

# National Tsing Hua University

## Fall 2023 11210IPT 553000

### Deep Learning in Biomedical Optical Imaging Report

GAO WEL LUN<sup>1</sup>

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

*Student ID: 110066511*

#### 1. 介紹

深度學習是屬於機器學習的一種，它透過模擬人腦的神經網路結構來解決複雜的任務和問題。深度學習的基本原理是透過多層次的神經網路學習輸入資料的表示，從而能夠對未見過的資料進行準確的預測或分類，並在各種領域取得了顯著的幫助，包括影像和語音辨識、自然語言處理、醫療診斷等。

本篇將使用深度學習，透過對癌組織的組織學影像進行分類來深入研究醫學影像分析領域。在資料集中的每個樣本為 150x150 像素的 RGB 影像，代表癌症組織學中常見的六種不同組織紋理之一，而本篇將使用深度學習中的其中一項技術遷移學習 (Transfer learning)，使用 ResNet18 和 ResNet50 兩個預先訓練好的模型來對每個圖形進行分類，並對其模型效能和分類表現進行評估。

#### 2. 遷移學習 (Transfer learning)

遷移學習是一種深度學習的技術，當遇到與任務直接相關性的資料不多，或是與任務不直接相關的資料卻不少的情況，通過一個來源任務中學習到的知識，來改善下一個有相關性任務或新目標領域的性能。遷移學習的基本目的是，已經學到的知識可以被運用到新的任務上，而無需從頭開始重新學習。以下幾個為遷移學習改進模型性能的應用：

1. **特徵提取**：使用預先在大規模數據上訓練好的模型，將其作為特徵提取器。這表示模型的前幾層被當作通用特徵提取器，這些通用特徵可以在新任務上派上用場。
2. **微調 (Fine-tuning)**：通過微調來調整模型以適應新任務。這意味著可以在新任務的數據上繼續訓練模型，但不是從隨機初始化開始，而是使用預先訓練模型的權重進行初始化。
3. 應用場景：
  - 數據不足情況：在目標任務上的標記數據相對有限時，能夠利用在源任務上學到的通用特徵。
  - 不同任務之間的相似性：當源任務和目標任務之間有一定相似性時，遷移學習可以將源任務上的知識轉移到目標任務上，從而提高模型性能。
  - 加速訓練：通過使用預訓練模型，可以加速新任務上的模型訓練，因為模型已經具有了通用的知識。

#### 3. 訓練模型 ResNet18 和 ResNet50 介紹

本篇使用的預先訓練模型分別是 ResNet18 和 ResNet50，這個兩個模型皆屬於 CNN(Convolution Neural Networks)神經網路架構，在圖像處理、視覺領域中擔任重要的模型。ResNet18 和 ResNet50 在 ResNet 系列中為兩個不同深度的模型，它們之間的差別在於深度和參數量。從模型結構和性能討論，ResNet18 是屬於輕量級的深度捲積神經網路，如字面上所示它擁有 16 層捲積層和 2 層全連接層，總共 18 層，使用深度殘差塊(Deep Residual Blocks)，並引入 skip connections，不僅增強了特徵的捕捉能力，還有助於解決梯度消失的問題，因此就性能方面有相當高的準確度。而 ResNet50 為 ResNet18 的加強版，屬於較深層次的深度捲積神經網路，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，由於模型的深度增加，參數量明顯的增多。而模型結構內部與 ResNet18 相似，但不同的是它內部殘差塊使用瓶頸結構，包含三個個卷基層，分別是 1×1、3×3、1×1，這種結構透過使用 1×1 的捲積核進行降維和升維，可顯著減少參數數量，提高網路的運算效率，並透過 3×3 的捲積核負責學習目標圖像特徵，使深度神經網路能夠學習到更複雜的特徵表示。性能上相比，ResNet18 屬於輕量級的結構模型，適用於計算資源有限的圖像任務分類。而 ResNet50 則屬於更深、更複雜的結構模型，可以更好的捕捉圖像的複雜特徵，相較於 ResNet18 有更好的效能。

本篇將使用 ResNet18 和 ResNet50 這兩個預先訓練好的模型，進行組織學影像的分類，探討這兩個模型的效能差異，其載入訓練模型如 Fig.1

<pre>[ ] import torch.nn as nn import torch.nn.functional as F from torchvision import models  model = models.resnet18(weights='IMAGENET1K_V1')  # ConvNet as fixed feature extractor (freeze parameters) # for param in model.parameters(): #     param.requires_grad = False  num_ftrs = model.fc.in_features  # Here the size of each output sample is set to 6. # Alternatively, it can be generalized to ``nn.Linear(num_ftrs, len(class_names))``. model.fc = nn.Linear(num_ftrs, 6) model = model.cuda() # print(model)</pre>	<pre>[ ] 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_ftrs = model.fc.in_features  # Here the size of each output sample is set to 6. # Alternatively, it can be generalized to ``nn.Linear(num_ftrs, len(class_names))``. model.fc = nn.Linear(num_ftrs, 6) model = model.cuda() # print(model)</pre>
--	--

Fig. 1. 預先訓練的模型程式碼，分別是(左) ResNet18 和(右) ResNet50。組織學影像要對六種不同組織紋理進行分類，須將全連接的輸出特徵數量改成 6。

#### 4. 訓練結果與性能比較

本篇使用了 ResNet18 和 ResNet50 兩個預先訓練模型，並分別用 ConvNet 作為固定特徵提取器和無部分進行訓練，來比對之間的差異，這邊將分成兩個部分進行解釋。

