Lane following on monocular visual features

Dzhavid Sadreddinov
Innopolis University, Russia
d.sadreddinov@innopolis.university

Diana Vostrova
Innopolis University, Russia
d.vostrova@innopolis.university

Alsu Khairullina
Innopolis University, Russia
a.khairullina@innopolis.university

***Abstract*—This project develops an autonomous driving agent using reinforcement learning in the CARLA simulator. The agent will rely on visual features: predicted semantic maps and depth maps from sensor. These features, combined with vehicle data like velocity and steering, will serve as input for training a Proximal Policy Optimization (PPO) agent. The goal is to improve the agent's decision-making and navigation by leveraging enhanced visual understanding of the environment.**

I. Introduction

The project focuses on developing a novel autonomous driving agent in the CARLA simulator using reinforcement learning, addressing both computational efficiency and decision-making accuracy. Autonomous driving has become a prominent research area, driven by the potential to revolutionize transportation systems, reduce accidents, and improve traffic efficiency. However, achieving robust and reliable autonomy in dynamic and unpredictable environments remains a significant challenge. Simulation platforms like CARLA provide a controlled environment to develop and test driving agents under varied scenarios, making them indispensable for research in autonomous navigation.

Unlike many existing approaches that rely on extensive sensor suites and real-time depth estimation, this project adopts a minimalist yet effective strategy by focusing primarily on RGB image data. By eliminating the need for training a depth estimation model, the system reduces computational overhead and simplifies the data pipeline, making it more applicable to real-world, resource-constrained scenarios. Despite the absence of explicitly estimated depth, the approach compensates by leveraging high-level features derived from RGB data, processed through a U-Net architecture for semantic segmentation. Semantic segmentation provides a pixel-level understanding of the environment, enabling the identification of critical elements such as roads, sidewalks, vehicles, and pedestrians.

To further optimize the decision-making process, a Variational Autoencoder (VAE) is employed to compress the semantic and depth map features into compact latent representations. This compression not only reduces the dimensionality of the input data but also ensures that essential environmental information is preserved for reinforcement learning tasks. Compact feature representation is crucial for reinforcement learning algorithms, as it reduces the complexity of the state space, leading to faster convergence and improved policy learning.

The driving strategies of the agent are optimized using the Proximal Policy Optimization (PPO) algorithm, a widely recognized reinforcement learning approach known for its stability and efficiency. PPO facilitates smooth policy updates, ensuring robust learning dynamics even in the face of complex input modalities. By combining visual features with navigational observations such as velocity and steering angle, the system enables the agent to perceive and act effectively in dynamic driving environments.

In addition to RGB-based methods, the project explores the incorporation of ground-truth depth data during inference. While the primary goal is to develop an RGB-driven system, analyzing the impact of ground-truth depth enhances our understanding of the trade-offs between computational simplicity and enriched spatial representation. Ground-truth depth data offers precise spatial information, which can improve obstacle detection and path planning but at the cost of increased computational complexity.

Overall, this project integrates cutting-edge techniques in computer vision and reinforcement learning to develop an autonomous driving agent capable of efficient and reliable navigation. By addressing key challenges such as feature representation, computational efficiency, and robust decision-making, the proposed approach contributes to advancing the state-of-the-art in autonomous driving research. Future directions include refining the integration of multimodal data, enhancing generalization across diverse driving scenarios, and exploring lightweight depth estimation techniques to further improve the balance between performance and efficiency.

II. Related Work

A. U-Net for Semantic Segmentation

Semantic segmentation has been a cornerstone in computer vision tasks, enabling pixel-level classification of images to derive structured representations of the environment. The U-Net architecture, originally designed for biomedical image segmentation [1], has gained prominence due to its encoder-decoder structure and the use of skip connections that effectively combine high-level abstract features with low-level spatial details. These attributes allow U-Net to excel in scenarios requiring precise spatial localization, such as autonomous driving.

In recent years, U-Net has been extensively adapted for real-time applications in autonomous vehicles. Studies have demonstrated its capability to accurately segment road elements such as lanes, sidewalks, vehicles, and pedestrians [**?**]. For example, advanced U-Net variants integrate dilated convolutions to capture multi-scale features, further enhancing their performance in complex urban environments. Additionally, the lightweight nature of U-Net makes it particularly suitable for computationally constrained settings, such as simulations in CARLA, where maintaining real-time processing capabilities is essential.

