

# Detection and Prediction of Lung Cancer Using Different Algorithms

N. Sudhir Reddy, V Khanaa,

**Abstract:** One of the major causes of cancer death is through Lung cancer. The overlapping of cancer cells acts as an impediment for its early detection. Identifying genetic and environmental factors plays a vital role in developing better techniques for its prevention. And in order to discover the anomalies in target images, time factor is paramount. In this work, decision tree algorithm is used for prediction of lung cancer wherein the important pattern with their corresponding weightage and score is studied. Processes such as preprocessing of images and feature extraction are done using Histogram Equalization and using neural network classifier does normal or abnormality check of the patient. If lung cancer is detected early, then it will reduce many treatment options, reduce risk of invasive surgery and increase survival rate. Therefore, an easy, cost effective and time saving method will produce promising result for detection and prediction of lung cancer.

**Keywords:** Data Mining, Decision Tree Algorithm, Early Detection, Histogram Equalization, Image Processing, Lung Cancer, Prediction.

## I. INTRODUCTION

To enhance the manual analysis, Image processing techniques acts as a good quality tool. In order to enhance the images during early detection and treatment stages, the Image processing technique is used in this proposed method. Assessment of Image quality and its improvement depends on the enhancement stage. In this stage, the pre-processing technique is used based on principal component analysis and Histogram Equalization. During the digital image analysis, classification is very important as it categorizes images in to groups as per their similarities. In proposed system, Histogram Equalization is used for preprocessing of images and feature extraction are done using Histogram Equalization and using neural network classifier does normal or abnormality check of the patient. Subsequently, survival rate of patient is predicted through the extracted features.

Many significant patterns such as Smoking, Environment, Alcohol, Obesity, Chronic Lung Disease, Balance Diet, and Mental trauma, Radiation Therapy, Tobacco and Genetic Risk are considered in this proposed early detection and prediction system. One of the most common causes (about 90%) of lung cancer is due to cigarette smoking as cigarette smoke contains more than 4,000 chemicals, many of which have been identified as carcinogenic. Other factors that contributes to lung cancer are environment pollution mainly air and excessive alcohol consumption.

Revised Manuscript Received on 14 September, 2019.

**N. Sudhir Reddy**, Research Scholar, Department of Computer Science and Engineering, Bharath Institute of Higher Education and Research (BIHER), Chennai, Tamil Nadu

**Dr. V Khanaa**, Professor, Dean Information, Department of Computer Science and Engineering, Bharath Institute of Higher Education and Research (BIHER), Chennai, Tamil Nadu

The risk of developing lung cancer is 20-25 times greater for someone whose per day consumption of cigarettes exceeds one pack than someone who abstains from smoking. Lung cancer begins in one or both lungs and occurs for out-of-control cell growth. There are other problems if Lung cancer extends to the brain affecting vision and weakness on one side of the body. Cough, coughing up blood, chest pain and shortness of breath are signs of primary lung cancers.

Chest Radiograph (x-ray), Computed Tomography (CT), Magnetic Resonance Imaging (MRI scan) and Sputum Cytology are some of the advanced techniques for diagnosing lung cancer. However, these techniques are not affordable for many and also time consuming. A new technique to diagnose the lung cancer in its early stages is the need of the hour as most of the aforementioned techniques are capable of detecting the lung cancer only in its advanced stages, reducing the patient's chance of survival. Hence, manual analysis can be enhanced through Image processing techniques.

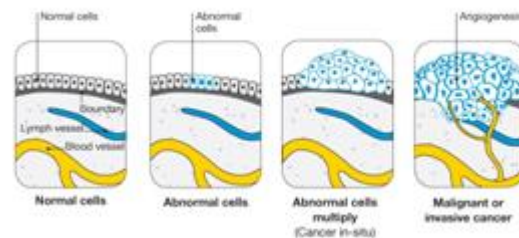


Fig.1: The beginning of cancer

This lung cancer risk prediction system using the significant pattern prediction tools will enhance the detection process. And thus, early prediction should play a significant role in the diagnosis process and for an effective preventive strategy.

