Pneumonia Detection using Convolutional Neural Network(CNN)

By: Temesgen Tesfay

September - 6th, 2021

Key fact:

- Pneumonia accounts for 15% of all deaths of children under 5 years old, killing 808 694 children in 2017.

Description of the Pneumonia Dataset

The dataset is organized into 3 folders (train, test, and validation) and contains subfolders for each image category (Pneumonia/Normal). There are 5,863 X-Ray images (JPEG) and 2 categories (Pneumonia/Normal). Chest X-ray imaging was performed as part of patients' routine clinical care. Chest x-ray images have been used to feed the model.

Data Augmentation

Data augmentation techniques deal with cases where the training data is limited. The techniques allow us to modify or even artificially synthesize more data thereby boosting the performance of a machine by reducing overfitting. The idea is to alter the training data with small transformations to reproduce the variations. These approaches that alter the training data in ways that change the array representation while keeping the label the same are known as data augmentation techniques. Some popular augmentations people use are grayscales, horizontal flips, vertical flips, random crops, color jitters, translations, rotations, and much more. By applying just a couple of these transformations to our training data, we can easily double or triple the number of training examples and create a very robust model.

- **Source**

  https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

  Description

  - The data has 1.2 GB of images, and files in .jpg containing the respective labels.
  - Classes labels, image is labeled as :
    - Pneumonia
    - Normal

The dataset has 5856 images, the images are classified among train, test, and validation. 89% allotted to train set, 10.6% to test set, and the remaining fraction is to the validation set. There is a data imbalance between the training and test, which might lead to model overfitting.

Exploratory Data Analysis(EDA)

Occurrence of Labels

The dataset contains over 5000 training images and additional 624 images for testing. Each image is stored in .jpg files The Distribution of the label instances in the training set can be seen in the bar graph below:

The distribution of dataset group between normal and pneumonia classes

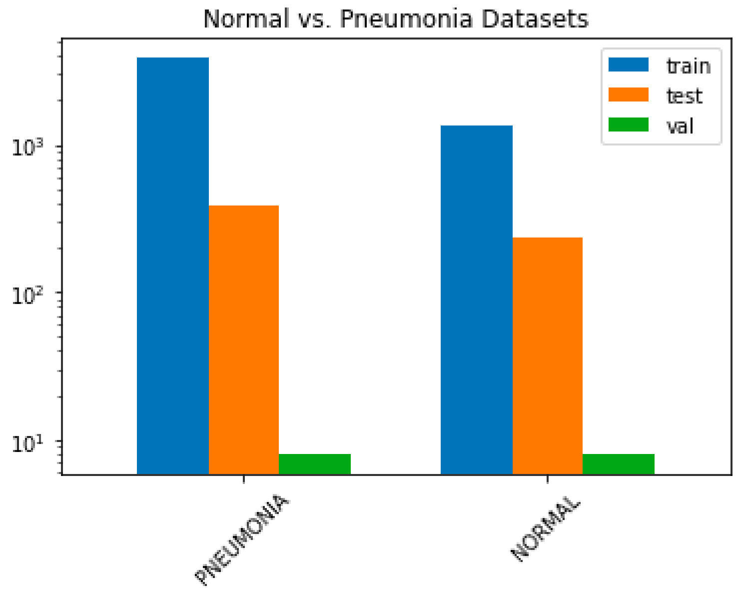

Fig 1.1 dataset proportion between normal and pneumonia

Fig 1.1 the sample size among the group has a huge difference hence, used log distribution to resize the bar graph for visualization. There is a clear indication of data imbalance.