The integration of semantic segmentation into autonomous systems has also led to the development of hybrid architectures that combine U-Net with attention mechanisms, improving the model's ability to focus on critical regions in the driving scene. These enhancements have proven particularly effective in scenarios with occlusions or varying illumination, which are common challenges in real-world driving tasks. Such advancements underline the adaptability and robustness of U-Net as a foundation for semantic feature extraction in autonomous driving.

B. VAE for Feature Compression

Variational Autoencoders (VAEs) have emerged as a versatile tool for encoding high-dimensional data into compact latent representations, playing a critical role in reducing computational complexity while preserving task-relevant information [?]. This capability is particularly beneficial in autonomous driving, where real-time decision-making necessitates efficient processing of sensory data.

In the context of autonomous navigation, VAEs have been used to compress semantic maps and depth data, transforming these high-dimensional inputs into lower-dimensional latent spaces suitable for downstream tasks. For instance, studies have shown that VAEs can effectively encode semantic segmentation outputs into meaningful latent representations, reducing storage and processing demands while retaining critical environmental details. Moreover, VAEs facilitate smoother integration with reinforcement learning algorithms by simplifying the state space, thereby enhancing training efficiency and stability.

Beyond feature compression, recent advancements in VAE architectures have introduced conditional and hierarchical VAEs, enabling more structured latent spaces that better capture the nuances of input data. These innovations are particularly relevant for autonomous driving, as they allow the model to encode not only static environmental features but also temporal dynamics, supporting more robust policy learning in reinforcement learning frameworks.

The robustness of VAEs has been further validated through applications in generative modeling, where compressed latent representations are used to reconstruct high-quality semantic maps and depth images. These findings demonstrate the dual utility of VAEs for both compression and generation, making them a critical component in developing scalable and efficient autonomous systems.

C. PPO for Controlling Agents

Proximal Policy Optimization (PPO) has become one of the most widely adopted reinforcement learning algorithms, particularly for tasks involving continuous control and dynamic environments [4]. PPO builds upon earlier policy gradient methods by introducing a clipped surrogate objective function, which ensures that policy updates remain within a bounded range. This mechanism mitigates instability issues commonly associated with traditional policy gradient methods, resulting in smoother and more reliable learning trajectories.

In the domain of autonomous driving, PPO has been extensively utilized to train agents for complex navigation tasks. Existing literature highlights its adaptability to diverse input modalities, including raw images, semantic features, and depth maps. For instance, studies have demonstrated PPO's ability to optimize driving policies using a combination of sensory inputs, enabling agents to navigate through urban environments with dynamic traffic and unpredictable obstacles.

One of the key advantages of PPO lies in its sample efficiency, which allows agents to achieve robust policies with fewer interactions with the environment. This attribute is particularly beneficial in simulation-based platforms like CARLA, where high-quality training data can be computationally expensive to generate. Additionally, PPO's compatibility with modular architectures enables seamless integration with feature extractors like U-Net and VAEs, further enhancing the agent's decision-making capabilities.

Recent advancements in PPO implementations have introduced multi-threaded and distributed learning frameworks, accelerating the convergence of policies for large-scale applications. These enhancements, combined with PPO's inherent stability, make it a suitable choice for developing autonomous driving agents capable of robust and adaptive navigation in dynamic environments. Furthermore, the integration of PPO with reward-shaping techniques has been shown to improve policy alignment with specific driving objectives, such as maintaining lane centering, regulating speed, and avoiding collisions, making it an indispensable tool for this project.

III. Methods

A. Data Collection

The dataset for this project was generated using the CARLA simulator, a high-fidelity open-source platform for autonomous driving research. The simulation environment included diverse urban layouts with variations in road topology, traffic density, weather conditions, and lighting. An autonomous vehicle was programmed to follow predefined safe trajectories in these environments, mimicking real-world driving scenarios. To increase the dataset's robustness, noise was deliberately added to the vehicle's location and orientation parameters. This perturbation simulated errors commonly encountered in real-world navigation, such as inaccuracies in GPS localization or minor deviations in steering.

