IMPROVED LUNG PATTERN CLASSIFICATION FOR INTERSTITIAL LUNG DISEASE USING DEEP LEARNING

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE
OF

MASTER OF TECHNOLOGY
IN
SIGNAL PROCESSING AND DIGITAL DESIGN

Submitted by:

KRITI PODDAR

2K17/SPD/05

Under the supervision of

DR. S. Indu

DEPARTMENT OF ELECTRONICS AND COMMUNICATION
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

JUNE,2019

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, Kriti Poddar, 2K17/SPD/05, of M.Tech, hereby declare that the project Dissertation Titled "Improved Lung Pattern Classification for Interstitial Lung Disease using Deep Learning" which is submitted by me to the Department of Electronics and Communication, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Place: Delhi (**KRITI PODDAR)**

Date:

ELECTRONICS AND COMMUNICATION ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation titled "Improved lung pattern classification for interstitial lung disease using deep learning" which is submitted by Kriti Poddar, Roll No. 2K17/SPD/05, Electronics and Communication Engineering, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the students under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: Delhi

Date:

DR. S. INDU

SUPERVISOR

Professor

Department of Electronics and Communication

Delhi Technological University, Delhi

ACKNOWLEDGEMENTS

My deepest gratitude is to my advisor, Dr. S. Indu. I have been amazingly fortunate to have an advisor who gave me the freedom to explore on my own, and at the same time the guidance to recover when my steps faltered. His patience and support helped me to finish this dissertation.

I would like to thank my family and friends and the one above all of us, the omnipresent God for the support they provided.

Finally, I would like to thank family of DTU.

KRITI PODDAR

ABSTRACT

In this research work, methods for classification and characterization of a computer aided diagnosis (CAD) system for interstitial lung disease have been executed. Although a lot of research has already been done in this regard, the growing popularity of the Deep learning techniques have evoked expectations that they might be applied in the field of image analysis as well. In the present work, a lot of methodologies have been applied on the chosen dataset and their results have been evaluated post which a network has been proposed. It consists of a layer of segmentation deploying the UNET architecture which is then connected and followed by a layer similar to that of CNN but with little modifications. It essentially consists of 5 convolutional layers and ReLU activations, followed by maxpooling which is then followed by 3 fully connected layers and a softmax layer. The last dense layer has got 5 outputs which form the classes to be considered. These are: healthy, ground glass opacity, emphysema, fibrosis and micronodules. In order to train and evaluate various network methodologies, a dataset of approximately 18400 image patches have been taken. A comparative analysis established the effectiveness of the proposed framework against all the other methods in a challenging dataset.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

ILD, that is, Interstitial Lung Disease essentially covers a lot of diseases related to the lungs which cause stiffening of the lungs and therefore makes it difficult to breathe. These diseases affect the interstitium of the lungs, that is a portion of the lung's anatomy. Interstitium is a network of lace like features consisting of tissues that extends throughout the area of the lungs. There are tiny blood vessels travelling to the interstitum which allow exchange of gas between blood and air that are present throughout the lungs. Normally, the fact is that interstitium is very thin and minute in structure to have caught notice in X-rays or CT scans. ILDs cause thickening of the interstitium. Reasons of thickening could be many such as scarring, inflammation or it could even be extra fluid(that is, edema). Other reasons can be exposure to harmful chemicals, medical treatments or due to certain medications. The causes may even be unknown for many cases. The damage caused by ILDs is often irreversible and progressive implying that it gets worse with time. However, certain medications can slow the deterioration, but, in other cases , lung transplant is generally recommended.

Some examples of ILDs include:

- Interstitial pneumonia: Interstitium can be affected by virus, fungi or bacteria. Mycoplasma pneumonia is the common bacteria that causes it.
- Idiopathic Pulmonary Fibrosis: A chronic and progressive form of scarring the reason for which stands unknown.

- Hypersensitivity Pneumonitis: An ILD caused due to inhaling dust, mold or any other irritants over quite a long period of time.

  Cryptogenic organizing pneumonia(COP): An ILD which similar to the condition of pneumonia but without any infection present. It is also called bronchiolitis obliterans with organizing pneumonia(BooP).

  Acute interstitial pneumonitis: This is a very severe ILD requiring life support.

  Desquamative interstitial pneumonitis: An ILD which is partially caused by smoking.
- Sarcoidosis: It is a condition causing ILD along with swollen lymph nodes and may sometimes even involve eye, heart, skin or nerve.
- Asbestosis: ILD caused due to exposure to asbestos.

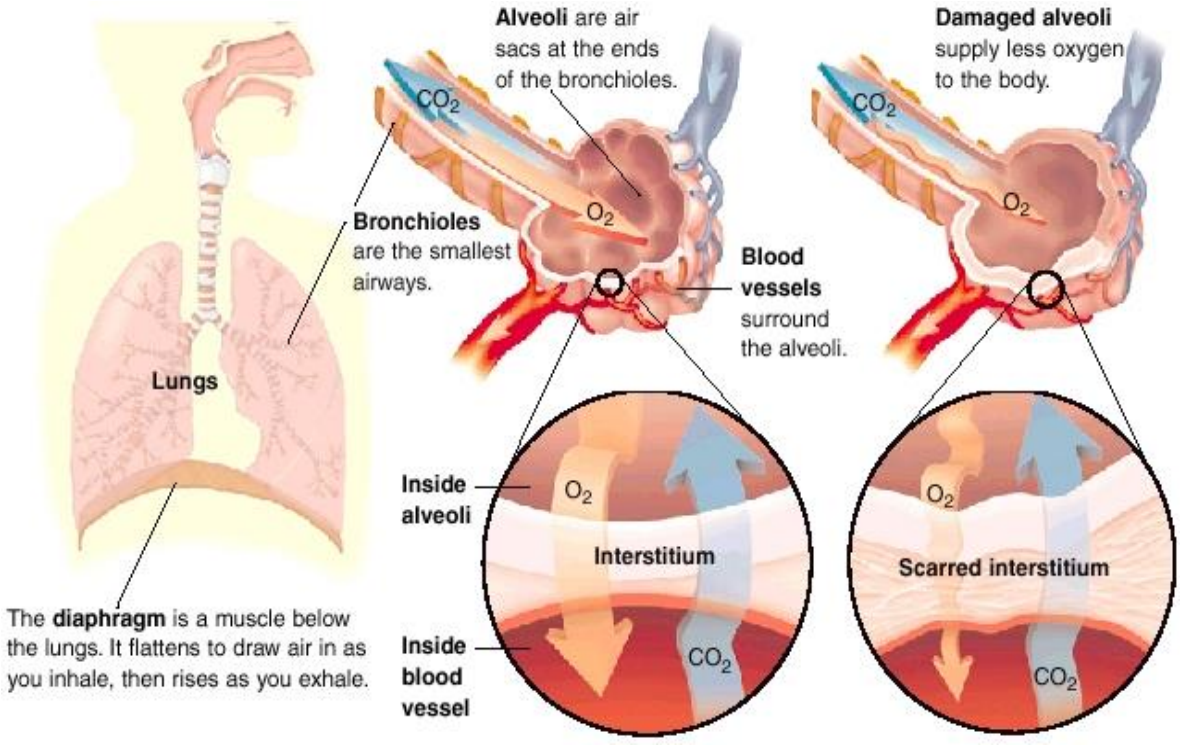


Fig 1.1. Lungs and some causes of ILD

1.1.1 Risk Factors:

Factors which can make one more prone to developing ILD includes:

- **Age**: Although this disease is more likely to affect adults, children may also develop this disorder.
- **Exposure to occupational and environmental toxins**: Working in mines, construction or even farming or in other area which is exposed to pollutants have high chances of damaging lungs. The risk of developing interstitial lung disease is increased in these scenarios. Some of these toxins/pollutants include:
    - Silica dust
    - Asbestos fibers
    - Grain dust
    - Bird and animal droppings
    - Indoor hot tubs
- **Gastroesophageal reflux disease:** Uncontrolled acid reflux or indigestion may also lead to developing interstitial lung disease.
- **Smoking:** A person with a history of smoking is more likely to develop interstitial lung disease. Active smoking makes the condition even worse, especially if emphysema is associated with it.
- **Radiation and chemotherapy**: Getting radiation treatments to chest (such as in cases of lung or breast cancer) with chemotherapy drugs increases the likeliness of developing such disease.

The symptoms of the developing disease are visible months or years after the treatment of initial stage.

- There are many drugs which can also cause deterioration to the lungs like:
  - **Chemotherapy drugs:** These are drugs designed to kill cancer cells like methotrexate (**O**trexup, Trexall, others). Cyclophosphamide can also cause damage to the lung tissue.
  - **Heart medications:** Some drugs which are meant to treat heart problems like irregular heartbeats, such as, amiodarone (Nexterone, Pacerone) or propranolol (Inderal, Innopran), may also harm lung tissue.
  - **Some antibiotics**: Antibiotics like nitrofurantoin (Macrobid, Macrodantin, others) and ethambutol (Myambutol) can also cause damage to lungs.
  - **Anti-inflammatory drugs.** Certain anti-inflammatory drugs, such as rituximab (Rituxan) or sulfasalazine (Azulfidine), can cause lung damage.

- Lung damage can also result from autoimmune diseases like:
  - Scleroderma
  - Rheumatoid arthritis
  - Sarcoidosis
  - Sjogren's syndrome
  - Mixed connective tissue disease Dermatomyositis and polymyositis

Some drugs like amiodarone, nitrofurantoin, bleomycin can rarely cause interstitial lung disease. Sometimes the reason for ILD may be unknown altogether.

ILD can be developed by anyone, that is, men and women of any age can be affected.

1.1.2 Symptoms:

Symptoms of ILD primarily includes shortness of breath. Almost everyone who develops this disease is likely to suffer from breathlessness which becomes even worse with time. Other symptoms may include:

- Cough which is mostly dry and nonproductive
- Breathlessness found mostly in people with C**O**P or B**OO**P.

In more general cases of ILD, the problem of breathlessness develops slowly over months. In case of interstitial pneumonias or acute interstitial pneumonitis, symptoms develop more rapidly (maybe in hours or in days).

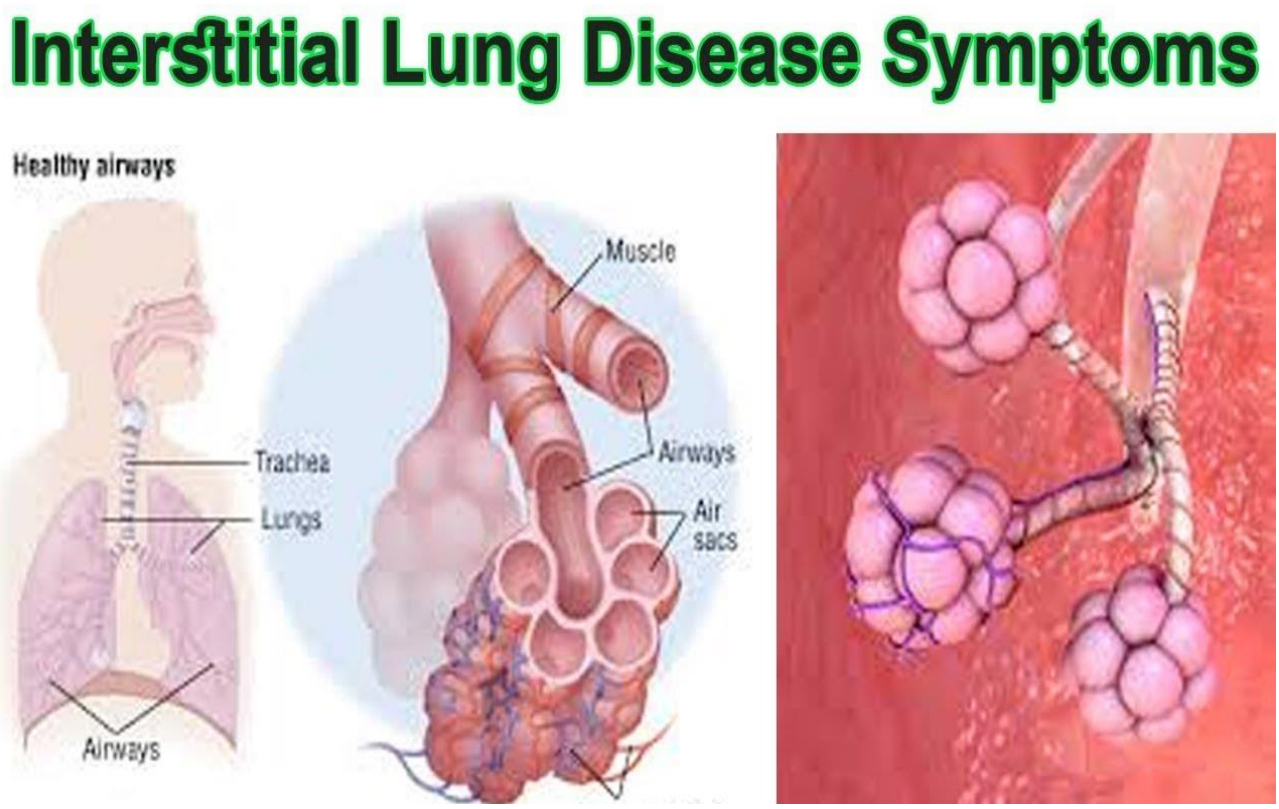


Fig 1.2. Symptoms of ILD

1.1.3 Complications:

ILDs may cause many life-threatening diseases to follow, like:

- **High blood pressure in lungs(pulmonary hypertension):** This condition is little different from the condition of normal blood pressure. Here, only the arteries in the lungs are affected. It starts when low level of oxygen or scar tissue restrict the smallest blood vessels, thus limiting flow of blood through the lungs. This leads to increase in pressure within pulmonary arteries. It is a very serious illness which becomes progressively worsens with time.

- **Right-sided heart failure (cor pulmonale):** As a matter of fact, lower right chamber, that is, right ventricle of heart is little less muscular than the one in the left side. A serious condition arises when the right ventricle of a person suffering with ILD has to pump harder than in the usual condition to propagate blood through the pulmonary arteries that are restricted. Eventually, due to extra strain, the right ventricle fails and this is a very serious consequence of pulmonary hypertension which in turn is a consequence of interstitial lung disease.

  Respiratory failure: When the end stage of chronic interstitial disease is reached, failure of respiration occurs as blood with rising blood pressure and low oxygen levels in the pulmonary arteries and right ventricle ultimately cause heart failure.

1.1.4 Diagnosis of Interstitial Lung Disease

People with ILD usually go to see a doctor due to the concern of breathlessness or constant cough. So, image tests of lungs are usually done in order to discover the problem.

Chest X-ray: It is the first or preliminary test in the diagnosis process of most people who come up with a problem of breathing. The X-ray films of the chest of such people usually shows lines which are fine in nature in the lungs.

Computed tomography (CT scan): A CT scan works by taking multiple images of the lungs and then combining these images to generate more detailed image of the lungs and the surrounding shapes and structures. ILDs are easy to be spotted on a CT scan.

High resolution CT scans: In case of ILD being suspected, improving the resolution or certain settings of the CT scanner improves the image of the interstitial. Hence this enhances the ability of CT scan to detect ILD.

Pulmonary function testing: The person suspected with ILD is made to sit in a plastic booth where he is asked to breathe through a plastic tube. The person due to lack of a good breath would have reduced total lung capacity and also would have a reduced ability to transfer oxygen from lungs into their blood.

Lung biopsy: Many times, examining a lung tissue only helps to determine which king of ILD the person is suffering from. There are various ways of examining lung tissue, called as lung biopsy.

- **Bronchoscopy:** In this, an endoscope is advanced through the mouth or nose into the airways which is capable of taking a sample of tissue of the lungs.
- **Video-assisted thoracoscopic surgery (VATS):** Small incisions are made and tools are inserted through them which can help a surgeon sample many areas of lung tissue.
- **Open lung biopsy (thoracotomy):** Few cases require proper treatment and surgery in which incision is made into the chest and lung biopsy is obtained.

1.1.5 Medications for ILD:

There are a handful of medications that can be prescribed to a person suffering with ILD.

- Oral Corticosteroids
- Immune suppressing or steroid sparing medications
- Anti-Fibrotic Medication
- Pirfenidone

- Nintedanib
- Oxygen Therapy
- Pulmonary Rehabilitation
- Lung Transplant: How far is ILD medication successful depends on the disease's response to medication. Some types of ILDs respond quickly while others may be too slow or may not respond at all. Symptoms, physiologic findings and X-ray, all need to be stabilized for treatment to be successful. Some ILDs progress even after treatment and become worse with ruining symptoms, physiological findings and X-ray. Sometimes, the symptoms worsen due to the treatment itself. This includes conditions such as pulmonary hypertension or heart failure. Other treatments may result in muscle weakness, osteoporosis and infection.

