

DIGITAL TWIN-BASED DRIVING SIMULATION FOR AUTONOMOUS DRIVING IN
DEVELOPING REGIONS*1Vinza Kiani, 2Muhammad Munwar Iqbal, 3Muhammad Farooq, 4Qamas Gul, 5Fareed Ahmad**1,2,4Department of Computer Science, University of Engineering and Technology, Taxila, Pakistan**3Department of Information Technology, University of the Punjab, Lahore**5Quality Operations Laboratory, Institute of Microbiology, University of Veterinary and Animal Sciences,
Lahore, Pakistan**vinzakiani044@gmail.com, munwariq@gmail.com, mfarooq@pucit.edu.pk, qamas.gul@uettaxila.edu.pk,
fareed.ahmad@uvas.edu.pk*DOI: <https://doi.org/10.5281/zenodo.21078349>**Keywords**

Digital Twin, Autonomous Driving, Developing Regions, Long-Short Term Memory, (LSTM), SUMO,

Article History*Received on 22 March 2026**Accepted on 11 June 2026**Published on 27 June 2026*

Copyright

Corresponding Author: *

Abstract

The current developments in the field of trajectory prediction have led to the realization that precision in motion forecasting is important in complex and non-structured traffic setups. This paper describes a Transformer-based architecture, which builds spatial-temporal motion patterns directly on GPS trajectories with the help of self-attention and trajectory-relative feature encoding, making it capable of modeling longer-range dependencies in recent patterns than recurrent networks. The multi-head attention mechanism is an effective way to improve fine-grained motion understanding and, at the same time, is able to run in real time. The trained model is deployed into a SUMO-based digital twin, making it possible to keep prediction simulation in sync at 0.1 second intervals. GeoLife experimental results indicate good performance, with a validation loss of 0.000443, an RMSE of 0.021, an ADE of 0.025, and 97.27% accuracy at a threshold of 0.05 in centimeter accuracy and higher than LSTM based baselines.

1. INTRODUCTION

Autonomous driving is among the most significant research fields in recent years. The current literature and systems focus on structured traffic situations, like those found in developed nations, where the roads have well-defined lane markings, traffic signs, and foreseeable driving patterns [1]. In developing countries, however, traffic tends to be unorganized, a mix of different types of vehicles, with no regular movement pattern, and insufficient road infrastructure. They reduce the ability of the current autonomous driving systems to perform in such real-world scenarios. However, the situation is considerably different in the traffic networks of most developing countries. The roads are usually unmarked with their lanes, the signage is usually incomplete and not standardized, intersections are often not controlled, and the movement of pedestrians or non-motorized vehicles is extremely unpredictable [2]. Besides, the traffic ecosystem is also heterogeneous; it includes cars, motorbikes, rickshaws, buses, bicycles, and carts pulled by animals, as well as pedestrians that share the same unorganized space. These conditions are characterized by unpredictable flow patterns, excessive concentration, and random maneuvers that are not provided in the assumptions of the

majority of current autonomous driving models [3]. Such systems that have been trained on structured data are therefore prone to poor performance and poor generalization, and their decision-making is also unstable in the real-world environment where no structure is provided to them. This gap brings to our attention the necessity to come up with autonomous driving frameworks that would specifically address the nuances of traffic in developing regions [4, 5].

Actual GPS trajectory data are employed in the real world to get real vehicle movement behavior. The model is able to implicitly learn driving patterns, changes in speed, and even trends in motion, which are hard to express in rules designed manually by designers. This solution is particularly appropriate to traffic situations in developing countries when the driving behavior can be considered heterogeneous and irregular in most cases [4].

The trajectory prediction problem is given as a supervised sequence learning problem in which previous vehicle locations are the input and predicted future location is the output. The formulation has greatly been used in earlier literature on trajectory prediction [5].

