JURNAL AI dan SPK
Jurnal Artificial Intelligent dan Sistem Penunjang Keputusan

Jurnal AI dan SPK : Jurnal Artificial Inteligent dan Sistem Penunjang Keputusan
Volume 3, No. 2, Tahun 2025
ISSN 3025-0927 (media online)
Hal 263-274

Implementasi Sistem Pengaturan Lampu Lalu Lintas Berbasis DRL untuk Mengoptimalkan Durasi Lampu Hijau

Roqman Firando[1], Saprudin[2]

[12]Fakultas Ilmu Komputer,Program Studi Teknik Informatika, Universitas Pamulang, Jl. Raya Puspiptek No. 46, Kel.Buaran, Kec. Serpong, Kota Tangerang Selatan. Banten 15310, Indonesia
Email: [1]firnandorfn@gmail.com*, [2]dosen00845@unpam.ac.id

Abstrak−Kemacetan lalu lintas di persimpangan jalan merupakan masalah serius di kota-kota besar, terutama akibat sistem kontrol lampu statis yang tidak adaptif. Sistem konvensional seringkali tidak mampu merespons kondisi lalu lintas *real-time*, mengakibatkan antrean panjang dan waktu tunggu yang tidak efisien. Oleh karena itu, penelitian ini mengimplementasikan sistem pengaturan lampu lalu lintas berbasis *Deep Reinforcement Learning (DRL)* untuk mengoptimalkan durasi lampu hijau secara adaptif. Pendekatan *DRL* dengan algoritma *Double Deep Q-Network (Double DQN)* diimplementasikan dalam simulasi *SUMO*. Agen *DRL* dilatih untuk menerima informasi *state* (jumlah antrean kendaraan dan fase lampu hijau yang aktif) dan menghasilkan *action* (memilih durasi fase hijau). Fungsi *reward* dirancang untuk meminimalkan "Tekanan" sistem (yaitu jumlah kuadrat antrean), yang berfungsi untuk menghukum ketidakseimbangan lalu lintas dan memandu agen belajar strategi optimal. Hasil pengujian komparatif pada berbagai skenario, khususnya skenario acak untuk uji generalisasi, menunjukkan *DRL* secara signifikan lebih unggul dari sistem kontrol waktu tetap (Statis-30s dan Statis-60s). *DRL* berhasil menurunkan rata-rata antrean hingga 36.76%, meningkatkan rata-rata kecepatan kendaraan hingga 29.47%, dan meningkatkan *throughput* hingga 2.70% (vs. Statis-60s). Terhadap Statis-30s, *DRL* unggul dengan penurunan rata-rata antrean 12.86%, peningkatan rata-rata kecepatan 6.05%, dan peningkatan *throughput* 2.86%. Keunggulan ini mencakup kemampuan bergeneralisasi yang kuat. Penelitian ini menyimpulkan pendekatan *DRL* dengan *Double DQN* adalah solusi efektif dan tangguh untuk mengoptimalkan durasi lampu hijau.

Kata Kunci: Sistem Lampu Lalu Lintas; Kemacetan; *Deep Reinforcement Learning; Double Deep Q-Network;* Lalu Lintas Adaptif

Abstract−*Traffic congestion at road intersections is a critical issue in major cities, mainly due to static traffic light control systems' inability to adapt to dynamic traffic volumes. This inability of conventional systems to respond in real-time leads to long queues and inefficient waiting times. Therefore, this research implements a Deep Reinforcement Learning (DRL)-based traffic light control system to optimize green light durations adaptively. The DRL approach, utilizing the Double Deep Q-Network (Double DQN) algorithm, is implemented in a SUMO environment. The DRL agent is trained to perceive its state (queue counts and the active green light phase) and produce an action (selecting green light phase durations). The reward function is designed to minimize overall system "Pressure" (sum of squared queue lengths), penalizing traffic imbalance and guiding the agent to learn optimal strategies. Comparative test results across diverse scenarios, particularly in the random scenario for generalization testing, convincingly demonstrate that the DRL agent outperforms fixed-time control systems (Static-30s and Static-60s). DRL successfully reduced average queue lengths by up to 36.76%, increased average vehicle speed by up to 29.47%, and increased throughput by up to 2.70% (vs. Static-60s). Against Static-30s, DRL also demonstrated better performance with a 12.86% reduction in average queues, a 6.05% increase in average speed, and a 2.86% increase in throughput. This improved performance includes strong generalization capabilities. This study concludes that the DRL approach with Double DQN is a highly effective and robust solution for optimizing green light durations.*

Keywords: *Traffic Light System; Congestion; Deep Reinforcement Learning; Double Deep Q-Network; Adaptive Traffic*

1. PENDAHULUAN

Kemacetan lalu lintas merupakan isu krusial yang melanda berbagai kota besar, membawa dampak negatif berupa keterlambatan perjalanan, peningkatan konsumsi bahan bakar, polusi udara, serta penurunan produktivitas (Astuti et al., 2024; Brilliant et al., 2024; Safira & Khuluqi, 2023). Persimpangan dengan lampu lalu lintas sering menjadi titik rawan kemacetan akibat pembagian waktu lampu yang tidak adaptif terhadap kondisi lalu lintas aktual (Astuti et al., 2024; Inayah et al., 2025; Maha, 2022). Sistem kontrol statis, dengan siklus waktu tetap, secara fundamental tidak mampu merespons dinamika lalu lintas yang berubah-ubah secara *real-time*. Kondisi ini menciptakan ketidakseimbangan waktu antara jalur yang kaku dan kebutuhan aktual di lapangan,

JURNAL AI dan SPK
Jurnal Artificial Intelligent dan Sistem Penunjang Keputusan
Jurnal AI dan SPK : Jurnal Artificial Inteligent dan Sistem Penunjang Keputusan
Volume 3, No. 2, Tahun 2025
ISSN 3025-0927 (media online)
Hal 263-274

mengakibatkan waktu tunggu yang tidak perlu pada jalur lengang dan penumpukan antrean parah di jalur padat.