  If ILD progresses despite therapy, lung transplantation may be recommended as an option. It replaces one or two lungs affected with ILD with healthy lungs which generally comes from a non-living donor.

1.2 ORGANIZATION OF DISSERTATION/THESIS

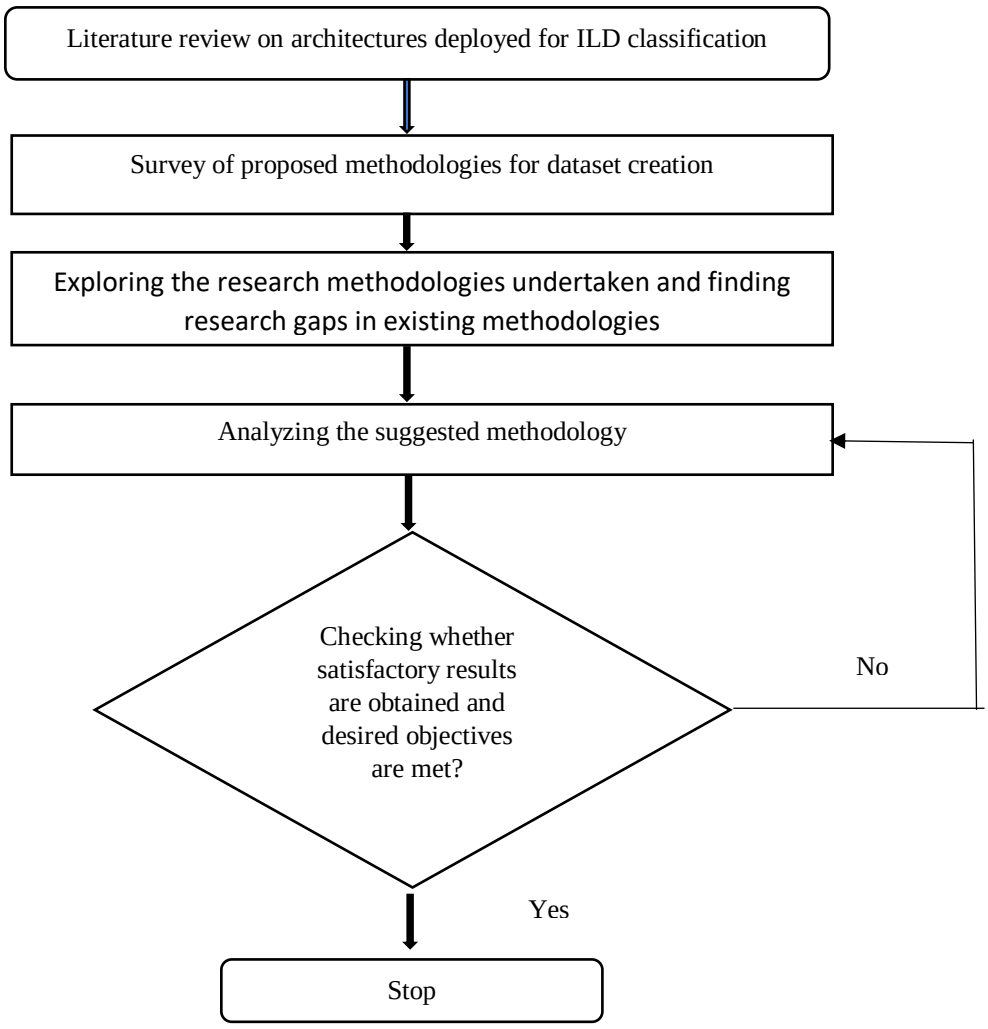


Fig 1.2 Thesis organization

1.2.1 Literature Review and Motivation:

Here, we intend to provide an overview of the studies that have been made on ILD pattern classification. We then move forward to convolutional neural network (CNN) which forms the basis for the proposed methodology [1].

Now, ILDs being generally considered as textural alterations in lung parenchyma, allowed for applying texture classification schemes on local regions of interest (ROIs) or volumes of interest (VOIs) again depending upon the 2D or 3D capabilities of the CT imaging modality employed. After pre-segmenting the lung fields and thereafter sliding a fixed-scale classifier over it, an ILD quantification map is expected to be generated which can then be used either by physicians or CAD systems to go forward for final diagnosis. Some of the main characteristics of such systems are:

1. chosen feature set and
2. the classification method.

Initially, the CAD systems for ILDs relied solely on classical feature extraction methods in order to describe the 2D textural alterations in the lung parenchyma. Some of these methods are first order gray level statistics, gray level co-occurrence matrices (GLCM), run-length matrices (RLM) and fractal analysis [2], [3]. Later, these features were merged and were collectively called as AMFM (Adaptive Multiple Feature Method).

AMFM could assess as many as 22 independent features that are textural in nature in order to classify pattern in a tissue [4]. In the related study, six tissue patterns were characterized like: honeycombing, ground glass, broncho vascular, nodular, emphysema like, and normal. Regional evaluation on lung slices were done involving 31x31 pixel regions of interest. In each of the regions of interest, evaluation was done on the most optimal subset of texture features to determine which of the six

characteristic patterns it resembled to. Validation of computer output was done against experienced observers in different settings. It was found that for regional tissue characterization, AMFM can be considered almost 100% reproducible and performs as well as experienced human observers who have been told the diagnosis of the patient.

AMFM was previously developed and demonstrated for computer-aided characterization of X-ray computed tomography lung slices. Training AMFM further can help to identify different lung tissue types using multiple texture features computed from the CT data. The previous methods which evaluated CT images of the lung drew information only from first-order textural feature like density. They, therefore did not exploit the complexity of lung parenchyma as shown and recorded in digital form by HRCT. AMFM was at first tested for differentiating between normal subjects and emphysematous subjects and it resulted to be superior as compared to all the previous methods. Following this, AMFM was further executed on a more complex four-subject group differentiation task involving normal subjects as well as subjects with emphysema, idiopathic pulmonary fibrosis (IPF), and sarcoidosis. AMFM performed better than the previous methods in this study as well. In these studies, AMFM was successfully able to recognize broad categories of diseases as opposed to more disease-independent tissue patterns.

AMFM was further developed to identify and differentiate normal lung pattern from parenchymal lung diseases. The patterns recognized by AMFM were commonly utilized in clinical practice to identify regional abnormalities in the lungs, thus, making AMFM acceptable as state of the art until new techniques for texture feature description were developed which gave new perspective to the problem of ILD identification. One of such systems tried to employ the filter bank approach to solve the problem of ILD identification and detection [5]. It is basically a CAD (computer aided diagnosis) system presented to automatically distinguish normal tissue from abnormal one in high resolution CT scans of chest acquired during daily clinical practice. From high resolution computed tomography scans of almost 116 patients, 657 regions of interest are extracted which either belong to the normal subjects or subjects with ILD. Here, a texture analysis approach which is principled in nature has been used which

tends to extract features to highlight local structures within the image with the help of a multi-scale filter bank. Later on, with help of various classifiers and feature subsets, the results are finally evaluated with ROC analysis.

Post filter bank approach, a new classification method was approached for five categories of lung tissues in HRCT images involving a feature based image patch approximation [6]. In order to increase the descriptiveness of the features involved, two new feature descriptors were designed namely the rotation invariant Gabor-local binary patterns (RGLBP) texture descriptor and multi-coordinate histogram of oriented gradients (MCHOG) gradient descriptor. Combining with the intensity features, labelling of each image patch is done based on its feature approximation from reference image patches. Following this, a new patch-adaptive sparse approximation (PASA) method is designed with some components like minimum discrepancy criteria for sparse-based classification, patch-specific adaptation for discriminative approximation, and feature-space weighting for distance computation. These markings are then cumulated as probabilistic estimations for region-level classification. The method is then evaluated on a publicly available ILD database which shows quite encouraging performance improvements over the other state-of -the-art methods.

Some of the feature extraction methods also used DCT-based filter bank [7] wherein after convolving the image with the filter bank, q-quantiles were computed for describing the distribution of the local frequencies that characterize the texture of the image. Then, final feature vector is constructed by adding the gray-level histogram values of the original image. For classifying the already described patches, random forest classifier is used.

Some methods made use of morphological filter in which the lungs were first segmented from the background using such a filter and a thresholding technique and then divided into many contagious regions of interest using a 32x32 matrix. Six physical measures were evaluated for each ROI which included the mean and the standard deviation of the CT value, air density components, line

components,nodular components,and multilocular components. Artificial neural networks were then employed to differentiate between seven different patterns including normal and other six patterns associated with diffuse lung disease. The method performed quite well with a sensitivity in each R**O**I being about 99.2% for ground glass opacities,100% for reticular lin ear opacities, 100% (98/98) for honeycombing, 88.0% (132/150) for nodular opacities, 95.8% (369/385) for emphysematous change, and 100% (43/43) for consolidation. The specificity in detecting a normal R**O**I was 88.1% (940/1067).

A novel algorithm was presented which classified diffuse lung disease into four patterns namely normal, emphysema, honeycombing and ground glass opacity by doing textural analysis of high resolution computed tomography images of lungs [8]. The algorithm incorporates two types of features. First, it uses scale-space features based on Gaussian derivative filters and second, it incorporates multidimensional multiscale features based on wavelet and contourlet transform of the original images. In order to model the output of filters and transforms and construct feature vectors, various statistical measures such as mean, standard deviation, and kurtosis along with generalized Gaussian density are used. The classifier used to evaluate the performance of the feature extraction scheme is a Multi-class multiple kernel learningo (m-MKL) classifier.

In another proposed method, a near-affine-invariant texture descriptor was derived from isotropic wavelet frames that characterized the patterns of lung tissue in high resolution computed tomography imaging [9]. Affine invariance is actually wished for as it enables to learn nondeterministic textures without a priori localizations, orientations, or sizes. After being combined with complementary gray level histograms,the method allows a global classificati on accuracy of 76.9% providing a balanced precision among the five classes of lung tissue.

Quantitative analysis of pulmonary emphysema using local binary patterns was also developed which aimed at improving quantitative measures of emphysema in CT images of the lungs

[10]. Standard methods of that time like the relative area of emphysema (RA) relied on a single intensity threshold of individual pixels, thus ignoring any interrelations between pixels. However, this paper focuses a lot on texture analysis, which provides a much richer representation of the local structure around the pixels. Severity of emphysema levels is measured by fusing output posterior probabilities of pixels by a classifier. Here, local binary patterns are used as texture features whereas joint LBP and intensity histograms are used for characterizing regions of interest (R**O**Is). A k nearest neighbor is then used for classification with a histogram dissimilarity measure as distance. A classification accuracy of 95.2% was achieved on a set of 168 manually annotated R**O**Is, comprising a class of three: normal tissue, centrilobular emphysema, and paraseptal emphysema. The measured severity of emphysema was found to be in good agreement with a pulmonary function test (PFT), thus achieving a correlation coefficient of up to $|r|=0.79$ in 39 subjects. The results were compared to RA and to a Gaussian filter bank, and texture-based measures correlated significantly better with PFT than did RA.

Some systems even made use of the ability of MDCT scanners to achieve isotropic 3D sub-millimeter resolution and also extended some of the already existing 2D feature sets into three dimensions [11]. The MDCT scans generally have the potential to allow for useful visualization, characterization, and quantification of the extent of diffuse lung diseases, such as the usual interstitial pneumonitis or idiopathic pulmonary fibrosis (UIP/IPF). Some of the works greatly objectify the need to standardize and improve the accuracy and repeatability of characterization of pulmonary diseases as well as of the quantification from such scans. Some of the approaches combine a texture matching method with histogram feature analysis.

In another research [12], an automated scheme was adopted for volumetric quantification of patterns within interstitial pneumonia (IP) which is essentially a subset of DPLD, that means, diffuse parenchyma lung disease. It obviously utilized a multidetector CT(MDCT) dataset and involved three important steps wherein initially a lung-filed segmentation was achieved by 3D

automated gray level thresholding which was then integrated with an edge characterizing wavelet preprocessing step followed by a texture based refinement of image border. The vessel tree volume is then identified and removed from the field of lung, thus resulting in lung parenchyma(LP) volume. Identification and characterization of IP patterns is then proposed as a three class pattern classification of LP into normal, ground glass and reticul ar pattern by utilizing k-nearest neighbor voxel classification and exploiting 3-D cooccurrence features. Volume overlap was then used to evaluate the performance of the proposed scheme in identifying and characterizing ground glass and reticular patterns.

Further investigation of 3D textural features' discriminating ability in diffuse lung disease quantification in MDCT was carried out [13]. Where the earlier schemes exploited textural features in combination with supervised classification algorithms, several 3D texture feature sets were then proposed. However, their ability to discriminate was not evaluated in a systematic manner, either in terms of individual feature sets or in conjunction to different classifiers. So in the associated work, four classification settings were combined with RLE feature set as well as with Laws feature set, which was first time employed for DLD characterization. Also, the combination of both the feature sets was tested using the same classification settings resulting in an accuracy above 98.6%.

One of the most recent studies on volumetric data suggested the application of 3D Riesz wavelet frames coupled with wavelet pyramids [14]. The sole aim of this method was to extract feature for differentiating between classical and atypical pattern of the usual interstitial pneumonia (UIP) in volumetric CT. The idea was to play with the bandwidth of steerable wavelets while still maintaining the time frame appropriately. On the other end, a family of maximally localized wavelet pyramids was designed in 3-D for a continuously adjustable radial bandwidth $[\Omega, \pi]$, $\Omega \in [\pi/4, \pi/2]$. The suggested wavelets were then merged with a rotational-covariant directional operator based on the Riesz transform. This transform is useful as it provides characterizations

of the organization image directions which are essentially independent of their local orientations. As a result, the effect of the bandwidth of the wavelet on classification performance was found to be really large. This kind of validated the importance of finding the minimum spatial support of wavelet which is required to leverage the wealth of morphological properties of tissue lying in the vicinity of the lung boundaries.

All the methods that have been described till now used hand-crafted features to describe lung tissue. One of the disadvantages of having such a feature is that it often fails to adapt to new data or patterns. Recently, studies are being adopted to learn schemes for feature extraction that would essentially result in feature sets customized to the training data. And would give useful results. Most of these are unsupervised techniques. In one of such works, texton-based classification is adopted [15]. It is mostly based on raw pixel representation and support vector machine having a radial basis function kernel for classification of emphysema in CT images of the lung. The proposed approach was tried and tested on 168 annotated ROIs which consisted of normal tissue, centrilobular emphysema, and paraseptal emphysema. The results obtained from this method clearly showed its superiority over other common techniques present in literature which included histogram of the filter responses based on Gaussian derivatives. The accuracy of the proposed method being 96.43% is also higher than that achieved through local binary patterns.

In another work [16], a wavelet transform that was based on Difference of Gaussians (DoG) was used so as to extract texture descriptors from a set of 90 images containing almost 1679 regions of manual annotations corresponding to various lung disorders associated with ILD. For future aggregation of tissues and to describe their pattern, visual words are used. Here various components like the progression scheme, number of visual words, as well as measurement of distance for clustering have been investigated. In order to describe patterns with high intra class variati ons such as that of healthy tissue, a sufficiently high number of visual words is required. According to the results, the system was able to learn wide intra class variations of healthy tissues as well as of abnormal lung tissue and provided quite a reliable assistance to the customers.

In the bag of words based method [17], six kinds of typical textural patterns were considered which were consolidation, ground-glass opacity, honeycombing, nodular, emphysema and normal tissue. Since these textural patterns are characterized both by CT values and by shapes, a set of statistical measure based local features was calculated from both CT values and the eigen values of Hessian matrices. The given method achieved a recognition rate of 95.85%, which was greater than the one with one global feature-based method and two other CT values-based bag of words methods.

Other class of methods used sparse representation models in which a set of atoms containing texture or textons are identified using k-means and k-SVD on the local patches that are already described. Since DLD patterns have got very high variety and complexity, the conventional methods are generally limited in performance when it comes to recognizing DLD patterns featured by geometrical information. Hence sparse representation based methods have been introduced to classify normal tissues and the five types of DLD patterns which include consolidation, honeycombing, emphysema, nodular and ground-glass opacity [18]. Both CT values and eigenvalues of Hessian matrices were used to calculate the local features. 2360 volumes of interest from almost 117 subjects were taken and then separated into two independent sets. **O**ne of the sets was used to optimize parameters whereas the other was used for evaluation. The textons that result from this make a problem-specific dictionary with every local structure in the image represented by the texton which is closest or by a linear combination of the entire set. A final global descriptor is prepared which consists of histogram of the image. The proposed method results in an overall accuracy of 95.4%. Hence experimental results prove that this method can be used to classify DLD patterns on HRCT images.

