Diagnosis of Chest Diseases in X-Ray images using Deep Convolutional Neural Network

[1]Arjun Choudhary [2]Abhishek Hazra, [3]Prakash Choudhary,
[1]Bhagwant University, Ajmer
[1,2,3]Department of Computer Science and Engineering
[2]Indian Institute of Technology(ISM), Dhanbad
[3]National Institute of Technology, Hamirpur
[1]arjun116a@gmail.com [2]abhishekhazra.18dr0018@cse.iitism.ac.in, [3]pc@nith.ac.in

Abstract— **Learning about Chest Diseases and their characterization is one of the most interesting research topics in recent years. With the various uses of medical images in hospitals, pathologies, and diagnostic centers, the size of the medical image datasets is also expanding expeditiously to capture the diseases in hospitals. Though a lot of researches have been done on this particular topic still this field is confusing and challenging. There are lots of techniques exist in literature to classify the medical images. The main drawback of traditional methods is the semantic gap that exists between the low-level visual information captured by imaging devices and high-level semantic information perceived by a human being. The difficulty of querying and managing the large datasets leads to a new mechanism called deep convolutional neural network. Deep learning techniques have recently achieved an impressive result in the field of computer vision along with Medical Engineering. In this paper, we proposed and evaluated a deep convolutional neural network, designed for classifying the Chest Diseases. The proposed model consists of Convolutional layers, ReLU Activations, Pooling layer, and Fully connected layer. Last full connected layer which is consists of fifteen output units. Each output unit will predict the probability of one of the fifteen diseases. A publicly available dataset called Chest X-Ray 14 which consists of fifteen classes named Atelectasis, Cardiomegaly, Effusion, Infiltration, Mass, Nodule, Pneumonia, Pneumothorax, Consolidation, Edema, Emphysema, Fibrosis, Pleural Thickening, Hernia and No Finding images used to train this model. Surprisingly this model gives a good result in multiclass classification. The average accuracy of 89.77% is achieved for the classification of different diseases. The comparative analysis shows the effectiveness of the proposed model. The proposed technique is best suited for classifying the multiclass medical images for different thorax diseases.**

Keywords— Convolutional Neural Networks (CNN), Deep Learning, Classification;

I. INTRODUCTION

According to the World Health Organization (WHO), "Chronic Diseases are the major Causes of death and disability in the world". India is the second largest country in the worldwide concerning population. More than 1.324 billion (2016) peoples from a different region and different community peoples are living together. But this country is facing a fact of the impact of different chronic diseases. According to the World Health Organization [4] in the next ten years, more than 16 million people will die due to critical diseases and every year it will increase by 18%. The different survey projected that death due to cancer is almost 8% and critical chronic diseases are almost 15%. One of the main reason for these diseases are overweight and uses of tobacco. Men and women both are affected in the twenty-first century. Statistics show that in 2005, 22% of men and 21% of women were suffering from overweight, but in 2015, 31% of men and 29% female are overweight. WHO says at least 8% heart diseases, diabetes, and 40% of cancer can be reduced by good diet and avoidance of tobacco. According to the Centers for Diseases Control and Prevention[5], top three causes of death in India are ischemic heart diseases 12%, chronic obstructive pulmonary diseases 11% and stroke 9% (approximately). These are the main inspiration and motivation to us for the focus of deep learning technology to medical science. In this paper, our main aim is to detect 14 different kinds of critical chest diseases by applying a Convolutional Neural Network.

In recent times, with the growth of computers advancement diagnostic and image acquisition devices, the medical image sets are also increasing exponentially. Clinical science and medical treatments are also benefited from the advance system of digital processing and content storing. Different hospitals and diagnostic centers are producing a large amount of data. Therefore deep analysis of the diseases is also a challenging issue. For analyzing these big data sets, several algorithms had been proposed. Deep learning based classification is one of the efficient techniques for analyzing medical images. Deep learning techniques are capable of achieving near to optimal solution in practical aspects due to its end-to-end learning capability. Survey [6, 10] shows that Deep Learning techniques are able to solve many real-life critical problems.

There are more than 200 types of thorax diseases. Which is often scaring like Fibrosis, Effusion, and Emphysema? These diseases reduce the blood circulation in the body and also reduce the breathing capability. Different studies show that more than 15 percent of the population suffer from these diseases. Although it is very difficult to recognize these diseases with normal symptom in preliminary stages. However, in some cases, it is a difficult task to recognize the images form low feature, smooth and low-resolution images. To resolve this problem we have used a new and large dataset called Chest X-Ray 14[18].

