

# National Tsing Hua University

## Fall 2023 11210IPT 553000

### Deep Learning in Biomedical Optical Imaging

#### Homework 4

GAO WEL LUN<sup>1</sup>

<sup>1</sup>Institute of Photonics Technologies, National Tsing Hua University, Hsinchu 30013, Taiwan

Student ID:110066511

#### 1. Task A

本篇使用的預先訓練模型分別是 RESNET50 和 ShuffleNetv\_V2\_X2\_0，探討這兩個模型的架構和效能，其訓練模型如 Fig.1

```
[1] import torch.nn as nn
import torch.nn.functional as F
from torchvision import models

model = models.resnet50(weights='IMAGENET1K_V1')

# ConvNet as fixed feature extractor (freeze parameters)
# for param in model.parameters():
#     param.requires_grad = False

num_fts = model.fc.in_features

# Here the size of each output sample is set to 2.
# Alternatively, it can be generalized to ``nn.Linear(num_fts, len(class_names))``.
model.fc = nn.Linear(num_fts, 2)
model = model.cuda()
```

```
[ ] import torch.nn as nn
import torch.nn.functional as F
from torchvision import models

model = models.shufflenet_v2_x2_0(weights='IMAGENET1K_V1')

# ConvNet as fixed feature extractor (freeze parameters)
# for param in model.parameters():
#     param.requires_grad = False

num_fts = model.fc.in_features

# Here the size of each output sample is set to 2.
# Alternatively, it can be generalized to ``nn.Linear(num_fts, len(class_names))``.
model.fc = nn.Linear(num_fts, 2)
model = model.cuda()
```

Fig. 1. 預先訓練的模型程式碼，分別是(左) RESNET50 和(右) ShuffleNetv\_V2\_X2\_0

RESNET50 和 ShuffleNetv\_V2\_X2\_0 是兩種不同的深度學習模型，從模型結構和性能討論，RESNET50 是屬於深度捲積神經網路，如字面上所示它擁有 48 層捲積層和 2 層全連接層，總共 50 層，並採用深度殘差塊(Deep Residual Blocks)，增強了特徵的捕捉能力，因此就性能方面有相當高的準確度，能在大規模的圖表分類中表現出色。但相反的該模型的深度和複雜度也相對較高，需要使用的參數也要更多，因此在訓練過程會消耗更多的時間和資源。而 ShuffleNetv\_V2\_X2\_0 是屬於較輕量級的深度學習，使用隨機區塊(ShuffleBlock)，其中包含逐點捲積(pointwise convolution)和組捲積(group convolution)，其目的是減少參數量和複雜度，因此在訓練過程消耗更少的時間和資源，相反的在性能方面會有所犧牲。

#### 2. Task B: Fine-tuning the ConvNet

本次的分類任務為 X 光胸腔照健康於否，辨別胸腔正常與否兩種類別，因此將預先訓練的模型導入這次深度學習架構中，分別是 ShuffleNetv\_V2\_X2\_0 和 RESNET50 兩種模型，並將全連接層輸出修改成二，其結果如 Fig.2 所示。

ShuffleNetv\_V2\_X2\_0 是屬於較輕量級的深度學習模型，主要是為減少訓練時間和電腦計算量，在訓練時間上和 RESNET18 一樣，花費了三分鐘左右完成了 30 次迭代的訓練，其訓練和驗證精準度和損失的結果，RESNET18 模型較好一些。而在新數據的測試中，兩者的準確都落在 78%。從結果來看，兩者模型在二元分類任務的性能很相似，若是在大規模的圖像分類任務中，RESNET18 會更具有優勢，由於它擁有 16 層捲積層和 2 層全連接層，較有深度的模型結構，能更好的捕捉圖像數據中的特徵。因此，

選擇使用 ShuffleNetV2 x2.0 或 ResNet-18 取決於個人需求。如果需要更高的準確度並且有足夠的運算資源，那麼 ResNet-18 比較合適。如果需要專注於模型大小和運算效率，或在資源受限的環境中工作，那麼 ShuffleNetV2 x2.0 會是更好的選擇。

