Pneumonia Prediction in Chest X-Rays

Ali Malenchik

Bellevue University

DSC680: Applied Data Science

Prof. Catherine Williams

November 20, 2021

Abstract

Pneumonia is a respiratory infection that affects 7% of the global population and results in approximately 700,000 child deaths annually (Hashmi). Early diagnosis of pneumonia is critical in preventing a patient's condition from worsening. Though radiologists typically use chest X-rays to diagnose the infection, this task is challenging, subjective, and potentially prone to human error.

In this project, a Convolutional Neural Network was created using pediatric patient X-ray images in order to predict Pneumonia diagnosis. The CNN model performed fairly well at predicting Pneumonia in X-ray images, at 91% accuracy.

Introduction

Background

Pneumonia is a respiratory infection that affects many individuals each year, even more so now due to Covid-19. Early diagnosis is vital to ensure prompt treatment and prevent the infection from becoming more severe and even resulting in death. Though chest X-rays are already used as part of diagnosis, it is a manual task that may be susceptible to subjectivity and experience of the doctor or radiologist. Additionally, when hospitals are overcrowded, resources may be limited or distracted. Using machine learning offers an opportunity for efficient, reliable diagnosis that could at the very least assist doctors in their decision-making process and therefore save lives.

Problem Statement

Can a machine learning model accurately detect Pneumonia in pediatric patient chest X-rays?

Scope/Assumptions

- As this project utilizes data pertaining to pediatric patients, the results may not be applicable to adult X-rays.
- The X-ray diagnosis labeled in the dataset was reviewed for accuracy by multiple experts and is considered to be reliable.

Methods

Data Source

The dataset used for this project is a collection of chest X-ray images found on Kaggle. The X-rays were gathered from pediatric patients aged one to five at the Guangzhou Women and

Children's Medical Center. The images are provided in JPEG format and can be classified as three different types: a normal chestX-ray, a chest x-ray demonstrating Bacterial Pneumonia, and a chest x-ray demonstrating Viral Pneumonia. The dataset does not include images that were low quality or unreadable, and the classifications were given and reviewed by multiple medical experts to ensure accuracy.

Data Import & Cleansing

The image data was imported from the individual directories (titled NORMAL and PNEUMONIA) using cv2's imread() function.

Exploratory Analysis

For exploratory analysis, three sample images from each category were chosen at random and printed using imshow(). From the samples we can observe that there is a discernable difference between an X-ray of a healthy chest compared to an X-ray of a chest with Pneumonia.

Figure 1

Normal X-ray Sample 1

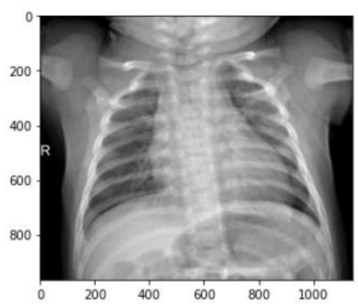
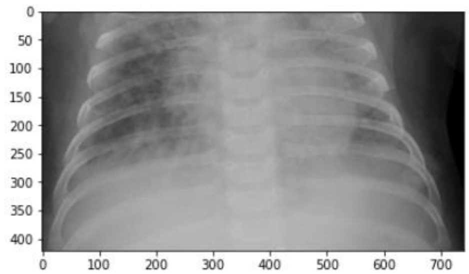


Figure 2

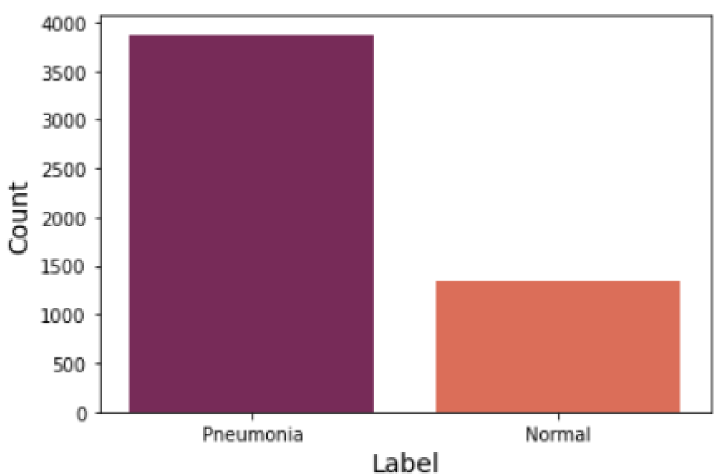
Pneumonia X-ray Sample 1



After data was organized, a count plot was created of the distribution of X-ray type. From the count plot we can see that the dataset is imbalanced, favoring X-rays of Pneumonia infection patients.

Figure 3

Count Plot of Chest X-ray Label



Feature Selection

For each of the datasets (train, test, and validation), labels were appended, separated from features (e.g. X_train, y_train), and converted to arrays to prepare for modeling. Features were normalized using reshape(). Finally, data was augmented using ImageDataGenerator() to apply random transformations on the training data to mimic variations and generate synthetic image data.

Model Deployment

The Convolutional Neural Network was built using Keras Sequential() to build the model layer by layer. Four Conv2D layers were added using the Rectified Linear activation function, followed by a Flatten layer, a Dense layer again using Rectified Linear activation, and finally a Dense layer using Sigmod activation. The model was compiled using RMSprop optimization and Binary Cross-Entropy for the loss function. Accuracy was chosen as the metric for accuracy scoring on the validation set.

