

**DEEP STRUCTURE BASED ON  
CONVOLUTIONAL NEURAL NETWORKS  
FOR IDENTIFICATION OF CHEST  
DISEASES**

**A THESIS SUBMITTED TO THE  
GRADUATE SCHOOL OF APPLIED  
SCIENCES  
OF  
NEAR EAST UNIVERSITY**

**By  
MOHAMMAD KHALEEL SALLAM  
MA'AITAH**

**In Partial Fulfillment of the Requirements for  
the Degree of Doctor of Philosophy  
in  
Computer Engineering**

**NICOSIA, 2018**

**MOHAMMAD K. S. MA'AITAH**

**DEEP STRUCTURE BASED ON CONVOLUTIONAL NEURAL**

**NEU**

**NETWORKS FOR DETECTION OF CHEST DISEASES**

**2018**



**DEEP STRUCTURE BASED ON  
CONVOLUTIONAL NEURAL NETWORKS FOR  
IDENTIFICATION OF CHEST DISEASES**

**A THESIS SUBMITTED TO THE  
GRADUATE SCHOOL OF APPLIED SCIENCES  
OF  
NEAR EAST UNIVERSITY**

**By  
MOHAMMAD KHALEEL SALLAM MA'AITAH**

**In Partial Fulfillment of the Requirements for  
the Degree of Doctor of Philosophy  
in  
Computer Engineering**

**NICOSIA, 2018**

**Mohammad Khaleel Sallam Ma'aitah: DEEP STRUCTURE BASED ON CONVOLUTIONAL NEURAL NETWORKS FOR IDENTIFICATION OF CHEST DISEASES**

**Approval of Director of Graduate School of Applied Sciences**

**Prof. Dr.Nadire CAVUS**

**We certify this thesis is satisfactory for the award of the degree of Doctor of Philosophy in Computer Engineering**

**Examining Committee in Charge:**

Prof. Dr. Rashad Aliyev

Department of Mathematics, EMU

Prof. Dr. Rahib Abiyev

Supervisor, Department of Computer Engineering, NEU

Assoc. Prof. Dr. Musbah Aqel

Department of Managment Information System, CIU

Assoc. Prof. Dr. Melike Sah Direkoglu

Department of Computer Engineering, NEU

Assoc. Prof. Dr. Kamil Dimililer

Department of Automotive Engineering, NEU

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name:

Signature:

Date:

## ACKNOWLEDGMENTS

This work would not have been possible without the support of the chairman of the Computer Engineering department and my supervisor Prof. Dr. Rahib Abiyev who have been supportive of my career goals and who worked actively to provide me with the protected academic time to pursue those goals. As my teacher and mentor, he has taught me more than I could ever give him credit for here. He has shown me, by his example, what a good scientist (and person) should be. I am grateful to all of those with whom I have had the pleasure to work during this and other related projects. Each of the members of my Dissertation Committee has provided me extensive personal and professional guidance and taught me a great deal about both scientific research and life in general. I would especially like to thank Assoc. Prof. Dr. Mustafa Menekay, Assist. Prof. Dr. Bengi Sonyel, Assist. Prof. Dr. Qais Almaaitah, and Dr. Abdulkader Helwan for their endless support. Nobody has been more important to me in the pursuit of this project than the members of my family. I would like to thank my parents, whose love and guidance are with me in whatever I pursue. They are the ultimate role models.

**To my parents...**

## ABSTRACT

Nowadays, artificial intelligence methods are being widely used for the identification of different diseases using their medical data. Misidentification of medical images can lead to fatal results and therefore the accuracy of the designed intelligent systems is very significant when creating a medical identification framework. Recently, different improvements are proposed for designing high accuracy models in the medical field. In this thesis, the design of deep learning structure is proposed for the identification of chest pathologies. The detection of chest diseases is highly required in the healthcare. The most common technique for identifying the chest diseases is the chest X-ray. False CXR interpretation, can result in sub-standard reports, misdiagnosis, confusion, and gaps in communication with primary care physicians. All of these severely negatively impact patient care, and can have life-changing consequences for patient. Chest X-ray can diagnose various diseases such as chronic obstructive pulmonary, pneumonia, asthma, tuberculosis, lung diseases. In this thesis, the identification of the chest pathologies in chest x-rays using deep learning approaches based on convolutional neural networks (CNN) is presented. The architecture of CNN and its design principle including learning algorithm are presented. The performance of the developed CNN is validated using chest dataset. Moreover, the performance of CNN in classifying chest diseases is compared with other machine learning techniques including backpropagation neural networks (BPNN) with supervised learning, competitive neural network (CpNN) with unsupervised learning. All the CNN, BPNN and CpNN models are trained and tested on the same chest X-ray database, and the performance of each network is discussed. The results of comparison in terms of accuracy, error rate, and training time of the employed networks are also presented.

**Keywords:** Artificial intelligence; deep learning; Convolutional neural networks; backpropagation neural networks; competitive neural network; chest pathologies; chest X-rays

## ÖZET

Günümüzde, tıbbi bilgi kullanılarak farklı hastalıkların tanımlanması için yapay zeka yöntemleri yaygın olarak kullanılmaktadır. Tıbbi görüntülerin yanlış tanımlanması ölümcül sonuçlara yol açabilir ve bu nedenle tasarlanan akıllı sistemlerin doğruluğu, bir tıbbi tanımlama çerçevesi oluştururken çok önemlidir. Son zamanlarda, tıp alanında yüksek doğrulukta modeller tasarlamak için farklı iyileştirmeler önerilmiştir. Bu tez çalışmasında, göğüs patolojilerinin tanımlanması için derin yassı yapı tasarımı önerilmiştir. Göğüs hastalıklarının saptanması sağlık hizmetlerinde çok gereklidir. Göğüs hastalıklarını tanımlamak için en yaygın teknik, göğüs röntgeni. Yanlış CXR yorumlaması, birinci basamak hekimiyle iletişimde alt standart raporlar, yanlış tanı, karışıklık ve boşluklarla sonuçlanabilir. Bunların hepsi hasta bakımını ciddi şekilde olumsuz etkiler ve hasta için hayatı değiştiren sonuçlar doğurabilir. Göğüs röntgeni kronik obstrüktif akciğer, pnömoni, astım, tüberküloz, akciğer hastalıkları gibi çeşitli hastalıkları teşhis edebilir. Bu tez çalışmasında, konvolüsyonel nöral ağlara (CNN) dayalı derin öğrenme yaklaşımları kullanılarak göğüs röntgenlerinde göğüs patolojilerinin tanımlanması sunulmuştur. CNN mimarisi ve öğrenme algoritması içeren tasarım prensibi sunulmuştur. Geliştirilmiş CNN'nin performansı, göğüs veri kümesi kullanılarak doğrulandı. Ayrıca, göğüs hastalıklarının sınıflandırılmasında CNN'nin performansı, denetimli öğrenme ile geri yayılma sinir ağları (BPNN), denetimsiz öğrenme ile rekabetçi sinir ağı (CpNN) dahil olmak üzere diğer makine öğrenme teknikleri ile karşılaştırılmıştır. Tüm CNN, BPNN ve CpNN modelleri aynı göğüs röntgeni veritabanında eğitilir ve test edilir ve her bir ağın performansı tartışılır. Karşılaştırma sonuçları, kullanılan ağların doğruluğu, hata oranı ve eğitim süresi açısından da sunulmuştur.

**Anahtar Kelimeler:** Yapay zeka; derin eğiklik; Konvolüsyon nöral ağlar; geri yayımlı; sinir ağları; rekabetçi sinir ağı; göğüs patolojileri; göğüs röntgeni

## TABLE OF CONTENT

<b>ACKNOWLEDGEMENT</b> .....	i
<b>ABSTRACT</b> .....	iii
<b>ÖZET</b> .....	iv
<b>TABLE OF CONTENT</b> .....	v
<b>LIST OF FIGURES</b> .....	viii
<b>LIST OF TABLES</b> .....	ix
<b>LIST OF ABBREVIATIONS</b> .....	x
<b>CHAPTER 1: INTRODUCTION</b> .....	1
1.1 Introduction .....	1
1.2 Significance of the Work.....	4
1.2 Thesis Overview.....	4
<b>CHAPTER 2: RADIOGRAPHY OVERVIEW</b> .....	5
2.1 Introduction .....	5
2.2 Chest X-rays .....	6
2.3 Chest Abnormalities .....	7
2.3.1 Pleural disease .....	7
2.3.2 Pneumothorax.....	8
2.3.3 Asbestos plaques .....	9
2.3.4 Pleural effusions .....	10
2.4 Features Extraction in Medicine .....	11
2.5 Image Processing.....	14
2.5.2 Image enhancement.....	14
<b>CHAPTER 3: LITERATURE REVIEW</b> .....	16
3.2 Review on Using Backpropagation Neural Networks in Medical Images Classification... 16	
3.3 Review on Using Unsupervised Learning in Medical Images Classification.....	17

3.4 Review on Using Deep Learning in Medical Images Classification ..... 19

    3.4.1 Review on using convolutional neural networks in medical images classification .. 19

**CHAPTER 4: DESIGN OF CNN BASED CHEST X-RAY PATHOLOGY IDENTIFICATION SYSTEM ..... 23**

    4.1 Overview ..... 23

    4.2 Deep Learning ..... 23

    4.3 Convolutional Neural Networks..... 24

    4.4 Understanding the Learning of Convolutional Neural Networks ..... 25

        4.4.1 Non linearity (ReLU) ..... 29

        4.4.2 Pooling layer ..... 29

        4.4.3 Fully connected layer (FC)..... 30

    4.5 Design of CNN Based Chest X-ray Identification System ..... 33

**CHAPTER 5: SIMULATION ..... 37**

    5.1 Overview ..... 37

    5.2 Simulations..... 37

        5.2.1 BPNN training..... 38

        5.2.2 Competitive neural network (CpNN)..... 40

        5.2.3 Convolutional neural networks ..... 42

    5.3 Results Discussion..... 42

**CHAPTER 6: CONCLUSION AND FUTURE WORKS..... 48**

    6.1 Conclusion..... 47

    6.2 Future Recommendations..... 47

**REFERENCES ..... 50**

**APPENDICES**

**Appendix 1: Backpropagation Neural Network Source Code ..... 58**

<b>Appendix 2:</b> Competitive Neural Network Source Code .....	61
<b>Appendix 3:</b> Convolutional Neural Network Source Code .....	63
<b>Appendix 4:</b> Learned Filters Visualizations .....	63
<b>Appendix 5:</b> Curriculum Vitae .....	70

## LIST OF FIGURES

<b>Figure 2.1:</b> Pleural thickening.....	8
<b>Figure 2.2:</b> Pneumothorax .....	9
<b>Figure 2.3:</b> Asbestos related pleural plaques .....	10
<b>Figure 2.4:</b> Pleural effusion .....	11
<b>Figure 2.5:</b> Medical image enhancement .....	15
<b>Figure 3.1:</b> Modified Alexnet proposed by Khan and Yong, (2017).....	21
<b>Figure 4.1:</b> Typical architecture of a convolutional neural network (CNN).....	25
<b>Figure 4.2:</b> Array of RGB Matrix.....	26
<b>Figure 4.3:</b> Example of a neural network with many convolutional layers.....	26
<b>Figure 4.4:</b> Image matrix multiplies kernel or filter matrix.....	27
<b>Figure 4.5:</b> Image matrix multiplies kernel or filter matrix.....	27
<b>Figure 4.6:</b> 3 x 3 Output matrix.....	28
<b>Figure 4.7:</b> Some common filters.....	28
<b>Figure 4.8:</b> ReLU operation.....	29
<b>Figure 4.9:</b> Max Pooling.....	30
<b>Figure 4.10:</b> After pooling layer, flattened as FC layer.....	30
<b>Figure 4.11:</b> Complete CNN architecture.....	31
<b>Figure 4.12:</b> The proposed Convolutional neural network for chest X-ray pathology identification.....	33
<b>Figure 4.13:</b> CNN final classification of chest X-rays with classes probabilities.....	34
<b>Figure 5.1:</b> Back propagation neural network.....	36
<b>Figure 5.2:</b> Learning curve for BPNN2.....	38
<b>Figure 5.3:</b> Competitive neural network.....	38

**Figure 5.4:** Learned filters: (a) Convolution layer 1, (b) Pooling layer 1.....43

## LIST OF TABLES

<b>Table 5.1:</b> Training parameters for backpropagation networks (32×32 input pixels).....	37
<b>Table 5.2:</b> Training parameters for competitive neural network (32×32 input pixels).....	39
<b>Table 5.3:</b> CNN training parameters.....	40
<b>Table 4.4:</b> Recognition rates for BPNNs on training and validation data (32×32 pixels)..	41
<b>Table 5.5:</b> Recognition rates for CpNNs on training and validation data (32×32 pixels)..	41
<b>Table 5.6:</b> Recognition rates for CNNs on training and validation data (32×32 pixels)...	42
<b>Table 5.7:</b> Performances of the BPNN, CpNN and CNN.....	42
<b>Table 5.8:</b> Results comparison with earlier works.....	44

## LIST OF ABBREVIATIONS

<b>ANN:</b>	Artificial Neural Network
<b>FFNN:</b>	Feedforward Neural Network
<b>NN:</b>	Neural Network
<b>CNN:</b>	Convolutional Neural Network
<b>CpNN:</b>	Competitive Neural Network
<b>BPNN:</b>	Back Propagation Neural Network
<b>MSE:</b>	Mean Square Error
<b>SEC:</b>	Second
<b>MIN:</b>	Minutes
<b>DCNN:</b>	Deep Convolutional Neural Network

# CHAPTER 1

## INTRODUCTION

### 1.1 Introduction

Chest radiography is still one the most economical and easy to use medical imaging technology. This technology allows the production of medical images of the chest, heart, lung, airways etc.. The interpretation of the chest X-ray images by trained radiologists can diagnose a big number of conditions and diseases such as pneumothorax, interstitial lung disease, heart failure, pneumonia, bone fracture, hiatal hernia and so on (Er et al., 2010).

