HW4: CNN

109511219 林錦樑

整個 classifier 的架構參考 PyTorch 官方教學

1. CNN

i. Baseline

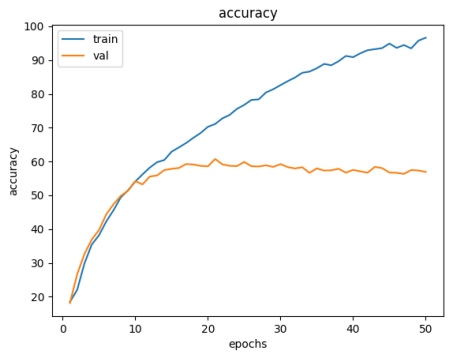
Baseline 的 CNN model 我使用了四層 convolution 與 max pooling，萃取出特徵後再使用五層 FC layer，其結構如下圖。

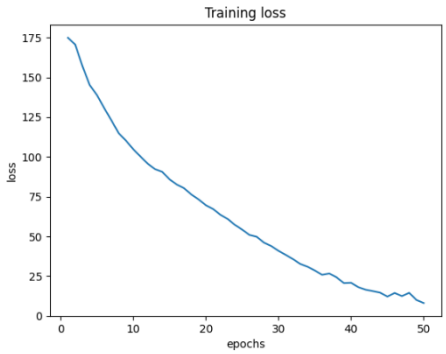
```python
class simpleCNN(nn.Module):
    def __init__(self):
        super().__init__()
        self.conv1 = nn.Conv2d(3, 6, 5) #124*124*6
        self.conv2 = nn.Conv2d(6, 12, 5) #58*58*12
        self.conv3 = nn.Conv2d(12, 16, 4) #26*26*16
        self.conv4 = nn.Conv2d(16, 20, 4) #10*10*20
        self.pool = nn.MaxPool2d(2, 2)
        self.fc1 = nn.Linear(20 * 5 * 5, 256)
        self.fc2 = nn.Linear(256, 128)
        self.fc3 = nn.Linear(128, 64)
        self.fc4 = nn.Linear(64, 32)
        self.fc5 = nn.Linear(32, 10)


    def forward(self, x):
        x = self.pool(F.relu(self.conv1(x))) #62*62*5
        x = self.pool(F.relu(self.conv2(x))) #29*29*5
        x = self.pool(F.relu(self.conv3(x))) #13*13*5
        x = self.pool(F.relu(self.conv4(x))) #5*5*20
        x = torch.flatten(x, 1) # flatten all dimensions except batch
        x = F.relu(self.fc1(x))
        x = F.relu(self.fc2(x))
        x = F.relu(self.fc3(x))
        x = F.relu(self.fc4(x))
        x = self.fc5(x)
        return x
```

Hyperparameters 的選擇上包括 batch size 為 256，optimizer 使用 Adam，learning rate 設為 0.001，總共訓練 50 個 epochs。資料的部分圖片輸入大小為 128*128，各種類圖片資料量如下表格。

類別	數量
butterfly	2012
cat	1528
chicken	2998
cow	1766
dog	4763
elephant	1346
horse	2523
sheep	1720
spider	4721
squirrel	1762

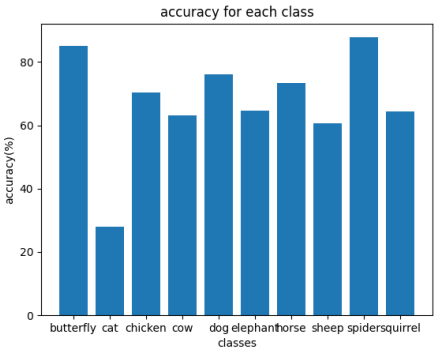
經過 50 個 epochs 後，其 accuracy 與 loss 趨勢如下圖。可以看出 loss 隨 epoch 增加持續下降，accuracy 在大概第 15 個 epoch 後，training accuracy 能持續上升，而 validation accuracy 在 55~60%震盪，可以推測訓練過程出現 overfitting 的現象。

