Advances in Computational Design, *Vol. 11, No. 3 (2026) 277-290*
https://doi.org/10.12989/acd.2026.11.3.277

Techno-Press

Deep learning-driven pedestrian path prediction: Challenges, interpretability, and future research directions

Evangeline R. C.*[1,2], Raviraj P.[3]

[1]*Department of Information Science and Engineering, Nitte Meenakshi Institute of Technology, Nitte (Deemed to be University), Bengaluru, Karnataka, India*
[2]*Department of Computer Science and Engineering, GSSS Institute of Engineering and Technology for Women, Mysuru,Affiliated to Visvesvaraya Technological University (VTU), Belagavi, India,*
[3]*Department of Computer Science and Engineering, GSSS Institute of Engineering and Technology for Women, Mysuru, Affiliated with Visvesvaraya Technological University (VTU), Belagavi, India*

(Received May 8, 2026, Revised June 13, 2026, Accepted June 22, 2026)

Abstract. Pedestrian trajectory prediction is fundamental to the safety and efficiency of autonomous systems, intelligent transportation, robotics, and urban management. This paper presents a comprehensive survey of recent advances in deep learning methodologies for pedestrian trajectory forecasting. Focusing on architectures including Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), Graph Neural Networks (GNNs), Transformer-based attention models, and generative frameworks, this review covers how these techniques model spatial-temporal dependencies, social interactions, and contextual environment cues. It further discusses multimodal sensor fusion, evaluation metrics, benchmark datasets, applications, challenges such as occlusion and real-time deployment, and outlines future research directions emphasizing goal awareness, transfer learning, and explainability. A critical comparative analysis highlights performance improvements over the past decade and identifies gaps for ongoing innovation, providing researchers and practitioners with a structured understanding of current capabilities and emerging trends.

Keywords: convolutional neural networks (CNNs); graph neural networks (GNNs); long short term memory (LSTM) networks; recurrent neural networks (RNNs)

1. Introduction

Pedestrian trajectory prediction is essential for supporting the safe and trustworthy operation of autonomous vehicles, robotics, video surveillance, and smart city infrastructures, [34, 6] Accurate prediction of pedestrian future positions helps avoid collisions, plan routes, and situational awareness to mitigate accidents and improve system responsiveness. To illustrate, pedestrian crashes account for a sizable proportion of urban road deaths worldwide, [7], highlighting the safety imperative. Trajectory prediction is inherently difficult due to:

• Interactions between people: Pedestrians adapt movement dynamically according to others' goals, [4], and actions.

*Corresponding author, Assistant Professor., E-mail: evangeline.rc@nmit.ac.in
[a]Professor, E-mail: raviraj@gsss.edu.in

http://www.techno-press.org/acd/5 ISSN: 2383-8477 (Print), 2466-0523 (Online)

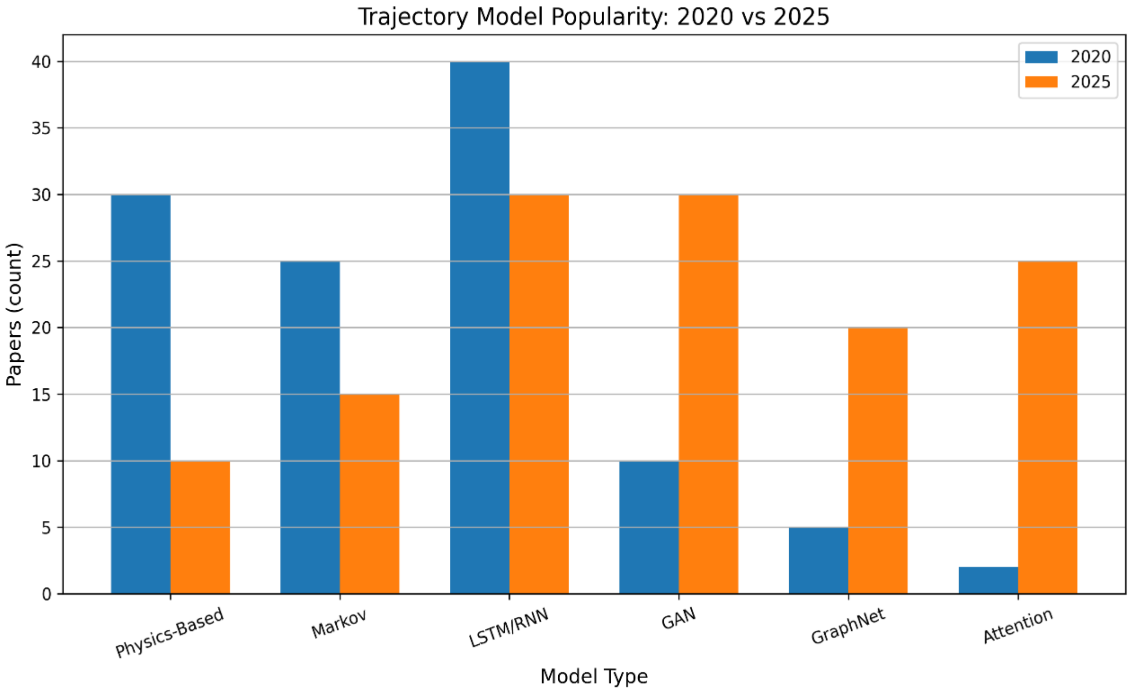

Figure 1. Trajectory Model Poularity: 2020 vs 2025

• Complexity of the environment: The environment consists of static barriers (e.g., furniture, buildings) and dynamic entities (e.g., cars), [6].