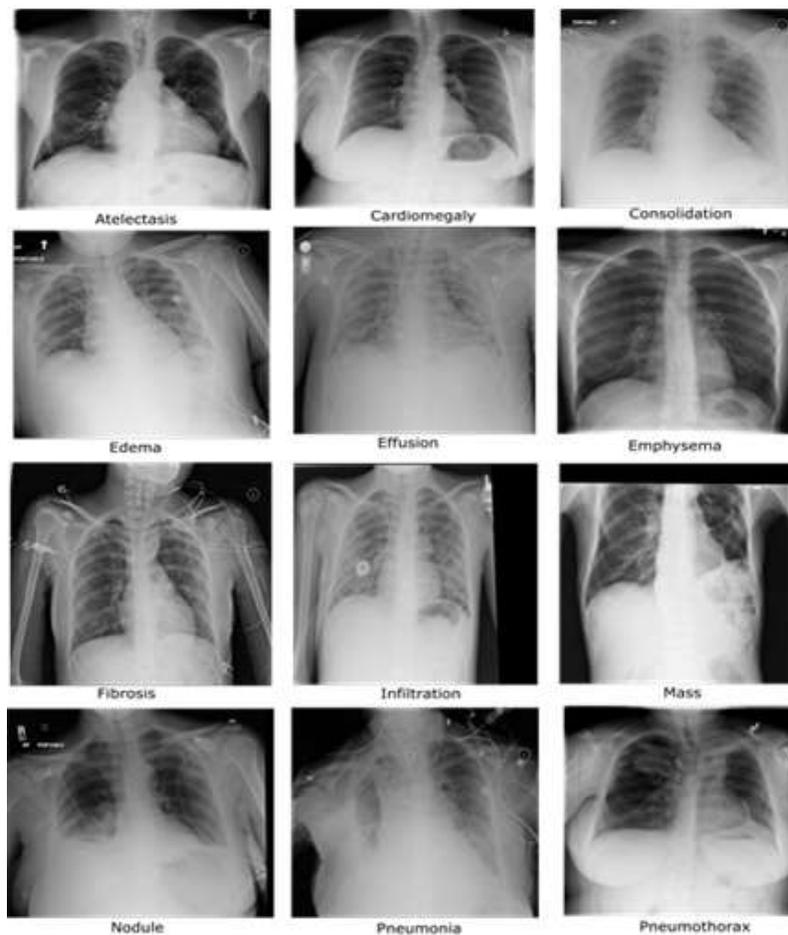
The identification of the chest X-ray abnormalities is still a tedious task for radiologists. Hence, the development of computer systems that helps radiologists in diagnosing the chest radiographs is in an immense need. Recently, deep neural networks have been extensively applied to solve various medical problems (El-Solh et al., 1999). Deep learning refers to machine learning models that have a deep structure granting them a great capability in obtaining the mid and high levels of abstractions from input raw data.

Deep neural networks, in particular convolutional neural networks (CNNs), have gained considerable interest by researchers in the medical field, due to its great efficacy in image classification (Krizhevsky et al., 2012, Albarqouni et al., 2016; Helwan et al., 2018). This motivated researchers to transfer the knowledge gained by these deep networks when trained on millions of images to address medical image diagnosis and classification tasks.

Accurate images classification has been achieved by deep learning based systems (Albarqouni et al., 2016; Helwan et al., 2018; Maa'itah and Abiyev, 2018). Those deep networks showed superhuman accuracies in performing such tasks. This success motivated the researchers to apply those networks on medical images for diseases classification tasks and the results showed that deep networks can efficiently extract useful features that distinguish different images classes (Avendi et al., 2016; Krizhevsky et al., 2012). Convolutional neural networks have been applied to various medical images diagnosis and classification due to its power of

extracting different level features from images systems (Albarqouni et al., 2016; Helwan et al., 2018; Maa'itah and Abiyev, 2018; Avendi et al., 2016; Krizhevsky et al., 2012).

Traditional networks have been also used in classifying medical diseases, however, their performance was not as efficient as the deep networks in terms of accuracy, computation time, and minimum square error achieved. In this work, traditional (shallow) and deep learning based networks are employed to classify most common thoracic diseases. Back propagation neural network (BPNN), competitive neural network (CpNN) and convolutional neural network (CNN) are examined in this study to classify 12 common diseases that may be found in chest X-ray, i.e, Atelectasis, Cardiomegaly, Effusion, Infiltration, Mass, Nodule, Pneumonia, Pneumothorax, Consolidation, Edema, Emphysema, Fibrosis (Figure 1.1).



**Figure 1.1:** Chest pathologies (Maa'itah and Abiyev, 2018)

The interpretation of a chest X-ray can diagnose many conditions and diseases such as pleurisy, effusion, pneumonia, bronchitis, infiltration, nodule, atelectasis, pericarditis, cardiomegaly, pneumothorax, fractures and many others.

Classifying the chest x-ray abnormalities is considered a tough task for radiologists. Hence, over the past decades, computer aided diagnosis (CAD) systems have been developed to extract useful information from X-rays to help doctors in having a quantitative insight about an X-ray. However, those CAD systems have not achieved a significance level to make decisions on the type of conditions of diseases in an X-ray. Thus, the role of them was left as visualization functionality that helps doctors in making decisions.

The aim of this thesis is the design of CXR identification system using a deep convolutional neural network (CNN). We explore the power of both traditional (supervised and unsupervised networks) and deep network in the classification of chest pathologies. The networks are all trained on the chest X-ray images and their performances are evaluated in classifying different chest diseases. The data used in obtained from the National Institutes of Health - Clinical Center (Wang et al., 2017) and it contains 112,120 frontal-view X-ray images of 30,805 unique patients.

## **1.2 Significance of the Work**

The problem of chest X-rays identification was proposed and presented by many previous researches. However, none of these researches employed a comparison of the unsupervised and supervised traditional networks, and deep models in solving this problem. Most of the related works are classifying chest X-rays using transfer learning models where pre-trained models are fine-tuned to classify chest x-rays. In this work, chest X-rays classification is discussed from different perspectives. Supervised conventional networks such as backpropagation neural network (BPNN), unsupervised traditional models competitive neural network (CpNN), and deep models such as convolutional neural networks (CNN) are all employed in this thesis, in order to investigate the problem chest X-rays classification.

Moreover, all of these different models are employed in order to find the optimum solution of this medical challenge in terms of accuracy, time, and error rates.

### **1.3 Thesis Overview**

This thesis is organized as follows:

**Chapter 1** is an introduction of the presented work in addition to identifying the problem statement of the thesis.

**Chapter 2** is brief explanation of the chest diseases that can be found in an X-ray.

**Chapter 3** is review of the three neural networks modalities used in this thesis, for the purpose of classifying the chest pathologies. This chapter reviews the traditional neural network basics such as backpropagation algorithm and supervised learning fundamentals. Moreover, the unsupervised learning and competitive neural networks are discussed in this chapter. Also, deep learning is presented in addition to the convolutional neural networks working principles.

**Chapter 4** presents a review of the use of soft computing tools such as neural networks, and deep learning in the field of medical image diagnosis and analysis, in particularly, chest X-ray identification.

**Chapter 5** discusses the simulation part of the work, in which the performance of the three employed models, during training and testing, are discussed.

Finally, **Chapter 6** shows a conclusion of the thesis in addition to listing some future recommendations that can be considered in order to further improvement of the work.

## CHAPTER 2

### RADIOGRAPHY OVERVIEW

#### 2.1 Introduction

In the territory of human services diagnostics, therapeutic image preparing has assumed a contributory part. From the different scopes of accessible radiological images created from ultrasound, x-Rays, attractive Resonance imaging, Computed Tomography, Positron Emission Tomography and so forth, every has its own particular technique for catching the images. Be that as it may, even in the wake of narrowing the focal point of the image catch, just a couple of segments of the radiological images are of clinical significance to the counseling doctor (Fushman et al., 2015). Be that as it may, there are different purposes behind which the pathologist and also radiologist trusted that image produced by such radiological test does not yield 100% exact data. For less minor types of the perilous ailment, such blunders may not make any difference much, but rather it does conceivably make a difference generally. In any case, presenting the patient to destructive radiological Rayss is restoratively not prudent and might be a significant costly issue for both specialist and patient. Consequently, from the previous decades utilization of image handling is progressively used to recognize the issues and settle it. The initial phase in such issue distinguishing proof is to perform image improvement. As though the image with clinical significance isn't upgraded it might possibly prompt exceptions in cutting edge investigation of medicinal information (Fushman et al., 2012).

Henceforth, image upgrade assumes a pivotal part in uncovering the malady with more data to the specialist or to the procedure of further investigation of the illness. This paper talks about the chest x-Rays images and proposed an answer with greater activity and lower computational cost for improving the chest x-Rayss. The radiological images particularly chest x-Rays images experiences following issues i.e. I) numerous outline, ii) nearness of rib confines (bones), iii) shadows of bosom in female subjects, iv) stomach and so on. In spite of

the fact that there are propelled adaptations of radiological images however chest x-Rays image is thought to be essential finding factor by the clinicians. Subsequently, if the chest x-Rays images are covered with different relics or issues, post diagnosis will dependably prompt anomalies. Subsequently, it is vital that chest x-Rays images ought to be appropriately pre-prepared even before subjecting it to propel investigation.

Therefore, this paper presents a very simple and cost effective chest X-rays classification system using deep networks for chest x-ray images with and without operations of enhancements.

## **2.2 Chest X-rays**

A chest X-Rays test is an extremely normal, non-obtrusive radiology test that creates an image of the chest and the inward organs. To deliver a chest X-Rays test, the chest is quickly presented to radiation from a X-Rays machine and an image is created on a film or into a computerized computer (Jaeger et al., 2014).

Chest X-Rays is additionally alluded to as a chest radiograph, chest roentgenogram, or CXR. Contingent upon its thickness, every organ inside the chest pit retains fluctuating degrees of radiation, creating distinctive shadows on the film. Chest X-Rays images are highly contrasting with just the brilliance or haziness characterizing the different structures. For instance, bones of the chest divider (ribs and vertebrae) may assimilate a greater amount of the radiation and along these lines, seem more white on the film.

Then again, the lung tissue, which is for the most part made out of air, will enable the majority of the radiation to go through, building up the film to a darker appearance. The heart and the aorta will seem whitish, however typically less brilliant than the bones, which are denser.

Chest X-Rays tests are requested by doctors for an assortment of reasons. Numerous clinical conditions can be assessed by this basic radiology test. A portion of the basic conditions recognized on a chest X-Rays include:

- pneumonia,
- enlarged heart,
- congestive heart disappointment,
- lung mass,
- rib cracks,
- fluid around the lung (pleural radiation), and
- Air around the lung (pneumothorax).

All in all, a chest X-Rays test is a straightforward, brisk, economical, and moderately innocuous system with insignificant danger of radiation. It is additionally broadly accessible.

## **2.3 Chest Abnormalities**

### **2.3.1 Pleural disease**

- The pleura and pleural spaces are just noticeable when abnormal
- There ought to be no noticeable space between the instinctive and parietal pleura
- Check for pleural thickening and pleural emanations
- If you miss a strain pneumothorax you hazard your patient's life – and in addition your outcome at finals!
- The pleura just wind up noticeable when there is an abnormality show. Pleural abnormalities can be unpretentious and it is critical to check precisely around the edge of every lung where pleural abnormalities are normally more effortlessly observed (Figure 1). A few illnesses of the pleura cause pleural thickening, and others prompt liquid or air assembling in the pleural spaces (Xue et al., 2015).



**Figure 2.1:** Pleural thickening (Xue et al., 2015)

### **2.3.2 Pneumothorax**

A pneumothorax happens when there is air caught in the pleural space. This may happen precipitously, or because of hidden lung illness. The most widely recognized reason is injury, with slash of the instinctive pleura by a broken rib (Figure 2).

On the off chance that the lung edge measures in excess of 2 cm from the inward chest divider at the level of the hilum, it is said to be 'substantial. If there is tracheal or mediastinal move far from the pneumothorax, the pneumothorax is said to be under 'pressure.' This is a restorative crisis! Missing a pressure pneumothorax may not just mischief your patient; it is likewise the snappiest method to fizzle the radiology OSCE at finals!



**Figure 2.2:** Pneumothorax (Xue et al., 2015)

### **2.3.3 Asbestos plaques**

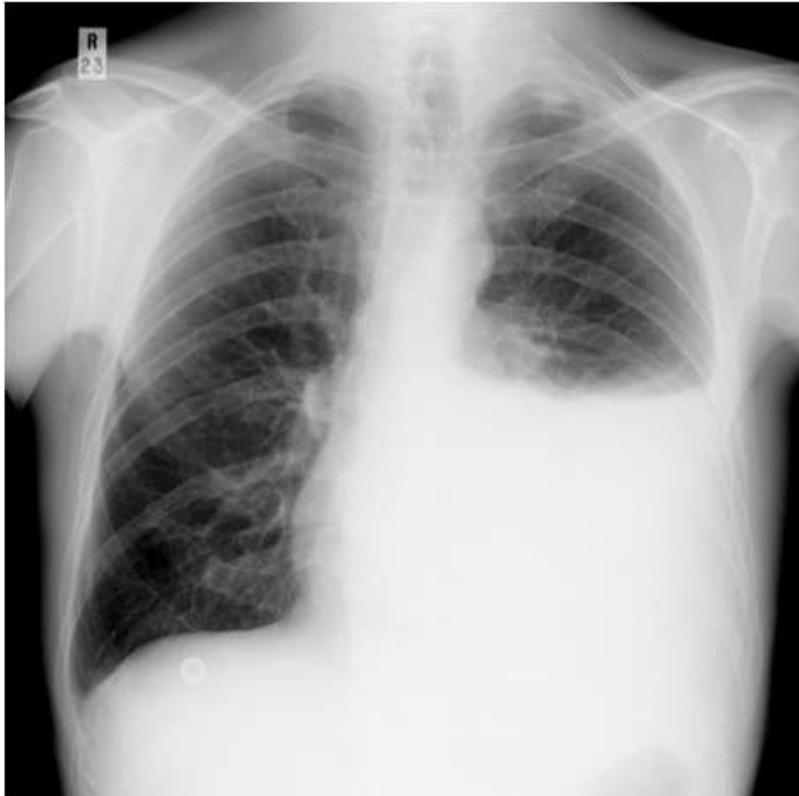
Calcified asbestos related pleural plaques have a trademark appearance, and are for the most part thought to be favorable. They are sporadic, all around characterized, and traditionally said to look like holly takes off (Candemir, et al., 2014).



**Figure 2.3:** Asbestos related pleural plaques (Candemir, et al., 2014)

#### **2.3.4 Pleural effusions**

A pleural emission or effusion is a gathering of liquid in the pleural space. Liquid accumulates in the most minimal piece of the chest, as per the patient's position. In the event that the patient is upright when the X-Rays is taken, at that point liquid will encompass the lung base framing a 'meniscus' – a sunken line clouding the costophrenic edge and part or the majority of the hemidiaphragm (Figure 4). On the off chance that a patient is recumbent, at that point a pleural radiation layers along the back part of the chest pit and ends up hard seeing on a chest X-Rays (Shiraishi et al., 2000).



**Figure 2.4:** Pleural effusion (Shiraishi et al., 2000)

## **2.4 Features Extraction in Medicine**

Pattern recognition is the process of developing systems that have the capability to identify patterns; while patterns can be seen as a collection of descriptive attributes that distinguishes one pattern or object from the other. It is the study of how machines perceive their environment, and therefore capable of making logical decisions through learning or experience. During the development of pattern recognition systems, we are interested in the manner in which patterns are modeled and hence knowledge represented in such systems. Several advances in machine vision have helped revamp the field of pattern recognition by suggesting novel and more sophisticated approaches to representing knowledge in recognition systems; building on more appreciable understanding of pattern recognition as achieved in the human visual processing.

Typical pattern recognition as the following important phases for the realization of its purpose for decision making or identification.

- Data acquisition: This is the stage in which the data relevant to the recognition task are collected.
- Pre-processing: It is at this stage that the data received in the data acquisition stage is manipulated into a form suitable for the next phase of the system. Also, noise is removed in this stage, and pattern segmentation may be carried out.
- Feature extraction/selection: This stage is where the system designer determines which features are significant and therefore important to the learning of the classification task.
- Features: The attributes which describe the patterns.
- Model learning/ estimation: This is the phase where the appropriate model for the recognition problem is determined based on the nature of the application. The selected model learns the mapping of pattern features to their corresponding classes.
- Model: This is the particular selected model for learning the problem, the model is tuned using the features extracted from the preceding phase.
- Classification: This is the phase where the developed model is simulated with patterns for decision making. The performance parameters used for accessing such models include recognition rate, specificity, accuracy, and achieved mean squared error (MSE).
- Post-processing: The outputs of the model are sometimes required to be processed into a form suitable for the decision making phase stage. Confidence in decision can be evaluated at this stage, and performance augmentation may be achieved.

