76 Analysis of various image processing approaches in detecting and predicting lung cancer

Mohd. Munazzer Ansari[1,a] *and Shailendra Kumar*[1,b]

[1]Department of Electronics Engineering, Integral University, Lucknow, Uttar Pradesh, India

Abstract

Image processing is one of the centre area that utilised in different domains. It is utilised to categorise cancer influenced areas in lung image. Early discovery of lung cancer can build the opportunity of endurance among people. Recognisable proof of cancer influenced areas in lung is basically started with image processing techniques. The techniques are followed by pre-processing having image enhancement along with noise removal using filtering techniques then Image segmentation with feature extraction. Finally, Classification of lung cancer and performance comparison of chronological data of lung cancer. Typically, Digital Image Processing follows numerous strategies to join various shapes in an image into a solitary unit. Despite the fact that Computed Tomography (CT) can be more proficient than X-ray. For detecting and predicting lung cancer, digital image processing techniques can be very much useful for radiologists. In this study, here MATLAB software is used and we discussed about an efficient study on performance analysis techniques in lung cancer detection and prediction by using different image processing methodologies.

Keywords: image processing, image segmentation, lung cancer, SVM and MATLAB

Introduction

Lung cancer is the significant reason for malignant growth passing on the world. The side effects of lung cancer come into light at the last stage. Therefore, it is exceptionally hard to recognise in its early phase. Thus, the passing rate is exceptionally high for lung cancer in correlation with any remaining sorts of disease. The two types of lung cancer which take place and spread in a sudden manner, are small cell lung malignancies (SCLC) and non-small cell lung tumours (NSCLC) [1]. The period of lung infection implies how much the development has spread in the lung. As per measurements led by world wellbeing association that consistently more than 7.6 million people passed-on of lung cancer. Also, the death-rate of lung cancer are relied upon to continue to ascend, to end up around 17 million worldwide in 2030 [2].

The information on Lung cancer includes numerous methods and technologies. A few Computer Aided Diagnosis (CAD) frameworks are formed for identifying lung cancer in its beginning phase. A CAD framework may comprise a few stages in detecting lung cancer. They are as follows: (1) Pre-processing or lung Segmentation, (2)

[a]munazzer@student.iul.ac.in; [b]skumar@iul.ac.in

Figure 76.1: Different steps in detecting lung cancer

Nodule detection, (3) Nodule Segmentation, (4) Feature Extraction, (5) Classification, (6) Prediction. Figure 76.1 shows the mean of the CAD framework in distinguishing lung disease. In its initial step, the images are pre-processed.

Literature Review

There are many methods for detecting and predicting lung cancer by implementing of image processing methodologies. In this study Table 76.1 shows various pre-processing and segmentation techniques since 2009-2021 and the literature survey based on that table. The major contribution in this field is summarised below.

Murphy, Keelin, (2009) [4] et al. described a plan for the programmed location of knobs in the field of thoracic (CT) - computed tomography examines has been introduced with broadly assessed. The implemented algorithm exploits the near-by image features of figure record and curvature to identify applicant constructions in the lung capacity also with application of different progressive KNN classifiers to decrease of bogus up-sides. They tried for knob identification framework and prepared it on 3 data sets removed obtained from an enormous scope trial screening study. De Oliveira Nunes, Éldman (2010) [5] et al. presents an involuntary technique of Medical image segmentation employed in the field of investigation of the (CNS) - Central Nervous System with the help of staggered thresholding dependent at histogram distinction. Our technique delivered a presentation of 88.6%, for the considered testing images, when the outcomes where contrasted and those given by a human expert.

Sharma, Disha, and Jindal (2011) [6] et al. They described (CAD) system for the application of finding lung cancer with the processing on CT images. In recent years there are fast growing research in the application field of medical diagnosis using image processing procedures implemented for findings. Initially, the concepts like Median Filter, Erosion, Outlining, Dilation and Features Extraction are explored and implemented on CT scanned pictures with the purpose of detection of the lung cancer. Using various algorithms of image segmentation, extracted cancer nodules occurred in the lung. After segmentation, rule-based techniques are used to identify for classification of cancer. Al-Tarawneh, Mokhled S. (2012) [7] et al. describes the position request of lung cancer for the males and females among Jordanians in 2008 shown that 356 examples of cellular breakdown in the lungs space indicating (7.7%) of all newly detected malignancy cases in 2008. Cellular breakdown in the lungs affected 13.1% i.e., 297 males 2.5% equal to 59 females and this ratio is defined as 5:1 that also explained that this abnormally exhibit in every second males and every tenth females.

