Implementing a Pulmonary Fibrosis Diagnostic System Using the Matlab Environment

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

2021 23rd International Conference on Control Systems and Computer Science (CSCS) | 978-1-6654-3939-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/CSCS52396.2021.00015

***Abstract*— The main objective of this paper is to implement an application using the Matlab working environment, in order to process medical images and to be able to use their analysis in the faster diagnosis of pulmonary fibrosis. To approach the proposed topic, algorithms for segmentation, active contour and image processing were studied, and then an optimal solution was implemented for image analysis, using the Matlab work environment. As final results, the chosen solution consists of several steps, which easily lead to the establishment and interpretation of CT images by a specialist. Each stage that makes up the application was tested individually, and then tested at the application level with all its stages included. The final solution was used for disease detection on many different CT images and the accuracy for detecting was very good.**

***Keywords*—pulmonary, segmentation, diagnosis, detection, CT image**

I. Introduction

The first description of fibrous interstitial lung disease dates back to the late 19th century, when G. E. Rindfleisch (1897) and P. von Hansemann (1898) described diseases with the names "Cirrhosis cystica pulmonum" and "Lymphangitis reticularis". More than 30 years later, pathologists Hamman and Rich first reported cases of fulminant, diffuse, interstitial pulmonary fibrosis, for which the term "Rich-Hamman syndrome" was later coined. The first modern classification of interstitial lung disease was developed by Liebow, Carrington and Gaensler in the 1960s and 1970s.

From these first descriptions, a wide differential diagnosis of the possible causes and associations of interstitial lung diseases has been described, so that today more than 150 different clinical images are differentiated, which can be associated with interstitial lung diseases [1].

A distinction is made between diseases of known cause (e.g., toxic-drug interstitial lung disease), granulomatous diseases (e.g., sarcoidosis) and special entities (e.g., Langerhans cell histiocytosis or lymphangioleiomyomatosis), and large group of pneumonias idiopathic interstitials. The latter comprises seven subspecialties, whereby idiopathic pulmonary fibrosis ith the histological pattern of common interstitial pneumonia is the most common form and, at the same time, presents the most unfavorable evolution of the disease. Due to the wide differential diagnosis and frequent association with systemic diseases and metabolic disorders, interstitial lung diseases are an important part of general internal medicine [2].

Despite the pronounced heterogeneity of the group of interstitial lung diseases, common features can also be determined, which on the one hand are due to altered physiological conditions, with restrictive ventilation disorders and limited gas exchange, and on the other hand, be observed in the final joint section of the irreversible remodeling of the lung parenchyma. In Fig.1, a classification of different types of diffuse parenchymal lung diseases was made [3].

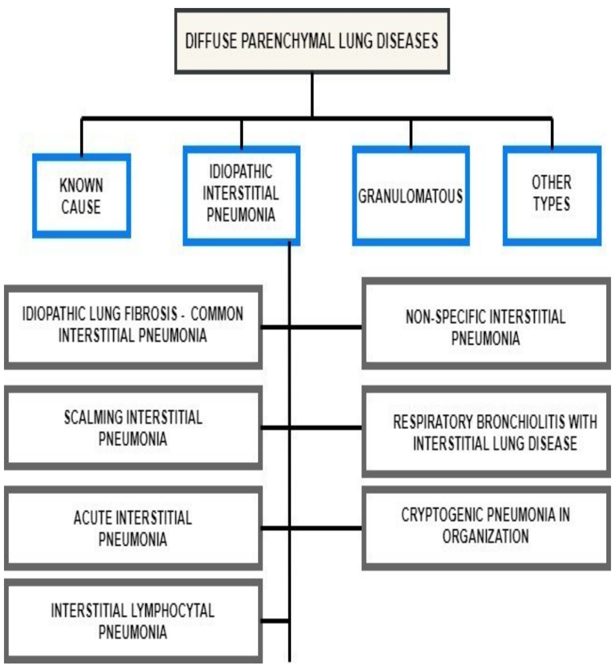

Fig. 1. Classification of lung diseases

2379-0482/21/$31.00 ©2021 IEEE
DOI 10.1109/CSCS52396.2021.00015

However, the different aspects of the disease in inflammation, granuloma formation, fibroproliferation and fibrotic scarring vary considerably between different disease entities, leading to a response therapeutically different.

