Identification Of Malignant and Begign Lung CT images using segmentation tools

1*Surekha Nigudgi,

Assistant Profeesor,ACS College Of Engineering,Bangalore, sursanju@gmail.com

2*Dr Channappa Bhyri,

Professor&Head of E&IE dept,PDA College Of Engineering,Kalaburgi, channubhyri@yahoo.com

Abstract

Discovering of cancer is the most extreme interesting investigation space for researcher in the early period.The method used to identify lung cancer with lung CT images in the begining stage so that the survival rate of the patient is high. The methods involve with several stages such as image aquizition,filtering techniques, segmentation techniques, morphological operations and classification.Here in this paper segmentation is perform to disintegrate an lung CT images where gradient magnitude technique is to extract certain essential features from the segmented images ,further extracted features are considered for support vector method to classify whether it is a malignant or begign lung CT images.

Keywords:Morphological operations,Gradient magnitude technique, malignant,begign, SVM,Lung CT Images.

Introduction

Cancer is one of the crucial reasons for non-accidental death. One of the highest among all cancers is the lung cancer. Cancer is a disease caused when cells divide out of control and spread into neighboring tissues.Cells go unpredictable development widely and unbalance to form malignant tumor and damage the surrounding cells. Lung cancer when it diagnosed and treated in early stage there will be less mortality rate associated with this disease. Cigarette smoking is one of main factor , which accounts that the 85% of the patients are with the lung cancer[16]Lung cancer,in nations like industrialized advances the social and cultural patterns have led to rising rates .[15]. Mortality rates of lung cancer among females have lesser than male, about 40 deaths per 100,000 among female and 90 deaths per 100,000 90 among males. These are similar when it is broken down into racial/ethical group[17].Imaging modalities are categorized in which images are generated.ultrasound,radiation such as x-rays,MRI,Computed tomography to develop Computer aided diagnosis. Here in this paper, Lung CT images are used.CT scan lung images are used. CT is a computerized tomography imaging in which is a of x-ray beam is positioned at a patient body and which is rotated around the body, to generate crossectional images or slices of the body.These slices are called as as tomographic images which contain a detailed information for further analysis.[18].

The scope of this proposed methodology is to design a system in which the lung CT images are taken as inputs and gives particular output. The method used are efficient in terms of accuracy, specificity and sensitivity. The proposed methodology consists of following steps such as: Acquisition of lung CT images ,removing unwanted noises with preprocessing steps, contrast stretching is performed., morphological operations is performed , segmentation is done with further extracted features are considered for support vector method to classify whether it is a malignant or begign from lung CT images.

Literature Survey

Ashwini P*, Sherin Antony, Kanchana V et al[1],Authors in this paper they have proposed the solid feature extraction technique to extract essential features from segmented images. Later the extracted features they are considered for multi support vector machine to classify it as cancer or non cancerous. Muzzamil Javaid et al[2], proposed a CAD lungs from CT scan are segmented using Intensity thresholding,for better segmentation results histogram analysis is done. Morphological operations like opening to remove impurities for segmentation outcomes and closing is used for juxtapleural nodules for segmented lung regions. K-means clustering is connected for the beginning location and division of potential nodules. These sectioned potential knobs are at that point isolated into six bunches on the premise of their thickness and rate network with lung dividers.Timor Kadir etal[3] Author proposed an diagram of the most lung cancer forecast approaches proposed to date and highlight a few of their relative qualities and weaknesses. We talk about a few of the challenges within the improvement and approval of such strategies and diagram the way to clinical appropriation. Rachid sammouda et al[4], authors proposed an updated technique of hopfield fabricated neural organize classifier is identified and the method which is used to disintegrate lung areas which means segmentation technique is used it from human chest computer tomography pictures.The images are acquired utilizing this computer tomography imaging techniques from normal subjects and the techniques is used to identify the diseases and diagnoise it.Ratishchandra Huidrom et al[5],proposed a method of segmentation which performs speed than the existing methods.Here the method uses optimal thresholding method the accuracy of the proposed method is speed as compared with the K-means method.Prenitha Lobo [et al], has presented segmentation and classification techniques used in detecting the lung cancer and and it has proposed an effective identification of lung tumor which gave 79.166% accuracy. Shahruk Hossain et al[7],has proposed an automated method for lung tumor detection and segmentation from lung ct images scan which is 3D from Radiomics dataset, further it also presents novel dilated hybrid -3D CNN for tumor segmentation. First segmetation of tumor is implemented and the slices are passed to the segmentation which extracts the features, secondly they are passed through the post processing block which cleans up through morphological operations. Muhammad Zia ur Rehman et al[8],has reviewed a systematic analysis of detection of lung cancer with a analysis of different trends and future challenges. A method has been propose to overcome these challenges and developing a method of computer aided detection (CAD) system.Authors also proposed that a team work is required among soft developers,technicians,physicians and othe related stakeholders to understand deeply about the issues which is particular needs of a computer aided detection systems and where auto matic techniques to overcome the upcoming challenges with image processing ,low cost of implementation and there will low cost implementation and security software assistance.

