**VICTOR BABEȘ UNIVERSITY OF MEDICINE
AND PHARMACY TIMIȘOARA
FACULTY OF MEDICINE
DEPARTMENT XIII – INFECTIOUS DISEEASE**

POPOVICIU ANA ADRIANA

PhD THESIS

COMPUTER-AIDED IMAGISTIC DIAGNOSIS IN DIFFUSE INTERSTITIAL LUNG DISEASES

A B S T R A C T

Scientific Coordinator
PROF. HABIL. OANCEA CRISTIAN IULIAN, MD PhD

**T i m i ș o a r a
2 0 2 2**

TABLE OF CONTENTS

RESEARCH PART

ABSTRACT

GENERAL PART

Pulmonary interstitial pathology presents a particularly complex problem caused by the diversity of pathological entities that make up this group. Although these entities have low individual prevalence and are considered rare diseases, they represent an essential part of pulmonary medicine practice; above 200 disorders can lead to interstitial lung injury, often with similar presentations, presenting difficulties in diagnosis and treatment, often requiring a multi-disciplinary approach to the case. The difficulty of diagnosing them can arise primarily in the imaging domain - a central analysis element. The broad spectrum of imaging patterns and the presence of a similar appearance of different types of tissue or, conversely, imaging variations of the same type of tissue/lesion create diagnostic variability among specialists in the field. This variability can lead, together with the rarity of cases and the difficulty of specifying a diagnosis, to grave confusion in classifying a patient in a disease pattern even after multi-disciplinary discussion sessions between medical specialists.

Because of both case management and treatment response variability, the gold for the pulmonologist is to identify the distinguished diffuse interstitial lung diseases (DILD) forms in the early stages as possible, without an invasive procedure such as a lung biopsy. Within this framework, high-resolution computed tomography (HRCT) remains the predominant method for ILD diagnosing due to lung tissue-specific radiation attenuation properties and maximum spatial resolution.

Idiopathic pulmonary fibrosis (IPF) is the most often encountered diffuse interstitial lung disease (DILD), a progressive fibrosing interstitial lung disease (PF-ILD), with a characteristically poor outcome and an amplified early death risk without treatment. There is particular interest in the pulmonology community for this pathology and, more recently, for all DILD with progressive fibrotic character.

In the last decade, substantial progress has been made in understanding pathogenic mechanisms, defining diagnostic criteria, and developing effective medication for treating DILDs. Computer-aided diagnosis (CAD) is often employed in DILD management during the diagnosis and treatment phases. However, CAD has several facets that need careful consideration: selecting the proper technique for the medical needs and conveying the answers in a manner accessible to persons without an information technology (IT) background. More than a few approaches to CAD for lung HRCTs are available or in development based on different techniques. Whether built on artificial intelligence, neural networks, or machine learning, these software applications fail to capture the dynamics of a pathology evolution. Therefore a math-based method of CAD image visualization could be the answer.

RESEARCH PART

CHAPTER 1: FACILITIES AND PITFALLS FOR USING CAD FOR DILD DIAGNOSIS

1.1 1.1. DEEP LEARNING IN INTERSTITIAL LUNG DISEASE-HOW LONG UNTIL THE DAILY PRACTICE

The artificial intelligence (AI) component's virtual subclass is machine learning comprising mathematical algorithms computer systems use to learn a specific task through experiences without specific human instructions. Deep learning (DL) is advanced, consisting of a multi-layer representation learning architecture. The representation activates the first layer of neurons through a sensor, which, in turn, activates the next layer by complex connections. Each layer processes the representation in a non-linear way, creating an increasingly complex schema, diverging from the general machine learning task-specific algorithm. DL'smajor advantage is that it can improve autonomously, without human input. From a usage standpoint, it can perform arbitrary parallel computation more efficiently than other algorithms. DL is used in visual object recognition, speech recognition, driving assistance, and language classification, among others. Complex neuronal networks (CNN) is an AI technique, the engine on which the DL runs.

Typical DILD patterns in HRCT images are reticulation (RE), honeycombing (HC), ground-glass opacity (GGO), consolidation (CD), micro-nodules (MN), emphysema (EM), or combinations of the above. The difficulty appears when the results are combined or inconclusive.

CNN's accuracy needs large image samples because normal lung or tissue categories could exhibit similar appearances, and significant variations might be seen between different subjects for the same tissue class. Thus, CNN's require large, balanced datasets and advanced algorithms, reflecting processing power and storage capacity requirements. A combination of multiple CNN can be the answer to reducing the costs spent on the social and healthcare aspects.