Untuk mengatasi tantangan adaptasi *real-time* ini, berbagai pendekatan telah dieksplorasi. Namun, solusi konvensional atau berbasis aturan sering kali kurang optimal dalam menghadapi fluktuasi mendadak dan pola lalu lintas yang kompleks. Sebaliknya, teknologi kecerdasan buatan, khususnya *Deep Reinforcement Learning (DRL),* menawarkan kemampuan unik untuk belajar dari interaksi dengan lingkungan dan mengoptimalkan kebijakannya secara mandiri tanpa pemrograman eksplisit (Liang et al., 2019; Liu & Li, 2023; Matsuo et al., 2022). Pendekatan ini sangat menjanjikan untuk pengaturan sinyal lalu lintas yang adaptif, memungkinkan sistem secara otonom mengoptimalkan strategi kontrolnya berdasarkan data historis dan merespons perubahan pola lalu lintas secara dinamis (Damadam et al., 2022; Li et al., 2021; Putra et al., 2024). Meskipun demikian, pemanfaatan *DRL* untuk optimisasi lampu lalu lintas di Indonesia masih relatif terbatas, sehingga menyoroti kebutuhan akan pengembangan sistem yang lebih responsif dan mampu menyesuaikan durasi lampu secara otomatis sesuai dengan kepadatan kendaraan.

Penelitian ini bertujuan untuk mengimplementasikan sistem pengaturan lampu lalu lintas berbasis *Deep Reinforcement Learning (DRL)* untuk mengoptimalkan durasi lampu hijau secara adaptif dalam lingkungan simulasi *SUMO (Simulation of Urban MObility).* Pendekatan ini menggunakan algoritma *Double Deep Q-Network (Double DQN)* yang terbukti efektif dalam menangani masalah keputusan secara bertahap. Fokus penelitian ini adalah merancang komponen utama seperti definisi *state* (informasi antrean dan fase aktif), *action* (pilihan durasi lampu), dan *reward function* (berbasis "Tekanan") untuk memandu agen dalam mengurangi ketidakseimbangan lalu lintas. Selain itu, penelitian ini akan mengevaluasi efektivitas optimisasi dan kemampuan generalisasi dari sistem *DRL* yang dikembangkan pada berbagai skenario lalu lintas dinamis, membandingkan kinerjanya secara terukur terhadap sistem kontrol statis.

Kontribusi penelitian ini adalah menyajikan bukti empiris tentang efektivitas *Double DQN* dalam optimisasi persimpangan tunggal yang terisolasi, terutama pada kondisi lalu lintas asimetris, hal ini membedakannya dari penelitian yang lebih berfokus pada optimisasi jaringan terkoordinasi (Damadam et al., 2022; Liu & Li, 2023). Penelitian ini juga memberikan data kinerja kuantitatif yang dapat menjadi dasar pertimbangan bagi pengembangan sistem transportasi cerdas (*ITS*) di masa depan.

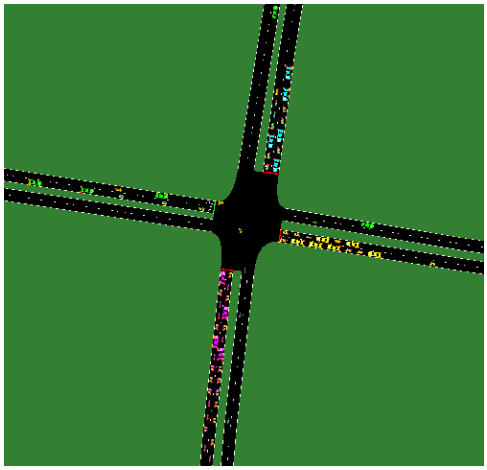
2. METODE PENELITIAN

Penelitian ini mengadopsi pendekatan eksperimen komputasional sistematis untuk mengevaluasi efektivitas algoritma *Deep Reinforcement Learning (DRL)* dalam optimisasi kontrol sinyal lalu lintas. Metodologi penelitian mencakup perancangan lingkungan simulasi, implementasi dan pelatihan agen *DRL*, serta evaluasi kinerja komparatif.

2.1 Desain Lingkungan Simulasi dan Baseline

Lingkungan eksperimen dibangun menggunakan simulator *SUMO (Simulation of Urban MObility)* versi 1.24.0 (Lopez et al., 2018), yang mendukung simulasi pergerakan setiap kendaraan secara individual (*mikroskopik*). Model persimpangan yang digunakan adalah persimpangan empat arah tunggal dan terisolasi, dengan dua lajur per jalur masuk untuk mengakomodasi volume tinggi. Skenario arus lalu lintas dirancang secara prosedural untuk merepresentasikan lima kondisi operasional yang berbeda (Seimbang, Jam Sibuk 2 Arah, Jam Sibuk 3 Arah, Arus Puncak 1 Arah, dan Skenario Acak untuk generalisasi). Durasi simulasi setiap episode adalah 7200 detik, dengan volume kendaraan mencapai 4.000 unit kendaraan, seperti praktik umum dalam penelitian simulasi lalu lintas yang sebanding (Coskun et al., 2018). Parameter simulasi diinformasikan oleh observasi lapangan untuk meningkatkan validitas.

JURNAL AI dan SPK
Jurnal Artificial Intelligent dan Sistem Penunjang Keputusan
Jurnal AI dan SPK : Jurnal Artificial Inteligent dan Sistem Penunjang Keputusan
Volume 3, No. 2, Tahun 2025
ISSN 3025-0927 (media online)
Hal 263-274

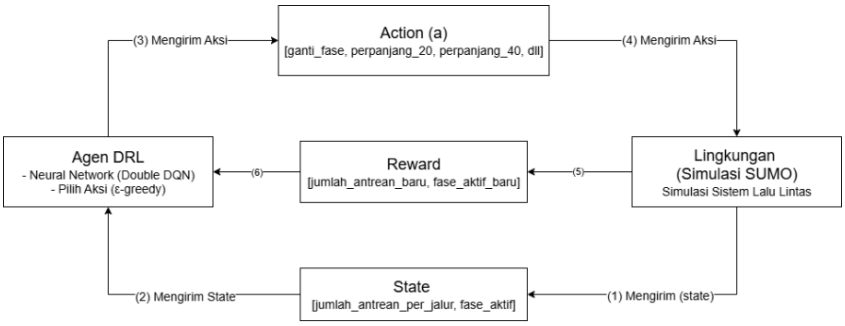


Gambar 1. Model Persimpangan Empat Arah dalam *SUMO GUI.*

Sebagai sistem kontrol pembanding (*baseline*), diimplementasikan dua konfigurasi kontrol lampu lalu lintas waktu tetap (*Fixed-Time Control*) yang beroperasi berdasarkan program statis: Statis-30s (dengan durasi fase hijau 30 detik) dan Statis-60s (dengan durasi fase hijau 60 detik). Kedua sistem ini juga mencakup transisi kuning 3 detik. *Baseline* ini merepresentasikan teknologi konvensional dan berfungsi sebagai tolok ukur objektif untuk memvalidasi kinerja agen *DRL*.

