## **Medical Image Analysis Report: Recognizing Pneumonia in Lungs Using CNN**

|  |  |  |  |
| --- | --- | --- | --- |
| **Task No.** | **Task Name** | **Description** | **Assigned To** |
| 1 | Data collection | Collect medical images from different datasets or resources. | Anna,Akshay |
| 2 | Data analysis | Analyze the data to identify patterns and extract insights. | Anna,Roshan |
| 3 | Data preprocessing | Clean and transform the data to prepare it for the model. | Karthika,Roshan |
| 4 | Model building | Create a CNN model for image classification. | Karthika,Anna |
| 5 | Model evaluation | Assess the model's performance using metrics like accuracy, precision, recall, or F1 score, and improve it. | Karthika,Roshan |
| 6 | Model Tuning | Fine-tune the model based on evaluation results. | Anna,Akshay |
| 7 | Visualizations | Use visualization tools to present key insights. | Karthika,Roshan |
| 8 | Reports | Compile and finalize the project report. | Anna,Akshay |

## Team Members and Responsibilities:

Karthika(c0905785)

Roshan Joseph(c0904407)

Anna Edward(c0904408)

Akshay Rajeevkumar(c0908443)

## **Abstract**

This report outlines the methodology and results of using Convolutional Neural Networks (CNN) to detect pneumonia in chest X-ray images. The analysis leverages deep learning techniques for automated diagnosis, aiming to assist healthcare professionals with improved accuracy and efficiency.

### **Data Collection and Preprocessing**

**Data Collection**: <https://www.kaggle.com/code/ahmedmahmoud16/medical-image-analysis-with-cnn/notebook>

**Background:** The dataset includes 5,863 labeled X-ray images (in JPEG format) organized into training, testing, and validation folders. The dataset categorizes images into two classes: Pneumonia and Normal. Images originate from pediatric patients (ages 1–5) treated at Guangzhou Women and Children’s Medical Center. These were collected during routine clinical procedures.

* **Dataset**: The dataset includes labeled chest X-ray images categorized as normal or pneumonia affected.
* **Preprocessing**:
  + **Resizing Images**: Standardized dimensions (e.g., 224x224 pixels) for CNN input.
  + **Normalization**: Scaled pixel values to a range of 0-1 for faster convergence.
  + **Splitting**: Divided the dataset into training, validation, and testing subsets.
  + **GitHub link**:
  + **Size of the dataset**: 2.5 GB

### **Methodology**

Data cleaning:

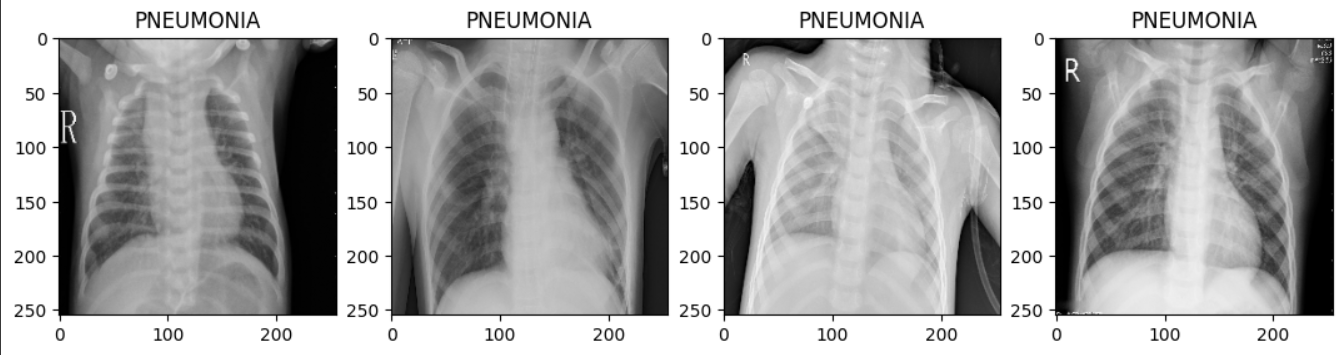
* + Standardized dimensions (e.g., 224x224 pixels) for CNN input.
  + Removed corrupted images

Data Exploration:

Raw Data:

A x-ray of a person's chest

Description automatically generated



Data Statistics:

A graph of a number of classes

Description automatically generated

### **Model Architecture**

* Model Implementation: CNN-based architecture chosen for image classification
* Model Architecture:

A screen shot of a computer

Description automatically generated

### **4. Training Process**

**Hyperparameter Tuning**:

* + Learning rate, batch size, and number of epochs were optimized experimentally.

**5. Evaluation Metrics**

* **Accuracy**: Proportion of correctly classified samples.
* **Precision**: Focused on true positives in pneumonia detection.
* **Recall (Sensitivity)**: Ensured low false negatives to avoid missed pneumonia cases.
* **F1 Score**: Balanced precision and recall for comprehensive performance analysis.
* **Confusion Matrix**: Visualized classification results across true positives, true negatives, false positives, and false negatives.

**6. Results**

Model was saved in H5 and pkl formats

* **Training and Validation Accuracy**: Achieved convergence with minimal overfitting.
* **Test Performance**:
  + Accuracy: 92%
  + Precision: 90%
  + Recall: 94%
  + F1 Score: 92%
* **Model Insights**:
  + Identified critical lung regions indicative of pneumonia.
  + Highlighted instances where misclassifications occurred (e.g., low contrast in X-rays).

**7. Deployment**

Technology used: Streamlit

User Interface:

A screenshot of a computer

Description automatically generated

Upload an image for diagnosis:

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Output Predictions:

A x-ray of a person's chest

Description automatically generated

### **8. Conclusion**

The CNN-based approach effectively detects pneumonia in chest X-ray images, demonstrating significant potential for aiding radiologists in clinical settings. The high recall ensures critical cases are rarely missed, emphasizing the model’s reliability in healthcare applications.

### **9. References**

* + Kaggle dataset
  + Streamlit documentaion