भारतीय सूचना प्रौद्योगिकी संस्थान, पुणे

Indian Institute of Information Technology, Pune

(An Institute of National Importance by an Act of Parliament)

Survey No. 9/1/3, Ambegaon Budruk, Sinhagad Institute Road, Pune - 411041

Enhancing the performance of DRL algorithm using channel attention mechanism.

M SAI DINAKAR REDDY	112215117
P. SHREYAS REDDY	112215136
SAMHITH KALARI	112215160
ANISH KUMAR S	112215022
B. HIMAVATH SAI	112215039

Supervised by :

Mrs. Anu Priya

Adjunct Assistant Professor

TABLE OF CONTENTS

INTRODUCTION

Our project works with various reinforcement learning models, such as SARSA, DQN, DDQN, and D3QN, alongside specialized attention mechanisms like channel attention, to enhance the agent’s learning and decision-making capabilities.

AIM

- To enhance path planning in reinforcement learning models by integrating attention mechanisms
- Optimizing navigation strategies to identify more efficient routes in complex environments.
- To improve the models' ability to prioritize relevant features, resulting in better decision-making, reduced exploration, and maximized cumulative rewards.

LITERATURE REVIEW

Sno	Journal & Publishers	Title	Methodologies	Remarks
1.	Journal:Neural Processing Letters Publisher: Springer Year: 2024	NAVS: A Neural Attention-Based Visual SLAM for Autonomous Navigation in Unknown 3D Environments.	The NAVS system employs neural attention and reinforced learning for smart navigation and divides tasks between high-level planning and quick obstacle responses.	This paper presents NAVS, a system that uses neural attention and reinforcement learning to help robots navigate new spaces.
2.	Publisher: Springer Year: 2024	Deep Q Network (DQN), Double DQN, and Dueling DQN. In: Deep Reinforcement Learning.	Deep Reinforcement Learning uses Policy Gradient MethodsQ-learning, DQN, Double DQN, Dueling Double DQN.	This chapter from the book explains about deep Reinforcement Learning that uses DQN, DDQN and D3QN.

Sno	Journal & Publishers	Title	Methodologies	Remarks
3.	Journal : Applied Soft Computing Publisher : Elsevier Year : 2024	Dynamic path planning via Dueling Double Deep Q-Network (D3QN).	Q-Learning, Deep Q-Learning, Temporal Difference Learning, DDQN, D3QN, PER(Prioritized Experience Replay), ROS(Robot Operating System)	This paper highlights D3QN-PER as an effective approach for dynamic path planning.
4.	Journal : Computational Visual Media. Publisher : Springer Year : 2022	Attention mechanisms in computer vision: A survey.	The methodologies used are channel attention, spatial attention,temporal attention,branch attention	This paper provides insights into attention mechanisms, especially channel attention, which can help RL path-planning model focus on critical features, enhancing efficiency.

Sno	Journal & Publishers	Title	Methodologies	Remarks
5.	Journal : Engineering Applications of Artificial Intelligence. Publisher : Elsevier Year : 2013	The combination of Sarsa algorithm and Q-learning, Journal of Engineering applications of artificial intelligence.	The methodologies used are sarsa algorithm and Q-learning.	This paper highlights how blending on-policy (SARSA) and off-policy (Q-learning) strategies can balance exploration and exploitation. This approach can enhance stability.
6	Journal : Autonomous Agents and Multi-Agent Systems. Publisher : Springer Year : 2022	Learning Multi-Agent Communication with Double Attentional Deep Reinforcement Learning, Journal of Autonomous Agents and Multi-Agent Systems.	The methodlogies used is double attentional mechanism.	This paper offers strategies to integrate attention layers for more efficient decision-making and collaboration, enhancing RL model performance in complex tasks.

Sno	Journal & Publishers	Title	Methodologies	Remarks
7	Journal:Neural Networks Publisher: Elsevier Year:2022	Generalised attention-weighted reinforcement learning.	This paper provides insights into applying attention mechanisms in reinforcement learning by dynamically weighting relevant features to improve decision-making.	This paper provides a robust framework by showcasing how attention-weighted mechanisms can dynamically emphasize critical features.
8	Publisher : Cornell university Year : 2023	Hierarchical Reinforcement Learning in Complex 3D Environments.	This paper uses a hierarchical reinforcement learning approach, breaking tasks into subtasks managed by high-level goal-setting policies and low-level action policies, enabling effective navigation and problem-solving in complex 3D environments.	Combining attention mechanisms with HRL's task decomposition might further improve decision-making by allowing the model to focus on critical spatial features at different levels of abstraction.

Sno	Journal & Publishers	Title	Methodologies	Remarks
9.	Journal:Energy Publisher: Elsevier Year: 2024	A comparative study of DQN and D3QN for HVAC system optimisation control, Journal of Energy.	It evaluates different Q-network architectures to determine the best configuration for each algorithm.	Performance metrics and algorithm insights from DQN vs. D3QN provides a strong foundation for selecting the optimal approach for efficient path planning.
10.	Publisher: arXiv Year : 2017	Towards Monocular Vision based Obstacle Avoidance through Deep Reinforcement Learning,	The methodologies used are deep reinforcement learning techniques, particularly a model-free approach to enable robots to navigate and avoid obstacles using only monocular vision.	This paper explores the use of a DRL model, a dueling double deep Q-network (D3QN), to allow robots to navigate and avoid obstacles using monocular vision.

PROBLEM STATEMENT

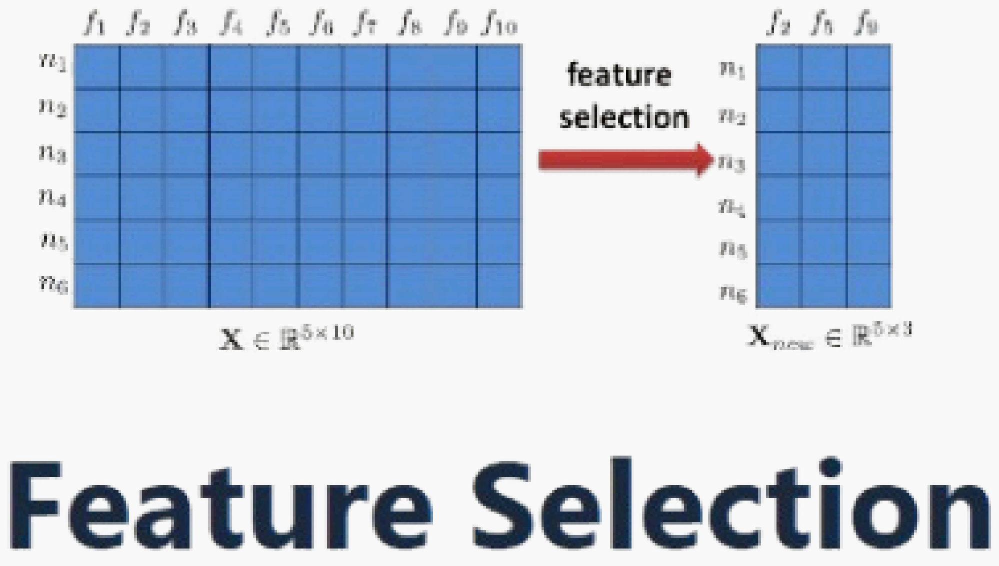


Fig 1: Feature selection

Src : aibook.in

- Standard DQNs often struggle to effectively capture important features within multi-dimensional state representations. This limitation can lead to reduced performance.
- Our project explores the integration of channel attention mechanisms within DQN models to highlight critical features during the decision-making process.

MOTIVATION

- Growing demand for autonomous systems across various real-world applications, including self-driving cars, robotic delivery systems.
- Practical need to evaluate the strengths and weaknesses of different reinforcement learning (RL) models, given their increased use in realistic scenarios.
- Interest in developing RL models that are adaptable and effective in complex, dynamic environments.

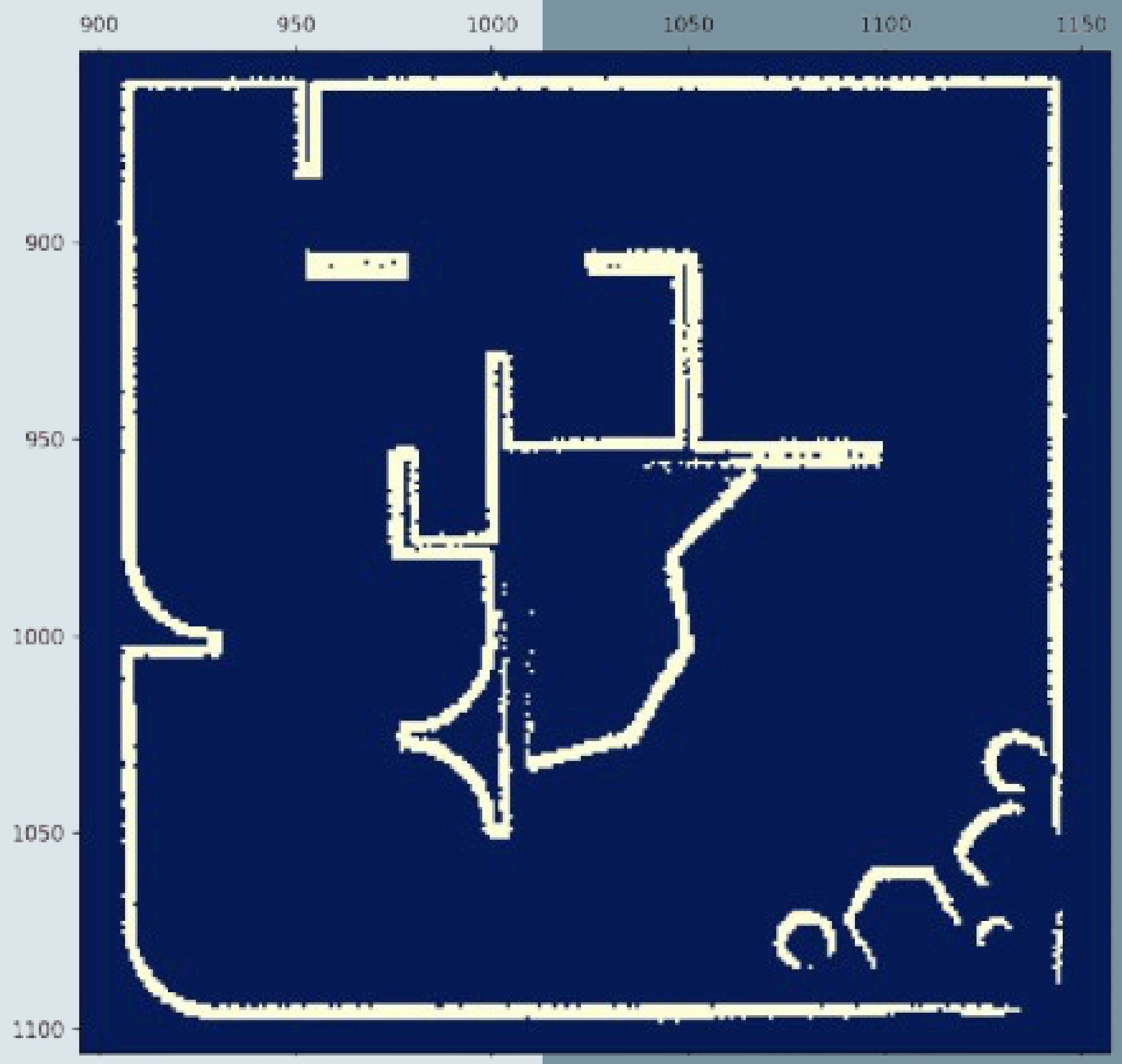


Fig 2 : A real time environment

Src : Stack Overflow

RESEARCH GAP

Efficiency and constraints:

- In real world applications where the input features are too many, the agent/robot takes too much time to make decisions and this is where feature selection comes in. Selecting important features and discarding others saves a lot of time and cost and this can be made possible using attention mechanisms.

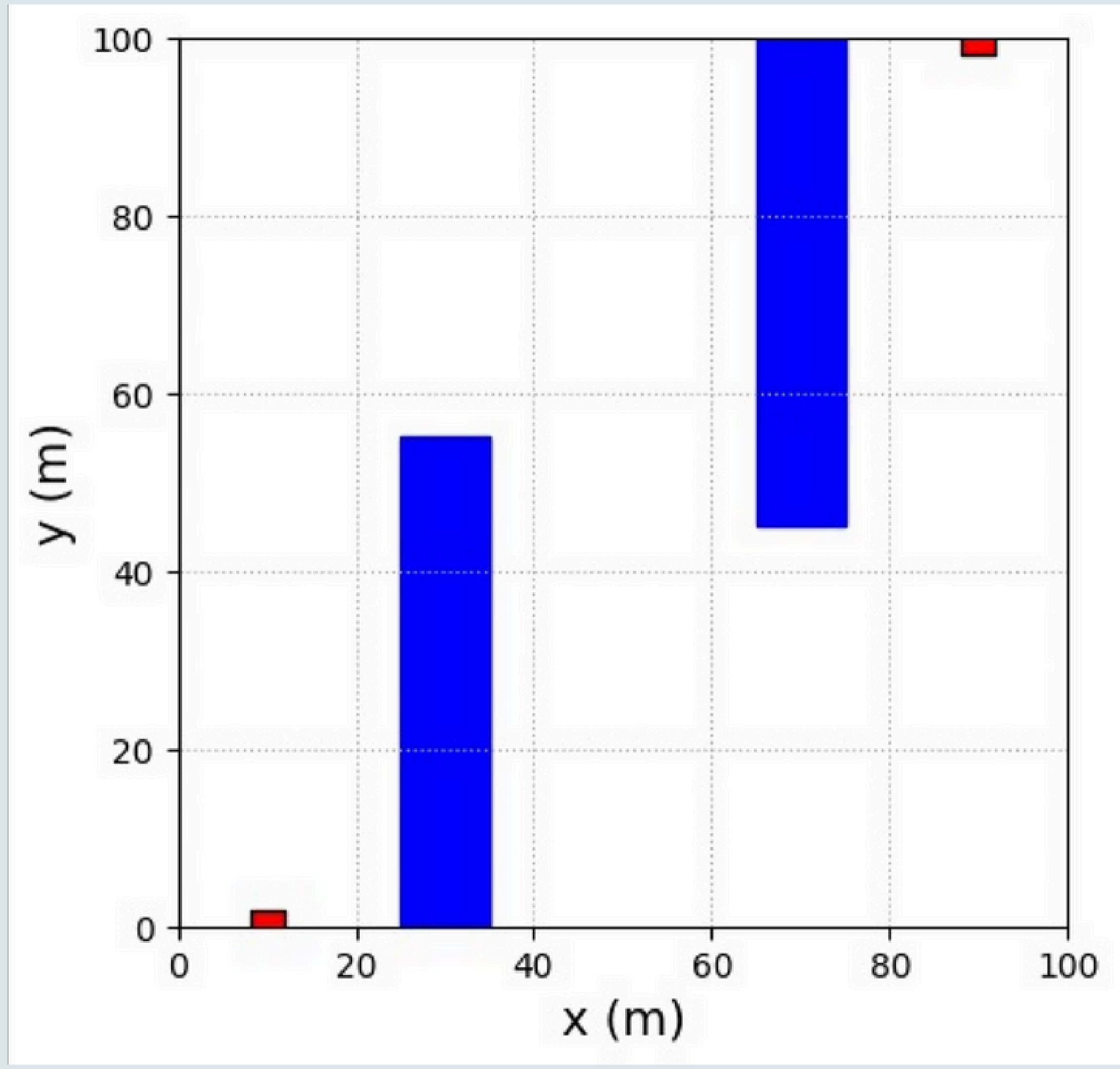
Under-explored Attention Mechanisms in RL :

- Although channel attention mechanisms are widely used in deep learning for their ability to focus on critical features, they have not been extensively applied in reinforcement learning.