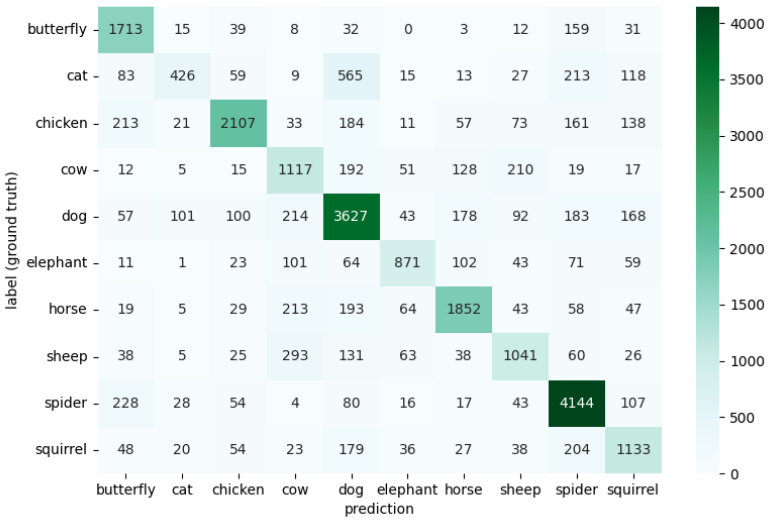




從全部資料得出的各類別的 accuracy 與 confusion matrix 來看，cat 的表現最差，常常誤分到 dog，其可能原因除了兩者相似之外，cat 的訓練資料最少也會導致 model 對 cat 的分類能力較差。而其他訓練圖片少於兩千張的類別，其準確率皆低於 70%；訓練圖片大於四千張的類別，其準確率可大於 85%，可見訓練圖片數量會影響模型分類表現。

```
Accuracy for class: butterfly is 85.1 %
Accuracy for class: cat   is 27.9 %
Accuracy for class: chicken is 70.3 %
Accuracy for class: cow   is 63.3 %
Accuracy for class: dog   is 76.1 %
Accuracy for class: elephant is 64.7 %
Accuracy for class: horse is 73.4 %
Accuracy for class: sheep is 60.5 %
Accuracy for class: spider is 87.8 %
Accuracy for class: squirrel is 64.3 %
```

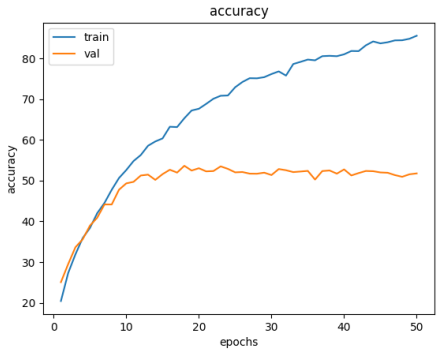


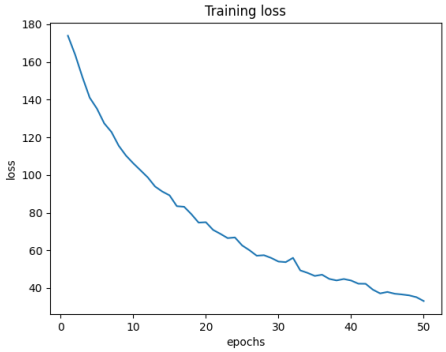


ii. L2 Regularization

由於出現 overfitting 的現象，我先利用 L2 Regularization 的方式限制參數大小，試著減緩 overfitting 的情形。在 optimizer 的參數上加上 1e-5 的 weight_decay，並將 learning rate 調為 0.005，訓練結果如下圖。

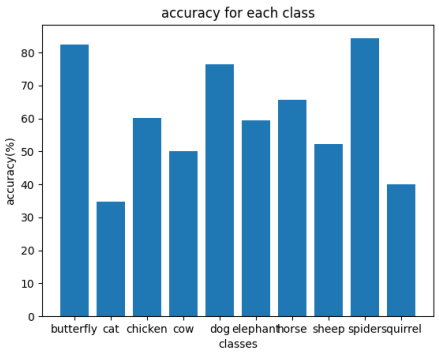
但其結果並不盡理想，不論是 training 或是 validation 的 accuracy 都降低了約 10%左右，loss 的部分則是由於加入了 Regularization term，有出現增加的情況。

