

Reinforcement Learning Controlled Variable Speed Limits in Urban Expressway Mixed Traffic

Alfayez Ahmad

Department of Computer Science & Engineering, Faculty of Engineering & Information Technology, Integral University, Lucknow, Uttar Pradesh, India
alfayezahmad@outlook.com

Mohd Sadat

Department of Civil Engineering, Faculty of Engineering & IT, Integral University, Lucknow, Uttar Pradesh, India
sadat.kh@gmail.com

Syed Aqeel Ahmad

Department of Civil Engineering, Faculty of Engineering & IT, Integral University, Lucknow, Uttar Pradesh, India
syedaqeel@iul.ac.in

Shrish Bajpai

Department of Electronics & Communication Engineering, Faculty of Engineering & IT, Integral University, Lucknow, Uttar Pradesh, India
shrishbajpai@gmail.com

Mehmet Ali Silgu

Department of Civil Engineering
Bartın University
Bartın, Turkey
masilgu@bartin.edu.tr

Abstract— Increase in Rapid urbanization has intensified traffic congestion, delays, and emissions on urban expressways, revealing the limitations of conventional control strategies. This study aims to develop a Reinforcement Learning (RL) based Variable Speed Limit (VSL) control framework using a Deep Q-Network (DQN) implemented in the Simulation of Urban Mobility (SUMO) environment to enhance traffic efficiency at an expressway merging section. Traffic data used was collected using a video camera recorder and radar speed gun, with vehicle trajectories extracted through the Traffic Data Extractor developed by IIT Bombay. The model implemented was calibrated and validated using field observations from two sites—one representing uninterrupted flow and the other in an on-ramp merging area. The simulation compared three configurations: a baseline case without control, a conventional rule-based VSL controller, and the proposed DQN-based VSL approach and the findings reveal that the DQN agent achieved a 18.8% reduction in total travel time compared to the baseline, while the rule-based VSL controller worsened performance by 21.7%. The learning-based controller effectively mitigated congestion, reduced shockwave formation, maintained higher average speeds, and improved travel time reliability under dynamic and stochastic traffic conditions demonstrating that a reinforcement learning-driven VSL system can significantly enhance both traffic flow efficiency and user-level reliability, outperforming traditional heuristic control strategies on urban expressways.

Keywords— Variable Speed Limit (VSL), Deep Q-Network (DQN), Reinforcement Learning (RL), Simulation of Urban Mobility (SUMO), Urban Expressway

I. INTRODUCTION AND LITERATURE REVIEW

The transportation system is a vital urban infrastructure system among which integrated traffic infrastructure, operational equipment, and a strategic management protocol compassing both real-time control and long-term planning traffic flow, and thus, dynamic travel demands efficiently regulate traffic flow. It is a determinant of economic vitality and social well-being as well. However, the rapid global

urbanization is a precipitated factor for a surge in vehicle ownership and thus chronic congestion is resulting in more accidents. The increasing number of problems with the mobility sector not only add inconvenience to the daily commuters but also bring the core functions of the urban environment to a standstill, thus presenting the urgent need for smart traffic managements. This paradigm shifts in computing technology has resulted in the emergence of Intelligent Transportation Systems (ITS) that traditionally span in multimodal domains such as air, rail, and marine transport [1].

The main composition of the ITS is built on four main components that arrange interdependent and build a closed-loop system: perception, prediction, simulation, and decision-making [2,3]. The very beginning of the process is traffic perception which brings large-scale acquisition from a heterogeneous suite of sensors such as sensors, cameras, and monitoring devices as well as data fusion that leads to a complete and updated traffic dynamics picture. The perceptual output that is perceptual output directly makes traffic prediction models possible, this is the results of advanced computational techniques that give an answer to what the future traffic demands will be and what the congestion states will be. Simultaneously, the high-fidelity digital twins of the transport network introduced the traffic simulation as it will modeling complex traffic patterns, with complex traffic conditions to them. The strategies mixing for application the measures within the scenario of the model are basically the targeted output approach of the actualization that gives the insight about what to do next in the terms of controlled traffic like adaptive signal timing or dynamic routing which in result lead to optimized network operation. The cyclical decision making is the typical feature of this structure; the implemented decisions will create different traffic data that will be perceived and the effect will be assessed together with the data used for further development. The application of this closed-loop cycle is with most of the ITS, which is the integrated, closed-loop cycle for their

highly efficient management, control, and optimization of modern transport networks.

