PPO-Based Intelligent Traffic Signal Management with Computer Vision Integration

Rupali Shankar Kale[1]; Dr. Manisha Bharati[2]

[1]Department of Technology SPPU, Pune, India
[2]Department of Technology SPPU, Pune, India

Publication Date: 2026/05/30

Abstract: Urban traffic congestion is a major challenge due to increasing vehicle density and the limitations of conventional fixed time traffic signal systems. Traditional signal controllers lack adaptability and fail to respond effectively to dynamic traffic conditions, especially in heterogeneous urban traffic environments such as Indian roads. This research proposes an intelligent adaptive traffic signal control framework integrating Deep Reinforcement Learning (DRL) with computer vision-based traffic perception. The proposed system employs Proximal Policy Optimization (PPO) as the primary learning algorithm and compares its performance with Deep Q-Network (DQN) and Advantage Actor-Critic (A2C). Traffic flow is simulated using the SUMO traffic simulator and controlled in real time through the TraCI interface. Vehicle density, lane occupancy, and congestion states are extracted using YOLOv8-based perception from rendered simulation frames. Experimental analysis demonstrates that the PPO-based adaptive controller significantly reduces average waiting time and queue length while improving traffic throughput compared to conventional fixed time systems and other reinforcement learning approaches.

Keywords: *Traffic Signal Control, Reinforcement Learning, PPO, YOLOv8, SUMO, TraCI, Intelligent Transportation Systems, Deep Learning.*

How to Cite: Rupali Shankar Kale; Dr. Manisha Bharati (2026) PPO-Based Intelligent Traffic Signal Management with Computer Vision Integration. *International Journal of Innovative Science and Research Technology*, 11(5), 2333-2339. https://doi.org/10.38124/ijisrt/26may1200

I. INTRODUCTION

Rapid urbanization and increasing vehicle ownership have resulted in severe traffic congestion in metropolitan cities. Conventional traffic signal systems operate using fixed timing schedules, which are incapable of adapting to real-time traffic fluctuations. These systems often lead to inefficient road utilization, prolonged waiting times, excessive fuel consumption, and increased emissions.

Artificial Intelligence (AI), particularly Reinforcement Learning (RL), has emerged as a promising solution for intelligent traffic management. Reinforcement learning enables an agent to learn optimal decision-making strategies through continuous interaction with the environment.

This research presents an adaptive traffic signal optimization framework that combines Deep Reinforcement Learning with computer vision-based traffic perception. The system utilizes YOLOv8 for vehicle detection and Proximal Policy Optimization (PPO) for intelligent traffic signal phase control.

➢ *Problem Statement*

Traditional traffic signal systems rely on static timing mechanisms that fail to adapt to dynamic traffic conditions. Such systems result in inefficient signal utilization, excessive queue formation, and increased travel delays.

There is a need for an intelligent adaptive traffic control system capable of learning from traffic behavior and optimizing signal timing decisions in real time.

➢ *Motivation*

The motivation behind this work lies in overcoming the inefficiencies of fixed-time traffic signal systems using AI driven adaptive decision-making.

- *Recent Advances in Reinforcement Learning and Computer Vision Make it Feasible to Design Traffic Systems Capable of:*

✓ Learning optimal traffic control strategies
✓ Detecting vehicles in real time
✓ Reducing congestion dynamically
✓ Improving transportation efficiency

II. LITERATURE REVIEW

➢ Li et al. [1] demonstrated deep reinforcement learning for adaptive traffic signal timing optimization.

- Wei et al. [2] proposed IntelliLight, a DRL-based traffic control framework using state-aware adaptive learning.
- Arel et al. [3] explored multi-agent reinforcement learning for traffic optimization.
- Van der Pol and Oliehoek [4] investigated coordinated deep reinforcement learning for traffic light control.
- Most prior approaches rely heavily on simulator-generated state representations and lack real-time visual traffic perception.

III. RESEARCH GAP

Existing traffic optimization systems predominantly depend on predefined simulator states rather than visual traffic understanding.

- *The Major Gaps Identified are:*

- Lack of integration between real-time computer vision and RL
- Limited PPO-based traffic signal optimization studies
- Insufficient training performance analysis using TensorBoard
- Limited comparison among modern RL algorithms

This work addresses these limitations by integrating YOLOv8-based vehicle perception with PPO-based adaptive traffic control.