2.2 Perancangan Model Deep Reinforcement Learning

Masalah optimisasi kontrol lalu lintas dimodelkan sebagai *Markov Decision Process (MDP).* Agen cerdas dirancang menggunakan algoritma *Double Deep Q-Network (Double DQN),* yang dipilih untuk mengatasi *overestimation bias* dan meningkatkan kestabilan pelatihan dibandingkan *DQN* standar (Boyko & Mokryk, 2024; Gao et al., 2025; HUANG et al., 2024). Interaksi dinamis antara agen dan lingkungan dalam kerangka *MDP* ini dapat divisualisasikan pada Gambar 2 di bawah ini. Diagram ini mengilustrasikan siklus umpan balik berkelanjutan yang menjadi dasar dari keseluruhan proses pembelajaran agen.



Gambar 2. *Cycle Diagram* Interaksi Agen-Lingkungan.

a. *State (S):* Merepresentasikan kondisi persimpangan sebagai vektor 8 elemen, terdiri dari 4 elemen jumlah kendaraan berhenti *(halting vehicles)* pada setiap jalur masuk dan 4 elemen *one-hot encoding* yang menunjukkan fase hijau yang sedang aktif (Kolat et al., 2023; Liang et al., 2019).

b. *Action (A):* Agen dapat memilih salah satu dari 5 aksi diskrit: Ganti Fase (mengatur durasi sisa ke 0 detik), Perpanjang 20 detik, Perpanjang 40 detik, Perpanjang 60 detik, atau Perpanjang 80 detik. Durasi hijau minimum 15 detik diterapkan untuk menjaga realisme operasional.

JURNAL AI dan SPK
Jurnal Artificial Intelligent dan Sistem Penunjang Keputusan
Jurnal AI dan SPK : Jurnal Artificial Inteligent dan Sistem Penunjang Keputusan
Volume 3, No. 2, Tahun 2025
ISSN 3025-0927 (media online)
Hal 263-274

c. *Reward Function (R):* Didesain untuk meminimalkan "Tekanan" sistem, dihitung sebagai jumlah kuadrat panjang antrean pada setiap jalur masuk.

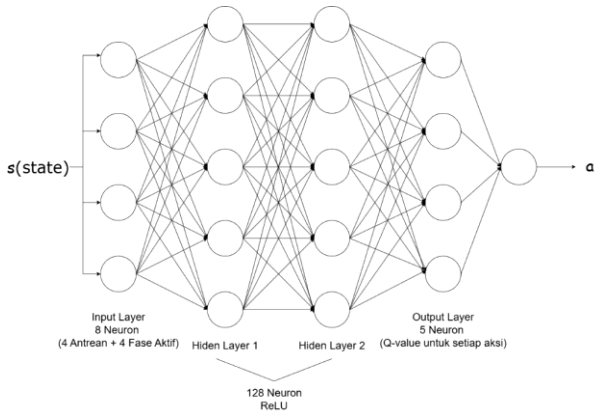
$$Pressure = \Sigma\ (Panjang\_antrean\_i)^2$$

Reward yang diterima agen pada setiap langkah waktu (t) adalah perubahan tekanan:

$$Reward(t) = Pressure(t\text{-}1) - Pressure(t)$$

Reward dibatasi (*clipped*) dalam rentang [-20000, 20000] untuk memastikan stabilitas pelatihan.

Arsitektur jaringan saraf (*neural network*) dari agen *DRL* adalah *Multi-Layer Perceptron (MLP)* dengan tiga lapisan utama: *Input Layer* (8 *neuron*, sesuai ukuran *state*), dua *Hidden Layers* (masing-masing 128 *neuron* dengan fungsi aktivasi *ReLU*), dan *Output Layer* (5 *neuron*, sesuai ukuran *action*, yang menghasilkan nilai *Q-value).*



Gambar 3. *Architecture* Diagram Model *Double Deep Q-Network*

2.3 Prosedur Eksperimen dan Evaluasi

Pelatihan agen *DRL* dilakukan menggunakan bahasa pemrograman *Python 3.13.2* dan *framework PyTorch*, dengan antarmuka *TraCI (Traffic Control Interface)* untuk komunikasi dua arah dengan *SUMO*. Parameter pelatihan *(hyperparameters)* diatur sebagai berikut:

Tabel 1. Parameter Pelatihan (*Hyperparameters*)

Parameter	Nilai
learning_rate	0.0001
gamma	0.95
epsilon_start	0.99
epsilon_decay	0.999
epsilon_min	0.0001
batch_size	64
replay_memory	20.000

JURNAL AI dan SPK
Jurnal Artificial Intelligent dan Sistem Penunjang Keputusan
Jurnal AI dan SPK : Jurnal Artificial Inteligent dan Sistem Penunjang Keputusan
Volume 3, No. 2, Tahun 2025
ISSN 3025-0927 (media online)
Hal 263-274

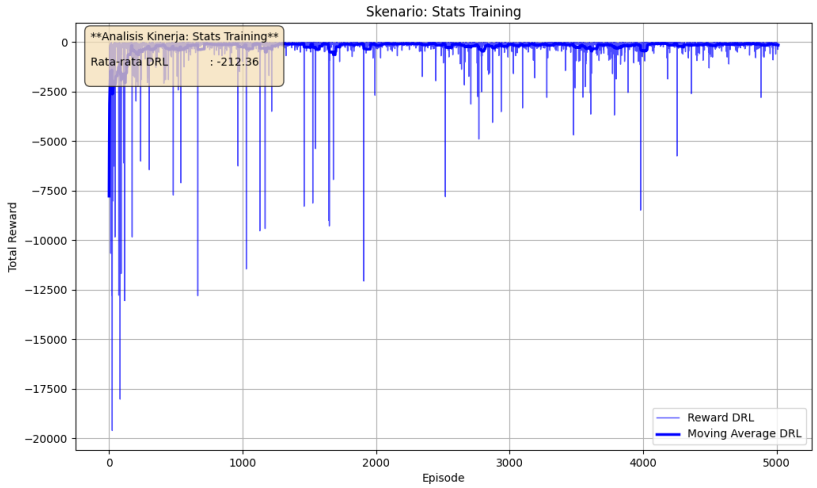
Pelatihan berlangsung selama total 10.000 episode, yang dibagi menjadi dua sesi pelatihan masing-masing 5.000 episode. Dalam setiap episode, skenario lalu lintas dipilih secara acak untuk mendorong generalisasi dan mencegah *overfitting*.

3. ANALISA DAN PEMBAHASAN

Bagian ini menyajikan dan membahas hasil eksperimen komputasional dari agen *Deep Reinforcement Learning (DRL)* dalam mengoptimalkan kontrol sinyal lalu lintas, membandingkannya dengan dua sistem kontrol statis (*Statis-30s dan Statis-60s*). Hasil-hasil utama mencakup proses pelatihan agen, evaluasi kinerja komparatif pada lima skenario lalu lintas, serta pembahasan implikasi penelitian.