The Distribution of Normal and Pneumonia images

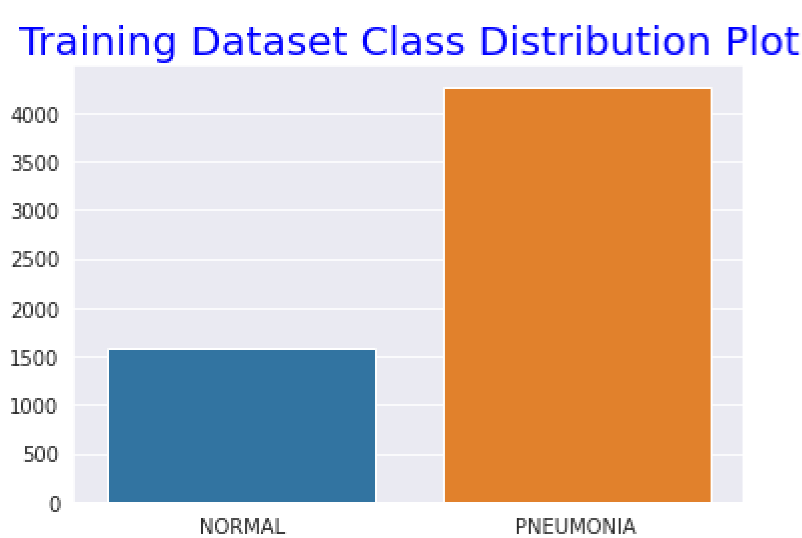

Fig 2.1 normal and pneumonia distribution

The training dataset is an imbalance to fed the model as input for convolutional neural network

A class of pneumonia versus normal lung images

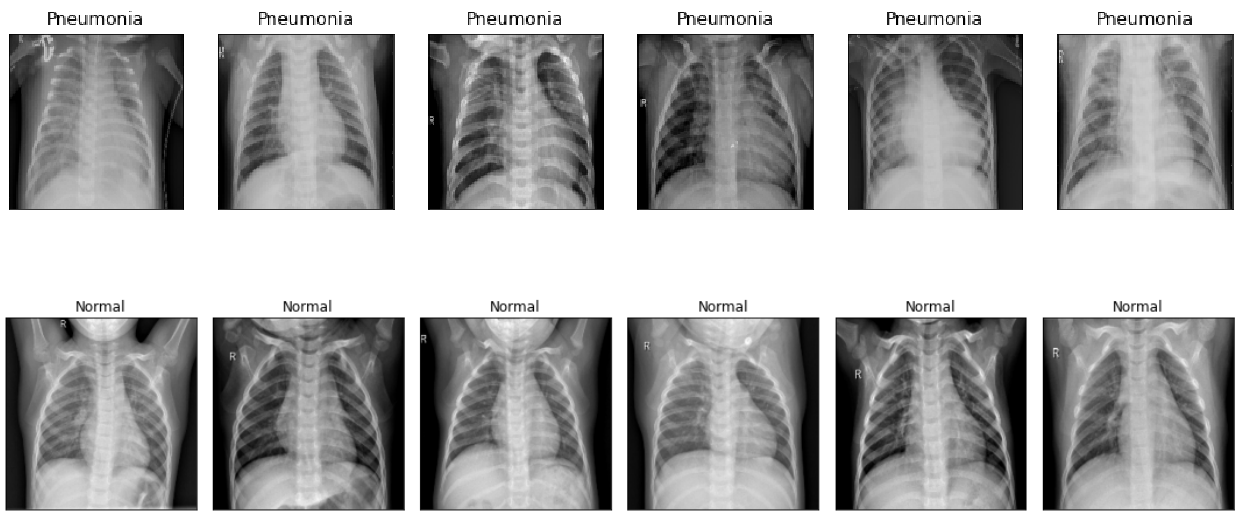

Fig 3.1 Sample images of pneumonia and normal chest-Xray image

Most of the pneumonia-positive lungs show a wide irregular whitish area versus the normal dark brown regular pattern.

Convolutional Neural Networks(CNN) Model

In deep learning, a convolutional neural network (CNN, or ConvNet) is a class of artificial neural networks, most commonly applied to analyze visual imagery. They are also known as shift invariant or space invariant artificial neural networks(SIANN), based on the shared-weight architecture of the convolution kernels or filters that slide along input features and provide a translation. Responses are known as feature maps. CNN is used in applications in image and video analysis, natural language processing, brain-computer interfaces, and financial time series.