The main objective of this paper is to develop an application using the Matlab environment in order to facilitate the analysis of a CT image and to be able to interpret and find the correct diagnosis to identify pulmonary fibrosis. This diagnosis is particularly difficult to detect and can be erroneous if it does not focus on several decisive factors.

The first description of sarcoidosis, as well as fibrous lung diseases, goes back to the histopathological observations of the 19th century. Since then, many different disease entities have been identified and are summarized under the generic term of diseases of diffuse lung parenchyma. What they have in common is the manifestation in the interstitium of the lungs, through which a wide spectrum is observed from the predominantly inflammatory processes to the purely fibrous processes [4].

Genetic engineering studies have fundamentally renewed the pathophysiological understanding of fibrous lung diseases. In the field of early detection and differential diagnosis, high-resolution computed tomography is a decisive advance. Regarding the treatment of pulmonary fibrosis, the first positive results for the high dose of N-acetylcysteine and for pirfenidone appear. An improved understanding of the pathophysiology gives hope that targeted therapeutic approaches can be successfully tested in the future [5].

II. Pathophysiology

Given the heterogeneity of the described disease group, it is not surprising that the pathophysiological bases are also very different. There are diseases with a predominantly inflammatory background, such as exogenous allergic alveolitis or involvement of the lung parenchyma in the context of collagenosis. On the other hand, a distinction must be made between diseases related to cigarette smoke, such as respiratory bronchiolitis with interstitial lung disease and squamous interstitial pneumonia. Numerous toxic reactions to drugs (e.g. bleomycin lung, amiodarone lung) are also characterized by pronounced mechanisms of inflammatory disease [6].

Given the broad differential diagnosis of interstitial lung disease, a correct internal diagnosis of each patient is required, which includes various aspects, including internal, private and occupational exposure to pollutants, medical history, family history and targeted searches for superordinate system diseases. High-resolution computed tomography of the lungs has played a central role in diagnosis in recent years. Technological progress in this area has led to a huge improvement in image quality, so that discovery models can now be delineated in high-resolution computed tomography, which can be attributed to appropriate clinical images [7].

For example, the detection of a basal sub pleural pattern of honeycomb with reticular interstitial growth and traction bronchiectasis combined with a lack of frozen glass is characteristic of common interstitial pneumonia. However, this pattern may appear either as an idiopathic disease, i.e. independent, or in connection with exogenous allergicalveolitis or collagenosis. Consequently, no definitive diagnosis can be derived from the outcome model of the high-resolution computed tomography, but rather the correct diagnosis can only be made by examining all the results of the clinical examination and the device [8].

The situation is similar to the so-called non-special interstitial pneumonia, which is often seen in patients with scleroderma, but which also occurs in exogenous allergic alveolitis or appears as an idiopathic disease, i.e. independent. Consequently, in addition to detecting the pulmonary fibrosis pattern, the exclusion or detection of an exogenous trigger or a major systemic disease is of crucial importance for the correct diagnosis of these clinical images. To this end, in addition to blood tests, bronchoscopy with bronchoalveolar lavage and transbronchial lung biopsy can provide important information about the underlying clinical picture or can help eliminate possible differential diagnoses [9].

Surgical biopsy as a gold standard in the diagnosis of interstitial lung disease has been questioned, given the complexity of different clinical images. Surgical lung biopsy can ultimately provide only one histological model, but not a definitive diagnosis. The latter results from the overview of clinical, functional, bronchoscopic and possibly histological analysis. Given the improved imaging quality of high-resolution computed tomography, pulmonary surgical biopsy is required only in approximately 20-30% of cases when diagnosing idiopathic pulmonary fibrosis. This is especially considered in patients in whom high-resolution computed tomography does not allow a clear assignment of finding a model, usually in the early stages of the disease.

In Fig. 2, the numerous granulomas of free giant cells (blue circles) can be seen which have been observed peri-bronchially and in the interstitium [10].

The subpleural and paraseptal parenchyma are usually most severely affected by the histopathologic changes. Hyperplasia of type 2 pneumocytes and bronchiolar epithelium are mostly associated with biological responses, such as inflammations, which most of the time consist of a patchy interstitial infiltrate of lymphocytes and plasma cells.