3.1 Hasil Pelatihan Agen *DRL*

Proses pelatihan agen *DRL* berlangsung secara *iteratif* selama total 10.000 episode, yang dibagi menjadi dua sesi pelatihan masing-masing 5.000 episode. Tujuan pelatihan adalah agar agen dapat secara otonom menemukan kebijakan (*policy*) optimal untuk meminimalkan "Tekanan" (representasi matematis dari kemacetan) di persimpangan. Gambar 4 menampilkan kurva belajar (*learning curve*) dari sesi pelatihan pertama sebanyak 5.000 episode, *memplot* nilai *Total Reward* kumulatif yang diperoleh agen per episode.



Gambar 4. Kurva Belajar *(Learning curve)* Agen DRL

Kurva belajar menunjukkan dinamika proses pembelajaran agen yang terbagi dalam beberapa fase. Pada Fase Eksplorasi Awal (Episode 0-200), kurva menunjukkan fluktuasi masif dengan *reward* rendah (-10.000 hingga -20.000), merepresentasikan tahap coba-coba acak. Fase Pembelajaran Awal (Episode 200-2000) menunjukkan perbaikan *reward* meskipun fluktuasi masih ada, mengindikasikan agen mulai menghindari kebijakan terburuk. Peningkatan kinerja yang lebih jelas terlihat pada Fase Pematangan (Episode 2000-4000) dengan frekuensi dan tingkat keparahan fluktuasi negatif berkurang. Akhirnya, pada Fase Stabilisasi Kinerja (Episode 4000-5000), kurva menjadi jauh lebih stabil dengan fluktuasi yang lebih sempit, mengindikasikan agen telah mencapai kinerja yang efektif dan stabil serta telah mempelajari kebijakan yang optimal. Stabilitas kinerja kurva ini secara fundamental membuktikan kemampuan *self-learning* agen *DRL*, memvalidasi representasi *state* dan *reward function* berbasis 'Tekanan' yang berhasil memandu agen untuk mengoptimalkan *reward* jangka panjang tanpa instruksi eksplisit.

3.2 Analisis Kinerja Komparatif per Skenario

Setelah pelatihan, model *DRL* dievaluasi secara *head-to-head* dengan sistem Statis-30s dan Statis-60s pada lima skenario lalu lintas yang identik secara *deterministik*, dengan setiap

JURNAL AI dan SPK
Jurnal Artificial Intelligent dan Sistem Penunjang Keputusan
Jurnal AI dan SPK : Jurnal Artificial Inteligent dan Sistem Penunjang Keputusan
Volume 3, No. 2, Tahun 2025
ISSN 3025-0927 (media online)
Hal 263-274

pengujian diulang 100 kali (n=100). Pembahasan di bawah ini menyajikan analisis mendalam untuk setiap skenario, dilengkapi dengan tabel perbandingan kinerja spesifik per skenario.

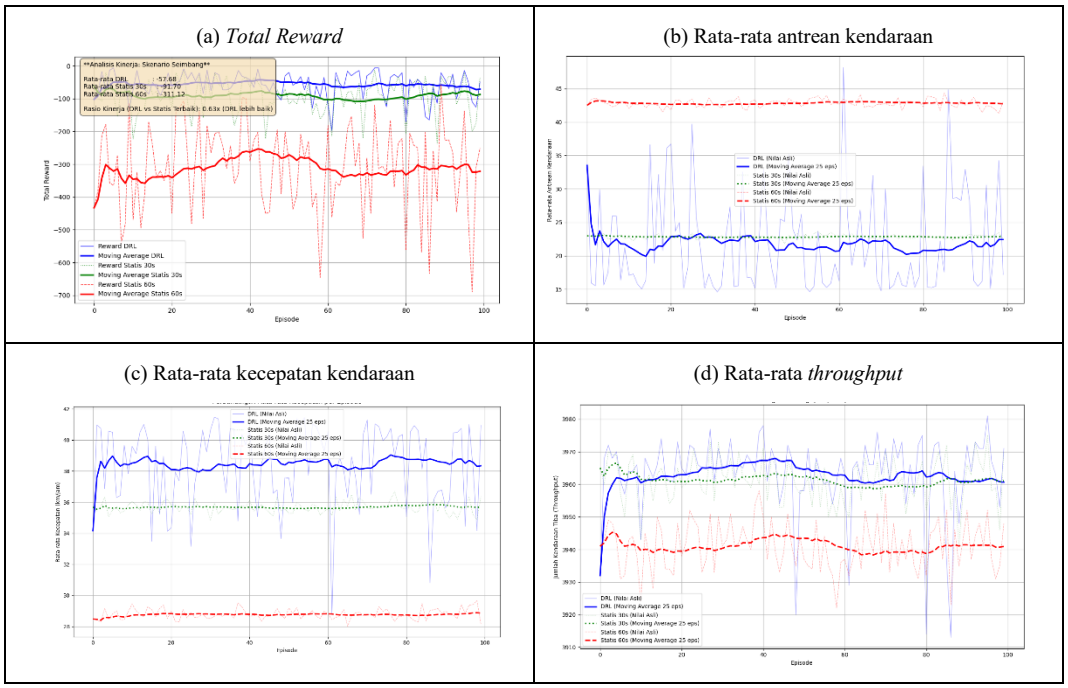
3.2.1 Skenario 1: Lalu Lintas Seimbang

Pada Skenario 1, yang merepresentasikan kondisi lalu lintas seimbang, agen *DRL* menunjukkan keunggulan kinerja yang jelas pada metrik-metrik efisiensi. Tabel 2 menyajikan perbandingan kinerja di skenario ini.

Tabel 2. Perbandingan Kinerja Skenario Lalu Lintas Seimbang (n=100)

Mode	*Total Reward*	Antrean Kendaraan (rata-rata)	Kecepatan (km/h) (rata-rata)	*Throughput* (rata-rata)
DRL	-57.68	21.91	38.42	3963.29
Statis-60s	-311.12	42.81	28.79	3941.04
Statis-30s	-91.70	22.82	35.67	3961.1

DRL (-57.68) menunjukkan *Total Reward* yang lebih baik dibandingkan Statis-30s (-91.70) dan Statis-60s (-311.12). Performa ini berakar dari kemampuan agen untuk melakukan optimisasi mikro secara berkelanjutan sebagai respons terhadap perubahan acak pada antrean. Efisiensi ini berdampak langsung pada kelancaran arus, tercermin dari rata-rata kecepatan *DRL* (38.42 km/jam) yang 7.71% lebih tinggi dari Statis-30s, dan rata-rata antrean yang lebih pendek (21.91 kendaraan). Visualisasi dinamis pergerakan *Total Reward* (Gambar 5(a)), antrean (Gambar 5(b)), kecepatan (Gambar 5(c)), dan *throughput* (Gambar 5(d)) selama 100 episode menunjukkan bahwa *DRL* secara konsisten berada pada rentang kinerja yang lebih tinggi.