##### 4.1 微調 ConvNet

使用的載入訓練模型如 Fig.1，微調調整全連接的輸出特徵數量改成 6，使得該模型能更好適應本次的組織學分類任務。

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 RESENT18 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，ResNet 50 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，ResNet 50 由於較深的神經網路結構，訓練花費時間為 ResNet 18 兩至三倍，其訓練和驗證精準度和損失的結果，ResNet 18 模型較好一些。而在新數據的測試中，ResNet 50 和 ResNet 18 的精準度達到 91%，如 Table1 所示。

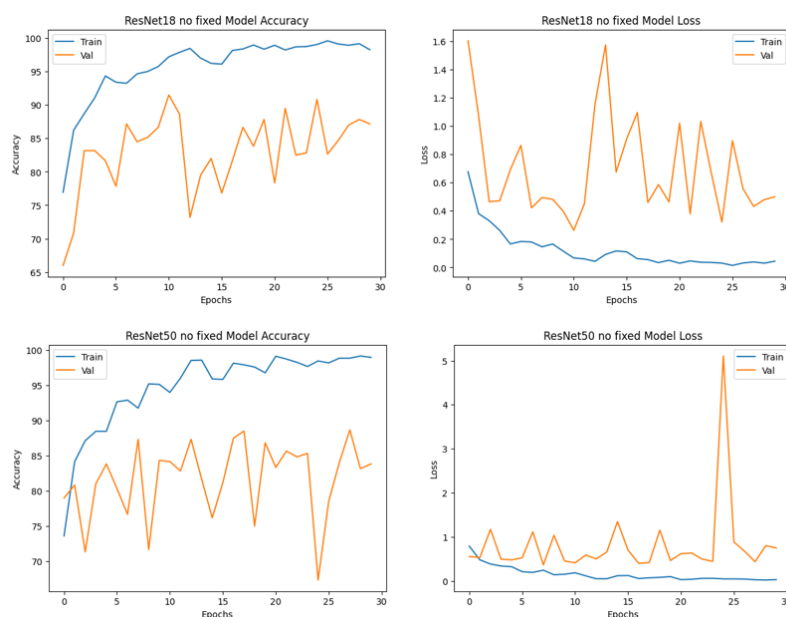


Fig. 2. 沒有固定特徵擷取器的訓練驗證精準度和損失結果。分別是(上) ResNet18 和(下) ResNet50 模型訓練結果。

#### 4.2 ConvNet 作為固定特徵提取器

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。其訓練結果如 Fig. 3 所示。在訓練時間上，和原先未凍結塗層的訓練時間相比，兩者的訓練時間有大幅度的減少，所短至原先的一半。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，兩者的訓練和驗證精準度有所提升，訓練和驗證精準度之間差距縮小，減少了過度擬合，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 86% 以上，如 Table1 所示。

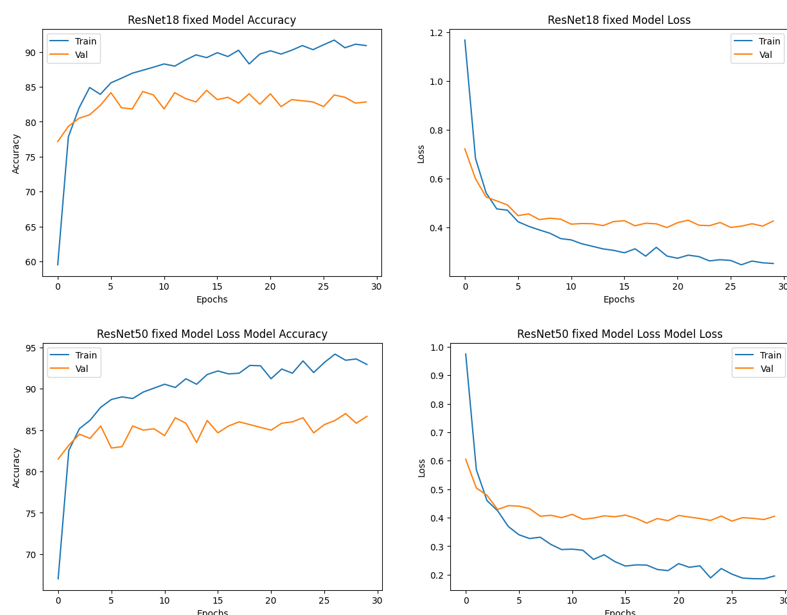


Fig. 3. 固定特徵擷取器的訓練驗證精準度和損失結果。分別是(上) ResNet18 和(下) ResNet50 模型訓練結果。

Table 1. ResNet18 和 ResNet50 在有無固定特徵擷取器模型訓練過後，在新數據中測試後的精準度結果。

	ResNet18	ResNet50
No fix	92%	91%
Fix	86%	88.3%

#### 4.3 微調 ConvNet 和 ConvNet 作為固定特徵提取器比較

4.1 微調 ConvNet 模型沒有凍結圖層，根據所給予的訓練數據，權重和參數的部分進行大幅度的更新，因此在訓練所消耗的時間相較更長，模型將被訓練成適用於組織學分類任務類型。4.2 ConvNet 作為固定特徵提取器模型則將圖層凍結，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，因此該模型適用於大數據圖形分類類型。在性能方面上，4.2 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。然而在測試新數據時，精準度相比於 4.1 模型較為低下。可能意味著任務類型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。

## 5. 結論

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。

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些微的布相關性，因此在凍結圖層的測試結果比沒凍結圖層的精準度低下。