Classification and predictions of Lung Diseases from Chest X-rays using MobileNet

Jeevitha Sivasamy[#1], T.S. Subashini[*2]
[#]*Department of Computer and Information Science,* [*] *Department of Computer Science and Engineering,Annamalai University*
[1]e-mail: jeevithamca06@gmail.com
[2]e-mail: rtramsuba@gmail.com

Abstract— Diagnosis of lung diseases cannot be analyzed only through the symptoms, but at the first level, CXR (Chest X-Rays) play in an important role on identifying the disease affected status. Various modalities such as CT, MRI, PET, etc., are there to analyze the lung diseases, but prominently the Chest X-rays acts as a first step of finding chest diseases due to the fact which is less expensive. The more advanced and effective Deep learning techniques are now emerging, is a boon of medical fields and for detecting the diseases more efficaciously. In this work, we proposed a Classification and prediction of Lung pathologies from Chest X-rays using use case MobileNet model (UCMobN). Through MobileNetmodel, multiple use cases can be designed. In this, the use case model is employed and datasets with the best accuracy of 86 %. In this, the use case model is employed and tested using NIH chest X-rays image datasets with the best accuracy of 86 %. We obtained an Area Under Curve (AUC) of 0.58 for Atelectasis, AUC of 0.67 for consolidation, 0.74 for Edema, AUC of 0.54for Effusion, AUC of 0.54 for Emphysema, AUC of 0.65 for Fibrosis, AUC of 0.57 for Infiltration and AUC of 0.51 for Mass.

Keywords— **MobileNet; classification; predictions; lung diseases; depthwise separable layers; Chest X-rays;**

I. INTRODUCTION

The various kind of lung diseases has affected many people worldwide. Pulmonary disease is that causes the lungs to become more vulnerable to certain physical problems and air pollution. Thus, causing impairment in lung function. Some lung diseases, similar to emphysema, asthma, pleural effusion, TB, others including aspiratory fibrosis, pneumonia, and lung malignant growth, are brought about by loss versatility in the lungs that creates a lessening in the total volume of air [16]. Lung problems are easy to spread so it is important to find out the problems and give right treatment for patients. For detection and diagnosis of lung issues the radiologist mainly handles Chest x-ray images. Radiologists can diagnose with chest x-ray and can detect illnesses and many other conditions. Such diseases are bronchitis, infiltration, atelectasis, pericarditis, fractures and many other diseases [1]. CAD system helps radiologists find extraordinary diagnosis with computer assistant. It has two types of algorithms CADx algorithms for distinguishing abnormal lesion into benign and malignant, CADe is that detects abnormal lesion. Finding and classifying the properties of lung diseases in chest x -ray images are a very complicated task because of complex images of lung diseases and morphologies of its diseases [2]. However, it is difficult to examine the conditions and illness of lung diseases in the chest X-ray images by the CAD and it could not achieve any significant level [3-5]. There are many numbers of work employed for the detection and diagnosis of lung diseases using AI techniques. such as probabilistic neural network, multilayer perceptron, learning vector quantization and Generalized Regression Neural Networks (RNN) have been used for diagnosis lung diseases. The detection of lung diseases such as TB, COPD, Pneumonia, are implemented in neural networks for diagnosis cancer [6].

Even if the above-mentioned works are effectively used inthe classification of diseases; their computation time, minimum square error, accuracy, performance are not as efficient [7,8] In spirit of these deep learning-based models have been employed to increase the accuracy of image recognition, detection, etc. These networks are used by the researchers in medical image diseases to accurately examine classification and to do many types of tasks in many fields and efficiently express many useful features [9-12].

