

SmartFlow with DecongesQN: Dynamic Traffic Optimization Using Deep Q-Network Reinforcement Learning

Neave Kallivalappil¹, Afif Deshmukh², Ashwini Pansare³, Kranti Wagle⁴, Ankita Amburle^{5,*}, Heenakausar Pendhari⁶
^{1, 2, 3, 4, 5, 6} Department of Computer Engineering, Fr. Conceicao Rodrigues College of Engineering, Mumbai, India
Email: ¹neave.mailbox@gmail.com, ²ashwinip@fragnel.edu.in, ³crce.iic@fragnel.edu.in, ⁴kranti@fragnel.edu.in, ⁵ankita.amburle@fragnel.edu.in, ⁶heenak.pendhari@fragnel.edu.in
*Corresponding Author

Abstract—SmartFlow revolutionizes urban traffic management through advanced technologies and adaptive strategies. Central to SmartFlow is DecongesQN, a Deep Q-Network (DQN) based reinforcement learning model that dynamically adjusts traffic signals to optimise flow and enhance efficiency in real-time. Utilising convolutional neural networks, DecongesQN accurately detects and tracks vehicles, enabling precise congestion assessment. This paper presents SmartFlow’s architecture, the integration of DecongesQN, simulation results, and real-world implementations, demonstrating its scalability and effectiveness. Beyond adaptive signal control, SmartFlow contributes to safer and more sustainable urban transportation networks. With its innovative approach, SmartFlow with DecongesQN emerges as a promising solution to address the challenges of modern urban mobility, paving the way for smarter and more resilient cities.

Keywords—Urban mobility, Deep reinforcement learning, Adaptive traffic signal control, traffic congestion, Road safety, Smart cities

I. INTRODUCTION

Traffic congestion has emerged as a critical challenge in urban areas. Rapid urbanization and economic development have led to increasing travel demands, overwhelming existing road infrastructure and causing congestion issues (Chen and Lam, 2015). Private transportation, improper management practices, and overall city development further worsen this problem (Subba, 2020). The high volume of traffic significantly impacts transportation systems, leading to delays, accidents, and environmental issues globally (Goncalves and Ribeiro, 2020). Salisu, Akanmu, Fasina and Sanni (2020) states that the relationship between land use and congestion is critical for managing urban areas’ traffic congestion. Wang, Chen, Li and Liu (2023) further states that poor planning in land use, poor road infrastructure, and disorganized public transport systems contributed to the onset of congestion. Factors such as traffic signal failures, lack of law enforcement, and ineffective traffic management exacerbate congestion (Rwakarehe, 2022). Traffic congestion not only affects transportation but also impacts the socio-economic well-being of road users in cities with high mobilization activities (Dongare, Huljute, Jadhav, Lad and Marawar, 2023). According to Statista (2021), the average commute distance in Mumbai by car was 12 kilometers

in 2019. Subsequently, Tom (2024) indicates that, for an average commute distance of 12 km, urban commuters spend an average of 237 hours per year stuck in traffic in Mumbai, with 110 hours attributed to congestion alone. This substantial amount of time not only leads to increased travel times but also results in significant fuel wastage and environmental harm due to emissions. In terms of emissions, the average commuter in Mumbai emits approximately 1,205 kg of CO₂ annually, with 375 kg attributed to congestion alone (Tom, 2024). This figure underscores the significant environmental impact of traffic congestion on a yearly basis, equivalent to the emissions that would require approximately 171 trees to absorb over the same period (Tom, 2024).

The escalating volume of vehicles on roads, driven by the increasing affordability of cars, exacerbates this problem, underscoring the limitations of conventional traffic management systems in efficiently regulating traffic flow. Traditional traffic management systems typically fail to manage the flow of traffic efficiently as the number of vehicles increases. Such alarming trends in urban mobility have propelled the discussion towards advanced technologies and intelligent systems that can address these challenges. Among these, Deep Reinforcement Learning (DRL) techniques, especially Deep Q-Networks (DQN), have emerged as powerful tools for traffic management. Inspired by principles of human decision-making, DRL algorithms offer adaptive solutions for optimizing traffic signal control, heralding a new era in urban transportation management.

As the years progress, the values associated with traffic congestion only increase exponentially. This calls for an urgent need for newer traffic management systems that can dynamically adapt to the changing needs of urban transportation.

In this paper, we introduce SmartFlow, an intelligent traffic management system designed to revolutionize urban mobility. At the heart is DecongesQN, which stands for Decongestion using Deep Q-Networks, a Deep Reinforcement Learning model that optimizes traffic signal timings in real-time. DecongesQN enables SmartFlow to adapt dynamically



Received: 9-4-2026
Revised: 30-6-2026
Published: 30-6-2026

to changing traffic conditions, improving flow and reducing congestion across urban networks.

A. Problem statement

Nowadays, urban traffic congestion persists as a widespread challenge, casting a big shadow over regular commutes and the greater environmental landscape. Traditional traffic signal control systems, designed to cope with the intricacies of contemporary conditions, frequently fail to deliver on their intended efficiency and instead generate longer travel times and ecological aftereffects. The vision of the SmartFlow project is sparked by the acute need to reimagine urban mobility into a determination to pioneer inventive and adaptable solutions.

In fact, it holds within it DecongesQN, having the transformational ability of solving the complexities associated with traffic management. The SmartFlow has been drawn to the DRL algorithm core, aiming to become a dynamic system and intelligent for the management of traffic with real-time changes in signal timings. The general objective here would be adding a new dimension of adaptability into the world of traffic control and orchestrating optimal improvements in traffic flow by promoting an ideal sustainable cause within urban transportation networks.

