DATA 403
Final Presentation

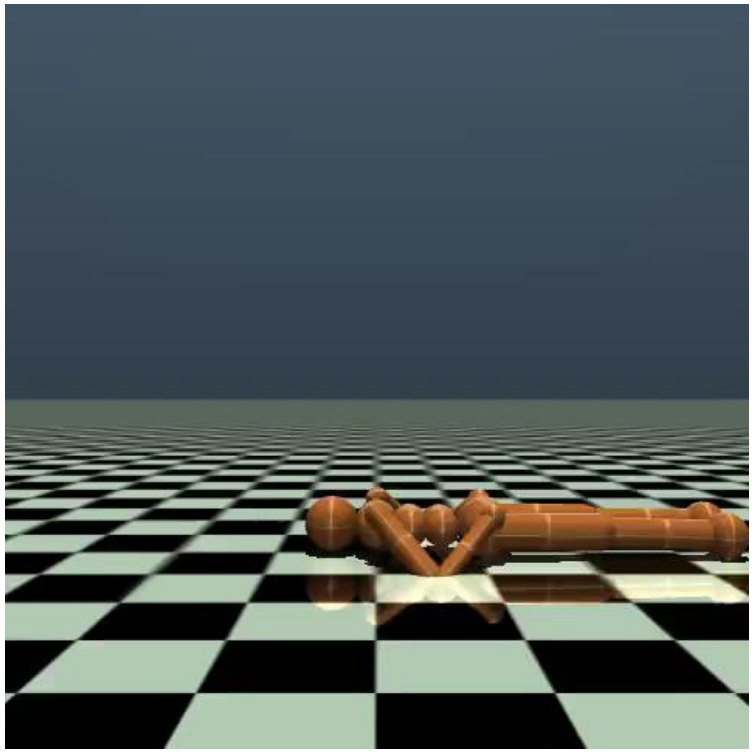
Student ID:2021320301

Name: Seong Ah Choi

Contents

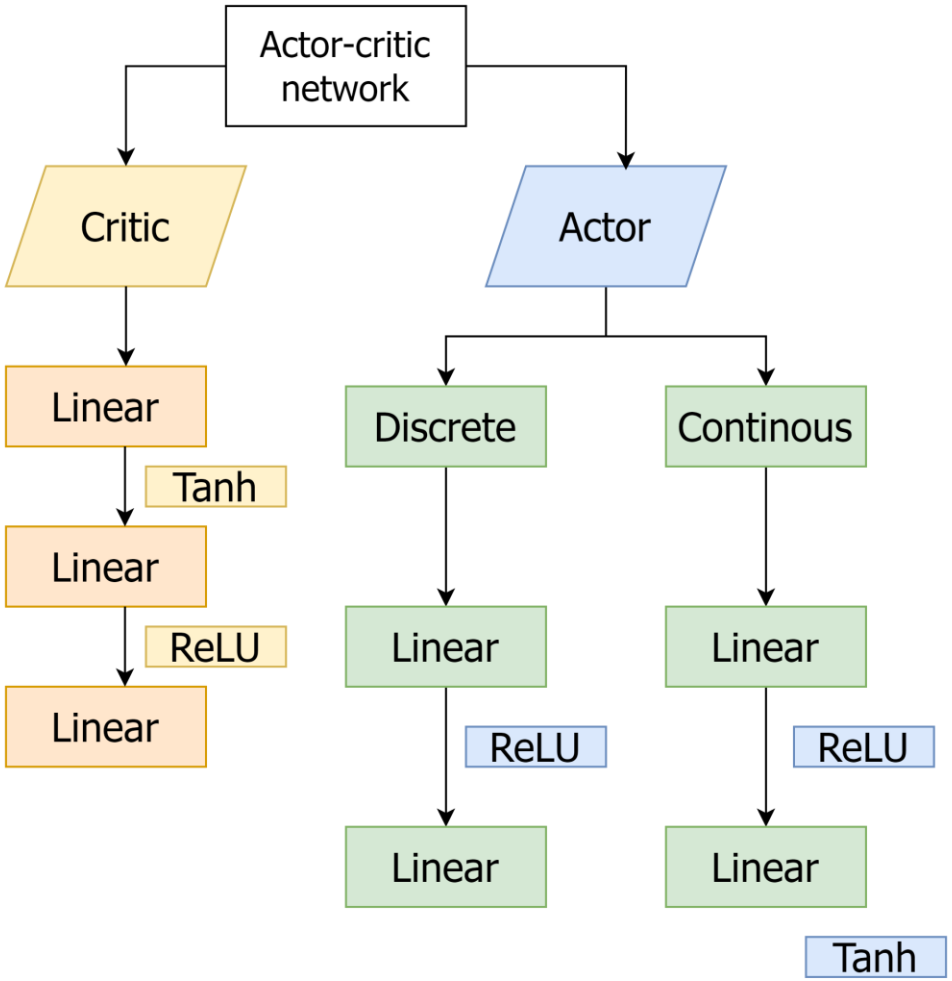
Introduction

Humanoid Standup



State	3D humanoid's state which is described by its joint's state
Action	Perform actions by selecting joint torques.
Reward	The goal is to make robot stand-up. It will be determined with the height of robot and control force of robot

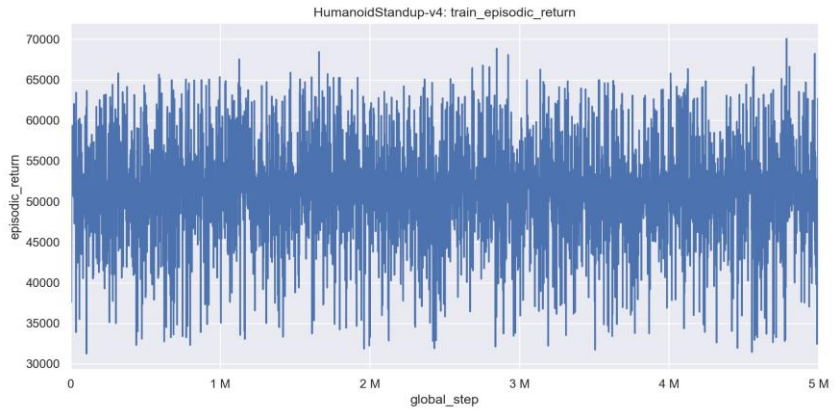
Method Actor-Critic Network

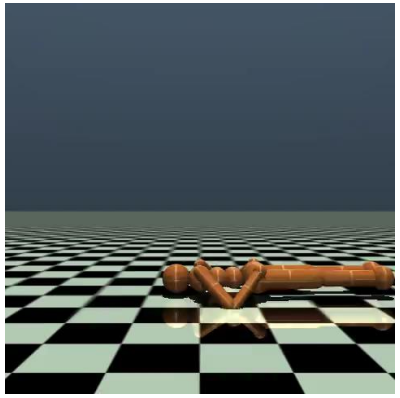


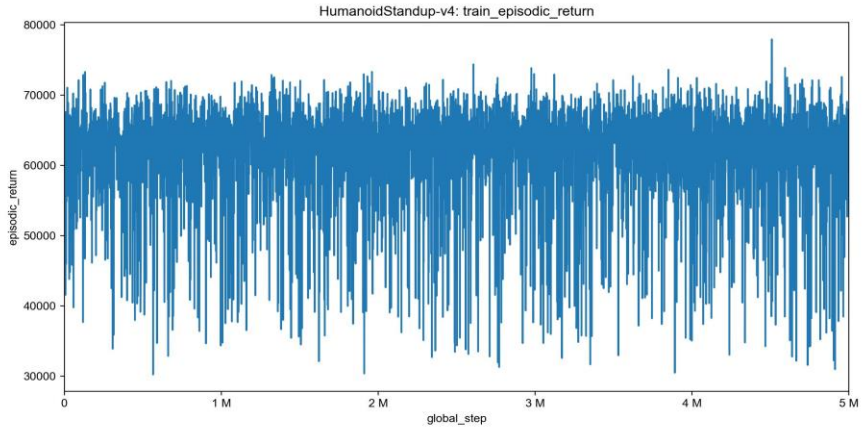
Experiments

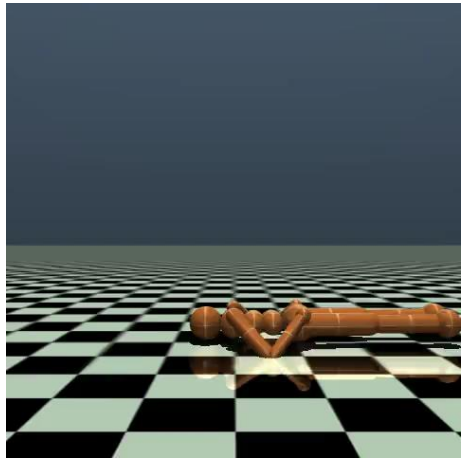
Parameter Experiment

After adding additional layers to the neural network, the average reward has increased.

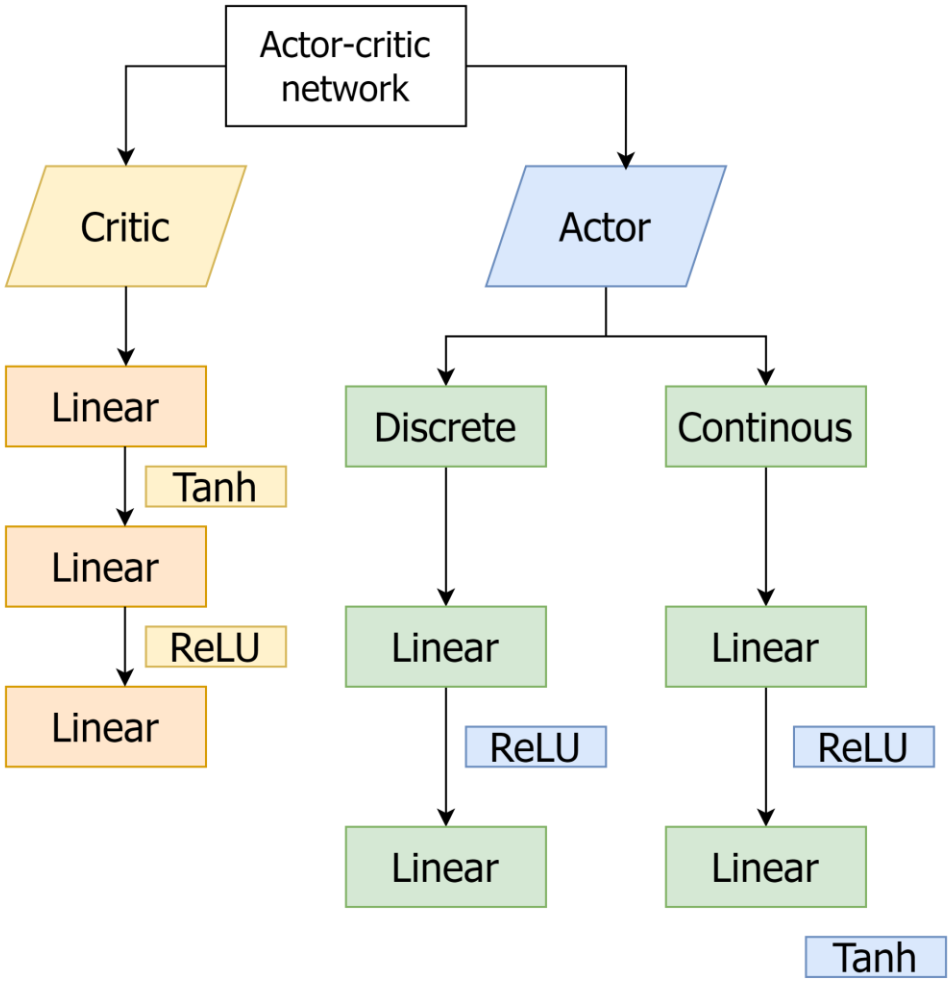






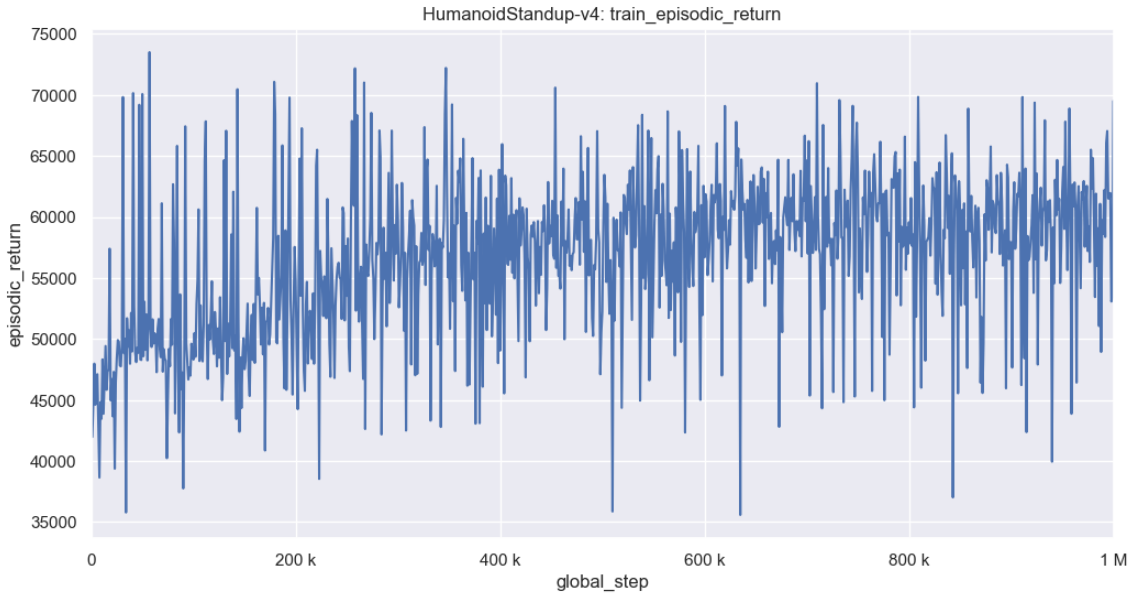


Method Actor-Critic Network



Experiments

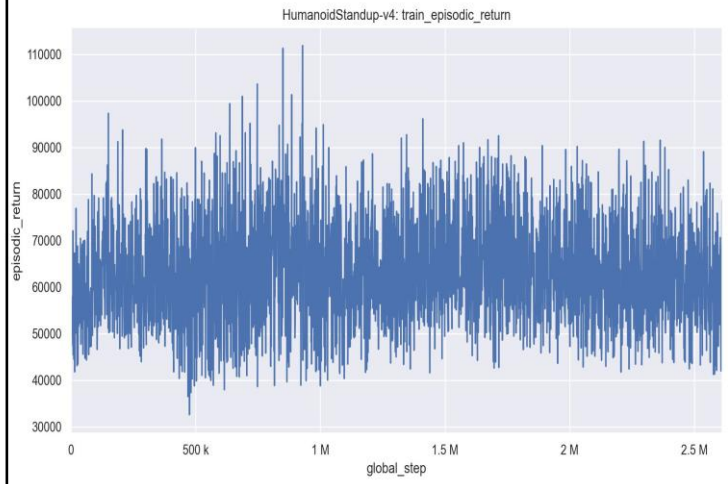
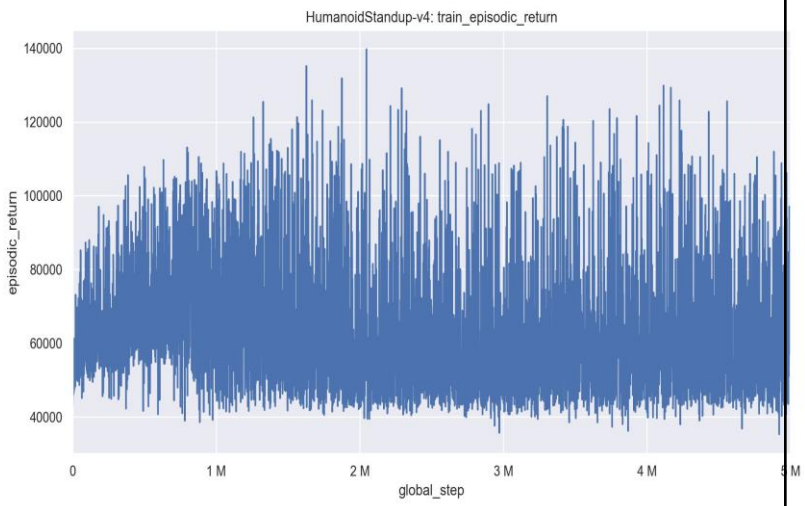
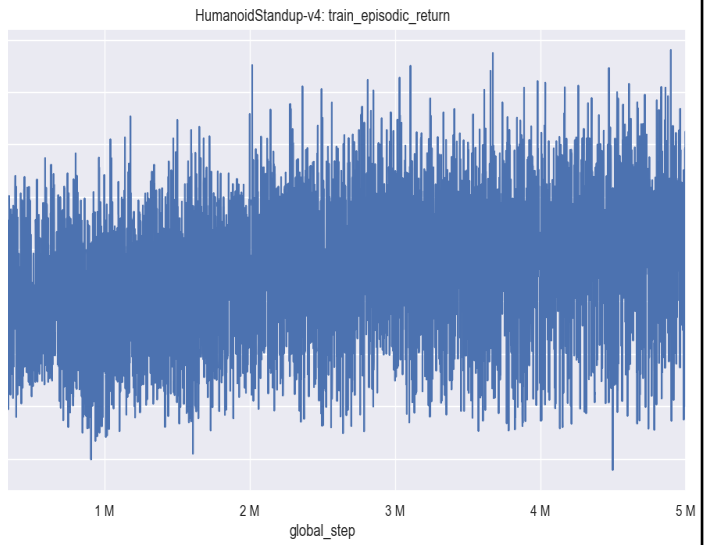
Parameter Experiment-smaller rl rate




```
ppo_config = OmegaConf.create({
    "anneal_lr": True,
    "update_epochs": 10, # The number of iteractions of ppo training
    "minibatch_size": 2048, #환경의 크기가 크대.
    "lr": 0.000015,
    "max_grad_norm": 0.5, #그레디언트 폭주 방지-최대 그레디언트 norm값
    "norm_adv": True,
    "clip_coef": 0.15,
    "ent_coef": 0.000015,
    "vf_coef": 0.5,
    "gamma": 0.96,
    "gae_lambda": 0.95,
    })
print(ppo_config)
print(ppo_config.minibatch_size)
```