Gambar 5. Perbandingan Kinerja Sistem pada Skenario Seimbang: (a) *Total Reward*, (b) Rata-rata antrean kendaraan, (c) Rata-rata kecepatan kendaraan, (d) Rata-rata *throughput.*

3.2.2 Skenario 2: Lalu Lintas Padat 2 Arah

JURNAL AI dan SPK
Jurnal Artificial Intelligent dan Sistem Penunjang Keputusan

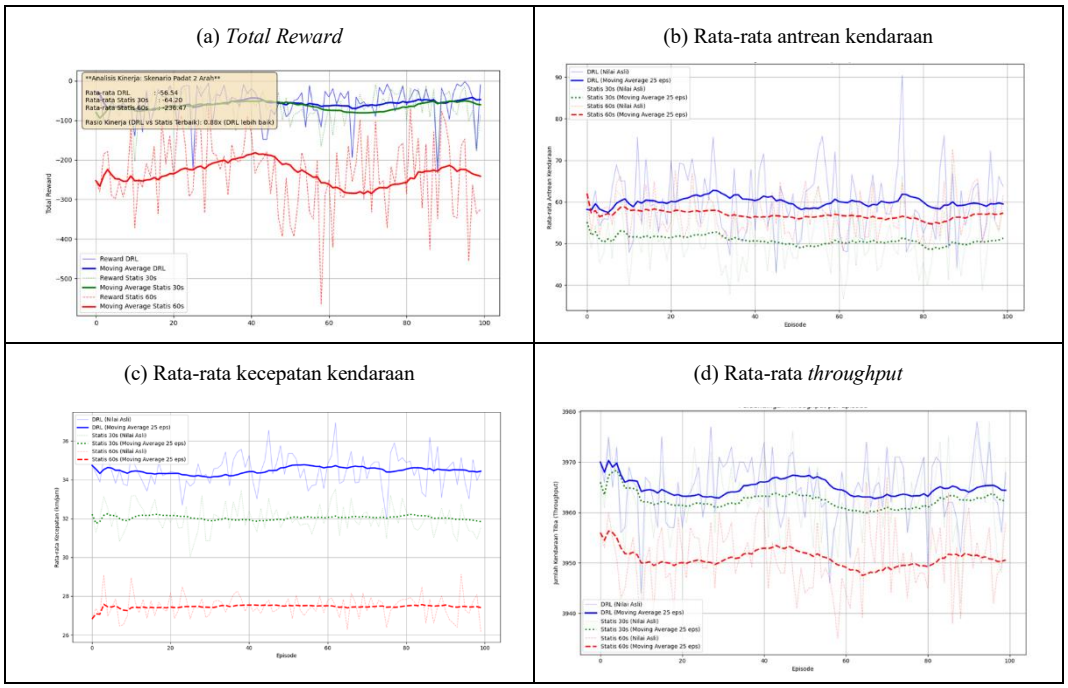
Jurnal AI dan SPK : Jurnal Artificial Inteligent dan Sistem Penunjang Keputusan
Volume 3, No. 2, Tahun 2025
ISSN 3025-0927 (media online)
Hal 263-274

Pada Skenario 2, yang dirancang untuk mensimulasikan kondisi lalu lintas padat dari dua arah, agen DRL kinerja strategis yang baik. Tabel 3 menyajikan perbandingan kinerja di skenario ini.

Tabel 3. Perbandingan Kinerja Skenario Lalu Lintas Padat 2 Arah (n=100)

Mode	*Total Reward*	Antrean Kendaraan (rata-rata)	Kecepatan (km/h) (rata-rata)	*Throughput* (rata-rata)
DRL	-56.54	60.05	34.45	3964.,7
Statis-60s	-236.47	56.85	27.46	3950.71
Statis-30s	-64.20	50.79	31.98	3962.15

Agen *DRL* berhasil mencapai *Total Reward* rata-rata tertinggi (-56.54), mengungguli Statis-30s (-64.20) dan Statis-60s (-236.47). Meskipun rata-rata antrean *DRL* (60.05 kendaraan) tercatat sedikit lebih tinggi dari Statis-30s (50.79 kendaraan), keunggulan *DRL* justru terlihat jelas pada rata-rata kecepatan tertinggi (34.45 km/jam) dan *throughput* terbanyak (3964.7 kendaraan). Ini mengindikasikan bahwa agen telah mempelajari strategi optimisasi global; tidak hanya bertujuan memperpendek antrean, tetapi memaksimalkan produktivitas persimpangan dengan menoleransi antrean yang sedikit lebih panjang demi meloloskan jumlah kendaraan terbanyak (Gambar 6(b), 6(c), 6(d)).



Gambar 6. Perbandingan Kinerja Sistem pada Skenario Padat dari 2 Arah: (a) *Total Reward*, (b) Rata-rata antrean kendaraan, (c) Rata-rata kecepatan kendaraan, (d) Rata-rata *throughput.*

