Advanced CNN for Pneumonia Detection in Chest X-Rays

HamidReza Saadi Dadmarzi
Department of Computer Engineering
Koç University
hdadmarzi23@ku.edu.tr

Ehsan Barkhordar
Department of Computer Engineering
Koç University
ebarkhordar23@ku.edu.tr

June 2024

1 Introduction to Pneumonia

Pneumonia is an infection that inflames the air sacs in one or both lungs. The air sacs may fill with fluid or pus, causing symptoms such as cough with phlegm or pus, fever, chills, and difficulty breathing. Various organisms, including bacteria and viruses, can cause the disease, leading to different types of pneumonia. Bacterial pneumonia is generally more severe, with symptoms like high fever and thick phlegm, often caused by the bacterium *Streptococcus pneumoniae* [1]. Viral pneumonia, on the other hand, tends to present milder symptoms similar to those of a cold, including a dry cough and muscle aches, and may result from viruses like the flu virus. Chest X-rays are crucial for diagnosing pneumonia and distinguishing between its bacterial and viral forms. In bacterial pneumonia, the X-ray shows white patches in the lungs where cells are fighting the bacteria, whereas viral pneumonia X-rays show more general changes, such as a diffuse, cloudy appearance in the lungs. This differentiation helps doctors decide on the best treatment approach [2].

2 Data Preparation

Data Sourcing and Labeling

The dataset used in this study is derived from the NIH Chest X-ray Dataset [3], which consists of images categorized by diagnosis (NORMAL and PNEUMONIA) and intended use (test, train, validation). The PNEUMONIA category is further divided into bacterial and viral pneumonia, identified by the filenames. This project focuses on binary classification: NORMAL vs. PNEUMONIA.

Images are organized in directories labeled according to their diagnosis, simplifying the labeling process during the training phase. Due to the large dimensions of the images (512 x 512 pixels, single channel), they are not loaded entirely into memory at once. Instead, paths to the images are stored in lists, and the actual image loading is managed on-the-fly by `ImageDataGenerator` objects during model training. This strategy efficiently handles memory usage and expedites the data preparation process.

Data Distribution Analysis

The distribution of diagnoses in the dataset is significantly skewed towards pneumonia, with approximately 73% of the samples classified under the PNEUMONIA category. This imbalance suggests the model will encounter more examples of pneumonia during training, which may influence its learning bias. It is crucial to consider the potential need for class weight adjustments to prevent the model from predominantly predicting the majority class. The table below illustrates the distribution of the categories:

Category	**Percentage**
NORMAL	27%
PNEUMONIA	73%

Table 1: Distribution of NORMAL and PNEUMONIA cases in the dataset.

2.1 Sample Images from the Dataset

To provide a clearer understanding of the dataset, here are examples of NORMAL and PNEUMONIA chest X-rays. These images offer a visual reference to the types of cases the model will be trained to recognize.