A multiscale sparse representation scheme was also developed which was based on wavelet and contourlet transforms [19]. It essentially described four patterns of diffused lung disease namely normal, emphysema, honey-combing and ground glass opacity based on HRCT images of lungs.

At first, sparse representation is used to train four discriminative dictionaries for four patterns respectively. After that, the classification phase begins in which a patch or ROI is assigned to the pattern which has got the minimum reconstruction error. The proposed method was tested against a collection of 89 slices from 38 patients, each slice being of size 512x512 and 16 bits/pixel in DICOM format. In the dataset, 73,000 ROIs were marked by experienced radiologists. This technique was essentially employed with 2-scale wavelet and [2 3] contourlet transform for classification of DLD. It was found to result in overall sensitivity of 91.05% and specificity of 97.01%.

Further developing the sparse representation, a texture classification method based on texton learned through SR was developed with new feature histogram maps in the classification of emphysema [20]. First, with the help of KSVD learning on every class of image patches in the training dataset, an overcomplete dictionary of KSVD is learnt. Also this stage makes use of a high pass filter so as to exclude patches in the smooth area in order to speed up the process of dictionary learning. In the second place, 3D joint-SR coefficients and intensity histograms are used to characterize regions of interest (ROIs) instead of the age old feature histograms which were constructed from SR coefficients of test images over the dictionary. Then comes classification which is performed using a classifier with distance as a measure of histogram dissimilarity. From 14 images of test subjects, almost four hundreds and seventy annotated ROIs are extracted, including 6 paraseptal emphysema(PSE) subjects, 5 centrilobular emphysema(CLE) subjects and 3 panlobular emphysema (PLE) subjects. These are essentially used to evaluate the effectiveness of the proposed method. It is also tested on 167 PSE, 20 CLE and 63 PLE ROIs which consist of mild, moderate and severe pulmonary emphysema.

The accuracy of this method was found to be around 74%, 89% and 88% for PSE, PLE and CLE respectively.

Another tool which was developed for extracting learned features is the restricted Boltzmann machine [21]. These are essentially artificial neural networks which are not just able to capture but also reproduce the statistical structure of the input and were employed for learning multiscale filters with their responses as output features. Although many feature vectors were already proposed before, they were quite complicated as well as domain specific. Recently the trend for automatic feature learning has become popular with the power to capture features without manual feature design. In the proposed paper, it has been suggested to create multi scale feature extractors that depend on an unsupervised learning algorithm. The image feature vectors are obtained by convolving the feature extractors with patches of the image. The image features that automatically result are data-adaptive as well as highly descriptive. A simple classification method can then be used to classify the image patches. A good thing about the proposed method is that it is a generic one and can easily be applied to different imaging domains. Evaluation was carried out on image patches to differentiate various patterns of lung tissue commonly seen in ILD which also demonstrated promising results.

Regardless of the features being handcrafted or learned, it is equally important to have a classifier that is able to optimally handle the properties of the feature space created. Literature contains many different approaches like the linear discriminant (LD), Bayesian classifiers, ANN, k-nearest neighbors, random forest and of course support vector machines with various kinds of kernels like linear, polynomial, or radial basis function. Furthermore, many works even utilized multiple kernel learning classifier (m-MKL) while in some of the works minimum reconstruction error served as classification criteria after the patch was reconstructed using some class-specific dictionaries.

Some recent techniques [22] have successfully made use of deep learning (DL) methods, especially of Convolutional Neural Networks after they performed impressively on large scale color image

classification. This approach is quite different from other feature learning methods in which data representation model is built in an unsupervised manner. CNNs rather learn features as well as train an ANN classifier simultaneously by constantly trying to minimize the classification error. Even though deep learning usually implies the use of not few but many learning layers, the initial attempts on lung CT images started with adopting quite shallow architectures.

In one of the works [21], RBM was modified to incorporate both feature extraction as well as classification of lung tissue, thus adopting some features of CNNs. RBM was introduced as convolutional classification, which is essentially a combination of the existing convolutional RBM and classification RBM, and this was used for discriminative feature learning. The classification accuracy of convolutional and non-convolutional classification RBMs was evaluated on two lung CT problems. It was found that features learnt by RBM way outperform their conventional RBM based feature learning, which is an unsupervised one, making use of only a generative learning objective and used filter banks. It has been shown in the proposed work that a combination of generative and discriminative learning are able to produce filters that result in higher classification accuracy. Weight sharing was used among the various hidden layers which ultimately had dense connection with the label or output neuron. The whole network was trained in a supervised way making use of contrastive divergence and gradient descent.

The authors of some other work [23] designed a convolutional layer with three dense layers and trained it from scratch. However due to the shallowness of the architecture, the descriptive ability of the CNNs could not be made use of. In other related work, 5 convolutional layers were used which were followed by two fully connected layers. The outputs of the final layer were completely different, corresponding to the ILD classes like healthy, reticulation, honeycombing, ground glass opacity and micronodules.

AlexNet was also used in [24] to classify whole lung slices after fine tuning it lung CT data. The authors of the work had to resize the input image as AlexNet basically works on natural color images of

size 224x224. After resizing the images, they generated three channels artificially by applying Hounsfield unit (HU) windows. However, the differences between color images of general type and medical images typically raised doubts about the transfer of knowledge between the two domains. Their classification of whole slices of lung could only provide a rough quantification of ILD.

1.2.2 Convolutional Neural Networks:

Convolutional Neural Networks are Artificial Neural Networks which are essentially feed-forward in nature. CNNs like neural networks are designed with neurons having learnable weights and biases. Each neuron receives several inputs from the neurons of the previous layer and then takes a weighted sum over them and thereby passes it through an activation function, responding with an output. The entire network is characterized by a loss function and all the techniques of reducing loss by backpropagating the errors is followed in CNNs as well. However, CNNs are quite different from neural networks in many ways. Some of the factors are:

CNNs work by operating over Volumes:

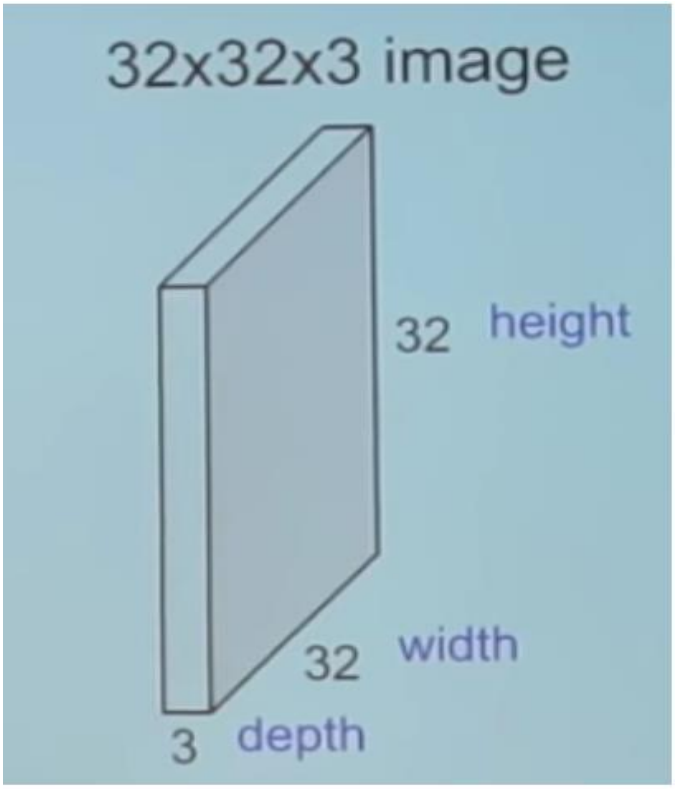


Fig 1.2.2 Image Patch

Unlike neural networks which take a vector as input, CNNs take multidimensional image as input. Before looking at other differences, we first try to understand what convolution means or how it is performed.

1.2.2.1 Convolution:

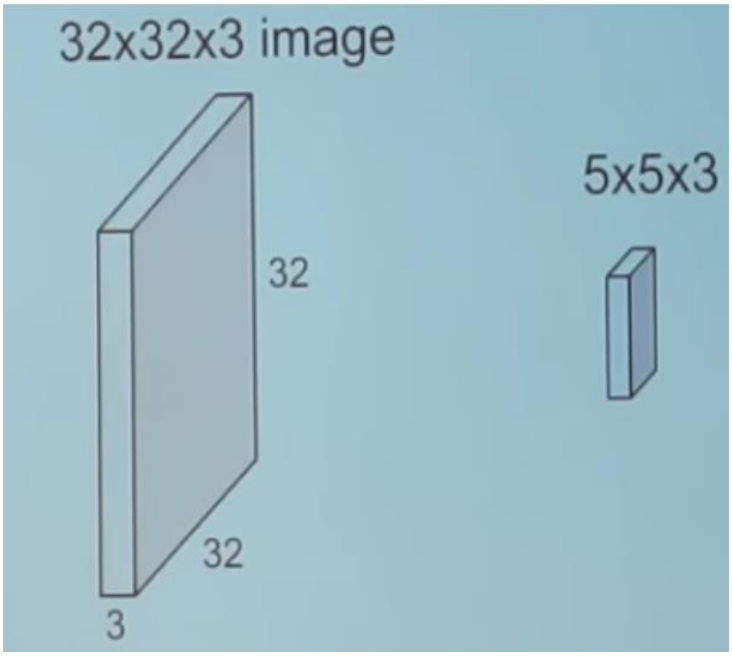


Fig 1.2.2.1 Convolving an image with a filter

As shown in the image (Fig 1.2.2.1), a 5x5x3 filter is taken and is slid over the input image where dot products of the filter and various chunks of the image is taken.

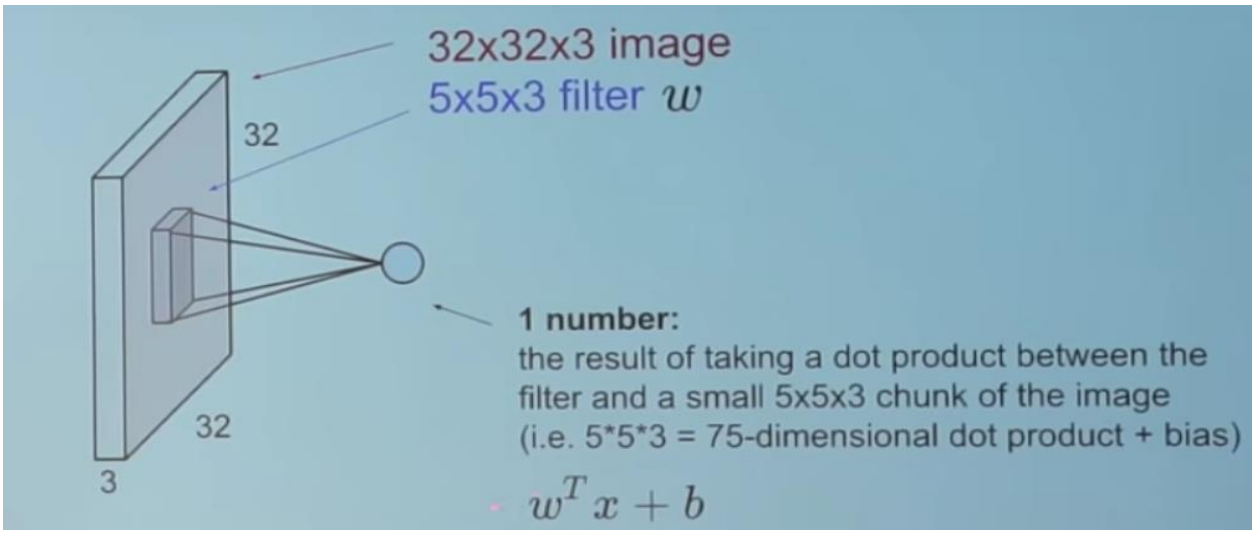


Fig 1.2.2.2. This is how it looks

For every product taken, the result is a scalar. Following happens when entire image is convolved with the filter:

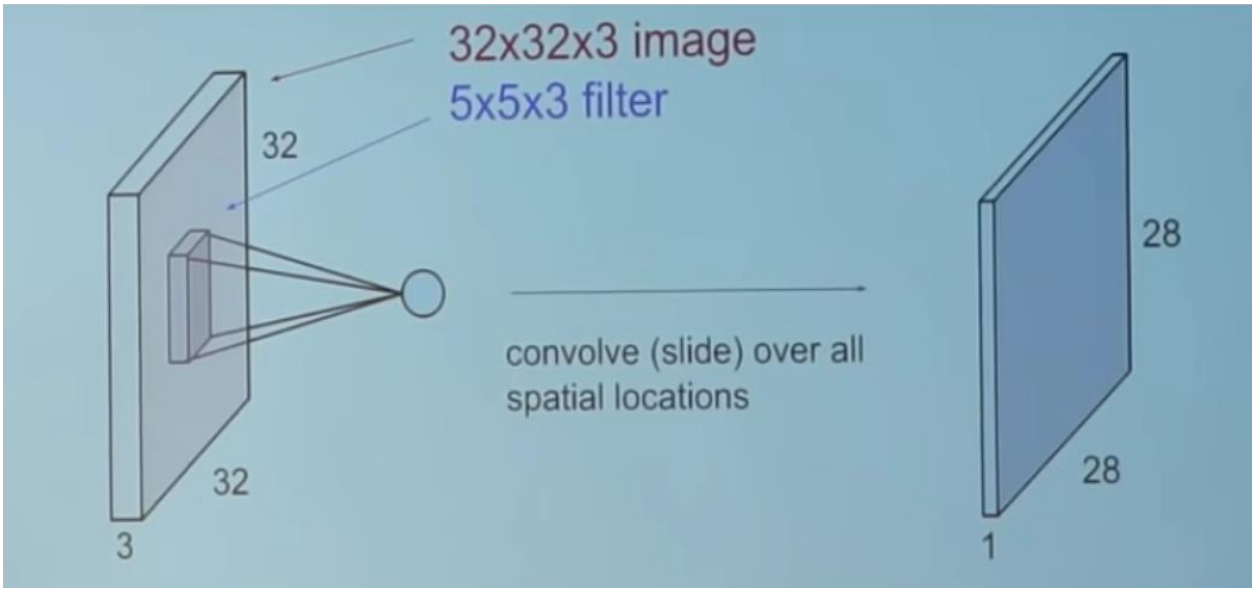


Fig. 1.2.2.3 Performing Convolution

Figure of 28 comes out in the output image as there are 28x28 unique positions where filter can be put on the image.

Now, coming back to CNNs, we observe that the convolutional layer is the main building block of a Convolutional Neural Network.

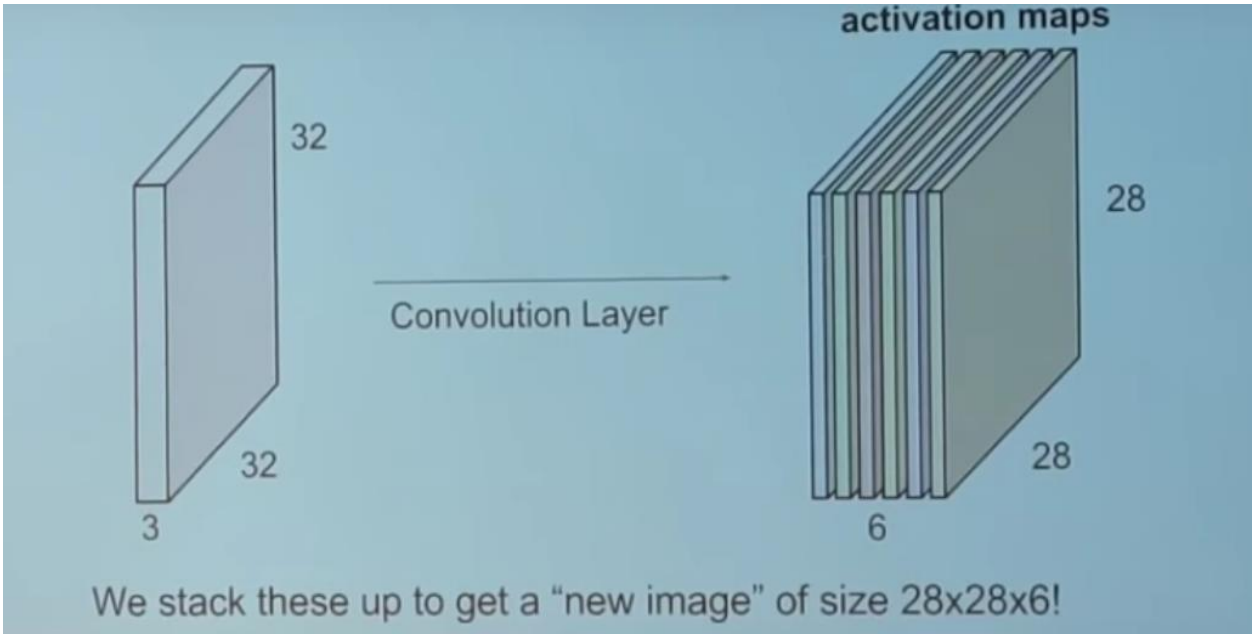


Fig 1.2.2.4. Convolutional Layer

In the convolutional layer, 6 independent filters are used, each being independently convolved with the image. In this way, we end up with 6 feature maps of shape 28*28*1.

