Implementing a System for Diagnosing Pulmonary Fibrosis using Hough Algorithm

Cristina-Maria Stancioi
Automation department
Technical University of Cluj-Napoca
Cluj-Napoca, Romania
stancioi_cristina@yahoo.com

Iulia Clitan
Automation department
Technical University of Cluj-Napoca
Cluj-Napoca, Romania
iulia.clitan@aut.utcluj.ro

Abrudean Mihai
Automation department
Technical University of Cluj-Napoca
Cluj-Napoca, Romania
mihai.abrudean@aut.utcluj.ro

Vlad Muresan
Automation department
Technical University of Cluj-Napoca
Cluj-Napoca, Romania
vlad.muresan@aut.utcluj.ro

Mihaela-Ligia Unguresan
Physics and chemistry department
Technical University of Cluj-Napoca
Cluj-Napoca, Romania
mihaela.unguresan@chem.utcluj.ro

2021 IEEE 15th International Symposium on Applied Computational Intelligence and Informatics (SACI) | 978-1-7281-9544-5/20/$31.00 ©2021 IEEE | DOI: 10.1109/SACI51354.2021.9465554

***Abstract*—The main objective of this paper by using image processing algorithms, mainly Hough algorithm, to develop a system which is able to analyze, process and put in evidence all the necessary features of a CT image for diagnosing pulmonary fibrosis. To approach the presented topic, image processing algorithms, image filtering, together with the Matlab work environment were combined and the optimal solution was finally implemented. The main idea behind the method is to identify two types of lines of the CT image which gives to the specialist the correct diagnostic by interpreting the obtained results. As final results, the chosen solution incorporates some crucial steps which have to be done in order to obtain the desired processed image with all the important details visible. The final solution was tested on different CT images and the medical specialist used it for detecting this type of disease, the detection being very easily mistaken.**

Keywords—fibrosis, pulmonary, Hough algorithm, diagnostic, CT image, disease.

I. INTRODUCTION

Idiopathic pulmonary fibrosis (IPF) is a special type of pulmonary fibrosis and is one of the rare diseases. How often the disease actually occurs is unclear. However, it is estimated that 4.6 -7.4 people per 100,000 people are affected in Europe. To date, there has been no approved drug treatment for IPF patients. For the treatment of mild to moderate IPF, two drugs are now available for this group of patients: pirfinidone (Esbriet) and nintedanib (OFEV). The new therapy options enable patients to survive longer. The drugs slow the progression of the disease, and lung function may improve [1].

The disease is called "idiopathic" because the cause o f its development is unknown. It is only known that in IPF a pathological change occurs in the normal healing process of the lungs after minor injuries at the cellular level. The small air sacs called "alveoli" are gradually replaced by fibrotic scar tissue. The scarring begins in the tissue between the alveoli. Usually this tissue is a thin layer of a few cells. However, the scarring in IPF means that the alveolar wall becomes thicker and thicker and the lungs become stiff. This prevents oxygen from entering the blood. Over time, the scarring of the lungs increases: the lungs become stiff and it becomes difficult to breathe [2].

Recognizing IPF can be challenging because symptoms can be similar to other more common conditions, such as asthma, chronic obstructive pulmonary disease or congestive heart failure. For most of those affected, the symptoms - exertional dyspnea (shortness of breath with low stress), dry cough, impaired lung function - steadily increase over time. In some patients, the disease progresses rapidly within a short period, others seem stable at first, but then the symptoms suddenly worsen.

The cause of idiopathic pulmonary fibrosis is not known. According to today's understanding, damage in the alveolar epithelium primarily leads to fibrotic remodeling processes. In addition to tobacco smoke consumption, numerous other inhaled noxious substances have been described as risk factors for the development of idiopathic pulmonary fibrosis (e.g. metal, wood, stone dust, plant and animal particles). In addition to sporadic forms of idiopathic pulmonary fibrosis, there are also rare familial forms.

High-resolution computed tomography is of central importance in the diagnostic algorithm. A histology acquisition (surgical lung biopsy) can only be considered if the findings are incongruent.

In Fig. 1 the difference between a normal lung (left) and an IPF lung (right) can be seen.

The authors focus on designing and implementing a system for diagnosing IPF using image processing techniques, mainly Hough algorithm. Such a system is intended as an aiding tool for doctors and is implemented using Matlab computer software.

