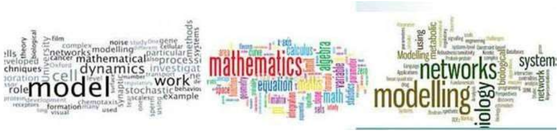

Journal, Advances in Mathematical & Computational Sciences
Vol. 12 No. 14, 2024 Series
www.isteams.net/mathematics-computationaljournal

Journal of Advances in Mathematical & Computational Sciences
An International Pan-African Multidisciplinary Journal of the SMART Research Group
International Centre for IT & Development (ICITD) USA

Available online at https://www.isteams.net/ mathematics-computationaljournal.info
CrossREF Member Listing - https://www.crossref.org/06members/50go-live.html

Deep Reinforcement Learning in Real Time Traffic Control Systems: A Constructive Review

[1]Ogbeide Oluwatosin Lara, [2]Akingbesote Boluwatife, [3]Olabode Olatunbosun, +Agbonifo Oluwatoyin & [5]Aliyu Ednad Olubunmi

[1] Department of Computer Science Adekunle Ajasin University, Akungba Akoko, Nigeria
[2] Department of Computer Science Adekunle Ajasin University, Akungba Akoko, Nigeria
[3] School of Computing, Federal University of Technology, Akure, Nigeria
[4] School of Computing, Federal University of Technology, Akure, Nigeria
[5]Department of Computer Science Adekunle Ajasin University, Akungba Akoko, Nigeria

ABSTRACT

Traffic congestion is a major factor currently affecting the global economy thereby costing billions of dollars in terms of the time lost in traffic, productivity and fuel wastage. Although, various researchers have made concerted effort to proffer solution to the problems of traffic congestion, for example, recent reinforcement learning research has also proposed multiple possible solutions. However, based on the literature, most of these reinforcement learning traffic control agents do not possess the real-time traffic control capabilities to efficiently deal with the complex and ever-changing nature of traffic on an average urban city road. This paper proposes a constructive review of Deep Reinforcement learning in Real Time Intelligent Traffic Control Systems. The concept is to provide clear understanding on deep reinforcement learning (DRL) traffic control agent, as well as discussing on the existing research and challenges. From our review evaluation, we observed that most of the existing studies on it have been based on off-policy DRL agents, However, there is need for a systematic on- policy DRL strategy towards addressing the shortcomings of the off- policy DRL agent for future work to prove the effectiveness of adopting on- policy DRL agents with optimized experience replay in intelligent traffic control will be carried out.

Keywords: Deep Reinforcement Learning, Real Time, Traffic Control Systems, Deep Deterministic Policy Gradient

Ogbeide O.L. Akingbesote, B., Olabode, O., Agbonifo, O. & 5Aliyu, E.O. (2024) Deep Reinforcement Learning in Real Time Traffic Control Systems: A Constructive Review.. Journal of Advances in Mathematical & Computational Science. Vol. 12, No. 1. Pp 1-14. dx.doi.org/10.22624/AIMS/MATHS/V12N1P1. Available online at www.isteams.net/mathematics-computationaljournal.

1. INTRODUCTION

An efficient traffic control system is a major pivot around which the economy of a city or country revolves. Traffic control systems are the most visible element of the urban infrastructure (Mcshane, 1999). This is because, all business transactions depends on transportation. Also, the modern businesses, industries, trades and general activities depend on transport and transport infrastructure. Movement of goods and services from place to place are becoming vital and inseparable aspects of global and urban economic survival(Oni, 2008)(Vanderschuren, 2006).

The rapid growth in the use of vehicles has led to overload urban roads causing traffic congestion. This creates an urgent need to operate our transportation systems with maximum efficiency. Man, nations, regions and the world would be severely affected in terms of all forms of development without an efficient traffic control system, which is a key factor for physical and economic growth(Azimian, 2011). Traffic congestion has become a major menace on urban roads as a result of inefficient traffic control systems.

Traffic congestion on an intersection may be caused by various factors like indiscipline drivers who show little respect for other road users. The common behavior in some cities is that many drivers try to cut a few seconds in order to reduce their journey times by forcing their way into intersections and by doing this they block the passage of other motorists. Also, traffic congestion maybe caused by the drivers' ignorance of the prevailing traffic conditions. The driver may not have knowledge of a particular routes traffic condition before entering into it and he may then decide to use another route in order to avoid being delayed. Another factor that may lead to traffic congestion is drivers neglecting traffic signal. There are three main traffic control signals, Red, yellow and Green. In most countries, the order in which traffic light changes is as follows:

- Red light on: indicates that the drivers should stop.
- Green light on: This indicates that the driver can start driving or keep driving.
- Yellow light on: drivers should stop when it is safe to, because the light is about to turn red.

