IIE Final project

以NSL進行臉部表情辨識

Presenter: 蕭詠仁
2020/12/24

Outline

1. 問題定義與介紹
2. 方法
3. 個案研究
4. 結果
5. 結論

問題定義與介紹

- 介紹:

- 現今人臉辨識已經被普遍的使用，也被大量的應用在生活當中。
- 目前人類情緒的辨識應用也越來越多，分類方法也不盡相同。
- 如醫學研究中可透過受測者表情來分辨當下的感受等。

問題定義與介紹

- 5W1H:

What	Where	When	Who	Why	How
利用人臉部表情進行分辨是否為快樂、悲傷、中性等7種表情。	人臉表情辨識可以運用在醫學研究、影像遊戲體驗及影片體驗評分等等。	需要判別人類情緒時。	醫學研究人員、遊戲設計人員與影片創作者等等。	透過臉部表情來識別使用者的體驗感受，針對辨識結果來進行所設計的環境或實驗是否要進行改善或調整。	利用NSL進行對抗性學習。

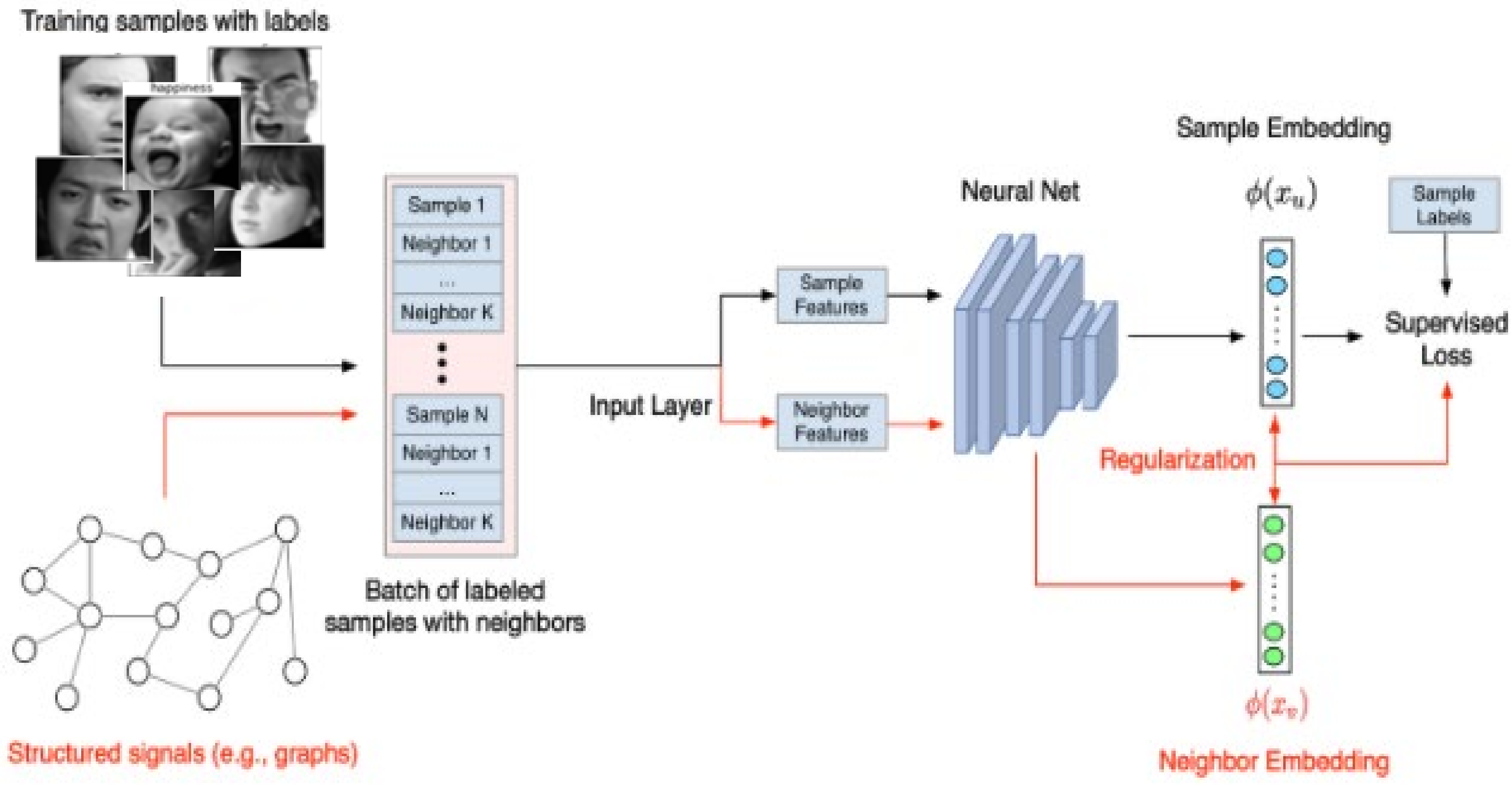
問題定義與介紹

- 目 的:

- 在過去人類情緒分類大多數研究利用Neural Network、SVM、AdaBoost等等[1]，這些方法為現在機器學習的主要架構。
- 因此本次專題主要動機為透過較新的神經網路架構進行對抗式學習來訓練人類表情分類，並期望達到較佳的分類結果。

方法

- 方法架構:

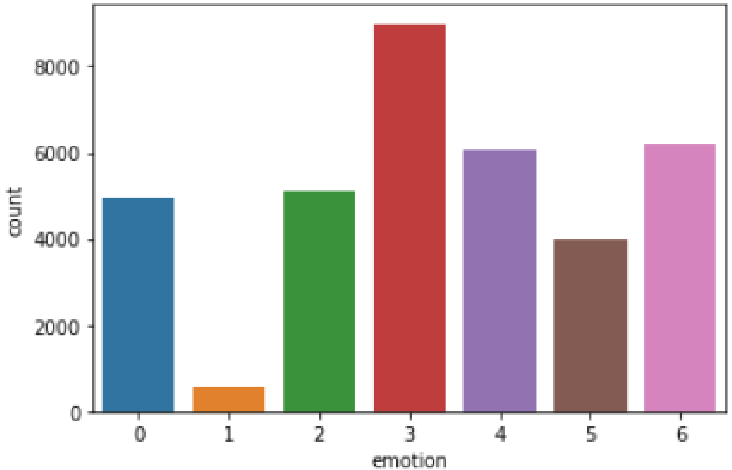


Training samples with labels
happiness
Sample 1
Neighbor 1
...
Neighbor K
Sample N
Neighbor 1
...
Neighbor K
Batch of labeled
samples with neighbors
Input Layer
Sample Features
Neighbor Features
Neural Net
Sample Embedding
$\phi(x_u)$
Sample Labels
Supervised Loss
Regularization
$\phi(x_v)$
Neighbor Embedding
Structured signals (e.g., graphs)

