National Tsing Hua University

11220IEEM 513600

Deep Learning and Industrial Applications

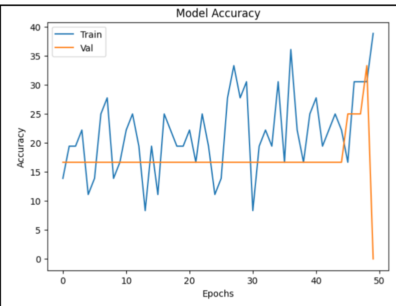
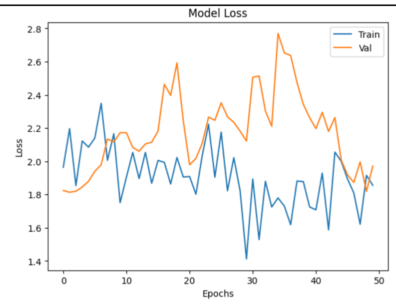
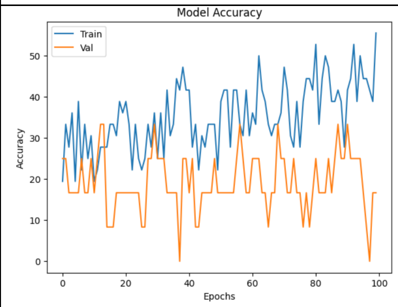
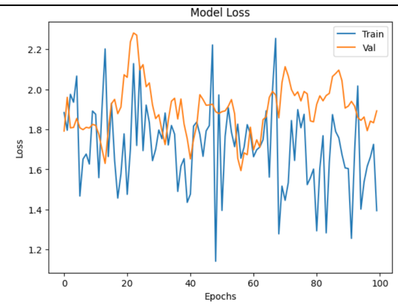
Homework 3

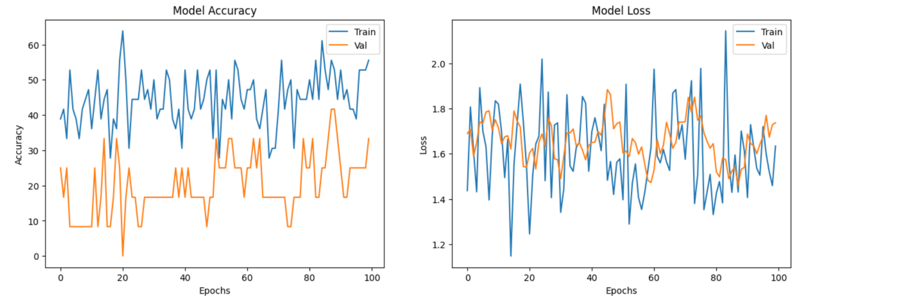
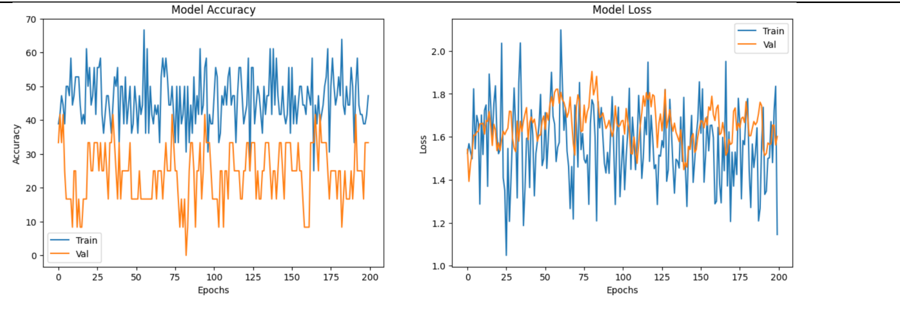
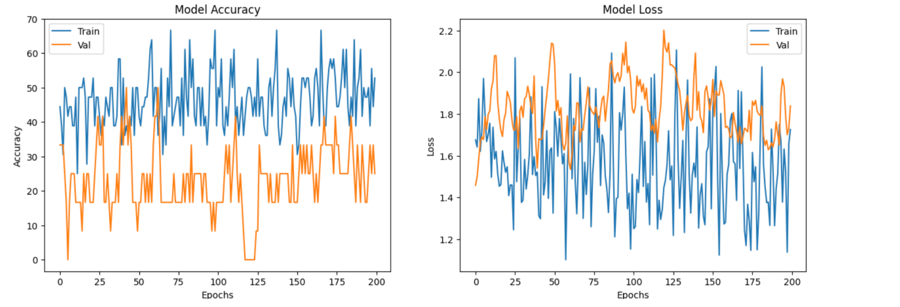
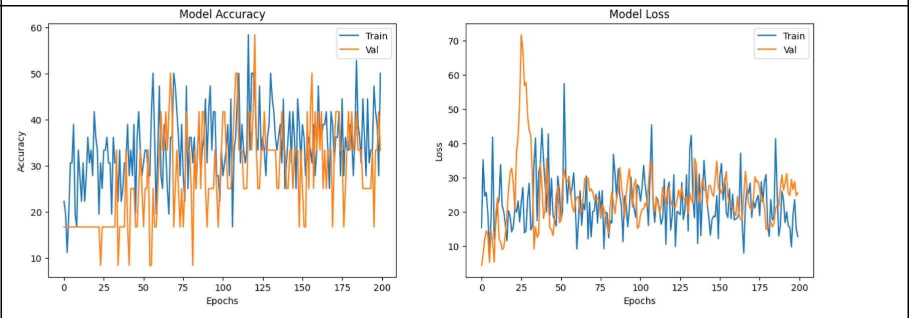
Name: 盧育琦 Student ID: 112034582

Due on 2024/04/11.
Note: DO NOT exceed 3 pages.

