

HW (2020/04/07)

曾宏鈞 06160485
徐友笙 05360365
蔡毓丞 06370136
盧君彥 05360153

本次作業主要目的是練習使用各種Optimizer以及Regularizer，並比較其效果。
Use the code, compare the prediction accuracy of the networks with/without :

Optimizer:	Regularizer:
<ol style="list-style-type: none">1. SGD2. RMSprop3. Adam	<ol style="list-style-type: none">1. L1, Norm Regularizer on weights2. L2 Norm Regularizer on weights3. Dropout4. Batch Normalization layers.5. Normal

總共12組組合，加上最基本的結構總共13組

=====

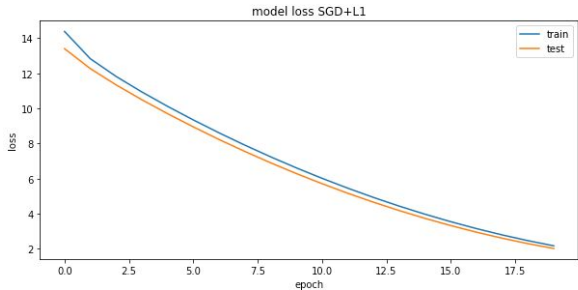
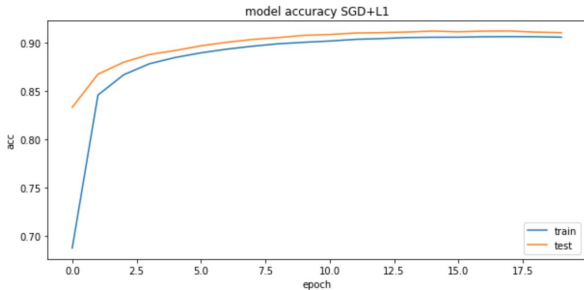
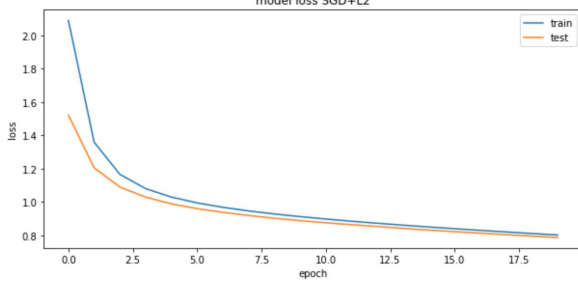
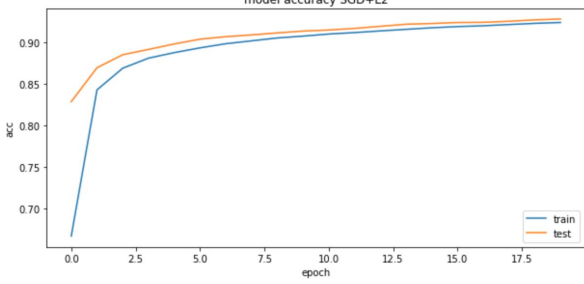
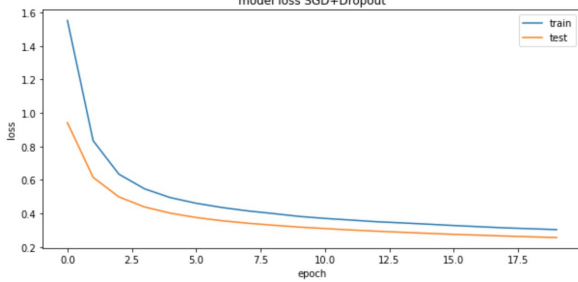
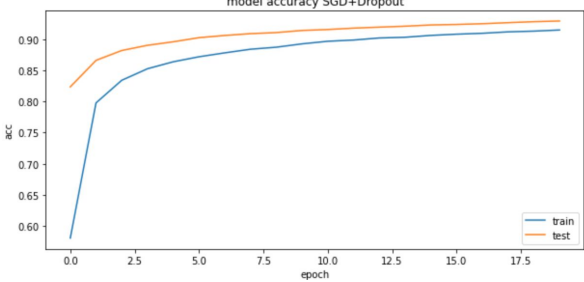
呈現報告內容:

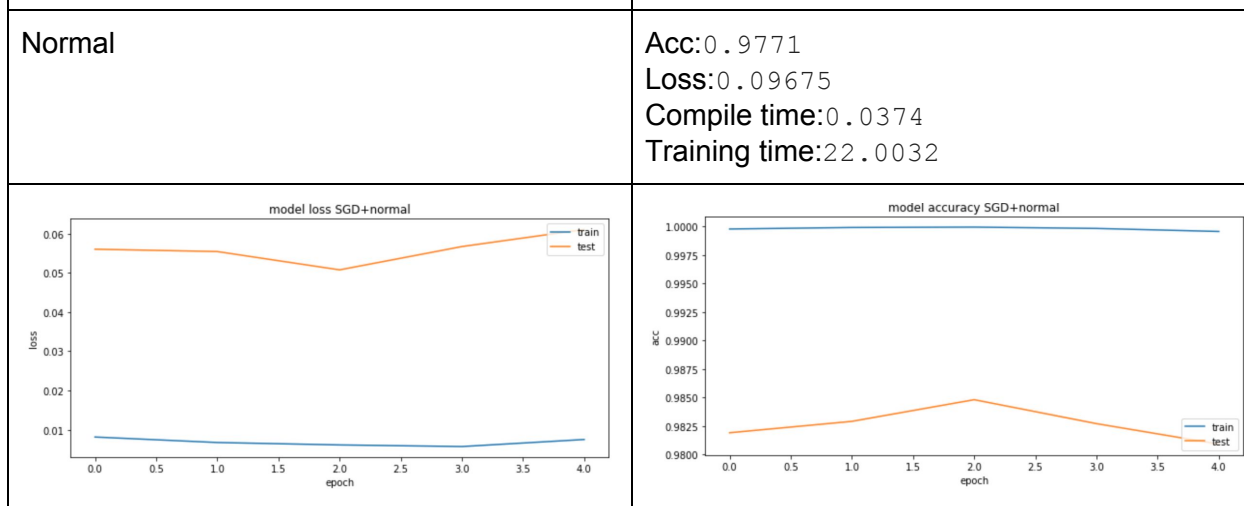
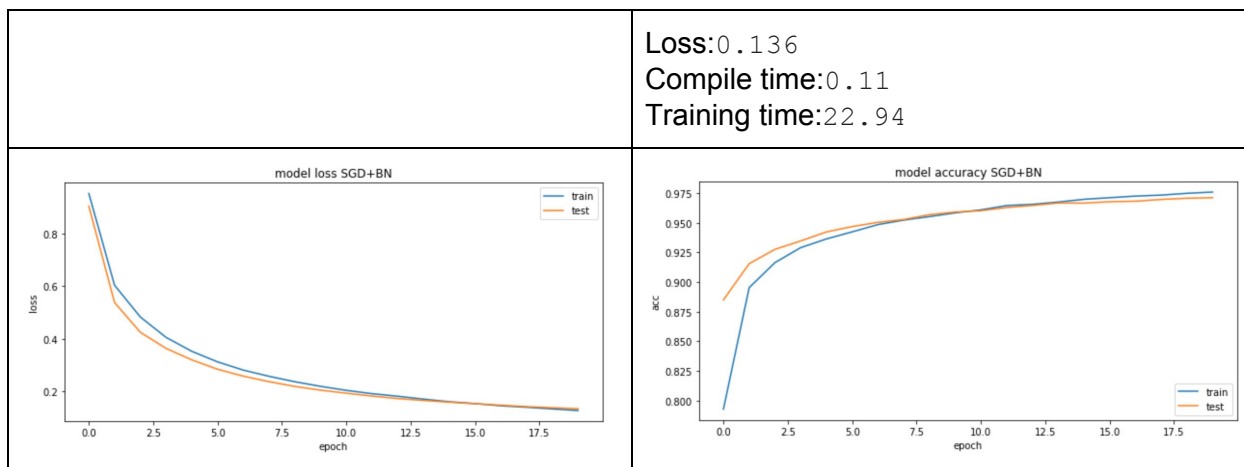
組合1: 準確率/截圖

