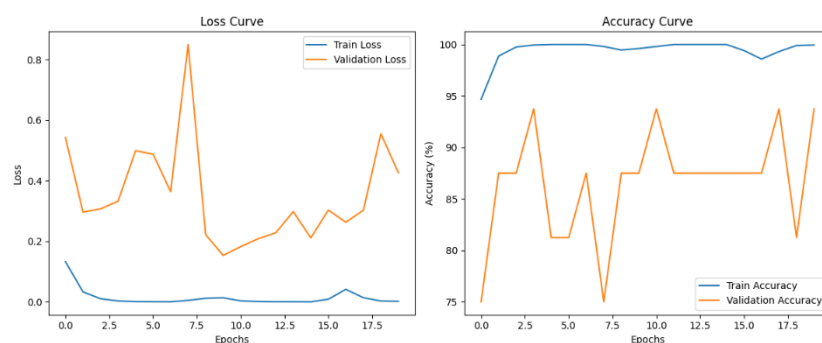


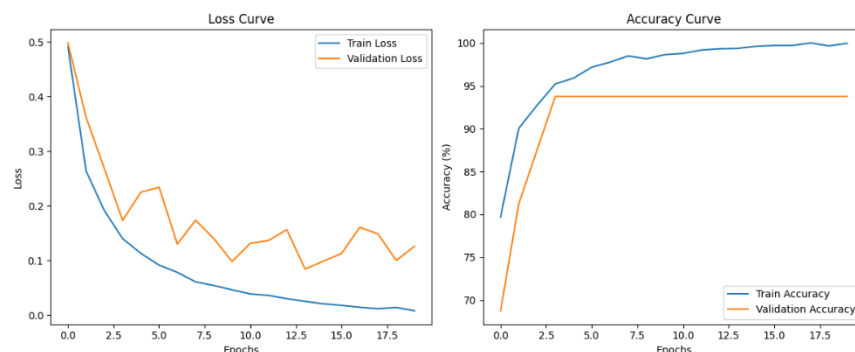
VGG 模型具有相對簡單且一致的架構，通過增加層數來增強特徵提取的深度，從而提高模型在複雜任務中的效能。尤其是 VGG19，在增強特徵提取和分類準確度上表現優異，並且由於其結構上的規律性，更便於理解和實現，因此成為許多應用場景中深度學習模型的理想選擇。

Task B

ResNeXt 模型：在 TaskB fine-tuning 中，我選擇了僅解凍第四層卷積層並凍結其餘卷積層，並且僅訓練全連接層。這樣的設計考量在於，ResNeXt 的卷積層已經在預訓練過程中學習到豐富的特徵，因此只解凍第四層以保留部分層的可調整性，同時保持模型的穩定性。這樣的微調設置使模型更好地適應二分類任務，並利用 ResNeXt 的分組卷積結構與模組化設計來增強泛化能力。根據訓練結果，最終測試準確率達到 88.44%，測試損失為 0.5356，從損失和準確率曲線可以看出訓練損失趨於穩定，驗證準確率也相對穩定，顯示出模型在微調過程中的穩定性和準確性。



VGG19 模型：在 Task B 中同樣凍結卷積層、僅訓練全連接層的設定下，VGG19 表現出穩定的特徵學習效果。由於 VGG19 的卷積層已學習到基礎的圖像特徵，因此將其固定並進行微調的策略能有效適應 2 類分類任務。VGG19 的簡單且統一的架構使其在全連接層的訓練過程中更具收斂性。最終測試準確率為 88.07%，測試損失為 0.2750。從 Loss Curve 和 Accuracy Curve 圖中可以看出，訓練與驗證損失逐漸減少且最終收斂，驗證準確率也達到穩定，這顯示 VGG19 在微調過程中具備良好的適應性和穩定性。



分析：從結果比較來看，ResNeXt 與 VGG19 的微調後準確率相近，但測試損失方面 VGG19 表現略優，顯示其模型在微調過程中較具穩定性，且能更好地適

應該任務的圖像特徵。ResNeXt 雖然在損失上略高，但其擴展性和模組化設計提供了更強的適應能力。

Task C

ResNeXt 模型：我們將 ResNeXt 模型轉變為固定特徵提取器，僅保留最後一層可訓練的全連接層，並凍結了其餘所有卷積層和全連接層。這樣的配置旨在利用 ResNeXt 模型的預訓練卷積層中學到的豐富特徵，將其作為固定特徵提取器來處理新任務中的二分類需求，而無需重新調整已提取的深層特徵。經過訓練，最終測試準確率為 88.96%，測試損失為 0.2879，從損失和準確率曲線圖中可觀察到模型在穩定過程中保持了良好的泛化性能，顯示出 ResNeXt 作為固定特徵提取器的效果。



