Intermediate report CIT

2019 Week 18

Work done in past few weeks:

* labelled the latest drilling session but testing needs to be done on the new data set.
* A material detection algorithm with good accuracy based on the data from AD10.
* Starting to build Reinforcement learning framework based the result from convergence test with supervised learning.
* Implemented a prioritized sampler based on the importance sampling weights and priorities ([prioritized experience replay](https://arxiv.org/pdf/1511.05952.pdf) based on expert demonstrations).
* Built a loss function based on the paper deep q-learning from demonstrations ([DQfD](https://arxiv.org/pdf/1704.03732.pdf)) to initially train on the expert demonstration data and then continue training on the drill setup.
* Tested the dqfd using the openAI gym lunarlander-v2 environment.

1. Using the already trained simple dqn algorithm as expert, the demonstration data needed for dqfd algorithm is collected.
2. Now the DQfD algorithm is pre-trained using just the demonstration data and after it is trained, it allowed run on the virtual environment directly.
3. The pre-trained DQfD algorithm started to perform well from the start, which can be seen from the accumulated reward of each episode.



Fig1: In this figure the dark blue represents the training of simple dqn's and the light blue represents the DQfD's accumulated rewards per each episode of lunar-landing virtual environment. (x-axis: number of episodes/sessions, y-axis: cumulative reward per episode)

In-progress for next 2 weeks:

* Implement a real-time and testable material detection algorithm.
* Build environment for drilling (reward function, terminal states, transitions etc.)
* Train the DQfD algorithm on the cleaned and pre-processed drill data and finally test it on the drill once everything looks good.