Y. Yang, Z, Fu, and SM. Naqvi, "Abnormal event detection for video surveillance using an enhanced two-stream fusion method," *Neurocomputing*, vol. 553, pp. 126561, 2023.

[84] J. Yang, Y. Chen, S. Du, B. Chen and J. C. Principe, "IA-LSTM: Interaction-Aware LSTM for Pedestrian Trajectory Prediction," *IEEE Transactions on Cybernetics*, vol. 54, no. 7, pp. 3904-3917, 2024.

[85] Y. Wang, C. Wen and X. Wu, "Fault detection and isolation of floating wind turbine pitch system based on Kalman filter and multi-attention 1DCNN," *Systems Science & Control Engineering*, vol. 12, no. 1, art. no. 2362169, 2024.

[86] C. Yang, H. Pan, W. Sun and H. Gao, "Social Self-Attention Generative Adversarial Networks for Human Trajectory Prediction," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 4, pp. 1805-1815, 2024.

[87] Y. Wang, C. Shen, J. Huang and H. Chen, Model-free adaptive control for unmanned surface vessels: a literature review, *Systems Science & Control Engineering*, vol. 12, no. 1, art. no. 2316170, 2024.

[88] B. Yang, C. He, P. Wang, C. -Y. Chan, X. Liu and Y. Chen, "TPPO: A Novel Trajectory Predictor With Pseudo Oracle," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 54, no. 5, pp. 2846-2859, 2024.

[89] B. Yu, K. Zhu, K. Wu and M. Zhang,"Improved OpenCL-Based Implementation of Social Field Pedestrian Model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 7, pp. 2828-2839, 2020.

[90] X. Yan, X. Zhong, Z. Yang, R. Zhang, W. Huang and Z. Wang, "Global Temporal Attention Optimization for Human Trajectory Prediction," *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Prague, Czech Republic, 2022, pp. 184-189.

[91] C. Zhang, J. Sprenger, Z. Ni and C. Berger, "Predicting Pedestrian Crossing Behavior in Germany and Japan: Insights into Model Transferability," *IEEE Transactions on Intelligent Vehicles*, early access.

[92] C. Zhang and C. Berger, "Pedestrian Behavior Prediction Using Deep Learning Methods for Urban Scenarios: A Review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 10, pp. 10279-10301, 2024.

[93] X. Zhou, W. Zhao, A. Wang, C. Wang and S. Zheng, "Spatiotemporal Attention-Based Pedestrian Trajectory Prediction Considering Traffic-Actor Interaction," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 1, pp. 297-311, 2023.

[94] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, "Informer: Beyond Efficient Transformer for Long Sequence Time-series Forecasting," *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2021, vol. 35, no. 12, pp. 11106-11115.

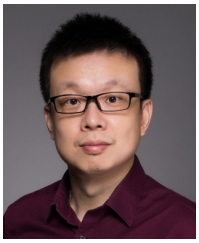
[95] H. Zhou, D. Ren, H. Xia, M. Fan, X. Yang and H. Huang, "AST-GNN: An Attention-based Spatio-Temporal Graph Neural Network For Interaction-Aware Pedestrian Trajectory Prediction," *Neurocomputing*, vol. 445, pp. 298–308, 2021.

[96] Y. Wu, X. Huang, Z. Tian, X. Yan and H. Yu, "Emotion contagion model for dynamical crowd path planning," *International Journal of Network Dynamics and Intelligence*, vol. 3, no. 3, art. no. 100014, 2024.

[97] R. Zhao, Q. Hu, Q. Liu, C. Li, D. Dong and Y. Ma, "Panic Propagation Dynamics of High-Density Crowd Based on Information Entropy and Aw-Rascle Model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 10, pp. 4425-4434, 2020.
[98] P. Zhang, J. Xue, P. Zhang, N. Zheng and W. Ouyang, "Social-Aware Pedestrian Trajectory Prediction via States Refinement LSTM," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 5, pp. 2742-2759, 2022.
[99] K. Zhu, Z. Wang, G. Wei and X. Liu, "Adaptive Set-Membership State Estimation for Nonlinear Systems Under Bit Rate Allocation Mechanism: A Neural-Network-Based Approach," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 11, pp. 8337-8348, 2023.
[100] K. Zhu, Z. Wang, Y. Chen and G. Wei, "Neural-Network-Based Set-Membership Fault Estimation for 2-D Systems Under Encoding–Decoding Mechanism," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 2, pp. 786-798, 2023.
[101] N. Zeng, P. Wu, Z. Wang, H. Li, W. Liu and X. Liu, "A Small-Sized Object Detection Oriented Multi-Scale Feature Fusion Approach With Application to Defect Detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1-14, 2022.
[102] H. Zhou, X. Yang, D. Ren, H. Huang and M. Fan, "CSIR: Cascaded Sliding CVAEs With Iterative Socially-Aware Rethinking for Trajectory Prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 12, pp. 14957-14969, 2023.
[103] P. Zhang et al., "Towards Trajectory Forecasting From Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 10, pp. 12550-12561, 2023.
[104] X. Zhong et al., "Visual Exposes You: Pedestrian Trajectory Prediction Meets Visual Intention," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 9, pp. 9390-9400, 2023.
[105] S. Zhang, M. Abdel-Aty, Y. Wu and O. Zheng, "Pedestrian Crossing Intention Prediction at Red-Light Using Pose Estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 3, pp. 2331-2339, 2022.
[106] W. Zhu, Y. Liu, M. Zhang and Y. Yi, "Reciprocal Consistency Prediction Network for Multi-Step Human Trajectory Prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 6, pp. 6042-6052, 2023.
[107] Y. Zhang, X. Shi, S. Zhang and A. Abraham, "A XGBoost-Based Lane Change Prediction on Time Series Data Using Feature Engineering for Autopilot Vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 19187-19200, 2022.