實驗心得

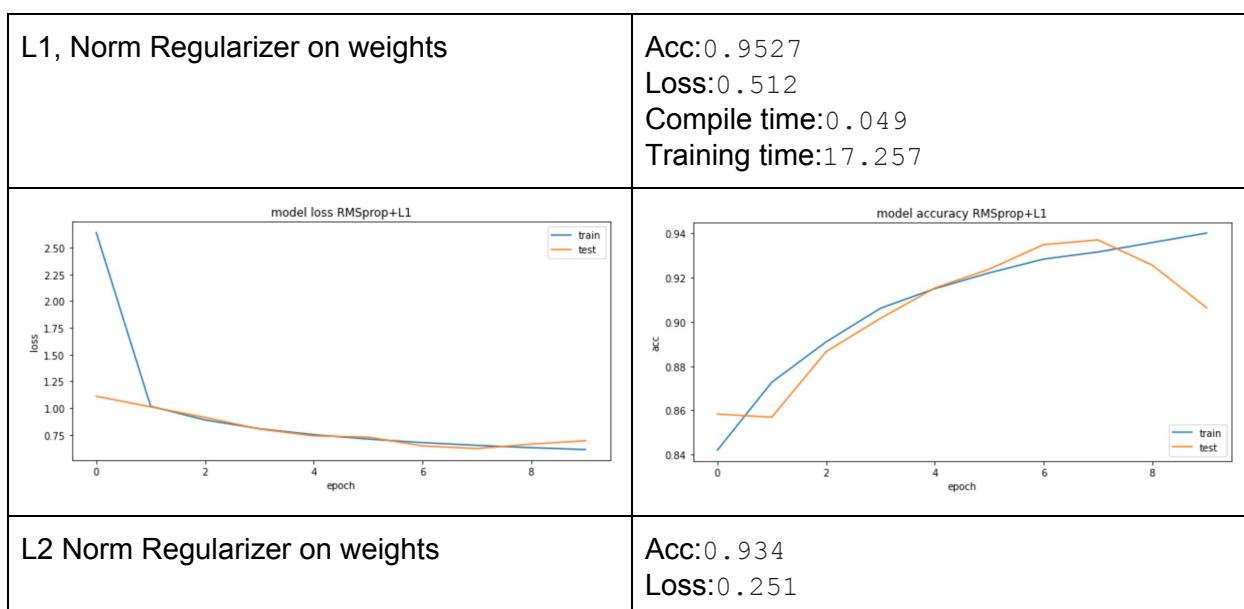
=====

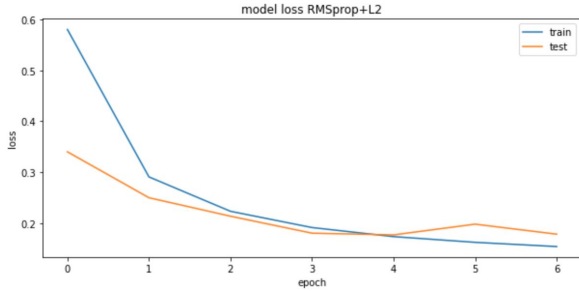
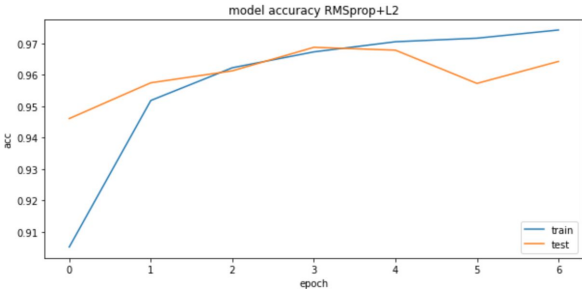
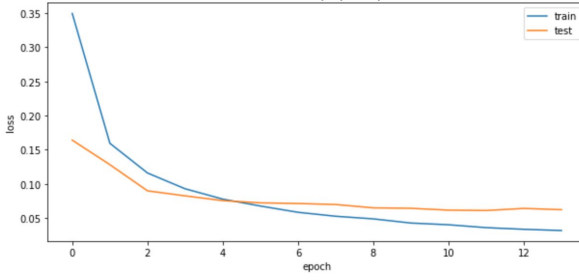
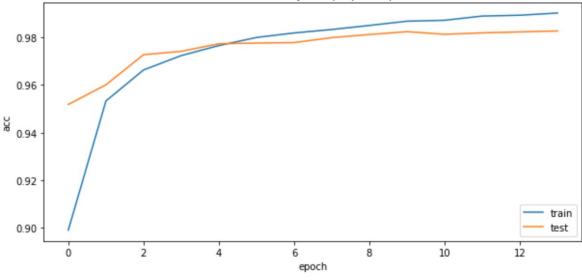
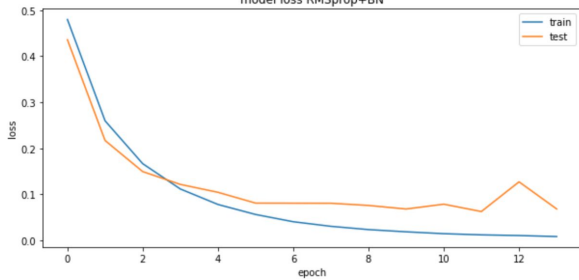
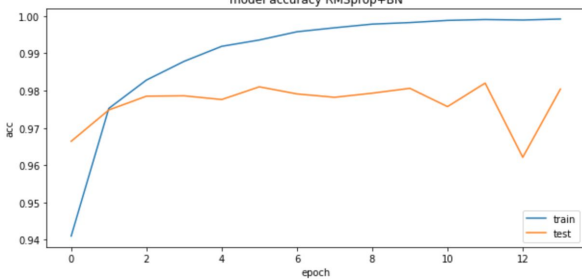
Optimizer:SGD