1.2. COMPUTER AIDED TEHNICS

The CAD algorithms are part of artificial intelligence (AI) since they imitate human thought. CAD can be broken down into two categories: learning and discovery. Both can be carried out under supervision or independently; nevertheless, the results of computer techniques are primarily data-based.

If we are looking to implement a diagnosis model that recognizes existing patterns, learning algorithms are a natural choice because the algorithms learn from the offered data. However, if a new diagnosis parameter is sought, the discovery section provides the algorithms. Reasoning algorithms are used when we need a consensus between inputs and rules, sometimes allowing for uncertainty or using statistical inferences. Therefore, these algorithms help implement a diagnosis or treatment algorithm consisting of clear rules that quantify the inputs.

As their name suggests, supervised machine learning consists of training a model by feeding it a set of input data together with the expected corresponding output values (which are known beforehand). The algorithm then generates a suitable model (formula) which fits the input data and can be used for the analyzing new input data

Two of the most common supervised learning purposes are linear regression and classification. With *unsupervised* machine learning, no training set tells the algorithm how it should generate results; instead, the algorithm is responsible for finding common denominators among data. They are mainly used for clustering purposes, anomaly detection, and neural networks. Some of the most popular unsupervised ML algorithms are K-means clustering, hierarchical clustering, DBSCAN clustering, and the hidden Markov model. *Reinforcement* learning (RL) algorithms are based on a trial-and-error approach. The learner is not told what to do but instead learns what good and bad actions are based on the rewards or penalties it gets according to its actions. Thus, it will always choose the moves which allow it to maximize the rewards.

These types of algorithms can be used in medical diagnosis in combination with medical imaging in cases where doctors might be dealing with a prolonged therapy process. Reinforcement learning algorithms have multiple applicability directions, such as DILD, dynamic treatment regimes, automated medical diagnosis, or more general domains.

With the aid of information technology (IT) procedures, CAD allows medical professionals to comprehend and employ distinct imagistic investigations. The objective is to increase the speed and accuracy of diagnosis, with IT as a supplement or even an independent diagnostic alternative.

CHAPTER 2: A NOVEL METHOD FOR LUNG IMAGE PROCESSING USING COMPLEX NETWORKS

2.1. MATERIALS AND METHODS

2.1.1. Lot Selection

To choose the eligible patients, we used ‘Dr. Victor Babes’ Infectious Diseases and Pneumoftiziology Clinical Hospital Timisoara database, stored in their private cloud archive. From more than 30000 imaging exams stored in Digital Imaging and Communications in

Medicine (DICOM) format, a total of 60 scans were selected based on the following inclusion criteria:

- 30 patients with CT exams and exploratory function tests with the diagnosis of DILD (diffuse interstitial lung disease);
- 30 patients with normal CT imaging were considered the control group.

2.1.2. Imaging Parameters

All examinations were performed with a General Electric (GE) Healthcare Optima 520 16 slices with 32 slices reconstruction. The scanner is a 0.5 mm × 16 detector row, allowing for an 8 mm total z-axis length. The slice is narrower than the recommended 1.5 mm by the Radiology Working Group of the Pulmonary Fibrosis Foundation to allow for better and smoother lesion detection and higher accuracy—both crucial aspects of DILD diagnosis. The spatial resolution (pixel spacing) for these settings is 0.74 mm.

The main criteria for analyzing image data were the tissue densities/opacities, which were determined by applying the Hounsfield scale's principles. The Hounsfield units (HU) are commonly used for quantitative analysis of radio density and tissue tightness, which help interpret CT scans. Image reconstruction relies on the tissue properties regarding X-ray beam penetration and attenuation to define a grayscale image system. These grayscale intervals vary between approximately −1000 HU (air) and 3000 HU (metals like silver and steel), according to the attenuation range of tissue absorption. A gray tone scale represents this transformation with the density of distilled water as a landmark, defined as zero HU. According to the HU intervals, each element of this lesional picture will have an equivalent.

2.1.3. Image Processing Algorithm

Each sample manually selected is then processed with the help of a Python-written program developed specifically for this purpose. Using a specialized CT library, pydicom, the DICOM slices are cropped to the pre-established size (65 × 65 pixels) around the interest area.