Thus, the SmartFlow initiative is specifically framed by inadequacies in conventional systems in satisfying the ever-changing demands of urban mobility. Out of the tension between the limitations of existing infrastructure and the burgeoning requirements for urban mobility, SmartFlow emerges as a beacon of innovation in the realm of traffic management. Deciding to transform the current situation, SmartFlow is poised to inject new meaning into traffic management with regards to responsiveness, effectiveness, and a firm commitment to environmental protection. The driving force behind SmartFlow is a deep sense of pursuing a holistic, far-reaching solution that would mark a new and transformational era in urban mobility.

B. Objectives

This paper aims to achieve the following objectives:

- Discuss the efficiency of DRL algorithms like Deep Q-Networks (DQN) in adaptive traffic signal control.
- Design a new concept called SmartFlow to innovate intelligent traffic management using the methodology of DRL.
- Introduce DecongesQN as the DQN component within SmartFlow, highlighting its role in dynamic traffic optimization.
- Perform a performance analysis of the proposed SmartFlow system with dynamic adaptation of traffic signal timings for congestion abatement and overall transportation efficiency enhancement.
- Validate the capability of SmartFlow in solving urban problems attributed to road congestion, such as inefficiency, time delay, and environmental degradation.

- Translate simulated results into practice by demonstrating the approach using a prototype—for an easy transition from theoretical concepts to real-life application.
- Contribute to further research development in the field of traffic management systems by proposing forward-looking answers aligned with the dynamic nature of urban environments.
- Provide insights and guidelines for future research in intelligent transportation systems, focusing on adaptive control in traffic signals.

By these means, this paper comprehensively aims at providing an understanding of how the SmartFlow system, powered by DecongesQN, is capable in real-world applications.

C. Significance

The significance of the project lies in its potential to solve the multifaceted problem of urban traffic congestion through the introduction of advanced technologies and innovative methodologies. Deep Reinforcement Learning (DRL) algorithms, such as Deep Q-Networks (DQN), promise a paradigm shift in metropolitan mobility with SmartFlow and DecongesQN. The significance of this research can be expressed along a number of dimensions as discussed below:

- **Reducing Traffic Congestion:** SmartFlow, with DecongesQN at its core, will significantly reduce traffic congestion in real time by dynamically changing signal timings of traffic signals. This is one of the biggest challenges in urban transportation systems, improving commute times, and reducing travel delays.
- **Environmental Impact:** SmartFlow reduces idling times while optimizing the flow of traffic, which reduces vehicle emissions in a direct effort toward environmental sustainability. This is to say that air quality improves with less CO₂ emission, countering global warming effects.
- **Economic Impact:** Saving time spent in traffic translates to an economic gain. Higher productivity and reduced costs for operation benefit commuters and businesses alike and translate to better economic growth.
- **Technological Advancement:** Implementing DRL into the traffic management problem can exemplify a practical implementation of advanced AI technologies in real-world problems. The implications of the project raise and remind with future implementation into urban infrastructure.
- **Scalability and Adaptability:** The methodology adopted by SmartFlow is such that it can scale and adapt to different city environments. It could be implemented based on local traffic dynamics of any city, thus making it an inclusive solution to this worldwide problem in urban transportation.
- **Policy and Planning:** Experience from SmartFlow's implementation will guide urban planning and

polycymak- ing. So, it is possible to look forward to better decision-making and strategic investments in this regard through data-driven approaches to traffic management.

D. Structure of the Paper

The paper is structured to provide a comprehensive analysis of the SmartFlow project, DecongesQN, and their implications for urban traffic management. Section 2 discusses related work, previous studies, and prevailing methodologies in the context of traffic management systems, particularly those involving adaptive signal control and DRL applications. Developments and limitations of these techniques set the stage for the SmartFlow project.

Section 3, Literature Survey, provides a thorough review of works on the application of DRL in traffic signal control; theoretical underpinnings, algorithmic developments, and practical implementations documented in academic and industry research provide a robust framework for the SmartFlow initiative.

Section 4, Findings and Discussion, outlines the key findings from the literature survey and related works, analyzing potential impacts and benefits that could be obtained through DRL-based traffic management systems and, more specifically, how SmartFlow and DecongesQN address previously identified gaps or improve existing solutions.

Section 5, Implementation and Results, details the SmartFlow system implementation process, highlighting the integration of DecongesQN into the DRL framework. Technical details about the prototype and the application of DecongesQN in traffic signal control are described. The outcomes of performance tests are presented to determine the proposed system's sufficiency for real-world operation. Finally, in Section 6, Conclusion and Future Works, we conclude with a summary of our study's findings, emphasizing the overall relevance of the SmartFlow project to the progress of urban traffic management. We identify limitations encountered during the study and suggest possible

directions for further research to improve and expand the SmartFlow system. This structured approach ensures that the information flows logically, providing coherence so the reader understands everything concerning the SmartFlow project, its innovative contributions, and its potential in revolutionizing urban traffic management systems.

II. LITERATURE REVIEW

Traffic signal control forms part of urban traffic management by aiming for optimization of the flow of vehicles through intersections to minimize congestion, delay, and improve the overall traffic efficiency. With ongoing urbanization and a growth in number of vehicles plying the roads, more sophisticated and intelligent traffic signal control systems are needed. The recent methodologies and results in the field of traffic signal control are surveyed, which can be divided into three major categories: Signal Timing Plans, Control Strategies, and Reinforcement Learning.

A. Signal Timing Plans

Signal timing plans are crucial parts of managing traffic flow in urban areas, aimed at optimizing the use of traffic signals so that congestion is avoided while efficiently utilizing the overall throughput of traffic flow.