[2]

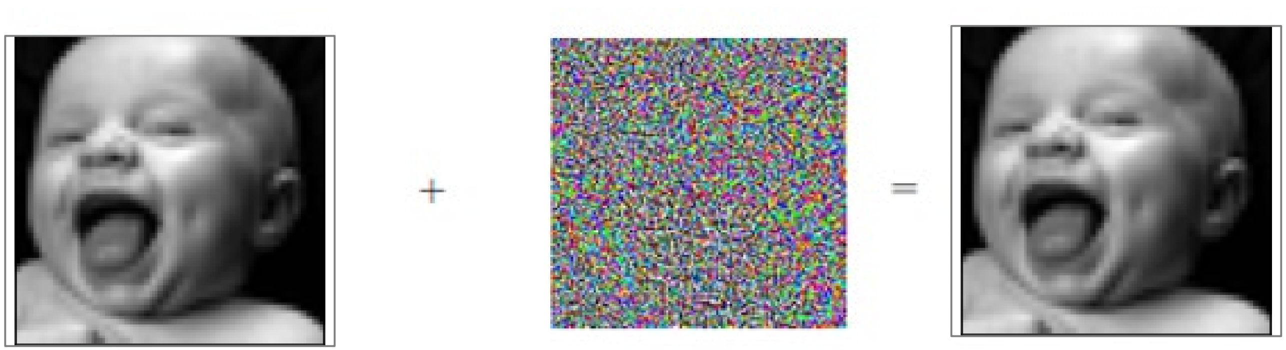
個案研究

- 數據:
- Data名稱:FER2013 (Facial Expression Recognition)
- 7種情緒:anger(0)、disgust(1)、fear(2)、happiness(3)、sadness(4)、surprise(5)、neutral(6)
- 共有35887張，Training80%(28709)、Validation10%(3589)、Test10%(3589)



個案研究

- Structured signals (Adversarial) :

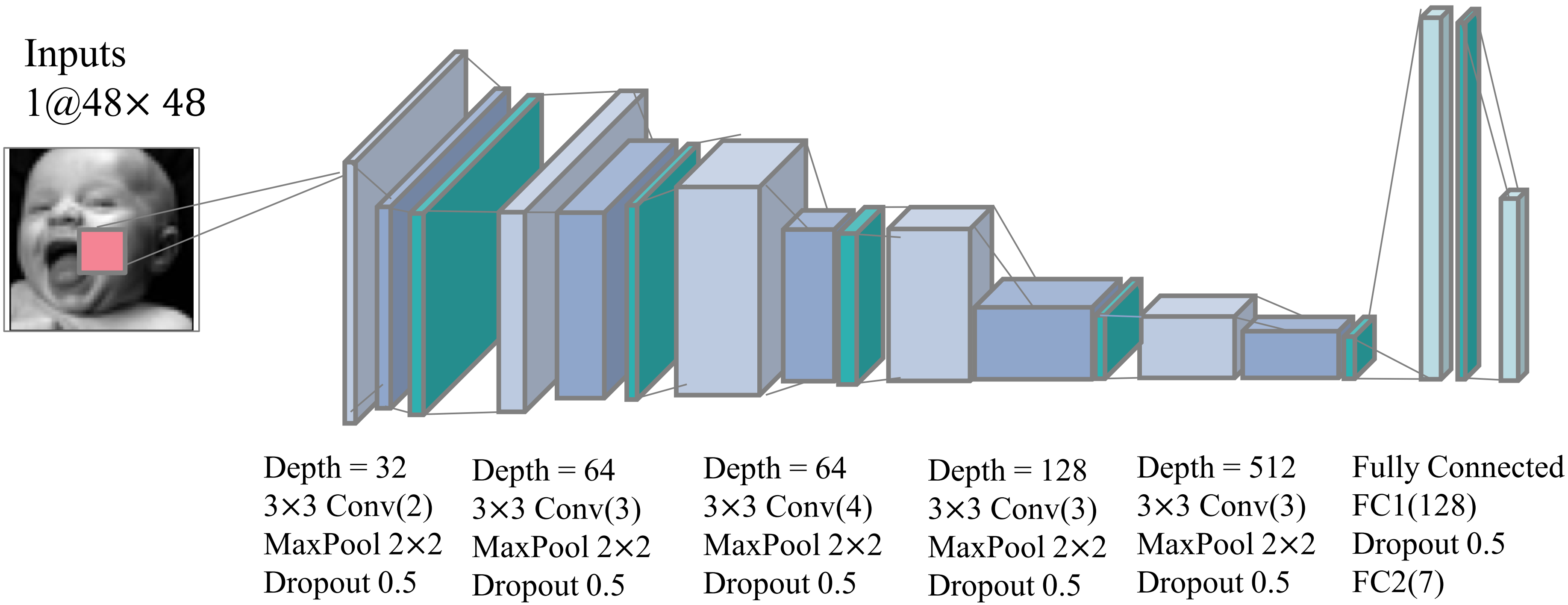


Happiness nematode Surprise

```
adv_config = nsl.configs.make_adv_reg_config(multiplier=0.3, adv_step_size=0.01)
```

個案研究

- CNN model:



個案研究

- Other function:

其他應用函數	名稱
各層kernel之初始化	he_normal
各層CNN之正規化	BatchNormalization
各層激活函數	elu
全連接層之激活函數	Softmax
優化函數	Adam(0.001)
自調整學習率	ReduceLROnPlateau
Loss	Categorical_crossentropy Adversarial loss

個案研究

- 資料前處理:

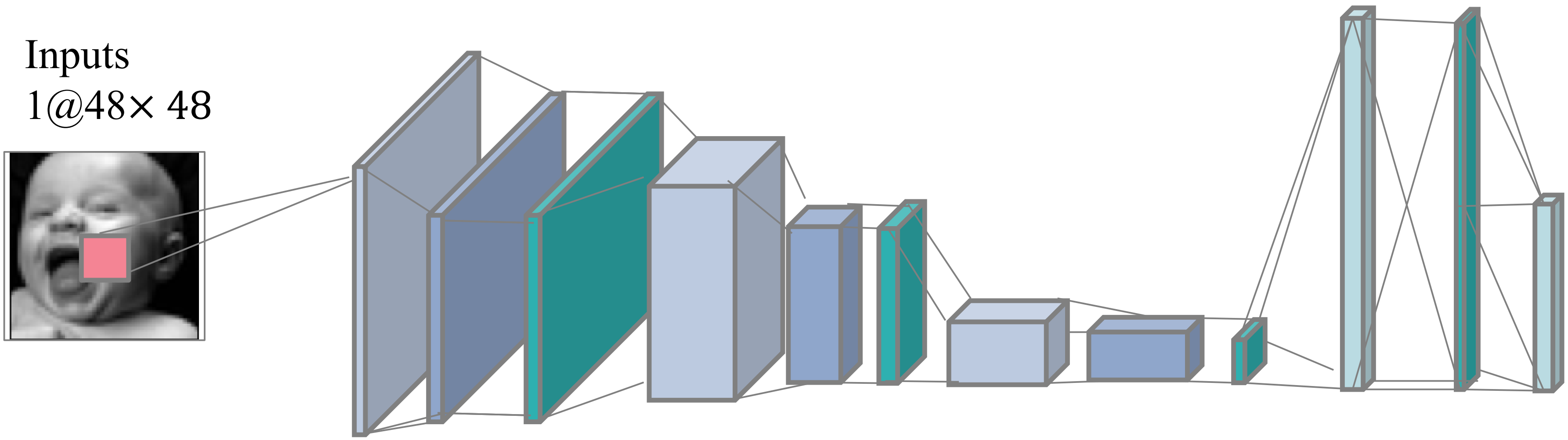
```
# Normalizing results, as neural networks are very sensitive to unnormalized data.
X_train = X_train / 255.
X_valid = X_valid / 255.
X_test = X_test / 255.
```

Data augmentation	參數設定
選轉	rotation_range=5
放大	zoom_range = 0.1
寬度改變	width_shift_range=0.1
高度改面	height_shift_range=0.1
垂直翻轉	vertical_flip=False
水平翻轉	horizontal_flip=False

個案研究

- 初步研究結果:

Inputs
1@48× 48



Depth = 64
5× 5 Conv(2)
MaxPool 2×2
Dropout 0.4

Depth = 128
3×3 Conv(3)
MaxPool 2×2
Dropout 0.4

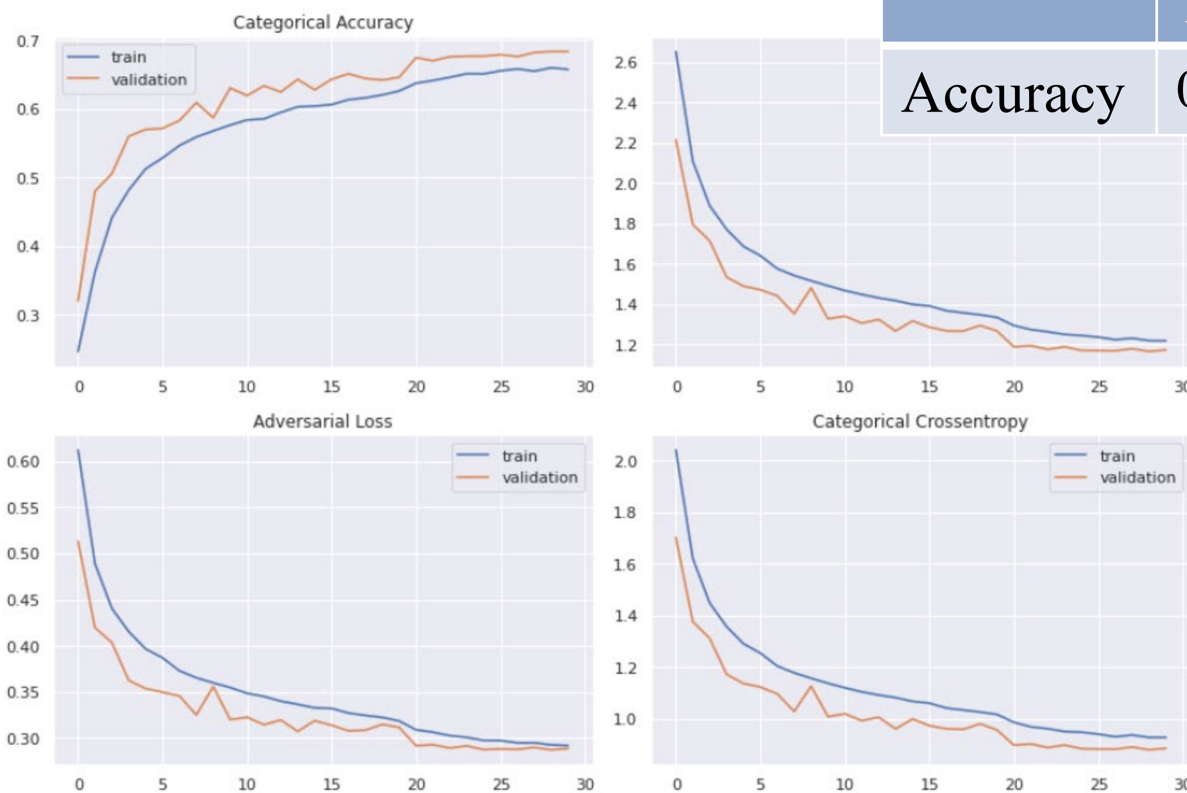
Depth = 256
3×3 Conv(4)
MaxPool 2×2
Dropout 0.5

Fully Connected
FC1(128)
Dropout 0.6
FC2(7)

個案研究

- 初步研究結果:

	Training	Validation
Accuracy	0.766	0.6882



個案研究-參數優化

調整部分	改變方式		Training	Validation
epochs	30→100	Accuracy	0.766 → 0.8035	0.6882 → 0.6980