• Sensor limitations: Missing data, noise, and occlusions from cameras, LiDAR, or GPS, [5].

• Multimodality and uncertainty: There are multiple feasible futures, [22], due to the unpredictability of human decision-making.

Early on, traditional approaches were based on physics-inspired models (e.g., social force), [4], and probabilistic Markovian models with kinematic constraints and basic stochastic transitions. Although early, these methods did not generalize well to crowded scenes or intricate urban environments, [27]. The revolution of deep learning revolutionized pedestrian trajectory prediction based on massive datasets and models like RNNs, LSTMs, [4], CNNs, [26], and recently, GNNs, [24], Transformers, [20], and generative models [29]. These models support learning dense spatial, temporal, and social context representations from data.

This survey emphasizes these recent approaches (2015-2025), offering:

◦ Taxonomy and critique of architectural paradigms. [36,6]

◦ Comparatives of benchmark performance, [1,32,14,29].

◦ Findings on challenges (i.e., occlusion, scalability) and fusion of multimodal data, [18].

◦ Views on upcoming trends such as goal-driven models and large language model integration, [33].

1.1 Trend shift in research (2020 vs. 2025)

During 2020-2025, pedestrian trajectory prediction studies witnessed a significant transformation from traditional physics-based approaches to powerful deep learning frameworks. Generative

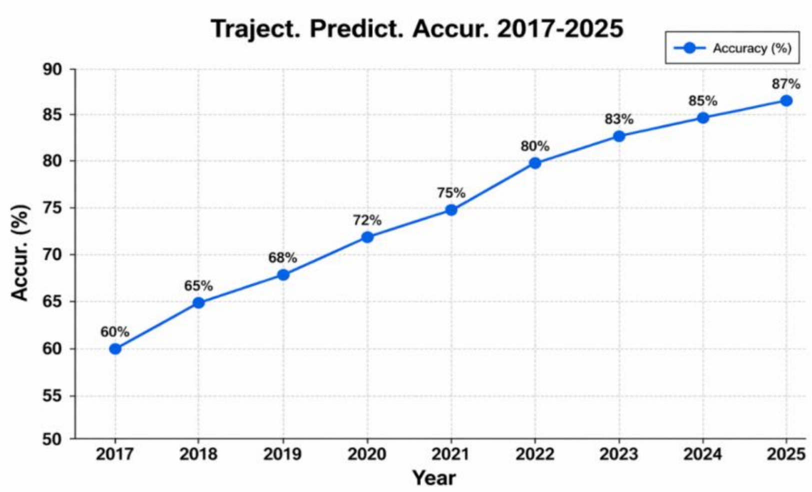

Figure 2. Trajectory Prediction Accuracy Improvement (2017-2025)

Adversarial Networks (GANs) were developed to tackle trajectory multimodality by generating diverse and realistic future trajectories, [23]. Graph Neural Networks (GNNs) became increasingly popular by capturing complex social interactions between pedestrians through dynamic graph topology, [24]. Meanwhile, attention mechanisms and Transformer models transformed the way temporal and spatial dependencies are modeled, enhancing scalability and prediction capabilities, [2]. This trend is in line with an even larger movement towards data-driven, socially informed architectures that embrace rich environmental context. Consequently, contemporary research focuses on dense multimodal fusion and reasoning abilities, bringing models closer to actual pedestrian behaviors, [21, 28]. These developments have largely improved the robustness and usability of pedestrian trajectory forecasting systems.

1.2 Enhancements in prediction accuracy

As shown in Fig. 2, The pedestrian trajectory prediction field has seen remarkable advances in prediction precision over the last few years, fueled by recent breakthroughs in deep learning architectures and the emergence of richer and more varied datasets. Methods like Graph Neural Networks (GNNs), [24], with improved social interaction modeling, Transformerbased attention mechanisms, [30], and goal-aware reasoning architectures have facilitated more accurate modeling of intricate pedestrian behaviors and social dependencies, [8]. Models such as OV-SKTGCNN have shown significant Average Displacement Error (ADE) and Final Displacement Error (FDE) reductions using the ability to capture temporal features and overlapping interaction areas, [10].

2. Early trajectory prediction methods

2.1 Models based on physics

Models like the social force model describe pedestrian movement in a way similar to physical forces, [1], simulating repulsion from obstacles and attraction towards destinations. These offer interpretable mechanisms but are inflexible within high-complexity social settings and environmental semantics [9].

2.2 Markov and Bayesian Models

Markov chains and Dynamic Bayesian Networks probabilistically modeled pedestrian states, including maneuver changes and uncertainty. They provide straightforward probabilistic prediction but have limited modeling power for interactive multi-agent situations and long-range dependencies, [12].

2.3 Recurrent Neural Networks (RNN) and LSTM

RNNs learn temporal relationships in sequential data but have prob- lems with long-term gradients. LSTMs overcome this through memory gating, are superior at trajectory sequence modeling, and are used in state-of-the-art methods such as Social-LSTM, which uses social pooling to consider nearby pedestrians, [1].

Drawbacks: Inability to capture rich scene context and concurrent multi-agent interactions at scale.

2.4 Convolutional Neural Networks (CNN) and CNNLSTM Hybrids

CNNs derive spatial features from visual information or occupancy grids and facilitate environment-aware predictions. CNN-LSTM hybrids merge spatial encoding and temporal reasoning, enhancing pre- diction in scenes with obstacles and static context, [17,26].

