APS360 Project Final Presentation - Group 60

Adam Roberge, Bill Jiang, Chris Kwon, Mitchell Souliere-Lamb

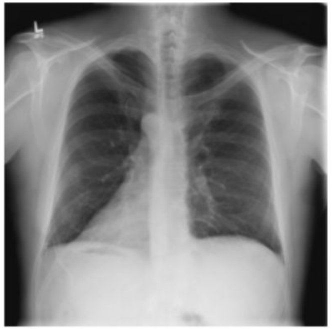
Problem Definition

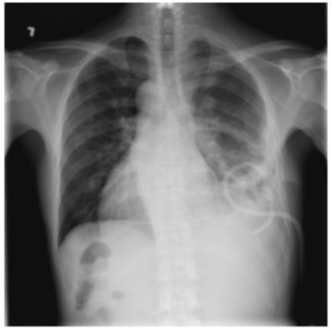
Problem Statement:

- **Objective:** Develop a deep learning model to classify chest X-ray images into various categories, including different types of lung diseases
- **Significance:** Assist radiologists in diagnosing diseases more accurately and efficiently, particularly in regions with a shortage of medical professionals

Importance of the Problem:

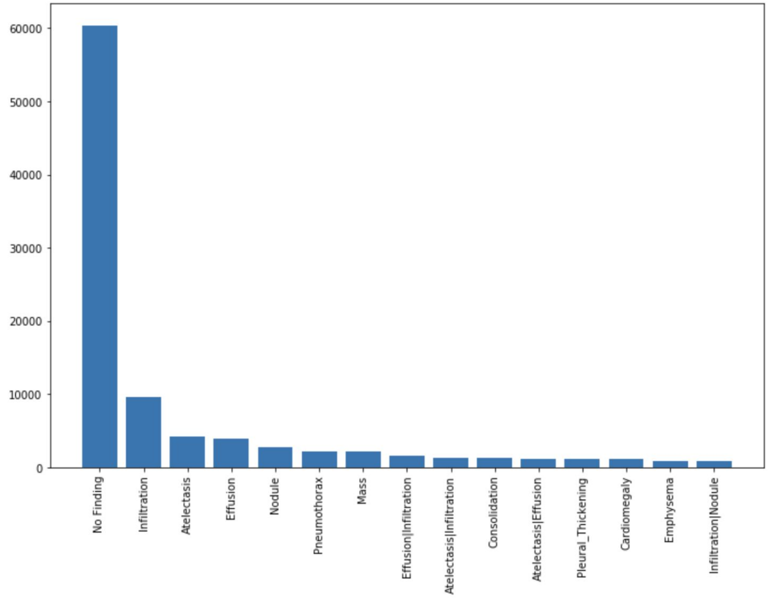
- **Enhancing Diagnostic Accuracy:** Utilize advanced machine learning techniques to improve the precision of diagnoses, potentially saving lives by catching diseases early





Data Overview

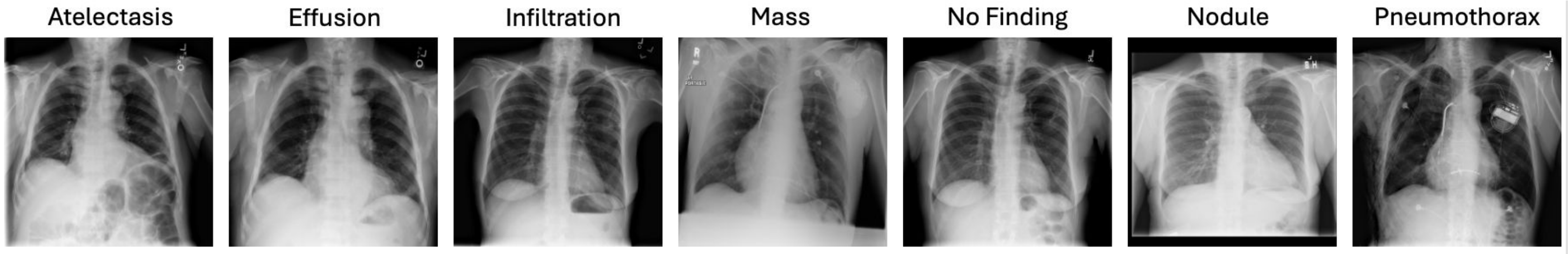
- NIH Chest X-ray Dataset
    - 112,000 images from the NIH dataset
    - Class imbalance
    - Mislabeled entries, missing data, inconsistencies in dataset's .csv files
- Data Cleaning
    - Class filtering, undersampling, and oversampling to achieve a uniform distribution of 12,159 images per class across 7 classes
    - **Final Dataset:** Narrowed down to 85,113 images