Intelligent Transportation Systems (ITS) are a set of advanced integrated technologies and applications with the main purpose of improving safety, network efficiency, and sustainability in transportation networks. These systems can be subdivided into core functional domains: real-time monitoring and traffic adaptive control with the help of Advanced Traffic Management Systems (ATMS); Advanced Traveler Information Systems (ATIS) with real-time data that the travelers make use of for good decision; Advanced Public Transportation Systems (APTS) advanced on time, reliable and user-friendly public transportation systems: Commercial Vehicle Operations (CVO) systems that streamlining freight and logistics as well as Connected and Automated Vehicle (CAV) technologies that leverage vehicle-to-everything (V2X) communication and automation to improve coordination and safety. These interconnected systems, through their integration, make a cyber-physical ecosystem which synergizes data from infrastructure, vehicles, and users to optimize operational decision-making and address the complex challenges of modern mobility.

Variable Speed Limit (VSL) systems are infrastructure-to-vehicle (I2V) communication systems [4] which allow the dynamic adjustment of posted speed limits in accordance with the traffic situation, accidents, or bad weather. VSL systems are an innovative traffic management [5] approach that transcends the usual static regulatory practices by utilizing real-time data inputs--namely the traffic speed, the traffic volume, and the weather on the road-- algorithmically generate and post the suitable speed limits. This method provides a continuous feedback loop, in which data from embedded sensors and road weather information systems (RWIS) is used to assess the prevailing conditions, enable the proactive adjustment of speed, and the direct communication of drivers via overhead or roadside variable message signs (VMS) [6]. By synchronizing the vehicle speeds with traffic flow and road conditions in real-time, VSL systems serve as a beneficial incident management tool, harmonizing traffic streams to suppress shockwave formation and congestion [8]. Thus, the dynamic control not only gets efficacy in single treatment but also operates multiple behaviors: it enhances traffic operations with the help of throughput optimization, it gets better safety by minimizing speed variance and also preemptively lowers speeds for hazards when needed, it helps the environment by minimizing the stop-and-go driving that is a source of increased fuel consumption and emissions on major roads. Worthy to mention is the fact that Variable Speed Limit (VSL) system has already shown to be successful in freeway traffic management, for example, in the Netherlands, the United Kingdom, Denmark, and Australia where it has been implemented with significant reductions on traffic congestion and emissions [8]. The primary algorithms which make these systems work are mainly predicated on real-time traffic variables, including occupancy, speed, and volume, whereby critical thresholds are pre-defined and then the speeds are actuate the dynamic adjustment. Out of the mechanistic goals, the attenuation of shockwave propagation during congested conditions is mainly targeted by strategically reducing the desired speed for the upstream traffic and therefore promoting flow stability. Although complex models have been produced for successful VSL optimization strategies like the METANET, many of those

that have been deployed in the real world are built on a operational algorithms based on specific metrics; for example, I-4 in Orlando, Florida, employs average occupancy thresholds while the M25 Motorway in England and the E6 in Mölndal, Sweden use flow-based algorithms. Research as a whole suggests that these systems lead to a real decrease in the Total Travel Time (TTT) across the network and a consequent harmonization of traffic flow [9].

II. TRAFFIC SIMULATION MODEL

The region we focused on is situated in Lucknow, the capital of Uttar Pradesh, North India. Establishing the Amar Shaheed Path expressway as a four-lane corridor connecting National Highway 56 and National Highway 30 was a step forward in the direction of urban congestion mitigation. The expressway is equipped with several on-ramps that connect the main arterial roads and residential areas. Uncontrolled merging at these ramps, particularly during peak hours, often leads to congestion on the mainline. The selected study section includes one such on-ramp from an arterial road to the main expressway, as illustrated in Figure 1. The traffic model of the section was imported from OpenStreetMap and edited in the netedit module of Simulation of Urban Mobility (SUMO) software (Figure 1). The traffic data for this study was collected using a video camera recorder and a radar speed gun between 7:00 am and 12:00 pm on a weekday. Video data were processed using the Traffic Data Extractor developed by IIT Bombay, which employs computer vision and machine learning for semi-automatic vehicle detection and counting. The extracted data, output in .xml format, facilitated direct use in the SUMO microsimulation platform.