The rest of the paper organized as follows. Section 2 shows the overview of Deep learning and CNN. Section 3 describes the methodology with Proposed CNN architecture. Experimental Results are shown in section 4 followed by a Conclusion and feature work is explained in section 5.

II. RELATED WORK

Computer-aided design for medical images had been in use since the last decades. From the general approach to Machine Learning approach and Machine Learning approach to the Deep Learning approach keeps the high acceleration because of its high performance. There are lots of techniques which have been used for classifying different diseases like Brain cancer, Skin diseases, chest diseases, and many other diseases. After increasing the number of datasets, Deep Convolutional Neural Network is one of the most frequently used approaches in not only for classifying the natural images but also for classifying different kinds of medical images.

A. *Deep Learning*

In Deep Learning, Learning is the process of improving behavior based on experience. Deep learning is a subfield of Artificial intelligence which is used to improve the performance of many Machine learning applications. Deep architecture with multiple processing units, having linear or nonlinear transformation function [10] attempt to model high-level abstraction present on data. This technique started in 1965, but it affects in recent times with the acceleration of GPU based computing power and non-linearity which allow deeper networks for better utilization [17]. Deep architecture with many hidden layers increases the performance of the artificial neural network. The early invention of backpropagation in the 1980s to today, it is still in use to retain the neural network [7]. Standard backpropagation algorithms are commonly used in image classification tasks. Recent Studies show that deep learning techniques are the most successful in medical engineering.

Classifying the medical images is one of the challenging and interesting tasks in recent times. Deep Learning techniques also used to classify the Interstitial Lung Diseases (ILDs) in [12]. They work on seven diseases between them six are infected tissue. In [2] restricted Boltzmann machine was used to analyze the Lung generated tomography (CT) diseases. They have introduced two different techniques for two datasets, Texture classification, and airway detection. In [15] multiple kernels based CNN network was used to classify the brain organs. In. [20]. Double instance framework was used, in the first phase, CNN is used for extraction and in the second phase extracted features are used for classification of images. They used a database using 12 classes of images having CT and MR images.

B. *Convolutional Neural Network*

The convolutional neural network is a feed-forward neural network mainly inspired by the human brain. CNN is capable of doing both special feature extraction and classification. CNN combines four types of layers: Convolutional layer, a Nonlinear Activation layer, subsampling layer, and fully connected layer.

A typical convolutional layer takes input and produces feature maps with the help of the kernel. Each neuron takes some input from the previous layer. Different neuron connection and overlapping with each other which increase the performance of better representation of images. Sharing of weights reduce the number of parameters. Each convolutional layer connected to a nonlinear activation layer which gives CNN more acceleration for understanding even more complex functions. Sometimes there is a chance of CNN overfitting. Pooling is a subsampling technique which often used to overcome overfitting and also reduces the number of parameters. Finally one or more numbers of fully connected layers are added to summarize the trained features that give the classified result. In this paper, we propose a deep end to end learning for classifying the chest diseases. This paper also shows the extensive capability of CNN.

Different authors follow the different architecture of CNN. In this paper, we have followed a basic and simple architecture which was described in LeNet-5[19]. By modifying some layers, kernels, cost function, optimizer and hyperparameters we have trained our model. In the proposed model one convolutional layer comprises several kernels which compute the different featured maps followed by an activation layer and pooling layer. The main reason for using the convolutional layer is to learn more complex functions. Theoretically number of convolutional layers increase the system capability. But in our model, we use only three convolutional layers. The last fully connected layer contains the encoded features of the input image. Let (p, q)be a location in the input image X^l_{pq} and i is the number of feature maps of the l^{th} layer. The feature map Z^l_{pqi} can be calculated as:

