

Implementasi Algoritma Deep Q-Network (DQN) pada Lampu Lalu Lintas Adaptif Berdasarkan Waktu Tunggu dan Arus Kendaraan

Ridho Amanda Putra¹, Yoanda Alim Syahbana², Ananda³

ridho22mttk@mahasiswa.pcr.ac.id¹, yoanda@pcr.ac.id², ananda@pcr.ac.id³

^{1,2,3} Politeknik Caltex Riau

Informasi Artikel

Diterima : 15 Agu 2024

Direvisi : 19 Sep 2024

Disetujui : 29 Okt 2024

Kata Kunci

Reinforcement Learning, Deep Q-Network, SUMO, Lampu Lalu Lintas Adaptif

Abstrak

Kemacetan lalu lintas sering terjadi akibat sistem buka tutup jalan selama perbaikan di satu sisi jalan, memaksa kendaraan bergantian melewati sisi yang tidak diperbaiki. Penggunaan lampu lalu lintas temporer dengan waktu tetap seringkali tidak efektif karena tidak memperhitungkan ketidakseimbangan arus kendaraan dari kedua sisi jalan. Untuk mengatasi masalah ini, penelitian ini mengimplementasikan algoritma Deep Q-Network (DQN) untuk mengoptimalkan durasi lampu lalu lintas berdasarkan waktu tunggu dari kendaraan dan arus kendaraan. Pengujian dilakukan menggunakan simulator SUMO, dengan fokus pada parameter epoch, exploration rate, dan discount factor yang mempengaruhi performa Agen DQN. Hasil menunjukkan bahwa DQN mencapai performa optimal saat dikonfigurasi dengan exploration rate 1 dan discount factor 0,9, setelah training selama 50 epoch dan diuji selama 10 epoch. Dalam kondisi ini, DQN terbukti lebih adaptif dalam mengatur lampu lalu lintas dibandingkan metode konvensional yang menggunakan waktu tetap untuk lampu hijau dan merah. Dengan nilai fairness dari DQN lebih kecil, DQN berhasil mengurangi kemacetan dan meningkatkan efisiensi lalu lintas secara keseluruhan.

Keywords

Reinforcement Learning, Deep Q-Network, SUMO, Adaptive Traffic Light

Abstract

Traffic congestion often occurs due to the alternating road closure system during repairs on one side of the road, forcing vehicles to take turns passing the unrepaired side. The use of temporary traffic lights with fixed timing is often ineffective because it does not account for the imbalance in traffic flow from both sides of the road. To address this issue, this study implements the Deep Q-Network (DQN) algorithm to optimize traffic light duration based on vehicle wait times and traffic flow. Testing was conducted using the SUMO simulator, focusing on the parameters of epoch, exploration rate, and discount factor that affect the performance of the DQN agent. The results show that DQN achieves optimal performance when configured with an exploration rate of 1 and a discount factor of 0.9, after training for 50 epochs and testing for 10 epochs. In this configuration, DQN proves to be more adaptive in managing traffic lights compared to conventional methods that use fixed timing for green and red lights. Although the fairness value of DQN is lower, it successfully reduces congestion and improves overall traffic efficiency.

A. Pendahuluan

Salah satu faktor penyebab kemacetan di jalan raya adalah sistem buka tutup jalan saat ada perbaikan di salah satu sisi jalan seperti diilustrasikan pada Gambar 1. Situasi ini mengharuskan kendaraan bergantian melewati sisi jalan yang tidak sedang diperbaiki. Untuk memastikan kelancaran arus kendaraan dari dua arah yang berlawanan, diperlukan pengaturan yang tepat. Penggunaan lampu lalu lintas yang temporer di jalan dapat membantu mengatur lalu lintas dengan lebih baik[1]. Lampu lalu lintas temporer ini biasanya masih beroperasi berdasarkan pengaturan waktu lampu hijau dan merah yang tetap. Hal ini tidak efektif karena lampu lalu lintas jenis ini tidak mempertimbangkan keseimbangan arus kendaraan dari kedua sisi. Ketika salah satu sisi jalan memiliki arus kendaraan yang lebih banyak, jenis lampu lalu lintas ini tetap menggunakan durasi lampu hijau dan merah yang sama. Sehingga, kemacetan terjadi ketika kendaraan harus menunggu lama untuk kendaraan dari arah sebaliknya.