The program consists of an algorithm meant to carry out the following steps:

1. Iterate over a set of HRCT slices (DICOM files);
2. For each one, crop out a 65 × 65 pixel area;
3. Analyze the selected area from 3 perspectives:

- Convert pixel gradient into a Hounsfield unit value according to the formula:

  HUv = rescaleSlope * PxGradient + rescaleIntercept,

- Where rescale slope and rescaleIntercept are constant values dependent on the CT equipment and embedded in the DICOM metadata, and PxGradient is the color code of a pixel;

- Isolate all emphysema-like tissue(E-considered to be equivalent to cyst by HU), GGO (Ground Glass Opacity), and consolidation densities(C-equivalent with reticulation) in the cropped image and leave out any other types of tissue;
- Separate each HU strip in the sample into a separate layer(E, GGO, C).

4. Generate complex networks out of each layer;
5. Analyze connectivity, closeness, and distribution of nodes (pixels);
6. Determine patterns of normal lungs and affected lungs.

2.2. RESULTS

Following the previously described method, all HRCTs (of both normal and affected lungs) were processed. Each individual steps are done for a single normal and DILD-affected patient.

2.2.1. Normal and DILD Case Sample Results

The first step is sample cropping into 65 × 65 pixels. The following steps imply splitting everything into layers and converting those layers into complex networks. First, the emphysema layer is examined. Next is the ground glass layer, where significant differences occur. Even though a visual inspection might evaluate the distributions as random, the network degree distribution shows an entirely different story: a logarithmic distribution for the normal process and a polynomial one for the IPF. The least was the consolidation layer.

The differences can be pretty distinctive at an individual level, and the entire image lot analysis presented the challenge of determining network metric relevance in a broader context. In order to measure the network invariant entropy, a metric based on degree sequences is usually preferred. However, the differences shown present the challenge of adding a measurement for the network size. Three metrics were selected due to their balance between metrics that measure network complexity and size: total count (the degree sum), average count (average degree), and maximum degree.

2.3. DISCUTION

The complex-network model based on HRCT lung imaging needs to be assessed as to how well that model fits known frameworks from network system science and medical science.

2.3.1. Network System Science

Results showcase a logarithmic distribution at the proper biological resolution (Rd = 4) for normal patients. Pathological lungs have an entirely different distribution, best fitted with a polynomial function, not a logarithmic one. The fit of various logarithmic and power

distributions was tested against relative distances between lung entities. Values less than 3 show a relatively similar fit, which is mathematically correct yet biologically incorrect because 1- and 2-pixel separation translate to a 0.74 mm to 1.48 mm gap, too small to be relevant.

The differences between the DILD-affected and normal networks are distinctive and can be further quantified if a simple standard deviation for all patient data series is computed. The results on all three measurements considered for the networks (maximum degree, total count, and average degree) for each HU band and the combined pathological HU bands prove a clear separation between the pathological and normal networks.

2.3.2. Medical Science

From a statistical perspective, comparing normal lungs with diseased lungs is challenging due to different DILD phenotypes and the relatively small lot size/disease class. To prove the method and model work overall, a t-test: two-sample assuming unequal variances was conducted, comparing normal to DILD samples. The results show that measured p is less than 0.05 (3.97×10^{-17},8.52×10^{-23}, and 5.31×10^{-9}) and observed t (10.49, 14.91, and 6.29) is more significant than critical t (1.98, 1.99, and 1.98), therefore rejecting the null hypothesis; i.e., being 95% confident that the differences between groups are not due to chance.

2.4. CONCLUSION

In this paper, a novel method of using complex networks to transform lung HRCT has been presented. The methodology section delves deeper into the algorithm steps and the justification of each chosen parameter. The sample size is justified by the anatomical bounds of the secondary pulmonary lobule; the radius influencing network connectivity is correlated with injury granularity, and the Hounsfield unit intervals depend upon the device and resolution. The results section presents in parallel the processing steps for two sample patients (a normal and a pathological one) and a whole-lot perspective. In the discussion section, the correctness of this model is justified from a system science perspective by using the degree distributions as the main system characterization tool.

Furthermore, the network measurement clusterization is described, showing that it results in evident disparities between the normal and pathological lots. From a medical science perspective, it is showcased how the chosen model reflects clinical data and how its low granularity presents an advantage in the diagnosis process. In the end, comparing this method with existing ones highlights its advantage: it offers a complex qualitative and quantitative measurement. Pitfalls of the proposed model include its inability to work alone yet and the relatively small lot on which it was tested, which will all need to be addressed in further research. In conclusion, the stated goal is considered to have been achieved by

showing how a complex network model can be used to transmute lung HRCT into a quantifiable and qualifiable structure that can enhance the DILD diagnosis.

