

Deep Learning (Homework 2)

0753420 郭家瑄

I. Convolutional Neural Network for Image Recognition

i. Data preprocessing

1. Preprocessing detail

將每張照片重新 resize 成 $w=100, h=100$ 的圖片。

2. Reason

檢視資料集後，發現每張圖片的 w, h 不盡相同，除此之外，資料集當中每張圖片的品質皆不同，有些圖片主體位在整張圖片的左上角，若是以 crop 裁切圖片，則必須判斷圖片中物體在哪裡然後再去切圖，過程過於繁瑣，成效不一定很好，因此選擇直接 resize 圖片，雖然有些圖片會出現變形狀況，但整體而言，經過 resize 的圖片中物體形狀還是具有一定程度的鑑別度，基本上不影響辨識。

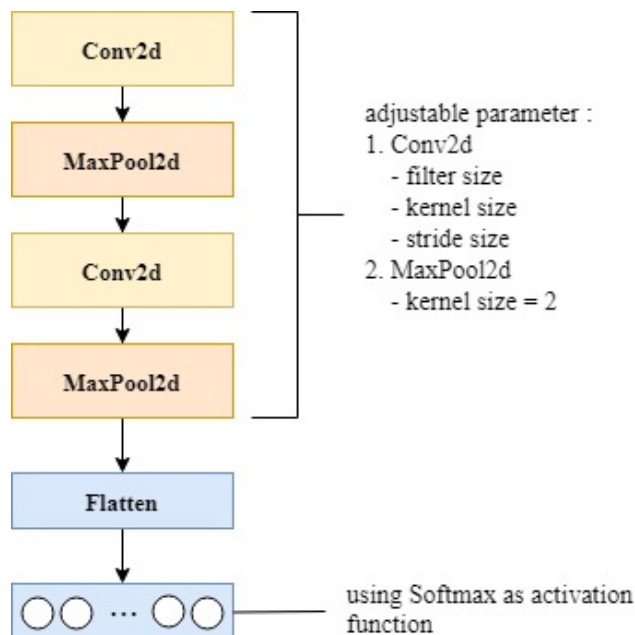
3. Code submission

(資料前處理程式碼位在 HW2_0753420_郭家瑄.py 中 class utilsTool 裡面的 dataPreprocessing() 函數)

ii. Model detail and learning figure

1. Network structure

針對 Model 的設計架構如 Figure 1 所示：

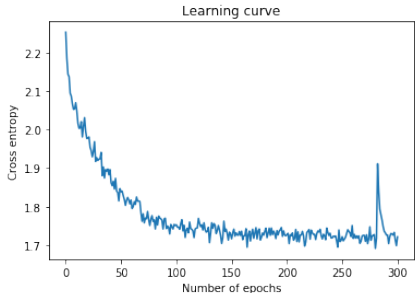
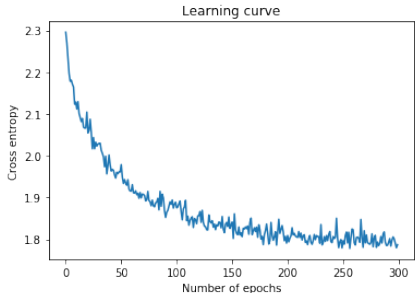
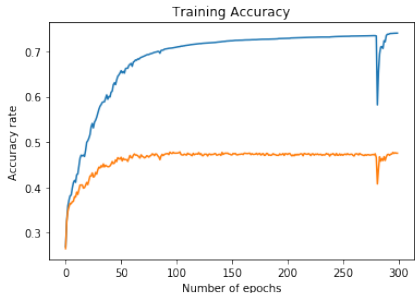
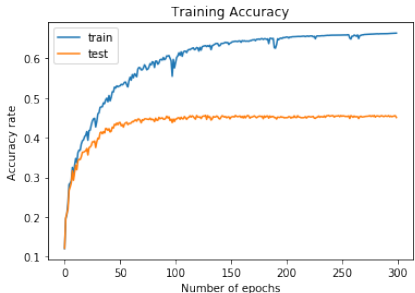


檢視了資料集之後，發現資料集只有 10,000 筆資料，因此我認為不用使用太過於複雜的模型。所以這邊只用了兩層 Convolution layer，與兩層 Pooling layer。Pooling 的方式則是使用 MaxPooling 的方式做處理。

Figure 1、Model structure

2. Compare with different hyperparameter

使用同上述的 Model structure，分別嘗試了兩種不同的超參數設定：

Model Structure	Setting 1		Setting 2	
Conv2d	Filter size	16	Filter size	16
	Kernel size	5	Kernel size	10
	Stride size	1	Stride size	2
MaxPool2d	Kernel size	2	Kernel size	2
Conv2d	Filter size	32	Filter size	32
	Kernel size	5	Kernel size	5
	Stride size	1	Stride size	2
MaxPool2d	Kernel size	2	Kernel size	2
Learning Curve	Setting 1		Setting 2	
Loss				
Accuracy				

iii. Result discussion

```

Accuracy of classes
dog       : 0.0 %
chicken   : 55.0 %
cat       : 55.5 %
cow       : 30.75 %
butterfly : 57.75 %
horse     : 50.75 %
sheep     : 43.75 %
elephant  : 57.5 %
spider    : 53.75 %
squirrel  : 46.25 %