Gambar 1. Studi Kasus Penelitian

Berbagai metode pengaturan lampu lalu lintas klasik untuk mengatasi kemacetan telah banyak dilakukan pada penelitian terdahulu [2]. Namun, penelitian tersebut masih memiliki kendala dalam kemampuan adaptabilitas terhadap kondisi lalu lintas yang sangat dinamis. Oleh karena itu, pendekatan machine learning dengan kemampuan belajar dari pola lalu lintas yang ada akan mampu mengatasi kelemahan dari metode klasik[3]. Dibandingkan dengan pengaturan berdasarkan waktu, kontrol lampu lalu lintas yang adaptif telah terbukti meningkatkan kualitas layanan lalu lintas kepada pengguna jalan serta mengurangi waktu perjalanan rata-rata lebih dari 10% [4].

Dalam machine learning, terdapat kemajuan yang signifikan dalam pengembangan Reinforcement Learning (RL), khususnya dalam mengatasi berbagai masalah [5]. Khusus dalam Intelligent Transport System (ITS), RL telah digunakan untuk mengoptimalkan berbagai aspek transportasi, seperti lalu lintas jalan raya, manajemen kendaraan, dan transportasi publik[6][7]. RL dapat digunakan untuk mengoptimalkan kontrol lampu lalu lintas dengan mempelajari pola lalu lintas dan mengatur durasi lampu yang sesuai untuk meminimalkan waktu tunda dan kepadatan lalu lintas [8][9].

Penelitian ini membahas implementasi algoritma Deep Q-Network (DQN) yang merupakan pengembangan dari metode Q-Learning dari RL dengan penambahan Deep Neural Network[10]. DQN digunakan untuk mengatur durasi lampu lalu lintas pada studi kasus yang telah diilustrasikan pada Gambar 1. Adaptabilitas dari DQN untuk mengatur durasi lampu lalu lintas dilakukan berdasarkan kondisi waktu tunggu dari kendaraan dan arus kendaraan. Pengujian performa DQN dilakukan dalam simulator lalu lintas Simulation of Urban Mobility (SUMO) [11]. Secara khusus, penelitian ini fokus pada pengujian parameter epoch,

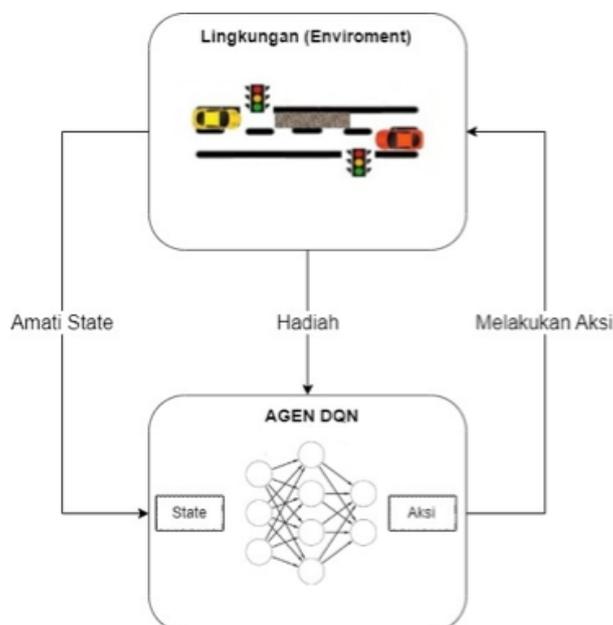
exploration rate, dan discount factor yang akan mempengaruhi perilaku dari agen RL dalam mengatur durasi lampu lalu lintas.