Traffic Control Signals determines traffic phase in order to manage the flow of traffic and reduce congestion. Traffic control signals are devices placed along, beside, or above a roadway to regulate the flow of traffic in order to guide or warn road users, which includes motor vehicles, motorcycles, bicycles, and pedestrians. A traffic signal controller allocates right-of-way at an intersection through a sequence of green signals(Al-mudhaffar, 2006).

There are different types of traffic signal control. These include fixed time signals and Vehicle Actuated control. With fixed time signals, the green and cycle times are predetermined and have fixed duration. Computations of delay, which were carried out for a variety of flows, saturation flows and signal settings. Based on this, a formula was deduced for the average of any signal on an approach to an intersection as reported byWebster (1966).

The theoretical fact is that:

$$d = \frac{c(1-\lambda)^2}{2(1-\lambda x)} + \frac{x^2}{2q(1-x)} - 0.65\left(\frac{c}{q^2}\right)^{1/3} * x^{(2-5\lambda)} \quad (1)$$

where:
d = Average delay per vehicle on the particular arm
q = Traffic flow
c = Cycle time
g = Effective green time
λ = Proportion of the cycle which is effectively green (i.e. g/c)
x = The degree of saturation (i.e. *q/*λs) (s = saturation flow)

Vehicle-actuated control uses information on current demands and operations. This is obtained from detectors within the intersection, to alter one or more aspects of the signal timing on a cycle(Mathew, 2014). Vehicle Actuated Signals by-cycle basis. Vehicle actuated control systems have also been used in systems like GLIDE(Burov et al., 2021), SCOOT, and SCAT(Stevanovic et al., 2009). Research has been undertaken to investigate Traffic signal control systems that can optimize traffic signal scheduling and timing, such as adjusting traffic phase splits, in order to ameliorate traffic congestions at moderately and heavily trafficked single or multiple intersections (Rasheed et al., 2020).

According to a number of recent studies, For example, (Clark, 2019)(Arnott & Small, 2014), they acknowledged that traffic congestion is currently costing the global economy billions of dollars in terms of the time lost in traffic, lost productivity and fuel wastage. Moving slowly than free-flow speeds wastes time and fuel, and that these time and fuel costs multiplied over millions of travelers in large urban areas add up to billions of dollars in congestion costs.

The issue of traffic congestion is not only affecting African countries, for example, the U.S. has an extremely busy transportation system that relied on personal vehicles to travel from work and home(Chris, 2019). The scenario is also the same in most urban areas around the world. The congestion growth trend in United States of America from the year 1982 to 2011 is shown in Figure 1.1.

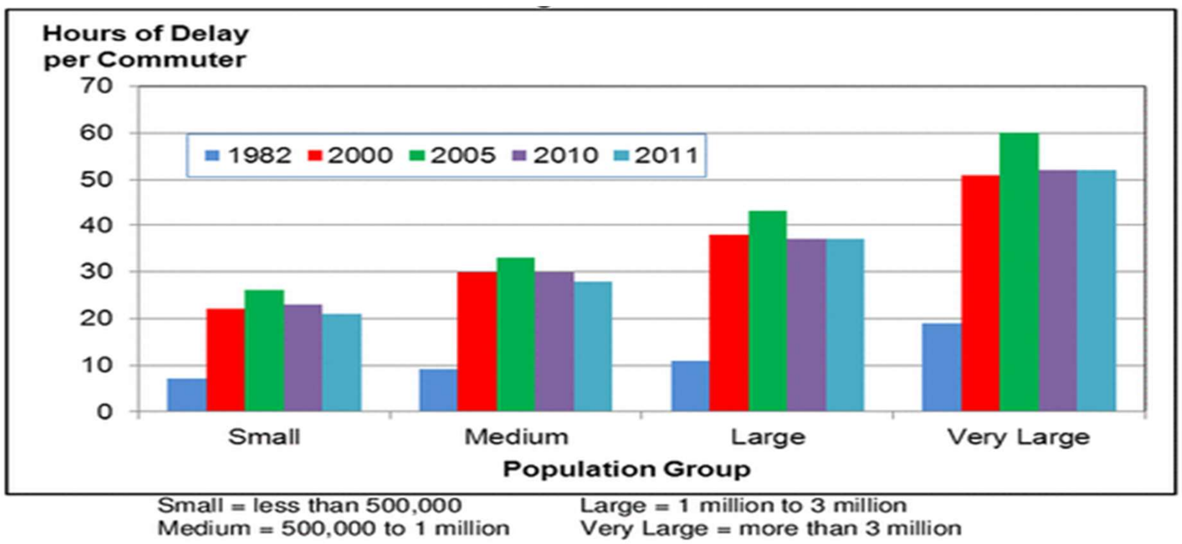

Figure 1.0: US traffic congestion growth trend(Source: (Chris, 2019))

There is obviously increase in traffic congestion growth between that period caused by rapid increase in the population of commuters. In recent years, the work of (Cheng et al., 2020)(Cao et al., 2017) show that the United States of America have adopted intelligent traffic control system as an efficient solution to traffic congestion. The U.S. Department of Transportation's Research and Innovative Technology Administration now offers cities and states in America with tools, training resources and other support to help get their intelligent traffic control systems up and running smoothly. However, traffic control has been quite poor in many developing countries among which Nigeria is inclusive despite the growth in transport demand and supply.

