# Double Dueling DQN with soft update and memory replay

The algorithm is based on 2 5-layer neural networks. One of these neural networks called local QNetwork learns by exploring and exploiting the environment and after every 5 episodes transfers 1% of its knowledge to the other neural network called target QNetwork. The weights are calculated by the difference between the returns actually received and the QNetwork Target prediction. It is worth mentioning that this neural network is internally divided into Q-Value which is a Value of a State and Advantage which is the advantage that each Action has. The final output is calculated by summing the outputs of both neural networks followed by subtraction by the average value of the Advantage neural network's output.

To eliminate the effect of correlation between the actions, states and rewards, a memory buffer of size 10000 is created and sampled randomly in batches of size 128.

The gamma discount factor is 0.99, which helps focus on long-term reward and the learning rate is 0.0001 which gives a lot of stability to the model.
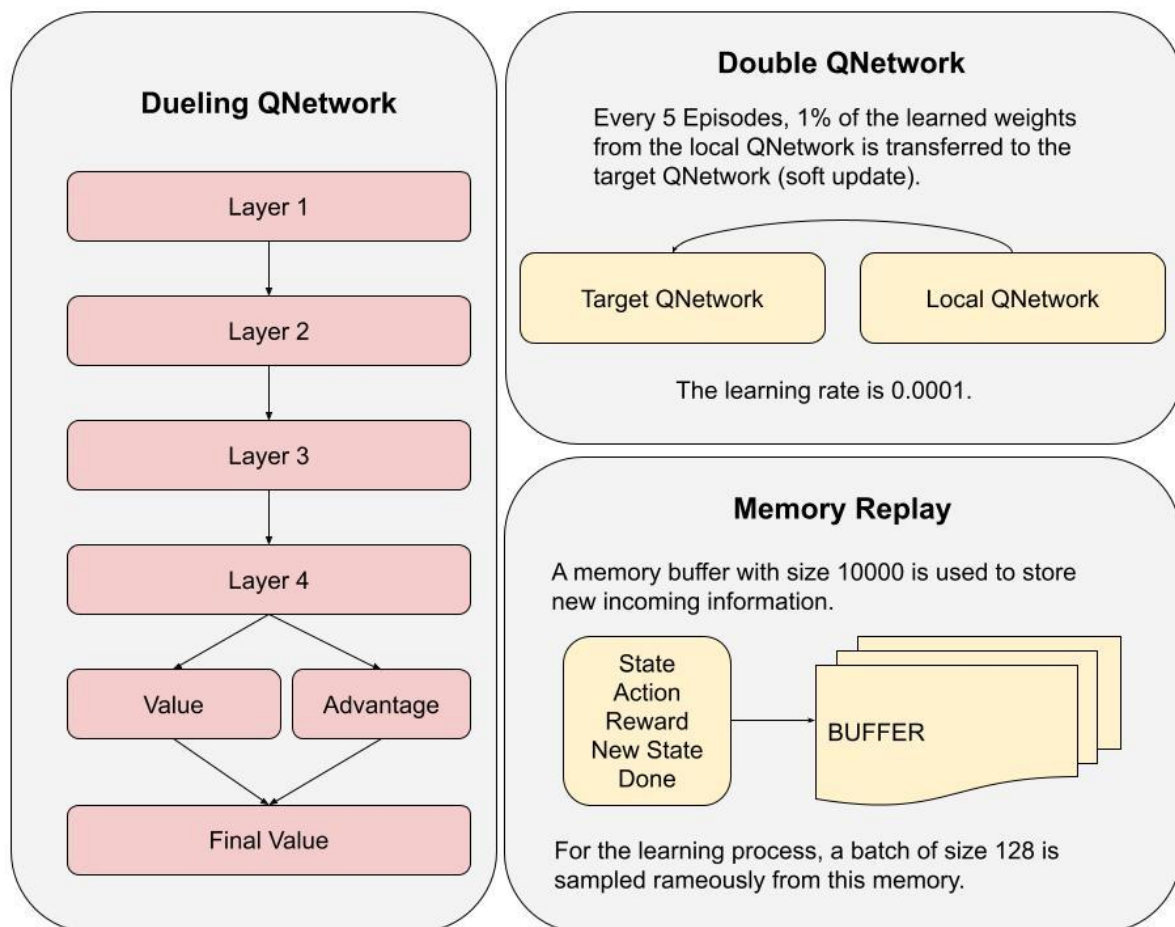
Ideas for Future Work

In order to improve the performance of this agent, several adjustments could be made. The main one is to add the screen image as part of the input via a convolutional neural network. Also a "Prioritized Experience Replay" could be implemented that would balance learning of states and actions that happen infrequently.