VISVESVARAYA TECHNOLOGICAL UNIVERSITY

JNANA SANGAMA, BELAGAVI-590014

Project Report on

"Reinforcement Learning For Self-Driving Car"

Submitted in partial fulfillment of the requirements of

B.E CSE (Artificial Intelligence & Machine Learning) degree award

Submitted by:

SRIVALLI Y

(USN: 3VC21CI052)

Under the Guidance of :

Prof &Hod: Dr.SAI MADHAVI D

Department of CSE (Artificial Intelligence & Machine Learning)

RAO BAHADUR Y. MAHABALESWARAPPA ENGINEERING COLLEGE

Approved by AICTE & Affiliated to VTU Belagavi

(NAAC A+)

Cantonment, BALLARI – 583104 Karnataka

2024 – 2025

VV SANGA'S

RAO BAHADUR Y. MAHABALESWARAPPA ENGINEERING COLLEGE

Cantonment, BALLARI – 583104

Department of CSE (Artificial Intelligence & Machine Learning)

CERTIFICATE

This is to Certify that the Project work entitled **"REINFORCEMENT LEARNING FOR SELF-DRIVING CAR"** is a bonafide work carried out by **Ms.SRIVALLI Y** bearing the **USN: 3VC21CI052** of **VII** semester inpartial fulfillment for the award of **Bachelor of Engineering** in **Department of CSE (Artificial Intelligence & Machine Learning) of Visvesvaraya Technological University, Belagavi** during the year 2024- 2025. The Project work has been approved as it satisfies the academic requirements in respect of Project work prescribed for the Bachelor of Engineering Degree.

Signature of Guide	Signature of Coordinator	Signature of Coordinator
Dr.SAI MADHAVI D	**Prof. GANGADHARA G H**	**Dr. SAI MADHAVI D**
Prof.&Hod, RYMEC, Ballari	**AP, RYMEC, Ballari**	**Prof. & HOD, RYMEC, Ballari**

Signature of HOD	Signature of Principal
Dr. SAI MADHAVI D	**Dr. HANUMANTHA REDDY**
Prof. & HOD, RYMEC, Ballari	**RYMEC, Ballari**

Name of the Examiners	Signature with Date
1.	
2.	

ACKNOWLEDGEMENT

The success and final outcome of this Project work required a lot of guidance and assistance from many people and we extremely privileged to have got this all along the completion of our Project. All that we have done is only due to such supervision and assistance and we will not forget to thank them.

We express our gratitude and respect to **Dr. T. Hanumanth Reddy** Principal, granting permission to carry out Project work.

We express our heartfelt gratitude to **Dr. Sai Madhavi D** Head of Department CSE (Artificial Intelligence & Machine Learning), Rao Bahadur Y. Mahabaleswarappa Engineering College, Ballari for being supportive for our work throughout the project.

We would like to sincerely thank our Project Coordinator's **Dr. Sai Madhavi D and Prof. Gangadhara G. H** for their valuable guidance, constant assistance, support, endurance and constructive suggestions for the betterment of the project, without which this project would not have been possible.

We would like to sincerely thank our Project Guide. **Prof. Dr.SAI MADHAVI D** for her timely guidance, constant support, motivation with good suggestions for the successful completion of this Project.

Last but not least we would like to thank my **Parents and Friends** for their moral support during critical phase of my work.

Project Associated by:

SRIVALLI Y
(3VC21CI052)

CHAPTER-1: INTRODUCTION

Driving a car requires precision and active attention. Any distraction to the driver could lead to accidents and losses. According to the data provided by Singapore Police Force, there were 94.34% of the road accidents led by human errors in year 2016 which could be avoided, for instance, failing to keep proper lookout, failing to have proper control and change lane without due care. Other road users are affected by the road accidents even though they follow the rules and regulations.

Lately, drivers divert their attention to their smartphones to answer calls, send messages or navigate to their destinations. It increases the chances of road accident when drivers are not focusing on the road condition.

The automotive industry is innovating on the smart features of car to further improve the safety and driving experience for users. The automatic braking system or collision avoidance system is one of the innovation that uses sensors such as radar to detect and mitigate collision. The innovation in preventive safety mechanism catalyses the development process of autonomous vehicle that can drive to the destination autonomously and then reduce the number of road accidents that were caused by human errors by prioritizing the safety of passengers and other road users including motorcyclists and pedestrians.