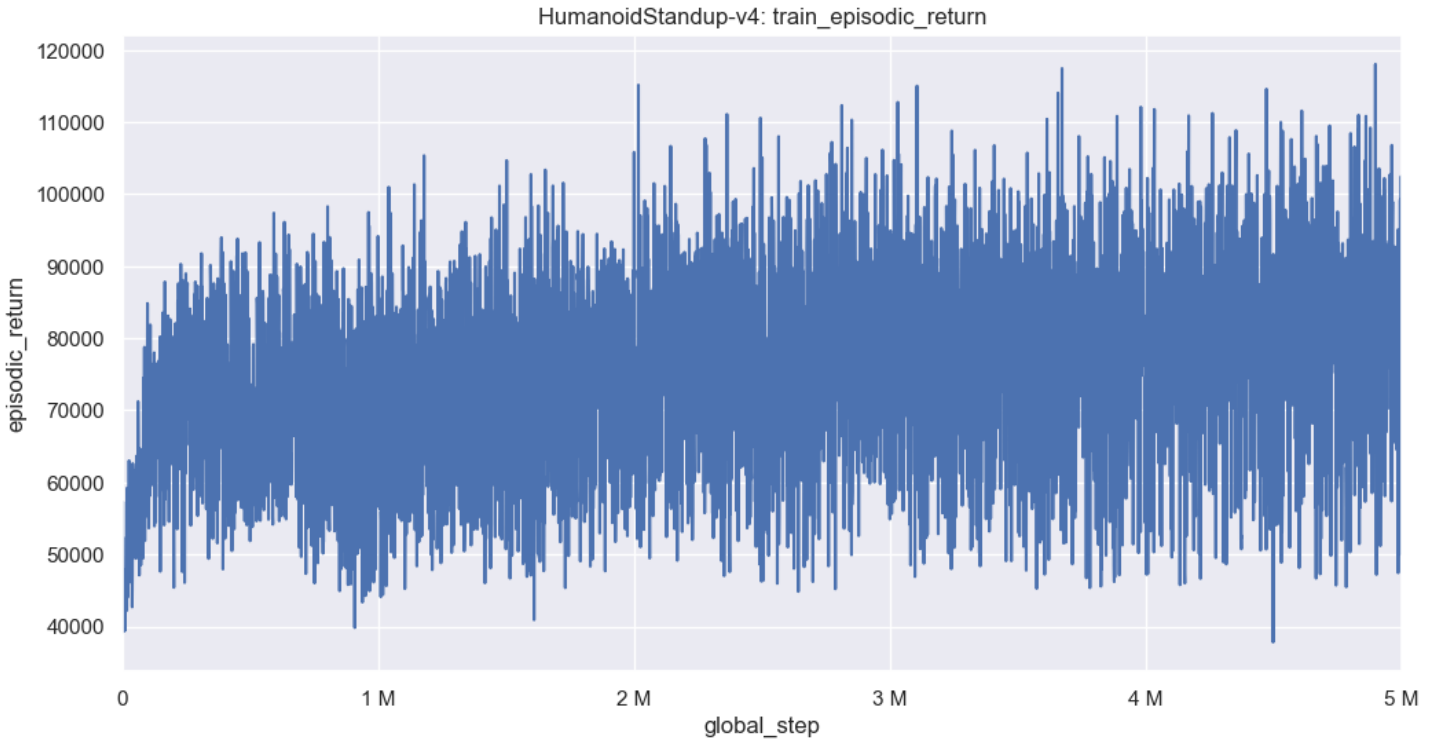
Experiments

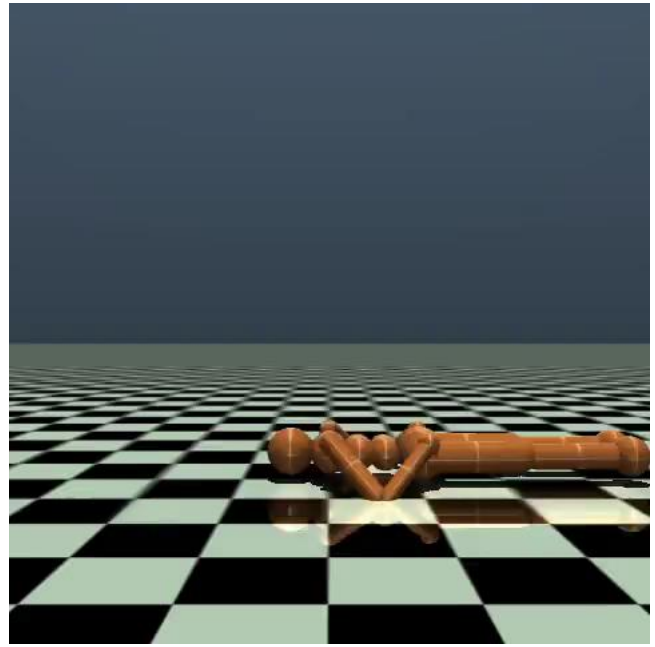
Parameter Experiment-about gamma value

Γ=0.96	Γ=0.97	Γ=0.99
		

Experiments

Performance





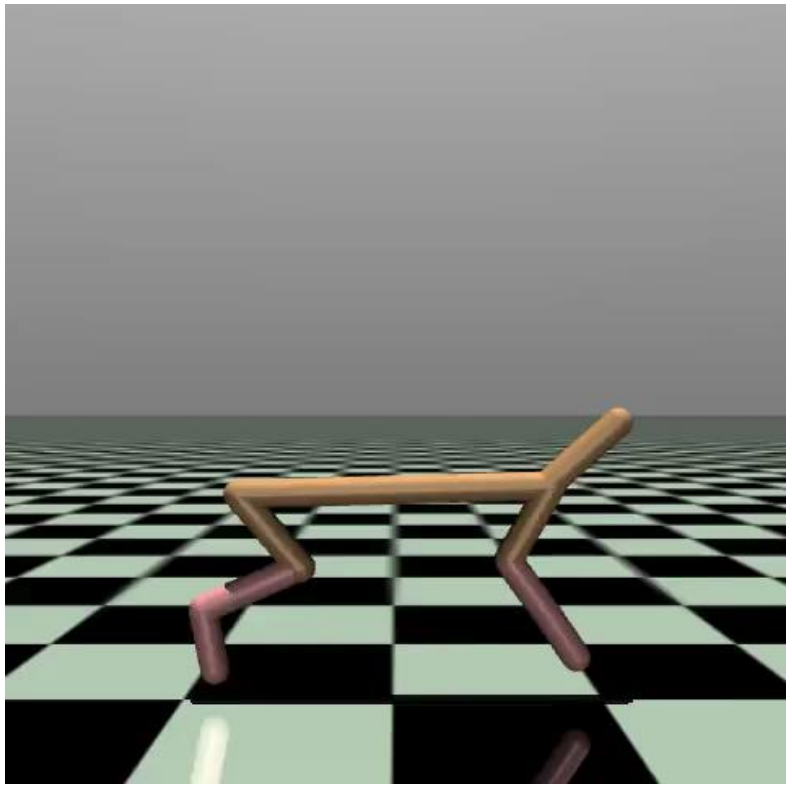
Experiments

Experimental setting

```
ppo_config = OmegaConf.create({
    "anneal_lr": True,
    "update_epochs": 15, # The number of iteractions of ppo training
    "minibatch_size": 64,
    "lr": 0.000015,
    "max_grad_norm": 0.55, #그레디언트 폭주 방지-최대 그레디언트 norm값
