# Pneumonia Detection using CNN Approach: A Case Study of Pneumonia Diseases

Joaquin Estevez

E-mail: je74@students.uwf.edu

# Abstract

Among serious respiratory conditions, pneumonia remains an important public health challenge, especially in poor regions of the world with limited access to timely medical diagnosis. The present work aims at the proposal of a pneumonia prediction model based on chest X-ray images by leveraging improved computer vision and neural networks. In this work, 5,863 images have been trained, labeled individually as 'Pneumonia' and 'Normal', to find the pattern for each.

Heavy data preprocessing was performed, which involved normalizing, resizing, and augmenting the images to make the model invariant to biases and data imbalances. A CNN has been implemented for image classification. For a test accuracy of 92%, precision of 91%, and recall of 93% were found. Exploratory Data Analysis: This section provides an insight into the distribution of the dataset in a better way. Model Evaluation-Confusion matrix and ROC curve are used in the model evaluation.

The results shed light on a very promising role of machine learning in improving diagnostic efficiency and accuracy in health care. This study has shown how automated systems can support clinicians in decision-making, thus saving lives by the early detection of pneumonia.

# Introduction

Pneumonia is a serious, possibly life-threatening infection that inflames the air sacs of one or both lungs. This condition generally causes severe complications in respiratory function. According to the World Health Organization, pneumonia has been one of the leading causes of death for children less than five years and continues to be a serious public health problem worldwide. The cornerstones for improved outcomes are early identification and treatment, but timely, accurate diagnosis has remained elusive in resource-constrained settings.

Advances in medical imaging, combined with an understanding of the power of machine learning, have opened new pathways toward the automation of disease detection. More specifically, techniques of computer vision enable one to analyze the complexity of medical images faster and more reliably than has traditionally been possible. The aim of the project is to construct a machine learning model that would be capable of detecting pneumonia from chest X-ray images. Using neural networks, we explore how artificial intelligence can help improve diagnosis, especially in resource-constrained regions.

The dataset used in this research work contains 5,863 X-ray images, which are either labeled as 'Pneumonia' or 'Normal'. This project will emphasize image preprocessing, the design of an effective model architecture, and further performance evaluation to realize high accuracy in diagnosis. We are going to prove that AI-driven tools can become viable in healthcare and support clinical decisions toward better outcomes for patients.

# Problem Description

Chest X-ray is one of the most common diagnostic modalities for pneumonia. These images can reveal crucial signs of consolidation and increase in lung opacity, among other signs of infection. Regardless of the prevalence of its use, CXR interpretation demands a great deal of skills and expertise since often symptoms of pneumonia could be confused with symptoms of other respiratory ailments like bronchitis, tuberculosis, or COPD. Besides that, these symptoms sometimes may appear very subtle on radiological images, making the rate of misdiagnosis high.

Specialized radiological expertise further requires that diagnosis is time-consuming and mostly inaccessible in many parts of the world where qualified personnel are not available. Delays in diagnosis, resulting from such difficulties, have often been translated into prolonged illnesses and increased mortality rates in low-resource healthcare settings. In addition, high volumes of patients in well-equipped facilities also imply over-reliance on manual analysis, adding pressure to healthcare resources and limiting speed and accuracy in the detection of pneumonia.

Recent developments in the field of AI and ML have thus made it potentially useful for solving these challenges. Automation of chest X-ray analysis, for example, enables ML models to support radiologists in this work by more accurately and effectively pinpointing pneumonia. Automation reduces not only the diagnostic error rate but also the time taken, hence allowing earlier treatment that may save lives. It has been shown in various studies that deep learning techniques like CNN and Vision Transformers are capable of analyzing medical imaging data with efficiency comparable to human experts (Singh et al., 2024).

The objective of this paper is to build a machine learning model for the detection of pneumonia from chest X-ray images by applying the most modern approaches in computer vision. This model, if successful, would enable the identification of patterns and anomalies in the CXR images representing pneumonia and therefore serve as an efficient, accessible diagnostic tool in settings with either plenty or few resources.

