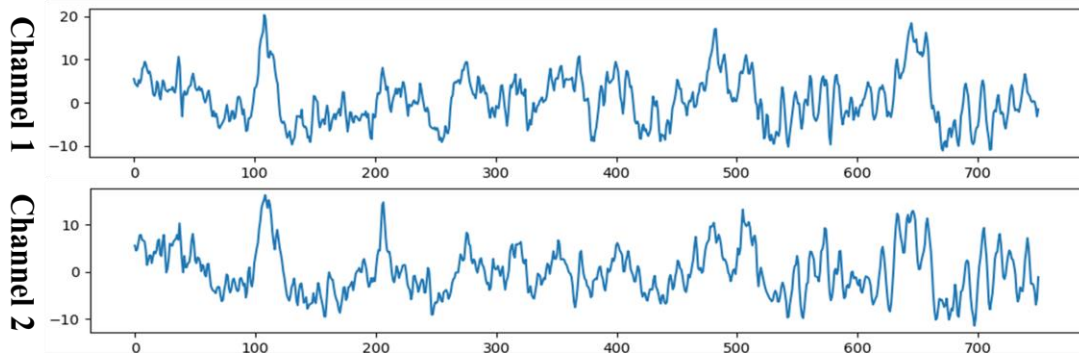


LAB 3 EEG Classification

謝宇恆 411551022

1. Introduction

本實驗藉由 EEGNet 以及 DeepConvNet 解決資料集:BCI competition 的分類問題。



2. Experiment Setup

a) EEGNet models:

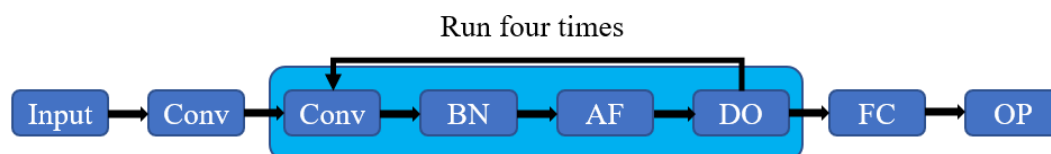
```
EEGNet(  
  (firstconv): Sequential(  
    (0): Conv2d(1, 16, kernel_size=(1, 51), stride=(1, 1), padding=(0, 25), bias=False)  
    (1): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)  
  )  
  (depthwiseConv): Sequential(  
    (0): Conv2d(16, 32, kernel_size=(2, 1), stride=(1, 1), groups=16, bias=False)  
    (1): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)  
    (2): ELU(alpha=1.0)  
    (3): AvgPool2d(kernel_size=(1, 4), stride=(1, 4), padding=0)  
    (4): Dropout(p=0.25)  
  )  
  (separableConv): Sequential(  
    (0): Conv2d(32, 32, kernel_size=(1, 15), stride=(1, 1), padding=(0, 7), bias=False)  
    (1): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)  
    (2): ELU(alpha=1.0)  
    (3): AvgPool2d(kernel_size=(1, 8), stride=(1, 8), padding=0)  
    (4): Dropout(p=0.25)  
  )  
  (classify): Sequential(  
    (0): Linear(in_features=736, out_features=2, bias=True)  
  )  
)
```

在此模型中，我們使用了 depthwise-separable convolution。depthwise-separable convolution 是 2017 年由 google 團隊提出的 MobileNet model [1]。隨著 DL 的快速發展，DL 模型變得深且巨大，該模型為了在維持其 performance 的情況下，同時降低其 convolution 的計算量。

b) DeepConvNet models:

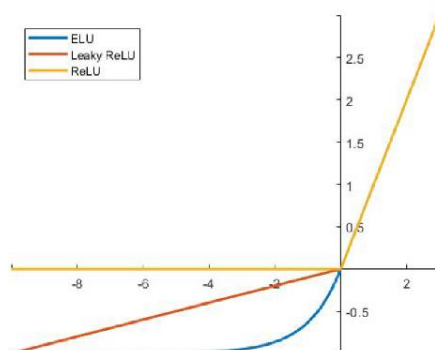
Layer	# filters	size	# params	Activation	Options
Input		(C, T)			
Reshape		(1, C, T)			
Conv2D	25	(1, 5)	150	Linear	mode = valid, max norm = 2
Conv2D	25	(C, 1)	$25 * 25 * C + 25$	Linear	mode = valid, max norm = 2
BatchNorm			$2 * 25$		epsilon = 1e-05, momentum = 0.1
Activation				ELU	
MaxPool2D		(1, 2)			
Dropout					p = 0.5
Conv2D	50	(1, 5)	$25 * 50 * C + 50$	Linear	mode = valid, max norm = 2
BatchNorm			$2 * 50$		epsilon = 1e-05, momentum = 0.1
Activation				ELU	
MaxPool2D		(1, 2)			
Dropout					p = 0.5
Conv2D	100	(1, 5)	$50 * 100 * C + 100$	Linear	mode = valid, max norm = 2
BatchNorm			$2 * 100$		epsilon = 1e-05, momentum = 0.1
Activation				ELU	
MaxPool2D		(1, 2)			
Dropout					p = 0.5
Conv2D	200	(1, 5)	$100 * 200 * C + 200$	Linear	mode = valid, max norm = 2
BatchNorm			$2 * 200$		epsilon = 1e-05, momentum = 0.1
Activation				ELU	
MaxPool2D		(1, 2)			
Dropout					p = 0.5
Flatten					
Dense	N			softmax	max norm = 0.5

此模型為傳統的 CNN 架構，網路的架構如下圖。Conv 為 Convolution Layer，BN 為 BatchNormalized Layer，AF 為 Activation Function，DO 為 Drop Out，FC 為 Fully connectted Layer，最後即為 Output。



c) Activation Function:

此實驗我們使用了三種不同的 Activation Function，分別是 ReLU、LeakyReLU 和 ELU。下圖為三種 Activation Function 會產生的輸出。



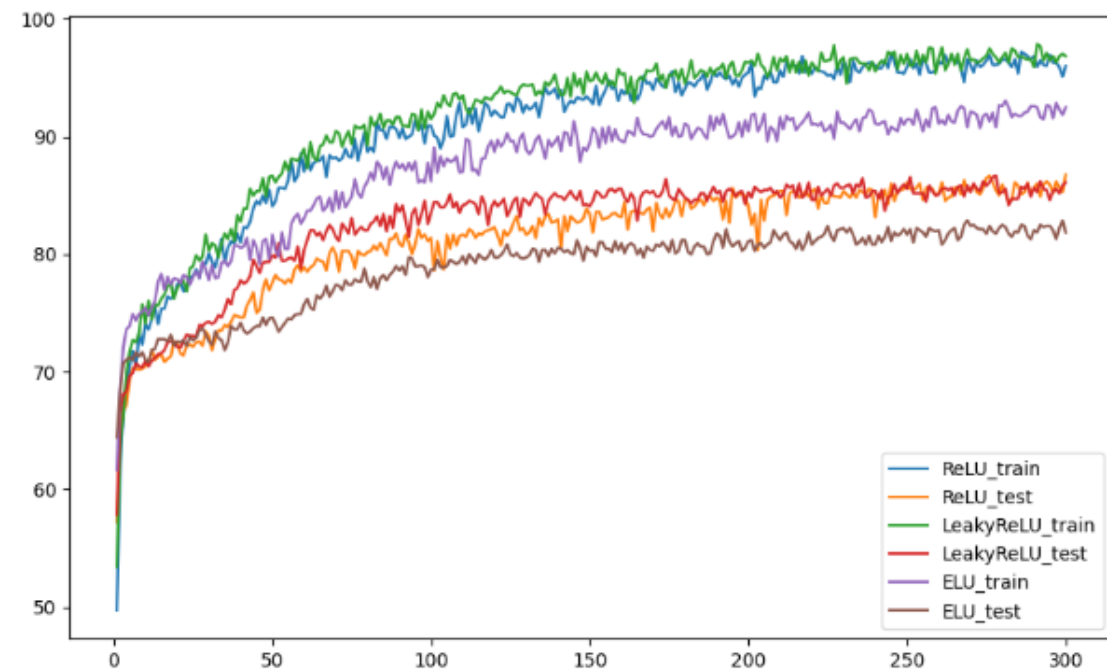
這三種 Activation Function 都有不同的優缺點，ReLU 的運算速度較快，但是其會有 dying ReLU Problem，因為其在 $X < 0$ 時梯度即為 0，因此無法調整權重。Leaky Relu 具有 Relu 的特點，但 $X < 0$ 時梯度不為 0，因此可以解決 dying ReLU Problem，但是其在 $X < 0$ 時為線性因此也不能在複雜的分類中使用。ELU 在 $X \geq 0$ 時與 Leaky ReLU 還有 ReLU 可以產生相同的輸出，在 $X < 0$ 時，則是會形成一個曲線，因此奇的分類效果應比其他兩個還好。