Zihan Jiang (Student Member, IEEE) received the B.Eng. degree from the Northeast Forestry University, Harbin, China, in 2022 and the M.Sc. degree from the University of Liverpool, UK, in 2024. He is pursuing a Ph.D. degree at Tongji University, Shanghai, China. His research interests include trajectory prediction and computer vision.

Chengxuan Qin received the MRes degree in Pattern Recognition and Intelligent Systems from the University of Liverpool in 2023. He is pursuing a Ph.D. degree at the University of Liverpool, China. His research interests include brain-computer interfaces, temporal signal analysis, and machine learning.



Rui Yang (Senior Member, IEEE) received the B.Eng. degree in Computer Engineering and the Ph.D. degree in Electrical and Computer Engineering from National University of Singapore in 2008 and 2013 respectively.

He is currently an Associate Professor in the School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, China, and an Honorary Lecturer in the Department of Computer Science, University of Liverpool, Liverpool, United Kingdom. His research interests include machine learning based data analysis and applications. He is the author or co-author of several technical papers and also a very active reviewer for many international journals and conferences. Dr. Yang is currently serving as an Associate Editor for Neurocomputing, Cognitive Computation, and IEEE Transactions on Instrumentation and Measurement.

Bingyu Shi (Student Member, IEEE) received the B.Eng degree from Northeast Forestry University in Harbin, China, in 2022, and is pursuing a Ph.D. degree at Northeast Forestry University. Her research interests include remote sensing data understanding and wildfire prediction.

Fuad E. Alsaadi received the B.Sc. and M.Sc. degrees in electronic and communication from King AbdulAziz University, Jeddah, Saudi Arabia, in 1996 and 2002, respectively and the Ph.D. degree in optical wireless communication systems from the University of Leeds, Leeds, U.K., in 2011. Between 1996 and 2005, he was with Jeddah as a Communication Instructor with the College of Electronics & Communication. He is currently an Associate Professor with the Electrical and Computer Engineering Department within the Faculty of Engineering, King Abdulaziz University, Jeddah, Saudi Arabia. He has authored or coauthored widely in the top IEEE Communications Conferences and Journals. His research interests include optical systems and networks, signal processing, synchronization and systems design. He was the recipient of the Carter Award, University of Leeds for the best Ph.D.

Zidong Wang (Fellow, IEEE) was born in Jiangsu, China, in 1966. He received the B.Sc. degree in mathematics from Suzhou University, Suzhou, China, in 1986, and the M.Sc. degree in applied mathematics and the Ph.D. degree in electrical engineering from the Nanjing University of Science and Technology, Nanjing, China, in 1990 and 1994, respectively.

From 1990 to 2002, he held teaching and research appointments in universities in China, Germany, and the U.K. He is currently a Professor of dynamical systems and computing with the Department of Computer Science, Brunel University London, Uxbridge, U.K. He has published more than 600 articles in international journals. His research interests include dynamical systems, signal processing, bioinformatics, control theory, and applications.

Prof. Wang is a member of the Academia Europaea and the European Academy of Sciences and Arts, an Academician of the International Academy for Systems and Cybernetic Sciences, a fellow of the Royal Statistical Society, and a member of the Program Committee of many international conferences. He holds the Alexander von Humboldt Research Fellowship of Germany, the JSPS Research Fellowship of Japan, and the William Mong Visiting Research Fellowship of Hong Kong. He serves (or has served) as the Editor-in-Chief for International Journal of Systems Science, Neurocomputing, and Systems Science and Control Engineering; and an Associate Editor for 12 international journals, including IEEE Transactions on Automatic Control, IEEE Transactions on Control Systems Technology, IEEE Transactions on Neural Networks and Learning Systems, IEEE Transactions on Signal Processing, and IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews.