

Proposal on Detection of Lung Cancer on Reduced Images Using Foggy K-Means Clustering Algorithm

S. Namitha¹, R. Preethi², M. Vinitha³, Dr. T. Kalaichelvi⁴, Dr. S. Hemalatha⁵, Dr. V. Subedha⁶

^{1,2,3}Research Scholar, Department of CSE, Panimalar Institute of Technology, Chennai, Tamil Nadu, India.

^{4,5,6}Professor, Department of CSE, Panimalar Institute of Technology, Chennai, Tamil Nadu, India.

ABSTRACT

Lung cancer is the second most common cancer in both women and men and is by far the most pivotal cause of cancer death among both men and women. Each year, a huge count of people die of lung cancer than of colon, breast, and prostate cancers combined. Early detection of lung cancer can increase the chance of survival among people. The overall 5-year survival rate for lung cancer patients increases from a meagre 14% to a decent 49% if the disease is detected at the right time. Although Computed Tomography (CT) can be more efficient than X-ray, however, problem was the time constraint in detecting the presence of lung cancer cells using the several diagnosis methods used. Hence, a lung cancer detection system using image processing is used to classify the present of lung cancer in a CT- images. In this project, MATLAB has been used through every procedure made. This involves image pre-processing, segmentation based on PPA, Foggy K-Means clustering and feature extraction by Neural network. The aim is to get more accurate results by using various enhancement and segmentation techniques.

Keywords : Dimensionality reduction, clustering, segmentation, PPA algorithm, histogram equalization, neural networks, Computed Tomography, Foggy k-means clustering

I. INTRODUCTION

Lung cancer is the most deadly cancer disease with an estimated 27% of all cancer deaths. A large part of patients with lung cancer undergoes radiotherapy. Tracking tumour motion poses a significant challenge for precise dose delivery in lung cancer radiotherapy, due to respiratory motion of up to 3.5 cm for primary lung tumours. If the patient's breathing motion is not correctly predicted, tumour miss might occur, or sensitive normal tissue might be undesirably exposed resulting in unwanted treatment toxicity. Advanced technologies of radiotherapy, intensity modulated radiotherapy and image guided radiotherapy, may offer the potentiality of precise radiation dose delivery for moving objects. However, they still need an additional function that is used to predict the precise position of the tumour against the subtle variations in real-time. Radiation measurement is ordinarily conveyed in 3 to 5 portions more than 5 to 12 days for early stage lung tumour utilizing stereotactic radiotherapy or 30 to 33

parts more than 6 to 7 weeks for more propelled sickness with each division enduring near 10 and 45 minutes

II. METHODS AND MATERIAL

A. Literature Survey

[1] Lung Cancer Detection by Using Artificial Neural Network and Fuzzy Clustering Methods.

This paper presents two division techniques, Hopfield Neural Network (HNN) and a Fuzzy C-Mean (FCM) grouping calculation, for sectioning sputum shading pictures to recognize the lung tumour in its initial stages. The division results will be utilized as a base for a Computer Aided Diagnosis (CAD) framework for early recognition of lung disease. In this review, 1000 sputum shading picture is utilized to test both strategies, and HNN has demonstrated a superior characterization result than FCM, the HNN prevailing with regards to separating the cores and cytoplasm areas. FCM flopped

in distinguishing the cores, rather it recognized just piece of it. Notwithstanding that, the FCM is not touchy to force varieties as the division blunder at joining is bigger with FCM contrasted with that with HNN.

[2] Clustering Of Lung Cancer Data Using Foggy K-Means.

In this flow look into work principle centre is to build up a novel way to deal with make precise groups of coveted constant datasets called Foggy K-implies bunching. The results in terms of more accurate clusters could be utilized by domain experts for their strategic planning. Need to use the consequences of foggy k-implies bunching for order of lung malignancy patients and concentrate the components that indicated incredible effect on lung disease.

[3] Feature Extraction and Principal Component Analysis for Lung Cancer Detection in CT scan Images.