Results

Multiple methods were employed to evaluate the model performance. First, the model loss and accuracy was calculated using the reserved test data set using the evaluate() function. The model accuracy was calculated at 90.71 and the loss at 46.25. The training and validation accuracy and loss were also plotted for each epoch, demonstrating that the training accuracy increased approximately 12 points over the course of the 10 epochs. The training loss decreased dramatically from approximately 25 to 2. The minimum loss occurred at both epoch 4 and 9.

Figure 4

Plot of Training and Validation Accuracy

Figure 5

Plot of Training and Validation Loss

Next, predictions were made on the test data and a classification report was created. From the classification report, we can observe the accuracy of the model is .91, the precision of the Normal class is .96, and the recall of the Pneumonia class is .98. The lowest metric in the classification report is the recall of the Normal class at .79.

Figure 6

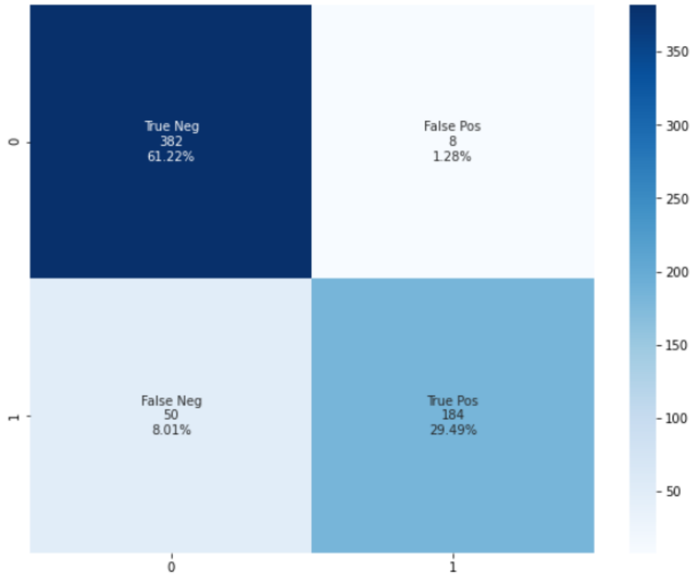
Classification Report for the Convolutional Neural Network Model

	precision	recall	f1-score	support
PNEUMONIA	0.88	0.98	0.93	390
NORMAL	0.96	0.79	0.86	234
accuracy			0.91	624
macro avg	0.92	0.88	0.90	624
weighted avg	0.91	0.91	0.90	624

A confusion matrix was also created and heatmap plotted to demonstrate the counts and percentages of each group. From the heatmap we can see that the true negatives make up 61.22% of the predictions, and the true positives make up 29.49% of the predictions. Only 8 false positives were found, whereas 50 false negatives were present in the predictions.

Figure 7

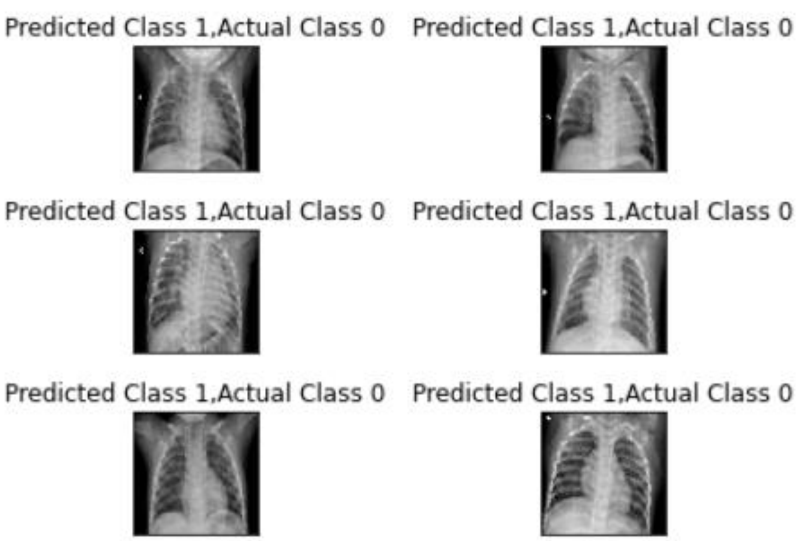
Heatmap of Confusion Matrix for the Convolutional Neural Network Model



Finally, six of the random sample X-rays that were incorrectly predicted were printed.

Figure 8

Sample Incorrect Predictions



Discussion

Conclusion

Based on the results of the model evaluation, the CNN model performed well at predicting the Pneumonia diagnosis using X-ray images. In order to determine if the model is accurate enough to replace clinical diagnosis by radiologists, research would need to be performed regarding the human error rate in order to compare.

Limitations/Challenges

One of the challenges in this project was understanding what exploratory data analysis could be performed on image data. Additionally, I struggled understanding what parameter values to choose in adding the CNN model layers, how many layers to add, etc. I referred to sample projects to understand the benefits of different functions and the impact of each layer.

Next Steps

To further improve the model performance, tuning should be performed on parameters such as activation function, loss function, optimization function, etc. Implementing model ensembles could also improve the performance of the model built in this project. Additional examples of healthy chest X-rays should be collected to balance the target classes and avoid the need for data augmentation. Also, a similar project could be performed on adult X-ray data to see if the results differ.

References

Hashmi MF, Katiyar S, Keskar AG, Bokde ND, Geem ZW. Efficient Pneumonia Detection in Chest Xray Images Using Deep Transfer Learning. Diagnostics (Basel). 2020;10(6):417. Published 2020 Jun 19. doi:10.3390/diagnostics10060417