Several studies have outlined various methods and tools that are utilized to optimize these plans, such as good strategies and existing gaps in research. Qadri, Go'kc,e and Ö ner (2020) reviewed traffic signal control methods. Key findings reveal that among the most cost-effective methods for improving traffic flow and alleviating congestion in urban areas is traffic signal timing optimization. The most widely utilized simulation tools in practice are VISSIM and SUMO, which have been applied to simulate conditions of traffic and test different possible signal timing options. Additionally, population-based metaheuristic algorithms are identified as dominant optimization methods for traffic signal control, proven effective in solving complicated traffic scenarios. Jamal, Rahman, Al-Ahmadi, Ullah and Zahid (2020) complement this understanding by optimizing signal timings at two isolated signalized intersections through Genetic Algorithm (GA) and Differential Evolution (DE). Results show that average vehicle delays decrease by a notable 15

Despite these advances, limitations have been observed. Qadri et al. (2020) identify the scarcity of works related to traffic signal timing problems due to sudden events of transient peak demand and at signalized roundabouts. There is also a significant shortage of research based on real-world data and its applications in real-world settings.

Jamal et al. (2020) note that their study was conducted only on isolated intersections, highlighting the need for further research addressing network-wide optimization problems. The impact on non-motorized modes such as pedestrians and cyclists has not been investigated. Their study used a one-objective optimization technique applying to minimize average vehicle delay, suggesting that the future research should apply multi-objective heuristic modeling with other performance metrics including the number of stops, energy consumption, and emissions.

B. Control Strategies

Strategies on control traffic signals are important to the management of urban traffic in that they target optimizing efficiency at intersections, reducing delays, and perfecting overall traffic flow. Over the years, several methodologies as well as varying technologies have been explored on addressing the ITSCP problem, each differently brought with advantages and limitations. According to Eom and Kim (2020) the traffic congestion problem has started growing; hence it strengthens research in ITSCP. Improved computational capabilities allow the application of sophisticated algorithms to large and realistic traffic networks. Connected vehicle technologies, including Vehicle-to-Infrastructure (V2I) and Vehicle-to-Vehicle (V2V) communication, have enabled collecting detailed traffic data for signal control improvement. Liang, Guler and Gayah (2020) proposed an approach to enhance delay equity by introducing vehicle delay thresholds into traffic signal control. Their results show effectiveness at any demand level and in partial penetration of connected vehicles. However, they find a trade-off between equity and intersection efficiency, with diminishing returns when delay threshold levels decrease.

Yu, Ma and Yang (2020) introduce a time-slot-based signal scheme model for fixed-time control at isolated intersections. Their model demonstrates improvements in intersection capacity and reductions in average vehicle delay

compared to conventional structures. They suggest an optimal time-slot number of 6 or 7 to balance solution optimality and model complexity.

Methodologies employed to tackle ITSCP are diverse. Eom and Kim (2020) categorize these into approaches including fixed-time, actuated, and adaptive traffic control strategies, analytic methods based on mathematical models, rule-based methods, genetic algorithms, simulation-based approaches, dynamic programming, multi-agent systems, game theory, neural networks, and reinforcement learning. Despite advancements, limitations persist. Eom and Kim (2020) highlight that most studies focus on simplified and small intersection networks, limiting generalizability to more complex real-world scenarios. There is a lack of consideration for diverse road user types, physical road properties, and varying traffic flow conditions. Scenarios beyond normal traffic conditions, such as accidents, and the implications of autonomous and connected vehicle technologies are often overlooked.

C. Reinforcement Learning

Reinforcement Learning (RL) is a promising approach to optimize traffic signal control, improving traffic flow and reducing congestion using complex neural network architectures and sophisticated algorithms. Many studies have demonstrated the potential of RL-based methods for various traffic scenarios, discussing strengths and weaknesses. Genders and Razavi (2016) designed a deep reinforcement learning agent to control traffic signals. A deep convolutional neural network approximated the action-value function. Discrete Traffic State Encoding, including vehicle presence, speed, and the current traffic phase, was used to define the state space. An epsilon-greedy exploration policy with decreasing epsilon value in Q-learning was performed, training and testing the agent using SUMO traffic micro-simulator. The deep reinforcement learning agent significantly outperformed a traditional neural network agent, achieving an 82 percent reduction in average cumulative delay, a 66 percent reduction in average queue length, and a 20 percent reduction in average travel time. The agent's policy might not ensure fairness, as it did not prioritize equal traversal opportunities for all vehicles. Balancing fairness and optimality could be achieved with an appropriate reward function. Building on dynamic adjustment in RL, Wei, Zheng, Yao and Li (2018) introduced a methodology combining offline data collection and online learning. Their neural network architecture included convolutional layers for image features, concatenation of image and traffic features, separate Q-value estimation branches for each light phase controlled by a "phase gate," and "memory palaces" for balanced experience replay. The IntelliLight method outperformed baseline methods, effectively adjusting to different traffic conditions, such as peak vs. non-peak hours, weekdays vs. weekends, and major vs. minor arterials. It learned to prioritize major arterials without prior knowledge by preferring to give green lights to directions with higher traffic volumes. IntelliLight could not capture multi-phase traffic lights and intricate road networks, and its effectiveness was never tested with real-field experiments.

Liang, Du, Wang and Han (2019) proposed a deep reinforcement learning model based on the Deep Q-Learning and Policy Gradient approach. They incorporated a reward function considering both traffic flow and delay to optimize

the traffic light timings. The model has highly efficiently controlled the traffic lights to be optimized in terms of flow and delays by dynamic adjustment based on real-time data. Further experiments for varied and complex scenarios have to be done to make this model robust and adaptable.

Applying a Deep Q-Network (DQN) to learn the optimal policy to control traffic signals, Zheng, Zang, Xu, Wei, Yu, Gayah, Xu and Li (2019) proposed a simple state and reward design. States are represented based on the number of vehicles on each lane and include the current phase of the signal while the reward function sums queue lengths across all lanes. Their RL approach realized state-of-the-art performance close to the performance of traditional transportation approaches under the steady traffic condition and performed better than these under complicated real-world traffic conditions. Like Liang et al. (2019), this research emphasized the need to test in different complex traffic scenarios. Paul and Mitra (2020) proposed a DRL agent for multiple intersections that adaptively controlled traffic signals. The adaptive agent was implemented as a policy gradient algorithm that extracted spatio-temporal data and actions over different neural network models to promote coordination of traffic flow and reduce congestion. The DRL-based traffic signal control system outperformed fixed-duration signal systems across all simulation metrics, achieving better coordination of traffic flows and reduced congestion. Scaling this up to more extensive networks is challenging.