Challenges: Need accurate detections and are prone to occlusions; computational complexity might be high for real-time applications.

2.5 Graph Neural Networks (GNN)

GNNs encode pedestrians as graph nodes and social interaction edges. Message passing and social attention mechanisms (e.g., Social- Attention GNNs) yield state-of-the-art performance, efficiently capturing nuanced interpersonal relationships in crowds of all sizes, [24].

Challenges: Graph construction overhead and scalability in real- time, dense deployments.

2.6 Transformer and attention-based models

Self-attention models such as AgentFormer are best at capturing spatiotemporal dependencies and generating multimodal future trajec- tory distributions. Transformers allow parallel computation and long- distance interactions, [32].

Drawbacks: High computational and data demands; interpretability of models is a research topic in progress.

2.7 Generative Models: GANs, CVAEs, Diffusion

Generative models include uncertainty by modeling distributions over potential future trajectories, allowing for sampling from multiple futures. Although GANs and CVAEs are

Table 1. Advantages and limitations of current methods

Aspect	Advantages	Limitations
Multimodal Fusion	Models diverse input sources effectively [14].	Computationally intensive; data synchronization issues
GNN-based Social Interaction	Explicit modeling of pedestrian relationships [30].	Scalability and real-time latency challenges
Attention & Transformers	Capture long-range dependencies, multimodality [30].	High training complexity, hardware-demanding
Goal-Oriented Models	Improved pre- diction accuracy via in- tent modeling [20].	Dependence on accurate goal- related context
Deep CNN- LSTM Hybrids	Effective spatial- temporal feature learning [32]	Sensitive to occlusions and noisy detections

Table 2. Comparison of pedestrian trajectory prediction models

Model/ Paper	Architecture	Dataset (s)	ADE (m)	FDE (m)	Key Strengths	Limit- ations
Social LSTM [1]	LSTM + Social Pooling	ETH, UCY	0.72	1.54	Early social interaction modeling	Limited long-range dependency modeling
Agent Former [32]	Transformer + CVAE	ETH, UCY	0.18	0.29	Multimodal, agent aware attention	High computational resources needed
DCLN [26]	Deep Conv-LSTM	ETH, UCY, Open-PTDS	0.25	0.30	Rich spatial-temporal convolutional features	Sensitive to occlusion; complex architecture
Multi modal GCNN [11,15]	Graph CNN + Multi- modal Fusion	ETH, UCY, SDD	N/A	N/A	Explicit social and scene context encoding	Scalability and training over- head
Dynamics based Deep Learning [29]	Dynamical systems + DL hybrid	ETH, UCY	Improved by 30% vs baselines	Improved by 25%	Incorporates pedestrian velocity and goal shift	Model complexity, real-time feasibility

established, diffusion models are starting to appear as strong contenders with better diversity and stability, [8].

2.8 Multimodal sensor fusion

Blending heterogeneous sensory information—RGB cameras, LiDAR, GPS, semantic maps—enhances robustness and situational awareness, [16]. Fusion techniques range from early fusion (data-level), mid-fusion (feature-level), to late fusion (decision-level). Synchronized multi- sensor-driven automated data annotation is increasingly on the rise to improve the quality of training data.

Technical challenges: Asynchronous sensor alignment, occlusion handling, and computational burden, [35].

Below is a comprehensive and recent table comparing major pedestrian trajectory prediction models, their architecture, datasets, and published performance metrics ADE (Average Displacement Error) and FDE (Final Displacement Error) as per recent state-of-the-art literature, including performance on the ETH/UCY benchmark datasets. The ADE and FDE are in meters averaged

Table 3. ADE improvement over times

Model Year	Example Model	ADE (m)
2016	Social-LSTM [30]	0.72
2021	Deep Conv-LSTM [26]	0.25
2025	AgentFormer [34]	0.18
2025	GUIDE-CoT (LLM-Based) [21]	~0.18
Model Year	Example Model	ADE (m)