Pembahasan penelitian ini dibagi menjadi empat bagian yang dimulai dengan bagian pendahuluan. Bagian kedua membahas implementasi metode DQN yang diujikan pada studi kasus. Hasil pengujian dan analisisnya dibahas pada bagian ketiga. Bagian keempat fokus pada kesimpulan dan saran penelitian di masa depan berdasarkan hasil implementasi yang telah didapatkan.

B. Metode Penelitian

B.1. Desain arsitektur RL

Arsitektur implementasi DQN pada studi kasus lampu lalu lintas adaptif direpresentasikan pada Gambar 2. Proses kontrol dimulai dari Agen DQN yang mengamati State di Environment. Environment merupakan simulator yang merepresentasikan kondisi studi kasus yang diimplementasikan pada SUMO. State yang diamati adalah waktu tunggu dan arus kendaraan pada kedua sisi jalan. Berdasarkan pengamatan, Agen melakukan aksi berupa pengaturan durasi lampu lalu lintas untuk kedua sisi jalan. Agen menentukan durasi lampu hijau yang memperbolehkan satu sisi jalan untuk melintas sekaligus menjadi durasi lampu merah yang tidak memperbolehkan sisi jalan lain untuk melintas. Kombinasi pengamatan dan aksi yang dilakukan Agen secara terus menerus akan dievaluasi berdasarkan fairness dari lalu lintas. Fairness didefinisikan sebagai selisih absolut waktu tunggu di kedua sisi jalan. Semakin kecil nilai fairness diartikan sebagai semakin adil DQN dalam menentukan sisi jalan yang boleh melintas. Parameter fairness umum digunakan pada penelitian sebelumnya sebagai ukuran performa dari kemampuan agent dalam mengatur kelancaran lalu lintas. Ini yang kemudian menjadi referensi hadiah (Rewards) sebagai umpan balik oleh Agen.



Gambar 2. Arsitektur DQN sebagai metode RL

B.2. Simulation of Urban Mobility (SUMO)

SUMO merupakan platform simulasi open source yang digunakan untuk merepresentasikan model lalu lintas di sebuah wilayah dengan berbagai elemen seperti jalan, simpang, kendaraan, dan lampu lalu lintas[12][13]. Penelitian ini menggunakan SUMO untuk merepresentasikan studi kasus yang ditampilkan pada Gambar 1. Gambar 3 merupakan tampilan SUMO yang digunakan untuk menyimulasikan studi kasus pada penelitian ini. Objek kuning merupakan mobil yang dapat bergerak pada jalan yang berwarna hitam. Mobil dapat bergerak dari kiri ke kanan dan kanan ke kiri. Namun, jalan yang memiliki sisi warna orange merepresentasikan area jalan sedang mengalami perbaikan. Area ini hanya dapat dilewati satu arah kendaraan pada satu waktu. Sebagai contoh, ketika mobil bergerak dari kiri ke kanan, maka mobil yang bergerak dari kanan ke kiri akan antri menunggu giliran melewati area warna orange. Ini merepresentasikan kondisi ketika lampu hijau untuk sisi kiri dan lampu merah untuk sisi kanan. Dataset yang digunakan untuk simulasi studi kasus dihasilkan dari fungsi random yang dibuat dari coding.