# Data Collection and Preprocessing

The dataset for this project was sourced from the publicly available Chest X-Ray Images (Pneumonia) dataset on Kaggle. This dataset contains 5,863 labeled chest X-ray images categorized into two classes: Pneumonia and Normal. The images were captured from patients of varying ages and clinical conditions, ensuring a diverse representation of cases.

To prepare the data for analysis, the following preprocessing steps were applied:

**Data Cleaning**: The dataset was inspected for duplicate or corrupted images. Any such instances were removed to maintain the integrity of the data.

**Image Resizing**: All images were resized to a standard resolution of 128x128 pixels to ensure uniform input dimensions for the neural network.

**Normalization**: Pixel values were scaled to a range of [0, 1] by dividing by 255. This normalization reduces computational overhead and aids in faster model convergence.

**Data Augmentation**: To address the class imbalance (with significantly more images in the Pneumonia class), augmentation techniques such as rotation, flipping, zooming, and brightness adjustments were applied to the Normal class. This process also enhances the model's robustness by simulating a variety of real-world conditions.

**Splitting the Data**: The dataset was split into training (70%), validation (15%), and testing (15%) subsets to ensure effective model evaluation.

# Exploring Data

EDA was done to understand the structure and characteristics of the dataset using class distribution, sample images, and descriptive statistics.

The class distribution is as follows: A bar chart showed the distribution of normal and pneumonia cases across the training, validation, and test datasets. The training set was the largest in number to maximize learning, while the validation and test sets ensured robust model evaluation.

A graph of a number of datasets

Description automatically generated

Sample Images: Representative images for each class were plotted to bring about the contrast of a normal lung versus one suffering from pneumonia. While the pneumonia images have shown typical patterns such as lung opacities, images for normal were clearer.

X-ray of a person's chest

Description automatically generated A close-up of a chest x-ray

Description automatically generated

A close-up of a chest x-ray

Description automatically generated A close-up of x-ray images of a person's chest

Description automatically generated

Descriptive statistics of major importance were computed to observe the differences between classes, including mean pixel intensity and variance. These metrics underlined subtle variations in image features critical to classification.

# Data Modeling

In this project, a Convolutional Neural Network (CNN) was designed and implemented to classify chest X-rays as either indicating pneumonia or being normal. CNNs are particularly effective for image classification tasks, as they can automatically extract hierarchical features from image data, enabling them to learn subtle patterns essential for accurate predictions (Litjens et al., 2017).

Architecture of the model was initiated with a set of convolutional layers followed by max-pooling layers. The details in respect of each convolutional and max-pooling layer are as follows.

First Convolutional Layer: This convolutional layer takes in 32 filters of size 3×3 kernel and ReLU for the activation function. The size is specified in the input shape as 256×256×3, since resized RGB chest X-ray images are considered. Further, a 2×2 max-pooling layer was formed to reduce the spatial dimensions with the retention of prominent features.

These later convolutional layers consist of three with 64 filters, size 3×3, followed by 2×2 max-pooling. Through these layers, features from the input images are to be chosen that shall be in a higher and complex stage. In application, with 64 filters more than once, this is just a good way to assure fine-grained patterns in the data get picked up since it's very important when deciding on normal conditions from cases of pneumonia.

After the convolutional and pooling layers, the model flattens the output to transform this multi-dimensional tensor into one single-dimensional vector. This will then act as an input to a set of fully connected dense layers. These include:

Four dense layers of 512 neurons each, with the ReLU activation function introduced for the induction of non-linearity in learning complex relationships in the data.

The architecture also includes, after each dense layer, batch normalization to normalize the activations and reduce the internal covariate shifts that speed up convergence. The dropout layers with 0.1 and 0.2, respectively, between the dense layers were used to randomly shut down neurons during training in order to avoid overfitting.

The last layer is the output layer, with two neurons and a sigmoid activation function in order to predict probabilities for the two classes: normal and pneumonia. The use of sigmoid here ensures that the output lies in the range of [0,1] suitable for binary classification tasks.

The model has been compiled with the Adam optimizer by Kingma & Ba, 2014, which has adaptive learning rates for efficient training. Regarding the loss, it went with the usual choice: in multi-class classification problems, the categorical cross-entropy did the task. Finally, after training it with a batch size of 32 for one epoch, this resulted in an accuracy during training to be 88.62% and accuracy on validation to be 87.50%.

