Assignment Task Detail (PneumoniaMNIST Data)

Pragya Kumari

June 8, 2025

1 Model for Transfer Learning

Used ResNet-50 pre-trained on ImageNet as base model

2 Evaluation Strategy

2.1 Chosen Metrics and Justification:

2.1.1 Accuracy

- Measures overall correctness: (TP + TN) / (TP + TN + FP + FN)
- Justification: Provides a general sense of model performance but can be misleading in imbalanced datasets (e.g., if 90% of cases are normal, a model predicting ”normal” always would achieve 90% accuracy).
- Limitation: Not sufficient alone for imbalanced medical datasets.

2.1.2 Precision

- Measures correctness of positive predictions: TP / (TP + FP)
- Justification: Critical to minimize false positives (misdiagnosing healthy patients as having pneumonia, which could lead to unnecessary treatments).
- Impact: High precision means fewer false alarms.

2.1.3 Recall

- Measures ability to detect true positives: TP / (TP + FN)
- Justification: Essential to minimize false negatives (missing actual pneumonia cases, which could be life-threatening).
- Impact: High recall means fewer missed diagnoses.

2.1.4 F1-Score (Harmonic Mean of Precision and Recall)

- Balances precision and recall: 2 * (Precision * Recall) / (Precision + Recall)
- Justification: Useful when both false positives and false negatives are costly (typical in medical applications).

2.2 Class Imbalance Handling

- Calculated class weights using class_weight.compute_class_weight(‘balanced’)
- Applied weights during training to give more importance to minority class
- Visualized class distribution through random samples
- Used metrics (precision/recall) that are more informative with imbalanced data

2.3 Measures Taken to Prevent Over-fitting

The code implements multiple strategies to prevent over-fitting:

2.3.1 Data Augmentation:

- Random horizontal flips
- Small rotations/zooms
- Brightness/contrast adjustments

Impact: Increases effective dataset size, reducing over-reliance on specific training examples.

2.3.2 Regularization:

- Dropout (0.5) in classification head
- L2 regularization could be added to dense layers

Effect: Randomly drops 50% of neurons during training, forcing the network to learn redundant features.

2.3.3 Early Stopping:

- Monitored validation loss with patience
- Restored best weights

Effect: Stops training if validation loss does not improve for 5 epochs, preventing overfitting to noise in training data.

2.3.4 Two-Phase Training:

- First trained only new layers (prevents destructive updates to pre-trained features early in training).
- Then fine-tuned with very low learning rate (ensures gradual adaptation without catastrophic forgetting).