The existing traffic simulation tools mainly depend on a set of defined behavioral guidelines

and do not extend to smoothly integrate data-friendly prediction models, which decrease an aptitude to mimic real-world driving behavior. This limitation underscores the fact that a single, data-driven digital twin system is necessary, which can combine the actual GPS trajectory data of the real world, machine-learning-based trajectory prediction, and real-time traffic simulation to enhance the realism and efficiency of autonomous driving studies in developed region countries. The proposed framework has several practical applications in Intelligent Transportation Systems, in Autonomous Driving and in the development of smart cities.

LITERATURE REVIEW

Alahi et al. [6] introduced the Social LSTM model, which uses recurrent neural networks with a social pooling mechanism to capture interactions between agents. It was an important step in the modeling of pedestrian behavior in a crowded setting, but it has problems with long-range dependencies. Song et al. [7] suggested a human mobility prediction neural network based on spatial-temporal. The model learns movement patterns using large-scale GPS information. It shows good results in urban mobility analysis. But it needs big datasets to train. Liu et al. [8] formulated a deep mobility model, which represents periodic movement patterns. The

model recognizes the repeatable human mobility behaviors. It enhances the accuracy of prediction of habitual motions. Nevertheless, it is not effective in non-regular situations. Wang et al. [9] suggested hybrid CNN LSTM model of trajectory prediction. The model is an integration of spatial feature extraction and temporal modeling. It achieves good performance on GPS datasets. Nonetheless, it has the disadvantage of being computationally expensive. A graph based mobility prediction model was presented by Feng et al. [10]. The model is a graph-based model that represents spatial relationships between sites. It enhances the predictive success with big mobility data. Nonetheless, it involves a complicated graph construction.

Deo et al. [11] introduced the convolutional social pooling methodology, which uses convolutional layers with LSTMs to capture the spatial and time dependence. The model also has increased computational complexity, but it increases the accuracy of prediction. Gupta et al. [12] created Social GAN, a generative adversarial model capable of generating several plausible future paths, which deals with the multimodality of prediction. Nevertheless, GAN-based methods are computationally inefficient and hard to train. Yuan et al. [13] proposed a hierarchical trajectory prediction model, which decouples world route

planning and local motion prediction. Such a stratified treatment enables the improved modelling of both long-term and short-term dependencies. Their model enhances the accuracy of predictions in complicated roads. It however needs more computational power to do hierarchical processing. Hong et al. [14] presented a lane-aware trajectory prediction framework that uses information on road geometry and lanes. The model enhances the accuracy of predictions since it matches the trajectories to the lane constraints. It is able to work under highway conditions. Nevertheless, it is dependent on lane information and therefore it is not effective in unstructured traffic [16].

Thus, there is an evident research gap in the development of an efficient, scalable, and integrated framework integrating further trajectory predictions with digital twin simulation, tailored to the dynamic and unpredictable traffic conditions of developing areas, and this study will fill this gap.

RESEARCH METHODOLOGY

The primary data source was the GeoLife GPS Trajectories dataset [16] that comprises millions of GPS trajectory records of pedestrians and vehicles (in the real-life urban setting). Despite the fact that the dataset was initially mapped in organized Chinese cities, it contains fine-grained movement traces that are rich and good enough to trace various mobility patterns. The data was further refined and enhanced with variation in speed, direction and time interval between trajectory points in order to provide a more realistic representation of unstructured traffic patterns in developing areas. This augmentation mimics the abnormal driving patterns that include changing lanes abruptly, following too closely, as well as mixed engagement between vehicles. In this way, the dataset was transformed to simulate chaotic and heterogeneous traffic states, which is why it is appropriate to train the Transformer model to simulate unstructured traffic and integrate digital twins. Sample Trajectory Data from GeoLife GPS Dataset is described in Table 1 and 2.

