Real-Time Dynamic Traffic Signal Controller

Date - 09-07-2024
By: Rakshitha Kalkura, Ayushi Kandpal, Hamees Muhammed

- ## Introduction

Due to rapid urbanization and the resulting increase in household car ownership, urban traffic congestion has become a significant obstacle. Traffic congestion not only wastes fuel but also increases harmful emissions, including greenhouse gasses like carbon dioxide and other harmful particles such as nitrogen oxides, which can harm human health. Existing studies indicate that the transport sector contributes to 23% of total CO2 emissions from fuel combustion, with road traffic accounting for about three-fourths of these emissions. Additionally, traffic congestion in urban areas can increase emissions by 40%. Therefore, mitigating traffic congestion is extremely urgent. One of the most effective approaches to achieve this is to control traffic signals more intelligently. In densely connected urban areas, the signal control strategies of intersections are highly correlated, making it crucial to address city-level signal control rather than focusing on a few regions separately. [1]

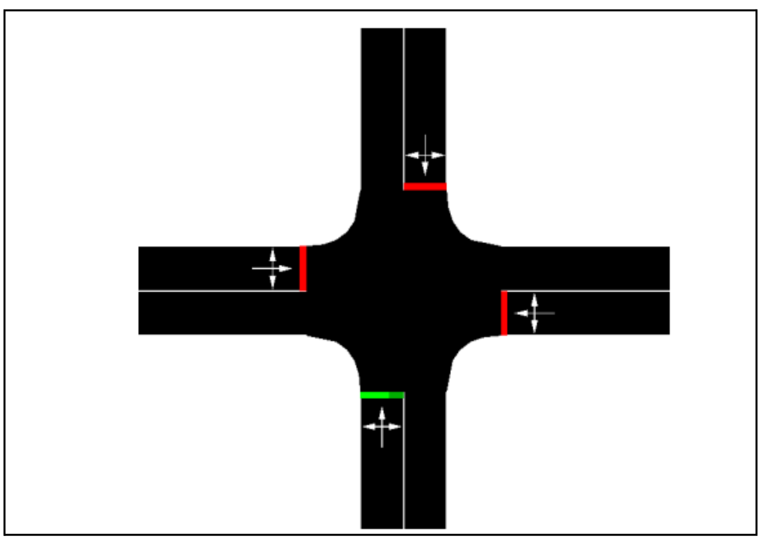

Fig 4.1: Four-way traffic lanes intersection.

The above figure is the pictorial representation of the problem statement provided to us. This report addresses these inefficiencies by proposing a dynamic and adaptive traffic signal control system tailored for a simple four-way traffic intersection. The system leverages Multi-Agent Reinforcement Learning (MARL) algorithms to optimize traffic flow based on real-time data inputs, such as traffic density, queue length, and the presence of emergency vehicles, detected using the YOLOv9 model.

Multi-agent reinforcement Learning, a branch of machine learning, allows multiple agents to learn optimal actions through interactions with their environment and each other [6]. By applying MARL algorithms, specifically a coordinated Q-learning approach, the proposed system can dynamically adjust signal timings to minimize overall waiting times and reduce congestion. The MARL-based controller prioritizes lanes with higher traffic densities and queue lengths under normal conditions and can swiftly adapt to prioritize lanes with emergency vehicles, thus ensuring a green light for those lanes. Waiting time will be used as the reward system for the MARL model, making it a key metric in evaluating the system's performance.

The proposed system architecture includes sensors and the YOLOv9 model to detect vehicles and collect real-time traffic data. This data is processed to analyze vehicle density, queue length, and the presence of emergency vehicles. A signal controller then executes the optimal signal timings determined by the MARL algorithm. Additionally, the system features a communication mechanism to detect emergency vehicles and adjust signal priorities accordingly. This comprehensive approach aims to enhance traffic flow efficiency at intersections, reduce waiting times, and ensure the prompt passage of emergency vehicles, demonstrating the potential of machine learning techniques in modern traffic management.

To evaluate and refine the proposed system, we utilize the Simulation of Urban Mobility (SUMO) simulator. SUMO is an open-source, highly portable, microscopic, and continuous multi-modal traffic simulation package designed to handle large road networks. By using SUMO, we can simulate complex traffic scenarios and assess the performance of our MARL-based traffic signal control system under various conditions. This enables us to validate the effectiveness of the proposed system in a controlled environment before deploying it in real-world intersections.