Gambar 3. SUMO untuk studi kasus penelitian

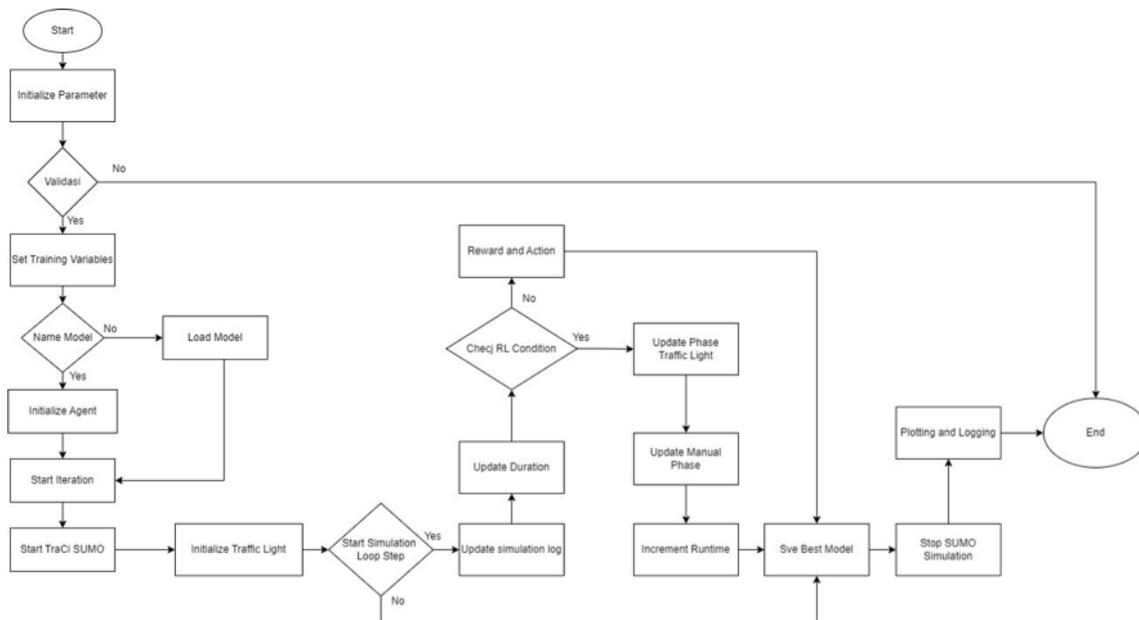
B.3. Deep Q-Network (DQN)

DQN adalah algoritma RL yang menggabungkan algoritma Q-Learning dengan Deep Neural Network (DNN)[10][14][15]. Q-Learning dikategorikan dalam metode model-free yang tidak membutuhkan informasi yang lengkap tentang kondisi Environment. Berbeda dengan pendekatan iteratif yang dikategorikan sebagai metode model-based, Q-Learning memungkinkan Agen untuk berada satu langkah ke depan[16]. Agen berada dalam keadaan tertentu, mengambil tindakan tertentu, dan menerima Value berdasarkan tindakan tersebut. Pada DQN, DNN digunakan untuk memodelkan nilai Value tersebut dengan menggunakan arsitektur Neural Network[14][16]. DQN cocok digunakan untuk studi kasus yang diuji pada penelitian ini karena kondisi lalu lintas sangat dinamis.

Implementasi yang mengintegrasikan DQN dengan SUMO dilakukan melalui IDE PyCharm menggunakan bahasa Python. Skrip Python dimulai dengan menginisialisasi simulasi, menjalankan algoritma DQN, dan menyimpan model pembelajaran DQN. Hasil simulasi berupa waktu tunggu dan arus kendaraan disimpan untuk digunakan lebih lanjut dalam pengukuran fairness dari keputusan Agen.

Gambar 4 menampilkan flowchart dari implementasi DQN untuk mengatur lalu lintas kendaraan pada SUMO. Pertama, inisialisasi dan validasi parameter dilakukan. Kemudian, variabel yang berkaitan dengan proses training dari Agen dimulai. Jika belum ada model yang tersedia, maka Agen akan mulai belajar dari awal, tanpa pengetahuan apapun terkait Environment. Namun, jika telah ada model yang tersedia, maka Agen dapat memanfaatkan informasi tersebut sebagai

pengetahuan awal. Kedua, Agen akan mulai mengatur lalu lintas berdasarkan trafik mobil yang ada di SUMO. Dalam pengujian ini, durasi eksperimen ditetapkan selama 1800 detik untuk satu epoch. Dalam setiap epoch, Agen akan menerima observasi kondisi Environment dan melakukan aksi mengatur durasi lampu lalu lintas. Di akhir epoch, reward dihitung berdasarkan perhitungan fairness. Jika fairness yang didapatkan lebih baik daripada fairness dari epoch sebelumnya, maka model RL diperbaharui.