$$Z^l_{pqi} = (W_i^{l^T} X^L_{pq} + S^l_i) \qquad (1)$$

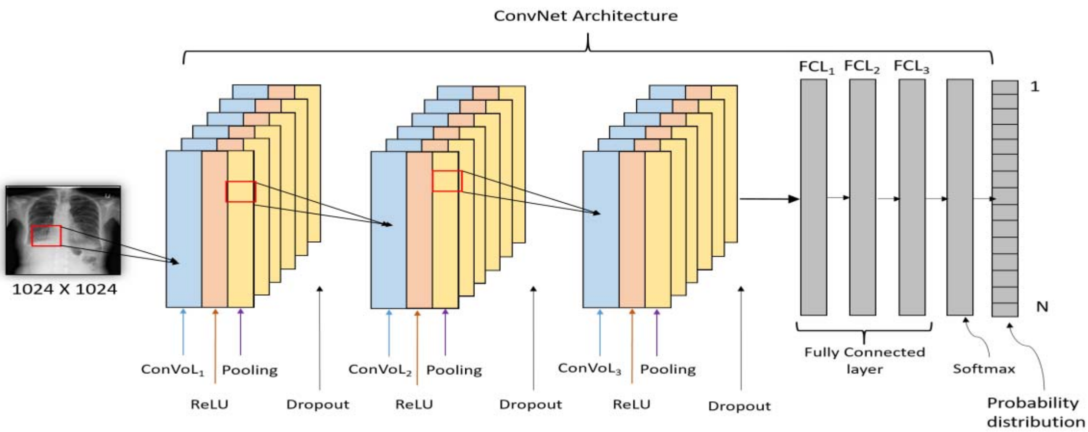

Figure 1. Proposed Architecture of convolutional neural network

Where W_i^l be the weight vectors and S_k^l be the bias of each filter in the l^{th} layer. The sharing of weights in every layer gives more power to the neural network for training. The nonlinear activation function denoted by λ (.) is used to understand the complex functions. The activation layer takes an input feature generated by convolutional layer, processes it and pushes into the pooling layer. In this model, we have used the ReLU activation function. This activation function can be derived and can be calculated as follows:

$$\lambda_{pqi}^{l} = \lambda\ (\ Z_{pqi}^{l}\ ) \tag{2}$$

$$\lambda_{pqi}^{l} = \max\ (\ W_i^T\ X_{pq} +\ S_i\ , 0) \tag{3}$$

$$Z_{pqi}^{l} = \text{pool}\ (\ \lambda_{abi}^{l})\ \ \forall (a, b) \in R_{pq} \tag{4}$$

Where R_{pq} is the nearest neighbor of location (p, q) the last fully connected layer is feed-forwarded into a Softmax activation layer which gives the probability of the output for each and every class corresponding to the input. The outcome of the Softmax activation can be calculated as:

$$s_{pqi}^{l} = \frac{e^{z^{[l]}}}{\sum_{n=1}^{N} e^{z^{[l]}}} \tag{5}$$

Training a neural network is one of the challenging tasks. By training the hyperparameters and using tuning sets for different parameters. The cost of the neural network can be calculated as:

$$\delta = \frac{1}{N} \sum_{b=1}^{N} l(\theta;\ Z^{(b)}, D^{(b)}) \tag{6}$$

Where θ is the set of all the parameters used in the neural network. Z^bdenotes the target output label whereas on is the output of the CNN. In this model, we use Adam optimizer to train the neural network.

III. Methodology

Most of the traditional methods are either complex or suffer from low accuracy. Here, our main objective is to make an efficient architecture which can overcome these two problems. In this segment, we first introduce the dataset which is used for the implementation of this model and then proposed CNN model is discussed.

A. Dataset

Analysis of Chest X-Ray is one of the challenging tasks in medical science. There are thousands of datasets available for chest X-Ray, but all the datasets are limited up to a few thousands of images [1]. We have used a significantly large dataset for our proposed work which is 27 times bigger than the old dataset. The dataset [14, 18] have a total of 112,120 frontal Chest X-Ray14 images taken from fourteen different pathologies each with resolution 1024×1024. These fourteen X-Ray images are taken from 30805 unique patients. In this Paper, a comparison with the old version of the Chest R-ray14 datasets has also been introduced.