- Decision: This is the stage in which the system supplies the identification predicted by the developed model.

There exist several approaches to the problem of pattern recognition such as syntactic analysis, statistical analysis, template matching, and machine learning using artificial neural networks.

Syntactic approach uses a set of feature or attribute descriptors to define a pattern, common feature descriptors include horizontal and vertical strokes, term stroke analysis; more compact descriptors such as curves, edges, junctions, corners, etc., which is termed geometric features analysis. Generally, it is the job of the system designer to craft such rules that distinguish one pattern or object from another. The designer is meant to explore attribute descriptors which are unique to identify each pattern, and where there seems to a conflict of identification rules such as can be observed in identifying Figure 6 and 9; they have same geometric feature descriptors save that one is the inverted form of the other, the system designer is meant to explore other techniques of resolving such issues (Yumusak and Temurtas, 2010).

Statistical pattern analysis uses probability theory and decision to infer the suitable model for the recognition tasks.

Template pattern matching uses the technique of collecting perfect or standard examples for each distinct pattern or object considered in the recognition task. It is with these perfect examples that the test patterns are compared. It is usually the work of the system designer to craft the techniques with which pattern variations or dissimilarities from the templates are measured, and hence determine decision boundaries as to accept or reject a pattern being a member of a particular class. Euclidean distance is a common used function to measure the distance between two vectors in n-dimensional space.

Template matching can either be considered as global or local depending on the approach and aim for which the recognition system is designed. In global template matching, the whole pattern for recognition is used to compare the whole perfect example pattern; whereas in local template matching, a region of the pattern for classification is used to compare a corresponding region in the perfect template.

Artificial neural networks, on the other hand, are considered intelligent pattern recognition systems due to their capability to learn from examples in a phase known as training. These systems have sufficed in lots of pattern recognition systems; the ease with which same learning algorithms can be applied to various recognition tasks is motivating.

In this approach, the designer is allowed to focus on determining features to be extracted for learning by the designed systems, rather than expending a huge amount of time, resources, and labour in understanding the whole details of the application domain; instead, the system learns relevant features that distinguish one pattern from the other (Yumusak and Temurtas, 2010).

## **2.5 Image Processing**

An image can be considered as a visual perception of a collection of pixels; where, a pixel can be seen as the intensity value at a particular coordinate in an image. Generally, pixels are described in 2D, such as  $f(x,y)$ .

The pixel values can vary in an image depending on the number gray levels used in the image. The range of pixels can be expressed as 0 to  $2^m$ , for an image with gray level of  $m$ . Image processing is a very important of computer vision, as image data can be suitably conditioned before machine learning.

### **2.5.2 Image enhancement**

Image processing has been extensively used in medicine. Image enhancement is always the most common process needed in this field. A medial image contains many parts and may have lot of noise. This makes it very tough for doctors to find the correct diagnosis of it. Image processing can be useful tool in this case as it helps in detecting and enhancing the images since all parts in image including noise differ from each other's in terms of brightness and intensities. Thus, in this work, image processing tools are used in order to enhance the chest X-ray images and remove the noise that may be found in them. This is done by using many

techniques for image enhancement such as filtering, histogram equalization, and intensity adjustment. An example of the working principle of the proposed algorithm is shown in Figure 5 (Yumusak and Temurtas, 2010).

In case of filtering, many filters can be used such as median, mean, Gaussian filters. For median filters, the images are filtered as some of them have noise artifacts which should be removed to enhance the quality of images. Median filter is a good technique for removing noise as it provides good rejection of the Salt and Pepper noise which is found in some medical images.

Moreover, image intensities adjustment can be also used for enhancing the quality of images. This technique involves the mapping of the pixels intensity distribution form one level to another level. To highlight the images more and more, the intensities of pixels are increased by mapping them into other values. This ended up with brighter images where the cells are clearer; including the cancerous cells.



**Figure 2.5:** Medical image enhancement (Yumusak and Temurtas, 2010)

## **CHAPTER 3**

### **LITERATURE REVIEW**

#### **3.1 Overview**

This chapter is a review of using machine learning techniques in medical image diagnosis and identification. Backpropagation neural networks and deep networks that are applied in solving medical image problems are shown in this chapter. Moreover, unsupervised learning based networks such as competitive neural networks applied in medical images analysis are also discussed.

#### **3.2 Review on Using Backpropagation Neural Networks in Medical Images Classification**

In a past work, (Cernazanu and Holban, 2012), described the segmentation of chest X-ray using convolutional neural network. In their work, they introduced image segmentation into bone tissue and non-bone tissue. The aim of their work was to develop an automatic or an intelligent segmentation system for chest X-rays. The system was established to have the capability to segment bone tissues from the rest of the image.

They were able to achieve the aim of the research by using a convolutional neural network, which was tasked with examining raw image pixels and hence classifying them into “bone tissue” or “non-bone tissue”. The convolutional neural networks were trained on the image patches collected from the chest X-ray images.

It was recorded in their work that the automatic segmentation of chest X-rays using the convolutional neural networks, and approaches suggested in their research produced plausible performance.

In another recent research, “lung Cancer Classification using Image Processing”, presented the application of some image processing techniques in the classification of patients chest X-rays into whether cancer is present or not (benign or malignant). In this work, it was shown that by

extracting some geometric features that are essential to the classification of the images such area, perimeter, diameter, and irregularity; an automatic classification system was developed.

Furthermore, in the same research, texture features were considered for a parallel comparison of results on the classification accuracy. The texture features used in the work are average gray level, standard deviation, smoothness, third moment, uniformity, and entropy. The back propagation neural network was used as the classifier, and an accuracy of 83% was recorded in the work (Patil and Kuchanur, 2012).

Schnorrenberg (1996) has suggested that a computer aided system that can estimate the malignancy probability of mammography lesion can assist the radiologists to decide patient management while improving the diagnostic accuracy. And since , various classifiers such as linear discriminants, rule based methods, and artificial intelligence (AI) are being investigated for building systems that can classify mass lesions in mammography by merging computer-extracted image features.

Andre et al., (2002) proposed a Kohonen's self-organizing map (SOM) which extracts and digitize the features from the mammograms. The whole system is ultimately based on artificial neural networks (ANN) where it offers segmented image data from SOM as an input to the MLP network for the diagnosis task. The performance of the system was not so good compared to the other state-of-the-art systems present, with only 60% of the cases were classified correctly, however the results obtained in this study indicate that the use of SOM to digitize mammograms is possible with an attempt to improve and optimize the system.

### **3.3 Review on Using Unsupervised Learning in Medical Images Classification**

Availability of labelled data for supervised learning is a major problem for narrow AI in current day industry. In imaging, the task of semantic segmentation (pixel-level labelling) requires humans to provide strong pixel-level annotations for millions of images and is difficult when compared to the task of generating weak image-level labels.

Unsupervised representation learning along with semi-supervised classification is essential when strong annotations are hard to come by. Unsupervised learning was used in many researches where output images are not labelled images or when labeling output images require a long time.

Shan et al., (2017) proposed a registration algorithm for 2D CT/MRI medical images with a new unsupervised end-to-end strategy using convolutional neural networks via direct deformation field prediction. The contribution of their algorithm is the development of an end-to-end CNN-based learning system under an unsupervised learning setting that performs image-to-image registration. Moreover, the Training of that CNN with additional data without any label which can further improves the registration performance. The presented method was capable of achieving a 100x speed-up compared to traditional image registration methods.

Dosovitskiy et al. (2015) propose an end-to-end fully convolutional neural net FlowNet for optical flow estimation in real time. FlowNet has an encoder-decoder architecture with skip connections. It predicts optical flow at multiple scales and each scale is predicted based on the previous scale. Compared with the nature of supervised learning of Flownet, an unsupervised architecture is utilized in this work to predict deformation field that aligns two images.

Jaderberg et al. (2015) propose the spatial transformer networks (STN) which focuses on class alignment. It shows that spatial transformation parameters (e.g. affine transformation parameters, B-Spline transformation parameters, deformation field, etc) can be implicitly learned without ground-truth supervision by optimizing a specific loss function (Jaderberg et al., 2015). STN is a fully differentiable module that can be inserted into existing CNNs, which makes it possible to cast the image registration task as an image reconstruction problem.

Wu et al. (2013) adopt unsupervised deep learning to obtain features for image registration. Though good performance is achieved, their method is a patch-based learning system and relies on other feature-based registration methods to perform image registration. Ren et al. (2017) and Yu et al. (2016) use the spatial transformer networks (STN) (Yu et al., 2016) and optical flow produced by a CNN to warp one frame to match its previous frame. The

difference between two frames after warping is used as the loss function to optimize the parameters of CNN. Their unsupervised methods do not require any ground-truth optical flow. Similarly, Garg et al. (2016) use an image reconstruction loss to train a network for monocular depth estimation. This work is further ameliorated by incorporating a fully differentiable training loss and left right consistency check (Godard et al., 2017). We follow the idea of these works to train a model for image-to-image registration in an unsupervised manner.

### **3.4 Review on Using Deep Learning in Medical Images Classification**

Neural networks have advanced at a remarkable rate, and they have found practical applications in various industries (Szegedy et al., 2015). Deep neural networks define inputs to outputs through a complex composition of layers which present building blocks including transformations and nonlinear functions (Abadi et al., 2016). Now, deep learning can solve problems which are hardly solvable with traditional artificial intelligence (LeCun et al., 2015). Deep learning can utilize unlabeled information during training; it is thus well-suited to addressing heterogeneous information and data, in order to learn and acquire knowledge. The applications of deep learning may lead to malicious actions; however the positive use of this technology is much broader.

#### **3.4.1 Review on using convolutional neural networks in medical images classification**

Back in 2015, it was noted that deep learning has a clear path towards operating with large data sets, and thus, the applications of deep learning are likely to be broader in the future (LeCun et al., 2015). A large number of newer studies have highlighted the capabilities of advanced deep learning technologies, including learning from complex data (Miotto et al., 2017; Wei et al., 2017), image recognition (Wei et al., 2017), text categorization (Song et al., 2016) and others. One of the main applications of deep learning is for medical diagnosis (Lee et al., 2017; Suzuk, 2017). This includes but is not limited to health informatics (Ravi et al., 2017). Biomedicine (Mamoshina et al., 2016), and magnetic resonance image MRI analysis (Liu et al., 2018). More specific uses of deep learning in the medical field are segmentation,

diagnosis, classification, prediction, and detection of various anatomical regions of interest (ROI). Compared to traditional machine learning, deep learning is far superior as it can learn from raw data, and has multiple hidden layers which allow it to learn abstractions based on inputs (Miotto et al., 2017). The key to deep learning capabilities lies in the capability of the neural networks to learn from data through general purpose learning procedure.

A convolutional neural network was proposed by (Avetisian, 2017) for the purpose of segmentation of medical images. The network trains from manually labeled images and can be used to segment various organs and anatomical structures of interest. The authors proposed an efficient reformulation of a 3D convolution into a series of 2D convolutions in different dimensions. A loss function that directly optimizes intersection-over-union metric popular in medical image segmentation field is also proposed. Experimentally, the authors showed that their designed convolutional neural network is capable of segmenting visually distinguishable anatomical structures on medical images.

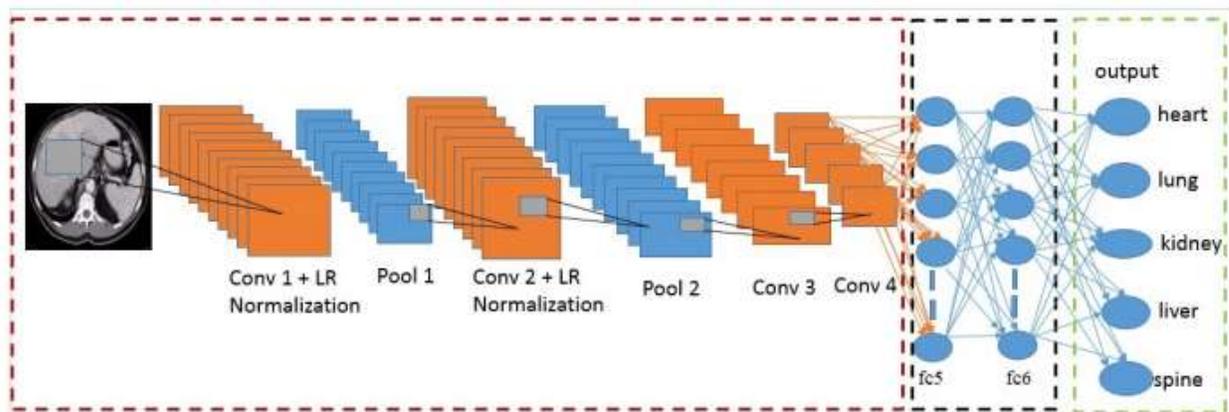
Moreover, the image classification task has been conducted in a single specific domain of anatomies and modalities, such as CT lung images (Jennifer et al., 2003), X-ray and CT images of different body parts i.e. skull, breast, chest, hand etc (Srinivas et al., 2015), breast ultrasound images (Ren et al., 2005). Although a variety of feature representation have been proposed for classifying medical images, these feature representations are domain specific that cannot be applied to other classes keeping in mind the variability in medical images.

In a this study done by (Khan and Yong, 2017), a Convolutional Neural Network (CNN) architecture for automatically classifying anatomy in medical images by learning features at multiple level of abstractions from the data obtained, was proposed. The authors of this paper aimed to present a comprehensive evaluation of the three milestone CNN architectures, i.e. LeNet, AlexNet and GoogLeNet for classifying medical anatomy images. The findings from the performance analysis of these architectures advocates the need of a modified architecture because of their poor performance for medical image anatomy classification. Hence, a modified Convolutional Neural Network architecture for classifying anatomies in medical

images was proposed. Our proposed model of the CNN architecture is a modification of the basic architecture of AlexNet (Krizhevsky et al., 2012).

This architecture contains four convolutional layers (conv) followed by two fully connected layers (fc). The first convolutional layer i.e conv1 subjected to local response normalization, with kernel size 11, which depicts that each unit in each feature map is connected to  $11 \times 11$  neighborhood in the input and stride of 4, which means after every four pixels perform the convolution on the input images.

The architecture of the proposed CNN used for medical image anatomy classification is as shown in Figure 3.1.



**Figure 3.1:** Modified Alexnet proposed by Khan and Yong, (2017)

Experiments were conducted with a machine incorporated with NVIDIA GeForce GTX 980M, using a data set that contains thousands of anonymous annotated medical imaging data. Anatomical images that are used in this experimentation consist of CT, MRI, PET, Ultrasound and X-ray modalities. This database contains images with various pathologies. For their experimental evaluation, the authors adopted 37198 images of five anatomies to train the CNN models. For testing, 500 images were used other than that in the training set, i.e. 100 images per anatomy. So a total of 37698 images were used in the experiments. The anatomies considered in our experiments were lung, liver, heart, kidney and lumbar spine.