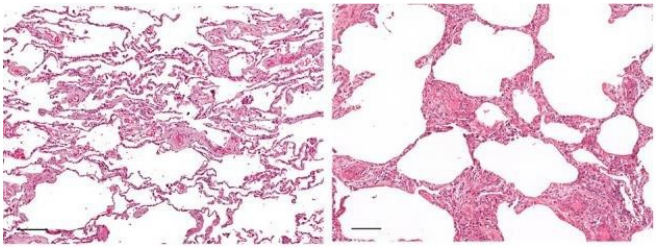

Fig. 1. The difference between a normal lung (left) and an IPF lung (right)

II. Causes and Classification

Patients with such clinical pictures often show very different symptoms. Over the course of the day, however, there is always a shortage of breath, initially only during exercise, later also at rest. Often there is a dry, irritating cough, sometimes whitish colored secretions are coughed up. The disease can worsen rapidly, but sometimes it only increases gradually over the years. If the air sacs are inflamed for a long time, scarring of the lungs occurs - pulmonary fibrosis.

- The known causes that can lead to pneumonia include: side effects of drugs (such as amiodarone or methotrexate);
- involvement of the lungs through a rheumatic disease;
- a connective tissue disease (collagenosis). There are also so-called granulomatous interstitial lung diseases, such as sarcoid or Wegener's disease [3].

Furthermore, allergic overreaction of the lungs caused by inhaled dust is a common cause of the disease. The “farmer's lung”, which is common in rural areas, is an allergic reaction to inhaled mold spores. Another group of IPFs includes rare interstitial lung diseases such as lymphangioleiomyomatosis, pulmonary histiocytosis, or eosinophilic pneumonia.

A further group is distinguished from these diseases: the idiopathic interstitial "pneumonia". These are lung diseases of unknown cause, with typical radiological and histological changes. The most common is idiopathic pulmonary fibrosis (IPF). The first specific medication was recently approved to treat this condition. All interstitial lung diseases have in common that they mostly show typical clinical, imaging and pathological changes in the lungs.

Several examinations are usually necessary for a thorough diagnosis of interstitial lung disease. First of all, pulmonary function examinations measure the degree of the impairment of lung function that has already occurred. Imaging procedures, in particular the computed tomographic chest examination, are of crucial importance. With it, clear indications of one or the other form of the disease can be obtained. In all cases, however, a bronchoscopy is necessary for further clarification [4].

During this examination, samples of the lungs are taken with small forceps and examined histologically. In the vast majority of cases, so-called bronchoalveolar lavage is also used. Here, the alveoli of a certain section of the lungs are rinsed and the cells obtained in this way are examined further cytologically. Serological blood tests can also be helpful in making a diagnosis. In some cases, a tissue sample is necessary to determine the cause. The biopsy is performed using the "keyhole technique", that is, minimally invasive using an endoscope.

Once it has been determined which interstitial lung disease the patient is affected by, immunosuppressive therapy is usually used. If a direct cause could be found, the patient must of course avoid contact with the causative drug or - in the case of a farmer's lung - with mold for life. In advanced disease, patients have to inhale oxygen.

The oxygen demand is regularly checked on an outpatient basis. In particularly severe cases, a lung transplant may also be necessary. After the preparatory examinations (stationary in the lung center), the patient is presented to the transplant center.

Since the classification of idiopathic interstitial pneumonia by the pathologist AA. Liebow over 30 years ago, the classification of interstitial accentuated lung diseases was modified several times (see Fig. 2) [5].

The Liebow classification from 1968 comprised 5 disease entities:

- common ("usual") interstitial pneumonia (UIP);
- desquamative interstitial pneumonia (DIP);
- Bronchiolitis obliterans with interstitial pneumonia (BIP);
- lymphocytic interstitial pneumonia (LIP);
- giant cell interstitial pneumonia (GIP).

III. Diagnosis Methods

Idiopathic pulmonary fibrosis does not primarily occur in smokers and is therefore not a tobacco smoke-associated interstitial lung disease in the narrower sense.

However, former or active smokers have an overall 1.6 - fold higher risk of developing idiopathic pulmonary fibrosis than non-smokers. The disease occurs with a prevalence of around 2– 43 per 100,000 people, predominantly in older adults. Idiopathic pulmonary fibrosis is defined as a specific form of chronic, progressive, fibrosing interstitial pneumonia of unknown cause [6].