- *Objectives*

- Develop an adaptive traffic signal control system using PPO
- Integrate YOLOv8 for real-time vehicle detection
- Compare PPO with DQN and A2C algorithms
- Reduce average waiting time and queue length
- Improve traffic throughput and control efficiency

IV. PROPOSED METHODOLOGY

- *System Architecture*

The proposed framework integrates traffic simulation, vehicle detection, reinforcement learning, and performance monitoring.

- *The Workflow Includes:*

✓ SUMO traffic simulation environment generation
✓ Frame acquisition from traffic simulation
✓ Vehicle detection using YOLOv8
✓ State representation extraction
✓ Reward computation
✓ PPO policy optimization
✓ Traffic signal action selection
✓ Performance logging using TensorBoard

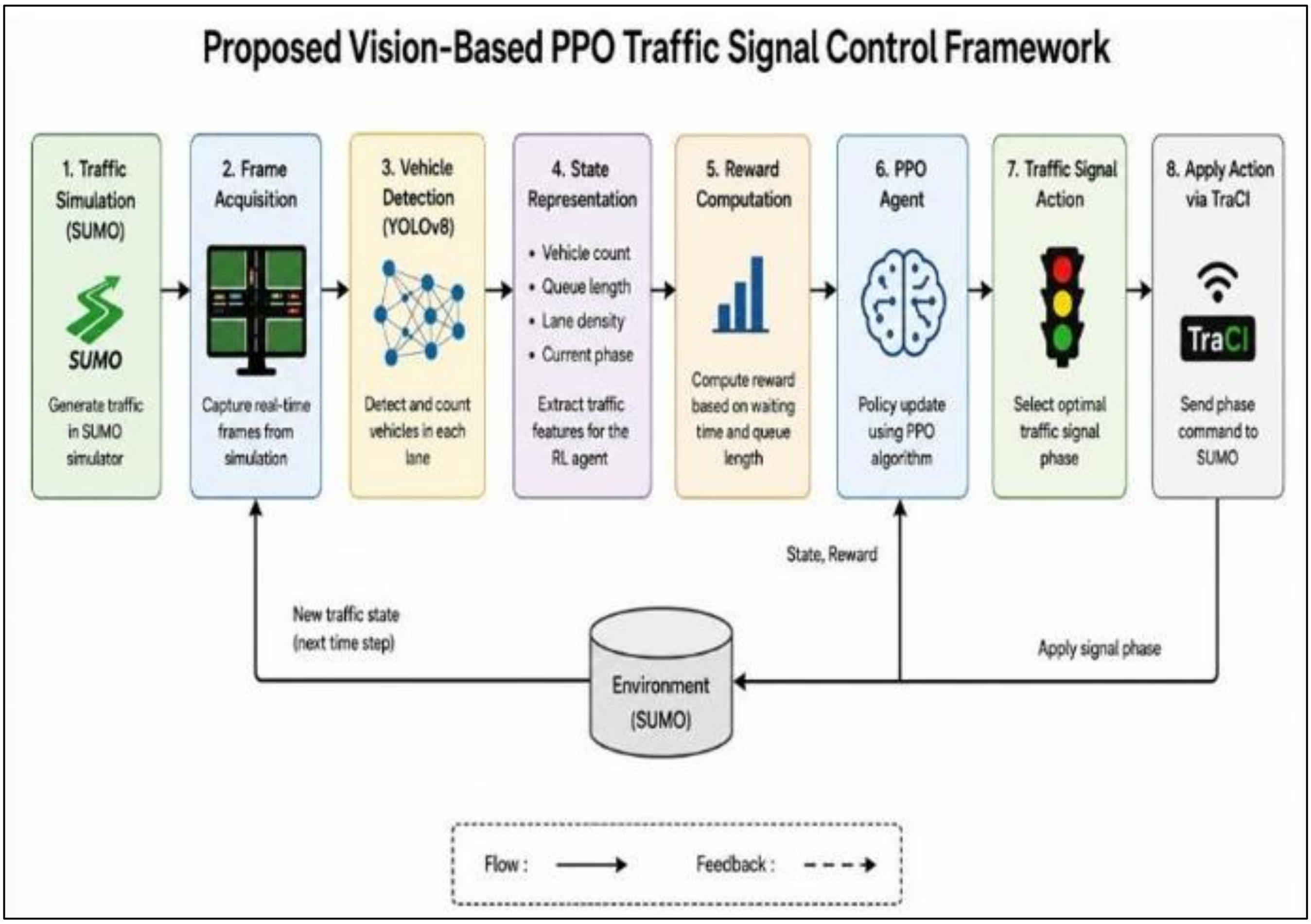


Fig 1 Architecture of the Proposed YOLOv8 + PPO Based Traffic Signal Control System

