

RADIOLOGY GUIDE: IMAGE CLASSIFIER FOR PNEUMONIA DETECTION DATA REPORT.

Business Understanding:

- Pneumonia is an infection that inflames the air sacs in one or both lungs. The air sacs may fill with fluid or pus (purulent material), causing cough with phlegm or pus, fever, chills, and difficulty breathing.
- Patients presenting with difficulty in breathing or presenting other respiratory symptoms for pneumonia in the emergency department are usually given a chest Xray. They have the advantage of lower radiation exposure, faster feasibility and better equipment portability compared to other imaging modalities such as computed tomography (CT). This diagnostic examination can provide supplemental and timely information regarding a patient's cardiopulmonary condition and probable changes from any infectious process. Studies have shown that with faster reporting of pneumonia in Chest radiographs, the median length of hospital stays is significantly shorter, the likelihood of receiving appropriate therapy is higher, and the probability of infectious spread is lower.
- However, the interpretation of CR examinations is variable and examiner-dependent. To increase the sensitivity and specificity of imaging patterns for pneumonia in Chest x-rays, deep learning (DL) algorithms must become more prevalent. Prior studies have shown that the use of artificial intelligence (AI) significantly improves the detection of pneumonia in Chest radiographs.
- Given the large number of examinations, reporting using AI can highlight Chest x-rays with abnormalities, helping to prioritize reporting by radiologists. Further, where Chest radiographs are initially evaluated by clinicians outside regular operations, AI can be of assistance. In this situation, a well-functioning evaluation of Chest x-rays by AI can significantly support clinicians' decision making.
- The target is to use algorithms to classify medical images for assistance in diagnosis, treatment planning, and disease monitoring. Our project aims to create an image classifier for pneumonia detection using machine learning techniques. Pneumonia is a

common and sometimes fatal respiratory illness, and early identification is critical for optimal treatment and patient outcomes. Our key objective is to build a strong classifier capable of correctly recognizing pneumonia in chest X-ray pictures using convolutional neural networks (CNN's) and sophisticated image processing methods.

Research Questions:

1. Which deep learning model architecture achieves the best performance in terms of accuracy, sensitivity, specificity in detecting pneumonia from chest X-ray images?
2. How can data augmentation techniques be employed to improve the generalizability and robustness of the deep learning model for pneumonia detection on unseen data?
3. Can transfer learning from pretrained models such as ResNet50v2 improve pneumonia detection performance?
4. How do pneumonia detection models perform compared to human radiologists, and what are the implications for clinical practice?

Problem Statement:

- Pneumonia is a serious respiratory illness that affects many people, especially kids and older folks. Getting a quick and accurate diagnosis is super important for making sure people get the right treatment and have better chances of getting better. But right now, diagnosing pneumonia using chest X-rays can be tough. It takes time for radiologists to look at the X-rays, and sometimes they might miss things or make mistakes. Plus, not every hospital has experts available all the time.
- A recent article from the National Library of Medicine even says that radiologists only get pneumonia diagnoses right about 60% of the time! [Link here](#) That's not great, especially when people's lives are on the line.
- To help tackle this problem and save lives, we're looking into using deep learning, which is like teaching computers to think and learn like humans do, to build a system that can automatically spot pneumonia in chest X-rays. Deep learning is cool because it can pick up on really complicated patterns from lots of X-ray pictures, and it's already changing the way doctors analyse medical images.
- Our goal is to create a tool that can help hospitals like KNH diagnose pneumonia faster and more accurately, especially during tough times like strikes when there

might be fewer experts around. We want to make a real difference in patient care and outcomes at Kenyatta National Hospital and beyond.

Objectives:

- **Main Objective:**
 - To use algorithm to develop an accurate and efficient image classifier for pneumonia detection using chest X-ray images for assistance in diagnosis.
- **Specific Objectives:**
 - To alleviate the strain on the hospital's diagnostic services.
 - To support clinicians' decision making.
 - Gather a comprehensive dataset of chest X-ray images containing both pneumonia-positive and pneumonia-negative cases.
 - Experiment with various architectures, hyperparameters, and optimization techniques to maximize the classifier's accuracy and efficiency.

Data Understanding:

- Data Source:
 - The dataset used for our project were sourced from <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>
 - It contains Chest X-ray images (anterior-posterior) which were selected from retrospective cohorts of paediatric patients of one to five years old from Guangzhou Women and Children's Medical Centre, Guangzhou.
 - All chest X-ray imaging was performed as part of patients' routine clinical care.
 - The dataset is organized into 3 folders (train, test, validation) and contains subfolders for each image category (Pneumonia/Normal).
 - There are 5,856 X-Ray images (JPEG) and 2 categories (Pneumonia/Normal).
 - The nearly 6000 images are classified into two categories: Normal or Pneumonia.
 - As provided by Kaggle, the images are divided into three subsets:
 - train - 5,216 images
 - test – 16 images
 - validation - 624 images

Data Preparation:

- Loading the data

- Cleaning the data

Feature Engineering:

External Data Source Validation:

Exploratory Data Analysis (EDA):

Modelling:

Conclusions:

Recommendation:

Future Improvement Ideas:

Deployment: