

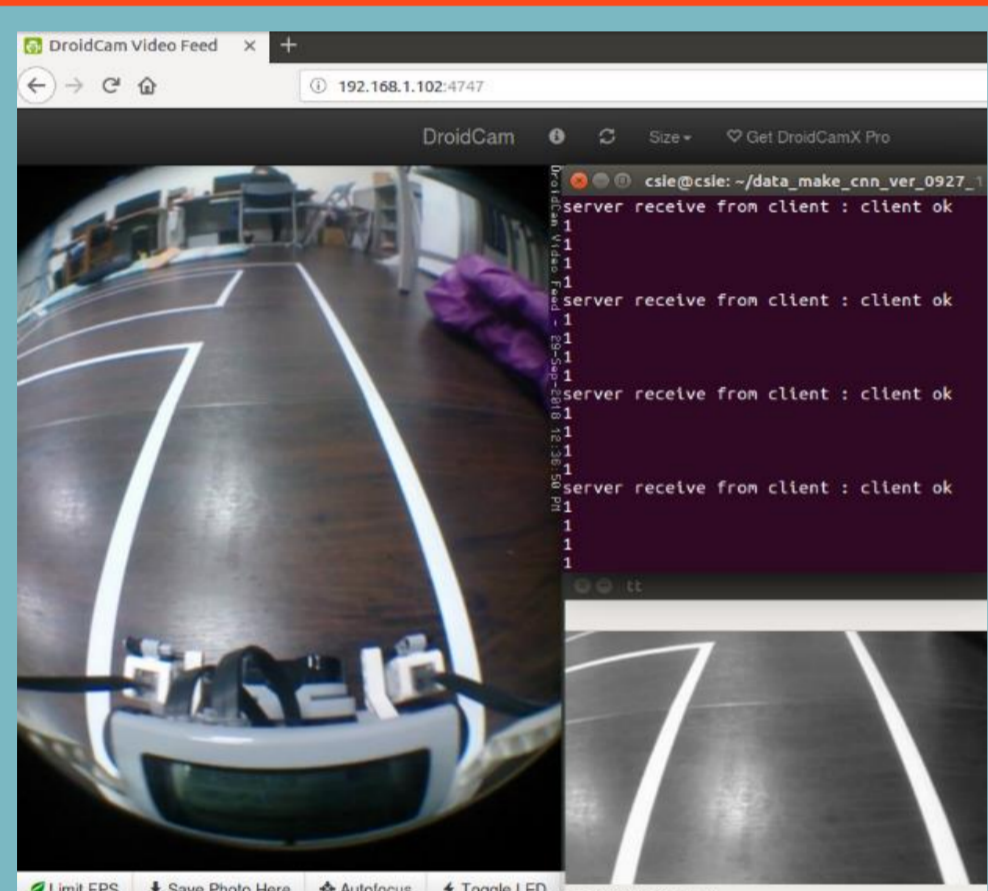
# 國立臺東大學理工學院 學習成果競賽

## 深度神經網路結合強化學習之自走車應用 Reinforcement Learning with Deep Neural Networks for Automatic Mobile Vehicle Applications

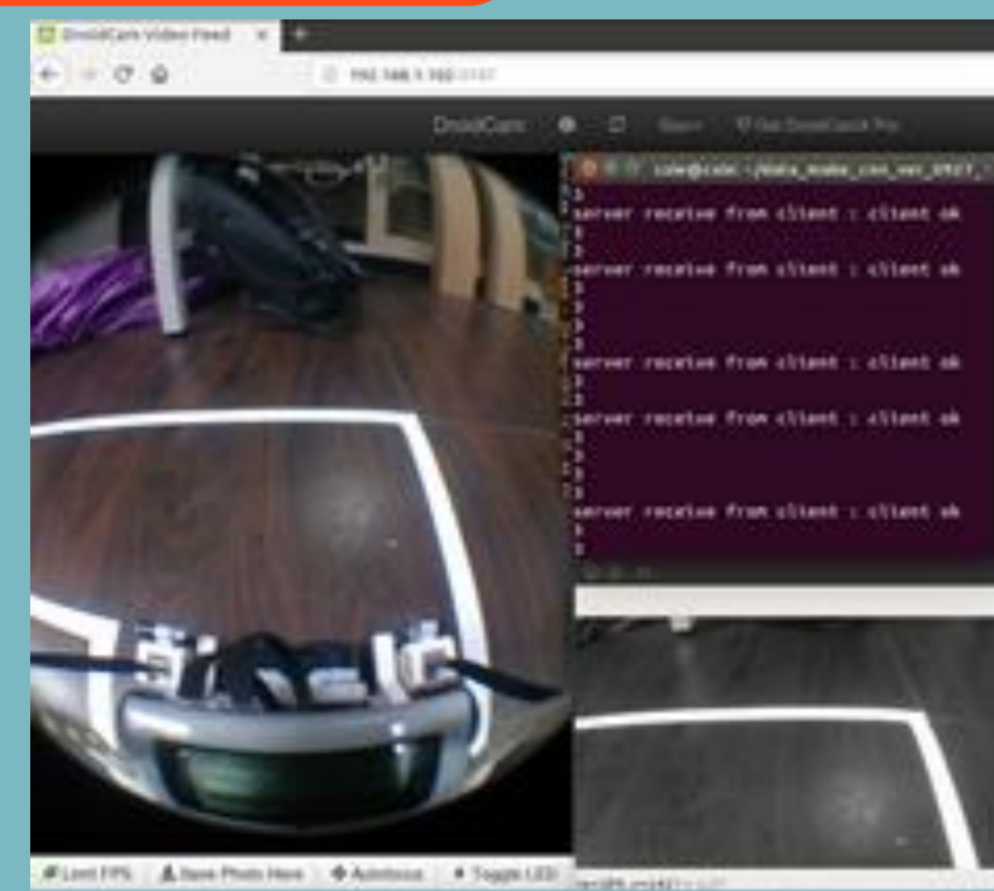
### 摘要

近幾年人工智慧受到越來越多人的關注，透過 Alpha Go 在圍棋的傑出表現、自動駕駛問世、機器人 Sophia 的對答如流都顯示出人工智慧應用的廣泛性與可發展性。