# Detection of Lung Diseases using Deep Learning

Mrs. Rasika Naik
Assistant Professor,
Vivekanand Education Society's
Institute of Technology
Mumbai, India
rasika.naik@ves.ac.in

Mr. Tejas Wani
Student, Vivekanand Education
Society's Institute of Technology
Mumbai, India
2016.tejas.wani@ves.ac.in

Ms. Sakshi Bajaj
Student, Vivekanand Education
Society's Institute of Technology
Mumbai, India
2016.sakshi.bajaj@ves.ac.in

Mr. Shiva Ahir
Student, Vivekanand Education
Society's Institute of Technology
Mumbai, India
2016shiva.ahir@ves.ac.in

Mr. Atharva Joshi
Student, Vivekanand Education
Society's Institute of Technology
Mumbai, India
2016.atharva.joshi@ves.ac.in

***Abstract* – There is an increasing risk on health due to changing environment, climate and lifestyle. India tops the world in deaths due to lung diseases. They were the second highest cause of deaths in India after heart disease in 2017, killing 1 million (958,000) Indians that year. Early diagnosis and treatment of lung diseases is critical to prevent complications including death. Chest X-ray is currently the best available method for diagnosis, playing a crucial role in clinical care. Using Deep Learning to predict lung diseases from Chest X-rays can be a lifesaving factor for an individual suffering from the disease. This is possible as the results can be predicted with a high percentage of accuracy instantly. This paper presents an effective way for expert diagnosis of lung diseases using Deep Learning. It focuses on creating a system for assistance of Radiologists in detection of lung diseases. This will especially benefit rural areas where radiologists aren't easily available. Our system connects radiology labs with the radiologists who can diagnose faster and better with our model.**

## I. INTRODUCTION

Lung diseases are considered to be frequently occurring medical anomalies around the globe. Many of the people suffer from various types of lung diseases in India. Genetics, infections and smoking are probable cause for such diseases. The lungs are vital organs that expand and relax many times each day to expel carbon dioxide and breathe oxygen. About 3 million deaths occur every year due to Lung diseases around the world. It is the leading cause of death among young people, especially children. This number can be reduced in a significant amount in the coming future. This could be possible if the type of disease is detected accurately as every other disease has other treatment. There are possibilities that one disease could be mistaken for another. Designing a typical Lung Disease detecting Deep Learning model is going to help rule out these possibilities of errors and help save lives when given proper treatments at appropriate time.

Today, Low dose computed tomography (LDCT) is popular in detecting lung cancer. LDCT has significantly reduced lung cancer deaths and it is advised. However it is an expensive method and everyone cannot afford it. So generally X-rays are used for diagnosis. Also people in rural areas have limited access to both doctors and specialists. X-rays being most widely used to detect disease, our aim has been to make diagnose X-rays to near perfection.

Convolutional neural network (CNN) is proven to be very effective in image recognition and classification tasks. DenseNet being a CNN based pre-trained model, it could be a great choice for training our model. For training this model, Densenet-121 has been used. While for security of Images, AES algorithm is found to be one of the best. Here, AES-128 is used for securing images considering the sensitivity of medical data.

## II. LITERATURE SURVEY

Y. Zhang et. al. (2017) [5] talks about AES cryptography in his paper. The plain image is divided into data blocks of 128 bits size each. The first block of plain image is permuted by an initial vector. Then, each block is encrypted sequentially using the cipher block chain method. The initial vector is generated and the cipher image is then transmitted for decryption through the public information channel. During the decryption process , AES uses the secret key and initial vector to decrypt the cipher image to obtain the original image. Simulation results show that this image cryptosystem is both secure and high-speed, which can be used as the comparison basement of newly proposed image cryptosystems based on chaotic systems.

Qiao Ke et. al. (2019) [17] in his paper proposes that the neuro-heuristic approach addresses small changes in the structure of lung tissues, which appear in pneumonia, sarcoidosis or cancer and some consequences that may appear after the treatment. After testing this approach, results obtained are showing high potential of this newly proposed method. Moreover this method is flexible, it also has a low computational burden**.**

Rakshit S. et. al. (2019) [18] states that Chest X-Rays are very popular source of medical images to find out lung and heart diseases and it is stated that previous works have been explored to understand the use of different pre-trained deep learning models like Resnet and Densenet to perform the classification. Here it is discussed about how the proposed model (network of Resnet18) is having few parameters for training and how it is having significantly good performance among the models which have been tested in the past.

Justin Ker et. al. (2018) [10] discussed how the machine learning algorithms can be used for analysis of medical images, significance of convolutional neural networks.. Here the significance of deep learning for detection of certain medical conditions is discussed.Also better use of neural networks for deep learning and emphasizing its clinical aspects is highlighted.

Rajpurkar et. al. (2017) [14] in his paper states about the Chexnet model which is used for prediction of lung diseases.The final output is the predicted probability of the presence of each pathology class.For interpretation of the network predictions, heatmap is generated to highlight the areas of the image which is having the disease infection using class activation mappings (CAMs). CAM is generated when an image is fed into the network which is fully trained and they extract the feature maps that are output of the final convolutional layer. With automation at the level of experts, it is hoped that this technology can improve healthcare system and where the access to skilled radiologists is limited, medical imaging with this technology can be helpful.

## III. SYSTEM FLOW

The input to the system is chest x-ray and the result is prediction of lung diseases. The system flow is as follows:

Patient visits a radiology lab to get Chest X-ray where there are no expert radiologists available. Radiology labs will send the encrypted X-ray to the radiologist via a secure cloud. Now an expert radiologist can access this image and feed it to our model for results. The output of this system will be a list of possible lung diseases with probabilities of each. Now radiologists could analyze the result and interpret it to get to a faster and more accurate decision. Radiologists can share this result with the patient as well as the consulting doctor. Doctors can analyze the results forwarded by the radiologist and prescribe appropriate treatment for the related disease.

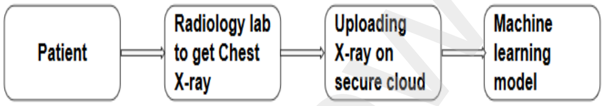


Fig. Desktop Application

Also, a mobile application is made which can be used by any person. Person can upload the image of his/her Chest X-ray to get an immediate response from the model. Here, there will not be an expert radiologist's involvement. This would totally be an individual based application which can only suggest if there is a need of worry.

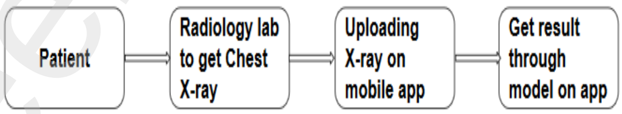


Fig. Mobile Application

The System can be used to detect fourteen common thoracic pathologies including Pleural Thickening, Consolidation, Edema, Cardiomegaly, Infiltration, Atelectasis, Pneumonia.

## IV. METHODOLOGY

**a. Dataset:**

NIH Chest X-ray : We have used NIH Chest X-ray dataset which is a large public dataset for chest radiograph interpretation, consisting of 112,120 chest radiographs of over 30,000 patients. It is a collection of chest radiographic examinations provided by National Institute of Health Care, America performed between October 2002 and July 2017, along with their associated radiology reports. We use these to train our model using deep learning.

**b. Data cleaning and preprocessing:**

Images with missing data about patient details, having poor quality or non-frontal view were removed. Total 2355 images were removed. After cleaning, CLAHE i.e Contrast Limited Adaptive Histogram Equalization technique was used for the image enhancement.

**c. Training of model:**

1. Loss function :

In DL, loss function (also known as objective function) is a mathematical formula that describes the penalty given to a model when it makes a wrong decision. We use the loss function to optimize the parameter of the model so that it fits better to the data. Lung disease classification can be seen as a classification problem, so we will use Binary Cross Entropy Loss as baseline loss to develop our objective functions.

2. Training settings

-DenseNet121 with pre-trained weight from ImageNet
-Adam optimizer with default hyper-parameters
-Initial learning rate = 0.0001 with decay by 10 scheduler
-Epochs = 100, batches = 500, batch size = 64

Note: **Batch size** : Smaller batch size adds more noise as they are not represented for the entire dataset. After hyper-parameter tuning we found that batch size of 64 images gives the best performance gain.
**Epochs** : Model is trained in the same setting with 100 longer epochs each time. The final conclusion was that after around 100 epochs, the performance has no significant improvement.

**d. Encryption of image**

AES 128 bit encryption is used for encryption of images to provide security. AES encryption is considered unbreakable. AES encrypted images are then uploaded on cloud. The key used for encryption and decryption is symmetric and is shared between the radiology lab and radiologist.

**e. Model evaluation**

Evaluation is done by seeing and comparing AUROC scores, training and testing accuracy, and confusion matrix.
Comparison of AUROC scores with different papers is given further in this paper.

## V. DENSENET-121

It is a fully convolutional network which has become one of the most successful deep learning frameworks for generic image classification and segmentation tasks. DenseNet pushed the idea of residual connection by iteratively concatenating all previous features. A DenseNet comprises many densely connected convolutional layers which are used for purpose of improving the information flow and gradients in network to converge better and mitigate gradient vanishing problems. Therefore, in many computer vision tasks, DenseNet has shown magnificently stronger capability of representation learning. We choose the publicly available DenseNet-121 model as backbone network. DenseNet consists of four consecutive dense blocks.

## VI. ADVANCED ENCRYPTION STANDARD

In this paper, we have used AES (Advanced Encryption Standard) algorithm for encryption of X ray images. AES functions by the use of a symmetric algorithm (i.e., the same key is used in encryption and decryption). In this paper, 128-bit key has been used for the purpose, which is shared between the radiology lab and the radiologist for the encryption and decryption of X ray images.

It should be noted that the 192-bit and 256-bit versions are theoretically more difficult to "crack" than AES 128-bit encryption, the difference is really moot in a practical sense. The EE Times points out that even using a supercomputer, a "brute force" attack would take one billion years to crack AES 128-bit encryption. So the fact that it would take at least an order of magnitude longer to crack AES 256-bit encryption. Though it is interesting but is of little practical relevance right now.

## VII. OBSERVATIONS AND RESULTS

**1) Desktop app for radiologist's assistance:**

An application made only for the radiologists. This will be used for their assistance where in our model makes predictions. The app as shown in the figure:

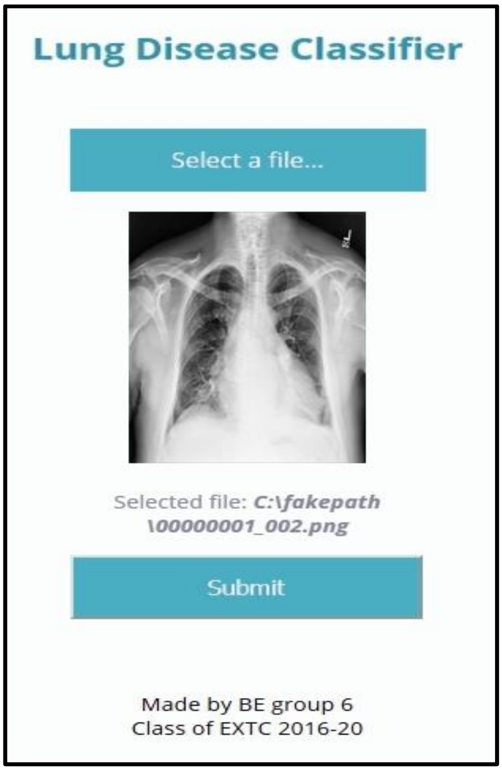


Fig. Desktop Application

| Disease | Probability |
|---|---|
| Cardiomegaly | 99.98 |
| Fibrosis | 62.25 |
| Effusion | 61.73 |
| Infiltration | 51.26 |
| Edema | 46.8 |
| Pneumonia | 45.27 |
| Atelectasis | 44.33 |

Fig : Predictions made

**2) Using Google cloud platform for making smartphone application:**

.