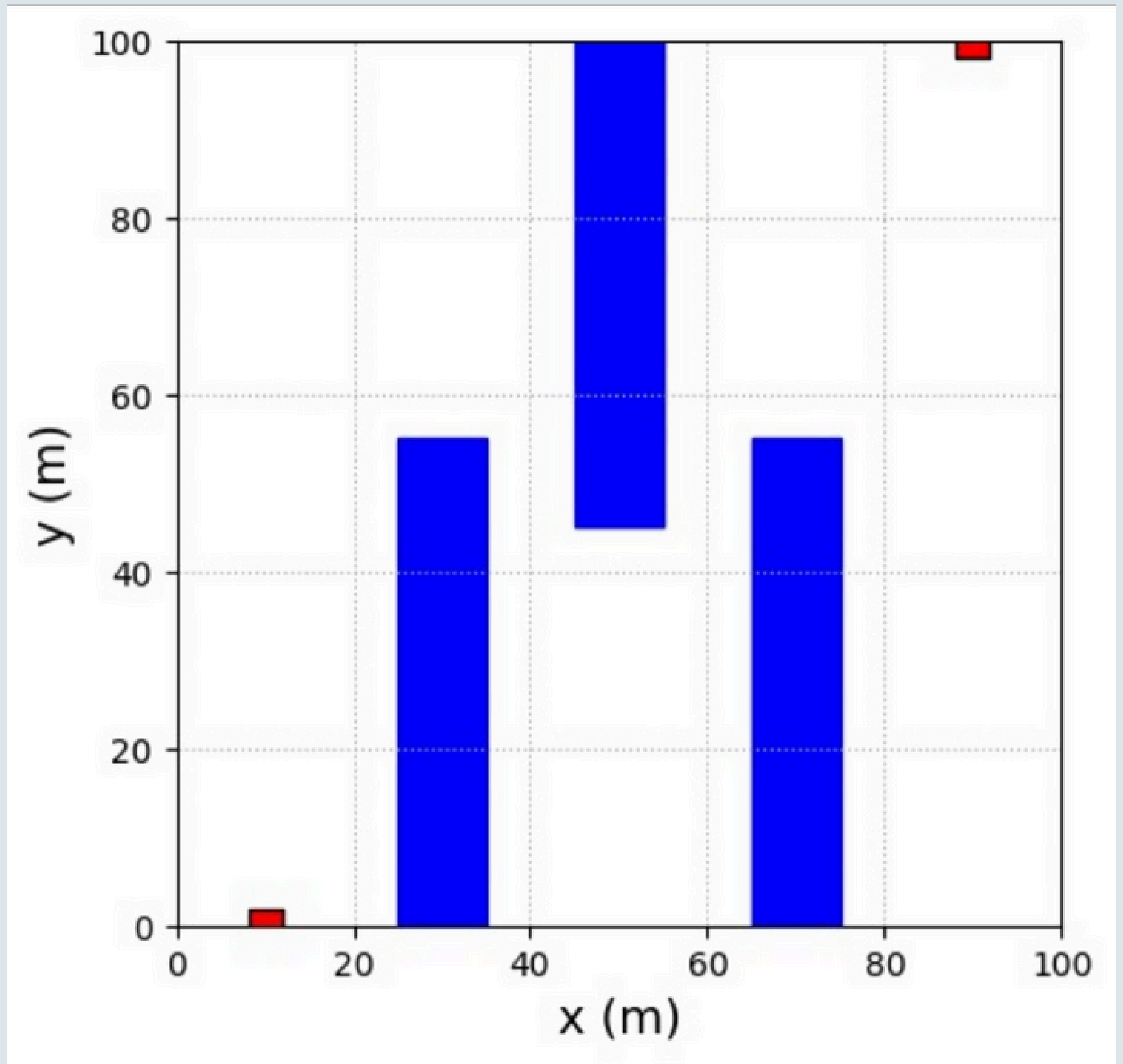
OBJECTIVES

- Evaluation of SARSA, DQN, DDQN and D3QN algorithms with each other on three different custom environments.
- Importance of feature selection and attention mechanisms in path planning.
- Implementing channel attention in the algorithms and checking their relevance on model performance.
- Comparing graphs of accumulative rewards, steps per episode and paths respectively to find variation in convergence of reward and learning rates

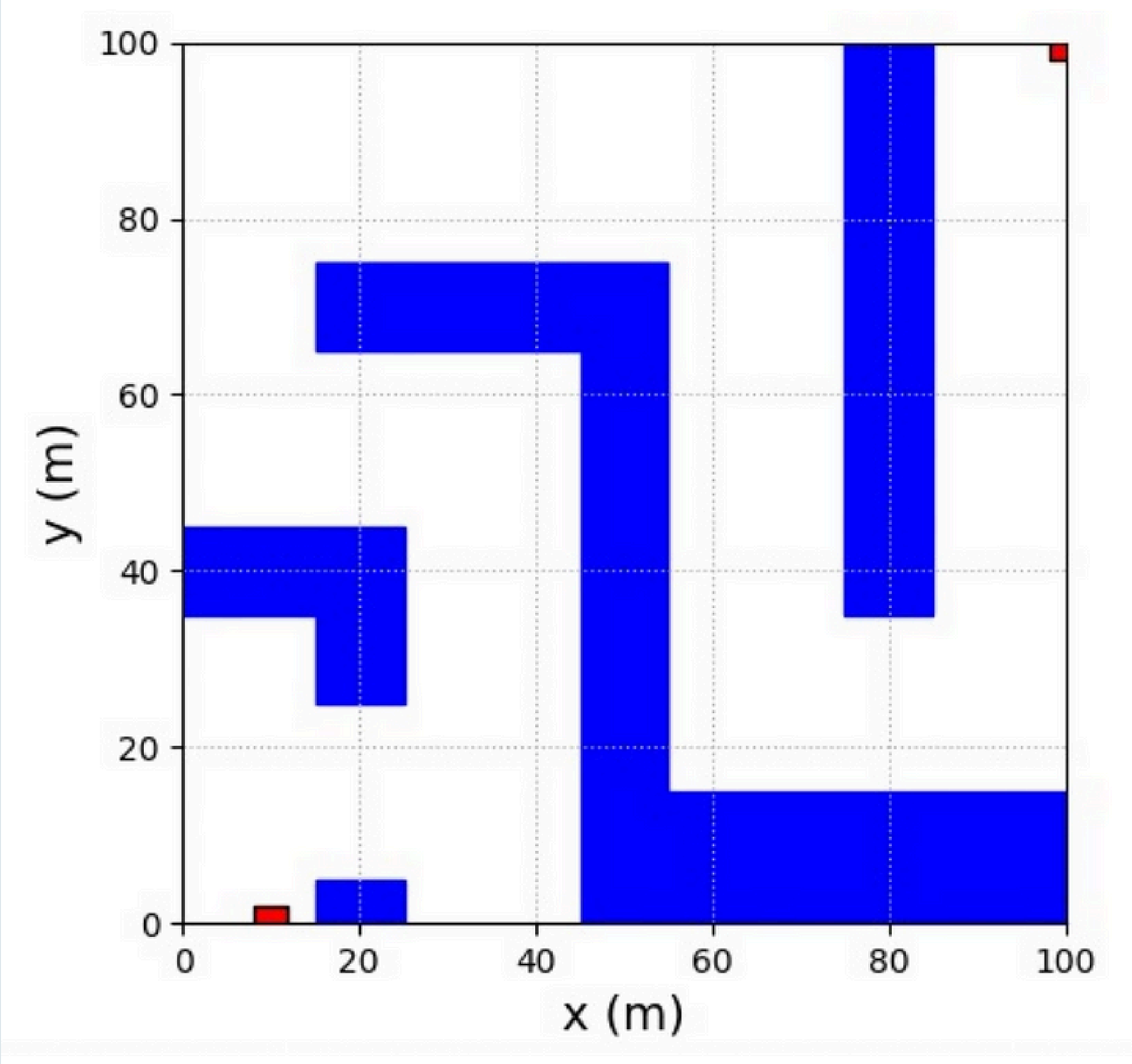
CUSTOM ENVIRONMENT DESCRIPTION



(a) 2 Obstacle Environment



(b) 3 Obstacle Environment



(c) Complex Environment

METHODOLOGY

Models used:

- SARSA

Three path planning models have been used to compare the results with and without channel attention.

- DQN
- DDQN
- D3QN

Epsilon Greedy Policy:

DQN, DDQN and D3QN use an epsilon greedy policy. The epsilon greedy policy balances out between exploring new actions and choosing the best known action. It chooses a random action with a probability ε and a known best action with a probability 1-ε where ε decreases over training.

METHODOLOGY

On policy approach:

- SARSA uses the on policy approach.
- It learns and updates its Q-values from the current action it is taking.
- It works well in environments which are dynamic.

Off policy approach:

- DQN, DDQN and D3QN use the off policy approach.
- It learns and updates its Q-values from the best actions it takes to find the most optimal path faster.
- It works well in environments which are predictable or static.

METHODOLOGY

Algorithms used:

SARSA:

State Action Reward State Action updates Q-values based on the action taken, following its current policy. This helps it to adapt well to environments with uncertainty, as it learns the values of actions that align with its exploration strategy.

DQN:

Deep Q Network uses deep neural networks to approximate Q-values helping an agent selecting a path that maximizes cumulative reward.

DDQN:

Double Deep Q Network improves upon DQN by using 2 networks(a primary and a target network) to reduce DQN's overestimation of Q-values. This dual network allows for more accurate values for Q-value calculation leading to more stable learning.

D3QN:

Dueling Double Deep Q Network combines the advantages of DDQN with a dueling architecture that separates the value of being in a state and the value of performing an action. This helps the model evaluate a state better even if the action does not impact the outcome significantly enhancing its performance in complex environments.

METHODOLOGY

Algorithms used:

Channel Attention:

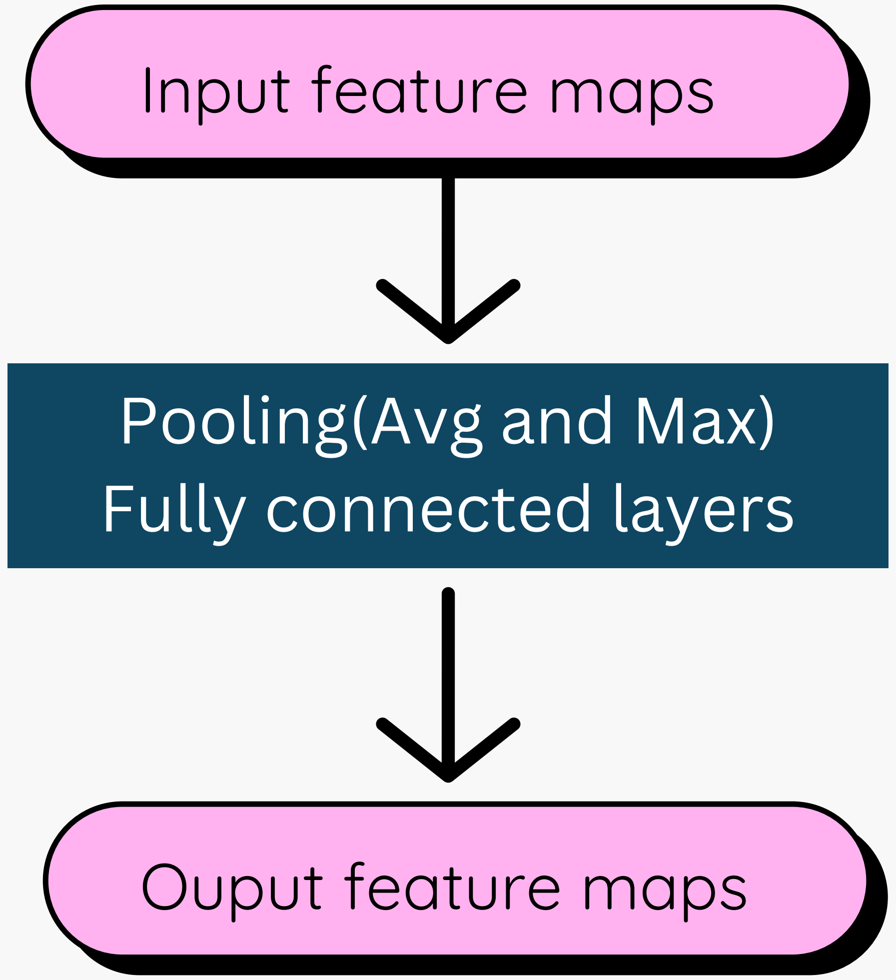


Figure 3 : Channel Attention

Objective: It prioritizes the relevant features for pathfinding.