H. Mahersia, M. Zaroug et al[9], analysed and overviewed that current techniques and approches have been investigated in the literature.It also gives the accuracy and performance of the existing approaches.CAD systems for lung cancer have proposed in several studies. Siddharth Bhatia et al[11], proposed deep learning residual of techniques with preprocessing stage to highlight lung regions which are exposed to cancer and extract features of UNet and ResNet models,later these features are fed into multiple classifiers i.e XGBoost and Random Forest.The accuracy achieved is 84% as compared with the earlier. A. Asuntha1, A.Brindha et al[12],proposed a model and proposed using PSO,genetic optimization and SVM algorithm, these techniques are used for feature extraction and classification, super pixel segmentation and Gabor filter is used for demising the medical images, later comparison is done between three medical images.

Methodology

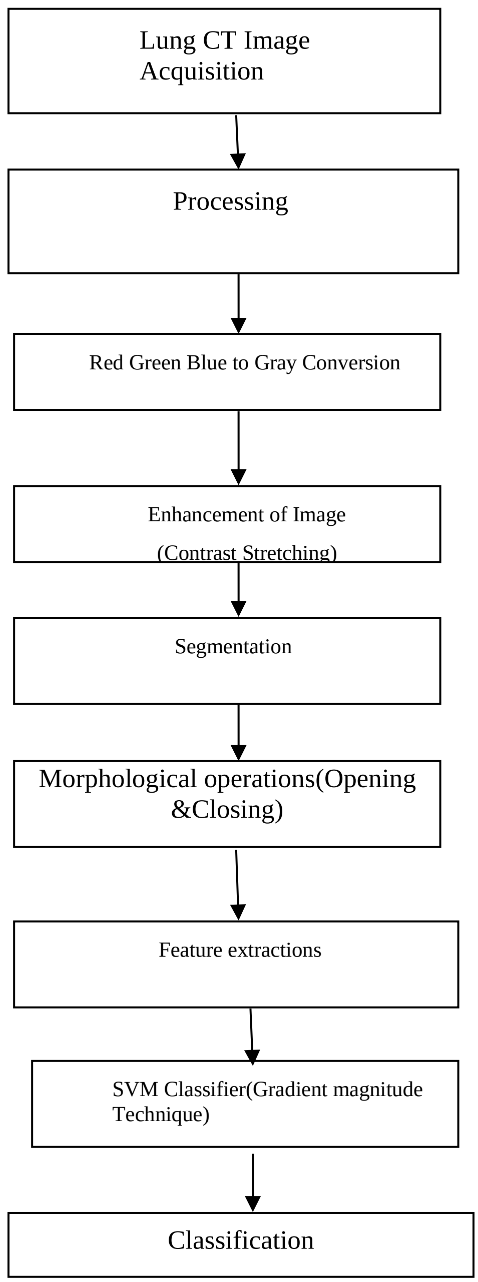


Figure 1:Flow diagram of proposed methodology

The above proposed methodology as shown in figure1. For experimental 20 images are taken from which images are of normal, malignant and begign lung CT images.