III. Diagnosis Methods

To diagnose interstitial pulmonary fibrosis, the doctor will have a detailed medical history to rule out other lung-related diseases or medical causes, and then perform a physical exam. Physical examination is not specific enough to diagnose interstitial pulmonary fibrosis, but it can identify some signs. The doctor will also listen to the patient's chest to determine if the lungs are making abnormal sounds when the patient is breathing. [11]

The physician may then order one or more of the following tests or diagnostic procedures (see Fig 3). A chest CT scan and sometimes a lung biopsy are required to make a diagnosis. Diagnostic tests for interstitial pulmonary fibrosis:

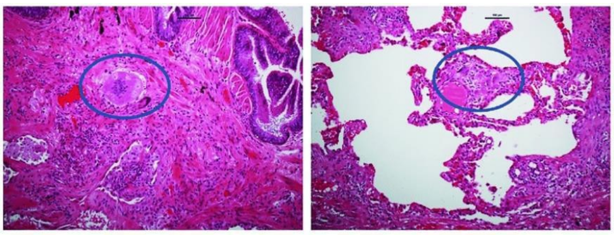

Fig. 2. Observed granulomas

- Chest imaging studies
- High resolution CT scans or CT scans: to rule out other diseases, allows early diagnosis (and, in some cases, definitive), assess the extent of the disease (changes in bronchial structure, scarring pattern in the entire lung).
- Lung biopsy
- Taking a small sample of lung tissue, usually through a small incision through the ribs with a thoracoscope: to directly examine lung tissue to help diagnose IPF or other lung diseases.
- Lung function test
- Use of a device to measure breathing capacity: to measure the degree of deterioration of lung function and to see if the disease has advanced.
- Oxygen desaturation study
- The patient walks for almost 6 minutes while his oxygen level is measured by a probe attached to the finger or forehead: to determine the need for oxygen and to see if the disease has advanced [12].
- Other laboratory tests: complete blood count, electrolytes, creatinine levels, liver function tests, blood arteries. To rule out other diseases, the changes in body functions have to be motorized over time.

IV. Implementation and Results

Medical image processing deals with the development of problem-specific approaches to improve medical data extracted from images for selective visualization as well as additional analysis. There are many topics in the medical field that refer to image processing: some emphasize the general applicable theory, and others focus on specific applications. In this area, the emphasis is mainly on image segmentation and multi-spectral analysis [13].

The methods that are applied in order to segment an image are:

- Mathematical models based methods represent the basics of computing in the medical field. This method is a fundamental technique for achieving scientific progress in any type of research: clinical, experimental, biomedical and even behavioral research. Nowadays, most of the medical images are difficult to interpret by using light or a microscope, but they can be quantified as various physical phenomena after being processed by a series of techniques.
- Statistical methods which take into consideration the pixels from an image. The pixels are labeled depending on the intensity of image distribution. A grayscale image is the simplest and most efficient image to segment, based on a gray threshold.
- Structural methods are based on the properties such as regions and edges of an image. There are algorithms which can be applied for edge detection, depending on the analyzed image. The disadvantage of such methods is the fact that they are sensitive to noise and artifacts [14].

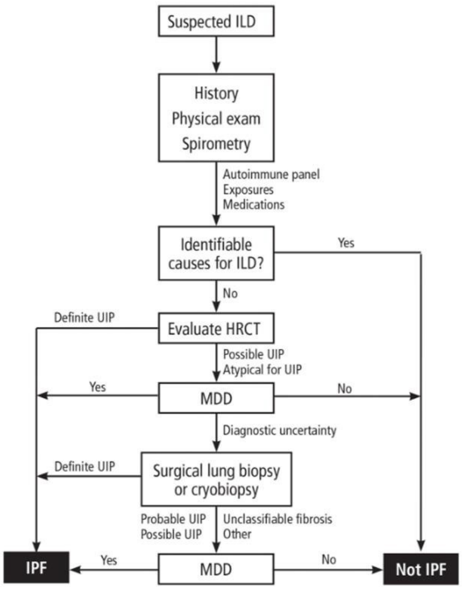

Fig. 3. Diagnostic algorithm

For the implementation of all project requirements, Matlab/Simulink was used as the implementation environment, because each point of the problem can be created and tested. In Fig.4, the main functionalities of the application are presented:

- Uploading the CT image to be processed in the application;
- Displaying the entire image;
- Cropping the image, practically selecting the area of interest on which the processing will be done; this step is very important because the segmentation is made faster on an smaller image;
- Using the toolbox from Matlab for image processing, the segmentation technique that it has will be used;
- Application of the Chan-Vese algorithm for segmenting the area of interest;
- Application of specific filters, for speckle filtering;
- Elimination of image noise;
- Obtaining a graph of the intensity of the pixels that make up the area of interest.