Now, it is interesting to know what happens if we have a number of convolutional layers in sequence.

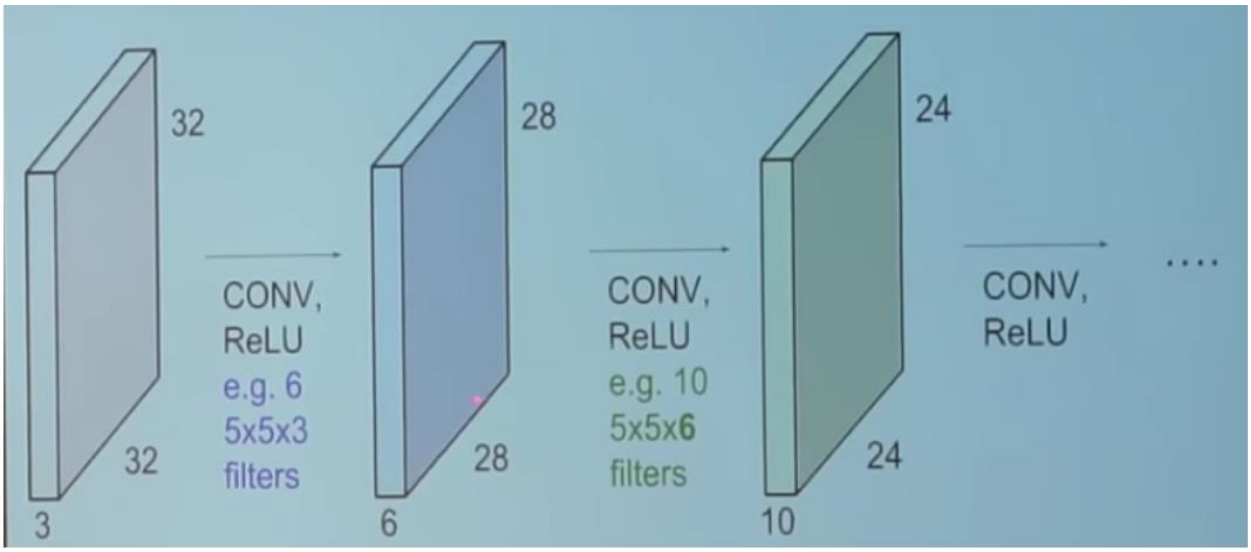


Fig 1.2.2.5. Convolutional layers in sequence

All the convolutional filters are randomly initialized which become our parameters of interest and are subsequently learned by the network.

Following is an example of trained network:

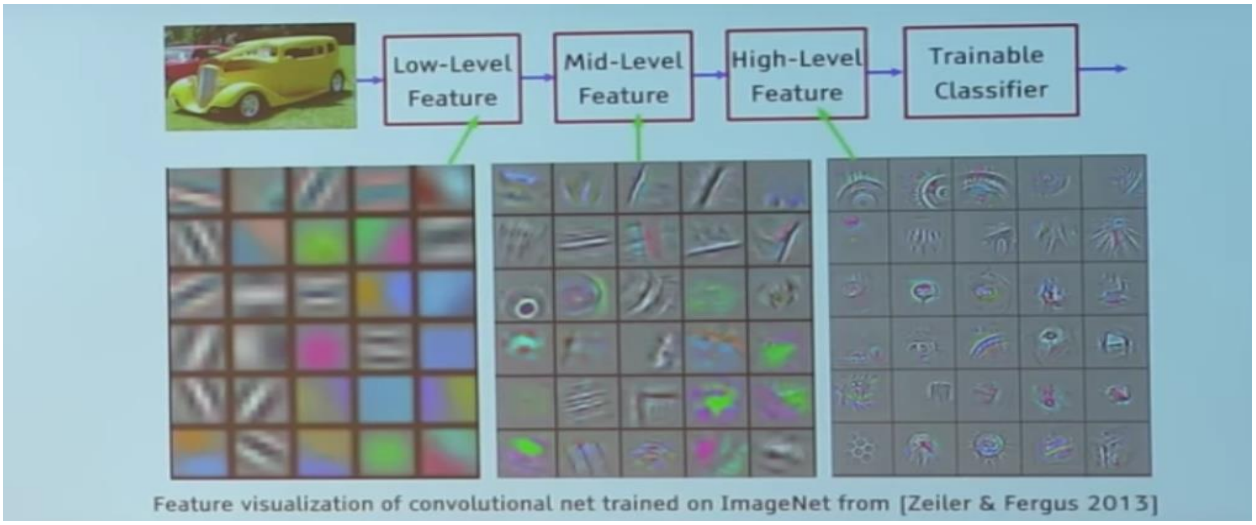


Fig 1.2.2.6. Filters in a trained network

Taking a look at the filters in the first layer (5*5*3 filters that we had taken), we see that they have tuned themselves and have become blobs of colored pieces and edges. As we move towards the other convolution layers, the filters are performing dot products with the input of the previous convolutional layers. So, essentially, smaller colored pieces or edges are being taken and larger pieces are being made out of them.

Taking a look at fig. 1.2.2.5, and imagining the 28*28*1 grid as a grid of 28*28 neurons, we find that for a particular feature map, each neuron is connected to only a small chunk of input image with all the neurons having the same connection weights.

Now, we come back to the differences between CNN and neural network.

CNN has concepts like parameter sharing and local connectivity

Parameter sharing implies sharing of weights by all the neurons in a particular feature map whereas local connectivity implies that each neuron is connected only to a subset of the input image, unlike that of neural network in which all the neurons are fully connected.

This really reduces the number of parameters in the entire system, thus, making the computation more effective.

1.2.2.2 Pooling Layers

A pooling layer is simply another building block of CNN.

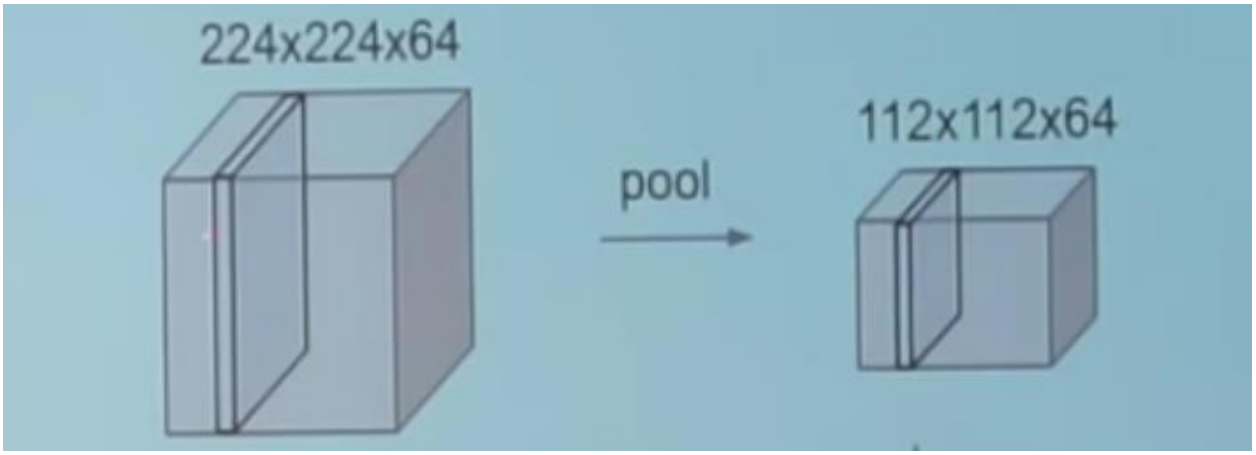


Fig 1.2.2.7. Pooling

Pooling is an operation to reduce the spatial size of representation so as to reduce the amount of parameters as well as the amount of computation in the network. Pooling layer works by operating independently on each feature map.

Max pooling is the most common approach used in pooling.

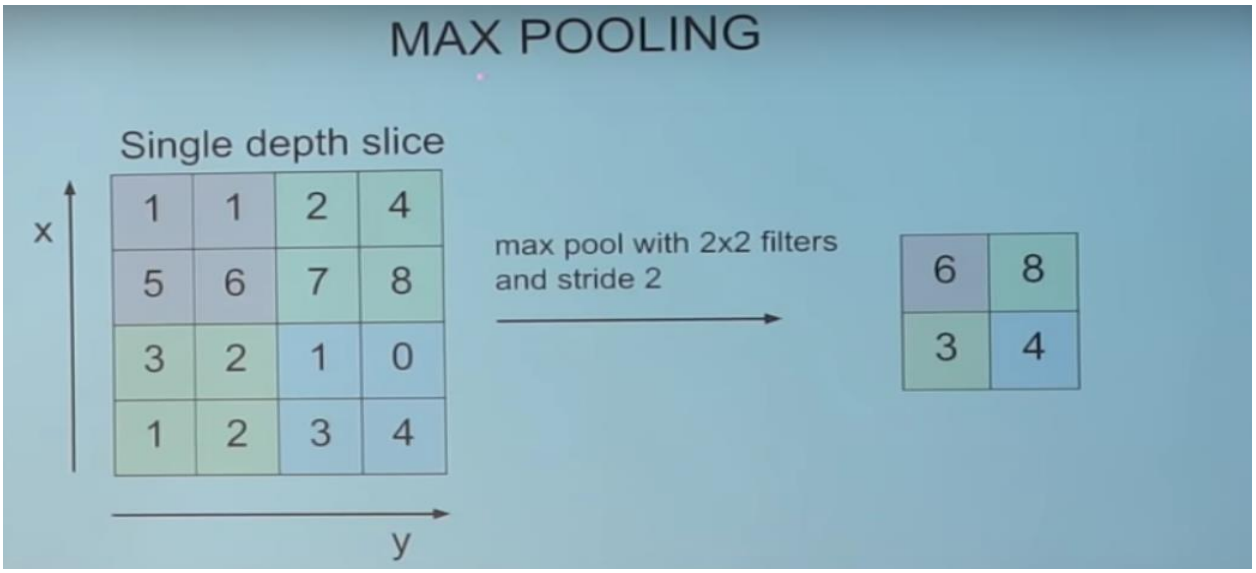


Fig 1.2.2.8. Max pooling

1.2.2.3 Typical Architecture of CNN:

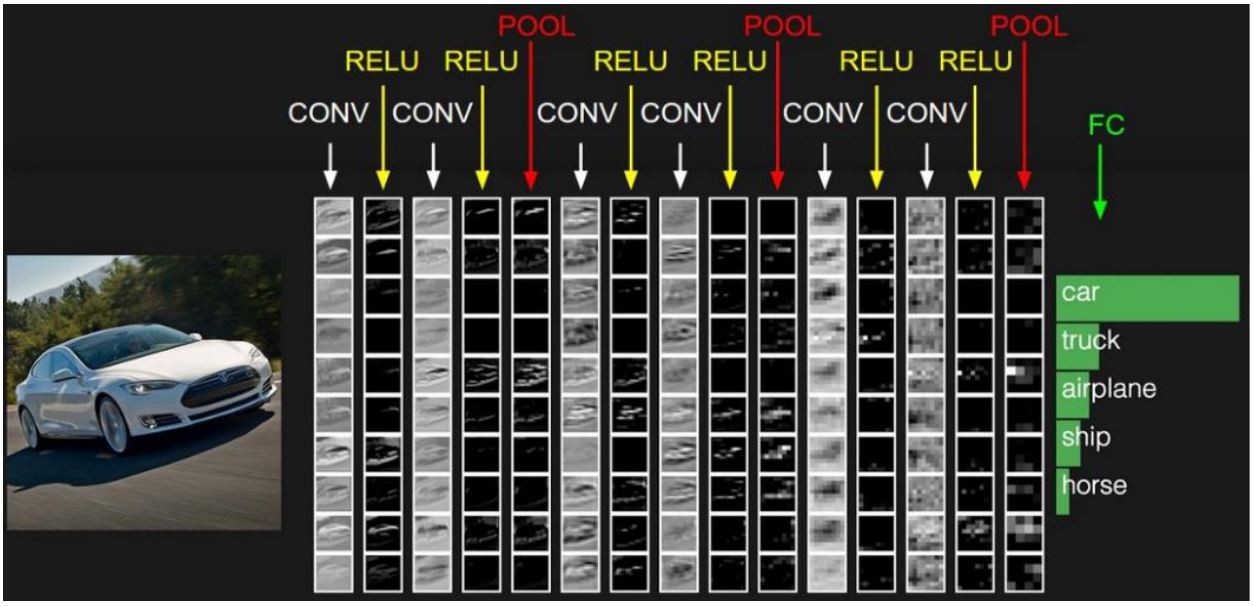


Fig 1.2.2.9. Typical architecture of CNN

CONV denotes the convolutional layer and POOL denotes the pooling layer. Similar to neural networks, non-linearity is introduced by RELU.

FC is the fully connected layer of neurons present at the end of the entire convolutional network. In a fully connected layer, neurons maintain connections with all the activations in the previous layer just as in case of regular Neural Networks which work in a similar way.

The basic concept of CNN remains the same although there may be slight variations in the architecture.

Although CNNs existed for a very long time (almost for decades), it was possible to train such deep neural networks only recently. This happened due to the extensive parallelization property of such networks, that is, their coupling with massively parallel GPUs, huge amounts of data involved and with several design tricks like the rectified linear unitso (ReLU) Krizhevsky et al. [25] in 2012 won the Large-Scale Visual Recognition Challenge by outperforming all the other competitors on a challenging dataset consisting of 1000 classes and 1.2 million images. The deep CNN architecture which was proposed was AlexNet, consisting of five convolutional layers followed by ReLU activations, some of which was followed by max-pooling layers along with three dense layers with a final 1000-way softmax layer. Training of the network was done with stochastic gradient descent (SGD) involving a momentum term and maximizing the objective of multinomial logistic regression.

One important fact about deep architectures is that it permits the learning of data representations in multiple levels, thus, allowing the recognition of even high level visual structures like cars or faces be recognized in the last layer by effectively combining the low-level features of the first layers like edges. However, it's not trivial to design a deep convolutional neural network as a large number of mutually dependent parameter values and algorithmic choices have to be made. Although a lot of research has been done

in recent years for classification of color images through deep CNNs, not much has been done to solve the problems of texture recognition and medical image analysis.

This motivates us to do proper research in this area by applying various methods and network architectures and see which gives us the best results. Also, applying proper pre-processing techniques is important prior to the application of different architectures so that it becomes easy and efficient for the network to do its task of classification and come up with high rates of accuracy and results.

CHAPTER 2

Research Methodologies:

First, we describe the dataset used in the study and then various architectures of CNN used. Such a description of the input data and the desired outputs helps us better understand the definition of the problem and therefore, understand the methods in a better way [26].

2.1 Data

Collection of sufficient number of cases, that are representative of the various realizations of the studied diseases is very important in order to understand the underlying mechanism of such diseases. Also, the quality of the acquired data is very important to carry out evaluations of the CAD systems in an unbiased way. Also, enhancing the disease therapy in clinical routine would require that sufficient library of cases are represented as close as possible to the hospital's population. This means, if possible, cases should be randomly chosen among the entire population. These requirements are really important to be fulfilled while building a diagnostic aid which is computerized, image-based and based on processing of medical image. For all this, a multimedia collection of very high quality and of cases containing annotated image series and associated clinical parameters is required so as to ensure that such a CAD system succeeds while being integrated into the clinical routine.

