

# Control of Swing Robot using Reinforcement Learning

Ho Seung Choi, Hee Jae Park

Department of Mechanical System Design Engineering, Seoul National University of Science and Technology

## ABSTRACT

본 연구는 실제 기구에 Actor-Critic 강화 학습 알고리즘을 적용하여 반복 학습을 통한 고난이도의 제어 기술을 기계 학습으로 대체하는 것의 가능성을 검증하는 것을 목표로 한다. 신경망 모델은 한 개의 은닉층을 공유하고 액터망과 크리틱망으로 분기되는 비선형 구조의 DNN을 설계하였다. 이를 위하여 비교적 간단한 동역학 모델을 가지는 1-DOF의 그네 타기 로봇과 시뮬레이션을 제작하고, 각각 신경망 모델을 학습한 뒤 이를 비교하였다. 시뮬레이션과 실제 기구 모두 적절한 정책을 학습하여 규칙적인 동작을 통해 최대 30°까지 올라가는 swing-motion을 만들어 낼 수 있음을 확인하였다.

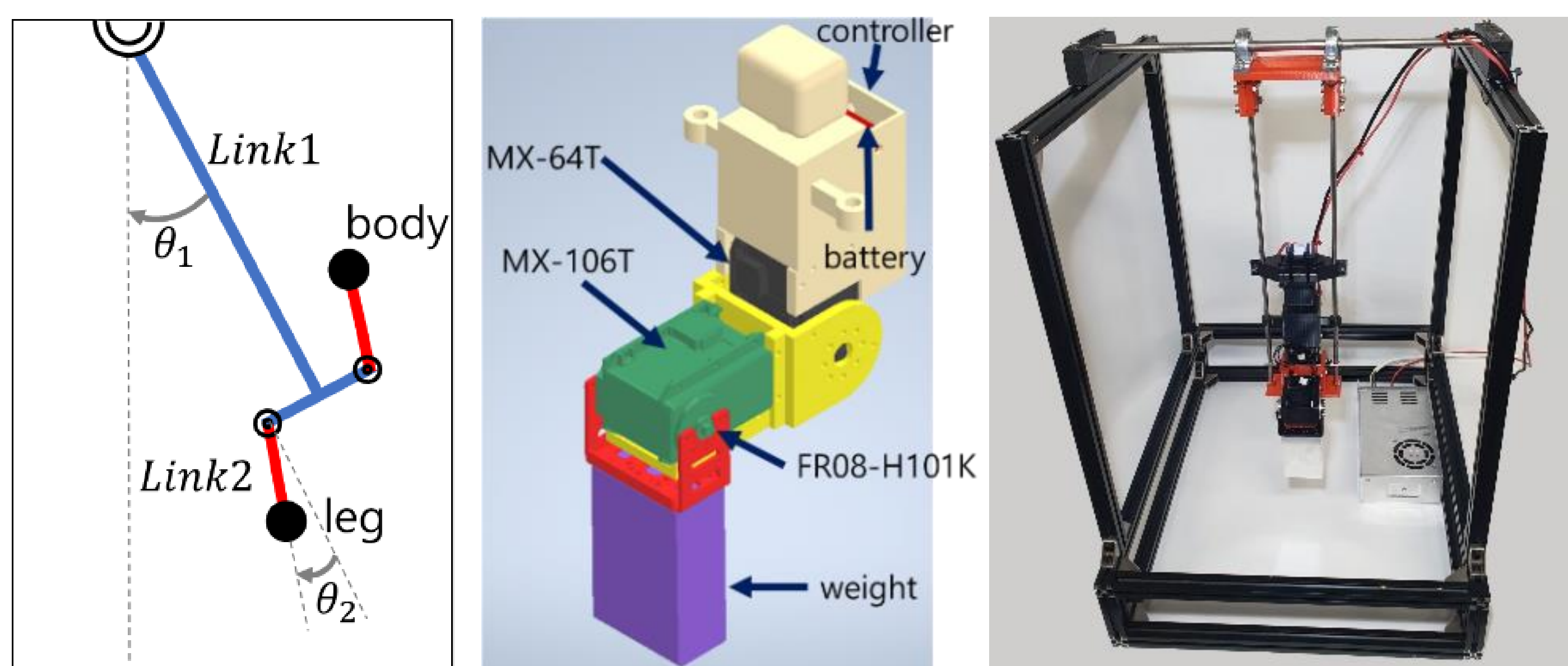
## Background

- ▶ 강화학습을 이용한 연구는 주로 기계 제어 분야에 집중되어 있다.
- ▶ 대부분의 연구가 시뮬레이션 상에서 그치고 실제 기구를 이용하여 학습하거나 성능을 검증한 사례가 적다.
- ▶ 본 연구는 실제 기구를 이용해서 학습하고 성능 평가를 통하여 연구의 효용성을 검증하는 것이 목표이다.