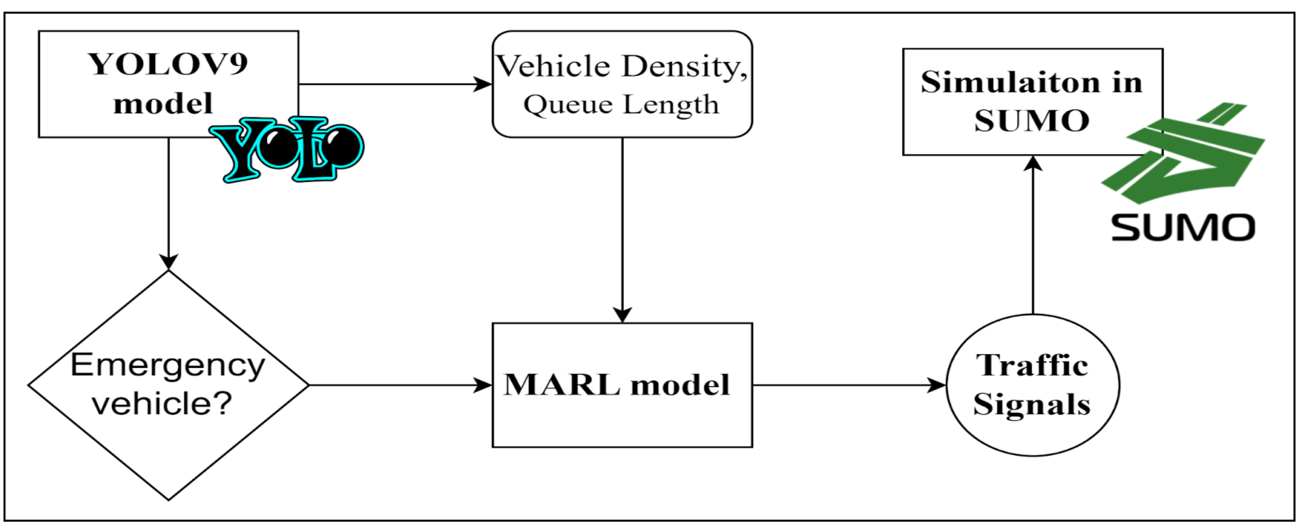

Fig 4.2: Block Diagram of the reinforcement learning process

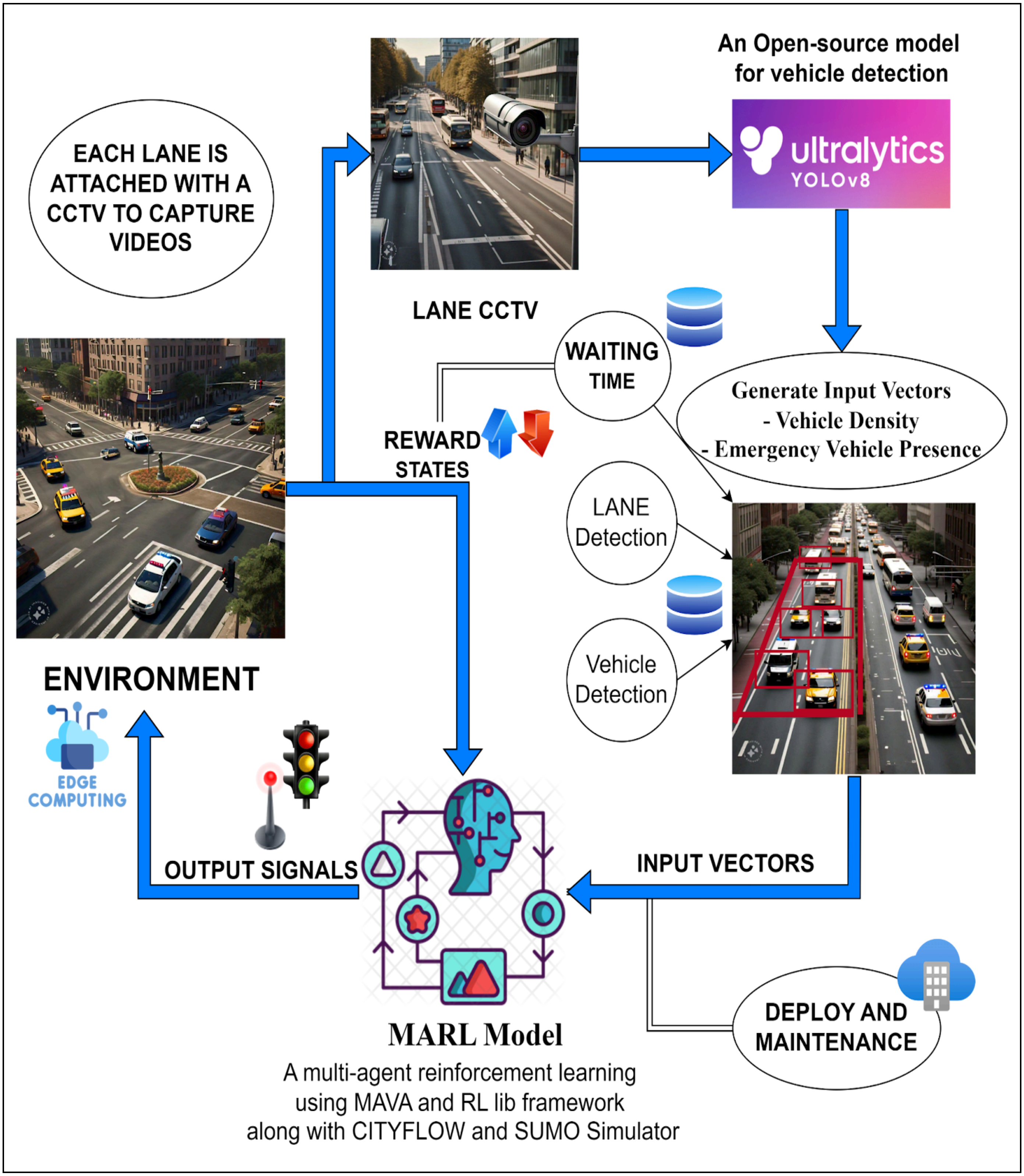

Fig 4.3: A pictorial representation of the Real-time Dynamic Traffic Signal Controller.

- **Literature Survey**

Sl. No	Title of the paper	Year of publication	Problem Statement	ML Technique used	Final Outcome
1	Deep Reinforcement Learning based Approach for Traffic Signal Control [1]	2021	A novel method for the TCS problem using AI (Reinforcement learning) is discussed. This paper proposes a new abstraction for the TSC problem: a combination of state representation, a brand new rewarding concept, and simplified action space.	1. RL training is formulated as a Markov Decision Process. 2. Control task is done using the Policy Gradient algorithm. 3. The training environment is modeled in SUMO.	The agent is tested in scenarios where one lane has traffic and the other doesn't; and the agent successfully provides more greenlight to the former lane - even though this a new scenario for it. Similarly, when both lanes had traffic, it alternated green light time between the lanes to keep the traffic symmetry.
2	Toward A Thousand Lights: Decentralized Deep Reinforcement Learning for Large-Scale Traffic Signal Control [2]	2020	In the paper, the problem of multi-intersection traffic signal control is tackled, especially for large-scale networks, based on RL techniques and transportation theories. MPLight is proposed as a solution.	1. For coordination, the concept of 'pressure' (derived from max pressure control theory) is used. 2. FRAP architecture is used as the base model. FRAP specially design a network architecture for learning the phase competition in traffic signal control problems. Following the base model, DQN	1. The proposed MPLight consistently outperforms all the other methods in the four different scenarios tested, leading to both the least travel time of passengers and the maximum throughput. 2. MPLight can handle traffic signal control for thousands of lights effectively and efficiently.

