



# LUNG OPACITY IDENTIFICATION USING MATHEMATICAL MODEL BASED ON DEEP LEARNING

Mahamudul Hashan Antor  
Department of Computer Science and Engineering,  
National University of Science and Technology MISIS, Moscow, Russia

**Abstract** - In this paper, “Identifying Lung Opacity Using a Mathematical Model Based on Deep Learning,” I presented a deep learning architecture for diagnosing lung disease during chest X-rays. I used the recently created RSNA lung disease dataset. This is the ImageNet dataset. This dataset is the first and so far the only dataset to specialize in the detection of constitutive areas of Lung Opacity on chest x-ray and will allow us to train neural networks specialized in classification and object detection tasks. I used the Convolutional Neural Networks (CNN) algorithm because CNN-based deep learning classification approaches have the ability to automatically extract high-level representations from big data using little preprocessing compared to other image classification algorithms. The Xception model was used for the classification tasks. At the same time, this project introduced learning strategies and prediction strategies using the Python and Keras deep learning library.

**Keyword** - RSNA, Deep Learning, Convolutional Neural Network, Xception Model, ImageNet Dataset, Object Detection.

## I. INTRODUCTION

Thoracic diseases are a big problem for human health. Although it affects millions of lives worldwide and causes significant morbidity and mortality worldwide, 90% of deaths associated with chronic obstructive pulmonary disease (COPD) occur in low- and middle-income countries. It's a global health problem that does not understand the social or cultural layers, causing millions of deaths every year.

In developing countries, the impact of this lung disease makes it one of the deadliest among children under 5 years of age, causing 15% of deaths recorded each year. In Spain, on the other hand, people over 65 are most at risk. There are 162 cases per year per 100,000 inhabitants, with a mortality rate of 5% for outpatients, 12% for hospitalization and up to 40% for the intensive care unit.

Lung opacity is an infectious disease that causes pneumonia. Our body's immune response to this infection leads to the fact that the lungs are filled with pus or other fluids that reduce the ability to hold air, so the patient may feel suffocation, coughing

and fever among other symptoms. Lung opacity spreads to the lung region of the human body. A chest x-ray is performed to diagnose this infection. Doctors use this x-ray image to diagnose or monitor the treatment of lung opacity conditions. This type of x-ray is also used to emphysema, lung cancer and tube placement, tuberculosis, and etc. The specific objectives are:

1. Help radiologists and medical experts in identifying gaps and slow changes among the many x-rays that might otherwise be missed.
2. Many people in developing countries do not have access to radiologists because of the high costs. This tool can help them read their x-ray images.
3. Create a foundation for the model to read more complex data, such as CT and MRI images, in the near future.

## II. RELATED WORKS

The chest x-ray is currently the best available method for diagnosing lung opacity, which plays a crucial role in clinical care and epidemiological studies. In developing these health systems, numerous documents support the distinction between two phases:

1. Preliminary Deep Training: In the early 2000s, the first computer detection (CAD) was created. These were systems built using traditional machine learning methods in which functions from the input data were selected and extracted completely manually and based on criteria that the engineer considered appropriate.
2. After Deep Training: Deep learning would bring a revolutionary breakthrough, changing the way the model interacts and learning to autonomously detect complex patterns and functions directly from input data. In the field of computer vision, an important milestone was achieved when, in 2012, the ImageNet deep learning model for large-scale visual recognition, ImageNet, based on convolutional networks, surpassed the performance achieved with other machine learning methods.

Recent improvements in deep learning models and having huge data sets helped algorithms Outperform medical staff in many medical imaging tasks such as classification of skin cancer, detecting hemorrhage, detecting arrhythmias, and detecting diabetic retinopathy. Automated diagnoses, including chest radiographs, have gained growing interest.