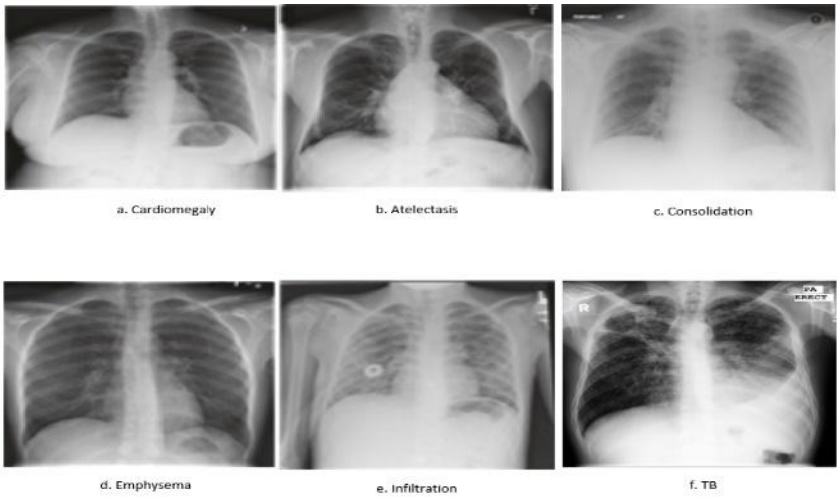


Fig 1. Lung diseases

Among these deep learning models CNN is the most commonly used methods namely ImageNet, GoogleNet, VGGnet, Inception, Xception etc., However, to get higher accuracy the common design offers becomes create deeper and complicated networks. With respect to the size and speed, we cannot greatly enhance the accuracy of networks [13-15,8]. Alongside merely the size and speed we cannot improve the precision of the system. The MobileNet model architecture is designed to create small, low-latency models for mobile and computer vision applications.

TABLE I. MobileNet Architecture

Type / Stride	Filter Shape	Input Size
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
Conv dw / s1	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
Conv dw / s2	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
Conv dw / s1	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
Conv dw / s2	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
Conv dw / s1	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
Conv dw / s2	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
$5\times$ Conv dw / s1	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
$5\times$ Conv / s1	$1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
Conv dw / s2	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$
Conv dw / s2	$3 \times 3 \times 1024$ dw	$7 \times 7 \times 1024$
Conv / s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$
Avg Pool / s1	Pool 7×7	$7 \times 7 \times 1024$
FC / s1	1024×1000	$1 \times 1 \times 1024$
Softmax / s1	Classifier	$1 \times 1 \times 1000$

This MobileNet deep learning method is can be achieved expecting output by fine-tuning the models according to our work [26]. Many research works focus the small network but not consider the speed of the network. MobileNet foremost focus on optimizing for latency and yet additionally yield small networks can run efficiently. [18]. It supports the any input size whether small or large its offering better performance.

This article focuses on the application of MobileNet model to classify the lung diseases on chest X-rays as either belonging to a healthy and precisely what kind of diseases affected in lungs. This work is organized into five sections: Section **II** deals with prior work and Section **III** deals with methodology**,** Experimental results and discussion is given in Section **IV** and finally Section **V** concludes the paper.

II. RELATED WORK

In [2] Various lung abnormalities are detecting and classifying such as lung nodules and lung diffuse by use of CNN for feature extractions and R-CNN for detecting lung diseases. Author yaniv Bar et al examine the power of deep learning approaches in chest radiograph data for pathologies detection. The datasets were trained by using the ImageNet model and mainly descriptors are DeCAF and PiCoDes (Picture Codes) are extracted using the implementation of Convolution Neural Network [19].

The depthwise separable convolution and ImageNet can have achieved slightly better performance at a lower computational cost for the image classification. MobileNet is built primarily from the depthwise separable convolution introduced in the paper [20]. MobileNet SSD [21] had been employed as an effective network in modern object recognition systems on the recent work. MobileNet feature extractors with Inception V1 and SSD are most precise of the faster network models. Alvin Rajkomar et al tested GoogLeNet along by image augmentation and pre-training on ImageNet to classifying CXR images either as lateral or frontal with good accuracy [22]. Recently explored using deep feature fusion method for classifying and detecting the lung nodules by CNN network models and it has given more hopeful results. The good performance was achieved using CNN methods with GIST features to classifying different pathologies in Chest X-rays images by Y. bar et al 2015 [23]. DCNNs [24] (Deep Convolutional Neural Network Work), AlexNet and GoogLeNet are used to evaluate and manifestations of detecting and classification of pulmonary tuberculosis on chest radiographs in [25].