Table 3.1: Samples of GeoLife GPS Dataset

Traj_id	Time	Latitude	Longitude	Altitude	Date	Time
0	0	39.984702	116.318417	492.0	08-10-23	02:53:04
0	1	39.984683	116.318450	492.0	08-10-23	02:53:10
0	2	39.984686	116.318417	492.0	08-10-23	02:53:15

0	3	39.984688	116.318385	492.0	08-10-23	02:53:20
0	4	39.984655	116.318263	492.0	08-10-23	02:53:25

Table 2: Dataset Statistics

Sr. No	Description	Value
01	Total training samples	12,264,374
02	Total validation samples	2,628,080
03	Total test samples	2,628,081
04	Sequence length	10
05	Number of features	4 (x, y, speed, heading)
06	Output dimensions	(x, y)

The processed data were split into training (70%), validation (15%), and test (15%) sets to enable model development and performance testing. The proposed framework is based on a closed-loop digital twin system that allows the simulation environment and prediction model to continuously share information with each other as shown in Figure 1. A closed-loop digital twin involves continuously collecting, analyzing, predicting and returning system states to the virtual environment to keep the system behavior synchronized between the real and virtual worlds. The SUMO traffic simulator is used as the virtual traffic environment, and the Transformer model is used as the predictive intelligence component in this research. SUMO provides the Vehicle

states, such as position, speed, heading, etc., for each step of the simulation via the Traffic Control Interface (TraCI). The Transformer model takes these observations and creates future trajectory predictions. These predicted trajectories are then displayed and analyzed in the simulation environment and a feedback loop is established between the simulator and the prediction model. The integration allows the digital twin to observe the current traffic situation but also predict future traffic flows, which can be used for real-time traffic analysis and autonomous driving applications. It is possible through this integration to have the predictive model and the simulation engine interacting in a bidirectional fashion:

- 1. Input Phase:** GPS sequential inputs or vehicle states are generated as simulation inputs to the DT.
- 2. Prediction Phase:** The Transformer model is a prediction of the subsequent spatial positions and velocities.
- 3. Visualization Phase:** SUMO displays the predicted trajectories and renders the digital environment with them as well.

- 4. Feedback Phase:** The system will get the results of a simulation (e.g., collisions, density, speed) and will reverse them back to the model to optimize or approve it.

As Figure 3.4 depicts, the conceptual design of the Digital Twin integration reflects a closed-loop interaction among data input, Transformer prediction, and visualization based on SUMO.

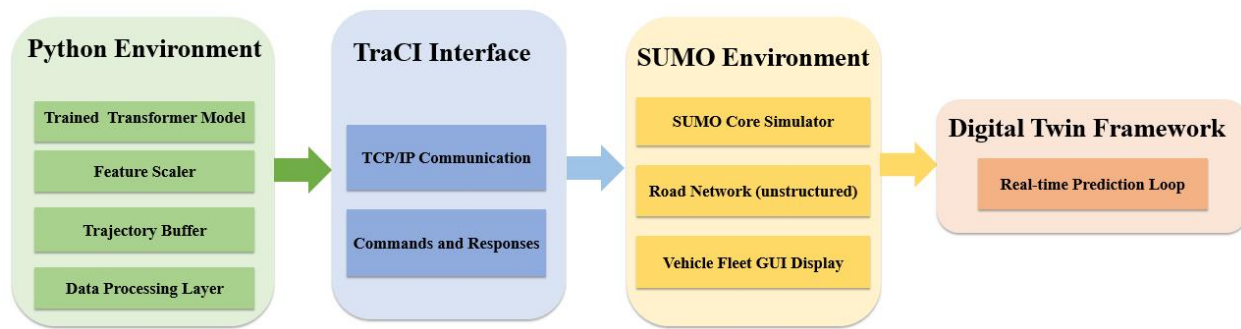


Figure 1: Real-time Digital Twin Integration Framework



The general workflow has four significant elements:

Data Interface Layer: In charge of collecting actual or recorded GPS data in CSV format in the GeoLife dataset. The interface takes the data and streams it to the model at a fixed interval to mimic a live data flow.

Prediction Engine: The model takes as input the trajectories and gives forecasted latitude-longitude returns. The predictions are provided as XML or CSV files that can be imported into SUMO routes.

Simulation Layer (SUMO): SUMO is based on these predicted movements to model the behavior of a multi-vehicle system. The speed, acceleration, and position of each vehicle are updated at discrete simulation time steps. Different parameters of driver behavior in the model of unstructured or mixed traffic conditions include aggressiveness, following distance, and so on, but not all possible combinations are represented [6].