The normal and pathological images were used, so that these frameworks should be generalized to classify any image of the same organ if it varies in shape or contrast. The dataset was tested with the three milestone architectures, i.e LeNet (LeCun et al., 1998), AlexNet (Krizhevsky et al., 2012) and GoogLeNet (Szegedy et al., 2015).

The summary of the comparative performance of the proposed CNN and the three milestone architectures in terms of runtime, training loss, validation accuracy and test accuracy. It was found that the proposed CNN outperforms other three milestone CNN architectures by having 81% accuracy while AlexNet achieved only 74%, followed by LeNet 59% and GoogleLeNet 45%. Moreover, it was seen that analyzing these visualizations of the networks learned filters clearly depict that the filters learned by LeNet and GoogLeNet are not distinguishable enough to depict the edge like features, that are supposed to be learned by the first convolutional layer as there is lot of noise in filters.

## **CHAPTER 4**

### **DESIGN OF CNN BASED CHEST X-RAY PATHOLOGY IDENTIFICATION SYSTEM**

#### **4.1 Overview**

This chapter discusses the design of the proposed system for the identification of chest X—rays pathologies. In this chapter, a brief explanation of deep learning is presented, in addition to detailed explanation of the convolutional neural network and its working principles.

#### **4.2 Deep Learning**

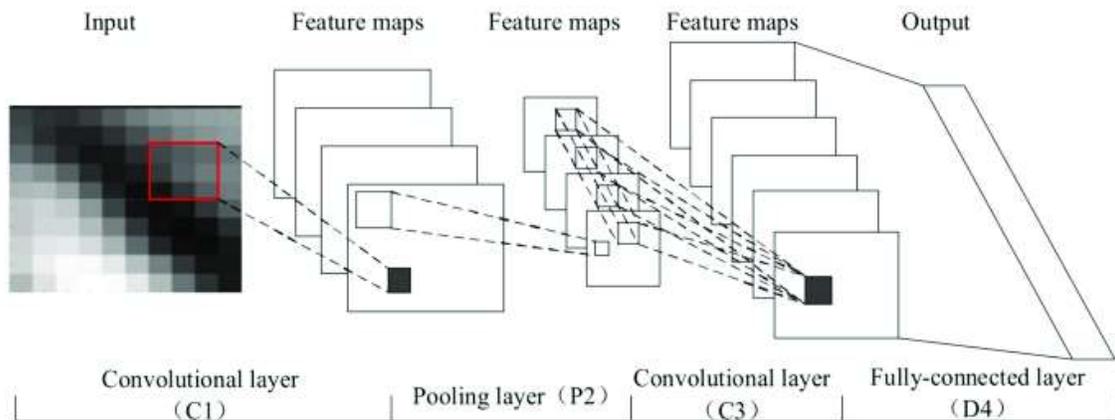
Deep convolutional neural network (DCNN) is a deep system that involves two-dimensional discrete convolution for image analysis to imitate human neural system activity, which has a comparable structure to the various leveled model of the visual nervous human system. This type of networks was first proposed by LeCun et al., (1994) which use the back propagation (BP) algorithm, as a feasible training technique of this network. Consequently, LeCun et al., (1998) employed a deep structure network and trained its parameters using the BP algorithm. The network achieved a great performance with high accuracy in recognizing the hand-written digits.

Generally, DCNN is a deep structure network in which mathematical operations such as convolution and subsampling occur within its hidden layers. These operations allow the network to learn different levels of features which results in an automatic extraction of deeply distinct and effective representation of the input data (Bengio et al., 2007). Moreover, DCNN coordinates a local connection mechanism in addition to a weight sharing aiming to reduce the number of learning parameters which consequently reduces the computations time and cost. This network has achieved significant performance in various areas where it was applied, i.e. computer vision (Avendi et al., 2016), biological computation (Wang et al., 2017), medical images classification (Maa'itah and Abiyev, 2018), etc.

The advancement in computer industries has motivated the researchers to further improve the performance of the CNN by making it deeper and more feasible. Therefore, a CNN of 19 layers was proposed and called VGG-Net (Simonyan and Zisserman, 2014), also Szegedy et al. (2015) proposed a 22 layers deep network named GoogLeNet which also includes an improvement in the architecture and working principles of the CNN by adding an inception module to it. Moreover, a CNN of 152 layers named ResNet (ResNet-152) was proposed by He et al. (2016).

### **4.3 Convolutional Neural Networks**

Deep learning is a machine learning method inspired by the deep structure of a mammal brain (LeCun et al., 2006). This method is characterized by a deep architecture in which multiple hidden layers are employed, which allows the abstraction of different levels of features. In 2006, Hinton et al. developed a new algorithm to train this deep architecture of neuron layers, which they called greedy layer-wise training (Hinton et al., 2006). This learning algorithm is basically seen as an unsupervised single layer greedily training where a deep network is trained layer by layer. Afterwards, this method became more effective and started to be used for training many later on proposed deep networks. One of the most powerful deep networks is the Convolutional neural network, a deep network comprised of many hidden layers performing convolution and subsampling in order to extract low to high levels of features of the input data (Krizhevsky et al., 2012; Rios and Kavuluru, 2015; Helwan et al., 2018). This network has shown a great efficiency in different areas where it was applied, i.e. computer vision (Krizhevsky et al., 2012), biological computation (Rios and Kavuluru), medical images classification (Helwan et al., 2018) etc... Basically, this type of networks consists of three main layers: convolution layers, subsampling or pooling layers, and full connection layers. Each type of layers is explained briefly in the following paragraph. Figure 3 shows a typical architecture of a convolutional neural network (CNN).

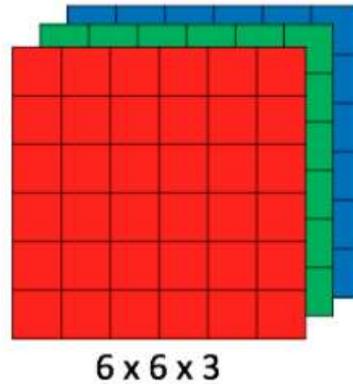


**Figure 4.1:** Typical architecture of a convolutional neural network (CNN)

#### 4.4 Understanding the Learning of Convolutional Neural Networks

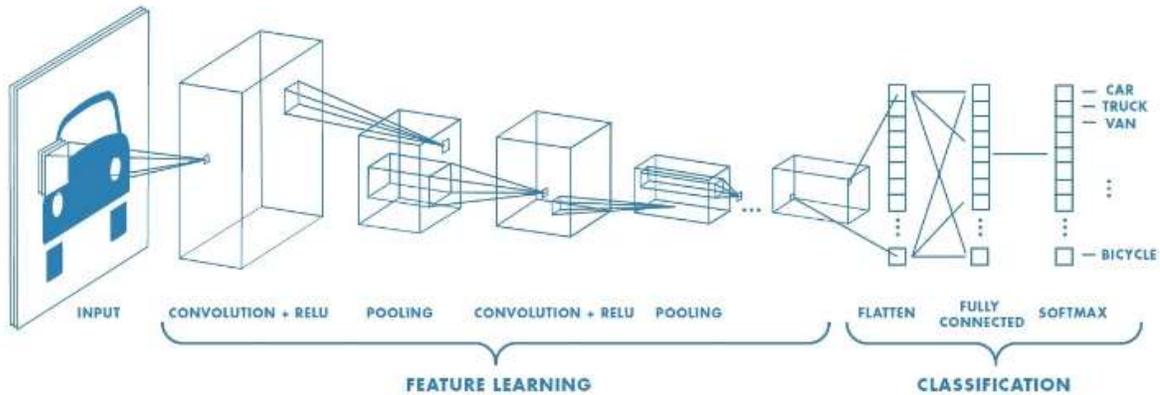
In neural networks, Convolutional neural network (ConvNets or CNNs) is one of the main categories to do images recognition, images classifications. Objects detections, recognition faces etc., are some of the areas where CNNs are widely used.

CNN image classifications takes an input image, process it and classify it under certain categories (Eg., Dog, Cat, Tiger, Lion). Computers see an input image as array of pixels and it depends on the image resolution. Based on the image resolution, it will see  $h \times w \times d$  ( $h$  = Height,  $w$  = Width,  $d$  = Dimension ). Eg., An image of  $6 \times 6 \times 3$  array of matrix of RGB (3 refers to RGB values) and an image of  $4 \times 4 \times 1$  array of matrix of grayscale image.



**Figure 4.2:** Array of RGB Matrix

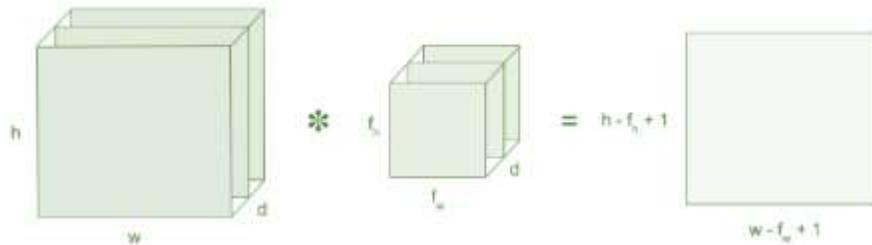
Technically, deep learning CNN models to train and test, each input image will pass it through a series of convolution layers with filters (Kernels), Pooling, fully connected layers (FC) and apply Softmax function to classify an object with probabilistic values between 0 and 1. The below figure is a complete flow of CNN to process an input image and classifies the objects based on values.



**Figure 4.3:** Example of a neural network with many convolutional layers

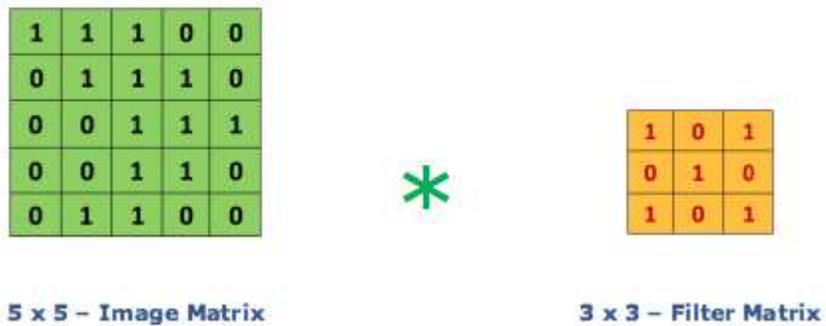
Convolution is the first layer to extract features from an input image. Convolution preserves the relationship between pixels by learning image features using small squares of input data.

It is a mathematical operation that takes two inputs such as image matrix and a filter or kernel



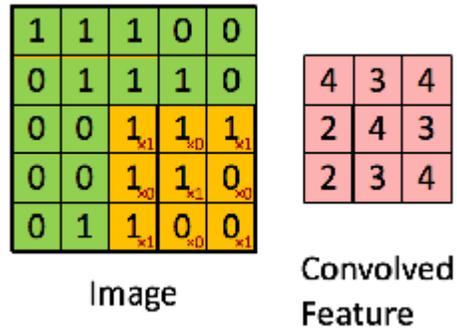
**Figure 4.4:** Image matrix multiplies kernel or filter matrix

Consider a 5 x 5 whose image pixel values are 0, 1 and filter matrix 3 x 3 as shown in below



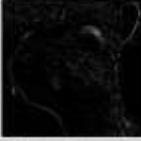
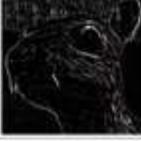
**Figure 4.5:** Image matrix multiplies kernel or filter matrix

Then the convolution of 5 x 5 image matrix multiplies with 3 x 3 filter matrix which is called “Feature Map” as output shown in below



**Figure 4.6:** 3 x 3 Output matrix

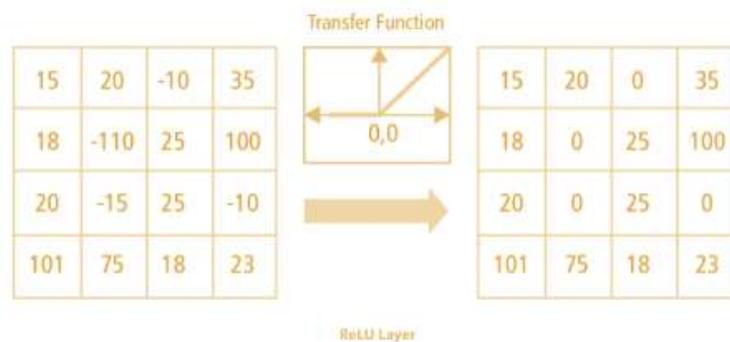
Convolution of an image with different filters can perform operations such as edge detection, blur and sharpen by applying filters. The below example shows various convolution image after applying different types of filters (Kernels).

Operation	Filter	Convolved Image
Identity	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
Edge detection	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	

**Figure 4.7:** Some common filters

#### 4.4.1 Non linearity (ReLU)

ReLU stands for Rectified Linear Unit for a non-linear operation. The output is  $f(x) = \max(0, x)$ . ReLU's purpose is to introduce non-linearity in our ConvNet. Since, the real world data would want our ConvNet to learn would be non-negative linear values.



**Figure 4.8:** ReLU operation

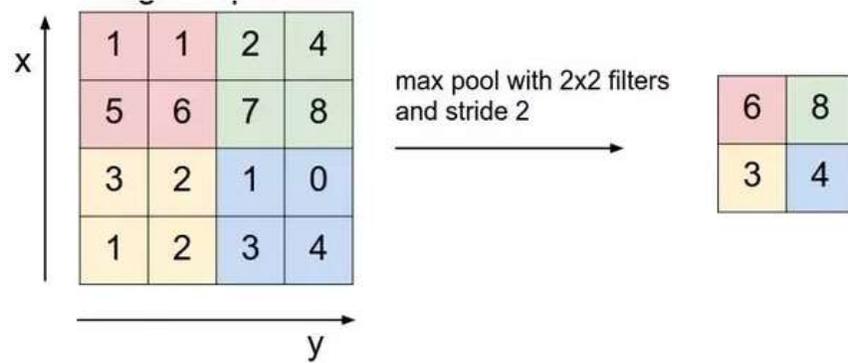
There are other non linear functions such as tanh or sigmoid can also be used instead of ReLU. Most of the data scientists uses ReLU since performance wise ReLU is better than other two.

