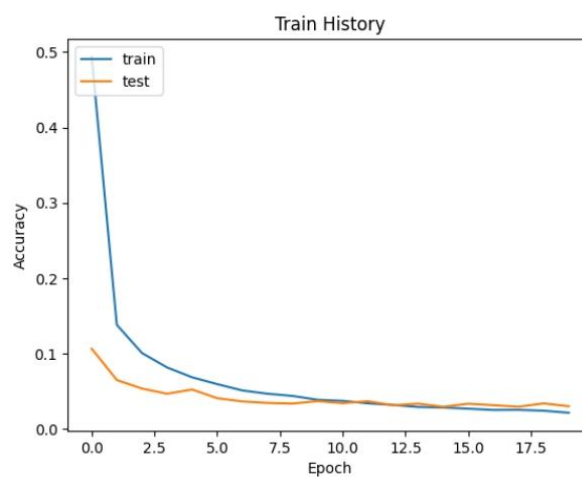
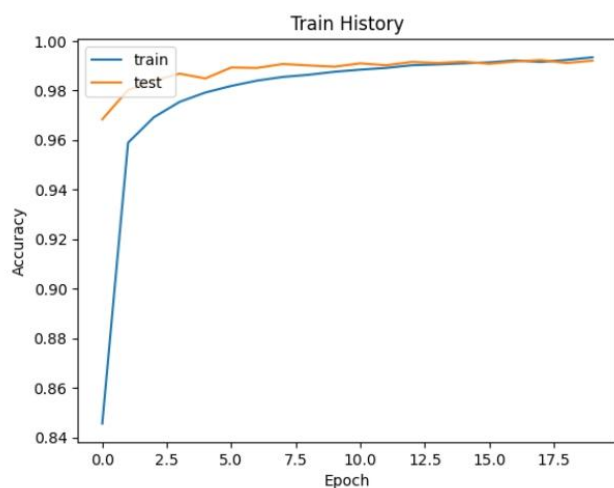


## Problem 1 - Handwritten Numeral Recognition

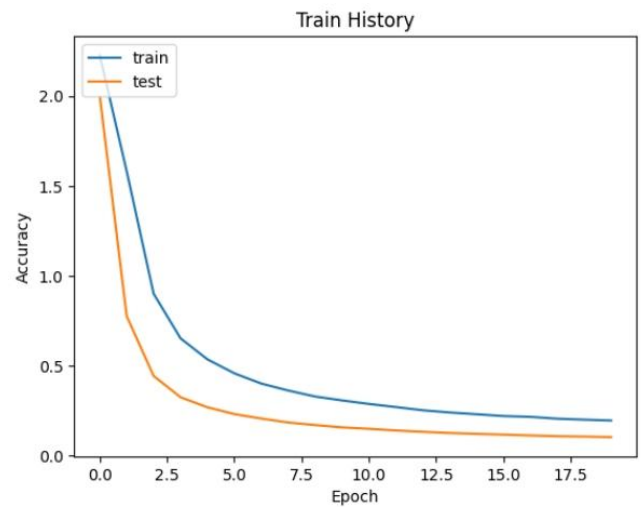
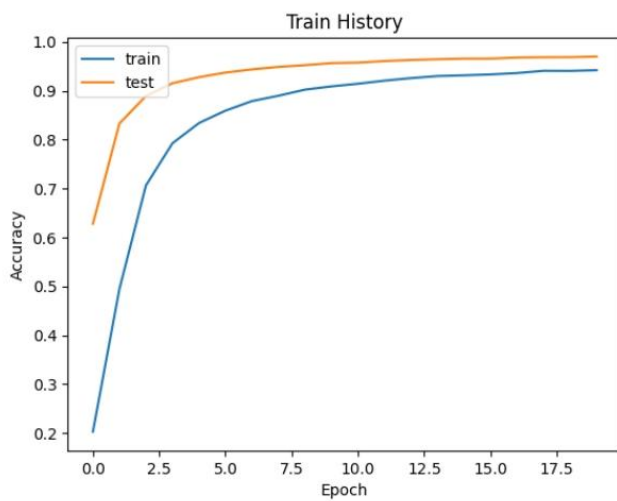
Epochs	Batch Size	Optimizer	Loss function	Accuracy	Run time
20	300	Adam	categorical cross entropy	train: 0.9922 val: 0.9920 test: 0.9928	25 sec
20	300	SGD	categorical cross entropy	train: 0.9428 val: 0.9698 test: 0.9691	30 sec
100	300	SGD	categorical cross entropy	train: 0.9800 val: 0.9874 test: 0.9902	202 sec
20	64	Adam	categorical cross entropy	train: 0.9944 val: 0.9924 test: 0.9925	83 sec