				is used to solve the multi-intersection signal control problem.	
3	Scheduling algorithm for adaptive traffic light under VANET scenario [3]	2022	In this paper, a scheduling algorithm based on the VANET scenario is proposed to analyze an isolated intersection as a case study. The proposed algorithm (Scheduling Traffic Light STL) aims to increase traffic flow by utilizing the excess green signal time to reduce the waiting time for mobile vehicles at the intersection.	A 'MaxGreen' value is set, and TrafficSpeed for each direction is read. 'DistanceGreenArea' (= MaxGreen/ TrafficSpeed) for each direction is calculated and all vehicles in the 'DistanceGreenArea' are read. Compute 'LastVehicleDistance' in each direction inside the 'MaxGreen' area and compute 'BestGreen' (=LastVehicleDistance / TrafficSpeed) for each direction. Schedule the current phase time by the lesser one among BestGreen and MaxGreen.	The total delay time of all vehicles in the traditional fixed algorithm is compared with the proposed STL algorithm. The results show a significant improvement in the delay time when the algorithm is implemented.
4	Emergency Vehicle Detection in Heavy Traffic using Deep ConvNet2D and Computer Vision [4]	2023	In this paper, a new method for detecting and classifying emergency cars is proposed. In the proposed system, the Deep ConvNet2D Algorithm is implemented instead of ANN. The proposed system is implemented in real-time and features high accuracy.	1. Data set is taken from Kaggle for the car. 2. CNN is used for model creation and training.	The comparison of accuracy between CNN and shows CNN has better accuracy (90%) compared to ANN (84%). So, it is better to use CNN.

5	Real-time Emergency Vehicle Event Detection Using Audio Data [5]	2022	In this paper, the focus is on detecting emergency vehicles using only audio data. Improved and quick detection can help in faster preemption of these vehicles at signalized intersections thereby reducing overall response time in case of emergencies. ELMs have been used in this work because of its simplicity and shorter run-time which can therefore be used for online learning.	The audio data is segmented into fixed lengths since the number of input features of the model needs to be consistent. Then the data is balanced using SMOTE. Two features were extracted such as MFCC and ZCR which were merged and then fed into ELM.	Five different models were trained alongside ELM such as K-Nearest Neighbors (KNN), Random Forest (RF), Support Vector Machine (SVM), Multilayer Perceptron (MLP) and CNN. In the time required by the models, ELM is 10 times better than the next-best performing model. The model with the highest here is ELM with an accuracy of 94.5%. For quick and reliable emergency vehicle siren sound detection, it is reasonable to use 10 or 100 neurons in an ELM.
6	Single Intersection Traffic Light Control by Multiagent Reinforcement Learning [6]	2022	In this paper, a traffic signal of an intersection is divided into four independent phases and then controlled by DQN(Revised QMIX) models respectively. Models can receive observations from their own angle of view instead of extracting features from the whole scene. We suppose that it is beneficial for learning better policy if agents could sense the environment more precisely.	Deploy QMIX to co-train the four DQN agents. Each agent receives the observation of its own phase The observation scope is 200 meters from the intersection. The global immediate reward is set as the difference between two vehicle queues of adjacent steps. If the queue at the current step is shorter than the last one, the immediate reward is positive and vv. The agents of the four phases share the same global reward but have their own	After 100 epochs of training, plot the average rewards of each episode, which clarifies the superior performance of our method (revised QMIX for cycle decomposition) compared to DQN (for the whole cycle). The rewards increase rapidly as training progresses. During the whole process, this method gets higher rewards than DQN.

				actions. Action=1 denotes that the light turns green and 0 vv. Add a constraint to TD error to prevent collision For execution, a method to compare the Q-value of action=1 of all agents and select the maximum Q-value's action is designed.	
7	Ambulance detection using image processing and neural networks [7]	2023	Ambulance Detection using Image Processing and Neural Networks is a vehicle detection and tracking system, which recognizes the ambulance amidst traffic congestion. Over the past few years, the range of vehicle usage of the road is growing each day that results in traffic congestion. Hence making it hard for the emergency vehicle to pass through the traffic at the earliest possible time.	The only piece of hardware used is the surveillance camera. YOLOv3 and CNN are used as they make work much more compatible with countries where road width is irrelevant for separate ambulances	Preparing a dataset for the program to train and assess was done, and it was used by the program to identify the Ambulance among all the vehicles. Upon sensing that the captured vehicle was an ambulance, we successfully got the program to turn that traffic light into green, while others were turned red.
8	Traffic sign recognition based on deep learning [8]	2022	In this paper, an experiment to evaluate the performance of the YOLOv5 is implemented, based on the custom dataset for Traffic Sign	YOLOv5 model and single shot multibox detector (SSD) are compared. The dataset of road signs was trained in both models and the	YOLOv5 model has an excellent visualization function in the result. The values for all the classes are all over 90.00%, which indicates the excellent TSR

			Recognition, which unfolds how the model for visual object recognition in deep learning is suitable for TSR through comparison with SSD as the objective of this paper.	TSR experiment was implemented in Google Colab.	performance of YOLOv5. For SSD, The accuracy of TSR in almost all classes reaches nearly 90.00%. Both YOLOv5 and SSD show good capabilities, but YOLOv5 performs a bit better.
9	A Comparative Study of Algorithms for Intelligent Traffic Signal Control [9]	2022	In this paper, methods have been explored to effectively optimize traffic signal control to minimize waiting times and queue lengths, thereby increasing traffic flow.	Round Robin Scheduler, a Feedback Control mechanism and two Reinforcement Learning techniques - Deep Q Network (DQN) and Advantage Actor-Critic (A2C) Are compared.	The results of testing the different traffic light agents in the three traffic scenarios for a time period of 5400 timesteps are compared. In RR, the agent performs well in uniform traffic scenarios and the queue length is above half only 33% of the time, but in non-uniform scenarios , it is above half ~80% of the time. The performance of MONOPOLY in uniform traffic scenarios displays a similarity in behavior to RR. However, under non-uniform traffic densities, the agent performs significantly better than the RR agent. The results of testing the DQN agent on the uniform traffic agent show no major improvement over RR and MONOPOLY. However, its performance in non-uniform traffic densities are improved. The plots of queue length

					of A2C over time were observed to be similar to DQN in all three scenarios. Though the queue lengths were the same, the training time was significantly lesser than DQN. The major advantage of A2C is the performance improvement achieved with lower training times, as the number of processes was increased.
10	Exploiting Stage Information for Prediction of Switching Times of Traffic Actuated Signals Using Machine Learning [10]	2022	This research explores using machine learning to improve traffic light switching time prediction for GLOSA systems. They test if including an estimated traffic light state (like a simplified picture of current signal combinations) as a feature improves prediction accuracy compared to just using signal data. They use XGBoost and Bayesian Networks to see if this additional information helps, and also investigate whether the actual next state or an estimated one works better, considering traffic patterns and volume.	They use two intersections with different traffic volumes and analyze data points including current signal state, time since last green light, detector data, and public transport information. The baseline model (R) uses only signal state and time information. An additional model (d(k)) incorporates the actual current traffic light stage based on the intersection's control system logic. To predict future stages, they use a combination of XGBoost (for predicting the next stage) and Bayesian Networks (for predicting the following two stages). The accuracy of these different models is	The compromised signal states and stages significantly enhance the prediction quality of switching time estimation of traffic lights. Whether the feature of the actual or following signal states respectively stage is best it depends on the individual signal group.