Pre-processing;

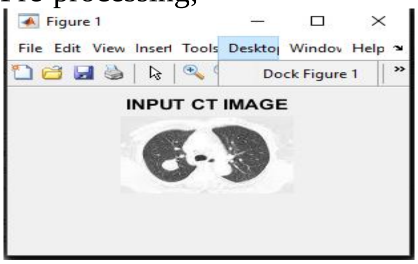


Figure 2:Input CT Image

The aim of the pre-processing is the improvement of the image that suppress the unwanted to remove artifact and denoise from the given CT imges or enhances some important image features for further preprocessing.

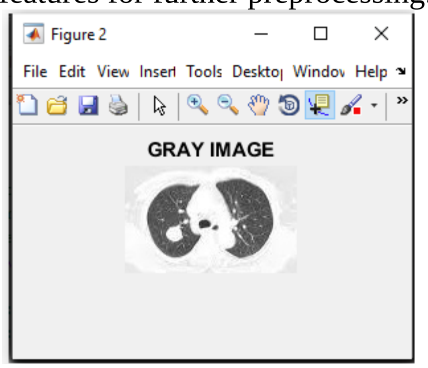
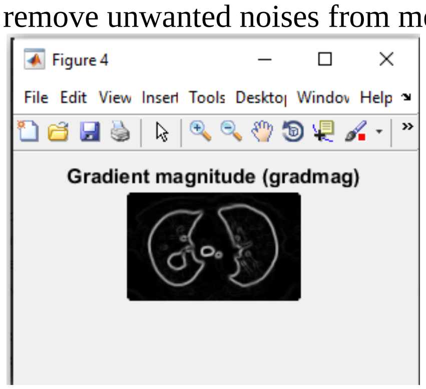


Figure3:Gray Image

The images which are acquired are converted to gray image and remove unwanted noises from median filter and it is enhanced by contrast stretching.



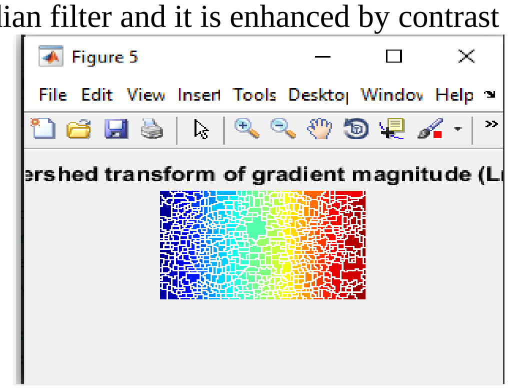


Figure4:Gradient Magnitude Figure5:Watershed transform of gradient magnitude

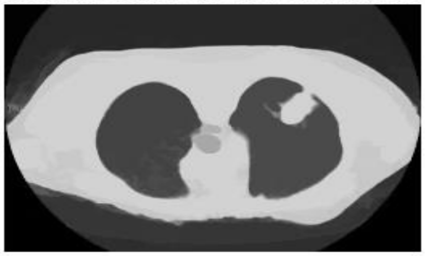


Figure 6:Morphological Operations

Segmentation is a process where it divides into smaller parts and as it increases accuracy, reliability, precision and decreases cost for computational detection[8].The enhanced lung image is then segmented by thresolding technique(method); here in segmentation first the Lung image is divided into left lung and right lung; for particular left lung and right lung cancer is identified. For region of interest in this proposed system the affected region is extracted using median filtering and for that morphological operations is used and perform,dilation and erosion.Here non-linear method is used and it is very effective removing noise and it preserves its edges[13], and in final stage features are extracted using SVM classifier with gradient magnitude technique where in 13 statistical parameters are identified out of which contrast,coorelation,energy,Homogenity,mean,standarddeviation,entropy,RMS,varaiance smoothness IDM has been identified to classify it as normal,malignant and begign.the formulae are as follows:

image contrast = max(grayImage(:)) - min(grayImage(:))

Homogeneity has the similrity between the largest element and smallest element in that region. energy=limit(z(t),t=infinity)

mean Intensity = mean(img(:))

Standard deviation

$$\sigma=\sqrt{\frac{\sum_{i=1}^{N}(Xi-\mu)^2}{N}}$$

Entropy

Entropy = −∑Ngi=1∑Ngj=1 P(i, j)log2[P(i, j)]

RMS:

$X_{rms}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}|X_n^2|}$

Variance: A variance is an image of the variances, that is the squares of the standard deviations, in the values of the input or output images.

variance =∑Ngi=1∑Ngj=1(i − μ)2 P(i, j)

Smoothness:

It is used as a method of smoothing images, reducing the amount of intensity variation between one pixel and the next resulting in reducing noise in images. Y = filter2(h,X) filters the data in X with the two-dimensional FIR filter in the matrix h.