第一次  
結果為sampl code的示範。



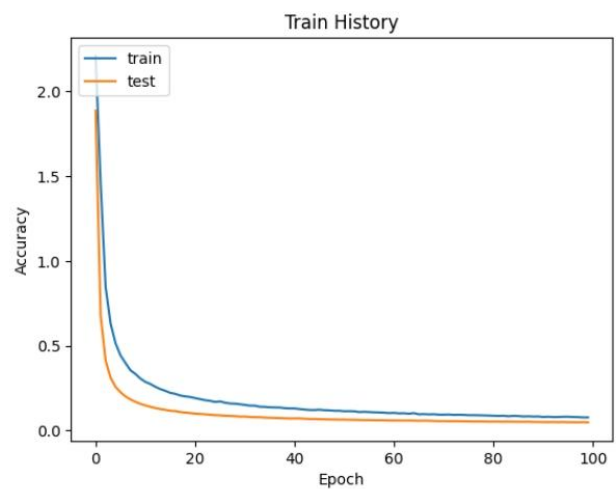
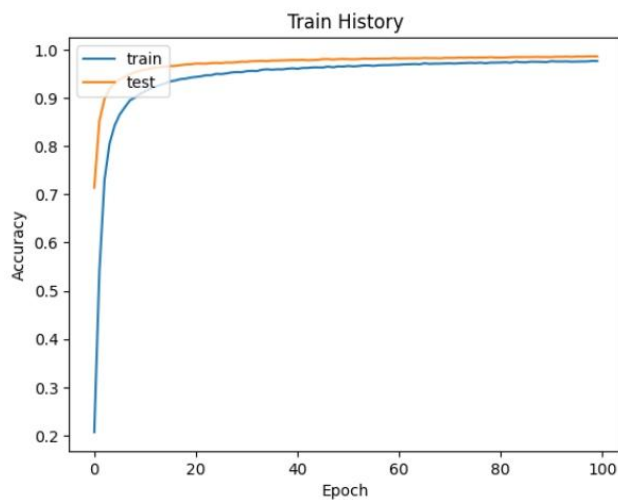
## 第二次

調整Optimizer, 改為SGD, 因為SGD沒有動量的概念, 所以花的時間會比較久, 且準確度沒有提升, 然而在圖中看起來還沒有overfitting, 因此第三次決定提高Epochs。



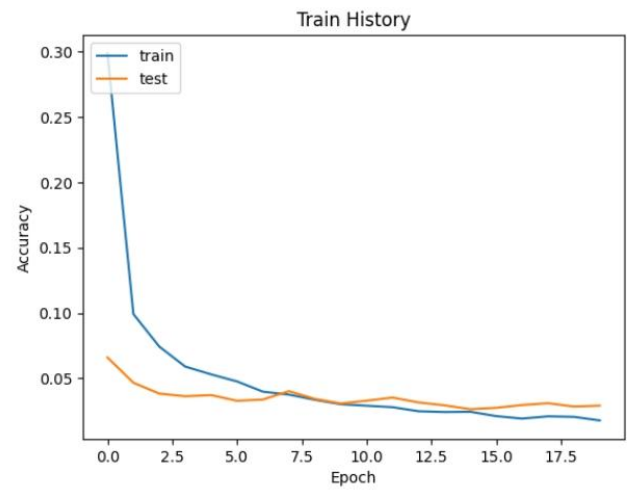
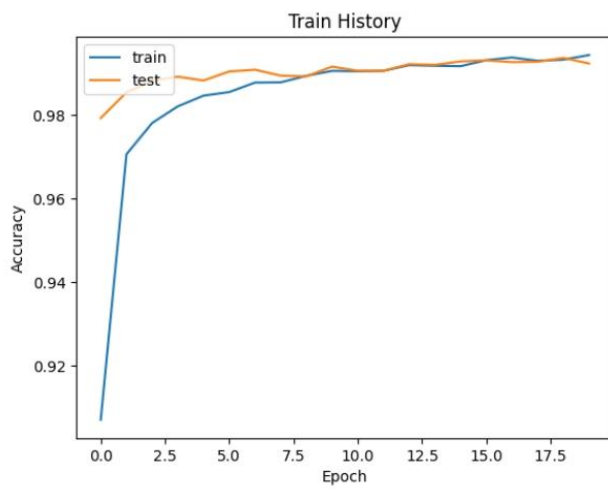
## 第三次