3.2.3 Skenario 3: Lalu Lintas Padat 3 Arah

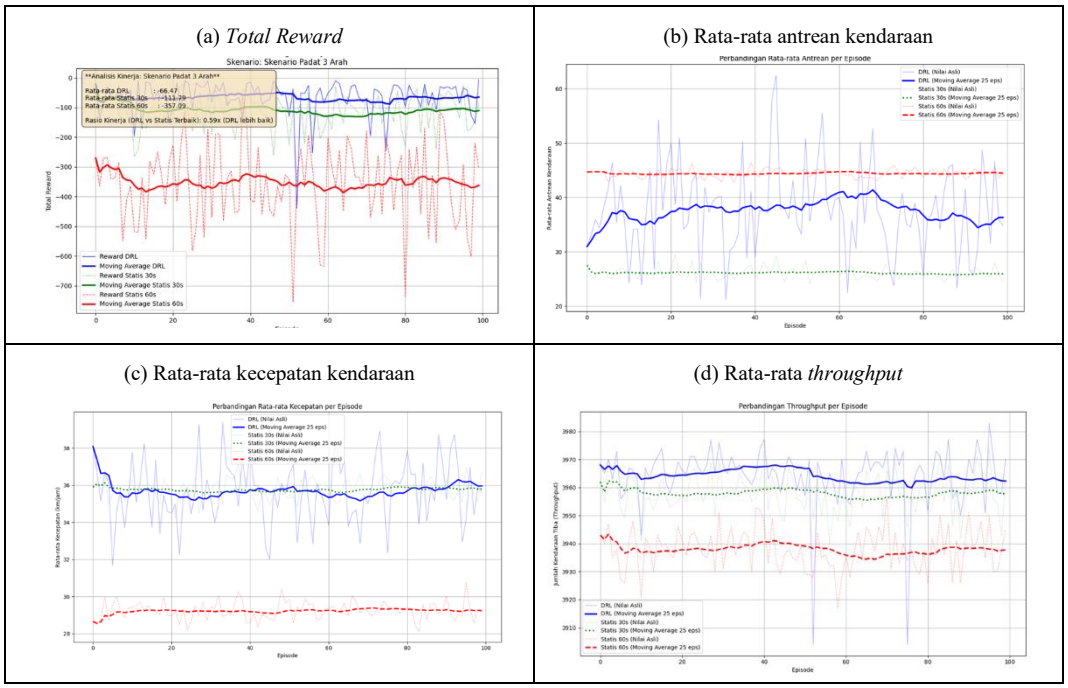
Pada Skenario 3, yang mensimulasikan kondisi volume yang tinggi parah dari tiga arah, agen *DRL* kembali membuktikan keunggulan strategisnya. Tabel 4 menyajikan perbandingan kinerja di skenario ini.

JURNAL AI dan SPK
Jurnal Artificial Intelligent dan Sistem Penunjang Keputusan
Jurnal AI dan SPK : Jurnal Artificial Inteligent dan Sistem Penunjang Keputusan
Volume 3, No. 2, Tahun 2025
ISSN 3025-0927 (media online)
Hal 263-274

Tabel 4. Perbandingan Kinerja Skenario Lalu Lintas Padat 3 Arah (n=100)

Mode	*Total Reward*	Antrean Kendaraan (rata-rata)	Kecepatan (km/h) (rata-rata)	*Throughput* (rata-rata)
DRL	-66.47	37.62	35.68	3964.23
Statis-60s	-357.09	44.40	29.24	3937.94
Statis-30s	-111.79	26.06	35.75	3957.92

DRL berhasil memperoleh *Total Reward* rata-rata tertinggi (-66.47), melampaui Statis-60s (-357.09) dan Statis-30s (-111.79) secara signifikan. Mirip dengan Skenario 2, Statis-30s berhasil mencapai rata-rata antrean terendah (26.06 kendaraan) dan kecepatan sedikit lebih tinggi (35.75 km/jam). Namun, kinerja yang lebih optimal dari *DRL* terletak pada *throughput* maksimal (3964.23 kendaraan). Ini mengindikasikan bahwa agen *DRL* mempelajari strategi yang lebih adaptif, tidak terjebak pada solusi lokal melainkan fokus pada tujuan global yaitu memaksimalkan produktivitas dan kelancaran arus total sistem persimpangan (Gambar 7(b), 7(c), 7(d)).



Gambar 7. Perbandingan Kinerja Sistem pada Skenario Padat 3 Arah: (a) *Total Reward*, (b) Rata-rata antrean kendaraan, (c) Rata-rata kecepatan kendaraan, (d) Rata-rata *throughput.*

3.2.4 Skenario 4: Lalu Lintas Padat 1 Arah

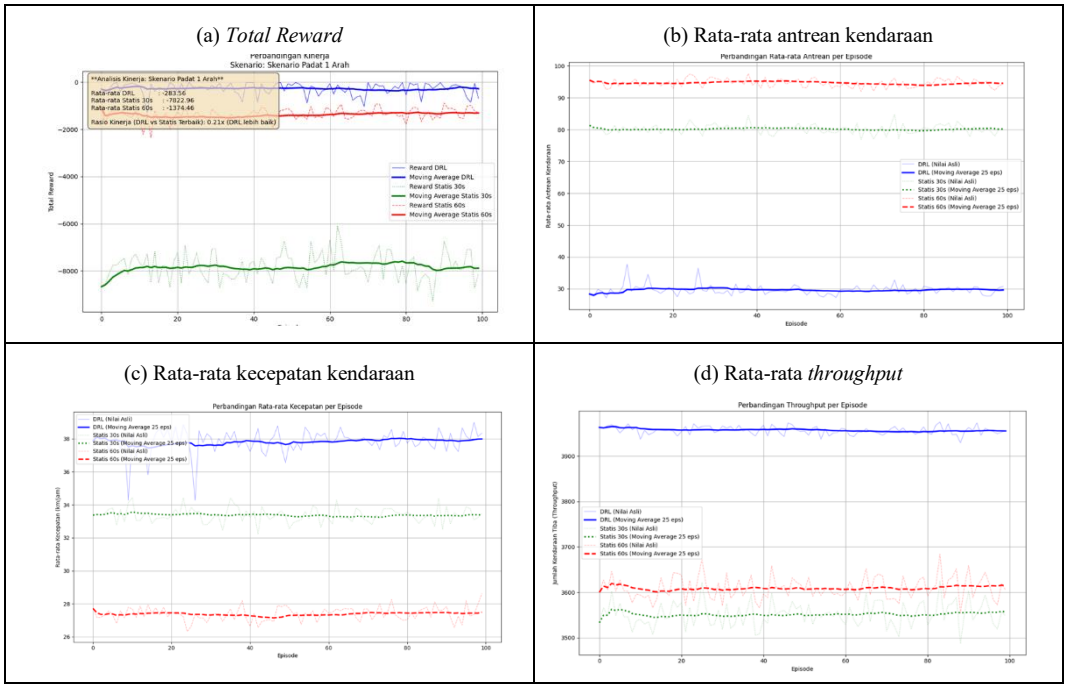
Dalam Skenario 4, yang dirancang untuk menguji respons sistem terhadap kondisi lalu lintas sangat padat dari satu arah utama, agen *DRL* menunjukkan kinerja yang konsisten lebih baik. Tabel 5 menyajikan perbandingan kinerja di skenario ini.