#### 4.4.2 Pooling layer

Pooling layers section would reduce the number of parameters when the images are too large. Spatial pooling also called subsampling or downsampling which reduces the dimensionality of each map but retains the important information. Spatial pooling can be of different types:

- Max Pooling
- Average Pooling
- Sum Pooling

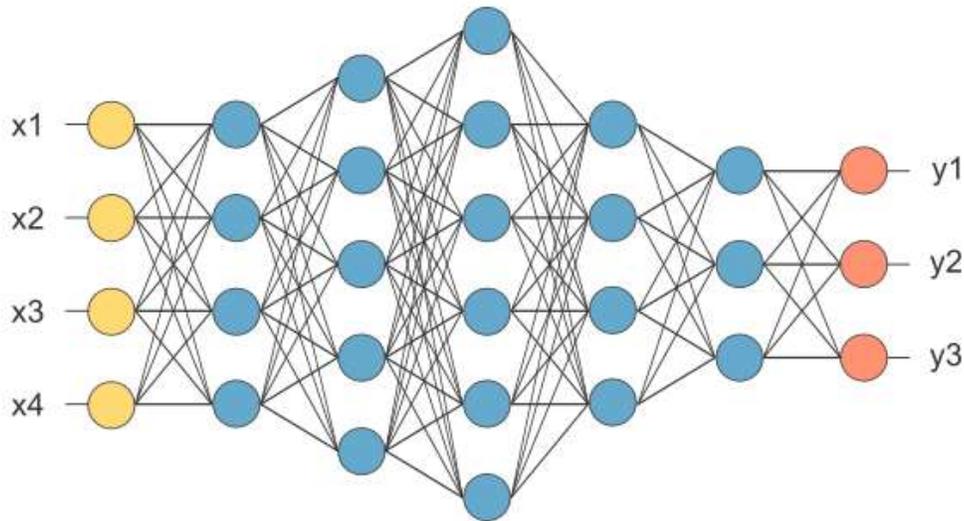
Max pooling take the largest element from the rectified feature map. Taking the largest element could also take the average pooling. Sum of all elements in the feature map call as sum pooling.



**Figure 4.9:** Max Pooling

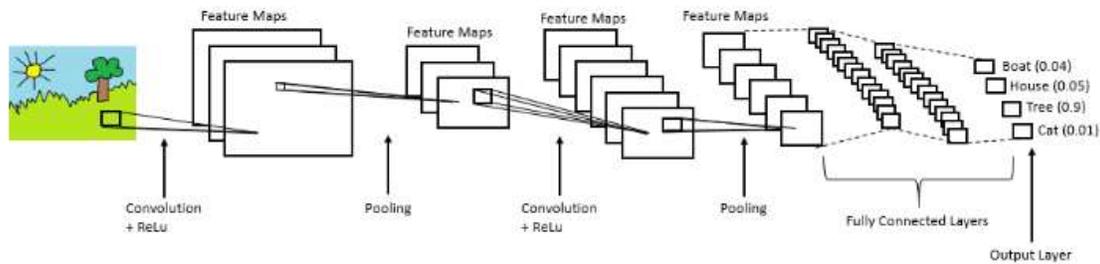
#### 4.4.3 Fully connected layer (FC)

The layer is called as FC layer; the matrix was flattened into vector and feed it into a fully connected layer like neural network.



**Figure 4.10:** After pooling layer, flattened as FC layer

In the above diagram, feature map matrix will be converted as vector ( $x_1, x_2, x_3, \dots$ ). With the fully connected layers, these features are combined together to create a model. Finally, we have an activation function such as softmax or sigmoid to classify the outputs as cat, dog, car, truck etc..



**Figure 4.11:** Complete CNN architecture

## 4.5 Stochastic Gradient Descent

Stochastic Gradient Descent is a batch technique training algorithm that is used in calculating the cost function in the convolutional neural network. Batch methods, such as limited memory Gradient Descent, which use the full training set to compute the next update to parameters at each iteration tend to converge very well to local optima. They are also straight forward to get working provided a good off the shelf implementation (e.g. minFunc) because they have very few hyper-parameters to tune. However, often in practice computing the cost and gradient for the entire training set can be very slow and sometimes intractable on a single machine if the dataset is too big to fit in main memory. Another issue with batch optimization methods is that they don't give an easy way to incorporate new data in an 'online' setting. Stochastic Gradient Descent (SGD) addresses both of these issues by following the negative gradient of the objective after seeing only a single or a few training examples. The use of SGD In the neural network setting is motivated by the high cost of running back propagation over the full training set. SGD can overcome this cost and still lead to fast convergence.

The standard gradient descent algorithm updates the parameters  $\theta$  of the objective  $J(\theta)$  as,

$$\theta = \theta - \alpha \nabla_{\theta} E[J(\theta)]$$

where the expectation in the above equation is approximated by evaluating the cost and gradient over the full training set. Stochastic Gradient Descent (SGD) simply does away with

the expectation in the update and computes the gradient of the parameters using only a single or a few training examples. The new update is given by,

$$\theta = \theta - \alpha \nabla_{\theta} J(\theta; x^{(i)}, y^{(i)})$$

with a pair  $(x^i, y^i)$  from the training set.

Generally each parameter update in SGD is computed w.r.t a few training examples or a minibatch as opposed to a single example. The reason for this is twofold: first this reduces the variance in the parameter update and can lead to more stable convergence, second this allows the computation to take advantage of highly optimized matrix operations that should be used in a well vectorized computation of the cost and gradient. A typical minibatch size is 256, although the optimal size of the minibatch can vary for different applications and architectures.

In SGD the learning rate  $\alpha$  is typically much smaller than a corresponding learning rate in batch gradient descent because there is much more variance in the update. Choosing the proper learning rate and schedule (i.e. changing the value of the learning rate as learning progresses) can be fairly difficult. One standard method that works well in practice is to use a small enough constant learning rate that gives stable convergence in the initial epoch (full pass through the training set) or two of training and then halve the value of the learning rate as convergence slows down. An even better approach is to evaluate a held out set after each epoch and anneal the learning rate when the change in objective between epochs is below a small threshold. This tends to give good convergence to a local optima. Another commonly used schedule is to anneal the learning rate at each iteration  $t$  as  $\frac{a}{b+t}$  where  $a$  and  $b$  indicate the initial learning rate and when the annealing begins respectively. More sophisticated methods include using a backtracking line search to find the optimal update.

One final but important point regarding SGD is the order in which we present the data to the algorithm. If the data is given in some meaningful order, this can bias the gradient and lead to

poor convergence. Generally a good method to avoid this is to randomly shuffle the data prior to each epoch of training.

#### 4.6 Design of CNN Based Chest X-ray Identification System

In this thesis, a convolutional neural network of three convolution layers is designed for the identification of 12 chest pathologies. This network architecture is explained as follows:

- *Convolution Layer:* In this layer, an input image of size R\*C is convolved with a kernel of size a\*a as shown in Figure 4. Each block of the input matrix is independently convolved with the kernel and generates a pixel in the new output image. The result of the convolution of the input image and kernel is used to generate the n output images of this layer. Generally, a kernel of the convolution matrix is referred to as a filter, while the output images obtained by convolving kernel and the input images are referred to as filter maps of size i\*i.

In each convolution layer, there is a bunch of n filters. These filters are convolved with the input image and the depth of the generated feature maps ( $n^*$ ) is equivalent to the number of filters applied in the convolution operation. Note that each filter map is considered as a specific feature at certain location of the input image.

The output of the convolution layer, denoted  $Y_i^{(l)}$ , is computed as

$$Y_i^{(l)} = B_i^{(l)} + \sum_{j=1}^{a_i^{(l-1)}} K_{i,j}^{(l-1)} * Y_j^{(l-1)}$$

Where  $B_i^{(l)}$  is the bias matrix;  $K_{i,j}^{(l-1)}$  is convolution filter or kernel of size  $a*a$ , that connects the  $j^{th}$  feature map in layer (l-1) with the  $i^{th}$  feature map in the same layer.

After the convolution layer, the outputs pass through an activation function named regularized linear units (ReLUs). This function is popularly used in deep learning models due to its help in reducing the interaction and non-linear effects. ReLU converts the output to 0 if it receives a negative input, while it returns the same input value if it is positive. Thus, it can be written as:

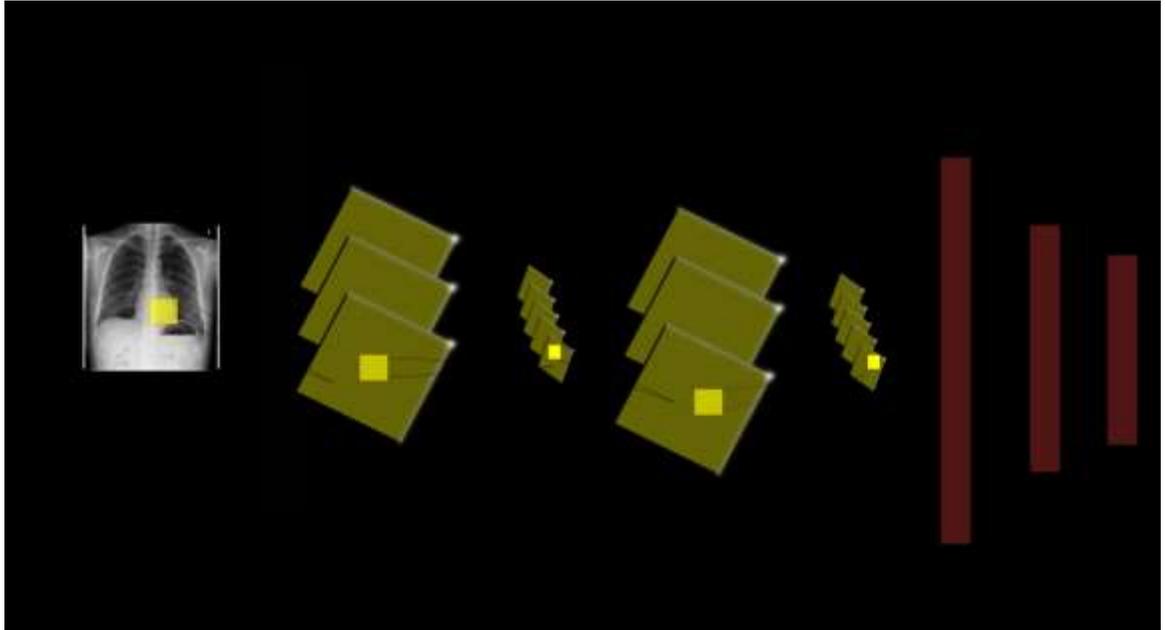
$$Y_i^{(l)} = \max(0, Y_i^{(l-1)})$$

Where  $Y_i^{(l)}$  is the output of the activation function, and  $Y_i^{(l-1)}$  is the input that it receives.

- *Subsampling Layer:* The main aim of this layer is to spatially reduce the dimensionality of the feature maps extracted in the previous convolution layer. To do so, a mask of size  $b \times b$  is selected as shown in Figure 4; and the subsampling operation between the mask and the feature maps in the convolution layer results in output images or filter maps of size  $i \times i$ , where  $i \times i$  is the size of feature maps of the convolution layer. Many subsampling methods were proposed such as averaging pooling, sum pooling, and maximum pooling. However, the most commonly used method is the Max pooling where the maximum value of each block is the corresponding pixel value of the output image. Note that a subsampling layer helps the convolution layer to tolerate rotation and translation among the input image.
- *Full connection:* the final layer of a CNN is a traditional feedforward network with one or more hidden layers and an output layer that uses Softmax activation function.

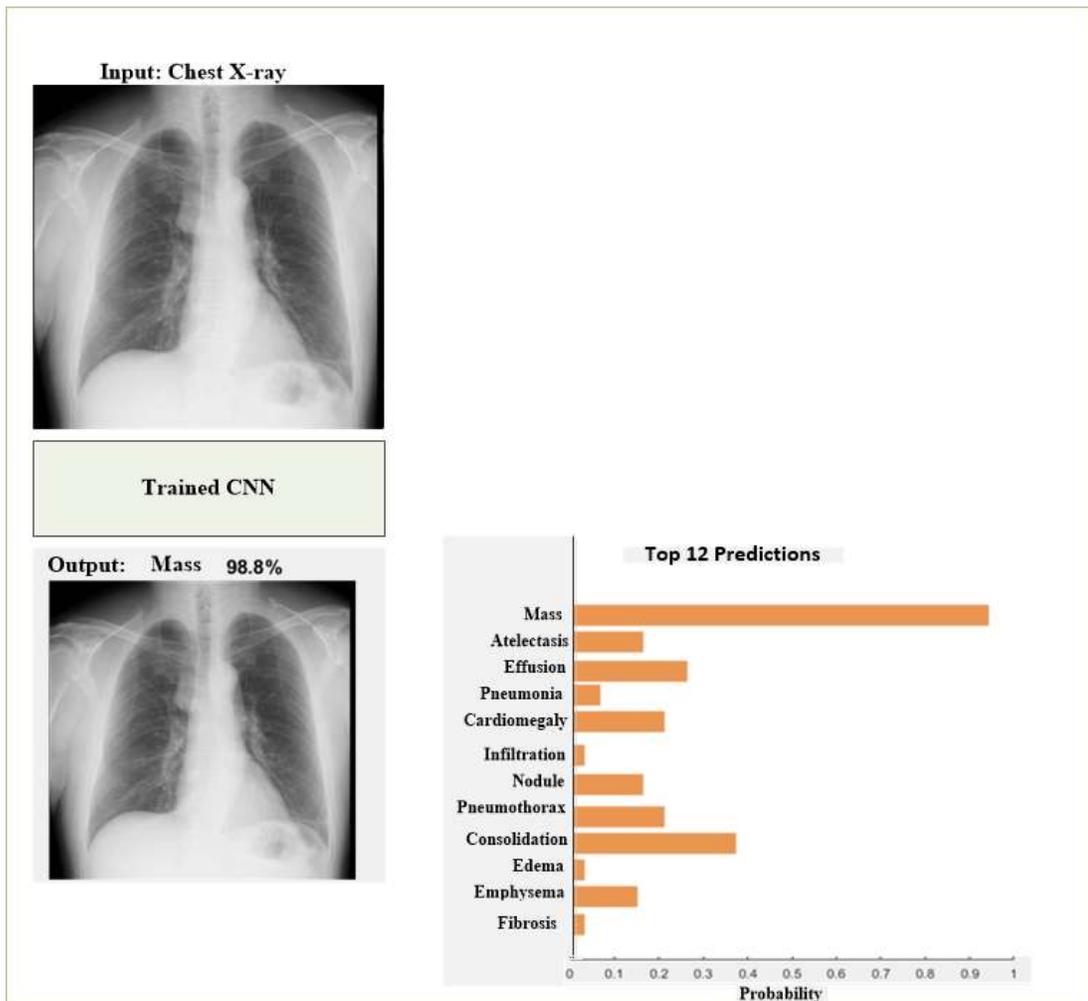
$$y_i^{(l)} = f(z_i^{(l)}) \text{ With } z_i^{(l)} = \sum_{j=1}^{m_i^{(l-1)}} w_{i,j}^{(l)} y_j^{(l-1)}$$

Where  $w_{i,j}^{(l)}$  are the weights that should be tuned by the complete fully connected layer in order to form a stochastic representation of each class, using a transfer function  $f$  which represents the non-linearity. Note that the non-linearity in the fully connected layer is built within its neurons, not in separate layers as in convolutions and pooling layers.



**Figure 4.12:** The proposed Convolutional neural network for chest X-ray pathology identification

After training, the network is expected to classify the type of chest pathology found in an image in terms of probability as shown in figure 4.12.



**Figure 4.13:** CNN final classification of chest X-rays with classes probabilities

## CHAPTER 5 SIMULATION

### 5.1 Overview

This chapter presents the simulation phase of the three employed models. It shows the training phase of each model in addition to the visualizations of the results and accuracies achieved. Moreover, in this chapter, the performance of each employed model during the testing phase is shown.

