LUNG CANCER IMAGING USING COMPUTER TOMOGRAPHY(CT) IMAGES

DIGITAL IMAGE PROCESSING: SWE1010

FINAL REVIEW

LAASYA YARLAGADDA - 19MIS0138

SABRINA MANICKAM - 19MIS0137

ABSTRACT

Lung cancer is by far the leading cause of cancer death among men and women, making up almost 25% of all cancer deaths. Early detection of lung tumour can help reduce deaths caused by lung cancers. We, therefore, aim at a system for detection of lung cancer in Computer Tomography (CT) images.

Image pre-processing techniques such as enhancement and restoration will be used. Image segmentation among others processes will be employed to effectively detect and mark the regions with lung cancer. In the later stages on this project, we aim to process, segment and correctly mark out a region of interest in CT images of lung cancer patients taken from the ELCAP Public Lung Image Database using Python.

LITERATURE REVIEW

1. Mahmudul Islam et al. [1] motivated by the recent progress of image processing and machine learning techniques in medical field have developed techniques using Gray level co-occurrence matrix-based texture image analysis and statistical parametric approach to assist doctors in detecting lung cancer. For feature extraction the approaches used are: Gray level co-occurrence matrix (GLCM) based texture image analysis and Statistical parametric approach.

2. Nidhi S. Nadkarni et al. [2] state that image processing techniques are used widely in medical fields for early-stage detection of lung tumour and present an automated technique for identification of lung cancer in CT scan images. Proposed method uses algorithms such as median filtering for image pre-processing followed by segmentation of lung region of interest using mathematical morphological operations. Support vector machine is used to classify CT images into normal and abnormal.

3. Wadood Abdul [3] works on automatic lung cancer detection and classification (ALCDC) system. Inspired by the outstanding deep learning (DL) progress in recognition related tasks, an ALCDC system for CT scan employing DL is introduced. They also use CNN to decide whether a tumour is benign or malignant.

4. D. Jayaraj et al. [4] propose a computer-based analysis tool has been developed for utilizing image processing approaches. In this paper, identification of lung cancer has been done using a new automated computer aided model. Four main stages namely pre-processing, segmentation, feature extraction and classification have been presented in this model.

5. Chethan K S et al. [5] proposed UNET architecture of CNN model is implemented. From these Deep Learning database, machine learning and deep learning algorithms happen to be the most employed for the implementation of 3D Convolution Neural Networks and TensorFlow.

6. Qurina Firdaus et al. [6] suggest a study aims to develop a lung cancer detection system based on CT-scan images. The proposed detection process has 4 stages, Pre-processing (improving the quality of CT-scan images), segmentation (identification and separation of cancer object), feature extraction (based on area, contrast, energy, and homogeneity), and lastly classification (cancer to benign or malignant).

7. Sachin Bhat et al. [7] proposed that in early stages of CT-scan images for lungs the classification and recognition task can be performed for detection of lung tumour. Also, the lung images were used from the Lung Image Database Consortium and Infections Disease Research Institute (LIDC/IDRI). The proposed methodology incorporated the use of Convolutional Neural Network (CNN) for the recognition and classification task. The overall comparison was made between the classical machine learning algorithms from the literature with standard evaluation metric with the results from this study. At the end it was observed that Deep CNN classifier outperformed all the other traditional classification algorithms.

8. Mehdi Hassan Jony [8] used grey level Co-occurrence Matrix technique for extraction of affected lung images. Also, it was purposed that the use of Support Vector Machine (SVM) for the detection of abnormal Lung images.

9. Shailesh S. Bhise et al. [9] designed Computer based diagnosis system to develop detection of lung cancer with the help of image processing techniques and artificial neural network (ANN) classifier. For the elimination of background and surrounding tissue, during pre-processing steps image enhancement techniques were used. Region of interest (ROI) is calculated using region-based segmentation algorithm. The desired nodule is extracted using circle fit algorithm. In the feature extracting step the features used were, ECD, Euler number, Area, Mean Intensity and Radius. The Artificial Neural Network (ANN) was trained using back propagation algorithm in the categorization stage.

10. Pramit Brata Chanda et al. [10] proposed pattern recognition for recognition of disease affected region in a better way. Otsu threshold based segmentation was used on abnormal images to segment the cancer affected regions. To change the threshold parameters dynamically adaptive threshold methods are used for segmentation. For the accurate detection of edges in the lung area gradient and sobel based detection was used.

11. Azmira Krishna et al. [11] proposed a better cancer detection technique compared to the earlier techniques by the use of Multi-layered Perceptron Back Propagation Neural Network (MLP-BPNN) which is based on Scale Invariant Feature Transform (SIFT). This showed a better accuracy rate of 89% than other techniques.

12. Vikul J. Pawar et al. [12] proposed Region of Interest (ROI) extraction based on the morphology operation, this can diminish misleading results. Convolution Neural Network can be used to achieve better possible results and also that computer aided detection has advantages of speed and accuracy.

13. Aisiri A P et al. [13] shows that cell breakdown in the lungs, all throughout the world, affects a large number of people since 2008 estimates. For productive treatment it is critical to evaluate early disclosure of cell breakdown. Few methods exist for the recognition of harmful cells. Here, process for division, for instance, watershed is performed for recognition of the infected cell and also to find better approach out of them.

14. M.Bikromjit Khumancha et al. [14] propose a study that aims to develop a lung cancer detection system based on CT-scan images. The proposed detection process has 4 stages, Pre-processing (improving the quality of CT-scan images), segmentation (identification and separation of cancer object), feature extraction (based on area, contrast, energy, and homogeneity), and lastly classification (cancer to benign or malignant).

15. B. Nirupriya et al. [15] propose work that aims to implement a methodology to extract the LN from the Computed-Tomography (CT) image with a considerable accuracy. This work implements Social Group-Optimization (SGO) and Kapur’s threshold (SGO+KE) to enhance the CT image. Later the Active-Contour (AC) and Watershed-Segmentation (WS) is executed to extract the LN. The merit of the proposed work is confirmed based on the performance measures attained with the proposed tool.

16. Hinisha K.V et al. [16] propose a system to detect lung nodule by using CT image. It consists of five stages, namely Pre-Processing, Lung Region Extraction, Nodule Segmentation, Feature Extraction, and Classification. Even the smallest nodules are identified in various slices. This can help radiologists and doctors to detect lung cancer in early stages. The CT images for the testing of this project are acquired from the Lung Image Database Consortium (LIDC).

17. Amitava Halder et al. [17] aimed to develop an automated computer-aided lung nodule detection system from HRCT images to provide a reliable second opinion to the radiologist and expert for further treatment. In this work a morphological filter aided Gaussian Mixture Model (GMM) is introduced for nodule segmentation and candidate detection. For nodule detection Support Vector Machine is used.

18. Soon Yee Chong et al. [18] aim with this research is to develop an image segmentation algorithm for nodule detection in computed tomography (CT) image. Performance of the obtained segmentation algorithm is checked through testing on the lung CT image.

19. Silvia Moreno et al. [19] present the state of the art of image processing techniques applied in the study of lung cancer, emphasizing in two main tasks: segmentation of nodules or tumours, and extraction of useful features for classification and prognosis of tumour evolution using Radiomics.

20. Qinhua Hu et al [20] propose an automatic segmentation of the lungs in CT images, using the Convolutional Neural Network (CNN) Mask R-CNN, to specialize the model for lung region mapping, combined with supervised and unsupervised machine learning methods (Bayes, Support Vectors Machine (SVM), K-means and Gaussian Mixture Models (GMMs)).

21. Caixia Liu et al. [21] propose an algorithm wherein an image decomposition-based filtering strategy is first introduced to denoise lung CT images while preserving their lung contours. Wavelet transformation combined with a group of morphological operations has been used to segment lung CT images. Further refining by a contour correction approach, built on a fast edge detection technique, is used to refine and smooth the segmented lung contours.

22. Amitava Halder et al. [22] suggest an adaptive morphology-based segmentation technique (AMST). It has been introduced by designing an adaptive morphological filter so as to provide better segmentation lung nodule regions. Adaptive structuring element (ASE) is applied by the adaptive morphological filters to identify possible nodule regions. It also improves nodule detection accuracy by minimizing false positives (FPs). Feature extraction is carried out next on the identified nodule regions. This study incorporates a support vector machine (SVM) for lung nodule detection.

23. R. Mohana Priya et al. [23] propose an automatic segmentation and classification of fused lung Computed Tomography (CT) and Positron Emission Tomography (PET) images is presented. This system includes four essential stages namely lung image fusion process; segmentation of fused lung CT/PET images; post and pre-processing of CT images followed by classification of fused lung images.

24. Prasad Dutande et al. [24] proposed a novel approach for segmentation, classification and detection of lung nodules from CT scan images. A maximum intensity projection method has been proposed for pre-processing the images.

25. Yuxin Zhou et al. [25] propose a technique that describes characterization methods for detecting cellular breakdown in the lungs. The lung images and their database in the basic three stages of pre-processing, division and highlight extraction stage to achieve greater quality and accuracy at the site of the cellular breakdown in the lungs. The Convolutional Neural Network providing precise order applications and strategy for detecting cellular breakdown in lung using channels and division methods is proposed.

26. Min Li et al.[26] study enclosed comparatively rare respiratory organ adenosquamous cancer (ASC) samples for the primary time, and projected a computer-aided identification technique supported histopathological pictures of ASC, respiratory organ epithelial cell cancer (LUSC) and little cell respiratory organ cancer (SCLC). Their experimental results verify the potential of the auxiliary identification model created by machine learning (ML) within the identification of LC.

27. Disha Sharma et al.[27] the most objective of the project is to develop a CAD (Computer motor-assisted Diagnosis) system for locating the first carcinoma nodules victimization the respiratory organ CT pictures and classify the nodules as Benign or Malignant. The machine-controlled laptop motor-assisted identification (CAD) system is projected during this paper for detection of carcinoma type the analysis of X-raying pictures. Then the segmentation algorithmic program is applied so as to sight the cancer nodules from the extracted respiratory organ image. Initially, the essential image process techniques like Erosion, Median Filter, Dilation, Outlining, and respiratory organ Border Extraction are applied to the CT scan image so as to sight the respiratory organ region. For experimentation of the projected technique, the CT pictures are obtained from a NIH/NCI respiratory organ Image information pool (LIDC) dataset that has the possibility to try and do the advised analysis.

28. R.Indumathi et al.[28] projected an image process technique to phase the carcinoma from CT image. This CAD system provides the radiologists and aid skilled to seek out the cancer at an early stage accurately and chop-chop. The input CT pictures are obtained from the respiratory organ Image information pool (LIDC).

29. Mohd Firdas et al.[29] projected technique includes the subsequent steps by victimization image process techniques: knowledge assortment, knowledge pre-processing, options choice and carcinoma classification. The central objective of this study is so to determine a picture process technique for extracting options of carcinoma from CT scan pictures. This paper revolves around the categorization of carcinoma Stages from CT Scan pictures victimization Image process and k-Nearest Neighbors. The pre-processing was done employing a median filter to get rid of noise contained within the pictures.

30. Diksha Mhaske et al.[30] projected a complicated computer-aided identification (CAD) system victimization deep learning algorithms that may with efficiency extract knowledge from CT scan pictures and supply precise and timely identification of carcinoma. This work focuses on utilizing the deep learning techniques, specifically Convolutional Neural Network (CNN) for feature extraction and repeated neural network (RNN-LSTM) for carcinoma classification and obtains high accuracy.

31. Zhitong Shanghai dialect et al.[31] projected a coarse-to-fine respiratory organ nodule segmentation technique by combining image improvement and a Dual-branch neural network. The projected image improvement technique improves the effectiveness of network learning, whereas the dual-branch neural network explores multi-view info. Second, Dual-branch network supported U-Net (DB U-Net) which may effectively explore info from each second slices and therefore the relationships between neighboring slices for a lot of precise and consistent segmentation. They preprocessed the image to boost the discrimination of the nodules and roughly find the lesion space so we are able to eliminate the noises from background and target learning the options round the boundaries.

32. Heng Yu et al.[32] projected AHHMM to investigate deep learning supported the historical medical care theme within the development of Non-Small Cell respiratory organ Cancers (NSCLC) machine-controlled radiation adaptation protocols that aim at optimizing native growth regulation at lower rates of grade a pair of RP2 radiation inflammation. Moreover, the system projected consists of many steps as well as getting the image, preprocessing, binarization, thresholding, and segmentation, extraction of options and detection of deep neural networks (DNN).

33. Gayatri Parasa et al.[33] projected a model to advance an automatic carcinoma sighting system in relatively high preciseness and provides widespread usage so as to detect carcinoma within the early stages itself, leading to life risk likelihood. machine-controlled technique with the mix of high boost filtering, Fuzzy C means that segmentation is projected that aids within the detection of carcinoma. The experimental results of the projected machine-controlled technique victimization Back Propagation Neural Network is tried to be higher than the usage of the algorithms like Support Vector Machine and Probabilistic Neural Network. Technique victimization Back Propagation Neural Network provides classification rate of eighty six and algorithms like Support Vector Machine and Probabilistic Neural Network provides classification rate of eighty fifth and eighty two.

34. R.Jenkin Suji et al.[34] conferred a typical thresholding and morphological operation based, mostly respiratory organ segmentation framework, applied to a number of the variants of thresholding techniques, viz. Fixed, bar chart based mostly, entropy supported 3D respiratory organ volume slice by slice and evaluates the standard of results obtained victimization 5 analysis metrics and presents the results and discussion on an equivalent.

35. Hela Mahersia et al.[35] projected a comparative classification technique supported SVM classifier, theorem regularization networks and ANFIS classifier, on the LIDC information, showing the hardiness of the approach projected. This analysis aims to determine a technique for the machine-controlled identification of respiratory organ nodules victimization image process and pattern recognition techniques. The projected pre- process system aims to delimit the respiratory organ tissue by deleting all the spare regions. The automated system that we tend to propose during this paper includes a pre-processing stage, a nodule characterization stage and a classification stage.

36. Arun B. Mathews et al.[36] centralized super pixels segmentation primarily based unvaried agglomeration (SSBIC) and advanced optimization methodology for precise segmentation of respiratory organ nodules. Within the future step the Super constituent Segmentation primarily based unvaried agglomeration (SSBIC) algorithmic rule is enforced in Associate in Nursing increased nodule image sequence for abnormal respiratory organ tissue prediction. Average segmentation time for nodule slice order is one.06s. The best classification accuracy is ninety seven by the Advanced GWO with ONN (AGWO-ONN) methodology and ninety seven.6% by the Advanced GWO with CNN (AGWO-CNN) methodology. Ultimately, the respiratory organ nodule pictures square measure procured by utilizing a complicated GWO with ONN (AGWO-ONN) and deep learning primarily based Advanced GWO with CNN (AGWO-CNN). The supreme intent of this analysis paper is the improvement of respiratory organ CT pictures to acknowledge the neoplasm with efficiency and small-scale abnormal nodule segmentation within the respiratory organ region.

37. Ariful Hoque et al.[37] planned a brand new framework for carcinoma diagnosing mistreatment numerous options extracted from computerized tomography pictures wherever totally different steps square measure used like improvement, median, filter, segmentation, feature extraction and support vector machine. computerized tomography (CT) pictures square measure accustomed to establish carcinoma at its early stage.

38. Mahmudul Islam et al.[38] was impressed by the recent success of image process and machine learning techniques in medical field and developed models mistreatment grey level co-occurrence matrix (GLCM) primarily based texture image analysis and applied math constant quantity approach for serving to doctors to sight carcinoma stages. For feature extraction purpose 2 approaches square measure used: grey level co-occurrence matrix (GLCM) primarily based texture image analysis and applied math constant quantity approach.

39. CMAK Zeelan Basha et al.[39] planned Associate in Nursing Automatic carcinoma Detection System mistreatment improved Haar moving ridge rework, Scale-Invariant Feature rework (SIFT), Back Propagation Neural Network (BPNN), and Watershed Segmentation. The validation results are planned to be ninety one correct compared to applying totally different algorithms.