Rani, J. (2013) [8] et al. proposed the system, during analyse the images might get adulterated by noise or the X-beam images generated noise. Various types of noise elimination Filters are fundamentally used and causes different errors produced in the process of image acquisition. Different cleaning trials are used for image advancement. Linear filter along with median filter has been employed to eliminate noise also

Table 76.1: Procedures used in pre-processing and segmentation since 2009-2021 with application and its improvement

References	Method Used	Dataset/Samples	Application	Improvements
Murphy, Keelin, et al. (2009) [4]	Region growing and morphological smoothing	750 training scans divided into 03 sets of 250 scans	Image Feature and local K-mean Classification	For creating a nodule detection system algorithm
De Oliveira Nunes, Éldman et al. (2010) [5]	Image Segmentation	30 images, in BMP format and with spatial resolutions from 512x512 pixels	Multilevel thresholding based on histogram difference	The extension of this methodology to different format of images
Sharma, Disha, and Jindal et al. (2011) [6]	Wiener Filter	1000 lung images	For Denoising and Signal detection	Detection of cancerous nodules during the study of clinical images of CT at 2.5-5.0 mm
Al-Tarawneh, Mokhled, S. et al. (2012) [7]	Gabor filter within Gaussian rules	Grey image usually contains 256 levels	Segmentation and Optical Character Recognition	To achieve more accuracy rate than 81.835 in thresholding
Rani, J., et al. (2013) [8]	Noise Removal using Filter	Images size of 3*3, 5*5 etc. for removal noise	Used to separate layer of image by removal of noise	Used non-linear filter to remove noise
Gajdhane, Vijay, A., et al. (2014) [9]	Grey Scale Image	X-ray, CT scan, MRI, PET images consider for technique	Used to change colour in grey	Increasing the number of images used in the process and MRI, X-ray, PET images are the types of images that offer the best results for detecting lung cancer
Onizawa, Naoya, et al. (2015) [10]	Gabor Filter	68 parallel stochastic Filters	Feature Extraction	Extended to the 2-D or 3-D Gabor filter by adding two scaled additions
Malik, Bhawana, et al. (2016) [11]	Image Processing Techniques and Classification	With 24 images and 17 images	By using non-local mean filter to remove Gaussian white noise	Prediction of disease in different stages using SVM with more than 95.12% accuracy rate
Al Zubaidi, Abbas K., et al. (2017) [12]	Classification	Histological images for large database	The CAD-FCM method used to diagnose lung cancer and diagnose	Radiology and histological imaging can look at future developments to find the right answer for serious cases

all the costs. All the outcomes supposed by the researchers proved that noise elimination is ideally performed by the median filter. Finally, both filters are used to apply for noise reduction in the images.

Gajdhane, Vijay A. (2014) [9] et al. in this they mentioned Early recognition of lung cancer can expand the report at endurance amongst individuals. The overall

five-year endurance degree for lung cancer affected people increases up to 49% when the infection is framed on schedule. General concept is CT may be more proficient than X-ray. Onizawa, Naoya, (2015) [10] et al. mentioned about the Gabor filter shows as incredible feature extraction capability from images, yet it requires huge computational power. Utilising stochastic calculation, a sin function utilised within Gabor filter is approached by taking advantage of a few stochastic tanh functions planned dependent with the help of state machine.

Malik, Bhawana, (2016) [11] et al. described some different to identify lung cancer at very early stage employing CT scan images obtained from Dicom. Al Zubaidi, Abbas K. (2017) [12] et al. this paper gives a wide survey to most significant calculations are experimented on Computer Aided Design-CAD exploring for lung material examined and featured all presentation of individually unmistakable by appropriate algorithm.