跟上這股潮流，本專題將利用即時影像結合深度神經網路與強化學習，利用手機鏡頭傳輸即時畫面到Server端，並由卷積神經網路(Convolutional Neural Networks)即時擷取目前道路的特徵，取得特徵資訊後做出決策，再利用強化學習(Reinforcement Learning)方法中的Policy Gradient更新權重來訓練，使自走車學習到畫面中的各項特徵，來達到自走車的成果。

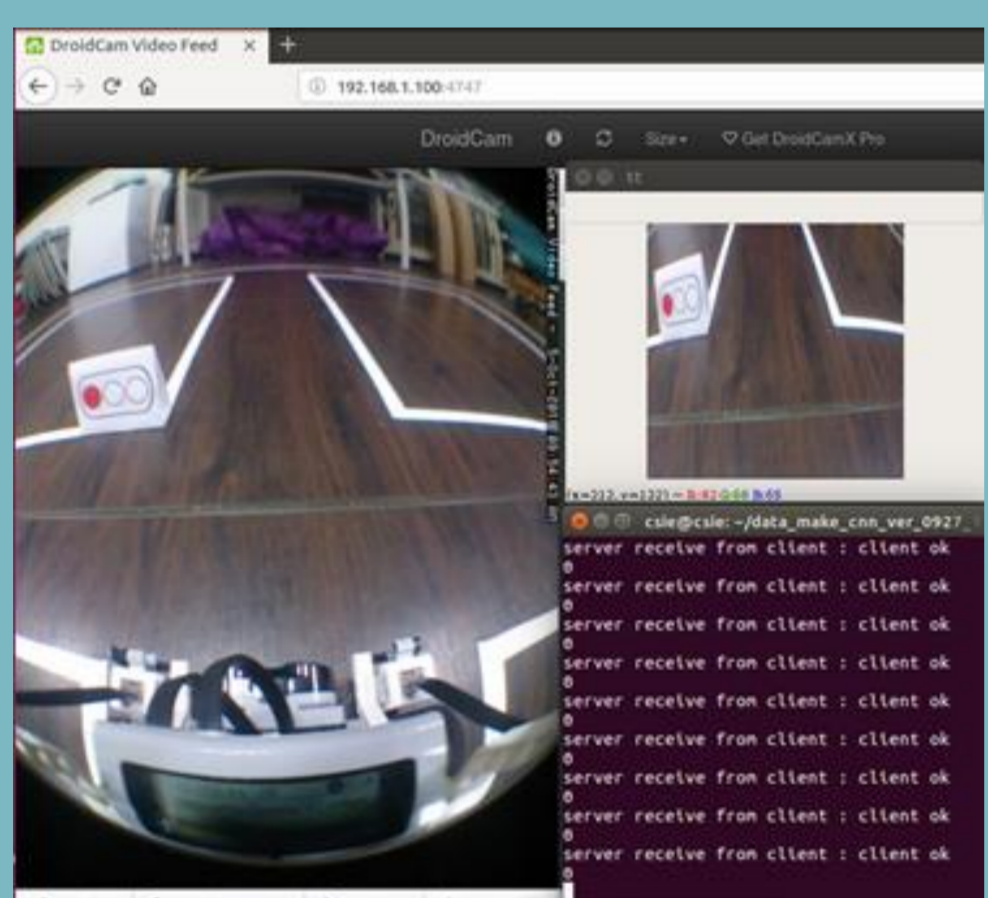
### Step1.擷取路況畫面(Server端)