調整部分	改變方式		Training	Validation
epochs	100→ 80	Accuracy	0.8035 → 0.7445	0.6980 → 0.6887
Batch size	50 → 42			
Conv layer	9 → 10			
Pool layer	3 → 5			
Dropout	0.4、0.5			
Augment	3 times (28691→ 86119)			

個案研究-參數優化

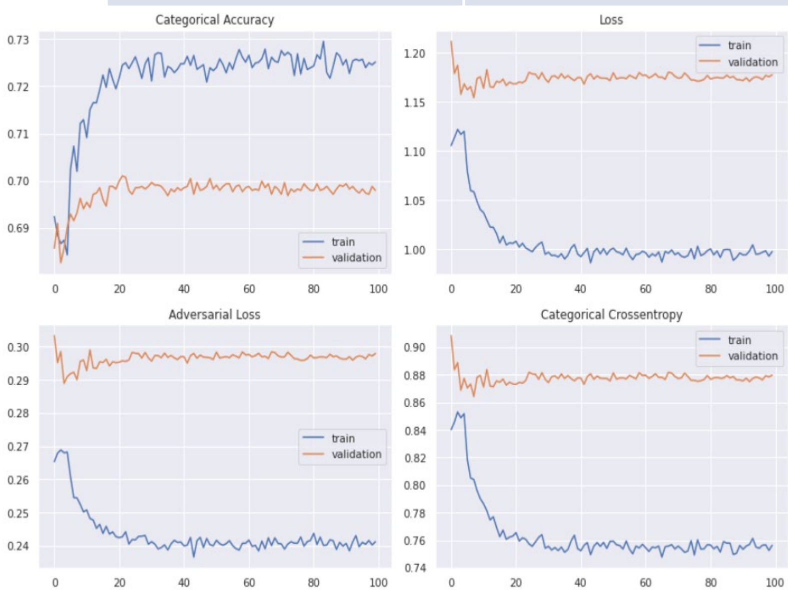
調整部分	改變方式		Training	Validation
epochs	80 → 50	Accuracy	0.7445 → 0.9468	0.6887 → 0.6992
Batch size	42 → 32			
Conv layer	10 → 15 (改變kernel size)			
Dropout	0.5			
Augment	86119 → 86139			

調整部分	改變方式		Training	Validation
epochs	50 → 32	Accuracy	0.9468 → 0.8868	0.6992 → 0.6982
Batch size	32 → 50			
Conv layer	15 → 15 (改變kernel size)			
Dropout	0.5			
Augment	86139 → 86141			

個案研究

- 參數優化:

調整部分	改變方式	原因	結果
epochs	100	從初步結果的學習圖來看，可以發現Model還有學習的空間，所以試著增加訓練步數	出現嚴重的over fitting的情況

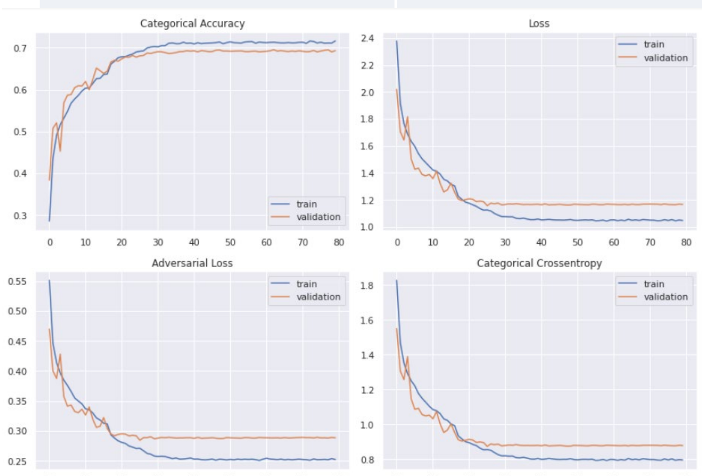


	Training	Validation
Accuracy	0.8035	0.6980

個案研究

- 參數優化:

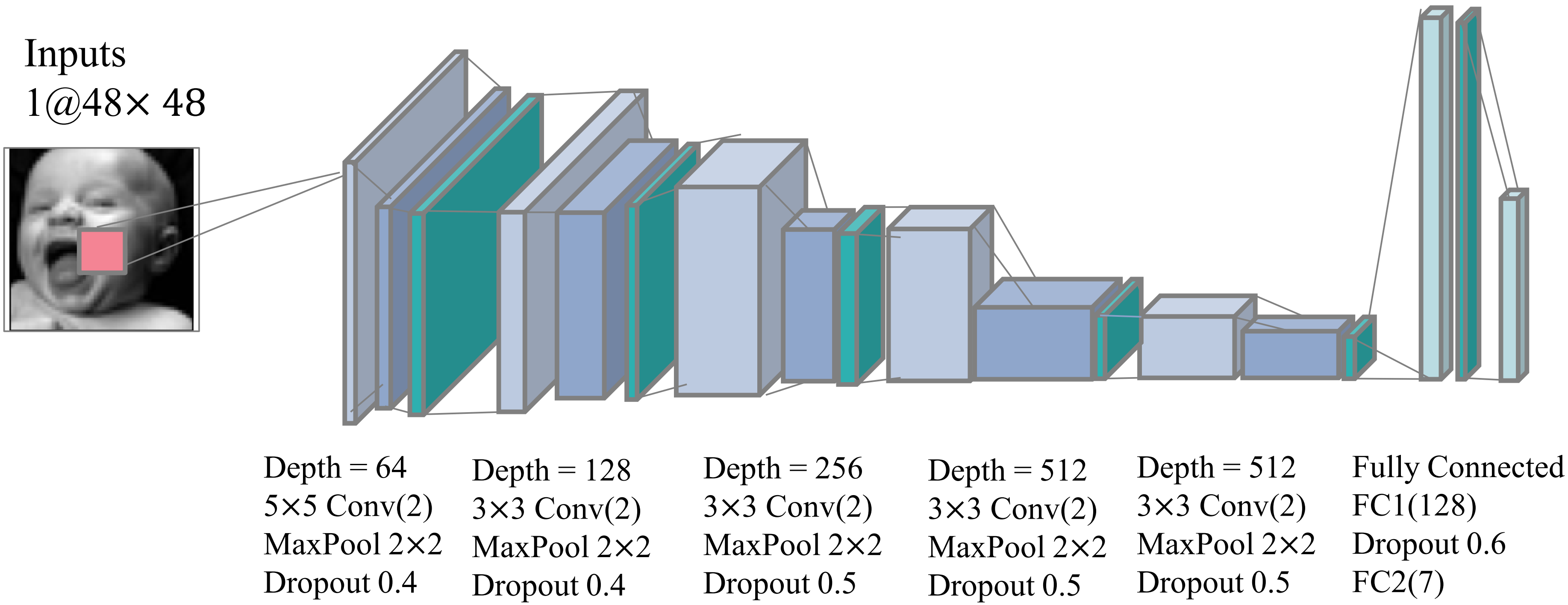
調整部分	改變方式	原因	結果
epochs	80	改善步數結果不佳 Over fitting情況	有改善嚴重over fitting 情況，但是訓練結果與第一次差不多
Model架構	加深		
增加資料增強數量	`steps_per_epoch=len(X_train)*4 / batch_size,`		