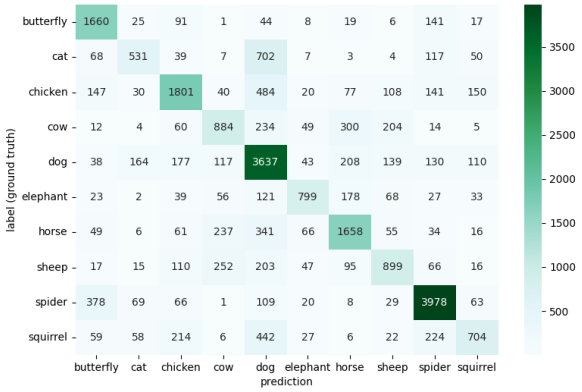




從全部資料得出的各類別的 accuracy 與 confusion matrix 來看，cat 的表現有變好，但誤分到 dog 的情況有增加，反倒是分類到其他類別的情況有減少。其他類別的準確率大部分下降，尤其以 squirrel 下降最多，最容易誤分到 dog。從以上實驗可看出 L2 Regularization 的方法無法提升模型準確度。

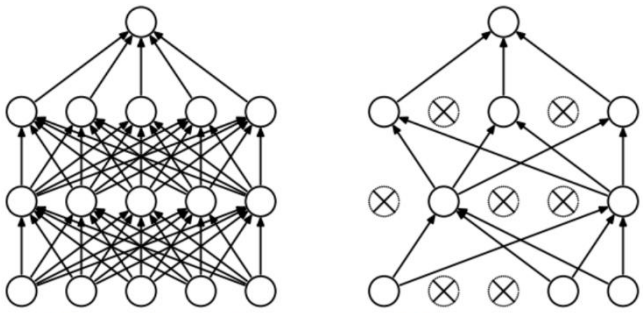
```
Accuracy for class: butterfly is 82.5 %
Accuracy for class: cat   is 34.8 %
Accuracy for class: chicken is 60.1 %
Accuracy for class: cow   is 50.1 %
Accuracy for class: dog   is 76.4 %
Accuracy for class: elephant is 59.4 %
Accuracy for class: horse is 65.7 %
Accuracy for class: sheep is 52.3 %
Accuracy for class: spider is 84.3 %
Accuracy for class: squirrel is 40.0 %
```





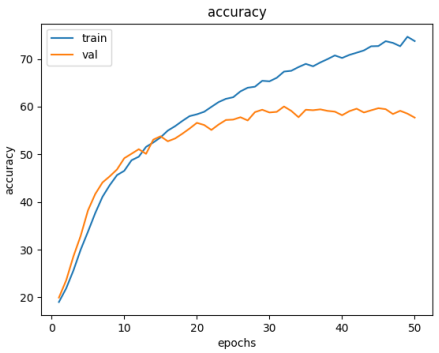
iii. Dropout

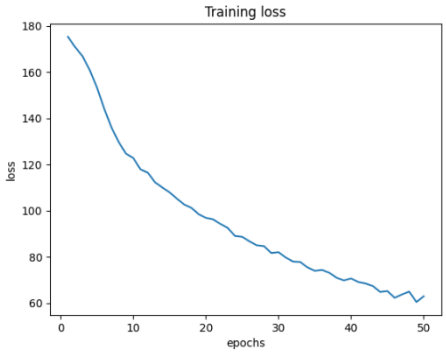
由於 L2 Regularization 無法解緩 overfitting 的問題，甚至帶來更差的結果，我決定改用 dropout 的方式來處理問題。Dropout 會在前向傳播時，依照設定的機率關閉神經元，降低對於特定神經元的依賴，如下圖所示。在實作上我在每層 Fully Connected Layer 都加上了機率為 0.3 的 dropout。



(a) Standard Neural Net　　(b) After applying dropout.