Each trajectory was sampled at regular intervals to capture synchronized RGB images, semantic segmentation masks, depth maps, and navigational variables such as velocity, throttle, and steering angle. The dataset was then partitioned into training (80%), validation (10%), and test (10%) subsets, ensuring a diverse distribution of environmental conditions across all splits. By simulating both structured urban roads and unstructured off-road scenarios, the dataset aimed to prepare the agent for handling a wide range of driving conditions.

B. Semantic Segmentation

To extract structured scene understanding from raw RGB inputs, a U-Net architecture was employed for semantic segmentation. The model was trained to classify each pixel in the image into one of 13 predefined classes, including roads, sidewalks, vehicles, pedestrians, vegetation, and road lines. Data augmentation techniques were applied during training to improve generalization, including random horizontal flipping, rotation, and color jittering. These augmentations simulated variations in camera angles and lighting, reducing the risk of overfitting.

The trained U-Net achieved a mean Intersection over Union (IoU) score of 0.7 across the target classes, reflecting its ability to accurately delineate critical elements in diverse scenarios. To optimize the model for real-time inference, the number of channels in the intermediate layers was reduced, decreasing memory usage and improving computational efficiency. This adjustment was particularly beneficial in the CARLA simulation environment, where maintaining low latency is critical for interactive applications.

The outputs of the U-Net model were used as input features for downstream tasks, providing the agent with a pixel-level understanding of the environment. These features were essential for distinguishing navigable areas from obstacles, supporting safe and efficient decision-making during reinforcement learning.

C. Feature Fusion and Compression

Semantic segmentation masks, while informative, are high-dimensional and computationally expensive to process in reinforcement learning pipelines. To address this challenge, a Variational Autoencoder (VAE) was utilized to compress these masks into compact latent representations. The VAE was trained on the semantic maps produced by the U-Net, encoding them into a latent vector of size 50. This dimensionality was chosen as a balance between preserving critical semantic information and reducing computational complexity.

The VAE's encoder-decoder architecture ensured that the compressed representations retained sufficient detail for downstream tasks while achieving a significant reduction in data size. Reconstruction loss was minimized during training to maintain the fidelity of the reconstructed segmentation masks, ensuring that key features such as lane boundaries and obstacles remained distinguishable.

In addition to semantic maps, depth data was also compressed using the VAE. While the primary focus of the project was on RGB-based methods, the inclusion of depth information provided an additional layer of spatial understanding. This fusion of compressed semantic and depth features created a unified state representation that enhanced the agent's decision-making process. By reducing the dimensionality of these inputs, the VAE enabled faster processing within the reinforcement learning framework.

D. Reinforcement Learning

The core of the project involved training an autonomous driving agent using the Proximal Policy Optimization (PPO) algorithm. PPO was chosen for its stability and sample efficiency, which are critical for training in simulation environments. The agent's state vector was constructed by concatenating the compressed semantic and depth features from the VAE with navigational observations such as velocity and steering angle. This multimodal input provided a comprehensive view of the environment, integrating visual and dynamic information.

The agent interacted with the CARLA environment to learn an optimal driving policy by maximizing a custom reward function. The reward function was designed to encourage behaviors aligned with safe and efficient driving, including:

1) **Lane Centering**: Positive rewards were assigned for maintaining proximity to the center of the lane, discouraging deviations.
2) **Speed Regulation**: Rewards were scaled based on adherence to a target speed range, penalizing excessive speeding or overly cautious driving.
3) **Collision Avoidance**: Penalties were applied for collisions with static or dynamic obstacles, encouraging safe navigation.

Episodes were terminated under specific conditions, including:

- Significant deviation from the lane center beyond a defined threshold.
- Exceeding predefined speed limits, reflecting unsafe driving behaviors.
- Collisions with other vehicles, pedestrians, or environmental obstacles.

These termination criteria ensured that the agent focused on learning safe and efficient driving behaviors while penalizing reckless or unsafe actions.

The PPO algorithm's clipped objective function played a crucial role in maintaining stability during training. By constraining policy updates within a predefined range, PPO prevented drastic changes that could destabilize learning. Additionally, a multi-threaded implementation of PPO was used to accelerate training, enabling the agent to interact with multiple instances of the CARLA environment in parallel.