				then compared using statistical tests.	
11	How Well Do Reinforcement Learning Approaches Cope With Disruptions? The Case of Traffic Signal Control [11]	2023	This research paper investigates the effectiveness of various traffic signal control methods, particularly Reinforcement Learning (RL), in handling disruptions within traffic networks. They compare two Reinforcement Learning (RL) methods (PressLight and GuidedLight) against other approaches like Analytics (self-organizing) and Random control.	The algorithm being described is **Q-learning**, a type of Reinforcement Learning (RL) technique. Markov Decision Process (MDP), a mathematical framework, is also used.	**1. Performance under disruptions:** Both RL methods struggled in congested scenarios like NY48, likely because the training signal wasn't informative in such situations. Analytic+ performed well across all scenarios, likely due to its use of short-term traffic predictions and established traffic flow principles. **2. Pre-training effectiveness:** Pre-training RL agents on simpler scenarios (like 4x4 homogeneous traffic) before deploying them in complex situations (like NY48) significantly improved their performance compared to not pre-training. **3. Pre-training challenges:** The pre-training scenario needs to be challenging enough to be informative but not so congested that learning becomes meaningless. **4. Meta-learning potential:** The study suggests that RL methods, particularly the Deep Q-Learning-based GuidedLight, might have surprising meta-learning abilities. Even without explicit techniques,

					pre-training in one scenario allowed the RL agent to adapt to unseen situations in other scenarios. **Environmental benefit:** Pre-training on smaller scenarios reduces computational resources needed for training compared to complex scenarios.
12	Boosted Genetic Algorithm Using Machine Learning for Traffic Control Optimization [12]	2022	To optimize the traffic signal timings in signalized urban intersections, under non-recurrent traffic incidents.	The optimization framework proposed combines the strengths of Genetic Algorithms (GA) and Machine Learning (ML) to improve traffic network management. This research proposes a hybrid approach for traffic signal control. It combines a Genetic Algorithm (GA) with a machine learning model (XGBT) to optimize traffic light timing. The GA searches for the best signal timings (decision variable) to minimize travel time (objective function), while the XGBT model predicts travel time efficiently.	This BGA-ML approach outperforms the original GA, significantly reducing travel time, especially during traffic disruptions.

13	Dynamic Traffic Control System with Reinforcement Learning Technique [13]	2020	This research proposes a new Deep Reinforcement Learning (DRL) model for traffic light control that adapts to real-world traffic patterns. Unlike prior research focused on simulations, this model is trained on a massive real-world traffic dataset from cameras. The authors plan to further improve performance by incorporating a feature extraction algorithm, with the ultimate goal of creating a practical DRL system to optimize traffic flow and reduce congestion in real-world transportation networks.	The agent uses five features (queue length, waiting time, phase, and vehicle count) to represent traffic conditions and a reward function to measure how well an action improves traffic flow. **Q-Network with Phase Gates:** The agent uses a special Q-network architecture to learn the best action for each traffic light phase. This network separates features into two convolutional layers and then combines them with additional traffic data (queue length, waiting time, phase, and vehicle count). Then, for each light phase, the network uses a separate process to estimate the best action based on the reward it expects. A "phase gate" controls which process is active depending on the current phase, ensuring the network considers the specific phase when making decisions. **Feature Extraction:** The system uses	1. **Peak hour vs Non-peak hour:** On the given day, there is more traffic on WE bearing than SN for more often than not, during which a perfect traffic light control strategy is relied upon to provide a longer time for WE guidance. And during top hours (around 7:00, 9:30 and 18:00), the policies gained from our technique give a longer time for the green light on WE than non-top hours. In the early morning, the vehicle appearance rates on SN are larger than the rates on WE, and our strategy consequently gives a longer time to SN. 2. **Weekday versus Weekend:** The policy gives fewer green lights on WE (more green lights on SN) during weekend daytime than it gives on weekdays. This is on the grounds that there is more traffic on SN than on WE during weekend daytime.

				Speeded-Up Robust Features (SURF) object recognition to extract additional features from camera data. SURF involves identifying key points in an image (like vehicles) and their descriptions. This might be used for future improvements but isn't part of the core decision-making process in this research.	
14	Adaptive Traffic Signal Control to Reduce Delay Time at a Single Intersection Point [14]	2021	This research tackles traffic congestion by proposing two new, simple methods for adaptive traffic light control at intersections. These methods leverage vehicle-to-infrastructure communication to gather real-time traffic data and dynamically adjust light timing or sequence. The goal is to minimize average wait times. Simulations show significant improvement, with the best policy reducing wait times by an average of 80% for all vehicles and over 60% for specific turn lanes.	Three signal control policies (SEQ, SEQDUR, and SEQ5) are being used which are designed to manage traffic light sequences and durations at a four-way intersection to minimize vehicle waiting times. The policies rely on monitoring the current state of traffic queues and making decisions accordingly.	The study evaluates the performance of three proposed traffic signal control policies (SEQ, SEQDUR, SEQ5) in reducing average waiting times at a four-way intersection. These policies are compared against a base case with a fixed 30-second green light duration and a static sequence. **Base Case:** The fixed 30-second green light duration results in the highest average waiting times across all vehicle categories. **SEQ Policy:** Reduces overall average waiting time by 16%. **SEQDUR Policy:** Reduces overall average waiting time by 63%. **SEQ5 Policy:** Reduces overall average waiting