As mentioned earlier, the existing techniques to diagnose lung cancer are not affordable for many and also time consuming. Also, it detects the cancer in its advanced stages ultimately reducing patients' chance of survival.

Therefore, the proposed system is designed to predict the lung cancer in its early stage depending on few factors and thresholding. Here, the time and cost required for excessive medical test is reduced as the number of rule for testing in this system is lesser. Another advantage is that, the proposed system is web based connecting patients from remote locations to directly communicate with doctors.

# Detection and Prediction of Lung Cancer Using Different Algorithms

Furnished below are the key objectives of this system:

- To bypass some testing rules not necessarily required.
- To make the system more time and cost efficient.
- To enhance the performance accuracy of the System.
- To use lesser number of attributes
- To detect during the initial stage.
- To increase the survival rate of the patient by more than 5 years.

The literature survey is conducted extensively in order to acquire enough relevant information concerning to topic of research.

## II. LITERATURE REVIEW

The aim of most of the earlier work is to develop an efficient system to predict cancer in its early stages.

In this paper, it was intended to study the different characteristics of information mining procedures as data processing ideas benefits in carcinoma classification. For applicable rule generation and classifications on diseases, it studies different data processing and hymenopteron colony improvement techniques. Furthermore, it provides the fundamental framework for enhancement in cancer diagnosis..

In this work, the preprocessing of the photographs is done using bar chart equalization. In order to decide the performance of the various classifiers, experimental analysis is formed with dataset wherein performance depends on the correct and incorrect classification of the classifier. Neural Network rule is implemented and performance comparison with other classification algorithms is done. It yields the highest TP Rate and lowest FP Rate clearly proving its efficiency as compare to alternative classifiers. Subsequent work from the same author relies on Feature Extraction and Principal element Analysis for detection of carcinoma in CT scan pictures. It utilizes a hybrid technique supported feature extraction and Principal element Analysis (PCA) [1].

The image quality and accuracy forms the key factors of the analysis. With the help of segmentation principles, associate degree increased region of the thing of interest is used as a basis for feature extraction. A normality comparison is formed by counting on general options. [4].

The primary target of the work is vital improvement in distinction of plenty with the suppression of background tissues. It is obtained by standardizing the parameters of the planned transformation. The manual analysis of the phlegm samples is not accurate, time taking and requires trained person to avert any errors during diagnosis. The segmentation results used as a base for a pc motor-assisted designation system for early detection of cancer enhances the probabilities of survival for the patient. In this work, physicist filter is used which enhances the medical pictures. [5].

In order to divide phlegm color pictures and detect carcinoma during early stages, this paper presents Hopfield Neural Network (HNN) and a Fuzzy C-Mean (FCM) bunch rule. The segmentation results are used as a base for a pc motor-assisted designation (CAD) system for early detection

of carcinoma which can improve the probabilities of survival for the patient. But segmentation results are inaccurate due to variation within the grey level and also the relative distinction among the photographs. Therefore, a thresholding technique is utilized as a pre-processing step. Subsequently, the quantitative procedures area unit supported the nuclear feature and proved the thresholding rule to be successful in extracting the nuclei and protoplasm regions. A thousand phlegm color pictures were used in this study to check each strategy. HNN proved to be more efficient in classification compared to FCM as it extracted the nuclei and protoplasm regions. An accuracy of 98% proves the efficiency of the thresholding classifier [2].

### 3. Steps involved

The following steps are involved in the diagnosis of carcinoma:

- Images captured
- Preprocessing of images
- Image segmentation
- Feature extraction
- Principle element analysis
- Neural network classifier
- Diagnosis result
- Prediction method
- Images captured or knowledge assortment

CT-Scan pictures of cancer and non-cancer patients are gathered. Usually it is taken from different diagnostic centers that have no connection. The digitized pictures area unit keeps within the DIACOM format with a resolution of eight bits per plane [1].