On one hand, the dataset should contain a means for developing computerized tools like automatic detection of abnormal pulmonary tissue types in HRCT, which are essentially high-resolution computed tomography images. Ground truth is also required to be detailed and a large number of cases are required to evaluate as well as compare the algorithms of medical image processing for some defined tasks, that is, benchmarks. There are some datasets like Lena, Brodatz, Iris1 which allow to qualitatively and quantitatively evaluate a large number of basic methods in image processing and also apply machine learning algorithms, thus establishing a

kind of de facto reference dataset. These datasets are popular as they reflect challenges of real life and therefore offer higher credibility of obtained results when compared to artificial datasets. These publicly available datasets can easily be reproduced by researchers and can be used to compare to the state of the art, which is an important element of science. Also, such datasets create opportunities for specialized teaching and learning. Knowledge base is constituted of cases with confirmed diagnosis which are used as diagnostic aid. In order to summarize, it can be said that multimedia library of cases is very useful for the following reasons:

- Teaching
- specialized descriptive studies
- training and testing pattern recognition techniques
- retrieving similar cases as diagnostic aid
- comparative performance analysis of medical image processing
- Methods

But, constructing a high-quality multimedia collection of cases is really expensive as well as time-consuming. This creates a bottleneck in studies involving image-based CAD systems. There are other relevant tasks which require a large amount of work with a wide range of skills like identifying the relevant cases, consulting the electronic health recordo (EHR) and archiving picture and communication systems(PACS) to gather the clinical parameters and image series, data entry as well as infrastructure and maintenance of database. The range of skills varies from medical knowledge to information technology (IT) expertise. So, in order to assess the high-quality data, many researchers and physicians are required to be involved in case selection process as well as in the delineation of regions of interest (R**OI**s) so as to cope up with inter and intra-observer variabilities where intra-observer variability is specifically important in radiology. Also, it is important to get the agreement of the ethics committee before starting any investigations. It contains the justification to access the content of EHR.

Depending upon the diseases that we are studying, there are sometimes cases which are rare and suddenly encountered even if it is taken from a large university hospital. These diseases, due to their sparsity, require reference databases to palliate the lack of experience. A major observation while studying the state-of-the art of texture-based CADs for analyzing lung tissue in thin-section computed tomography (CT) is the fact that measured performance of statistical significance lacks as small number of cases are usually used to evaluate CAD systems.

We already saw that ILDs are gradual alteration of lung parenchyma which leads to dysfunction in breathing. More than 150 histological diagnoses are regrouped under ILD which are associated with disorders of lung parenchyma. Although mechanisms as well as factors of the diseases vary from one disease to another, the exact cause of many ILDs is still not known. It is very frequent that diagnosis of a person affected with ILD does not lead to any specific findings. The diagnosis of such pathologies is established depending upon the complete history of the patient, also involving physical examination, laboratory tests, pulmonary function testing (PFT) and visual findings on chest X-ray.

Images are really important as they confirm the diagnosis and also, the patients need not require surgical lung biopsy when clinical and radiographic impression hint towards safe diagnosis. The first imaging examination to be used is chest radiograph. It is used because of its low cost weak and exposure to radiation also providing a quick overview of the entire chest. However, chest radiographs are not that reliable as they come out to be normal in more than 10% of the patients suffering with some form of ILD. They can provide a confident diagnosis in only 23% of cases involving lung diseases in general. Hence it is very likely for this information to arouse suspicions towards an ILD, with HRCT imaging of chest acquiring a rapid and accurate visual assessment of lung tissue. Indeed, the three-dimensional form of HRCT data avoids the superposition of organs and also provides an accu -rate assessment of the patterns as well as distribution of the lung tissue with sub-millimetric resolution. This really made it an imaging protocol with the standard of gold for

diagnosis of diffuse pulmonary parenchymal disease. Details of the most common histological diagnoses of ILDs are depicted in fig 2.1.1.

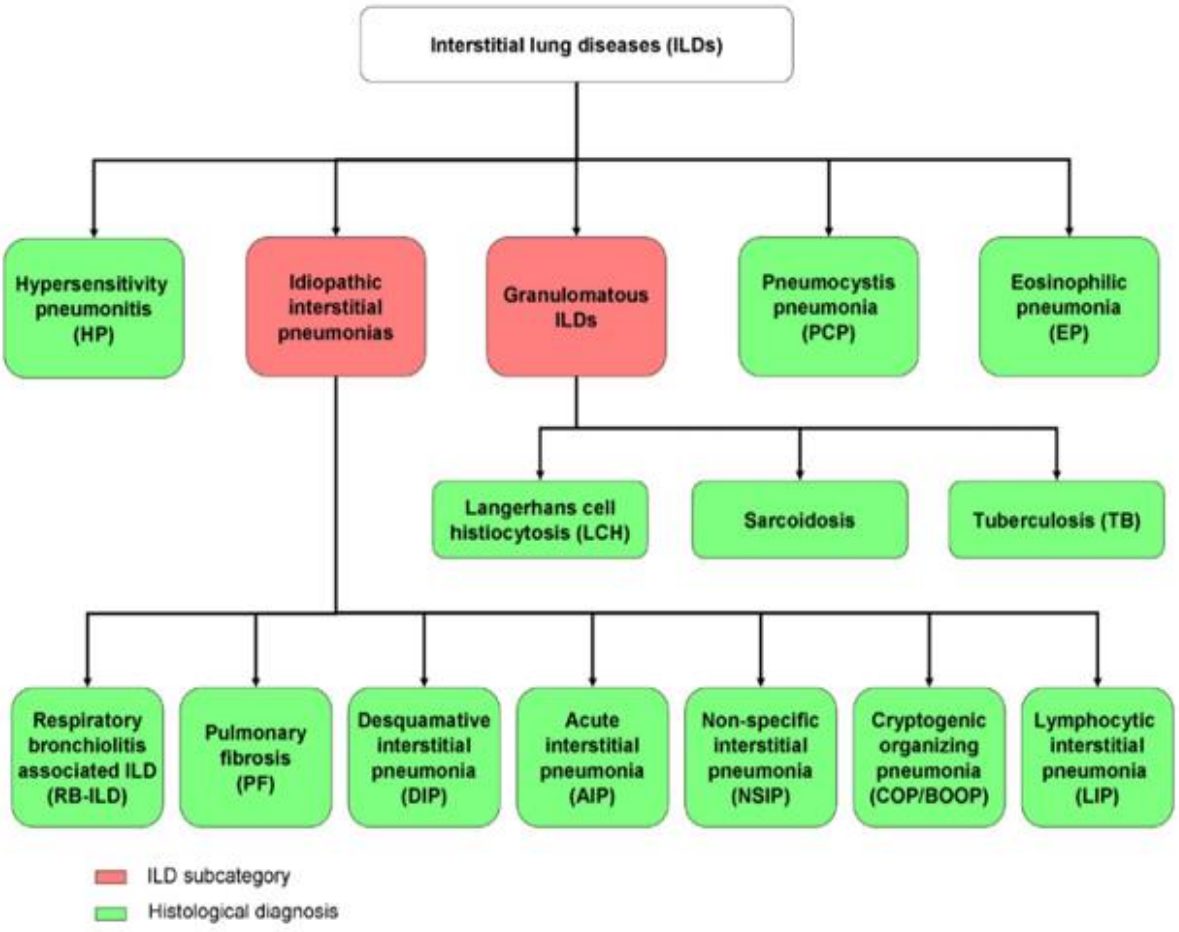


Fig 2.1.1 Most common histological diagnosis of ILD according to [20,23]

2.2 High resolution computed tomography of the lungs:

The first image series with a relatively low resolution was created by a commercial CT scanner in the year 1972. The image series had a resolution of 80x80 pixel in each of the axial slices. X-ray attenuation relates to the value of each pixel numerically which is expressed in Hounsfield Units (HU). Multiple detectors in modern scanners use a scanning mode which is helical in nature and can provide a 3D array of isotropic voxels having a sub millimetric resolution. Such a protocol is called multidetector computed tomography (MDCT). During this protocol, the patient is exposed to a very high amount of radiation dose which is also one of the main drawback of the protocol. Sub-millimetric resolution is required to assess and differentiate between a healthy and pathological lung tissue. But then, some areas can be skipped between the thin sections so as to limit the exposure of radiation. HRCT is the protocol and it is also the gold standard imaging protocol for diagnosing ILDs. Following is the table which lists for HRCT scanning protocol:

Table 2.2.1

HRCT scanning protocol

Slice thickness	1-2 mm
Spacing between slices	10-15 mm
Scan time	1-2 s
Lung Shape	Inspirium
Contrast agent	None
Axial pixel matrix	512 x 512
X, y spacing	0.4-1 mm

HRCT is in fact more apt than magnetic resonance imaging (MRI) for assessing appearance of the lung tissue visually. Also, because of the fact that other tissue types have a low density of protons.

MRI is sensitive to inflammatory changes of pulmonary parenchyma only. Following shows a comparison between HRCT and MRI.

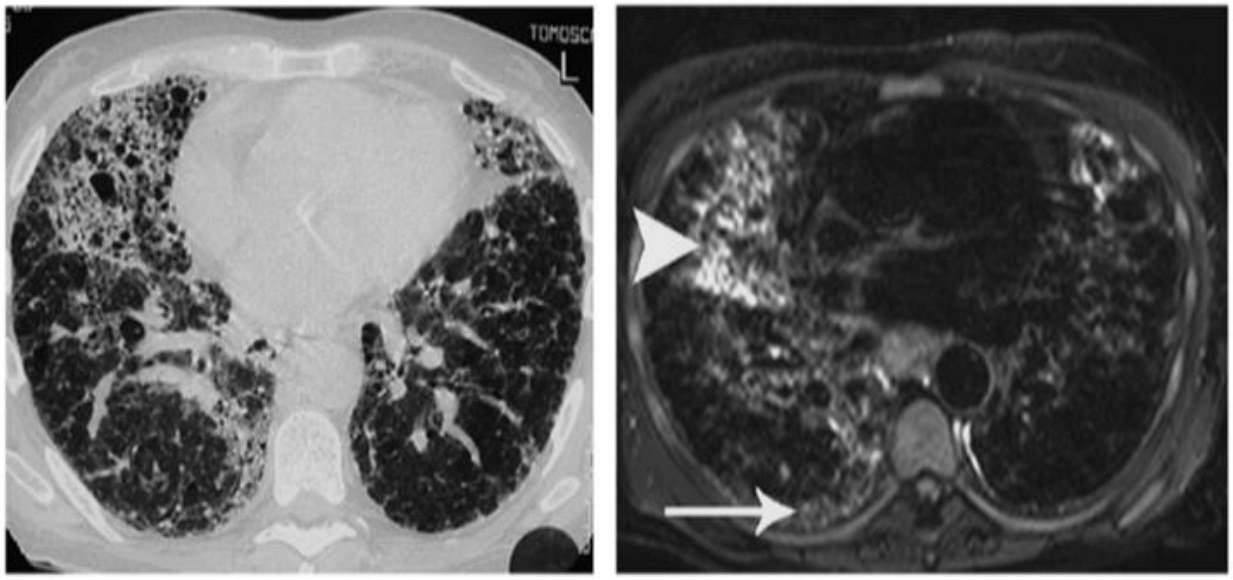


Fig 2.2.1. HRCT and MRI for visual comparison and assessment of lung tissue. In case of MRI image, (b) the signal of the infected lung area (shown with an arrowhead) is 120% higher than that of the fibrous tissue area (depicted with an arrow)

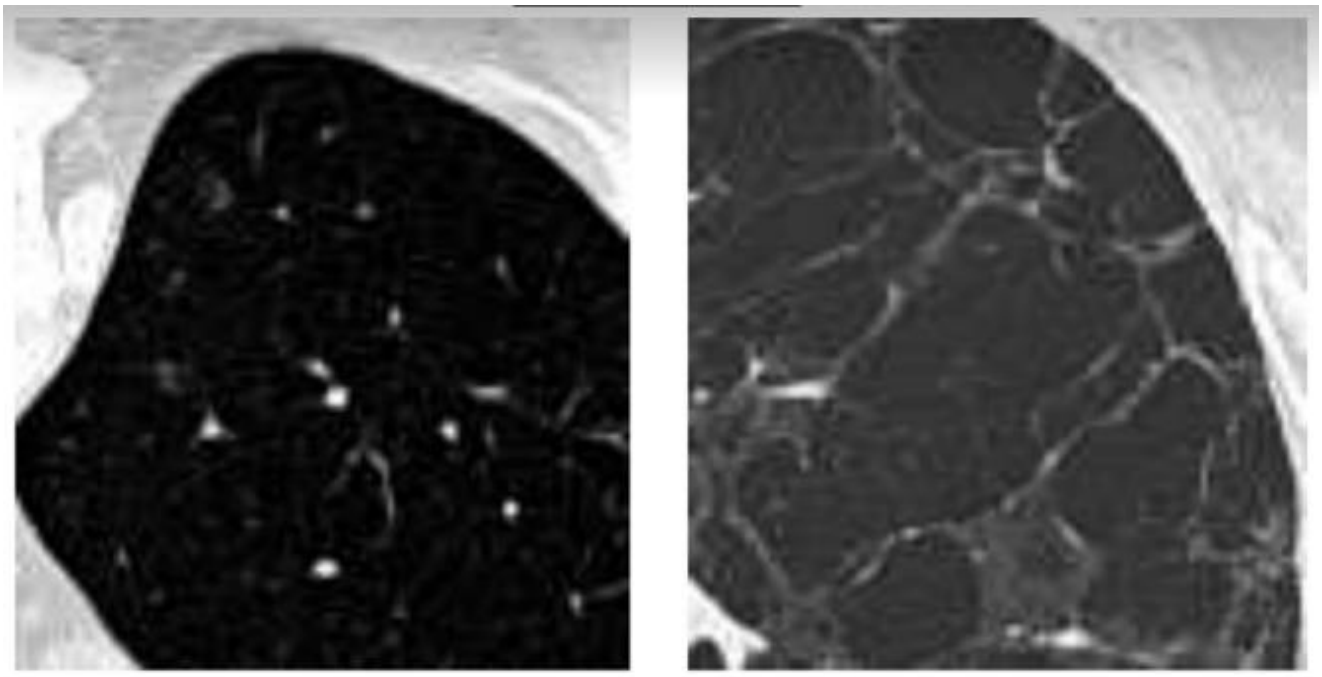
Table 2.2.2

Histological diagnosis	HRCT lung tissue patterns	Predominance
Hypersensitivity pneumonitis (HP)	Ground glass, emphysema, fibrosis	Diffuse
Pneumocystis pneumonia (PCP)	Ground glass, crazy-paving, cysts, pneumothorax	Central , perihilar
Eosinophilic pneumonia (EP)	Ground glass, consolidation, crazy-paving	Peripheral , apex
Langerhans cell histiocytosis (LCH)	Cysts , ground-glass, micronodules, reticulation	Apex
Sarcoidosis	Micronodules, consolidation, macronodules, ground-glass, fibrosis (end-stage)	Peribronchovascular, subpleural , peripheral
Tuberculosis (TB)	Micronodules(military), tree-in-bud, consolidation	Diffuse
Respiratory bronchiolitis associated ILD (RB-ILD)	Ground-glass, emphysema	Diffuse , centrilobular
Pulmonary fibrosis (PF)	Fibrosis, bronchiectasis, ground-glass	Peripheral , subpleural , basal, posterior
Desquamative interstitial pneumonia (DIP)	Ground-glass, emphysema , fibrosis (uncommon)	Subpleural

Acute interstitial pneumonia (AIP)	Ground-glass, consolidation	Basal , diffuse
Non-specific interstitial pneumonia (NSIP)	Ground-glass, consolidation, reticulation, fibrosis (uncommon)	Peripheral , basal
Cryptogenic organizing pneumonia (C**O**P)	Ground-glass, consolidation (patchy), macronodules , bronchial wall thickening, crazy-paving	Peribronchovascular, subpleural
Lymphocytic interstitial pneumonia (LIP)	Ground-glass, micronodules	Peribronchovascular, subpleural

2.3 Lung tissue patterns associated with ILDs in HRCT:

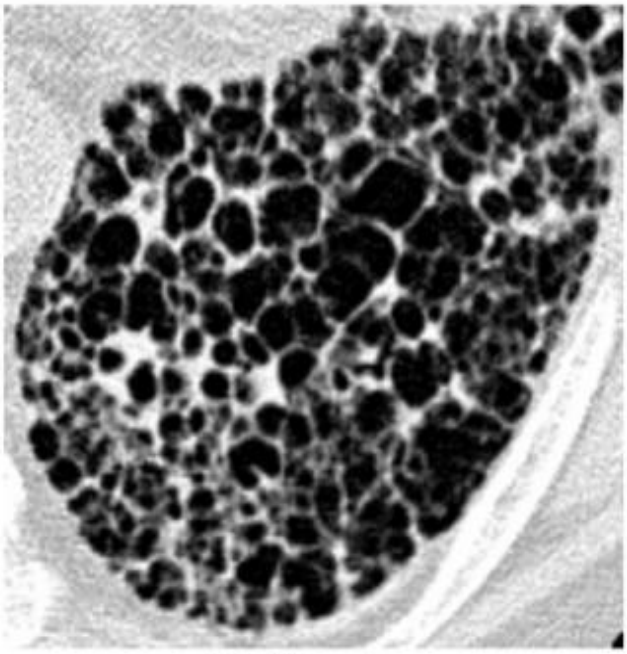
In order to establish the differential diagnosis of ILD, information contained in the appearance and quantification of lung tissue types is very important. The table above lists 13 common histological diagnoses of ILDs, its associated HRCT findings and the specific region of the lungs where ILD is prevalent. The following figure depicts the visual aspects of most common patterns of lung tissue. The taxonomy used for their depiction often connects to their texture properties. Fibrosis is used in the present work to denote histological diagnosis of "pulmonary fibrosis" and it includes traction bronchiectasis, reticulation, honeycombing and architecture distortion. As it can be inferred from Table 2, the most encountered patterns are the ground-glass ones and are therefore non-specific. This calls for clinical context and other findings of HRCT required to provide orientation to the dia