Self-driving car is the future development in automotive industry. The breakthrough in performance of parallel processing hardware and innovation in neural network design allows an artificial intelligence(AI) agent to control a vehicle autonomously. An AI agent is more capable of sustained attention and focus than humans. However, building an autonomous vehicle is a long-standing goal and complex task. Driving requires three main capabilities including recognition, prediction and planning[1].

First of all, an AI agent requires to recognize the environment with sensors. Convolutional neural networks (CNN) are the most successful neural network model that is scalable and able to extract important spatial features from spatial data. Moreover, an AI agent has to predict the future states of the environment and perform best action at the moment.

Furthermore, the planning component integrate the ability of model to understand the environment (recognition) and its dynamics (prediction) to plan the future actions to avoid unwanted situations (penalties) and drives safely to its destination (rewards). The video feed from camera and output of Light Detection and Ranging(LIDAR) are mainly used as the visual feed to AI agents.

1.1 Background

An autonomous vehicle requires to carry out multiple tasks to drive on the road with dynamic traffic condition. The tasks include following the road marking, keeping a safe distance with vehicles, reacting to emergency situation accordingly and navigating to the destination. Udacity has partnered with Didi Chuxing to organize the first self-driving car challenge in year 2017 to come up with the best way to detect obstacles using camera and LIDAR data. It could ensure self-driving cars prioritize the safety of road users and pedestrians. Besides that, Udacity has organized another challenge to predict the steering angle with deep learning based on the image inputs from a camera mounted to the windshield.

However, the quality of the driving skill and its efficiency in getting passengers to the destination are often neglected while most of the researches and designs focus on the basic features of self-driving car. The driving skill of driving agent becomes the priority to deliver the best transportation service compared to a manned vehicle. For instance, a self-driving caragent should be capable to switch to the optimal lane which maximize its speed rather than follow slow-moving vehicles in a predefined lane all the time. It is important because passengers expect to reach its destination by taking the shortest journey and shortest time along with safety.

1.2 Objective

The objective of this project is to design and implement a self-driving car-agent to maximize the its speed. A simulation environment software is required to simulate and visualize the actual traffic environment. A neural network model is essential to build a self-driving car-agent that interact with complex traffic condition. Reinforcement learning technique is used to train the model to adapt to the dynamic environment with infinite action-state sequences. The agent is required to learn the optimal policy to choose the optimal available option which maximizes the speed of vehicle.

1.3 Scope

The scopes of this project are to maximize the speed of a self-driving car on an expressway with seven lanes in a simulated environment. The simulated environment is built with pygame framework in Python environment. The autonomous vehicle is assumed to function perfectly to steer along the road. It simulates the road conditions where a self-driving car-agent on a 3 seven-lane roads with the other non-agent-controlled cars with several human driving behavior’s.

1.4 Existing Systems

Autonomous cars rely on a combination of hardware and software to navigate and operate without human intervention. Several existing systems power these vehicles, each incorporating advanced technologies such as AI, computer vision, and sensor fusion.

1. Key Systems in Autonomous Cars

(a) Perception System

- Uses sensors (LiDAR, cameras, radar, ultrasonic) to detect surroundings.
- Identifies obstacles, pedestrians, traffic signals, and road conditions.
- Example: Tesla's Autopilot uses a vision-based system, while Waymo relies on LiDAR.

(b) Localization & Mapping

- Uses GPS, IMU (Inertial Measurement Unit), and HD Maps for positioning.
- SLAM (Simultaneous Localization and Mapping) helps in real-time environment mapping.
- Example: Waymo's HD mapping provides centimeter-level accuracy.

(c) Planning & Decision Making

- Uses AI algorithms and deep learning to make driving decisions.
- Considers factors like road rules, traffic, and pedestrian movement.
- Example: NVIDIA's DRIVE platform uses deep neural networks for planning.

(d) Control System

- Converts decisions into actions like steering, acceleration, and braking.
- Uses drive-by-wire technology for electronic control of vehicle functions.

2. Levels of Autonomous Systems (SAE Levels 0-5)

- Level 0: No automation (manual driving).
- Level 1: Driver assistance (adaptive cruise control, lane-keeping).
- Level 2: Partial automation (Tesla Autopilot, GM Super Cruise).
- Level 3: Conditional automation (Mercedes-Benz Drive Pilot).
- Level 4: High automation (Waymo, Cruise – in geofenced areas).
- Level 5: Full automation (No human intervention – still in development).