Class	Original Number of Images	Final Number of Images
Atelectasis	5,389	12,159 (+6,770)
Effusion	4,763	12,159 (+7,396)
Infiltration	9,627	12,159 (+2,468)
Mass	2,842	12,159 (+9,317)
No Finding	54,379	12,159 (-42,220)
Nodule	3,687	12,159 (+8,472)
Pneumothorax	2,942	12,159 (+9,217)

Data Processing

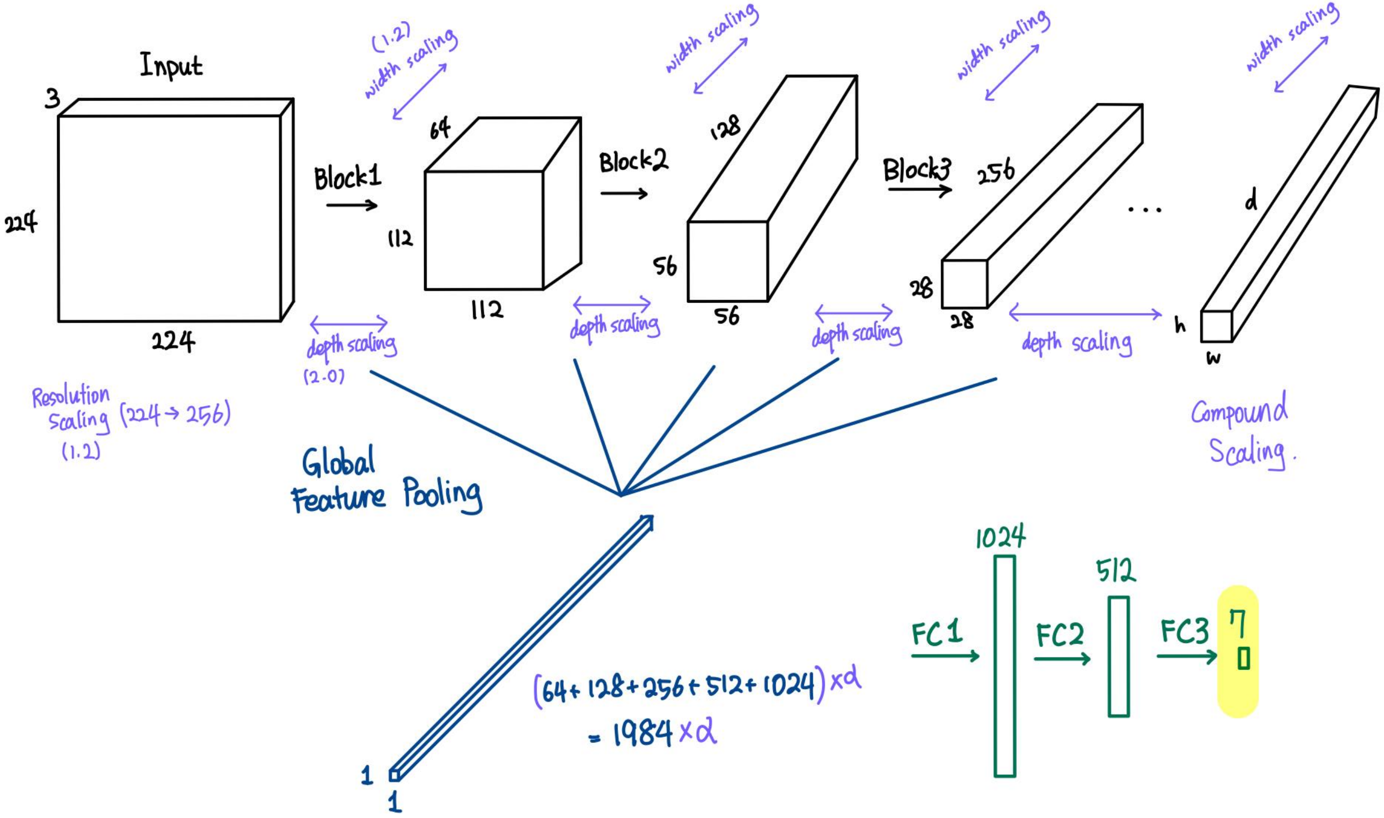
- CSV File Generation
  - Reorganized images into class-specific folders and updated the .csv file to reflect accurate, single-label classifications.
- Class Mapping and Dataset Creation
  - Standardized label representation by mapping class names to unique integer indices
- Data Augmentation
  - Resizing, cropping, flipping, rotation, color jittering, affine transformations, and random erasing
  - Training (70%) [*with augmentation*] ; validation (15%) & testing (15%) [*without augmentation*]