Kuang, Zheng, Li and Gao (2021) constructed a reinforcement learning model based on historical traffic flow data. A dual-objective reward function is proposed, minimizing vehicle delay while optimizing signal timing alignment with traffic demand. Control phase combination and elimination of rare states were used to decide the state and action spaces for the Q-table, simplifying the method to optimize algorithm complexity and enhance intersection capacity with less computational complexity. The paper addresses inefficiency in learning and resource consumption associated with existing RL approaches by optimizing the state space to improve storage and training time with similar control effects. A framework of deep reinforcement learning based on the approach of consistent state and reward design (CSR D) was proposed by Bouktif, Cheniki, Ouni and El-Sayed (2023). The state vector is defined relative to the number of vehicles, queue length, or waiting time along with corresponding reward functions and penalty processes. The discrete action space comprises four phases of traffic signals, where Double DQN with Prioritized Experience Replay is used as a learning architecture. The approach of state and reward design led to stable training and sharp convergence to optimal policies, outperforming state-of-the-art RL techniques for traffic signal control. The effect of green phase duration has not been considered; such a setting requires further optimization and validation in real scenarios.

III. MATERIALS AND METHODS

A. Data Collection

The concept of SmartFlow is based on training DecongesQN, a Deep Q-Network (DQN), by leveraging traffic simulations from Eclipse Simulation of Urban Mobility (SUMO) as the basis of the source of data. SUMO (Lopez, Behrisch, Bieker-Walz, Erdmann, Floettero, Hilbrich,

(Lu \ddot{c} ken, Rummel, Wagner and Wiessner, 2018) is a realistic and flexible simulation platform for urban traffic flows, which allows modeling of complex interactions between traffic movements, road networks, traffic signals, and weather conditions. The training data set for DecongesQN was developed with SUMO, which emulated different traffic cases to reflect realistic complexity with the variations in traffic density, road configuration, and the dynamic demand change. Vehicle trajectories, states of the traffic signals, and other pertinent information derived from SUMO simulations provided a basis for reinforcement learning. This simulated traffic environment allows DecongesQN to develop optimal policies, endowing SmartFlow with the ability to tackle diverse problems in urban traffic management with efficiency and adaptability.

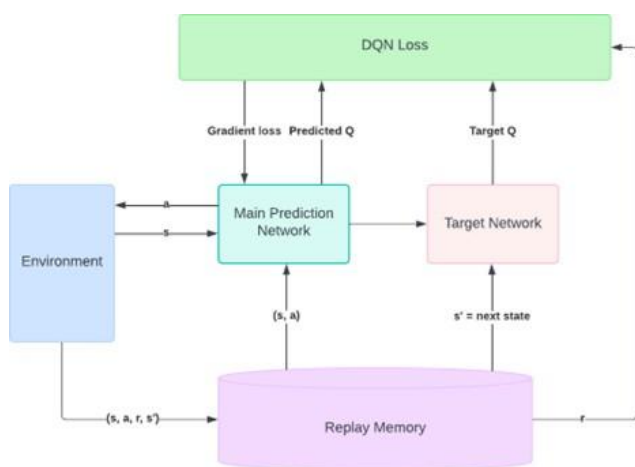


Fig. 1. DecongesQN Architecture with Target Network

B. Reinforcement Learning with DecongesQN

An essential component of the SmartFlow architecture is Reinforcement Learning (RL), implemented through DecongesQN. This DQN model enables self-learning decision-making to optimize traffic signal control. DecongesQN learns from interactions between the traffic signal control agent and the dynamic urban environment simulated with Eclipse SUMO.

It consists of an agent, which monitors the traffic system state at each point in time, as depicted by the SUMO simulator. More information regarding the state representation can be found in the following subsection. Then it acts by changing signal timing to impact the flow of traffic, and DecongesQN attempts to find the optimal policy through trial-and-error to maximize cumulative rewards, about congestion and waiting time.

A target network in the architecture of DecongesQN, as depicted in Figure 1, is designed to improve learning stability. This helps stabilize the process of training by minimizing the possibility of oscillations and helps to speed convergence; these aspects strengthen and accelerate the efficiency of the reinforcement learning framework for SmartFlow's urban traffic management system. Via adaptive learning mechanisms, the DecongesQN evolves to effectively manage complex scenarios while providing

potential advantages in the enhancement of RL algorithms in enhancing the efficiency of urban traffic management systems.

C. State Representation in DecongesQN

In the traffic management framework of SmartFlow, we use the state representation in DecongesQN to capture the essential nature of urban traffic dynamics. For the purpose of traffic-condition modeling, we separate the lanes on a road into two groups: straight and right-turn lanes and left-turn lanes. For example, in a 4-lane road, three of them would make a lane group and the remaining a separate group.