### Preprocessing of images

With an aim to improve the interpretability or sensitivity of knowledge enclosed, image Pre-processing stage during this system commences with image improvement. This is done in order to provide higher input for alternative programmed image process techniques. Abstraction domain strategies and frequency domain strategies are the two major categories into which Image improvement techniques are divided. Subsequently for alternative image process, the image improvement techniques area unit is utilized as pre-process tools and the quantitative measures will validate that techniques area unit are most fitted.

In the pre-processing stage, the initial image is not disturbed but there is selective elimination of the redundancy in scanned pictures. This is very important within the designation of carcinoma. To strengthen the superiority of every image, they are preprocessed with the Histogram-equalization.

### III. IMAGE SEGMENTATION

| S.No | Study of                 | No. of sample images | Method used   | Performance Results      |
|------|--------------------------|----------------------|---|--------------------------|
| 1.   | Armato et al.[2]         | 24 datasets          | Iterative thresholding and morphological                | RmsD=0.54 (0.8 pixel)    |
| 2.   | Yimet al. [26]           | 10 subjects          | Region growing, connected component                     | RmsD=1.2 pixel           |
| 3.   | Puetal.[19]              | 20 datasets          | Thresholding  | FP/GT=0.43%, FN/GT=      |
| 4.   | Gaoet al.                | 8 subjects           | Thresholding  | DSC=                     |
| 5.   | Johnet al.               | 60 slices            | Multilevel thresholding                                 | -                        |
| 6.   | Itaiet al [11]           | -                    | Snake deformable model                                  | Qualitative assessment   |
| 7.   | Silveria et              | 1 slice              | Level set   | Qualitative assessment   |
| 8.   | Annan giet al.[1]        | 1130 image           | Shape-based deformable Model                            | DSC=0.88                 |
| 9.   | Shi et al. [20]          | 247 image            | Shape-based model and PCA                               | OM=0.92, AD = 1.78 pixel |
| 10.  | Sun et al. [23]          | 30 scans             | Active Shape Models and optimal surface finding methods | DSC=0.975                |
| 11.  | Besbes and Paragios      | 247 image            | Shape-based deformable Model with boosted image         | OM=0.9474, AD=1.39 pixel |
| 12.  | Sofka et al. [22]        | 260 scans            | Graph-based shape with iterative surface deformation    | SCD=1.95                 |
| 13.  | Hua et al. [10]          | 15 scans             | Graph-search  | Sen.=0.986, Spec.=       |
| 14.  | Kockelkorn et al. [14]   | 22 scans             | Prior training, statistical Classifier usin             | OM=0.96, AD=1.68mm       |
| 15.  | El-Ba et al.             | 10 image datasets    | Statistical MGRF model                                  | Accu.=0.968              |
| 17.  | Dai et al. [6]           | 23 scans             | 2D wavelet  | RmsD=0.52mm              |
| 18.  | Yankel evitz et al. [25] | Not specified        | Graph-cuts and GMM                                      | DSC=0.984                |

### IV. FEATURE EXTRACTION

In order to find and segregate the most well liked parts or shapes of an inputted image, algorithms associate degree strategies are used during image options Extraction stage. The binarization and GLCM are the two strategies to predict the presence of carcinoma. The respective strategies area unit

supported facts that powerfully associated with respiratory organ anatomy and knowledge of respiratory organ CT imaging.

#### A. Binarization Approach

The binarization approach is employed for detection of cancer. In order to validate whether the cells are normal or abnormal in the respiratory organ, this approach extracts the amount of white pixels and check them against some threshold. Subsequently, validation is done on whether a range of white pixels of a brand new image is a smaller amount than the edge. Depending on the quantity of white pixels, it indicates whether the image is normal or if the quantity of the white pixels is larger than the edge, it specifies that the image in abnormal.

#### B. GLCM (Grey Level Co-occurrence Method)

The gray levels in a picture (that is the brightness value) can be tabulated using GLCM. The gray-level co-occurrence matrix from image in MATLAB is formed as the beginning of this process.