B. Proposed CNN Architecture

The proposed architecture consists of twelve layers. Three Convolutional layers, three activation layers, three pooling layers, then followed by three fully connected layers. The output of the fully connected layer is feed-forwarded into a Softmax function which gives the probability of each and every class for a given input. The output vector size is used in this case is 1×15 as 15 is the number of output classes. It can be observed from figure1 that convolutional layers are coded as ConVoLi and the suffix number denotes the ith layer. Similarly, fully connected layers are also coded in FCL1, FCL2, and FCL3. In the first convolutional layer, 16 kernels each with size 3×3. Before each convolutional layer, we add 1 bit for managing the dimension. The output of the ConVoL1 is pushed into a nonlinear activation function which gives the faster execution and helps the neural network to understand the complex functions. The activation function which is used in this case is the rectified linear unit (ReLU). The performance of ReLU activation function is faster than other activation functions. The output of the ConVoL1 which was produced by the activation layer is fed into ConVoL2. Here 32 kernels each with 3×3 and stride size with 1is used. Again the output of this ConVoL2 is passed through ReLU activation function. As the model is deep, there may be a chance of overfitting which may reduce the performance of the neural network. Pooling a subsampling technique which is widely used in many neural networks. In this case, we use max pooling which not only reduces the number of parameters but also tries to reduce the overfitting problem. The output of ConVoL2 is finally passed to another convolutional layer. The output of the dropout layer feeds into the ConVoL3 where the filter size is 64 with 3×3 kernel each. After the first, the second and third convolutional layer dropout regularization is used to prevent the computational cost. Whenever the length of the neural network is high and the number of parameters is also more, then sometimes the performance of the neural network goes down which also increases the chance of overfitting. Dropout is a cost affecting technique to address the overfitting problem. The simple idea behind this scheme is to drop the number of units from neuron while training. This technique surprisingly works very well if the network is large and the performance of execution is slow. All the weights of the ConVoL3 layer are connected to a fully connected layer. In this model, three fully connected layers are in use. Each fully connected layer computationally expensive compared to other layers. Finally, all the weights pass through a Softmax function which gives the probability of the output for each and every class. Figure 1 shows the proposed CNN architecture with different parameters.

IV. Experimental REsult

The entire implementation was done in windows10 OS using python as the programming language. The whole experiment was coded in Keras framework using Theano as the backend. The machine with Intel(R) Core(TM) i7-4790 CPU @3.60GHz 12GB RAM is used for implementing the model.

A. Training.

The research group of NIH. USA [18] did tremendous work and collected 112,120 frontal view Chest X-Ray images from 14 different pathologies for detecting various diseases. We split our data into the three different categories accordingly training, validation and testing set. Total 70% of the images are used for training, 10% of the images for validation and 20% of the images are used for testing these diseases. Before training, we resize our images into 224 ×224 and also normalize the images by the mean and standard deviation. Color shifting and augmentation are also used for acquiring better accuracy. Finally, Adam optimizer and cross entropy have been used for minimizing the loss. The loss the function can be calculated as:

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{k=1}^{K} y_k^{(i)} \log h_\theta\left(x^{(i)}\right)_k + \left(1 - y_k^{(i)}\right)\log\left(1 - h_\theta\left(x^{(i)}\right)_k\right)\right] \tag{7}$$

Where m is the number of training sample and K defines the number of output dimension. λ is a regularization Parameter, $Y_k^{(i)}$ is i[th] training output for the k[th] output node. $h_\theta(x^{(i)})_k$ denotes the value of the hypothesis at output k with weight and training input i.

B. Model Evaluation

In This Segment, the output of the proposed model has been examined. We also show how the tuning of the hyperparameter increases the performance of the neural network. The performance of our model with Adam optimizer [9] and categorical cross-entropy as the cost function gives the optimal output for detecting several diseases. Table 2 shows some comparison with different existing techniques. For the implementation of simplicity and improving performance, we have used 32 is the batch size and dropout percentage is 50% of all the times. A constant of the 3×3 kernel is used for every filter. To reduce the number of parameters max-pooling with kernel 2×2 is used. ReLU nonlinear activation which gives the understanding of more complex patterns and also gives acceleration to the neural network. The final output layer contains 15 different classes which are the number of diseases. A most popular well-known technique called early stopping is also used to reduce the overfitting problem and also gives the maximal output from this model. Table 1 shows some of the testing results with their corresponding desired output label versus actual output level.

Figure 2 shows the comparison of training accuracy and validation accuracy with different epochs. The performance of the model gives around 90% training accuracy and 89% testing accuracy. From this figure, it is also clear that this architecture reduces both training and validation loss with the increase in the number of epochs.