By using Google Cloud Application Engine, the code for the frontend of the website was deployed. For frontend JavaScript framework was used. A docker image was built which included the backend code and the trained model. This image was pushed into Google registry. A new Google Cloud run service was created with this image. The request to get the prediction by the user is then handled by Cloud Run service. Image was stored in Cloud storage. It is as shown in the figure under.

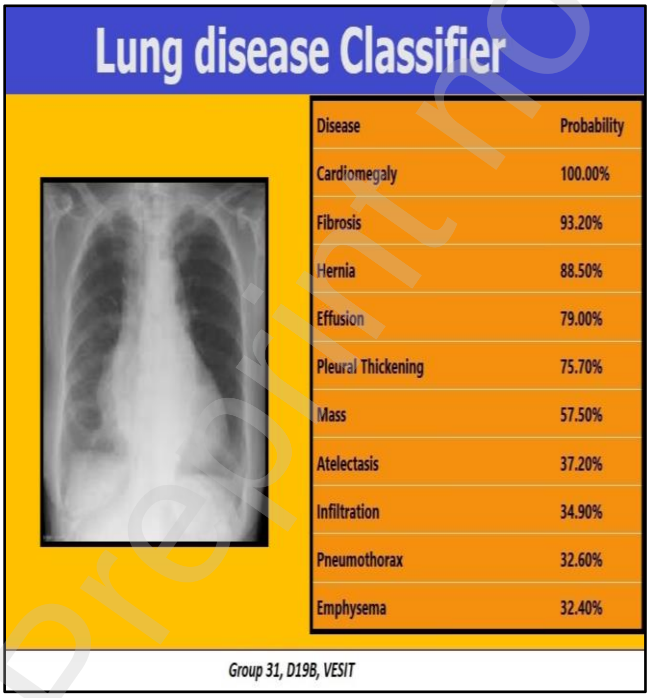


Fig : Predictions made by the app

**3) AUC score comparisons for different models:**

From the table displayed above, we can observe the result clearly mentioning the AUC scores for the models telling which one is the best for different classes (lung diseases). Average AUC of 0.843 was achieved.

| Diseases | Wang et. al. | Yao et. al. | Rajpurkar et. al. | Our model |
|---|---|---|---|---|
| Atelectasis | 0.716 | 0.772 | 0.8094 | 0.831 |
| Cardiomegaly | 0.807 | 0.904 | 0.9248 | 0.918 |
| Consolidation | 0.708 | 0.788 | 0.7901 | 0.813 |
| Edema | 0.835 | 0.882 | 0.8878 | 0.899 |
| Effusion | 0.784 | 0.859 | 0.8638 | 0.883 |
| Emphysema | 0.815 | 0.829 | 0.9371 | 0.911 |
| Fibrosis | 0.769 | 0.767 | 0.8047 | 0.826 |
| Hernia | 0.767 | 0.914 | 0.9164 | 0.943 |
| Infiltration | 0.609 | 0.695 | 0.7345 | 0.712 |
| Mass | 0.706 | 0.792 | 0.8676 | 0.859 |
| Nodule | 0.671 | 0.717 | 0.7802 | 0.789 |
| Pleural Thickening | 0.708 | 0.765 | 0.8062 | 0.782 |
| Pneumonia | 0.633 | 0.713 | 0.768 | 0.765 |
| Pneumothorax | 0.806 | 0.841 | 0.8887 | 0.88 |
| **Mean** | **0.738** | **0.8027** | **0.841** | **0.843** |

Fig: Comparison of AUROC scores

## VIII. CONCLUSION

The development of a CAD (Computer Aided Diagnosis) which can help the radiologist to make better decisions faster is very important. In future, this system can be extended in rural areas where getting urgent medical help is difficult. Doctors, radiologists and patients can be added as users so that patients can take expert advice from them at any point. Concept of verified users will also be added. Currently the system is developed for detecting lung diseases. In future, more types of lung disease detection can be added in the system. Also the system can be extended to detection of other medical issues also by developing the new machine learning models. Moreover the approach for localization of lung diseases in Chest X-ray images improvements can be made in algorithms to generate the activation maps. Also accuracy of the deep learning model can be significantly improved by training on improved datasets.

## IX. References


[1] M. Anthimopoulos et. al. "Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network", *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1207-1216,2016. Available: 10.1109/tmi.2016.2535865.

[2] J. Tong, Z. Da-Zhe, W. Yung, Z. Xin-Hua and W. Xu, "Computer-Aided Lung Nodule Detection Based on CT Images", *2007 IEEE/ICME International Conference on Complex Medical Engineering*, 2007.

[3] P. Hattikatti, "Texture based interstitial lung disease detection using convolutional neural network", *2017 International Conference on Big Data, IoT and Data Science (BID)*, 2017.

[4] R. Pich, S. Chivapreecha and J. Prabnasak, "A Single, Triple Chaotic Cryptography Using Chaos in Digital Filter and Its Own Comparison to DES and Triple DES", *2018 International Workshop on Advanced Image Technology (IWAIT)*, 2018.

[5] Y. Zhang, W. Hou and X. Li, "A fast image encryption scheme based on AES", *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*, 2017.

[6] N. Kumar, "IoT Architecture and System Design for Healthcare Systems", *2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon)*, 2017.

[7] A. Sterling, "Frontend JavaScript: NodeJS and Angular Tools for JSON-LD", *2019 IEEE 13th International Conference on Semantic Computing (ICSC)*, 2019.

[8] G. Huang, Z. Liu, L. v. d. Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 2261-2269.

[9] S. Anakal and S. P, "Clinical decision support system for chronic obstructive pulmonary disease using machine learning techniques", *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)*, 2017.

[10] J. Ker, L. Wang, J. Rao and T. Lim, "Deep Learning Applications in Medical Image Analysis," in *IEEE Access*, vol. 6, pp. 9375-9389, 2018.

[11] T. Guo, J. Dong, H. Li and Y. Gao, "Simple convolutional neural network on image classification," *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)(*, Beijing, 2017, pp. 721-724.

[12] W. G. Hatcher and W. Yu, "A Survey of Deep Learning: Platforms, Applications and Emerging Research Trends," in *IEEE Access*, vol. 6, pp. 24411-24432, 2018.

[13] M. Faisal, S. Bashir and Z. Khan, "An Evaluation of Machine Learning Classifiers and Ensembles for Early Stage Prediction of Lung Cancer", *(ICEEST)*, 2018.

[14] Pranav R. et al., CheXNet: Radiologist level pneumonia detection on chest X-rays with deep learning, Dec.2017, [online] Available: https://arxiv.org /abs/1711.05225.

[15] X. Wang et. al. , ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases, Dec. 2017, https://arxiv.org/abs/1705.02315.

[16] H.-C. Shin et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures dataset characteristics and transfer learning", *IEEE Transactions in Medical Imaging*, May 2016.

[17] Qiao Ke, Zhang, J., Wei, W., Połap, D., Woźniak, M., Kośmider, L., & Damaševĭcius, R. (2019). A neuro-heuristic approach for recognition of lung diseases from X-ray images. Expert Systems with Applications. doi:10.1016/j.eswa.2019.01.060

[18] S. Rakshit, I. Saha, M. Wlasnowolski, U. Maulik, and D. Plewczynski, ‘‘Deep Learning for Detection and Localization of Thoracic Diseases Using Chest X-Ray Imagery’’, in Artificial Intelligence and SoftComputing, pp. 271–282, 2019.