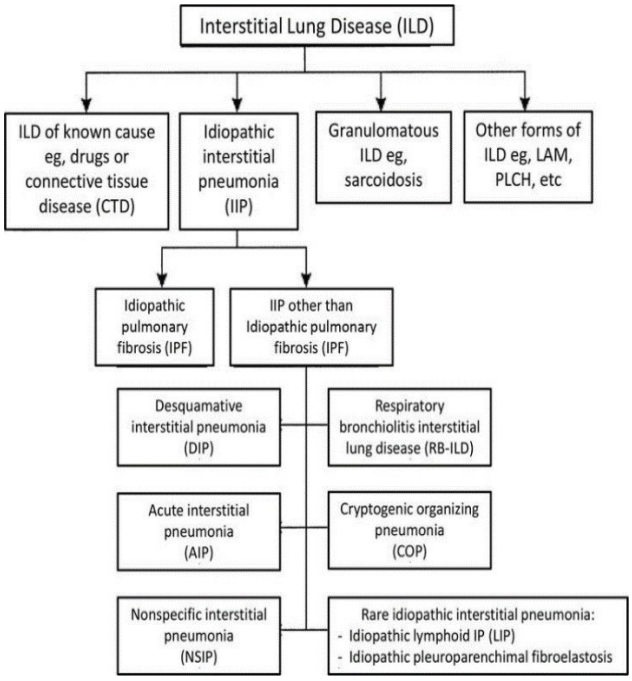

Fig. 2. The classification of interstitial accentuated lung diseases

It is limited to the lungs and is associated with the histopathological and / or radiological pattern of a so-called UIP ("usual interstitial pneumonia"). Idiopathic pulmonary fibrosis is a diagnosis of exclusion because the UIP pattern can also be present in other interstitial lung diseases (pulmonary asbestosis, chronic exogenous allergic alveolitis and interstitial lung diseases that are triggered by medication or are associated with an underlying rheumatological disease).

IPF manifests itself predominantly in the 50th to 60th year of life with progressive dyspnea and a non-productive cough. Men get sick more often than women (2:1). The pulmonary function test usually shows a restrictive ventilation disorder with a reduced lung volume. In the chest x-ray there are diffuse interstitial markings and in the HR-CT there are pronounced peripheral and subpleural reticulonodular shadows as well as a disruption of the lung architecture with an emphatically subpleural honeycomb-shaped lung remodeling. The prognosis for the disease is poor, and without a lung transplant, most patients die 5-10 years after diagnosis from respiratory failure (see Fig. 3) [7].

An essential histological characteristic is the heterogeneity of the lung changes with areas of pronounced subpleurally pronounced alveolar septal fibrosis and a honeycomb-shaped lung remodeling ("honeycomb fibrosis") next to sections of a largely preserved lung architecture ("spatial heterogeneity"). In addition to a widening of the alveolar septa due to increased collagen fibers, there are small-focal areas with loose granulation tissue rich in fibroblasts ("fibroblastic foci"; "temporal heterogeneity") [8]. The honeycomb-shaped air spaces are often lined with respiratory epithelia and goblet cells ("bronchiolisation") and show retention of secretion. In addition, there is a very differently pronounced interstitial accentuated infiltrate of lymphocytes, plasma cells and