40. Shahruk Hossain et al.[40] planned a way within which to phase the tumors, the chosen slices square measure passed to the segmentation model that extracts feature maps from every 2nd slice mistreatment expanded convolutions then fuses the stacked maps through 3D convolutions - incorporating the 3D structural info gift within the CT scan volume into the output. 3D respiratory organ CT scans from the NSCLC-Radiomics Dataset.

41. Q. Wei et al.[41] given a moving ridge algorithmic rule that permits automatic identification of the left and right oblique fissures, similarly semi-automatic identification of the horizontal fissures. Tested on eight, half-dozen and half-dozen stacks of identical CT pictures for the left oblique, right oblique and horizontal fissures, severally, the algorithmic rule yielded Associate in Nursing accuracy of seventy seven.1 – 93.6% with strict analysis criteria. This algorithmic rule took a two-stage approach: (a) accommodative fissure sweeping to search out fissure regions; and (b) moving ridge rework to spot the fissure locations and curvatures inside these fissure regions.

42. Mustafa Alam et al.[42] planned a completely unique patch-based multi-atlas methodology with 3 main steps: a) atiny low set of atlases is chosen by examination the target image with a bigger set of atlas pictures employing a size-shape primarily based feature vector, b) respiratory organ nodules square measure designated employing a patch-based methodology, wherever every constituent of a target image is labeled by examination the image patch, focused by the constituent with patches from Associate in Nursing atlas library and selecting the foremost probable labels per an outlined highest match criterion and Laplacian of Gaussian blob detection methodology is employed to search out the divided space of the respiratory organ nodule.

43. Chinmayi Thallam et al.[43] concerned examination numerous classification and ensemble models like Support Vector Machine(SVM), K Nearest Neighbour (KNN), Random Forest(RF), Artificial Neural Networks (ANN) and a hybrid model, option classifier.

44. Kyamelia Roy et al.[44] developed the exactness Associate in Nursing verify specific words for detection of respiratory organ cancer at an earlier stage employing a mixture of medicine image process technique and information Discovery in knowledge. The image of the lungs that's obtained from the

CT(Computed Tomography) scan pictures is being pre-processed and therefore the segmentation is being exhausted the Region Of Interest(ROI), Random Forest methodology are applied to classify the distinct options. Salience improvement has been done and mistreatment of SURF(Speeded Up strong Features) algorithmic rule options like entropy, correlation, energy and variance are extracted from the salience increased pictures with the assistance of a SVM (Support Vector Machine) Classifier.

45. B. Hemalatha et al.[45] planned a model within which at the start, the CT scan image is pre-processed for removing the unwanted signals and smoothing them by using Improved Kaun Filter (IKF). At last, the neoplasm has been categorised by Elman Neural Network (ENN) and weights square measure optimized with PSO and compared the accuracy results with SVM, RBFN and ANFIS.

46. R.Sathishkumar et al.[46] exaggerated the distinction of the image with the CLAHE effort technique .Then it's divided with the assistance of stochastic process segmentation methodology. In segmentation the 3 methods can happen if the ROI of the image is divided then the border correction is completed.

47. K.Karthick et al.[47] planned a model within which the classification of Machine Learning primarily based Reinforcement Learning Algorithms may be a mathematical relation and likelihood models from Medical Image process classify pictures into affected and non-affected components which may be analyzed through the networking methodology mistreatment the web of Things (IOT), and it'll store or monitor and show.

48. Prenitha Lobo et al.[48] reviewed classification and segmentation techniques employed in police work on the carcinoma neoplasm and evaluated the performance of every approach. It's a planned way for effective identification of carcinoma. The planned methodology gave79.166% accuracy.

49. Varsha Prakash et al.[49] given a comprehensive analysis of various carcinoma detection techniques for predicting the nodule as benign or malicious. It absolutely was divided into four sub-sections like pre-processing techniques, segmentation and nodule extraction, nodule classification and therefore the final sub-section CNN primarily based approach is divided into three sub-sections namely: knowledge augmentation, nodule detection and classification.

50. S.Ziyad et al.[50] was galvanized to investigate the preprocessing stage that includes a distinction sweetening stage of respiratory organ pictures. During this regard, the performance of various distinction sweetening strategies is compared for respiratory organ image obtainable within the public LIDC info victimisation normal distinction analysis metrics. A CAD solely detects the pneumonic nodule within the parenchyma and alternative near regions, whereas CADx not solely identifies the pneumonic nodules however conjointly classifies them as malignant and benign.

LITERATURE SURVEY:

1) Lung Image Database Consortium Image collection (LIDC-IDRI)

Reference No	Paper Name	Image acquisition or datasets	Image Enhancement	Image Restoration	Image Segmentation	Features Extraction	Classifiers	Quality Metrics	Results	Remark
1	Analysis of CT Scan Images to Predict Lung Cancer Stages Using Image Processing Techniques	Lung Image Database Consortium Image collection (LIDC-IDRI)	Gabor filter	Median filtering	Global thresholding and Otsu's technique	Gray Level Co-occurrence Matrix based Texture analysis and Statistical Parametric approach	Support vector machine Nearest Neighbours Random Forest, Naive Bayes:	Accuracy, precision, recall	No correlation ship found among features so the obtained features are remained for training.	Aimed at improving accuracy in detection of lung cancer
2	Detection of Lung Cancer in CT Images using Image Processing	Lung Image Database Consortium Image collection (LIDC-IDRI)	Contrast adjustment	Median filtering	Morphological opening operation, border operation on opened image	Extraction of three geometrical features of lung tumor: Area, Perimeter, eccentricity	SVM classifiers	Accuracy	The proposed CAD system is capable of detecting nodules of lungs, larger than 25 mm of diameter, which	Aimed at providing automatic detection of lung cancer using CT Images

									means they are able to detect nodules in the lungs at the initial stage.	
3	An Automatic Lung Cancer Detection and Classification (ALCDC) System Using Convolutional Neural Network	Lung Image Database Consortium Image collection (LIDC-IDRI)	-	-	CT image is reduced to 28× 28,	s 8296 nodules (4329 benign and 3967 malignant), each 2-D CT slice is then used as one training sample	Convolutional Neural Network (CNN)	Accuracy, sensitivity, specificity	Proposed ALCDC system gives an accuracy of 97.2%	Aimed at providing automatic detection of lung cancer using CT Images
4	Random Forest based Classification Model for Lung Cancer Prediction on Computer Tomogra	Lung Image Database Consortium Image collection (LIDC-IDRI)	Median filer	Gaussian Filter	Watershed segmentation technique	Area, eccentricity, diameter, mean intensity and centroid are extracted	RF based Classification	Accuracy, sensitivity, specificity	Presented model attains superior classification performance by attaining the maximum accuracy of 89.90,	Aimed at creation of new model to detect lung cancer

	phy Images								sensitivity of 90.85 and specificity of 88.32 respectively	
7	Convolutional Neural Network approach for the Classification and Recognition of Lung Nodules	Lung Image Database Consortium Image collection (LIDC-IDRI)	Individual images are resized to 40*40 pixels	-	16, 32 and 64 size filters applied	-	Convolutional Neural Networks	Precision and recall	This model has a precision of 98.11% and recall of 94.97%	Used CNN to classify lung nodules
12	Evaluation on on Lung Cancer Detection Using Computer Assisted Diagnosis (CAD) System	Lung Image Database Consortium Image collection (LIDC-IDRI)	Median Filter, Histogram equalization	-	Thresholding, morphological and hybrid techniques	SVM, K-means	Convolutional Neural Network (CNN)	Specificity, Sensitivity and Accuracy	This research paper presents the techniques for detection of lung-cancer disease system through Computer-Aided Diagnosis(CAD) system,	This research paper presents the techniques for detection of lung-cancer disease system through Computer-Aided Diagnosis(CAD) system,

15	Lung Nodule Detection using Soft-Computing based Imaging Practice	Lung Image Database Consortium Image collection (LIDC-IDRI)	SGO assisted Kapur’s Thresholding	-	Active Counter and watershed	Binary mask-based technique	-	Sensitivity, specificity, accuracy, precision and F1score	The outcome of this tool helped to attain an average accuracy of >99% on the considered lung CT scan database.	The proposed work aims to develop a lung disease appraisal scheme to examine the nodule present in lungs
16	Lung Nodule Identification	Lung Image Database Consortium Image collection (LIDC-IDRI)	Median filtering	-	Thresholding, nodule segmentation	Area, perimeter, circularity, eccentricity. intensity, Location, No.of pixels in each direction, Diameter	SVM classifier	Accuracy, specificity, sensitivity	The accuracy of classification is 90.25%.	This project proposes an automated system for early detection of lung cancer using CT images
17	Morphological Filter Aided GMM Technique for Lung Nodule Detection	Lung Image Database Consortium Image collection (LIDC-IDRI)	Median filter	-	Thresholding, Gaussian Mixture Model	Morphology and intensity-based features	SVM classifier	Sensitivity, specificity, accuracy	Detected the lung nodules with an overall sensitivity, specificity and accuracy of 89.77%, 86.92% and 88.24% respectivel	An automated CAD framework has been developed for nodule detection

									y.	
18	Segmenting Nodules of Lung Tomography Image with Level Set Algorithm and Neural Network	Lung Image Database Consortium Image collection (LIDC-IDRI)	Median filter	-	Level set method (LSM)	Centroid, major and minor axis length and area of the nodule	Artificial Neural Network (ANN)	Accuracy and false positive rate (FPR)	The result shows that the proposed segmentation method can effectively segment the lung nodule.	This paper proposed segmentation method to segment the suspected nodule from the lung image
19	Study of Medical Image Processing techniques applied to Lung Cancer	Lung Image Database Consortium Image collection (LIDC-IDRI)	-	-	Click and Grow algorithm, Graph Cut method	Morphological, statistical features	Support Vector Machines (SVM), Convolution Neural network (CNN)	Accuracy	An average classification accuracy of 81% was achieved with an average AUC of 86.3%.	This paper presents a review of image processing techniques applied to the study of lung cancer(Spec. lung nodule segmentation, and feature extraction for tumor classification and prognosis

22	An adaptive morphology-based segmentation technique for lung nodule detection in thoracic CT image	Lung Image Database Consortium Image collection (LIDC-IDRI)	adaptive morphological filter	-	Adaptive morphology-based segmentation technique (AMST)	Morphological, texture and intensity-based	Support vector machine (SVM)	Sensitivity, specificity detection, accuracy	It has been observed that the proposed automated computer-aided detection system has achieved overall classification performance indices with 94.88% sensitivity, 93.45% specificity and 94.27% detection accuracy with 1.8 FPs/scan on LIDC/IDRI dataset and 91.43% sensitivity, 90.45% specificity, 92.83% accuracy with 3.2	Proposed CAD system presented in this paper outperforms the other state-of-the-art methods for automatic nodule detection from the HRCT image.

									FPs/scan on a private dataset	
24	LNCDS: A 2D-3D cascaded CNN approach for lung nodule classification, detection and segmentation	Lung Image Database Consortium Image collection (LIDC-IDRI) and purely independent Indian Lung CT Image Database (ILCID) clinical dataset.	Maximum intensity projection technique	-	SquExUNet segmentation model	Maximum intensity projection (MIP) technique to extract nodules	2D-3D cascaded CNN strategy	Sensitivity, Dice coefficient	We achieved a We achieved a Dice-Coefficient metrics of 0.80 for segmentation of nodule and 90.01% Sensitivity for nodule detection. metrics of 0.80 for segmentation of nodule and 90.01% Sensitivity for nodule detection.	Demonstrated an unique approach for classification, detection and segmentation of lung nodules on voluminous CT scan data of lung cancer patients
28	Segmentation of Lung Cancer from CT Image- A Compara	Lung Image Database Consortium Image collection (LIDC-IDRI)	Median Filtering	Median Filter	Adaptive image threshold using local first-order statistics, Select contiguous	Texture features	SVM	Accuracy, sensitivity	68 CT images were taken for segmentation and was observed that Mark-	Five different methods of segmentation and their comparisons

	tive Analysis				image region with similar grey values using flood-fill technique, Binary image segmentation using fast marching method, Segmentation based on grey scale intensity difference, Marker-Controlled Watershed Segmentation				controlled Watershed segmentation had highest accuracy of 92%	
30	Deep Learning Algorithm for Classification and Prediction of Lung Cancer using CT Scan Images	Lung Image Database Consortium Image collection (LIDC-IDRI)	Median Filtering	Kernel Filter	OTSU thresholding	Using CNN methods	RNN-LSTM		Accuracy is 97% which was the highest so far.	Aim was to develop advanced computer aided system using deep learning algorithm to classify and predict lung cancer
31	Coarse-to-Fine Lung Nodule	Lung Image Database Consortium Image	Histogram equalization, Contrast Enhancement	Median Filter	global-threshold binarization, lesion	Dual-Branch based on U-Net(2D and 3D)	CNN	Positive Prediction Value, Sensitivity	The Dice coefficients of nodule segmentati	Aim to develop a nodule segmentation

	Segmentation in CT Images with Image Enhancement and Dual-Branch Network	collection (LIDC-IDRI) SHCH			localization with region growing, Dual-Branch Network for Nodule segmentation				on on the LIDC dataset and our own dataset are 83.16% and 81.97% respectively,	combining image enhancement and DBNN
35	CAD system for lung nodules detection using wavelet-based approach and intelligent classifiers	Lung Image Database Consortium Image collection (LIDC-IDRI)	Morphological operators	Gaussian filter	Thresholding	3-level redundant wavelet transform	ANFIS,SVM, Bayesian neural networks	Very negative (VNE), Negative (NE), Zero (Z), Positive (PO) and very positive (VPO), confusion matrix	97.88% with Neuronal classifier, 97.27% with the SVM classifier, and 94.24% with the ANFIS classifier.	First research to establish a comparison between SVM, Bayesian neural networks and ANFIS classifiers.
36	Automatic Detection of Segmentation and Advanced Classification	Lung Image Database Consortium Image collection (LIDC-IDRI),In-house dataset	Anisotropic diffusion filtering (ADF) with the unsharp masking technique	Gaussian filter	Ostu,Fuzzy C-means clustering, Watershed, Proposed SSBIC	Shape and size	Advanced GWO with ONN and Advanced GWO with CNN.	Peak signal-to-noise ratio (PSNR), Mean Square Error (MSE) and Structural Similarity	Advanced GWO with CNN has auspiciously engendered 97.6% of accuracy.	Aim was use to try advanced classification algorithm and achieve maximum accuracy

	Algorithm							Index Method (SSIM).		
38	Analysis of CT Scan Images to Predict Lung Cancer Stages Using Image Processing Techniques	Lung Image Database Consortium Image collection (LIDC-IDRI)	Gabon Filter	Median Filter	OTSU thresholding, global thresholding	Gray Level Co-occurrence Matrix based Texture analysis: Convexity, Mean Intensity Value, Area Statistical Parametric approach:Mean,Standard Deviation, Higher Order Moments,	Support vector machine, K Nearest Neighbors, Random Forest, Naive Bayes	Accuracy, Confusion Matrix,Precision,Recall,F1 Score	Performance Comparison Table, highest accuracy 78.95% with 0.77 precision and 0.83 recall using Support Vector Machine (SVM) in the Statistical parametric approach of feature selection.	A good comparison was done between different ML classifiers which in future can give good results with more data
39	An Effective and Robust Cancer Detection in the Lungs with BPNN and	Lung Image Database Consortium Image collection (LIDC-IDRI)	Median Filter	Median Filter	Watershed Segmentation	Scale-Invariant Feature Transform(SIFT), Dense key point detector values	Backpropagation Neural Networks, Haar wavelet transform	Bag of Visual words histogram, ROC curve	Accuracy is 91%	Suitable for conducting lung cancer detection as a Computer-Aided Detection. (CAD) method.