一樣選擇使用SGD並提升Epochs, 準確率確實較第二次有所提升, 但需要耗時的訓練也跟著增加。



#### 第四次

訓練，最終選擇改變Batch size，因為批次進入模型訓練的數量降低，所花的時間也相對高，而準確率沒有特別顯著的升高。



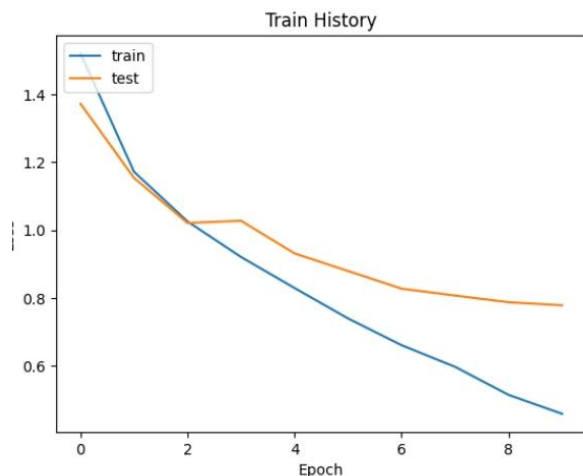
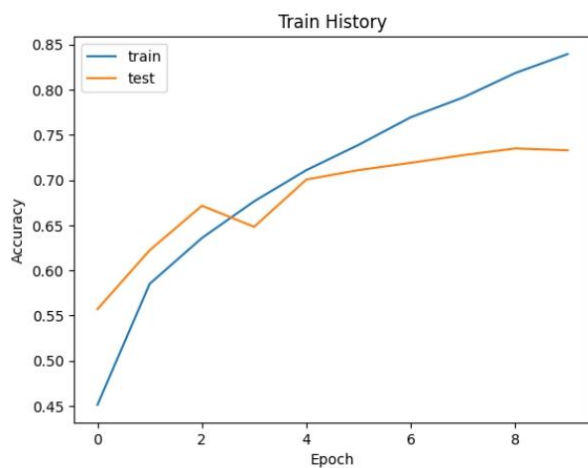
## Problem 2 - Image Recognition

### 1. Discuss those modifications affect its performances

Epochs	Batch Size	Optimizer	Loss function	Accuracy	Run time
10	128	Adam	categorical cross entropy	train: 0.8393 val: 0.7328 test: 0.7241	36 sec
50	128	Adam	categorical cross entropy	train: 0.9705 val: 0.7459 test: 0.7306	24 sec
50	256	Adam	categorical cross entropy	train: 0.9716 val: 0.7447 test: 0.749	131 sec
50	128	Adam	categorical cross entropy	train: 0.8723 val: 0.7762 test: 0.7698	209 sec
500	128	SGD	categorical cross entropy	train: 0.9143 val: 0.7978 test: 0.7884	1535 sec