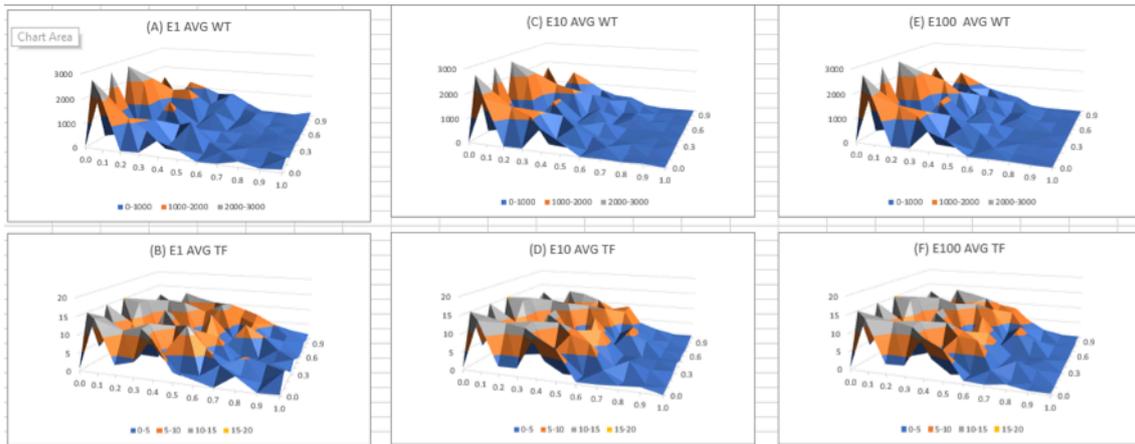


Gambar 4. Flowchart sistem

C. Hasil dan Pembahasan

Eksperimen implementasi DQN dimulai dengan melatih Agen untuk mengatur lalu lintas pada studi kasus untuk 50 epoch. Setelah proses training 50 epoch selesai, evaluasi fairness dilakukan pada epoch ke 51, 60, dan 150 (testing epoch 1, 10, dan 100). Setiap eksperimen ini akan memvariasikan exploration rate dan discount factor untuk menganalisis perilaku pembelajaran dari Agen. Kedua parameter ini divariasikan mulai dari 0 sampai 1 dengan kenaikan berkala 0,1. Berdasarkan hasil eksperimen, 1 epoch membutuhkan rata-rata waktu proses 23 detik, 10 epoch membutuhkan rata-rata 222 detik, dan 100 epoch membutuhkan 2.084 detik. Evaluasi fairness juga dilakukan untuk membandingkan pembelajaran agen berdasarkan observasi waktu tunggu (WT) dan arus kendaraan (TF). WT dihitung dalam satuan detik yang merepresentasikan lama waktu tunggu kendaraan untuk menunggu melintas. TF dihitung dalam satuan unit kendaraan yang merepresentasikan jumlah kendaraan yang melintas.

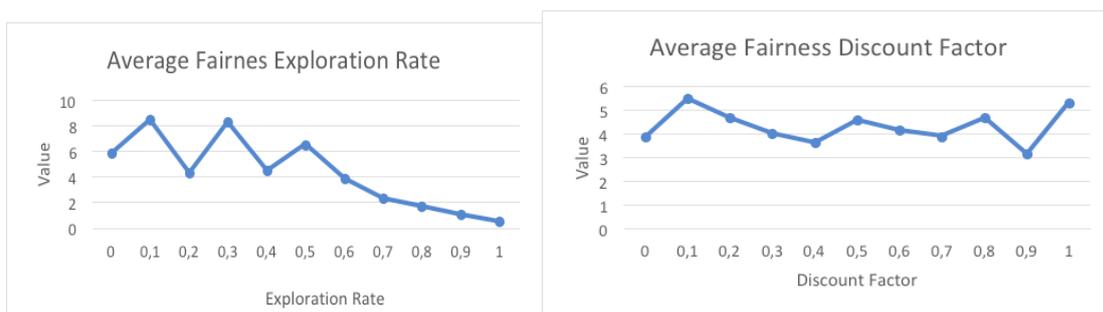
Eksperimen ini telah dilakukan dan hasil eksperimen ditampilkan pada Gambar 5. Setiap sub gambar menampilkan hasil WT dan TF dalam grafik tiga dimensi. Sumbu x merepresentasikan variasi exploration rate, sumbu y merepresentasikan variasi discount factor, sedangkan sumbu z merepresentasikan rata-rata WT dan TF pada masing-masing percobaan.