The application consists of a graphical user interface that incorporates various functions of the Matlab work environment responsible for image processing. This application mainly includes functions related to image analysis and segmentation, along with the development of segmentation, contour algorithms and problem area detection.

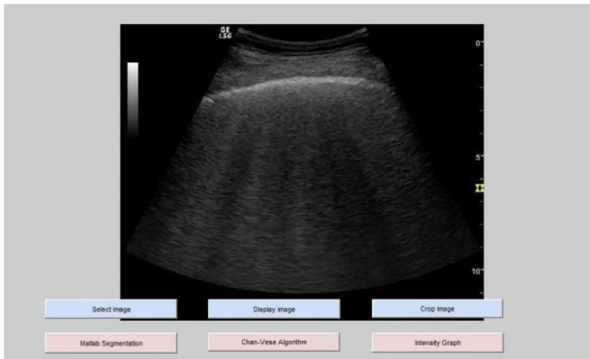

Fig. 4. Graphical user interface

In this application there are presented the buttons with the provided options:

- The first button is for selecting an image to be analyzed / processed;
- The image display option opens the selected image in a new window;
- Next, use the "Crop image" button. With its help, an area of interest is chosen from the original image that will be processed or analyzed using this application (see Fig. 5). This functionality allows the user to choose only the part that interests him, not making sense to process the entire image. Also, the results obtained in the area of interest are much more objective and practical;
- After choosing the area of interest, it can be analyzed by using the tools provided by Matlab. One possibility of segmentation would be to use the "Segmentation Matlab" option, using the application integrated in Matlab "Image segmenter". Here we also have the possibility to manually or automatically select a threshold as well as other options for refining the desired selection. You can export the image obtained from segmentation or you can save the options used as Matlab code;
- The user will be able to manually process the image using this option, to prepare the image to be analyzed in more detail;
- Using the Matlab Toolbox, one can refine and process the original image, but at the same time it can be applied a series of operations to establish and identify the pleural area, which depending on its thickness can state a diagnosis. The Matlab segmentation technique can be used to highlight the pleural area as described in Fig 6.
- After selecting this area manually or automatically, one can choose the number of iterations performed to segment the image.
- In this case, the number of iterations is 100.
- The better the pleural area is highlighted, the more accurately the identification will be made. This part of the process is very important, because once identified, some computations can be made in order to establish the length and thickness of the pleura which is used in diagnosis of the disease.

