Chest X-ray Classification: Normal vs Pneumonia

Deep Learning Project Report

(Spandana Damannagari, Sunaina Makkena)

1. Introduction & Objective

This project demonstrates how deep learning can be used for medical image classification. It focuses on classifying chest X-ray images as either Normal or Pneumonia using a ResNet50 model implemented in PyTorch. The objective is to build an accurate and interpretable model capable of aiding radiologists in identifying pneumonia cases efficiently.

2. Deep Learning Concepts Used

- Supervised Learning: Uses labeled data to learn mappings between input (X-ray) and output (Normal/Pneumonia).
- Transfer Learning: Employs a pre-trained ResNet50 model to leverage existing visual features.
- Fine-Tuning: Initially freezes pre-trained layers, then unfreezes for further optimization.
- Data Augmentation: Applies transformations like flips and rotations to prevent overfitting.
- Normalization: Ensures pixel intensity distribution matches pretrained model expectations.
- Cross-Entropy Loss: Measures the prediction error for classification tasks.
- Optimizers: Adam and SGD with momentum for efficient training.
- Automatic Mixed Precision (AMP): Reduces memory use and increases speed on GPU.
- Gradient Clipping: Prevents exploding gradients during updates.
- Early Stopping: Stops training when performance no longer improves.
- Grad-CAM: Provides interpretability by visualizing focus areas.
- Learning Rate Finder: Helps determine optimal learning rate range.
- Hyperparameter Sweep: Evaluates multiple configurations for best results.

3. Project Workflow (Step-by-Step Explanation)

3.1 Setup & Imports

The notebook begins by importing necessary libraries (torch, torchvision, sklearn, matplotlib). It detects GPU availability and initializes mixed precision training for efficiency. AMP ensures computations run faster with reduced precision where safe.

3.2 Dataset Preparation & Augmentation

The dataset is organized into train, validation, and test splits, each containing subfolders for Normal and Pneumonia. The notebook automatically handles upload and extraction if not found locally. Augmentation introduces variability, improving model robustness.

3.3 DataLoaders

DataLoaders batch images for training and validation. Shuffling ensures model generalization. Pin memory is used only for GPU systems to improve data transfer speeds.

3.4 Model Architecture (ResNet50)

A pre-trained ResNet50 from ImageNet is used. The convolutional backbone is initially frozen to retain learned features. The final fully connected layer is replaced with a two-class output head for binary classification.

3.5 Training Phase

Training occurs in two stages: (1) only the classifier head is trained, (2) the entire network is fine-tuned with a lower learning rate. Gradient clipping and early stopping are used to stabilize and optimize training. The best-performing model is saved.

3.6 Evaluation & Metrics

Evaluation uses accuracy, precision, recall, F1-score, and AUC. The confusion matrix visualizes true vs. predicted outcomes, while the ROC curve shows classification confidence.

3.7 Grad-CAM Explainability

Grad-CAM highlights the regions in X-rays that most influenced the model's decision, improving interpretability. It overlays a heatmap on the image, showing attention areas (e.g., infected lung regions).

3.8 Learning Rate Finder

A learning rate range test determines the best learning rate by observing loss changes over exponentially increasing learning rates. The point before loss explosion suggests the ideal starting LR.

3.9 Hyperparameter Sweep

This grid search tests combinations of optimizer types, learning rates, weight decay, and batch sizes to identify optimal training configurations based on validation accuracy.

4. Results & Observations

The fine-tuned ResNet50 achieved high classification accuracy with strong AUC scores, confirming transfer learning's effectiveness. Grad-CAM visualizations demonstrated the model's focus on meaningful lung regions, validating interpretability. Data augmentation proved essential in reducing overfitting.

5. Conclusion & Future Work

This deep learning project successfully classifies chest X-rays into Normal or Pneumonia categories. The workflow integrates best practices such as transfer learning, fine-tuning, mixed precision, and explainability. Future enhancements include expanding to multi-disease detection and deploying as a web-based medical aid tool.