### III. DATASET

Using the National Institutes of Health (NIH) chest x-ray dataset, which contains more than 112,000 chest x-rays from more than 30,000 unique patients issued by the Radiological Society of North America (RSNA). This dataset provided me with the annotation data needed to train and evaluate my detection and classification architecture. Since there are too many categories, I have drop-down categories that have less than 5000 cases. For the discovery task, I randomly split the data set into training, validation, and testing. I used 80% (32,000 images) for training and 20% (8,000 images) for testing.

ImageNet is formally a project aimed at labeling and categorizing images into nearly 22,000 distinct object categories for deep learning purposes. The main objective of this ImageNet dataset to train a model, which can classify an Input Images into 1000 separate object categories. ImageNet Optimization configuration:

- Optimizer: SGD
- Momentum: 0.9
- Initial learning rate: 0.045
- Decrease Learning Rate: Decrease the learning rate by 0.94 every 2 epochs.

### IV. CONVOLUTIONAL NEURAL NETWORK

Convolutional neural network (ConvNets or CNNs) is one of the main categories to do images recognition, images classifications, Objects detections, recognition faces etc.

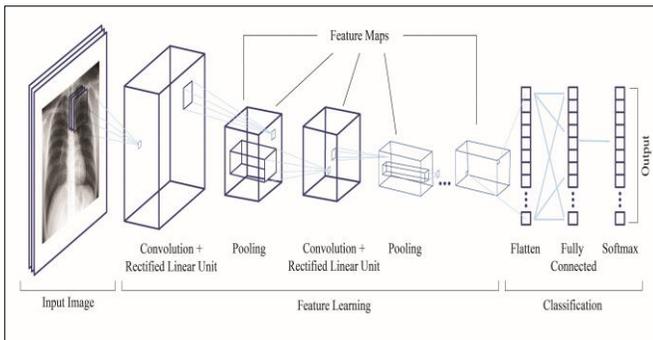


Figure 1: Convolutional Neural Network Architecture

CNN is a special type of Feed-Forward Artificial Neural network, where neurons receive input for transforming to hidden layers. Every hidden layer is fully connected to all neurons of the previous layer. It has four main layers:

1. Convolution Layer: Convolution is the first layer to extract features from an input image. The convolutional layer uses a matrix of filters over an array of image pixels.
2. Rectified Linear Unit: The purpose of the rectifier function is to increase the non-linearity of the network. Like sigmoid function. The rectifier function removes all the black elements, keeping only positive value. Negative pixels to zero.

3. Pooling Layer: It basically reduces the dimensionality of the feature map or reduce the number of parameters when the images are too large. The max pool takes the largest item from the revised feature map.
4. Fully Connected Layer: The output of the last Pooling Layer acts as input for the Fully Connected Layer. The output of both convolution and pooling layers are 3D volumes (Matrix), but a fully connected layer expects a 1D vector of numbers (Vectors). So, I'm smoothing the result of the last vector merge layer. Anti-aliasing is simply the conversion of a three-dimensional volume of numbers to a one-dimensional vector.

### V. MODIFIED DEPTHWISE SEPARABLE CONVOLUTION

Modified depth-separable convolution is point convolution. The original depth-separable convolutions first perform channel-wise convolution and then perform  $1 \times 1$  convolution. But the modified depth-separable convolution first performs  $1 \times 1$  convolution and then channel spatial convolution.