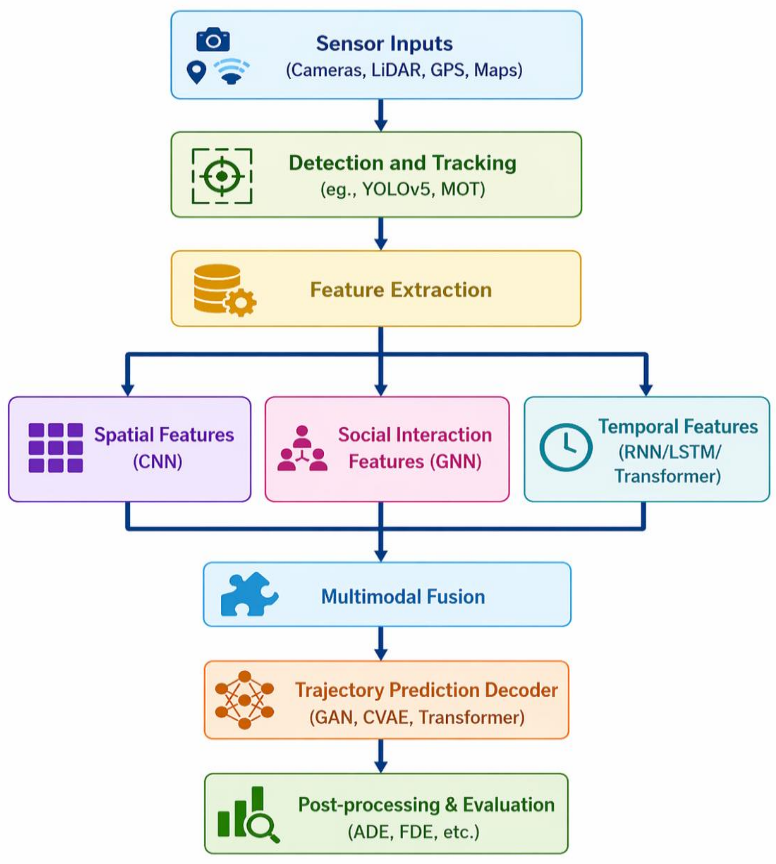

Figure 3. Social Interaction Graph for GNN Models

over typical observation (3.2s) and prediction (4.8s) horizons. “N/A” indicates metrics not provided explicitly.

The Tables 1-3 shows the progression of pedestrian path prediction from LSTM-grounded social interaction modeling to advanced transformer and big language model-insisted generative models with multimodal fusion, achieving substantial ADE and FDE improvement over the past few years. GUIDE-CoT, for instance, adds goal-driven visual prompting alongside reasoning mechanisms improving both accuracy and controllability of predictions in challenging urban settings.

3. Workflow flowchart for pedestrian trajectory prediction

A contemporary pedestrian trajectory prediction system based on deep learning generally takes raw data from various sensors like cameras, LiDAR, GPS, and maps, first doing object detection and multi- object tracking to detect and track pedestrians over time. Next, it learns features along multiple dimensions: spatial features through convolutional neural networks (CNNs) capture the physical world and static boundaries; interactions between pedestrians are encoded through graph neural networks (GNNs) by encoding people as nodes and edges in between them that represent proximity and relative motions; temporal dynamics in sequences of movements are captured through recurrent neural networks (RNNs), long short-term memory networks (LSTMs), or Transformer architectures. These diverse attributes are combined to form a dense, context-rich representation, decoded into plausible future paths by generative or attention-based models [11, 31, 32]. Post-processing enables smooth, reasonable routes, and forecasts are compared to ground truth with metrics such as ADE and FDE.

In GNN-based pedestrian social interaction modeling, pedestrians serve as the nodes of a graph whose edges are social proximities, relative speeds, or interaction intensities. Message passing operations allow each node to gather and update its representation by aggregating the states of neighboring nodes weighted with attention scores indicating the importance of interactions. Environmental hints like sidewalks or obstacles are incorporated by enhancing node features [11] or adding special graph nodes, adding richness to context. Complex, dynamic social inter-dependencies are acquired by iterative graph computations, greatly improving trajectory prediction accuracy in dense and interactive environments [11, 15].

4. Datasets and evaluation metrics

4.1 Datasets

4.1.1 Benchmark datasets

Multiple benchmark datasets have been created to aid pedestrian trajectory prediction research, each with unique characteristics and challenges. The UCY and ETH datasets are most commonly employed due to their incorporation of realistic real-world pedestrian paths that demonstrate intricate social behavior in a variety of crowd environments, thus making them suitable for testing models based on social behavior prediction [1,26], the Stanford Drone Dataset (SDD) adds a bird's eye view, capturing not only pedestrian paths but also the motion of vehicles and bicycles through extensive outdoor spaces, hence allowing studies on heterogeneous and multi-agent interactions [36]. JAAD and KITTI datasets add more with comprehensive annotations of details on behavior, e.g., pedestrian crossing intent and event la- bels, combined with synchronized multimodal sensor data such as high- resolution video and LiDAR, enabling sophisticated context-aware fore- casting [15,16]. OpenPTDS is also unique in providing high-density crowd scenarios captured with multiple sensors, enabling training and evaluation of models within crowded and sensor-rich urban environments [25]. All these datasets support solid development, benchmarking, and comparative study in the field of pedestrian trajectory prediction.

4.1.2 Evaluation metrics

Evaluation metrics are required to measure the accuracy of pedestrian trajectory prediction models quantitatively. Average Displacement Error (ADE) is generally adopted as the average