C. Comparison

To evaluate this model comparison between our models with some other recent works related to this field is done. Though it is not possible to make a direct comparison as there is no such standard database available for this particular task. With the best of our knowledge, some criteria are made to make a comparison in the relevance of accuracy. Table 2 shows a comparison of the proposed model with different architecture. From this table, it is obvious to say that our proposed model created a benchmark for classifying different chest disease.

Table 1. Images with different labels with predicted class labels

	
Actual Class: Atelectasis, Cardiomegaly, Emphysema, Mass, Pneumothorax	**Actual Class:** Effusion, Fibrosis, Mass, Nodule, Pleural Thickening
Predicted Class: Atelectasis, Effusion, Infiltration	**Predicted Class:** Mass, Nodule, Pleural Thickening
	
Actual Class: Atelectasis, Effusion, Infiltration, Pneumonia	**Actual Class:** No finding
Predicted Class: Infiltration, Pneumonia, Mass	**Predicted Class:** No finding

D. Classification Performance

The performance of this model is evaluated in the context of Precision, recall and f1 measure.

$$Average\ Precision = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TP_i + FP_i} \tag{8}$$

$$Average\ Recall = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TP_i + TN_i} \tag{9}$$

$$Accuracy = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \tag{10}$$

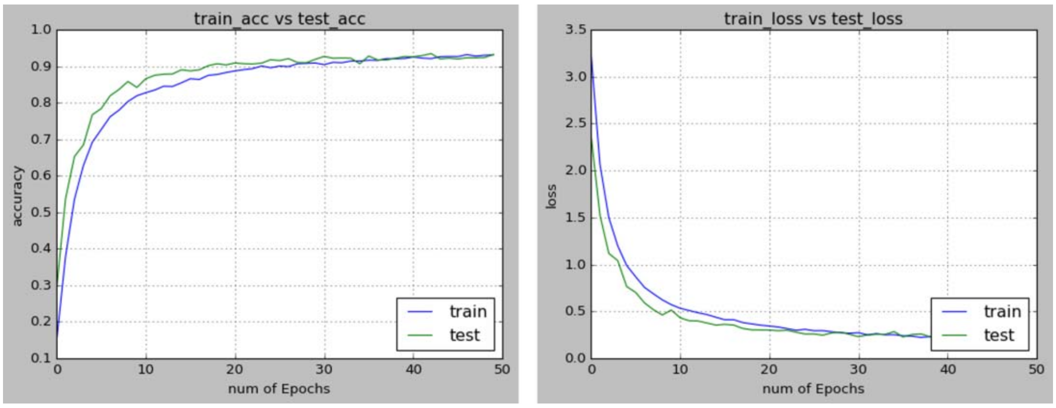

Figure 2. Output plot of Loss and Accuracy with the different epoch in training and validation

$$F1\ measures = 2 \times \frac{TP_i}{TP_i + TN_i} \quad (11)$$

Where TP denotes the number of images which are correctly classified. FP is the number of images which are misclassified to some other classes. TN is the number of images which are correctly classified as the image does not belong to this class. Similarly, FN is the number of images which belongs to some classes but misclassified to another class. So it is clear from the above equation that accuracy is the ratio between correctly classified observations to total sample observations. Precision gives the correctly predicted observation among the total correct observation. A recall is the identification of correct output among all the total correct output and F1 Score provides the weighted average output of recall and precision. Here N is the number of classes which is 15 in this case. K-fold cross-validation is used to test the training accuracy. The average Precision, recall and F1 measures for the proposed model is 91%, 89%, and 90% respectively. Table 2 shows that our proposed Model performs better than the previous state-of-the-art results.

Table2: Comparison of the proposed model with different models

Method	Techniques	Favg	Accuracy
Gangeh[3]	SVM-RBF	0.7127	0.7152
Sorensen[11]	KNN	0.7322	0.7333
Anthimopoulos[13]	RF	0.7786	0.7809
Li[16]	CNN	0.6657	0.6705
LeNet[19]	CNN	0.6783	0.6790
Alexnet[17]	CNN	0.7582	0.7609
Kingam[9]	CNN	0.7804	0.7800
Proposed Method	**CNN**	**0.9014**	**0.8977**