- *Reinforcement Learning Formulation* State Representation:

$St = \{vehicle\ count,\ queue\ length,\ lane\ density\}$ Action Space:

$At = \{Traffic\ Signal\ Phase\ Selection\}$ Reward Function:

$Rt = -(w1 \times WaitingTime + w2 \times QueueLength)$ where:

- $w1$ = waiting time weight
- $w2$ = queue length weight VIII. ALGORITHM

Algorithm 1 PPO-Based Traffic Signal Control

```
Initialize PPO policy network
Initialize value network
for each episode do
    Reset SUMO environment
    for each timestep do
        Capture current traffic state
        Detect vehicles using YOLOv8
        Extract queue and density features
        Select traffic phase action using PPO
        Apply selected phase through TraCI
        Observe reward
        Store transition
    end for
    Update PPO policy using collected trajectories
end for
```

Table 1 Performance Comparison

Method	Waiting Time (s)	Queue Length	Throughput
Fixed-Time	78.5	24.3	Low
DQN	52.1	16.8	Medium
A2C	48.6	15.2	Medium
PPO	35.4	10.6	High

V. EXPERIMENTAL SETUP

Experiments were conducted using a four-way traffic intersection modeled in SUMO.

- *Implementation Environment:*

- SUMO Traffic Simulator
- Python
- Stable-Baselines3
- TensorBoard
- YOLOv8
- OpenCV
- TraCI Interface Algorithms evaluated:
- Fixed-Time Controller
- DQN
- A2C
- PPO

VI. RESULTS AND DISCUSSION

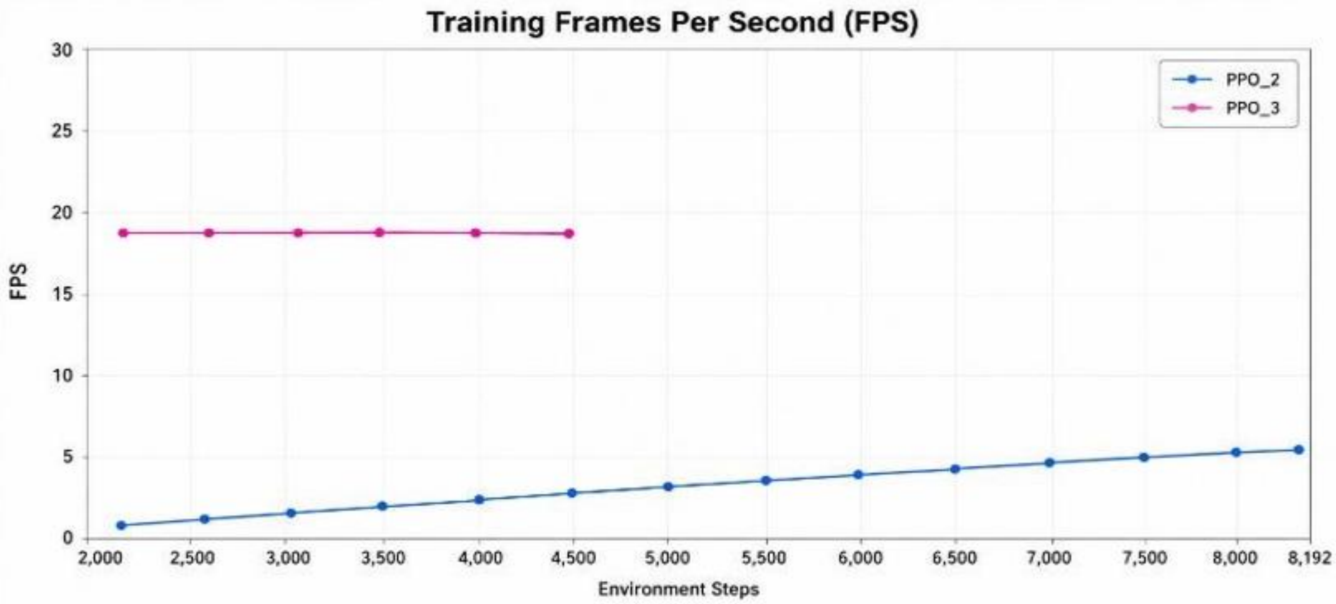
- *Training Performance*

TensorBoard was used for analyzing reinforcement learning training metrics including FPS, training loss, policy gradient behavior, entropy loss, and optimization stability.