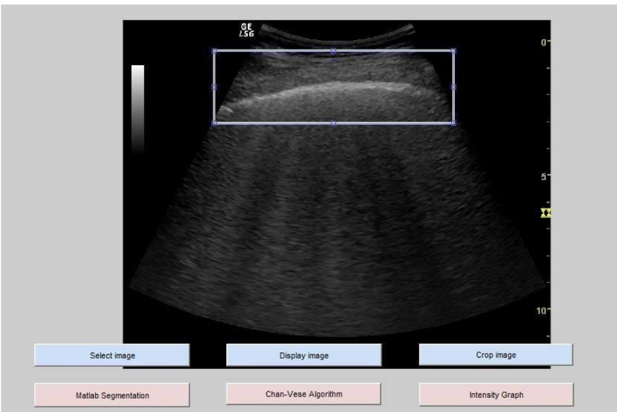

Fig. 5. Selecting area of interest

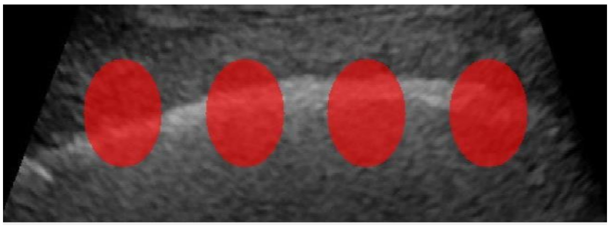

Fig. 6. The image after segmentation

- By using Matlab Toolbox, the segmentation should be done without any issues and without being necessary to stop the segmentation process. After this process is finished, the specialist can identify and study the image that was obtained (see Fig. 7).
- In order to compare the results obtained by using the segmentation from Matlab, another segmentation method was used. From active contour algorithms like Snake Algorithm, Gradient vector flow algorithm, Chan-Vese Algorithm, the most suited was the last one because of its efficiency and accuracy in processing CT images. The other algorithms can also detect the contour of an object in an image, but have several inefficiencies like insensitivity to noises, false contour detection in high complex objects;
- Automatic segmentation of medical images is challenging because medical images are complex and rarely have simple linear characteristics. In addition, the result of the segmentation algorithm is affected by the partial volume;
- Although some of the algorithms had been proposed within the subject of clinical photo segmentation, clinical photo segmentation is still a complicated and tough problem. Furthermore, it can be taken into consideration artificial intelligence as gear to optimize those simple strategies to gain correct segmentation results;
- Therefore, all these aspects must be considered when choosing a segmentation algorithm.
- It is a flexible and powerful method which can be used for all types of medical images, including some that can be difficult to segment (e.g. brain images).

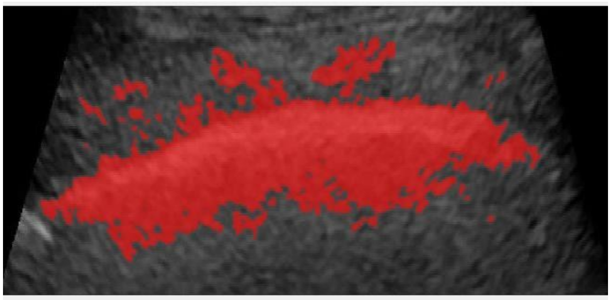

Fig. 7. The image after segmentation

- Edge segmentation is the most common edge detection method by highlighting the boundary that separates different areas. The edge detection process is based on mark gaps such as grayscale and color, and these edges usually represent the boundaries between objects. This method divides the image along the edges. Based on the gradient function (derivative function), many edge detection operators can be used: Prewitt, Sobel, Roberts (first derivative type) and Laplace (second derivative type), Kenny, Marr-Hilklrat limit detector. In addition, in the edge-based segmentation method, the edges need to be constructed by combining the identified edges into edge chains in the process;

- Textural capabilities of a CT image are essential from photograph segmentation and class factor of view. The goal of texture primarily based totally segmentation approach is to subdivide the image into small pieces having distinct texture properties, at the same time to take the areas that have already been segmented by one or other different approach;

- In the case of statistical methods, the texture is determined by a set of statistically features, which are represented as vectors in a multi-dimensional feature space. The statistical information can be based on first-order, second-order, or higher-order gray-scale statistical information. Use probabilistic or deterministic decision-making algorithms to assign feature vectors generated from templates in this way to their specific classes.

- A collection of the most important methods that help in the process of segmentation are described below together with the advantages and disadvantages of the studied methods (see Fig. 8)[15];

Lung Segmentation Class	Application	Advantages	Disadvantages
Thresholding-based methods	Identification and segmentation of well-defined normal structures and isolated lesions such as tumors, cavities, and nodules	Basic, intuitive, fast, least expensive computationally	Fails to deal with attenuation variations, fails to categorize pathologic classifications
Region-based methods	Normal structures, regions with minimal noise, minimal abnormality	Fast, works well with more-subtle attenuation variations	Fails to segment regions with moderate to high levels of abnormality or when pathologic condition abuts adjacent structures
Shape-based methods (atlas-based and model-based methods)	Abnormal pathologic conditions that defy segmentation of normal anatomy	For a well-conceived representative template (atlas or model), segmentation accuracy can be high	Representative training features difficult to create, computationally expensive, performance highly dependent on the feature set and training data
Neighboring anatomy–guided methods	Identification and classification of pleural effusions or atelectasis	Works well for cases in which attenuation-based matrices fail	Computationally expensive, severe pathologic condition could throw it off (eg, opacification of entire hemithorax)
Machine learning–based methods	Delineation of pathologic conditions with signature textured patterns such as GGO, consolidation, and crazy-paving pattern	Works well to identify ill-defined diffuse pathologic conditions, categorizes pathologic classifications such as GGO, consolidation, and septal line	Computationally expensive, no good separation among classes of pathologic conditions