	Watershed Segmentation									
42	Lung Nodule Detection and Segmentation Using a Patch-Based Multi-Atlas Method	Lung Image Database Consortium Image collection (LIDC-IDRI)	morphological operators	Median Filtering	Atlas based segmentation method	K-NN algorithm, Area, Convex area, Perimeter of the actual shape, Convex perimeter, Convex perimeter, Major axis length, Minor axis length, Circularity, Centroid,	SVM, Bayesian Networks	Normalize Euclidian distances, sensitivity	Sensitivity is 100%	First to use a patch based multi-atlas method in lung nodule selection.
43	Early-Stage Lung Cancer Prediction Using Various Machine Learning Techniques	Lung Image Database Consortium Image collection (LIDC-IDRI)	Histogram equalization, grey-scale	Median Filter, Max Pooling filtering and Convolution Filtering	Multiple Gray Level thresholding	Gray Level Co-occurrence method	Support Vector Machine, Random Forest, K-Nearest Neighbour, Artificial Neural Network, Voting Classifier	Accuracy, Confusion Matrix	Support Vector Machine 95 Random Forest 97.5 K-Nearest Neighbour 97 Artificial Neural Network 95.99 Voting Classifier	To detect ling cancer at early stages and save people in today's sophisticated world
46	Detection of Lung Cancer using	Lung Image Database Consortium Image collection	CLAHE Equalization technique	Gaussian Filter	Random Walk Segmentation	binarization and GLCM, correlation, entropy, variance, contrast, dissimilarity and	SVM, K-NN Algorithm	Level of prediction	Images	SVM to decide if the images are carcinogenic or not

	SVM Classifier and KNN Algorithm	(LIDC-IDRI)				energy				
49	Survey on Lung Cancer Detection Techniques	Lung Image Database Consortium Image collection (LIDC-IDRI)	CLAHE equalization method, Wiener Filter, Adaptive Neighborhood Contrast Enhancement (ANCE)	Garbon Filter, Median Filter, Laplacian Filter, spherical enhancement filters and contour filters.	EM (expectation-maximization), Feed-Forward Neural Network, Greyscale thresholding, thresholding Otsu, histogram analysis, optimal, iterative, clustering, multispectral and multi-thresholding, maximum correlation	SURF(Speeded Up Robust Features) algorithm descriptors	2D and 3D CNN, K-NN classifier	Accuracy, Level of prediction	Images	Huge dataset with large annotated sample requirement makes obstacle for applying deep learning in medical field.
50	Performance Evaluation of Contrast Enhancement Techniq	Lung Image Database Consortium Image collection (LIDC-IDRI)	Contrast Stretching, Histogram Equalization, Frequency Based Techniques, Spatial	Gaussian Filter	Otsu thresholding algorithm and region growing algorithm	ROI Extraction	Deep feature neural network and support vector machine	MSE, PSNR, RMSE, UIQI, SSIM, PCC	Images, comparison of different metrics for contrast enhancement methods	AHE is recommended for lung contrast enhancement.

	ues in Computed Tomography of Lung Images		Frequency Domain based Techniques							

2) NIH/NCI Dataset

Reference No	Paper Name	Image acquisition or datasets	Image Enhancement	Image Restoration	Image Segmentation	Features Extraction	Classifiers	Quality Metrics	Results	Remark
10	Effective And Reliable Lung Segmentation of Chest Images With Medical Image Processing And Machine Learning Approaches	NIH/NCI Datasets	Gabor filtering	-	Thresholding, Edge with morphological segmentations,	Otsu Thresholding	K-Means Clustering	Classification Accuracy, statistical analysis as PSNR, MSE	Precision, Recall values are providing better results as more than 90 percent accuracy rates for classifying the lung diseases	The paper is aimed at finding the accuracy of their proposed model to detect lung cancer.
27	Identifying Lung Cancer Using Image Processing Techniques	NIH/NCI Lung Image Database Consortium, DICOM format	K-nearest Neighbour, Weiner Filters, (local) thresholding, region growing, edge detection, and ridge detection, morphological	Gaussian Filtering, Anisotropic Filtering	Sobel Edge Detection Method, Region Growing algorithm	Area of the interest, Calcification, Shape and Size of nodule, Contrast Enhancement	Neural networks or Markov random field modelling	Accuracy, Sensitivity	Accuracy is 80%, 90% sensitivity with 0.05 false positives per image	This study showed different image processing techniques to detect lung cancer by surgeons and radiologists

			operations, fitting of geometrical models or functions and dynamic programming.							

3) Kaggle Dataset

Reference No	Paper Name	Image acquisition or datasets	Image Enhancement	Image Restoration	Image Segmentation	Features Extraction	Classifiers	Quality Metrics	Results	Remark
5	Segmentation And Prediction from Ct Images for Detecting Lung Cancer	Kaggle and LUNA (Lung Nodule Analysis)	CT scans and save them as dimensional array for neural network and saving the processed data as list	-	Morphological closing: Dilation followed by erosion	-	Convolutional Neural Networks	Accuracy	Were successfully able to train data set into UNET model segment and predict the presence of cancer cells in lungs	Worked on initial stages of training data set to identify cancer using initial image processing
14	Lung Cancer Detection from Computed Tomography (CT) Scans using Convolutional Neural Network	Kaggle Data Science Bowl 2017 competition data	Each voxel rescaled as a volume:1mm^3,each image are converted to Hounsfield units (HU)	-	Binarization, Clear Border, Labelling the Image, Erosion, Closure Operation, Filling Holes	CNN	3D CNN	Accuracy, precision and recall	The model achieves precision and recall of 89.24% and 82.17% respectively.	3D structure of nodules is used to extract features by 3D CNN to predict lung cancer

45	Automatic Detection of Lung Cancer Identification using ENNPSO Classification	Kaggle website: DICOM format	Kaun Filter	Median Filter	Active Contour method	Texture features, size, shape, and geometric features like mean, median, irregularity, concavity, convexity, area and perimeter has been extracted.	Elman Neural Network (ENN) , SVM , RBFN , ANFIS	Accuracy, Specificity, Recall, F1 Score, Precision, Performance Matrix Graph	Accuracy of ENM is 93%, Overall performance matrices for lung cancer detection method	Future research on this new algorithm can be increased with more images.

4) <u>Local Hospitals Datasets</u>

Reference No	Paper Name	Image acquisition or datasets	Image Enhancement	Image Restoration	Image Segmentation	Features Extraction	Classifiers	Quality Metrics	Results	Remark
6	Lung Cancer Detection Based On CT-Scan Images with Detection Features Using Gray Level Co_x0002_Occurrence Matrix (GLCM) and Support Vector Machine (SVM) Methods	CT-Scan images that have been taken from the hospital	Gray Scale and thresholding	-	Find contour method	Gray Level Co-occurrence Matrix (GLCM) features:	Support Vector Machine (SVM)	area (mm), contrast, energy, entropy, and homogeneity of extracted features and accuracy	From the system trial, the accuracy level based on the system decision in determining the diagnosis of lung cancer is benign or malignant was 83.33%.	This paper discusses the development of a CT-Scan based image-based lung cancer detection system
11	Efficient Computerized Lung Cancer	Lung images of 300 are collected from the Rajiv Gandhi	High Boost filtering	-	-	Scale Invariant Feature Transform (SIFT)	Clustered using K-means, grouped using Bag	Accuracy, specificity, sensitivity	Graph showing how better the proposed	Proposed system suits computerized lung cancer

	Detection Using Bag of Words	Cancer Institute and Research Centre, Delhi as a dataset out of which 100 images are used for testing and 200 images are used for training.					of Words, MLP Based BPNN		methodology gives better accuracy than other techniques	detection as this gives better accuracy compared to all existing systems
26	Research on the Auxiliary Classification and Diagnosis of Lung Cancer Subtypes Based on Histopathological Images	94 patients from a Cancer Hospital in Xinjiang from October 2011 to October 2020.	Adaptive histogram equalization algorithm (AHE)	Gaussian Filtering	Convolutional Neural Network	Huinvari-ant moments, GLGCM, wavelet transform, GLCM, LBP,GLDS and Markov random field, CNN	Relief-Support Vector Machines, Back Propagation (BP) Neural Network Classi-fiers and K-Nearest Neighbor (KNN), Decision tree, Naive Bayesian (NB) classifier, CNN	Receiver operating characteristic (ROC), Area under the curve (AUC)	Relief-SVM classification model achieves best classification performance regardless of whether it is compared with CNN models. Comparison of the results of the RELIEF-SVM model and classifier.	Study has a considerable guiding function for doctors with less experience, which can provide objective reference results and relieve the heavy work pressure of doctors.

29	Classification of Lung Cancer Stages from CT Scan Images Using Image Processing and k-Nearest Neighbour	Advanced Medical and Dental Institute, University Sains, Malaysia, DICOM format	Median Filtering	Median Filter	Manually using ImageJ and Adobe Photoshop CS6 Software, 99.70% thresholding	Area, Perimeter and Centroid	K-nearest Neighbours	Accuracy	KNN method has a high accuracy of 98.15% which proved that it has the potential to determine the stages of lung cancer	Establish an image processing technique for lung cancer using K-NN method
33	Computer Aided Lung Cancer Detection Based on Statistical Features	200 CT Scan images are collected from Yashoda hospitals, Hyderabad.	Garbon Filter	High-boost filter	Fuzzy C-means Algorithm	Statistical features	Support vector machine, back propagation neural network using multi layered perceptron and Probabilistic Neural Network	Activation function	BPNN has accuracy of 86%, SVM 85%, PNN 82%	BPNN can be used to detect the early stages of lung cancer and reduce deaths
41	Segmentation of Lung Lobes in Isotropic	60-100 CT images from local clinics	Wiener filter	Wiener filter	OTSU thresholding	Wavelet transform, adaptive Fissure Sweeping	SVM	Accuracy, RMS, worst case error	Accuracy of 77.1 – 93.6%	Automatic identification of left and right fissures and semi-

	CT Images Using Wavelet Transformation									automatic identification of horizontal fissures
47	Analysis of lung cancer detection based on the machine learning algorithm and iot	Clinical Centres	Median Filter	Median Filter	Gaussian Filter	Bilateral Filter	SVM (Support Vector Machine) Algorithm, Decision Tree Algorithm, Machine Learning based Reinforcement Learning Algorithms	Accuracy, Sensitivity	Accuracy of Machine Learning based is 92%	IOT along with ML is used for storage of output results and better quality of images on monitors
48	Classification and Segmentation Techniques for Detection of Lung Cancer from CT Images	Oncology centre in Mangalore, Karnataka	Contrast Limited Adaptive Histogram Equalization	Garbon Filter	Thresholding, Fuzzy k means, Otsu thresholding, Watershed ,Histogram+ Threshold+ Morphological operations,Self Organizing Map (SOM),Optimalthresholding, region	Texture features like contrast, correlation, energy, sum variance, maximum probability, dissimilarity were extracted from the image, Gray level co-occurrence matrix	SVM, BPN, ANN, Neuro Fuzzy ,CNN, Linear classifier, CNN with voting system, Bayesian classifier, Calculate	Accuracy Recall Precision	Accuracy 79.166% Recall 83.33% Precision 76.92% of proposed method Summary of the different approaches	Even though lot of segmentation methods and classifiers are used there are still issues that need to be solved

					growing, Gray level threshold+ binary lung filling, Background elimination		tumor area, feed forward BPN, Parsen, Naïve Bayes+ Genetic Algorithm			

5) Miscellaneous

Reference No	Paper Name	Image acquisition or datasets	Image Enhancement	Image Restoration	Image Segmentation	Features Extraction	Classifiers	Quality Metrics	Results	Remark
8	Detection of Lung Cancer from CT Scan Images using GLCM and SVM	IMBA Home (VIA-ELCAP Public Access)	Gabor Filter	-	Marker-Controlled Watershed Approach	Grey Level Co-Occurrence (GLCM) , Area = A = (Ai ,j , X ROI[Area] = i , Y ROI [Area] = j), Perimeter = P =(Pi ,j , X edge[P] = i, Y edge[P] = j),Eccentricity = Length of Major Axis/ Length of Minor Axis	Support Vector Machine (SVM)	Confusion matrix	categorizing the images into normal (non-cancerous) or abnormal (cancerous)	Aimed at detecting stage of cancer using GLCM and SVM
9	Early Stage Lung Cancer Diagnosis using ANN Classifier	Total 20 Cancer images and 20 Normal images	Median filter, Histogram equalization, Morphological operation	-	Identified seed point is selected manually; neighbourhood criterion includes growing of the region till the lung edge is identified.	Circle fit algorithm, Area, Mean Intensity, Diameter, Euler Number, Equivalent Circular Diameter	Artificial Neural Network (ANN)	Prediction in percentage as benign or malignant lung nodule. The threshold value decided is 50%, accuracy	29 desired nodules are detected in cancer images and 16 desired nodules are detected in normal images	Aimed at detecting cancer nodules using CAD system
13	Implementation of Image Enhance	E.g.: Camera, X-Beam, CT Sweep, etc the image	Gabor Filter	Auto Enhancement Algorithm,	Thresholding Approach	Staggered thresholding by Otsu's procedure	-	Mean PSNR	Gabor Filter produces the most	This paper compares Gabor Filter, DWT and

	ment and Image Segmentation In Disease Diagnosis (Lung Cancer)	that is acquired is common.		Discrete Wavelet Transform, PSNR					enhanced image.	Auto Enhancement techniques.
20	An effective approach for CT lung segmentation using mask region-based convolutional neural networks	Set of 39 lung computed tomography (CT) exams (DICOM format)	No pre-processing was applied	No pre-processing was applied	Mask R-CNN	-	Support Vectors Machine (SVM), K-means Classifiers	Position adjustment, size adjustment, Dice's coefficient, Sensitivity and specificity, Accuracy	The classification with the K-means and GMM methods was tested with 2–10 clusters and the best results were using only two clusters	Proposed a new robust method for automatic lung segmentation in Computed Tomography (CT) scans using Mask R-CNN with supervised and unsupervised machine learning models acting as kernels.
21	Automatic lung segmentation based on image decomposition	Eight types of lung images of ILDS	Image decomposition -based filtering	-	Thresholding, Multi-resolution analysis based on wavelet transform, Contour correction	Contour correction approach,	-	Over-segmentation rate (OR), under_x0002_segmentation rate (UR), Dice	Achieved an averaged Dice similarity coefficient of 98.04% and	Algorithm can correctly segment lung tissues from lung CT images and is helpful for radiologists'

	and wavelet transform							similarity coefficient (DSC), Jaccard's similarity index (JSI), Error (ERR) and modified Hausdorff distance (MHD)	Jaccard's similarity index of 94.91% on lung CT image segmentation compared with ground truths.	diagnosis of lung diseases
23	An efficient image segmentation and classification of lung lesions in pet and CT image fusion using DTWT incorporated SVM	Reference Image Database to Evaluate therapy Response (RIDER) lung PET-CT database	Decomposed by Dual Tree m-band Wavelet Transform (DTWT).	-	Clustering-based thresholding method	Morphological operations	Hybrid classifiers like Support Vector Machine (SVM)	Accuracy	The performance of the system has a higher classification accuracy of 99% using SVM classifier.	An efficient method for classification and segmentation of fused CT/ PET lung image is presented.
32	Deep Learning Assisted Predict of Lung	the datasets has been taken from http://diagnijmegen.nl/.	Image Histogram	Median Filter	K-mean algorithm, Markov decision process,	Area, 9 selected features, including SNPs, cytokines, PET, and miRNA, which has been deemed	Bayesian Network, deep learning neural	Area under curve, Mean accuracy analysis,	Accuracy is 96.67%	Aim to come up with deep learning methods to predict lung

	Cancer on Computed Tomography Images Using the Adaptive Hierarchical Heuristic Mathematical Model				Adaptive Hierarchical Heuristic Mathematical Model (AHHMM)	to reflect the state variables as important predictors.	network (DITNN).	determination of efficiency ratio, probability of survival rate, the loss function of AHHMM, Misclassification ratio		cancer
34	Comparison of Some Image Thresholding Techniques with Morphological Operations and Quantitative Evaluation for Lung Segmentation	Finding and Measuring Lungs in CT Data dataset	Histogram equalization	Median filtering, wiener filtering, smoothing filters	Fixed, Otsu, Yen Triangle Minimum, Mean Li Thresholding	False-positive values	3D tensor filtering	F1-score, Accuracy score, Precision score, Recall score and Jaccard score.	Fixed thresholding out performed others with a performance of 16 applied on four 25% scan volumes	Aim to compare different segmentation methods by different researchers

37	Automated Detection of Lung Cancer Using CT Scan Images	A lung cancer dataset	Contrast stretching	Median filtering	OTSU thresholding	Area, circularity and solidity	SVM	True positive images, true negative images, false positive images, false negative images and the accuracy of those datasets.	Accuracy is 95%	Automated detection using CT images' was developed
40	A pipeline for lung tumor detection and segmentation from ct scans using dilated convolutional neural networks	300 patients from the NSCLC-Radiomics dataset	Morphological operators	Median Filter	Hybrid-3D dilated fully convolutional neural network based on the LungNet architecture, proposed model	Training curve, feature maps	Binary classifiers	Average, median dice coefficient	Average-65.7%, Dice coefficient-70.39% in proposed model. 62.67% and 66.78% respectively in LungNet	CNN methods were compared with proposed model and proposed model outperformed them
44	A Comparative	Collecting pictures of CT scan of	Saliency Enhancement	Median Filter	Watershed Segmentation	SURF(Speeded Up Robust Features) algorithmic-relation	Support Vector Machine,	Accuracy Sensitivity Specificity	Accuracy 94.5% Sensitivity	SVM has shown best result when

	study of Lung Cancer detection using supervis ed neural network-	the person who is using				,contrast of the dataset provided, homogeneity, mean, energy, SD(Standard Deviation), Entropy of the dataset, Root Mean Square(RMS), Variance, smoothness of the images provided, skewness, kurtosis, IDM(Image Difference Measure).	Random Forest Algorithm	Recall	74.2% Specificity 77.6% Recall 66.3%	compared to Random Forest Algorithm, Accuracy can be increased if more images are provided

DATASET

For this study we have accessed the ELCAP Public Lung Image Database for lung CT scans. Our collected database contains a total of 556 lung images in DICOM format.

