Detection of Lung Cancer Cells using Image Processing Techniques

Gawade Prathamesh Pratap[1] and R.P. Chauhan[2]
[1,2]NIT Kurukshetra–136119, India
E-mail: [1]pg1911991@gmail.com

***Abstract*—Lung cancer keeps on changing on various medical factors depending on topographic areas. The identification of Lung cancer at initial stages is of extreme importance if it is intended to degrade high mortality rate. The worldwide lung screening program focuses to imagine PET/CT examinations amongst most matured gatherings at danger to upgrade the early location rate. In spite of the fact that utilization of obtrusive procedures, side effects scarcely show up until infection is propelled making it troublesome for radiologist to recognize sores. Every year, the American Cancer Society appraises the quantities of new growth cases and passing that will happen in the world in the present year and aggregates the latest information on tumor frequency, mortality, and survival. Genuine and precise information is the basis of disease control initiatives. More than 3/4th of the illness is identified with tobacco utilization. Furthermore, hereditary components, presentation to ecological poisons, second hand smoking expand illness quickly. Cures including chemotherapy, radiotherapy, surgery, epidermal open medications raise survival rate and personal satisfaction. This strategy is more about diagnosing at ahead of schedule and critical stages with keen computational procedures with different noise elimination by segmentation strategies and calculations which is the root idea of digital image processing. Location of CT pictures received from cancer research organizations is investigated utilizing MATLAB.**

Keywords—Lung Cancer; MATLAB; American Cancer society; Noise Removal; Segmentation; Mortality Rate

I. INTRODUCTION

Tumor is a noteworthy general health issue worldwide and is the second driving reason for death all over the world, regardless of age. Cancer can be clarified as uncontrolled cell development having capacity to spread everywhere throughout the body. Our body contains red blood cells (RBC's) as well as white blood cells (WBC's). The main function is to supply fresh oxygen (O2) to all parts of the body with the help of blood flow, due to which blood appears red. [1] In the lungs, tissue receives oxygen (O2) because of RBC's only.

The hereditary substance of erythrocytes has high centralization of hemoglobin. The cell film comprises of proteins and lipids which is spine of physiological cell capacity. They don't contain any imperative piece of cell, which incorporates hemoglobin. Around 20 lakhs new RBC's are created every second. [2] The cells are delivered in the bone marrow and turn all through the body for around 4 months to and fro in arteries and veins. Every revolution takes around 20 seconds. Around 75% of the cells and lion's share (majority concentration) grouping of blood in the human body are red platelets [3, 4]. Youthful Dutch scholar Jan Swammerdam portrayed it impeccably with an early magnifying instrument in 1658.

Lung disease, additionally called as carcinoma [5] is a dangerous lung tumor characterized by uncontrolled cell development in tissues of the lung.[6] If left undiscovered, this development enters past the lung by the procedure of 'metastasis'- which produces tumour (as growth cells) from introductory site of infection to another piece of body.[7] Most malignancies otherwise called primary lung cancers, are carcinomas which begin itself in the lungs.[8] There are two primary sorts are carcinoma 1)small-cell lung tumour and 2)non-small cell lung cancer. [9] The most widely recognized side effects are coughing(counting blood coughing), weight reduction, breath shortness, and little as wee as serious chest pains.[10] These growth cells aggravates the cycle of RBC generation and debasement. Inside they change the structure and chemical composition of plasma membrane i.e. the external structure of the cell in such a way that these RBC's do not die as per the corresponding lifetime. Accordingly the RBC count continues expanding and more number of cell gets accumulated bringing about shortening of veins and supply routes and extreme blasting. These results into blood through cough and so on.

Image processing is utilized to break down pictures at the most reduced level gave any quality. These operations don't amplify probability of data content, however they diminish it if entropy is a data measure. The principle necessity of preparing is to enhance pixel power by changing over from discrete to computerized picture, sectioning to pixels, completing numerical operations on pixels, and recreating of picture with better quality. [11] Pre-processing of CT pictures is the underlying stride in picture examination took after by division handle and finished with some morphological operations are connected to recognize the tumor spots/cells in the picture. Likewise it can be utilized to decide the measure of spreading of malignancy i.e. what proportion of lung is influenced with disease. The morphological operations are essentially connected by looking at the size and state of the malignancy cell with ordinary cell, and after that the contaminated cells pictures showed onto dim scale picture

978-1-4673-8587-9/16/$31.00 ©2016 IEEE

with greatest intensity (255). The sequencing of the paper is as follows. Section II describes the prescribed methodology, in detail with mathematical operations performed on the image using various filtering methods followed by segmentation. Section III shows the results of the developed methodology using a MATLAB image. Finally, section IV defines the conclusion and application of the method in diagnostics of various diseases.

II. METHODOLOGY

The algorithm is proposed as is given in figure.1 as follows:

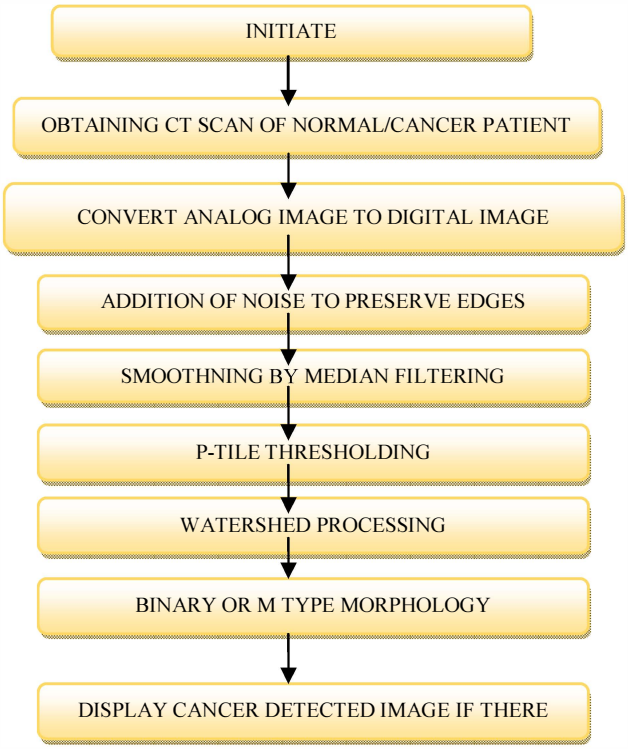

Fig. 1: Idea for Cancer Detection

The design is explained below as follows:

A. *Input Image*

For any kind of cancer, firstly image of interior organs of the body ought to be acquired. CT scan otherwise called multiple angle X-beam computed tomography makes utilization of X-beam for catching the pictures from different points and union these pictures to create cross sectional tomographic picture of specific zones of filtered tissues i.e. it permits the individual to see the status inside body without noninvasive techniques [12].

The lungs are the prime vital organs of breath in humans including different creatures. In well evolved creatures and major type of vertebrates, two lungs are situated on either side of the heart close to spine. Their part is to take oxygen from the atmosphere and move it into the circulation system, and to give out carbon dioxide from the circulatory system into the air. Humans have 2 lungs, right and left. They are arranged inside the thoracic cavity of chest. The right lung being greater than the left, imparts space to the heart. The lungs together weigh around 1.5 kg. Plural sac in which lungs are encased permits internal and external dividers to slip over each other without more grinding. This sac encases every lung furthermore separates every lung into segments called lobes. The right lung has 3 lobes and the left has two. These lobes are further grouped into bronchopulmonary sections and lobules. The lungs have an exceptional blood supply, getting deoxygenated blood from the heart for accepting oxygen (i.e. aspiratory flow) and a recognized supply of oxygenated blood (the bronchial dissemination).

The tissue of the lungs can be struck by various illnesses, including pneumonia and lung malignancy. Chronic disorders, for example, ceaseless discouraging aspiratory infection and emphysema (harming alveoli's in the lungs) can be connected to smoking or presence to toxic substances. Illnesses, for example, bronchitis can likewise include the respiratory tract. The picture of influenced lungs and typical lungs is very distinctive and effectively differentiable. These CT pictures are changed over to greyscale pictures.

B. *Grayscale Image*

In the digitized world, a greyscale image is a computerized/digital image, in which the estimation of every pixel is an individual example, i.e., it conveys just power or intensity i.e. white or black in terms of display. Pictures of this kind, otherwise called white (maximum amplitude) and dark (minimum amplitude) pictures, comprise selective shades of dim.

Grayscale pictures are the aftereffect of measuring the amplitude of light at every pixel in a single band of the light spectrum. They can likewise be obtained from a full coloured picture. The explanation for picking grayscale picture is even least pixel power is additionally useful in recognizing changes in the cells. In fact, a dim shading is one in which the R, G, B planes have equal intensity, the intensity level represented as a number from decimal 0 to 255. For each pixel in a RGB grayscale picture, G = B = R. The amplitude differs in extent with the number speaking to the brightness levels of the RGB hues. Dark is spoken to by R = G = B = 0 and white is spoken to by R = G = B = 255.

C. *High Pass Filter*

As the name proposes, it passes the frequency over certain cutoff frequency and weakens all the frequency beneath the cutoff frequency. A high pass filter is essentially utilized for making bright images by sharpening. It is done when contrast is improved between the nearby regions with expansion or reduction in

amplitude level. A high pass filter sets high edge cutoff to get data of a picture while cutting the low frequency information. The premise of the high pass filter is intended to build the intensity of the middle pixel with respect to neighboring pixels. The kernel function generally contains a single value at its center, which is completely surrounded by other values. The values might be characterized as far as positive or negative.

D. *Median Filtering*

It is a nonlinear digital filter used to remove some noise in the image. To detect some edge in the image, firstly noise should be removed up to some threshold value and then edge removal is performed. Hence the median filter is placed before edge detector. Its main feature is it removes noise without edge removal. Median filter is same as that of averaging filter, in which each output image pixel is set corresponding to the average value of neighboring pixel of the input image. The median filter is more sensitive to mean values and less sensitive to extreme values of pixel which helps in noise reduction [13].

E. *Threshold Segmentation*

Edge division is one of the most effortless division strategy. The pixels are partitioned relying on their intensity levels. There are different sorts of segmentation on the basis of various parameters like threshold values of pixels, edge based, region based, clustering etc. and so forth. It relates greyscale picture to binary picture, additionally called as mapping. After this operation, picture is separated into 2 pixel values just, 0 and 1. If there is a picture which contains dark pixel on bright background or vice-versa, then thresholding can be utilized to isolate the structure. Likewise to set a specific edge esteem, numerous sub calculations can be utilized e.g. histogram estimation, optimal thresholding, iterative thresholding, K implies grouping. In K implies bunching, grayscale picture is isolated into K fragments i.e. K-1 edge values, subsequently lessening change. Numerous pictures which are made of pixels contains more than one quality e.g. RGB [14]. If we isolate these pixel values for R, B and G, they are called channels.

F. *Watershed Algorithm*

It can be explained using a practical idea. Consider a surface immersed into the lake with a hole at minima, so that water will start filling through that hole and will go on increasing. If 2 such surfaces very close to each other are placed then a point will come where the water will overlap and mix from both the surface. At that point only, dams are built so that water does not mix. These dams are watershed lines and by the process of filling water, surface separation can be done. There are many methods to carry out this algorithm. One of the most common watershed algorithms was introduced by Otsu called 'Otsu watershed algorithm'.

This algorithm applies only to grayscale image [15].

G. *Morphological Operations*

As per Wikipedia, mathematical morphology is a procedure to assess sectioned structures/images in light of random functions and variables, set hypotheses and so forth. It is entirely connected to computerized pictures as it were. E.g. binary morphological operations investigates a specific grayscale picture with straightforward, predefined shapes, and finishing up how this shape fits into the picture gave or what part of the picture gets missed because of this shape [16].

III. Experimental Results and Analysis

For implementation of this above method, real time images of cancer patients have been obtained from institutes. Since cancer in CT image have more intensity level than background, it becomes very easy to extract image from CT image for further analysis.

