# Generative Adversarial Imitation Learning for Quadrupedal Locomotion using Unstructured Expert Demonstrations

Siddhant Gangapurwala, Ioannis Havoutis (*DRS, Oxford Robotics Institute*)

## Introduction

Robotic locomotion tasks heavily rely on control solutions for considerably simplified mechanical models of the robotic system in development. Despite being regarded as a highly complex problem, locomotion tasks performed by animals and humans can be considered to be near-optimal and highly efficient. For this reason, our work focusses on the use of expert demonstrations performed by humans to infer an underlying policy that directs the expert behaviour, and further proposes the use of the learned policy for execution of locomotion tasks. This work extends upon the technique used by authors of [1] on generative adversarial imitation learning (GAIL) to quadrupedal locomotion. The work also demonstrates the feasibility of using a learned policy over states that are not included in the set of expert training examples thereby making the algorithm suitable for execution in environments that are significantly different from the training environments.

## Problem Statement

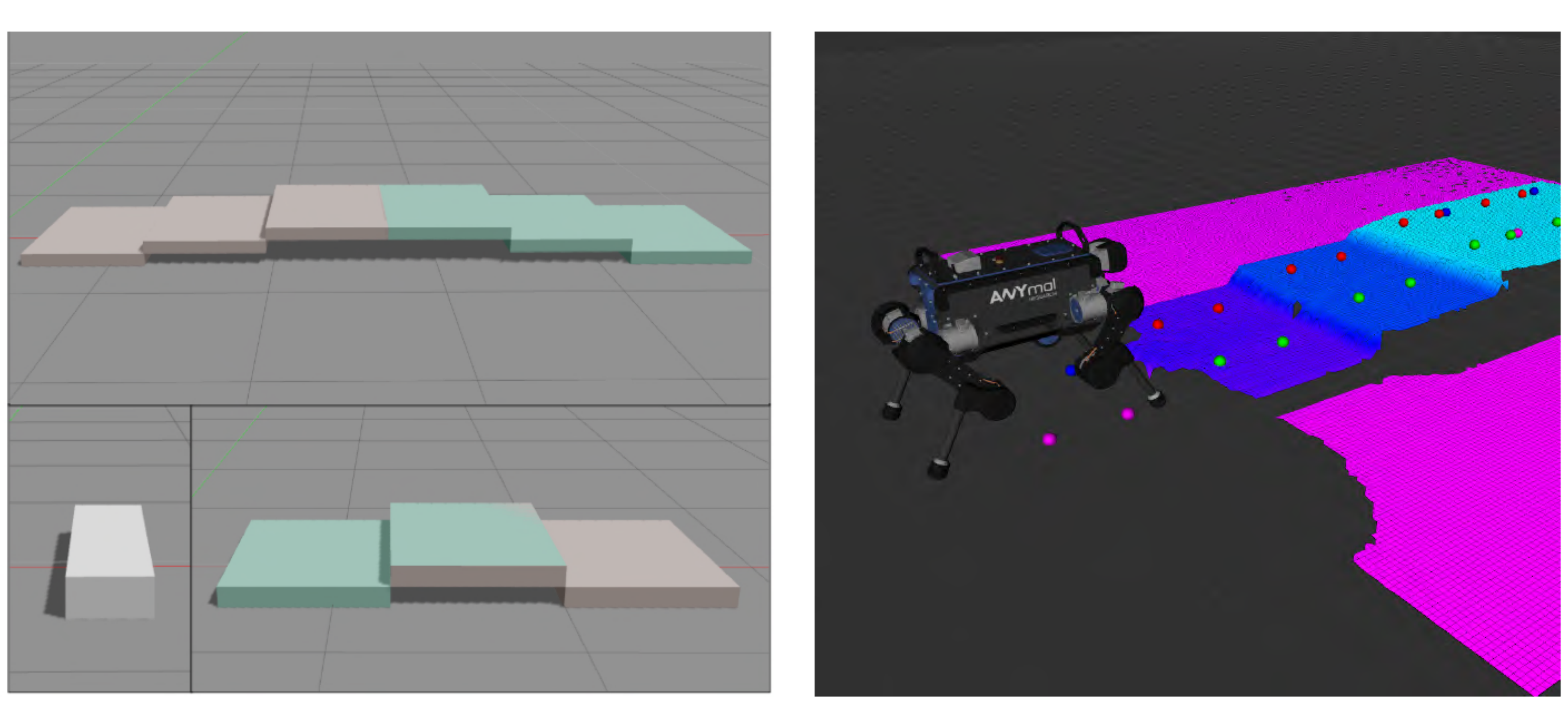
*Planning sequence of footstep placements for ANYmal quadruped [2] using elevation map and robot state as observations in order to ascend and descend a wide range of staircases using generative adversarial imitation learning approach by providing expert demonstrations for inference.*