Alam, Janee, (2018) [13] et al. they focused on the detection as well as prediction of cancer. This paper projected an actual and explained lung cancer detection technique with accurate prediction procedure with the assessment by multi-class SVM-Support Vector Machine classifier. Suren, (2018) [14] et al. focused on the CAD (computer aided diagnosis) techniques utilising for processing of various image with different algorithm of machine learning implementation. Projected prototype identifies obtained image's malignancy of 92% precision more than existing model and achieve accuracy of 86.6%. In general, they achieved development in their projected agenda in comparison with current best model. But this planned method is not able to structure numerous stages like I, II, III, IV stages of disease. Dev, Chethan, (2019) [15] et al. proposed the technique explained a better technique of computer-aided classification applied on automated tomography pictures of lungs. In this work, planned method on MATLAB for execution of each methodology.

Pawar, Vikul J. (2020) et al. [17], in this author mentioned about the Computer Aided Diagnosis (Computer aided design) structures requires the pre-processing and feature extraction obtained by X-beams methods, (CT)- Computer Tomography yield to separate the Lungs facts in human body. Manju, B. R., V. Athira (2021) et al. [18].

Methodology

The process of detection and prediction of lung cancer using image processing technique has been built by MATLAB. In this study proposed method implemented in MATLAB by various data of lung cancer. In each stage of algorithm classification image pre-processing, enhancement, segmentation, followed by feature extraction has been done.

Image Pre-processing

Pre-processing is the softening of unwanted distortions or the augmentation of particular visual highlights in preparation for future processing. Pre-processing is required to eliminate unwanted regions in the image and is sometimes used to enhance image features such as lines, boundaries, and surfaces. It is expected to minimise the effects of distortion found on the imaging gadget, such as fluctuations in light, to eliminate blueness, and at the same time, it is required to eliminate unwanted regions in the image and is sometimes used to enhance image features such as lines, boundaries, and

Table 76.2: Various noise type and its application in image processing

S.No.	*Noise Type*	*Application*
1	Gaussian Noise	Outside the Normal Distribution values, not seen in the image
2	Salt and Paper Noise	Tiny black and white pointe randomly appears in the image
3	Poisson Noise	Noise is appear due to non-linear response of image detectors
4	Impulse Noise	Appears in EM- electromagnetic interference, scratches in the disk

Table 76.3: Different filtering techniques

S.No.	*Techniques*	*Method Used*	*Application*
1	Gabor Filter	Linearing Filtering technique used Gaussian function	Face recognition and vehicle verification etc.
2	Auto enhancement or Weiner Filter	Automatic recognition and mathematical functionality such as mean calculations and variations	Denoising and speech recognition etc.
3	Fast Fourier transform (FFT)	Algorithm takes a signal from a specific space or time and divides it into its frequency forms	Ultrasound images and MR (Magnetic Resonance) images etc.

surfaces. In medical imaging all sorts of filtering techniques can be applied depending on the noise present in the image [19]. Detail list is given in Table 76.2.

Image Enhancement

A technique used to improve the quality of image is known as image enhancement (Table 76.3). The main goal of image enhancement is to make image better by improving the quality of image. It can be divided into two categories.

i. Spatial Domain method.
ii. Frequency Domain method.

Image Segmentation (Table 4)

Table 76.4: Types of segmentation and its algorithm used in detecting lung cancer

S.No	*Techniques*	*Method Used*	*Application*
1	Thresholding	Histogram thresholding, Otsu's thresholding, fast matching thresholding	Nodule segmentation, medical image segmentation
2	Watershed transform	Morphological operations	Medical CT data collection
3	Edge detection	Sobel-algorithm, common edge detection, Canny Prewitt, Robert's, fuzzy logic methods	Computer vision, machine vision
4	Region growing	Seeded or unseeded growing	Detection of cardiac disease, Delineate tumour volumes
5	Clustering method	K-mean clustering	Vector quantisation, Features- extraction
6	Manual segment	Photoshop software	Volumetric measurement in MRI

Table 76.5: Types of classification techniques

S.No.	*Classification Techniques*	*Method used*	*Application*
1	Bag Classifier	Language Processing and information retrieval (IR)	Computer vision
2	Naive Bayes Classifier	Probability Model using Bay's equation	Sophisticated Classification techniques
3	K-NN Classifier	Non-parametric method	Classification and Regression
4.	Adaboost Classifier	Low training error	To Multi Class and Regression Problem
5.	SVM Classifier	Binary Classification Method	Data Classification, Speech recognition, Cancer Diagnosis & Prognosis
6.	ELM Classifier	Least square method	Pattern Recognition, forecasting and Diagnosis etc.

