

Homework #5 Report

土木所交通組 R12521502 陳冠言

Describe the Deep Q-Network.

Deep Q-Network (DQN) 是一種用於解決強化學習問題的深度學習模型，DQN將 Q Learning 的概念與深度學習相結合，其中 Q Learning 是一種基於價值的強化學習方法，其核心是學習一個價值函數 $Q(s, a)$ ，代表在狀態 s 下執行行動 a 所期望的未來獎勵。DQN 的主要架構包括狀態表示 (State Representation)、行動評估 (Action Evaluation)、行動選擇策略 (Action Selection Policy)、優化目標 (Optimization Objective) 以及經驗回放 (Experience Replay)。

而 DQN 的關鍵創新之處為引入經驗重放 (experience replay) 的概念，主要是在訓練過程中，模型不是直接使用當前的經驗來更新網絡，而是將這些經驗存儲在一個記憶體中並隨機抽取過去的經驗來學習。而另一個創新之處為建立目標網絡 (target network)，在 DQN 中主要有兩個神經網絡：一個是正在訓練的網絡，另一者為固定的目標網絡，有助於穩定訓練過程，因為目標網絡的參數通常會比較慢更新。

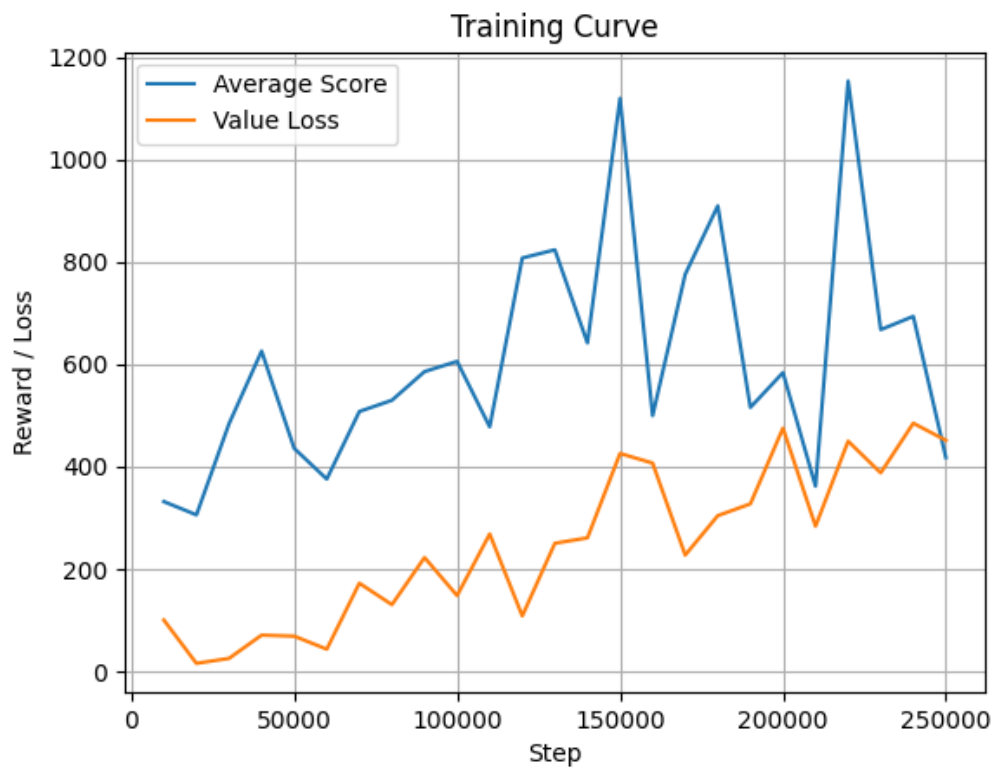
Describe the architecture of your PacmanActionCNN.

模型主要包括三層的 Convolutional Layers，並對此三層 layers 都做 Batch Normalization，然後在第二層的部分加入 Max Pooling，最後用 Fully Connected Layers 來輸出。那在 forward 的過程中，各層都有透過激活函數 (ReLU) 進行非線性轉換，並將最後一層的輸出作為模型的預測結果返回。

那在設計的過程中有發現幾個特點，說明如下：

1. 若 Convolutional Layers 設計得太過複雜或是神經元設置過多會導致結果變差或是出現 error。
2. 模型加入 Batch Normalization Layers 後，有提高模型的穩定性讓輸出結果不會差距過大且有變好的趨勢
3. 加入 Max Pooling 後提取最重要的特徵，可讓模型分數更高但加入過多或是插入在不適當的位置時會適得其反。
4. 有嘗試過雙層、三層和更多層的全連接層，但發現三層的效果最好，多層可能會導致訓練時間過長。

Plot your training curve, including both loss and rewards.



Show screenshots from your evaluation video.