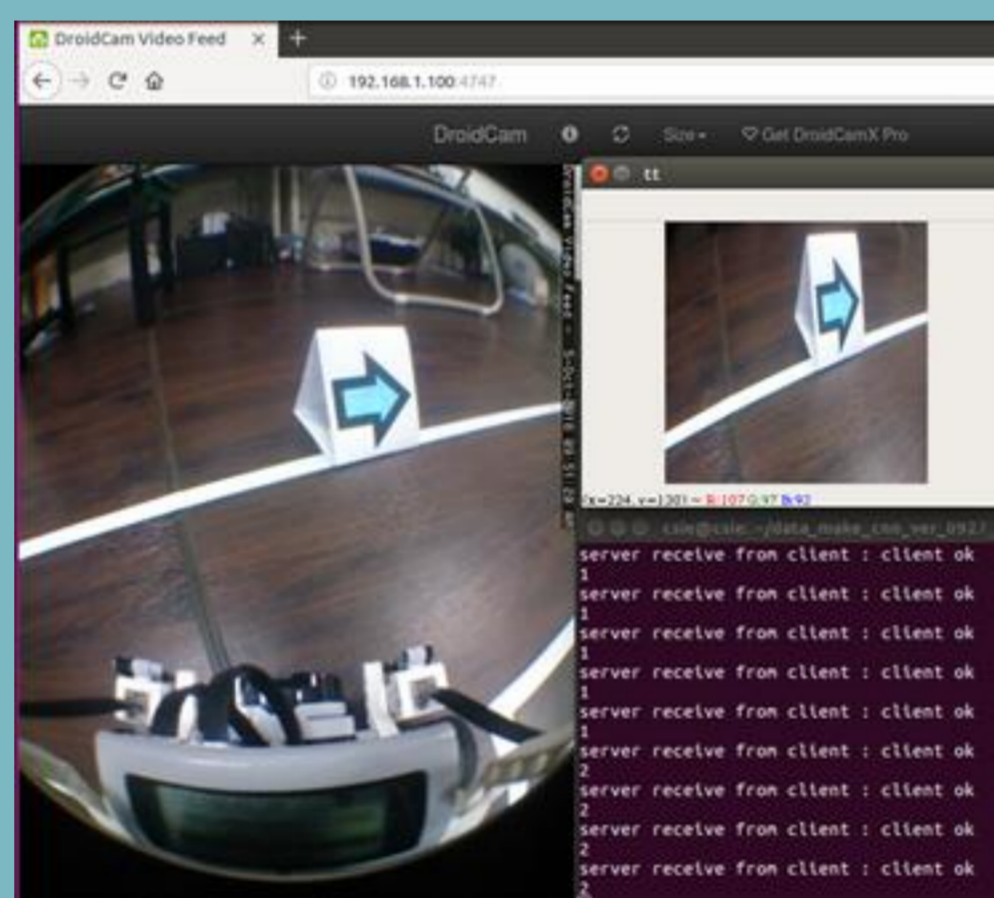
當擷取到直走的畫面



當擷取到轉彎的畫面



當遇到紅燈號誌的畫面



當遇到右轉號誌的畫面

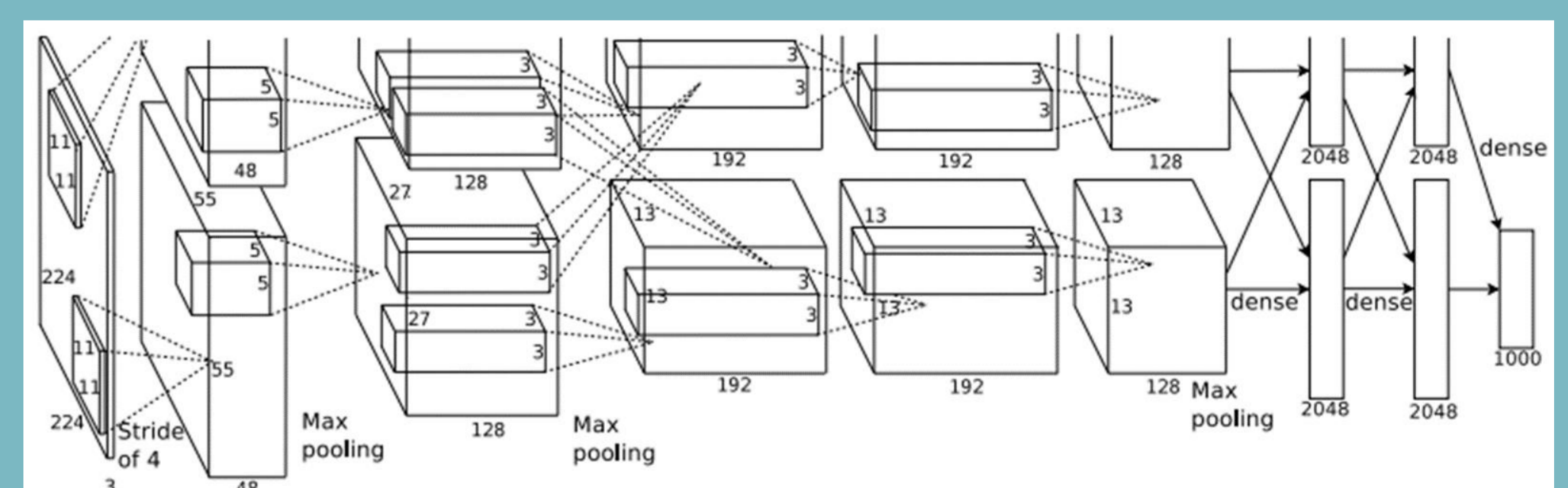
### Step2.傳送至CNN做決策

#### 卷積神經網路(CNN):

卷積神經網路與傳統多層感知網路差異在於多了卷積及池化，也因為多了這兩層結構，讓影像處理及辨識加快了不少。

#### Alexnet:

網路架構總共有8層Layers，前5層為卷積網路，後3層為全連接層。在第一&二的全連接層加入 dropout。這種方式使得網絡強制以部分神經元來表示當前的圖片，很大程度上降低過擬合。