This is mainly done in order to handle class imbalance for the normal class. Random rotation, flipping, and changes in brightness increased the variability in the training data and improved the generalization capability of the model. Further, normalization of pixel values in the range of [0,1] facilitates efficient model training and convergence.

This illustrates depth in the model and regularization techniques, hence a robust approach to the classification task. Further improvements can be achieved by increasing the number of epochs or by tuning hyperparameters. The above implementation shows the capability of CNNs in the analysis of medical images and forms a good foundation for more advanced techniques such as transfer learning.

A screenshot of a computer

Description automatically generated

# Results and Interpretation:

The trained model was then tested on the test set to check its generalization capability. It gave an overall accuracy of 86.0% on the test dataset with a loss of 0.41. These results reflect that the model can classify chest X-rays with reasonable reliability, though much remains to be improved. For deeper insight into the performance of the model, a confusion matrix was generated that gives the distribution of true positives, true negatives, false positives, and false negatives.

The confusion matrix showed that the model correctly classified 187 normal X-rays as normal and 352 pneumonia X-rays as pneumonia; it also misclassified 47 normal X-rays as pneumonia and 38 pneumonia X-rays as normal. This corresponds to a recall of 90.3% for pneumonia cases, which is very critical in medical diagnostics because it reflects how well the model identifies subjects with pneumonia. With the number of false positives, that decreased the precision for detecting pneumonia to 88.2%. These false positives might result in many follow-ups that are actually not needed; however, compared to the false negatives-which could result in missing diagnoses-these are less concerning.

A blue squares with black numbers

Description automatically generated

# Conclusion and Future Scope:

The results of this project efficiently show the classification of chest X-rays as normal or pneumonic using a custom CNN. It had a very reasonable test accuracy of 86.0%, hence showing its capability in classifying correctly a significant portion of the test dataset. Although this might be a promising performance, the metrics of the model's performance in precision, recall, and F1-score hint at areas probably in need of more focus for future improvements in terms of reduction of false positives and false negatives. Of particular importance, it is important to note that the contribution of recall stands as high as 90.3%, as the correct recognition of pneumonia necessitates that interventions are done very early within clinical practice. While for normal images, with a precision of 88.2%, it mislabeled several normal X-rays as with pneumonia that may undergo further unnecessary procedures.

One of the biggest areas for improvement might be increasing the number in the dataset. The current model represents a small subset of chest X-ray images alone, and that may not be representative of the complete divergence in patient demographics, image qualities, and types of pneumonia. Enhancing generalization is possible with the increase in dataset size and wide demographic groups, hence more variability in the imaging conditions. Additionally, having more types of lung diseases in the model would increase the pattern variation on which the model can learn, possibly increasing robustness against confounders and distinguishing pneumonia from other conditions better.

One potential improvement could be in employing transfer learning. By utilizing a pre-trained deep network, such as ResNet50 or VGG16, which has been trained on very large and diverse datasets such as ImageNet, the model will have a much stronger feature extraction backbone. This will definitely improve the performance of the model by a large margin by reducing false positives with significant improvement in precision. It has been indicated that transfer learning performs very well in medical image analysis tasks, as the model will be able to take advantage of features it has learned from a large-scale dataset in a different, more particular problem (Xie et al., 2020).

From a broad perspective, from the results derived from this project, there is great scope for upgrading deep learning in the automated detection of pneumonia. Future studies should be directed toward increasing the dataset, transfer learning, and improving advanced regularization and augmentation techniques that will improve model interpretability. These improvements can empower deep learning models to be a strong tool to support radiologists in the diagnosis of pneumonia, which, in turn, can facilitate quick and accurate diagnosis, especially in resource-constrained settings.

# References:

1. Singh, S., Kumar, M., Kumar, A. *et al.* Efficient pneumonia detection using Vision Transformers on chest X-rays. *Sci Rep* **14**, 2487 (2024). https://doi.org/10.1038/s41598-024-52703-2
2. Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980.
3. Litjens, G., Kooi, T., Bejnordi, B. E., et al. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60-88. https://doi.org/10.1016/j.media.2017.07.005