CHAPTER 3: ENHANCING IMAGISTIC INTERSTITIAL LUNG DISEASE DIAGNOSIS BY USING COMPLEX NETWORKS

3.1. MATERIALS AND METHODS

3.1.1. Lot Selection

From the private "Dr. Victor Babes," Infectious Diseases and Pneumoftiziology Clinical Hospital Timisoara National Fibrosis Center database were selected 65 DILD patients with multiple scans and 31 normal lung patients. Inclusion and exclusion criteria were established and respected. For each patient's physiological data (age, sex, smoking status), pulmonary function tests (PFT)–like forced vital capacity (FVC) by spirometry performed and diffusing capacity of the lungs for DLco together with HRCT annotations were investigated. Patients' quantitative dynamic HRCT images were also provided, and four pneumology specialists reviewed their case history.

The selected primary lesions were reticulation and consolidation (defined together as band C), ground glass opacity (band GGO) as well as emphysema, and cysts (defined together as band E). These lesions have precise imagistic absorption rates that permit grouping.

The HRCT region of interest was marked by a radiologist with high experience in the imagistic diagnosing of DILDs (10+ years), who collaborated with the other specialists' inputs. The selected imagistic elements were typical for IPF (29 patient—44.62%), NISIP (16 patients— 24.62%), OP (8 patients—12.3%), S (8 patients—12.3%) and HP (4 patients—6.15%).

From the DICOM format, three complex networks were generated for each selected region of interest, one for each pathologically relevant Hounsfield unit (HU) interval: E for emphysema and cysts, GGO for ground glass opacity, and C for consolidation and reticulations. The HU transformation scale is device-specific and based on this implementation.

3.1.2.Selecting relative measurements

A CN can be characterized by many metrics and should reflect the underlying biological processes and their dynamic evolution. Since the underlying purpose of this paper can be biologically translated into a way to measure lesions and their expansion. Therefore, the selected measurements to reflect interconnectedness and size are maximum degree number (the number of the maximum connection in the network for a singular node), total degree count (how many connections are in the network), and average degree count (the

average number of connections per node–how sparse the network is). A network node can represent either a singular pixel.

3.1.3. Results

3.1.3.1. Case Reports

This section presents sample locations from two very different patients put through the analysis process(UIP+ emphysema and NSIP). A patient classified as a typical UIP following a heated discussion among our fibrosis center specialists presents an untypical honeycombing pattern, which may skew the diagnosis towards probable UIP. However, age and sex leaned heavily toward the final diagnosis. Therefore, this case and an NSIP pattern were tested to detect the capability of the studied algorithm.

3.1.3.2. Progression Speed

The defined relative speed on each HU band and each CN parameter was analyzed with a t-test versus DLco relative variation. The lot on which this test was performed is the entire lot, normal and DILD patients. It should be noted that, while the maximum degree can also be analyzed since the measurement searched for is progression, peak singular lesion is less relevant. The null hypothesis is retained for all but one of the selected series. The E band's average count VS DLco test rejects the null hypothesis.

3.1.3.3. Testing for Early Detection

To search for early detection, the lot was grouped into cases considered normal and cases with incipient DILD and fairly good functional parameters (GAP-ILD 0–3 points, DLco values between 70 and 85%). The DLco values were chosen as an interval centered on the lower normal limit (80%) to allow the inclusion of early impairment in alveolar-capillary membrane. The cases were analyzed on the same three axis

3.1.4. Discussion

Two levels of axial HRCT slices (superior and basal lung region) are presented, selected to showcase debatable UIP pattern+ emphysema (CPFE phenotype) imagistic progression. Imagistic interpretation for the progression starts with the initial t0 point, which, in the superior lung region, indicates fine reticulation presence, bullous emphysema, and slight subpleural honeycombing cysts; the basal region is marked with sparse reticulation and honeycombing lesions.

According to the HU ranges, reticulations and consolidations have similar values, yet in this specific context, the values are interpreted as reticulations. The CN model offers data for relative variation speed on each layer in the selected areas. This speed is specific to a

selected site and reflects a relative variation in characteristics over a time period. It is not an absolute value; its meaning is related to the swiftness of change, highlighting rapidly deteriorating areas. Since the algorithm behind the CN conversion considers lesions as small as 3 mm, by default, the speed is more granular than the human eye.