The problem statement was slightly modified from the original statement in which, instead of planning footstep placements, obtaining joint trajectories was proposed. However, the higher dimensionality associated with solving the problem for 12 joints in the quadruped resulted in trajectories which were not smooth, and thus, not suitable for robotic control.

## Approach

### *A. Expert Demonstrations and Trajectory Data Collection*

For the purpose of demonstrations, various world models were created in Gazebo, some of which are shown in Figure 1 (a). The expert data collected was in the form of state-action pairs wherein the state consisted of a 200x200 elevation map matrix and an 8-dimensional vector consisting of the robot end effector positions relative to the robot's base.



(a) (b)

**Figure 1:** (a) World models created using Gazebo for demonstrations. (b) Example of expert footstep placement plan used as demonstration for GAIL algorithm.

The expert demonstrations were given by preselecting the positions of the each of the legs' end effector for each set of steps. The locations of these pre-planned footsteps (for one of the demonstrations) are as represented in Figure 1 (b).

### *B. Learning a Policy from Expert Demonstrations*

The GAIL algorithm was used to train a policy from the expert data collected in the previous step. The training was done in a manner such that the previous four state vectors were used to form the state representation vector as in Figure 2. Having obtained the data, the GAIL algorithm with the losses defined in [1] were used to train the neural networks with a PPO step for generator policy optimization. The training was slow especially because the simulations in Gazebo had to be run every time new samples were required. The time consumed in generating the expert trajectories was significantly high.