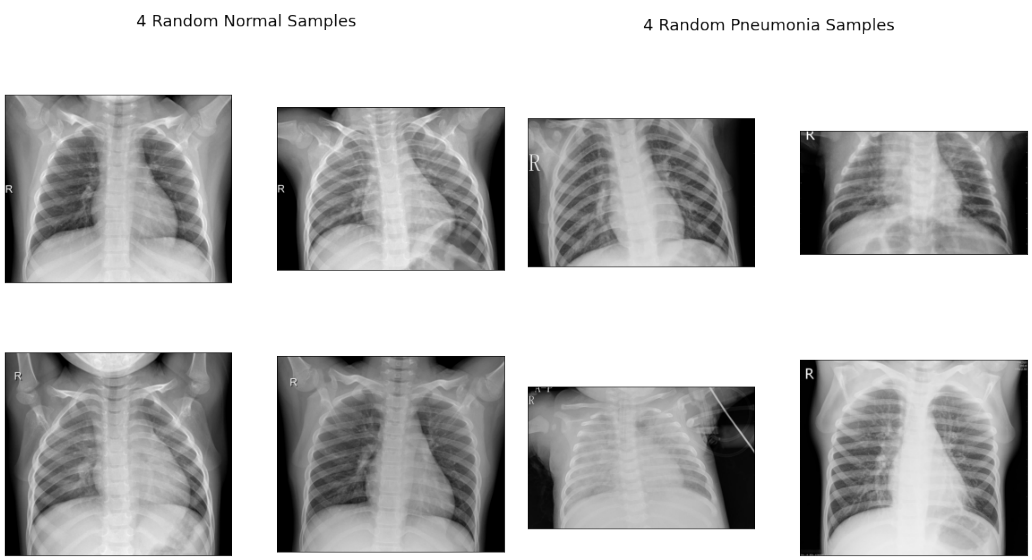

Figure 1: Sample NORMAL chest X-rays.

3 Training Phase

3.1 Model Architecture

This section outlines the Convolutional Neural Network (CNN) architecture developed for efficient processing of 512x512 grayscale images to classify them as normal or pneumonia-infected lung X-rays.

3.1.1 Input Layer

- **Image Input**: The network accepts input images of size 512x512 pixels, each in a single-channel (grayscale). This design is specifically tailored to the characteristics of typical medical X-ray images.

3.1.2 Convolutional Blocks

- **Structure**: The network features seven convolutional blocks. Each block includes:
  - A convolutional layer that uses ReLU activation to introduce non-linearity.
  - A 2x2 max pooling layer to reduce spatial dimensions by half, thus lowering computational demands and enhancing feature detection.
  - A dropout layer with a rate of 0.1 to prevent overfitting, ensuring the model's generalizability.

3.1.3 Global Average Pooling

- **Feature Compression**: Following the final convolutional block, a global average pooling layer aggregates the features from each feature map into a single value per map, dramatically reducing the number of parameters and hence the complexity of the model.

3.1.4 Output Layer

- **Prediction**: The architecture concludes with a dense output layer that uses a sigmoid activation function. This setup provides the probability that an X-ray image indicates pneumonia, facilitating binary classification.

3.1.5 Compilation

- **Optimization and Loss**: The model is compiled with the Adam optimizer, employing binary crossentropy as the loss function. This choice ensures a balanced focus on precision and recall, crucial for medical diagnostic accuracy.

3.1.6 Training Configuration

- **Data Split**: The dataset is divided into 70% training, 15% validation, and 15% test sets using scikit-learn's `train_test_split` method. This separation ensures a comprehensive evaluation of the model across unseen data.

- **Epochs and Batching**: The training process consists of multiple epochs with incremental processing of image batches. This methodical approach allows for the effective optimization of network weights.

3.2 Data Augmentation

To further enhance the model's ability to generalize and reduce overfitting, data augmentation techniques are applied exclusively to the training data. These techniques include rotations, translations, and zoom to simulate a variety of imaging conditions. The `ImageDataGenerator` object in Keras handles these transformations dynamically during model training, thereby enriching the dataset variability without the need for additional data storage. Augmented images are displayed below to illustrate the transformations applied.