A fully connected CNN model was created using Keras on top of TensorFlow. The base model of CNN architecture was created using ResNet50V2 Keras applications, and its parameters are as follows:

- Image resolution input shape = (height = 224, width, 22, and 3 channels = (220,220,3)
- Training set size: 5216
- Validation set size: 16
- Test set size: 624
- Batch size = 32
- Early Stopping: Monitoring validation loss with the patience of 4
- Metric: Recall or Sensitivity, and F1 score

The model's Training and Validation accuracies and losses during training are shown in the plots below:

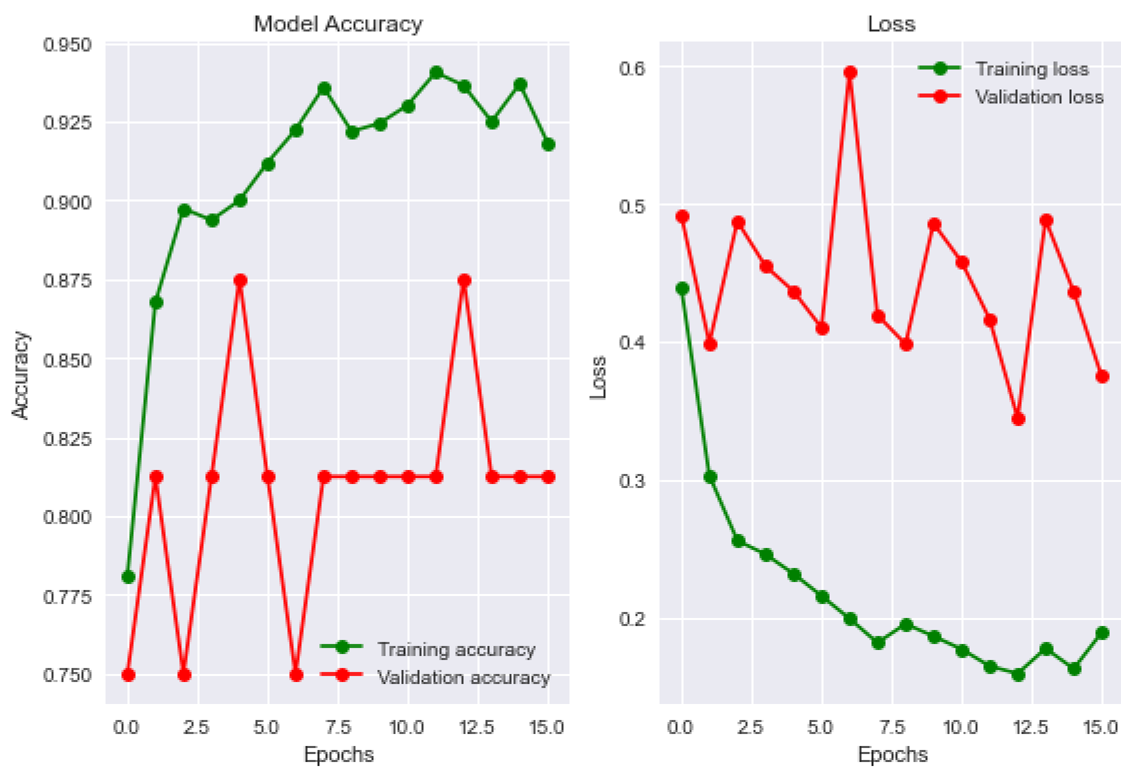

Fig 4.1 Model Accuracy vs Loss Distribution

Confusion Matrix

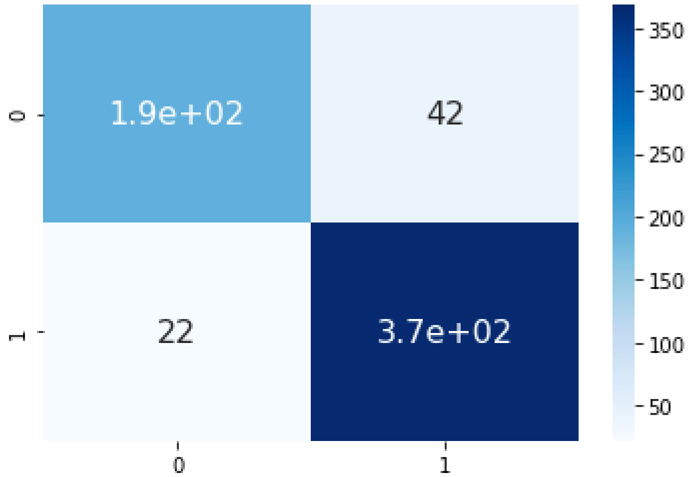

Fig 5.1 Confusion Matrix

The confusion matrix classification report metrics return the following:

	precision	recall	f1-score	support
0	0.90	0.82	0.86	234
1	0.90	0.94	0.92	390
accuracy			0.90	624
macro avg	0.90	0.88	0.89	624
weighted avg	0.90	0.90	0.90	624

The focal point of this project is to minimize false negatives or incorrectly classified pneumonia patients. Therefore, Recall is the most useful metric to check model performance. The project did experiments to improve recall outcomes. Some of the experiments implemented:

- Implemented fully connected convolutional neural network
- Executed various regularization techniques
- Transfer Learning from imagenet

The outcomes are reviewed with the next most important metric, the F score, which is the appropriate metric of imbalance data in classification. However, the type of F_score chosen for the model relies on what the target outcome is aiming to. The project is focused to perform better in a recall, hence F_beta greater than one is the appropriate approach to enhance recall. Hence using beta equals two, F_score raised to 93.5 from the harmonic mean returned 0.89.

Precision, recall, and F0.071 Measure

Fig 6.1 Precision, recall, and F0.935 measure value as a function of model threshold

Several other parameters and models were tested to check model performance to improve baseline CNN classification. Various regularization methods such as relu, batch-normalization, pooling, and dropout techniques were implemented on the subsequent fully connected layer to check the model performance, and the sigmoid activation function to the last layer of binary outputs.