In the subsequent stage, the GLCM is normalized

Where:  $i$  is that the row range and

$J$ : is that the column range

And thus the texture events from the GLCM can be found [1].

#### C. Masking Approach

This approach has high dependency on numerous area unit that looks like white connected areas within ROI (lungs). With the increase in the percentage of the area unit inside lungs, the cancer increases. With the integration of the aforementioned two approaches, it is concluded that if the range of black pixels in the image is bigger than white pixels then that image is traditional. But if the range of black pixels in the image is smaller than white pixels then that image is considered to be having an anomaly.

#### D. PCA (Principle element Analysis)

The information in image can be regularized with the help of this technique. Generally, in Real-world knowledge sets the variables exhibit some associations. These kind of area unit associations are either linear or a minimum making them feasible for common examination techniques. Principal element analysis is one such method that rotates the initial knowledge to new coordinates. The information thereby becomes uniform. [1].

PCA comprises of the following processes:

- The mean and variance of the options within the image are calculated.
- The sample mean are subtracted from all the observations.
- Then divided by the sample variance that scales and centers the information.

## Detection and Prediction of Lung Cancer Using Different Algorithms

- The coefficients of the key components and their applicable changes area unit are then calculated by finding the Manfred Eigen operate of the sample variance matrix.
- For the key components, the matrix holds the coefficients and diagonal parts store the modification of the relevant principal constituents.
- The diagonal is mined.
- The maximum info content is arrived through the differences in knowledge that is necessary for higher classification [1].

### V. NEURAL NETWORK CLASSIFIER

The tool used here is supervised feed-forward back-propagation neural network ensemble. This tool differs in many aspects from other classifiers like theorem and k – nearest neighbor classifiers. In the later two classifiers, the linear knowledge is necessitated to accurately figure. But neural network operates for non- linear knowledge because of which it is simulated on the reflection of biological neurons and network of neurons.

The detection accuracy can be enhanced by training the neural network with broader selection of input file. This would mean that the system might be influenced with a tiny low set {of knowledge of knowledge of information} or massive set of comparable data. Neural network classifier requires an outsized set of information for coaching. But coaching together is time consuming that is an impediment to attain stability. Sotraining them as biological neural network transmitting signals at a speed equivalent to electrical signals can enhance the speed drastically.

Neural network comprises of three layers namely the Input layer, internal hidden layer and output layer area unit. In the input layer, the nodes of area unit are coupled with variety of nodes within the internal hidden layer. Linking of nodes within the internal hidden layer might take place either with another internal hidden layer or with associate degree output layer. Further, the output layer comprises of one or more response variables.

The final Steps comprises of the followingin Neural Network Classifier:-

- Feed-forward back propagation network is created.
- The accessible samples and cluster outlined for it are trained with neural network.

From the results of neural network, the area unit is determined. The samples are then trained in network classification rate and calculated using mathematical formulas.

#### A. Support Vector Machines (SVMs)

For the purpose of classification, the algorithms such as support vector machines area unit supervised learning models with associated learning algorithms thatexamine knowledge and its patterns are used. Assuming that the pattern classifier using a hyper plane to segregate two types of patterns supported given examples  $i=1n$ . Where (i) could be a vector within the input area  $I=Rk$  and  $y(i)$  denotes the category index taking worth one or zero.

Binary categories are categorized by support vector machine. They find and employ a class boundary (the hyper plane) increasing the margin within the given coaching knowledge. The support vectors are the coaching knowledge samples on the hyper planes close to the category boundary area unit. And the distance between the support vectors and the category boundary hyper planes is known as margins. The SVM area unit supported the idea of call planes that outline call boundaries.