Link to dataset: **https://veet.via.cornell.edu/cgi-bin/datac/cformXP.cgi**

CONCLUSION

To treat lung cancer is patients, it is essential to correctly detect the cancerous lung nodules. So, we would be using the following steps for detection

1. For image enhancement and restoration, we will use image smoothening and edge enhancement technique (majority of papers used filtering techniques such as Median filter or Gabor filter).
2. For image segmentation, we will use segmenting techniques such as thresholding.
3. For classification we will use different CNN models for better results (many papers used SVM and pre-trained CNN models).

<u>REFERENCES</u>

[1] M. Islam, A. H. Mahamud and R. Rab, "Analysis of CT Scan Images to Predict Lung Cancer Stages Using Image Processing Techniques," *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 2019, pp. 0961-0967, doi: 10.1109/IEMCON.2019.8936175.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/8936175

[2] N. S. Nadkarni and S. Borkar, "Detection of Lung Cancer in CT Images using Image Processing," *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 2019, pp. 863-866, doi: 10.1109/ICOEI.2019.8862577

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/8862577

[3] W. Abdul, "An Automatic Lung Cancer Detection and Classification (ALCDC) System Using Convolutional Neural Network," 2020 13th International Conference on Developments in eSystems Engineering (DeSE), 2020, pp. 443-446, doi: 10.1109/DeSE51703.2020.9450778.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/9450778/

[4] D. Jayaraj and S. Sathiamoorthy, "Random Forest based Classification Model for Lung Cancer Prediction on Computer Tomography Images," *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2019, pp. 100-104, doi: 10.1109/ICSSIT46314.2019.8987772.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/9225486

[5] K. S. Chethan, S. Vishwanath, R. V. Patil and K. A. Vijetha, "Segementation and Prediction from CT Images for Detecting Lung Cancer," *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2020, pp. 1-6, doi: 10.1109/ICCCNT49239.2020.9225486.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/9225486

[6] Q. Firdaus, R. Sigit, T. Harsono and A. Anwar, "Lung Cancer Detection Based On CT-Scan Images With Detection Features Using Gray Level Co-Occurrence Matrix (GLCM) and Support Vector Machine (SVM) Methods," *2020 International Electronics Symposium (IES)*, 2020, pp. 643-648, doi: 10.1109/IES50839.2020.9231663.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/9231663

[7] S. Bhat, R. Shashikala, S. Kumar and K. Gururaj, "Convolutional Neural Network approach for the Classification and Recognition of Lung Nodules," *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2020, pp. 1310-1314, doi: 10.1109/ICECA49313.2020.9297626.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/9297626

[8] M. H. Jony, F. Tuj Johora, P. Khatun and H. K. Rana, "Detection of Lung Cancer from CT Scan Images using GLCM and SVM," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), 2019, pp. 1-6, doi: 10.1109/ICASERT.2019.8934454.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/8934454

[9] S. S. Bhise and S. R. Khot, "Early Stage Lung Cancer Diagnosis using ANN Classifier," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021, pp. 22-27, doi: 10.1109/ICAIS50930.2021.9395952.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/9395952

[10] P. B. Chanda and S. K. Sarkar, "Effective And Reliable Lung Segmentation Of Chest Images With Medical Image Processing And Machine Learning Approaches," 2020 IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI), 2020, pp. 1-6, doi: 10.1109/ICATMRI51801.2020.9398450.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/9398450

[11] Krishna, P. C. S. Rao and C. Zeelan Basha, "Efficient Computerized Lung Cancer Detection Using Bag of Words," 2020 7th International Conference on Smart Structures and Systems (ICSSS), 2020, pp. 1-4, doi: 10.1109/ICSSS49621.2020.9202039.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/9202039

[12] V. J. Pawar and P. Premchnd, "Evaluation on Lung Cancer Detection Using Computer Assisted Diagnosis (CAD) System," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 2021, pp. 812-817, doi: 10.1109/ICACCS51430.2021.9441852.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/9441852

[13] P. A. Aisiri, B. Ramya, D. J. Chaithanya and S. Anitha, "Implementation of Image Enhancement and Image Segmentation In Disease Diagnosis (Lung Cancer)," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 2021, pp. 151-156, doi: 10.1109/ICACCS51430.2021.9441930.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/9441930

[14] M. B. Khumancha, A. Barai and C. B. R. Rao, "Lung Cancer Detection from Computed Tomography (CT) Scans using Convolutional Neural Network," 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2019, pp. 1-7, doi: 10.1109/ICCCNT45670.2019.8944824.

https://ieeexplore.ieee.org/document/8944824

[15] B. Nirupriya, P. Atilakshmy, G. Jayashree, P. Deepa and S. Arunmozhi, "Lung Nodule Detection using Soft-Computing based Imaging Practice," 2020 International Conference on System, Computation, Automation and Networking (ICSCAN), 2020, pp. 1-4, doi: 10.1109/ICSCAN49426.2020.9262323.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/8944824

[16] K. V. Hinisha and A. Lijiya, "Lung Nodule Identification," *2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, 2019, pp. 451-455, doi: 10.1109/ICICICT46008.2019.8993218.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/8993218

[17] Halder, S. Chatterjee and D. Dey, "Morphological Filter Aided GMM Technique for Lung Nodule Detection," 2020 IEEE Applied Signal Processing Conference (ASPCON), 2020, pp. 198-202, doi: 10.1109/ASPCON49795.2020.9276715.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/9276715

[18] S. Y. Chong, M. K. Tan, K. B. Yeo, M. Y. Ibrahim, X. Hao and K. T. K. Teo, "Segmenting Nodules of Lung Tomography Image with Level Set Algorithm and Neural Network," 2019 IEEE 7th Conference on Systems, Process and Control (ICSPC), 2019, pp. 161-166, doi: 10.1109/ICSPC47137.2019.9067987.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/9067987

[19] S. Moreno, M. Bonfante, E. Zurek and H. S. Juan, "Study of Medical Image Processing Techniques Applied to Lung Cancer," 2019 14th Iberian Conference on Information Systems and Technologies (CISTI), 2019, pp. 1-6, doi: 10.23919/CISTI.2019.8760888.

http://ieeexplore.ieee.org.egateway.vit.ac.in/document/8760888

[20] Qinhua Hu, Luís Fabrício de F. Souza, Gabriel Bandeira Holanda, Shara S.A. Alves, Francisco Hércules dos S. Silva, Tao Han, Pedro P. Rebouças Filho,An effective approach for CT lung segmentation using mask region-based convolutional neural networks,Artificial Intelligence in Medicine,Volume 103,2020,101792,ISSN 0933-3657.

http://www.sciencedirect.com.egateway.vit.ac.in/science/article/pii/S1746809420301889

[21] Caixia Liu, Mingyong Pang,Automatic lung segmentation based on image decomposition and wavelet transform, Biomedical Signal Processing and Control,Volume 61,2020,102032,ISSN 17468094.

http://www.sciencedirect.com.egateway.vit.ac.in/science/article/pii/S1746809420301889

[22] Amitava Halder, Saptarshi Chatterjee, Debangshu Dey, Surajit Kole, Sugata Munshi,An adaptive morphology based segmentation technique for lung nodule detection in thoracic CT image,Computer Methods and Programs in Biomedicine, Volume 197,2020,105720,ISSN 01692607.

http://www.sciencedirect.com.egateway.vit.ac.in/science/article/pii/S0169260720315534

[23] R. Mohana Priya, P. Venkatesan,An efficient image segmentation and classification of lung lesions in pet and CT image fusion using DTWT incorporated SVM,Microprocessors and Microsystems, Volume 82,2021,103958,ISSN 0141-9331.

http://www.sciencedirect.com.egateway.vit.ac.in/science/article/pii/S014193312100137X

[24] Prasad Dutande, Ujjwal Baid, Sanjay Talbar,LNCDS: A 2D-3D cascaded CNN approach for lung nodule classification, detection and segmentation,Biomedical Signal Processing and Control, Volume 67,2021,102527,ISSN 1746-8094.

http://www.sciencedirect.com.egateway.vit.ac.in/science/article/pii/S1746809421001245

[25] Yuxin Zhou, Yinan Lu, Zhili Pei,Accurate diagnosis of early lung cancer based on the convolutional neural network model of the embedded medical system,Microprocessors and Microsystems, Volume 81,2021,103754,ISSN 0141-9331

http://www.sciencedirect.com.egateway.vit.ac.in/science/article/pii/S0141933120308991

[26] M. Li et al., "Research on the Auxiliary Classification and Diagnosis of Lung Cancer Subtypes Based on Histopathological Images," in IEEE Access, vol. 9, pp. 53687-53707, 2021, doi: 10.1109/ACCESS.2021.3071057.
http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9395071

[27] Sharma, D., & Jindal, G. (2011). Identifying Lung Cancer Using Image Processing Techniques.
https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1059.8042&rep=rep1&type=pdf#:~:text=Initially%2C%20the%20basic%20image%20processing,from%20the%20extracted%20lung%20image

[28] R. Indumathi, R. Vasuki,"Segmentation of Lung Cancer from CT Image- A Comparative Analysis",Materials Today: Proceedings,2021,
http://www.sciencedirect.com.egateway.vit.ac.in/science/article/pii/S2214785321026390

[29] M. Firdaus Abdullah, S. Noraini Sulaiman, M. Khusairi Osman, N. K. A. Karim, I. Lutfi Shuaib and M. Danial Irfan Alhamdu, "Classification of Lung Cancer Stages from CT Scan Images Using Image Processing and k-Nearest Neighbours," 2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC), 2020, pp. 68-72, doi: 10.1109/ICSGRC49013.2020.9232492.
http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9232492

[30] D. Mhaske, K. Rajeswari and R. Tekade, "Deep Learning Algorithm for Classification and Prediction of Lung Cancer using CT Scan Images," *2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBEA)*, 2019, pp. 1-5, doi: 10.1109/ICCUBEA47591.2019.9128479.
http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9128479

[31] Z. Wu, Q. Zhou and F. Wang, "Coarse-to-Fine Lung Nodule Segmentation in CT Images with Image Enhancement and Dual-Branch Network," in *IEEE Access*, vol. 9, pp. 7255-7262, 2021, doi: 10.1109/ACCESS.2021.3049379.
http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9314141

[32] H. Yu, Z. Zhou and Q. Wang, "Deep Learning Assisted Predict of Lung Cancer on Computed Tomography Images Using the Adaptive Hierarchical Heuristic Mathematical Model," in IEEE Access, vol. 8, pp. 86400-86410, 2020, doi: 10.1109/ACCESS.2020.2992645.
http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9086786

[33] G. Parasa, M. B. Anusha, Y. Tondepu, C. Z. Basha and G. Kalyan Chakravarthy, "Computer Aided Lung Cancer Detection Based on Statistical Features," *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2020, pp. 182-188, doi: 10.1109/ICIRCA48905.2020.9183051.
http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9183051

[34] R. J. Suji and S. S. Bhadauria, "Comparision of Some Image Thresholding Techniques with Morphological Operations and Quantitative Evaluation for Lung Segmentation," 2021 2nd International Conference for Emerging Technology (INCET), 2021, pp. 1-4, doi: 10.1109/INCET51464.2021.9456110.
http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9456110

[35] H. Mahersia, H. Boulehmi and K. Hamrouni, "CAD system for lung nodules detection using wavelet-based approach and intelligent classifiers," 2020 17th International Multi-Conference on Systems, Signals & Devices (SSD), 2020, pp. 173-178, doi: 10.1109/SSD49366.2020.9364203.
http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9364203

[36] A.B. Mathews and M. K. Jeyakumar, "Automatic Detection of Segmentation and Advanced Classification Algorithm," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020, pp. 358-362, doi: 10.1109/ICCMC48092.2020.ICCMC-00067.
http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9076423

[37] Hoque, A. K. M. A. Farabi, F. Ahmed and M. Z. Islam, "Automated Detection of Lung Cancer Using CT Scan Images," 2020 IEEE Region 10 Symposium (TENSYMP), 2020, pp. 1030-1033, doi: 10.1109/TENSYMP50017.2020.9230861. http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9230861

[38] M. Islam, A. H. Mahamud and R. Rab, "Analysis of CT Scan Images to Predict Lung Cancer Stages Using Image Processing Techniques," 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2019, pp. 0961-0967, doi: 10.1109/IEMCON.2019.8936175. http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=8936175

[39] C. Z. Basha, B. Lakshmi Pravallika, D. Vineela and S. L. Prathyusha, "An Effective and Robust Cancer Detection in the Lungs with BPNN and Watershed Segmentation," 2020 International Conference for Emerging Technology (INCET), 2020, pp. 1-6, doi: 10.1109/INCET49848.2020.9154186. http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9154186

[40] S. Hossain, S. Najeeb, A. Shahriyar, Z. R. Abdullah and M. Ariful Haque, "A Pipeline for Lung Tumor Detection and Segmentation from CT Scans Using Dilated Convolutional Neural Networks," ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 1348-1352, doi: 10.1109/ICASSP.2019.8683802. http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=8683802

[41] Q. Wei, Y. Hu, G. Gelfand and J. H. MacGregor, "Segmentation of Lung Lobes in Isotropic CT Images Using Wavelet Transformation," 2019, 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2010, pp. 5551-5554, doi: 10.1109/IEMBS.2019.4353604. http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=7881305

[42] M. Alam, G. Sankaranarayanan and V. Devarajan, "Lung Nodule Detection and Segmentation Using a Patch-Based Multi-Atlas Method," 2016 International Conference on Computational Science and Computational Intelligence (CSCI), 2016, pp. 23-28, doi: 10.1109/CSCI.2016.0012. http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=7881305

[43] C. Thallam, A. Peruboyina, S. S. T. Raju and N. Sampath, "Early Stage Lung Cancer Prediction Using Various Machine Learning Techniques," 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp. 1285-1292, doi: 10.1109/ICECA49313.2020.9297576. http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9297576

[44] K. Roy et al., "A Comparative study of Lung Cancer detection using supervised neural network," 2019 International Conference on Opto-Electronics and Applied Optics (Optronix), 2019, pp. 1-5, doi: 10.1109/OPTRONIX.2019.8862326. http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=8862326