Model Architecture

- **Base Model**
  - Enhanced MultiLevelUNet built upon traditional U-Net structure
  - Tailored for complex medical image classification tasks
  - Combines CNN principles with advanced scaling techniques inspired by EfficientNet
- **Key Enhancements and Components:**
  - **Scaling** of width, depth, and resolution from EfficientNet to optimize capacity and computational efficiency
  - Vision Transformer-inspired **attention layers** were conducted but eventually not included
  - **Core Components**: Double Convolution Blocks, Residual Blocks with skip connections, downsampling path, bottleneck structure to enhance feature extraction and learning efficiency

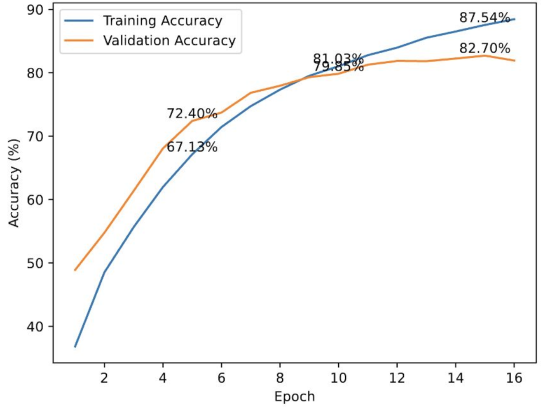
Demonstration

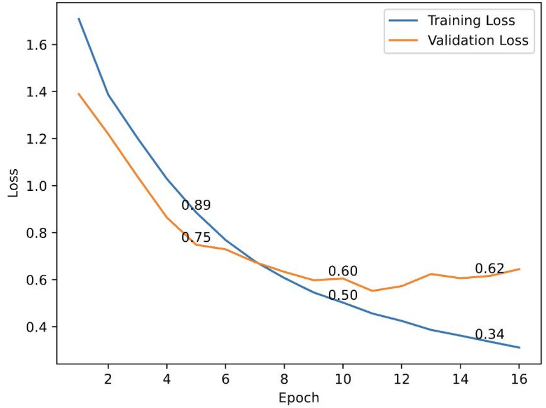


Quantitative Results

Efficient MultiLevelUNet

- ❖ Training accuracy peaked at **87.54%**
  - ➢ *Strong learning capabilities*
- ❖ Validation accuracy reached **82.70%**
  - ➢ *Good generalization to unseen data*
- ❖ Training loss reduced from **1.6 to 0.34**
  - ➢ *Effective error minimization*
- ❖ Validation loss decreased to **0.62**
  - ➢ *Well-balanced model*

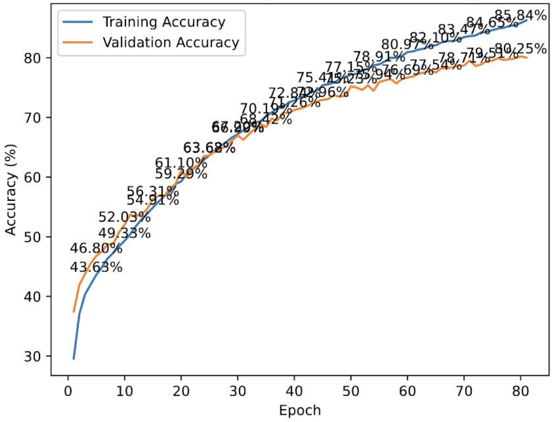




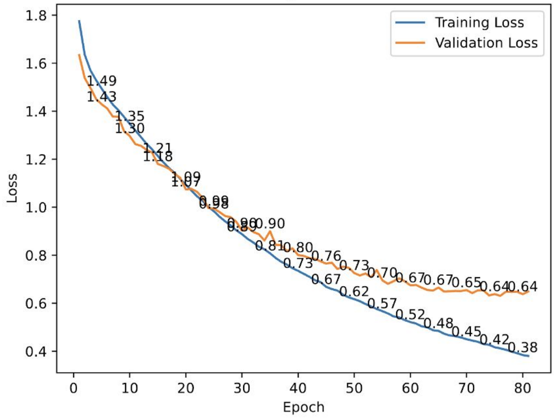
ResNet18 Baseline

- ❖ Training accuracy of **85.84%**
- ❖ Validation accuracy of **80.25%**
  - ➢ *Slightly lower than MultiLevelUNet, but showing good generalization*
- ❖ Training loss improved significantly to **0.38**
  - ➢ *Efficient learning*
- ❖ Validation loss stabilized at **0.64**
  - ➢ *Steady model convergence*