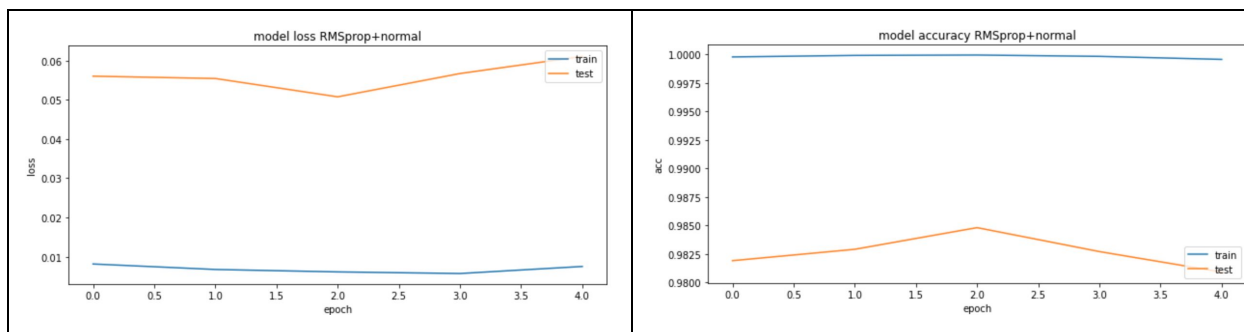
L1, Norm Regularizer on weights	Acc:0.909 Loss:2.023 Compile time:0.054 Training time:16.766
	
L2 Norm Regularizer on weights	Acc:0.9283 Loss:0.789 Compile time:0.044 Training time:16.352
	