Feedback and Synchronization Layer: An output loop is set up between SUMO and the Python environment (e.g., travel time, average speed,

collisions). This allows the system to verify simulation results against predicted values, dynamically adjust system parameters, and enhance the relevance of the digital twin model.

The real-time synchronization process consists of a series of steps based on the digital twins, which are based on previous digital twin-based traffic control frameworks. As shown in

3.4.1 SUMO GUI Visualization Framework

A tailored framework for the SUMO GUI was developed to visualize vehicle motion with the

Transformer in real time. The GUI shows the early unstructured traffic configuration, online trajectory forecasting and live visualization of digital twin. TraCI facilitates the synchronization of updates and ensures that the state of every vehicle is updated at every step of the simulation. This interactive visualization enables users to pause, replay, and dynamically analyze traffic flow patterns and prediction accuracy.

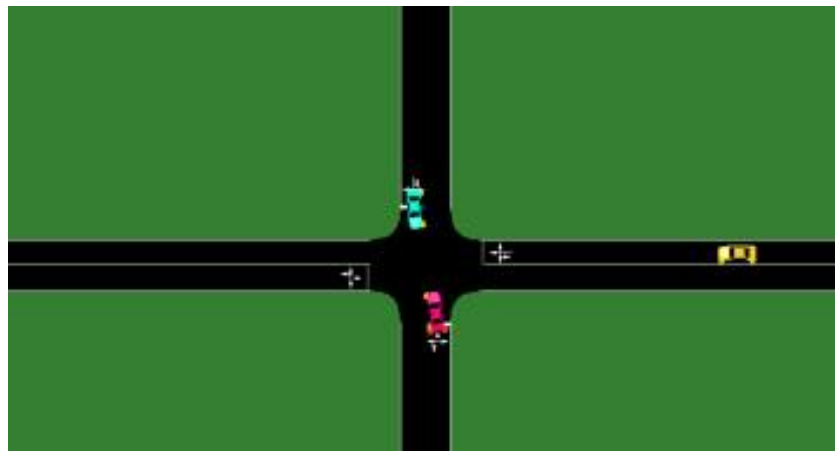


Figure 2: SUMO GUI Visualization Vehicle Predictions

Figure 2 shows the real-time connection between the SUMO simulation environment and the Transformer-based prediction model using TraCI. The figure showcases the real-time position of several vehicles in an unstructured traffic environment, and their predicted trajectories. The state of individual vehicles is updated in real time, facilitating smooth interaction between the real

(simulation) and digital twins. It also shows how the system enables users to interactively explore traffic dynamics, observe the prediction model's behavior, and assess its performance in real time. This showcases the effectiveness of the proposed approach in offering an interactive and scalable digital twin environment for research in autonomous driving.

RESULTS AND DISCUSSION

The proposed digital twin-based framework of trajectory prediction implementation is described. Real-world GPS trajectory data are used to conduct experimental studies to quantify the behavior of vehicles in unstructured traffic. The data is first preprocessed and then applied to train a deep learning model based on a Transformer to predict the future vehicle paths. The trained model is next connected to the SUMO traffic simulator, which creates a digital twin environment where the prediction can communicate in real-time with the simulation. The implementation will enhance the accuracy of the trajectory prediction and make the traffic simulation more realistic in developing areas.

4.1 Experimental setup and Evaluation

The fundamental step in this study is model training and evaluation, during which the provided Transformer model is trained on spatial-temporal associations between sequence trajectories and then tested for predictive quality. The aim is to tune the model parameters so that

the predicted vehicle positions are highly accurate in real-world, uncoordinated traffic.

1.1 Experimental Setup

All testing was done in Python 3.10 and PyTorch 2.1 on a CPU-based workstation with an Intel Core i7 (2.6 GHz) and 16 GB of RAM. Although the model does not use GPU acceleration, it is optimized for efficient computation due to its small number of trainable parameters (around 108,000) and mini-batch processing. The small-footprint design enables training and inference on a regular computer without specialized hardware, making the proposed architecture suitable for low-resource settings and emerging markets.

