Pneumonia Detection on Chest X-Ray with Deep Learning

EXECUTIVE SUMMARY

The Problem

The goal of this project is to detect the presence of Pneumonia on Chest-X-rays using Deep Learning Techniques and to evaluate how well the model performs in achieving the objective. According to the CDC, Pneumonia was responsible for the deaths of 47,601 people in the U.S in 2020 [1]. Globally, it killed more than 740,000 children younger than 5 years old in 2019. November 12 is observed as World Pneumonia Day [1] to raise global awareness of the disease. Statistics Canada identified Pneumonia as the cause of death for 5,931 people and ranked it 8 as a leading cause of death [2]. This project aims to reduce mortality caused by Pneumonia by applying Deep Learning to detect the presence of Pneumonia on X-ray images. It will facilitate the rapid referral of patients needing urgent medical intervention.

Background

Pneumonia is an infection that inflames the lungs. The air sacs in the lungs may become filled with pus or fluid causing cough, fever and breathing difficulties. Viruses, bacteria, and fungi can cause Pneumonia. It can range in severity from mild to life-threatening and is most serious for adults 65 years and older, children younger than 5 years old, people who have ongoing medical conditions and people who smoke.

Pneumonia can be detected through Blood tests, Chest X-ray, Pulse oximetry and Sputum test. Chest X-rays are currently the best available method for diagnosing pneumonia. Still, detecting pneumonia in chest X-rays is a challenging task that relies on the availability of expert radiologists, even though it can still be difficult for radiologists because its appearance in X-ray images can overlap with other diagnoses [3]. This presents a challenge that interests the world of Data science and Machine Learning. The problem is a classification one, and Machine Learning can be deployed to solve it. Chest X-ray imaging technology was widely used for the early detection of COVID-19 pneumonia and deep learning methods have recently shown remarkable results in detecting COVID-19 on chest X-rays [4]. Also, research has shown that in the case of COCID-19 detection, the combination of deep learning and chest X-rays could be faster and less expensive than polymerase chain reaction (PCR), the current gold standard for COVID-19 detection [4].

Details of the dataset

The image dataset was collected from Kaggle. However, the original data itself was gotten from Guangzhou Women and Children's Medical Center, Guangzhou, China. The images were selected from cohorts of pediatric patients of one to five years old. The images being collected as standard of care in the Hospital was used to train an AI system [5]. Before use, it was screened by chest radiologists, graded by two expert physicians, and checked by a third expert [5].

Summary of cleaning and preprocessing

A total of 5856 images were collected and split into train (4842 images), test (604 images) and validation (420) set. The following preprocessing techniques were performed on the images: Data Augmentation, Grayscale conversion, Normalization, and Image standardization

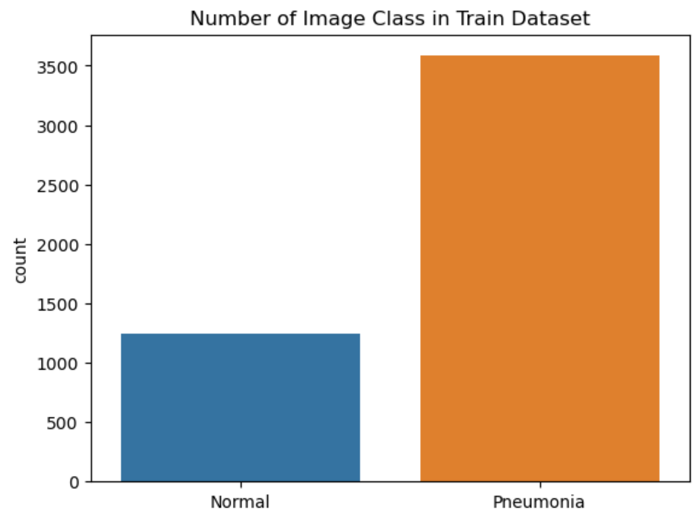

Fig 1: Chart showing number of images in the train set. The dataset is unbalanced, and this can cause bias in the model. This was corrected by applying image augmentation, a technique used to increase the number of images in the train set.

Due to the imbalance in the train dataset, data augmentation was used to increase the number of images. Augmentation also applies distortions to the images. This helps to make the model more robust. The augmentation techniques applied include rotation, shifting, flip.