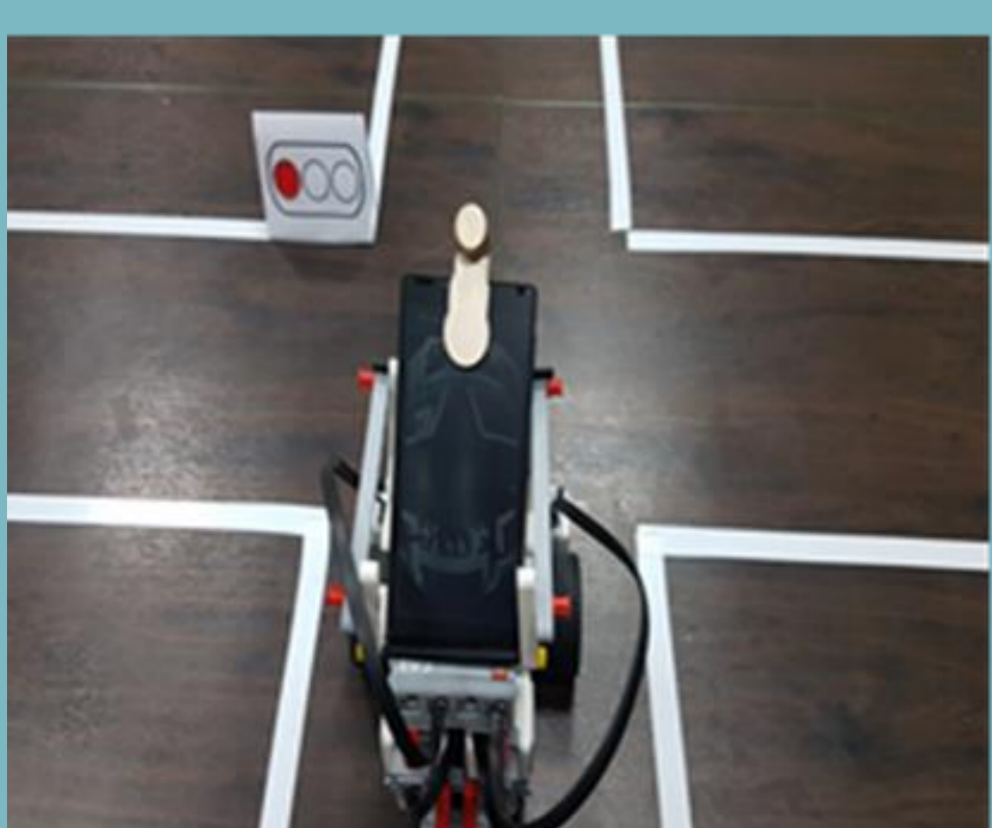
### Step4.傳送決策給EV3執行(實際環境)