Transfer Learning:

Transfer learning (TL) is a research problem in machine learning (ML) that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem. For example, knowledge gained while learning to recognize cars could apply when trying to recognize trucks. This area of research bears some relation to the long history of psychological literature on the transfer of learning, although practical ties between the two fields are limited. From the practical standpoint, reusing or transferring information from previously learned tasks for the learning of new tasks has the potential to significantly improve the sample efficiency of a reinforcement learning agent.
If the dataset image is a subset of Imagenet, where thousands of images are already pre-trained. Therefore, similar objects efficiently identify with a few layer convolutional network because the most relevant features and their weights are stored in the Imagenet. This reduces the computation cost significantly and performs sometimes better but might depend on the model performance.

The sample of the first filter image in fifth block layers

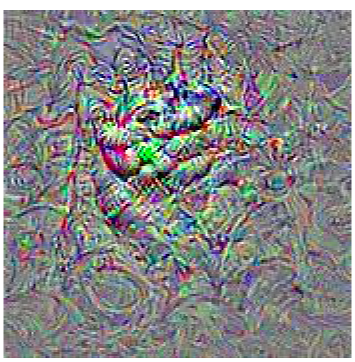

Fig 7.1 Images from a feature from the first index of the last convolutional layer

Conclusion

-
- The model performs well to identify the labeled images. The data used Chest-Xrays images to detect pneumonia symptom images seen on the lung beneath the chest or diaphragm make it impossible to get the pixels from the second deep layer because feature extraction is responsible for collecting pixels from the outermost visible edges, corners, and plane part of the chest and ribs. The ResNet50V2 Keras application for transfer learning performs better to image classification.
- The insufficient data sample size coupled with the imbalance in training and testing contributed to model biases. The score could be better as more data is fed to the model to minimize overfitting and bias problems.
- The model is applied both in deep learning convolutional neural networks and transfer learning just adding a single layer on top of the base model and various regularization techniques. The pre-trained images from the imagenet in transfer learning not only perform better but, it minimized the computational cost.