The CN model's relative speed on the E layer presents an increase in follow-up in the year 1 and year 2, yet the magnitude between the superior and basal slices is very different. The superior region is almost 10 times faster deteriorating than the basal slice, quantifying the superior lobe's emphysema lesion extension and honeycombing cyst layers increase compared with the basal lobe in which emphysema is not very well. C layer increases on the superior and basal slices, presenting the pathological process of lesion progression with lung architectural distortion, reticulation, and multi-layer variate size cyst. The model detects minor variations in the GGO, especially in the basal plane, suggesting a probable acute substrate in that area. This image is highly annotated by lung experts, with GGO difference imperceptible. Studying the patient's data, the symptoms from follow-up year 1 are inexplicably slightly exacerbated, yet they are not so in follow-up year 2. This confirms the CN relative speed light variation and its ability for early detection. Functional parameter relative variation is almost zero in both follow-up years, defining a stationary functional status, underling the early detection of the proposed CN model.

Then, the imagistic axial lung HRCT lesion evolution in an NSIP pattern case is evaluated. On the E band, relative speed expresses a marked increase in the emphysema focus points numbers (total count), with only a medium increase in their intensity (average), for both sample sites, clearly explained by the buildup in honeycombing cysts layers. GGO in t_0 shows a slight increase in the follow-up sample, corresponding with the imaging slice HRCT interpretation. The C layer displays only on superior regions a slight increase, reflected by the well-defined multi-layer cysts and their defining walls. Functional parameters have almost no variation underlying the premature detection of the proposed CN model.

Screening the entire lot, results support the state that the CN algorithm accurately and quantitatively characterizes DILD progression. The fact that most of the statistical comparisons between DLco and CN measurements variation show relevant similarities. The only exception is comparing the average count and DLco on the E band. Some patients classified as normal have chronic obstructive lung pathology in a clinical compensation status and/or are active or former smokers. Since the CN measurements reflect biological terms, the number of the E-layer regions of interest is the same, but the regions' median intensity is statistically relevant and higher than its corresponding functional parameter variance.

The statistical testing between the borderline and normal groups warrants further exploration. On the E layer, there is no statistical difference between the early diagnostic and

normal sets; therefore, the CN model does not allow early detection on this layer. From a biological perspective, early DILD diagnosis with emphysema phenotype is almost identical to smokers' emphysema lesions, as confirmed by the results. On the GGO layer, there is a statistical difference, the null hypothesis is rejected, and the proposed model is successful in early DILD detection. On the C band, the maximum degree and total count detect early DILD, yet the average count does not. Pathologically, the proposed model accurately detects well-defined consolidation lesions and does not successfully differentiate diffuse early consolidations with blurred edges in their early stages. Consequently, the hypothesis that the CN algorithm allows early detection is accurate on the GGO, primarily true on the C layer and false on the E layer.

Previous CAD approaches, like the ones that implement simple mathematical based techniques in one or more dimensions or more complex machine and deep learning algorithms, do not provide a way to objectively assess the aggressive aspect of a lung disease that can serve as an indicator for the commencement of the antifibrotic protocol. The proposed speed measurement does not assess the disease severity, yet it assesses its aggressive aspect. For example, a simple <insert disease here> in its early stages can progress rapidly, and the measured speed is high. In this paper, the superior region, although it has a less severe aspect, deteriorates faster, which is quantified by the speed measurement. A more severe aspect can be pretty stationary, a sign that there is another factor to be considered (the medication is working, the phenotype is slowly progressive, the disease is remissive, or it shifted towards other areas).

3.1.5. Conclusions

To successfully deal with DILDs, two issues need to be solved, well known by all the practicians: early detection and accurate progression evaluation. So far, the traditional medical and computer-based approaches based on artificial intelligence, machine learning, etc., have come up short even though some diseases, such as IPF, critically need efficient solutions. This paper aimed to explore whether a CN-based computer-aided diagnosis can successfully provide the required data to manage DILDs.

In order to do so, two hypotheses were tested: the first one explored progression, and the second one was early detection. For progression, the CN CAD was an almost complete success. Its fine accuracy in testing lesions as small as 3 mm, allowed correlation with the clinical status beyond the granularity of standard functional tests. The only problem was on the E band for the average count measurement type, yet the other five measurement axis easily offsets this.

For early detection, the inflammation GGO layer proved to be key. Inflammation and fibrosis are the two typical DILD states, and the CN algorithm performed well on both GGO

and C-defined HU bands. This showcases the practical abilities of this algorithm type, particularly well-suited to DILDs, not filled so far by any other tools, such as, for example, Caliper.

As pitfalls, the CN algorithm has a considerable run-time, growing exponentially proportional to the analyzed window. It also needs prior lung segmentation, which can be obtained through other CAD or manually. In conclusion, this algorithm should be incorporated in a much larger CAD, combining the faster machine learning segmentation and pattern detection capabilities with the slower yet accurate CN local analysis.