3. Existing Autonomous Car Companies & Technologies

- Tesla Autopilot & FSD (Full Self-Driving): Vision-based AI.
- Waymo (Google): LiDAR-heavy approach with HD mapping.
- Cruise (GM): Robotaxis operating in select cities.
- Mobileye (Intel): Advanced driver-assistance systems (ADAS).
- NVIDIA DRIVE: AI-powered autonomous vehicle platform.

1.5 Proposed Systems

A proposed system for an autonomous car aims to improve upon existing technologies by increasing safety, accuracy, and efficiency. It integrates advanced AI, sensor fusion, V2X (Vehicle-to-Everything) communication, and improved decision-making algorithms to achieve full automation (SAE Level 5).

1. Key Enhancements in the Proposed System

(a) Advanced Sensor Fusion

- Combines LiDAR, RADAR, Cameras, and Ultrasonic Sensors with AI for better object detection.
- Uses thermal imaging to enhance night and fog vision.

(b) High-Precision Localization & Mapping

- Uses 5G-based GPS & Real-Time Kinematic (RTK) positioning for centimeter-level accuracy.
- Integrates SLAM (Simultaneous Localization and Mapping) for real-time adjustments.
- Dynamic HD maps with AI-driven updates to account for road changes.

(c) AI-Driven Decision-Making

- Implements deep reinforcement learning (DRL) to adapt driving behavior dynamically.
- Uses predictive AI models to anticipate pedestrian and vehicle movements.
- Applies edge computing for faster real-time decision-making without cloud dependency.

(d) V2X Communication (Vehicle-to-Everything)

- Vehicle-to-Vehicle (V2V): Communicates with nearby vehicles for coordinated movement.
- Vehicle-to-Infrastructure (V2I): Interacts with traffic lights, road signs, and smart city systems.
- Vehicle-to-Pedestrian (V2P): Detects pedestrians and cyclists for improved safety.

(e) Enhanced Cybersecurity & Privacy

- Uses blockchain for data integrity and secure communication between vehicles.
- Implements AI-driven anomaly detection to prevent hacking attempts.

(f) Sustainable Energy & Efficiency

- Optimizes battery usage in electric vehicles through AI-powered energy management.
- Uses solar-assisted power systems to increase energy efficiency.

2. Proposed Architecture of the System

- Perception Layer: Sensors collect data (LiDAR, Radar, Cameras, GPS).
- Processing Layer: AI algorithms analyze sensor data for object detection, and mapping.

- Decision-Making Layer: Reinforcement learning AI decides acceleration, braking, and lane changes.
- Communication Layer: V2X communication exchanges data with vehicles and infrastructure.
- Control Layer: Drive-by-wire system executes vehicle control commands.

3. Expected Benefits

- Higher Safety: AI-based predictive analysis reduces accidents.
- Better Efficiency: 5G, edge computing, and V2X improve traffic flow.
- Scalability: Works in complex urban environments.
- Sustainability: Optimized for electric vehicles with minimal energy waste.

CHAPTER-2: LITERATURE SURVEY

1. A Comprehensive Survey on Autonomous Vehicle Technologies

Summary: This survey explores key technologies enabling self-driving cars, including machine learning, sensor fusion, computer vision, and control systems. It discusses different levels of automation, challenges in real-world deployment, and future trends.
Reference: Thrun, S. (2010). "Toward Robotic Cars." Communications of the ACM, 53(4), 99–106.

2. Sensor Fusion and Perception in Autonomous Vehicles

Summary: This study reviews different sensors used in autonomous vehicles, such as LiDAR, radar, cameras, and IMUs. It also examines sensor fusion techniques and their role in perception, object detection, and scene understanding.
Reference: Alam, K., Sarker, M. R., & Kowadlo, G. (2014). "A Review of Sensor Fusion Techniques in Autonomous Vehicles." Journal of Robotics and Automation, 2(3), 45–57.

3. Deep Learning Applications in Autonomous Driving

Summary: This paper reviews the role of deep learning in perception, decision-making, and control. It highlights CNNs for object detection, RNNs for trajectory prediction, and reinforcement learning for behavior planning.
Reference: Bojarski, M., Testa, D., & Dworakowski, D. (2016). "End to End Learning for Self-Driving Cars." arXiv preprint arXiv:1604.07316.

4. Challenges and Ethical Considerations in Autonomous Driving

Summary: The study discusses ethical dilemmas (e.g., decision-making in crash scenarios), legal issues, cybersecurity threats, and public trust in self-driving cars.
Reference: Bonnefon, J. F., Shariff, A., & Rahwan, I. (2016). "The Social Dilemma of Autonomous Vehicles." Science, 352(6293), 1573–1576.