ResNet18 Training and Validation Accuracy



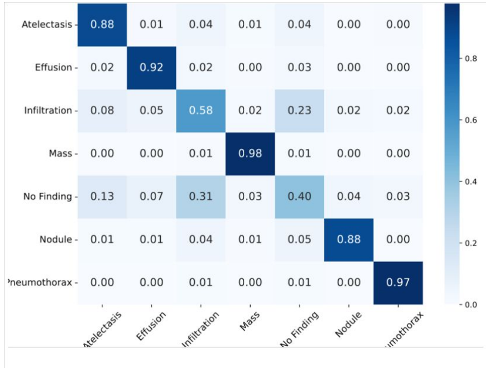
ResNet18 Training and Validation Loss

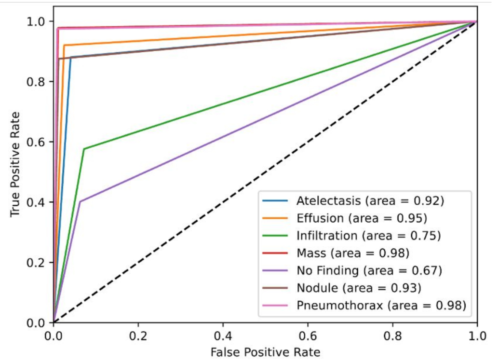


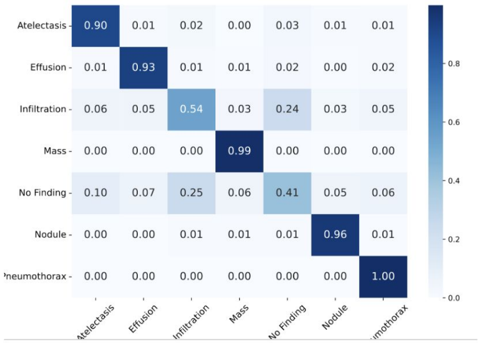
Qualitative Results

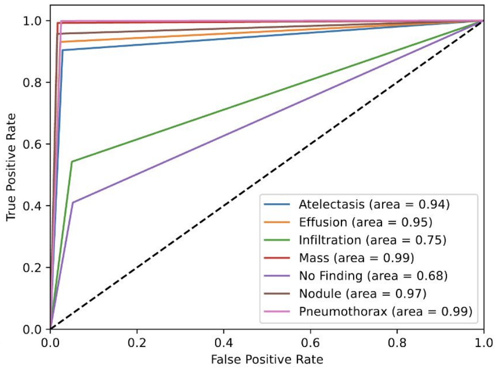
Efficient MultiLevelUNet

- ❖ **Confusion Matrix**
  - ➢ Excellent accuracy in classifying 'Mass' and 'Pneumothorax' with minimal errors
  - ➢ Struggles with 'Infiltration', showing confusion with similar conditions like 'No Finding'
- ❖ **ROC Curve**
  - ➢ Strong discriminative ability with AUC values near or above 0.95 for most classes,
  - ➢ Infiltration' has a lower AUC of 0.75









ResNet18 Baseline

- ❖ **Confusion Matrix**
  - ➢ Demonstrates high accuracy for 'Mass' and 'Pneumothorax'
  - ➢ Faces challenges with 'Infiltration'
- ❖ **ROC Curve**
  - ➢ The ROC curve shows strong performance with AUC values above 0.90 for most classes
  - ➢ Similar to the Efficient MultiLevelUNet, 'Infiltration' has a lower AUC

Takeaways

1. Importance of Data Quality
   - ❑ Label accuracy and balance of medical image data directly influence model performance
2. Model Complexity vs. Performance
   - ❑ Explored the impact of incorporating advanced features like **attention mechanisms**, which may not always lead to improved outcomes
   - ❑ Gained insight into optimizing model architecture using **scaling techniques** that enhanced computational efficiency without compromising on accuracy
3. Generalization and Overfitting
   - ❑ Implemented strategies such as **early stopping and rigorous validation** to prevent overfitting, ensuring that the model generalizes well to new, unseen data
   - ❑ Crucial in maintaining the balance between training and validation performance
4. Ethical Considerations and Impact
   - ❑ Recognized the **ethical implications** of deploying machine learning models in healthcare, particularly the consequences of potential misdiagnoses
   - ❑ Ssing machine learning tools as assistive technologies rather than standalone solutions

Thank You for Listening !