					time by 80%.
15	CycLight: learning traffic signal cooperation with a cycle-level strategy [15]	2024	This study introduces CycLight, a novel cycle-level deep RL approach for network-level adaptive traffic signal control system. CycLight adopts a cycle-level strategy, optimizing cycle length and splits simultaneously using Parameterized DQN algorithm.	PAMDP is utilized to accurately capture key traffic flow features while minimizing computational load. CycLight adopts the cycle-level TSC logic, leveraging PDQN agents to perform discrete-continuous hybrid actions. Given an environmental state, a hybrid action is obtained so as to interact with the environment. Then the collected transition sample is stored in the replay buffer, after which the policy learning is performed using the data sampled.	The average waiting time values is extracted the training curve, reveals that CycLight and Advance CycLight demonstrate standard deviations of 25.65 and 26.97, respectively. While single PDQN and MAADDPG exhibit standard deviations of 34.03 and 42.54, respectively. It is evident that both CycLight and Advance CycLight converges to the most optimal and stable policy with a narrow deviation range.
16	Deep Reinforcement Learning for Traffic Light Control in Intelligent Transportation Systems [16]	2023	In this paper, deep reinforcement learning to control traffic lights is investigated, and both theoretical analysis and numerical experiments show that the intelligent behavior “greenwave” emerges naturally from a grid road network, which is proved to be the optimal policy in an avenue with multiple cross streets.	Two DL algorithms are analyzed in different scenarios - In a single road intersection, the DQN algorithm is verified to deliver a thresholding policy; and in a grid road network, the deep deterministic policy gradient (DDPG) algorithm is adopted. Optimality for ‘greenwave’ policy in grid networks is also	In the comparisons of the policies derived by the DQN algorithm, the fixed-cycle policy, and the optimal policy, the result of the DQN policy coincides with that of the optimal policy. The training time using DQN is very short. As the car arrival/passing ratio increases, the queue length also increases. The policy learned by DQN shows better effects than the

				achieved.	fixed-cycle policy. DDPG demonstrates emergent "greenwave" patterns, showcasing its ability to learn beneficial traffic flow structures from passive observations.

- **Proposed Methodology**

The proposed dynamic traffic signal control system integrates YOLOv9 for vehicle detection, Multi-Agent Reinforcement Learning (MARL) for policy optimization, and SUMO for traffic simulation. YOLOv9, deployed at the intersection, continuously monitors traffic to detect vehicles and identify emergency vehicles, providing data on traffic density and queue length for each lane. This data is fed into MARL agents, which use Q-learning algorithms to develop optimal traffic signal policies. The implementation begins with deploying YOLOv9 to collect traffic data. This data is aggregated to form state representations for MARL agents, which model traffic scenarios to make informed decisions. Each lane is assigned an agent that learns optimal signal timings based on traffic conditions and emergency vehicle presence. The agents are trained in the SUMO simulation environment, refining their policies iteratively through continuous feedback.

Various scheduling strategies are evaluated, including Fixed-Time Control, Actuated Control, and Adaptive Control. A Q-learning-based Adaptive Control strategy is chosen for its real-time adaptability, efficiency, and automated prioritization of emergency vehicles. This strategy dynamically adjusts signal timings based on real-time traffic conditions to reduce waiting times and congestion. The trained MARL agents and their policies are implemented in the SUMO simulation environment, where they are tested against traditional traffic signal control methods under various traffic scenarios. Performance metrics such as overall waiting time, queue length, and response time for emergency vehicles are measured to ensure the effectiveness of the proposed system. Finally, the MARL-based traffic signal control system is deployed at the actual intersection. Continuous monitoring and real-time adjustments are made based on feedback, ensuring the system remains efficient and responsive. Real-time data is collected to evaluate the system's performance, and necessary adjustments are made to enhance traffic management and safety, thus optimizing traffic flow and reducing overall waiting time. This comprehensive approach ensures that the system is robust, adaptable, and capable of handling dynamic traffic conditions effectively.

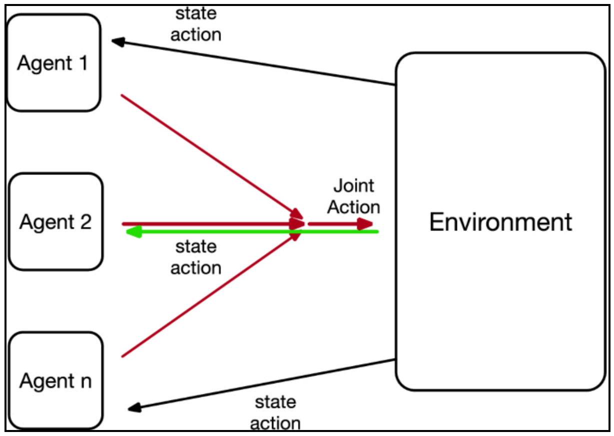

Fig 4.3: A pictorial representation of MARL

- **System Architecture**

The proposed dynamic traffic signal control system is composed of three main components: YOLOv9 for vehicle detection, Multi-Agent Reinforcement Learning (MARL) for policy optimization, and SUMO for traffic simulation. The YOLOv9 model is deployed at the intersection to continuously monitor traffic, detect vehicles, and identify emergency vehicles. The output from YOLOv9 includes traffic density and queue length for each lane, which serves as crucial input for the MARL agents.

The MARL system uses the inputs from YOLOv9 to learn optimal traffic signal policies [6]. Each lane is assigned an individual agent responsible for determining the best signal timings based on traffic density, queue length, and emergency vehicle presence. These agents are trained using Q-learning algorithms within a simulated environment provided by SUMO, which allows for the evaluation and refinement of traffic signal strategies.

- **Implementation Steps**

The implementation process begins with deploying the YOLOv9 model at the intersection for continuous traffic monitoring and data collection. The collected data on vehicle density, queue length, and emergency vehicle presence is then aggregated to form state representations for the MARL agents. These state representations are crucial for the MARL system to accurately model the traffic scenario and make informed decisions.

In the MARL phase, agents are designed for each lane to learn optimal traffic signal policies. The state space is defined as a combination of traffic density, queue length, and emergency vehicle presence, while the action space includes possible traffic signal timings such as green, yellow, and red durations. A reward function is designed to minimize overall waiting time, prioritize emergency vehicles, and ensure fair traffic distribution. The agents are trained using Q-learning algorithms in the SUMO simulation environment, where they iteratively update their policies based on feedback.

- **Scheduling Strategies and Algorithm Selection**

Various scheduling strategies are evaluated, including Fixed-Time Control, Actuated Control, and Adaptive Control. A Q-learning-based Adaptive Control strategy is selected for its ability to dynamically adjust signal timings in real-time, offering flexibility, efficiency, and effective prioritization of emergency vehicles. The benefits of Q-learning include its adaptability to changing traffic patterns, reduction in overall waiting time and congestion, and automated prioritization of emergency vehicles [3].