5. Safety and Reliability in Autonomous Vehicles

Summary: This survey examines the reliability of autonomous driving systems, covering software validation, real-world testing, and failure analysis. It also reviews accident case studies involving self-driving cars.

Reference: Koopman, P., & Wagner, M. (2017). "Autonomous Vehicle Safety: An Interdisciplinary Challenge." IEEE Intelligent Transportation Systems Magazine, 9(1), 90–96.

6. V2X Communication and Its Role in Autonomous Vehicles

Summary: This literature survey focuses on vehicle-to-everything (V2X) communication, including V2V (vehicle-to-vehicle) and V2I (vehicle-to-infrastructure). It discusses protocols, latency challenges, and security concerns in connected autonomous driving.
Reference: Abbas, R., Clancy, T., & Gopalakrishnan, K. (2018). "A Survey on V2X Communication for Autonomous Vehicles." IEEE Access, 6, 35089–35112.

7. Motion Planning and Control for Autonomous Vehicles

Summary: This paper reviews different motion planning algorithms like A*, RRT (Rapidly-exploring Random Trees), and optimization-based approaches. It also discusses vehicle dynamics, trajectory generation, and real-time control techniques.
Reference: Paden, B., Čáp, M., Yong, S. Z., Yershov, D., & Frazzoli, E. (2016). "A Survey of Motion Planning and Control Techniques for Self-Driving Urban Vehicles." IEEE Transactions on Intelligent Vehicles, 1(1), 33–55.

8. Human-Machine Interaction in Autonomous Vehicles

Summary: This study explores how autonomous vehicles interact with human drivers, pedestrians, and passengers. It includes topics such as user trust, explainability of AI decisions, and transition of control in semi-autonomous systems.
Reference: Merat, N., Madigan, R., & Nordhoff, S. (2018). "Human Factors, User Requirements, and User Acceptance of Ride-Sharing in Automated Vehicles." Transportation Research Part A: Policy and Practice, 122, 116–127.

9. Edge Computing for Autonomous Vehicles

Summary: This paper examines how edge computing enhances real-time decision-making in autonomous vehicles. It discusses reducing latency in perception and control through decentralized processing near the vehicle.
Reference: Hou, X., Li, Y., Chen, M., Wu, D., Jin, D., & Chen, S. (2016). "Vehicular Fog Computing: A Viewpoint of Vehicles as the Infrastructures." IEEE Transactions on Vehicular Technology, 65(6), 3860–3873.

10. Cybersecurity Threats and Countermeasures in Autonomous Vehicles

Summary: This survey explores various cybersecurity challenges in autonomous vehicles, including hacking, spoofing attacks on sensors, and software vulnerabilities. It also reviews security frameworks and encryption techniques.
Reference: Petit, J., & Shladover, S. E. (2015). "Potential Cyberattacks on Automated Vehicles." IEEE Transactions on Intelligent Transportation Systems, 16(2), 546–556.

11. Energy Efficiency and Sustainability in Autonomous Driving

Summary: This paper reviews how autonomous vehicles impact fuel consumption, electric vehicle integration, and smart traffic management. It discusses how AI-driven driving can optimize energy efficiency and reduce carbon emissions.
Reference: Wadud, Z., MacKenzie, D., & Leiby, P. (2016). "Help or Hindrance? The Travel, Energy, and Carbon Impacts of Highly Automated Vehicles." Transportation Research Part A: Policy and Practice, 86, 1–18. 12. Machine Learning and AI for Autonomous Driving

12. Machine Learning and AI for Autonomous Driving

Summary: This survey explores the use of artificial intelligence in self-driving cars, focusing on supervised and reinforcement learning techniques for perception, decision-making, and control. It also reviews dataset availability and AI model performance evaluation.
Reference: Grigorescu, S., Trasnea, B., Cocias, T., & Macesanu, G. (2020). "A Survey of Deep Learning Techniques for Autonomous Driving." Journal of Field Robotics, 37(3), 362–386.

13. Autonomous Vehicles in Smart Cities

Summary: This paper reviews the integration of autonomous vehicles into smart city infrastructure, covering traffic management, urban planning, and environmental impact. It discusses how AVs interact with IoT-based smart transportation systems.
Reference: Talebpour, A., & Mahmassani, H. S. (2016). "Influence of Connected and Autonomous Vehicles on Traffic Flow Stability and Throughput." Transportation Research Part C: Emerging Technologies, 71, 143–163.