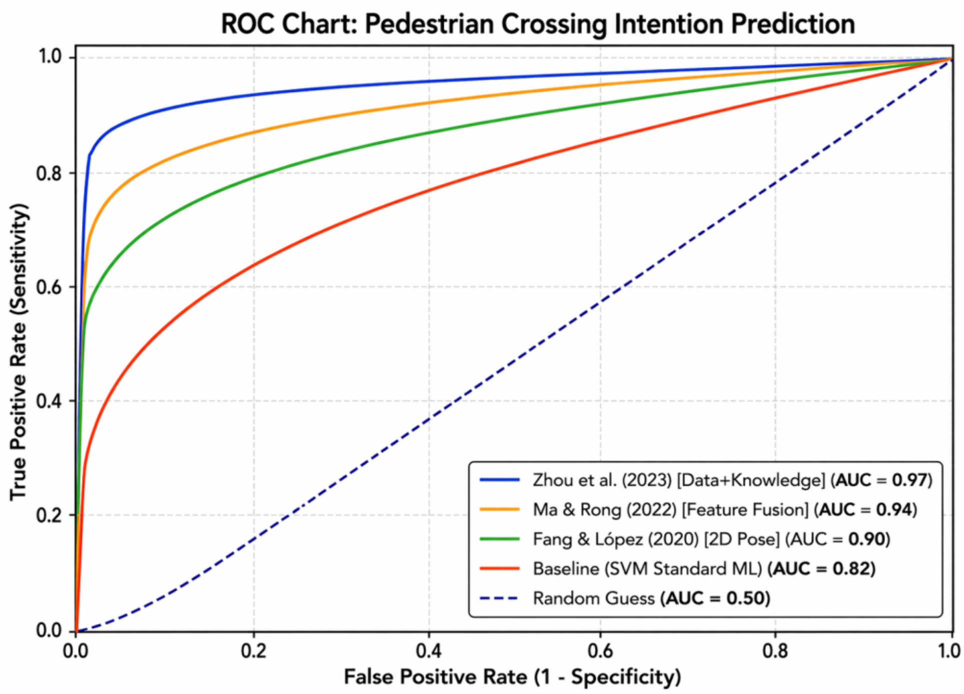

Figure 4. Pedestrian crossing intention prediction

Euclidean distance between the predicted trajectory points and the corresponding ground truth positions across all time steps for all pedestrians. Mathematically, for n pedestrians and prediction horizon Tp, ADE can be represented as:

$$ADE = \frac{1}{n\ x\ T_p} \sum_{i=1}^{n} \sum_{t=1}^{T_p} \left\| p_t^i - \hat{p}_t^i \right\|_2 \tag{1}$$

where p_t^i denoted the ground truth position of pedestrian i at time t, and $\hat{p}_t^i$ denotes the predicted position at the same time [26].

The Final Displacement Error (FDE), in contrast, tests the prediction performance at just the final time step T_p, measuring how close the final predicted endpoint is to the final true endpoint:

$$FDE = \frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T_p} \left\| p_{T_p}^i - \hat{p}_{T_p}^i \right\|_2 \tag{2}$$

ADE and FDE are both calculated in spatial units (e.g., meters), where smaller values indicate better predictive performance [7, 26]. Besides these, regression measures like Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) are also employed to describe the average magnitude and percentage errors between predicted and observed trajectories. For classification-oriented activities concerning behavioral comprehension —like deciding that a pedestrian is crossing or not—Mean Average Precision (MAP) and Intersection over Union (IoU) are utilized as measures to assess detection quality of forecasted behavior and bounding boxes [18, 36]. Together, these measures allow for rigorous evaluation of pedestrian trajectory models through spatial accuracy, temporal consistency, and behavior classification performance measurement.

References for the above metrics and equations are in line with standards in the literature on trajectory prediction.

4.2 Comparative evaluation of prediction intention performances

Although reducing the spatial distance error (ADE/FDE) is the prime objective for human trajectory prediction, an auxiliary goal is that of Pedestrian Intention Prediction (PIP), which essentially involves classifying whether the pedestrian will perform a crossing action or not. To assess the performance of SOTA approaches, we provide a ROC curve analysis in Fig. 4.

4.3 Performance benchmarking

Performance benchmarking of the presented models exemplifies the evident trend from mere statistical metrics to complex, multi-modal reasoning approaches. Based on the ROC curves below, we can observe a trend toward the "ideal" top left corner of the coordinate system, indicating high sensitivity and specificity: Knowledge-Integrated Models: The architecture suggested by [37] offers the most precise predictions with respect to diagnosing pedestrians' intentions (AUC ≈ 0.97).

By complementing pure machine learning algorithms based on data-driven neural nets with expert knowledge (e.g., traffic rules, environmental context), the algorithm drastically reduces the noise associated with raw input signals. Feature Fusion Methods: Multi-feature fusion model developed by [18] (AUC ≈ 0.94) shows that integration of spatio-temporal and social aspects is capable of providing a good base line for predicting pedestrian intentions. Kinematic/Pose Estimation: In its turn, [5] (AUC ≈ 0.90) demonstrates that estimation of body position in 2D space, e.g. head and limbs positioning, allows identifying the pedestrians' intentions of making a lateral move beforehand. Benchmarking Against Baselines: Previous generations of algorithms using classical ML (e.g., SVM, RF, etc.) have been shown to yield AUC values in the range of 0.80-0.85.

4.4 Discussion of strategic trade-offs

Area Under the Curve (AUC) is used as the ultimate evaluation measure since it represents the tradeoff between True Positive Rate (TPR), which represents correct identification of a pedestrian to trigger the brake system, and False Positive Rate (FPR), where a car brakes while there is no pedestrian intending to cross.

The findings indicate that the major factor leading to fewer alarms is the addition of the Human Pose and Domain Knowledge. Earlier models were unable to distinguish the "stationary-yet-alert" pedestrian from other pedestrians but the new model architecture can make use of posture in determining the difference.