This is due to the fact that these countries still use traffic control systems which cannot efficiently deal with the complexity of the current situation on roads. This is because the current situation on major commercial roads is quite complex considering the volume of commuters on our major roads especially during rush hours. Therefore, there is need for intelligent traffic control system. This paper presents a constructive review of Real Time Intelligent Traffic Control Systems as well as discussing on the existing research and challenges. In addition, the paper provides clear understanding on how deep reinforcement learning (DRL) traffic control agent can be used to approach the solution. in such way to produce the real time effect,

The organization of the paper is as follows. Section 2 discusses the background of intelligent traffic control system and the perspectives of different scholars in that area. In Section 3, emphasizes on Deep reinforcement learning agent in intelligent traffic control and the paper ends with conclusions in section 4.

2. RECENT APPROACHES IN INTELLIGENT TRAFFIC CONTROL SYSTEM.

Intelligent Traffic Control System allows people to get more from transport networks, with greater safety and with less impact on the environment. Quite a number of researchers have propounded theories and made contributions on the application of intelligent traffic control systems in different countries around the world. Scholars around the world have viewed and explained Intelligent Traffic Control Systems from different perspectives. In (Ajay and Chandra, 2009), the authors developed an intelligent traffic control agent using radio frequency identification (RFID). The system was able to improve traffic flow and safety, it was fully automated saving costly of constant human involvement. Also, (Yousef & Shatnawi, 2013) simulated and evaluated an Intelligent traffic light flow control system.

Wireless Sensor network was used to control traffic signals. Also, an intelligent traffic controller is developed to control the operation of the traffic infrastructure supported by the WSN. The controller embodies traffic system communication algorithm (TSCA) and the traffic signals time manipulation algorithm (TSTMA). (Oladipo, 2015) developed a hybrid methodology, the Structured Systems Analysis and Design Methodology (SSADM) and the Fuzzy-Logic based Design Methodology to monitor and control road traffic in a Nigerian city. It was observed that the fuzzy logic control system provided better performance in terms of total waiting time as well as total moving time of the vehicles. All these authors have made one or more contributions to the existing knowledge. However, most of the aforementioned approaches could only control traffic to some extent; they couldn't perform efficiently in an extremely congested and unpredictable traffic(Koukol et al., 2015). Also, the large-scale implementation of RFID-based vehicular traffic management in smart and connected communities requires a stable power supply. This makes it unaffordable and expensive to implement in developing communities in the world(Matthews et al., 2017).

These traffic control systems were classified under the umbrella of Intelligent traffic control systems. The question is, are they really intelligent systems? What is an intelligent system? *Intelligent systems* are *systems* that can perceive, create action, and learn in an autonomous fashion, i.e., without external supervisory intervention for an extended amount of time (Rosenblat and Kneese, 2014). These intelligent traffic control systems mentioned earlier using RFID, WSN and Fuzzy logic cannot function without external supervisory intervention for an extended period of time. For any intelligent traffic control system to be able to cope with increasing present day traffic demand, it must be able to perceive its environment, respond automatically to its perception of its environment and save information on the past states of its environment in order to learn from its past experience and interact better with its environment in the future, this is where the deep reinforcement learning agent comes in.

The work of (Wei, 2018)(Wang et al., 2019)(Gustafsson, 2019)(Casas, 2017) emphasized on the application of reinforcement learning algorithms hand in hands with Deep Learning in traffic control. Deep learning (DL) is an aspect of AI based on neural networks with powerful generalizing ability to retrieve highly abstract structure or feature from the real environment (Arnold et al., 2019). Deep Reinforcement Learning (DRL) combines the perceiving ability of DL with the decision making ability of Reinforcement Learning (Zhao et al., 2017)(François-lavet et al., 2018).

Deep reinforcement learning (deep RL) is a subfield of machine learning that combines reinforcement learning (RL) and deep learning. Deep learning is learning from an image training set and then applying that learning to a new data set, The DNN integrates feature extraction, and classification (or prediction) process in a single framework by using information-dense input datasets(Greguri and Vuji, 2020a). Convolutional Neural Network is the most used DNN in traffic control systems for Feature extraction and classification. CNN extract images by learning the basic shapes in the first layers and evolving to learn features of the images in the deeper layers, resulting in more accurate image classification. CNN takes input images as matrices of pixels as input and processes it through convolutional layers, this is illustrated in Figure 1.0.