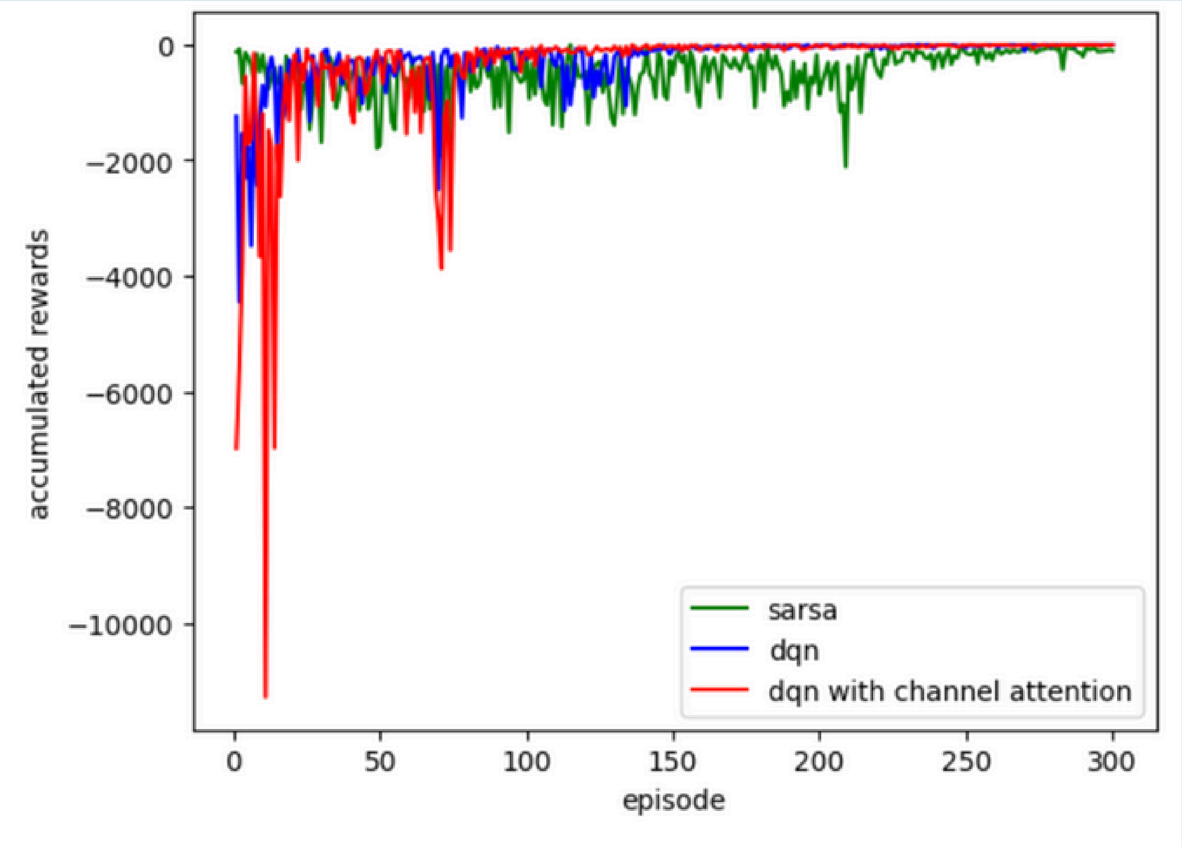
Input and Processing: Pools input data for each input channel into two average pool and max pool. The avg pool has the overall information from the particular channel while the max pool contains the most prominent features from the channel.

Advantages: It focuses on the critical information such as obstacles, optimal pathways, target, etc.

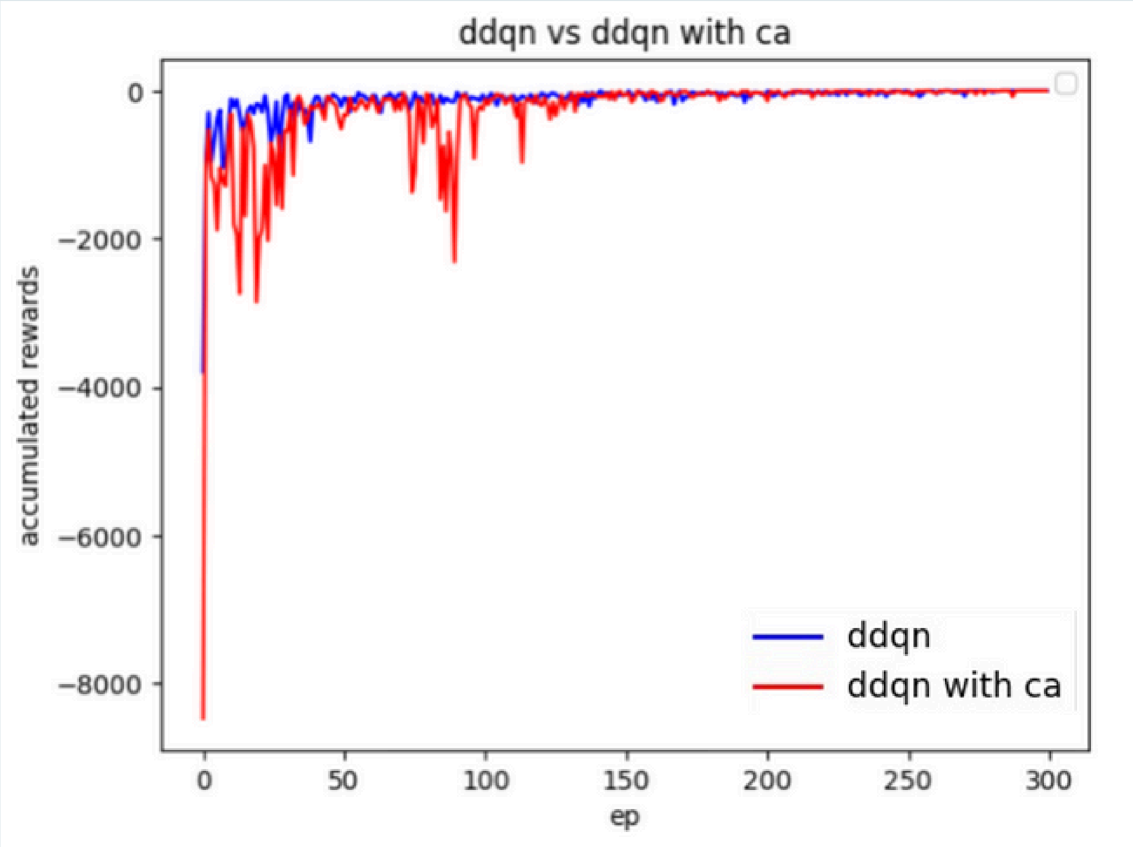
Channel attention boosts accuracy and efficiency in path planning by prioritizing the important features.

RESULTS

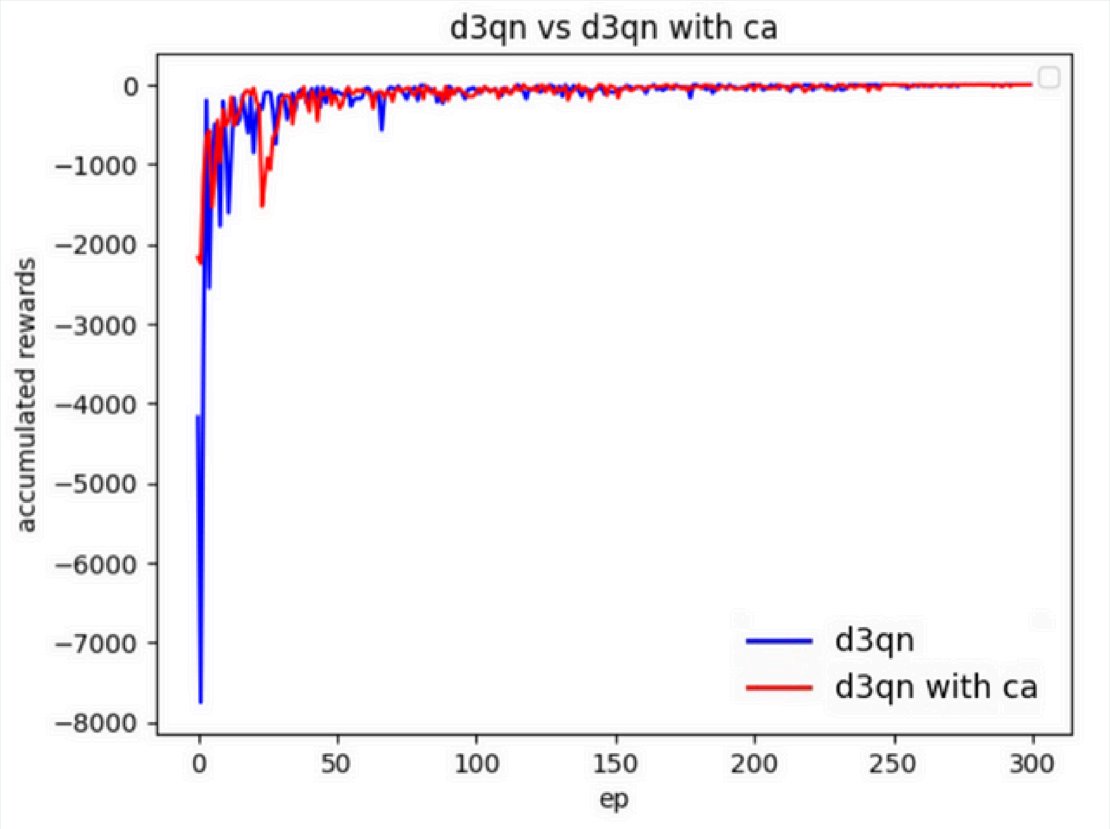
Comparison graphs : Accumulated rewards vs Episodes for 2 obstacle environment