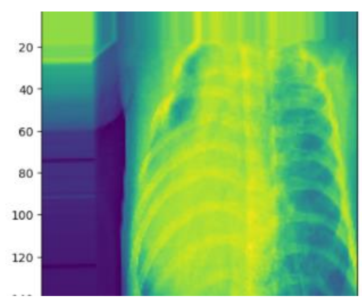

Fig 2: Augmentation has not been applied to the image on the left. The image on the right has augmentation(shifting) applied. Shifting shifts image pixels horizontally or vertically.

Grayscale conversion converts the images to black and white; it reduces computation complexity in the model.

Normalization was applied to convert the pixel values of the images to a predefined range 0,1.

Image standardization is a method that scales and preprocesses images to have similar heights and widths. It-rescales data to have a standard deviation of 1 and a mean of 0.

Insights, Modelling and Results

Four different Convolutional Neural Network models were built. The model that gave the best evaluation on the test data was chosen. The models had similar architecture but had different parameters set and trained to obtain the best result.

Models	Test loss	Test accuracy
Model 1	0.826	0.821
Model 2	0.414	0.849
Model 3	0.830	0.685
Model 4	0.540	0.857

Table 1: Comparison of the key indices for model selection.

In Machine Learning, loss is the penalty for bad prediction, it indicates how bad the model predicted a given example. The lower the loss, the better the model. A perfect model would have a loss of zero. Model 2 was the best model trained because it returned the least loss on the test data.

	precision	recall	f1-score	support
Normal (Class 0)	0.81	0.75	0.78	214
Pneumonia (Class 1)	0.87	0.91	0.89	390
accuracy			0.85	604
macro avg	0.84	0.83	0.83	604
weighted avg	0.85	0.85	0.85	604

The classification report evaluates the performance of the model in predicting unseen data. Precision is the accuracy of positive predictions, how many predictions the model got right out of all the predictions made. The model predicted a normal X-ray image 81% of the time and a Pneumonia X-ray image 87% of the time.

Recall has to do with the truth. It tells us the percentage of positive cases that were correctly identified. The model correctly identified a normal X-ray 75% of the time and a Pneumonia image 91% of the time.

The F1 is a weighted average of precision and recall, with the best score being 1.0 and the worst 0.0. The model has a F1 score of 0.83 and an accuracy of 0.85.

The support is the number of actual occurrences of the class in the dataset. The test dataset had 214 occurrences of Normal images and 390 Pneumonia images.

Findings and Conclusions

The model has on overall accuracy of 85%. It performed better at predicting Pneumonia X-ray images than Normal X-ray images. While it is good that the model predicts Pneumonia X-ray images with high accuracy it is also important that it gets better at also identifying Normal X-ray images. This would prevent misdiagnosing of Pneumonia.

The model's performance can be attributed to the number of images in the train dataset: there were more Pneumonia images than Normal images in the dataset, so the model performed better at identifying images positive for Pneumonia.

Potential next steps will be to improve the model's performance by increasing the number of Normal images in the train data set, in addition to improving the model's overall accuracy.
The model's performance will be compared with a pre-trained model.

Convolutional Neural Network models are usually referred to as a black box. However, Gradcam can be used to understand how a model makes predictions. Gradcam would be used to understand how the model made predictions.

References

[1] CDC, "Pneumonia Can Be Prevented—Vaccines Can Help," 30 September 2022. [Online]. Available: https://www.cdc.gov/pneumonia/prevention.html. [Accessed 20 November 2022].

[2] Statistics Canada, "Leading causes of death, total population, by age group," 20 11 2022. [Online]. Available: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310039401. [Accessed 20 11 2022].

[3] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, R. L. Ball, C. Langlotz, K. Shpanskaya, M. P. Lungren and A. Y. Ng, "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," arXiv:1711.05225v3, 2017.

[4] M. Elgendi, M. U. Nasir, Q. Tang, D. Smith, J.-P. Grenier, C. Batte, B. Spieler, W. D. Leslie, C. Menon, R. R. Fletcher, N. Howard, R. Ward, W. Parker and Ni, "The Effectiveness of Image Augmentation in Deep Learning Networks for Detecting COVID-19: A Geometric Transformation Perspective," *Frontiers in Medicine,* pp. 1-12, 2021.

[5] D. Kermany, M. Goldbaum, W. Cai, M. A. Lewis, H. Xia and K. Zhang, "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning," *Cell,* vol. 172, no. 5, pp. 1122-1131, 2018.