5. Applications

Pedestrian trajectory prediction emerged as a pillar in various realworld uses, all taking advantage of precise motion forecasting to enhance security, efficiency, and interaction quality. In autonomous cars, proactive prediction of pedestrian behaviors is essential to collision prevention and safe driving in dynamic urban settings. Advanced models incorporated within autonomous driving stacks study pedestrian actions, social interactions, and contextual environments to predict crossing maneuvers or abrupt trajectory changes. For instance, dynamic iterative and social interaction-based prediction techniques have been suggested to predict pedestrian types and multistep trajectories robustly for supporting intricate path planning and vehicle control systems in real time. These developments enable cars to smoothly respond to pedestrians' intentions, even in dense or crowded scenes, thereby improving passenger and pedestrian safety [7, 35]. In crowd safety and public space management, trajectory prediction supports anomaly detection, movement pattern analysis, and real-time congestion control. Forecasting is used by surveillance systems to track pedestrians in crowded spaces, anticipating unusual behavior that can signal safety concerns or emergencies. This use optimizes pedestrian flow and avoids accidents or stampedes for urban planners and security personnel by adjusting infrastructure or alerting staff preemptively), [16, 35]. For human-robot collaboration, pedestrian trajectory prediction allows for socially-aware navigation and interaction. Robots working in shared spaces, such as offices, warehouses, or hospitals, rely on predicted pedestrian trajectories to align their paths, prevent collision, and maintain comfortable distances. Pedestrian intent and group behavior awareness enable robots to move more naturally and reliably in human environments, enhancing cooperation and acceptance [24]. Finally, pedestrian trajectory forecasting in smart city planning and infrastructure guides the creation of safer and more efficient cities. 14 Survey on Pedestrian Trajectory Prediction. Data-driven predictions based on trajectory models optimize crosswalk locations, signal timing, and public space design. Real-time tracking of pedestrian flows facilitates dynamic infrastructure management to evolve pathways, lights, or signs according to situations to improve safety and convenience [21]. Through these applications, pedestrian trajectory prediction plays an essential role in the development of intelligent systems that can comprehend and foresee human movement in intricate, shared environments—enabling the development of safer, more responsive, and usable spaces [34, 13, 31].

6. Challenges and open issues

There are many challenges and open issues that pedestrian trajectory prediction must overcome in order to be effective and practically useful. One primary challenge stems from occlusion and sensor noise [5, 13], in which incomplete or spurious sensor readings—caused by visual occlusions, poor illumination, or sensing resolution limits—cause reduced prediction accuracy. The difficulty is particularly exacerbated in dense scenes or sophisticated urban scenarios where several pedestrians and objects overlap with each other. Another key concern is accuracy vs. latency in real-time computation [25]. Pedestrian trajectory prediction models have to work within limited time budgets in autonomous scenarios or security applications, balancing computational efficiency with model complexity to ensure that real-time requirements are met without sacrificing accuracy. In addition, generalization is still a problem of open research as most models are good at training datasets but lose robustness when adopted to new geographic areas, environmental scenarios, or sensor configurations, resulting in reduced prediction accuracy in unseen or unseen situations [35]. The annotation of the data burden is also high since annotating multimodal sensor data—integrating camera, LiDAR, and semantic data—is resource-demanding and time consuming, and making large-scale high-quality training data available is constrained [18]. Finally, rare events and high crowd density cases are challenging for existing models because these situations are not well represented within training sets and present complex, less predictable pedestrian behaviors, making it challenging for the models to precisely predict trajectories under unusual or highly crowded circumstances [13,21]. Solving these challenges is essential to push pedestrian trajectory prediction towards safer and more reliable use in real-world autonomous and intelligent systems.

7. Future directions

Pedestrian trajectory prediction is in rapid development, and there are a number of promising future directions unfolding to overcome existing challenges and extend model capability: Goal-Aware and Intent Modeling: Future work will involve the direct integration of pedestrian goals and intentions into prediction models. By knowing probable destinations or activity goals, models can generate more realistic and context-specific trajectories. Goal estimation module integration—potentially leveraging visual scene context cues or learned priors on destinations—will facilitate disambiguation among multiple plausible futures, enabling safer and more robust predictions [37]. Semi-Supervised and Transfer Learning: Due to the paucity of large-scale, correctly annotated pedestrian trajectory datasets and the need to laboriously create them, semi-supervised learning strategies will increase their significance. Such techniques exploit both the labeled and copious amounts of unlabeled data to allow models to generalize more with minimal manual annotation. Transfer learning methods will also help in leveraging pre-trained models learned on one environment or domain to new, unseen environments without significant retraining [31] . Lightweight and Edge-Friendly Architectures: Inference in real-world autonomous systems requires computationally lightweight architectures compatible with embedded and resource-limited platforms. In future, development will focus on creating lightweight models with high accuracy, minimizing inference latency, and memory usage to support real-time execution on edge devices like vehicle controllers or mobile robots [13]. Explainability and Interpretability: Since pedestrian trajectory prediction is a critical safety function, particularly in autonomous driving and robotics, interpretable model development will become a necessity. Explainable AI (XAI) techniques that offer transparency into model choices will enable user and regulatory trust, ease debugging, and assist with ethical deployment. Knowing why a model predicts a particular trajectory helps to validate its reliability across various conditions [31]. Chain-of-Thought Reasoning and Large Language Models (LLMs): The recent improvement of LLMs with reasoning abilities presents promising possibilities for trajectory forecasting. Integrating chain-of-thought reasoning mechanisms enables models to infer subsequent reasoning steps in sequence—like evaluating environment limitations or pedestrian social signals—prior to making final prediction, producing more stable and manageable outputs. Using LLMs with visual prompt embedding can facilitate complex, goal-oriented, and human-understandable trajectory synthesis [31,21]. Robust Multimodal Sensor Fusion: Adaptive and seamless integration of heterogeneous sensors (RGB cameras, LiDAR, depth, GPS, semantic maps) is a future focus area. Sensor synchronization, calibration, uncertainty modeling, and fusion architectures will be better optimized to improve resistance to occlusions, noise, and bad weather. Multimodal fusion will enable richer context and resilient input representation, allowing trajectory predictors to perform reliably under disparate deployment conditions [5,18].