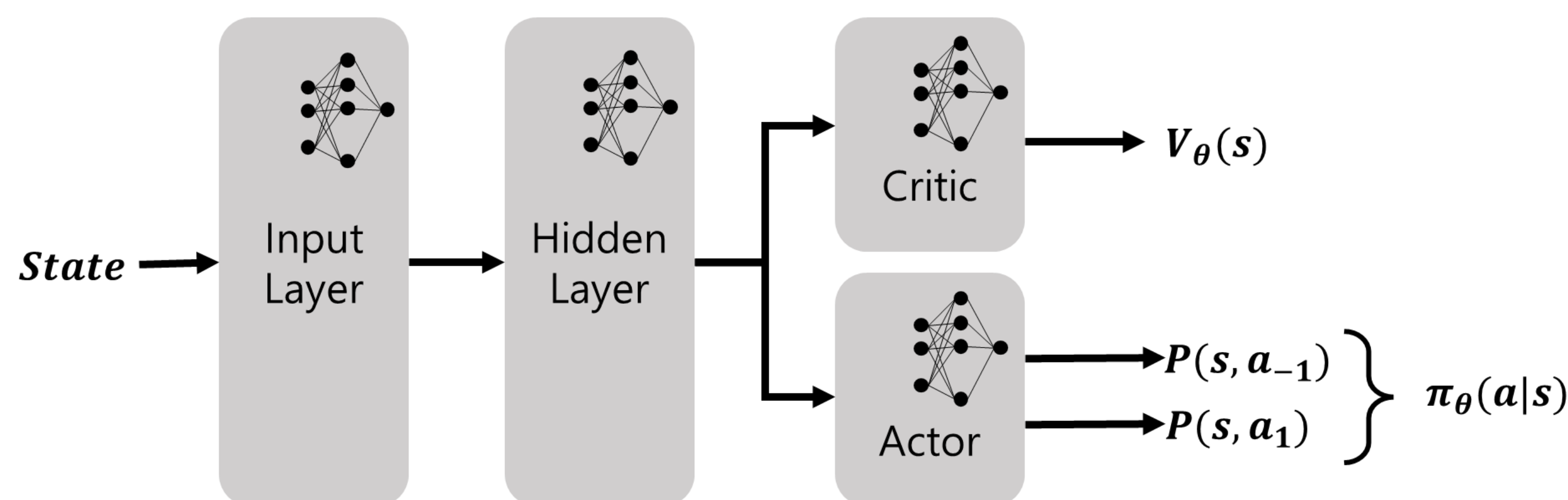
## Research Question

- 01 실제 기구를 이용한 학습의 가능성
- 02 시뮬레이션을 이용한 학습과 실제 기구를 이용해서 한 학습의 성능 차이

## Robot Design



## Neural Network Model



### Definition of Observation Space

1	2	3	4
$\theta_1$	$\theta_2$	$\dot{\theta}_1$	$\dot{\theta}_2$

### Definition of Action Space

-1	1
CCW (10°)	CW (-100°)

