The Image Classification of Pneumonia vs Normal X-ray Images Using ResNet-18

1 Introduction

Pneumonia is a life-threatening respiratory disease that is a leading cause of death in children under the age of five worldwide. Early detection of pneumonia through medical imaging, particularly chest X-ray (CXR), plays a crucial role in timely diagnosis and treatment. Recently, machine learning models, particularly Convolutional Neural Networks (CNNs), have shown good potential in automating the detection of pneumonia from X-ray images, significantly improving diagnostic accuracy. This report presents the implementation of a deep learning-based model, specifically ResNet-18, for the classification of chest X-ray images into two categories: Pneumonia and Normal. The task is divided into two sub-tasks:

Task 1.1: Training ResNet-18 from scratch (random initialization).

Task 1.2: Fine-tuning a pre-trained ResNet-18 model on the target X-ray dataset.

2 Implementation Details

Task 1.1 - Training ResNet-18 from Scratch

For this task, ResNet-18 model was initialized from scratch (i.e., without pre-trained weights) and trained on the on the chest X-ray dataset. The architecture used for this task is based on the ResNet-18 model from the torchvision library. The dataset used for the implementation is publicly available at: Chest X-Ray Images (Pneumonia)

Model Architecture: The ResNet-18 architecture consists of 18 layers. The original final layer of ResNet-18 was modified to have two output units (since binary classification is being performed: Pneumonia vs Normal). The modified layer is:

model.fc = nn.Linear(model.fc.in_features, 2)

Learning Rate: The learning rate for training from scratch is set to 0.001. This is a commonly used learning rate for training deep learning models. The Adam optimizer was used to adjust the learning rate during training.

Batch Size: A batch size of 32 was used for training. This provides a balance between memory usage and training efficiency.

Epochs: The model was trained for 10 epochs. Early stopping or further tuning could be applied to prevent overfitting or underfitting, but for this report, 10 epochs were implemented, and it took more than seven hours to complete.

Loss Curves for Training and Validation: To monitor the model's training progress, the training and validation loss curves were plotted for both sub-tasks, as shown in Figure 1 and 2. These curves help diagnose whether the model is underfitting or overfitting.

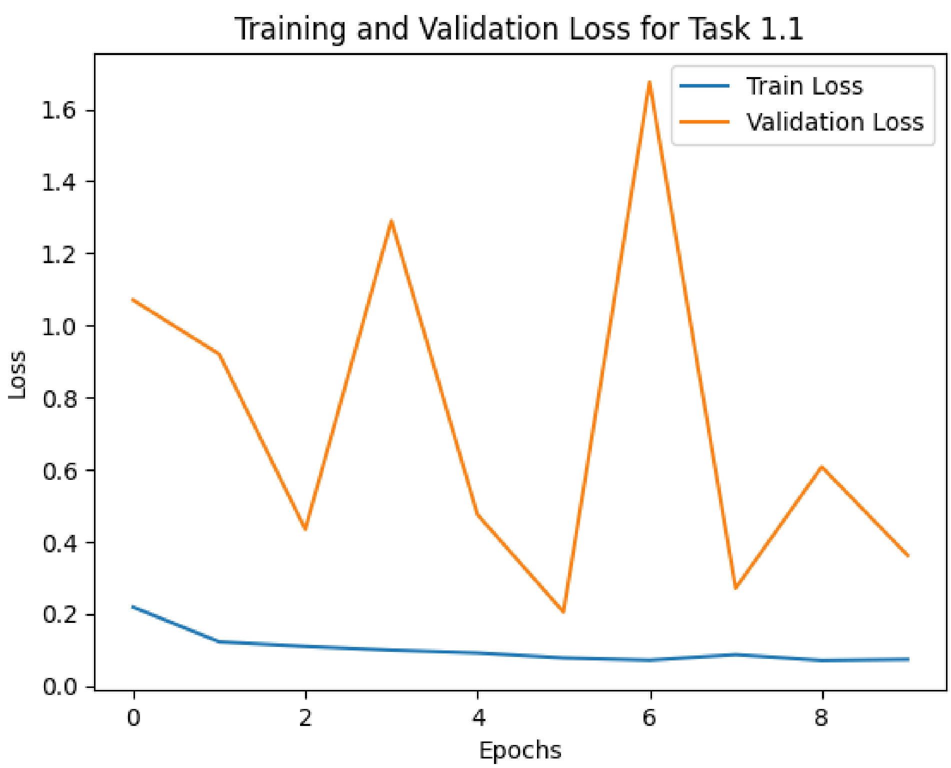

Figure 01: Training and Validation for Task 1.1

Task 1.2 - Fine-Tuning Pre-trained ResNet-18

In Task 1.2, a pre-trained ResNet-18 model (pre-trained on the ImageNet dataset) was used and it was fine-tuned on the chest X-ray dataset. Fine-tuning allows the model to adapt the learned features from ImageNet to the X-ray classification task.