Fig 1. Section for the study and corresponding model in SUMO

We employed speed data collection radar guns working in the continuous mode for each vehicle class. INDO-HCM

(2017) describes Free-Flow Speed (FFS) as the average speed under free-flow conditions when traffic is uninterrupted flow conditions, which is generally seen when the headways are more than eight seconds. The dataset contained four vehicle categories—Two-Wheelers, Cars, Light Commercial Vehicles (LCVs), and Buses each with unique physical and speed characteristics. Two-Wheelers (1.8 m × 0.6 m) recorded an FFS of 80 km/h, a mean speed of 45.82 km/h (Standard Deviation (SD) = 8.67 km/h), and observed speed range of 31–66 km/h. LCVs and Cars (4 m × 1.6 m) had an FFS of 80 km/h, mean speed of 56.62 km/h (SD = 12.14 km/h), and limits between 34–76 km/h. Buses (7 m × 2.5 m) showed a lower FFS of 60 km/h, mean speed of 38.5 km/h (SD = 4.5 km/h), and limits between 34–43 km/h.

We calibrated several key parameters in SUMO (see Table 1) including acceleration (0.78 m/s² for 2Ws; 0.80 m/s² for other vehicles), deceleration (0.83 m/s² for 2Ws; 1.00 m/s² for Cars, LCVs, and Buses), and maximum emergency deceleration (10 m/s²) and minimum time headway (1.0 s) but the vehicle dimensions varied from compact Two-Wheelers (1.87 m × 0.64 m) to large Buses (10.10 m × 2.5 m) where driver imperfection parameter (σ) was set at 0.40 for 2Ws and 0.50 for the other vehicles, which shows behavioral variability. The maximum desired speed (v_{\max}) was set to 19 m/s for Cars, 14 m/s for 2Ws, and 12 m/s for LCVs and Buses. A minimum gap of 2.5 m was applied uniformly across vehicle categories [11].

Table 1: Key calibration parameters

Parameters	2W	Cars	LCV	Bus
Acceleration (m/s ²)	0.78	0.80	0.80	0.80
Deceleration (m/s ²)	0.83	1.00	1.00	1.00
Vehicle Length (m)	1.87	4.58	6.10	10.10
Vehicle width (m)	0.64	1.77	2.10	2.5
Emergency deceleration (m/s ²)	10	10	10	10
Minimum Time Headway (s)	1.00	1.0	1.0	1.0
Driver’s Imperfection (Sigma)	0.40	0.50	0.50	0.50
Maximum Speed, v_{limit} (m/s)	14	19	12	12
Minimum gap (m)	2.5	2.5	2.5	2.5

In general, the combination of the measured data and calibrated parameters has allowed the proper representation of mixed traffic flow dynamics on urban expressways under the two conditions, free-flow, and ramp-influenced. For the car, the following model IDM [10] was used and for default lane changing model implemented in SUMO was used.

A. Scenarios

The study assessed three simulation scenarios in total. The first scenario simulated a uncontrolled scenario in place in the merging section thus it is a baseline scenario for performance comparison. During the second, a rule-based Variable Speed Limit (VSL) control system was implemented, and in the third scenario is the application of the reinforcement learning-based control strategy, motivated by the stochastic nature of traffic flow dynamics i.e. the adaptive algorithms are expected to perform better in different traffic conditions. For the second scenario, a feedback-based control loop was used [12,13,14].

The control law of the feedback controller is expressed as: $b(k) = b(k - 1) + K_I(\hat{q}_{VSL}(k) - q_{VSL}(k))$ (1)

Here, b indicates the VSL control rate defined as the ratio of posted speed limit to free-flow speed K_I , which is the Integral gain of the controller. Additionally, $\hat{q}_{VSL}(k) - q_{VSL}(k)$ stands for a flow tracking error. The constraints for the controller are: $b(k) \in [b_{\min}, 1]$ and $\hat{q}_{VSL}(k) \in [\hat{q}_{\min}, \hat{q}_{\max}]$. [12,13]