The working methodology is as follows:

1) Pre-processing grayscale image using salt and pepper noise

The proposed work flow is given in figure 2

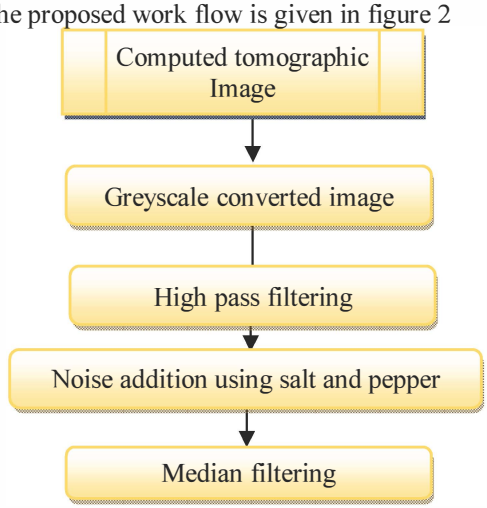

Fig. 2: Flowchart for Initial Processing of CT Scan Image

a. The CT scanned Image with cancer infected lung as seen a spot in left lung from patient's point of view is obtained from the cancer institute is given in fig. 3

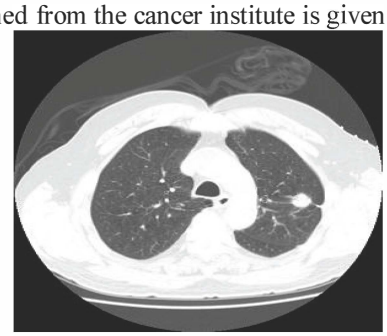

Fig. 3: CT image of Cancer Infected Lungs

b. The CT image is converted to grayscale image to perform mathematical operations shown in figure 4. In terms of digital, greyscale is a unique, reflexive and unit distance code. Changing from bit to bit is only of unit distance so probability of error is minimum.

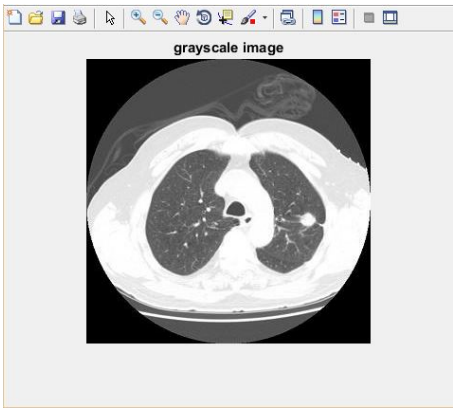

Fig. 4: Greyscaled Image

c. Now it is ready to carry out further mathematical operations, the image is passed through a high pass filter to enhance the information needed shown in figure 5 but there is problem of edge preserving.

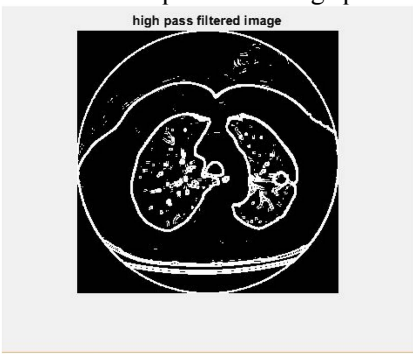

Fig. 5: High Pass Filtering

d. Further improvement of the image for proper diagnosis is carried out using median filtering. For this we added salt and pepper noise to minimize noise from the image, so that the image is distortion less shown in figure 6. Above damaged edge is restored due to salt and pepper pixel spots as seen below.

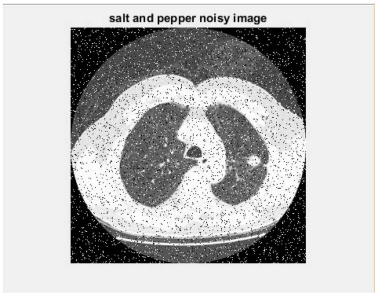

Fig. 6: Addition of Salt and Pepper Noise

e. Final preprocessing is done by passing salt and pepper noise through a median filter which will allow to remove noise completely from the image while restoring edges shown in figure 7.

