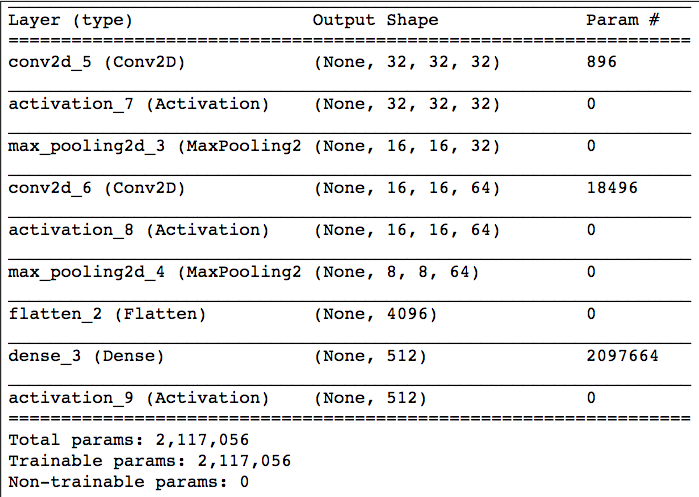
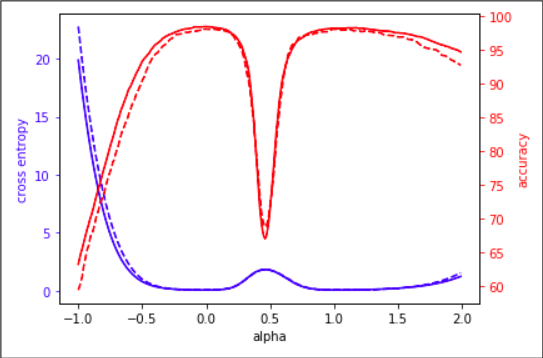
* + - Describe the settings of the experiments (e.g. which task, what training approaches) (0.5%)

使用MNIST資料集，model採用CNN，結構如下：



在該task中使用的兩個model的batch\_size分別為50及1000。兩個model都訓練到準確率接近99%

* + - Plot the figures of both training and testing, loss and accuracy to the number of interpolation ratio. (1%)



其中虛線為testing的結果 ; 實線為training的結果

* + - Comment your result. (1%)
    - 在這次的實驗中，比較難看出large batch size的sharpness，反而，當考慮進alpha=-1,2時，small batch size的流域較小。與課程投影片及助教相同的是，在從small batch size與large batch size之間的變化劇烈，或許可以理解成『既使在同一個model structure下，在不同batch size也會找到不同的local minimum』。