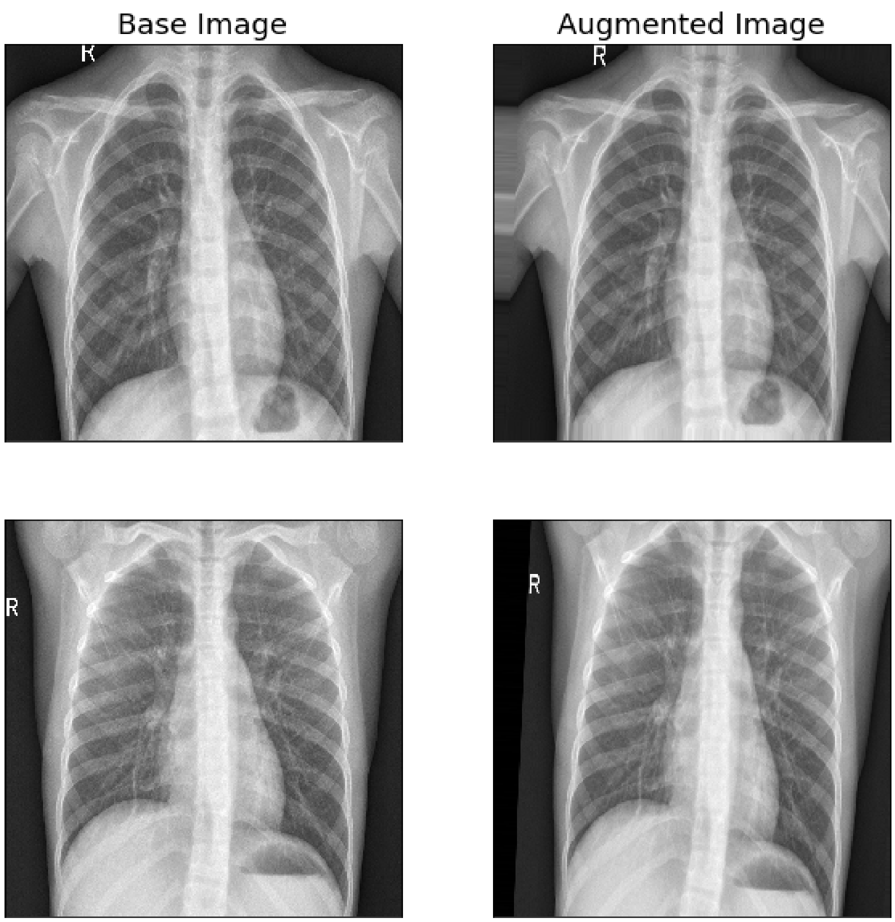

Figure 2: Sample of augmented training images.

Model Training Results

The following figures illustrate the performance of the convolutional neural network under various configurations. These include different dropout rates, the presence or absence of image augmentation, and the impact of learning rate decay. Each subplot provides insights into the model's accuracy and loss metrics over training epochs, aiding in understanding the effect of regularization techniques and training strategies on model generalization.

The confusion matrix below illustrates the performance of the model in distinguishing between Normal and Pneumonia cases based on X-ray images. Each cell represents the count of predictions made by the model compared to the true labels.

Table 2: Updated Confusion

True Label \Prediction	**Normal**	**Pneumonia**
Normal	222	14
Pneumonia	38	558

This matrix helps in assessing the model's ability to correctly classify cases as either Pneumonia or Normal, which is crucial for reliable medical diagnostics.

Appendix: Model Performance Plots

The figures in this appendix present the performance metrics of various convolutional neural network configurations tested during this study. These metrics include accuracy and loss plots across different epochs, illustrating the model's behavior with varying numbers of convolutional layers, the use of image augmentation, and the application of regularization techniques such as dropout. This comprehensive evaluation helps in understanding how each configuration contributes to the model's ability to generalize and perform under different conditions. The findings from these tests guided the selection of the optimal model architecture described in the main sections of this paper.

References

[1] Harsh Sharma et al. "Feature extraction and classification of chest x-ray images using cnn to detect pneumonia". In: *2020 10th international conference on cloud computing, data science & engineering (Confluence)*. IEEE. 2020, pp. 227–231.

[2] Shagun Sharma and Kalpna Guleria. "A deep learning based model for the detection of pneumonia from chest X-ray images using VGG-16 and neural networks". In: *Procedia Computer Science* 218 (2023), pp. 357–366.

[3] Xiaosong Wang et al. "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 2097–2106.

