Intisari

Robot manipulator adalah subjek yang sangat signifikan dalam industri, karena aplikasinya yang luas dan kemampuannya untuk disesuaikan dengan kebutuhan melalui modifikasi pada *end-effector*. Salah satu tantangan utama dalam penelitian pergerakan *robot manipulator* adalah kinematika yang terbagi menjadi dua jenis: kinematika maju dan kinematika balik. Kinematika balik, terutama pada *robot manipulator* M-DoF studi kasus *arm manipulator*, menjadi fokus utama penelitian ini. Hambatan utama dalam menyelesaikan kinematika balik termasuk keberadaan singularitas, solusi yang tidak unik atau solusi lebih dari satu, dan kompleksitas yang meningkat dengan jumlah derajad kebebasan.

Secara umum, solusi kinematika balik dibagi menjadi dua jenis: solusi analitik seperti *closed-form solution* dan solusi numeris seperti *Pseudo-inverse*, *Neural Network*, dan ANFIS. Penelitian ini akan menggunakan algoritma *Deep Reinforcement Learning* (DRL), khususnya *Deep Deterministic Policy Gradient* (DDPG), untuk menyelesaikan kinematika balik pada *arm* robot manipulator. Kelebihan DDPG, terutama dalam penggunaan kebijakan *deterministic*, membantu dalam stabilitas pembelajaran dan menangani tindakan dalam domain *continuous* dan berdimensi tinggi. Penelitian ini mengembangkan algoritma DDPG (*improved*-DDPG) dengan parameterisasi *noise*. DDPG menggunakan parameter *noise Ornstein-Uhlenbeck* (OU), sedangkan pengembangan DDPG (*improved*-DDPG) dimodifikasi menggunakan parameter *noise* berupa *Gaussian noise* dan *OU-Gaussian noise*. Dengan pengembangan parameterisasi noise tersebut, diharapkan dapat meningkatkan kinerja eksplorasi algoritma DDPG. Hasil pengembangan algoritma DDPG (*improved*-DDPG) akan digunakan pada M-DoF *arm manipulator*.

Hasil implementasi algoritma DDPG dan pengembangan algoritma DDPG (*improved*-DDPG) menggunakan parameterisasi *Gaussian noise* dan *OU-Gaussian noise* berhasil menyelesaikan kinematika balik pada M-DoF PUMA 560. Hasil koordinat *joint* pada *arm manipulator* 6 DoF PUMA 560 dalam 1.000 episode yaitu $\boldsymbol{q}_{z-DDPG}, \boldsymbol{q}_{z-G}, \boldsymbol{q}_{z-OUG}$ yang sama yaitu [0,0000 0,7854 3,1416 0,0000 0,7854 0,0000]. Nilai tersebut mendekati nilai posisi original yaitu pada nilai dari solusi *closed-form* yaitu $\boldsymbol{q}_z = [0{,}0000\ 0{,}0000\ 0{,}0000\ 0{,}0000\ 0{,}0000\ 0{,}0000]]$. Pengembangan algoritma DDPG (*improved*-DDPG) ini terbukti memiliki kemampuan eksplorasi yang lebih baik dari algoritma DDPG original. Hal ini diindikasikan dengan jumlah *action* yang dihasilkan oleh *improved*-DDPG pada 100 episode adalah 2.401 data sedangkan DDPG original adalah 241. Nilai koordinat *joint* yang dihasilkan oleh improved-DDPG dengan *Gaussian noise* saat 100 episode menghasilkan nilai paling dekat dengan *closed-form*.

Kata kunci—*Robot manipulator* M-DoF, 6 DoF PUMA 560, *Deep Reinforcement Learning* (DRL), *Deep Deterministic Policy Gradient* (DDPG), Parameterisasi *Noise*.

Abstract

The *robot manipulator* is a highly significant subject in the industry due to its broad applications and adaptability through modifications to the end-effector. One of the primary challenges in research on the movement of robot manipulators is its kinematics, which is divided into two types: forward kinematics and inverse kinematics. Inverse kinematics, particularly in the context of M-DoF (Multi-Degree of Freedom) case study *arm manipulator*s, is the main focus of this research. The principal obstacles in solving inverse kinematics include the presence of singularities, non-unique solutions or multiple solutions, and increased complexity with the number of degrees of freedom.

Generally, solutions to inverse kinematics are categorized into two types: analytical solutions such as closed-form solutions, and numerical solutions such as Pseudo-inverse, Neural Networks, and ANFIS (Adaptive Neuro-Fuzzy Inference System). This study employs a Deep Reinforcement Learning (DRL) algorithm, specifically Deep Deterministic Policy Gradient (DDPG), to solve the inverse kinematics of the arm robot manipulator. The advantages of DDPG, particularly its use of deterministic policies, aid in learning stability and managing actions in continuous and high-dimensional domains. This research develops an improved-DDPG algorithm with noise parameterization. While DDPG utilizes Ornstein-Uhlenbeck (OU) noise parameters, the improved-DDPG incorporates Gaussian noise and OU-Gaussian noise parameters. It is anticipated that this noise parameterization enhancement will improve the exploratory performance of the DDPG algorithm. The developed improved-DDPG algorithm will be applied to an M-DoF *arm manipulator*.

The implementation results of the DDPG algorithm and the improved-DDPG algorithm using Gaussian noise and OU-Gaussian noise parameterization successfully solves the inverse kinematics of the M-DoF PUMA 560 *arm manipulator*. The joint coordinates of the 6 DoF PUMA 560 *arm manipulator* over 1.000 episodes, $\boldsymbol{q}_{z-DDPG}, \boldsymbol{q}_{z-G}, \boldsymbol{q}_{z-OUG}$ are identical: [0.0000 0.7854 3.1416 0.0000 0.7854 0.0000]. These values are close to the origin pose or the closed-form solution, $\boldsymbol{q}_z = [0.0000\ 0.0000\ 0.0000\ 0.0000\ 0.0000\ 0.0000]$.

The development of the improved-DDPG algorithm demonstrates superior exploratory capabilities compared to the original DDPG algorithm. This is evidenced by the number of actions generated by improved-DDPG over 100 episodes being 2.401, whereas the original DDPG produced 241. The joint coordinate values generated by improved-DDPG with Gaussian noise in 100 episodes are closest to the closed-form solution.

Keywords— M-DoF Robot Manipulator, 6 DoF PUMA 560, Deep Reinforcement Learning (DRL), Deep Deterministic Policy Gradient (DDPG), Noise Parameterization.