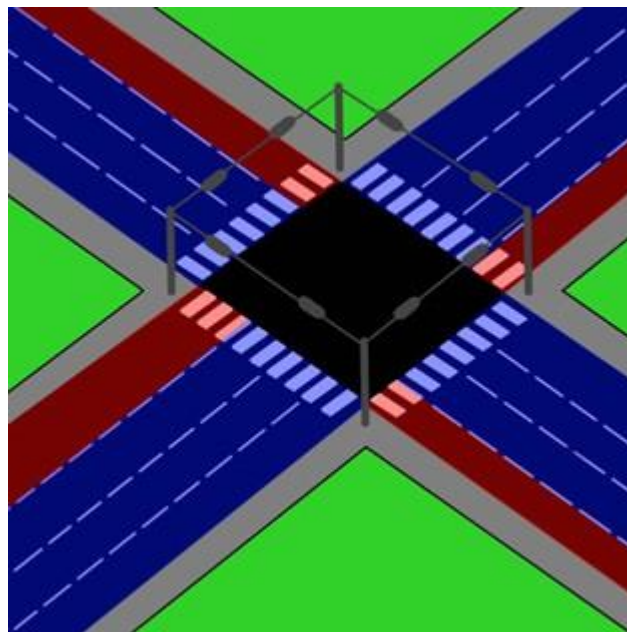


Fig. 2. 4-way intersection with lane groups highlighted

The detailed configuration of lane groups is clarified visually in Figure reffig:road-marked. This division shall enable the description of traffic dynamics in a rather sophisticated way, especially in intersections. DecongesQN takes a state input in the form of an array representing cumulative waiting times of vehicles within each lane group. For a 4-way intersection, the state consists of an array of 8 elements. This state representation provides a better balance of model complexity and performance. During our experiments, several other state representations were also experimented with, including vehicle queue length, vehicle density, and assigning weights to different types of vehicles. However, the cumulative waiting times turned out to be the more desirable and efficient state representation. This simplicity, that has a direct correlation with the reward function, improves the learning efficiency of the model as portrayed by Salah Bouktif's results (Bouktif et al., 2023). The selected state representation supports the aim of the system in optimizing traffic signal control in order to decrease waiting times and enhance urban mobility.

D. Reward Function in DecongesQN

The main purpose of DecongesQN in SmartFlow is the reduction of wait times at intersections, hence optimizing the flow of traffic in urban areas. The reward function works as a

negative reward directly proportional to the cumulative waiting time of vehicles within the specific lane groups. The system works by the principle that optimal behavior for the model is when it takes actions that reduce this negative reward, thus reducing wait times. The reward R_t at time step t is formally defined as the negative of the cumulative waiting times in the system.

The aspiration to maximize the accumulation of rewards over time motivates the learning process of DecongesQN. A positive cumulative reward corresponds to successful traffic control of signals, aligning with the negative reward structure designed to encourage proactive traffic management and improve mobility in urban environments.

E. Policy

DecongesQN adopts an ϵ -greedy strategy for its policy. That is, it balances exploitation and exploration when choosing actions: with a probability of $1 - \epsilon$, it selects the optimal action based on current knowledge, while with a probability of ϵ , it explores other possible actions.

This approach enables the model to discover potentially rewarding actions while still favoring those known to yield high rewards. The value of ϵ decreases over time, allowing the model to gradually transition from exploration to exploitation as it gains better understanding of the traffic environment.

Such an adaptive strategy enhances the learning capability of the model, enabling it to respond effectively to the dynamic nature of traffic and support improved long-term decision-making.

F. Neural Network

The neural network structure used in DecongesQN is relatively simple yet efficient. It consists of several layers and is designed to learn traffic signal control policies adequately. This includes some hidden layers that may learn complex patterns and an output layer which will produce Q-values for every possible action. The best action is the one with the highest Q-value. This architecture uses historical experiences in the replay buffer and maintains the explorative properties with exploitative properties using an epsilon-greedy policy. The simplicity of the architecture leads to fast convergence and quicker training with data from traffic simulation. Please refer to Table 1 for a general view of the neural network architecture.

G. Model Training of DecongesQN

The DecongesQN model in SmartFlow is trained using SUMO simulations to generate a wide range of traffic scenarios and congestion patterns. This allows DecongesQN to develop versatile and adaptive policies for traffic signal control. The basic design of DecongesQN is based on Deep Q-Learning (DQN), a family of popular reinforcement learning

TABLE I. Architecture of Neural Network in DecongesQN

Layer (type)	Output Shape	Param #
input1 (InputLayer)	[(None, 8)]	0
dense (Dense)	(None, 400)	3600
dense1 (Dense)	(None, 400)	160400
dense2 (Dense)	(None, 400)	160400
dense3 (Dense)	(None, 400)	160400
dense4 (Dense)	(None, 400)	160400
dense5 (Dense)	(None, 4)	1604

algorithms for solving complex decision-making tasks. DecongesQN applies neural networks as approximators of the Q-function to predict expected future rewards for all actions conceivable from a particular state. We employ a target network in DecongesQN to make training more stable and convergent. The target network is updated by periodically copying the parameters from the primary Q-network which avoids oscillation or divergence of the Q-value estimations in training.

Specifically, Q-values denote the future reward one would expect to obtain after taking a given action at a given state. The model learns an optimal policy concerning which decisions about actions enhance these Q-values. In this manner of progressive learning, the model becomes more refined so that it makes decisions on traffic signal control that are rational in nature, thereby decreasing waiting time and enhancing the movement of traffic. The training procedure repeats the updates of the Q-values.

The model updates the Q-values iteratively, thereby approximating the optimal action-value function. Through the discount factor γ , a balance is maintained between immediate rewards and future rewards. Fundamentally, this guides the model toward selecting actions that maximize cumulative rewards over time steps, which in turn helps optimize traffic flow by reducing waiting times.

The use of distributed training and the integration of DecongesQN with a target network contribute to the development of a powerful and adaptable traffic management system capable of performing effectively under diverse urban traffic conditions.

