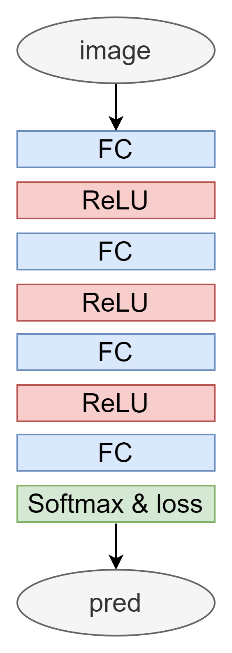
My network



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？