III. METHODOLOGY

In this chapter, we propose to achieve the classification and prediction of Lung diseases in Chest X-rays using by Keras framework. This proposed work we have identified and classified 40000 chest X-rays images from the NIH dataset. The MobileNet model, and additionally CNN layers are employed in this work to predict and classifying the chest thoracic diseases in Chest X-ray images.

MobileNet is an associate design that is additionally appropriate for mobile and embedded based vision applications it takes responsibility for lack of computing power. Nowadays deep learning methods has been used not only assist for computer vision based but also includes various applications such as Robotics, Internet Of Things (IOT), Natural Language Processing (NLP), Medical image processing applications fields. The proposed work is depicted in figure 2.

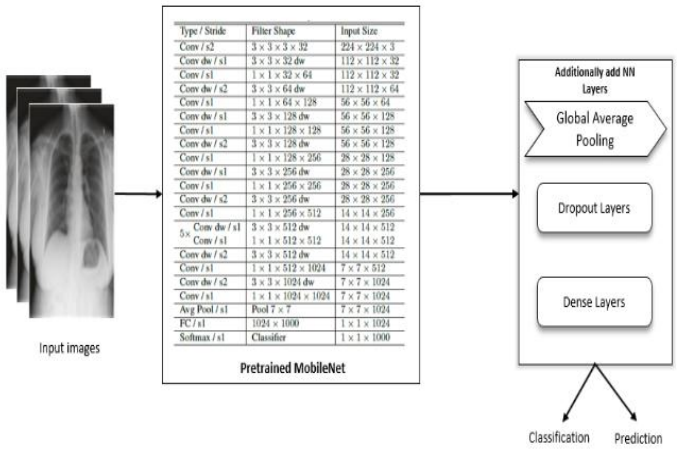


Fig 2. Block diagram of Proposed work architecture.

This architecture has depthwise separable convolution that considerably reduces the number of parameters and it leads to the lightweight neural network and it differs from the normal convolution. The normal convolution is replaced by depthwise convolution has a single filter is followed by pointwise convolution that is termed as depthwise severable convolution [26].

A. Depthwise Separable Convolution

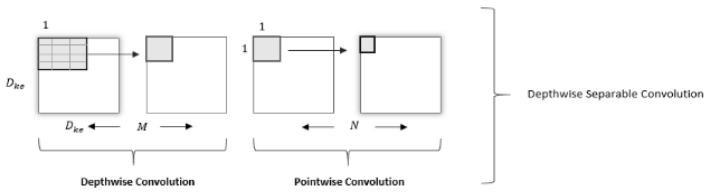


Figure 3. Depthwise separable Convolution layer in MobileNet

The MobileNet has two steps are: Depthwise and Pointwise convolutions, these are performing after finished down sampling in each feature maps separately. First Depthwise convolution which is performing the filtering

stage. Second Pointwise convolution which is perform the combining stage. Depthwise convolution applies to single convolution at a time in number of input channels. Pointwise convolution is applied to create a linear output combination of the depthwise layer and it has 1X1 filter convolution in the MobileNet. The Main difference between the MobileNet architecture and a traditional Convolution Neural Networks is in place of single 3X3 convolution followed by ReLU and Batch Normalization. MobileNet split up the convolution into a 3X3 depthwise convolution kernel and a 1X1 pointwise convolution in each input channels separately.