	Training	Validation
Accuracy	0.7445	0.6887

個案研究

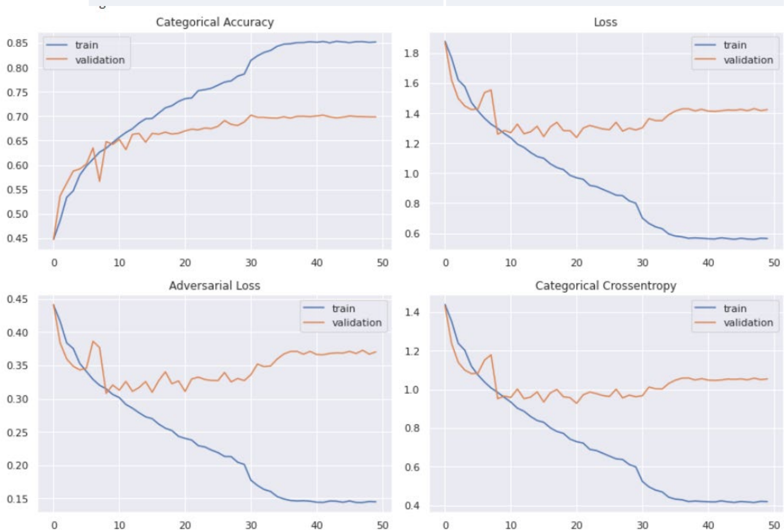
- CNN model:



個案研究

- 參數優化:

調整部分	改變方式	原因	結果
epochs	50	Model趨於穩定，進行最後微調整，但是仍有overfitting的情況	在訓練第20步開始出現較嚴重的over fitting，但是訓練結果有提升
Model架構	微調整		

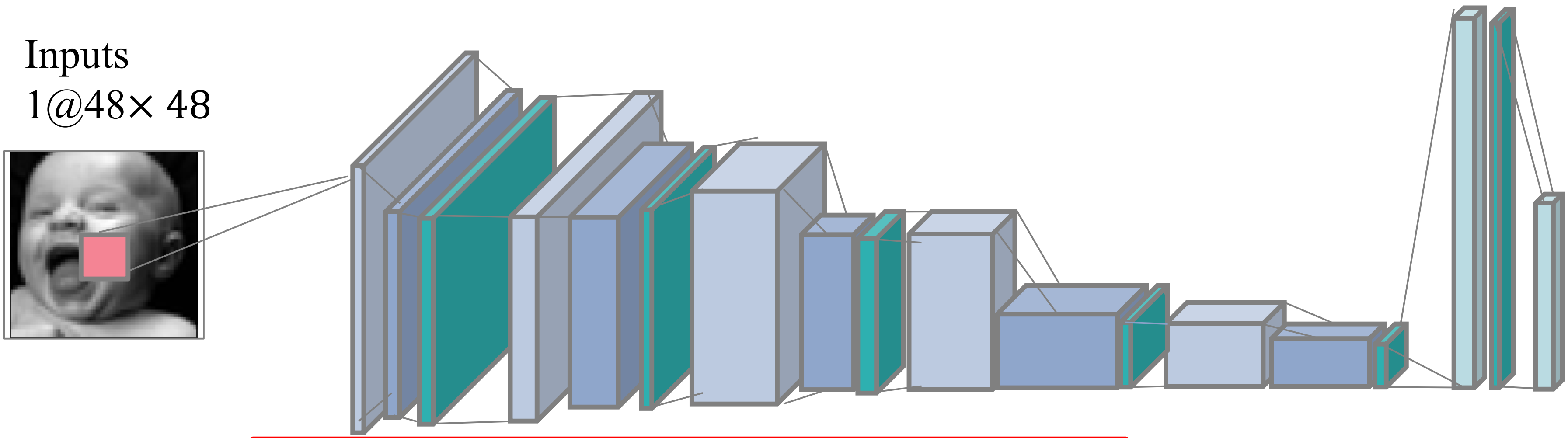


	Training	Validation
Accuracy	0.9468	0.6992

個案研究

- CNN model:

Inputs
1@48× 48

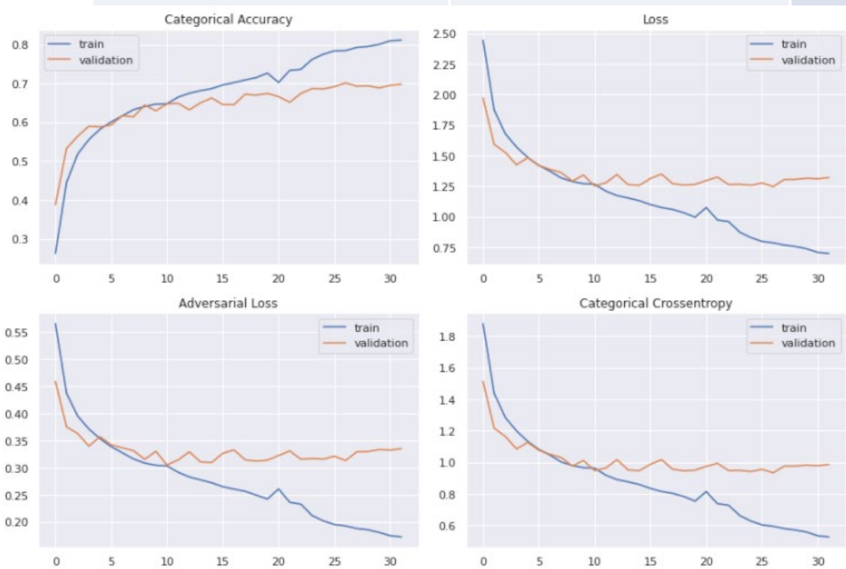


Depth = 64	Depth = 64	Depth = 128	Depth = 128	Depth = 512	Fully Connected
3× 3 Conv(2)	3×3 Conv(3)	3×3 Conv(4)	3×3 Conv(3)	3×3 Conv(1)	FC1(128)
MaxPool 2×2	MaxPool 2×2	MaxPool 2×2	MaxPool 2×2	2× 2 Conv(2)	Dropout 0.5
Dropout 0.5	Dropout 0.5	Dropout 0.5	Dropout 0.5	MaxPool 2×2	FC2(7)
				Dropout 0.5	