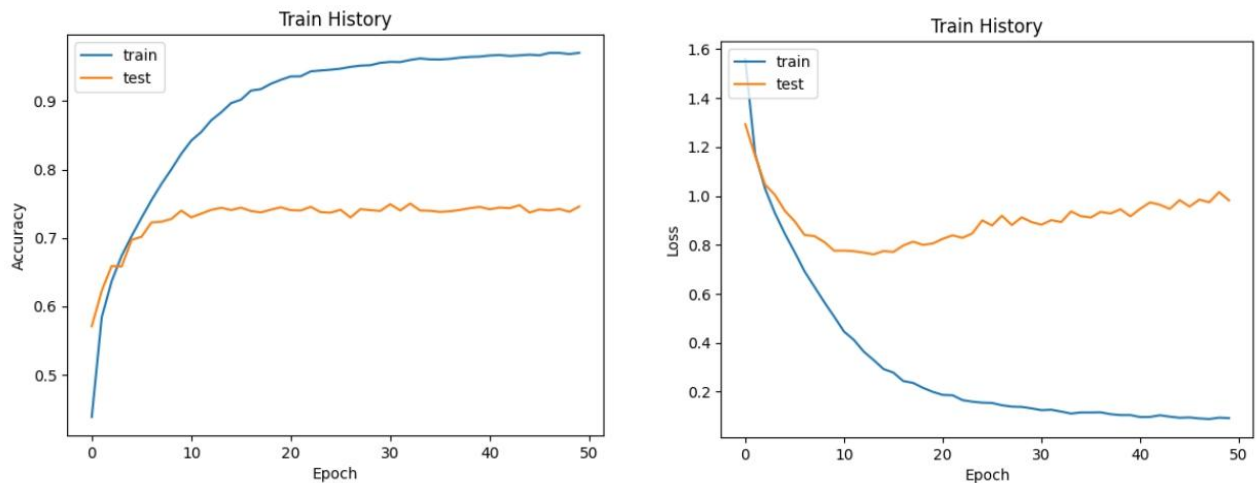
第一次

從第一次結果圖可以發現，無論是Accuracy或是Loss都還沒有收斂，因此選擇在下次增加epoch增加訓練次數，讓模型能完全擬合



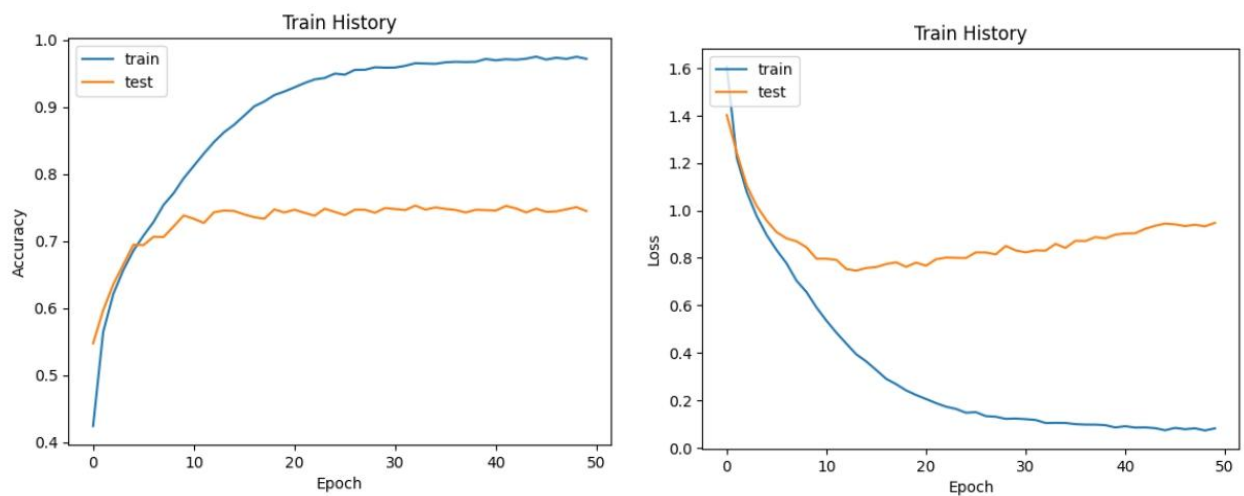
## 第二次

改變Epoch後，模型確實比先前更加擬合，但又可以發現Training的loss正常下降並降至低點，但Validation的loss卻上升，代表過度擬合(Overfitting)，因此選擇增加Batch size