### 5.2 Dataset

The networks are all trained on the chest X-ray images and their performances are evaluated in classifying different chest diseases. The data used in training and testing the employed models are obtained from the National Institutes of Health - Clinical Center (Wang et al., 2017). This database is a public and it contains 112,120 frontal-view X-ray images of 30,805 unique patients. Images of this database are originally of 1064\*1064 pixels, however, they were all resized to 32\*32 pixels for faster processing. Note that this size was selected as it fastens the learning of the network and preserves the good and useful features of the image. Table 5.1 shows the dataset description and the division of data into training and testing for each network

**Table 5.1:** Dataset description

	<b>Total number of images</b>	<b>Training</b>	<b>Testing</b>
BPNN	1000	620	380
CpNN	1000	620	380
CNN	112,120	70%	30%

### 5.3 Simulations

In this work, the two employed models are trained and tested using Matlab environment. The networks were simulated on a Windows 64-bit desktop computer with an Intel Core i7 4770 central processing unit (CPU) and 8 GB random access memory. Note that there was no graphical processing unit (GPU) available in the used desktop.

In this section, the simulations of the above networks are described. Note that the BPNN and CpNN networks are trained using 620 out of 1000 images and the rest is used for testing. The CNN is trained using 70% of 120,120 available data and 30% are used for testing as shown in table 5.1. The input images are of size 32×32 for the sake of reducing computation cost.

Loss and accuracy of each model were calculated as follows:

$$\text{Loss} = -(1/n) \sum_{i=1}^n \log(P(CC)) \quad (1)$$

$$\text{Accuracy} = \frac{CC}{T} \quad (2)$$

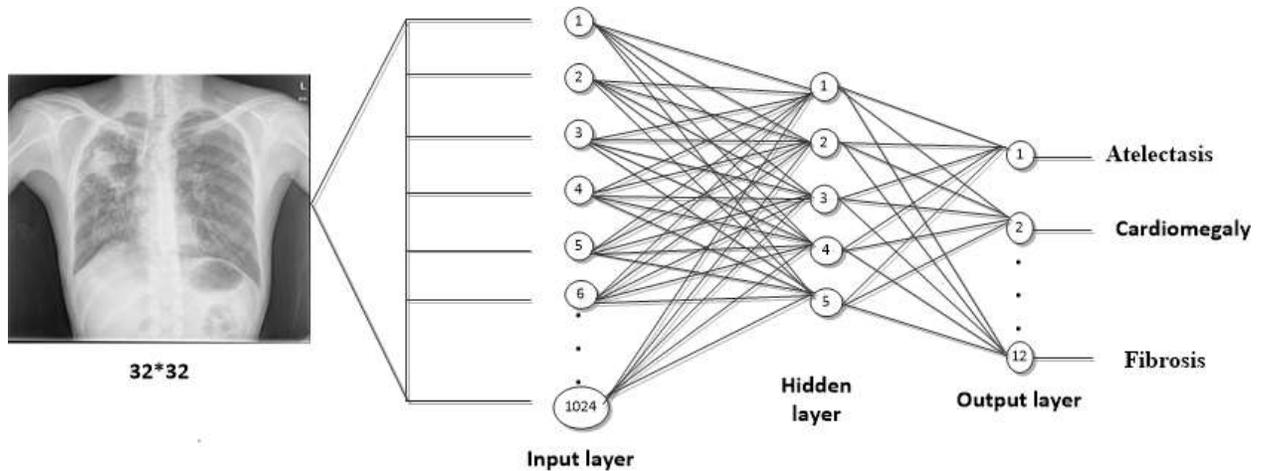
Where P(CC) is the probability of the correctly classified images, n is the number of images, and T is the total number of images during the training and/or testing phases.

#### 5.3.1 BPNN training

Back propagation neural networks are very important and useful in pattern recognition problems; it is based on a supervised learning algorithm. The successful training of back propagation networks is generally a heuristic process (trial and error), in order to obtain network parameters which produce good results.

Hence, in this paper, several experiments were conducted such that significantly important results can be obtained. The learning parameters varied include, the number of hidden neurons, the learning rate, and momentum rate.

The architecture of the designed back propagation neural network for the  $32 \times 32$  images is described below in Figure 5.1.



**Figure 5.1:** Back propagation neural network

An image from the database representing an X-ray chest was used in the figure to clearly show the design and operation of the network.

Since, the back propagation network uses a supervised learning algorithm, it is therefore necessary that the training data be labelled. The training data used in this work have been labelled according to the 12 categories present in the classification task. It is the aim to develop an intelligent classification system for chest X-rays diseases.

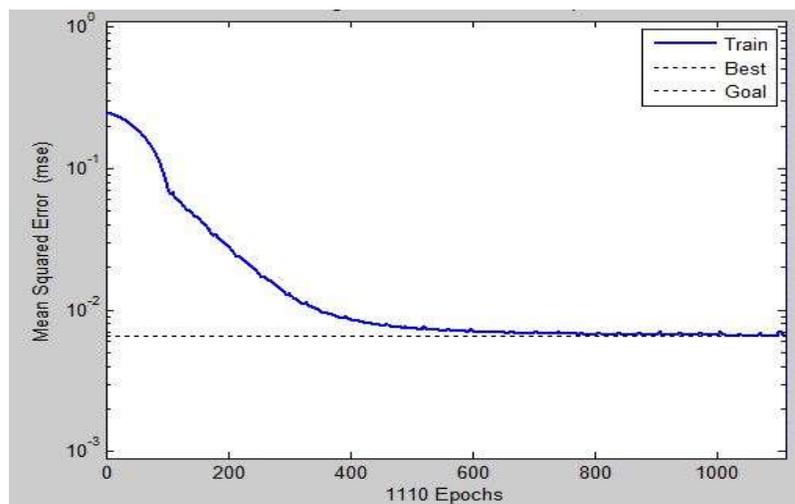
Different number of hidden neurons, learning rate, and momentum rates were experimented with during the training phase of the network as can be seen in Table 5.2 as BPNN1, BPNN2, BPPN3, and BPNN4.

Since there are 12 classes for the classification task, 12 neurons have been used in the output layer of the network.

**Table 5.2:** Training parameters for backpropagation networks (32×32 input pixels)

Networks	BPPN1	BPNN2	BPNN3	BPNN4
Training samples	620	<b>620</b>	620	620
Hidden neurons	20	<b>35</b>	45	60
Learning rate	0.010	<b>0.0045</b>	0.300	0.15
Momentum rate	0.040	<b>0.0072</b>	0.0504	0.0619
Activation function	Sigmoid	<b>Sigmoid</b>	Sigmoid	Sigmoid
Epochs	1000	<b>1110</b>	1256	1374
Training time (secs)	148	<b>156</b>	184	193
Mean Squared Error	0.0077	<b>0.0025</b>	0.0056	0.0096

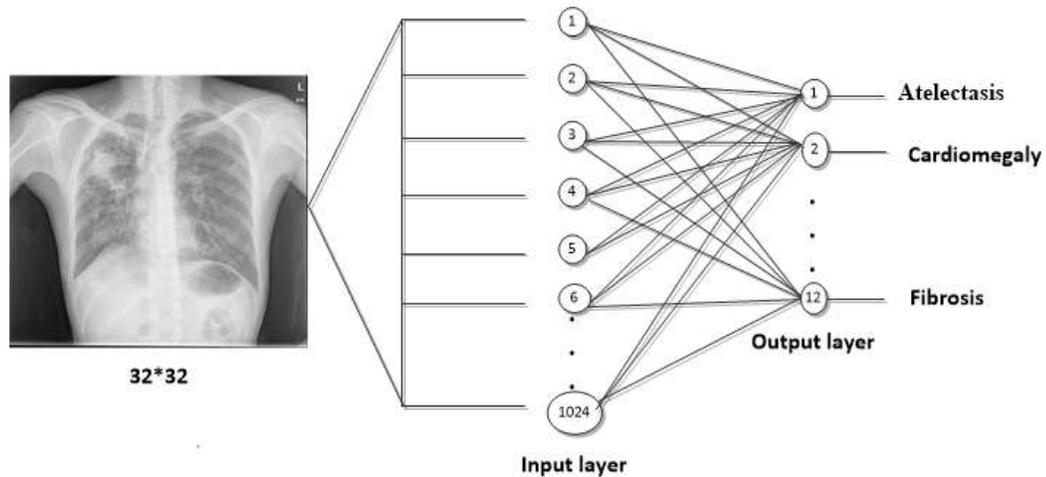
The learning curve for BPNN2, which is the network with lowest achieved MSE is shown below in Figure 5.2.

**Figure 5.2:** Learning curve for BPNN2

### 5.3.2 Competitive neural network (CpNN)

In this paper, an unsupervised learning algorithm using competitive neural network was also considered, leveraging on the fact that such networks do not need manual labelling of training data, hence saves considerable time and cost.

Figure 5.3 shows the architecture of the network as used in this paper.



**Figure 5.3:** Competitive neural network

The competitive neural network has only two layers as can be seen in the figure above, the input and output layers. The processed images are fed as inputs to the network, and the output neurons learn unique attributes or patterns in the images that differentiates one class of images from the other. The number of input neurons remain 1024 (input image pixels), and the number of output neurons 12 (number of output classes).

The training parameters for the network as used in this paper can be seen in Table 5.3 below.

**Table 5.3:** Training parameters for competitive neural network (32×32 input pixels)

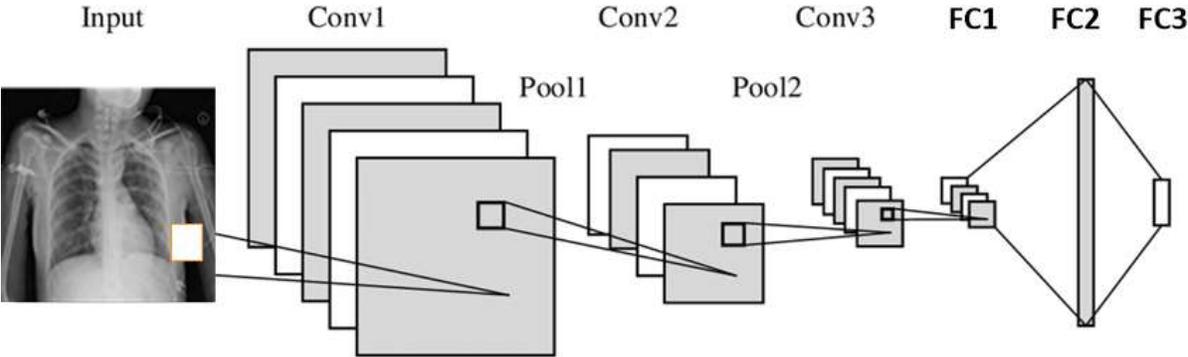
Networks	CpNN1	CpNN2	CpNN3
Training samples	<b>620</b>	620	620
Learning rate	<b>0.0036</b>	0.05	0.1
Maximum Epochs	<b>1000</b>	2000	4000
Training time (secs)	<b>300 secs</b>	434 secs	468

The table shows the different competitive networks designed and trained in this study using  $32 \times 32$  pixels images as inputs. Since the network uses an unsupervised learning algorithm, hence, there is no mean squared error goal to minimize or mean squared error goal.

**5.3.3 Convolutional neural networks**

In this section, the model architecture and training parameters for the convolutional network employed to learning the chest X-ray medical conditions are shown. The suitable learning parameters are determined through extensive experiments for the network optimization purposes. Note that out of the obtained 120,120 images, 70% are used for training and 30% are used for validating the trained network.

Table 5.4 shows the learning parameters of the CNN; where, “Conv” represents a convolution layer, “BN” represents batch normalization, “FM” represents feature maps and “FC” represents fully connected layer. Note that all a convolution filter of size  $3 \times 3$  is used in all convolution operations with padding; while all pooling operations are performed using max pooling windows of size  $2 \times 2$ . Moreover, the input images of the network are of size  $32 \times 32$ . Note that the architecture of this network in addition to the number of filters and their sizes were all based on the typical architecture of the CNN (Figure 5.4).



**Figure 5.4:** Convolutional neural network

The size of available training data and used system specifications for constructing a learning model were also considered. Thus, dropout training schemes and a batch normalization and were employed; which have been shown to improve model generalization (Abiyev and Maa'itah, 2018; Helwan et al., 2018). Note that a mini-batch optimization of size 100 via gradient descent is employed. In addition, a learning rate of 0.001 and 40000 iterations are used for training the CNN model.

**Table 5.4:** CNN training parameters

<b>Layers</b>	<b>Description</b>	<b>Values</b>
Input layer	Input image	32x32x1 images with 'zerocenter' normalization
Hidden Layer 1	Conv1 + BN+ReLU	16 feature maps of size 10×10
Hidden Layer 2	Pool1 Conv2 + BN+ReLU	2×2 kernel size with stride of 2 32 feature maps of size 10×10
Hidden Layer 3	Pool2 Conv3 + BN+ReLU	2×2 kernel size with stride of 2 64 feature maps of size 10×10
Classification layer	FC Softmax	2 fully connected layers 12 units

### 5.3 Results Discussion

The overall proposed networks for chest X-rays diseases classification are tested using 30% of the data. Table 5.5 shows the recognition rates obtained for the backpropagation networks using 32×32 pixels as the input image size.

**Table 5.5:** Recognition rates for BPNNs on training and validation data (32×32 pixels)

<b>Network models</b>	<b>Training data (70%)</b>	<b>Validation data (30%)</b>
BPNN1	92.74%	87.42%
<b>BPNN2</b>	<b>99.19%</b>	<b>89.57%</b>
BPNN3	97.32%	84.36%
BPNN4	98.10%	85.24%

It can be seen from the table above that though all the trained backpropagation neural networks (BPNNs) have different performance on both the training and test databases. BPNN2 achieved the highest recognition rate on both the training and test data compared to the other networks, i.e., 99.19% and 89.57%, respectively.

Competitive neural networks which rely on an unsupervised learning algorithm were also trained and tested in this work for the same classification task. These networks are faster to train considering that they have no desired outputs and therefore no error computations and back pass of error gradients for weights update, as it obtains in the back propagation networks.

The tables describing the simulation results of the competitive networks for different learning rate and number of maximum epochs are given below as Table 5.6.

**Table 5.6:** Recognition rates for CpNNs on training and validation data (32×32 pixels)

<b>Network models</b>	<b>Training data (70%)</b>	<b>Validation data (30%)</b>
CpNN1	84.21%	81.40%
<b>CpNN2</b>	<b>85.23%</b>	<b>84.71%</b>
CpNN3	86.57%	76.25%

From the table above, it can be seen that CpNN2 has the highest recognition rates on both the training and test data when the trained networks were simulated. Furthermore, it can be seen that CpNN3, though, has a higher recognition rate than CpNN2 on the training data, its performance on the test data is lower compared to CpNN2. i.e. it can be stated that CpNN3 has lower generalization power as compared to CpNN2.

Furthermore, the convolutional neural network (CNN) designed for this classification task is also tested using 30% of the available chest X-ray images and the results are shown in Table 5.7.

**Table 5.7:** Recognition rates for CNNs on training and validation data (32×32 pixels)

<b>Network model</b>	<b>Training data (70%)</b>	<b>Validation data (30%)</b>
CNN	100%	92.4%

Overall, the performance of the three employed networks in terms of recognition rate, training time, and reached mean square error (MSE) is described in Table 5.8.