    "norm_adv": True,
    "clip_coef": 0.2,
    "ent_coef": 0.0001,
    "vf_coef": 0.5,
    "gamma": 0.99,
    "gae_lambda": 0.95,
    })
print(ppo_config)
print(ppo_config.minibatch_size)
```

Introduction

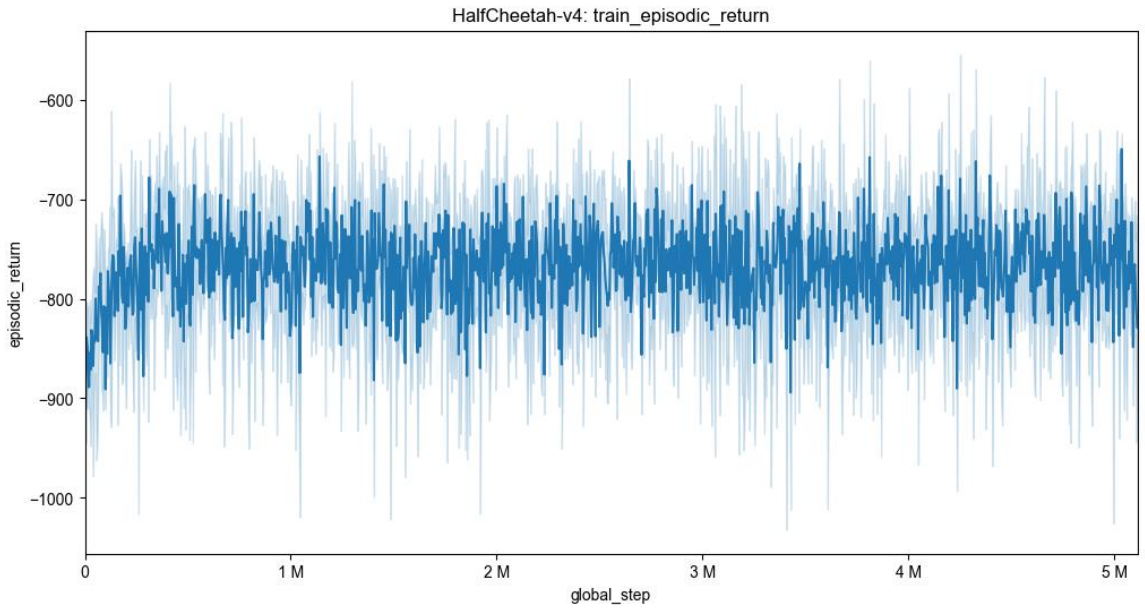
Half cheetah



State	3D robot with four legs which is described by its velocity and its joints' state
Action	Perform actions by selecting joint torques.
Reward	The goal is to make robot go ahead.