Gambar 5. Grafik hasil testing DQN dengan variasi parameter: A) WT pada epoch 1, B) TF pada epoch 1, C) WT pada epoch 10, D) TF pada epoch 10, E) WT pada epoch 100, F) TF pada epoch 100

Hasil eksperimen menunjukkan perbedaan signifikan antara epoch 1 dan epoch 10. Namun, perbandingan antara epoch 10 dan 100 tidak menunjukkan perbedaan signifikan. Sehingga, optimalisasi hasil dengan mempertimbangkan waktu proses menunjukkan epoch 10 sebagai hasil pembelajaran Agen terbaik. Perbandingan hasil fairness dari eksperimen juga menunjukkan bahwa Agen menghasilkan keputusan lebih baik dengan mengobservasi TF daripada WT.

Fokus pada epoch 10 dan observasi TF, titik optimal fairness dari kombinasi exploration rate dan discount factor didapatkan pada exploration rate 1 dan discount factor 0.9 sebagaimana ditampilkan pada Gambar 6. Hasil ini menunjukkan Agen mampu menghasilkan keputusan pengaturan lalu lintas terbaik ketika Agen diperbolehkan untuk eksplorasi semua kemungkinan keputusan dan tidak bergantung pada keputusan terbaik saja. Selanjutnya, Agen juga mampu menghasilkan keputusan terbaik ketika Agen lebih memprioritaskan Rewards jangka panjang di masa depan.



Gambar 6. Hasil eksperimen untuk epoch 10: A) exploration rate, B) discount factor

Pengujian juga dilakukan dengan metode konvensional di mana lampu lalu lintas diatur memiliki durasi tetap dan berulang untuk membandingkan apakah penggunaan DQN lebih baik dibandingkan metode konvensional. Dari hasil pengujian metode konvensional, diperoleh nilai fairness 0,84. Sementara itu, dengan menggunakan DQN, nilai optimal diperoleh pada tingkat eksplorasi 1 dan

discount factor 0,9, dengan nilai fairness dari TF sebesar 0,79. Dari hasil tersebut, dapat disimpulkan bahwa penggunaan DQN lebih optimal dalam meminimalkan kemacetan dibandingkan dengan metode konvensional.

D. Simpulan

Penelitian ini telah berhasil mengimplementasikan algoritma Deep Q-Network (DQN) pada studi kasus lampu lalu lintas adaptif berdasarkan waktu tunggu dan arus kendaraan. Berdasarkan eksperimen penelitian yang telah dilakukan didapatkan hasil bahwa Agen mencapai titik optimal dalam pengambilan keputusan saat dikonfigurasi dengan exploration rate 1 dan discount factor 0,9. Ini dicapai saat Agen telah training selama 50 epoch dan testing sampai 10 epoch setelah training. Performa DQN juga telah dibandingkan dengan lampu lalu lintas konvensional yang memiliki durasi lampu hijau dan merah yang tetap. Hasil perbandingan menunjukkan DQN memiliki nilai fairness yang lebih kecil.

Selanjutnya, penelitian dapat dikembangkan dengan mempertimbangkan faktor kecepatan kendaraan yang melintas. Hal ini merepresentasikan kelancaran lalu lintas yang diatur. Selain itu, DQN juga dapat mempertimbangkan pembobotan dalam penentuan Rewards.

E. Ucapan Terima Kasih

Kami menyampaikan terima kasih kepada seluruh mahasiswa dan dosen jurusan magister terapan teknik komputer Politeknik Caltex Riau yang terlibat langsung dalam penelitian ini. Terima kasih juga kepada teman-teman yang terlibat secara tidak langsung dan juga orang tua yang selalu mendukung dalam penelitian ini.