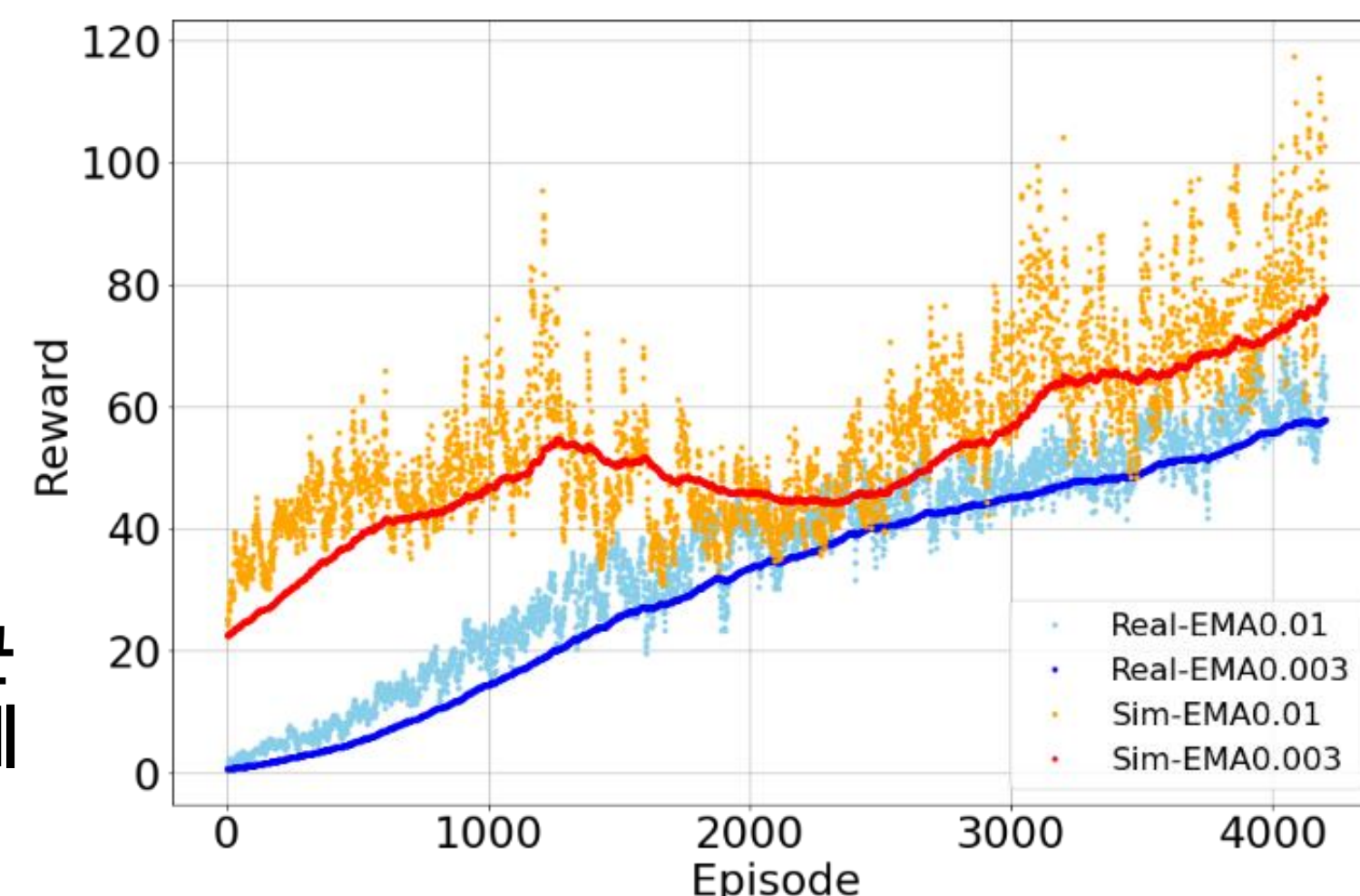
$$reward = \begin{cases} |\sin(\theta_1)|, & \text{sgn}(\dot{\theta}_1) = \text{sgn}(action) \\ -|\sin(\theta_1)|, & \text{sgn}(\dot{\theta}_1) \neq \text{sgn}(action) \end{cases}$$