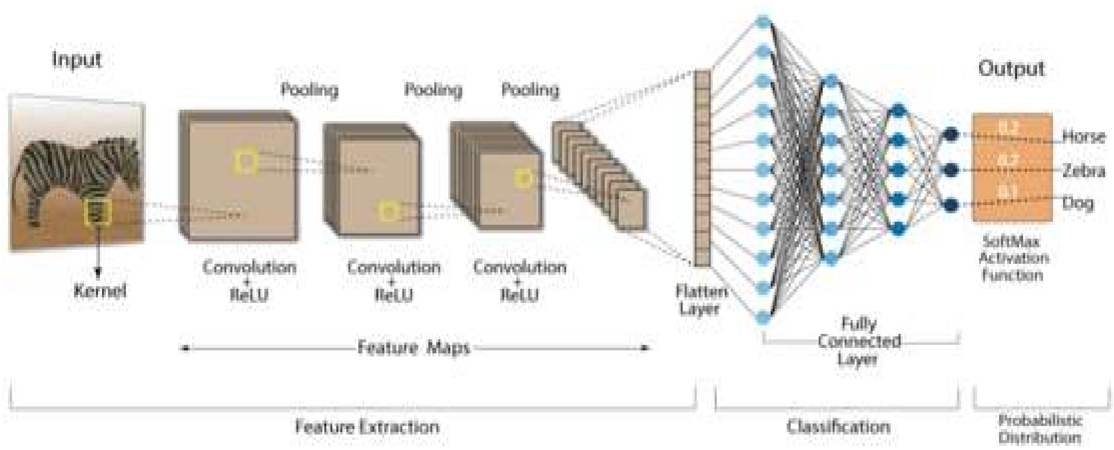

Fig 2: Image classification using deep learning and tensorflow (Source:(Yadav et al., 2022))

The image is processed in the convolutional layers then pooling layers extracts specific features from the image. This is done for several convolutional layers and pooling layers till we arrive at a flatten size of the image then the flattened image will then be input into the fully connected layer which finally classifies the images through various trained neurons. Reinforcement learning is dynamically learning by adjusting actions based in continuous feedback to maximize a reward. Reinforcement learning is learning how to map situations to actions that maximize a numerical reward signal.

The components of RL are the agent, the environment, state, action, reward and policy. Agent represents the "solution", that is a computer program which does the decision making (actions) and it solves complex decision-making problems under uncertainty. States are a representation of the traffic environment. In reinforcement learning, an agent needs to interact with the environment (either physical or simulated) by performing actions to obtain rewards. The agent's goal is to maximize its rewards and learns by adjusting its policy (the agent's strategy) based on the rewards. An RL algorithm can either be model free or model based. For model-based RL algorithm, the policy is based on the use of a machine learning models like random forest, gradient boost, neural networks, and others while for model-free RL algorithm ,the policy is guided by the use of non-ML algorithms.

There are three major model-free deep RL approaches which includes methods based on value functions, methods based on policy search, and methods which employ both value functions and policy search, namely actor-critic (AC) approach. More specifically, value function methods, for instance, Q-learning and state-action-reward-state-action (SARSA) algorithms, attempt to find a policy that maximizes the return by maintaining a set of estimates of expected returns for some policies (Zhao and Eskenazi, 2016). The agent's behavior is described by a function (policy) that maps any given perceived sequence to an action. The policy is the core of what the agent learns. A policy defines the way the agent behaves in a given time.

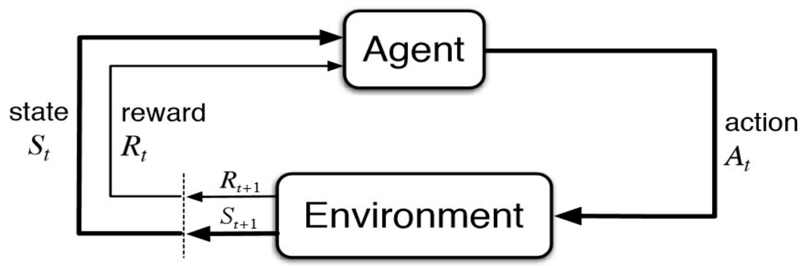

Figure 2.1: Reinforcement Learning Illustration (Source: (Barto, 2014))

An agent can be an off-policy reinforcement learning agent or an on-policy reinforcement learning agent. An off-policy RL. Off-policy learning allows the use of older samples (collected using the older policies) in the calculation. To update the policy, experiences are sampled from a buffer which comprises experiences and interactions that are collected from its own predecessor policies. In on-policy RL, the agent interacts with the environment to collect the samples and uses the latest learned policy.

3.0 DEEP REINFORCEMENT LEARNING AGENT IN INTELLIGENT TRAFFIC CONTROL.

Deep neural network empower Reinforcement learning directly deal with high dimensional states like images. A common example of DRL agent applied in traffic control is the Deep Q-Network(DQN) agent. Figure 1 shows the framework of a DQL agent. The traffic data from the traffic environment serves as input into the deep convolutional neural network and is trained using Q-learning with experience replay. DQN overcomes unstable learning through experience replay and target network. Experience replay stores experiences including state transitions, rewards and actions, which are necessary data to perform Q learning, and makes mini-batches to update neural networks.