III. RESULTS AND DISCUSSION

A. Overall Performance

One of the most important aims in the real-world traffic control is to ensure that the road system is as efficient as possible, that is to say, to serve the required traffic demand through the road network with the minimum total delay. In our scenario there were three simulation scenarios, each case processing a total of 9,272 vehicles which is the same for all of them, thus providing a fair baseline for the performance comparison and making it possible to attribute the differences in the controller only to the strategies used. Given identical throughput, the figure 8 demonstrates discrepancies in terms of network efficiency. In the Baseline scenario, drivers spent an average of 103.5 seconds to complete their trip while the Rule-based VSL controller, typically applied with the objective of improving traffic flow, increased the trip duration by 21.7% (Table 2) to 126.0 seconds indicating that it was the static heuristic logic that likely triggered the early speed reductions that hindered the mainline traffic flow more than it helped in the merging. This performance degradation relative to the no-control baseline is the same as what previously described was found in traffic control literature which said that the VSL is a very sensitive calibration and that the poorly tuned heuristic is frequently counterproductive by inducing artificial bottlenecks formation or by underutilizing downstream capacity [15,16,17]. On the other hand, the DQN agent was able to achieve an average of only 84.1 s instead of the 103.5 s Baseline thus showing an 18.8% reduction in average travel time which reflects the ability of the RL to learn a control policy better.

Table 2: Performance comparison

Control Scenario	Total Time Spent (vehicle-hours)	Improvement vs. Baseline (%)	Total Throughput (Vehicles)
Baseline (No Control)	266.57	—	9272
Rule-Based VSL	324.52	-21.7%	9272
DQN Agent VSL	216.60	+18.8%	9272

Although the overall vehicle processing was the same (9,272) in all cases, the DQN agent realized this performance by consuming significantly less time, thus proving that its efficiency gain came from better traffic regulation instead of variations in demand levels.

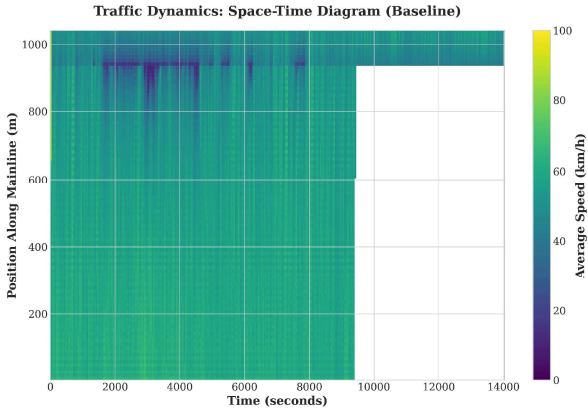


Fig 2. The vehicle space-time distribution under the baseline scenario

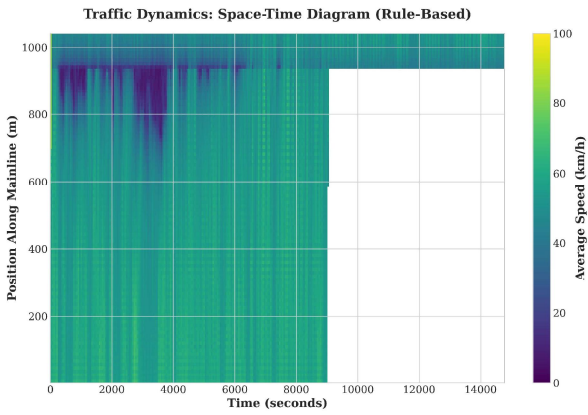


Fig 3. Rule-based space-time distribution of vehicles

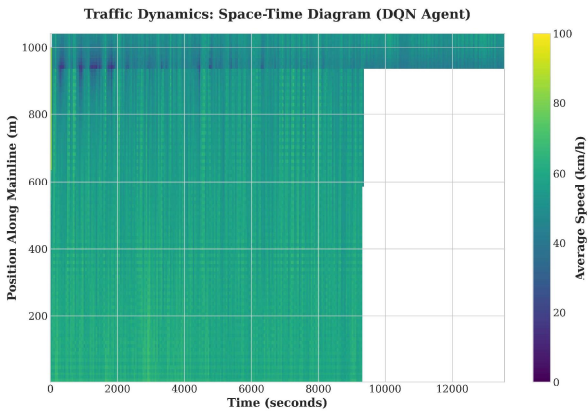


Fig 4. Space-time distribution of vehicles for DQN-Agent based scenario

B. Traffic Flow Dynamics

The outcomes of the traffic flow behavior analysis denoted that the DQN agent demonstrated an effective congestion mitigation and stop-and-go shockwaves standing in the way of the traffic to become more fluid. In the baseline scenario, stop-and-go waves were signaled by the repeated low-speed areas. Under the Rule-Based control, these disturbances intensified, reflecting the limitations of static, heuristic control logic. On the other hand, the DQN agent achieved smoother traffic flow through the learning of timely and adaptive speed adjustments and the prevention of traffic breakdown mainly that come with the high traffic.