個案研究

- 參數優化:

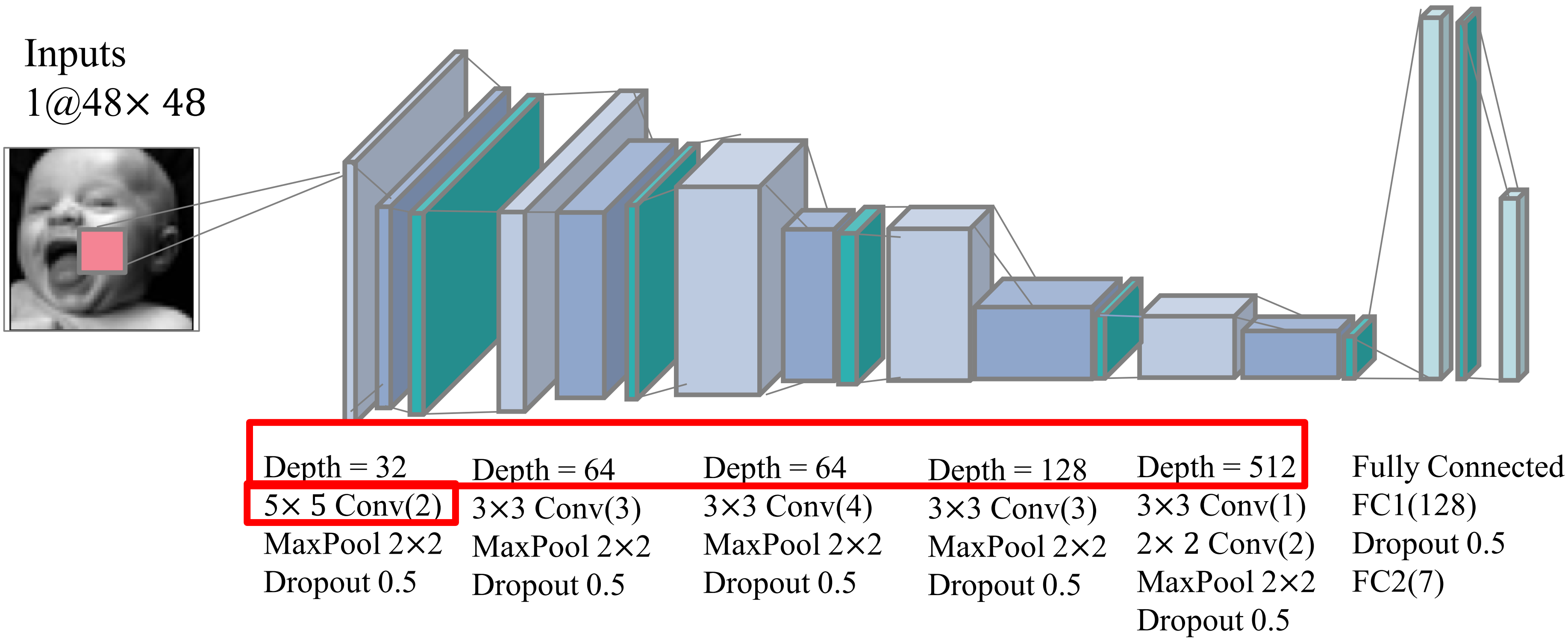
調整部分	改變方式	原因	結果
epochs	32	改善後的Model還是有Overfitting的情況發生	結果顯示有降低Overfitting的效果，但是在後面的訓練還是有過擬合的趨勢
Model架構	微調整		



	Training	Validation
Accuracy	0.8868	0.6982

個案研究

- CNN model:

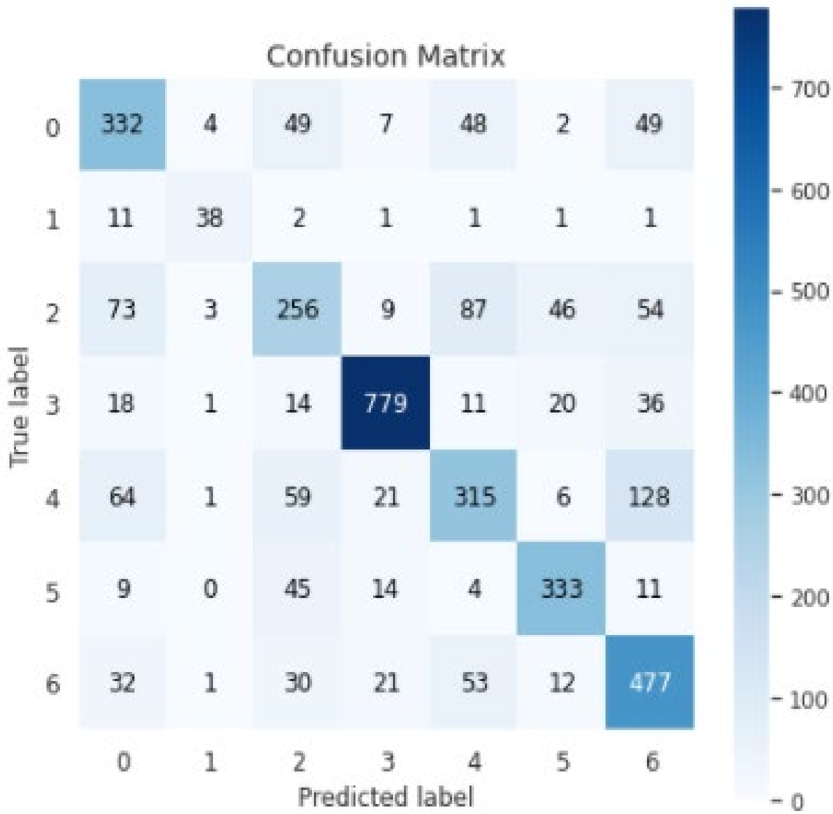


結果

- 結果:

```
total wrong validation predictions: 1059
accuracy: 70.493%
```

	precision	recall	f1-score	support
0	0.62	0.68	0.64	491
1	0.79	0.69	0.74	55
2	0.56	0.48	0.52	528
3	0.91	0.89	0.90	879
4	0.61	0.53	0.57	594
5	0.79	0.80	0.80	416
6	0.63	0.76	0.69	626
accuracy			0.70	3589
macro avg	0.70	0.69	0.69	3589
weighted avg	0.71	0.70	0.70	3589



