

ABSTRACT

Navigation with mobile robots is currently widely used to assist human tasks. However, navigation with mobile robots generally has several challenges such as performing navigation tasks in hard-to-reach environments and being performed by avoiding obstacles. Conventional methods generally rely only on obstacle maps and cannot learn autonomously. Mapless navigation methods based on reinforcement learning (RL) are widely used to learn mobile robot navigation systems. However, the RL method requires a suitable reinforcement learning reward function setting technique, and a simple reinforcement learning reward function may not be easy to determine for the robot to perform navigation tasks.

One of the learning from demonstration (LfD) methods that uses concepts based on demonstration data is Adversarial Inverse Reinforcement Learning (AIRL). However, the Standard AIRL method has a drawback in terms of generalization of the reward to navigation tasks with mobile robots. Modified Adversarial Inverse Reinforcement Learning (Modified-AIRL) using an additional reward function named `goal_reward` used for mobile robot navigation tasks is proposed. Modified AIRL has the advantage of being robust to environmental dynamics in the presence of `goal_reward` so that the mobile robot can perform the navigation task efficiently. Since it is robust to dynamic changes and good at generalizing the reward function, it performs well when searching for the optimal policy. This method is intended for the navigation task of a mobile robot to reach a target point with avoiding obstacles task.

The results were evaluated by comparing the proposed method with variations of the regular AIRL method using the evaluation metrics of success rate, trajectory length and travel time. It is found that the proposed method, Modified-AIRL, is sometimes less superior to other methods in terms of travel time values in some scenarios, trajectory length in some scenarios and the level of difference with the demonstration data. However, in specific scenarios, it can excel in testing the length of the path traveled and the length of time traveled in a static environment. In addition, the overall success rate performance when using the Modified-AIRL (PPO-AIRL+SAC) method can achieve a success rate of 100% slightly faster than other method variations and achieve the highest success rate in the moving obstacle scenario with a value of 49 out of 50 tests. This makes the proposed Modified-AIRL (PPO-AIRL+SAC) method possible to use for navigation tasks to successfully reach a position.

Keywords :Navigation, Mobile Robot, Learning from Demonstration, Adversarial Inverse Reinforcement Learning

INTISARI

Navigasi dengan robot bergerak saat ini banyak digunakan untuk membantu tugas manusia. Namun, navigasi dengan robot bergerak umumnya memiliki beberapa tantangan seperti melakukan tugas navigasi di lingkungan yang sulit terjangkau dan dilakukan dengan menghindari rintangan. Metode konvensional umumnya hanya mengandalkan peta rintangan dan tidak dapat belajar secara otonom. Metode navigasi tanpa peta berbasis *reinforcement learning* (RL) banyak digunakan untuk mempelajari sistem navigasi robot bergerak. Namun, metode RL memerlukan teknik pengaturan fungsi *reward* yang sesuai, dan fungsi *reward* yang sederhana mungkin tidak mudah ditentukan agar robot dapat melakukan tugas navigasi.

Salah satu metode *learning from demonstration* (LfD) yang menggunakan konsep berdasarkan data demonstrasi yaitu *Adversarial Inverse Reinforcement Learning* (AIRL). Namun metode AIRL standar memiliki kelemahan dalam hal generalisasi *reward* pada tugas navigasi dengan robot bergerak. *Modified Adversarial Inverse Reinforcement Learning* (Modified-AIRL) dengan menggunakan fungsi *reward* tambahan yang dinamakan *goal_reward* digunakan untuk tugas navigasi robot bergerak diusulkan. Modified AIRL memiliki keunggulan yang kuat terhadap dinamika lingkungan dengan adanya *goal_reward* sehingga robot bergerak dapat melakukan tugas navigasi dengan efisien. Karena kuat terhadap perubahan dinamis dan baik dalam hal generalisasi fungsi *reward*, maka kinerja akan baik saat mencari *policy* yang optimal. Metode ini ditujukan untuk tugas navigasi pada robot bergerak mencapai titik target dengan menghindari rintangan.

Hasil dievaluasi dengan membandingkan metode yang diusulkan dengan variasi metode AIRL biasa dengan menggunakan metrik evaluasi tingkat keberhasilan, panjang lintasan dan waktu tempuh. Hasil pengujian didapatkan bahwa metode yang diusulkan yaitu Modified-AIRL terkadang kurang unggul dari metode lainnya dalam hal nilai waktu tempuh pada beberapa skenario, panjang lintasan pada beberapa skenario dan tingkat perbedaan dengan data demonstrasi. Namun, dalam skenario yang spesifik dapat unggul pada pengujian panjang lintasan yang ditempuh dan lama waktu yang ditempuh di lingkungan statis. Selain itu, kinerja tingkat keberhasilan saat *policy optimization* dengan metode Modified-AIRL (PPO-AIRL+SAC) secara keseluruhan dapat mencapai tingkat keberhasilan 100% sedikit cepat dari variasi metode lainnya dan mencapai tingkat keberhasilan tertinggi pada skenario rintangan yang bergerak dengan nilai 49 dari 50 kali pengujian. Hal ini menjadikan metode Modified-AIRL (PPO-AIRL+SAC) yang diusulkan dapat memungkinkan digunakan untuk tugas navigasi untuk dapat mencapai suatu posisi dengan sukses.

Kata kunci – Navigasi, Robot Bergerak, Learning from Demonstration, Adversarial Inverse Reinforcement Learning