The below representation shows the Theory of SVM:

Given coaching knowledge  $(x_i, y_i)$  for  $i = \text{one} \dots N$ , with  $x_i, R_d$  and Lolo  $\rightarrow$

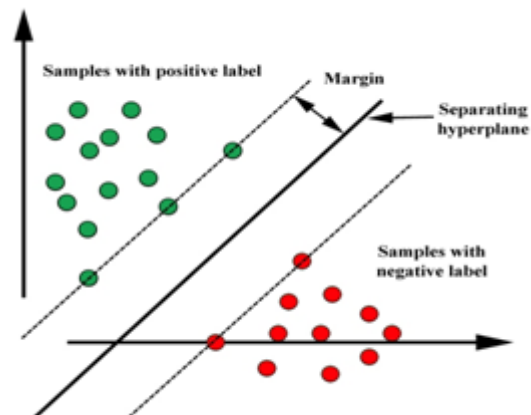
Every input  $x$  is ab initio mapped into the next dimension feature area  $F$ , by  $z = \phi(x)$  through a nonlinear mapping  $\phi$ :

$$R_n \rightarrow F.$$

$$f(x) = W^T x + b$$

$W$  is that the traditional to the road, and  $b$  the bias.  $W$  is thought because the weight vector.

Assets of objects having completely different category memberships are segregated by decision plane. For knowledge classification, SVM can be considered as a very valuable method. In coaching and testing knowledge comprising of some knowledge instances, classification task is generally involved. One target value/class labels and numerous other attributes/features are an integral part of every instance within the coaching set.





VI. DIAGNOSIS RESULT

| S. No | Study of                     | Parameters used                       | Method used  | Performance Results (Accuracy) |
|-------|------------------------------|---------------------------------------|--|--------------------------------|
| 1     | Kawata, Yoshiaki et al. [13] | Curvature Index                       | Linear Discriminant classifier                     | $A_z=0.94$                     |
| 2     | McNitt-Gray et al. [17]      | shape, size, attenuation, and texture | Stepwise discriminant Analysis                     | Accuracy=90.3%                 |
| 3     | Cristobal et al. [5]         | nodule border smoothness              | Neural network classifier                          | $A_z= 0.871$                   |
| 4     | Gurney and Swensen [62]      | nodule border smoothness              | Bayesian classifier                                | $A_z= 0.894$                   |
| 5     | Matsuki, Yuichi et al.       | Shape-based                           | Neural network                                     | $A_z= 0.951$                   |
| 6     | Henschke et al. [9]          | -                                     | Statistical-multiple-object detection and location | -                              |
| 7     | Lo et al. [15]               | Radiographical                        | Neural network                                     | $A_z= 0.89$                    |
| 8     | Suzuki et al. [23]           | -                                     | MTANN classifier                                   | $A_z= 0.89$                    |
| 9     | Chenet et al. [4]            | -                                     | ANN ensemble                                       | $A_z= 0.91$                    |
| 10    | Nakurama et al. [18]         | Subjective Features                   | Neural network                                     | $A_z= 0.854$                   |

VII. PREDICTION METHOD

Cancer at a very advanced stage does not have a remedy and death becomes inevitable. Therefore, detecting the carcinoma at an early stage plays a vital role within the designation method. In this paper, an efficient carcinoma prediction system that supports data processing is planned. The risk prediction system proposed in this work is expected to enhance the detection process. Many risk factors such as genetic, age, diet, gender, medical history, sedentary lifestyle, obesity, environment, mental trauma, heart condition, consumption of excessive alcohol, tobacco and anti-allergic medicine, actinotherapy and chronic respiratory organ diseases are all taken into consideration. Call tree algorithm is used for the purpose of prediction that is illustrated in the figure below. The probabilities of patient can be predicted by the system that will give hope to the patients in future.

VIII. CONCLUSION

It can be concluded that, by using a method that integrates neural network classifier, binarization and GLCM, the accuracy of lung cancer detection process can be highly enhanced. This will subsequently help in the reduction of time required for cancer detection, ultimately making the system cost efficient. Further, if the patient is not detected with the lung cancer the system will proceed for the prediction process. Most state-of-the-art medical facilities are not available to patients in the remote regions but this proposed system would be web based. Therefore, even patients from remote areas with Internet connectivity can reap the benefits of this system thereby making the proposed system reach out to the wider public.