Pre-trained Model: The ResNet-18 model pre-trained on ImageNet was loaded using:

```
pretrained_model = models.resnet18(pretrained=True)

pretrained_model.fc = nn.Linear(pretrained_model.fc.in_features, 2)
```

Freezing Layers: All layers except the final fully connected layer were frozen to prevent updates to the weights during fine-tuning. This reduces the computational cost and prevents overfitting:

```
for param in pretrained_model.parameters():
    param.requires_grad = False
pretrained_model.fc.requires_grad = True
```

Learning Rate: For fine-tuning, a smaller learning rate of 0.0001 was used to prevent drastic updates to the pre-trained weights, allowing the model to adjust more gradually to the target task.

Batch Size and Epochs: The batch size and number of epochs were kept the same as Task 1.1, with 32 for the batch size and 10 epochs for training.

Fine-Tuning Pre-trained Model

The fine-tuned model exhibited faster convergence and lower validation loss than the model trained from scratch, demonstrating the advantages of using a pre-trained model.

Figure 02: Training and Validation for Task 1.2

3. Classification Accuracy on Test Set

The model's performance was evaluated on the test set, which consists of X-ray images labeled as either "Pneumonia" or "Normal." The overall classification accuracy on the test set for the model trained was calculated, and then for pneumonia and normal cases separately. The accuracy for Pneumonia is 91.45%.

4. Failure Cases and Misclassifications

To further analyze the model's performance, several misclassified images were visualized. These images were examined to understand the model's failure cases. Additionally, to better understand the model's decisions, Class Activation Mapping (CAM) technique was used to visualize the important regions of the X-ray images that contributed to the model's classification decision. This is particularly useful for analyzing misclassifications and understanding what the model is focusing on.

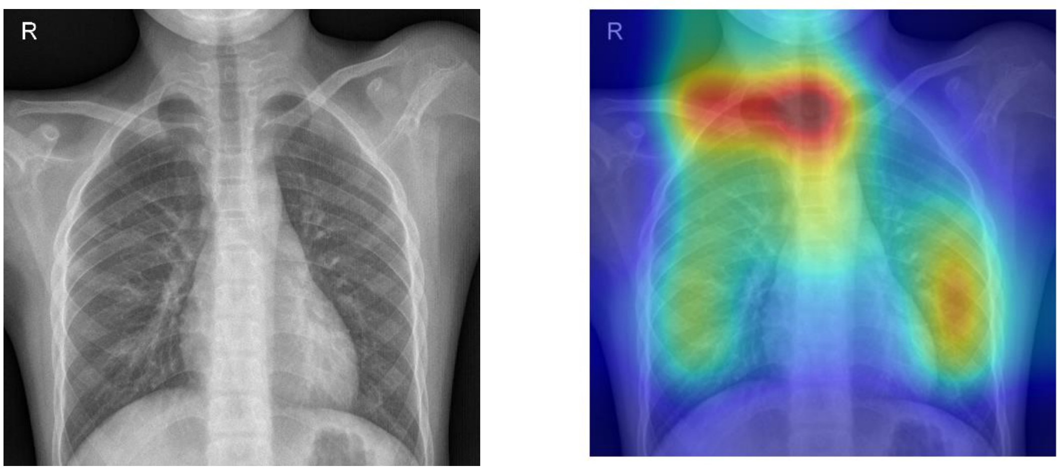

Figure 03: Original vs CAM (normal)

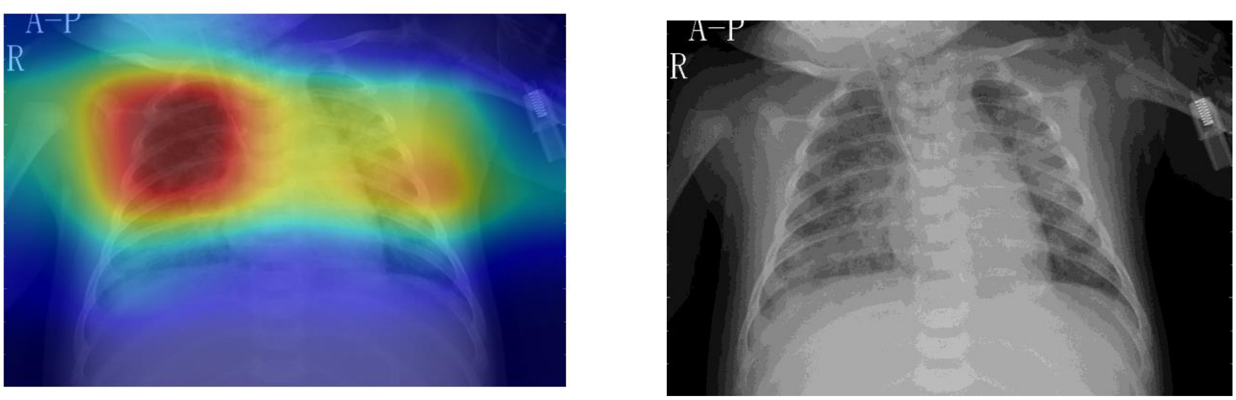

Figure 04: Original vs CAM (pneumonia)

5. Conclusion

In this report, we implemented two approaches for classifying chest X-ray images into Pneumonia and Normal categories using ResNet-18. The accuracy for pneumonia class was 91.45%, however, more training and processing is required in order to improve the accuracy and keep it up to the mark.