IV. LEGEND AND PARAMETERS



Fig. 3. Top view of the prototype

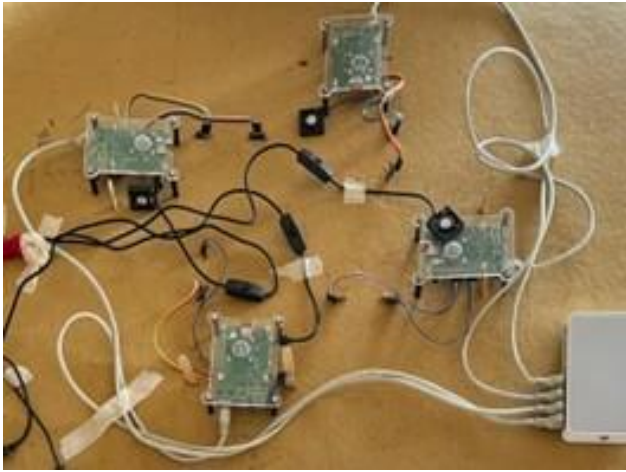


Fig. 4. Bottom view of the prototype

- ϑ, ϑ^* : Weights of the primary network Q and target network Q_{target}
- D : Replay memory with capacity N
- M, T : Total number of episodes and maximum time steps per episode
- ϵ : Exploration rate (decays over time)
- γ : Discount factor (range: 0 to 1)
- K : Minibatch size for gradient descent
- C : Number of steps after which the target network is updated
- $L(\vartheta)$: Loss function used for updating network weights

The training process involves interacting with the SUMO simulation environment, collecting experiences, and learning from them. The use of replay memory D enables the model to learn from a diverse set of experiences, thereby breaking the correlation between consecutive samples and improving learning stability.

The target network Q_{target} provides stable targets during training, helping to prevent oscillations or divergence in the learning process. By periodically updating the target network with the weights from the primary network, the algorithm ensures stability and convergence toward an optimal policy.

Through this iterative learning process, DecongesQN refines its policy to make effective decisions for traffic signal control, ultimately reducing waiting times and improving overall traffic flow.

A. Real-world Traffic State Collection

In the actual implementation of SmartFlow, computer vision is used to retrieve vital data about traffic status. Cameras are mounted across intersections to monitor the movement of vehicles across the lanes. The state collection process involves using a tailor-made object detection model with the integration of the popular SORT (Simple Online and Realtime Tracking) algorithm. This helps track vehicles across multiple frames with a high degree of proficiency and

assigns identifiers to each vehicle. This enables precise tracking of movement and waiting times by specific vehicles in the system.

The system calculates cumulative waiting times for each lane group, such as straight+right and left in a divided road. Waiting times are the total time spent by vehicles waiting within the lanes considered.

These calculated cumulative waiting times are converted to an array of eight values representing the 4-way intersection, encapsulating the waiting times for every lane group. This array is the input to DecongesQN, allowing it to make informed decisions on traffic signal control based on actual traffic dynamics.

B. Prototype Description

The SmartFlow prototype demonstrates the practical implementation of DecongesQN for adaptive traffic signal control. Utilizing the Raspberry Pi 4B platform, combined with custom-built hardware components, we showcase real-time traffic optimization. In this subsection, we describe in detail the SmartFlow prototype setup: hardware components, system architecture, and the operational workflow used in its design.

C. Hardware Components

The SmartFlow prototype is built on a sturdy base, serving as the foundation for the entire system. The top surface of the board is painted with road markings to simulate a realistic traffic environment, as illustrated in Figure 3. Mounted on the underside of the cardboard base are four Raspberry Pi 4B microcomputers, each equipped with 8GB of RAM, collectively referred to as the worker Pis. Each worker Pi is equipped with an OV5647 5MP 1080P IR-Cut Camera, strategically positioned to capture live traffic data (Figure 4). These cameras are securely mounted onto the cardboard base using custom mounts crafted from ice cream sticks, providing both stability and flexibility in positioning. They capture the feed from the road opposite their placement. The camera feeds of these cameras are visible in Figure 5. The system also includes a dedicated server Raspberry Pi 4B with 4GB of RAM, referred to as the master Pi, serving as the central processing unit for traffic signal control and model deployment.



Fig. 5. The prototype with RC cars

D. System Architecture

The SmartFlow prototype assumes a four-way intersection and leverages edge computing IoT. It implements the custom SmartFlowVision algorithm on the worker Pis. SmartFlowVision employs lane group segmentation and custom object tracking to monitor traffic conditions. The algorithm features a dividing line that separates lane groups, enabling independent monitoring and analysis.

At each worker Pi node, the custom-built traffic detection script utilizes the SORT algorithm to detect and track RC cars within the camera's field of view. Following detection, the script assigns a timer to each detected car, facilitating individualized monitoring of waiting times. Subsequently, the script calculates the cumulative waiting time (CWT) for vehicles in both lane groups and streams this tuple to the master Raspberry Pi. The data is in the format $(CWT_{left}, CWT_{right})$.

E. Operational Workflow

The operational workflow of the SmartFlow prototype is as follows:

- **Data Collection:** Each worker Pi captures live traffic data using its integrated camera and processes it using the SmartFlowVision algorithm.
- **Traffic Analysis:** The SmartFlowVision algorithm segments lane groups, detects vehicles, and calculates the cumulative waiting time for each lane group.
- **Data Transmission:** The waiting time data from each worker Pi is transmitted to the master Pi, where it is aggregated into an 8-element Python list representing the waiting times for all lane groups.
- **Model Inference:** The aggregated waiting time data is fed into DecongesQN deployed on the master Pi. The model outputs an action number between 0-7, corresponding to the traffic signal phase to be activated.
- **Traffic Signal Control:** The master Pi remotely controls the traffic signal phases based on DecongesQN's output, using the Raspberry Pi's Remote GPIO communication functionality. The traffic lights, connected to the worker Raspberry Pis, change their states accordingly, allowing vehicles in the designated lane groups to proceed.

Dynamic Adjustment: DecongesQN continues to monitor traffic conditions, making further decisions based on the streaming data, thereby ensuring adaptive and efficient traffic signal control in real-time. This integrated approach enables the SmartFlow prototype to dynamically optimize traffic flow, mitigate congestion, and enhance overall transportation efficiency in urban environments.

V. RESULTS

For comparative analysis, we simulated three variations of the traditional method: with green phase durations set to 30s, 45s, and 60s, each followed by a 4-second yellow phase. In these simulations, identical routes and vehicular densities were maintained for consistency. Results revealed that with a 30s green phase, the maximum cumulative queue

length peaked at 30 cars, with a corresponding maximum cumulative waiting time of 1,300s.