VGG19 模型：在 VGG19 模型中，我們同樣選擇了僅訓練最後一層全連接層，並凍結所有其他層以達到固定特徵提取的效果。這樣的設定是基於 VGG19 的層次結構中各層已提取出有效特徵，最終測試準確率達到 86.25%，測試損失為 0.3218。



分析：ResNeXt

在測試準確率和損失方面表現出更好的泛化能力，這歸因於其分組卷積和模組化設計，使模型能夠捕捉更豐富的特徵。而 VGG19 由於層次結構較為單一，難以提取細微差異。兩者在訓練中都因凍結大部分層，僅訓練分類層，避免過度擬合，但 ResNeXt 的表現曲線更穩定。

Task D

在 Task B 和 Task C 中，ResNeXt 和 VGG19 模型經歷了兩種不同的遷移學習方法，即全層微調與固定特徵提取的比較，顯示出顯著的性能和適應性差異。

在 Task B 中，兩個模型均進行了全層微調。ResNeXt 的結構設計增強了多樣化的特徵提取能力，取得了 88.44% 的測試準確率；然而，由於分組卷積結構帶來的複雜性，ResNeXt 的測試損失較高。而 VGG19 則表現出較低的測試損失 (0.2750)，準確率達到 88.07%。這說明 VGG19 的簡單且一致的卷積架構在全層微調下更能保持穩定的分類性能。

在 Task C 中，兩個模型只進行最後全連接層的微調，以固定特徵提取器的方式應用。ResNeXt 在此設定下展現出更強的穩定性和泛化能力，其測試準確率上升至 88.96%，測試損失也顯著降低至 0.2879，顯示其在固定特徵提取的任務中適應性較佳。相較之下，VGG19 的測試準確率降至 86.25%，測試損失也略有上升，反映出其在僅依賴前層特徵提取時缺乏 ResNeXt 的適應力。

整體來看，當進行全層微調時，VGG19 的穩定性較優；但在固定特徵提取方式下，ResNeXt 展現了更佳的泛化性和適應性。這表明 ResNeXt 的結構更適合在固定特徵提取任務中應對多變的資料，而 VGG19 的效能則更依賴其層層卷積的穩定提取特徵能力。

Task E

在實驗過程中，我們嘗試在測試資料集上提升 ResNeXt 和 VGG19 模型的性能，但遇到了顯著的困難。這一現象背後有幾個可能的原因：

首先，**over-fitting** 是造成測試資料集表現不理想的主要原因之一。在 Task B 中，我們對模型進行了全層微調，這使模型能夠高度適應訓練資料。然而，過度的調整也可能讓模型學習到訓練資料中的細微差異和噪音，而這些特徵並不適用於測試資料。當模型在測試集上進行預測時，由於這些無法泛化的特徵，導致性能提升受到限制。

其次，**訓練資料與測試資料之間的分佈差異**也是一個重要因素。由於測試資料可能包含與訓練資料不同的圖像特徵或噪音，模型難以對測試資料作出準確預測。尤其是在 VGG19 模型的 Task C 固定特徵提取設定下，我們僅微調最後的全連接層，這意味著模型在依賴既有的卷積特徵上進行預測，當測試資料

的特徵分布偏離訓練資料時，性能自然會受到影響。

另外，**模型架構的適應性和靈活性**也是值得關注的因素。在 ResNeXt 中，分組卷積的結構增強了模型的適應能力，使其在 Task C 固定特徵提取的情況下仍能展現相對穩定的表現。然而，這一結構也增加了模型的複雜性，導致在訓練全層微調 (Task B) 時更容易過擬合。因此，當我們嘗試在測試資料集上提高性能時，ResNeXt 在不同設定下的複雜性反而成為了阻礙因素。而 VGG19 的單一卷積結構在特徵提取上穩定性較佳，卻也因此在此類任務中難以進一步適應測試資料的多樣性，表現相對受限。

總結來說，在測試資料集上難以提升性能的原因，主要來自於 over-fitting、資料分布差異以及模型架構適應性三方面的限制。這些因素共同導致了模型在訓練資料上效果不錯，但一旦應用於測試資料集，性能提升的空間變得有限。