## Method

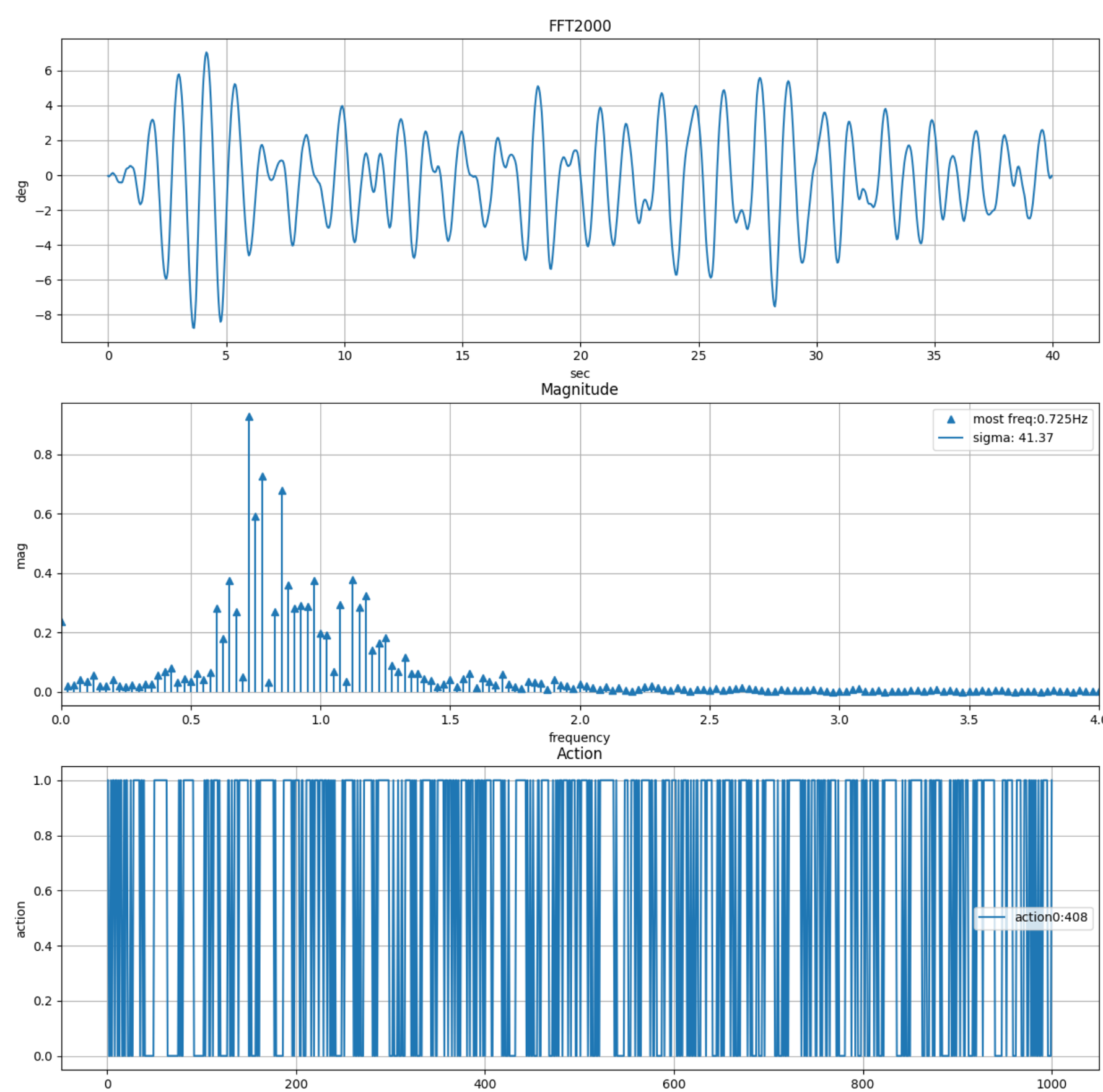
- ▶ OpenAi Gym의 Acrobot 라이브러리를 수정하여 시뮬레이션을 구현하였다.
- ▶ agent의 action은 다리의 목적 각도만 나타내는 이산 행동이지만, 실제로는 목적 각도와 현재 각도의 차이에 따라서 다른 토크를 입력하도록 하여 연속적인 토크 입력을 줄 수 있도록 하였다.
- ▶ 128개의 노드를 가지는 공통 신경망을 가지고, 총 1027개의 파라미터를 가지는 신경망을 설계하였다.
- ▶ Reward는 그네가 수평에 가까울 수록 큰 보상을 얻는다. 그네와 다리가 같은 방향으로 동작하면 양의 보상을, 반대는 음의 보상을 준다.
- ▶ 한 에피소드당 1000 step을 행동하고, step간 interval은 40ms이다. Learning Rate는  $1.2 \times 10^{-4}$ 를 사용하였다.
- ▶ 학습의 종료 조건은 한 에피소드 동안의  $\theta_1$ 으로 FFT를 수행하여 지배적인 주파수가 검출이 되고 30° 이상 올라가는 것이 연속적으로 가능할 때이다.