- **System Simulation and Evaluation**

The trained MARL agents and their policies are implemented in the SUMO simulation environment to evaluate performance under various traffic scenarios. Performance metrics such as overall waiting time, queue length, and response time for emergency vehicles are measured and compared against traditional traffic signal control methods. This step ensures that the proposed system is effective and can adapt to real-world conditions.

Finally, the MARL-based traffic signal control system is deployed at the actual intersection. Continuous monitoring and adjustments are made based on real-world feedback, ensuring the system remains efficient and responsive. Real-time data is collected to evaluate the system's performance, and necessary adjustments are made to improve traffic management and safety.

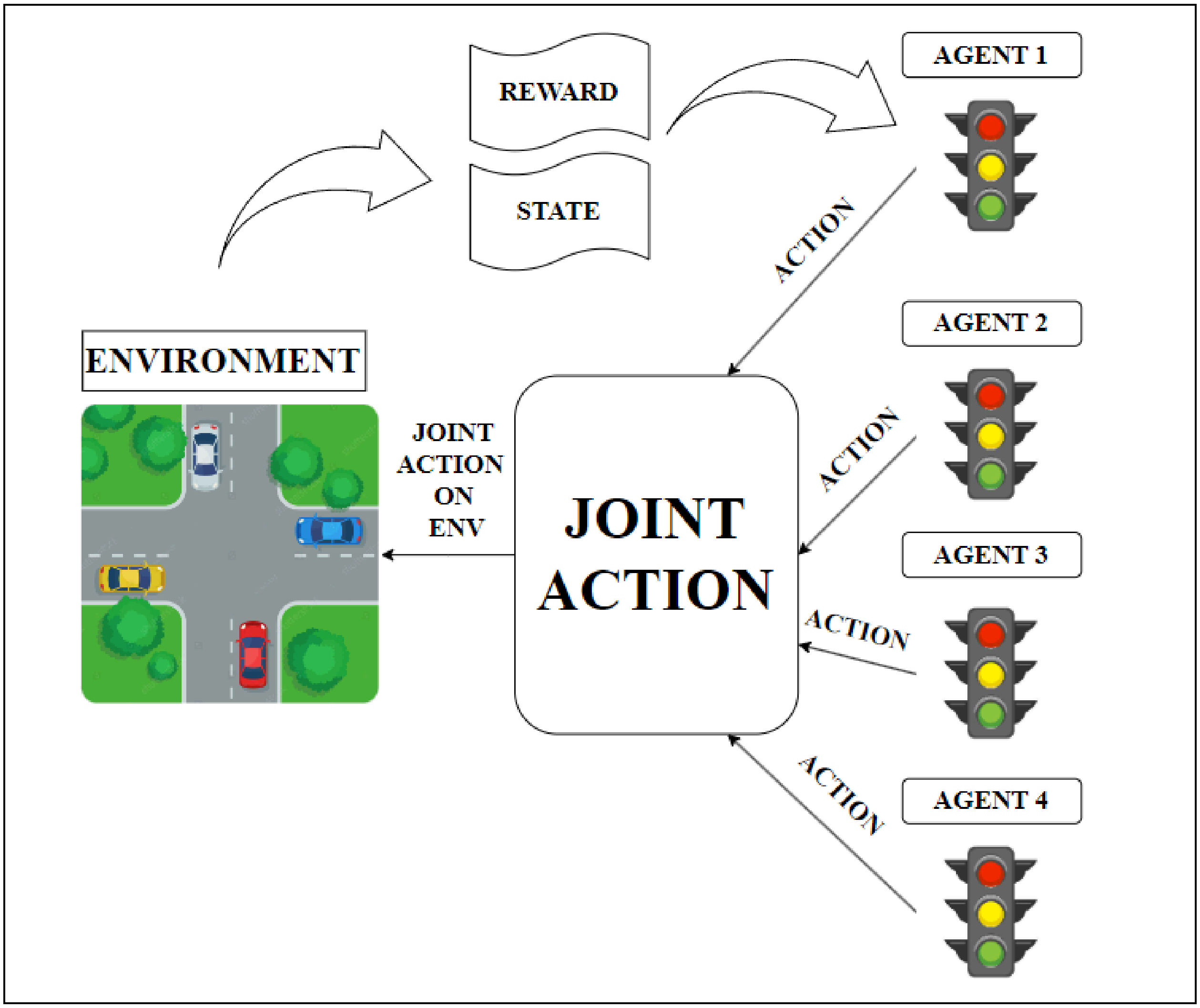

Fig 4.4: A pictorial representation of the complete process for TSC (Traffic signal controller).

In our dynamic traffic signal control system, the action-state-reward framework serves as the backbone for optimizing traffic flow and efficiently managing intersections. This framework, rooted in reinforcement learning principles, enables our system to make informed decisions in real time.

Action: Our system employs the Multi-Agent Reinforcement Learning (MARL) algorithm to determine actions based on the current intersection state. These actions primarily involve adjusting the duration of green signals for each lane. For instance, MARL might extend the green signal for lanes with high traffic density or prioritize emergency vehicles by immediately granting them passage.

State: The state of the intersection encapsulates crucial traffic conditions derived from YOLOv9 vehicle detection:

- **Traffic Density**: Measured through real-time analysis of vehicle presence in each lane.
- **Queue Length**: Estimated based on the number of vehicles waiting at any given time.
- **Emergency Vehicle Presence**: Detected promptly by YOLOv9, triggering an immediate state change to prioritize these vehicles.

Utilizing these state variables, MARL dynamically evaluates the current traffic scenario and determines optimal actions to enhance traffic flow and minimize congestion.

Reward: After each signal cycle, our system evaluates its performance using key metrics such as waiting times, queue lengths per lane, and emergency vehicle response times. Positive rewards are assigned for reducing waiting times, efficiently managing queues, and prioritizing emergency vehicles promptly. Conversely, negative rewards may occur for prolonged waiting times or inefficiencies in queue management.

This feedback loop allows our system to learn and improve continuously, adapting its strategies based on past outcomes to optimize future traffic control decisions effectively. By integrating these components, our traffic signal control system not only enhances operational efficiency but also lays a foundation for adaptive and responsive urban traffic management solutions.