(a) *healthy*

(b) *emphysema*

(c) *ground glass*



(d) *fibrosis*

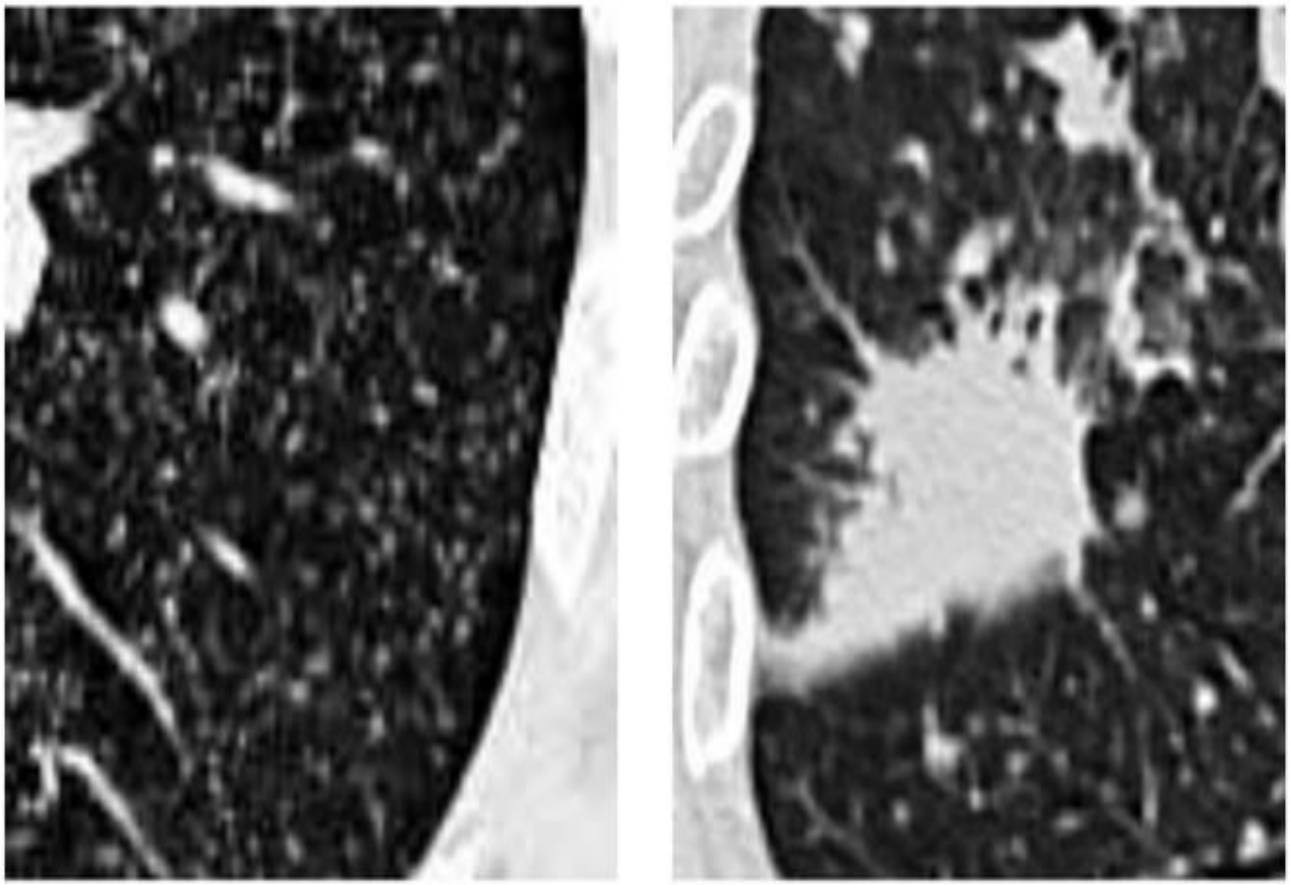


Fig 2.3.1. Visual aspects of the most common lung tissue patterns in HRCT of patients suffering with ILDs . (a) Healthy (b) emphysema (c) ground glass (d) fibrosis (e) micronodules (f) consolidation

2.4 Interpretation of HRCT image series:

HRCT image representation is really a tough task even for trained radiologists of chest and the lung specialists . It requires a lot of reading time, effort as well as experience to correctly interpret the three-dimensional form of data. Most of the times, the process of interpretation involves comparison of a

case with similar types of images found in textbooks or by comparing with similar collections of images often organized by pathology. In order to do this, it is important for the radiologists to have a guess or pre-knowledge of the disease suspected which might be present in the image and miss the true pathology shown. Certain medical services like emergency radiology service requires radiologists to have recourse to a wide diversity of imaging modalities like CT, conventional projection radiography, MRI, functional imagingo (functional MRI (FMRI), positron emission tomography (PET)) as well as ultrasound applied to various organs like the brain ,breast , liver, vascular and skeletal systems , colon, chest and kidney. They are required to provide radiological report at first with quick ideas on diagnosis. This might also result in errors due to omission or due to confusion of pathological lung tissue of varied types . Also, it is very fundamental for the context to be interpreted correctly. For example, a healthy tissue may have visually different aspects due to factors of age or due to smoking history of the patient. Similarly, the ground glass findings are unspecific without parameters that are clinical and complementary.

2.5 Existing databases of CT imaging of the lung

For testing and developing lung CADs from trustworthy and reliable datasets for the detection of nodules of lung on CT scans,the lung imaging database c onsortium (LIDC) was formed consisting of five academic institutions throughout United States. Along with including healthy images in the database, it also included pathologic CT images having nodules with annotations and primary clinical data of the patient. In order to agree upon the definition of nodules and to finalize the criteria for inclusion in database, expert radiologists from five different institutions were considered together. The classification of the nodules was done into three classes based on their diameters , that is, “nodules<3 mm”, “nodules>3 mm” and “non-nodule>=3mm”.Annotations of five radiologists are available. After a session of blinded annota -tions, radiologists had to access the annotations of fellow radiologists or colleagues and thus modify their own findings in a retrospective manner. The number of nodules agreed upon by all the four radiologists was 33.8% for blinded session and 45.8% after the review was done.

The database could be accessed online from the national biomedical image archive (NBIA). However, it was unfortunate that the LIDC database did not contain any ILD cases as it only focused and had CT imaging nodules. Also,a small database of annotated nodules in CT imag ing of lung is available publicly having a purpose of comparing CAD performance. This database is available to be downloaded online.

The National Heart, Lung and Blood Institute (NHLB) laid similar efforts but focused on lung tissue by creating the lung tissue research consortium(LTRC). It aimed at improving the management of diffuse lung diseases by developing a better understanding of the biology related to chronic obstructive pulmonary diseases (C**O**PDs) as well as of fibrotic ILDs which included idiopathic pulmonary fibrosis (IPF). Also, enrollment of controlled cases was done. Creation of an open data set was aimed at. It essentially contained clinical, radiological and histological data. The lung tissue patterns were then described in a structured report. But, there were no delineated regions of interest present in the image series. Hence, no ground truth was present to evaluate computerized categorization of lung tissue. HRCT image series as well as metadata could be freely accessed after obtaining approval of LTRC data coordinating center which would effectively be based on the submission of a concept sheet. The concept sheet was meant to describe the aim of study by making use of a standardized format. Then, on the basis of LTRC data, disease patterns were quantized and this was proposed as means to depict the extent of pulmonary disease.

Another architecture was proposed which was web-based teleradiological framework to acquire different cases of lung diseases with annotated regions already proposed. Tools are provided by the Learning Medical Image Knowledge (LMIK), which is essentially a collaborative platform. It provides tools to the clinicians so as to delineate regions of interests (R**O**Is) in HRCT imaging of the chest. A central database is then used to anonymize the cases which can later be queried by researchers who are authorized for evaluations such as CAD and could also be used by radiologists for teaching purposes.

However, it is unfortunate that no public access is mentioned to the case repository. Also, no recent reports have been found on LMIK activities since the year 2003.

Apart from these efforts, there is no large dataset with annotated regions available to be used for ground truth for the purpose of either evaluation or comparison of computerized categorization of lung tissue in HRCT.

2.6 Methodology:

This section includes details such as scope of the database and also the different steps involved in the collection of data. Various methods are described which result from repeating the refinement iteratively in the duration of 38 months of data collections.

2.6.1 Scope of the database:

Ahead of collecting any cases, database was defined or limited to have a scope so that the set of cases is consistent and the objective of having computerized diagnosis aid for ILDs is fulfilled. This section contains details of the selection parameters of the histological diagnosis that needs to be included in the database and also the associated clinical parameters.

Selection of the histological diagnoses

A number of histological diagnoses which are considered as the most frequent causes of lung parenchymal disorders were chosen with the help of a collaboration between the Service of Emergency Radiology and the Service of Pneumology of the University Hospitals of Geneva (HUG). The agenda was to analyze closely and retrospectively enough cases representative of the 5 most common and frequent forms of ILDs from EHR at HUG collected during four years of Talisman project. Although there were rare cases of healthy diagnoses, yet they were included as much as possible to serve as control cases.

2.6.2 Selection of clinical parameters:

Depending upon each of the pathologies, the clinical parameters that were found to be most discriminative were kept so as to establish differential diagnosis. A domain-specific literature was used to carry out the selection process and it also included the knowledge bases of computer-based diagnosis decision support systems. Experts of the field like the radiologists, lung specialists and medical informatics research group (Service of Medical Informatics, HUG) reviewed the selected parameters and standardized units and data formats iteratively. Certain parameters which were not present in EHR (such as race) were removed. The list was extensively reviewed and various modifications were also made based on the parameters that were available in EHR, 159 fields or parameters or attributes were selected to characterize or signify a group of ILDs. In order to take note of the clinical parameters, an HTML was used. The standard units and for numerical data and laboratory tests were kept in accordance to the formats used in electronics patient record at HUG.

2.6.3 Data Collection:

The selection of the cases at first, annotation of the images as well as entry of data are resulted as part of regular discussions that involve not only the radiologists, but also, the computer scientists and the research physicians during four years of project. Several steps that are taken to construct a multi-media dataset ranging from selection of cases to data entry are described in the following sections in detail.

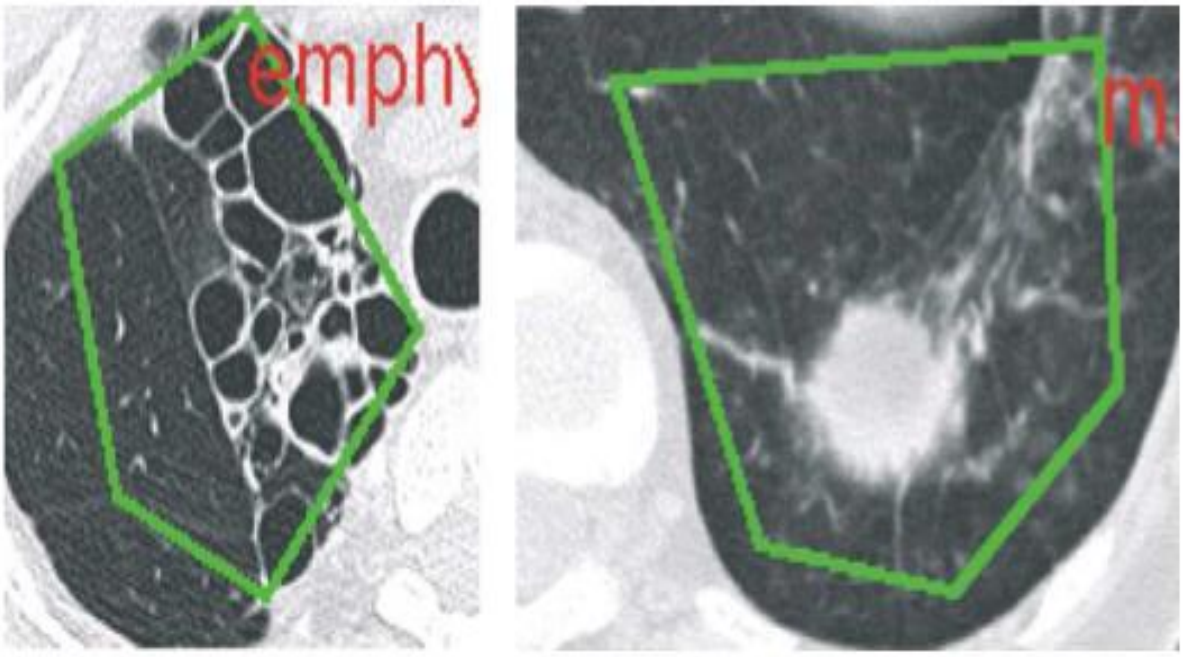


Fig 2.6.1. An example of inappropriate delineation of ROIs where it contains as much healthy tissue as pathological, thus, introducing noise in the training data.

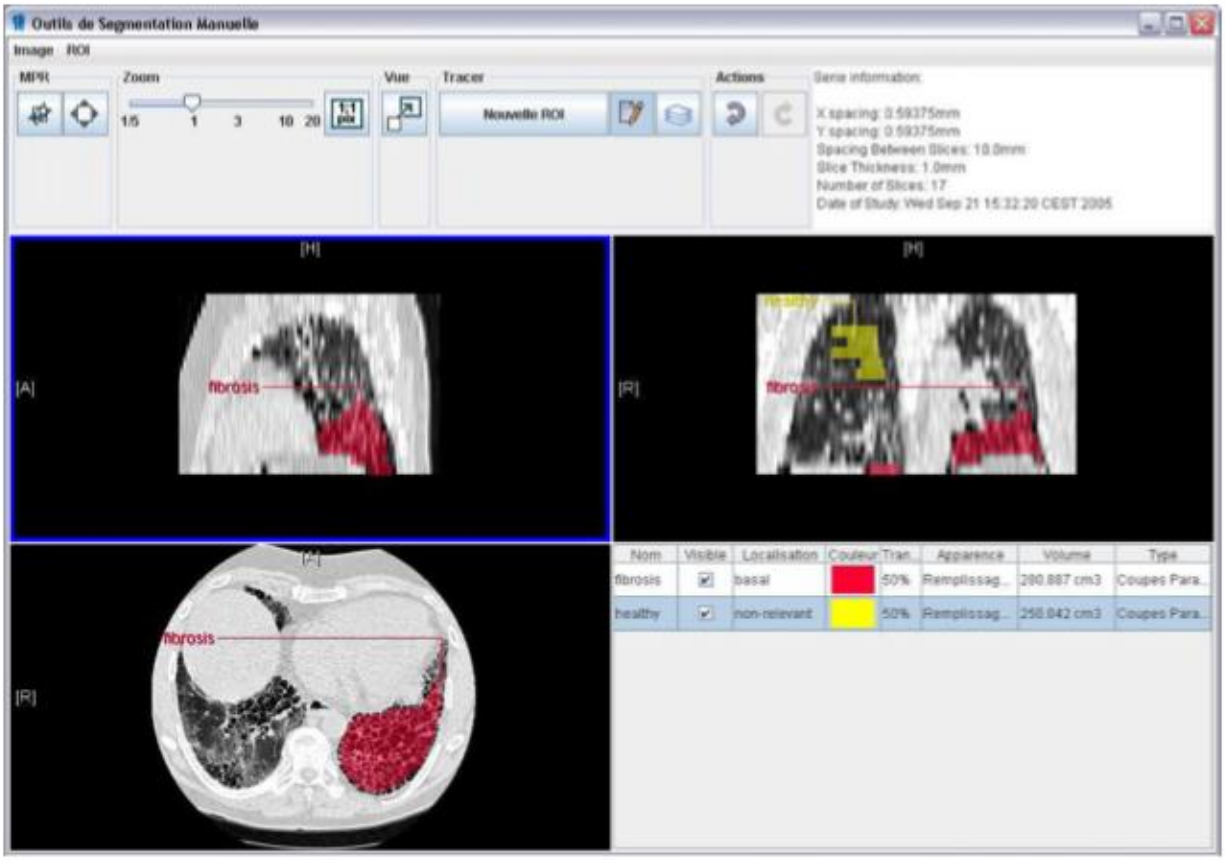


Fig 2.6.2. A screenshot of the graphical tool showing annotation of image regions

2.7 Selection of cases at the University Hospitals of Geneva

About 18400 samples or images were derived from a raw list of some 1266 patients who underwent a thorax CT within a stay in the pneumology service between 2003 and 2008.These were extracted from data repository of EHR by the helpdesk at the medical informatics service.