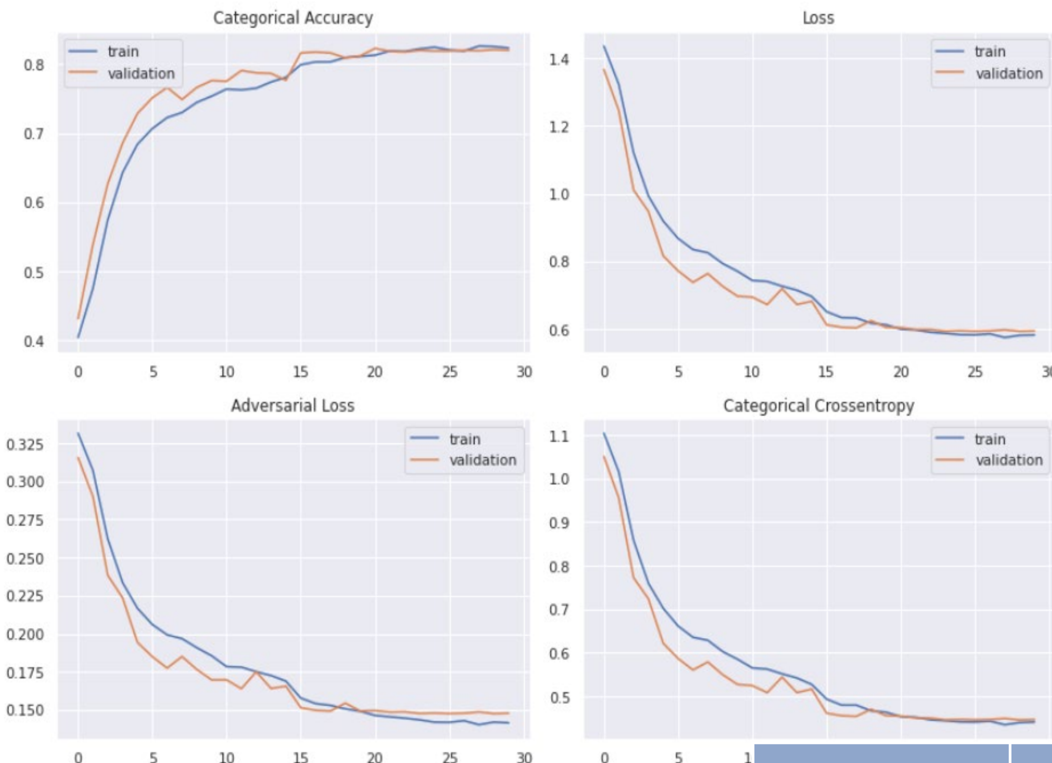
結果

- 結果:

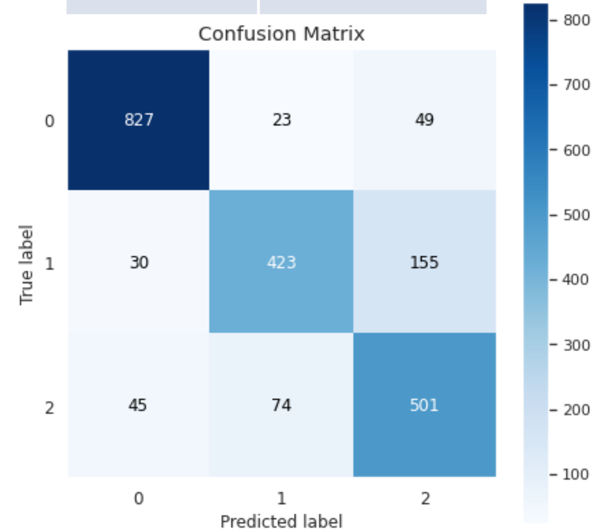
Algorithm	Test accuracy	Hybrid Model	Other Function	Pre-Trained
Zhang[3]	64.15%			✓
Ionescu, R. T., Popescu, M., & Grozea, C. [4]	67.484%			✓
Gao& Ma [5]	65.20%			✓
Raksarikorn, T., & Kangkachit, T [6]	71.69%		✓	✓
Sheng, et al [7]	73.28%	✓	✓	
Proposed method	70.493%			

結果

- 結果:



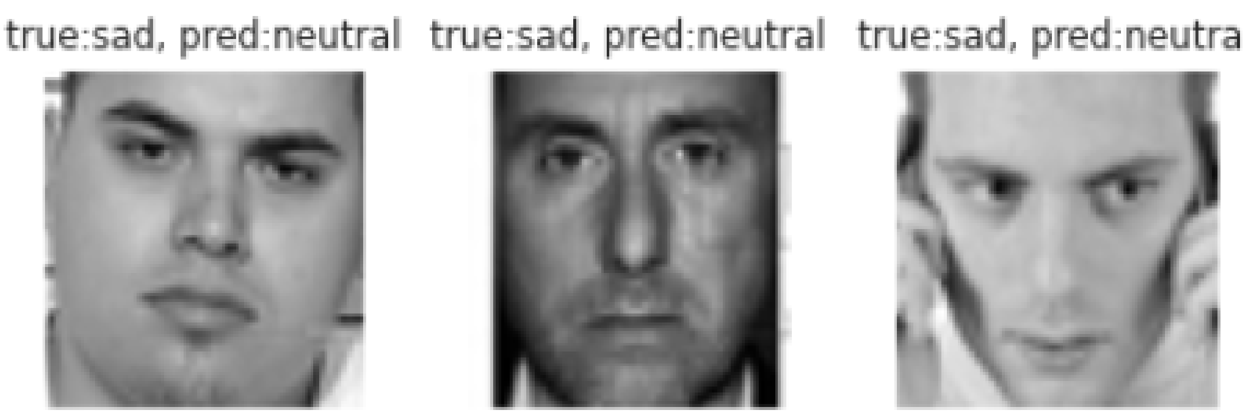
類別	情緒
3	happiness
4	sadness
6	neutral



	Training	Validation	Test
Accuracy	0.8551	0.8295	0.8305

結論

- 結果:
  - 模型架構及參數設定大多數參考網路上以訓練好的CNN模型，以vgg16當作模板來進行架構改變
  - 但是結果卻沒有想像中來的好，還有待加強
  - 在未來可能可以增加資料訓練張數或是利用GAN等方式，來加強辨識率