Here, the visualization of macroscopic congestion mitigation is achieved through the analysis of the severity and persistence of stop-and-go shockwaves. The figure 2 and figure 3 display the widespread, thick bands of congestion (dark violet) starting at the bottleneck and moving to a further upstream area. The aforementioned, persistent queues result in delays in the clearance of the network. The figure 4, in turn, indicates a much lower level of congestion severity with shockwaves that are narrower and dissipate more quickly. The appearance of the "white void" or the complete network clearance state came earlier in the DQN simulation time frame than in the case of the other two scenarios, which is clear visual evidence of the idea that the agent's smoother flow led to the clearing of total traffic demand more rapidly and efficiently.

VSL systems are effective only when they utilize the bottleneck management principle properly and with a control input variable that initiates the upstream speed control to avert the flow breakdown at the merge point. The figure 5 and figure 6 illustrate this dynamic at the bottleneck located between the 900 m and 1000 m marks. The Rule-based VSL controller, in not very efficient management of the bottleneck, had a speed drop to a minimum of approximately 23 km/h and a peak occupancy of 0.13 which is much more than a baseline case with a slightly better stability and the DQN agent which displayed an even superior performance, by maintaining 42 km/h at the minimum speed and the peak occupancy approximately 0.115. That is to say, by the density-based control mechanism, DQN agent stabilized traffic flow more effectively through the critical section.

C. Vehicle-Level Performance

The DQN agent, from an individual driver's perspective, is more reliable and gives a improved user-level performance. Travel time distribution indicates that in addition to the DQN scenario reaching lower mean travel times, the very much reduced variance in travel times suggests that the drivers were subject to fewer unexpected or extreme delays, thus, it is a substantial improvement in the travel time reliability. On the other hand, the significant variability that was introduced by the Rule-based VSL controller in many vehicles experiencing extreme travel time delays was quite different (Figure 6). For them, travel time reliability (low variance in travel times) becomes a major factor just like the average speed. A system with high average speed but low reliability is undesirable. The figure 9 illustrates a serious issue in the Rule-Based method regarding reliability; the distribution of the Rule-Based method reveals high variance with a long tail that reaches beyond the point of 800 seconds which means that a lot of automobiles were subject to even extreme travel time delays, whereas the DQN agent, in addition, demonstrates a highly concentrated distribution, meaning the highly consistent performance of almost all vehicles with the same travel time. This enhanced user experience is also shown in the Figure 7, which illustrates that under Rule-Based control, vehicles spent almost 25% of their time going at speeds less than 40 km/h that were undesirable, as compared to the 8% time that DQN-controlled vehicles operated within this speed range. DQN-controlled vehicles spent proportionally more time in higher speed ranges.

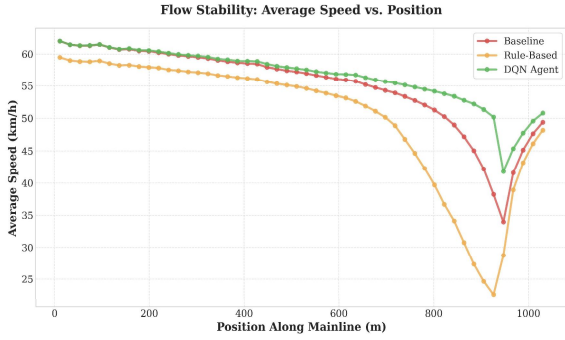


Fig 5. Flow stability

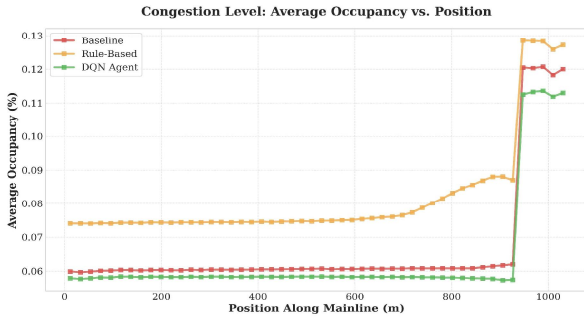


Fig 6. Congestion Level

The results obtained from the cumulative speed distribution analysis have corroborated these findings, which revealed through the graphs that the vehicles controlled by the DQN agent spent a greater proportion of time at higher speeds and took a lesser amount of time at the low speeds, stating the fact of smoother and reduced stop-and-go behavior (Figure 6).

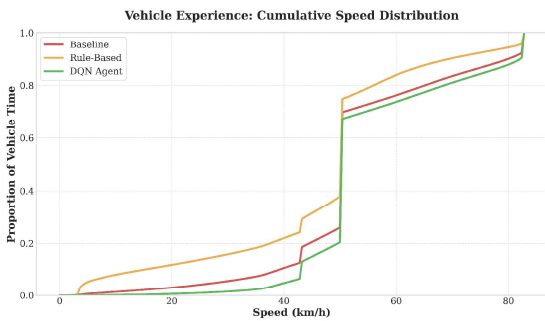


Fig 7. Vehicle-level performance

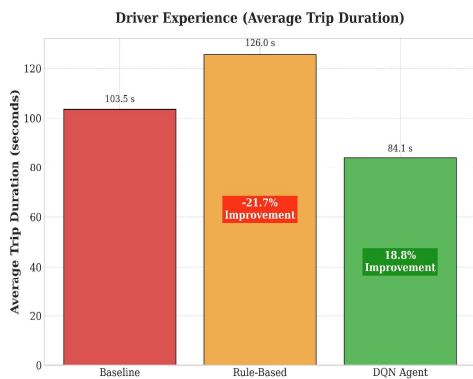


Fig 8. Driver Experience

D. Contextual Robustness

A real-world deployable controller should definitely show operational robustness, which implies the ability to perform under non-ideal and changing traffic conditions instead of being limited to just static, idealized conditions.

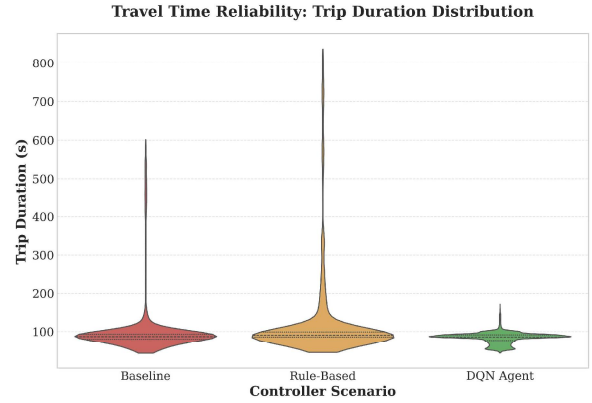


Fig 9. Illustration of reliability of algorithms

In general, given the results DQN-based control framework demonstrates a substantial improvement in both system-level efficiency and user-level experience, outperforming traditional rule-based methods and unmanaged baseline conditions under dynamic traffic environments. In future work, VSL can be integrated with ramp metering using Reinforcement Learning algorithms to further enhance the traffic flow on the urban expressway.

Table 3. Reliability of Algorithms

Control Scenario	Median Trip Duration (s)	95th Percentile Trip Duration (s)	Qualitative Reliability
Baseline (No Control)	~90	~580	Low - Large spread and significant extreme delays.
Rule-Based VSL	~95	~800	Very Low - Highest variability and most severe extreme delays.
DQN Agent VSL	~85	~140	High - Tight distribution with very few outliers or major delays.

ACKNOWLEDGMENT

The authors would like to thank the Integral University Lucknow, India for providing manuscript number: IU/R&D/2025-MCN0004194 for the present research.

REFERENCES

- [1] Mohamed Elassy, Mohammed Al-Hattab, Maen Takruri, and Sufian Badawi, "Intelligent transportation systems for sustainable smart