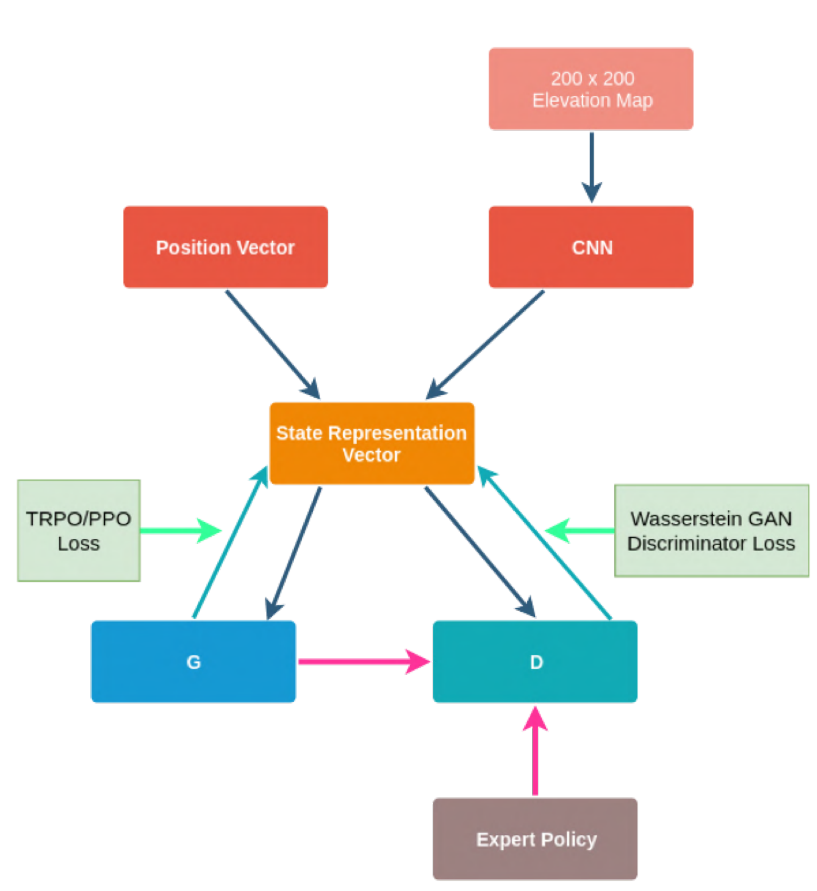
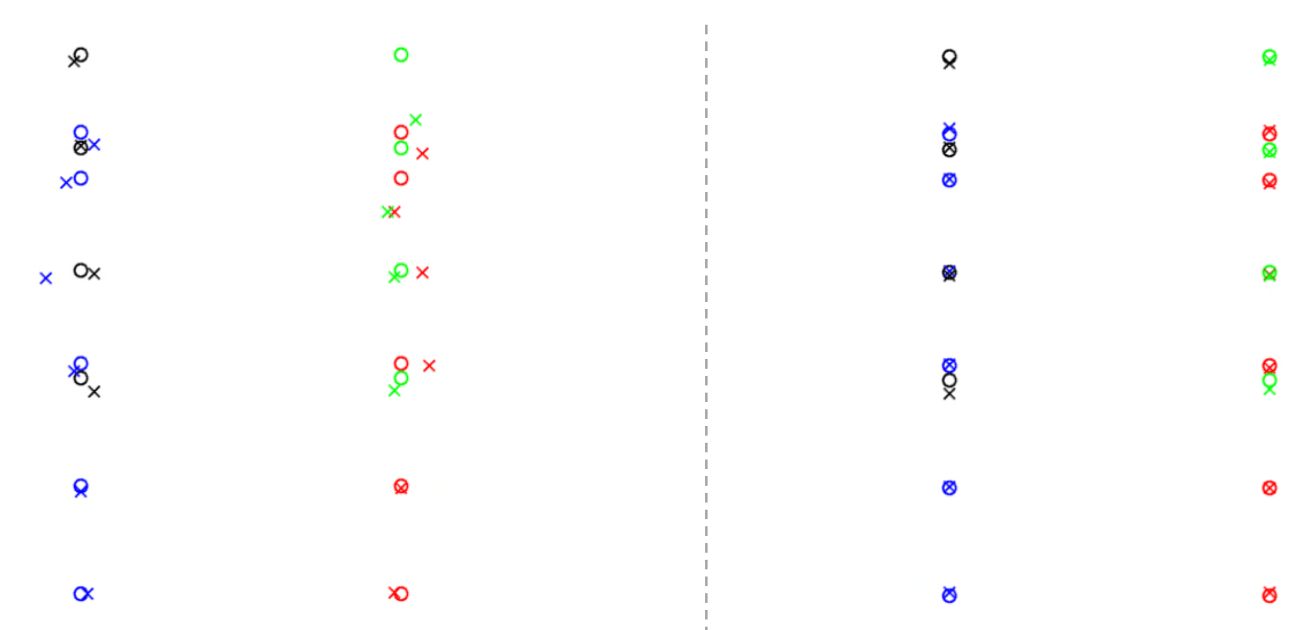



**Figure 2:** The flow of the algorithm implemented for imitation learning using generative adversarial networks (GANs).

### *C. Executing the Learned Policy*

Upon training the generator by the method shown in Figure 2, the policy was tested for its performance on Gazebo worlds that had not been included in the training examples.



(a) (b)

**Figure 3:** The figures represent the footstep placement for a certain test Gazebo world. The circular markers represent the expert's footstep placement, while the cross markers represent the output of the GAIL algorithm. (a) Observed output for obtaining joint trajectories using GAIL. (b) GAIL for footstep placement planning.

## Observations and Results

Through the experiments conducted with GAIL for world models that had not been a part of the training set, it was observed that the algorithm performed well for new environments. It successfully managed to ascend and descend 4 out of 7 test staircases. Out of the other 3 staircases, the algorithm managed to ascend the staircase for 2 world models, however, failed to descend. This was largely due to the fact that the steps were not observable. The laser scanner failed to detect the steps that were right underneath the scanner. For the other 1 test staircase, the algorithm failed to plan footsteps that could result in a stable ascent. It was observed that along the edges of the steps, the algorithm generated foot sequences that were extremely close to the edges. In 3 test cases, the footstep plan was quite close to resulting in the quadruped flipping over.