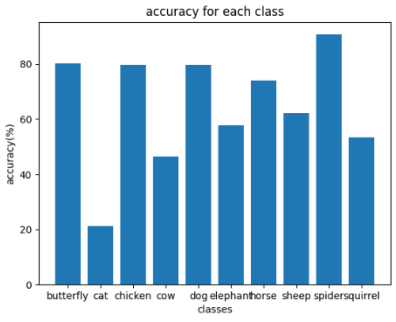
加入 dropout 後的實驗結果如下，validation 的 accuracy 回到跟 baseline 差不多水準，但 training accuracy 有下降，loss 也有上升，可能需要經過更多 epoch 的訓練才能得出較 baseline 好的 model。

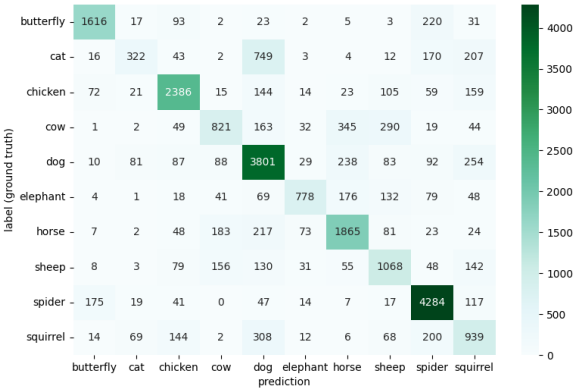




從全部資料得出的各類別的 accuracy 與 confusion matrix 來看，進步與退步的類別各占一半。其中將 cat 誤分到 dog 的狀況變得更加嚴重。

```
Accuracy for class: butterfly is 80.3 %
Accuracy for class: cat   is 21.1 %
Accuracy for class: chicken is 79.6 %
Accuracy for class: cow   is 46.5 %
Accuracy for class: dog   is 79.8 %
Accuracy for class: elephant is 57.8 %
Accuracy for class: horse is 73.9 %
Accuracy for class: sheep is 62.1 %
Accuracy for class: spider is 90.7 %
Accuracy for class: squirrel is 53.3 %
```





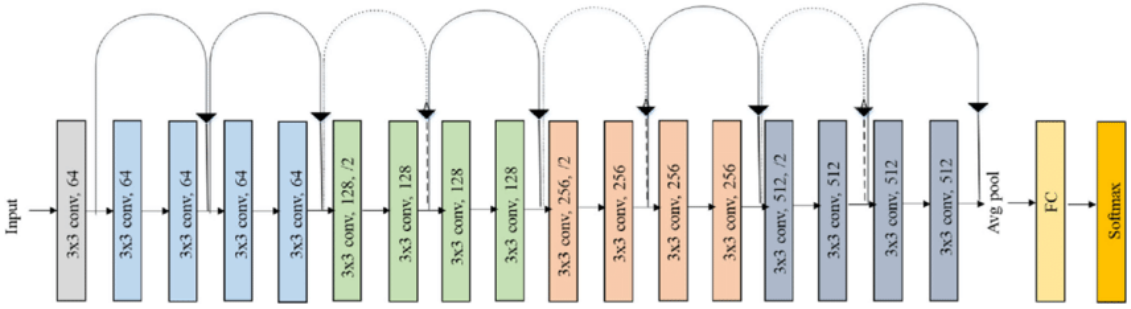
從以上結果來看，L2 Regularization 與 dropout 都無法解決 overfitting 的問題，甚至會帶來更差的結果。我推測可能的原因來自訓練圖片不夠多，導致一些處理 overfitting 的方法不夠有效。也有可能嘗試改變這些方法的 hyperparameter 後，能帶來更好的效果。

2. ResNet18

i. Baseline

由於可能因訓練資料不足，而造成前面使用處理 overfitting 的方法來提升 validation accuracy 的效果都不盡理想，我嘗試使用一個更大的模型來同時提升 training 和 validation accuracy，這邊以 ResNet18 來進行實驗。