REFERENCES:

- Annangi, Pavan, Sheshadri Thiruvankadam, Anand Raja, Hao Xu, XiWen Sun, and Ling Mao. "A region based active contour method for x-ray lung segmentation using prior shape and low level features." In 2010 IEEE international symposium on biomedical imaging: from nano to macro, pp. 892-895. IEEE, 2010.
- Armato, Samuel G., Maryellen L. Giger, Catherine J. Moran, James T. Blackburn, Kunio Doi, and Heber MacMahon. "Computerized detection of pulmonary nodules on CT scans." *Radiographics* 19, no. 5 (1999): 1303-1311.
- Besbes, Ahmed, and Nikos Paragios. "Landmark-based segmentation of lungs while handling partial correspondences using sparse graph-based priors." In *Biomedical Imaging: From Nano to Macro*, 2011 IEEE International Symposium on, pp. 989-995. IEEE, 2011.
- Chen, Hui, Yan Xu, Yujing Ma, and Binrong Ma. "Neural network ensemble-based computer-aided diagnosis for differentiation of lung nodules on CT images: clinical evaluation." *Academic radiology* 17, no. 5 (2010): 595-602.
- Cristobal, Gabriel, and Rafael Fonolla Navarro. "Space and frequency variant image enhancement based on a Gabor representation." *Pattern recognition letters* 15, no. 3 (1994): 273-277.
- Dai, Shuangfeng, Ke Lu, Jiyang Dong, Yifei Zhang, and Yong Chen. "A novel approach of lung segmentation on chest CT images using graph cuts." *Neurocomputing* 168 (2015): 799-807.
- El-Ba, Ayman, Georgy Gimel'farb, Robert Falk, Trevor Holland, and Teresa Shaffer. "A new stochastic framework for accurate lung segmentation." In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 322-330. Springer, Berlin, Heidelberg, 2008.
- Gao, Qixin, ShengJun Wang, Dazhe Zhao, and Jiren Liu. "Accurate lung segmentation for X-ray CT images." In *Natural Computation*, 2007. ICNC 2007. Third International Conference on, vol. 2, pp. 275-279. IEEE, 2007.
- Henschke, Claudia I., David F. Yankelevitz, Ion Mateescu, Dean W. Brettle, Timothy G. Rainey, and Fred S. Weingard. "Neural networks for the analysis of small pulmonary nodules." *Clinical imaging* 21, no. 6 (1997): 390-399.
- Hua, Panfang, Qi Song, Milan Sonka, Eric A. Hoffman, and Joseph M. Reinhardt. "Segmentation of pathological and diseased lung tissue in CT images using a graph-search algorithm." In *Biomedical Imaging: From Nano to Macro*, 2011 IEEE International Symposium on, pp. 2072-2075. IEEE, 2011.
- Itai, Yoshinori, Hyoungeop Kim, Seiji Ishikawa, Shigehiko Katsuragawa, Takayuki Ishida, Katsumi Nakamura, and Akiyoshi Yamamoto. "Automatic segmentation of lung areas based on SNAKES and extraction of abnormal areas." In *Tools with Artificial Intelligence*, 2005. ICTAI 05. 17th IEEE International Conference on, pp. 5-pp. IEEE, 2005.
- John, Jibi, and M. G. Mini. "Multilevel Thresholding Based Segmentation and Feature Extraction for Pulmonary Nodule Detection." *Procedia Technology* 24 (2016): 957-963.