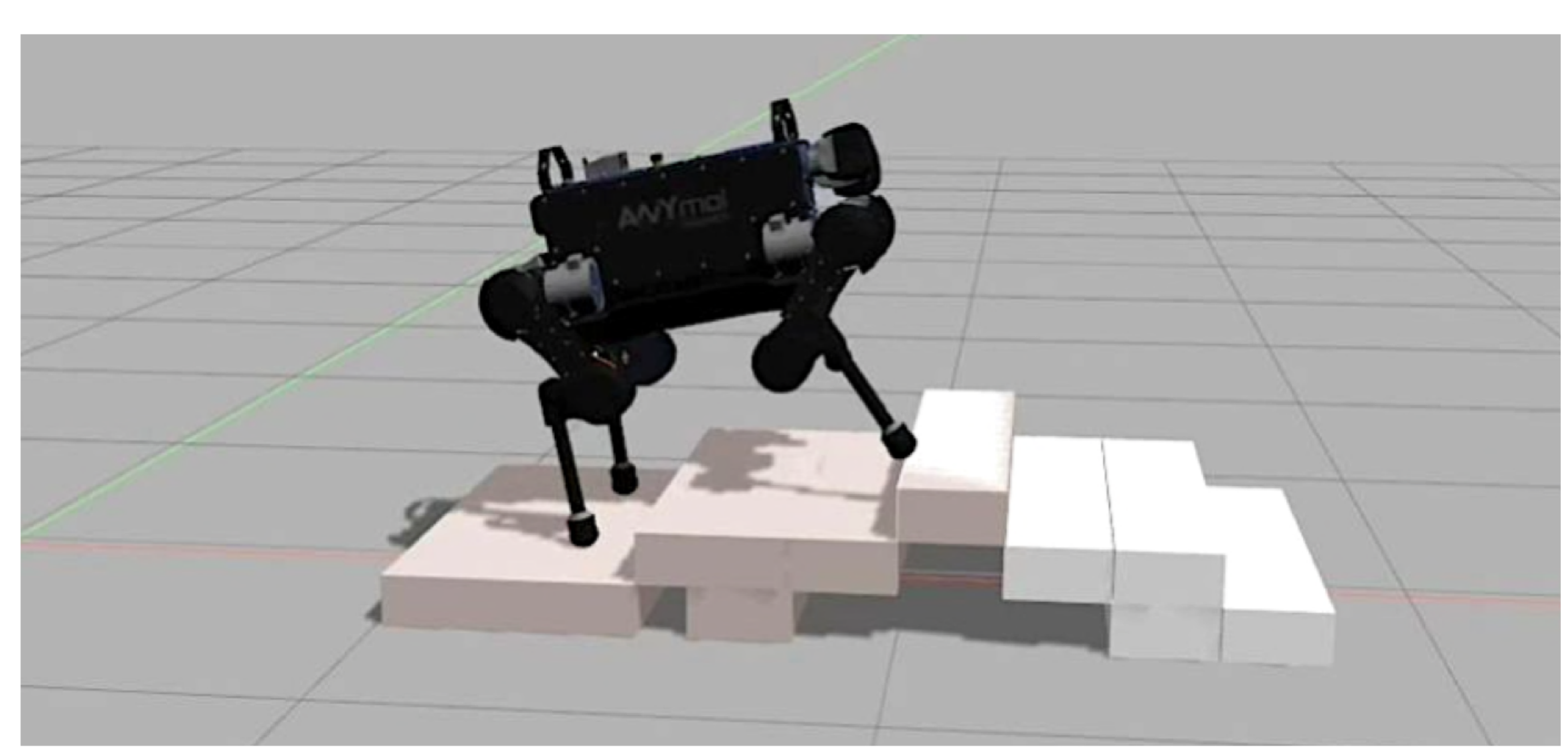


**Figure 4:** ANYmal successfully ascending and descending a test environment using the GAIL strategy,

## References

[1] Jonathan Ho and Stefano Ermon. "Generative adversarial imitation learning". In: Advances in Neural Information Processing Systems . 2016, pp. 4565–4573.
[2] Marco Hutter, Christian Gehring, Dominic Jud, et al. "Anymal-a highly mobile and dynamic quadrupedal robot". In: Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on . IEEE. 2016, pp. 38–44.

Report: jalark.com/gail
Contact: siddhant@jalark.com