CNN Techniques for Medical Purposes

Project Report

4 / 12 / 2022

Submitted by Group 17

Tanmay Agarwal	ID - 2020A7PS2057H
Akshat Khaitan	ID - 2020A7PS2055H
Shubhankar Vivek Shastri	ID - 2020A7PS2054H
Dhruv Agrawal	ID - 2020A7PS2193H

Birla Institute of Technology and Science, Pilani
Hyderabad

Project Report for the course Deep Learning

Supervisor Dr. Paresh Saxena

Introduction

At this time, Deep Learning and Machine learning techniques are extensively used in almost all the fields a man can think of. Medicine is no exception. In fact, Deep Learning techniques have been used in molecular biology to discover a whole new paradigm of research. Not only research, but Data Science has a direct impact on day-to-day medical purposes. In this project, we explore one such domain, i.e, Chest X-Rays. We will use a CNN technique to analyze a patient's chest X-Rays to know if they are suffering from any fatal disease, which in general, radiologists also cannot analyze clearly just by looking at the X-Ray. We have cited four research papers that have different CNN techniques used to classify some diseases or problems in any organ.
A CNN is a neural network architecture for deep learning algorithms which is specifically used for image recognition and tasks that involve the processing of pixel data. The CNN-based deep neural system is widely used in the medical classification task. CNN is an excellent feature extractor, therefore utilizing it to classify medical images can avoid complicated and expensive feature engineering.
CNN has been used in many fields of medicine for example it helps detect heart development anomalies and monitors the progression of the same. It also has been used in detecting anomalies in MRI, CAT, and CT scans, and building on this we are going to make a CNN architecture improvement on the same.
After analyzing each paper, we implemented a research paper known as Chex-Net and then used our knowledge gained from reading the rest papers to improve the model to get better prediction accuracy.

Original Chex-Net Paper

Earlier, there was an already implemented Chex-net paper that would classify if a patient is suffering from pneumonia by studying the X-Ray diagnosis of the patient. So the new Chex-Net paper improved the results of the same and further implemented a new classifier to diagnose a patient with 14 different diseases, which demonstrated a groundbreaking result as the accuracy of the model was better than that of the radiologists, making it one of the superhuman intelligence AI.

The paper was implemented on the X-Rays dataset known as the ChestX-Ray 14 Dataset. The paper's importance lies in the fact that detecting pneumonia via chest X-rays is a challenging task that depends primarily on the availability of expert radiologists. The results achieved by the model were groundbreaking, as it could lead to state-of-art medical diagnosis in rural areas as well, which lacked such expert

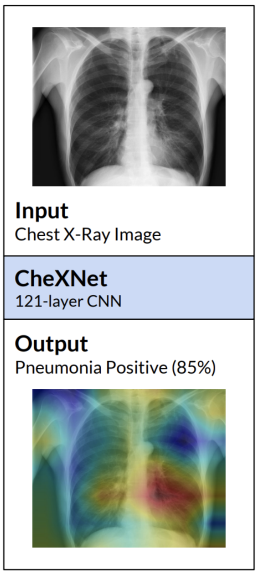


Figure 1. CheXNet is a 121-layer convolutional neural network that takes a chest X-ray image as input, and outputs the probability of a pathology. On this example, CheXnet correctly detects pneumonia and also localizes areas in the image most indicative of the pathology.

radiologists. The Chex-Net model is a 121-layer convolutional neural network (DenseNet) that inputs a chest X-ray and outputs the probability of the patient having pneumonia by generating a heatmap as the single output of the model, as can be seen in Figure 1. The model already proves to be an advancement as it performs better as compared to the general predictions by a radiologist if tested on an F1 metric.

Pneumonia detection is a binary classification task where the input is a frontal chest X-ray image and the output to be generated is y ε {0,1} where 1 signifies the presence of pneumonia and 0 otherwise. The fully connected layer is replaced with a single output, after which a sigmoid nonlinearity is applied. The weights applied to the network had been taken from a pre-trained model on ImageNet.