Experiments

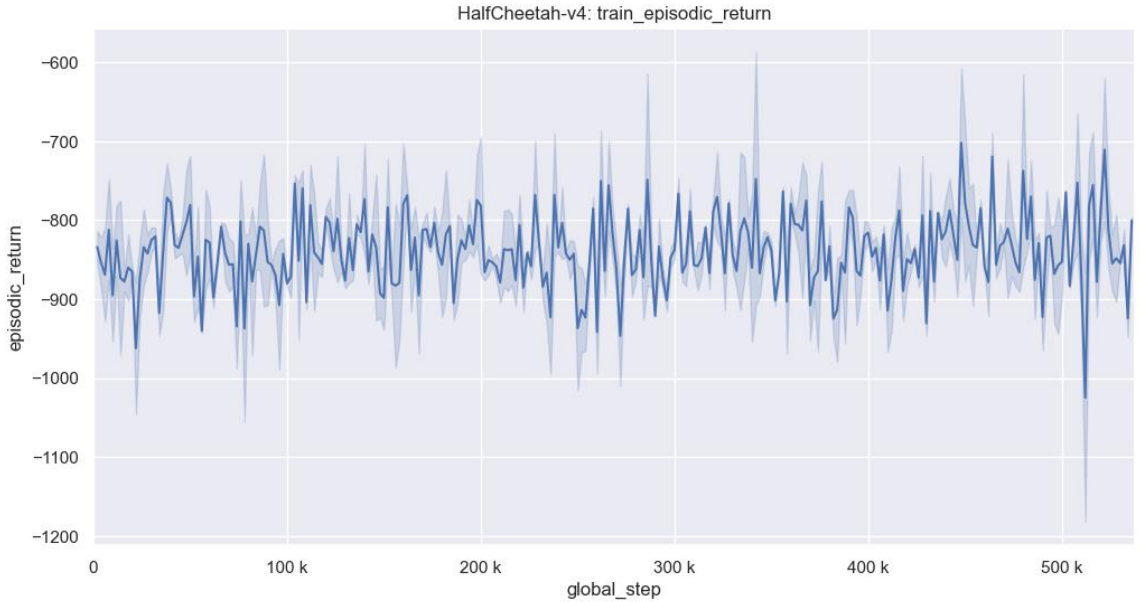
Parameter Experiment-with larger entropy coefficient




```
ppo_config = OmegaConf.create({
    "anneal_lr": True,
    "update_epochs": 10, # The number of iteractions of ppo training
    "minibatch_size": 64,
    "lr": 0.00001,
    "max_grad_norm": 0.5, #그레디언트 폭주 방지-최대 그레디언트 norm값
    "norm_adv": True,
    "clip_coef": 0.2,
    "ent_coef": 0.0055,
    "vf_coef": 0.5,
    "gamma": 0.99,
    "gae_lambda": 0.95,
    })
print(ppo_config)
print(ppo_config.minibatch_size)
```

Experiments

Parameter Experiment-with larger rl rate




```
ppo_config = OmegaConf.create({
    "anneal_lr": True,
    "update_epochs": 15, # The number of iteractions of ppo training
    "minibatch_size": 64,
    "lr": 0.0003,
    "max_grad_norm": 0.55, #그레디언트 폭주 방지-최대 그레디언트 norm값
    "norm_adv": True,
    "clip_coef": 0.2,
    "ent_coef": 0.01,
    "vf_coef": 0.5,
    "gamma": 0.99,
    "gae_lambda": 0.95,
    })
print(ppo_config)
print(ppo_config.minibatch_size)
```