(a) Sarsa vs Dqn vs Dqn-CA



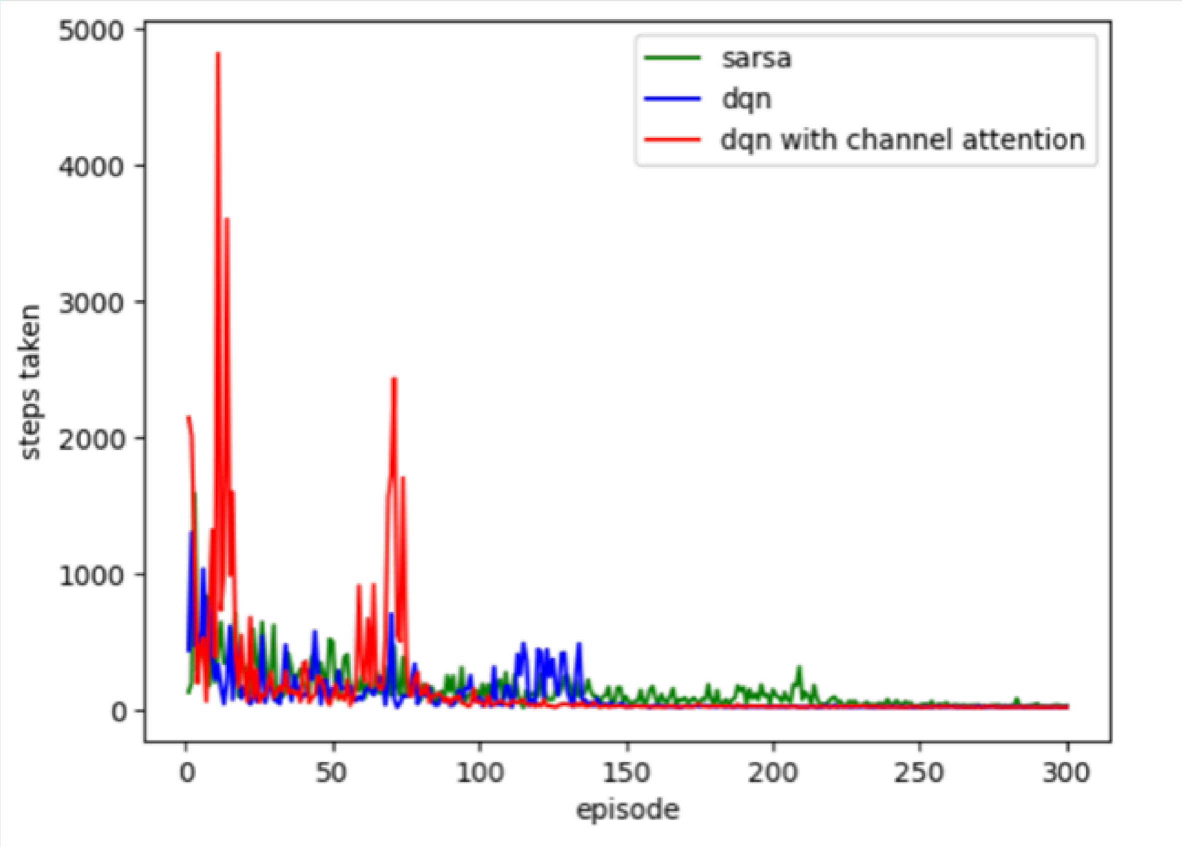
(b) Ddqn vs Ddqn-CA



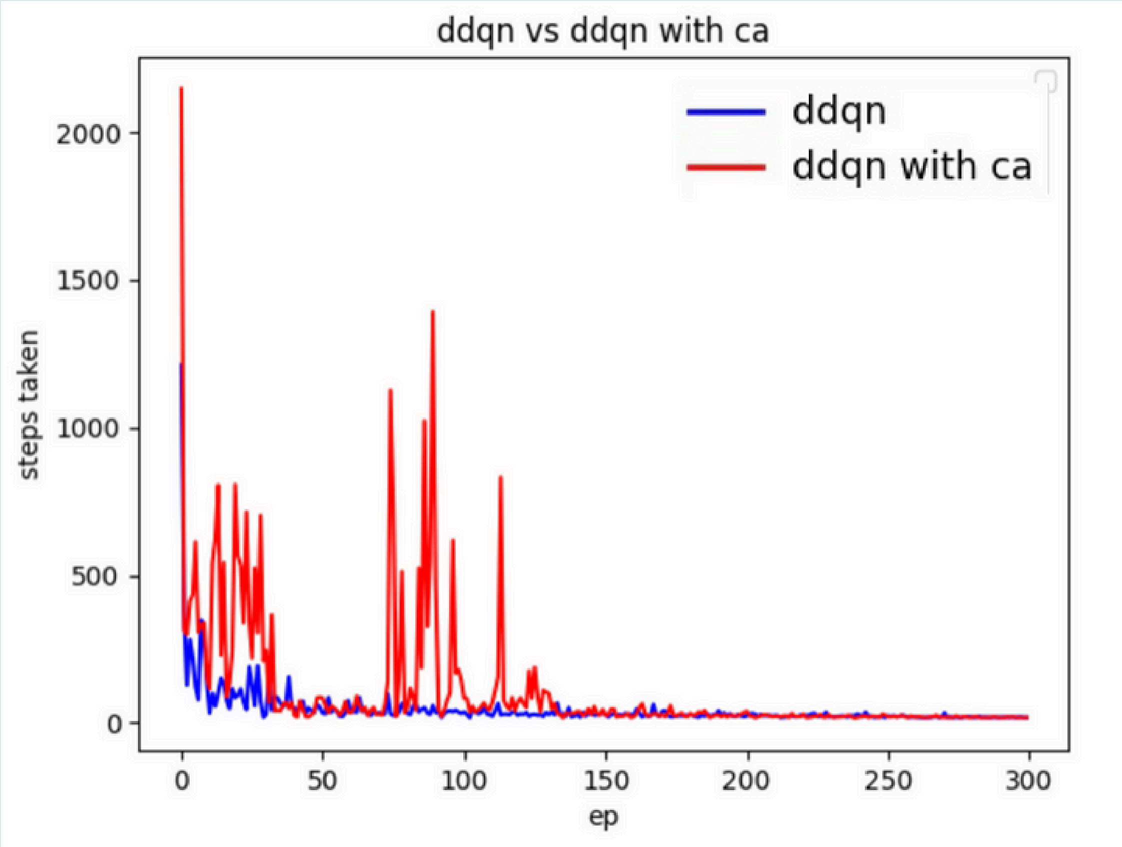
(c) D3qn vs D3qn-CA

RESULTS

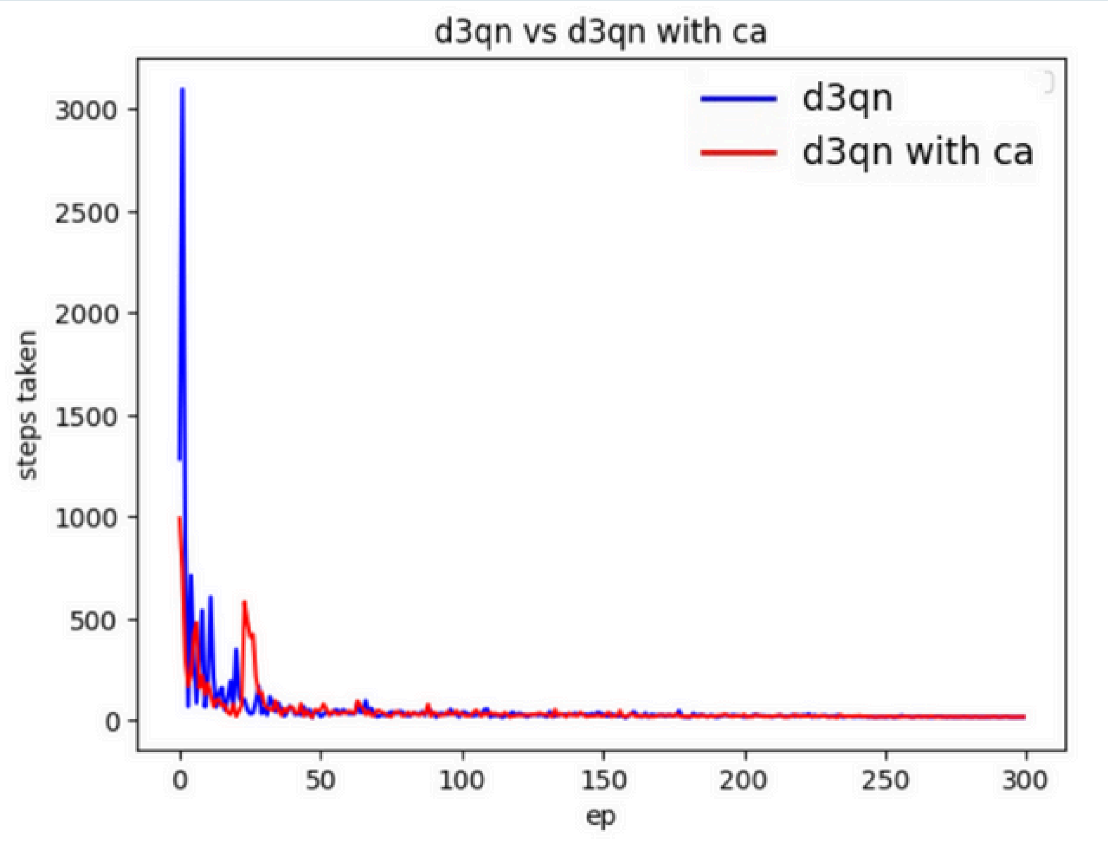
Comparison graphs : No of steps vs Episodes for 2 obstacle environment



(a) Sarsa vs Dqn vs Dqn-CA



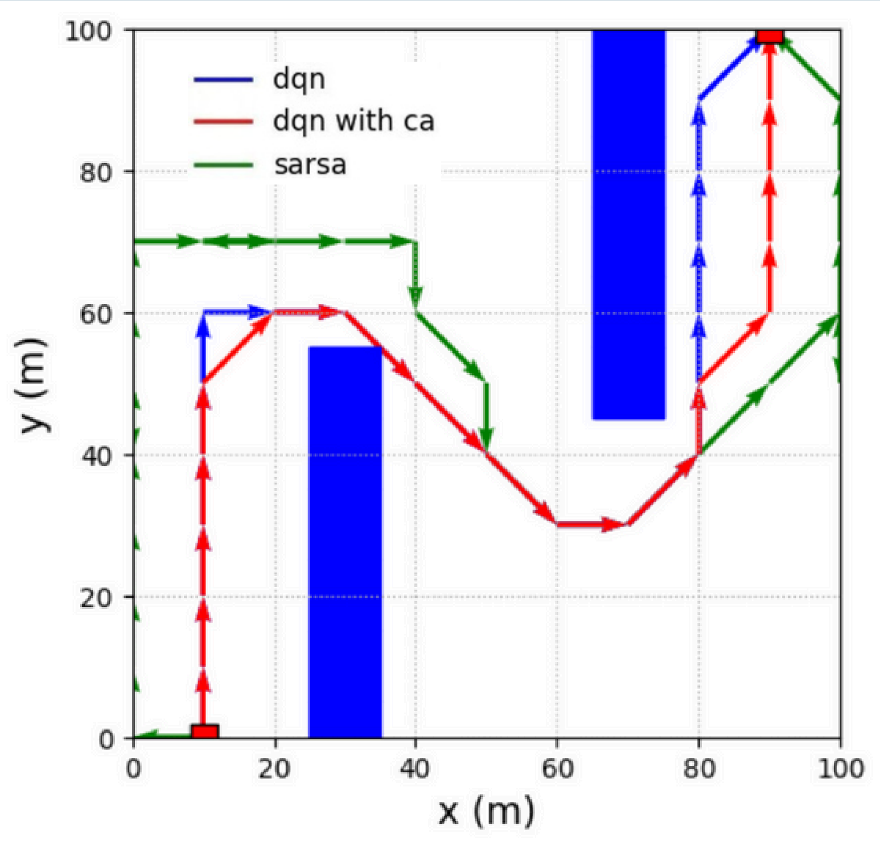
(b) Ddqn vs Ddqn-CA



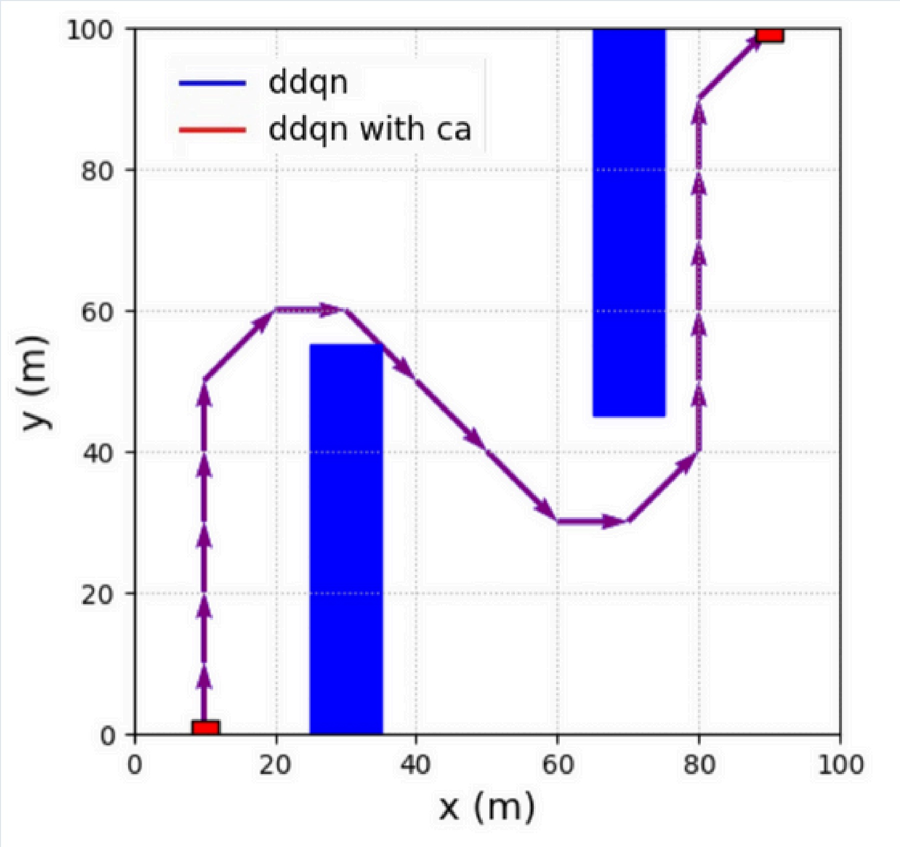
(c) D3qn vs D3qn-CA

RESULTS

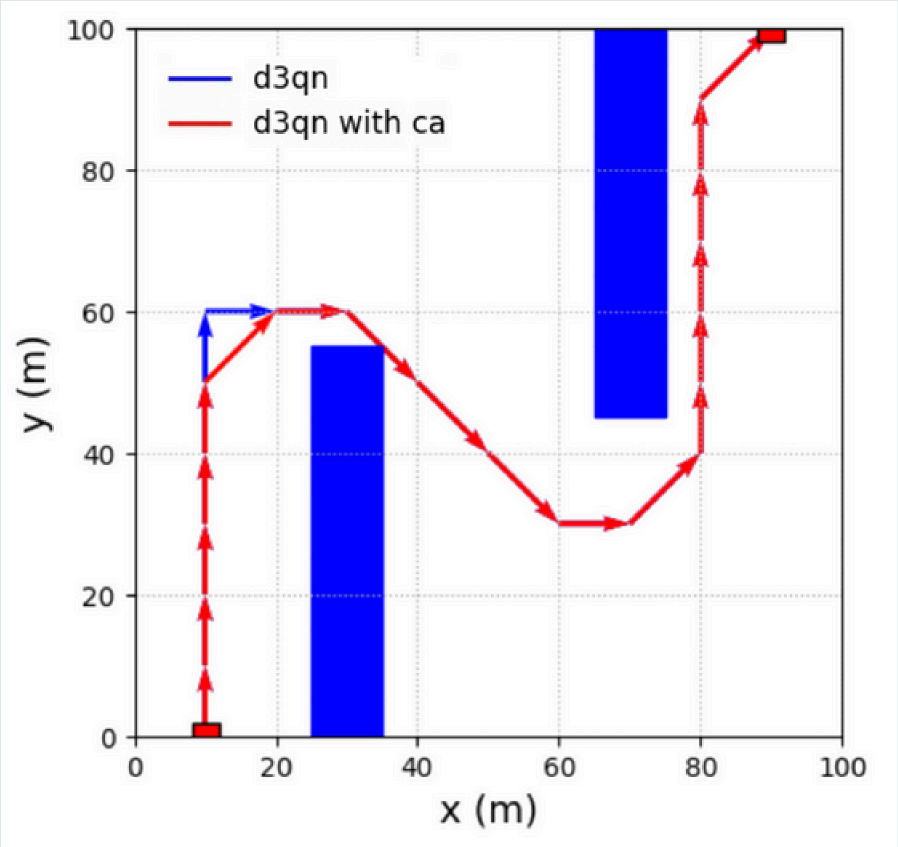
Comparison graphs : Paths taken for 2 obstacle environment