- *Observations:*

- ✓ PPO training remained stable throughout learning
- ✓ Policy optimization converged smoothly
- ✓ Loss reduction indicates effective learning
- ✓ FPS remained consistent, indicating efficient simulation execution

- *FPS Performance*



Run ↑	Min	Max	Start Value	End Value	Δ Value	Δ %	Start Step	End Step
PPO_2	0.84	5.40	3.976	5.40	0.144	3.62%	8,192	8,192
PPO_3	19.00	19.00	19.00	19.00	0.00	0.00%	4,096	4,096

Fig 2 Training FPS Performance of PPO Agent in SUMO Environment

- *PPO Policy Optimization Metrics*

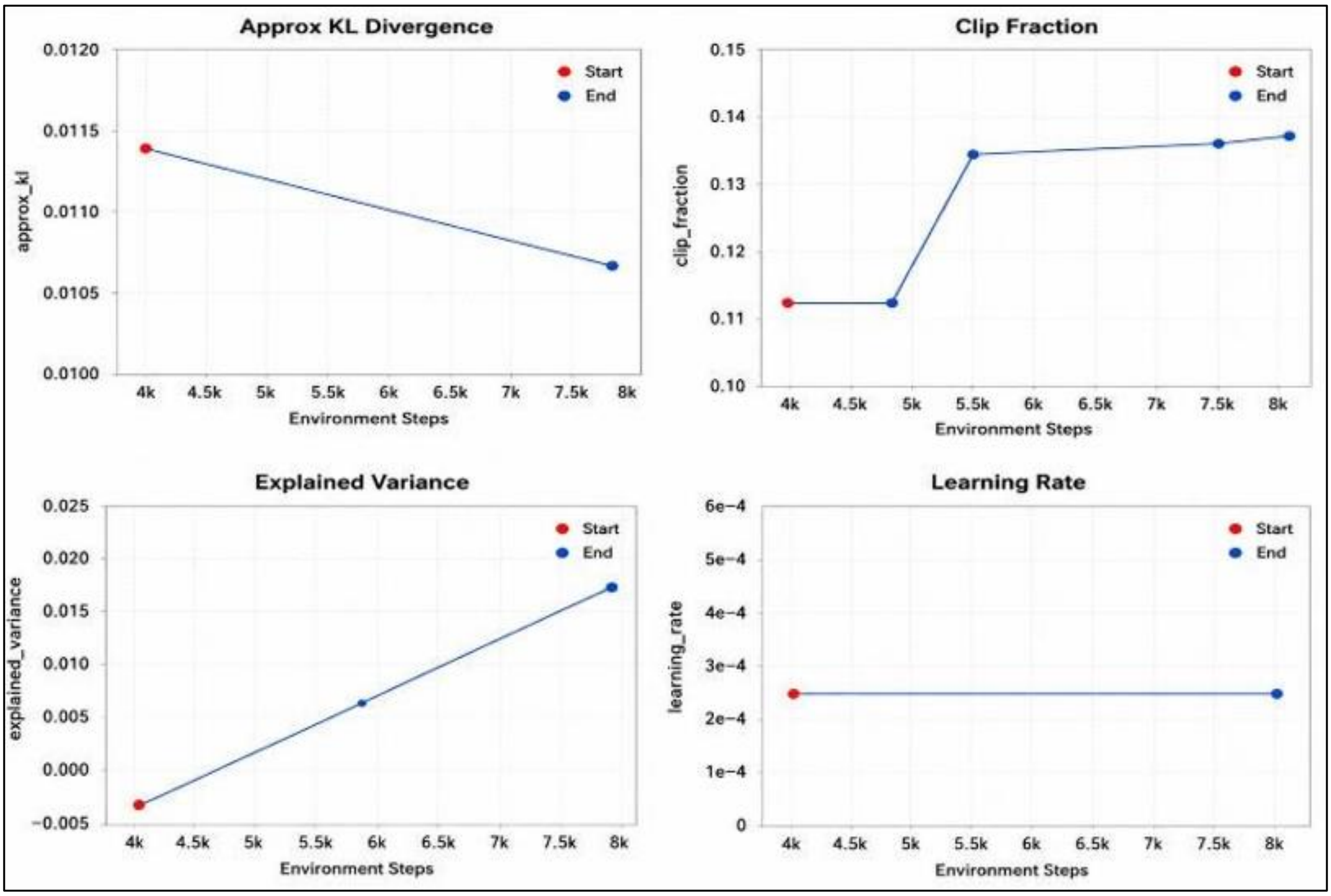


Fig 3 PPO Policy Optimization Metrics: KL Divergence, Clip Fraction, Explained Variance, and Learning Rate

- *PPO Training Stability Metrics*

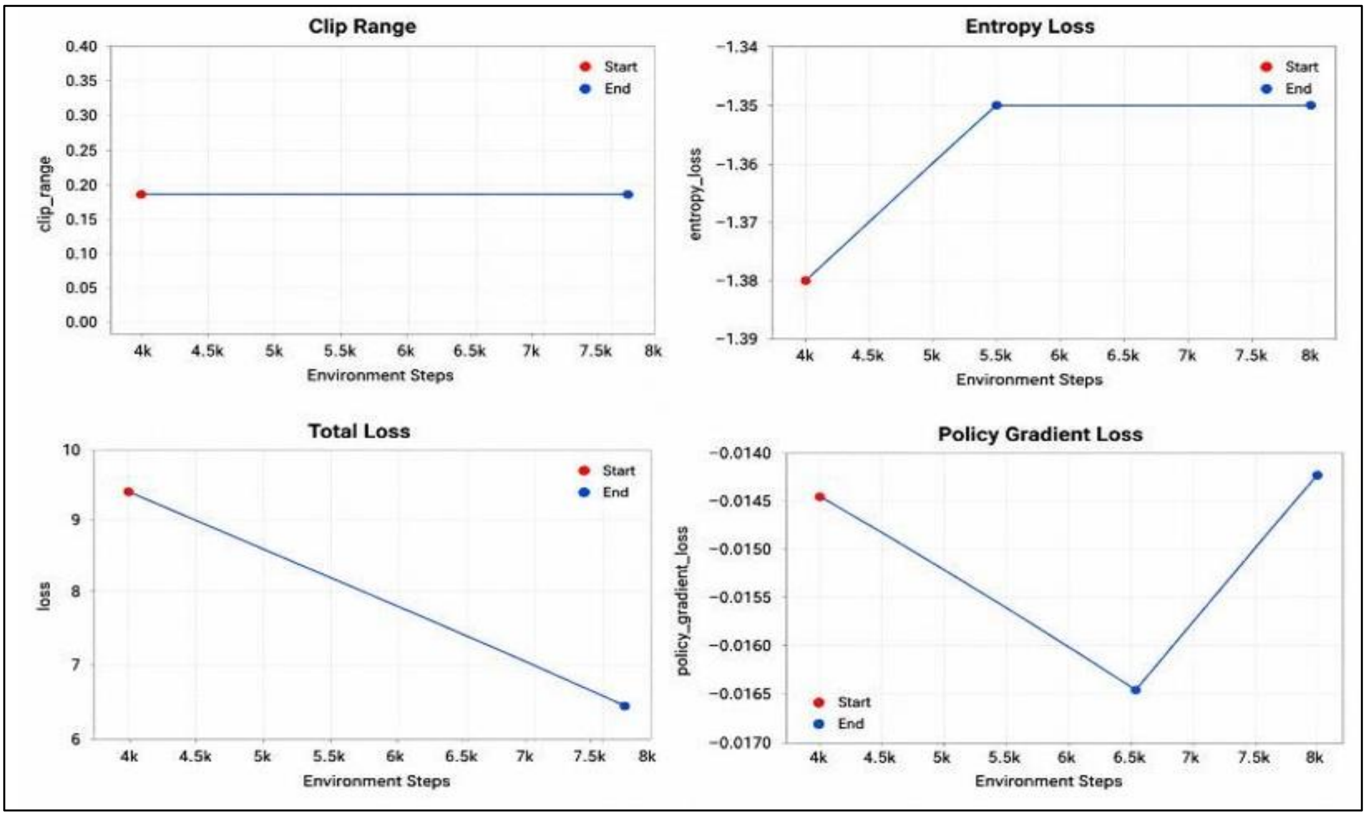


Fig 4 PPO Training Stability Metrics: Clip Range, Entropy Loss, Total Loss, and Policy Gradient Loss