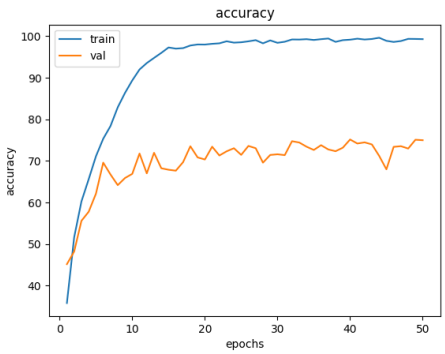
ResNet18 包含 17 層的 convolutional layer 與 1 層的 FC layer，共 18 層，其中由兩層 convolutional layer 所形成的 residual block 可以透過增加一條加法路線的方式，解決梯度消失的問題。整體架構如下圖，其中上方連線部分即為 residual block 的加法路線。

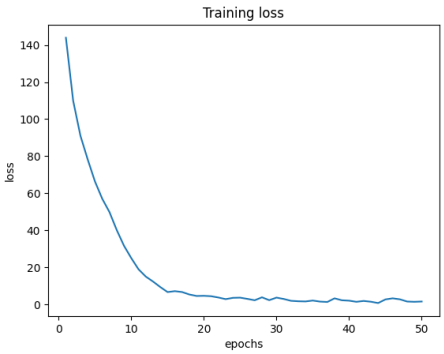


整個 model 的部分是經過一個 ResNet18 萃取出特徵後，再使用四層 FC layer 進行分類任務，其結構如下圖。

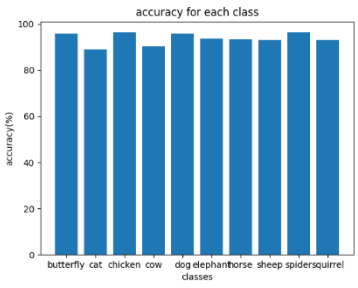
```python
class ResNet(nn.Module):
    def __init__(self):
        super().__init__()
        #self.resnet18 = models.resnet18(weights='IMAGENET1K_V1')
        self.resnet18 = models.resnet18()
        self.fc = nn.Linear(self.resnet18.fc.in_features, 512)
        self.fc2 = nn.Linear(512, 128)
        self.fc3 = nn.Linear(128, 32)
        self.fc4 = nn.Linear(32, 10)
        self.resnet18.fc = self.fc
        #self.dropout = nn.Dropout(0.3)

    def forward(self, x):
        x = self.resnet18(x)
        #x = self.dropout(x)
        x = F.relu(self.fc2(x))
        #x = self.dropout(x)
        x = F.relu(self.fc3(x))
        #x = self.dropout(x)
        x = self.fc4(x)
        return x
```

從整體 accuracy 來看，training accuracy 比一般簡單的 CNN 上升速度還快，validation accuracy 也增加了約 10%左右。Loss 的收斂速度也增加許多。從個別的 accuracy 來看，整體表現大幅提升，各類別 accuracy 都有大於 88%，且各類別表現較為平均。可以看出使用更大的模型能帶來更好的效果，但仍舊存在 overfitting 的問題。

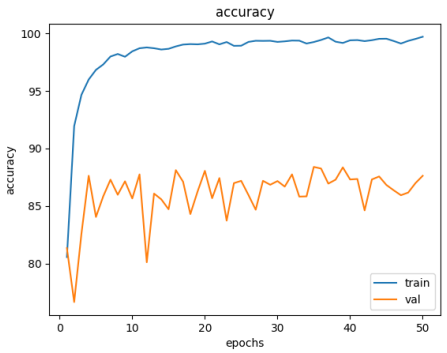


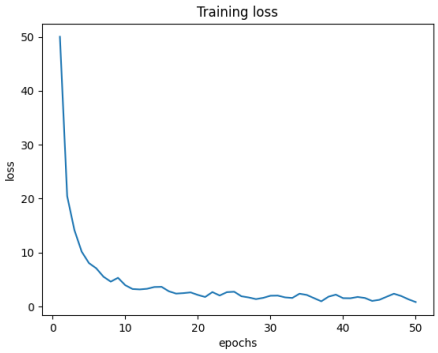



```
Accuracy for class: butterfly is 95.9 %
Accuracy for class: cat   is 88.9 %
Accuracy for class: chicken is 96.3 %
Accuracy for class: cow   is 90.4 %
Accuracy for class: dog   is 95.8 %
Accuracy for class: elephant is 93.6 %
Accuracy for class: horse is 93.3 %
Accuracy for class: sheep is 93.0 %
Accuracy for class: spider is 96.3 %
Accuracy for class: squirrel is 93.2 %
```