CHAPTER-3: SYSTEM REQUIREMENT SPECIFICATIONS

3.1 FUNCTIONAL REQUIREMENTS

Functional requirements for a self-driving car define the essential tasks and behaviors the system must perform to ensure safe and efficient autonomous operation. These include:

1. Perception & Sensing

- Detect and classify objects (vehicles, pedestrians, cyclists, road signs, lane markings, traffic signals).
- Use sensors (LiDAR, radar, cameras, ultrasonic sensors) to gather real-time environmental data.
- Detect road conditions (potholes, weather impact like rain or snow).

2. Localization & Mapping

- Continuously determine vehicle position using GPS, IMU, and sensor fusion.
- Maintain and update high-definition (HD) maps for navigation.
- Identify road lanes, intersections, and landmarks for better path planning.

3. Path Planning & Decision Making

- Plan a safe and optimal route from the source to the destination.
- Make real-time driving decisions (lane changes, overtaking, merging, turning, stopping at signals).
- Adapt to dynamic road conditions (accidents, detours, traffic congestion).

4. Control & Motion Execution

- Execute acceleration, braking, and steering commands smoothly.
- Maintain lane discipline and follow speed limits.
- Ensure smooth handling in emergency scenarios (sudden braking, obstacle avoidance).

5. Communication & Human Interaction

- Communicate with other vehicles (V2V) and infrastructure (V2I) for traffic updates.
- Notify passengers of critical decisions (e.g., emergency braking).
- Allow manual override and ensure seamless transition between autonomous and manual modes.

6. Safety & Compliance

- Comply with traffic rules and regulations.
- Detect and respond to emergencies (ambulances, law enforcement, roadblocks).
- Implement fail-safe mechanisms (safe stopping in case of system failure).

7. Cybersecurity & Data Management

- Ensure data encryption and secure vehicle-to-cloud communication.
- Protect against hacking and unauthorized access.
- Log and analyze driving data for continuous improvement.

3.2 NON-FUNCTIONAL REQUIREMENTS

Non-functional requirements (NFRs) for a self-driving car define the quality attributes, constraints, and overall system performance expectations. These ensure reliability, safety, and user experience beyond just functional capabilities.

1. Performance

- Real-time Processing: The system must process sensor data and make driving decisions within milliseconds.
- Low Latency: Decision-making (braking, steering, obstacle avoidance) should happen with minimal delay.
- High Accuracy: Object detection, localization, and path planning must achieve near 100% accuracy.

2. Safety & Reliability

- Fail-Safe Mechanisms: The system should transition to a safe state in case of sensor failure, software errors, or cyber-attacks.
- Redundancy: Critical components (braking, steering, computing) must have backups.
- System Uptime: The autonomous system should operate with at least 99.99% availability.

3. Security

- Data Encryption: Secure all communications (V2V, V2I, cloud) to prevent hacking.
- Access Control: Restrict unauthorized modifications or control of vehicle functions.
- Anomaly Detection: Identify and respond to cyber threats in real-time.

4. Scalability & Maintainability

- Over-the-Air (OTA) Updates: Support remote software updates without disrupting functionality.
- Modular Architecture: Components should be upgradable or replaceable without affecting the whole system.
- Self-Diagnosis: Ability to detect and report hardware or software issues proactively.

5. User Experience (UX) & Human Interaction

- Smooth Transitions: Ensure seamless handoff between autonomous and manual driving.
- Clear Communication: Provide intuitive alerts and updates to passengers (e.g., route changes, emergency stops).
- Comfortable Ride: Optimize acceleration, braking, and turns for a smooth experience.

6. Compliance & Regulatory Standards

- Legal Compliance: Must adhere to regional traffic laws and safety regulations (e.g., ISO 26262 for automotive safety).
- Ethical Decision-Making: Ensure ethical handling of unavoidable crash situations.
- Testing & Certification: Pass rigorous validation under various weather, road, and traffic conditions.

CHAPTER-4: ALGORITHMS

4.1 REINFORCEMENT LEARNING

Reinforcement learning is one of the learning approach for neural networks. Reinforcement learning implements differently compared to supervised learning and unsupervised learning. Reinforcement learning makes an agent to generate the optimal policy in an environment and maximizes the reward [2]. An agent can learn the rewards given by taking action in respective state and subsequently learn the proper action for each state without having any predefined rules and knowledge about the environment [3]. The agent learns by trials and errors, or in another words, the actions with higher reward are reinforcement while actions with penalties are avoided in the future. It is different with supervised learning because supervised learning requires huge number of labelled dataset to train the model [4]. The training process is repeated with same set of data until the model converges. It requires efforts in labelling the data and the process is error-prone[5]. Reinforcement learning is also different with unsupervised learning because reinforcement learning aims to maximize the reward while unsupervised learning does not.