Algorithm 1 Training Algorithm for DecongesQN

- 1: Initialize primary network Q with random weights θ
- 2: Initialize target network Q_{target} with weights $\theta^- = \theta$
- 3: Initialize replay memory D to capacity N
- 4: **for** episode = 1 to M **do**
- 5: Initialize state s_0
- 6: **for** $t = 1$ to T **do**
- 7: With probability ϵ , select a random action a_t
- 8: Otherwise, select $a_t = \arg \max_a Q(s_t, a; \theta)$
- 9: Execute action a_t in SUMO simulation
- 10: Observe reward r_t and next state s_{t+1}
- 11: Store transition (s_t, a_t, r_t, s_{t+1}) in D
- 12: Sample a random minibatch of transitions (s_j, a_j, r_j, s_{j+1}) from D
- 13: Compute target y_j for each minibatch transition:

$$y_j = \begin{cases} r_j & \text{if } s_{j+1} \text{ is terminal} \\ r_j + \gamma \max_{a'} Q_{\text{target}}(s_{j+1}, a'; \theta^-) & \text{otherwise} \end{cases}$$

- 14: Perform a gradient descent step on loss $L(\theta)$:

$$L(\theta) = \frac{1}{K} \sum_j (y_j - Q(s_j, a_j; \theta))^2$$

- 15: **if** $t \bmod C == 0$ **then**

- 16: Update target network weights:

$$\theta^- \leftarrow \theta$$

- 17: **end if**

- 18: Update state $s_t \leftarrow s_{t+1}$

- 19: Decay exploration rate ϵ

- 20: **end for**

- 21: **end for**

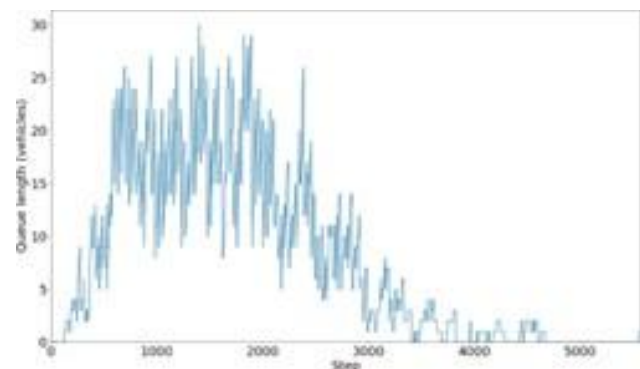


Fig. 6. Traditional method with 30s green phase

Increasing the green phase duration to 45s and 60s resulted in higher queue lengths of 45 and 51 cars, respectively, accompanied by escalating cumulative waiting times of 2,392s and 4,313s. These values are detailed in Table 2, and the cumulative waiting time plots are visualized together in Figure 6. On the other hand, the SmartFlow system with DecongesQN clearly shows improvements, with a maximum cumulative queue length of 15 cars and a maximum cumulative waiting time of 234s. Most importantly, on average, a long phase duration of 60s with the traditional method shows a longer waiting time and a larger queue length than the SmartFlow system, indicating that the system improves the flow of traffic and reduces congestion. It is found that cumulative waiting time reduces by 94.57%. Overall, these results show the potential of SmartFlow with DecongesQN in optimizing traffic signal

control and overall traffic flow, making it a solution that can address the challenges of urban traffic management.

TABLE II. Comparison of evaluation statistics

Method	Cumulative Waiting Time (s)	Cumulative Queue Length (vehicles)
Traditional (30s)	1,300	30
Traditional (45s)	2,392	45
Traditional (60s)	4,313	51
SmartFlow with DecongesQN	234	15

VI. DISCUSSION

Strategic, SmartFlow approach to traffic management with the DecongesQN power is quite an interesting consideration for urban environments.

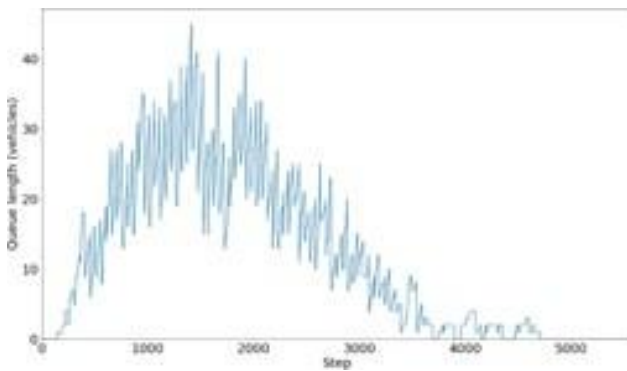


Fig. 7. Traditional method with 45s green phase

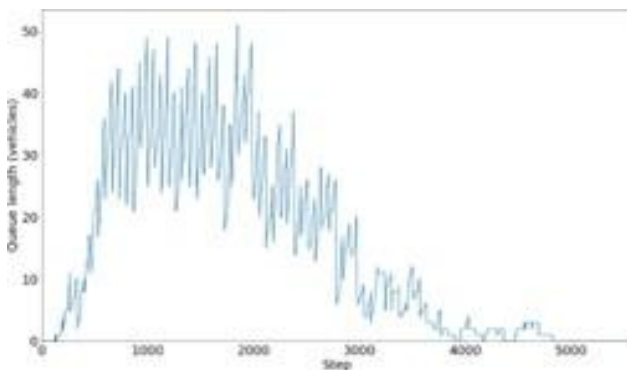


Fig. 8. Traditional method with 60s green phase

The SmartFlow system offers multifaceted management of the traffic that is in-built in towns with many important implications. It can potentially reduce the fuel consumption and the environmental impact of idling vehicle engines.