(a) Sarsa vs Dqn vs Dqn-CA



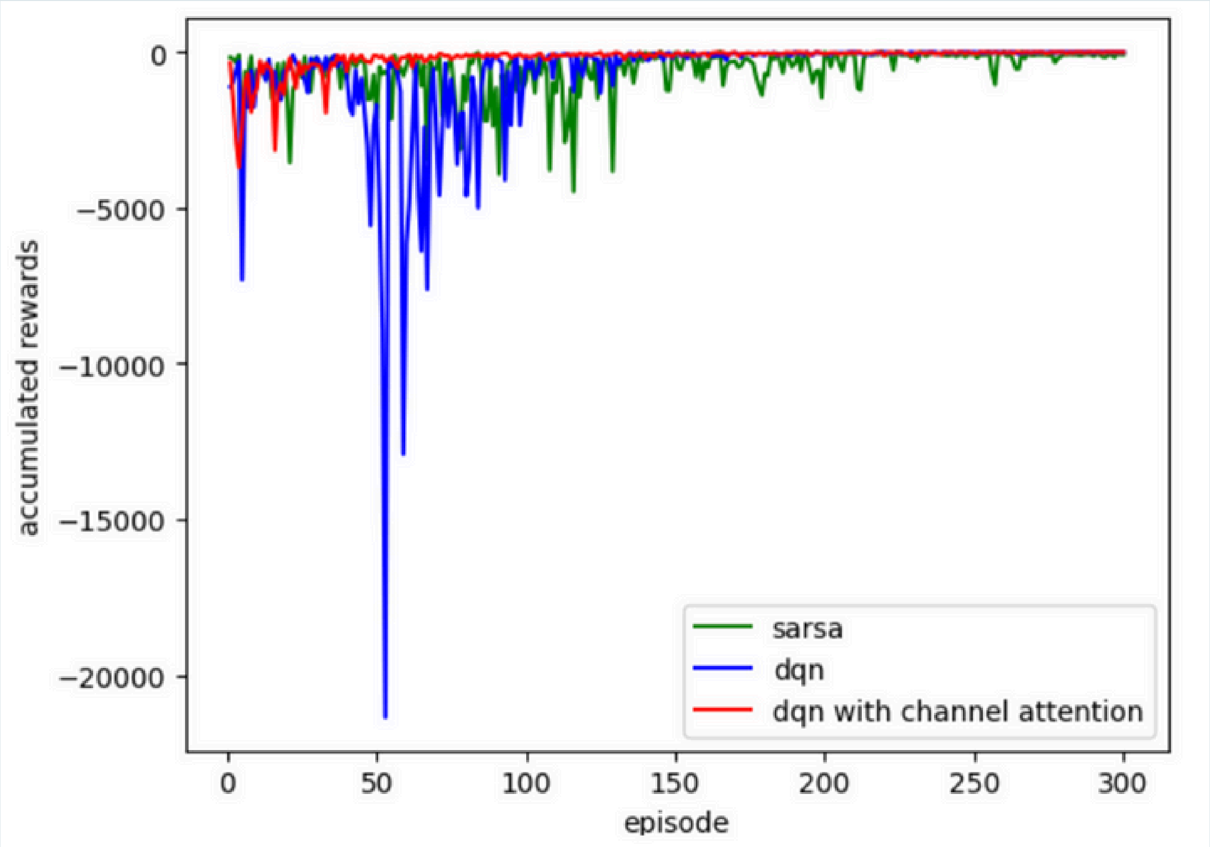
(b) Ddqn vs Ddqn-CA



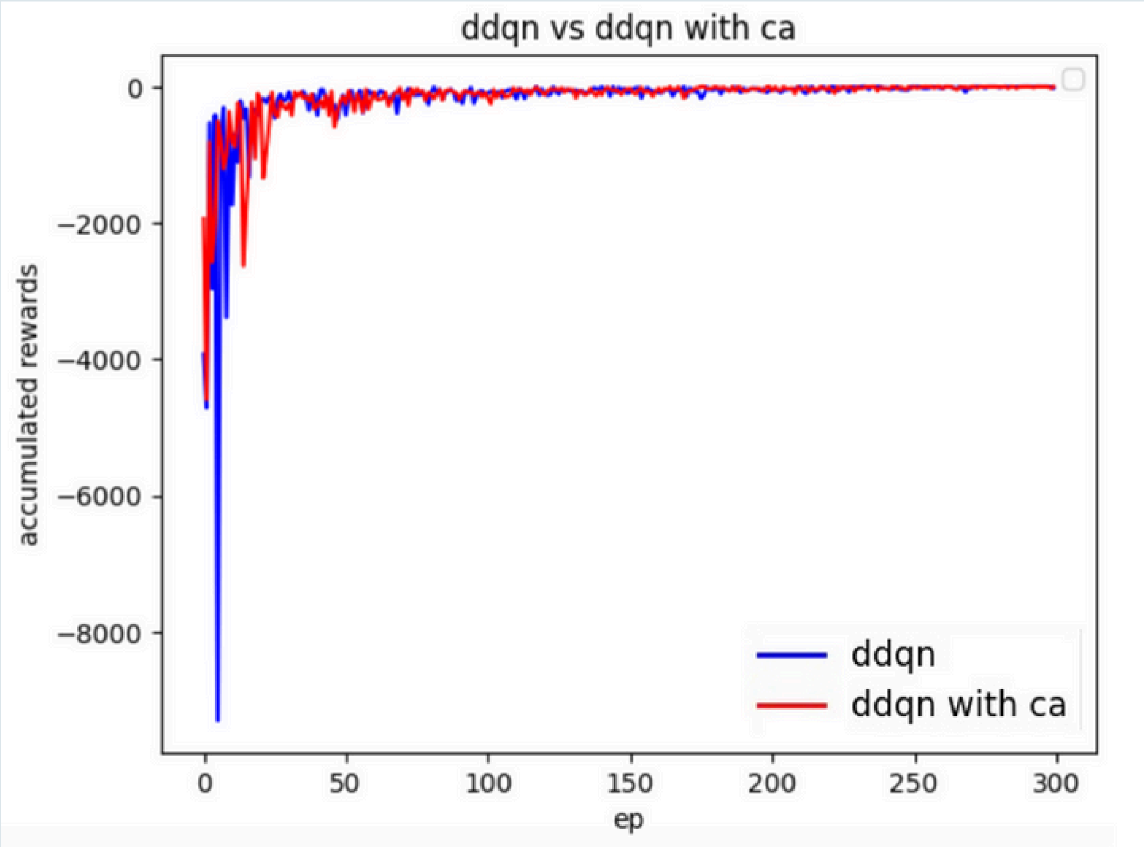
(c) D3qn vs D3qn-CA

RESULTS

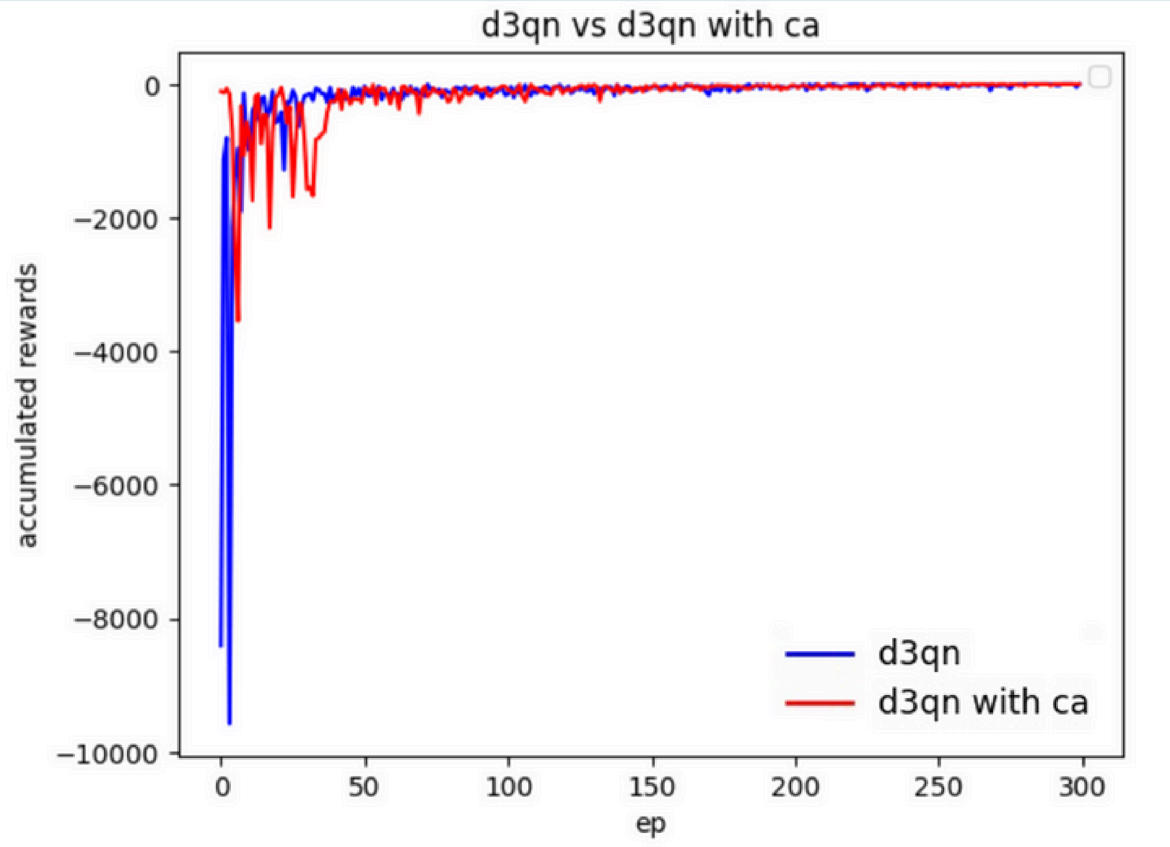
Comparison graphs : Accumulated rewards vs Episodes for 3 obstacle environment



