DLCV HW4

R08945011 張祐祥

Problem 1

1. Describe the architecture & implementation details of your model.

   Training episode : 50

   Distant function : support data 所求之中心與 query data 點對點相乘後的總合

   Optimizer : Adam

   Learning rate: 起始 0.5 之後每 10 個 epoch 乘上 0.5

   Data augmentation : 標準化

   Meta train : 5-way 5-shot , 30 query data per way

   Meta teat : 5way 1shot , 15 query data per way

   accuracy on validation set : 45.55 +- 0.82

2. Please report and discuss the accuracy of the prototypical network using 3 different distance function

   在本題中三種 distance function 在 1shot 時差異並沒很大，若再增加 Epoch 應該可以達到同樣效果，只是收斂速度稍微不同，其中 Euclidean distance 收斂速度相對較慢

   under 5 way 1 shot:

   Euclidean distance : 39.32 +- 1.02%

   cosine similarity : 40.01 +- 1.2%

   parametric function: 42.06 +- 0.79%

   parametric function 為計算每組 support data 所求之中心與 query data 點對點相乘後的總合

3. When meta-train and meta-test under the same 5-way K-shot setting, please report and compare the accuracy with different shots. (K=1, 5, 10)

   在不同 shot 中當數量大於五之後效果沒有太大差異，其中差異較大為 1 shot，與 5 shot，5 shot 可以達到較高的正確率

   Using dot function:

   K = 1 : 43.65+- 2.10%

   K = 5 : 45.55 +- 0.82

   K = 10 : 44.01 +- 0.92

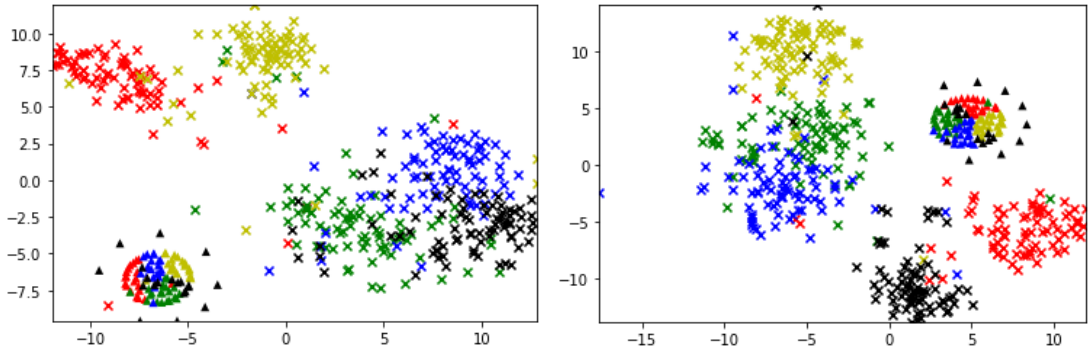
Problem 2

1. Describe the architecture & implementation details of your model.

```
Hallucinator(
    (linear1): Linear(in_features=3200, out_features=2400, bias=True)
    (relu1): ReLU()
    (drop1): Dropout(p=0.5, inplace=False)
    (linear2): Linear(in_features=2400, out_features=2400, bias=True)
    (relu2): ReLU()
    (drop2): Dropout(p=0.5, inplace=False)
    (linear3): Linear(in_features=2400, out_features=1600, bias=True)
    (relu3): ReLU()
)
```

Training episode : 50
Distant function : support data 所求之中心與 query data 點對點相乘後的總合
Optimizer : Adam
Learning rate: 起始 0.5 之後每 10 個 epoch 乘上 0.5
Data augmentation : 標準化
Meta train : 5-way 5-shot , 30 query data per way, 100-augmentation
Meta teat : 5way 1shot , 15 query data per way, 100-augmentation
accuracy on validation set : 47.59 +- 0.79%

2. Visualize the real and hallucinated data in the latent space



3. Report and compare the accuracy with different number of hallucinated data. (M=10, 50, 100)

M = 10 : 44.1 +- 1.1%
M = 50 : 44.26 +- 0.62%
M = 100 : 47.59 +- 0.79%

當 M 較小時初期訓練速度，而當 M 較大可以達到較好的正確性，但當 M 過大則會產生反效果同時拖慢訓練速度

4. Discuss what you've observed and learned from implementing the

data hallucination model.
在 hallucinator 的設計上，對於正確率有滿大的影響，尤其是加入 dropout layer 降低 overfitting 可以大幅提升正確率。而由於額外加入合成的資料，此方法相較於第一題達到收斂所需要的 epoch 比較少。

Problem 3

1. Describe the architecture & implementation details of your model.

```
Hallucinator (
  (linear1): Linear(in_features=3200, out_features=2400, bias=True)
  (relu1): ReLU()
  (drop1): Dropout(p=0.5, inplace=False)
  (linear2): Linear(in_features=2400, out_features=2400, bias=True)
  (relu2): ReLU()
  (drop2): Dropout(p=0.5, inplace=False)
  (linear3): Linear(in_features=2400, out_features=1600, bias=True)
  (relu3): ReLU()
)

discriminator(
  (linear1): Linear(in_features=1600, out_features=512, bias=True)
  (relu1): ReLU()
  (drop1): Dropout(p=0.5, inplace=False)
  (linear2): Linear(in_features=512, out_features=64, bias=True)
  (relu2): ReLU()
  (drop2): Dropout(p=0.5, inplace=False)
  (linear3): Linear(in_features=64, out_features=1, bias=True)
  (relu3): ReLU()
  (drop3): Dropout(p=0.5, inplace=False)
)
```

此題是額外加入 Discriminator loss 來訓練 Hallucinator，其中先用 Meta train 的所有真實資料與 Hallucinator 產生的所有資料計算 Discriminator loss 並將其更新，再以更新的 Discriminator 計算出 Hallucinator 的 generation loss 結合原先 meta learning loss 來更新 Hallucinator

Training episode : 50
Distant function : support data 所求之中心與 query data 點對點相乘後的總合
Optimizer : Adam

Learning rate: 起始 0.5 之後每 10 個 epoch 乘上 0.5
Data augmentation : 標準化
Meta train : 5-way 5-shot , 30 query data per way, 100-augmentation
Meta teat : 5way 1shot , 15 query data per way, 100-augmentation
accuracy on validation set : 21.32%