The standard convolution working in two steps one is filtering and another one is combining inputs into new outputs set. This factorization reducing the computation and size of the model. The depthwise separable convolution facilitates the separation of filtration and combining steps and further more uses to reduce the model size and complexity. These are the main factors that contribute to a significant reduction in computational cost. A usual convolution layer takes input as

$$D_{fm} \times D_{fm} \times M$$

Where as fm is Feature map and fm produces $D_{fm} \times D_{fm} \times N$ N is number of output channel separately. D_{fm} is the spatial height and width of each input feature map. M is the number of input channel depthwise. D_{ofm} is spatial width and height of the output feature map and N is number of output channel (depth). The standard convolution is parameterized by convolution kernel $D_{ke} \times D_{ke} \times M \times N$ where is D_{ke} is the spatial dimension of the kernel. Computational Cost for the standard convolution is:

$$D_{ke} . D_{ke} . M . N . D_{ke} . D_{ke} . M$$

The MobileNet can use various width factors and input layers size and it reduce the inference cost of on use case devices. This model specifies the interaction of the output channel kernel size $D_{ke} \times D_{ke}$ and their feature map $D_{fm} \times D_{fm}$ and also create the depthwise convolution to break the interaction between the output channels and kernel size. In our proposed customized model two steps there are filters and combination are split to use by the factorized convolution (depthwise separable convolution) for the substantial decrease of computational cost.

The depthwise separable convolution initially perform the just beginning step in a depthwise spatial convolution. Depthwise separable convolution shall write as (one kernel per input channels).

$$\hat{O}_{k,l,m} = \sum_{i,j} \widehat{ke}_{i\,j\,,m} . fm_{ke+i-1,l+j-1,m}$$

ke is filtered kernel size of $D_{ke} . D_{ke} . M$, where the m_{th} filter $\hat{O}$ filtered output feature map. The computational cost of depthwise convolution $D_{ke} . D_{ke} . M . D_{fm} . D_{fm}$. Depthwise convolution is not used to create a new feature, an additional layer is therefore required so that the linear combination of the output of the depth change can be computed via a 1×1 convolution in order to generate new features of chest X-rays images. The combination of pointwise 1×1 convolution and depthwise convolution used to reduce computational cost is called the depthwise separable convolution. It was introduced in [20][26]. By using the depthwise separable convolution give the accuracy and minimum complexity deep neural network.

B. Preparation for training data:

In this work, MobileNet deep neural network used with the supervised learning algorithm. We have randomly divided the entire dataset into three group, i.e. training (80%), validation and testing (20%) and made a single vector (disease_vec) with the 0/1 outputs for the disease status.
In the preprocessing steps the datasets are loading and transforming images randomly by using TensorFlow and Keras packages. The images batch size is 32 for the training.

- Proposed model UCMobN were trained in Tensorflow and Keras.
- First, we given inputs (size 224x224) to the MobileNet as a base additionally Global Average Pooling(GAP) layer 2D(2X2) filter is performed for spatial data and reduce the dimension of data, Dropout layers are randomly select neurons, and finally Dense layer is FC (Fully Connected) layer for all neurons connected to the next output nodes.
- Next the learnable parameters such as activation function (ReLU) and Sigmoid) is applied for obtained feature map and loss function, optimization algorithms are deployed to classify images.
- After the training the model has been created as shown below table 2.

TABLE 2. Output of the Proposed Use case trained model architecture

Layer (type)	Output Shape	Param #
mobilenet_1.00_128 (Model)	(None, 4, 4, 1024)	3228288
global_average_pooling2d_3 (	(None, 1024)	0
dropout_5 (Dropout)	(None, 1024)	0
dense_5 (Dense)	(None, 512)	524800
dropout_6 (Dropout)	(None, 512)	0
dense_6 (Dense)	(None, 10)	5130

Total params: 3,758,218
Trainable params: 3,736,330
Non-trainable params: 21,888

We proposed a model performing under different **hyper-parameters** namely filter size, number of kernels in all input channels, number weights, pooling. Activation function translates the input range and calculates the loss function with the help of forward propagation on the training images. Batch Normalization (BN) and ReLU are applied after each convolution channels and for preventing over fitting problems while epochs and data augmentation techniques are implemented (ImageDataGenerator). GAP layers act as a more depth type of reduce dimensionality.