- *Entropy / Stability Metrics*

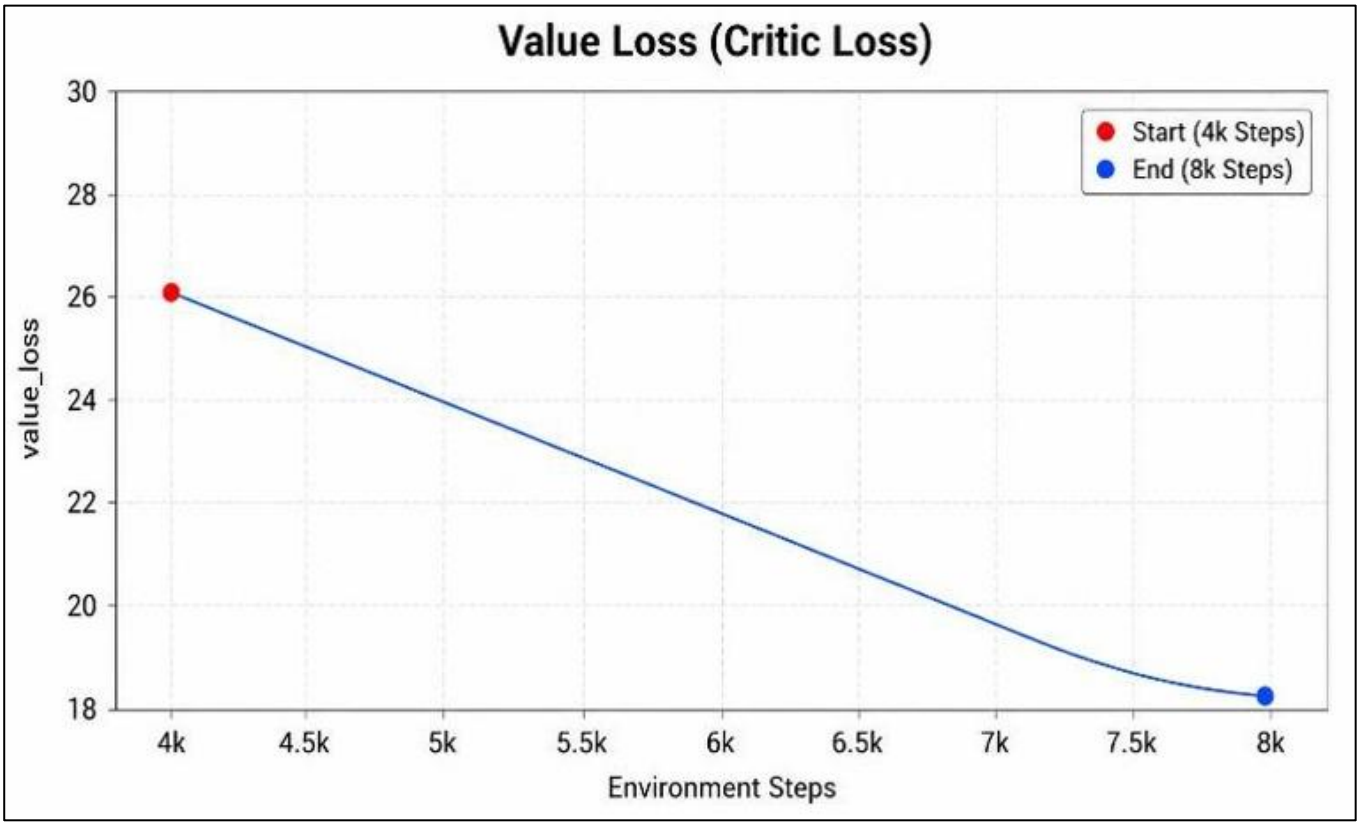


Fig 5 PPO Value Function Loss Convergence during Training

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may1200

- *Simulation Output*

The simulation demonstrates adaptive signal switching based on detected traffic density and learned PPO decision making.

```
Loading done.
Simulation started with time: 0.00
```

Fig 6 SUMO Simulation Output with Adaptive Traffic Signal Control

- *Quantitative Results*

The results clearly show that PPO outperforms other methods in minimizing waiting time and queue length while maximizing throughput.