Feature Extraction

Subsequently segmentation, a segmented lung nodule is used for feature extraction. To diagnose lung cancer, we need to find the key features in the image. In the image processing techniques, there are three main types of focus areas for this study are as following:

a. Firstly, the structural element.
b. Secondly, the elements of mathematical i.e. statistical texture.
c. Third is the features of spectral.

Classification

Classification of images is an essential errand that seeks to clarify the image all in all. By allocating a mark, the intention is to recognise the image. Image classification as a rule mentions to images where just one object shows up and is inspected. Then again, object recognisable proof requires both classification and localisation errands and is utilised to analyse more common sense occurrences in which an image may have a few objects. Here the errand is to characterise lung nodules as malignant or benevolent. Different classification procedures are recorded beneath in Table 76.5 alongside the outcomes acquired [23].

Tools Required

Experiments of proposed method is based on MATLAB software and are implemented on the following:

- 8GB RAM per worker is recommended.
- MATLAB 9.9.
- Processor: Intel I-7 with CPU speed 400-450MHz.
- Operating System: Windows 10.
- TCP port are required for MATLAB workers.
- Database of lung cancer images got from the IMBA Home.

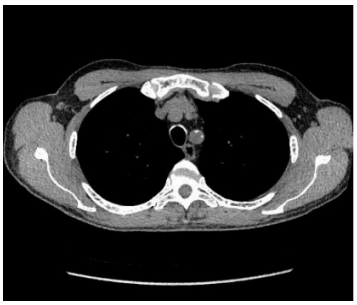

Figure 76.2a: Input CT-scan image

Figure 76.2b: Filtered image

Figure 76.2c: Thresholding

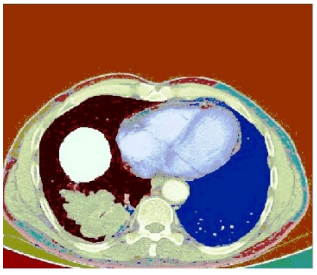

Figure 76.2d: Watershed transform image

Discussion

This study presents an assessment among the most recent research paper in the field of the lung cancer. This performance-based paper, a comprehensive analysis has been done in the predicting and detecting the lung cancer. The main step of image processing in detection and prediction of lung cancer are as follows: image pre-processing followed by image acquisition, noise detection, filtering, image enhancement analysis viewed in two domains, and further by image segmentation techniques using different algorithm analysis, then segmented image is gone under feature extraction techniques and finally classifications of images in lung cancer. In the field of image processing, initially the CT scan/ MRI image processed for cancer detection. There are some captured images showed in the analysis using image processing techniques. In the pre-processing phase of analysis, the input CT- image shown in Figures 76.2(a–d) has been processed to improve the quality.

The CT image is converted into grey scale to mathematical operations then pre-processing is done by filtering which allow to remove noise completely from the image. Figure 76.2b shows the filtered image. MATLAB can be used for filtering operation. Figure 76.2c shows thresholding.

Conclusion & Future Work

Lung cancer is a widely recognised form of malignant growth, affecting a large number of people worldwide. If lung cancer is detected early enough, it can be managed. Image Processing Mechanisms are very important in the medical field for recognising and estimating various diseases, including lung cancer. An accurate system can be recommended by recognising the main stages of Image Processing to predict and detect

dangerous lung tumours, which consist mainly of noise removal, image enhancement, image segmentation, feature extraction, and feature classification to determine whether a knobble or nodule is normal or abnormal. Many of the structure were designed by a researcher and many of them are discussed in this paper. The main objective of this paper is to design the system for detecting and predicting lung cancer nodule by various classifiers such as ANN, SVM etc. For future work, I look forward to proposing a more effective lung cancer detection and predictive system for image processing techniques for features extraction of AI algorithm such as shark scalp enhancement, to be separated using a (LGBM) light gradient magnification machine.

References

[1] Krishnaiah, V., Narsimha, G., Subhash, C. (2013). Diagnosis of lung cancer prediction system using data mining classification techniques. *Int J Comp Sci Inform Technol.* 4(1):39–45.