1. (10 points) Download the MVTec Anomaly Detection Dataset from Kaggle ([here]()). Select one type of product from the dataset. Document the following details about your dataset: **wood**
   - Number of classes: **6**
   - Types of classes: **color, combined, hole, liquid, scratch, good**
   - Number of images used in your dataset: **79(test)+307(train)=386**
   - Distribution of training and test data: **0.8, 0.2**
   - Image dimensions**: 3*1024*1024**

2. (30 points) Implement **4** different attempts to improve the model's performance trained on the dataset you choose in previous question. Ensure that at least one approach involves modifying the pre-trained model from TorchVision. Summarize the outcomes of each attempt, highlighting the best performing model and the key factors contributing to its success. You may also need to describe other hyperparameters you use in your experiment, like epochs, learning rate, and optimizer. (Approximately 150 words.)

Epoch: 50 Lr: 1e-3 Optimizer: Adam Train_loss: 1.86 Train_acc: 38.89% Val_loss: 1.97 Val_acc: 0% Test_acc: 33.33%	 
Epoch: 100 Lr: 1e-3 Optimizer: Adam Train_loss: 1.39 Train_acc: 55.56% Val_loss: 1.89 Val_acc: 16.67% Test_acc: 33.33%	 

Epoch: 100 Lr: 1e-3 Optimizer: SGD Train_loss: 1.63 Train_acc: 55.56% Val_loss: 1.74 Val_acc: 33.33% Test_acc: 41.67%	
Epoch: 200 Lr: 1e-3 Optimizer: SGD Train_loss: 1.14 Train_acc: 47.22% Val_loss: 1.60 Val_acc: 33.33% Test_acc: 41.67%	
Epoch: 200 Lr: 1e-3 Optimizer: Adam Train_loss: 1.73 Train_acc: 52.78% Val_loss: 1.84 Val_acc: 25% Test_acc: 50%	
Epoch: 200 Lr: 1e-1 Optimizer: Adam Train_loss: 12.79 Train_acc: 50% Val_loss: 25.56 Val_acc: 33.33% Test_acc: 58.36%	

觀察以上發現，訓練次數增加通常對模型的性能有所改善，但也使 model_loss 變得較不穩定。而提高學習率不一定可以增加模型效能，反而會導致訓練不穩定或過度擬合。另外，當 optimizer 為 SGD 時，相較於 Adam 更適合用在訓練次數少的時候。

3. (20 points) In real-world datasets, we often encounter long-tail distribution (or data imbalance). In MVTec AD dataset, you may observe that there are more images categorized under the 'Good' class compared to images for each defect class. (Approximately 150 words.)

   i 長尾分佈是指在數據集中，某些類別或數據點的頻率相較於其他來得非常低，這導致了數據集的分佈在頻率較低的端部延伸出一條長尾。

ii 《Learning from Imbalanced Data》(2020)，提出了一種名為「Re-weighting Loss」的解決方案來應對數據不平衡問題。該方法通過調整每個樣本的損失函數的權重，使少數類別的樣本被賦予更大的權重，從而平衡類別之間的重要性。這種方法可以應用於 MVTec AD 資料集中，使缺陷類別的影像在訓練中更加重要，從而改善模型在少數類別上的性能。

4. (20 points) The MVTec AD dataset's training set primarily consists of 'good' images, lacking examples of defects. Discuss strategies for developing an anomaly detection model under these conditions. (Approximately 100 words.)
   - Data Augmentation：使用各種數據增強技術，如旋轉、翻轉、縮放、添加噪聲等，來生成更多的合成缺陷範例。這有助於模型更好地學習並辨識不同類型的缺陷。
   - Semi-Supervised Learning：利用已知的「良好」影像作為訓練集，然後使用未標記的影像來進行半監督學習，從而擴展訓練數據集並提高模型的泛化能力。
   - Generative Adversarial Networks，GANs：使用 GANs 生成逼真的缺陷影像，以補充訓練集中的缺陷範例。這樣可以幫助模型學習更多關於缺陷的特徵。
   - Class Imbalance Handling：使用類別不平衡處理技術，如權重調整、過採樣或者欠採樣，來平衡訓練集中「良好」影像和缺陷範例之間的類別分佈。

5. For the task of anomaly detection, it may be advantageous to employ more sophisticated computer vision techniques such as object detection or segmentation. This approach will aid in identifying defects within the images more accurately. Furthermore, there are numerous open-source models designed for general applications that can be utilized for this purpose, including YOLO-World (website) and SAM (website). (Approximately 150 words.)

   i **影像標籤**：每張影像都需要標記其包含的物件或缺陷的位置和類別信息。**適當的樣本數量**：資料集應包含足夠的樣本以涵蓋各種不同類型的缺陷和場景。**多樣性和覆蓋性**：資料集應該包含各種不同大小、形狀和位置的缺陷，以便模型能夠在各種情況下進行準確檢測和分割。

   ii 因為它們已經在大型資料集上進行了預訓練，具有良好的通用性和泛化能力。透過微調，可以將這些模型的特徵提取能力和檢測能力遷移到自訂的資料集上，從而實現更準確的異常檢測。此外，開源模型如 YOLO-World 和 SAM 已經經過廣泛測試和改進，提供了許多方便的工具和功能，使得在特定應用中進行微調變得相對容易。