

■ Describe the Deep Q-Network

Deep Q-Network (DQN) 是一種結合了傳統 Q 學習演算法與深度學習技術的增強學習方法。DQN 利用深度神經網路來近似 Q 函數，這是一個預測在給定狀態下採取特定行動所能獲得的總回報的函數。這種網絡能夠處理高維度的輸入數據，使其能夠應用於例如視覺資料的複雜環境中。

DQN 的關鍵創新包括使用「體驗重放」(experience replay) 來打破數據間的相關性，以及「固定 Q 目標」(fixed Q-targets) 來穩定學習過程。這些技術幫助 DQN 在多種任務上達到了優異的性能，特別是在玩 Atari 2600 視頻遊戲等複雜環境中表現出色。

■ Describe the architecture of your PacmanActionCNN

PacmanActionCNN 架構

第一層卷積層 (Conv1):

輸入通道數：由 `state_dim` 決定，這是輸入狀態的維度。

輸出通道數：16

卷積核大小：8x8

步長 (Stride): 4

這層主要用來從遊戲狀態中提取特徵。

第二層卷積層 (Conv2):

輸入通道數：16

輸出通道數：32

卷積核大小：4x4

步長：2

進一步提取遊戲特徵，增加了網絡的深度來捕捉更複雜的特徵。

第三層卷積層 (Conv3):

輸入通道數：32

輸出通道數：32

卷積核大小：3x3

步長：1

這層專注於捕獲更精細的特徵細節。

全連接層 (Fully Connected Layer - Head):

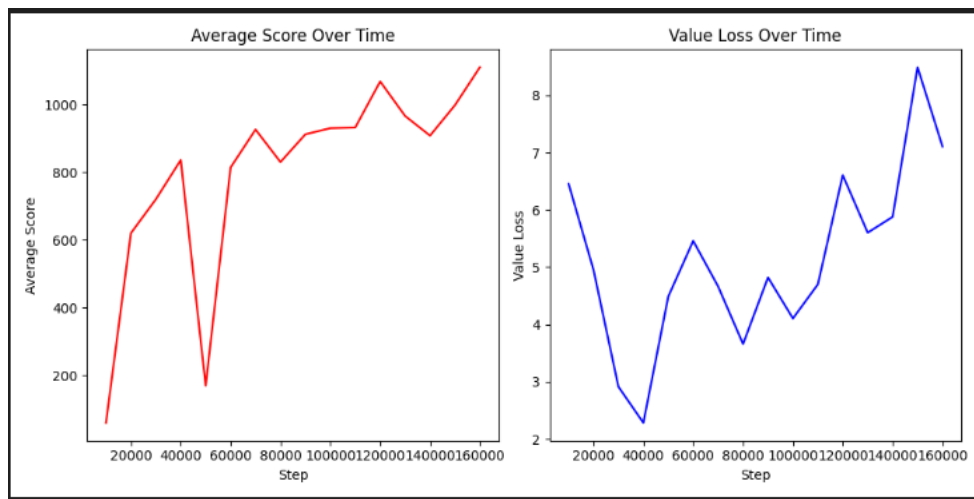
輸入維度：通過卷積層之後，根據卷積操作的計算得出的維度。具體數字取決於前面卷積層處理後的輸出大小。

輸出維度：`action_dim`，即根據遊戲中可能的動作數量決定。

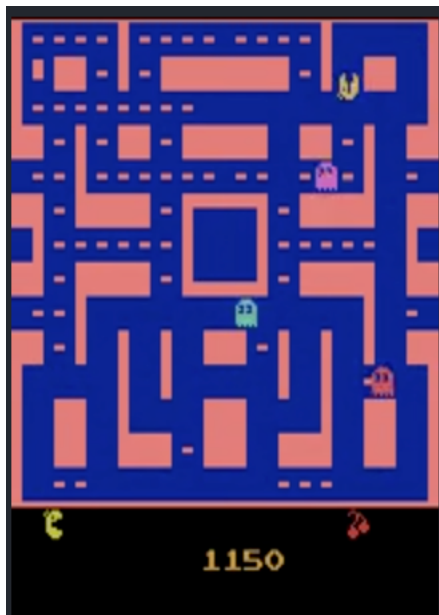
這層將前面卷積層的輸出展平並進行最終的分類或回歸，以決定應採取的動作。

整個網絡透過這些層來處理遊戲的視覺輸入，並輸出一個決策動作。這個架構適用於處理基於圖像的輸入並作出反應，是遊戲 AI 中常見的卷積神經網絡設計。

- Plot your training curve, including both loss and reward



- Show screenshots from your evaluation video



補充：助教您好，由於學生的電腦非常老舊，顯卡為 GTX 1050Ti，故跑得真的非常非常慢，從截止日前大約 3~5 天開始跑，中間還會因為電腦過熱自動關機，導致學生需要重跑，故到今天已經遲交的最後期限，還是只跑了 160000/250000 個 epoch 而已，分數有持續在上升，但就是沒有達到最低標準 1500。

在這次之前的作業，都大約 1~3 天可以寫完，學生並不知道這次會花這麼久時間，雖然晚開始也是學生的問題，且作業也已經延長了，但還是希望助教能夠斟酌扣分><，造成助教的不便非常抱歉！