(a) Sarsa vs Dqn vs Dqn-CA



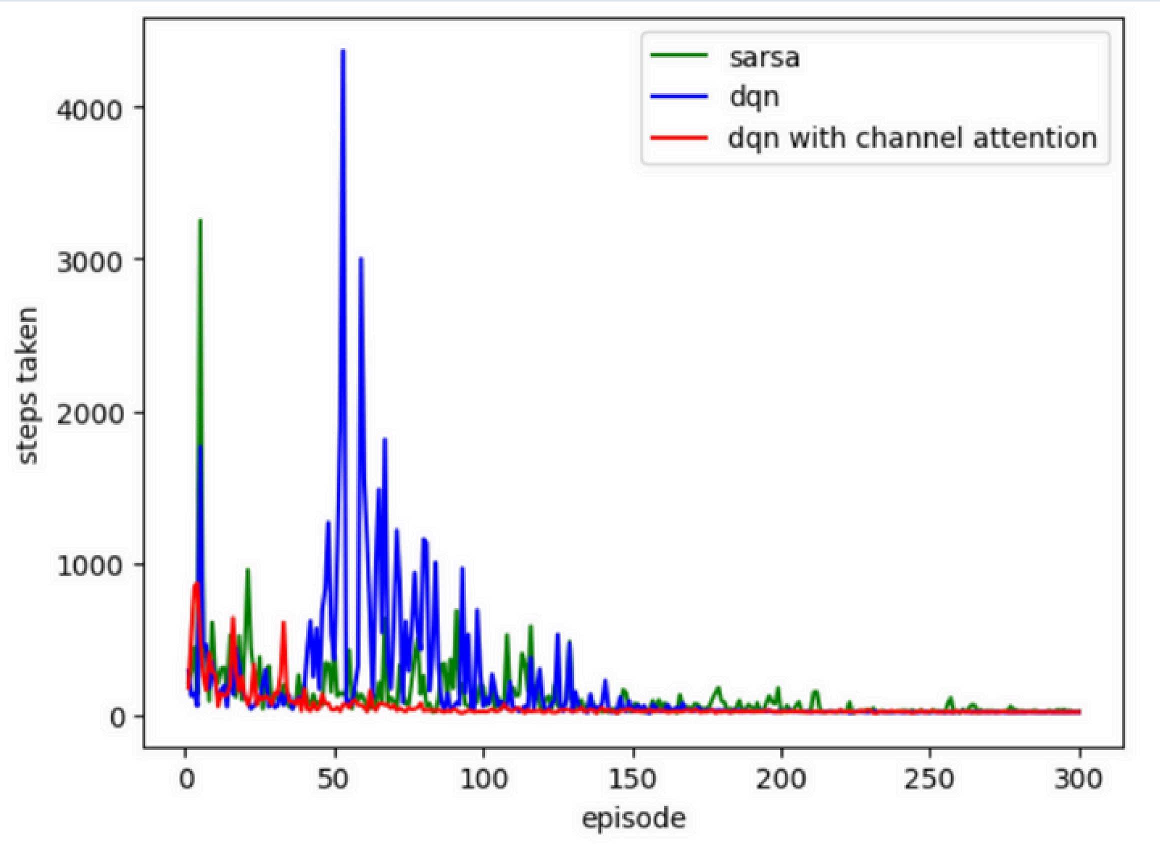
(b) Ddqn vs Ddqn-CA



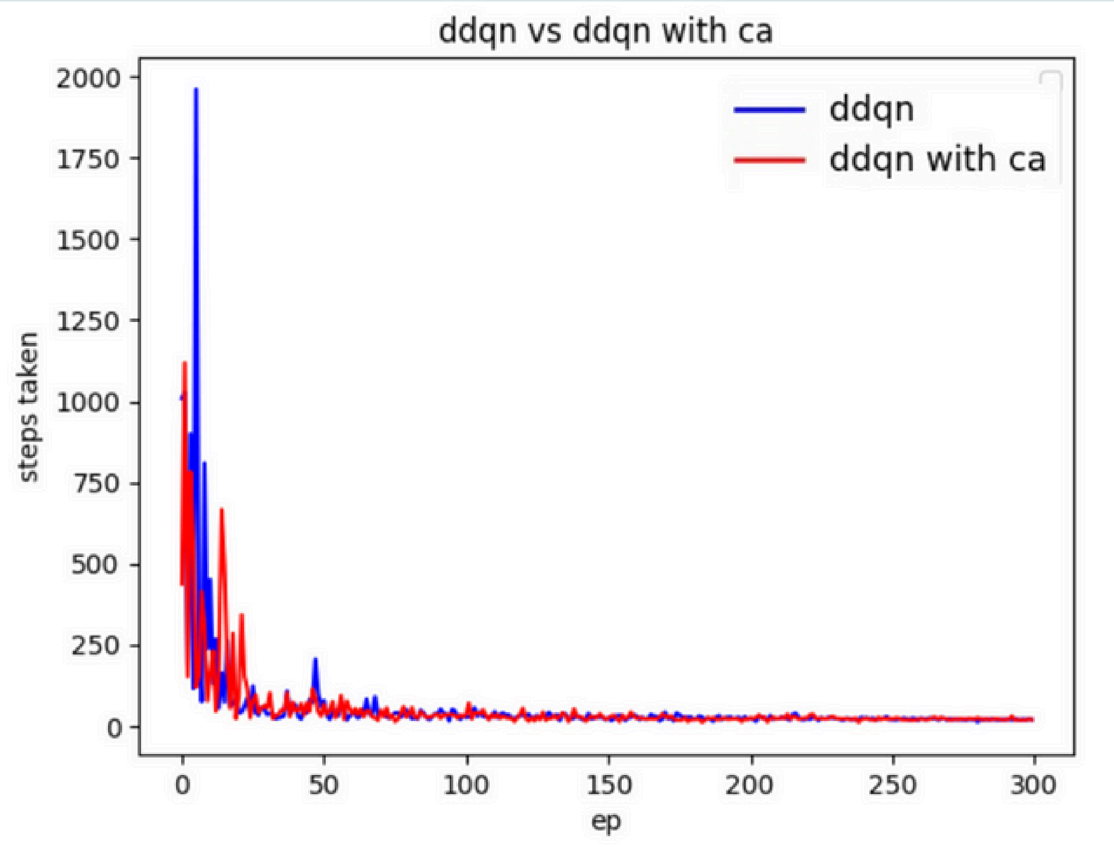
(c) D3qn vs D3qn-CA

RESULTS

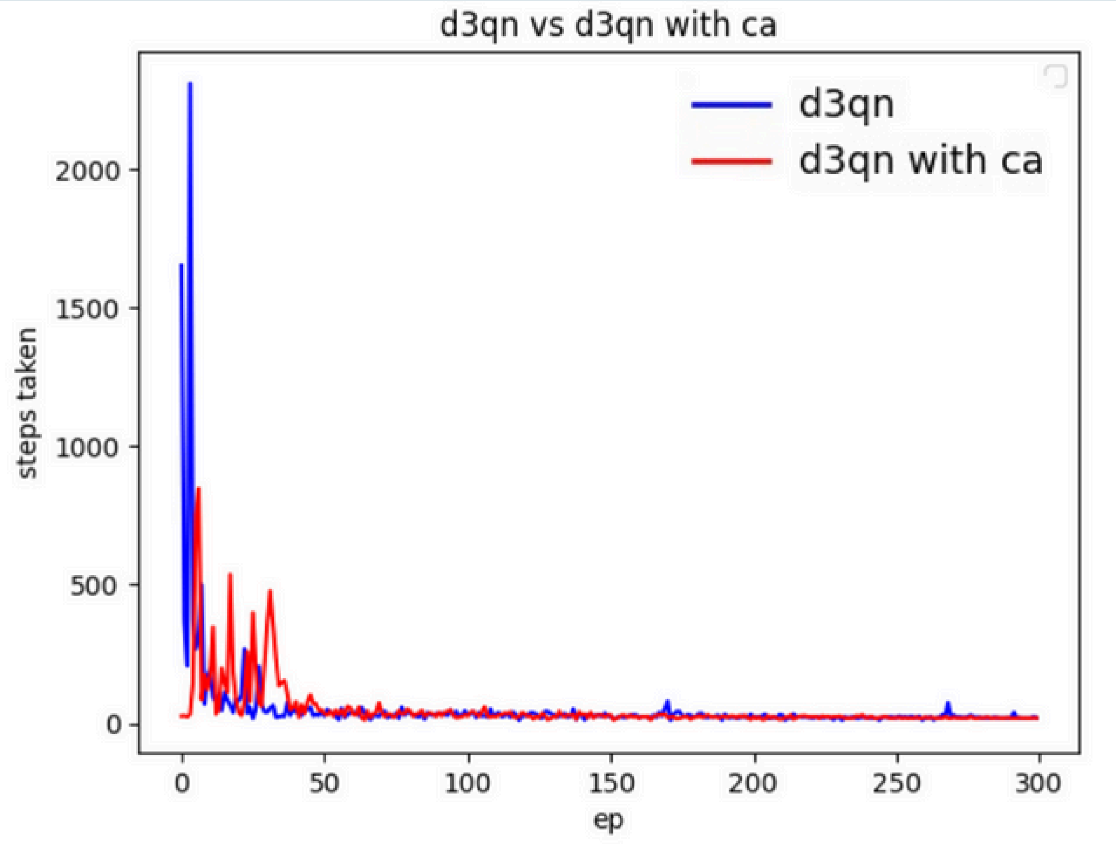
Comparison graphs : No of steps vs Episodes for 3 obstacle environment



(a) Sarsa vs Dqn vs Dqn-CA



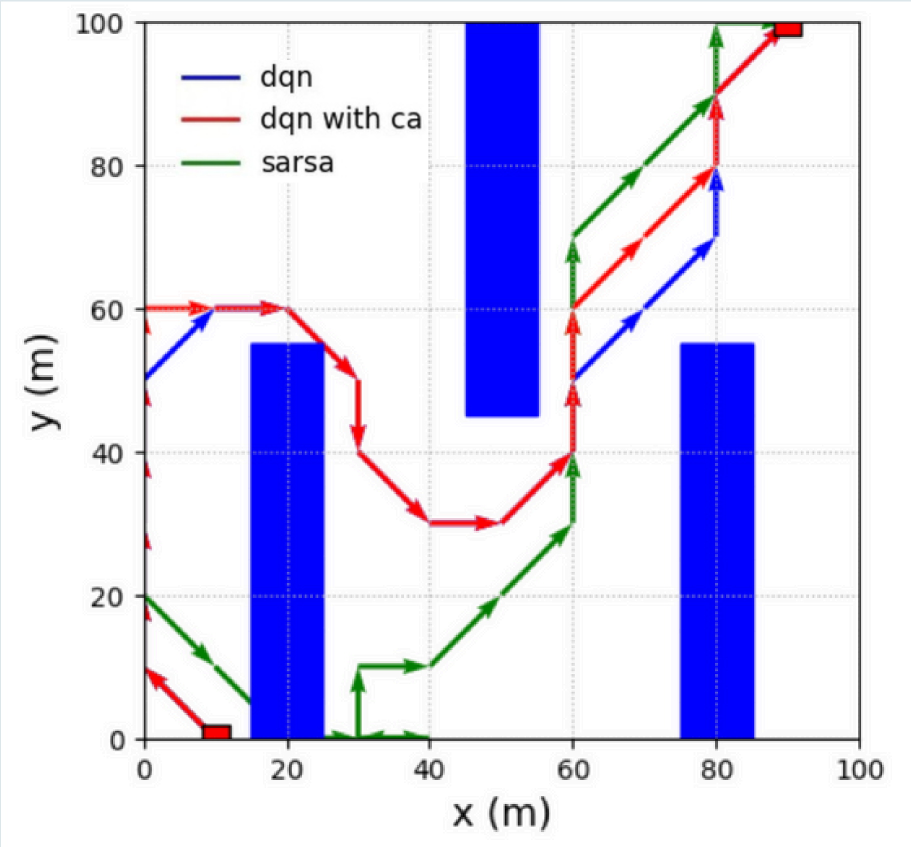
(b) Ddqn vs Ddqn-CA



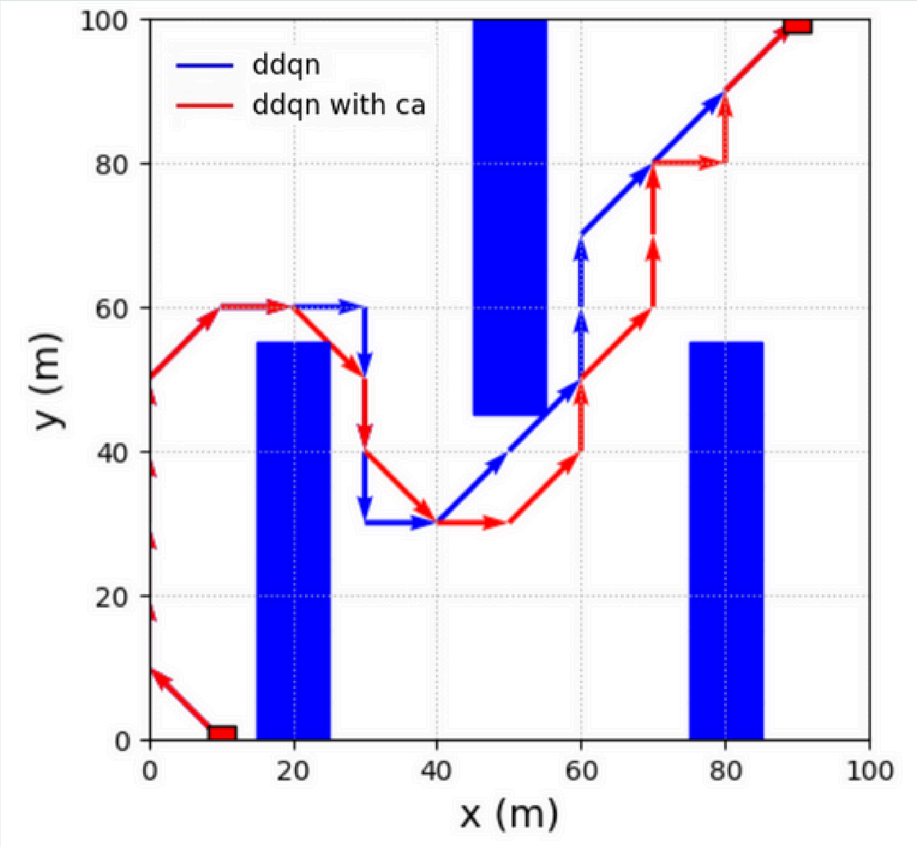
(c) D3qn vs D3qn-CA

RESULTS

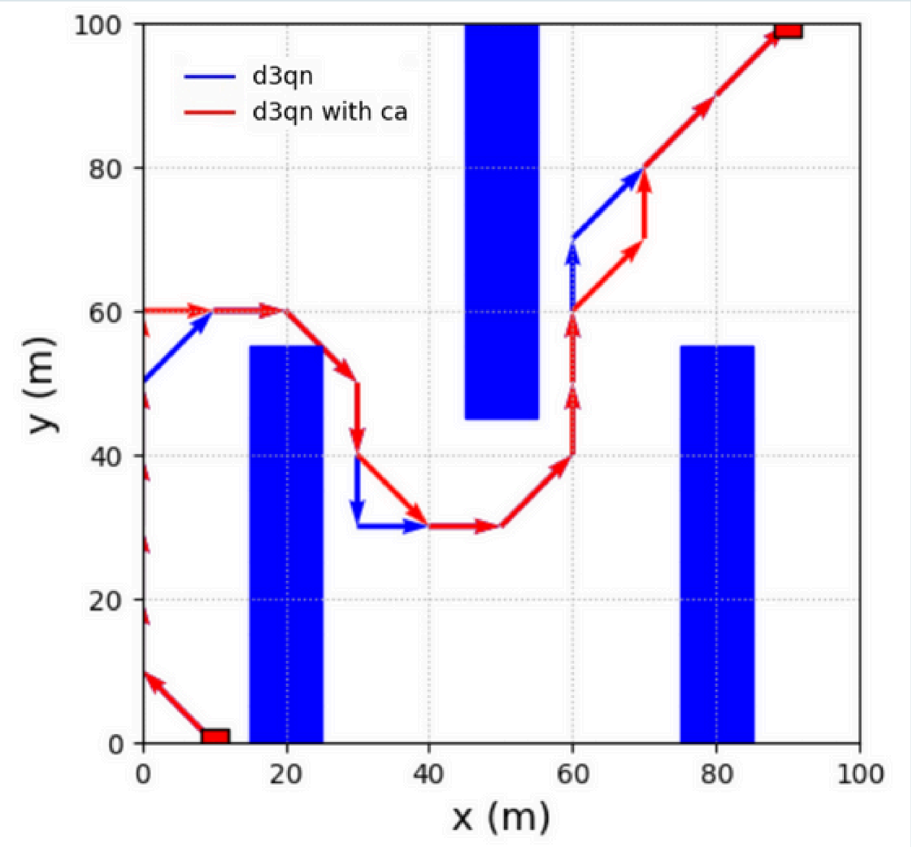
Comparison graphs : Paths taken for 3 obstacle environment