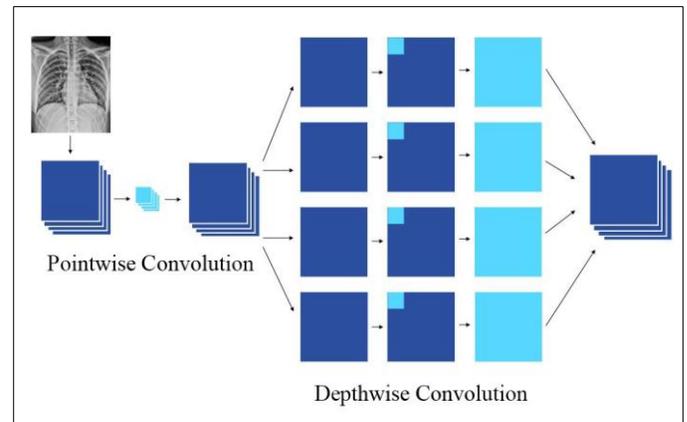


Figure 2: Depthwise Separable Convolution

Xception offers an architecture that consists of Depthwise Separable Convolution, Max pooling blocks, all of which are associated with shortcuts, as in ImageNet implementations. The specifics of Xception is that Depthwise Convolution is not followed by Pointwise Convolution, but the order is reversed. A modified deep separable convolution is a pointwise convolution followed by a deep convolution. This modification is motivated by the initial module in Inception-v3 in that the  $1 \times 1$  convolution is performed first before any  $n \times n$  spatial convolutions.

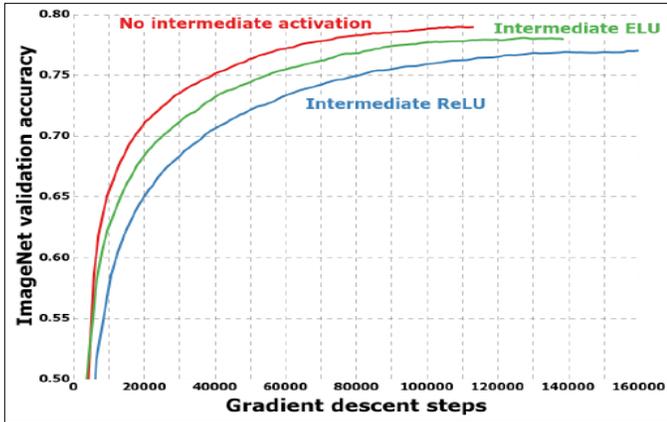


Figure 3: Different Activation Units

In original Model, there is non-linearity after first operation. Modified depthwise separable convolution, there is NO intermediate non-linearity (ReLU). Without any intermediate activation, it has the highest accuracy compared with other ELU or ReLU function.

### VI. XCEPTION MODEL

There have been many algorithms in the development of deep learning using convolutional neural networks (CNN). The Xception model is one of them. This model, developed by Francois, surpassed previous records in speed and accuracy inherent in the architecture of Inception. The architecture of the Xception model is based on depth-separable convolution layers and consists of three main sections: Entry Flow, Middle Flow, and Exit Flow.

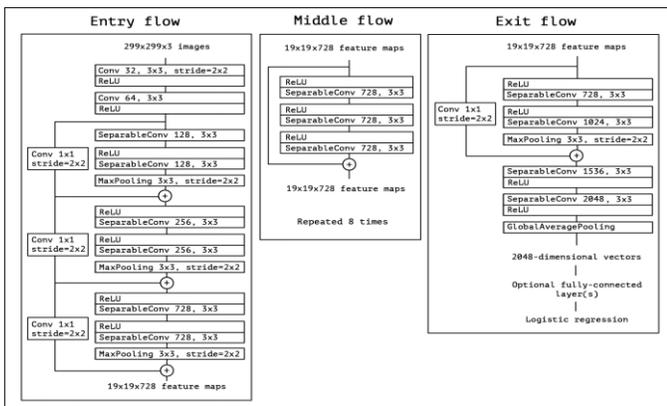


Figure 4: Xception Architecture

The Xception architecture has 36 convolutional layers that form the basis for network feature extraction. This architecture is very easy to define and change and takes only 30-40 lines of code using a high-level library like TensorFlow. All networks were implemented using the TensorFlow framework.