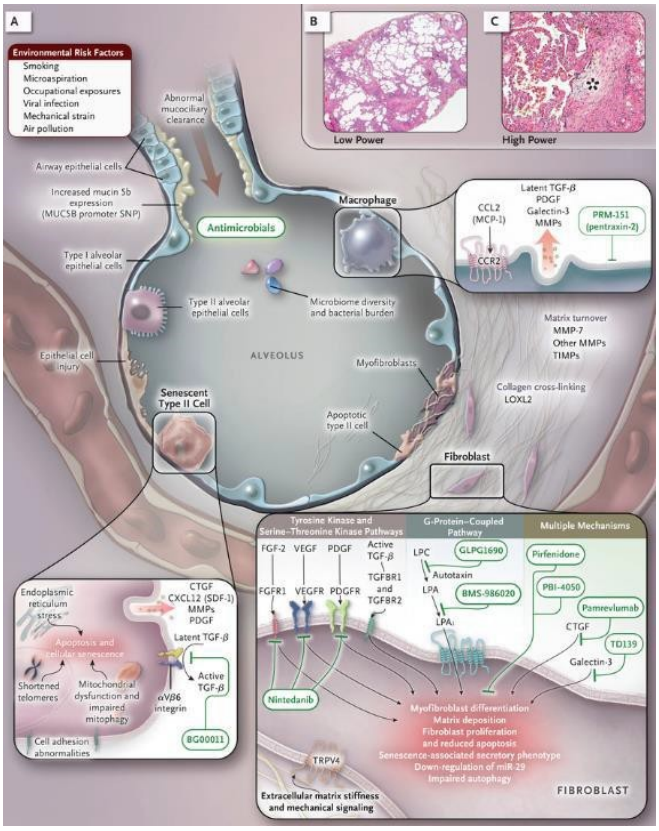

Fig. 3. Factors influencing the occurrence of IPF

histiocytes. Focal hyperplasias of type II pneumocytes and smooth muscles are found ("muscular cirrhosis measuring"). In the acute episode ("accelerated IPF"), there is histologically a combination of UIP and diffuse alveolar wall damage ("diffuse alveolar damage" / DAD) [9].

IV. Implementation and Results

In order to analyze and process the CT images, the Hough algorithm will be used in the next steps.

Hough's basic principle states that, the possible positions of a shape in the image space represents the parameter space in which the project image data is projected, then search for accumulation points in the parameter space, which correspond to the most likely positions of the shape in the CT image. Even though some variants have been released until recently, it is astonishing that, in most cases, the projection is slightly reduced on the contour or the main points of the image space.

Furthermore, it is almost always done using one of two of these methods: one-to-many (one point in the image space votes for a multidimensional surface in the parameter space) or many-to-one (a set of points in the space).

An *n*-dimensional binary image *I* being a subset of R_n, an analytical form can be defined by using a parametric equation like the following(1):

$$Ca_0 = \{x \in R_n;\ \Phi\ (x, a_0)\ =\ 0\} \qquad (1)$$

Where *x* is the spatial variable and $a_0 \in R_n$ is a constant parameter.

Taking a certain $x_0 \in R_n$, the set $Dx_0 = \{a \in R_m$; Φ $(x_0,\ a) = 0\}$, where *a* is the parametric variable, is a surface in the space of the *m*-dimensional parameters, which is the projection, or dual form of the point x_0. The Hough transformation of image *I* relative to ϕ is represented by the sum of all I projections. The most representative forms in *I* are finally detected by searching for the maxima of $\Gamma\varphi I$, represented also in Fig. 4. For purposes of uniform quantification, it would rather have the polar parametric equation to be used for line detection [10].

The spatial variable is *x = (x, y)*, the parametric variable is *a = (θ, ρ)*, and the parametric equation (2) is:

$$x\ cos\theta\ +\ y\ sin\theta\ =\ \rho. \qquad (2)$$

The dual form *D (x, y)* is a sinusoidal curve.

In practice, both the image and the parameter space are quantified. The Hough algorithm is applied to a binary image made of thin and regular curves, obtained using a contour detection algorithm [11].

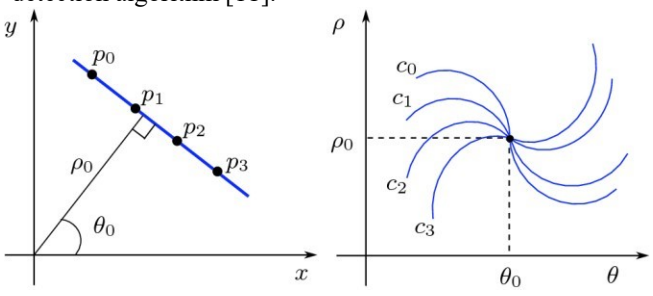

Fig. 4. Representation of lines and curves in Hough space

The transformation is calculated using one of two dual methods:

- One-to-many projection (divergent transformation), which consists of scanning each x_0 pixel of the binary image, then increasing $\Gamma\phi I$ over the entire Dx_0 surface as defined above.
- Many-to-one projection (convergent transformation), which consists in considering each m-tuple of pixels $\{x_i\}$ $1 \leq i \leq ma$ of the binary image so that there is a unique a_0 in the parameter space, so that for all x_i, ϕ, $(x_i, a_0) = 0$) and then only $\Gamma\phi I$ (a_0) increases.