(a) Sarsa vs Dqn vs Dqn-CA



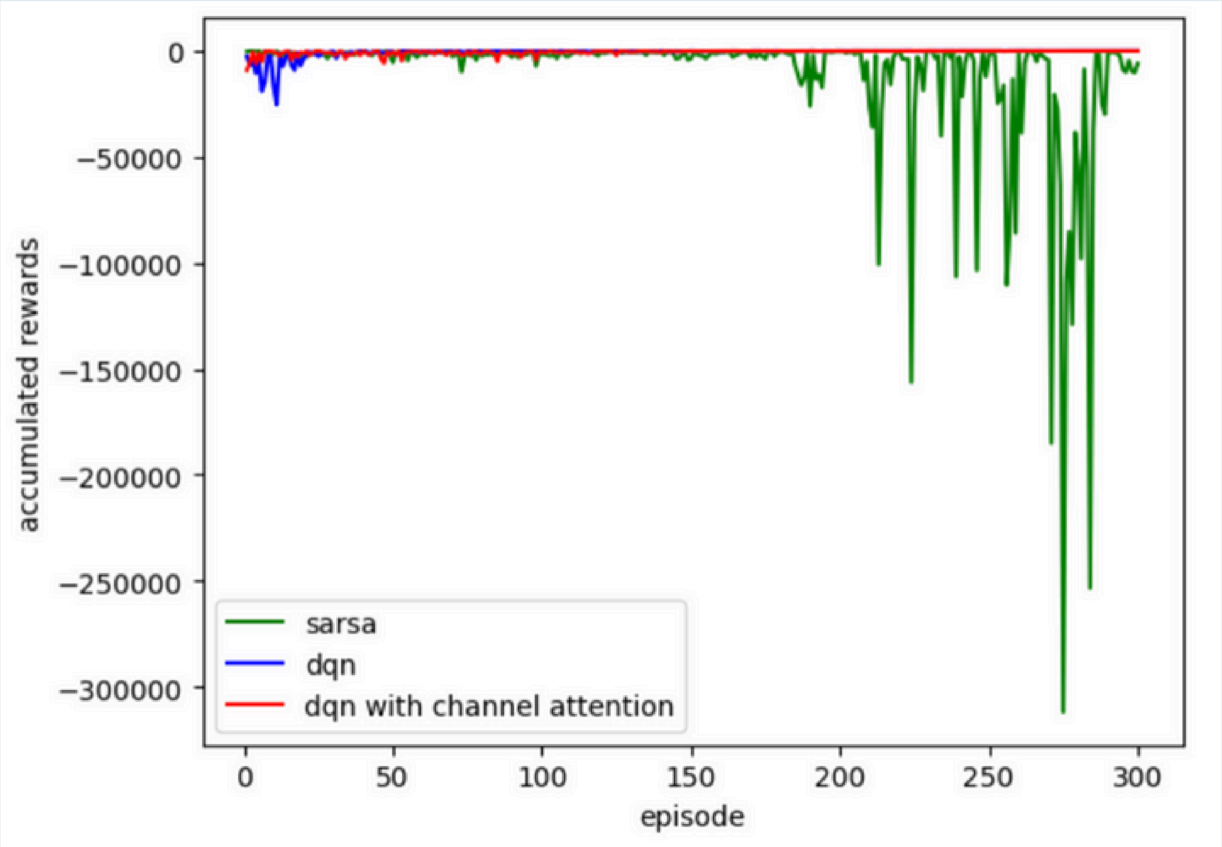
(b) Ddqn vs Ddqn-CA



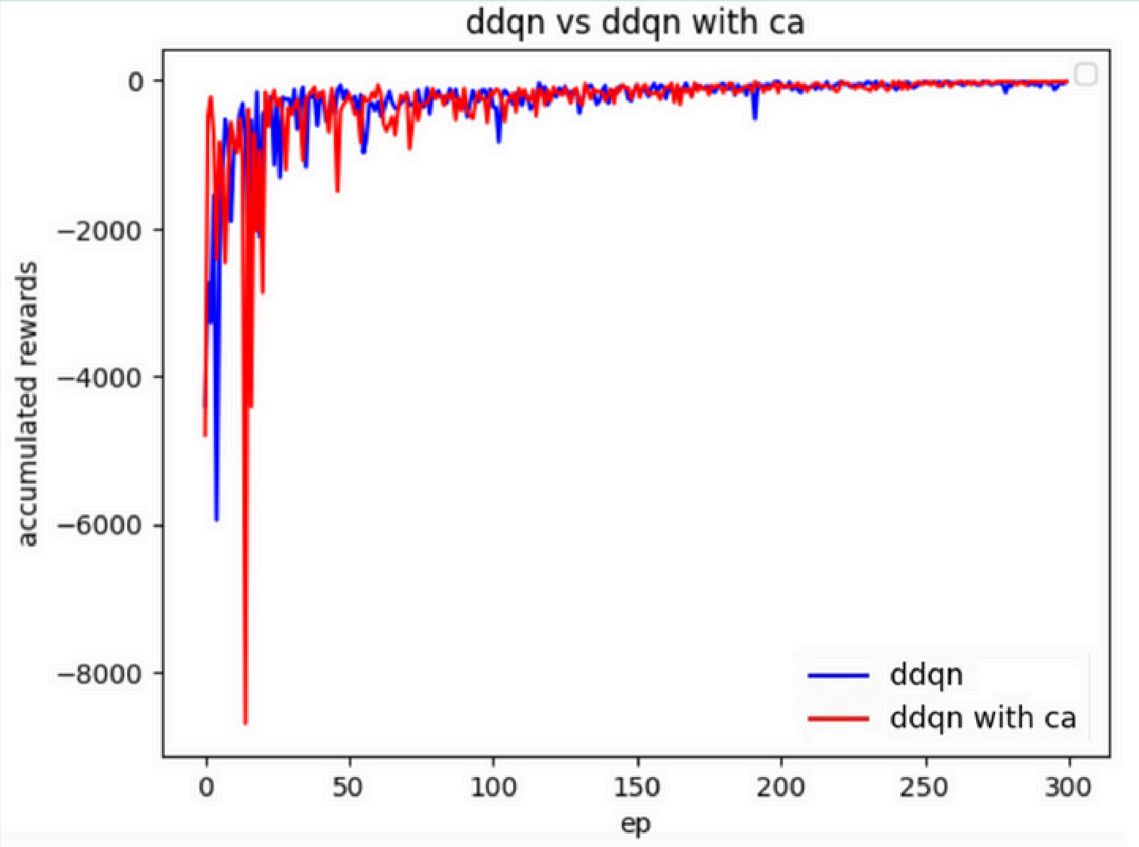
(c) D3qn vs D3qn-CA

RESULTS

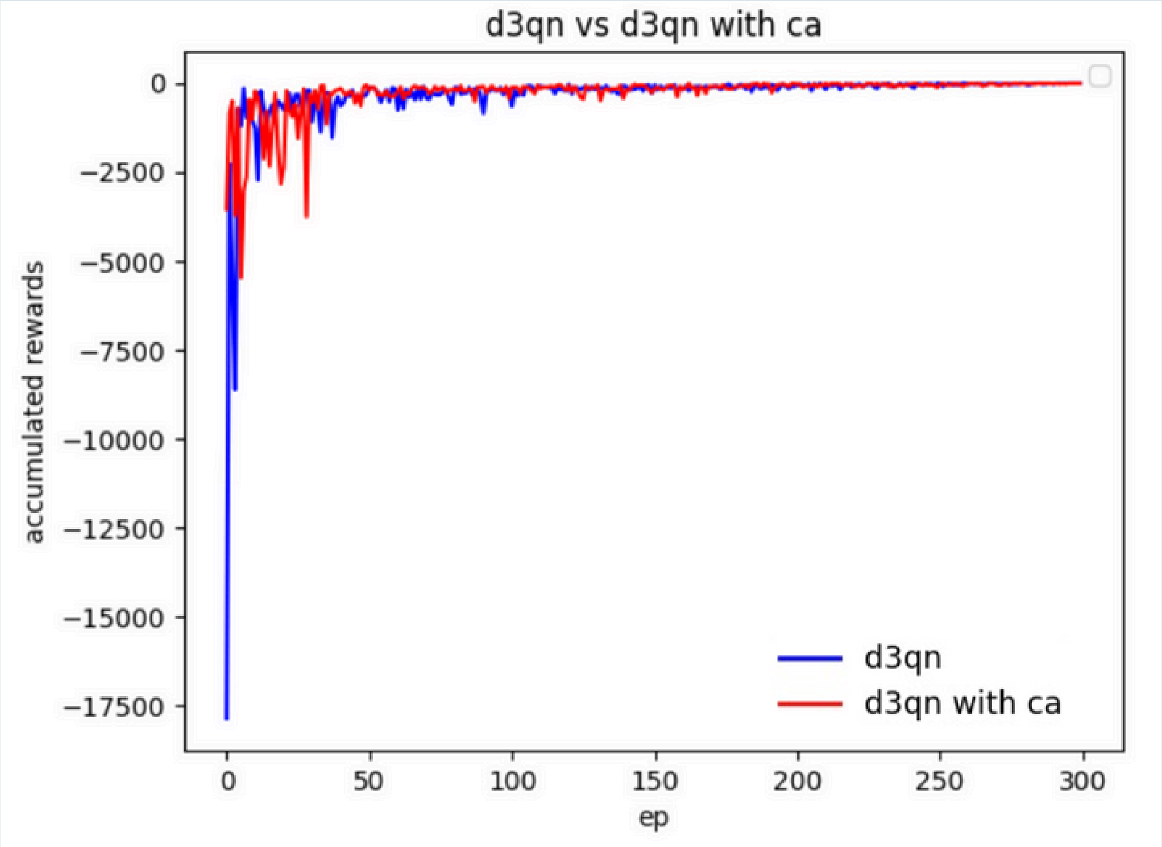
Comparison graphs : Accumulated rewards vs Episodes for Complex environment



(a) Sarsa vs Dqn vs Dqn-CA



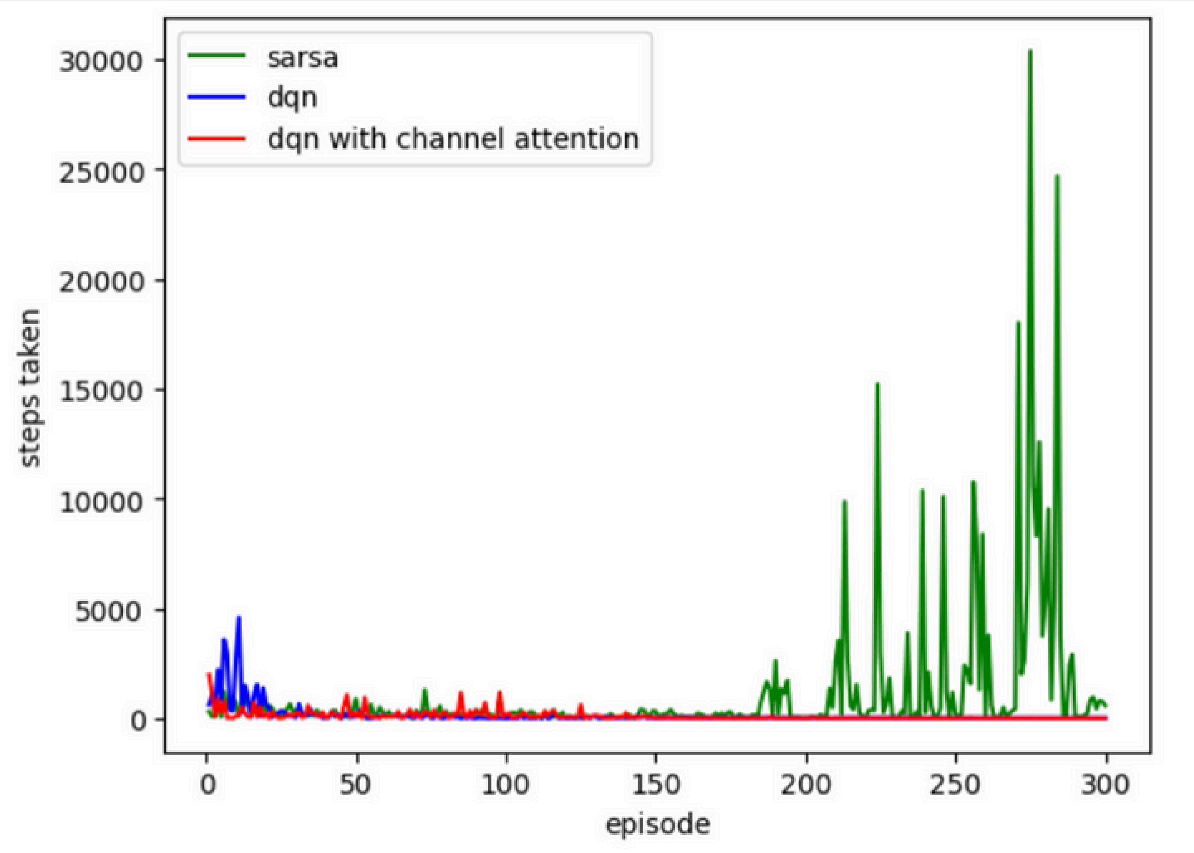
(b) Ddqn vs Ddqn-CA



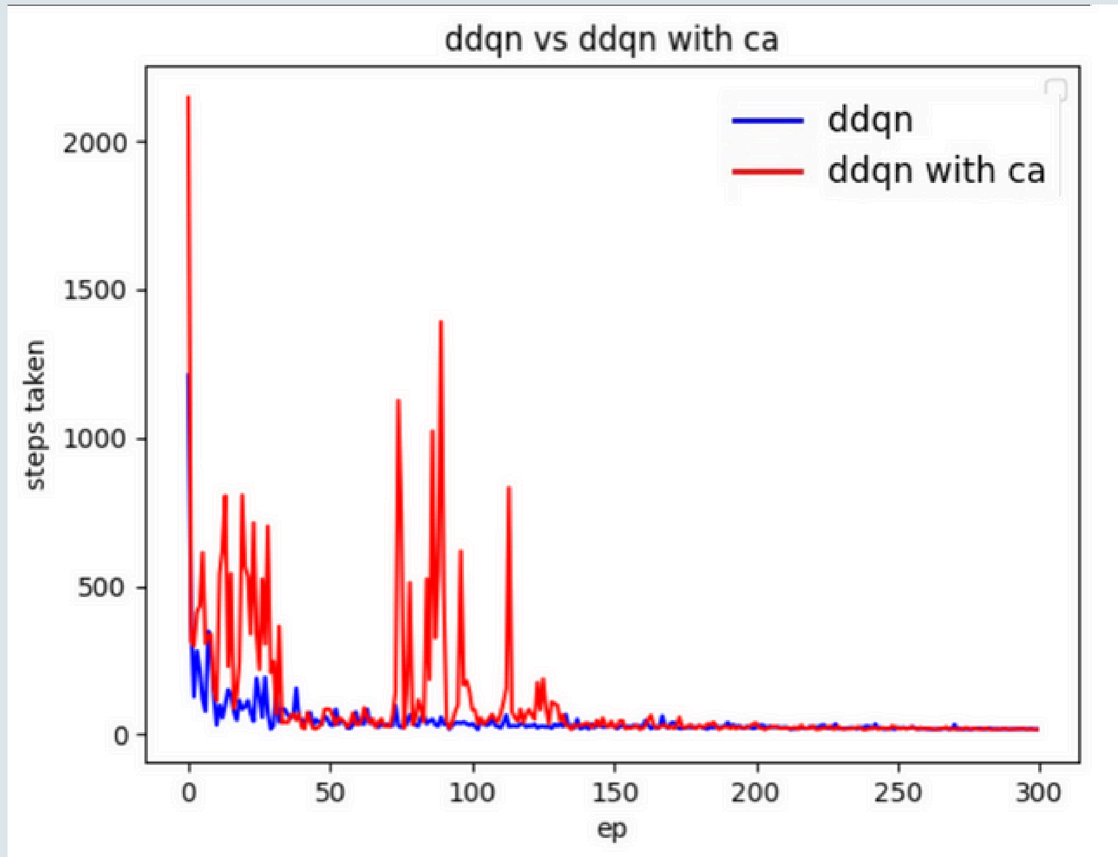
(c) D3qn vs D3qn-CA

RESULTS

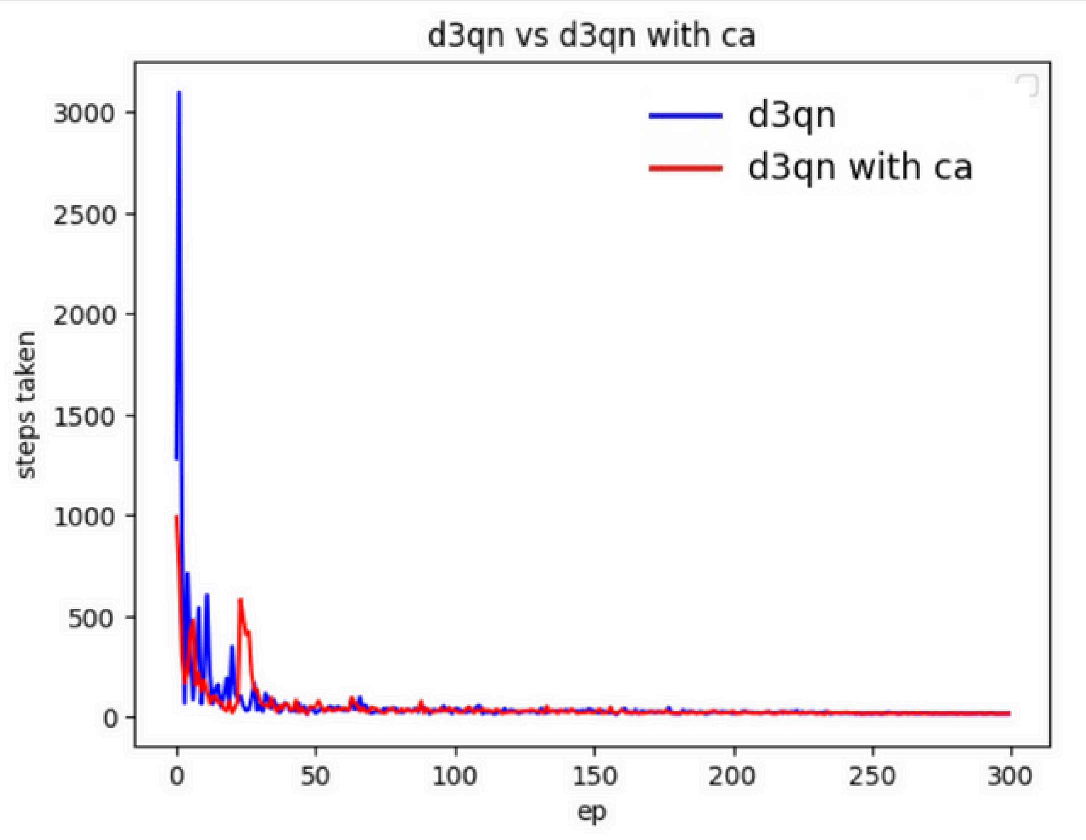
Comparison graphs : No of steps vs Episodes for Complex environment