F. Referensi

- [1] B. Wang, Z. He, J. Sheng, and Y. Chen, "Deep Reinforcement Learning for Traffic Light Timing Optimization," *Processes*, vol. 10, no. 11, 2022, doi: 10.3390/pr10112458.
- [2] H. Wei, G. Zheng, V. Gayah, and Z. Li, "Recent Advances in Reinforcement Learning for Traffic Signal Control," 2021. doi: 10.1145/3447556.3447565.
- [3] T. Pan, "Traffic light control with reinforcement learning," *Appl. Comput. Eng.*, vol. 43, no. 1, pp. 26–43, 2024, doi: 10.54254/2755-2721/43/20230804.
- [4] X. Liang, X. Du, G. Wang, and Z. Han, "A Deep Reinforcement Learning Network for Traffic Light Cycle Control," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1243–1253, Feb. 2019, doi: 10.1109/TVT.2018.2890726.
- [5] D. Silver *et al.*, "A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play," *Science (80-.)*, vol. 362, no. 6419, pp. 1140–1144, 2018.
- [6] I. Moumen, J. Abouchabaka, and N. Rafalia, "Adaptive traffic lights based on traffic flow prediction using machine learning models," *Int. J. Electr. Comput. Eng.*, vol. 13, no. 5, pp. 5813–5823, 2023, doi: 10.11591/ijece.v13i5.pp5813-5823.
- [7] X.-Y. Liu, M. Zhu, S. Borst, and A. Walid, "Deep Reinforcement Learning for Traffic Light Control in Intelligent Transportation Systems," vol. XX, no. Xx,

- pp. 1–17, 2023, [Online]. Available: <http://arxiv.org/abs/2302.03669>
- [8] B. Koziarski Michał and Cyganek, “Impact of low resolution on image recognition with deep neural networks: An experimental study,” *Int. J. Appl. Math. Comput. Sci.*, vol. 28, no. 4, pp. 735–744, 2018.
- [9] K.-L. A. Yau, J. Qadir, H. L. Khoo, M. H. Ling, and P. Komisarczuk, “A survey on reinforcement learning models and algorithms for traffic signal control,” *ACM Comput. Surv.*, vol. 50, no. 3, pp. 1–38, 2017.
- [10] S. Park, E. Han, S. Park, H. Jeong, and I. Yun, “Deep Q-network-based traffic signal control models,” *PLoS One*, vol. 16, no. 9 September, pp. 1–14, 2021, doi: 10.1371/journal.pone.0256405.
- [11] Y. H. Chow, K. J. A. Ooi, M. A. S. Bhuiyan, M. B. I. Reaz, and C. W. Yuen, “Computation and Optimization of Traffic Network Topologies Using Eclipse SUMO,” *Ann. Emerg. Technol. Comput.*, vol. 6, no. 4, pp. 31–37, 2022, doi: 10.33166/AETiC.2022.04.004.
- [12] K. Shingate, K. Jagdale, and Y. Dias, “Adaptive traffic control system using reinforcement learning,” *Int. J. Eng. Res. Technol*, vol. 9, 2020.
- [13] P. A. Lopez *et al.*, *Microscopic Traffic Simulation using SUMO*, vol. 2018-Novem. 2018. doi: 10.1109/ITSC.2018.8569938.
- [14] M. R. T. Fuad *et al.*, “Adaptive Deep Q-Network Algorithm with Exponential Reward Mechanism for Traffic Control in Urban Intersection Networks,” *Sustain.*, vol. 14, no. 21, 2022, doi: 10.3390/su142114590.
- [15] H. Joo and Y. Lim, “Traffic signal time optimization based on deep q-network,” *Appl. Sci.*, vol. 11, no. 21, Nov. 2021, doi: 10.3390/app11219850.
- [16] R. Ducrocq and N. Farhi, “Deep Reinforcement Q-Learning for Intelligent Traffic Signal Control with Partial Detection,” *Int. J. Intell. Transp. Syst. Res.*, vol. 21, no. 1, pp. 192–206, 2023, doi: 10.1007/s13177-023-00346-4.