**Report of COMP219 Ass2 Task2**

**Maizhen Ning 201376369**

**How to run:**

After you decompressing this project, please make sure all the files are in one folder. My development is under Ubuntu16.04.

**Libraries & Environment:** Ubuntu16.04, numpy, tensorflow-gpu-1.12.0, CUDA9, OpenAI Gym, OpenAI Universe, OpenCV3. (Please make sure your computer is ready for openai universe, it may need install the docker but during my development, I don’t need to activate docker before run the code. But you may need to activate your environment.)

Run the code: Open a new terminal and cd to the path of this folder. If you want keep training please run training.py and please use ‘sudo’ command because it may need authority to activate docker of OpenAI (e.g. sudo python training.py). If you want run test please use testing.py (e.g. sudo python testing.py).

The performance is not good enough, so please give algorithm several times to see it is actually working, thank you!

\*Sometimes the first time of activate OpenAI after turn on computer may get an error, that is fine, please just run again. I list some error report during my test at the end.

If files be damaged, please go visit: <https://github.com/ning-mz/COMP219-Ass2>

The code will be released be public after DDL of Assignment.

**Description:**

This algorithm is based on DQN with CNN to achieve both deep reinforcement learning and deep learning. The game is SpaceInvaders-v0 in the OpenAI universe environment. I used this environment to get data from game and do the action of game. The network be saved in folder had 400000 steps training. However, it is not good enough and please see the disadvantage part for details.

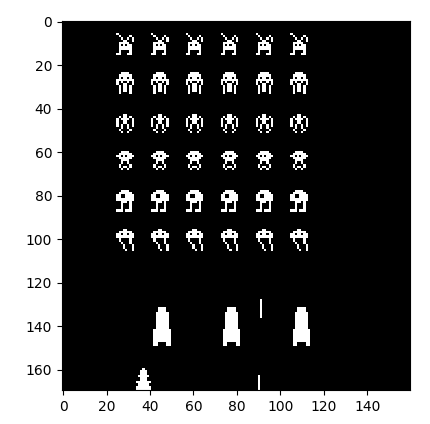
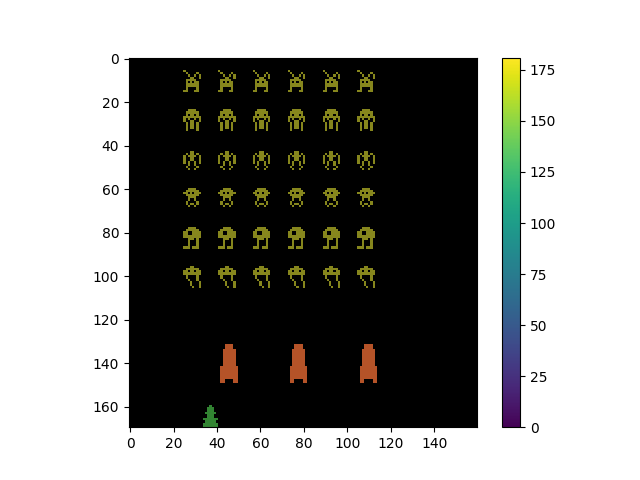
I did some changes of information the get from environment to help train the model. The reward from environment be changed. The original reward data only give reward when shoot the target, I added the negative reward when spaceship die. I also changed reward when spaceship do nothing to make the algorithm become more energic.

The DQN is based on the model that on the Nature, and CNN algorithm is changed from LeNet. There are 3 convolutional layers, 1 pool layer and 2 dense layers. The reason why I only used 1 pool layer after the first convolutional layer is that the bullet in this game is difficult to get, it moves fast and shining on the way, it will be easy to lose this important image information during processing.

The first 10000 steps are recording into batch for training, the next 10000 steps are for explore. So, after 20000 steps model will start training and the epsilon will reduce during these stages. The algorithm will start get the large control of action and keep training.

**Simple Flow:**

1. Start game and get image data from environment and do the preprocess of image.
2. Get the action command from algorithm and call step() function to play the game.
3. Get the return information from environment (including image data).
4. CNN layers do the image process with 3 convolutional layers and 2 dense layers.
5. DQN algorithm each time get 64 data from batch to train the network
6. Loop 2-5



**Disadvantage:**

You will see the algorithm is not good enough after 400000 steps training, I think there are some reasons.

Bullet is still difficult to avoid, it may need to keep optimize the CNN layers. But the original game fps may need to be higher to show the response of action.

Because the original reward from environment only give reward when shoot target successfully, I adjusted the reward. So, the reward adjustment function may need optimize.

CNN layers need further adjustment even I have tried several hyper parameters of kernel, but I think it can get better.

**Error List:**

1. UnknownError (see above for traceback): Failed to get convolution algorithm. This is probably because cuDNN failed to initialize, so try looking to see if a warning log message was printed above.

**Solve:** Your GPU resource are using by other process, please clear or just reboot.

2. raise error.Error('{}/{} environments have crashed! Most recent error: {}'.format(len(self.crashed), self.n, errors))

universe.error.Error: 1/1 environments have crashed! Most recent error: {'0': 'Rewarder session failed: connection was closed uncleanly (WebSocket connection upgrade failed (502 – BadGateway))'}

**Solve:** This is always happen when the first time run openai after turn on computer, just run the python file again, it will not happen again.