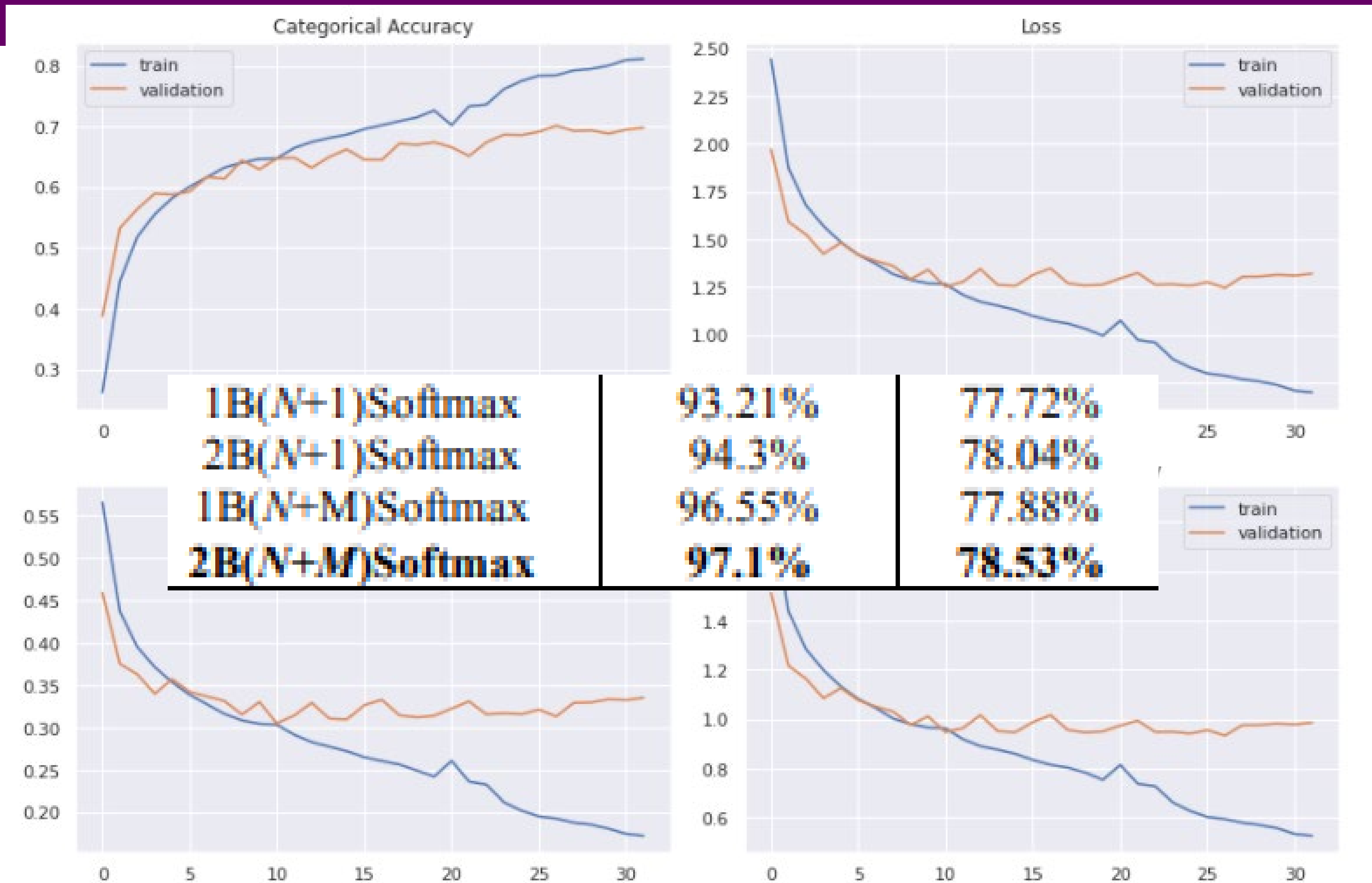


參考資料

- 參考資料:

- [1]http://nccur.lib.nccu.edu.tw/bitstream/140.119/37112/9/100309.pdf
- [2] https://blog.tensorflow.org/2019/09/introducing-neural-structured-learning.html
- [3] Zhang, J. (2020, October). Movies and Pop Songs Recommendation System by Emotion Detection through Facial Recognition. In *Journal of Physics: Conference Series* (Vol. 1650, No. 3, p. 032076). IOP Publishing
- [4] Ionescu, R. T., Popescu, M., & Grozea, C. (2013, June). Local learning to improve bag of visual words model for facial expression recognition. In Workshop on challenges in representation learning, ICML.
- [5] Gao, H., & Ma, B. A Robust Improved Network for Facial Expression Recognition.
- [6] Raksarikorn, T., & Kangkachit, T. (2018, July). Facial Expression Classification using Deep Extreme Inception Networks. In 2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE) (pp. 1-5). IEEE.
- [7] Sheng, M., Zhang, L., Yan, L., Wang, C., Li, M., Xia, H., & Zhang, Y. (2020, August). Facial Expression Recognition Based on Sparse Autoencoder and Shallow Convolutional Neural Network. In 2020 15th International Conference on Computer Science & Education (ICCSE) (pp. 669-674). IEEE.

Thank You for Your Listening



1B(*N*+1)Softmax	93.21%	77.72%
2B(*N*+1)Softmax	94.3%	78.04%
1B(*N*+M)Softmax	96.55%	77.88%
2B(*N*+*M*)Softmax	**97.1%**	**78.53%**