[45] B. Hemalatha, S. Yuvaraj, K. V. Kiruthikaa and V. Viswanathan, "Automatic Detection of Lung Cancer Identification using ENNPSO Classification," 2019 International Conference on Advances in Computing and Communication Engineering (ICACCE), 2019, pp. 1-6, doi: 10.1109/ICACCE46606.2019.9079963. http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9079963

[46] R. Sathishkumar, K. Kalaiarasan, A. Prabhakaran and M. Aravind, "Detection of Lung Cancer using SVM Classifier and KNN Algorithm," 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), 2019, pp. 1-7, doi: 10.1109/ICSCAN.2019.8878774. http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=8878774

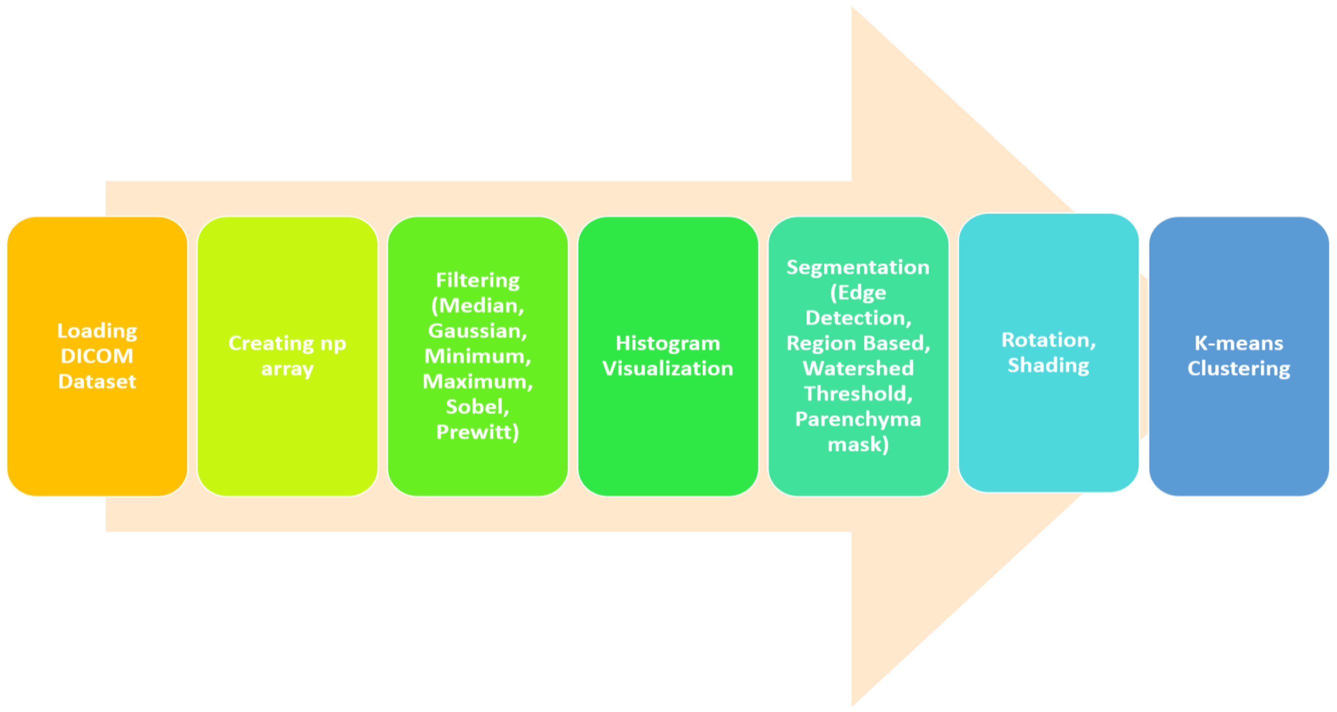
[47] K. Karthick, S. Rajkumar, N. Selvanathan, U. K. B. Saravanan, M. Murali and B. Dhiyanesh, "Analysis of Lung Cancer Detection Based on the Machine Learning Algorithm and IOT," 2021 6th International Conference on Communication and Electronics Systems (ICCES), 2021, pp. 1-8, doi: 10.1109/ICCES51350.2021.9489084.
http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9489084

[48] P. Lobo and S. Guruprasad, "Classification and Segmentation Techniques for Detection of Lung Cancer from CT Images," 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), 2018, pp. 1014-1019, doi: 10.1109/ICIRCA.2018.8597273.
http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=8597273

[49] V. Prakash and P. Smitha Vas, "Survey on Lung Cancer Detection Techniques," 2020 International Conference on Computational Performance Evaluation (ComPE), 2020, pp. 800-803, doi: 10.1109/ComPE49325.2020.9200019.
http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9200019

[50] S. Ziyad, V. Radha and V. Thavavel, "Performance Evaluation of Contrast Enhancement Techniques in Computed Tomography of Lung Images," 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), 2019, pp. 1-4, doi: 10.1109/I2CT45611.2019.9033602.
http://ieeexplore.ieee.org.egateway.vit.ac.in/stamp/stamp.jsp?tp=&arnumber=9033602

Architecture Model:



Hardware and software requirements:

- CPU: 2 x 64-bit 2.8 GHz 8.00 GT/s CPUs
- RAM: 8 GB
- Storage: 300 GB. (600 GB for air-gapped deployments.)
- Operating system – Windows
- Anaconda-Spyder

IMPLEMENTATION

Loading necessary packages:

```
import numpy as np
import os
import copy
from math import *
import matplotlib.pyplot as plt
from functools import reduce
import pydicom
from skimage import measure, morphology
from skimage.morphology import ball, binary_closing
from skimage.measure import label, regionprops
from scipy.linalg import norm
import scipy.ndimage
from ipywidgets.widgets import *
import ipywidgets as widgets
import plotly
from plotly.graph_objs import *
import chart_studio.plotly as py
from scipy import ndimage, misc
import cv2
from skimage import data
from skimage.exposure import histogram
from skimage import feature
from skimage import data
```

```
from skimage.exposure import histogram
from skimage.feature import canny
from skimage.filters import sobel
from skimage.segmentation import watershed
from scipy import ndimage as ndi
```

Loading DICOM Data:

```
def load_scan(path):
    slices = [pydicom.dcmread(path+'/' + s) for s in
              os.listdir(path)]
    slices = [s for s in slices if 'SliceLocation' in s]
    slices.sort(key = lambda x: int(x.InstanceNumber))
    try:
      slice_thickness = np.abs(slices[0].ImagePositionPatient[2]-slices[1].ImagePositionPatient[2])
    except:
      slice_thickness = np.abs(slices[0].SliceLocation-slices[1].SliceLocation)
    for s in slices:
        s.SliceThickness = slice_thickness
    return slices
def get_pixels_hu(scans):
    image = np.stack([s.pixel_array for s in scans])
    image = image.astype(np.int16)
    image[image == -2000] = 0
     # Convert to Hounsfield units (HU)
    intercept = scans[0].RescaleIntercept
    slope = scans[0].RescaleSlope
```

```
    if slope != 1:
        image = slope * image.astype(np.float64)
        image = image.astype(np.int16)
    image += np.int16(intercept)
    return np.array(image, dtype=np.int16)
path = r"C:\Users\user\OneDrive\Desktop\data"
patient_dicom = load_scan(path)
patient_pixels = get_pixels_hu(patient_dicom)
#sanity check
plt.imshow(patient_pixels[53], cmap=plt.cm.bone)
```

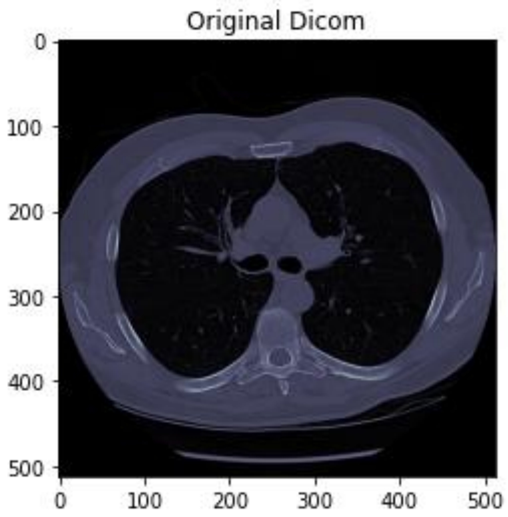
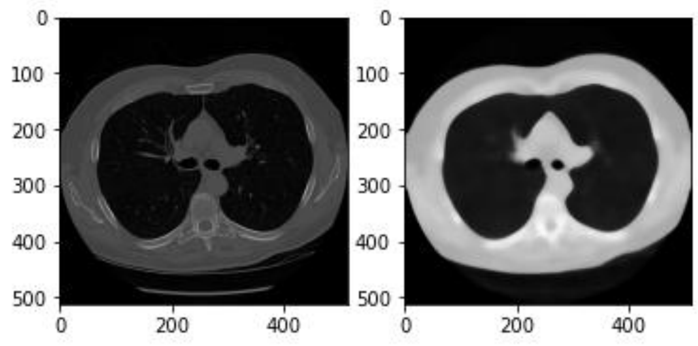


Image Enhancement using filtration:

What Is Median Filtering?

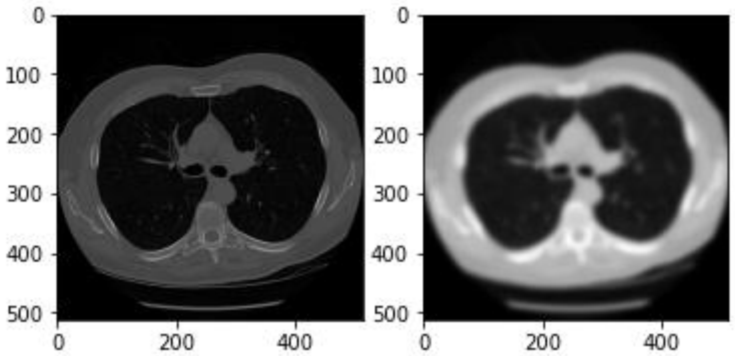
Image noise can be briefly defined as random variations in some of the pixel values of an image. Median filtering is excellent at reducing this type of noise. The filtering algorithm will scan the entire image, using a small matrix (like the 3x3 depicted above), and recalculate the value of the center pixel by simply taking the median of all of the values inside the matrix.

```
#MEDIAN
fig = plt.figure()
plt.gray()
ax1 = fig.add_subplot(121)  # left side
ax2 = fig.add_subplot(122)  # right side
ascent = patient_pixels[53]
result = ndimage.median_filter(ascent, size=20)
ax1.imshow(ascent)
ax2.imshow(result)
plt.show()
```



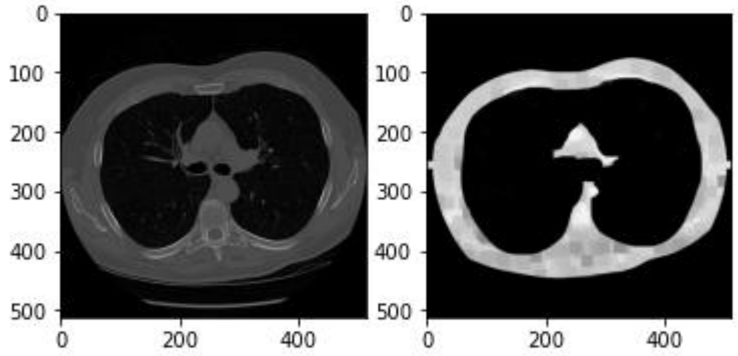
Gaussian filtering g is used to blur images and remove noise and detail.

```
#GAUSSIAN
fig = plt.figure()
plt.gray()
ax1 = fig.add_subplot(121)  # left side
ax2 = fig.add_subplot(122)  # right side
ascent = patient_pixels[53]
result = ndimage.gaussian_filter(ascent,sigma=5)
ax1.imshow(ascent)
ax2.imshow(result)
plt.show()
```

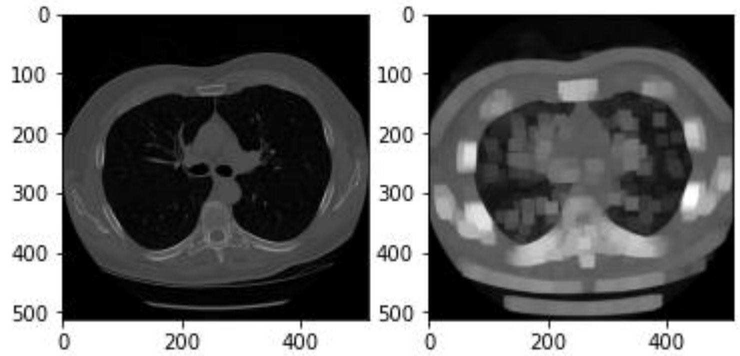



```
"""Minimum"""
fig = plt.figure()
plt.gray()
ax1 = fig.add_subplot(121)  # left side
ax2 = fig.add_subplot(122)  # right side
ascent = patient_pixels[53]
```

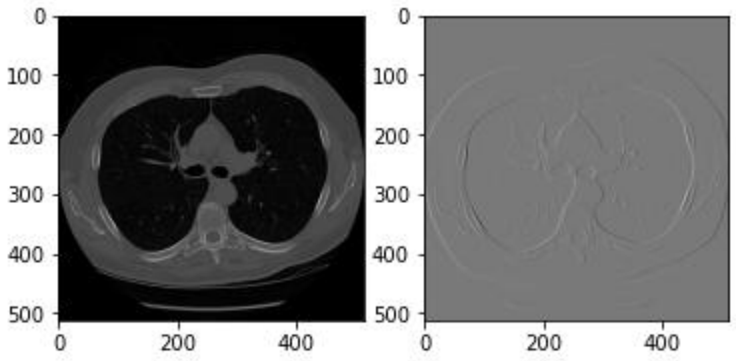
```
result = ndimage.minimum_filter(ascent,size=20)
ax1.imshow(ascent)
ax2.imshow(result)
plt.show()
```




```
"""MAXIMUM"""
fig = plt.figure()
plt.gray()
ax1 = fig.add_subplot(121)  # left side
ax2 = fig.add_subplot(122)  # right side
ascent = patient_pixels[53]
result = ndimage.maximum_filter(ascent,size=20)
ax1.imshow(ascent)
ax2.imshow(result)
plt.show()
```

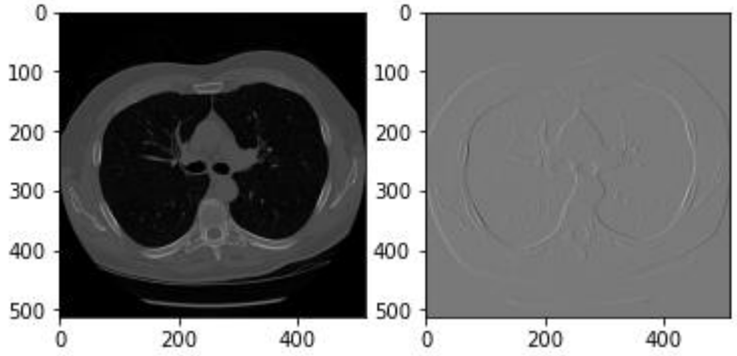


```
"""SOBEL"""
fig = plt.figure()
plt.gray()
ax1 = fig.add_subplot(121)  # left side
ax2 = fig.add_subplot(122)  # right side
ascent = patient_pixels[53]
result = ndimage.sobel(ascent)
ax1.imshow(ascent)
ax2.imshow(result)
plt.show()
```