Kurtosis:

$\gamma=\frac{1}{N}\sum_{i=1}^{N}\frac{Xi-u^4}{\sigma}$ -3

IDM(Inverse difference Moment) $=\sum^{Ng-1}{}_{i=0}\sum^{Ng-1}{}_{j=0}$ 1 /1+(i−j)2 p(i, j)

Classification:

Classification is a process where here it is into malignant,begin and normal from lung CT images,for this process suport vector machine is used and it a machine learning algorithm,which can be used for classification to find the hyperplane which will help to differentiate datapoints.From these data points which are falling on either side of hyperplane where it can be classified to different classes.In this step the scope is to identify types of cancer is is a malignant begign and noemal based on sensitivity,specificity and accuracy.

Sensitivity =TP/TP + FN
TP=True Positive.FN=False negative
Specificity =TN/TN + FP
TN=True Negative,FP=False positive

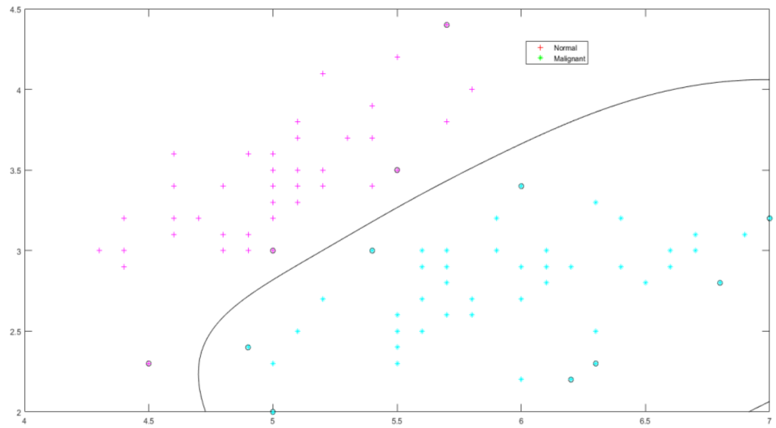


Figure 7:SVM Classifier

Results:

Figure 8:Shows a set of malignant images

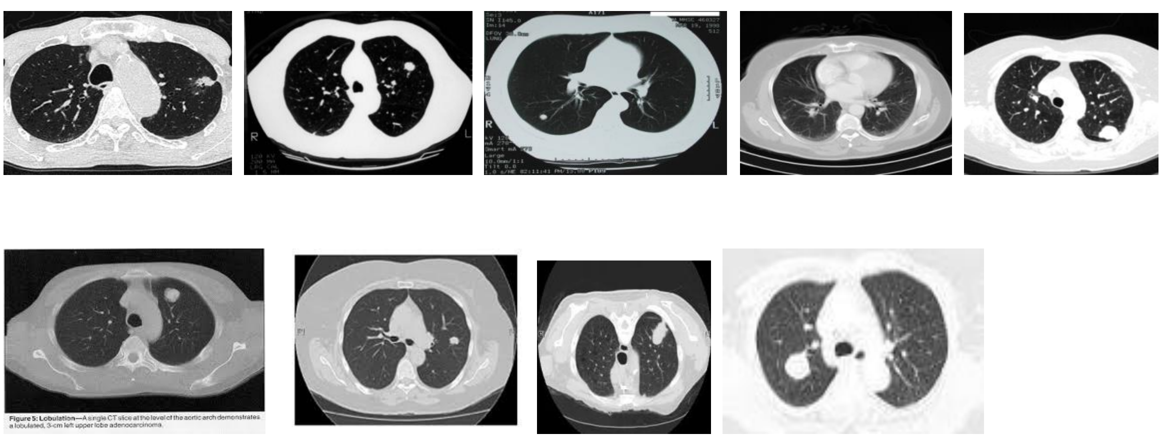


Figure 9:Shows a set of Begign images

Figure 10:Shows a set of normal images

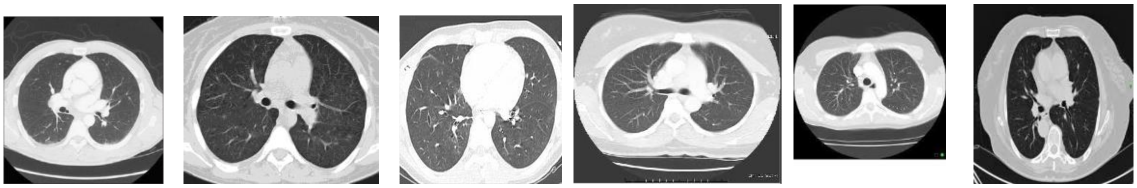


Table1:Shows the support vector machine(SVM) parameters

Features	Normal Images/1			Malignant Images/2			Benign Images/3		
	1	2	3	1	2	3	1	2	3
Contrast	0.1852	0.1898	0.1817	0.1228	0.2129	0.1402	0.136	0.1727	0.2571
Correlation Index	0.0911	0.1316	0.0715	0.2006	0.1867	0.248	0.1191	0.1381	0.1852
Energy	0.7967	0.8194	0.9204	0.9159	0.9034	0.9212	0.8683	0.8802	0.8898
Homogeneity	0.9452	0.9509	0.9779	0.9769	0.9718	0.9782	0.9645	0.968	0.9678
Mean	0.0029	0.0017	0.001	0.0018	0.003	0.0023	0.0011	0.0022	0.0026
Standard deviation	0.0771	0.0801	0.0611	0.0573	0.0668	0.0606	0.0651	0.0686	0.0762
Entropy	3.3203	2.9404	2.4485	1.5433	1.5501	1.1941	1.7795	2.6656	1.7569
RMS	0.0772	0.0801	0.0611	0.0574	0.0668	0.0606	0.0651	0.0687	0.0762
Variance	0.0059	0.0064	0.0037	0.0033	0.0045	0.0037	0.0042	0.0047	0.0058

Smoothness	0.9261	0.886	0.9365	0.972	0.9707	0.9696	0.9063	0.9531	0.9486
Kurtosis	7.7963	9.4719	54.4222	41.0523	48.6739	40.8132	14.5392	24.819	40.2894
Skewness	0.7071	0.76	3.8035	2.4703	3.605	2.6496	0.8829	1.9758	3.0082
IDM(Inverse Difference Moment)	0.9897	0.5833	5.047	2.7892	0.5046	4.1863	-0.3542	0.1121	2.0853

Table2:List of parameters of Coorelation,entropy,Kurtosis and Skewness

Parameters	Normal			Malignat			Begign		
Correlation Index	0.0911	0.1316	0.0715	0.2006	0.1867	0.248	0.1191	0.1381	0.1852
Average	0.098			0.2117			0.147		
Entropy	3.3203	2.9404	2.4485	1.5433	1.5501	1.1941	1.7795	2.6656	1.7569
Kurtosis	7.7963	9.4719	54.4222	41.0523	48.6739	40.8132	14.5392	24.819	40.2894
Skewness	0.7071	0.76	3.8035	2.4703	3.605	2.6496	0.8829	1.9758	3.0082

Discussion

For experimentation 20 image are taken from which images are of normal lung ct images and images are of malignant and begign.The images acquired are converted to gray image and then unwanted noises is removed using median filter and then the image is enhanced by contrast stretching. This is the enhancement method which improves image by stretching the range of intensity values. The image is resized to 250*250.The resized image is segmented right and left to determine the cancer nodules in the lungs. This phase will help to identify the region in the lung nodule that can help to identify the normal, malignant and begign.

