1 Experiment

First, just to check that stuff works, the convolutional autoencoder identical to the convolutional layer of the policy network in the Nature paper was used. It worked very well, but the feature dimension was 3136. Also the Q learning from this was in fact slower, altough that could be due to the fact that the policy network was smaller as it was just 2 FC layers.