A known disadvantage of classic HT is their calculation cost. If p is the number of voting pixels in the binary image, m is the size of the parameter space and k is the average number of samples per the size of the parameter space, the complexity of the projection is proportional to pkm - 1 for first-to-many turns to many-to-one. The most popular approaches to reducing complexity are to decrease the number of voting pixels by randomly choosing a subset of them. For the one-to-many transformation, such an approach is known as TPH or the Hough probabilistic transformation. In the case of many-to-one transformations, it is called THR, or randomized Hough transform. Classical HT, TPH and THR were compared from a qualitative and computational perspective.

Using local derivatives to improve or accelerate HT has been done previously, it has been proposed for O'Gorman and Clowes lines and for differentiated Shapiro curves. In his review article, Maître explicitly mentioned the one-to-one HT for lines and circles. Although, these approaches have never been used directly on gray images, only on binary curves. Gradient information was used to detect the line for accelerating HT or to control the voting process to improve progressive TPH. Valenti and Gevers presented an efficient ocular center location algorithm based on a voting scheme using isophoto curvature estimation. Lately, Yao and Yi estimated the radius by using the contour curvature and then accelerate the HT circle. However, all of these approaches still reduced the voting pixels to a previously calculated thin contour [12].

Kesidis and Papamarkos suggested an HT reversible gray level by using the pixel gray level directly. In their circle detector, Atherton and Kerbyson applied a collection of convolutions to the gradient image, which is equivalent to the HT weighting by the gradient. A dense vote was performed using the gradient and without calculating the contours in order to detect the line by Dahyot. But it used an estimated density core to spread the votes in the parameter space, resulting in a higher computational cost than the classical approach. In conclusion, of the few existing one-to-one HTs, none is a dense method and none of the existing dense approaches is a one-to-one projection [13].

The framework proposed in this article, which combines the two, is a systematic approach based on estimating the gradient and curvature of the gray image, achieving, for all pixels of the image space, a vote in the parameter space and the weight of the vote by the meaning of the derived measure. The direct calculation at the gray level eliminates the dependence on the limits and parameters of the contour detection algorithm and, finally, although the number of voting points increases by an order of magnitude, the complexity is actually lower, because the estimation of derivatives has a calculation cost. which is smaller or comparable to the contour detection algorithm and the one-to-one voting process a constant complexity.

Ultrasound skills are based not only on the ability to visualize real images of the disorder, but also on the analysis of various artifacts. Based on the basic physical principle of sound waves, the remarkable potential in the study of lung parenchyma is its ability to detect changes in parenchymal density resulting from the loss of alveolar air with or without the increase of interstitial fluids. Scanning a normally ventilated lung allows visualization of the entire chest wall from the skin to the pleura. Under the pleura, due to the thin interlobular septum and no change in the alveolar space, the air creates an interface with the pleura and the tissues of the chest wall due to acoustic impedance differences. The image corresponding to the normal lung is obtained by artifacts that represent the reflections of the wall thoracic below the pleural line ("mirror effect" of the lung or A-lines) [14].

Type A lines represent horizontal lines constantly visualized in the area of the artifacts and represent multiple reflections of the pleural line. Some authors use the term "Multiple A-lines" due to the increase in the amount of subpleural air. The evaluation of the REIT is performed by detecting and quantifying the B lines. B-lines are long, well-defined, flowing from the pleural line (see Fig. 5).

Patients with heart failure have multiple diffuse B lines due to thickening of the interlobular septum through water, as the first sign of pulmonary congestion. An increased amount of extravascular water in the alveolar compartment has a direct correlation with the number of B Lines [15].

Type B lines are generated by an air-fluid mixture, which occurs when the interlobular septum surrounded by subpleural alveoli full of subpleural air becomes edematous. Three or more visible B lines between two ribs define B lines or lung missiles. With a CT scan as a source, the lung missiles appear completely sensitive and specific for which they demonstrate thickened subpleural interlobular septa and areas of ground glass (i.e. interstitial edema). The algorithm used to remove Lee filter-based noise is defined by applying a spatial filter to each pixel. Based on an assumption that the noise is white with the average pixel unit, the Lee multiplicative filter looks for the best estimate for each pixel. If the option to crop the area of interest is used, no more it is necessary to apply the refined Lee filter, which pays special attention to the edges of the image [16].