The 13 statistical parameters features for lung CT images are taken for experimentation using sequential forward selection algorithm in SVM classifier is applied. Here 2 classes (Class1 as cancer in fig 11,class2 as normal in fig 12) is used. An support vector machine is a classifier which differentiate the data to separate all data points of one class from those of other class by finding the best hyperplane.Performance is measured by accuracy, specificity ,sensitivity as shown below;

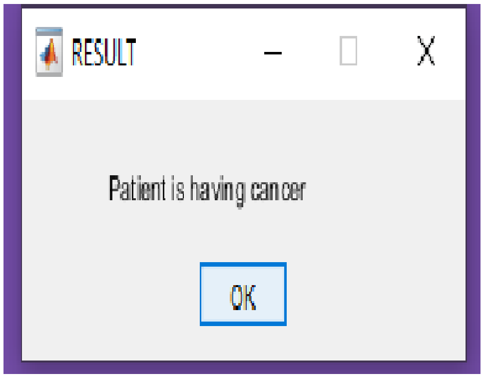


Figure11:Class1 as Cancer

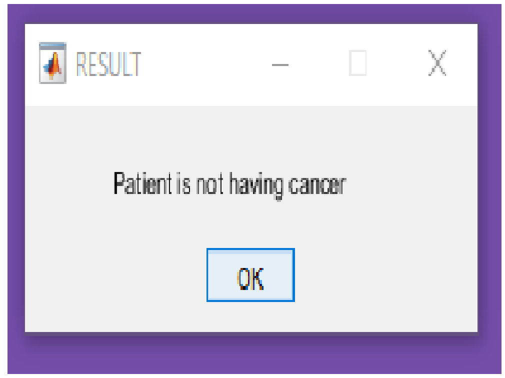


Figure12:Class2 as Normal

Calculation:

Accuracy Analysis is performed for one parameter.

From the above Table2: 0.2 value is a threshold which is fixed by taking the average of normal,malignant and begign.
If 0.2> then it is malignant,
If 0.2<then it is a normal or begign.

Coorelation

A set of 20 images are taken for experimentation out of which normal, begign and malignant. Only 9 images are taken and are tabulated in table1 from which the specificity, sensitivity and accuracy are calculated one in normal is true negative and one in malignant also true negative and remaining 7 images are true positive so

Sensitivity = True positive /True pasitive +False Negative

=7/9=77%

Specificity=True Negative/True positive +false negative

=2/9=22%

Accuracy=True positive /Total number of images

=7/9=77%

sensitivity =77%

specificity = 22%

accuracy =77%.

Table 3:Performanace of SVM classifier

Dataset	Accuracy	Specificity	Sensitivity
Lung Images	90%	22%	90%

Conclusion

A Set of 20 images of which normal,begign and malignant lung CT images are taken for experimentation. The 13 statistical parameters like contrast, correlation index,Energy,Homogenity,mean,standarddeviation,entropy,RMS,variance, smoothness,kurtosis,skewness,IDM(inverse,difference moment). The proposed methodology is to classify normal,begign and malignant tumors correctly by SVM classification and handle for better tumor classification[13]. The proposed method is providing 90% accuracy but, the 100% accuracy of the method can be concluded by considering more images.

References

1)Ashwini P*, Sherin Antony, Kanchana V Department of Computer Science, Amrita School of Arts and Sciences, Amrita Vishwa Vidyapeetham Campus, Mysuru, Karnataka, India" Categorizing the stages of lung cancer using Multi SVM Classiʕier" International journal of research in pharmaceutical sciences ,10(3), 2323-2328 ISSN: 0975-7538,2019, Pharmascope.org © 2019.

2) Muzzamil Javaid a,, Moazzam Javid b, Muhammad Zia Ur Rehman a,Syed Irtiza Ali Shah a " A novel approach to CAD system for the detection of lung nodules in CT images" computer methods and programs in biomedicine 135 (2016) 125–139, journal homepage: www.intl.elsevierhealth.com/journals/cmpb. cmpb.2016.07.0310169-2607/© 2016.

3) Timor Kadir1, Fergus Gleeson2 " Lung cancer prediction using machine learning and advanced imaging techniques" Translational Lung Cancer Research, Vol 7, No 3 June 2018, Transl Lung Cancer Res 2018;7(3):304-312.