Experiments

Parameter Experiment-Reducing the number of hidden units

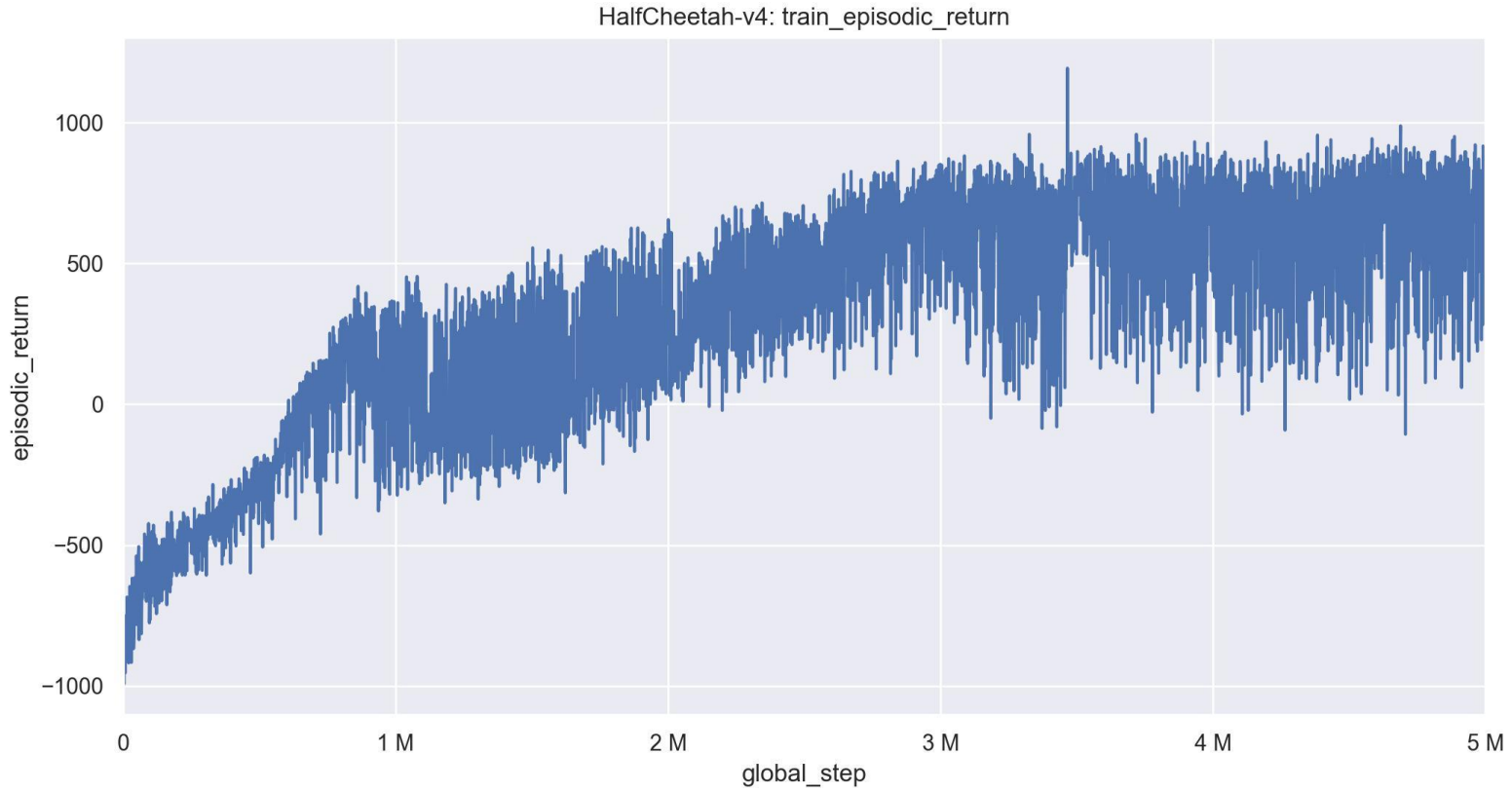
```
self.critic = nn.Sequential(
    nn.Linear(self.state_dim, 4),
    nn.Tanh(),
    nn.Linear(4, 2),
    nn.ReLU(),
    nn.Linear(2, 1)
)
```

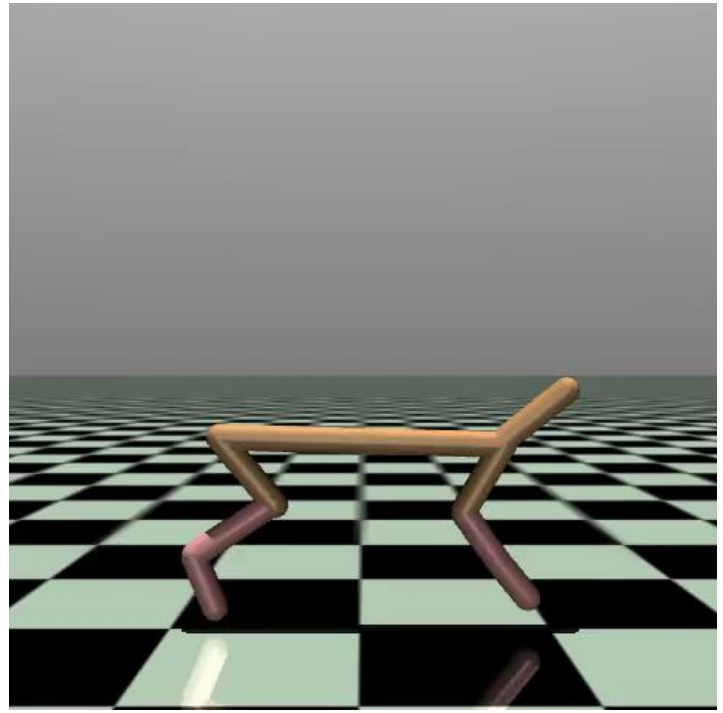
Hidden unit number

64	4
Overfitting occurred Bad learning	Much better learning

Experiments

Performance





Conclusion

Overfitting

Too many hidden units can lead to overfitting.

RL rate

High RL rate can make learning unbalanced

Gamma value

Low gamma value can make learning immediate

Entropy coefficient

High entropy coefficient can make Learning unstable