End-to-End Autonomous Driving using Proximal Policy Optimization

Sai Harshitha Peddi (018222087), Venkata Satya Charan Naga Deep (019127238), Niharika Janapati (017553744), Bhargava Phaneendra Kumar Suryadevara (018176379)

Finalized Specifications:

The goal of the project is to develop an end-to-end autonomous driving agent using Proximal Policy Optimization (PPO), a stable on-policy reinforcement learning algorithm introduced by Schulman et al. [1]. The agent will be trained in the CARLA simulator [2], which provides realistic urban driving conditions. The system will learn to control steering, throttle, and braking directly from sensory inputs such as RGB camera images or latent representations.

Our PPO implementation will include clipped objectives, value functions, generalized advantage estimation (GAE), and entropy regularization. The training process will use reward shaping to encourage route completion, lane-following, and compliance with traffic rules, while penalizing unsafe maneuvers such as collisions or lane deviations. To enhance robustness, we will implement curriculum learning (progressing from simple to complex driving tasks) and domain randomization (varying weather, lighting, and traffic conditions).

Due to the high computational cost of reinforcement learning from visual data, we will utilize the SJSU High-Performance Computing Cluster (HPC) for GPU-accelerated model training and evaluation.

Abstract:

Autonomous driving has emerged as a key research area within Artificial Intelligence, combining advances in computer vision, reinforcement learning, and control systems. Developing a safe and reliable autonomous driving system requires learning from continuous, high-dimensional sensory data and making real-time decisions in complex environments. However, direct real-world training poses significant safety, cost, and ethical challenges, making simulated learning environments essential.

This project explores an end-to-end reinforcement learning approach to autonomous driving using Proximal Policy Optimization (PPO) [1] within the CARLA simulator [2]. PPO provides a robust framework for policy-gradient optimization by balancing stability and efficiency through clipped objectives and advantage estimation. Our model will learn directly from visual input, using raw RGB images or compressed latent representations, to control steering, throttle, and braking actions.

To improve performance and generalization, we will design a custom reward function that incentivizes smooth lane-following, safe driving, and efficient route completion. The project also introduces curriculum learning, enabling the agent to gradually master increasingly complex scenarios, and domain randomization to expose it to diverse environmental conditions such as varying weather, lighting, and traffic density.

Through systematic training and evaluation, the project aims to produce a robust autonomous driving policy capable of navigating unseen environments safely. The results are expected to provide valuable insights into the scalability of deep reinforcement learning algorithms for real-world driving applications and contribute to the broader understanding of stable policy optimization in continuous control tasks.

Outline of Final Report:

The final report will begin with an Introduction and Motivation section, providing context for autonomous driving as an application of artificial intelligence and discussing the relevance of reinforcement learning approaches. It will also define the problem statement and outline the project objectives.

The Background and Related Work section will review existing literature and methods in deep reinforcement learning, including previous applications of PPO and other algorithms in the CARLA simulator, highlighting gaps our project aims to address.

The Methodology section will describe the technical details of the PPO algorithm, including policy and value network architectures, loss functions, and training mechanisms. It will also explain the CARLA simulation setup, reward function design, curriculum learning process, and domain randomization strategies used to improve robustness and generalization.

Next, the Data and Experimental Setup section will describe the datasets and environment configurations used in training, along with preprocessing steps such as image normalization and frame stacking. The Results and Evaluation Metrics section will present key performance metrics, including success rate, lane deviation, and collision frequency, supported by visualizations of driving behavior.

Finally, the Discussion and Conclusion section will summarize findings, analyze limitations, and suggest directions for future improvements, such as integrating sensor fusion or hybrid reinforcement-imitation learning approaches.

Data Sources:

The primary data source for this project will be the CARLA Simulator [2], an open-source platform that provides urban driving environments and rich sensor data, including RGB images, depth maps, semantic segmentation, and LiDAR. The simulator supports variable weather, lighting, and traffic, enabling the generation of diverse driving experiences for reinforcement learning.

In addition to simulator-generated data, the following CARLA-based datasets will be used for training and evaluation:

1. CARLA Autonomous Driving Dataset (CARLA AD Dataset) – includes RGB images, semantic segmentation labels, and control signals for urban driving.
2. NoCrash Benchmark Dataset – evaluates driving models under different traffic and weather conditions, testing robustness and generalization.
3. CARLA Leaderboard Evaluation Suite – a standardized benchmark with predefined routes and driving tasks for consistent model comparison.

We will begin exploration by setting up the CARLA environment, testing small-scale simulations in Town01 and Town03, and generating initial samples to validate the observation and reward signals. Later stages will involve adjusting simulation parameters for diversity and performing preprocessing steps such as image normalization, frame stacking, and reward scaling to prepare data for PPO training.

Repository setup:

A GitHub repository has been created to store and manage all project-related files, including code, experiment logs, and reports. The repository serves as the central collaboration hub for version control and progress tracking. It will include separate folders for source code, trained models, documentation, and experiment outputs.

Repository Link: https://github.com/bhargava-suryadevara/End-to-End-Autonomous-Driving-using-Proximal-Policy-Optimization

The repository will be updated regularly with code commits, training logs, and documentation to ensure reproducibility and continuous monitoring of project milestones.

References:

[1] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal Policy Optimization Algorithms," arXiv preprint arXiv:1707.06347, 2017.

[2] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An Open Urban Driving Simulator," arXiv preprint arXiv:1711.03938, 2017.