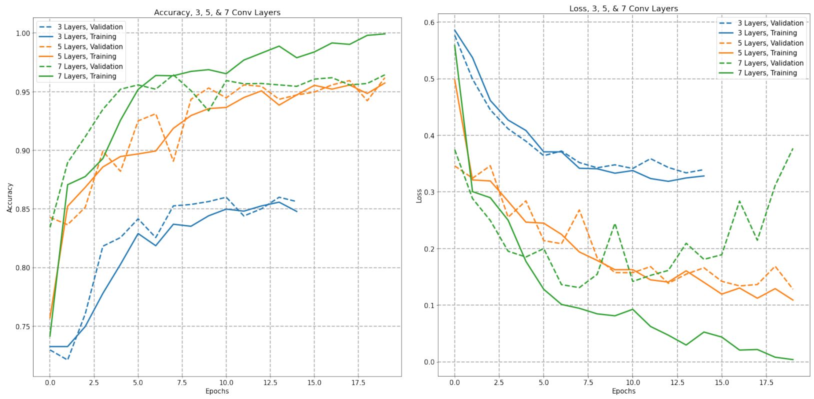

Figure 3: Accuracy 3, 5, & 7 Conv Layers and Loss, 3, 5, & 7 Conv Layers

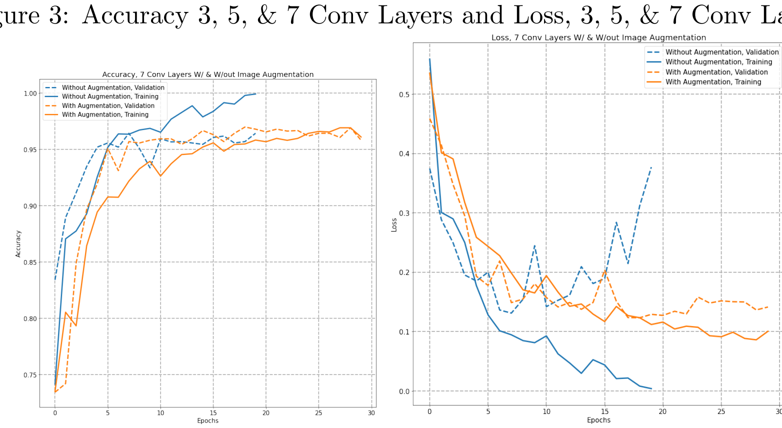

Figure 4: Accuracy and Loss 7 Conv Layers W & Without Image Augmentation

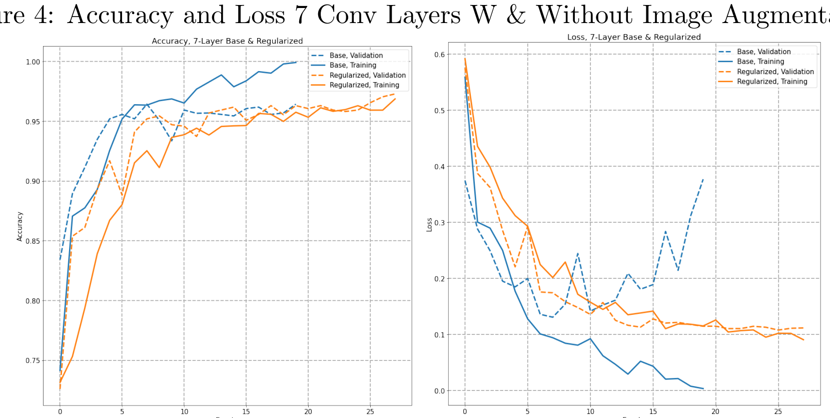

Figure 5: Accuracy and Loss 7-Layer Base & Regularized

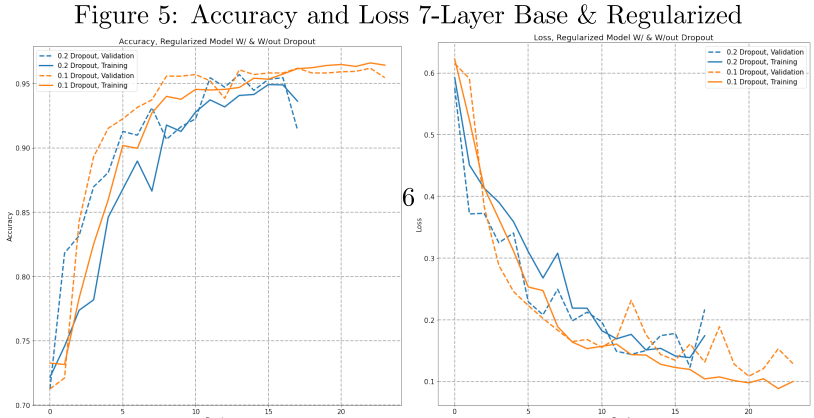

Figure 6: Accuracy and Loss Regularized Model W & Without Dropout