JURNAL AI dan SPK
Jurnal Artificial Intelligent dan Sistem Penunjang Keputusan
Jurnal AI dan SPK : Jurnal Artificial Inteligent dan Sistem Penunjang Keputusan
Volume 3, No. 2, Tahun 2025
ISSN 3025-0927 (media online)
Hal 263-274

Tabel 5. Perbandingan Kinerja Skenario Lalu Lintas Padat 1 Arah (n=100)

Mode	*Total Reward*	Antrean Kendaraan (rata-rata)	Kecepatan (km/h) (rata-rata)	*Throughput* (rata-rata)
DRL	-283.56	29.66	37.83	3956.5
Statis-60s	-1374.5	94.65	27.39	3609.3
Statis-30s	-7823	80.16	33.36	3551.73

Agen *DRL* mencatatkan *Total Reward* (-283.56) yang secara signifikan lebih baik dibandingkan Statis-60s (-1374.5) dan Statis-30s (-7823). Pada skenario ini, *DRL* menunjukkan kinerja yang lebih baik pada semua metrik efisiensi: mengurangi rata-rata antrean lebih dari 68% (menjadi 29.66 kendaraan), menghasilkan kecepatan rata-rata tertinggi (37.83 km/jam), dan mencapai *throughput* tertinggi (3956.5 kendaraan). Keberhasilan ini berakar pada kemampuan adaptif *DRL* untuk secara cerdas mengalokasikan waktu hijau ke arah yang paling membutuhkan (Gambar 8(b), 8(c), 8(d)).



Gambar 8. Perbandingan Kinerja Sistem pada Skenario Sangat Padat dari 1 Arah: (a) *Total Reward*, (b) Rata-rata antrean kendaraan, (c) Rata-rata kecepatan kendaraan, (d) Rata-rata *throughput.*

3.2.5 Skenario 5: Skenario Acak (Uji Generalisasi)

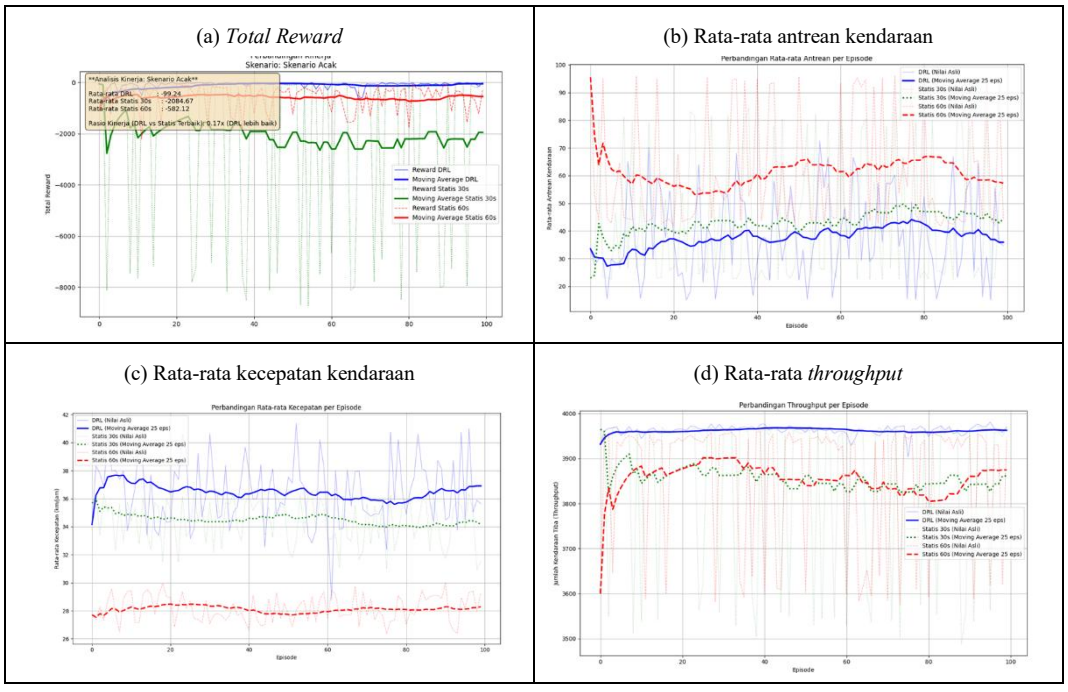
Skenario 5 merupakan puncak dari pengujian, dirancang khusus untuk mengevaluasi kemampuan generalisasi model *DRL* pada kondisi lalu lintas yang berubah-ubah secara acak. Tabel 6 menyajikan perbandingan kinerja di skenario ini.

JURNAL AI dan SPK
Jurnal Artificial Intelligent dan Sistem Penunjang Keputusan
Jurnal AI dan SPK : Jurnal Artificial Inteligent dan Sistem Penunjang Keputusan
Volume 3, No. 2, Tahun 2025
ISSN 3025-0927 (media online)
Hal 263-274

Tabel 6. Perbandingan Kinerja Skenario Acak (n=100)

Mode	*Total Reward*	Antrean Kendaraan (rata-rata)	Kecepatan (km/h) (rata-rata)	*Throughput* (rata-rata)
DRL	-99.24	38.20	36.46	3963.11
Statis-60s	-582.12	60.41	28.16	3858.81
Statis-30s	-2084.7	43.84	34.38	3852.98

Hasil menunjukkan kinerja yang konsisten lebih baik dari agen *DRL* dengan *Total Reward* rata-rata -99.24, yang melampaui Statis-60s (-582.12) dan Statis-30s (-2084.7). *DRL* mencatatkan panjang antrean rata-rata terendah (38.20 kendaraan), kecepatan rata-rata tertinggi (36.46 km/jam), dan *throughput* tertinggi (3963.11 kendaraan). Gambar 9(a), 9(b), 9(c), dan 9(d) secara visual menegaskan stabilitas sistem *DRL*; meskipun dihadapkan pada kondisi acak, kurvanya secara konsisten berada pada rentang *reward* yang tinggi dengan fluktuasi minimal. Ini sangat kontras dengan Statis-60s dan Statis-30s yang menunjukkan volatilitas tinggi dan kinerja yang sangat bergantung pada kondisi spesifik. Keberhasilan *DRL* dalam skenario ini adalah bukti penting dari ketangguhan model, menunjukkan bahwa ia tidak hanya menghafal solusi, tetapi telah 'belajar' prinsip-prinsip dasar manajemen lalu lintas yang efisien.



Gambar 9. Perbandingan Kinerja Sistem pada Skenario Acak: (a) *Total Reward*, (b) Rata-rata antrean kendaraan, (c) Rata-rata kecepatan kendaraan, (d) Rata-rata *throughput*.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan dan mengimplementasikan sistem kontrol sinyal lalu lintas adaptif berbasis *Deep Reinforcement Learning (DRL)* menggunakan algoritma *Double Deep Q-Network (Double DQN)* dalam lingkungan simulasi *SUMO*. Perancangan komponen utama seperti representasi *state, action,* dan *reward function* berbasis 'Tekanan' terbukti efektif dalam memandu agen untuk mempelajari kebijakan optimal. Proses pelatihan menunjukkan stabilitas dan

JURNAL AI dan SPK
Jurnal Artificial Intelligent dan Sistem Penunjang Keputusan
Jurnal AI dan SPK : Jurnal Artificial Inteligent dan Sistem Penunjang Keputusan
Volume 3, No. 2, Tahun 2025
ISSN 3025-0927 (media online)
Hal 263-274

kinerja yang semakin optimal, memvalidasi pendekatan *Double DQN* untuk manajemen lalu lintas yang dinamis.