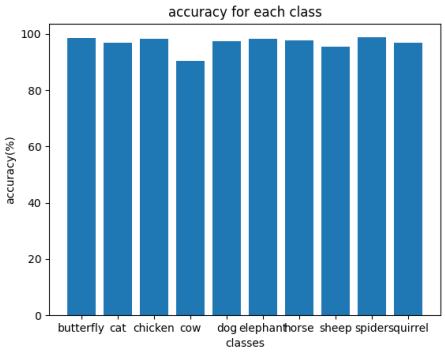


Pretrained model 的部分是使用了 IMAGENET1K_V1 這個 weight，從整體與個別的 accuracy 來看皆有提升，validation 的 accuracy 增加了 10%～15%，各類別的 accuracy 均能大於 90%，可以看出使用 pretrained model 能對這個分類任務帶來更好的效果。



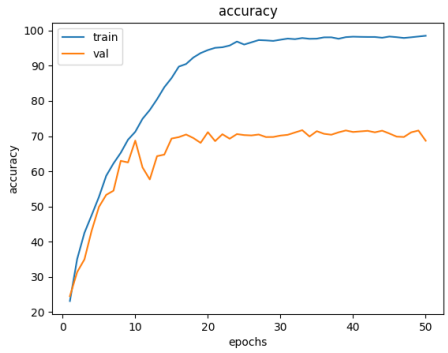


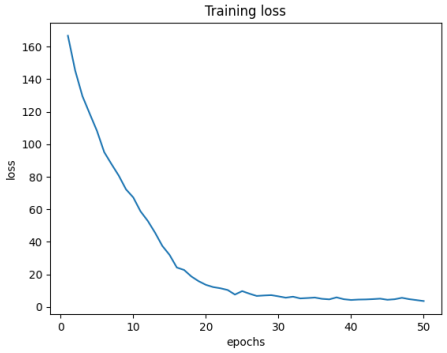

```
Accuracy for class: butterfly is 98.4 %
Accuracy for class: cat   is 96.9 %
Accuracy for class: chicken is 98.2 %
Accuracy for class: cow   is 90.4 %
Accuracy for class: dog   is 97.5 %
Accuracy for class: elephant is 98.1 %
Accuracy for class: horse is 97.6 %
Accuracy for class: sheep is 95.5 %
Accuracy for class: spider is 98.8 %
Accuracy for class: squirrel is 96.9 %
```



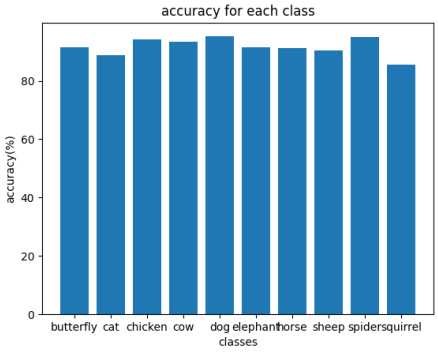
ii. L2 Regularization

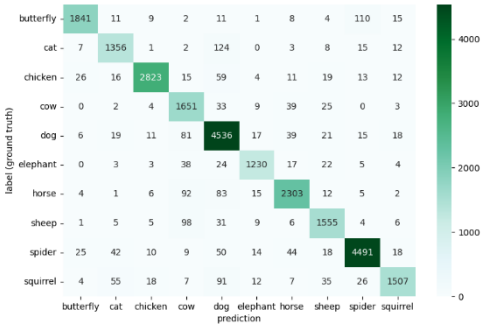
以下為使用 L2 Regularization 後的 ResNet18 model 的表現。從 accuracy 曲線、各類別 accuracy 與 confusion matrix 來看，皆無太大變化。



