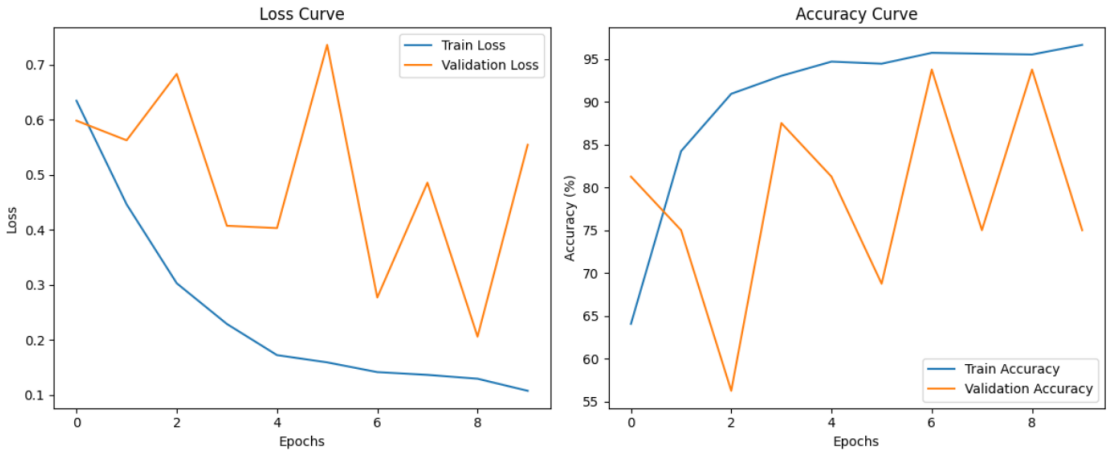
HW3

Hyperparameter Tuning

張晉綸

113003813

1.

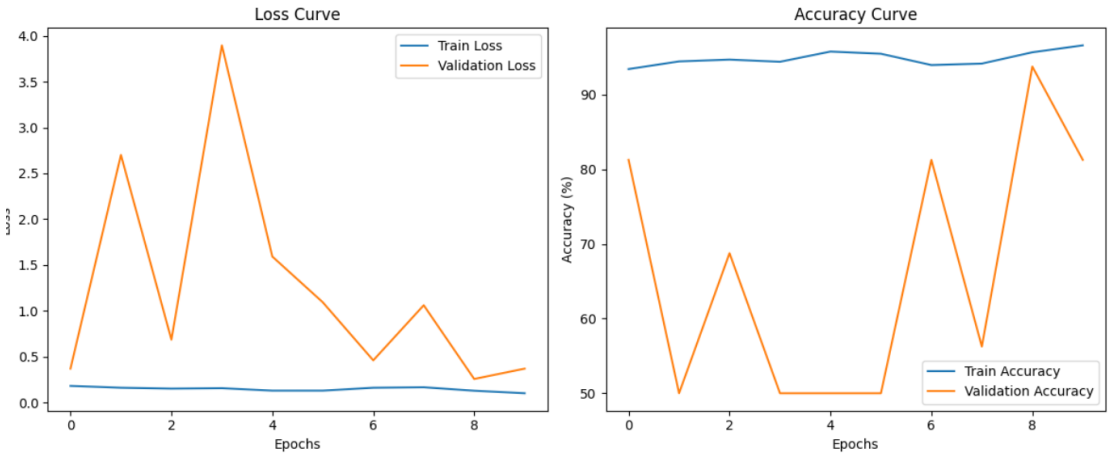



```
model = nn.Sequential(
        nn.Flatten(),
        nn.Linear(256 * 256 * 1, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.5),

        nn.Linear(64, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.5),

        nn.Linear(64, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.5),

        nn.Linear(64, 1),
        nn.Sigmoid()
)
```

```
lr = 0.001
weight_decay = 0.001
epochs = 10
optimizer = optims.Adam(model.parameters(), lr=lr, weight_decay = weight_decay)
lr_scheduler = optims.lr_scheduler.ReduceLROnPlateau(optimizer, factor=0.1, patience=5, mode='max')
criterion = nn.BCELoss()
```

```
Test Accuracy: 79.43%
Test Loss: 0.4777
```

2.