The target network technique fixes parameters of target function and replaces them with the latest network every thousand steps. Considering dynamic characteristics of the actual traffic environment in most urban cities, deep reinforcement learning agent based traffic control approach can be applied to get optimal real-time traffic control. This is evident in the work of some researchers that modeled the performance of deep reinforcement learning traffic control agent. (Zhou et al., 2014), modeled the performance of an on-line SARSA(λ)-based traffic signal light optimization model that is capable of maintaining real time traffic signal timing control policy more effectively. However, SARSA(λ)-based is a general learning model for a long-term prediction of the dynamic system, this feature makes it difficult for SARSA(λ) agent to determine an optimal action in a real time traffic environment at a faster rate. Also, SARSA(λ) is sensitive to different models of reward and initialization.

In some cases the asymptotic performance can be significantly reduced(Grzes and Kudenko, 2008). In (Casas, 2017), the author(s) proposed an Urban traffic light control system applying Deep Deterministic Policy Gradient (DDPG) algorithm to handle large input spaces. The result shows that DDPG was able to better scale larger networks than classical tabular approaches like Q-learning. However, the DDPG agent was only able to address the curse of dimensionality regarding the traffic light control domain partially (Greguri and Vuji, 2020b)(Hausknecht et al., 2016). Also, off-policy deep reinforcement learning algorithms are computationally expensive(Shin and Kang, 2019)(Fujimoto et al., 2019). In (Gao et al., 2017), the author proposes an adaptive traffic signal control system using a deep reinforcement learning algorithm that automatically extracts all useful features (machine-crafted features) from raw real-time traffic data and learns the optimal policy for adaptive traffic signal control.

Experience replay and target network mechanisms were adopted to improve the algorithms stability. The simulation results showed that the algorithm reduced vehicle delay by up to 47% and 86% when compared to another two popular traffic signal control algorithms, longest queue first algorithm and fixed time control algorithm, respectively. However, the complexity of the system serves as a major loophole(Xu et al., 2018). Perez-murueta et al., (2019) developed a vehicle redirection system to avoid congestion that uses a model based on deep learning to predict the future state of the traffic network. The model uses the information obtained from the previous step to determine the zones with possible congestion, and redirects the vehicles that are about to cross them. Alternative routes are generated using the entropy-balanced k Shortest Path algorithm (EBkSP).

The system used information obtained in real time by a set of probe cars to detect non-recurrent congestion. Prediction and redirection however is not the efficient solution in reduction of traffic congestion. Redirecting vehicles will only lead to traffic congestion at other points in the traffic (Greguri and Vuji, 2020b). The work of (Kim and Jeong, 2020) went slightly further into the areas of cooperative traffic signal control with traffic flow prediction (TFP-CTSC) for a multi-intersection. To deal with multi-intersection efficiently, agents share their traffic information with other adjacent intersections. However, this cooperative method cannot be efficient if the agents for individual intersections are not learning efficiently. Vidali et al., (2019) made use of Deep neural network and Q-learning, a combination of two aspects widely adopted in the field of reinforcement learning.

The Q-learning involves assigning a value, called the Q-value, to an action taken from a precise state of the environment. The deep neural network approximates the Q-learning function. It selects the action with the highest value given the current state by doing these, the best traffic efficiency is achieved. However, this system is not dynamic enough to cope with the increasing traffic demand. Also, real traffic data were not used so this poses a question on how efficient this system can be in real-time. (Paras et al., 2020) considered an intelligent traffic control system made use of a Deep Q-network with experience replay.

Borges et al., (2021)developed a model based on deep Q-learning traffic control models that can handle high-volume traffic data and synchronize traffic lights in an urban network in real time. Li et al., (2021) considered a collaborative computation offloading in a time-varying edge-cloud network, and formulated an optimization problem with considering both delay constraints and resource constraints, aiming to minimize the long-term system cost. The authors then utilized a DQN-based approach to solve the optimization problem.

Q-learning is off-policy: regardless of the policy being followed, it always estimates the optimal Q-function. It utilizes previously collected data, without additional online data collection control. Most recently(Zeinaly et al., 2023) in their work, developed a reliable deep reinforcement learning based controller for a highly dynamic environment and investigated the resilience of these controllers to a variety of environmental disruptions. The agent is trained using deep Q-learning and experience replay. The model is evaluated in the traffic micro-simulator SUMO. The simulation results demonstrate that the proposed method is effective at shortening queues when there is disruption. The works of these authors have contributed greatly to the body of knowledge especially in the area of traffic optimization in order to reduce traffic congestion. However, the work of (Zhou et al., 2014)only applied SARSA learning without integrating it with Deep learning. This will make it unable to interact with the traffic environment at real time.