```

這邊使用 Setting 2 的訓練出來的模型做預測，去對 val 資料集做預測，整理出 model 對各個類別的預測準確率做比較。這邊我覺得比較不合理的是，dog 類別的準確率竟然低至 0%，不過我認為我的 model 還沒有訓練完，loss 到最後一個 epoch 跑完，還是高達 1.8，因此 model 預測效能不好是可預期的。

以下也使用 Setting 2 的訓練出來的模型做預測，並畫出預測結果做探討：

pred: horse,
truth: horse



pred: chicken,
truth: elephant



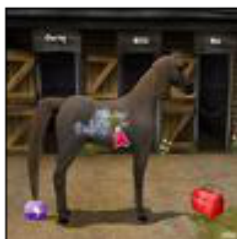
pred: cat,
truth: cat



pred: cat,
truth: dog



pred: squirrel,
truth: horse



pred: butterfly,
truth: butterfly



pred: squirrel,
truth: squirrel



pred: elephant,
truth: elephant



pred: squirrel,
truth: spider

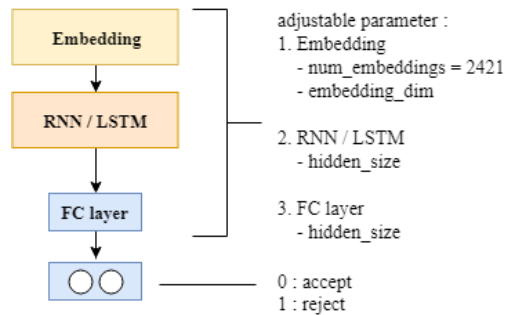


這邊顯示 9 張模型預測的結果與實際 label 的差異。其中我覺得挺多張圖片的預測效果蠻差的，例如正中間的馬預測成松鼠、最右下角的蜘蛛也預測成松鼠。不過也有幾張是有預測對的，我想若是未來要提升預測準確率，可以先從增加模型參數量著手，或是再多加幾層 Convolution layer。

II. Recurrent Neural Network for Prediction of Paper Acceptance

i. Standard RNN

1. Model Structure



經過資料前處理後，dictionary 的總數量是 2,421。因此 Embedding num_embeddings 參數固定是 2421。其他參數則可以任意更改。

2. Parameter setting and learning curve

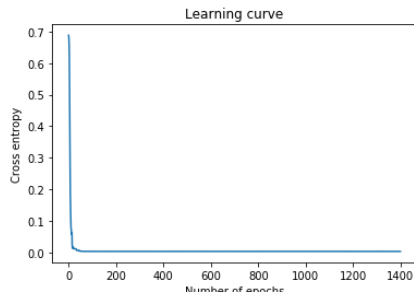
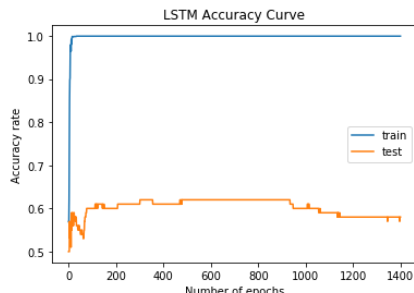
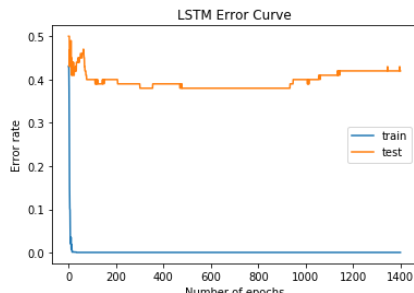
Model Structure	Setting 1	
Embedding	embedding_dim	100
RNN	hidden_size	64
FC layer	hidden_size	32
Learning Curve	Setting 1	
Loss		
Accuracy Rate		
Error Rate		

ii. LSTM

1. Model Structure

(同 RNN，只是把 RNN 層換成 LSTM)

2. Parameter setting and learning curve

Model Structure	Setting 1	
Embedding	embedding_dim	100
LSTM	hidden_size	64
FC layer	hidden_size	32
Learning Curve	Setting 1	
Loss		
Accuracy Rate		
Error Rate		

iii. Discussion

從上圖 RNN 的 Accuracy Curve 跟 LSTM 的 Accuracy Curve 中，可以明顯看出 LSTM 在訓練過程中較為穩定，最後的 model 效果也比較好，Testing data 的準確率可以到 0.6，而 RNN 大約只能在 0.5 左右，且震盪幅度也比較大，除此之外，也可以觀察到 LSTM 的收斂速度也比較快。