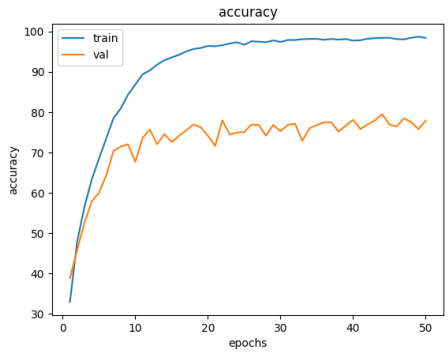


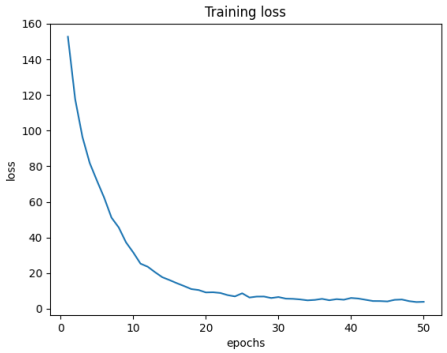

```
Accuracy for class: butterfly is 91.5 %
Accuracy for class: cat   is 88.7 %
Accuracy for class: chicken is 94.2 %
Accuracy for class: cow   is 93.5 %
Accuracy for class: dog   is 95.2 %
Accuracy for class: elephant is 91.4 %
Accuracy for class: horse is 91.3 %
Accuracy for class: sheep is 90.4 %
Accuracy for class: spider is 95.1 %
Accuracy for class: squirrel is 85.5 %
```



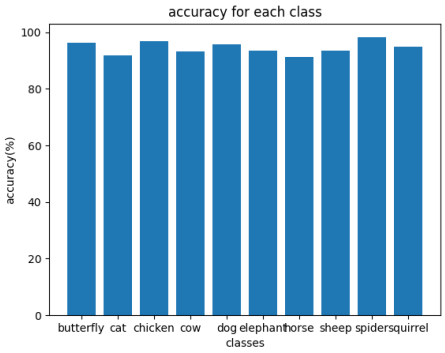


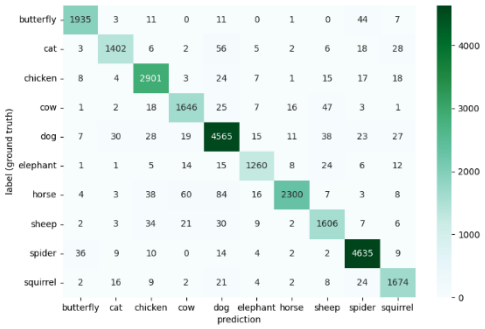
以下為使用 L2 Regularization 後的 pretrained ResNet18 model 的表現。從 accuracy 曲線來看，validation accuracy 掉了約 10%左右。從各類別 accuracy 與 confusion matrix 來看，大部分類別的 accuracy 皆下降。





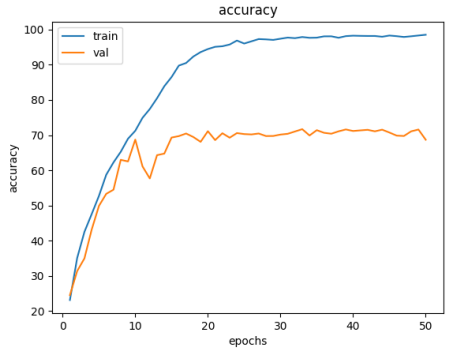

```
Accuracy for class: butterfly is 96.2 %
Accuracy for class: cat   is 91.8 %
Accuracy for class: chicken is 96.8 %
Accuracy for class: cow   is 93.2 %
Accuracy for class: dog   is 95.8 %
Accuracy for class: elephant is 93.6 %
Accuracy for class: horse is 91.2 %
Accuracy for class: sheep is 93.4 %
Accuracy for class: spider is 98.2 %
Accuracy for class: squirrel is 95.0 %
```