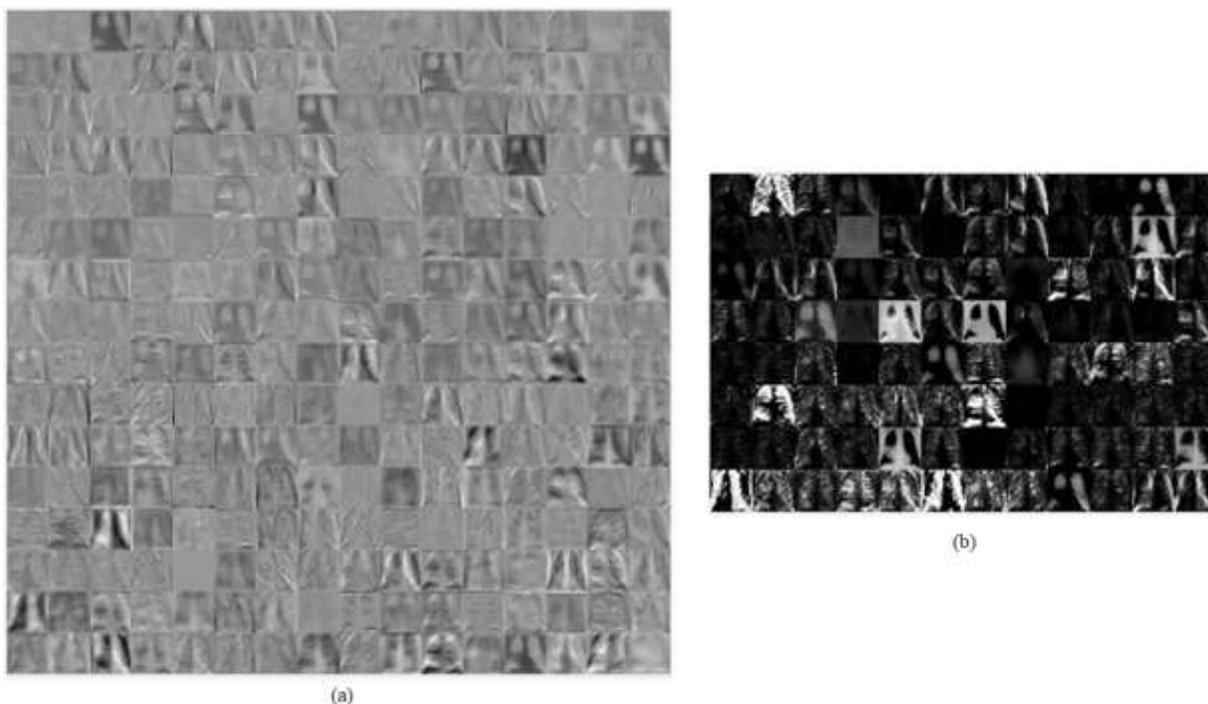
**Table 5.8:** Performances of the BPNN, CpNN and CNN

<b>Network models</b>	<b>Training time</b>	<b>Recognition rate</b>	<b>Reached MSE</b>	<b>Maximum number of iterations</b>
BPNN2	630 secs	80.04%	0.0025	5000
CpNN2	300 secs	89.57%	0.0036	1000
CNN	2500 secs	92.4%	0.0013	40000

As seen in Table 5.5 and Table 5.6, the networks behave differently during training and testing, and this is obviously due to the difference in the structures, working principles, and training algorithms of the three employed networks. Also, as seen in Table 5.8, the CNN has achieved the highest recognition rate during training and testing compared to other employed networks. In contrast, this outperformance of CNN over other networks requires longer time and larger number of iterations than that of BPNN2 and CpNN2. Moreover, it is seen that the three networks have achieved a low MSE whereas; the CNN scored the lowest (0.0013). Furthermore, it is noted that time needed for the CNN to converge is roughly higher than that of BPNN2 and CpNN2. Consequently, this is due to the deeper structure of a convolutional neural network which normally requires long time, in particularly, when the number of inputs is large. Nonetheless, this deep structure is the main factor of achieving a higher recognition

rate compared to other shallow networks such as BPNN and recurrent networks such as CpNN.

Figure 5.5 shows the learned filters or kernels at convolution layer 1 (a) and at pooling layer (b) of the convolution neural network (CNN). This shows the extraction of different levels of features of chest X-ray images in both convolution and pooling layer 1.



**Figure 5.5:** Learned filters: (a) Convolution layer 1, (b) Pooling layer 1

A comparison of the developed networks with some earlier works is shown in Table 5.8. Firstly, it is seen that shallow (traditional) networks (BPNN and CpNN) couldn't achieve high recognition rates compared to other deep networks, which is obviously due to their deficiency in extracting the important features from input images. Moreover, it is noticed that the proposed deep convolutional neural network (CNN) is achieved a higher recognition rate than other earlier research work such as CNN with GIST features (Bar et al., 2015).

The transfer learning based networks are also used for chest X-rays classification such as VGG16 (Islam et al., 2017) and VGG19 (Islam et al., 2017). They have gained lower generalization capabilities compared to the proposed network. Although, these pre-trained models have a very powerful features extraction capabilities as they were trained using a huge database, ImageNet (Russakovsky et al., 2012). Note that we compared the researches that provided explicitly achieved accuracies. The obtained results can show that by applying deep CNNs to the problem of chest X-rays diseases is promising, in a way that similar or confusing diseases could be correctly classified with good recognition rates.

**Table 5.8:** Results comparison with earlier works

Parameters	CNN	BPNN2	CpNN2	CNN with GIST(Bar et al., 2015)	VGG16(Islam et al., 2017)	VGG19 (Islam et al., 2017)
Number of images	120,120	1000	1000	637	8100	8100
Accuracy	92.4%	80.04%	89.57%	92%	86%	92%

## **CHAPTER 6**

### **CONCLUSION AND FUTURE WORKS**

#### **6.1 Conclusion**

In this thesis, deep learning represented by convolutional neural network (CNN), shallow network represented by backpropagation neural network (BPNN), and competitive neural network (CpNN) is carried out for the classification of the chest X-ray diseases. The designed CNN, BPNN and CpNN were trained tested using the chest X-ray images containing different diseases. Several experiments were also carried out during training of these networks using different learning parameters and number of iterations. In both backpropagation and competitive networks, it was observed that input image of size  $32 \times 32$  pixels produced a good performances on the achieved recognition rates; and the average, the backpropagation networks outperformed the competitive networks on recognition rates. Moreover, the competitive networks did not require manual labelling of training data as it was carried out for the backpropagation network; hence, it can be as a compensation for the lower classification accuracy obtained from the networks.

Furthermore, A CNN was also trained and tested on the same data used for training and testing the BPNN and CpNN. After convergence, it was noticed that the CNN was capable of gaining a better generalization power than that achieved by BPNN and CpNN, although required computation time and number of iterations was roughly higher. This outperformance is mainly due to the deep structure of CNN which grants it the power of extracting different level features, which results in better generalization capability.

#### **6.2 Future Recommendations**

In medical image analysis and processing, a most common issue is that the number of available data for research purposes is limited and small. Hence, training a fully deep network

structure like CNN with small number of data may result in Overfitting, which is usually the reason of low performance and generalization power.

Transfer learning is one solution of this problem, by sharing the learned parameters of effective and well-trained networks on a very large dataset. The concept of transfer learning is the use of a pre-trained model that is already trained on large datasets, and transfer its pre-trained learning parameters, in particular weights, to the target network model. The last fully connected layers are then trained with initial random weights on the new dataset. Note that, although the dataset is different than the on the network was trained on, the low-level features are similar. Thus, the parameter's transfer of the pre-trained model may provide the new target model with a powerful feature extraction capability and reduce its training computations speed and memory cost.

## REFERENCES

- Abadi, M., Chu, A., Goodfellow, I., McMahan, H.B., Mironov, I., Talwar, K., Zhang, L. (2016). Deep Learning with Differential Privacy. In *Proceedings of the International ACM SIGSAC Conference on Computer and Communications Security* (pp. 308–318). Vienna, Austria.
- Abas, A. R. (2013). Adaptive competitive learning neural networks. *Egyptian Informatics Journal*, 14(3), 183-194.
- Abiyev, R. H., & Ma'aitah, M. K. S. (2018). Deep Convolutional Neural Networks for Chest Diseases Detection. *Journal of healthcare engineering*, 3(5), 23-34.
- Ashizawa, K., Ishida, T., MacMahon, H., Vyborny, C. J., Katsuragawa, S., Doi, K., et al. (2005). Artificial neural networks in chest radiography: Application to the differential diagnosis of interstitial lung disease. *Academic Radiology*, 11(1), 29–37.
- Avendi, M. R., Kheradvar, A., & Jafarkhani, H. (2016). A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Medical image analysis*, 30(7), 108-119.
- Avetisian, M. Volumetric Medical Image Segmentation with Deep Convolutional Neural Networks.
- Balakrishnan, M. Mesbah, G. Molenberghs. (2013). *Advances in Statistical Methods for the Health Sciences*. Boston: Birkhäuser.
- Bar, Y., Diamant, I., Wolf, L., Lieberman, S., Konen, E., & Greenspan, H. (2015). Chest pathology detection using deep learning with non-medical training. In *Proceedings of the IEEE 12th International Symposium on Biomedical Imaging (ISBI)* (pp. 294-297). Berlin: Germany.

- Barreto, G. A., Mota, J. C. M., Souza, L. G. M., Frota, R. A., Aguayo, L., Yamamoto, J. S., & Macedo, P. E. (2004). Competitive neural networks for fault detection and diagnosis in 3G cellular systems. *Lecture notes in computer science*, 2(8) 207-213.
- Bengio, Y., Lamblin, P., Popovici, D., & Larochelle, H. (2007). Greedy layer-wise training of deep networks. *Advances in neural information processing systems*, 3(6), 153-160.
- Candemir, S., Jaeger, S., Palaniappan, K., Musco, J. P., Singh, R. K., Xue, Z., & McDonald, C. J. (2014). Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration. *IEEE transactions on medical imaging*, 33(2), 577-590.
- Cernazanu-Glavan, C., & Holban, S. (2013). Segmentation of bone structure in X-ray images using convolutional neural network. *Advanced Electrical and Computer Engineering*, 13(1), 87-94.
- Cilimkovic, M. (2015). Neural networks and back propagation algorithm. Institute of Technology Blanchardstown, Blanchardstown Road North Dublin, 15.
- Cohen, M. A., & Grossberg, S. (1983). Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. *IEEE transactions on systems, man, and cybernetics*, (5), 815-826.
- Demner-Fushman, D., Antani, S., Simpson, M., & Thoma, G. R. (2012). Design and development of a multimodal biomedical information retrieval system. *Journal of Computing Science and Engineering*, 6(2), 168-177.
- Demner-Fushman, D., Kohli, M. D., Rosenman, M. B., Shooshan, S. E., Rodriguez, L., Antani, S., & McDonald, C. J. (2015). Preparing a collection of radiology examinations for distribution and retrieval. *Journal of the American Medical Informatics Association*, 23(2), 304-310.

- Dosovitskiy, A., Fischer, P., Ilg, E., Hausser, P., Hazirbas, C., Golkov, V., & Brox, T. (2015). FlowNet: Learning optical flow with convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2758-2766). Santiago: Chile.
- El-Solh, A. A., Hsiao, C. B., Goodnough, S., Serghani, J., and Grant, B. J. B. (1999). Predicting active pulmonary tuberculosis using an artificial neural network. *Chest*, *116*, 968–973.
- Er, O., Yumusak, N., & Temurtas, F. (2010). Chest diseases diagnosis using artificial neural networks. *Expert Systems with Applications*, *37*(12), 7648-7655.
- Er, O., Yumusak, N., & Temurtas, F. (2010). Chest diseases diagnosis using artificial neural networks. *Expert Systems with Applications*, *37*(12), 7648-7655.
- Garg, R., BG, V. K., Carneiro, G., & Reid, I. (2016, October). Unsupervised cnn for single view depth estimation: Geometry to the rescue. In *Proceedings of the International European Conference on Computer Vision* (pp. 740-756). Amsterdam: The Netherlands.
- Godard, C., Mac Aodha, O., & Brostow, G. J. (2017, July). Unsupervised monocular depth estimation with left-right consistency. In *Proceedings of the International Conference on Computer Vision and Pattern recognition* (pp. 7-14). Honolulu: Hawaii.
- He, K., Zhang, X., Ren, S., Sun, J. (2015). Deep residual learning for image recognition. (2015). In *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770-778). Hynes Convention Center: Massachusetts.
- Helwan, A., El-Fakhri, G., Sasani, H., & Uzun Ozsahin, D. (2018). Deep networks in identifying CT brain hemorrhage. *Journal of Intelligent & Fuzzy Systems*, (Preprint), 1-1.
- Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, *18*(7), 1527-1554.

- Islam, M. T., Aowal, M. A., Minhaz, A. T., & Ashraf, K. (2017). Abnormality Detection and Localization in Chest X-Rays using Deep Convolutional Neural Networks. *arXiv preprint arXiv:1705.09850*.
- Jaderberg, M., Simonyan, K., & Zisserman, A. (2015). Spatial transformer networks. In *Proceedings of the International Conference on Advances in neural information processing systems* (pp. 2017-2025). Montréal: Canada.
- Jaeger, S., Karargyris, A., Candemir, S., Folio, L., Siegelman, J., Callaghan, F., & Thoma, G. (2014). Automatic tuberculosis screening using chest radiographs. *IEEE transactions on medical imaging*, 33(2), 233-245.
- Jason, J. Y., Harley, A. W., & Derpanis, K. G. (2016, October). Back to basics: Unsupervised learning of optical flow via brightness constancy and motion smoothness. In *Proceedings of the International European Conference on Computer Vision* (pp. 3-10). Amsterdam: The Netherlands.
- Jennifer, G.D., Carla, B., Avi, K., Lynn, S.B., and Alex, M. (2003). Unsupervised feature selection applied to content-based retrieval of lung images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(3), 373–378.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 5(7), 1097-1105.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- Lee, J.G., Jun, S., Cho, Y.W., Lee, H., Kim, G.B., Seo, J.B., Kim, N. (2017). Deep Learning in Medical Imaging: General Overview. *Korean Journal of Radiology*, 18, 570–584.
- Liu, J., Pan, Y., Li, M., Chen, Z., Tang, L., Lu, C., Wang, J. (2018). Applications of deep learning to MRI images: A survey. *Big Data Mining Annals*, 1, 1–18.

- Mamoshina, P., Vieira, A., Putin, E., Zhavoronkov, A. (2016). Applications of Deep Learning in Biomedicine. *Molecular Pharmacology*, 13, 1445–1454.
- Miotto, R., Wang, F., Wang, S., Jiang, X., Dudley, J.T. (2017). Deep learning for healthcare: Review, opportunities and challenges. *Brief Bioinformatics*, 2(9), 12-25.
- Oyedotun, O. K., & Khashman, A. (2017). Deep learning in vision-based static hand gesture recognition. *Neural Computing and Applications*, 28(12), 3941-3951.
- Patil, S. A., & Kuchanur, M. B. (2012). Lung cancer classification using image processing. *International Journal of Engineering and Innovative Technology*, 2(3).
- Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., Yang, G.-Z. (2017). Deep learning for health informatics. *IEEE Journal of Biomedical Health Informatics*. 21, 4–21.
- Ren, Z., Yan, J., Ni, B., Liu, B., Yang, X., & Zha, H. (2017, February). Unsupervised Deep Learning for Optical Flow Estimation. In *Proceedings of the Thirty-First International Conference on Artificial Intelligence* (pp. 1-7). California: USA.
- Rios, A., & Kavuluru, R. (2015, September). Convolutional neural networks for biomedical text classification: application in indexing biomedical articles. In *Proceedings of the 6th ACM Conference on Bioinformatics, Computational Biology and Health Informatics* (pp. 258-267). Atlanta: USA
- Shan, S., Guo, X., Yan, W., Chang, E. I., Fan, Y., & Xu, Y. (2017). Unsupervised End-to-end Learning for Deformable Medical Image Registration. *arXiv preprint arXiv:1711.08608*.
- Shiraishi, J., Katsuragawa, S., Ikezoe, J., Matsumoto, T., Kobayashi, T., Komatsu, K. I., & Doi, K. (2000). Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. *American Journal of Roentgenology*, 174(1), 71-74.