IV. GITHUB LINK

Source code

V. EXPERIMENTS AND EVALUATION

To rigorously evaluate the performance and robustness of the proposed autonomous driving system, extensive experiments were conducted in the CARLA simulator. The experiments were designed to address key aspects of the system, including data diversity, feature utilization, and decision-making efficiency. Each experimental stage provided insights into the model's capabilities and limitations, guiding refinements to the architecture and training process.

A. Dataset Augmentation and Diversity

To ensure the robustness of the dataset, noise was systematically introduced to the vehicle's location and rotation parameters. This augmentation simulated real-world conditions, such as navigation errors, sensor inaccuracies, and unpredictable maneuvers, creating scenarios where the vehicle deviated from ideal driving trajectories. Such perturbations were critical for testing the model's resilience to in-domain variations and its ability to generalize to unseen situations.

The collected data encompassed a comprehensive set of features, including synchronized RGB images, semantic segmentation masks, depth maps, and navigation variables such as throttle, velocity, and steering angle. By capturing these modalities across diverse environmental conditions—including varying weather, lighting, and traffic scenarios—the dataset was tailored to reflect the dynamic nature of real-world driving environments. These measures aimed to improve the model's generalization and ensure its applicability to novel, complex scenarios.

B. Depth Estimation Experiments

The MiDaS depth estimation model, known for its scale-invariant monocular depth estimation capabilities [3], was initially trained on the collected dataset. It achieved a promising mean absolute error (MAE) of 0.06323, demonstrating its ability to predict depth with high precision. However, the model's inference time proved to be a significant bottleneck, particularly when integrated into the reinforcement learning pipeline. The high computational cost of depth estimation limited the agent's ability to interact with the environment in real time, hindering the efficiency of policy optimization.

As a result, the reliance on MiDaS was discontinued, and the experiments shifted to focus on semantic segmentation features, which offered a more computationally efficient alternative. While depth estimation was excluded from real-time processing, the option to append ground-truth depth maps directly from sensors remained available. This hybrid approach provided a valuable fallback for scenarios requiring enhanced spatial awareness, such as obstacle avoidance and tight-space navigation.

C. Agent Training and Evaluation Metrics

The autonomous driving agent was trained using the Proximal Policy Optimization (PPO) algorithm, leveraging a state vector formed by combining compressed semantic segmentation features with navigation variables. To assess the agent's performance, a set of evaluation metrics was established:

- **Navigation Accuracy**: Measured by the agent's ability to maintain lane centering and follow predefined trajectories across varied driving conditions.
- **Collision Rate**: Calculated as the frequency of collisions with static obstacles, dynamic agents, or environmental boundaries.
- **Smoothness of Control**: Evaluated based on the standard deviation of steering, throttle, and braking inputs, with smoother control indicative of better policy optimization.
- **Inference Time**: Measured to ensure that the system met real-time requirements, particularly when processing high-dimensional inputs.

The experiments were conducted across multiple test scenarios, including urban environments, highway settings, and unstructured terrains. These scenarios featured varying levels of complexity, such as dynamic traffic, pedestrian crossings, and construction zones, to test the agent's adaptability.

D. Ablation Studies

To evaluate the contributions of individual components to the overall system performance, ablation studies were conducted. These studies systematically removed or modified key elements, such as:

- **Semantic Features**: Experiments were performed using raw RGB inputs instead of semantic segmentation maps to quantify the impact of structured scene understanding.
- **Depth Information**: The influence of ground-truth depth maps was assessed by integrating them into the state vector and comparing the results with RGB-only and semantic-only configurations.
- **Compressed Representations**: The effect of VAE-based feature compression was analyzed by replacing compressed features with high-dimensional raw inputs, examining the trade-offs in computational efficiency and policy performance.

The results from these studies highlighted the critical role of semantic segmentation in enabling efficient decision-making, with semantic-only configurations outperforming RGB-only inputs in navigation accuracy and collision avoidance. The addition of ground-truth depth maps provided marginal improvements in spatial reasoning but incurred a significant computational cost, aligning with the decision to prioritize semantic features in real-time applications.