Edge detection makes it possible to significantly reduce the amount of data in an image. If the lines could be defined

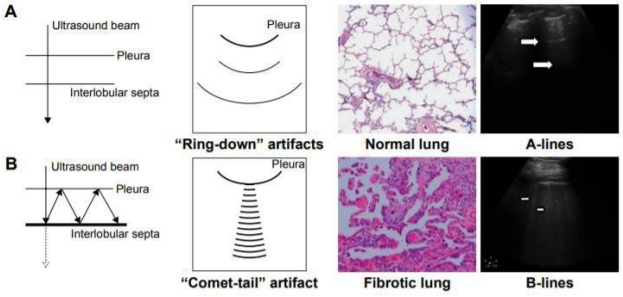

Fig. 5. The representation of A lines and B lines

by their characteristic by equations, the amount of data would be further reduced. Hough's algorithm was initially developed to recognize lines and was later generalized to cover arbitrary shapes, so to detect type B lines, they are described in (3) as follows:

$$r = x \cdot cos\theta + y \cdot sin\theta$$
$$y = -cos\theta sin\theta \cdot x + rsin\theta \quad (3)$$

where, r is the distance from the line to the origin, and θ is the angle of the line.

The Hough space for lines has these two dimensions: θ and r, and a line is represented by a single point, corresponding to a unique set of parameters (θ_0, r_0).

An important detail for the Hough transformation is the mapping of unique points. The basic idea is that a point is mapped to all the lines that can pass through that point. This gives a sinus-like line in the Hough space [17].

The algorithm for detecting type B lines can be divided into the following steps:

- Edge detection, for example using a Lee or Kuan filter; if the image allows, you can also use the image segmentation provided by the Matlab work environment, but in this case it will be done manually;
- Mapping the edge points to the Hough space and storing them properly to allow later their evaluation and interpretation;
- Interpretation of data obtained in the previous step to produce lines of infinite length. The interpretation is performed by a threshold value and possibly other constraints;
- Conversion of infinite lines into finite lines.

The finished lines can then overlap over the original image, but the level of overlap accuracy is not exactly 1:1.

To detect type B lines, the original image was uploaded into Matlab work environment. Because not all the CT images are easy to analyze and directly find the correct diagnostic, they have to be firstly processed and then the detection will be easier and visible. In this case, the Lee and Kuan filters were studied and they are the most suited for this type of images. They are used in order to remove the noise from the CT image and after that the identification of the edges is better and faster.

Hough's algorithm which was described above was implemented and used for the detection of type B lines. After detecting the B type lines, another “lines” remained from the previous image processing. They were eliminated from the final image, so that just the important and helpful details remained visible, as presented in Fig. 6.

Interpretation of results after each of the steps mentioned below will be done by a specialist who will decide if the proposed algorithm needs further modifications or if the detection of type B lines has to be done in another way which will then be the most appropriate.

V. CONCLUSIONS

The presented system has been designed so that for a medical specialist the processed CT image is easier to interpret, and thus the diagnosis is as accurate as possible. The ability to process ultrasound images and get a good result, considerably reduces the time of analysis longer by specialized people and also the diagnostic is the correct one, because of the details of the final image which are crucial for detecting the IPF.

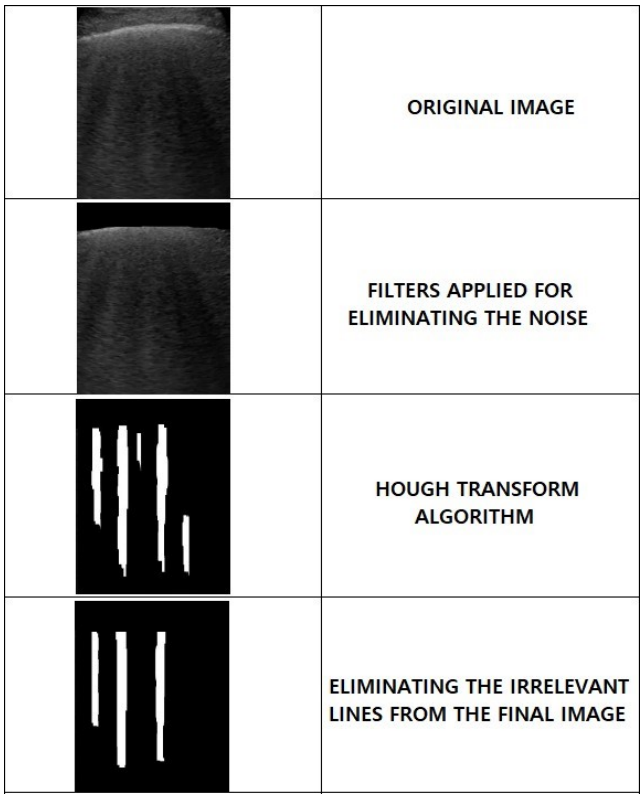

Fig. 6. The CT image processing stages

After analyzing other image processing algorithms, the best suited was Hough algorithm. The chosen algorithm was studied, tested and modified in order to meet the project requirements.