CNN and Deep Learning Techniques used:

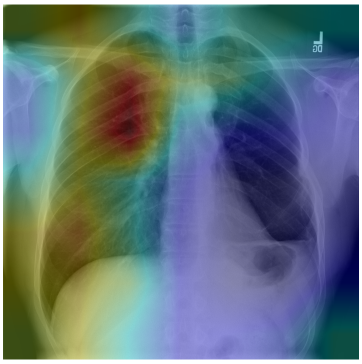
1. During preprocessing of the input data, the images were downscaled to 224 x 224 and normalized.
2. It used a 121-layer Dense convolutional neural network that had more than 10 million parameters to be trained, and after that, a flattened neural network was used with sigmoid classification.
3. The loss function was the weighed binary cross-entropy loss:
   $loss = -w_{+} y log(\widehat{y}) - w_{-}(1 - y) log(1 - \widehat{y})$ where
   $w_{+} = Negative\ cases/(Negative\ cases + Positive\ cases)\ and$
   $w_{-} = Positive\ cases/(Negative\ cases + Positive\ cases)$
4. The loss function was optimized by the well-known Adam Optimizer with its standard parameter setting β1 = 0.9 and β2 = 0.999 using mini-batches of size 16.
5. Initial learning rate of 0.001 was taken, which is decayed by a factor of 10 once the validation loss attains a constant value after certain epochs. The model with the lowest validation loss is chosen.
6. DenseNets simplify the connectivity pattern between layers introduced in other architectures. They solve the problem by ensuring maximum information (and gradient) flow. To do it, they simply connect every layer directly with each other. Instead of drawing representational power from extremely deep or wide architectures, DenseNets exploit the potential of the network through feature reuse.
7. The original paper, while training, made use of transfer learning as well, as it got the weights

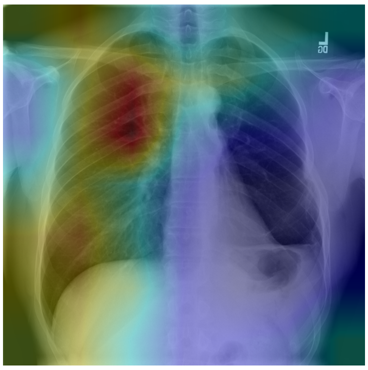
Outcome

The outcome of the paper was it not only classified the images on the basis of pneumonia but also classified images for 14 different pathology classes. The result of this implementation was pretty much more accurate than the actual practicing radiologists could predict as the F1 score for the ChexNet model is statistically significantly higher as compared to the radiologist's performance.

Results of our implementation of the Paper

The AUROC score is the measure of performance of the classification problems at various threshold settings, and this is a metric that we tried to measure the performance of the implementation of the paper as it gives a good idea of the performance for a multi-classification problem.
The original paper gave the AUROC of: 0.841
The implementation that we did gave the total AUROC score of: 0.845.
Also, the original paper gave output as a heatmap for a specific patient, which showed where the problem or the infection of disease is contained, and here is a comparison of the heatmap generated by the original paper and our implementation of the same.





As a result, we can clearly state that the two pictures are actually similar.
Here is the classification table of the implemented Chex-Net and our implementation of the same.

Pathology Classes	Original Chex-Net result	Our implementation result
Atelectasis	0.8094	0.8094
Cardiomegaly	0.9248	0.9165
Effusion	0.8638	0.8870

Infiltration	0.7345	0.7143
Mass	0.8676	0.8597
Nodule	0.7802	0.7873
Pneumonia	0.7680	0.7745
Consolidation	0.8887	0.8726
Consolidation Edema	0.7901	0.8142
Edema	0.8878	0.8932
Emphysema	0.9371	0.9254
Fibrosis	0.8047	0.8304
Pleural Thickening	0.8062	0.7831
Hernia	0.9164	0.9104

Our Improved Model

We added another layer of classification with a linearly activated layer and then followed by the sigmoid activation function. Which improved the AUROC score of our model by a decent margin as the score became 0.8742.
Why did it work?
We took inspiration from different papers that we have cited for the project and came up with this model as we taught that training our softmax with a linear layer followed by another sigmoid classifier can improve the result as the neural networks will get more information from the previous layer and it will train on it a bit better and actually gave us better results for the same. Also, up to a certain point in deep learning, adding more layers will not hurt your accuracy. Also, with this implementation, our computational cost was not affected by a considerable margin as the parameters for the last two layers were pretty less in comparison to the 10 million parameters we had already trained for the whole network hence the time taken for this model did not increase by that much time. So after a lot of trial and error and different kinds of experiments, we got the best result for this implemented model, and this is our improved version of the Chex-Net paper. Here are the results:

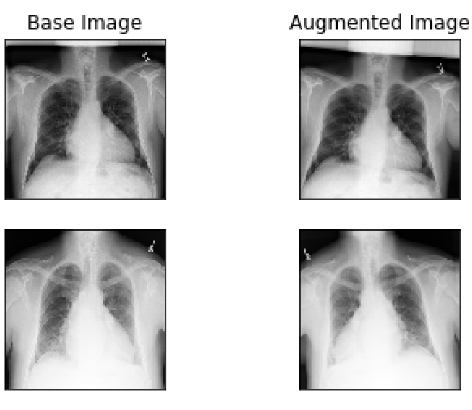
Chex-Net AUROC Score	Our implemented AUROC	Our improved Model
0.841	0.845	0.868

Below is the table showing the prediction accuracy of each pathology class:

Pathology Classes	Original Chex-Net result	Our implementation result of Chex-Net	Our Improved Model
Atelectasis	0.8094	0.8094	0.8311
Cardiomegaly	0.9248	0.9165	0.9220
Effusion	0.8638	0.8870	0.8891
Infiltration	0.7345	0.7143	0.7146
Mass	0.8676	0.8597	0.8627
Nodule	0.7802	0.7873	0.7883
Pneumonia	0.7680	0.7745	0.7820
Consolidation	0.8887	0.8726	0.8844
Consolidation Edema	0.7901	0.8142	0.8148
Edema	0.8878	0.8932	0.8992
Emphysema	0.9371	0.9254	0.9342
Fibrosis	0.8047	0.8304	0.8385
Pleural Thickening	0.8062	0.7831	0.7914
Hernia	0.9164	0.9104	0.9206

Experiments Done for our model:

- Firstly with the model, we tried to use 2 different optimizers apart from the already implemented one which is ADAM we tried RMS-prop and Ada-grad which we tried only for a small batch of images because of the computing time and we got the best results for the adam optimizer.
- After We found that the best optimizer was ADAM, we tried hyper-parameter tuning for the parameters **β**1 and **β**2, and we finally saw that the original parameters that were used in the original paper were giving the best result, so we did not change the parameters.
- We tried Data Augmentation of some images to have more data to train our model here is a sample of the same where we turned the image angle by 45 degrees and also made a mirror image which you can see



in the augmented image 1 and 2 respectively. After Data Augmentation we got a bit better results which gave the AUROC of: 0.85

- As we had used the DenseNet and a well-shuffled dataset we did not face any overfitting issues, but still we tried to implement the same and did not get any better results for our test dataset.
- We also experimented by adding a relu layer instead of softmax which was followed by a softmax layer and it resulted in a slightly better result which indicated us towards our finally implemented layer.

Main Findings:

1) We found that Dense-Net Architecture gives us a better result than either a CNN or Res-Net neural network architecture for such large and delicate data.
2) We also found that adding a classification layer can actually help us with the prediction in some cases (especially for this one).
3) We found that data augmentation which adds more data to our training data set actually helped with training and improving the accuracy of our model.
4) ADAM optimizer also works better when there is a good amount of data compared to other optimizers such as RMS-Prop and Ada-grad.

Accomplishments of the Project:

1. We accomplished the task of implementing an advanced real-world research paper with the same results.
2. We improved the accuracy of the original network by a decent margin.
3. We made a classifier that actually predicted if a patient has a chest disease only by reading the X-ray images and gave the result of 14 different diseases a person can suffer from.
4. Implemented one of the rare examples where our neural network gave results better than humans specifically radiologists could do.

Papers Cited:

1. CheXNet:

[1711.05225] CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning (arxiv.org)

2. CNN for Coronary Artery Disease:

[2001.08593] CNN-CASS: CNN for Classification of Coronary Artery Stenosis Score in MPR Images (arxiv.org)

3. Covid using Chest XRay:

COVID-19 Detection using Convolutional Neural Network Architectures based upon Chest X-rays Images | IEEE Conference Publication | IEEE Xplore

4. CNN For Heart Disease Prediction

(PDF) Prediction of Heart Disease using CNN (researchgate.net)

Contributions of the Group Members:

Member Name	Contribution
Tanmay Agarwal	Implemented the original paper and experimentation
Akshat Khaitan	Data preprocessing and Data Augmentation
Dhruv Agrawal	Implemented the new model
Shubhankar Vivek Shastri	Implemented the new model and defined error metrics for the old and new model.

Common Contribution: Literature Survey, Brainstorming the proposed changes in the original model, and Report Drafting.

GitHub link - Tan141101/DL_Project: Chex-Net implementation and improvisation (github.com)