4.2 CONVOLUTION NUERAL NETWORK

The efficiency and learning ability of brain have inspired the innovation of the convolutional neural networks (CNN). The ability of animal to perceive features from complex visual map and generate responses has biological guided the construction of CNN [11]. The local correlation can be extracted spatially by allowing a local connectivity pattern between neighboring cells of adjacent layers. Figure 1 shows that each filter is applied across the entire input map and generate a feature map. The output of the feature map is connected to the adjacent layer with features extracted. Each feature map explores unique features regardless of the location across visual field.

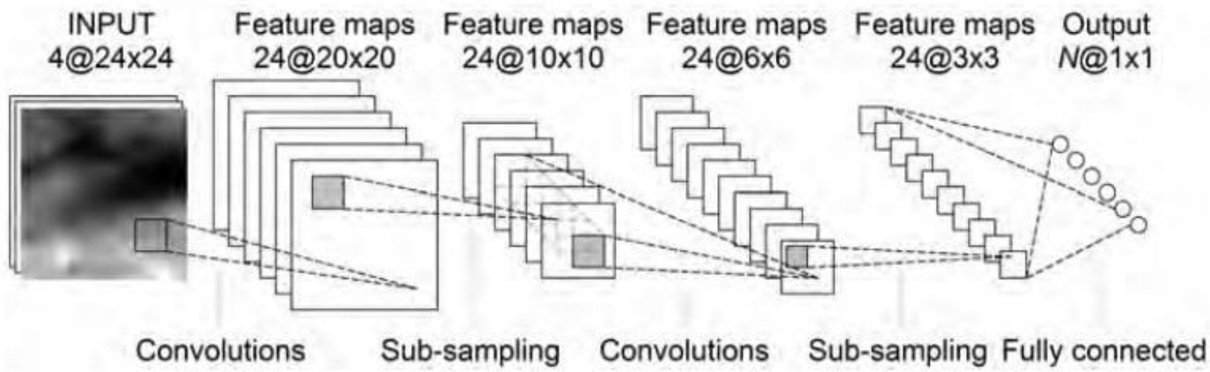

Figure 1: The illustration of convolutional layers and fully-connected layers forming convolutional neural network

4.3 DEEP Q-LEARNING NETWORKS(DQN)

Neural network is utilized as the function approximator of the Q-learning algorithm to form the Deep Q-learning network (DQN). The objective of DQN is to minimize the Mean Square Error (MSE) of the Q-values. Besides that, experience replay technique uses first-in-first-out(FIFO) method to keep experience tuples in replay memory for every time step. The replay memory stores experience tuples for several episodes to ensure that the memory holds diversified experiences for different states. The experience tuples are sampled from replay memory randomly during Q-learning updates. The implementation of experience replay can remove correlations in the observation sequence and smoothing over changes in the data distribution by randomizes over the data [13]. In practice, the most recent N experience records are sampled uniformly and randomly from replay memory. The random sampling gives equal probability to all experience tuples.

4.4 DOUBLE DEEP Q-LEARNING(DDQN)

In DQN, a batch of experience tuples are sampled in each learning iteration. The Q-learning network is updated based on the difference of the estimated future rewards. The sampling process is random and the frequent update of the network destabilizes the learning process. The network is facing difficulties to reach convergence. Thus, a separate target network, Q^ is used to improve the stability of the main neural network [14]. The target network Q^ is used to estimate the discounted future rewards given the subsequent state Q^,12 and action Q^,12 at unit time step Q^. The target network Q^ is synchronized with the main network Q^ for every C Qlearning updates on the main network.

4.5 MARKOV DECISION PROCESS

Markov Decision Process (MDP) can make a sequence of decisions based on the utilities of environment state. In which a set of possible environment states(S), a set of possible actions(A), discount factor(�), reward function (R), and state transition probability(ℙ). Environment states S are sequence of states with initial state s0. Actions A are available actions for environment states S. Reward function R is a map of immediate reward given (state, action) pair. State transition probability ℙ is the probability of transition from states S1. … Sn. However, there are infinite state sequences in the environment and calculation of transition probability for all state sequences is not feasible.