Table 2 Quantitative Results

Method	**Waiting Time (s)**	**Queue Length**	**Throughput**
Fixed Time	78.5	24.3	Low
DQN	52.1	16.8	Medium
A2C	48.6	15.2	Medium
PPO	35.4	10.6	High

VII. IMPLEMENTATION AND REPRODUCIBILITY

The complete implementation, source code, and experiment setup are publicly available for reproducibility:

https://github.com/rupalikale15/Traffic-Signal-Optimization-Using-Reinforcement-Learning Researchers can reproduce the framework by installing SUMO, Stable-Baselines3, YOLOv8 dependencies, and executing the training pipeline.

VIII. CONCLUSION

This research proposed an intelligent adaptive traffic signal optimization framework combining YOLOv8-based traffic perception with PPO reinforcement learning.

The proposed approach demonstrated significant improvements over conventional fixed-time control and alternative reinforcement learning methods. Major contributions include:

- Vision-based traffic state perception
- PPO-driven adaptive signal optimization
- Comparative RL performance analysis
- Real-time simulation integration with SUMO Future work includes:
- Multi-intersection traffic optimization
- Real-world traffic deployment
- Multi-agent reinforcement learning
- Integration with IoT smart city infrastructure

REFERENCES

[1]. L. Li, Y. Lv, and F.-Y. Wang, “Traffic signal timing via deep reinforcement learning,” *IEEE/CAA Journal of Automatica Sinica*, vol. 3, no. 3, pp. 247–254, 2016.

[2]. H. Wei et al., “IntelliLight: A reinforcement learning approach for intelligent traffic signal control,” in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2018, pp. 2496–2505.

[3]. I. Arel, C. Liu, T. Urbanik, and A. G. Kohls, “Reinforcement learning-based multi-agent system for network traffic signal control,” *IET Intelligent Transport Systems*, vol. 4, no. 2, pp. 128–135, 2010.

[4]. E. Van der Pol and F. A. Oliehoek, “Coordinated deep reinforcement learners for traffic light control,” in *Proc. 15th Int. Conf. Autonomous Agents and Multiagent Systems (AAMAS)*, 2016.

[5]. J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” *arXiv preprint arXiv: 1707.06347*, 2017.

[6]. G. Zheng, Y. Wang, and X. Wang, “Learning phase competition for traffic signal control,” in *Proc. 29th ACM Int. Conf. Information and Knowledge Management (CIKM)*, 2020, pp. 1965–1974.

[7]. T. Chen et al., “PPO-based multi-intersection traffic signal control with graph attention networks,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 12345–12356, 2022.

[8]. Q. Guo et al., “Vision-based reinforcement learning for adaptive traffic signal control using YOLO and deep RL,” *Expert Systems with Applications*, vol. 215, 2023.

[9]. X. Liang et al., “Deep reinforcement learning for traffic signal control with computer vision,” *Transportation Research Part C: Emerging Technologies*, vol. 128, 2021.

[10]. W. Genders and S. Razavi, “Using a deep reinforcement learning agent for traffic signal control,” *arXiv preprint arXiv: 1611.01142*, 2016.

[11]. V. Mnih et al., “Asynchronous methods for deep reinforcement learning,” in *Proc. 33rd Int. Conf. Machine Learning (ICML)*, 2016, pp. 1928–1937.

[12]. Ultralytics, “YOLOv8 Documentation,” 2023. [Online]. Available: https://docs.ultralytics.com.

[13]. P. A. Lopez et al., “Microscopic traffic simulation using SUMO,” in *Proc. 21st IEEE Int. Conf. Intelligent Transportation Systems (ITSC)*, 2018, pp. 2575–2582.

[14]. Z. Wu et al., “YOLOv8-based real-time vehicle detection for intelligent traffic management,” *Sensors*, vol. 24, no. 5, 2024.