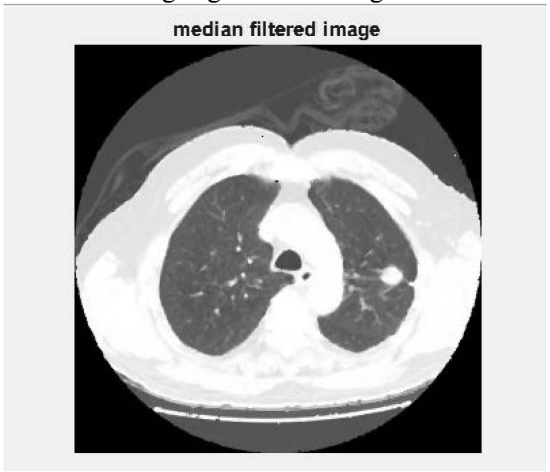

Fig. 7: Median Filtered Image

2) *The median filtered passed through various segmentations shown in figure 8*

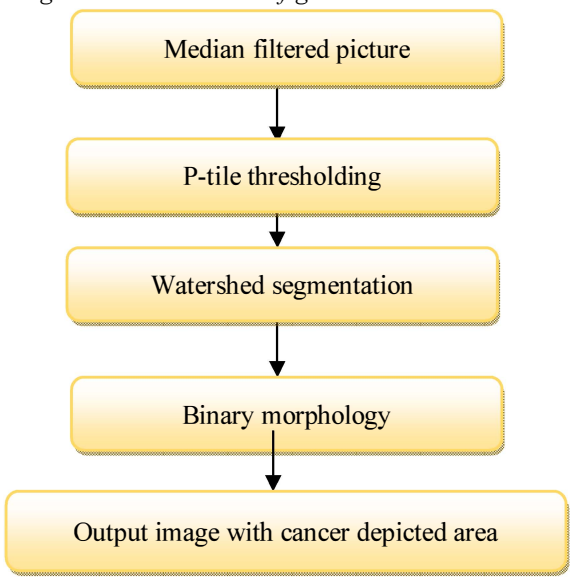

Fig. 8: Segmentation Process

f. The segmented image is shown as follows in figure 9. There are various threshold detection methods like P- tile thresholding, optimal thresholding, mixture modelling, and adaptive thresholding. All the above methods are automated. For cancer cell detection P-tile thresholding is used because in this process object is with lesser/greater intensity as compared to background and subjugates a definite known percentile 1/p from the total image area (example: printed text sheet). We set the threshold by finding the intensity level such that 1/p image pixels are below this value

Fig. 9: Threshold Segmented Image

g. The main branch of this process is watershed algorithm. It will identify the areas affected by cancer in the infected lung with the help of which we could differentiate cancerous lung tissue with normal lung tissue as shown in figure 10.

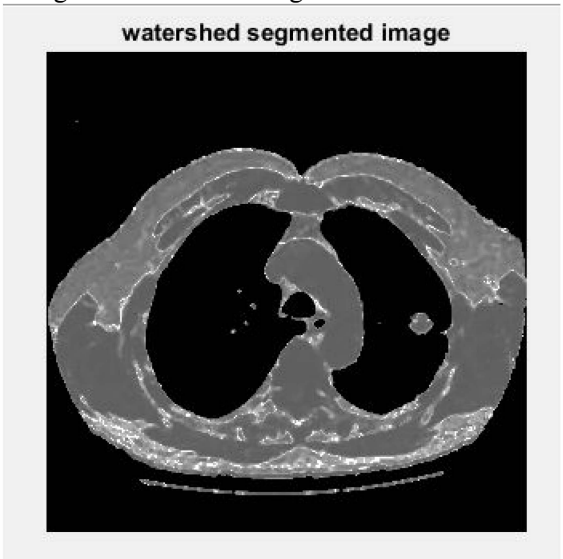

Fig. 10: Watershed Segmentation

h. The output of the watershed image contains pixels of infected lung tissues which are not labeled. These pixels are in the form of watershed lines. Figure 11 shows the output image after successful Morphological operations. This figure clearly represents left lung being more infected by cancer with background stating the cancer region as compared to right lung of the given CT scan image. Also pink marks with the help of binary morphology indicating cancer region affected in both the lungs can be seen in following image