- Simonyan, K., Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv:1409.1556.
- Song, J., Qin, S., Zhang, P. (2016). Chinese text categorization based on deep belief networks. In *Proceedings of the 15th International Conference on Computer and Information Science Computer and Information Science (ICIS)*, (pp. 26–29). Okayama: Japan.
- Srinivas, M., Naidu, R., Sastry, C.S., and Mohan, C.K. (2015). Content based medical image retrieval using dictionary learning. *Neurocomputing*, 168, 880–895.
- Suzuki, K. Overview of deep learning in medical imaging. (2017). *Radiological Physics and Technology*, 10, 257–273.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 1-9). Hynes Convention Center: Massachusetts.
- Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017, July). Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 3462-3471). Honolulu: Hawaii.
- Wei, J., He, J., Chen, K., Zhou, Y., Tang, Z. (2017). Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems*, 69, 29–39.
- Xue, Z., You, D., Candemir, S., Jaeger, S., Antani, S., Long, L. R., & Thoma, G. R. (2015). Chest x-ray image view classification. In *Proceedings of the 28th International Symposium on Computer-Based Medical Systems (CBMS)* (pp. 66-71). Brazil: Ribeirão Preto.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.

## **APPENDICES**

**APPENDIX 1**  
**BACKPROPAGATION NEURAL NETWORK SOURCE CODE**

```
% Solve a Pattern Recognition Problem with a Neural Network
```

```
% Script generated by NPRTOOL
```

```
%
```

```
% This script assumes these variables are defined:
```

```
%
```

```
inputs = input_data;
```

```
targets = Target_data;
```

```
% Create a Pattern Recognition Network
```

```
numHiddenNeurons = 35;
```

```
net.trainParam.lr=0.00045;
```

```
net.trainParam.mc=0.0072;
```

```
net = newpr(inputs,targets,numHiddenNeurons);
```

```
% Train the Network
```

```
[net,tr] = train(net,inputs,targets);
```

```
% simulate network on train data
```

```
%target matrix t
```

```
%target max indices
```

```
target=input_target;
```

```
[M,I_t]=max(target);% row vector
```

```
%dimensions of target matrix
```

```
[u,v]=size(target);
```

```

%actual output matrix sim_net
sim_in=input_train;
sim_net=sim(net,sim_in);
[N,I_sim_net]=max(sim_net);% row vector
%comparison of target and actual outputs
result = I_t==I_sim_net;% row vector
%sum of all elements,1s, to know how many corrects
corrects=sum(result);
%recognition rate,
w=double(corrects*100/v); %let recognition rate be w
fprintf('recognition rate on train data is %d\n',w);
% simulate network on test data
%target matrix t

%target max indices
target=test_expected_output;
[M,I_t]=max(target);% row vector
%dimensions of target matrix
[u,v]=size(target);

%actual output matrix sim_net
sim_in=test_input
sim_net=sim(net,sim_in);
[N,I_sim_net]=max(sim_net);% row vector
%comparison of target and actual outputs
result = I_t==I_sim_net;% row vector
%sum of all elements,1s, to know how many corrects
corrects=sum(result);

```

```
%recognition rate,  
w=double(corrects*100/v); %let recognition rate be w  
fprintf('recognition rate for test data is %d\n',w);  
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

## APPENDIX 2

### COMPETITIVE NEURAL NETWORK SOURCE CODE

```
% Create inputs X.
bounds = [0 1; 0 1]; % Cluster centers to be in these bounds.
clusters = 8; % This many clusters.
points = 10; % Number of points in each cluster.
std_dev = 0.05; % Standard deviation of each cluster.
x = nngenc(bounds,clusters,points,std_dev);

% Plot inputs X.
plot(x(1,:),x(2:,:),'+r');
title('Input Vectors');
xlabel('x(1)');
ylabel('x(2)');

net = competlayer(8,.1);
net = configure(net,x);
w = net.IW{1};
plot(x(1,:),x(2:,:),'+r');
hold on;
circles = plot(w(:,1),w(:,2),'ob');

x1 = [0; 0.2];
y = net(x1)
%%%%%%%%%%%%%%

net = competlayer(2);

%Now you have a network, but you need to train it to do the classification job.

%The first time the network is trained, its weights will initialized to the centers of the input
ranges with the %function midpoint. You can check see these initial values using the number
of neurons and the input %data:

wts = midpoint(2,p)
biases = initcon(2)

%Training

%Now train the network for 500 epochs. You can use either train or adapt.
net.trainParam.epochs = 500;
```

```
net = train(net,p);
```

%Note that train for competitive networks uses the training function trainru. You can verify this %by executing the following code after creating the network.

```
net.trainFcn
```

```
a = sim(net,p);
```

```
ac = vec2ind(a)
```

```
net.IW{1,1}
```

### APPENDIX 3

## CONVOLUTIONAL NEURAL NETWORK SOURCE CODE

```
dataFolder = 'D:\data1\ChestXray-NIHCC';
imds = imageDatastore(dataFolder,...
    'IncludeSubfolders',true,'LabelSource','foldernames');
    % creating datastore with files from my folder
trainingNumFiles = 12;
rng(1) % For reproducibility
% categories = {'Atelectasis', 'Cardiomegaly','Effusion', 'Infiltration','Mass',
'Nodule','Pneumonia', 'Pneumothorax','Consolidation', 'Edema','Emphysema', 'Fibrosis'};
% imds = imageDatastore(fullfile(dataFolder, categories), ...
%   'LabelSource', 'foldernames');
% % imds.ReadFcn = @(imds)imresize(imread(imds),[1024 1024])

[trainingImages,validationImages] = splitEachLabel(imds,0.8,'randomized');
% if iscell( imds )
%   try
%       data = cat(4, imds{:});
%   catch e
%       throwVariableSizesException(e);
%   end
% else
%   data = imds;
% end
%%%%%%%%%%
%%%%%%%%%%
numTrainImages = numel(trainingImages.Labels);
idx = randperm(numTrainImages,16);
```

```

figure
for i = 1:16
    subplot(4,4,i)
    I = readimage(trainingImages,idx(i));
    imshow(I)
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
numClasses = numel(categories(trainingImages.Labels));
% layers = [
%   imageInputLayer([1024 1024 1])
%
%   convolution2dLayer(10,16,'Padding',1)
%   batchNormalizationLayer
%   reluLayer
%
%   maxPooling2dLayer(2,'Stride',2)
%
%   convolution2dLayer(10,32,'Padding',1)
%   batchNormalizationLayer
%   reluLayer
%
%   maxPooling2dLayer(2,'Stride',2)
%
%   convolution2dLayer(10,64,'Padding',1)
%   batchNormalizationLayer
%   reluLayer
%   fullyConnectedLayer(12)
% %   fullyConnectedLayer(74935)

```

```

% softmaxLayer
% classificationLayer];
layers = [ ...
    imageInputLayer([1024 1024 1],'Normalization','none')
    convolution2dLayer(6,20)
    reluLayer
    maxPooling2dLayer(2,'Stride',2)
    fullyConnectedLayer(12)
    softmaxLayer
    classificationLayer];

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
options = trainingOptions('sgdm',...
    'MaxEpochs',1, ...
    'ValidationData',validationImages,...
    'ValidationFrequency',30,...
    'Verbose',false,...
    'Plots','training-progress');
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
net = trainNetwork(trainingImages,layers,options);

predictedLabels = classify(net,validationImages);
valLabels = validationImages.Labels;

accuracy = sum(predictedLabels == valLabels)/numel(valLabels)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

% Get the network weights for the second convolutional layer
w1 = net.Layers(2).Weights;

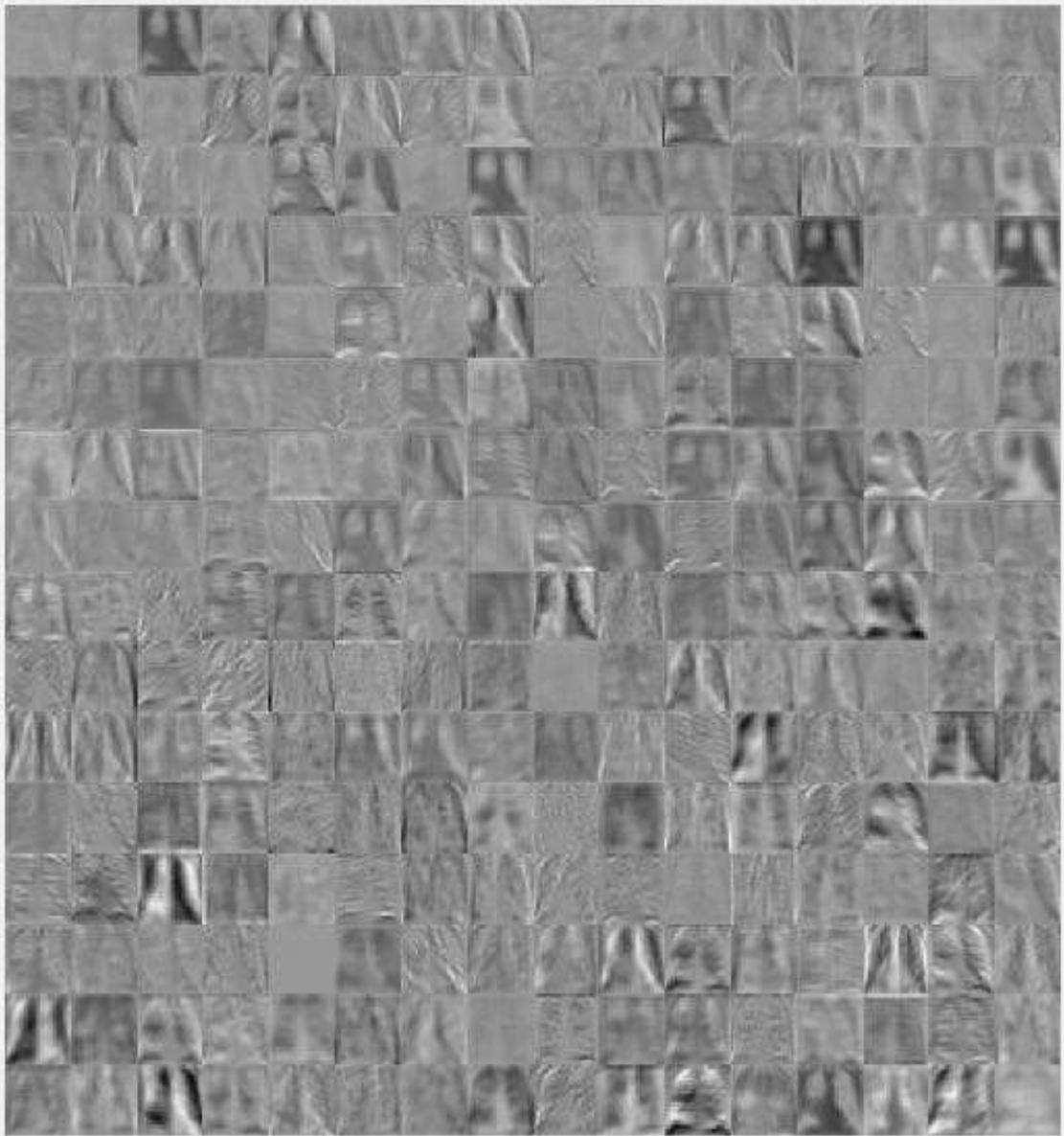
% Scale and resize the weights for visualization
w1 = mat2gray(w1);
w1 = imresize(w1,5);

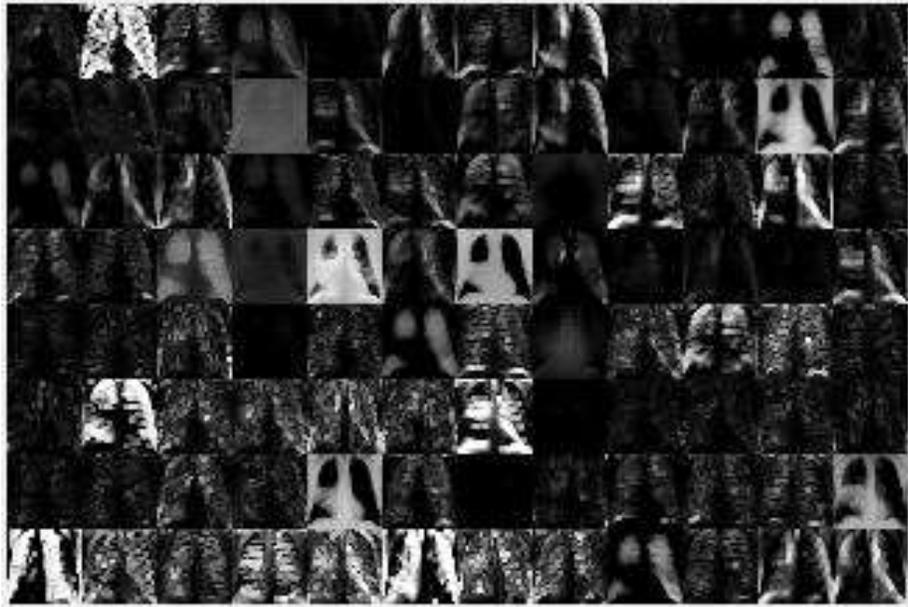
% Display a montage of network weights.
figure
montage(w1)
title('First convolutional layer weights')
%
%%%%%%%%%%
%%%%%%%%%%
%
% im = imread('H2.jpg');
% imshow(im)
% imgSize = size(im);
% imgSize = imgSize(1:2);
% net.Layers
% act1 = activations(net,im,'pool1','OutputAs','channels');
% sz = size(act1);
% act1 = reshape(act1,[sz(1) sz(2) 1 sz(3)]);
% montage(mat2gray(act1),'Size',[8 12])
% pause
% act1ch32 = act1(:, :, :, 32);
% act1ch32 = mat2gray(act1ch32);
% act1ch32 = imresize(act1ch32,imgSize);
% imshowpair(im,act1ch32,'montage')

```

```
% pause
% act5 = activations(net,im,'conv1','OutputAs','channels');
% sz = size(act5);
% act5 = reshape(act5,[sz(1) sz(2) 1 sz(3)]);
% montage(imresize(mat2gray(act5),[48 48]))
```

**APPENDIX 4**  
**LEARNED FILTERS VISUALIZATIONS**





**APPENDIX 5**  
**CURRICULM VITAE**