## Conclusion

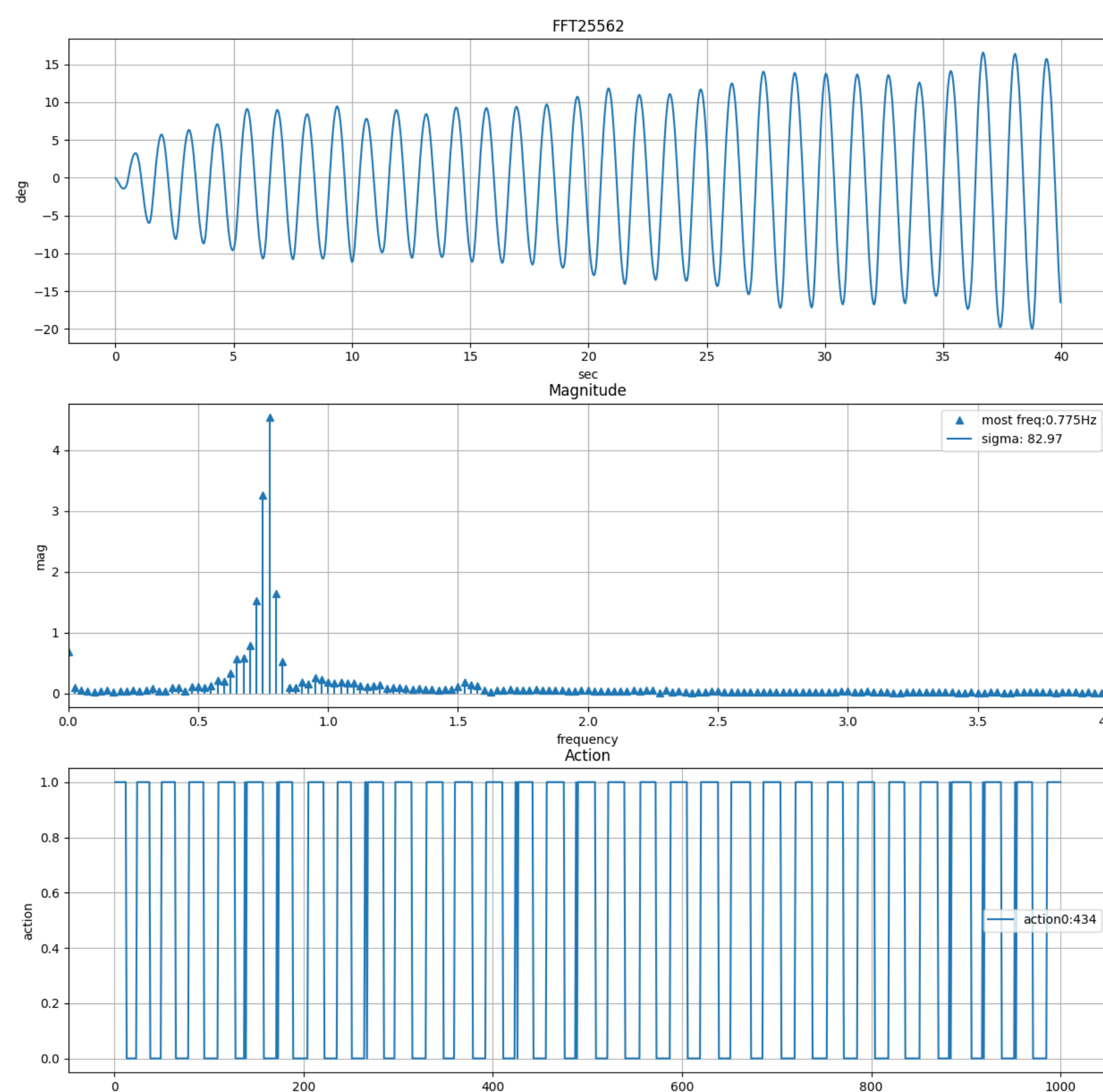
- ▶ 시뮬레이션은 약 15°, 실제 기구는 약 30° 이상 상승 가능성을 확인하였다.
- ▶ 학습은 시뮬레이션이 더 잘 되었지만 실제로 검증한 성능은 실제 기구를 이용한 것이 더 높았다.
- ▶ 지배적인 주파수는 0.775Hz로 평균 주파수 대비 80~100배 이상 강하게 검출되었다.



### Before Training at Simulation



### After Training at Simulation



### After Training at Real World