The main focus was on detecting the type B lines from a CT image and put them in evidence. In order to obtain this, the raw image was processed and some filters were needed to be applied for eliminating the noise and other unnecessary details from the CT image.

This system is very useful for CT images that contain a lot of noise and are very hard to analyze just by looking at it. The importance of detecting the correct diagnostic is crucial in some cases, so that by using this system, there will be no erroneous diagnostic.

As further development, an application can be designed and implemented to assure the total effectiveness of the above presented steps by putting them all together and make the entire system be more user friendly for the doctors.

REFERENCES

[1] J. Behr, U. Costabel "Interstitielle Lungenerkrankungen - historische Entwicklung, Status quo und Ausblick" in Pneumologie, Munchen, Thieme, 2018;

[2] Dunnill M S. "Pulmonary fibrosis, Invited review" Histopathology 1990;

[3] Lacasse Y, Fraser R S, Fournier M, Cormier Y, "Diagnostic accuracy of transbronchial biopsy in acute farmer's lung disease" Chest, 1997;

[4] Liebow AA (1968) New concepts and entities in pulmonary disease. In: Liebow AA, Smith DE (eds) The lung. Wiliams & Wilkins, Baltimore, pp 332-365

[5] Müller KM (1986) Lung Findings and Smoking Habits - Pathological Anatomy. In: Geisler LS (Ed) Smoking and the respiratory tract - Preventive and therapeutic aspects. Verlag for applied sciences, Munich, pp. 65–84

[6] Friedrichs K H, Otto H, Fischer M. "Gesichtspunkte zur Faseranalyse in Lungenstäuben", Arbeitsmed Sozialmed Präventivmed 1992;

[7] David J. Lederer, Fernando J. Martinez, “Idiopathic Pulmonary Fibrosis”, 2018;

[8] Braun FM, Johnson TRC, Sommer WH, Thierfelder KM, Meinel FG, "Chest CT using spectral filtration: radiation dose, image quality, and spectrum of clinical utility", Eur Radiol, 2015;

[9] Goldin J, Elashoff R, Kim HJ, Yan X, Lynch D, Strollo D, “Treatment of scleroderma-interstitial lung disease with cyclophosphamide is associated with less progressive fibrosis on serial thoracic highresolution CT scan than placebo: findings from the scleroderma lung study”, Chest, 2009;

[10] Antoine Manzanera, Thanh Phuong Nguyen, Xiaolei Xu, “Line and circle detection using dense one-to-one Hough transforms on greyscale images”, 2016;

[11] Gruden JF, Panse PM, Leslie KO, Tazelaar HD, Colby TV, “UIP diagnosed at surgical lung biopsy, 2000-2009: HRCT patterns and proposed classification system”, AJR Am J Roentgenol, 2013;

[12] Simon Just Kjeldgaard Pedersen, "Circular Hough Transform", Aalborg University, Vision, Graphics, and Interactive Systems, November 2007;

[13] Gerig G, Klein F, “Fast contour identification through efficient Hough transform and simplified interpretation strategy”, 1986;

[14] Tcherakian C, Cottin V, Brillet PY, Freynet O, Naggara N, Carton Z, “Progression of idiopathic pulmonary fibrosis: lessons from asymmetrical disease”, Thorax, 2011;

[15] Souza CA, M ¨uller NL, Lee KS, Johkoh T, Mitsuhiro H, Chong S, “Idiopathic interstitial pneumonias: prevalence of mediastinal lymph node enlargement in 206 patients”, AJR Am J Roentgenol, 2006;

[16] Zisheng Li, Peifei Zhu, Takashi Toyomura, Yoshimi Noguchi, “Automatic Image Analysis and Recognition for Ultrasound Diagnosis and Treatment in Cardiac, Obstetrics and Radiology”, 2018;

[17] Hiremath, Prema T. Akkasaligar, Sharan Badiger, “Speckle Noise Reduction in Medical Ultrasound Images”, 2013;