Only those cases in which HRCT (without having contrast agent, 1mm slice thickness) were analyzed retrospectively. Those cases from pediatrics were left aside. Each of the remaining cases were

diagnosed and were retraced in the EHR on the basis of reports, clinical history and clinical examinations.

Initially, the consultation report of pneumology and the discharge summary were put under revision to decide whether it is possible for the cases to contain an ILD. After the reports of CT scans were found to be consistent with that of clinical reports, clinical history and laboratory studies including pulmonary function testing, pathology, hematologic test, only then the cases were considered as candidates. Cases in which the histological diagnosis resulted into one of the 13 listed in fig. 2.3.1 from biopsy, bronchoalveolar lavageo (BAL) were included and selected. After the radiographic impression was found to be consistent with verified diagnosis, the case was kept and considered for further annotation sessions with the radiologists. A very selective procedure was carried out to retain only those cases where the confidence interval was found to be really high so that more time is attributed to the annotation sessions and more focus is put on radiological aspects only.

The time taken to decide whether a case has to be considered for further annotation or not varied from 15 mins to 2 hrs. Average time for each case was approximately 1 hr. In the early months of the project, usually more time was taken as the methodology was not properly established. Ruling out a case was comparatively easy than to decide whether it should be kept or not as it was easy to reject an ILD diagnosis or to discover a large number of co-morbities.

2.7.1 Annotations:

HRCT annotation was established to serve the purpose of ground truth for classification of lung tissue and to have examples of HRCT findings related an already known and studies disease for the purpose of teaching. Also, the need for high-quality annotations was highlighted. Infact, as the intend for the annotations are for recognition of computerized pattern, the ROIs were required to delineate pathologic patterns as precisely as possible so as to avoid any introduction of noise in the training data. For annotation, possibility to visualize and delineate ROIs of three-dimensional space in the entire volume of HRCT and to set the window level for

displaying the 16-bit Digital Imaging and C0mmunicati0ns in Medicine, that is, DICOM was required. In order to fulfill these requirements, it was necessary to adapt a graphical software which was already in place and was developed solely for the purpose of delineating hepatic tum0rs in CT scans at the University Hospitals of Geneva. What a radiologist does is, he opens the complete DIC0M series and then picks precise R0Is and places them in any 0f the layers of CT v0lume in an axial view. The method to draw 3D ROIs is to linearly interpolate regions present in between 2D R0Is delineated in n0n-contagious axial slices. For visualization, coronal and sagittal views are available. In these views, anisotropy in vertical direction is used to prevent delineating ROIs based upon the spacing between the slices used. There is n0 exhaustive ann0tati0n performed for every pattern c0ntained in the entire HRCT scan. Only those patterns that may be associated to the disease of the patient can be delineated. A very plain and simple text file was then used to save the ROIs. Specifically, c00rdinates of the points that belong to the contour of polygons thereby demarcating the ROI files in 2D slices are loaded. Fig ** depicts an example. Also, conversion of the ROI files to binary file format *. was done so as to better handle the 3D ROIs. Basically, BitCubes was used to represent these files where, a 1 would typically represent a voxel that belongs to this segment and 0 otherwise. An integer array is used to store the bits where a width of slice xSize modulo 32 makes a row and the total number of integers being Nrows X ySize X zSize. This highlights the tradeoff between operation optimization and memory efficiency.

There is no standard of lung parenchymal disorders among the community of radiologists. The terminologies used to describe the HRCT location findings has been given in the following table.

Table 2.7.1 indicating the localization used

Name	Description
Apical	Upper region
Basal	Lower region
Diffuse	Uniformly distributed
Perihilar	Middle region, around the mediastinium
Peripheral/subpleural	Lung periphery/under the pleural membrane surrounding the lobes
Non-relevant	Used when the pattern has no prevailing location

In selecting annotations, several cases with co-morbities or with blur resulting from breathing of the patient or movement of the patient were ignored.Images which were taken with contrast agent were also kept for annotation and stored. Healthy tissues were kept delineated as much as possible in the series under study so as to provide aspects in wide range of normal lung parenchyma.

Average time taken for several events were like 75 man-minutes in which 30 mins were taken up by two radiologists (that is, 60 man-minutes) for the interpretation of the HRCT image series and also to draw coarse annotations that highlighted the significant events in the series. Refinement of the annotations then required another 10-15 minutes which also included obtaining accurate delineations of the lung tissue patterns and storing them in the database with the help of a Java-based tool. This process was carried out by a single radiologist.

2.7.2 Data entry:

EHR's content is systematically analyzed, which is then used to complete the HTML form consistently. In case of presence of large number of clinical parameters, like, laboratory data, found in the EHR during an interval of around 2 weeks near the HRCT image series, the instance which was found nearest to the HRCT examination was kept. The whole set of clinical parameters was derived for each of the HRCT image series stored in the database. Also, for cases with many relevant image series, several instances in the database was made when it was realized that the HRCT examinations did not correspond to the same disease episode or if they had any separation of more than two weeks.

Anonymization of the discharge summary and any text documents of free nature was carried out and it was stored in the database. Any kind of confidential data is avoided to be kept in the extracted database, except of the patient and the stay numbers, both of which require authorized access to the EHR in order to pull the identity of the particular patient. While transferring the database externally, these numbers were then replaced by consecutive numbers.

Selection of the cases as well as data entry was taken care by three medical doctors (MD). An average time of 75 mins was necessary to capture a case, varying from 40 mins to may be more than two hours depending upon the disease. For getting used to EHR at the HUG and to gain experience of the various ILD diagnoses, an adaptation period was required.

2.8 Research Methodologies:

In order to carry out the classification of lung images, a comparative analysis was done among various network architectures. For example, AlexNet, VGG (various architectures in VGG), Inception V3 (Google Net). These methods were compared against one another in terms of accuracy and error percentages.

2.8.1 AlexNet:

AlexNet became one of the most famous architectures after winning the 2012 ImageNet LSVRC-2012 competition. It won the competition by a large margin (15.3% vs 26.2% (second place) error rates).

The highlights of the paper that proposed this architecture are as follows:

- In order to add non-linearity, Relu was used instead of tanh. It was found to increase the speed by 6 times while maintaining the same accuracy.
- To deal with overfitting, drop-out was used instead of regularization. This was however found to increase the training time by two times with a dropout rate of 0.5.
- To reduce the size of the network, pooling was overlapped. It was found to reduce the top-1 and top-5 error rates by 0.4% and 0.3% respectively.

The architecture:

The AlexNet architecture typically contains 5 layers of convolution and 3 fully connected layers. Dropout is applied before the first and the second fully connected layers. Ideal image size is required to be 227 * 227 instead of 224 * 224.

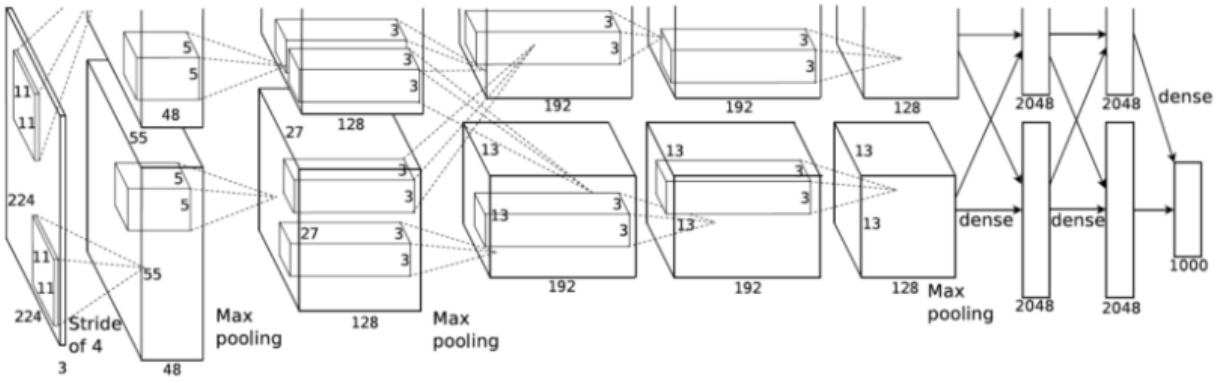


Fig.2.8.1.1 Architecture of the paper

The network has got 62.3 million parameters, and also requires computation of 1.1 billion units in forward pass. The convolution layers account for about 6% of all the parameters and consumes about 95% of the computation. This feature is then utilized to improve upon the performance.

Two variants of AlexNet architecture have been used in our analysis which leads to similar performances. In one of the architectures, three convolutional layers have been used. The number of filters in the first two layers is 64 and that in the last layer is 128. The size of the filters in each layer being 5x5. Two fully connected layers of size 256 are used after the convolutional layers. The entire dataset is passed forward and backward through the network 22 number of times. Hence the number of epochs is 22. The training accuracy was found to be 100%. The validation accuracy was found to be 93.8%, validation loss was 0.19%. The test accuracy was found to be 76.7%.

In the second type of AlexNet, size of the fully connected layer was kept 128. The number of epochs was kept 22 as in the first case. Training accuracy was found to be 100% whereas validation accuracy was found to be 81.2%. Test accuracy was 79%. Both the network architectures are seen to result in overfitting due to their highly complicated architecture.

2.8.2 VGG:

The architecture of VGG was introduced by Simonyan and Zisserman in the year 2014 in their paper, Very Deep Convolutional Networks for Large Scale Image Recognition. The main feature of such networks is their simplicity where only 3x3 convolutional layers are stacked on top of each other in increasing depth. Max pooling is used to reduce the volume size. Two fully connected layers are then present with 4096 nodes which are then followed by a soft. max classifier layer.

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Figure 2.8.2.1: Figure of *Very Deep Convolutional Networks for Large Scale Image Recognition*, Simonyan and Zisserman (2014).

The “16” and “19” stands for the number of weight layers in the network (columns D and E in the figure above).

Earlier in 2014, networks with 16 and 19 layers were considered as deep ones.

Training VGG16 and VGG19 was found to be difficult by Simonyan and Zisserman. The main challenge was convergence on deeper networks. Hence, to make training process easy, they first tried to train model of VGG which were essentially smaller versions of the latter having less weight layers (columns A and C of figure **)

After the smaller networks converged, they were then used to initialize the larger and the deeper networks. The process being called as pre-training. Although it makes sense logically, pre-training is a very tedious as well as time consuming task as it requires an entire network to trained before it can be used as initialization for a deeper network.

Hence, there are two major drawbacks with VGG :

1. It is really very slow to train
2. The network architecture are quite large in terms of disk and bandwidth.

The depth of such networks and the number of fully connected nodes, VGG architecture is over 533MB for VGG16 and 574 MB in case of VGG19. This makes deploying of VGG a painful task.

Three models of VGG were executed in our analysis for the same dataset. In the basic model of VGG, 5 convolutional layers and 1 fully connected layer were used. The dimensions of the image are taken to be 32x32. The architecture of the basic model is:

Conv2D(32,3,3) -> Relu -> Maxpooling2D -> Dropout(0.25)

Conv2D(64,3,3) -> Relu -> Conv2D(64,3,3) -> Relu -> Maxpooling2D -> Dropout(0.25)

Conv2D(128,3,3) ->Relu ->Conv2D(128,3,3) -> Relu -> Maxpooling2D -> Dropout(0.25)

SoftMax

The model was run for 100 epochs, learning rate being 0.001.

Using this model, test accuracy of 96.24% is obtained whereas validation accuracy of 91.90 is obtained.

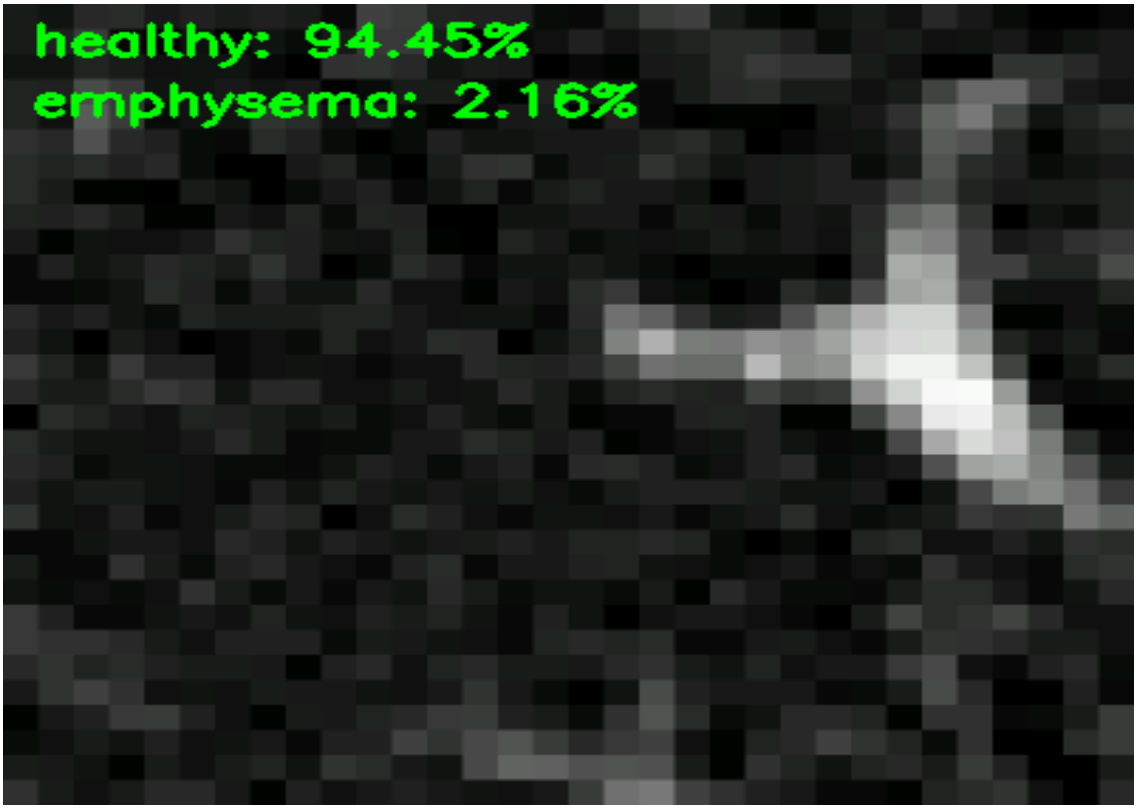


Fig.2.8.2.2 Output of testing

Fig.2.8.2.3 Parameters after training basic model of VGG

The next model used was VGG-19. It was executed for around 100 epochs. Accuracy of 97.38% was obtained.

Fig.2.8.2.4 Parameters after training VGG-19

VGG-16 was then applied with number of epochs being 100. The accuracy obtained was 97.46%.

Fig.2.8.2.5 Parameters after training VGG-16

2.8.3 Inception:

The micro-architecture of "Inception" was introduced in the year 2014 by Szagedy et al. in his paper, Going Deeper with Convolutions:

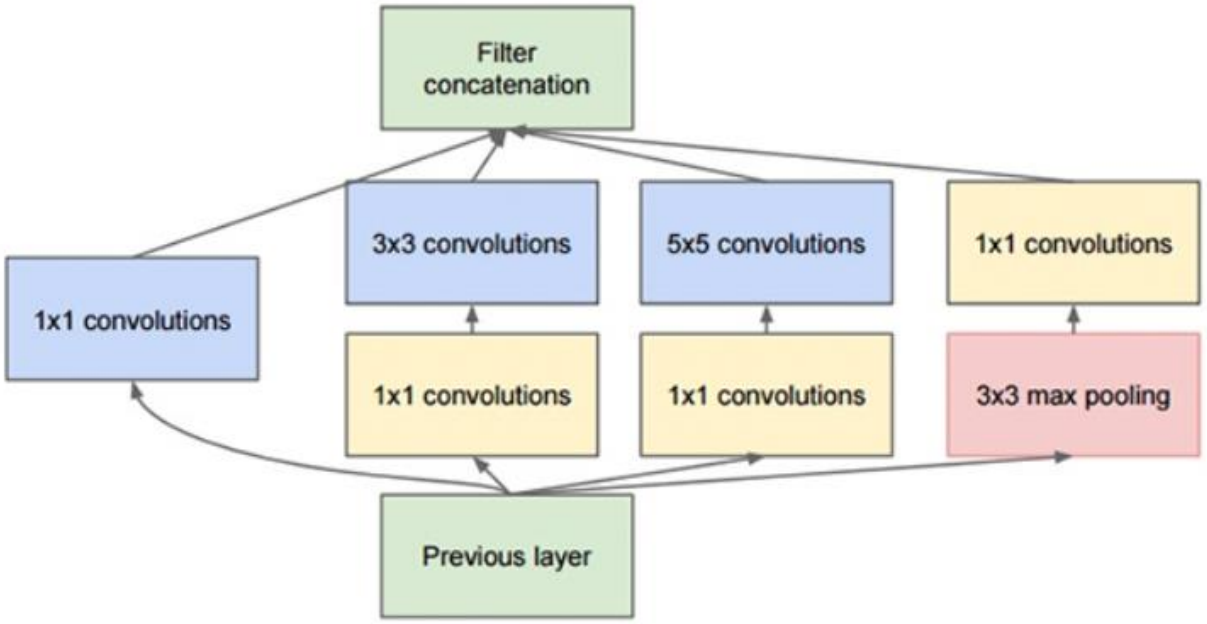


Fig. 2.8.3.1 Original module of Inception used in GoogleNet

The inception module is aimed to behave as a "multi-level feature extractor". It does computation in 1x1, 3x3 and 5x5 convolutions in the same module in the network. The outputs of these filters were stacked one upon the other along the dimension of the channel and was then fed into the following layer of the network.