Also, (Li et al., 2021; Yen et al., 2020) went a step ahead applying DDPG and DQN algorithm with experience replay which are off policy deep reinforcement learning algorithms. Off policy signal control model does not interact with the traffic environment at real time (Tan et al., 2022). One of the challenges observed by the researcher in is on how to integrate a mechanism for utilizing past experiences of traffic conditions using the normal traffic pattern of that vicinity and still control the road traffic in that environment at real-time. If this is not done, it will have considerable effect on journey times and in the predictions of upcoming traffic conditions. This will require a deep learning agent that judges by the result they produce. This brought about the idea of Real Time Deep SARSA Reinforcement Learning replay agent for the Control of Traffic in Nigeria.

Deep Reinforcement learning agents are well known for its application in control systems due to its ability to perceive its environment, interact, reason and learn from through it interaction to its environment. However, DRL approaches which have been proffered in traffic control so far have not provided an intelligent traffic control that is dynamic to change based on the queue of vehicles on each lane at an intersection and based on the use of past traffic information of the specified road. The Deep SARSA learning incorporated with experience replay can be applied in achieving this at real time.SARSA learning will be integrated to Deep Reinforcement Learning (DRL)framework. Similar to Deep Q-Network in(Mnih et al., 2015) , given the current state s , the action a is selected by ε -greedy method. Then the next state s′ and the reward r will be observed. The current state-action value is Q (s ,a) .

SARSA learning uses a Q-table to store values for each state-action pair. So in DRL based on SARSA, the current optimal state-action can be estimated by

$$Q^* (s, a) = E [r + \gamma Q (s', a') | s,a],$$

where
a ′ is the next action selected by ε -greedy.

Note that in SARSA learning action is taken using epsilon greedy policy and also while updating the Q-value, action is selected using the ε -greedy policy. ε -greedy policy selects action randomly with the highest estimated reward.

Similarly, in deep SARSA learning, the value function approximation is with the convolution neural network (CNN). The input of the CNN will be the raw images from the road traffic simulation and the output will be the Q values of all actions.The Deep SARSA replay agent adopts SARSA (on-policy algorithm) on-policy: at every update, it aims to estimate the Q-value of the current policy being followed, deep neural networks (CNNs), and experience replay memory. The overall framework of the Deep SARSA signal control model is shown in Figure 2.

Deep SARSA learning model for traffic signal control is achieved by the sequences of States, actions and corresponding rewards. Generally, the States is described by features on the traffic environment like queue length, delay time, number of stops and vehicle density at a given time t with state s_t. When the agent has observed the state s_t, the need arises for the agent to choose the appropriate an action $a_t \in A$. The set of all possible actions is denoted A = {NSG, EWG, NSLG, EWLG} which is defined by a pair of compass direction and set of movement priorities as expressed in work of(Genders & Razavi, 2018). The update the Q-values for state-action pairs is based on the estimate of reward. This is due to the fact that on-policy learning estimates rewards for state-action pairs by assuming that the same policy will be followed.

4. CONCLUSION

An efficient traffic control system should be able to cope with the rising transport demand and supply in order to eliminate traffic congestion. To address the traffic signal control problem, recent reinforcement learning research has proposed multiple possible solutions. This paper has made a constructive review of Deep Reinforcement learning in Real Time Intelligent Traffic Control Systems. Our objectives are to provide clear understanding on deep reinforcement learning (DRL) traffic control agent, as well as discussing on the existing research and challenges. From our review evaluation, we observed that most of the existing studies on it have been based on off-policy DRL agents,

We have been able to discuss some of the existing approaches and reached our observation are that most of this reinforcement learning traffic control agents do not possess the real-time traffic control and online learning capabilities to cope with the complex and changeable nature of traffic on an average urban city road. We therefore suggest that for future research on on-policy DRL traffic control agent integrated with optimized experience replay. This DRL traffic control agent can be adopted in achieving a traffic control system that can operate at real-time and also use past information to interact with the traffic environment considering factors like busiest working days in a week and rush hours.