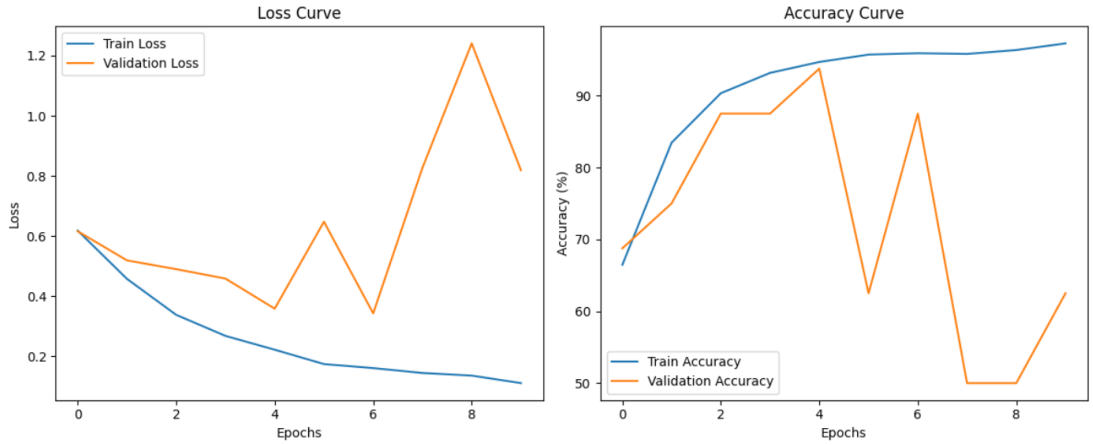



```
model = nn.Sequential(
        nn.Flatten(),
        nn.Linear(256 * 256 * 1, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.5),

        nn.Linear(64, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.5),

        nn.Linear(64, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.5),

        nn.Linear(64, 1),
        nn.Sigmoid()
)
```

```
# hyperparameter
lr = 0.002
weight_decay = 0.001
epochs = 10
optimizer = optims.Adam(model.parameters(), lr=lr, weight_decay = weight_decay)
lr_scheduler = optims.lr_scheduler.ReduceLROnPlateau(optimizer, factor=0.1, patience=5, mode='min')
criterion = nn.BCELoss()
```

```
Test Accuracy: 84.90%
Test Loss: 0.3603
```

3.

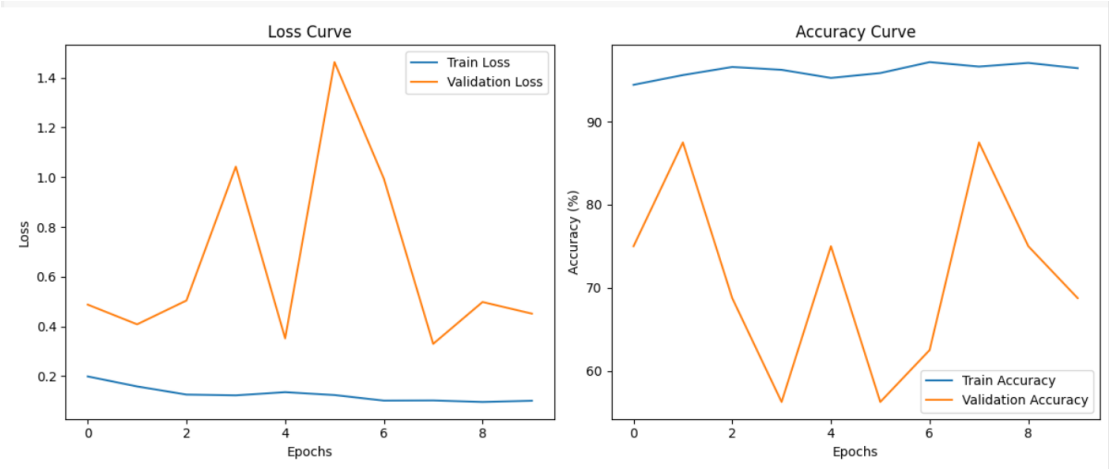



```
model = nn.Sequential(
        nn.Flatten(),
        nn.Linear(256 * 256 * 1, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.5),

        nn.Linear(64, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.5),

        nn.Linear(64, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.5),

        nn.Linear(64, 1),
        nn.Sigmoid()
)
```

```
# hyperparameter
lr = 0.0005
weight_decay = 0.001
epochs = 10
optimizer = optims.Adam(model.parameters(), lr=lr, weight_decay = weight_decay)
lr_scheduler = optims.lr_scheduler.ReduceLROnPlateau(optimizer, factor=0.05, patience=5, mode='min')
criterion = nn.BCELoss()
```

```
Test Accuracy: 81.30%
Test Loss: 0.3927
```

4.