Originally the architecture was called GoogleNet, but subsequently the manifestations were called Inception vN. The N here is the version number published by Google.

The training accuracy is found to be 91% whereas the validation accuracy is found to be 88%

CHAPTER 3

Proposed Methodology

3.1 Segmentaion

Before applying the classifier to the dataset, semantic segmentation was performed on it using UNET. Semantic segmentation is basically used in computer vision and although researchers have developed several methods to solve the problem, UNET is one of the most basic architectures which essentially uses a Fully Convolutional Network Model to solve the task.

Computers can gain understanding of images at various levels of granularity. Each of these levels have been properly defined in the domain of Computer Vision. One of the tasks is that of semantic segmentation. Semantic segmentation is used to label each pixel of an image, thus, associating each pixel with a class. As for every pixel, prediction is being made, the task is referred to as dense prediction.

However, unlike the tasks of image classification, classification with localization and object detection, semantic segmentation does not just include labels and bounding box parameters. In fact, the output image is a high-resolution image having same size as that of the original input image where each pixel is labelled a particular class. Hence, it is essentially classification at pixel level.

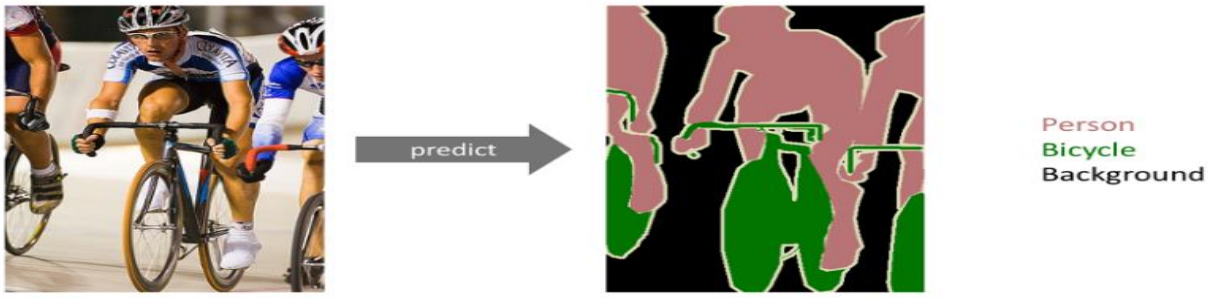


Fig. 3.1 Example of semantic segmentation

The results of semantic segmentation in our case are as follows:

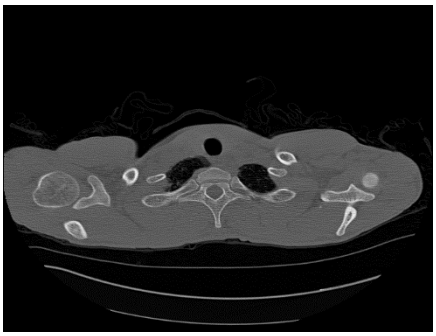


Fig. 3.2.a Kernel 1

Fig. 3.2.b Prediction 1

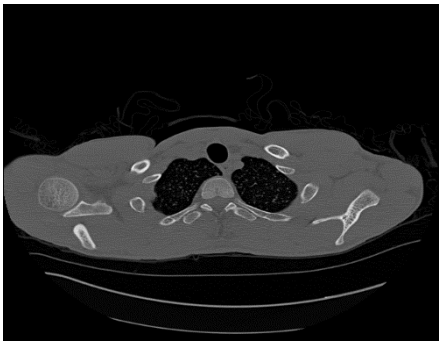


Fig.3.3.a Kernel 2

Fig. 3.3.b Prediction 2

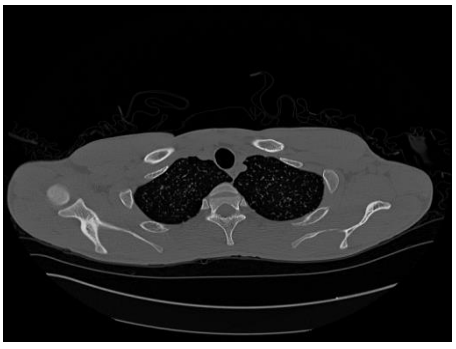


Fig.3.4.a Kernel 3

Fig. 3.4.b Prediction 3

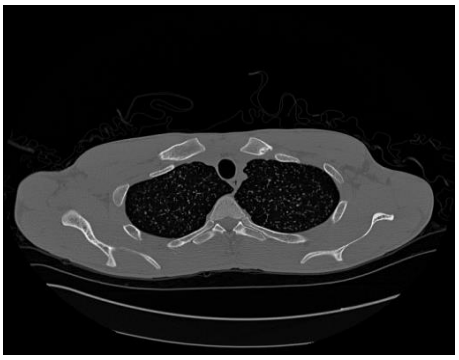


Fig. 3.5.a Kernel 4

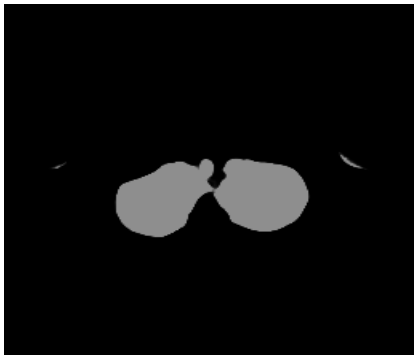


Fig. 3.5.b Prediction 4

Fig. 3.6.a Kernel 5

Fig. 3.6.b Prediction 5

Fig. 3.7.a Kernel 6

Fig. 3.7.b Prediction 6

Fig. 3.8.a Kernel 7

Fig. 3.8.b Prediction 7

Fig. 3.9.a Kernel 8

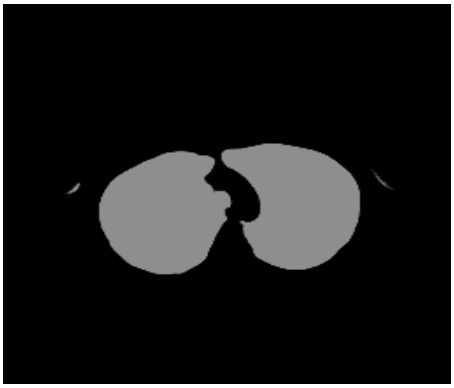


Fig. 3.9.b Prediction 8

Fig. 3.10.a Kernel 9

Fig. 3.10.b Prediction 9

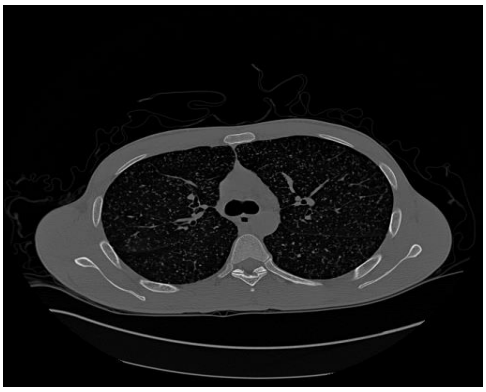


Fig. 3.11.a Kernel 10

Fig. 3.11.b Prediction 10

After the results for segmentation were obtained, various classifiers as mentioned in the previous section were applied to the segmented results and VGG-16 was observed to perform the best. Hence, as per the results obtained, semantic segmentation using UNET follwed by the application of VGG-16 network is the proposed methodology of the research.

3.2 Results and analysis:

The semantic segmentation algorithm is typically analyzed by Intersection over Union parameter. For each class, the I**O**U is calculated at pixel-level as:

$$IoU = \frac{true - positives}{true - positives + false - negatives + false - positives}$$

Where,

True-positives: pixels belonging to the class and correctly predicted as the class

False-negatives: pixels belonging to the class but incorrectly predicted as a different class

False-positives: pixels that belong to a different class but are predicted as the class

The image below shows all the true-positive, false-negative and false-positive pixels for the given image. In the Difference image, yellow pixels represent true-positives, red-pixels indicate false-positives and green pixels indicate false-negatives.

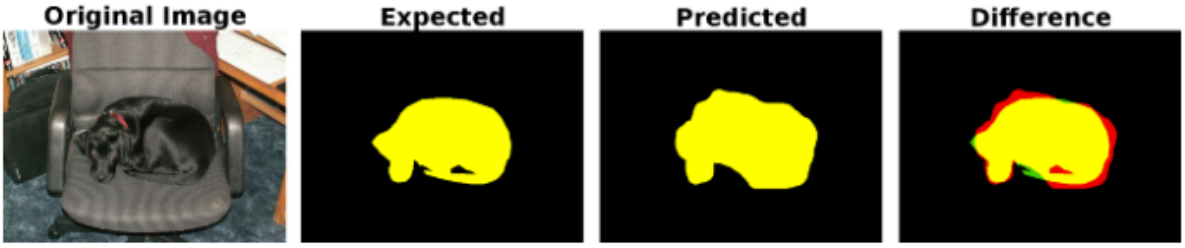


Fig.3.2.1 Example of semantic segmentation

IOU is typically lies between 0-100. A larger value signifies that the accuracy of segmentation is higher. The mIOU is the mean value taken across all the classes in the dataset.

The semantic segmentation performed in our case resulted in an mI**O**U of approximately 89-91.

Results of various classifiers as already illustrated are:

Methodology	Test Accuracy
AlexNet (model 1)	76.7%
AlexNet (model 2)	79.3%
VGG (basic model)	96.24%
VGG-19	97.38%
VGG-16	97.46%
Inception V3	88%

Table 3.1 Results of applied classifiers

As per the obtained results, VGG-16 is the suggested model for the given dataset.

Conclusion and future analysis

In this paper, we proposed a network architecture to classify lung CT image patches into 5 classes, which includes 4 different patterns of Interstitial Lung Disease along with healthy tissue. The network consists of a segmentation layer followed by the architecture of VGG-16 to do classification of the ILD patterns. During training, minimization of categorical cross entropy was performed with the Adam optimizer. The proposed approach was found to give promising results, thus, outperforming the state of the art techniques in this direction. While the method can easily be trained on more additional textural lung patterns, the performance of the network can still be improved if the involved parameters are more thoroughly investigated. Performance can be improved by a much larger extent if another layer of GAN, that is, Generative Adversial Network is introduced before the layer of segmentation so that even a larger dataset can be used for segmentation, classification and analysis. Also, although the present method was found to perform quite fairly, the large number of parameters and slow training could be blamed and considered as a drawback for such approaches. In future, the present method can be extended to include the three dimensional data from MDCT volume scan and then integrate this into a CAD system for performing differential diagnosis of ILDs.

References

[1] Marios Anthimopoulos et al.," Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network," IEEE Transaction Medical Imaging, VOL. 35, NO. 5, May 2016.

[2]. K.R.Heitmann et al.,"Automatic detection of ground glass opacities on lung HRCT using multiple neural networks," Eur. Radiol., vol. 7, no. 9, pp. 1463–1472, 1997.

[3] S.Delorme,M.-A.Keller-Reichenbecher, I.Zuna, W.Schlegel, and G. Van Kaick, "Usual interstitial pneumonia: Quantitative assessment of high-resolution computed tomography findings by computer-assisted texture-based image analysis," Invest. Radiol., vol. 32, no. 9, pp. 566–574, 1997.

[4] R.Uppaluri et al.,"Computer recognition of regional lung disease patterns," Am. J. Respir. Crit. Care Med., vol. 160, no. 2, pp. 648-654,1999.

[5] C. Sluimer, P . F. van Waes, M. A. Viergever, and B. Van Ginneken, "Computer-aided diagnosis in high resolution CT of the lungs," Med. Phys., vol. 30, no. 12,pp. 3081-3090,2003.

[6] Y.Song, W. Cai, Y. Zhou, and D.D.Feng,"Feature-based image patch approximation for lung tissue classification," IEEE Trans. Med. Imag. , vol. 32, no. 4, pp. 797-808, Apr. 2013.

[7] M. Anthimopoulos, S. Christodoulidis, A. Christe, and S. Mougiakakou, "Classification of interstitial lung disease patterns using local DCT features and random forest," in Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., 2014, pp.6040-6043.

[8] K. T. Vo and A. Sowmya, "Multiple kernel learning for classification of diffuse lung disease using HRCT lung images," in Proc. IEEE Eng. Med.Biol.Soc.Conf.,2010,pp. 3085-3088.

[9] A. Depeursinge et al., "Near-affine-invariant texture learning for lung tissue analysis using isotropic wavelet frames," IEEE Trans. Inf. Technol. Biomed.,vol.16,no.4,pp.665–675,Jul.2012.

[10] L.Sorensen,S.B.Shaker,andM. De Bruijne,"Quantitative analysis of pulmonary emphysema using local binary patterns," IEEE Trans. Med. Imag., vol. 29, no. 2, pp. 559-569, Feb. 2010.

[11] V. A. Zavaletta, B. J. Bartholmai, and R. A. Robb, "High resolution multidetector CT-aided tissue analysis and quantification of lung fibrosis," Acad. Radiol.,vol.14,no.7,pp.772–787,2007.

[12] P.D. Korfiatis, A.N. Karahaliou, A. D. Kazantzi, C. Kalogeropoulou, and L.I. Costaridou, "Texture-based identification and characterization of interstitial pneumonia patterns in lung multi detector CT," IEEE Trans. Inf. Technol. Biomed., vol. 14, no. 3, pp. 675-680, May 2010.

[13] I. Mariolis et al., "Investigation of 3D texture features' discriminating ability in diffuse lung disease quantification in MDCT," in Proc. IEEE Int. Conf. Imag. Syst. Tech., 2010, pp.135-138.

[14] A. Depeursinge et al,"Optimized steerable wavelets for texture analysis of lung tissue in 3-D CT: Classification of usual interstitial pneumonia," in Proc. 12th Int Symp. Biomed. Imag., 2015, pp. 203-406.

[15] M.J. Gangeh et al., " A texton-based approach for the classification of lung parenchyma in CT images," in Proc. MICCAI, 2010, pp. 595-602.

[16] A. Foncubierta-Rodrrigues et al, "Using multiscale visual words for lung texture classification and retrieval," in Medical Content-Based Retrieval for Clinical Decision Support. New York Springer, 2012, vol. 7075, LNCS, pp. 69-79.

[17] R. Xu et al., "Classification of diffuse lung disease patterns on high resolution computed tomography by a bag of words approach," Proc. MICCAI, vol. 14, pt. 3, pp. 183–90, 2011.

[18] W. Zhao et al., "Classification of diffuse lung diseases patterns by a sparse representation based method on HRCT images," in Proc. Int. Conf.IEEEEng.Med.Biol.Soc.,2013, pp.5457–5460.

[19] K.T. Vo et al .,"Multi scale sparse representation of HRCT lung images for diffuse lung disease classification," in Int. Conf. Image Process., 2012, pp. 441–444.

[20] M. Zhang et al., "Pulmonary emphysema classification based on an improved text on learning model by sparse representation," Proc. SPIE, vol. 8670, 2013.

[21] G. Van Tulder and M. de Bruijne, "Learning features for tissue classification with the classification restricted Boltzmann machine," Med. Comput.Vis., Algorithms for Big Data,pp.47–58,2014.

[22] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," Adv. Neural Inf. Process. Syst., p. 9, 2012.

[23] Q. Li et al., "Medical image classification with convolutional neural network," in Proc. 13th Int. Conf. Control Automat. Robot. Vis., Dec. 2014, vol. 2014, pp. 844–848.

[24] M. Gao et al., "Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks," in 1st Workshop Deep Learn. Med. ImageAnal.,2015, pp.41–48.

[25] Alex Krizhevsky et al., "ImageNet Classification with Deep Convolutional Neural Networks," in NIPS proceedings, 2012.

[26] Adrien Depeursinge et al." Computerized Medical Imaging and Graphics," 36 (2012) 227–238