(a) Sarsa vs Dqn vs Dqn-CA



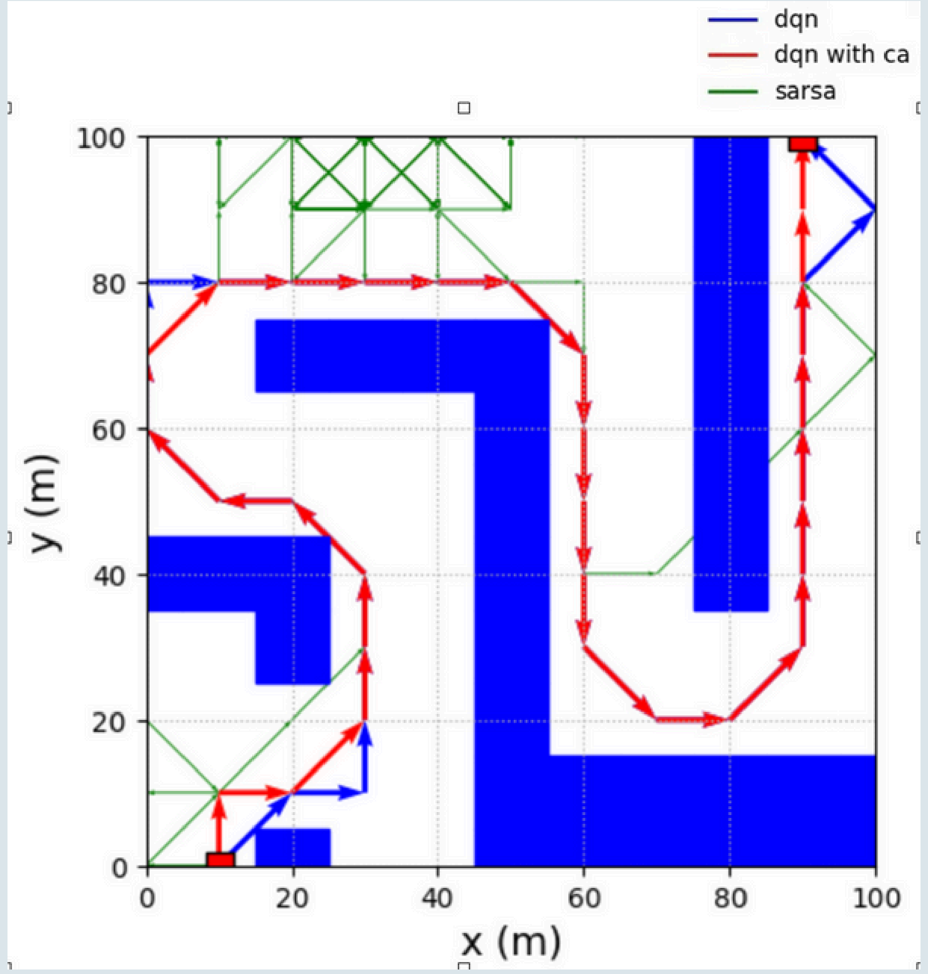
(b) Ddqn vs Ddqn-CA



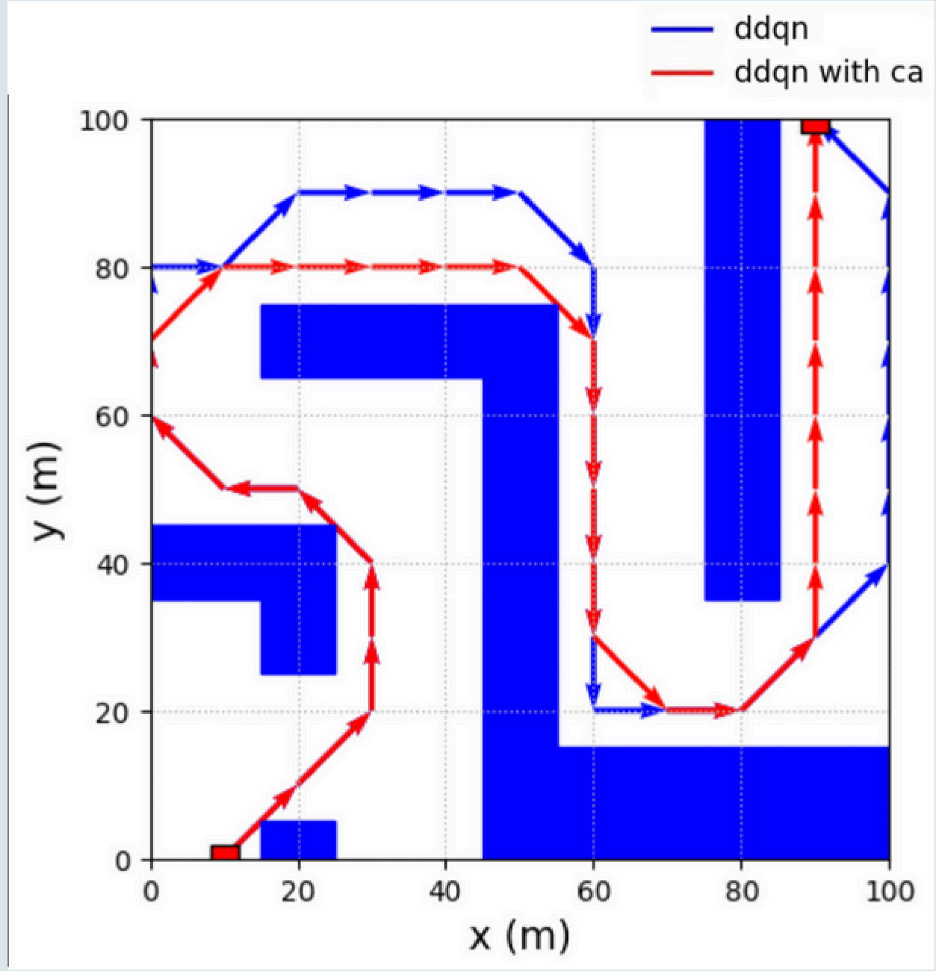
(c) D3qn vs D3qn-CA

RESULTS

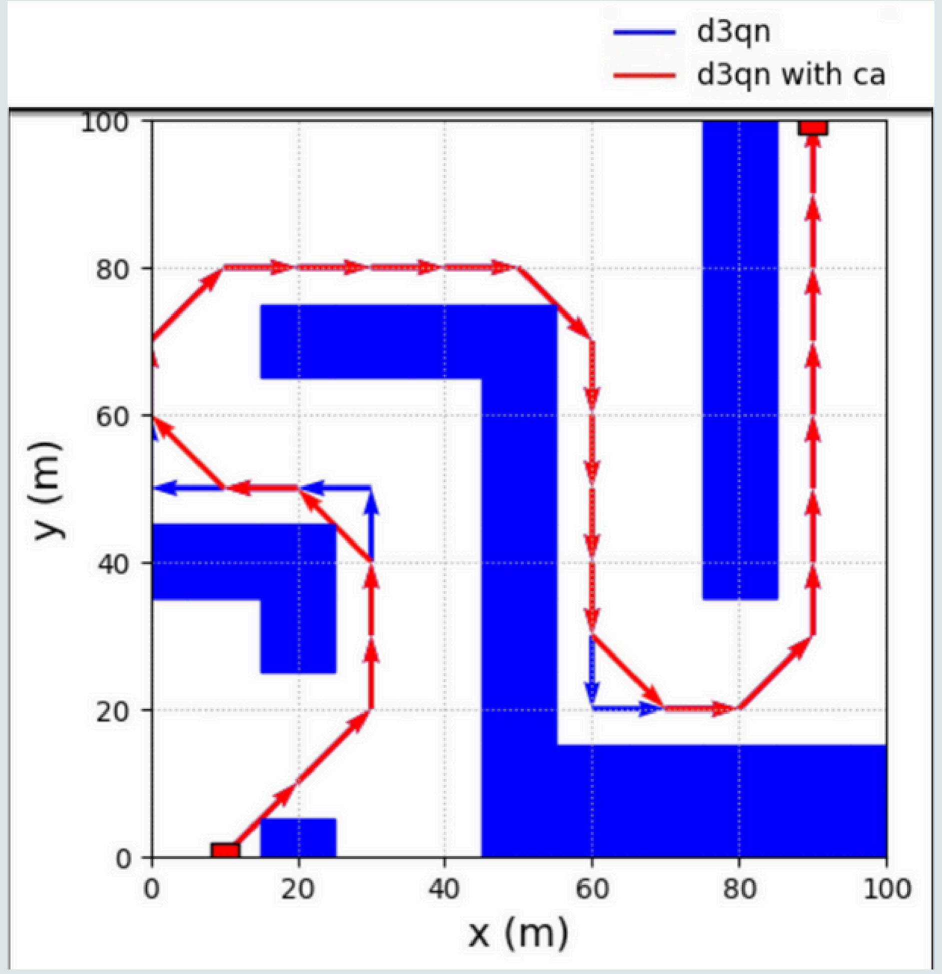
Comparison graphs : Paths taken for Complex environment



(a) Sarsa vs Dqn vs Dqn-CA



(b) Ddqn vs Ddqn-CA



(c) D3qn vs D3qn-CA

CONCLUSION

- Incorporating channel attention mechanism into models like D3QN, DDQN and DQN led to enhanced performance of models.
- Better path efficiency, higher cumulative rewards, and improved convergence stability are results of this integration.
- Channel attention significantly boosts efficiency and reward optimization in complex reinforcement learning tasks.

FUTURE WORK

- **3D environment :** Exploring in 3d environments present opportunities to test reinforcement learning models in realistic settings with increased complexity. Key areas include:
- **Model Training :** Implementing attention mechanisms on other algorithms such as PPO and DDPG for navigation and obstacle avoidance to maximise rewards and stability.
- **Additional Attention Mechanisms :** Implementing spatial and goal-based attention from camera feeds to improve situational awareness and focused decision-making.
- **Real-World Applications :** Aligning these models with practical uses in robotics, virtual reality, and automated navigation to evaluate their scalability in complex environments.

REFERENCES

1) Yu Wu, Niansheng Chen, Guangyu Fan, Dingyu Yang, Lei Rao, Songlin Cheng, Xiaoyong Song, Yiping Ma (2024). NAVS: A Neural Attention-Based Visual SLAM for Autonomous Navigation in Unknown 3D Environments. Journal of Neural Processing letters.

2)Sewak, M. (2019). Deep Q Network (DQN), Double DQN, and Dueling DQN. In: Deep Reinforcement Learning. Springer, Singapore.

3) Mehmet Gök, Dynamic path planning via Dueling Double Deep Q-Network (D3QN) with prioritised experience replay, Journal of Applied Soft Computing, Volume 158, 2024, 111503,ISSN 1568-4946.

4) Guo M-H, Xu T-X, Liu J-J, Liu Z-N, Jiang P-T, Mu T-J, Zhang S-H, Martin RR, Cheng M-M, Hu S-M (2022). Attention mechanisms in computer vision: a survey. Journal of Computational Visual Media.

5) Yin-Hao Wang, Tzuu-Hseng S. Li, Chih-Jui Lin, Backward Q-learning: The combination of Sarsa algorithm and Q-learning, Engineering Applications of Artificial Intelligence,Volume 26, Issue 9,2013,Pages 2184-2193,ISSN 0952-1976.
6) Mao, H., Zhang, Z., Xiao, Z. et al. Learning multi-agent communication with double attentional deep reinforcement learning. Auton Agent Multi-Agent Syst 34, 32 (2020).
7) Lennart Bramlage, Aurelio Cortese, Generalised attention-weighted reinforcement learning, Journal of Neural Networks, Volume 145, 2022, Pages 10-21, ISSN 0893-6080
8) Bernardo Avila Pires, Feryal Behbahani, Hubert Soyer, Kyriacos Nikiforou, Thomas Keck, Satinder Singh (2023). Hierarchical Reinforcement Learning in Complex 3D environments.
9) Haosen Qin, Tao Meng, Kan Chen, Zhengwei Li, A comparative study of DQN and D3QN for HVAC system optimisation control, Journal of Energy, Volume 307, 2024, 132740, ISSN 0360-5442.
10)Xie, L., Wang, S., Markham, A., & Trigoni, N. (2018). Towards monocular vision based obstacle avoidance through deep reinforcement learning. arXiv preprint arXiv:1706.09829.

Thank You