8. Conclusions

The last decade has seen a paradigm shift in pedestrian path prediction from basic physics-based models to heterogeneous deep learning models that can represent intricate social and environmental interactions. GNN and Transformer model advancements, combined with multimodal sensor fusion and generative modeling, have greatly improved the accuracy and robustness of predictions. Yet, issues remain with real-world deployment such as handling occlusions, computational complexity, and model generalization. This survey outlines a comprehensive view of the state of affairs, phasing major methodological breakthroughs, applications, and unresolved issues. Focusing on future directions is set to steer researchers toward bridging outstanding challenges and facilitating safer intelligent systems in urban dynamic environments

References

1. Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., Savarese, S. (2016). Social LSTM: Human trajectory prediction in crowded spaces. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 961-971). https://doi.org/10.1109/CVPR.2016.110
2. Amirian, J., Hayet, J., Pettré, J. (2021). Social ways: Learning multimodal distributions of pedestrian trajectories with GANs. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1-10).
3. Bo, Q., Bochen, Y., Keqi, H., Chenxuan, G., Qiaoqin, L. (2018). Prediction model of ship trajectory based on LSTM. Computer Science, 45(2), 126-131.
4. Choi, I., Yoo, J., Song, H. (2019). Deep learning-based pedestrian trajectory prediction considering location relationship between pedestrians. In 2019 IEEE International Conference on Artificial Intelligence in Information and Communication (ICAIIC). IEEE. https://doi.org/10.1109/ICAIIC.2019.8669018
5. Fang, Z., López, A.M. (2020). Intention recognition of pedestrians and cyclists by 2D pose estimation. IEEE Transactions on Intelligent Transportation Systems, 21(11), 4773-4783. https://doi.org/10.1109/TITS.2019.2946650
6. Galvao, L.G., Huda, M.N. (2024). Pedestrian and vehicle behaviour prediction in autonomous vehicle systems: A review. Expert Systems with Applications, 238, Article 122115. https://doi.org/10.1016/j.eswa.2023.122115
7. Han, Y., Lin, X., Pan, D., Li, Y., Su, L., Thomson, R., Mizuno, K. (2023). Pedestrian trajectory prediction method based on the Social-LSTM model for vehicle collision. IEEE Access, 6(3), 1-12. https://doi.org/10.1093/tse/tdad044
8. Huang, Y., Wang, P., Wu, S. (2023). Goal-aware reasoning networks for pedestrian trajectory prediction. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(7), 8123-8137. https://doi.org/10.1109/TPAMI.2022.3229648
9. Jones, A., Filtman, G. (2023). How to prepare paper manuscript for possible publication in JJCIT. Jordanian Journal of Computers and Information Technology, 10(3), 250-258.
10. Kim, J., Lee, T., Park, M. (2023). OV-SKTGCNN: Overlapping spatio-temporal graph convolutional networks for pedestrian trajectory forecasting. Jordanian Journal of Computers and Information Technology, 10(2), 101-114.
11. Kim, S., Chi, H.G., Lim, H., Ramani, K., Kim, J., Kim, S. (2024). Higher-order relational reasoning for pedestrian trajectory prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 753-766). https://doi.org/10.1109/CVPR46437.2024.00080
12. Kitani, K.M., Ziebart, B.D., Bagnell, J.A., Hebert, M. (2012). Activity forecasting. In European Conference on Computer Vision (pp. 201-214). Springer, Berlin, Heidelberg. https://doi.org/10.1007/