E. Observations and Insights

Several key observations emerged from the experimental evaluations:

- **Computational Efficiency**: The reliance on semantic segmentation features and compressed representations significantly reduced inference time, ensuring the system's suitability for real-time deployment.
- **Generalization Capability**: The diverse dataset and noise augmentations improved the agent's ability to handle in-domain variations, with robust performance observed across all test scenarios.
- **Depth Information Trade-offs**: While depth data enhanced performance in specific tasks, its exclusion during real-time inference highlighted the feasibility of achieving comparable results with semantic features alone.
- **Policy Robustness**: The PPO algorithm demonstrated consistent stability and adaptability, particularly in scenarios involving dynamic traffic and complex navigation tasks.

F. Future Directions for Experimentation

The insights gained from these experiments suggest several avenues for future research. These include exploring lightweight depth estimation models that balance computational cost and accuracy, incorporating adversarial training to enhance robustness against adversarial inputs, and investigating hybrid reinforcement learning architectures to further optimize policy performance.

By systematically evaluating the proposed system under diverse conditions and configurations, these experiments provide a comprehensive understanding of its capabilities and limitations, laying the groundwork for future advancements in autonomous driving technologies.

VI. Analysis and Observations

A. Model Efficiency

One of the key improvements in our approach was replacing the depth estimation model with a more efficient RGB-based methodology. By eliminating the need for real-time depth estimation during inference, we reduced computational overhead significantly. This not only accelerated the agent's learning process but also enhanced the overall system's responsiveness. Despite this simplification, the performance of the autonomous agent remained on par with or slightly better than models that relied on depth estimation, indicating that the model can operate efficiently without compromising its decision-making capabilities.

B. Feature Representation

A critical part of our approach was the use of a Variational Autoencoder (VAE) to compress both semantic and depth map features into compact state vectors. This compression allowed the agent to efficiently process these features while maintaining a rich representation of the environment. The VAE successfully encoded the high-dimensional input data into lower-dimensional latent spaces, enabling the agent to capture essential environmental information for decision-making. This representation not only improved the PPO agent's learning stability but also reduced the complexity of the state space, contributing to faster convergence and more reliable performance during training.

C. Ground-Truth Depth

Incorporating ground-truth depth data during inference provided marginal improvements in the agent's performance, particularly in scenarios requiring precise spatial understanding, such as obstacle avoidance and path planning. However, the additional depth information did come at the cost of increased computational complexity. The depth data, while enhancing feature richness, introduced additional processing steps and storage requirements, which made the system less efficient overall. Despite these trade-offs, the slight improvement in the agent's decision-making ability, especially in dynamic environments, suggests that ground-truth depth could be a valuable addition in situations where performance gains justify the added computational load.

VII. Conclusion

This project successfully developed an autonomous driving agent in the CARLA simulator, utilizing reinforcement learning techniques to enhance decision-making and navigation. By relying primarily on RGB image data and semantic segmentation, along with compressed feature representations via a Variational Autoencoder (VAE), we were able to achieve efficient agent training and stable learning dynamics with the Proximal Policy Optimization (PPO) algorithm.

The decision to replace depth estimation with RGB-based approaches proved effective in reducing computational overhead without compromising the agent's performance. The VAE provided a powerful tool for encoding complex environmental information into compact state vectors, further improving the learning process. Additionally, while the integration of ground-truth depth data during inference offered marginal improvements in the agent's decision-making ability, it also introduced additional complexity, suggesting that its benefits might not justify the added computational cost in most scenarios.

Overall, the results highlight the feasibility of leveraging simplified visual features, such as semantic segmentation and RGB images, for autonomous driving tasks, achieving efficient performance without the need for real-time depth estimation. Future work could explore further optimizations to refine the balance between computational efficiency and the use of richer sensory data like depth, potentially incorporating more advanced feature fusion techniques to improve robustness and performance in complex driving environments.

References

[1] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *MICCAI*. U-Net GitHub

[2] Dosovitskiy, A., Ros, G., Codevilla, F., López, A., & Koltun, V. (2017). CARLA: An open urban driving simulator. *Conference on Robot Learning (CoRL)*. CARLA Website

[3] Ranftl, R., Bochkovskiy, A., & Koltun, V. (2020). MiDaS: Monocular depth estimation using scale-invariant learning. *arXiv preprint*. MiDaS GitHub

[4] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal Policy Optimization Algorithms. *arXiv preprint*. PPO Paper