V. Conclusion

Diagnosis of chest diseases in X-Ray images is one of the critical research challenges from the past decades. Several machine learning, feature learning, and pattern analysis techniques are introduced to address this issue. But most of the techniques have failed to provide good results in practical aspects. In this paper, A deep end-to-end Convolutional Neural Network which improves the performance of computer vision in the field of Medical Engineering is proposed here. In this model, we use a non-parametric weighting scheme which shows the behavior of visual recognition as well as feature extraction technique in the layer by layer. Results obtaining from this scheme outperforms the state-of-the-art in the computer vision field. Here, we classify 15 different diseases including 14 infected and one healthy organ. Along with with this a novel structure for classifying the low-level features for these diseases is also proposed. The proposed architecture also takes care of the bias-variance problem. Moreover, the early stopping technique is used to reduce the loss as well as the overfitting problem. This model also has the capacity to offer the insight feature of the convolutional neural network. Learning the neural network for a specific task is a promising challenge for medical image classification and our proposed model gives a clear idea of learning the weights for this task. Our work is still in progress. In this recent feature, our aim is to predict the person's age from the skeleton.

Acknowledgment

The author would like to thank the research group of NIH U.S. National Library of Medicine [10] to provide the standard Chest X-Ray 14 database.

References

[1] Bounding boxes for ~1000 images (BBox_List_2017.csv): Image Index, Finding Label, Bbox[x, y, w, h]. [x y] are coordinates of each box's top left corner. [w h] represent the width and height of each box.

[2] G. van Tulder and M. de Bruijne, "Combining Generative and Discriminative Representation Learning for Lung CT Analysis With Convolutional Restricted Boltzmann Machines," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1262-1272, 2016.

[3] Gangeh M.J., Sørensen L., Shaker S.B., Kamel M.S., de Bruijne M., Loog M. A Texton-Based Approach for the Classification of Lung Parenchyma in CT Images Lecture Notes in Computer Science, vol 6363, (2010). Springer, Berlin, Heidelberg

[4] http://www.who.int/respiratory/en/

[5] https://www.cdc.gov/

[6] Jiuxiang Gu, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Gang Wang, Jianfei Cai, Tsuhan Chen, Recent advances in convolutional neural networks,

Pattern Recognition, Volume 77, 2018, Pages 354-377, ISSN 0031-3203, https://doi.org/10.1016/j.patcog.2017.10.013.,

[7] WAN, Ji; WANG, Dayong; HOI, Steven C. H.; WU, Pengcheng; ZHU, Jianke; ZHANG, Yongdong; and LI, Jintao. Deep learning for content-based image retrieval: A comprehensive study. (2014). MM '14: Proceedings of the 22nd ACM International Conference on Multimedia: November 3-7, 2014, Orlando. 157-166.

[8] K. Simonyan and A. Zisserman "Very Deep Convolutional Networks for Large-Scale Image Recognition", in International Conference on Learning Representations 2015

[9] Kingma, Diederik & Ba, Jimmy. "Adam: A Method for Stochastic Optimization". International Conference on Learning Representations. (2014)

[10] Li. Deng and D. Yu, "Deep Learning," Signal Processing, vol. 7, pp. 3-4, 2014.

[11] L. Sorensen, S. B. Shaker, and M. de Bruijne, "Quantitative Analysis of Pulmonary Emphysema Using Local Binary Patterns," in IEEE Transactions on Medical Imaging, vol. 29, no. 2, pp. 559-569, Feb. 2010.

[12] M. Anthopoulos, S. Christodoulidis, L. Ebner, A. Christe, and S. Mougiakakou, "Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network," in IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1207-1216, May 2016. doi: 10.1109/TMI.2016.253 5865

[13] M. Anthimopoulos, S. Christodoulidis, A. Christe and S. Mougiakakou, "Classification of interstitial lung disease patterns using local DCT features and random forest," 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, pp. 6040-6043, 2014

[14] Open-i: An open access biomedical search engine. https//openi.nlm.nih.gov.

[15] P. Moeskops et al., "Automatic segmentation of MR brain images with a convolutional neural network," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1252-1261, 2016.

[16] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng and M. Chen, "Medical image classification with the convolutional neural network," 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), Singapore, pp. 844-848. 2014

[17] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in Aistats, 2010, pp. 249-256.

[18] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, pp. 3462-3471. 2017

[19] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov 1998

[20] Z. Yan et al., "Multi-Instance Deep Learning: Discover Discriminative Local Anatomies for Bodypart Recognition," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1332-1343, 2016.