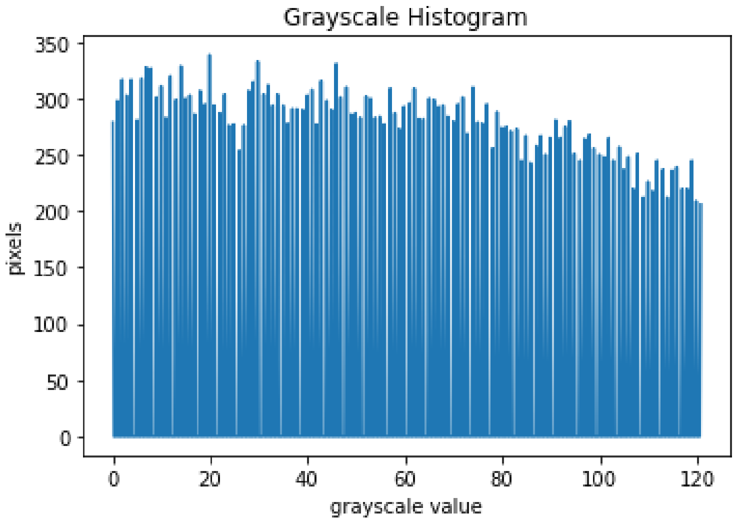

```
"""PREWITT"""
fig = plt.figure()
plt.gray()
ax1 = fig.add_subplot(121)  # left side
ax2 = fig.add_subplot(122)  # right side
ascent = patient_pixels[53]
result = ndimage.prewitt(ascent)
ax1.imshow(ascent)
```

```
ax2.imshow(result)
plt.show()
```

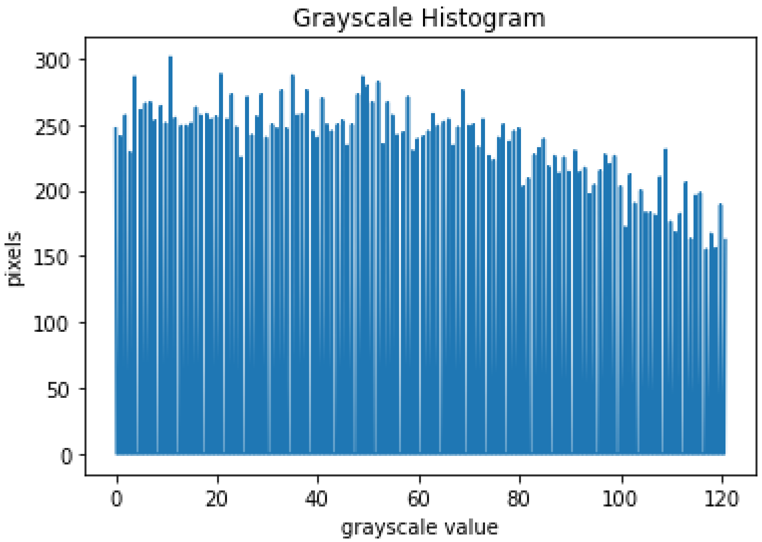


<u>Histogram Visualization:</u>

```
histogram, bin_edges = np.histogram(patient_pixels[53], bins=512, range=(0,121))
plt.figure()
plt.title("Grayscale Histogram")
plt.xlabel("grayscale value")
plt.ylabel("pixels")
plt.plot(bin_edges[0:-1], histogram)  # <- or here
plt.show()
```

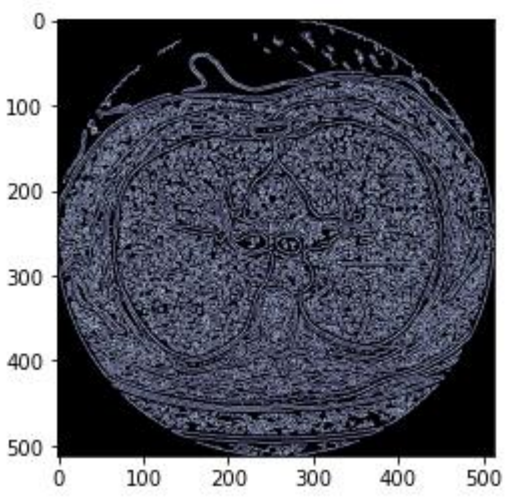



```
histogram, bin_edges = np.histogram(patient_pixels[18], bins=512, range=(0,121))
plt.figure()
plt.title("Grayscale Histogram")
plt.xlabel("grayscale value")
plt.ylabel("pixels")
plt.plot(bin_edges[0:-1], histogram)  # <- or here
plt.show()
```



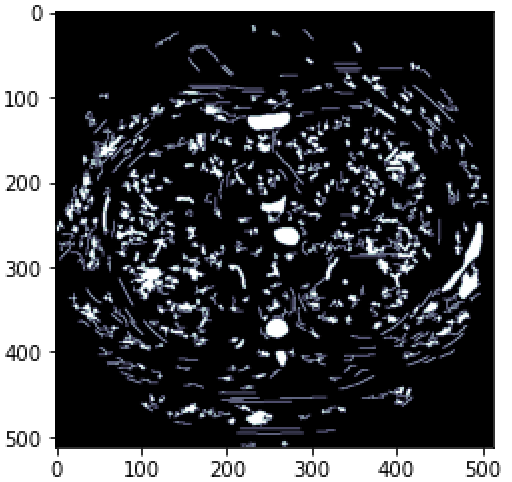
Segmentation

```
"""EDGE BASED SEGMENTATION"""
lung= patient_pixels[53]
hist, hist_centers = histogram(lung)
edges = canny(lung/512.)
plt.imshow(edges,cmap=plt.cm.bone)
```

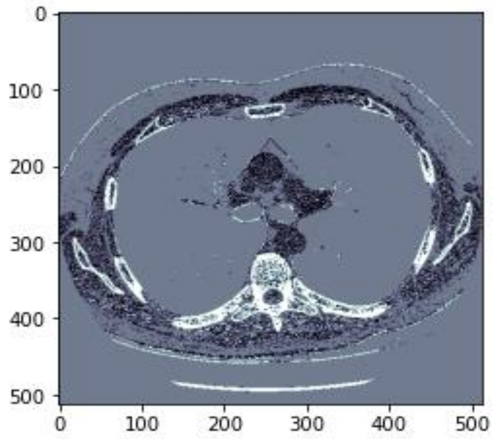



```
fill_lung = ndi.binary_fill_holes(edges)
label_objects, nb_labels = ndi.label(fill_lung)
sizes = np.bincount(label_objects.ravel())
mask_sizes = sizes > 20
mask_sizes[0] = 0
lung_cleaned = mask_sizes[label_objects]
```

```
plt.imshow(lung_cleaned,cmap=plt.cm.bone)
```

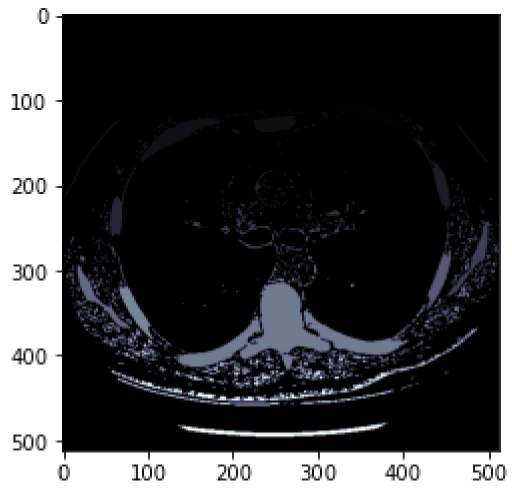


```
"""Region-based segmentation"""
markers = np.zeros_like(lung)
markers[lung < 30] = 1
markers[lung> 150] = 2
elevation_map = sobel(lung)
plt.imshow(elevation_map,cmap=plt.cm.bone)
markers = np.zeros_like(lung)
markers[lung < 30] = 1
markers[lung > 150] = 2
plt.imshow(markers,cmap=plt.cm.bone)
```



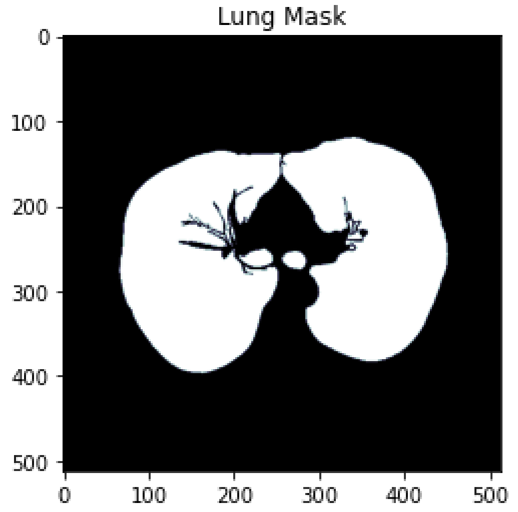
"""Watershed transform based"""

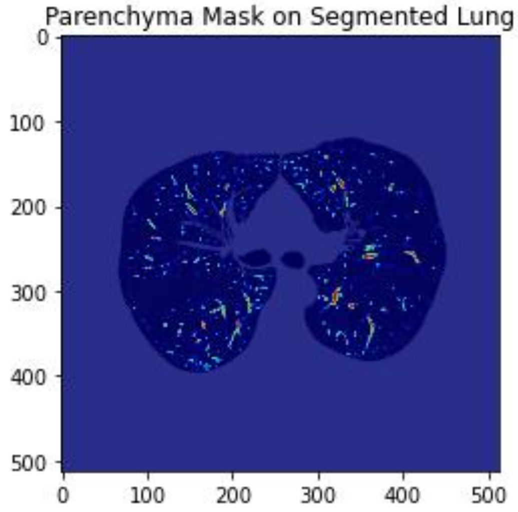
```
segmentation = watershed(elevation_map, markers)
plt.imshow(segmentation,cmap=plt.cm.bone)

segmentation = ndi.binary_fill_holes(segmentation - 1)
labeled_lung, _ = ndi.label(segmentation)
plt.imshow(labeled_lung,cmap=plt.cm.bone)
```



```
“””Lung Mask Segmentation”””
def largest_label_volume(im, bg=-1):
    vals, counts = np.unique(im, return_counts=True)
    counts = counts[vals != bg]
    vals = vals[vals != bg]
    if len(counts) > 0:
        return vals[np.argmax(counts)]
    else:
        return None
def segment_lung_mask(image, fill_lung_structures=True):
    binary_image = np.array(image >= -700, dtype=np.int8)+1
    labels = measure.label(binary_image)
    background_label = labels[0,0,0]


    binary_image[background_label == labels] = 2
        for i, axial_slice in enumerate(binary_image):
            axial_slice = axial_slice-1
            labeling = measure.label(axial_slice)
            l_max = largest_label_volume(labeling, bg=0)

            if l_max is not None:
                binary_image[i][labeling != l_max] = 1
    binary_image -= 1 #Make the image actual binary
    binary_image = 1-binary_image # Invert it, lungs are now 1

    # Remove other air pockets inside body
```

```
    labels = measure.label(binary_image, background=0)



 l_max = largest_label_volume(labels, bg=0)
    if l_max is not None: # There are air pockets
        binary_image[labels != l_max] = 0

    return binary_image
# get masks
segmented_lungs = segment_lung_mask(patient_pixels,
                fill_lung_structures=False)
segmented_lungs_fill = segment_lung_mask(patient_pixels,
                    fill_lung_structures=True)
internal_structures = segmented_lungs_fill - segmented_lungs
# isolate lung from chest
copied_pixels = copy.deepcopy(patient_pixels)
for i, mask in enumerate(segmented_lungs_fill):
    get_high_vals = mask == 0
    copied_pixels[i][get_high_vals] = 0
seg_lung_pixels = copied_pixels
# sanity check
plt.imshow(seg_lung_pixels[7], cmap=plt.cm.bone)
```

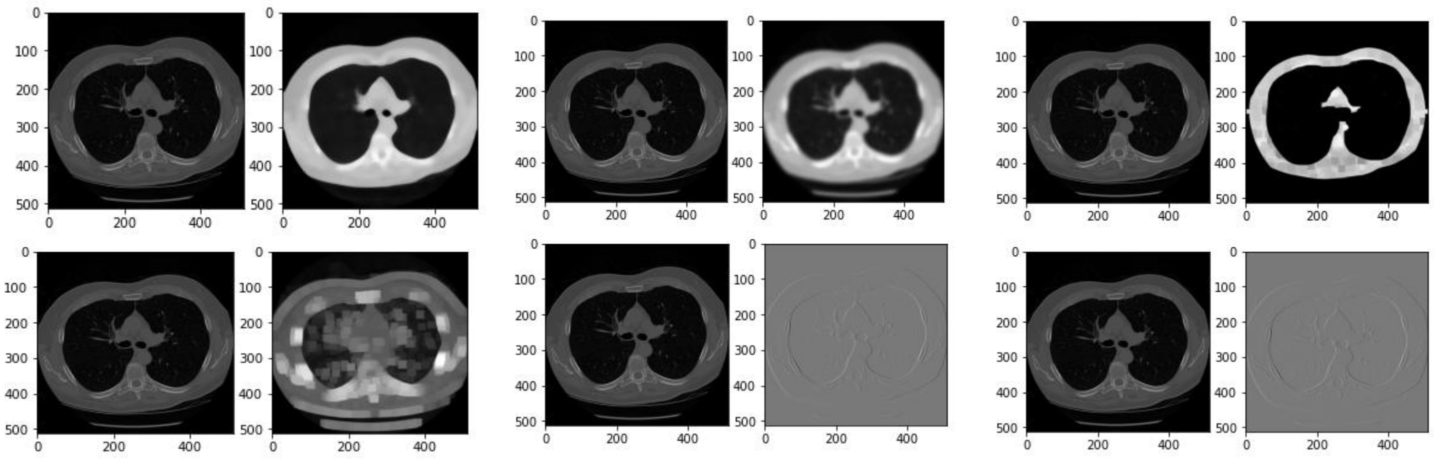




Comparison of results using various methods

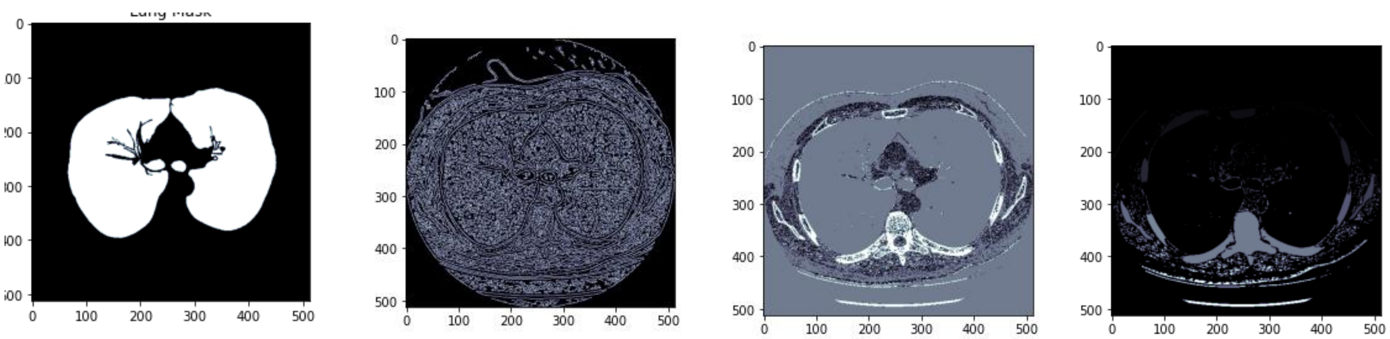
Filtration Results

Amongst the filters used i.e. Median, Gaussian, Minimum, Maximum, Sobel, Prewitt, Gaussian filter gave the best result.



Segmentation results

Segmentation based on Hounsfield Units (HU Units) gave the best result.