RESNET50 是屬於深度捲積神經網路，主要是強化特徵的捕捉能力，它和 RESNET18 一樣是由多層捲積層組成，然而不同它擁有 48 層捲積層和 2 層全連接層，總共 50 層，由於更深的網路結構，RESNET50 具有更多的參數，這意味著它可以提供更強大的特徵表示能力。這對於大規模影像分類和其他複雜任務非常有用。然而在本次的二元分類任務中，RESNET50 由於較深的神經網路結構，訓練時間花費 10 分鐘完成了 30 次迭代的訓練，其訓練和驗證精準度和損失的結果，RESNET18 模型較好一些。而在新數據的測試中，RESNET50 的準確則有 81%，和 RESNET18 相比有較高一些的精準度。總結來說，ResNet-50 適用於需要更高準確性和更強大特徵提取能力的任務，但它也需要更多的運算資源。ResNet-18 更適合資源受限的環境，其中運算資源有限，但效能仍可滿足。選擇哪個模型應根據個人的任務需求和可用資源來決定。

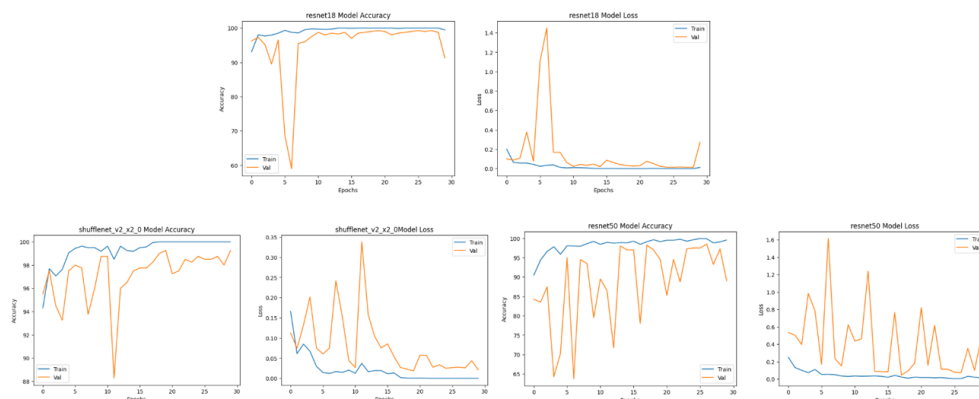


Fig. 2. 訓練和驗證精準度和損失的結果。分別是(上) RESNET18、(左) ShuffleNetv\_V2\_X2\_0 和(右) RESNET50 模型訓練結果。

### 3. Task C: ConvNet as Fixed Feature Extractor

延續 Task B 的訓練模型，透過凍結預先訓練模型除了最後一層外的所有層，將所選模型轉為固定特徵提取器，也就是利用 `requires_grad` 設定 `False`，以防止它們在訓練期間更新，這將凍結所有圖層。凍結所有圖層主要用途為包含：

- 減少 GPU 計算量: 凍結圖層可以大幅減少前向傳播和反向傳播的計算開銷，因為在訓練期間不需要更新這些圖層的權重。
- 避免過度擬合: 過度擬合是指模型過於複雜，適應了訓練資料的雜訊，導致在新資料上的效能下降。凍結圖層有助於維持模型的一般性，可以減少過度擬合的風險。
- 提高模型泛化能力: 確保模型保留了在大規模資料上學到的通用特徵，這有助於模型更好地。
- 更快的訓練速度: 由於凍結圖層可以減少反向傳播的計算量，因此模型的訓練速度更快。

透過凍結 RESNET18 、ShuffleNet\_V2\_X2\_0 和 RESNET50 模型的所有圖層，其訓練結果如 Fig. 3 所示。在訓練時間上，三者的訓練時間有大幅度的減少，RESNET18 和 ShuffleNet\_V2\_X2\_0 所減至 1 分鐘，而 RESNET50 則縮減至 4 分鐘，和原先未凍結的訓練時間所短至一半。由於凍結圖層可以在訓練期間不需要更新這些圖層的權重，因此能大幅減少前向傳播和反向傳播的計算容量，減少參數的更新變動，在模型的訓練速度上更快。在訓練和驗證精準度和損失的結果方面，三者的精準度有所提升，且損失方面有所降低，由於凍結了圖層，使模型保留了在大規模資料上學到的通用特徵，並維持模型的一般性，因此在性能方面所提升，三者在新數據的測試中，精準度都達到 80% 以上。