$$H_{imp} * W_{imp} * D_{imp}$$

Where H_{imp} is height, W_{imp} weight and D is dimension. By handling all HW average values, GAP layers decrease by a single number $(1 \times 1 \times D)$ of each H*W feature map. 1024 neurons dropout with regularization rate of 0.5 during the training. Dense layer extracts the features from the convolution layer and down sampled by the GAP layer. In a Fully connected layer (Dense layer) every input node is connected to the output node. The Adam optimizer algorithm finds the individual learning rate in every parameter for calculating the loss by binary_crossentropy and MAE (Mean Absolute Error). We measure the performance of the classification and prediction probability value between 0 and 1.

IV. EXPERIMENTAL RESULTS

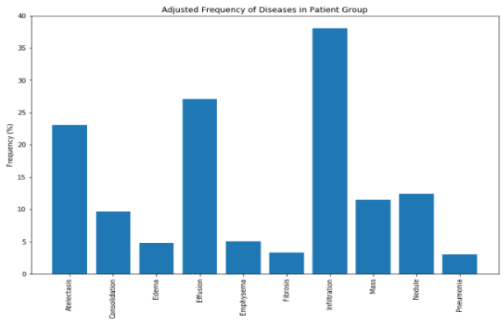
A. *Datasets*

We use the NIH Chest X-rays dataset which is released by Wang et al [28**].** It is comprised of 1,12,120 Chest X-rays image of unique patients. The Natural Language processing used for creating up to 14 different lung diseases was labeled for each data [27]. In this work, we focus on Recognition and Prediction of Chest X-rays pathologies using MobileNet use case model (UCMobN). We have used 10 different diseases labeled data after clean category process.

```
# keep at least 1000 cases
MIN_CASES = 1000
all_labels = [c_label for c_label in all_labels if all_xray_df[c_label].sum()>MIN_CASES]
print('Clean Labels ({})'.format(len(all_labels)),
      [(c_label,int(all_xray_df[c_label].sum())) for c_label in all_labels])
```

Clean Labels (10) [('Atelectasis', 11559), ('Consolidation', 4667), ('Edema', 2303), ('Effusion', 13317), ('Emphysema', 2516), ('Fibrosis', 1686), ('Infiltration', 19894), ('Mass', 5782), ('Nodule', 6331), ('Pneumonia', 1431)]

(a)



(b)Fig 4**. (a)** Cleaned labels disease categories and **(b)** frequency chart

Above images, as shown the 10 different categories are spilt into training sets and validation sets are Prepare for Training. Then create the data generates for transforming images randomly is called data augmentation. Totally we have taken 40000 images among them 30000 for training and 10000 for Testing and validation part, with Input Image size is 224×224. After the training process, the model is created and parameters are learned which is shown in table 2. In the first-round, epochs for training model is easy to obtain results which is shown in fig 5. We get the binary accuracy as 0.8359 in the first round of training. After continuing the training process, we achieved 0.8614 which is 86 % accuracy in 5th epoch. we used 5 epoch and 100 steps per epoch

Fig 5. Binary accuracy in first round of training

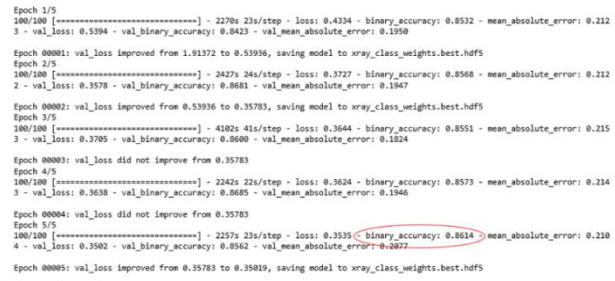


Fig 6.Performance of the proposed work

The proposed work was able to predict how much percentage of diseases affected in given data. ROC curve in Fig 7. is plotting for the false positive rate against true positive rate of predicting images also, We obtained an AUC for Atelectasis (AUC 0.58), consolidation (AUC 0.67), Edema (AUC 0.74), Effusion (AUC 0.65), Emphysema (AUC 0.54), Fibrosis (AUC 0.65), Infiltration (AUC 0.57), Mass (AUC 0.51), Nodule (AUC 0.58), and Pneumonia (AUC 0.62) for predicts certain diagnosis.