The Transformer was implemented in PyTorch in a compact configuration to balance computational efficiency and prediction accuracy. The hyperparameters used in this study are summarized in Table 3 in which lightweight design enables the model to remain computationally efficient while maintaining high predictive accuracy, achieving an optimal balance between real time performance and modeling precision.

Table 3: Transformer Model Parameters

SR.No	Parameter	Value
01	Input Dimension	4 features
02	Hidden Dimension	64
03	Attention Heads	4

04	Encoder Layers	2
05	Feed-Forward Dimension	256
06	Dropout Rate	0.1
07	Output Dimension	2 (x, y)

4.2 Quantitative Results and Analysis

In this section, the overall training behavior of the Transformer model is described, including convergence patterns and generalization properties observed during training. In this model, we define the training behavior of the Transformer-based trajectory prediction model in general, as well as its convergence behavior and overall generalization performance. Even though the model was set to train for 20 epochs, it achieved its best validation loss at epoch 8 as shown in Figure 3. Beyond this, no significant

improvement was observed, and further training would not lead to improvements in time to compute or performance. Thus, training was terminated at the optimal epoch when it had been attained as shown in Table 3. It can be seen that the loss curves converge quickly and stabilize, and the model learns trajectory patterns within the initial few epochs. The fact that the validation loss is always low and smooth further supports the model's strong generalization, indicating that it extracted the underlying spatio-temporal structure of the data without overfitting.

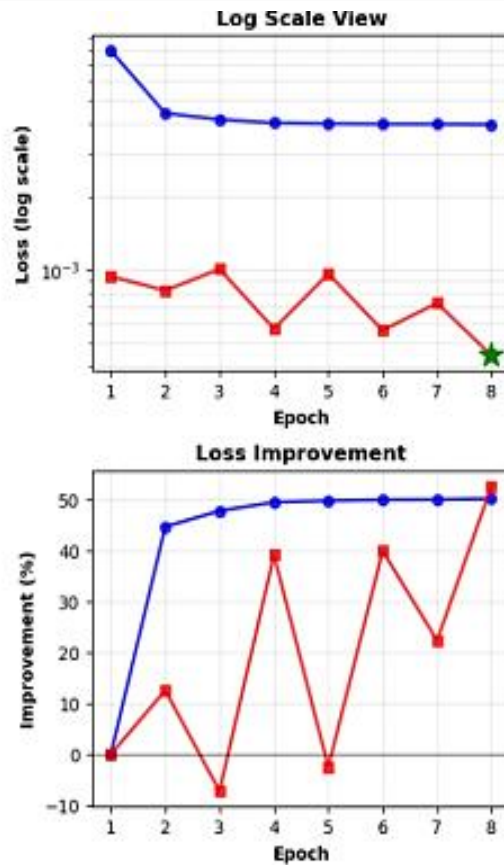


Figure 3: Training and Validation improvements and logscale view

Table 3: Training Performance Metrics

Metric	Initial Value	Final Value	Improvement (%)
Training Loss	0.007964	0.003959	50.3%
Validation Loss	0.000936	0.000443	52.7%

The gradual decrease is indicative of stable dynamics in optimization and an efficient relationship between learning rate, batch size, and model capacity. The gap between the final training and the validation losses was used to measure generalization as Generalization gap = -

0.003516. The minor difference suggests that during training, this model acquired consistent, generalizable trajectory dynamics rather than memorizing the training data. On the whole, the training phase demonstrates that the Transformer architecture is very effective for predicting short-

term vehicle trajectories, achieving fast convergence and high generalization despite its small size. Figure 4 proves that validation loss was

lower in all epochs in comparison to training loss, which is a peculiar but positive feature that shows the strength of the model.