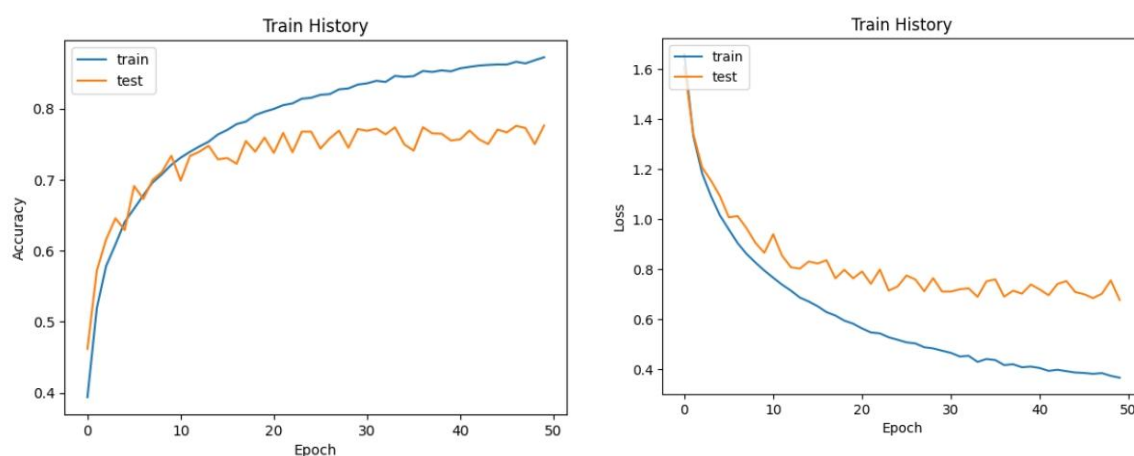
## 第三次

改變Batch size後，Overfitting 的狀況確實有些微改善，但仍然存在且結果也沒有提升，因此選擇增加層數，解決Overfitting問題



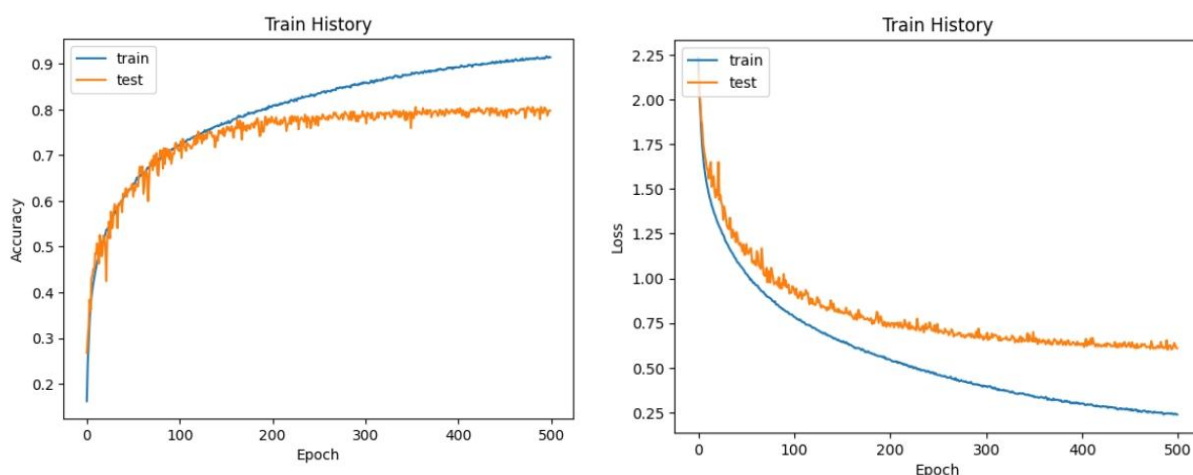
## 第四次

增加層數後(額外加入卷積層3與池化層3), 在Validation的時候loss沒有再出現下降後又上升的狀況, overfitting的問題確實被解決, 準確率也有所提升



## 第五次

改變Optimizer並增加Epoch, 選擇使用SGD, 雖然所花時間較多, 但在Testing的準確率確實增加, 選擇SGD缺點是所需迭代次數較多且耗時比較長



## 2. Explain the function of convolution layer, pooling layer and fully connected layer in CNN model.

Convolution layer透過指定一個尺寸的卷積核, 由上而下滑動來萃取圖像中的特徵。

Pooling layer 是一種非線性形式的降採樣方式, 用來降低參數量、降低維度, 同時又保留重要的特徵, 增加運算效率, 減少過度擬合的狀況。

Fully Connected Layer目的是為了實現分類, 經過前面兩層做出的提取特徵並減少參數後, 剩下來的特徵資訊輸入至fully connected layer做判斷屬於哪類。