3. Experimental Results:

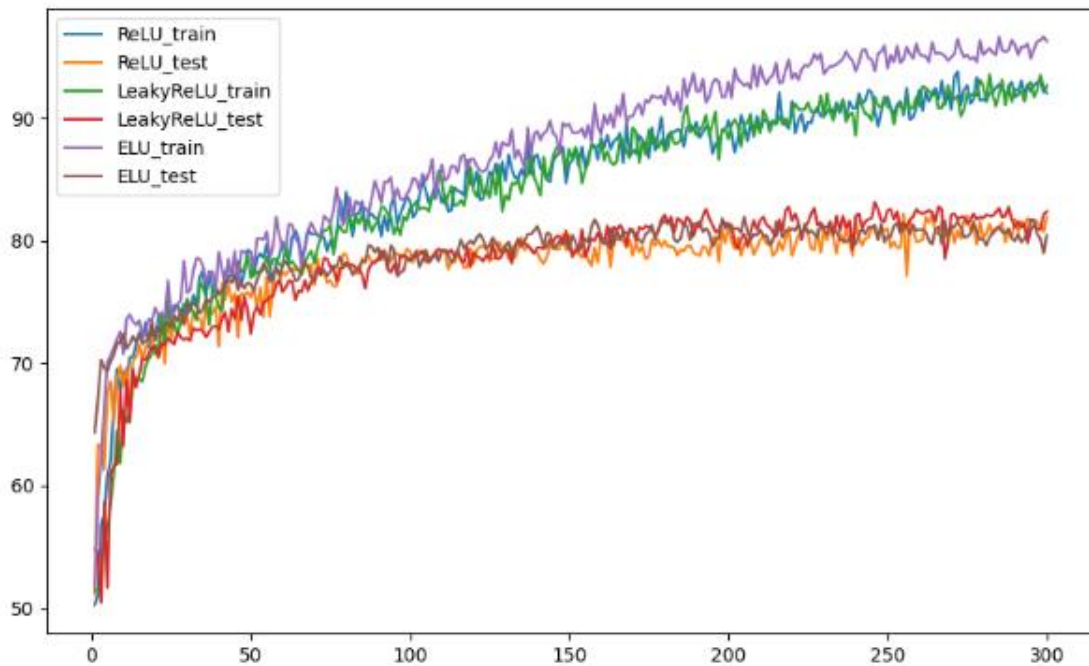
在此實驗中，Epoch 皆設定為 300，Accuracy 都可達到 82% 以上。

	ReLU	Leaky ReLU	ELU
EEGNet	87%	87%	83%
DeepConvNet	82%	83%	82%

EEGNet:



DeepConvNet:



4. Discuss:

- a) Activation function 中的 ELU 有一個設定值 α ，default 為 1，可以用來設定其飽和點。ELU 的式子計算如下：
$$\text{ELU} = \begin{cases} x, & x \geq 0 \\ \alpha \times (e^x - 1), & x < 0 \end{cases}$$
。我將不同的 α 丟入 ELU 去訓練此模型，但是得出的 Accuracy 皆約等於 82%，在此分類模型中， α 的影響力不大，但是在不同的分類器中可能會有不同的影響，因此其為可以探討的一個問題。

Reference:

- [1] Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." *arXiv preprint arXiv:1704.04861* (2017).