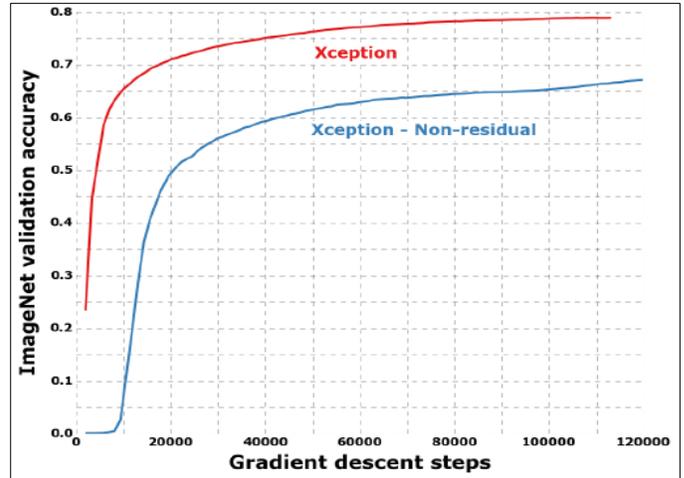


Figure 5: Validation Accuracy Against Gradient Descent

Xception architecture has residual links. In this picture above, we can see that the accuracy is much higher when using residual connections. Residual compounds can pass multiple layers, not just one. In short, residual connections are used to simplify the optimization of deeper networks.

### VII. CLASSIFICATION

I used a binary classification, where a single binary output {0,1} that will represent the presence or absence of lung Opacity.

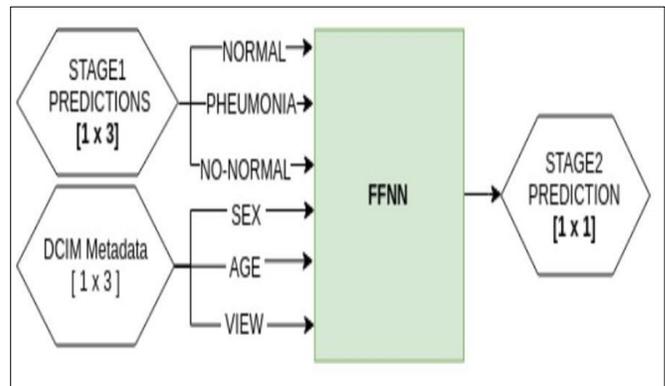


Figure 6: Binary Classification

Also, I used End-to-end Network It can identify presence of lung Opacity. The input divided by 6 features: 3 extracted from DICOM file (sex, age, view) and 3 corresponding to the probabilities of the normal, lung Opacity, and non-normal classes from classification model. As the input vector is having 6 features, I have to arrange the input layer 6 nodes. These nodes implement rectified linear unit and their weights will be initialized by the initialization technique.

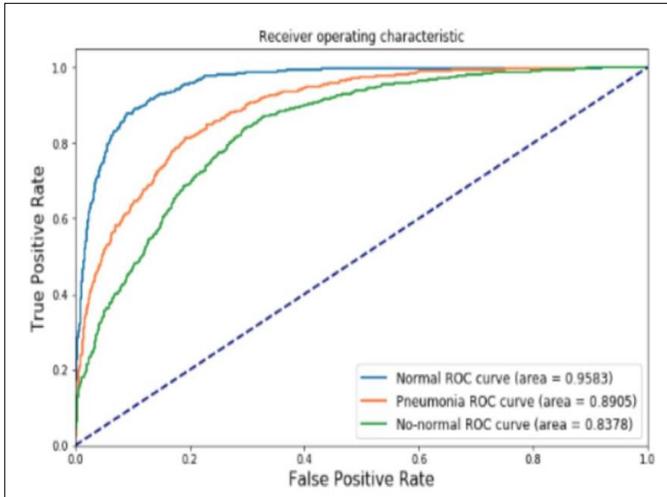


Figure 7: ROC Curve for Each Classes

The ROC curve is a graphical graph that shows the performance of any binary classifier system as a discrimination threshold. The discrimination threshold will give the result as a probability, it will be class 0 or 1. The activity diagram is a specific description of the use case processes, and the flowchart is a graphical diagram that represents a sequence of steps to solve a problem. The system Working Process is presented below:

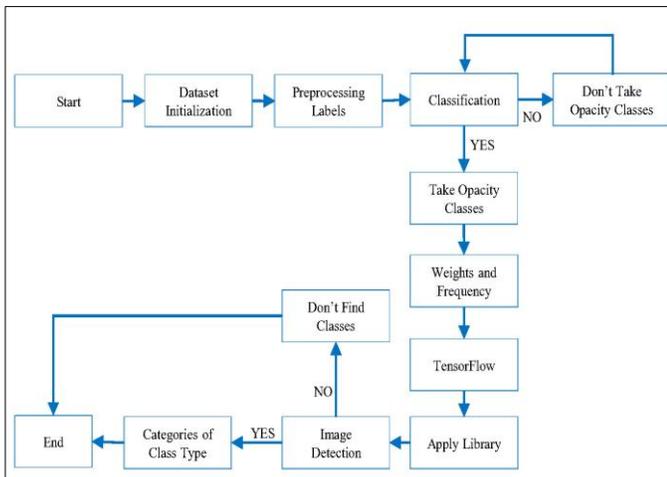


Figure 8: Working Process

### VIII. RESULT AND DISCUSSION

Usually, the lungs are full of air. When someone has lung opacity, the air in the lungs is replaced by another material - fluids, bacteria, cells of the immune system, etc. This is why areas of opacity are areas that are gray. When we see them, we understand that lung tissue in this area is probably not healthy.

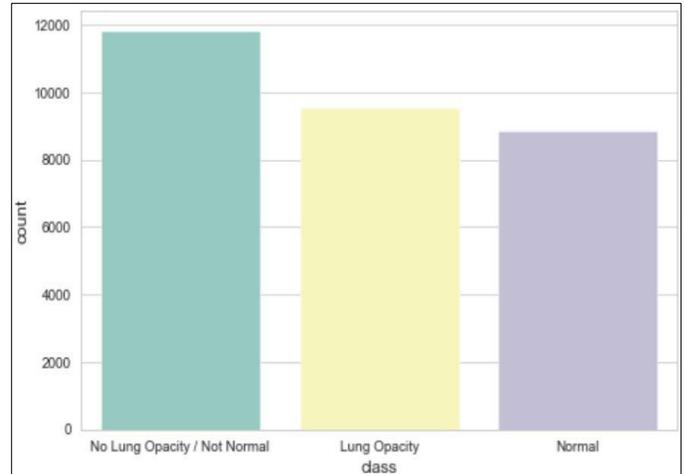


Figure 9: Label of Opacity Classes

Each line in the CSV file contains the patient's identifier, target 0 or 1 for the absence or presence of lung obstruction, as well as restrictions on certain x, y coordinates and width and height. I select several images with Target = 1. I would like to represent images with superimposed rectangles. To do this, I first need to analyze the entire data set with Target = 1 and collect all the coordinates of the windows showing the opacity of the lungs in the same image.

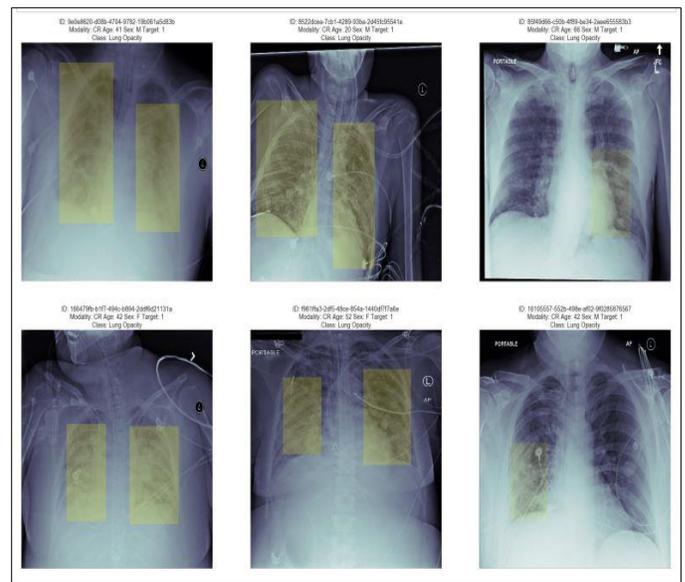


Figure 10: Output Images

This success is supported by a number of factors:

1. Increased availability of digitized data.
2. Increased computing power, increasingly powerful GPUs and enhanced access to computing power through IaaS services.
3. The rapid development of algorithms and methods that significantly improved the originally proposed technology.



## IX. FUTURE WORKS

My future aimed at increasing efficiency by improving its characteristics. Also, my idea about the Improvement of this project:

1. Use the same identifier for x-rays of the same patient. Currently, due to the privacy policy, each sample is identified with a unique identifier, even if they are received from the same patient.

2. For recording, each patient has at least one frontal x-ray and one lateral x-ray. Currently, this data set has only front-facing images, losing very valuable information and with the help of which professionals usually have to make their diagnoses.

## X. CONCLUSION

In this article, my goal is to propose a deep learning approach for classifying lung opacity from chest x-rays using transfer learning. I used the recently created data set, RSNA Pulmonary Disease Detection Dataset, consisting of several medical samples presented in format and the necessary annotations for training and optimization of each of the controlled models that make up my implementation. I used three main modules. This are preprocessing module, a classification module, and an object detection module. Early diagnosis and treatment of lung opacity is critical to preventing complications, including death. When carrying out approximately two billion procedures per year, chest radiography is the most common imaging tool that is used in practice and is crucial for screening, diagnosis and treatment of various diseases, including lung opacity.

## XI. ACKNOWLEDGEMENT

I want to thank NUST "MISIS" staff for their support and commitment to my prosperity as a student. Also, I want to thank my thesis supervisor, Kalitin Denis, for giving me the chance to deal with such a vital point. His help, learning, and specialized skill have been enormously significant to this extend, as well as to my prosperity as an engineer.

## XII. REFERENCES

- [1] F. Chollet. (2017). "Xception: Deep learning with depthwise separable convolutions". doi: 10.1109/CVPR.2017.195.
- [2] Y. Tan, L. H. Schwartz, and B. Zhao. (2013). "Segmentation of lung lesions on CT scans using watershed, active contours, and Markov random field". doi: 10.1118/1.4793409.
- [3] S. S. Sun, H. Li, X. R. Hou, Y. Kang, and H. Zhao. (2007). "Automatic segmentation of pulmonary nodules in CT images". doi: 10.1109/ICBBE.2007.206.
- [4] K. He, X. Zhang, S. Ren, and J. Sun. (2016). "Deep residual learning for image recognition". doi: 10.1109/CVPR.2016.90.
- [5] Y. Kim. (2014). "Convolutional neural networks for sentence classification". doi: 10.3115/v1/d14-1181.

- [6] J. Davis and M. Goadrich. (2006). "The relationship between precision-recall and ROC curves". doi: 10.1145/1143844.1143874.
- [7] J. Read, B. Pfahringer, G. Holmes, and E. Frank. (2011). "Classifier chains for multi-label classification". doi: 10.1007/s10994-011-5256-5.
- [8] T. Ojala, M. Pietikäinen, and T. Mäenpää. (2000). "Gray scale and rotation invariant texture classification with local binary patterns". doi: 10.1007/3-540-45054-8\_27.
- [9] H. Xu, X. Su, Y. Wang, H. Cai, K. Cui, and X. Chen. (2019). "Automatic Bridge Crack Detection Using a Convolutional Neural Network". doi: 10.3390/app9142867.
- [10] R. A. Pooley. (2005). "AAPM/RSNA physics tutorial for residents: fundamental physics of MR imaging". doi: 10.1148/rg.254055027.
- [11] Z. Q. Zhao, P. Zheng, S. T. Xu, and X. Wu. (2019). "Object Detection with Deep Learning: A Review". doi: 10.1109/TNNLS.2018.2876865.
- [12] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan. (2010). "Object detection with discriminatively trained part-based models". doi: 10.1109/TPAMI.2009.167.
- [13] C. P. Papageorgiou, M. Oren, and T. Poggio. (1998). "General framework for object detection". doi: 10.1109/iccv.1998.710772.