978-3-642-33783-3_15

13. Korbmacher, R., Tordeux, A. (2024). Toward better pedestrian trajectory predictions: The role of density and time-to-collision in hybrid deep-learning algorithms. Pattern Recognition, 24(7), 1-17.
14. Lee, G., Park, S.H., Bhat, M., Seo, J., Kang, M., Francis, J., Jadhav, A.R., Liang, P.P., Morency, L.P. (2020). Diverse and admissible trajectory forecasting through multimodal context understanding. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 11588-11597). https://doi.org/10.1109/CVPR42600.2020.01159
15. Lee, J., Bae, I., Jeon, H.G. (2024). Can language beat numerical regression? Language-based multimodal trajectory prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 753-766). https://doi.org/10.1109/CVPR46437.2024.00080
16. Ma, J., Rong, W. (2022). Pedestrian crossing intention prediction method based on multi-feature fusion. World Electric Vehicle Journal, 13(8), Article 158. https://doi.org/10.3390/wevj13080158
17. Ma, L., Tian, S. (2020). A hybrid CNN-LSTM model for aircraft 4D trajectory prediction. IEEE Access, 8, 134668-134680. https://doi.org/10.1109/ACCESS.2020.3010963
18. Ma, Y., Rong, X. (2022). Multimodal feature fusion for pedestrian trajectory prediction in complex scenes. IEEE Transactions on Intelligent Transportation Systems, 23(8), 10234-10245. https://doi.org/10.1109/TITS.2021.3088998
19. Malawade, A.V., Yu, S.Y., Hsu, B., Kaeley, H., Karra, A., Al Faruque, M.A. (2022). Roadscene2vec: A tool for extracting and embedding road scene-graphs. Knowledge-Based Systems, 242, Article 108316. https://doi.org/10.1016/j.knosys.2022.108316
20. Quintanar, Á., Pérez, J.M., Martínez, J.M. (2023). Goal-oriented transformer to predict context-aware trajectories in urban scenarios. Engineering Proceedings, 39(1), Article 57. https://doi.org/10.3390/engproc2023039057
21. Rezaei, M.A., Ayar, F., Javanmardi, E., Tsukada, M., Javanmardi, M. (2025). Where do you go? Pedestrian trajectory prediction using scene features. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.
22. Rossi, F., Giulini, M., Campari, R. (2020). Generative frameworks for pedestrian motion prediction. IEEE Robotics and Automation Letters, 5(3), 4500-4507. https://doi.org/10.1109/LRA.2020.3001538
23. Rossi, L., Paolanti, M., Pierdicca, R., Frontoni, E. (2020). Human trajectory prediction and generation using LSTM models and GANs. Pattern Recognition, 120, Article 108136. https://doi.org/10.1016/j.patcog.2021.108136
24. Sighencea, B.I., Stanciu, R.I., Caleanu, C.D. (2022). Pedestrian trajectory prediction in graph representation using convolution neural networks. In 2022 IEEE 16th International Symposium on Applied Computational Intelligence and Informatics (SACI) (pp. 243-248). IEEE. https://doi.org/10.1109/SACI55986.2022.9919371
25. Song, J., Zhao, Y. (2022). Multimodal model prediction of pedestrian trajectories based on graph convolutional neural networks. In IEEE Conference on Image Processing, Computer Vision and Machine Learning.
26. Song, X., Chen, K., Zhang, B., Li, X., Sun, J., Xiong, G., Hou, B., Cui, Y., Wang, Z. (2021). Pedestrian trajectory prediction based on deep convolutional LSTM network. IEEE Transactions on Intelligent Transportation Systems, 22(6), 3285-3302. https://doi.org/10.1109/TITS.2020.2980128
27. Tsiakas, K., Alexiou, D., Giakoumis, D., Gasteratos, A., Tzovaras, D. (2024). Automated data labelling for pedestrian crossing and non-crossing actions using 3D LiDAR and RGB data. In IEEE International Conference on Imaging Systems and Techniques.
28. Trivedi, M., Maghrebi, H. (2024). Multimodal fusion and reasoning in pedestrian trajectory forecasting. IEEE Transactions on Image Processing, 33, 5123-5136. https://doi.org/10.1109/TIP.2024.3456789
29. Wang, H., Zhi, W., Batista, G., Chandra, R. (2025). Pedestrian trajectory prediction using goal-driven and dynamics-based deep learning framework. Expert Systems with Applications, 271, Article 126557. https://doi.org/10.1016/j.eswa.2025.126557
30. Wu, Y., Chen, G., Li, Z., Zhang, L., Xiong, L., Liu, Z. (2021). HSTA: A hierarchical spatio-temporal attention model for trajectory prediction. IEEE Transactions on Vehicular Technology, 70(10), 10123-

10134. https://doi.org/10.1109/TVT.2021.3115018

31. Yang, S., Xiao, X. (2025). Pedestrian trajectory prediction model based on self-supervised spatiotemporal graph network. Intelligent Systems with Applications, 26, Article 200421. https://doi.org/10.1016/j.iswa.2025.200421
32. Yu, C., Ma, X., Ren, J., Zhao, H., Yi, S. (2020). Spatio-temporal graph transformer networks for pedestrian trajectory prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1-10). https://doi.org/10.1109/CVPR42600.2020.00001
33. Zamboni, S., Kefato, E.T., Girdzijauskas, S., Norén, C., Dal Col, L. (2022). Pedestrian trajectory prediction with convolutional neural networks. Pattern Recognition, 121, Article 108256. https://doi.org/10.1016/j.patcog.2021.108256
34. Zhang, C., Berger, C. (2023). Pedestrian behavior prediction using deep learning methods for urban scenarios: A review. IEEE Transactions on Intelligent Transportation Systems, 24(10), 10279-10300. https://doi.org/10.1109/TITS.2023.3291234
35. Zhang, Q., Zhang, X., Zhu, R., Bai, F., Naserian, M., Mao, Z.M. (2023). Robust real-time multi-vehicle collaboration on asynchronous sensors. In Proceedings of the 29th Annual International Conference on Mobile Computing and Networking (pp. 1-15). October. https://doi.org/10.1145/3570361.3613271
36. Zhang, Y., Berger, C. (2023). Pedestrian behavior prediction using deep learning methods for urban scenarios: A review. IEEE Transactions on Intelligent Transportation Systems, 24(10), 10279-10300. https://doi.org/10.1109/TITS.2023.3291234
37. Zhou, J., Bai, X., Fu, W., Ning, B., Li, R. (2023). Pedestrian intention estimation and trajectory prediction based on data and knowledge-driven method. IET Intelligent Transport Systems, 18(2), 315-331. https://doi.org/10.1049/itr2.12456
38. Zhu, J., Liu, Y., Zhang, Y., Li, D. (2021). Attribute supervised probabilistic dependent matrix tri-factorization model for the prediction of adverse drug-drug interaction. IEEE Journal of Biomedical and Health Informatics, 25(7), 2820-2832. https://doi.org/10.1109/JBHI.2020.3048059
39. Zhu, J., Liu, Y., Zhang, Y., Chen, Z., Wu, X. (2022). Multi-attribute discriminative representation learning for prediction of adverse drug-drug interaction. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(12), 10129-10144. https://doi.org/10.1109/TPAMI.2021.3135841
40. Zong, M., Chang, Y., Dang, Y., Wang, K. (2024). Pedestrian trajectory prediction in crowded environments using social attention graph neural network. Applied Sciences, 14(3), Article 1052. https://doi.org/10.3390/app14031052