Melalui serangkaian pengujian komparatif pada lima skenario lalu lintas yang berbeda, agen *DRL* secara konsisten menunjukkan kinerja yang lebih baik dibandingkan sistem kontrol waktu tetap (Statis-30s dan Statis-60s). *DRL* berhasil mengatasi ketidakcocokan antara alokasi waktu hijau yang kaku dengan volume lalu lintas yang berubah-ubah, menyesuaikan durasi lampu hijau secara *real-time* untuk mengurangi kemacetan. Strategi yang dipelajari agen *DRL* cenderung mengutamakan optimisasi global seperti *throughput* dan kecepatan rata-rata, meskipun terkadang menoleransi antrean yang sedikit lebih panjang pada jalur tertentu, yang pada akhirnya menghasilkan *Total Reward* yang lebih tinggi.

Pada skenario acak, yang dirancang untuk menguji kemampuan generalisasi, agen *DRL* menunjukkan peningkatan kinerja yang signifikan. *DRL* berhasil menurunkan rata-rata antrean hingga 36.76%, meningkatkan rata-rata kecepatan kendaraan hingga 29.47%, dan meningkatkan *throughput* hingga 2.70% dibandingkan dengan sistem Statis-60s. Bahkan terhadap Statis-30s, *DRL* tetap unggul dengan penurunan rata-rata antrean 12.86%, peningkatan rata-rata kecepatan 6.05%, dan peningkatan *throughput* 2.86%. Hasil ini membuktikan efektivitas dan ketangguhan pendekatan *DRL* dengan *Double DQN* sebagai solusi untuk optimisasi sinyal lalu lintas adaptif.

REFERENCES

Astuti, S., Manalu, R., Simamora, E., Ulina, E., Hombing, B., Maulana, J., Sarah, K., Siahaan, A., Manalu, M. E., & Ekonomi, I. (2024). Analisis Faktor Penyebab Kemacetan Lalu Lintas Di Simpang Empat Unimed Mmtc Medan Analysis Of Factors Caused By Traffic Construction At The Unimed Middle Mixation Of Mmtc Medan. https://jicnusantara.com/index.php/jiic

Boyko, N., & Mokryk, Y. (2024). Optimizing Traffic at Intersections With Deep Reinforcement Learning. *Journal of Engineering*, *2024*(1), 6509852. https://doi.org/10.1155/2024/6509852

Brilliant, M. G., Meilala, R. R. S., & Herwanis, D. (2024). Manajemen Transportasi: Kerugian Transportasi Akibat Kemacetan Lalu Lintas di Aceh. *Sammajiva: Jurnal Penelitian Bisnis Dan Manajemen*, *2*(4), 42–53.

Damadam, S., Zourbakhsh, M., Javidan, R., & Faroughi, A. (2022). An Intelligent IoT Based Traffic Light Management System: Deep Reinforcement Learning. *Smart Cities*, *5*(4). https://doi.org/10.3390/smartcities5040066

Gao, S., Ding, J., Fu, L., & Wang, X. (2025). *Controlling Underestimation Bias in Constrained Reinforcement Learning for Safe Exploration*. https://proceedings.mlr.press/v267/gao25b.html

HUANG, F., HE, Y., ZHANG, Y., DENG, X., & JIANG, W. (2024). Controlling underestimation bias in reinforcement learning via minmax operation. *Chinese Journal of Aeronautics*, *37*(7), 406–417. https://doi.org/10.1016/j.cja.2024.03.008

Inayah, S., Yusuf, I. A. W., & Farihin, A. (2025). Faktor Penyebab Kemacetan Lalu Lintas di Kecamatan Gedebage Kota Bandung. *Sosiosaintika*, *3*(1), 11–20.

Kolat, M., Kővári, B., Bécsi, T., & Aradi, S. (2023). Multi-Agent Reinforcement Learning for Traffic Signal Control: A Cooperative Approach. *Sustainability*, *15*(4). https://doi.org/10.3390/su15043479

Li, Z., Yu, H., Zhang, G., Dong, S., & Xu, C.-Z. (2021). *Network-wide traffic signal control optimization using a multi-agent deep reinforcement learning*. https://doi.org/10.1016/j.trc.2021.103059

Liang, X., Du, X., Wang, G., & Han, Z. (2019). A Deep Reinforcement Learning Network for Traffic Light Cycle Control. *IEEE Transactions on Vehicular Technology*, *68*(2). https://doi.org/10.1109/TVT.2018.2890726

Liu, D., & Li, L. (2023). A traffic light control method based on multi-agent deep reinforcement learning algorithm. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-36606-2

Lopez, P. A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flotterod, Y. P., Hilbrich, R., Lucken, L., Rummel, J., Wagner, P., & Wiebner, E. (2018). Microscopic Traffic Simulation using SUMO. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, *2018-November*. https://doi.org/10.1109/ITSC.2018.8569938

Maha, E. (2022). Analisis faktor-faktor pendorong penyebab terjadinya kemacetan di kawasan Pajus Padang Bulan Medan. *Jurnal Samudra Geografi*, *5*(1), 38–42.

Matsuo, Y., LeCun, Y., Sahani, M., Precup, D., Silver, D., Sugiyama, M., Uchibe, E., & Morimoto, J. (2022). Deep learning, reinforcement learning, and world models. *Neural Networks*, *152*, 267–275. https://doi.org/10.1016/j.neunet.2022.03.037

JURNAL AI dan SPK
Jurnal Artificial Intelligent dan Sistem Penunjang Keputusan

Jurnal AI dan SPK : Jurnal Artificial Inteligent dan Sistem Penunjang Keputusan
Volume 3, No. 2, Tahun 2025
ISSN 3025-0927 (media online)
Hal 263-274

Putra, R., Syahbana, Y., & Ananda, A. (2024). Implementasi Algoritma Deep Q-Network (DQN) pada Lampu Lalu Lintas Adaptif Berdasarkan Waktu Tunggu dan Arus Kendaraan. *The Indonesian Journal of Computer Science*, *13*. https://doi.org/10.33022/ijcs.v13i5.4372

Safira, E., & Khuluqi, S. F. (2023). Analisis Tingkat Kemacetan Dan Faktor Penyebab Kemacetan Lalu Lintas Di Jalan Sultan Hamid Ii Kecamatan Pontianak Selatan.