Rotation:

```
img = cv2.imread(path)
(h,w) = img.shape[:2]
center = (w/2, h/2)
angle45 = 45
angle90 = 90
scale = 1.0

M = cv2.getRotationMatrix2D(center, angle45, scale)
M2 = cv2.getRotationMatrix2D(center, angle90, scale)
```

```
abs_cos = abs(M[0,0])
abs_sin = abs(M[0,1])

abs_cos = abs(M2[0,0])
abs_sin = abs(M2[0,1])

bound_w = int(h * abs_sin + w * abs_cos)
bound_h = int(h * abs_cos + w * abs_sin)

M[0, 2] += bound_w/2 - center[0]
M[1, 2] += bound_h/2 - center[1]

M2[0, 2] += bound_w/2 - center[0]
M2[1, 2] += bound_h/2 - center[1]

rotated45 = cv2.warpAffine(img, M, (bound_w,bound_h))

for i in range(7):
	rotated45 = cv2.warpAffine(rotated45, M, (bound_w,bound_h))

rotated90 = cv2.warpAffine(img, M2, (bound_w,bound_h))

for i in range(3):
	rotated90 = cv2.warpAffine(rotated90, M2, (bound_w,bound_h))

cv2.imshow('Rotated by 45 8 times', rotated45)
cv2.imshow('Rotated by 90 4 times', rotated90)
cv2.imwrite('rotated45.jpg',rotated45)
```

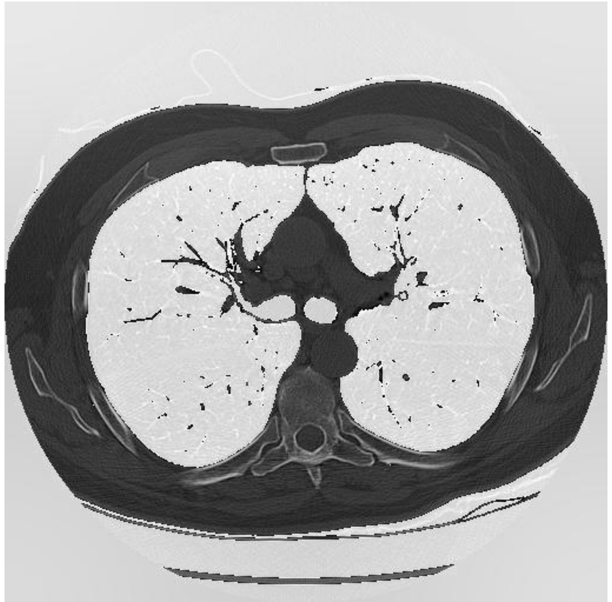
```
cv2.imwrite('rotated90.jpg',rotated90)
cv2.waitKey(0)
cv2.destroyAllWindows()
```

Rotated by 45 degree

Rotated by 90 degree

Shading correction:

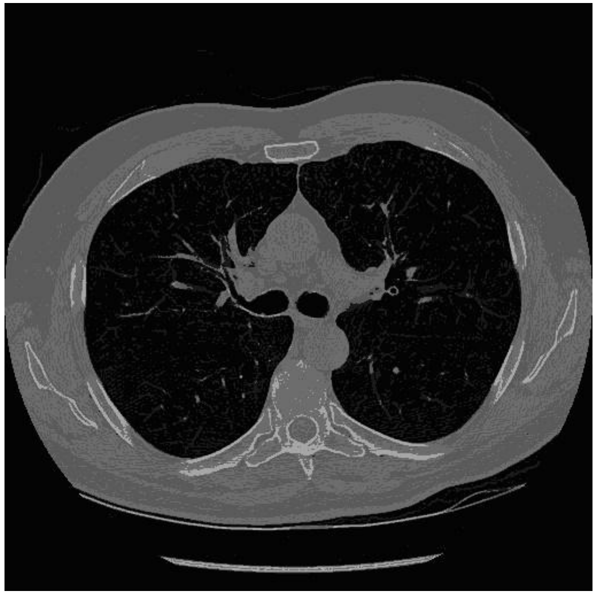
```
img = cv2.imread(path)
grayImg = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
cv2.imshow('Gray Image', grayImg)
cv2.waitKey(0)
filtersize = 513
gaussianImg = cv2.GaussianBlur(grayImg, (filtersize, filtersize), 128)
cv2.imshow('Converted Image', gaussianImg)
cv2.waitKey(0)
newImg = (grayImg-gaussianImg)
cv2.imshow('New Image', newImg)
cv2.imwrite('Converted.png', newImg)
cv2.waitKey(0)
cv2.destroyAllWindows()
```



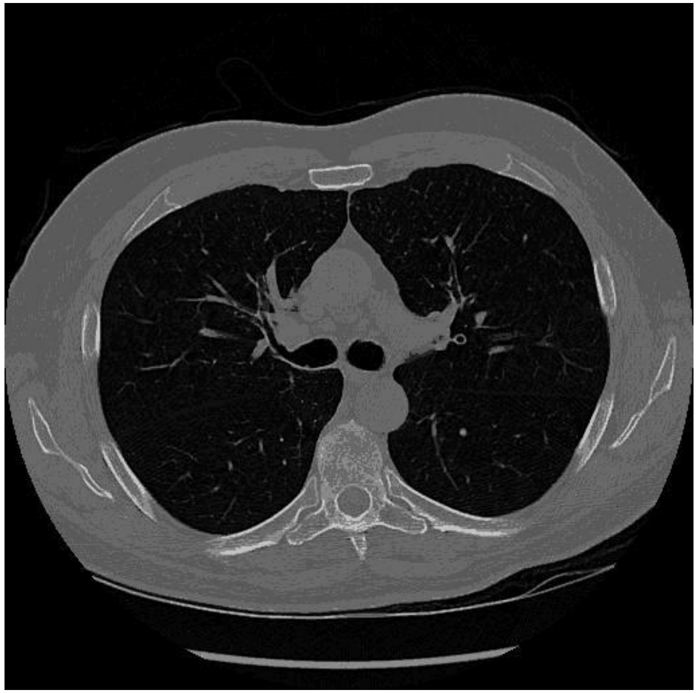
K-Means Clustering:

```
import numpy as np
import cv2
import matplotlib.pyplot as plt

path=r'C:\Users\user\OneDrive\Desktop\review3DIPData\1-053.jpg'

img = cv2.imread(path)
Z = img.reshape((-1,3))

Z = np.float32(Z)

criteria = (cv2.TERM_CRITERIA_EPS + cv2.TERM_CRITERIA_MAX_ITER, 10, 1.0)

K = 3
ret,label,center=cv2.kmeans(Z,K,None,criteria,10,cv2.KMEANS_RANDOM_CENTERS)

center = np.uint8(center)
res = center[label.flatten()]
res2 = res.reshape((img.shape))

cv2.imshow('k=3',res2)
plt.imshow(res2,cmap=plt.cm.bone)
cv2.imwrite('k=3.jpg',res2)
```

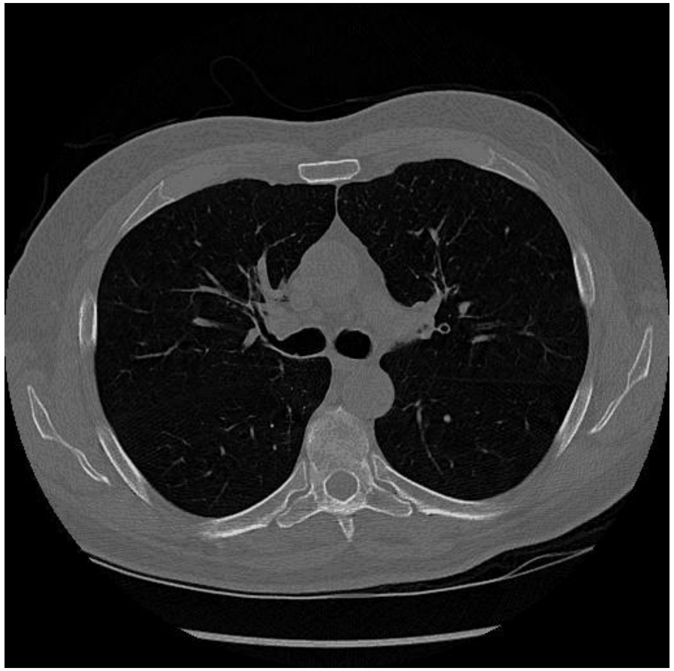
```
K = 5
ret,label,center=cv2.kmeans(Z,K,None,criteria,10,cv2.KMEANS_RANDOM_CENTERS)

center = np.uint8(center)
res = center[label.flatten()]
res2 = res.reshape((img.shape))

cv2.imshow('k=5',res2)
cv2.imwrite('k=5.jpg',res2)
```



```
K = 8
ret,label,center=cv2.kmeans(Z,K,None,criteria,10,cv2.KMEANS_RANDOM_CENTERS)

center = np.uint8(center)
res = center[label.flatten()]
res2 = res.reshape((img.shape))

cv2.imshow('k=8',res2)
cv2.imwrite('k=8.jpg',res2)
cv2.waitKey(0)
cv2.destroyAllWindows()
```



```
K = 15
ret,label,center=cv2.kmeans(Z,K,None,criteria,10,cv2.KMEANS_RANDOM_CENTERS)

center = np.uint8(center)
res = center[label.flatten()]
res2 = res.reshape((img.shape))

cv2.imshow('k=15',res2)
cv2.imwrite('k=15.jpg',res2)
cv2.waitKey(0)
cv2.destroyAllWindows()
```



Classification:

```
from tensorflow.keras.layers import Input,Lambda,Dense,Flatten
from tensorflow.keras.models import Model
from tensorflow.keras.preprocessing import image
from tensorflow.keras.preprocessing.image import ImageDataGenerator from
tensorflow.keras.models import Sequential
import numpy as np
from glob import glob
```

```
import matplotlib.pyplot as plt
IMAGE_SIZE = [224,224]
train_path = "Lung_cancer/Train/"
from keras.preprocessing.image import ImageDataGenerator
train_datagen =
ImageDataGenerator(rescale=1./255,horizontal_flip=True,zoom_range=0.2,validatio
n_split=0.15) training_set = train_datagen.flow_from_directory(
train_path,target_size=(224,224), batch_size=32,class_mode='categorical',
subset='training')
validation_set = train_datagen.flow_from_directory(
train_path,target_size=(224,224), batch_size=32,class_mode='categorical',shuffle =
True, subset='validation')
from tensorflow.keras.applications import VGG19
from tensorflow.keras.layers import GlobalAveragePooling2D,Dropout
## We are initialising the input shape with 3 channels rgb and weights as
imagenet and include_top as False will make to use our own custom inputs
mv = VGG19(input_shape=IMAGE_SIZE+[3],weights='imagenet',include_top=False)
for layers in mv.layers:
layers.trainable = False
# if u want to add more folders and train then change number 4 to 5 or 6 based
on folders u have to train
x = Flatten()(mv.output)
prediction = Dense(2,activation='softmax')(x)
model = Model(inputs=mv.input,outputs=prediction)
model.summary()
import tensorflow as tf
class myCallback(tf.keras.callbacks.Callback):
def on_epoch_end(self,epoch,logs={}):
```

```
if(logs.get('loss')<=0.05):
print("\nEnding training")
self.model.stop_training = True
# initiating the myCallback function
callbacks = myCallback()
Training Model:
def create_model():
model.add(ZeroPadding3D(0, 1, 1)))
model.add(Convolution3D(64, (3, 3, 3), activation = 'relu'))
model.add(MaxPooling3D((2, 2, 2), strides = None, padding = 'valid))
model.add(ZeroPadding3D(0, 1, 1)))
model.add(Convolution3D(128, (3, 3, 3), activation = 'relu'))
model.add(ZeroPadding3D(0, 1, 1)))
model.add(Convolution 3D(128, (3, 3, 3), activation = 'relu'))
model.add(ZeroPadding3D(0, 1, 1)))
model.add(Convolution3D (256, (3, 3, 3), activation = 'relu'))
model.add(ZeroPadding3D(0, 1, 1)))
model.add(Convolution3D (256, (3, 3, 3), activation = 'relu'))
model.add(ZeroPadding3D(0, 1, 1)))
model.add(Convolution3D (256, (3, 3, 3) activation = 'relu'))
model.add(MaxPooling3D((2, 2, 2), strides = None, padding = 'valid))
model.add(ZeroPadding3D(0, 1, 1)))
model.add(Convolution3D(512, (3, 3, 3), activation = 'relu'))
model.add(ZeroPadding3D(0, 1, 1)))
model.add(Convolution 3D(512, (3, 3, 3), activation = 'relu'))
model.add(ZeroPadding3D(0, 1, 1)))
model.add(Convolution3D(512, (3, 3, 3), activation = 'relu'))
model.add(MaxPooling3D((2, 2, 2), strides = None, padding = 'valid))
```

model.add(ZeroPadding3D(0, 1, 1)))
model.add(Convolution3D(512, (3, 3, 3), activation = 'relu'))
model.add(ZeroPadding3D(0, 1, 1)))
model.add(Convolution3D(512, (3, 3, 3), activation = 'relu'))
model.add(ZeroPadding3D(0, 1, 1)))
model.add(Convolution 3D(512, (3, 3, 3), activation = 'relu'))
model.add(MaxPooling3D((2, 2, 2), strides = None, padding = 'valid))
model.add(Flatten())
model.add(Dense(4096, activation = 'relu'))
model.add(Dropout(0.5))
model.add(Dense(4096, activation = 'relu'))
model.add(Dropout(0.5))
model.add(Dense(nclasses, activation = 'softmax'))

Output Checking:

# import the necessary packages
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.applications import VGG16
from tensorflow.keras.layers import AveragePooling2D
from tensorflow.keras.layers import Dropout
from tensorflow.keras.layers import Flatten
from tensorflow.keras.layers import Dense
from tensorflow.keras.layers import Input
from tensorflow.keras.models import Model
from tensorflow.keras.optimizers import Adam
from tensorflow.keras.utils import to_categorical
from sklearn.preprocessing import LabelBinarizer
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report

```
from sklearn.metrics import confusion_matrix
from imutils import paths
import matplotlib.pyplot as plt
import numpy as np
import argparse
import cv2
import os
from tensorflow.keras.preprocessing import image
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.models import load_model
from tensorflow.keras.preprocessing import image
import numpy as np
from tensorflow.keras.preprocessing.image import load_img, img_to_array
from tensorflow.keras.models import load_model
from tensorflow.keras.preprocessing import image
import numpy as np
classes = {0:"normal:-{ normal }",1:"lungcancer:-{ lungcancer} "}
# dimensions of our images
img_width, img_height = 224,224
# load the model we saved
model = load_model('VGG-19.h5')
# predicting images
#img = image.load_img('lungcancer/Train/lungcancer/3.png', target_size=(img_width,
img_height)) image =
load_img('lungcancer/Train/Normal/3.jpeg',target_size=(224,224))
image = img_to_array(image)
image = image/255
```

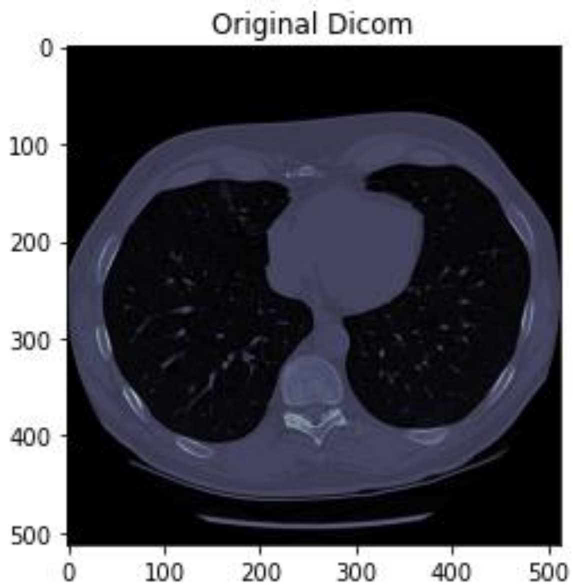
Gaussian Filter:

```
img_median = cv2.GaussianBlur(img,(5,5), 0)
res = np.concatenate((img, img_median), axis=1)
Histogram equalisation :
equ = cv2.equalizeHist(res)
res = np.hstack((img, equ))
Image Segmentation by Thresholding :
thresh1 = cv2.threshold(res, 120, 255, cv2.THRESH_BINARY)
final_image = thres1
final_image = np.expand_dims(equ,axis=0)
model = load_model('VGG-19.h5')
result = np.argmax(model.predict(final_image))
prediction = classes[result]
cv2.imshow(prediction, final_image)
```

Result: The final accuracy received on predicting over the remaining dataset is 94.38% using CNN algorithm