## Detection and Prediction of Lung Cancer Using Different Algorithms

13. Kawata, Yoshiki, Noboru Niki, Hironobu Ohmatsu, Masahiko Kusumoto, Ryutaro Kakinuma, Kensaku Mori, Hiroyuki Nishiyama, Kenji Eguchi, Masahiro Kaneko, and Noriyuki Moriyama. "Computerized analysis of 3-D pulmonary nodule images in surrounding and internal structure feature spaces." In *Image Processing, 2001. Proceedings. 2001 International Conference on*, vol. 2, pp. 889-892. IEEE, 2001.
14. Kockelkorn, Thessa TJP, Eva M. van Rikxoort, Jan C. Grutters, and Bram van Ginneken. "Interactive lung segmentation in CT scans with severe abnormalities." In *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, pp. 564-567. IEEE, 2010.
15. Lo, Shih-Chung Benedict, Li-Yueh Hsu, Matthew T. Freedman, Yuan Ming Fleming Lure, and Hui Zhao. "Classification of lung nodules in diagnostic CT: an approach based on 3D vascular features, nodule density distribution, and shape features." In *Medical Imaging 2003: Image Processing*, vol. 5032, pp. 183-190. International Society for Optics and Photonics, 2003.
16. Matsuki, Yuichi, Katsumi Nakamura, Hideyuki Watanabe, Takatoshi Aoki, Hajime Nakata, Shigehiko Katsuragawa, and Kunio Doi. "Usefulness of an artificial neural network for differentiating benign from malignant pulmonary nodules on high-resolution CT: evaluation with receiver operating characteristic analysis." *American Journal of Roentgenology* 178, no. 3 (2002): 657-663.
17. McNitt-Gray, Michael F., Eric M. Hart, Nathaniel Wyckoff, James W. Sayre, Jonathan G. Goldin, and Denise R. Aberle. "A pattern classification approach to characterizing solitary pulmonary nodules imaged on high resolution CT: preliminary results." *Medical physics* 26, no. 6 (1999): 880-888.
18. Nakamura, Katsumi, Hiroyuki Yoshida, Roger Engelmann, Heber MacMahon, Shigehiko Katsuragawa, Takayuki Ishida, Kazuto Ashizawa, and Kunio Doi. "Computerized analysis of the likelihood of malignancy in solitary pulmonary nodules with use of artificial neural networks." *Radiology* 214, no. 3 (2000): 823-830.
19. Pu, Jiantao, Justus Roos, A. Yi Chin, Sandy Napel, Geoffrey D. Rubin, and David S. Paik. "Adaptive border marching algorithm: automatic lung segmentation on chest CT images." *Computerized Medical Imaging and Graphics* 32, no. 6 (2008): 452-462.
20. Shi, Yonghong, Feihu Qi, Zhong Xue, Liya Chen, Kyoko Ito, Hidenori Matsuo, and Dinggang Shen. "Segmenting lung fields in serial chest radiographs using both population-based and patient-specific shape statistics." *IEEE Transactions on Medical Imaging* 27, no. 4 (2008): 481-494.
21. Silveira, Margarida, Jacinto Nascimento, and Jorge Marques. "Automatic segmentation of the lungs using robust level sets." In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 4414-4417. IEEE, 2007.
22. Sofka, Michal, Jens Wetzl, Neil Birkbeck, Jingdan Zhang, Timo Kohlberger, Jens Kaftan, Jérôme Declerck, and S. Kevin Zhou. "Multi-stage learning for robust lung segmentation in challenging CT volumes." In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 667-674. Springer, Berlin, Heidelberg, 2011.
23. Sun, Shanhui, Christian Bauer, and Reinhard Beichel. "Automated 3-D segmentation of lungs with lung cancer in CT data using a novel robust active shape model approach." *IEEE transactions on medical imaging* 31, no. 2 (2012): 449-460.
24. Suzuki, Kenji, Feng Li, Shusuke Sone, and Kunio Doi. "Computer-aided diagnostic scheme for distinction between benign and malignant nodules in thoracic low-dose CT by use of massive training artificial neural network." *IEEE Transactions on Medical Imaging* 24, no. 9 (2005): 1138-1150.
25. Yankelevitz, David F., Anthony P. Reeves, William J. Kostis, Binsheng Zhao, and Claudia I. Henschke. "Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation." *radiology* 217, no. 1 (2000): 251-256.
26. Yim, Yeny, Helen Hong, and Yeong Gil Shin. "Hybrid lung segmentation in chest CT images for computer-aided diagnosis." In *Enterprise networking and Computing in Healthcare Industry, 2005. HEALTHCOM 2005. Proceedings of 7th International Workshop on*, pp. 378-383. IEEE, 2005.