Fig. 8. The image after segmentation

- One of the most effective method it was proved to be by using artificial intelligence tools;

- In the Chan-Vese segmentation algorithm, the initialization of the degree-set features is a tough problem. In the proposed new segmentation algorithm, the initialization isn't always required. And every step is easy and effortlessly achieved. In step one, there are numerous algorithms to get the easy model of the unique CT image, and within the second step, it is calculated at each grey degree and after deciding which locates the top of the line from the grey degrees. The proposed approach is extra green than the Chan-Vese approach

- Compared with the previous presented segmentation methods, the proposed approach is extra simple, green, and flexible. First, we separate the segmentation tactics into smoothing the initial CT image and segmenting it into regions, and within the second step, the mass of the computation system includes operators and logical operators. Step one is simply to get a few easy variations of the initial CT image, so the second one is unbiased of the end result of step one, and the second one needs to have an end result;

- Regarding snake algorithms the situation is differently drawn. In the sense that the snake moves to correspond to the maximum value of the scene edge map, this segmentation strategy is usually based on edges. A region-based snake method can be implemented using fast algorithms to segment objects in a CT image. The best algorithm in terms of maximum probability is based on the calculation of statistical data of the inner and outer regions, so it can be applied to different types of random fields that can describe the input image.

- However, if there are multiple objects in the CT image, the snake model must be adapted to determine the appropriate contour of each object. In addition, when tracking multiple objects in consecutive frames, pre-initialized serpentine contours may produce incorrect results due to subtle topological changes.

- Because of the complexity of the algorithm, the image should be grayscale and also the larger the image is, the more computational time is needed, but this issue was already solved at the previous step, where just the important and the area of interest is cropped from the entire image and just that part of the image will be processed;

- Depending on the obtained results at first try, the algorithm has been customized for the CT image and the parameters of the functions that were used were calculated in order to obtain the best detection in as fewas possible iterations;

- The approach of this method of segmentation is considered a “modern” one, because the computation behind the method is based on calculus and partial differential equations;

- Also for less number of iterations, a refinement of the algorithm was needed;
- The main objective of Chan-Vese algorithm is to minimize an energy functional $F(a_1,a_2,A)$, defined by equation (1) as follows:

$$F(a_1, a_2, A) = \mu \cdot Length(A) + v \cdot Area(inside(A))$$

$$+\lambda_1 \int_{inside(A)} |i_0(x, y) - a_1|^2\, dxdy \qquad (1)$$
$$+\lambda_2 \int_{outside(A)} |i_0(x, y) - a_2|^2\, dxdy$$

Where:

- i_0: the CT image;
- A: a piecewise parameterized curve;
- a_1: average pixels' intensity inside A;
- a_2: average pixels' intensity outside A;
- $\mu,v,\lambda_1,\lambda_2>0$: fixed parameters which were determined for the CT image;

The length of the curve (A), the area inside the curve and the average intensities (a_1,a_2) are computed by using partial differential equations which are influenced by the curvature of the curve and the motion of the curve.[16]

By applying the Chan-Vese algorithm on the CT image, after 20 iterations, the contour of the interested pleura was obtained and it can be seen in Fig.9.

The results obtained by the last method were better than by using the Matlab Toolbox and also the computation time was less than 1 minute.

For each functionality a separate script was implemented, and the test images were passed one by one through each, and then, the final product was made to be evaluated at the system level. The latter has undergone further changes in order not to overlap the parameters of each functionality, but also to put them all together.

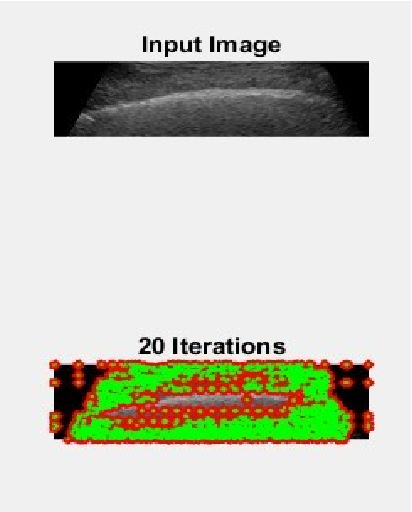

Fig. 9. Chan-Vese Segmentation

- The "Intensity Graph" option opens a new window in which the image is presented as a 3D graph, displaying the third dimension, which represents the intensity of the pixels in the image, as can be seen in Fig. 10.