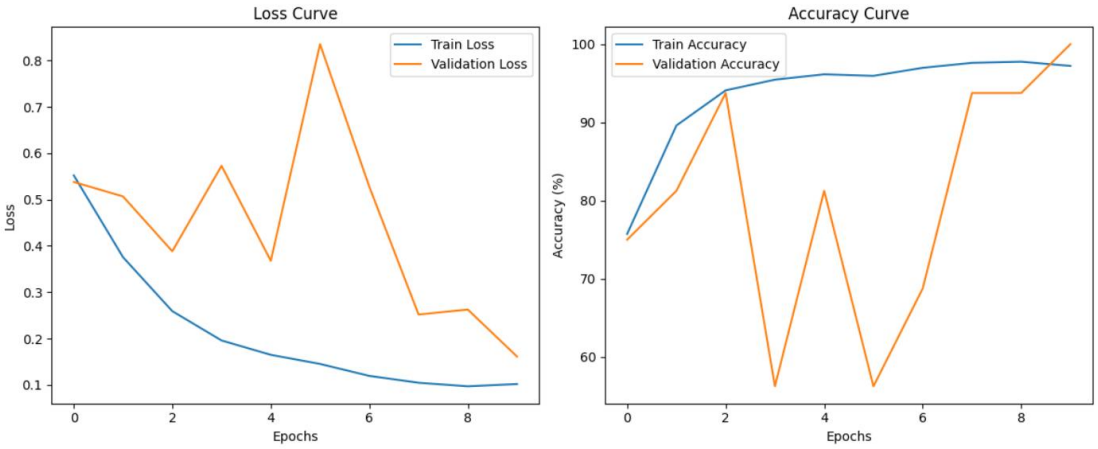



```
model = nn.Sequential(
        nn.Flatten(),
        nn.Linear(256 * 256 * 1, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.5),

        nn.Linear(64, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.5),

        nn.Linear(64, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.5),

        nn.Linear(64, 1),
        nn.Sigmoid()
)
```

```
# hyperparameter
lr = 0.0005
weight_decay = 0.002
epochs = 10
optimizer = optims.Adam(model.parameters(), lr=lr, weight_decay = weight_decay)
lr_scheduler = optims.lr_scheduler.ReduceLROnPlateau(optimizer, factor=0.1, patience=10, mode='min')
criterion = nn.BCELoss()
```

```
Test Accuracy: 81.30%
Test Loss: 0.3927
```

5.




```
model = nn.Sequential(
        nn.Flatten(),
        nn.Linear(256 * 256 * 1, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.5),

        nn.Linear(64, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.4),

        nn.Linear(64, 64),
        nn.BatchNorm1d(64),
        nn.ReLU(),
        nn.Dropout(0.3),

        nn.Linear(64, 1),
        nn.Sigmoid()
)
```

```
# hyperparameter
lr = 0.0005
weight_decay = 0.001
epochs = 10
optimizer = optims.Adam(model.parameters(), lr=lr, weight_decay = weight_decay)
lr_scheduler = optims.lr_scheduler.ReduceLROnPlateau(optimizer, factor=0.1, patience=2, mode='min')
criterion = nn.BCELoss()
```

```
Test Accuracy: 83.96%
Test Loss: 0.3849
```

本次的超參數調整共做了五種版本，主要調整為 learning rate, weight decay, learning rate scheduler 內的 factor、patience 與 mode 這幾個參數。

Training set 的表現都不錯，最終結果幾乎都在 95%以上，validation set 因為數量較少，所以在每個 epoch 呈現較起伏的情況，但最終在第五組參數的設定下，validation set 的準確率也是高於 95%。

最終 test set 的準確率都可以到達 80%以上，以目前的設置及超參數的變動選擇，相信已經是不錯的表現了。