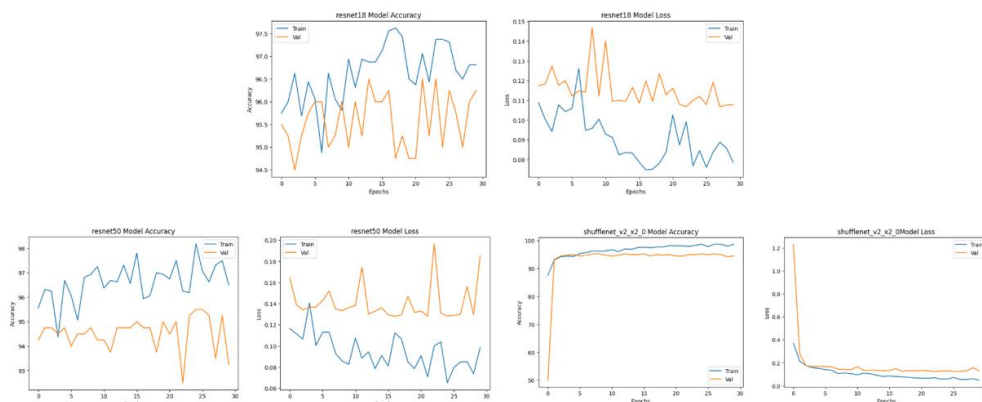


Fig. 3. 透過凍結所有圖層，訓練和驗證精準度和損失的結果。分別是(上) RESNET18、(左) RESNET50 和 (右) ShuffleNet\_V2\_X2\_0 模型訓練結果。

#### 4. Task D: Comparison and Analysis

Task B 和 Task C 的差異，在於凍結圖層的有無。由於 Task B 模型沒有凍結圖層，權重和參數的部分，根據所給予的訓練數據，進行大幅度的更新，因此在訓練所消耗的時間相較更長。而 Task C 模型則將圖層凍結，意味著模型保留了在大規模資料上學到的通用特徵，大幅減少前向傳播和反向傳播的計算開銷，因此在訓練期間不需要更新這些圖層的權重，訓練所消耗的時間相較更短。

在性能方面上，Task C 模型保留了在大規模資料上學到的通用特徵，有助於提升模型的泛化性和維持模型的一般性，因在訓練和驗證的精準度和損失比較，如 Fig. 2 和 Fig. 3 所示，有凍結圖層的精準度相對較好，訓練和驗證之間的變異數相對減小一點，減少了過度擬合的跡象。總結來說，Task C 圖層凍結的模型相比 Task B 沒有圖層凍結，訓練花費時間更短，性能上更好。

然而在某些任務中，當任務與預訓練模型的底層特徵較不相關時，凍結圖層可能不是最佳策略。在這種情況下，可能需要微調一些圖層，以適應特定任務的需求。凍結圖層與否依據任務和可用的資料來決定。而本次的分類任務與預訓練模型的底層特徵有一定的相關性，因此在訓練上有良好的表現。

#### 5. Task E: Test Dataset Analysis

難以提高精準度在新數據上的原因，可能包含：

- 資料偏移：訓練資料和測試資料之間可能存在資料分佈的偏移。這意味著測試資料與訓練資料在統計特性上有所不同，導致模型難以泛化到測試資料。
- 過度擬合：模型在訓練資料上表現良好，但在測試資料上表現很糟糕，可能是過度擬合造成的。模型可能在訓練資料上過度擬合雜訊或細微差異，而這些雜訊和差異在測試資料中不再存在。
- 數據品質：測試數據本身可能包含雜訊、錯誤標籤或缺失數據，這些因素都會降低模型在測試數據上的效能。
- 樣本不平衡：測試資料中可能存在類別不平衡問題，其中一些類別的樣本數量明顯少於其他類別。這可能導致模型在少數類別上表現不佳。

可能是以上原因，導致在新數據測試中，精準度無法再進一步提升的原因。