Based on the values of the resulted pixels' intensity from the CT image, a medical specialist can identify and diagnose with a greater level of accuracy the stage of the disease.

To avoid the discrepancy between different pulmonary diseases, the total value of surface highlighted in yellow can be computed by using the intensity graph.

V. Conclusions and Further Development

The application has been designed so that a specialist can make the necessary adjustments to the image, and thus the diagnosis is as accurate as possible. But, without a doubt, the ability to process ultrasound and get a result, considerably reduces the time of analysis longer by specialized people, and at the same time the way the application is organized, allows the user to highlight several important parameters in order to diagnosis.

As directions for future development, with the help of this application it will be possible to implement a more precise algorithm for identifying and analyzing images in the medical field in order to diagnose patients with pulmonary fibrosis.

To correctly identify, more images and experimental data will need to be considered to eliminate any type of false diagnosis.

For this topic, an implementation based on machine-learning would be useful, because this type of diagnosis can be easily confused with other types.

The next step will be to automate this technique, by developing a computer tool that is easy to use clinically, able to provide an objective assessment of the pleural line. This would reduce inter- and intra-observer variability and create a unique quantification system to standardize diagnostic and monitoring scores. Such a methodology, supported by artificial intelligence, will make a robust end product. The potential benefits of fast data collection without high costs and patient risks are intuitive. The utility and clinical significance for the diagnosis and monitoring of lung disease, in an era fraught with the challenge of pandemic infectious interstitial diseases (such as COVID-19), are easy to guess.

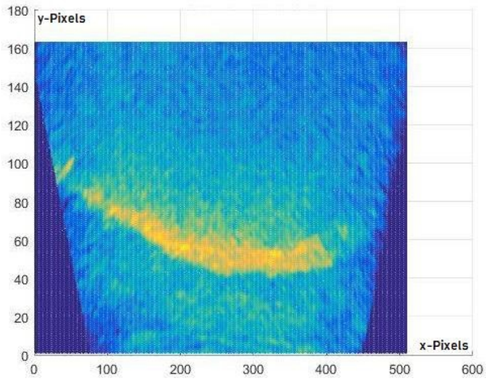

Fig. 10. Intensities Graph

References

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

[2] David J. Lederer, Fernando J. Martinez, "Idiopathic Pulmonary Fibrosis", 2018;

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

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

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

[6] Sennekamp J., " Neue Kenntnise über die exogen-allergische Alveolitis", Pneumologie, 1992;

[7] 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;

[8] Diana Manolescu, Lavinia Davidescu, Daniel Traila, Cristian Oancea, Voicu Tudorache, "The reliability of lung ultrasound in assessment of idiopathic pulmonary fibrosis", 2018;

[9] Irving WL, Day S, Johnston ID, "Idiopathic pulmonary fibrosis and hepatitis C virus infection", Am Rev Respir Dis, 1993;

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

[11] Miyake Y, Sasaki S, Yokoyama T, Chida K, Azuma A, Suda T, "Occupational and environmental factors and idiopathic pulmonary fibrosis in Japan", Ann Occup Hyg, 2005;

[12] Parry EM, Alder JK, Qi X, Chen JJ, Armanios M., "Syndrome complex of bone marrow failure and pulmonary fibrosis predicts germline defects in telomerase", Blood, 2011;

[13] Dr. J. Thirumaran, S. Shylaja, "Medical Image Processing – An Introduction", 2014;

[14] 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;

[15] A. Mansoor, U. Bagci , B. Foster, Segmentation and Image Analysis of Abnormal Lungs at CT: Current Approaches, Challenges, and Future Trends, 2015;

[16] Pan et. al, "Efficient Implementation of the Chan-Vese Models Without Solving PDEs," in International Workshop On Multimedia Signal Processing, 2006, pp. 350-353.