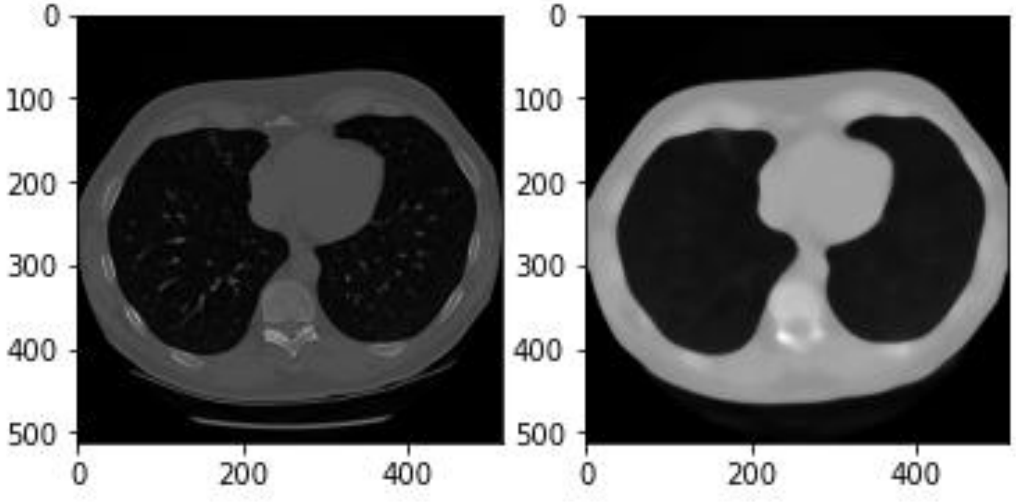
Test Case 2:

Sample Input image:

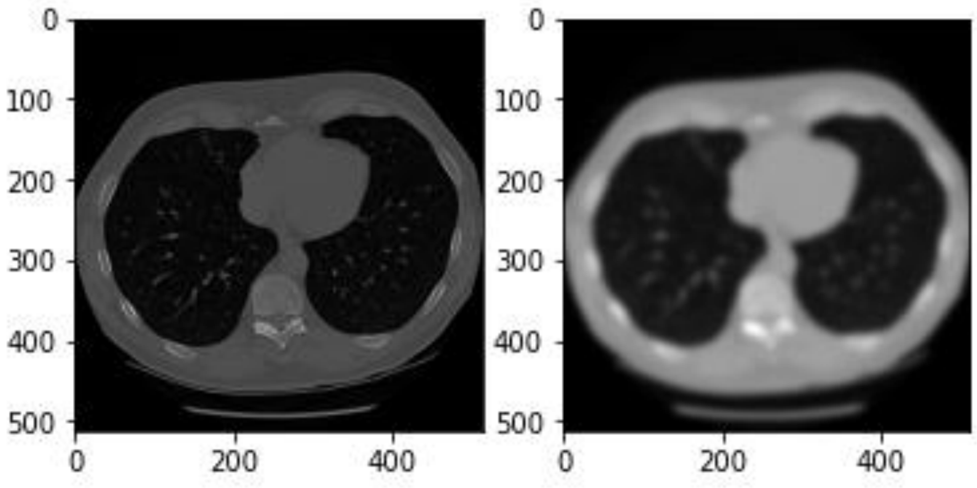


Output Received:

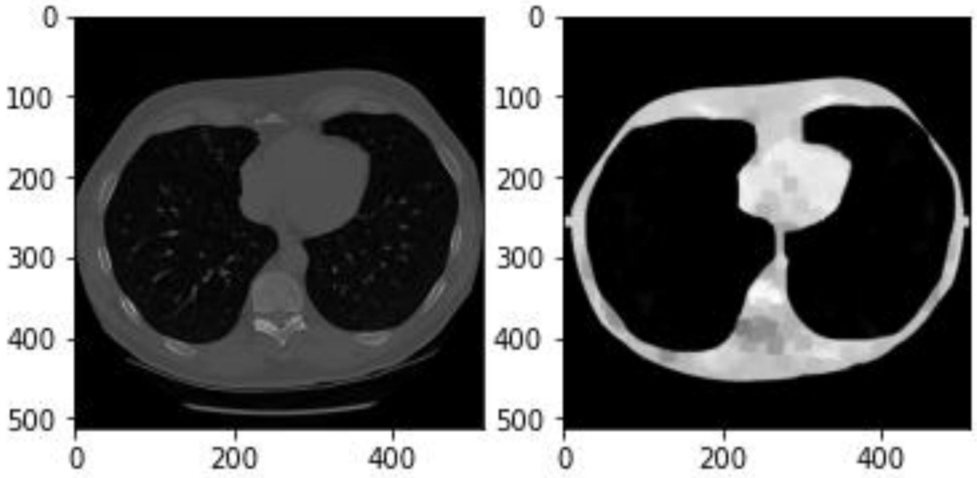
Median Filtering:



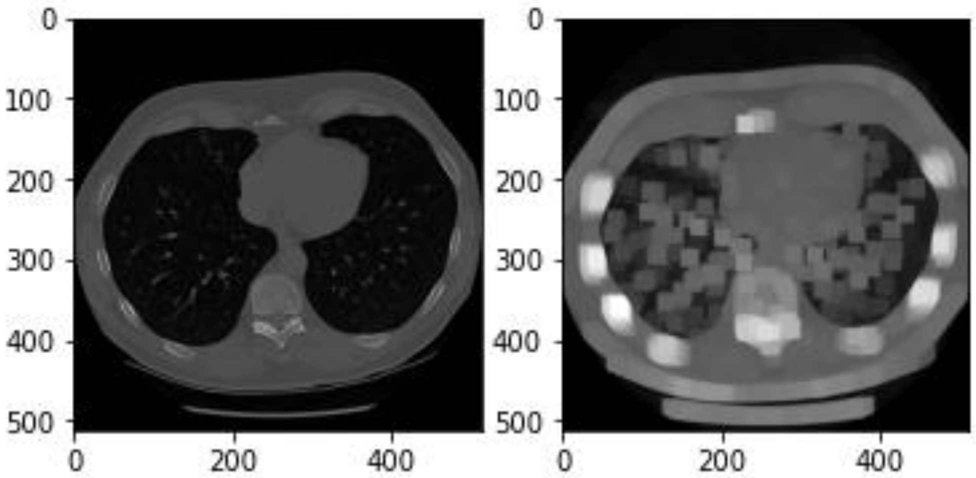
Gaussian Filtering:



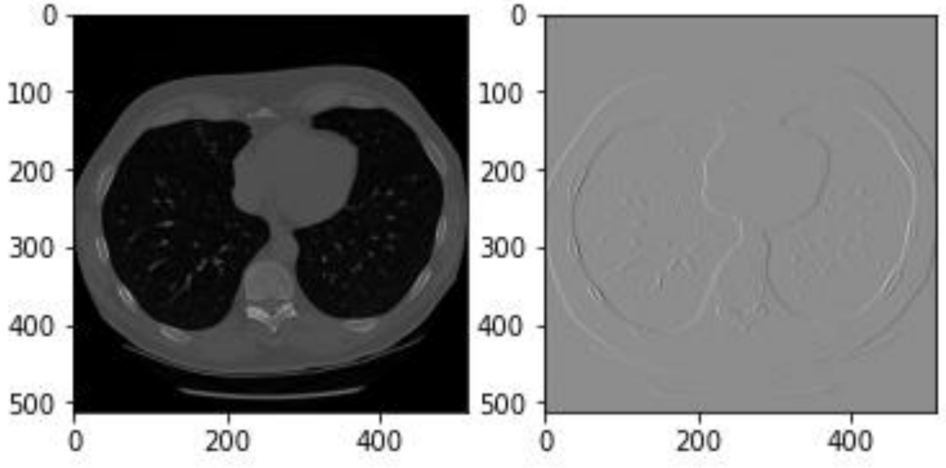
Minimum Filtering:



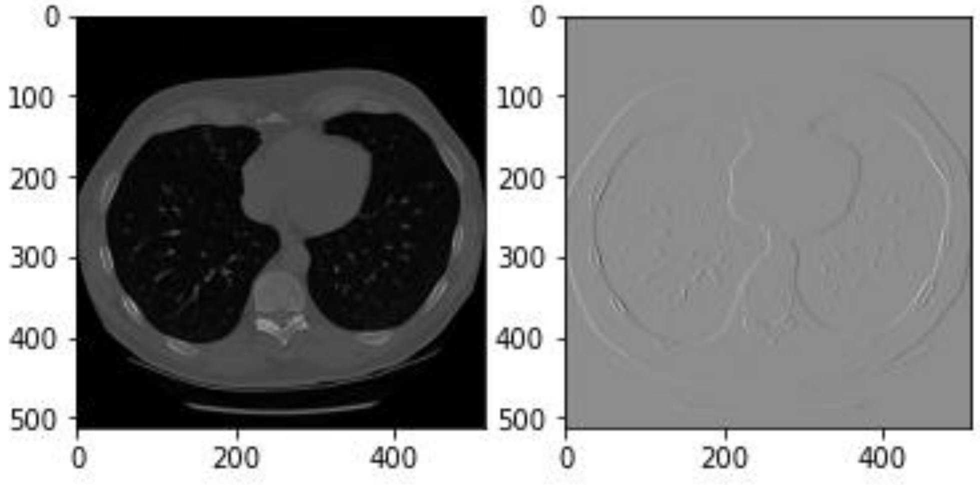
Maximum Filtering:



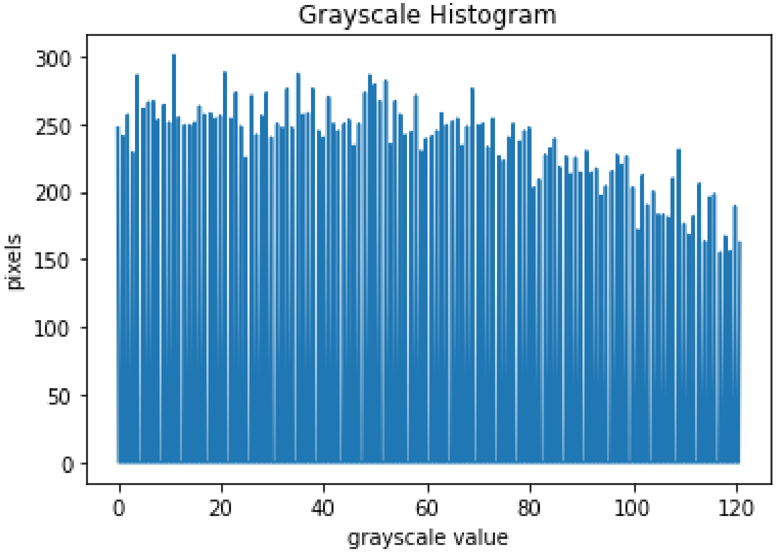
Sobel Filtering

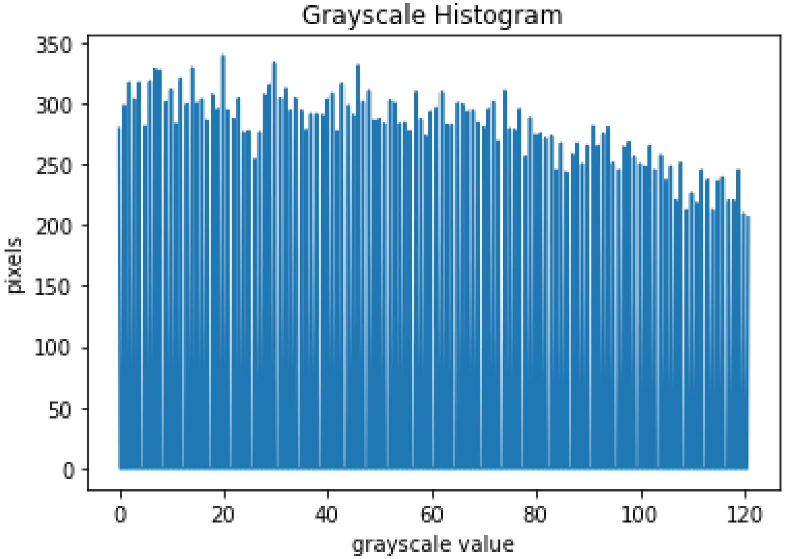


Prewitt Filtering:



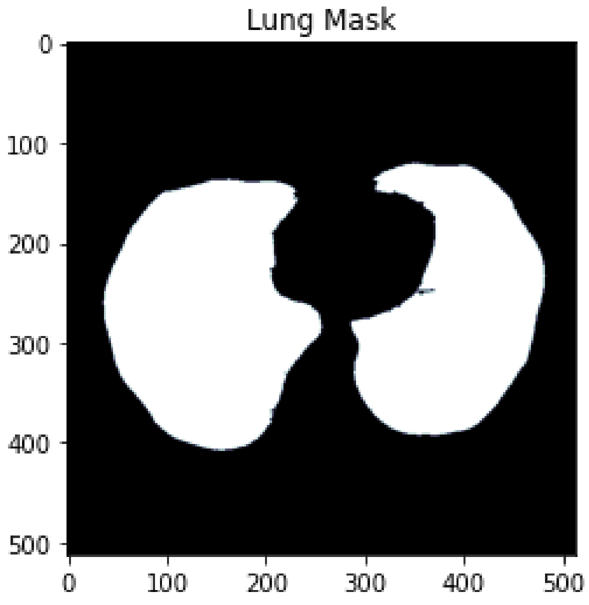
Histogram Visualization:

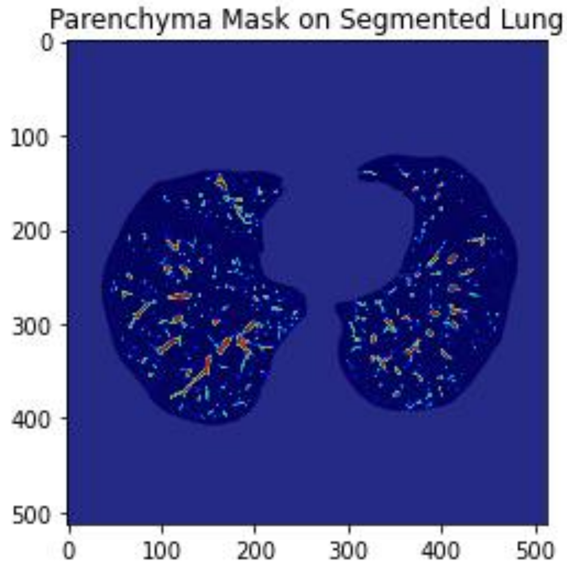




Segmentation

Hounsfield Units





K-NN clustering

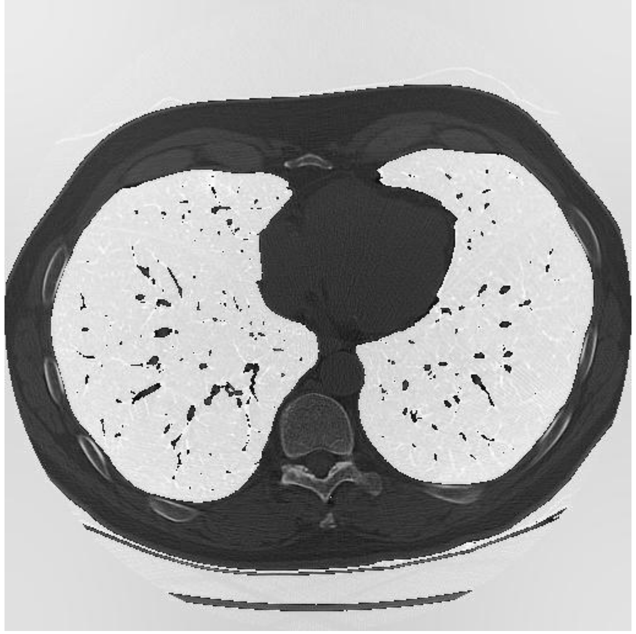
K=3

K=5

K=8

K=15

Shading Correction:



Rotation:

Rotated by 45 degree

Rotated by 90 degree