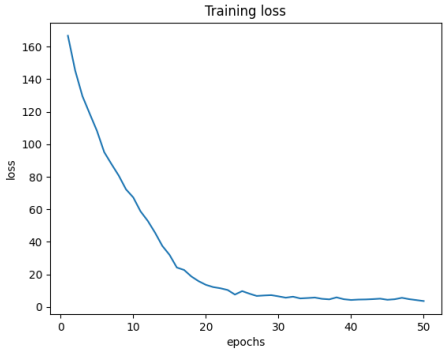




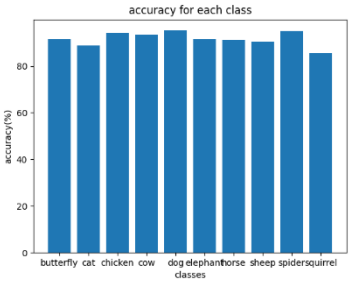
iii. Dropout

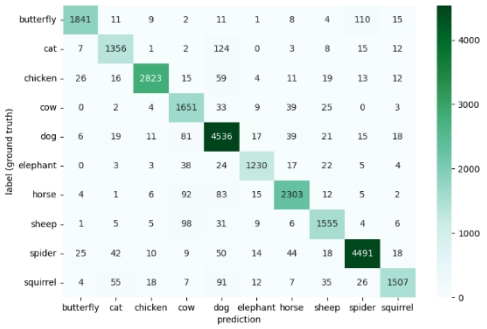
以下為使用 Dropout 後的 ResNet18 model 的表現。從 accuracy 曲線、各類別 accuracy 與 confusion matrix 來看，皆無太大變化。



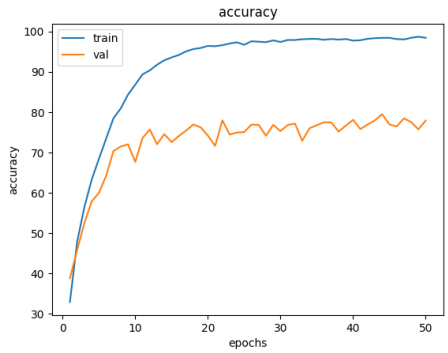


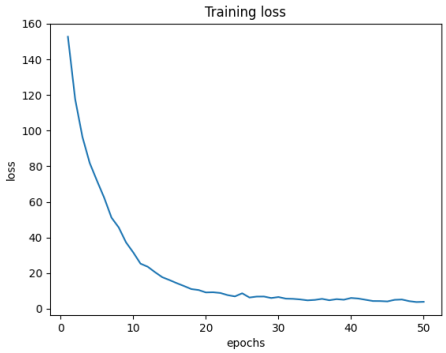

```
Accuracy for class: butterfly is 96.4 %
Accuracy for class: cat   is 88.4 %
Accuracy for class: chicken is 95.8 %
Accuracy for class: cow   is 93.8 %
Accuracy for class: dog   is 95.7 %
Accuracy for class: elephant is 94.9 %
Accuracy for class: horse is 90.7 %
Accuracy for class: sheep is 92.3 %
Accuracy for class: spider is 97.2 %
Accuracy for class: squirrel is 91.3 %
```





以下為使用 Dropout 後的 pretrained ResNet18 model 的表現。從 accuracy 曲線來看，validation accuracy 掉了約 10%左右。從各類別 accuracy 與 confusion matrix 來看，無太大變化。






```
Accuracy for class: butterfly is 97.1 %
Accuracy for class: cat   is 93.5 %
Accuracy for class: chicken is 98.7 %
Accuracy for class: cow   is 91.8 %
Accuracy for class: dog   is 97.7 %
Accuracy for class: elephant is 97.1 %
Accuracy for class: horse is 97.9 %
Accuracy for class: sheep is 95.2 %
Accuracy for class: spider is 99.5 %
Accuracy for class: squirrel is 94.5 %
```