REFERENCES

1. Ajay, S., and Chandra, A. (2009). Intelligent Traffic Control System using RFID. *IEEE Potentials*, *September 2014*. https://doi.org/10.1109/MPOT.2009.932094
2. Al-mudhaffar, A. (2006). *Impacts of Traffic Signal Control Strategies*. Royal Institute of Technology Stockholm, Sweden.
3. Aleksandar Stevanovic and Martin, P. (2009). SCOOT and SCATS: A Closer Look into Their Operations. *88th Annual Meeting of the Transportation Research Board. Washington DC*, *500*(801), 18.
4. Arnold, L., Rebecchi, S., and Chevallier, S. (2019). *An Introduction to Deep Learning*. *January 2011*.
5. Arnott, R., and Small, K. (2014). The Economics Of Traffic Congestion. *Journal of Transport,Economics and Policy.*, *32*(August), 93–112.
6. Azimian, A. (2011). *Design of an intelligent traffic management system*. University of Dayton.
7. Borges, F, Fonseca, A. P., and Garcia, R. C. (2021). Deep reinforcement learning model to mitigate congestion in real-time traffic light networks. *MDPI Infrastructures*, *6*(10), 1–25. https://doi.org/10.3390/infrastructures6100138
8. Burov, M., Arcak, M., and Kurzhanskiy, A. (2021). *Speed Advisory System Using Real-Time Actuated Traffic Light Phase Length Prediction*. 1–10. http://arxiv.org/abs/2107.10372
9. Cao, Z., Jiang, S., Zhang, J., and Guo, H. (2017). A Unified Framework for Vehicle Rerouting and Traffic Light Control to Reduce Traffic Congestion. *IEEE Transactions on Intelligent Transportation Systems*, *18*(July), 1958–1973. https://doi.org/10.1109/TITS.2016.2613997
10. Casas, N. (2017). *Deep Deterministic Policy Gradient for Urban Traffic Light Control*. *2*(arrXiv:1703.09035), 1–38.
11. Cheng, A., Pang, M., and Pavlou, P. A. (2020). Mitigating Traffic Congestion : The Role of Intelligent Transportation Systems. *The London School of Economics and Political Science*, 1–56.
12. Chris, G. (2019). *4 Ways Cities Are Using Smart Technology To Control Traffic Congestion » Traffic Safety Resource Center*. https://www.trafficsafetystore.com/blog/4-ways-cities-are-using-smart-technology-to-control-traffic-congestion/
13. Clark, B. (2019). How traffic jams cost the US economy billion dollars every year. *CNBC*.
14. François-lavet, V., Henderson, P., Islam, R., Bellemare, M. G., François-lavet, V., Pineau, J., and Bellemare, M. G. (2018). *An Introduction to Deep Reinforcement Learning*. https://doi.org/10.1561/2200000071.Vincent
15. Fujimoto, S., Meger, D., and Precup, D. (2019). *Off-Policy Deep Reinforcement Learning without Exploration*.

16. Gao, J., Shen, Y., Liu, J., Ito, M., and Shiratori, N. (2017). *Adaptive Traffic Signal Control : Deep Reinforcement Learning Algorithm with Experience Replay and Target Network*. *1*, 1–10.
17. Genders, W., and Razavi, S. (2018). Evaluating reinforcement learning state representations for adaptive traffic signal control. *Procedia Computer Science*, *130*(January), 26–33. https://doi.org/10.1016/j.procs.2018.04.008
18. Greguri, M., and Vuji, M. (2020a). Application of Deep Reinforcement Learning in Traffic Signal Control : An Overview and Impact of Open Traffic Data. *MDPI Applied Sciences*, *MI*. https://doi.org/10.3390/app10114011
19. Greguri, M., and Vuji, M. (2020b). Application of Deep Reinforcement Learning in Traffic Signal Control : An Overview and Impact of Open Traffic Data. *MDPI, Applied Sciences*, *MI*, 1–25. https://doi.org/10.3390/app10114011
20. Grzes, M., and Kudenko, D. (2008). *Robustness Analysis of SARSA (lambda): Different Models of Reward and Initialisation . Robustness Analysis of SARSA (λ): Different*. *September*. https://doi.org/10.1007/978-3-540-85776-1
21. Gustafsson, R. (2019). *Machine Learning for Traffic Control of Unmanned Mining Machines Using the Q-learning and SARSA algorithms Maskininlärning för Trafikkontroll av Obemannade Gruvmaskiner Med användning av algoritmerna Q- learning*.
22. Hausknecht, M., Stone, P., and Mc, O. (2016). *On-Policy vs . Off-Policy Updates for Deep Reinforcement Learning*.
23. Kim, D., and Jeong, O. (2020). *Cooperative Tra ffi c Signal Control with Tra ffi c Flow Prediction in Multi-Intersection*.
24. Koukol, M., and Marek, L., (2015). *Fuzzy Logic in Traffic Engineering : A Review on Signal Control*. *2015*.
25. Li, Y., Xu, S., and Li, D. (2021). Deep Reinforcement Learning for Collaborative Computation Offloading on Internet of Vehicles. *Wireless Communications and Mobile Computing*, *2021*, 1–13. https://doi.org/10.1155/2021/6457099
26. Mathew, T. (2014). Vehicle Actuated Signals. *Transportation Systems Engineering*, 1–19.
27. Mathew, T., and Rao, K. (2007). Traffic signal design-I 41.1 Overview. In *Introduction to Transportation and Engineering* (pp. 1–10). https://nptel.ac.in/courses/105101087/downloads/Lec-41.pdf
28. Matthews, V. O., Ajala, A. O., Popoola, S. I., and Atayero, A. A. (2017). *Smart Vehicular Traffic Management System using RFID Technology*. *I*, 7–10.
29. Mcshane, C. (1999). The origins and globalization of traffic control signals. *Journal of Urban History*, *25*(3), 379–404.

30. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., and Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533. https://doi.org/10.1038/nature14236
31. Nasuha, N., Azlan, N., and Rohani, M. (2018). *Overview Of Application Of Traffic Simulation Model*. *03006*, 1–6.
32. Oladipo, F. O. (2015). Design and Simulation of an Intelligent Traffic Control System. *International Journal of Advances in Engineering and Technology*, *1*(5), 47–57.
33. Oni, A. (2008). *Intelligent traffic control system*.
34. Perez-murueta, P., Gomez- Espinosa, A., Cardenas, C., and Gonzalz-Mendoza. (2019). Deep Learning System for Vehicular Re-Routing and Congestion Avoidance. *Applied Science*, 1–14.
35. Rasheed, F., Yau, K. A., and Member, S. (2020). *Deep Reinforcement Learning for Traffic Signal Control : A Review*. *8*. https://doi.org/10.1109/ACCESS.2020.3034141
36. Rosenblat, A., and Kneese, T. (2014). *Understanding Intelligent Systems*.
37. Ryabokon, Y. (2017). The Method of Determining the Number of Phases in the Traffic Light Cycle on the Allowable Intensity of Conflicting Flows. *Transportation Research Procedia*, *20*, 571–577. https://doi.org/10.1016/j.trpro.2017.01.092
38. Salter, R. J. (1989). Introduction to Traffic Signal Control. *Traffic Engineering*, *877*, 79–85. https://doi.org/10.1007/978-1-349-10800-8_20
39. Shin, S., and Kang, Y. (2019). Obstacle Avoidance Drone by Deep Reinforcement Learning and Its Racing with Human Pilot. *Applied Science*, *doi:10.339*(1), 1–22. https://doi.org/10.3390/app9245571
40. Tan, J., Yuan, Q., Guo, W., Xie, N., Liu, F., Wei, J., and Zhang, X. (2022). Deep Reinforcement Learning for Traffic Signal Control Model and Adaptation Study. *MDPI*, 1–21. https://doi.org/https://doi.org/10.3390/ s22228732
41. Vanderschuren, M. (2006). *Intelligent Transport Systems for South Africa Impact Assessment Through Microscopic Simulation in the South African Context* (Issue August 2006).
42. Vidali, A., Crociani, L., Vizzari, G., and Bandini, S. (2019). *A Deep Reinforcement Learning Approach to Adaptive Traffic Lights Management*. *Woa*, 42–50.
43. Wang, S., Xie, X., Huang, K., Zeng, J., and Cai, Z. (2019). Deep reinforcement learning-based traffic signal control using high-resolution event-based data. *Entropy*, *21*(8), 1–16. https://doi.org/10.3390/e21080744
44. Wei, H. (2018). IntelliLight : A Reinforcement Learning Approach for Intelligent Traffic Light Control. *International Conference on Knowledge Discovery&DataMining*. https://doi.org///doi.org/10.1145/3219819.3220096

45. Xu, Z., Tang, J., Meng, J., Zhang, W., Wang, Y., Liu, C. H., and Yang, D. (2018). *Experience-driven Networking : A Deep Reinforcement Learning based Approach*. 1–9.
46. Yadav, A., Rais,I., Kumar,M., Sharma, A. and Kushwaha A., Image Classification using Deep Learning and Tensorflow , International Journal for Research in Applied Science & Engineering Technology, 10(5) . https://doi.org/10.22214/ijraset.2022.43385
47. Yen, C. C., Ghosal, D., Zhang, M., and Chuah, C. N. (2020). A Deep On-Policy Learning Agent for Traffic Signal Control of Multiple Intersections. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems, ITSC 2020*. https://doi.org/10.1109/ITSC45102.2020.9294471
48. Yousef, K. M. A., and Shatnawi, A. M. (2013). Intelligent Traffic Light Flow Control System Using Wireless Sensors Networks. *Journal of Information Science and Ngineering*, *26*, 753–768.
49. Zeinaly, Z., Sojoodi, M., and Bolouki, S. (2023). *A Resilient Intelligent Traffic Signal Control Scheme for Accident Scenario at Intersections via Deep Reinforcement Learning*. 1–27.
50. Zhao, D., Wang, H., Shao, K., and Zhu, Y. (2017). Deep reinforcement learning with experience replay based on SARSA. *2016 IEEE Symposium Series on Computational Intelligence, SSCI 2016*. https://doi.org/10.1109/SSCI.2016.7849837
51. Zhou, X., Zhu, F., Liu, Q., Fu, Y., and Huang, W. (2014). A Sarsa(λ)-based control model for real-time traffic light coordination. *The Scientific World Journal*, *2014*. https://doi.org/10.1155/2014/759097