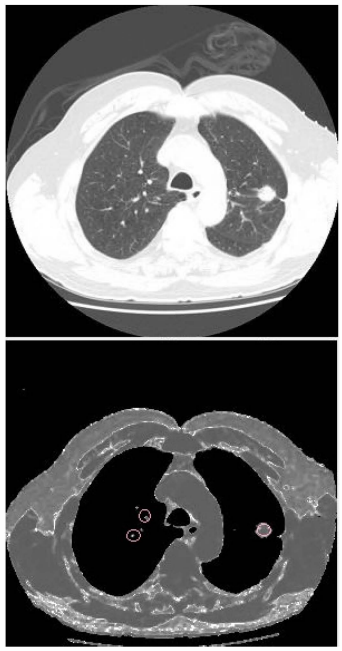

Fig. 11: Output Image with Input Image after Morphological Operations

IV. Conclusion and Future Application

The above technique can be handled in two stages 1) Processing of distortion input image utilizing filter and segmentation 2) Morphological operations on CT picture. The growth influenced lungs locale can be seen in the last algorithm process for particular CT information image. The proposed strategy can likewise be connected to identify some other malignancy like breast cancer, skin malignancy and so forth. Also it finds place in medical research as well.

References

[1] Erich Sackmann, Preface to Volume 1A from Cells to Vesicles: Introduction and Overview, In: R. Lipowsky and E. Sackmann, Editor(s), Handbook of Biological Physics, North-Holland, 1995, Volume 1, Pages ix-x, ISSN 1383-8121, ISBN 9780444819758, http://dx.doi.org/10.1016/S1383-8121(06)80016-3

[2] Hassan Lemjabbar-Alaoui, Omer Ul Hassan, Yi-Wei Yang, Petra Buchanan, Lung cancer: Biology and treatment options, Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, Volume 1856, Issue 2, December 2015, Pages 189-210, ISSN 0304-419X, http://dx.doi.org/10.1016/j.bbcan.2015.08.002. Pierigè F, Serafini S, Rossi L, Magnani M (January 2008). "Cell-based drug delivery". Advanced Drug Delivery Reviews 60 (2): 286–95. doi:10.1016/j.addr.2007.08.029. PMID 1799750.

[3] F. Pierigè, S. Serafini, L. Rossi, M. Magnani, Cell-based drug delivery, Advanced Drug Delivery Reviews, Volume 60, Issue 2, 14 January 2008, Pages 286-295, ISSN 0169409X, http://dx.doi.org/10.1016/j.addr.2007.08.029. "Lung carcinoma: Tumors of the lungs". Merck Manual Professional Edition, Online edition. Retrieved 15 August 2007 in press

[4] Carlos H. Villa, Aaron C. Anselmo, Samir Mitragotri, Vladimir Muzykantov, Red blood cells: Supercarriers for drugs, biologicals, and nanoparticles and inspiration for advanced delivery systems, Advanced Drug Delivery Reviews, Available online 3 March 2016, ISSN 0169-409X, http://dx.doi.org/10.1016/j.addr.2016.02.007.

[5] http://www.cancer.org/cancer/lungcancer-non smallcell/detailedguide/non-small-celllung-cancer-survival-rates.

[6] Hisao Asamura, Kari Chansky, John Crowley, Peter Goldstraw, Valerie W. Rusch, Johan F. Vansteenkiste, Hirokazu Watanabe, Yi-Long Wu, Marcin Zielinski, David Ball, Ramon Rami-Porta, The International Association for the Study of Lung Cancer Lung Cancer Staging Project: Proposals for the Revision of the N Descriptors in the Forthcoming 8th Edition of the TNM Classification for Lung Cancer, Journal of Thoracic Oncology, Volume 10, Issue 12, December 2015, Pages 1675-1684, ISSN 1556-0864, http://dx.doi.org/10.1097/JTO.0000000000000678.

[7] Andrew L. Laccetti, Sandi L. Pruitt, Lei Xuan, Ethan A. Halm, David E. Gerber, Prior cancer does not adversely affect survival in locally advanced lung cancer: A national SEER-medicare analysis, Lung Cancer, Available online 31 May 2016, ISSN 01695002, http://dx.doi.org/10.1016/j.lungcan.2016.05.029.