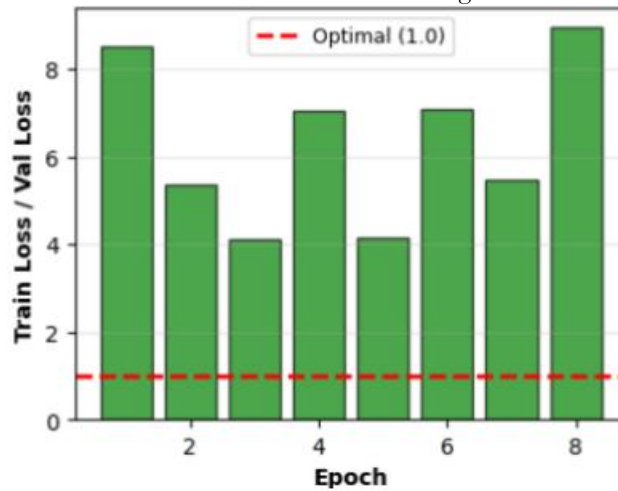


Figure 4: Depicts the Loss Ratio Generalization

This scatter plot which shown in Figure 5 compares the X- and Y-direction prediction errors for each test point. The majority of the points are very close to the origin (0,0), indicating that the errors in predicting both X and Y are very small.

The dispersion is also evenly distributed, indicating that the model does not favor a single axis and predicts horizontal and vertical movement equally well.

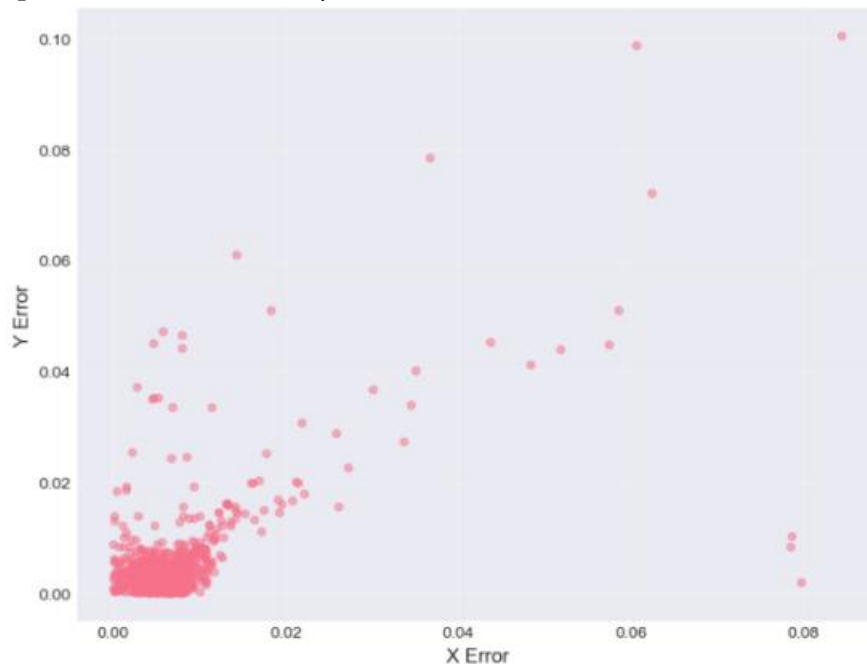


Figure 5: X vs. Y Dimensional Error Correlation Plot

4.3.1 Accuracy Analysis

The model was tested in various spatial thresholds as illustrated in table 4.

Table 4: Accuracy at Distance Thresholds

Threshold (units)	Accuracy (%)	Interpretation
< 0.01	82.85%	High precision
< 0.02	93.50%	Very accurate
< 0.05	97.27%	Highly reliable
< 0.10	99.66%	Excellent coverage
< 0.20	99.95%	Near-perfect

This precision is appropriate in digital twin simulation, microscopic traffic simulation, and autonomous driving. The model predicts the positions of vehicles at varying distance thresholds in the future. The model’s high accuracy at small thresholds (0.01 to 0.05 units) indicates it is reliable for making very precise short-term predictions, which is crucial for tasks such as real-time digital twins, SUMO microscopic traffic simulation, decision-making, and collision avoidance in autonomous driving.

