Vision-Based Interactive Mapping and Navigation in Dynamic Environments using Semantic Segmentation and Reinforcement Learning

Mosroor Mofiz Arman(1921079642)
Electrical & Computer Engineering,
North South University
Dhaka, Bangladesh
mosroor.arman@northsouth.edu

Conventional mapping and navigation systems typically rely on static maps and predefined paths, which limit their effectiveness in dynamic environments. Dynamic obstacles, such as moving vehicles, pedestrians, and environmental changes, present significant challenges for autonomous systems, leading to safety concerns and operational limitations. Vision-based approaches offer a promising alternative by leveraging real-time sensory data to perceive and understand the surrounding environment. However, the complexity of dynamic scenes requires advanced techniques for scene understanding, obstacle detection, and adaptive navigation. The research addresses these challenges by combining semantic segmentation for scene understanding with reinforcement learning for adaptive decision-making, enabling autonomous systems to navigate safely and efficiently in dynamic environments.

I. Introduction

The proliferation of autonomous systems in various domains, such as robotics, transportation, and surveillance, has underscored the need for robust mapping and navigation solutions capable of operating effectively in dynamic environments. Traditional approaches often struggle to adapt to the challenges posed by dynamic obstacles, changing lighting conditions, and unpredictable scene configurations. To address these challenges, the research proposes the development and implementation of a vision-based interactive mapping and navigation system utilizing semantic segmentation and reinforcement learning techniques. By integrating computer vision with machine learning, the proposed system aims to enhance adaptability, robustness, and real-time decision-making capabilities in dynamic environments.

II. Objectives

The primary objectives of the research are as follows:

- Develop a robust semantic segmentation framework for scene understanding and dynamic obstacle detection in real-time.
- Integrate reinforcement learning algorithms to enable adaptive navigation and decision-making based on semantic scene understanding.
- Implement a real-time interactive mapping system capable of updating the environment representation dynamically.
- Evaluate the performance of the proposed system through extensive simulations and real-world experiments in various dynamic environments.

III. Methodology

The proposed research will follow a systematic methodology comprising the following key steps:

A. Data Collection and Preprocessing

- Gather diverse datasets containing images and corresponding ground truth annotations for semantic segmentation.
- Preprocess the collected data to ensure consistency and compatibility with the proposed system.

B. Semantic Segmentation Framework

- Design and train a deep neural network architecture for semantic segmentation, leveraging state-of-the-art techniques such as convolutional neural networks (CNNs) and attention mechanisms.
- Fine-tune the segmentation model to accurately detect and classify dynamic obstacles in the environment while maintaining real-time performance.

C. Reinforcement Learning Integration

- Develop reinforcement learning algorithms, including Deep Q-Networks (DQN) or Proximal Policy Optimization (PPO), to facilitate adaptive navigation.
- Train the reinforcement learning agent to navigate through dynamic environments while maximizing predefined rewards, considering factors such as obstacle avoidance, efficiency, and safety.

D. Real-Time Interactive Mapping System

- Implement an interactive mapping system capable of integrating real-time sensor inputs and updating the environment representation dynamically.
- Integrate the semantic segmentation and reinforcement learning components into the mapping system to enable seamless navigation while continuously updating the map.

E. Performance Evaluation

- Conduct comprehensive evaluations of the proposed system using both simulated environments and real-world scenarios.
- Assess the system's performance in terms of mapping accuracy, navigation efficiency, adaptability to dynamic changes, and robustness to environmental uncertainties.

IV. Expected Contributions

The research is expected to make several significant contributions to the field of autonomous systems and robotics:

- A novel vision-based mapping and navigation framework tailored for dynamic environments, leveraging semantic segmentation and reinforcement learning techniques.
- Insights into the efficacy of combining computer vision and machine learning for enhancing adaptability, robustness, and real-time decision-making in autonomous systems.
- Practical guidelines for implementing interactive mapping and navigation systems in real-world applications, addressing challenges related to dynamic environments and dynamic obstacle avoidance.

V. Conclusion

In conclusion, the research proposal outlines a comprehensive investigation into the development of a vision-based interactive mapping and navigation system for dynamic environments. By leveraging semantic segmentation and reinforcement learning techniques, the proposed system aims to overcome the challenges associated with dynamic scene understanding and adaptive decision-making. The outcomes of the research are expected to significantly advance the state-of-the-art in autonomous systems and pave the way for practical applications in various domains, including robotics, transportation, and surveillance.

References

[1] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 4, pp. 834-848, 2018.

[2] V. Mnih et al., "Human-level control through deep reinforcement learning," in Nature, vol. 518, no. 7540, pp. 529-533, 2015.

[3] D. Scherer, A. Müller, S. Behnke, and M. Souiai, "Evaluation of pooling operations in convolutional architectures for object recognition," in International Conference on Artificial Neural Networks, pp. 92-101, Springer, Berlin, Heidelberg, 2010.