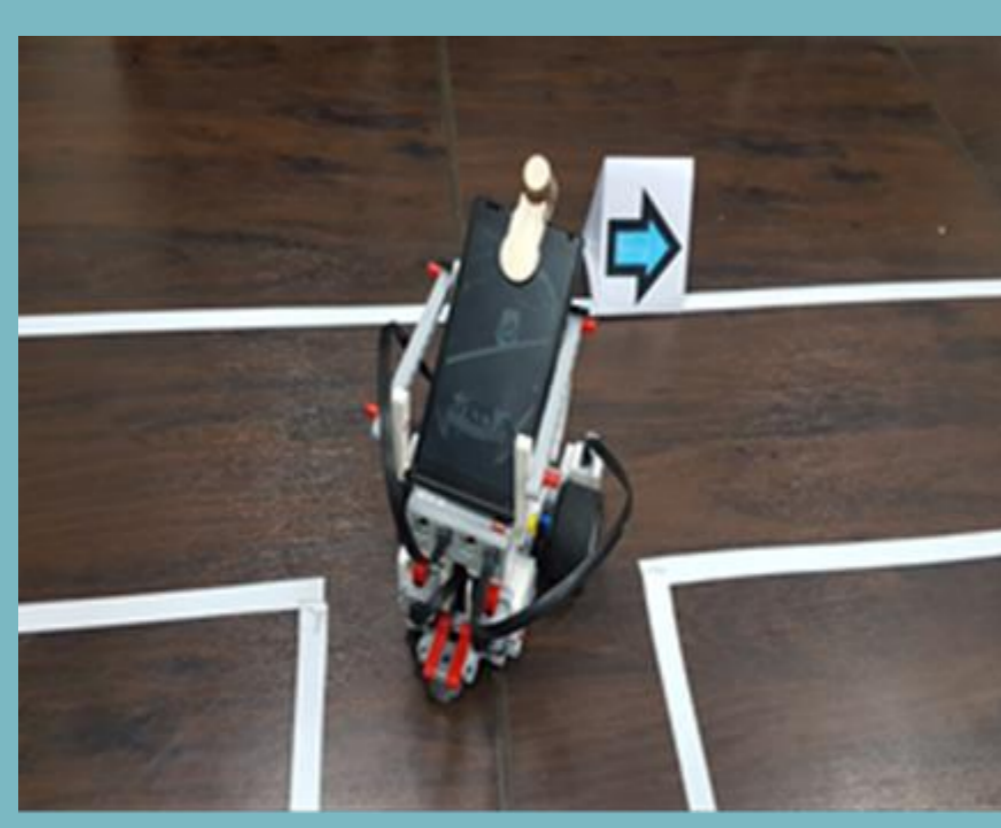
直走時修正於道路中央



轉彎時避免壓道路邊緣



當遇到紅燈號誌即停止

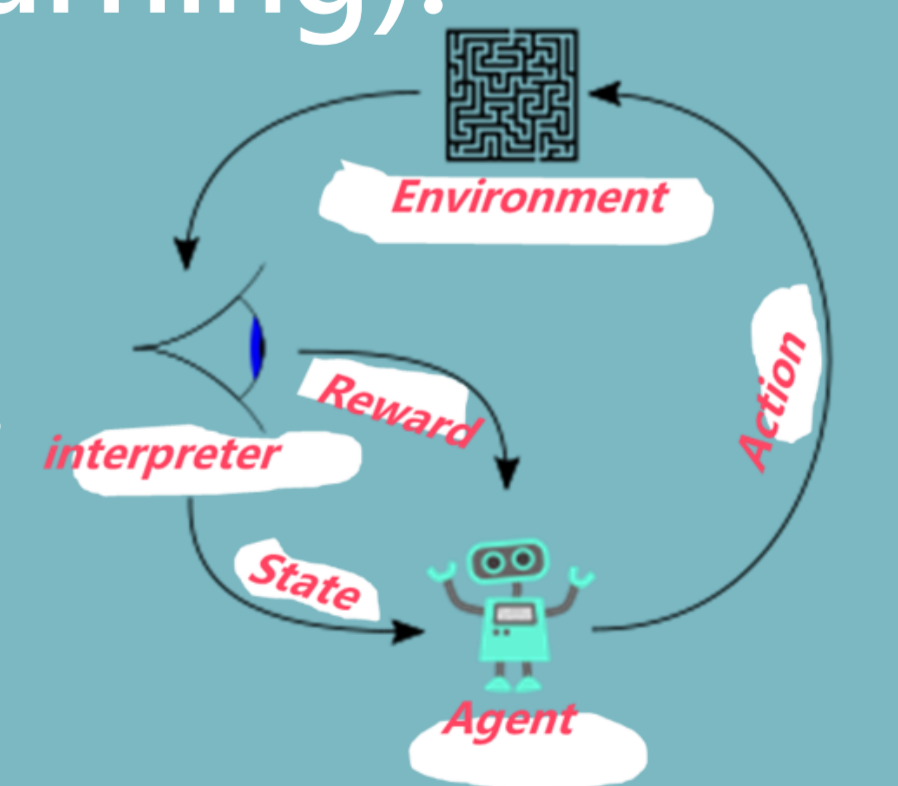


路口若遇右轉號誌即右轉

### Step3.經由強化學習修正

#### 強化學習(Reinforcement Learning):

強化學習是機器學習裡面的一個領域，主要建立在基於環境的回饋下行動，其特點是從無到有在環境裡面做出各種動作，從環境裡的回饋進而做出修正。



#### Policy Gradient:

Policy Gradient是強化學習中的一個方法，它是Policy-based方法，輸入為環境訊息，而它輸出為動作的機率，適合在連續分布上選取動作。

```

input: policy parameterization  $\theta_h$ .
1 repeat
2   generate policy variation  $\Delta\theta_1$ .
3   estimate  $\hat{J}_j \approx J(\theta_h + \Delta\theta_1) = \langle \sum_{k=0}^H a_k r_k \rangle$  from roll-out.
4   estimate  $\hat{J}_{ref}$ , e.g.,  $\hat{J}_{ref} = J(\theta_h - \Delta\theta_1)$  from roll-out.
5   compute  $\Delta \hat{J}_j \approx J(\theta_h + \Delta\theta_1) - \hat{J}_{ref}$ .
6   compute gradient  $\mathbf{g}_{FD} = (\Delta\Theta^T \Delta\Theta)^{-1} \Delta\Theta^T \Delta \hat{J}$ .
7   until gradient estimate  $\mathbf{g}_{FD}$  converged.
return: gradient estimate  $\mathbf{g}_{FD}$ .
    
```