Implementing the action-state-reward system in your dynamic traffic signal control system provides significant advantages. Firstly, it enhances **adaptability** by continuously monitoring and responding to real-time traffic conditions. This allows the system to dynamically adjust signal timings, effectively minimizing congestion and maximizing traffic flow at intersections. Secondly, the system improves **efficiency** by strategically prioritizing actions based on factors such as traffic density, queue lengths, and the presence of emergency vehicles. By allocating green signal durations accordingly, it optimizes intersection management to ensure smoother traffic flow. [11]

Moreover, the **learning and improvement** capabilities of the system are crucial. Utilizing feedback loops and reinforcement learning techniques, such as MARL, enables the system to evolve. By analyzing past outcomes and performance metrics like waiting times and emergency vehicle response, MARL refines its decision-making processes. This iterative learning not only enhances current traffic control strategies but also prepares the system to handle future traffic scenarios more effectively. Together, these features make the action-state-reward system a powerful tool for creating adaptive, efficient, and continually improving traffic signal control solutions.

Algorithm: Dynamic Traffic Signal Control Using MARL

1. Initialize:

- YOLOv9 for vehicle detection.
- MARL agents for each lane.
- Initial traffic signal timings.

2. Loop (continuous operation):

a. Data Collection:
  - Detect and count vehicles per lane using YOLOv9.
  - Identify emergency vehicles.

b. State Representation:
  - Calculate traffic density and queue length for each lane.
  - Update state variables (traffic density, queue length, emergency vehicle presence).

c. Action Selection (MARL):
  - MARL agents decide green signal durations based on the current state.
  - Prioritize lanes with emergency vehicles.

d. Signal Adjustment:
  - Adjust signal timings per MARL decisions.

e. Execute Actions:
  - Implement adjusted signal timings.

f. Reward Calculation:
  - Measure total waiting time and queue length per lane.
  - Measure emergency vehicle response time.
  - Generate rewards:
    - Positive for reduced waiting times and efficient queue management.
    - Positive for prompt emergency vehicle prioritization.
    - Negative for increased waiting times or inefficient management.

g. Learning and Improvement:
  - Update MARL policies based on reward signals.
  - Refine strategies using traffic outcome feedback.

3. End Loop (continuous real-time operation).

Table 4.1: Algorithm for dynamic TSC using MARL.

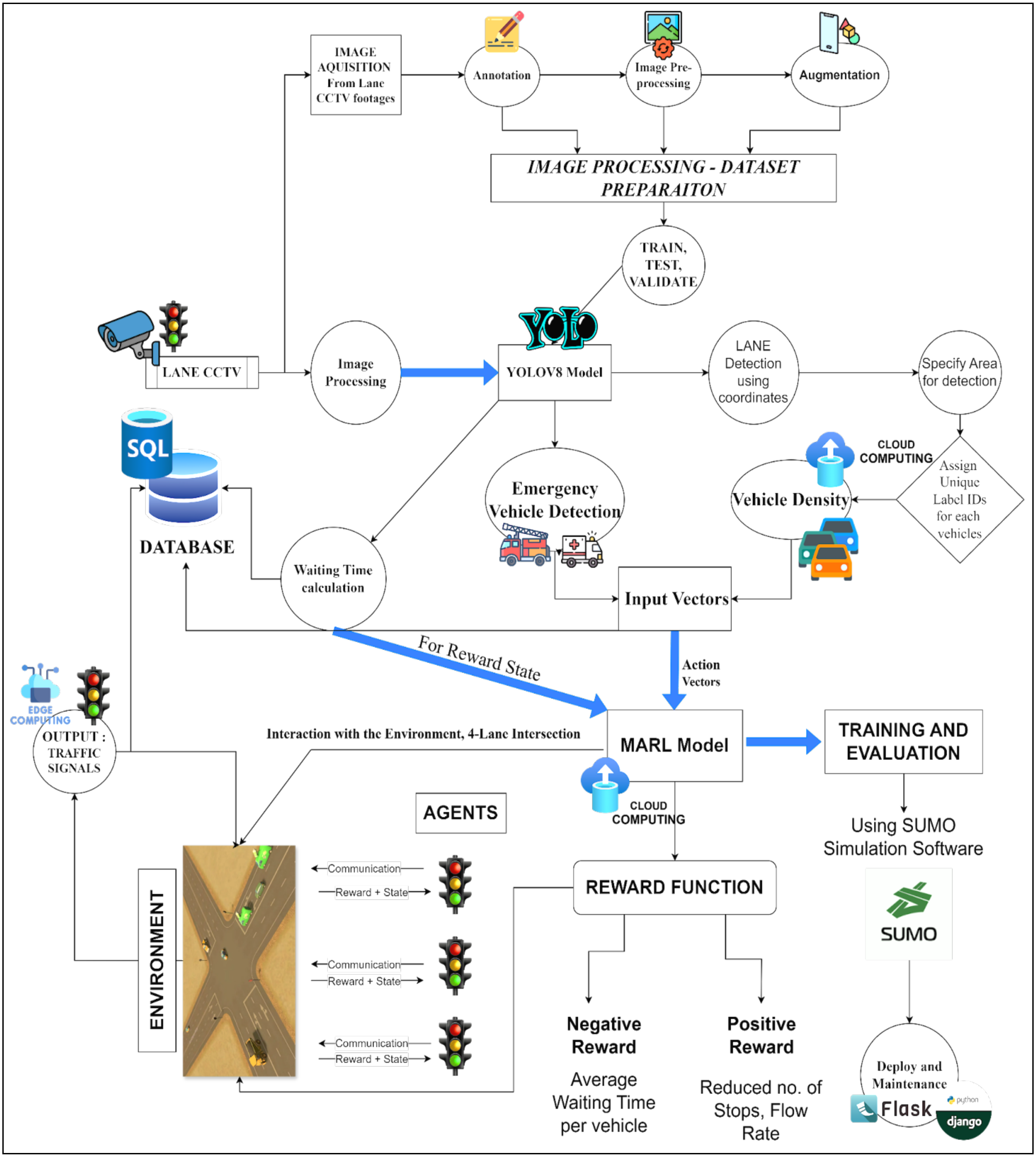

Fig 4.5: A detailed Flow analysis of the proposed Real-time dynamic TSC.