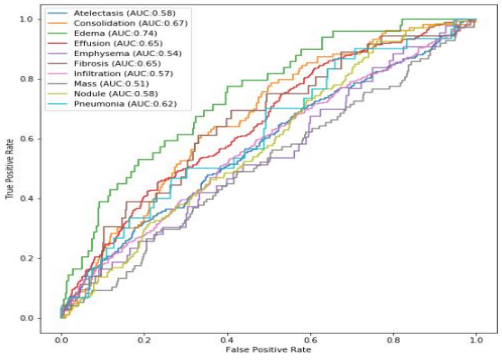


Fig 7. ROC plotting for pathologies predictions

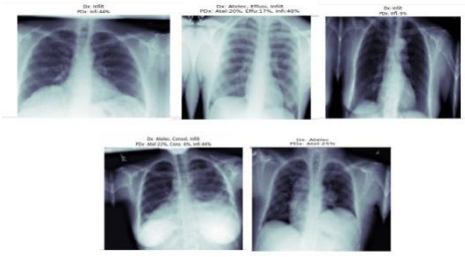


Fig 8. Prediction of lung pathologies

In the above depicted fig. 8 image, some examples are shown as associated prediction output images. Through these images, we can see how much percentage of lung disease are affected or spreaded.

V. CONCLUSIONS

In this work, the efficacy of proposed method to classify and predict the Chest X-rays was experimented. The areas of pathological irregularities enable the model to focus adapting on the pathologically abnormal regions. Hyper-parameters and learnable parameters are updated according to the loss value, through back propagation with Adam optimization algorithm. The performance of the model is very good, where the training accuracy is 86.14 % and the validation accuracy is 85.62 %. The results reveal that, deep networks could be successfully employed for lung diseases classification and prediction with good accuracy.

VI. FUTURE WORK

We have successfully recognized and predicted 10 thoracic diseases (Atelectasis, Effusion, Infiltration, Edema, Consolidation, fibrosis, emphysema, Mass, Nodule and Pneumothorax) using chest X-rays dataset provided by NIH.

In let near future we may collect Chest X-ray images from any local hospitals to predict the promised outcome and to achieve+_-_ better results by using another new CNN and deep neural network models or use case model with best optimization algorithm.

REFERENCES:

[1]. Rahib H. Abiyev and Mohammad KhaleelSallamMa'aitah, "Deep Convolutional Neural Networks for Chest Diseases Detection", Hindawi Journal of Healthcare Engineering, Volume 2018, pp. 1-10, 2018. https://doi.org/10.1155/2018/4168538.

[2] Shoji Kido, Yasusi Hirano, Noriaki Hashimoto, "Detection and Classification of Lung Abnormalities by Use of Convolutional Neural Network (CNN) and Regions with CNN Features (R-CNN)", IEEE, 2018.

[3] A. A. El-Solh, C.-B. Hsiao, S. Goodnough, J. Serghani, and B. J. B. Grant, "Predicting active pulmonary tuberculosis usingan artificial neural network," Chest, vol. 116, no. 4, pp. 968–973, 1999.

[4] K. Ashizawa, T. Ishida, H. MacMahon, C. J. Vyborny,S.Katsuragawa, and K. Doi, "Artificial neural networks in chest radiography: application to the differential diagnosis of interstitial lung disease," Academic Radiology, vol. 11, no. 1,pp. 29–37, 2005.

[5] A.M. Dos Santos, B. B. Pereira, and J. M. de Seixas, "Neural networks: an application for predicting smear negative pulmonary tuberculosis," in Proceedings of the Statistics in the Health Sciences, Li`ege, Belgium, 2004.

