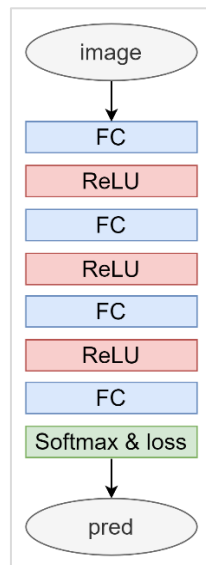


My network



How to improve the accuracy

一開始我使用 FC+ReLU+FC+Softmax 的架構，但我發現不管怎麼調超參數，train 的準確度到 97%就很極限了。於是我選擇使用如上方示意圖的架構(即多疊幾層)，train 的準確度才能來到 100%。

Loss function

由於此次是影像分類，所以我採用 [Cross entropy](#)。需注意的是 cross entropy 中有 log 項，遇到值為 0 會出問題，導致最後的 loss 為 NaN。所以在遇到值為 0 時，我將其改成 $1e-8$ ，就巧妙避免了這個問題。

Activation function

至於為什麼我要使用 [ReLU](#) 作為 [activate function](#)。拋開什麼梯度消失梯度爆炸不說，就是一點：ReLU 是最簡單的非線性函數，forward 和 backward 都不需要進行加減乘除運算，算起來快，對於只使用 CPU 訓練，何樂而不為呢？