4.6 Q-LEARNING

Q-learning is a reinforcement learning mechanism that compare the expected utility of available actions given a state. Q-learning can train a model to find the optimal policy in any finite MDP.

4.7 ALGORITHM

Reinforcement learning is the learning approach used in this project. Algorithm from [8] is used to interact with the simulator and use the surrounding environment states as the input of the algorithm. The S number of the most recent states St, at time step t are supplied to as input. The available action options are Accelerate (A), Decelerate (D), Left (L), Right (R) and Maintain (M). Epsilon greedy exploration is implemented to maximize the exploration and ensure agent experiences more possible states in the state domain. An action at, is randomly chosen from the action options as the agent's decision when the probability is lower than the current epsilon-greedy probability. The epsilon-greedy probability is uniformly decreases with number of processed states. Otherwise, an action with highest Q-value among the other options is chosen. The chosen action is then relayed to the simulator and the resulting state St+1,12 and reward rt, is recorded. Reward rt, is the difference of score at time step t and t− 1. The experience tuple (st,at,rt,st+1) is appended to the replay memory D with capacity N.

The learning algorithm starts at F frames so that the learning process does not over fit the limited experiences in the replay memory D. Experience replay is deployed to apply Q-learning update gradually. Learning process samples the most recent N experience rows in the replay memory. The goal of the reinforcement learning agent is to perceive the environment in the simulator and the optimal action that maximizes the future reward is chosen. The optimal Qvalue function Q*(s,a) can be modelled by a deep learning model as the nonlinear function approximator of the Q-learning algorithm. Neural networks are used to approximate the optimal Q-learning network. Thus, the parameterized neural network is used to approximate the optimal Q-learning network.

Frame-skipping technique is adopted during training. The agent reacts with the surrounding environment every kth frame to increase the response time of agent's reaction. In this project, the actions for skipped frames are determined in Table 3. Action at kth frame Action for skipped frames are determined in the following table.

Action at k^{th} frame	Action for skipped frames
Accelerate (A)	Accelerate (A)
Decelerate (D)	Accelerate (A)
Left (L)	Maintain (M)
Right (R)	Maintain (M)
Maintain (M)	Maintain (M)

Table 3: Actions for skipped frames with respect to action at k^{th} frame.

CHAPTER-5: SYSTEM DESIGN

5.1 MODEL ARCHITECTURE

The baseline model follows the network design listed. As shown in Figure 4, the baseline network takes the environment mapping of four most recent states as input, which is a set of 4×36×3 state. The state consists of Boolean values, in which true (1) if there is obstacle in the vision field and false (0) if there is not. The first hidden layer convolves 16 filters of 3×3 and stride 2. Rectified Linear Unit (ReLU) is applied to the outputs of the first hidden layer. The outputs of first hidden layer are flatten and passed to the fully-connected layer. The fully connected linear layer comes with one output for five actions.

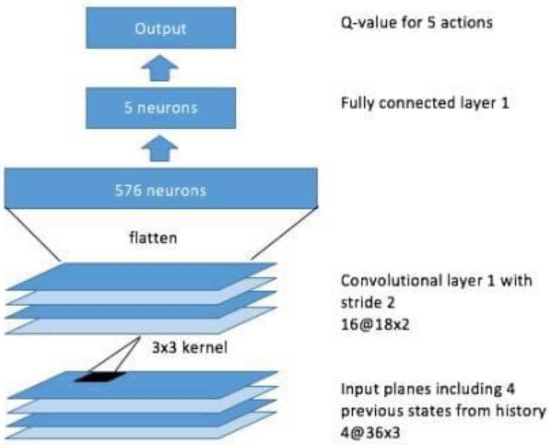

Figure 4: Baseline model with CNN architecture with one convolutional layer and one fully-connected layer.

The proposed model is implemented as shown in Figure 5. The network is enhanced with additional convolutional layer and fully-connected layer. The second convolutional layer compresses the output from the first convolutional layer to extract important features from the inputs. It passes the important features to fully-connected layer. The two fully-connected layers of the model brings the ability to approximate nonlinear function of action-value function. The additional sub-network is added to accept action stack as input. Action stack is a stack of action memory which stores the four most recent actions taken. The rational is to improve the stability of the agent from switching lane at high frequency.