Dropout	Acc:0.9308 Loss:0.2542 Compile time:0.046 Training time:16.856
	
Batch Normalization layers.	Acc:0.971



Optimizer:RMSprop

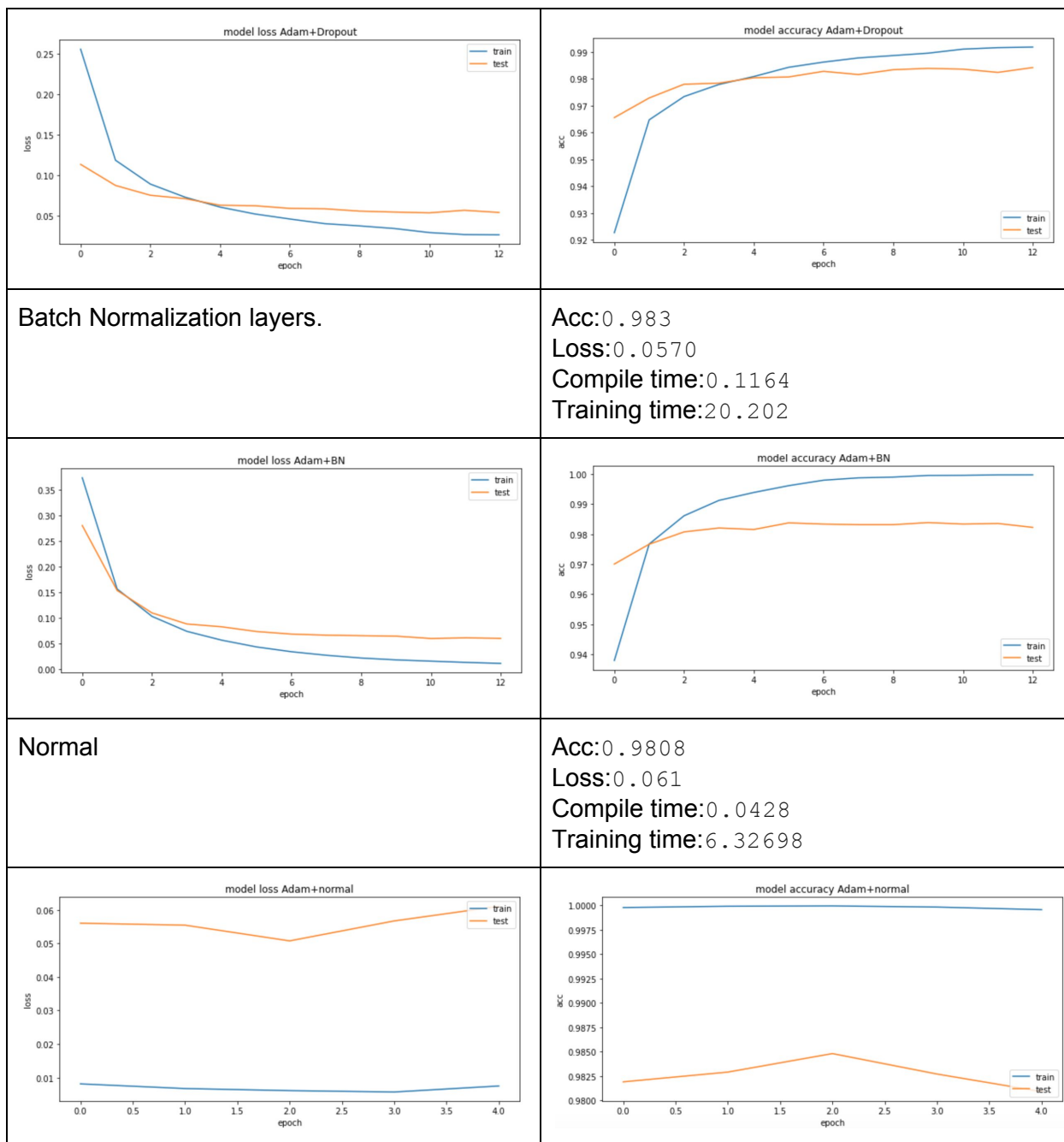


	<p>Compile time:0.045 Training time:7.284</p>
	
<p>Dropout</p>	<p>Acc:0.9815 Loss:0.0641 Compile time:0.048 Training time:10.931</p>
	
<p>Batch Normalization layers.</p>	<p>Acc:0.9808 Loss:0.0654 Compile time:0.121 Training time:13.864</p>
	
<p>Normal</p>	<p>Acc:0.9782 Loss:0.0738 Compile time:0.0377 Training time:4.9947</p>



Optimizer:Adam

L1, Norm Regularizer on weights	Acc:0.94809 Loss:0.3823 Compile time:0.047 Training time:14.838
<p>model loss Adam+L1</p> <p>loss</p> <p>epoch</p> <p>train</p> <p>test</p>	<p>model accuracy Adam+L1</p> <p>acc</p> <p>epoch</p> <p>train</p> <p>test</p>
L2 Norm Regularizer on weights	Acc:0.972 Loss:0.1537 Compile time:0.0484 Training time:9.640
<p>model loss Adam+L2</p> <p>loss</p> <p>epoch</p> <p>train</p> <p>test</p>	<p>model accuracy Adam+L2</p> <p>acc</p> <p>epoch</p> <p>train</p> <p>test</p>
Dropout	Acc:0.982 Loss:0.0542 Compile time:0.0505 Training time:14.238



實驗心得

這次的實驗主要在看看不同的Optimizer跟Regularizer對模型所造成的影響，以訓練20個epoch時間來看的話，Regularizer使用L2的時間相較其他都短，使用Batch Normalization所花的時間最多。但相對來說Batch Normalization會使模型準確度最後較高，dropout的結果也不錯，有可能是當沒有神經元鏈結時，減少了一些錯誤的權重加成，以不同的Optimizer來說，Adam與RMSprop準確度差不多。然而在各個Regularizer的部分都是L1最大，以SGD來說就高達2。

但觀察train及test的資料狀況，同樣20個epoch下，RMS很快就有over-fitting的問題，可能可以減少epoch數來避免模型過度訓練。

觀察正常的對照組normal反而訓練幾個epoch就有很好的效果，使用的時間也是裡面最少的，但因test已經高於train，因此減少epoch能獲得更好的準確率。