[8] Yuki Nakajima, Hirohiko Akiyama, Hiroyasu Kinoshita, Maiko Atari, Mitsuro Fukuhara, Yoshihiro Saito, Hiroshi Sakai, Hidetaka Uramoto, Case report of two patients having successful surgery for lung cancer after treatment for Grade 2 radiation pneumonitis, Annals of Medicine and Surgery, Volume 5, February 2016, Pages 1-4, ISSN 2049-0801, http://dx.doi.org/10.1016/j.amsu.2015.11.003.

[9] Ajay Aggarwal, Grant Lewison, Saliha Idir, Matthew Peters, Carolyn Aldige, Win Boerckel, Peter Boyle, Edward L. Trimble, Philip Roe, Tariq Sethi, Jesme Fox, Richard Sullivan, The State of Lung Cancer Research: A Global Analysis, Journal of Thoracic Oncology, Available online 21 March 2016, ISSN 1556-0864, http://dx.doi.org/10.1016/j.jtho.2016.03.010.

[10] Tomoyuki Hishida, Etsuo Miyaoka, Kohei Yokoi, Masahiro Tsuboi, Hisao Asamura, Katsuyuki Kiura, Kazuhisa Takahashi, Hirotoshi Dosaka-Akita, Hideo Kobayashi, Hiroshi Date, Hirohito Tada, Meinoshin Okumura, Ichiro Yoshino, Lobe-specific nodal dissection for clinical stage I-II non-small cell lung cancer: Japanese multi-institutional retrospective study using a propensity score analysis, Journal of Thoracic Oncology, Available online 29 May 2016, ISSN 1556-0864, http://dx.doi.org/10.1016/j.jtho.2016.05.014

[11] D. Brian Larkins, William Harvey, Introductory computational science using MATLAB and image processing, Procedia Computer Science, Volume 1, Issue 1, May 2010, Pages 913-919, ISSN 1877-0509, http://dx.doi.org/10.1016/j.procs.2010.04.100.

[12] Gaëlle Coureau, L. Rachid Salmi, Cécile Etard, Hélène Sancho-Garnier, Catherine Sauvaget, Simone Mathoulin-Pélissier, Low-dose computed tomography screening for lung cancer in populations highly exposed to tobacco: A systematic methodological appraisal of published randomised controlled trials, European Journal of Cancer, Volume 61, July 2016, Pages 146-156, ISSN 0959-8049, http://dx.doi.org/10.1016/j.ejca.2016.04.006.

[13] Kesari Verma, Bikesh Kumar Singh, A.S. Thoke, An Enhancement in Adaptive Median Filter for Edge Preservation, Procedia Computer Science, Volume 48, 2015, Pages 29-36, ISSN 1877-0509, http://dx.doi.org/10.1016/j.procs.2015.04.106.

[14] Shuangfeng Dai, Ke Lu, Jiyang Dong, Yifei Zhang, Yong Chen, A novel approach of lung segmentation on chest CT images using graph cuts, Neurocomputing, Volume 168, 30 November 2015, Pages 799-807, ISSN 0925-2312, http://dx.doi.org/10.1016/j.neucom.2015.05.044.

[15] Hassan Masoumi, Alireza Behrad, Mohammad Ali Pourmina, Alireza Roosta, Automatic liver segmentation in MRI images using an iterative watershed algorithm and artificial neural network, Biomedical Signal Processing and Control, Volume 7, Issue 5, September 2012, Pages 429-437, ISSN 1746-8094, http://dx.doi.org/10.1016/j.bspc.2012.01.002.

[16] Hamid Hassanpour, Najmeh Samadiani, S.M. Mahdi Salehi, Using morphological transforms to enhance the contrast of medical images, The Egyptian Journal of Radiology and Nuclear Medicine, Volume 46, Issue 2, June 2015, Pages 481-489, ISSN 0378-603X, http://dx.doi.org/10.1016/j.ejrnm.2015.01.004.