Model Card - Pneumonia Detection with X-rays using ResNet-18

Model Details

- Developed in 2015 by researchers at Microsoft Research
- Introduces residual connections to mitigate vanishing gradients in deep networks
    - Model Type: Image classifier (Convolutional Neural Network)
    - Architecture: ResNet-18 (pretrained on ImageNet, final layer replaced)
    - Framework: PyTorch
- Has been compared to MobileNet and ResNet50. ResNet18 has shown the best results
- Training Algorithms & Parameters:
    - Optimizer: Adam (learning rate=1e-4, weight decay=1e-4)
    - Loss: CrossEntropyLoss
    - Scheduler: StepLR (step size=7, gamma=0.1)
    - Epochs: 20
    - Batch size: 32
    - Data Augmentation: RandomResizedCrop, RandomHorizontalFlip (train only)

Intended Use

- Assist radiologists in detecting pneumonia from chest X-ray images
- Intended to be used by healthcare professionals, researchers, and medical imaging developers
- Primary Use Case: Automated classification of chest X-ray images into:
    - NORMAL
    - PNEUMONIA_BACTERIA
    - PNEUMONIA_VIRUS
- Users:
    - Medical researchers, radiologists, and healthcare professionals for research and decision support
- Out-of-Scope Uses:
    - Not intended for direct clinical diagnosis or as a substitute for professional medical judgment

Factors

- X-Ray Image Quality
- Virus and Bacteria types of Pneumonia

Metrics

- Accuracy
- Reported Test Accuracy:
    - 82.97% on 3-classification
    - 96.94% on 2-class-classification
- Confusion Matrix
- Precision, Recall, F1 Score, Matthews Correlation Coefficient (MCC)

Training Data

- Chest X-RAY Images (Pneumonia)
- Data is split into Normal, Pneumonia Bacteria, and Pneumonia Virus
- Train/Test/Val split: (70/15/15)
- Preprocessing:
    - Normalize Images
    - Random cropping of images to 244x244

Evaluation Data

- Chest X-RAY Images (Pneumonia)
- Preprocessing:
    - Normalize Images
    - Images resized to 244x244

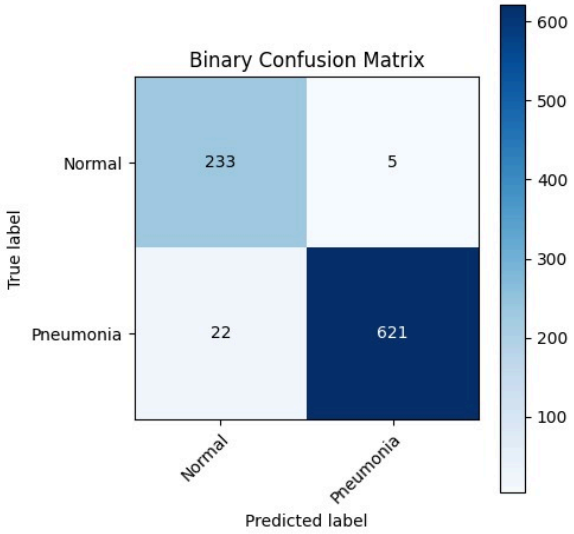
Ethical Considerations

- The diagnoses for the images were graded by two expert physicians before being cleared for training the AI system. In order to account for any grading errors, the evaluation set was also checked by a third expert
- Designed for diagnostic support, not standalone diagnosis

Caveats and Recommendations

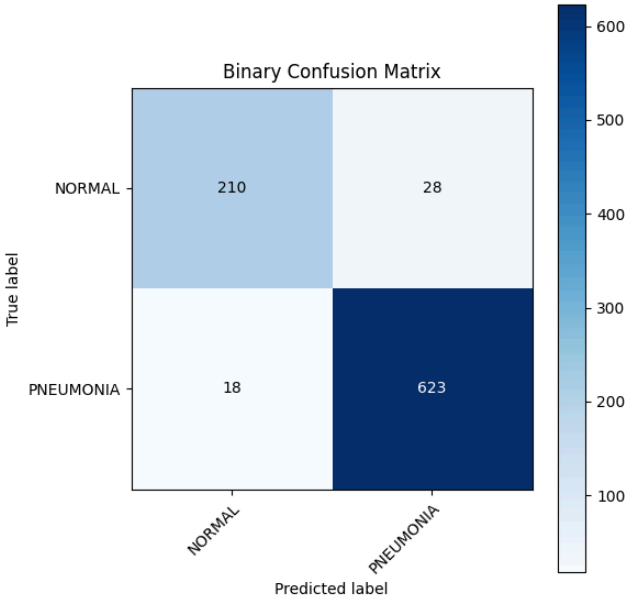
- No data about patient age, sex or other factors was used in training, which could improve the model.
- No data was provided about the calibration of the instrument before taking the pictures, which could create noise and unclarity on the images
- Should be used as an assistant system for a professional and not as a final diagnosis

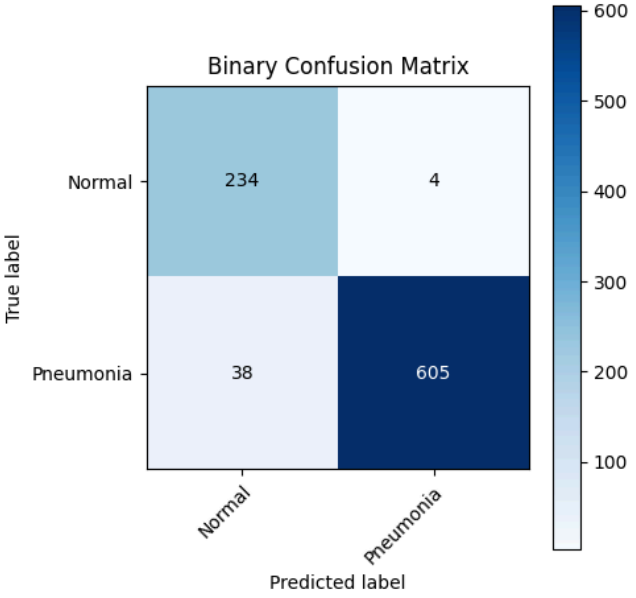
Quantitative Analyses



Binary Accuracy	0.9694
Three-Class Accuracy	0.8297
Precision	0.9920
Recall (Sensitivity)	0.9658
Specificity	0.9790
F1 Score	0.9787
Matthews Correlation Coefficient (MCC)	0.9251

ResNet-18





MobileNet-v2 **ResNet-50**

- ResNet-50 achieved slightly less accurate results: 95.23% accuracy, MobileNet-v2: 94,77% accuracy