[2] Dignam, J. J., Huang, L., Ries, L., Reichman, M., Mariotto, A., Feuer, E. (2009). Estimating cancer statistic and other-cause mortality in clinical trial and population-based cancer registry cohorts. *WileyInterScience*[Online]. 115(22). Available:http://onlinelibrary.wiley.com/doi/10.1002/cncr.24617/epdf.

[3] Sevani, A., et al. (2018). Implementation of image processing techniques for identifying different stages of lung cancer. *Int J Appl Engg Res*. 13(8):6493–6499.

[4] Murphy, K., et al. (2009). A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification. *Med Image Anal.* 13(5):757–770.

[5] De Oliveira, N. É., Maria, G. P. (2010). Medical image segmentation by multilevel thresholding based on histogram difference. *17th International Conference on Systems, Signals and Image Processing*.

[6] Sharma, D., Gagandeep, J. (2011). Identifying lung cancer using image processing techniques. *Int Conf Comput Tech Artif Intel (ICCTAI)*. 17.

[7] Al-Tarawneh, M. S. (2012). Lung cancer detection using image processing techniques. *Leonardo Elec J Prac Technol.*11(21):147–158.

[8] Rani, J. (2013). Noise removal in medical images using filters. *Int J Engg Res Technol.* 2:1013–1016.

[9] Gajdhane, V. A., Deshpande, L. M. (2014). Detection of lung cancer stages on CT scan images by using various image processing techniques. *IOSR J Comp Engg (IOSR-JCE)*. 16(5):28–35.

[10] Onizawa, N., et al. (2015). Gabor filter based on stochastic computation. *IEEE Signal Process Lett.* 22(9):1224–1228.

[11] Malik, B., et al. (2016). Lung cancer detection at initial stage by using image processing and classification techniques. *Lung Cancer* 3(11).

[12] Al Zubaidi, A. K., et al. (2017). Computer aided diagnosis in digital pathology application: Review and perspective approach in lung cancer classification. *2017 Annual Conference on New Trends in Information & Communications Technology Applications (NTICT)*. IEEE.

[13] Alam, J., Sabrina, A., Alamgir, H. (2018). Multi-stage lung cancer detection and prediction using multi-class svm classifie. *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*. IEEE.

[14] Makaju, S., et al. (2018). Lung cancer detection using CT scan images. *Procedia Computer Science* 125:107–114.

[15] Dev, C., et al. (2019). Machine learning based approach for detection of lung cancer in DICOM CT image. *Ambient Communications and Computer Systems*. Springer, Singapore, pp. 161–173.
[16] Abdullah, M. F., et al. (2020). 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)*. IEEE.
[17] Tripathi, S. L., Patel, G. S. (2020). Design of low power Si0.7Ge0.3 pocket junction-less tunnel FET using below 5 nm technology. *Wireless Pers Commun*. 111:2167–2176. https://doi.org/10.1007/s11277-019-06978-8 ISSN: 0929-6212.
[18] Suman, L. T., Raju, P., Vimal Kumar, A. (2019). Low leakage pocket junction-less DGTFET with bio-sensing cavity region. *Turk J Elec Engg Comp Sci*. 27(4):2466–2474. DOI:10.3906/elk-1807-186.
[19] Verma, S. B., Saravanan, C. (2019). Performance analysis of various fusion methods in multimodal biometric. *Proceedings of the International Conference on Computational and Characterization Techniques in Engineering and Sciences, CCTES*. pp. 5–8.
[20] Saravanan, C., Satya Bhushan, V. (2015). Touchless palmprint verification using shock filter SIFT I-RANSAC and LPD IOSR. *J Comp Engg*. 17(3):2278–8727.
[21] Pawar, V. J., et al. (2020). Lung cancer detection system using image processing and machine learning techniques. *Cancer* 3.
[22] Manju, B. R., Athira, V., Athul, R. (2021). Efficient multi-level lung cancer prediction model using support vector machine classifier. *IOP Conference Series: Materials Science and Engineering*. 1012(1). IOP Publishing.
[23] Al-Tarawneh, M. S. (2012). Lung cancer detection using image processing techniques. *Leonardo Elec J Prac Technol*. 11(21):147–158.