[6] O. Er, N. Yumusak, and F. Temurtas, "Chest diseases diagnosis using artificial neural networks," Expert Systems with Applications, vol. 37, no. 12, pp. 7648–7655, 2010.

[7] G. Litjens, T. Kooi, E. B. Bejnordi et al., "A survey on deep learning in medical image analysis," Medical Image Analysis, vol. 42, pp. 60–88, 2017.

[8] S. Albarqouni, C. Baur, F. Achilles, V. Belagiannis, S. Demirci,and N. Navab, "Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images," IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1313–1321, 2016.

[9] G. E. Hinton, S. Osindero, and Y. W. Teh, "A fast learning algorithm for deep belief nets," Neural Computation, vol.18, no. 7, pp. 1527–1554, 2006.

[10] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," in Proceedings of Advances in Neural Information Processing Systems, pp. 153–160, Vancouver, BC, Canada, 2007.

[11] M. R. Avendi, A. Kheradvar, and H. Jafarkhani, "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI," Medical Image Analysis, vol. 30, pp. 108–119, 2016.

[12] H.-C. Shin, K. Roberts, L. Lu, D. Demner-Fushman, J. Yao, and R. M. Summers, "Learning to read chest X-rays: recurrent neural cascade model for automated image annotation," Cornel University library, 2016, https://arxiv.org/abs/1603.08486.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Proceedings of Advances in Neural Information Processing Systems, pp. 1097–1105, Lake Tahoe, Nevada, USA, 2012.

[14] J.Bouvrie, "Notes on Convolutional Neural Networks" , http://cogprints.org/5869/1/cnn_tutorial.pdf.

[15] A. Rios and R. Kavuluru, "Convolutional neural networks for bio-medical text classification: application in indexing bio-medical articles," in Proceedings of the 6th ACM Conference on Bio-informatics, Computational Biology and Health Informatics, pp. 258–267, ACM, Atlanta, GA, USA, 2015.

[16] https://www.niehs.nih.gov/health/topics/conditions/lung-diseases/index.cfm

[17] Dataset: https://www.kaggle.com/nih-chest-xrays/data

[18] J. Jin, A. Dundar, and E. Culurciello. Flattened convolutional neural networks for feedforward acceleration. arXiv preprintarXiv:1412.5474, 2014.

[19] Yaniv Bar, IditDiamant, Lior Wolf, Hayit Greenspan, "Deep learning with non-medical training used for chest pathology identification".

20] L. Sifre. Rigid-motion scattering for image classification. Ph.D thesis, Ph. D. thesis, 2014.

[21] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z.Wojna, Y. Song, S. Guadarrama, et al. "Speed/accuracy trade-offs for modern convolutional object Detectors" arXiv preprint arXiv:1611.10012, 2016.

[22] Changmiao Wang, Ahmed Elazab, Jianhuang Wu, Qingmao Hu, "Lung nodule classification using deep feature fusion in chest radiograph" Journal of Computerized Medical Imaging and Graphics, 2017, http://dx.doi.org/10.1016/j.compmedimag.2016.11.004.

[23] Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, and H. Greenspan, "Chest pathology detection using deep learning with non-medical training," in Proceedings of Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium, pp. 294–297, Brookly, NY, USA, 2015.

[24] Mark Cicero, Alexander Bilbily et al, "Training and Validating a Deep Convolutional Neural Network for Computer-Aided Detection and Classification of Abnormalities on Frontal Chest Radiographs", Vol 52, No.5, pp.281-287, Journal of Investigative Radiology, 2017.

[25] Paras Lakhani, BaskaranSundaram, "Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks", Thoracic Imaging, pp.1-9, 2017.

[26] Andrew G. Howard Menglong Zhu Bo Chen Dmitry Kalenichenko, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications.

[27] Pranav Rajpurkar, JeremyIrvin, Kaylie Zhu, "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning" 2017.

[28] Xiaosong Wang1, Yifan Peng 2, Le Lu et al, "ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases", Provided by Computer Vision Foundation (CVF), pp.2097-2106, 2017.