- ## Compared Methods

Multi-agent reinforcement Learning (MARL) distinguishes itself from fuzzy logic, deep Q-learning, and genetic algorithms by its ability to handle complex environments with multiple autonomous agents [9]. Unlike fuzzy logic, which relies on linguistic rules and lacks adaptability in dynamic scenarios, MARL learns optimal behaviors through interactions with the environment and other agents, achieving robustness and scalability. Compared to deep Q-learning, MARL excels in decentralized decision-making by coordinating multiple agents' actions simultaneously, leveraging joint action policies to maximize collective rewards. Genetic algorithms, while effective for optimization, may struggle with high-dimensional action spaces and coordination among agents, where MARL's decentralized execution and communication protocols shine. MARL's capacity to learn collaborative strategies through interactions and competition between agents represents a significant advancement in tackling real-world challenges requiring adaptive and cooperative behaviours. [11]

The below table shows the comparative analysis of different algorithms and MARL.

Table 4.2: Comparison of different techniques with MARL [9]

ML Technique	Focus	Comparison to MARL
Decentralized RL	Large-scale traffic networks	Similar to ours in using RL, but uses pressure concept and FRAP architecture. MARL can potentially learn more complex coordination strategies.
Rule-based algorithm	Isolated intersection, VANET scenario	A simpler approach, with limited adaptability. MARL can learn optimal strategies based on real-time traffic conditions.
Deep Q-learning with Fuzzy Logic	Real-time control based on traffic flow	Combines fuzzy logic rules with deep learning. MARL focuses solely on learning through reinforcement, potentially adapting better to changing traffic patterns.

- **Novelty of the Solution**

Our approach breaks new ground in traffic management by combining several innovative elements:

1. **Multi-Agent Reinforcement Learning (MARL) for Lane Traffic Analysis:** We introduce a framework where individual vehicles act as agents, learning optimal lane-changing and speed control strategies based on real-time conditions. This enables dynamic lane management, optimizing traffic flow across all lanes.
2. **Synergy with Secure Blockchain Communication:** Blockchain technology safeguards the integrity of data exchanged between vehicles and infrastructure. This fosters trust and collaboration within the MARL framework:
3. **Holistic Traffic Management:** This combined approach addresses a wider range of traffic challenges:
    - **Reduced Congestion:** MARL optimizes lane usage for smoother traffic flow.
    - **Improved Efficiency:** Collaborative lane changes and coordinated speeds minimize bottlenecks.
    - **Enhanced Safety:** Secure communication promotes safer driving behavior.
    - **Emergency Vehicle Prioritization:** The framework can prioritize emergency vehicles for faster passage.

The proposed dynamic traffic signal control system represents a significant advancement in traffic management by integrating multi-agent reinforcement learning (MARL) with Q-learning scheduling algorithms and utilizing YOLOv9 for real-time vehicle detection. This approach stands out for its innovative use of MARL, which allows each traffic signal to operate as an independent agent capable of learning optimal policies through interaction with the environment. The Q-learning algorithm further enhances this system by providing a robust method for scheduling and adjusting traffic signal timings based on learned policies. By feeding real-time data on vehicle density, queue length, and emergency vehicle detection from the YOLOv9 model into this MARL framework, the system dynamically adjusts traffic signals to optimize flow, reduce congestion, and prioritize emergency vehicles effectively.

The lane prioritization in this system is particularly effective due to the precise and real-time data provided by YOLOv9, which enables the detection of vehicle types and traffic conditions with high accuracy. The incorporation of Q-learning ensures that the system continuously improves its decision-making process by learning from historical traffic patterns and current conditions. This results in a more responsive and adaptive traffic signal control, significantly reducing overall waiting time and queue lengths and ensuring that emergency vehicles can navigate intersections with minimal delay. The use of SUMO for simulation validates the efficacy of this approach, demonstrating its potential to improve urban traffic management substantially through intelligent and adaptive control mechanisms.

- **References**

[1] Kovari Balint, Tettamanti Tamas, Becsi Tamas - "Deep Reinforcement Learning based approach for Traffic Signal Control", 2021

[2] Chacha Chen, Hua Wei, Nan Xu, Guanjie Zheng, Ming Yang, Yuanhao Xiong, Kai Xu, Zhenhui Li - "Toward A Thousand Lights: Decentralized Deep Reinforcement Learning for Large-Scale Traffic Signal Control" , 2020

[3]Belal R. Alwan, Muntaser A. Salman - "Scheduling Algorithm For Adaptive Traffic Light Under VANET Scenario" , 2022

[4]Shankar K, Madhavan V, Muhammad Tariq P R, Sanjay Subramanian K - "Emergency Vehicle Detection in Heavy Traffic using Deep ConvNet2D and Computer Vision", 2023

[5] Zubayer Islam, Mohamed Abdel-Aty - "Real-time Emergency Vehicle Event Detection Using Audio Data" , 2022

[6] Tongchun Du, Bo Wang, Liangchen Hu - "Single Intersection Traffic Light Control by Multi-agent Reinforcement Learning", 2022

[7] K Agrawal, M K Nigam, S Bhattacharya, Sumathi G - "Ambulance detection using image processing and neural networks, 2021

[8] Yanzhao Zhu, Wei Qi Yan - "Traffic sign recognition based on deep learning", 2022

[9] Hrishit Chaudhuri, Vibha Masti, Vishruth Veerendranath, S Natarajan - "A Comparative Study of Algorithms for Intelligent Traffic Signal Control", 2022

[10] Lena Elisa Scheegans, Kevin Heckmann, Robert Hoyer - "Exploiting Stage Information for Prediction of Switching Times of Traffic Actuated Signals Using Machine Learning" , 2022

[11] Marcin Korecki, Damian Dailisan, Dirk Helbing - "How Well Do Reinforcement Learning Approaches Cope With Disruptions? The Case of Traffic Signal Control" , 2023

[12] Tuo Mao, Adriana-Simona Mihaita, Fang Chen, Hai L. Vu - "Boosted Genetic Algorithm Using Machine Learning for Traffic Control Optimization", 2021

[13] V. Jyothirmai, N. K. Kameswara Rao - "Dynamic Traffic Control System with Reinforcement Learning Technique", 2020

[14] Cansu Cav, Nilgun Fescioglu Unver, Aslinur Colak - "Adaptive Traffic Signal Control to Reduce Delay Time at a Single Intersection Point", 2021

[15] Gengyue Han, Xiaohan Liu, Xianyue Peng, Hao Wang, Yu Han - "CycLight: learning traffic signal cooperation with a cycle-level strategy", 2024

[16] Xiao-Yang Liu, Ming Zhu, Sem Borst, Anwar Walid - "Deep Reinforcement Learning for Traffic Light Control in Intelligent Transportation Systems" , 2023