- Perception (Sensors & Data Fusion): Detects surroundings.
- Localization & Mapping: Determines vehicle position.
- Perception & Object Detection: Identifies objects and road conditions.
- Decision-Making & Path Planning: Plans optimal routes and actions.
- Control & Actuation: Executes vehicle movements.
- Communication & V2X: Interacts with external systems.
- Cybersecurity & Safety: Ensures secure and fail-safe operations.

This architecture enables a self-driving car to navigate autonomously while ensuring safety, efficiency, and compliance with traffic regulations.

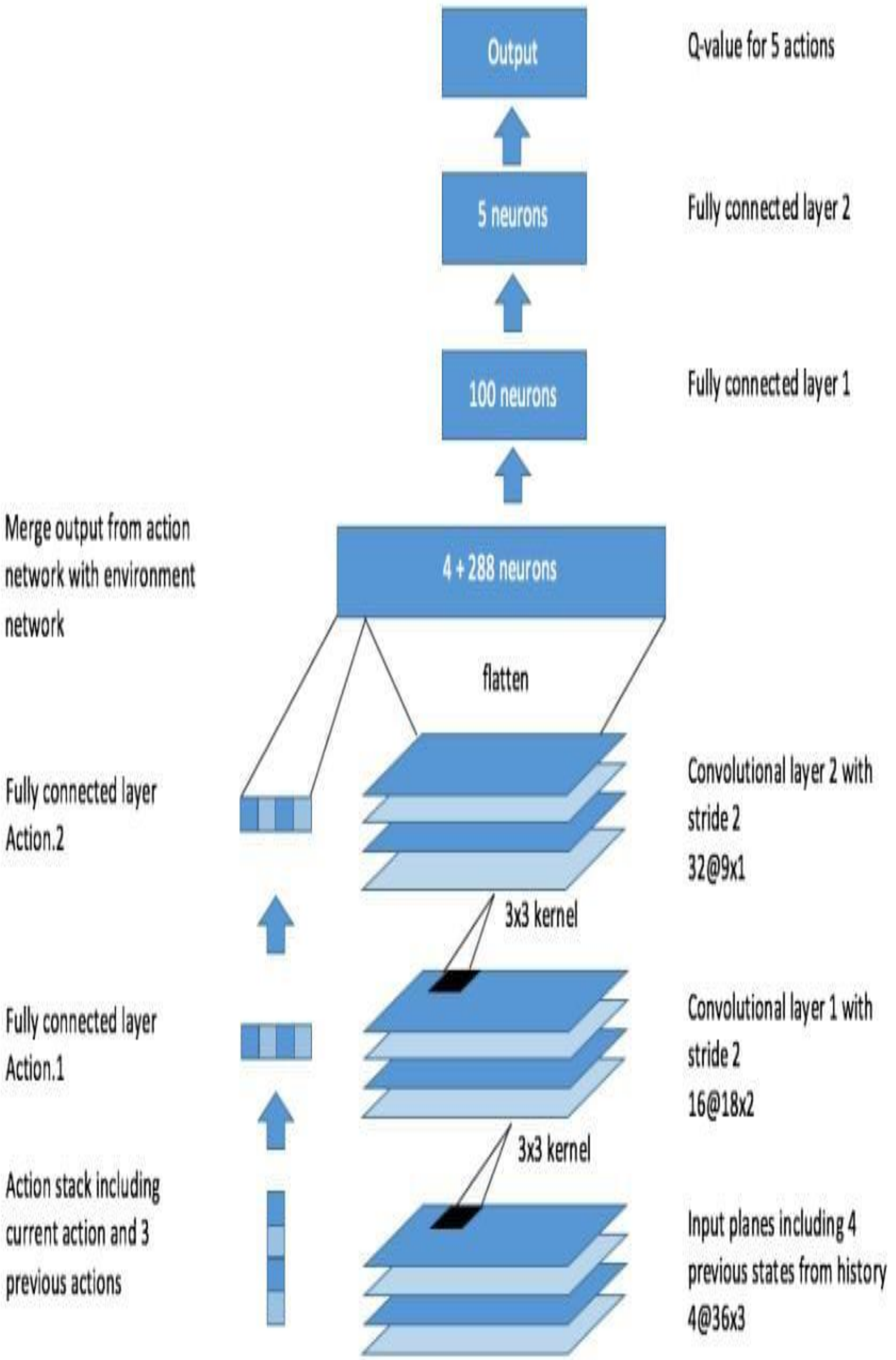

Figure 5: Proposed model with CNN architecture with 2-layer fully-connected for action stack and two convolutional layer and two fully-connected layer.