- cities", *Transportation Engineering*, Vol. 16, pp. 100252, 2024. doi : 10.1016/j.treng.2024.100252.
- [2] Mingxing Peng, Kehua Chen, Xusen Guo, Qiming Zhang, Hui Zhong, Meixin Zhu, Hai Yang, "Diffusion models for intelligent transportation systems: A survey." *IEEE Transactions on Intelligent Transportation Systems*, pp. 1-18, 2025. doi : 10.1109/TITS.2025.3613178.
- [3] Christian Creß, Zhenshan Bing, and Alois C. Knoll, "Intelligent transportation systems using roadside infrastructure: A literature survey." *IEEE Transactions on Intelligent Transportation Systems*. Vol. 25, no. 7, pp. 6309-6327, 2023. doi : 10.1109/TITS.2023.3343434.
- [4] K. Hemant Kumar Reddy, Rajat Shubhra Goswami, Diptendu Sinha Roy. "A deep learning-based smart service model for context-aware intelligent transportation system." *The Journal of Supercomputing*, Vol. 80, no. 4, pp. 4477-4499, 2024. doi : 10.1007/s11227-023-05597-2.
- [5] Ruhul Amin Khalil, Ziad Safelnasr, Naod Yemane, Mebruk Kedir, Atawulrahman Shafiqurrahman, Nasir Saeed. "Advanced learning technologies for intelligent transportation systems: Prospects and challenges." *IEEE Open Journal of Vehicular Technology* Vol. 5, pp. 397-427, 2024. doi : 10.1109/OJVT.2024.3369691.
- [6] Y. Rajkumar, S. V. N. Santhosh Kumar, "A comprehensive survey on communication techniques for the realization of intelligent transportation systems in IoT based smart cities." *Peer-to-peer networking and applications*, Vol. 17, no. 3, pp. 1263-1308, 2024. doi : 10.1007/s12083-024-01627-9.
- [7] Aidil Redza Khan, Mohd Faizal Jamlos, Nurmadiha Osman, Muhammad Izhar Ishak, Fatimah Dzaharudin, You Kok Yeow, and Khairil Anuar Khairi. "DSRC technology in Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) IoT system for Intelligent Transportation System (ITS): A review." *Recent trends in mechatronics towards industry 4.0: selected articles from iM3F 2020, Malaysia*, pp. 97-106, 2021. doi : 10.1007/978-981-33-4597-3_10.
- [8] Jiuling Jin, Helai Huang, Ye Li, Yuxuan Dong, Gongquan Zhang, Jiguang Chen, "Variable speed limit control strategy for freeway tunnels based on a multi-objective deep reinforcement learning framework with safety perception." *Expert Systems with Applications*, Vol. 267, pp. 126277, 2025. doi : 10.1016/j.eswa.2024.126277.
- [9] Ellen F. Grumert, Andreas Tapani, Xiaoliang Ma, "Characteristics of variable speed limit systems." *European transport research review*, Vol. 10, no. 2, pp. 21, 2021. doi : 10.1186/s12544-018-0294-8.
- [10] Sadat, Mohd, Syed Aqeel Ahmad, and Mehmet Ali Silgu. "Mixed traffic modelling: An overview of car following and lane change models." *AI and Machine Learning Impacts in Intelligent Supply Chain* (2024): 209-225. DOI: 10.4018/979-8-3693-1347-3.ch014
- [11] Sadat, Mohd, et al. "A study on environmental impact of slow moving electric vehicles using microsimulation on Lucknow urban road with an on-ramp." *Environmental Health Insights* 18 (2024): 11786302241231706. doi.org/10.1177/11786302241231706
- [12] Müller, E. R., Carlson, R. C., Kraus, W., & Papageorgiou, M. (2015). Microsimulation analysis of practical aspects of traffic control with variable speed limits. *IEEE Transactions on Intelligent Transportation Systems*, 16(1), 512-523.
- [13] Sadat, Mohd, et al. "Smart ramp metering techniques An overview." *Advances in Science, Engineering and Technology* (2025): 50-54.
- [14] Sadat, M., Ahmad, S. A., & Silgu, M. A. (2025). Microscopic modelling of uncontrolled on-ramp at urban expressway under mixed traffic conditions. In *Advances in Science, Engineering and Technology* (pp. 55-59). CRC Press.
- [15] Hegyi, A., De Schutter, B., & Hellendoorn, J. (2005). Model predictive control for optimal coordination of ramp metering and variable speed limits. *Transportation Research Part C: Emerging Technologies*, 13(3), 185-209.
- [16] Abdel-Aty, M., Dilmore, J., & Dhindsa, A. (2006). Evaluation of variable speed limits for real-time freeway safety improvement. *Accident Analysis & Prevention*, 38(2), 335-345.
- [17] Carlson, R. C., Papamichail, I., Papageorgiou, M., & Messmer, A. (2010). Optimal motorway traffic flow control involving variable speed limits and ramp metering. *Transportation Science*, 44(2), 238-253.