4) Rachid Sammouda"Segmentation and Analysis of CT Chest Images for Early Lung Cancer Detection"2016 Global Summit on Computer & Information Technology, 978-1-5090-2659-3/17 $31.00 ©[2017] IEEE DOI 10.1109/GSCIT.2016.29.

5) Ratishchandra Huidrom, Yambem Jina Chanu, Khumanthem Manglem Singh " A Fast Automated Lung Segmentation Method for the Diagnosis of Lung Cancer" Proc. of the 2017 IEEE Region 10 Conference (TENCON), Malaysia, November 5-8, [2017].

6) Prenitha Lobo ,Sunitha Guruprasad " Classification and Segmentation Techniques for Detection of Lung Cancer from CT Images " Proceedings of the International Conference on Inventive Research in Computing Applications (ICIRCA 2018) IEEE Xplore Compliant Part Number:CFP18N67-ART; ISBN:978-1-5386-2456-2.

7) Shahruk Hossain Suhail Najeeb Asif Shahriyar Zaowad R. Abdullah M. Ariful Haque," A pipeline for lung tumor detection and segmentation from ct scans using dilated convolution neural networks" ICASSP 2019, 978-1-5386-4658-8/18/$31.00 ©[2019] IEEE.

8) Muhammad Zia ur Rehman∗, Muzzamil Javaid, Syed Irtiza Ali Shah, Syed Omer Gilani,Mohsin Jamil, Shahid Ikramullah Butt" An appraisal of nodules detection techniques for lung cancer in CT images" https://doi.org/10.1016/j.bspc.2017.11.017 1746-8094/© [2017] Elsevier Ltd. M. Zia ur Rehman et al. / Biomedical Signal Processing and Control 41 (2018) 140–151 1.

9) H. Mahersia1, M. Zaroug1,L. Gabralla2 "Lung Cancer Detection on CT Scan Images: A Review on the Analysis Techniques"(IJARAI) International Journal of Advanced Research in Artificial Intelligence, Vol. 4, No.4, [2015],www.ijarai.thesai.org.

10) Ashutosh Kumar Dubey*, Umesh Gupta and Sonal Jain," Epidemiology of lung cancer and approaches for its prediction: a systematic review and analysis"Dubey et al. Chin J Cancer (2016) 35:71 DOI 10.1186/s40880-016-0135-x pp2-13,chinese journal of cancer.

11) Siddharth Bhatia, Yash Sinha and Lavika Goel."Lung Cancer Detection:A Deep Learning Approach"© Springer Nature Singapore Pte Ltd. [2019] J. C. Bansal et al. (eds.), Soft Computing for Problem Solving, Advances in Intelligent Systems and Computing 817, https://doi.org/10.1007/978-981-13-1595-4_55.

12) A. Asuntha1, A.Brindha1, S.Indirani1, Andy Srinivasan2"Lung cancer detection using SVM algorithm and optimization techniques"ISSN: 0974-2115 www.jchps.com Journal of Chemical and Pharmaceutical Sciences, October - December [2016], JCPS Volume 9 Issue 4.

13) Keziah T A1 , Haseena P2" Lung Cancer Detection Using SVM Classifier and MFPCM Segmentation" International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 05 Issue: 04 | Apr-[2018] ,www.irjet.net p-ISSN: 2395-0072,pp3114-3118.

14) https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/

15) https://pubmed.ncbi.nlm.nih.gov/30741509/

16) Suresh S. Ramalingam, MD1; Taofeek K. Owonikoko, MD, PhD2; Fadlo R. Khuri, MD3" Lung Cancer: New Biological Insights and Recent Therapeutic Advances" CA CANCER J CLIN 2011;61:91–112.

17) Lindsey A. Torre , Rebecca L. Siegel , and Ahmedin Jemal " Lung Cancer Statistics " © Springer International Publishing Switzerland 2016 1 A. Ahmad, S. Gadgeel (eds.), Lung Cancer and Personalized Medicine, Advances in Experimental Medicine and Biology 893, DOI 10.1007/978-3-319-24223-1_1.

18) https://www.nibib.nih.gov/science-education/science-topics/computed-tomography-ct.