CONCLUSION AND FUTURE WORK

The proposed model in combination with traffic simulation, can yield better prediction accuracy and simulation realism in unstructured traffic conditions. This paper proposed an effective In this study, we proposed a novel and efficient Transformer-based approach for short-term vehicle trajectory prediction, embedded in a digital twin simulation environment using SUMO. The developed framework effectively integrates machine learning with simulation tools, facilitating a real-time interaction between the prediction model and the digital twin using TraCI.

Our performance suggests that it outperforms conventional methods like LSTM, Social-LSTM and graph-based models, which may fail to capture long-term and spatial dependencies. Additionally, the Transformer model captures spatio-temporal interactions in trajectory data without recurrence, enabling enhanced performance and efficiency. These future research directions will help build a more reliable, scalable and smart digital twin-based system for next-generation autonomous driving and smart transportation systems. Extend the framework to multi-step trajectory prediction to enable long-

term forecasting and autonomous driving decision-making.

REFERENCES

- [1]. Xu, R., Li, J., Zhang, S., et al. Interactive trajectory prediction for autonomous driving based on Transformer. *Mech Sci.* 2025;16:87-97.
- [2]. Rui Xu, Jun Li, Shiyi Zhang, Lei Li, Hulin Li, Guiying Ren, and Xinglong Tang. Interactive trajectory prediction for autonomous driving based on transformer. *Modeling and Simulation*, 16:87-102, 2025.
- [3]. Usman Ahmad, Mu Han, Alireza Jolfaei, Sohail Jabbar, Muhammad Ibrar, Aiman Erbad, Houbing Herbert Song, and Yazeed Alkhrijah. A comprehensive survey and tutorial on smart vehicles: Emerging technologies, security issues, and solutions using machine learning. *IEEE Transactions on Intelligent Transportation Systems*, 25(11):15314-15341, 2024.
- [4]. Pranav Chaudhari, Raghavendra Achar, and Sanjay Singh. Enhancing lane recognition in autonomous vehicles using cross-layer refinement network. *IEEE Access*, 12:117650-117664, 2024.
- [5]. Miao Xu, Zhi Liu, Bingyi Wang, and Shengyan Li. A survey of autonomous driving trajectory prediction: Methodologies, challenges, and future prospects. *Machines*, 13(9), 2025.
- [6]. Vibha Bharilya and Neetesh Kumar. Machine learning for autonomous vehicle's trajectory prediction: A comprehensive survey, challenges, and future research directions. *Vehicular Communications*, 46:100733, 2024.
- Chengdong Wang, Jianming Wang, Wenbo Gao, and Lei Guo. Siat: Pedestrian trajectory prediction via social interaction-aware transformer: C. wang et al. *Complex & Intelligent Systems*, 11(8):335, 2025.
- Song, X., et al. Deep mobility prediction using spatial-temporal neural networks. *Proceedings of AAAI Conference on Artificial Intelligence*, 30(1): 305-312, 2016.
- Liu, Y., et al. Predicting human mobility with deep learning. *IEEE Transactions on Knowledge and Data Engineering*, 30(10): 1920-1933, 2018.
- Wang, H., et al. CNN-LSTM hybrid model for trajectory prediction. *IEEE Access*, 7: 152243-152252, 2019.
- Feng, J., et al. Graph-based mobility prediction using spatial dependencies. *Proceedings of IJCAI*, pages 389-395, 2018.
- Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., and Savarese, S. Social LSTM: Human trajectory prediction in crowded spaces. *Proceedings of CVPR*, pages 961-971, 2016.
- Yuan, Y., et al. Hierarchical trajectory prediction model. *IEEE Transactions on Intelligent Transportation Systems*, 21(3): 1152-1163, 2020.
- Gupta, A., Johnson, J., Fei-Fei, L., Savarese, S., and Alahi, A. Social GAN: Socially acceptable trajectories with generative adversarial networks. *Proceedings of CVPR*, pages 2255-2264, 2018.
- Hong, J., et al. Lane-aware trajectory prediction in structured roads. *IEEE Intelligent Vehicles Symposium*, pages 1-6, 2019.
- Yifan Zhao, Lin Xu, and Chen Wei. Gps trajectory-based traffic flow prediction. *IEEE Access*, 10:10234-10245, 2022.