SmartFlow encourages fuel efficiency via rigorous optimization in signal timings while minimizing waiting time. In this way, air pollution and greenhouse gases are curbed. One major environmental concern is the reduction in air quality around red lights that idle engines create. The efficient management of a flow by traffic signals under SmartFlow results in less wastage of fuel as well as less carbon

dioxide (CO₂) and nitrogen oxide emissions, resulting in a sustainable solution to the problem.

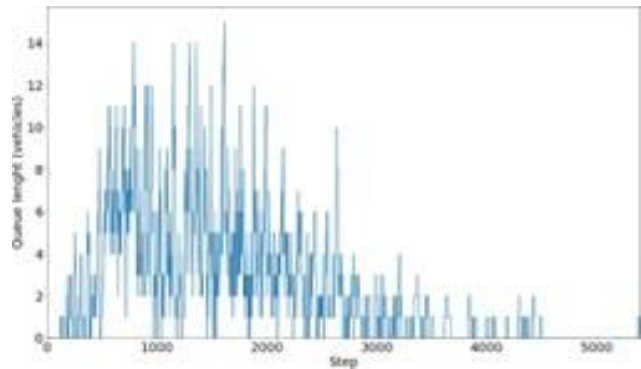


Fig. 9. SmartFlow system with DecongesQN

Beyond environmental benefits, SmartFlow affects the general flow of traffic in cities. Its adaptive adjustment of signal timings and sequences based on real-time conditions brings about smoother traffic flow, thereby decongesting roads and reducing travel time for commuters.

The system promises to be a very effective solution for the nearer, immediate future of traffic management and a potential solution for more distant possibilities of urban development while being environmentally conscious. It aligns with the broader-scale goals of developing smarter, greener, more livable cities by optimizing traffic signals.

VII. CONCLUSION AND FUTURE SCOPE

SmartFlow with DecongesQN introduces a new paradigm for traffic management to be realized through Reinforcement Learning of dynamic signal optimization. The promising results with positive implications reported here might become a revolutionary solution to urban traffic problems, reducing waiting times, improving commuter experiences, and promising considerable environmental benefits due to reduced fuel wastage and emissions.

A. Expansion to Multimodal Transport

Further developments of SmartFlow are going to take into account multimodal transportation management, taking into account, in addition to vehicular traffic, pedestrians and cyclists along with public transport. The holistic approach in traffic management leads to more inclusive and efficient urban mobility.

B. Collaboration with Vehicle-to-Everything (V2X) Technologies

Interaction with newly emerging V2X (Vehicle-to-Everything) technologies can provide new avenues for SmartFlow. The obtained states of the camera streams are prone to errors since the SORT algorithm could fail to track vehicles that are obscured from view, such as a car behind a bus. SmartFlow becomes a more dynamic and responsive traffic management system due to smooth interaction with connected vehicles and infrastructure. Besides reducing chances of errors, overall user experience is improved—a major stride toward effective and reliable urban traffic management.

C. Public Awareness and Engagement

Public education and engagement will be important steps toward developing favorable support for the deployment of SmartFlow. Community engagement, pilot programs, and feedback gathering can foster a sense of ownership and participation, making the system a more effective and accepted traffic management solution. In essence, SmartFlow with DecongesQN heralds the beginning of a new era in traffic management, providing ample opportunities to innovate and collaborate. The evolutionary process of the technology promises to create ever more intelligent, sustainable, and resilient urban environments.

REFERENCES

- [1] TomTom Traffic Index, "Mumbai Traffic Report," 2024. [Online]. Available: <https://www.tomtom.com/traffic-index/mumbai-traffic/>
- [2] S. Bouktif, A. Cheniki, A. Ouni, and H. El-Sayed, "Deep reinforcement learning for traffic signal control with consistent state and reward design approach," *Knowledge-Based Systems*, vol. 267, p. 110440, 2023.
- [3] B. Y. Chen and W. H. K. Lam, "Smart transportation: Theory and practice," *Journal of Advanced Transportation*, vol. 50, pp. 141–144, 2015.
- [4] T. Dongare, D. Huljute, P. Jadhav, A. Lad, and S. Marawar, "A review on traffic management and road analysis of Porwal Road," *International Journal for Research in Applied Science and Engineering Technology*, vol. 11, pp. 881–884, 2023.
- [5] M. Eom and B. I. Kim, "The traffic signal control problem for intersections: A review," *European Transport Research Review*, vol. 12, 2020.
- [6] W. Genders and S. Razavi, "Using a deep reinforcement learning agent for traffic signal control," arXiv:1611.01142, 2016.
- [7] L. A. P. J. Goncalves and P. J. G. Ribeiro, "The impact of theering road conclusion to the city of Guimaraes," *WSEAS Trans. Environment and Development*, vol. 16, pp. 11–22, 2020.
- [8] A. Jamal, M. T. Rahman, H. M. Al-Ahmadi, I. Ullah, and M. Zahid, "Intelligent intersection control for delay optimization," *Sustainability*, vol. 12, p. 1896, 2020.
- [9] L. Kuang, J. Zheng, K. Li, and H. Gao, "Intelligent traffic signal control based on reinforcement learning," *ACM Trans. Internet Technology*, vol. 21, pp. 1–24, 2021.
- [10] X. Liang, X. Du, G. Wang, and Z. Han, "A deep reinforcement learning network for traffic light cycle control," *IEEE Trans. Vehicular Technology*, vol. 68, pp. 1243–1253, 2019.
- [11] P. A. Lopez *et al.*, "Microscopic traffic simulation using SUMO," in *Proc. IEEE Int. Conf. Intelligent Transportation Systems (ITSC)*, 2018, pp. 2575–2582.
- [12] A. Paul and S. Mitra, "Deep reinforcement learning based traffic signal optimization," in *Proc. IEEE ANTS*, 2020.
- [13] H. Wei, G. Zheng, H. Yao, and Z. Li, "IntelliLight: A reinforcement learning approach for traffic light control," in *Proc. ACM SIGKDD*, 2018, pp. 2496–2505.