In this paper features are extracted using principal component analysis and Histogram Equalization is used for pre-processing of the images. Image quality and accuracy is the core factors of this research, image quality assessment as well as enhancement stage were adopted on low pre-processing techniques based on Histogram Equalization. The covariance matrix is difficult to be evaluated in an accurate manner which would be a disadvantage of principal component analysis.

[4] Lung Cancer Data Analysis by K-Means and Farthest First Clustering Algorithms.

The very famous two partition based algorithms namely k-Means and Farthest First are implemented. A comparative analysis of clustering algorithms is also carried out using two different datasets. The k-Means algorithm is widely used in several studies for grouping data. It may produce more uniform grouping between one cluster to another. The performance of the partitioning based algorithms were analysed using the only selected three attributes from the total number of attributes of input dataset.

[5] K-Means Clustering Using Fuzzy C-Means Based Image Segmentation For Lung Cancer.

This review gives another way to deal with K-implies bunching procedure (K-CT) coordinated with Fuzzy C-implies calculation for lung division. It is trailed by limit and level set division stages to give a precise district developing recognition. This technique can get benefits of the K-means clustering for lung lesion segmentation in the aspects of minimal computation time. In addition, it can get advantages of the Fuzzy C-means in the aspects of accuracy. The k-means clustering technique representation can include information about the background in order to improve image classification results.

[6] A Novel Approach for Lung Cancer Detection Using Morphological Operations.

Using image processing techniques like adaptive k-means clustering and by using the Fuzzy c-means threshold for segmentation, feature extraction, area of interest is separated. In this paper an adaptive k-means clustering has been used which clusters the pixels very accurately. All possible locations of the neighbour are strictly depending on the grid type. As a consequence, not all directions can be taking under consideration - only these which describe location of neighbours in given grid.

[7] Adaptive Detection of Pulmonary Nodules in Ct Images by Segmentation and Classification Using Matlab.

In this paper, a new decision based technique called new adaptive median filter is presented which shows better performance than those already being used. In this work, advanced classification techniques based on Least Squares Support Vector Machines (LS-SVM) are proposed and applied to CT image slices classification using features derived from slices. The limitation in using Least Squares support machines is speed and size, both in training and testing.

[8] Lung Cancer Detection Using Neural Networks.

This research focuses upon image quality and accuracy. There has been done a classification based upon artificial neural networks which is more satisfactory than other current classification methods. Best

segmented image is found out on the basis of the entropy value. Image with higher entropy values is the best image which gives more information. Such images are with higher entropy values are much efficient for further processing. "Black box" nature, greater computational burden, proneness to over fitting, and the empirical nature of model development.

[9] Detection of Lung Cancer Using Back Propagation Neural Networks and Genetic Algorithm.

A powerful learning model that is (BPNN) is used for classification which would classify the digital X A ray-images, MRI's, etc. as cancerous or non-cancerous. Further Genetic Algorithm will be used that would extract feature on the basis of the fitness function. Whenever you deal with huge amounts of data and you want to solve a supervised learning task with a feed-forward neural network, solutions based on back propagation are much more feasible. The reason for this is that for a complex neural network, the number of free parameters is very high. The process is time-consuming and involves operations of high complexity. BPTT has difficulty with local optima.

[10] Lung Cancer Diagnosis By using Fuzzy Logic.

In this paper, there is proposed a diagnosis system to detect lung cancer based on fuzzy logic and neural network, there has been used neural network to classify the normal and abnormal images, in the abnormal result, use other parameters (symptoms) as input to fuzzy logic system to find the case of the patient (affected or not) depending on the membership function of inputs. Allows you to model in a more intuitive way complex dynamic system. Expanding rough approximations into fuzzy environment which help to obtain solutions for various real time problems. Lots of optimization problems don't have very critical solutions. In other words, tuning of system parameters will in general only slightly increase the system performance.

B. Methodology Used

- Histogram exploit is employed to boost distinction.
- Segmentation method is predicated on principal pattern analysis (PPA) formula.

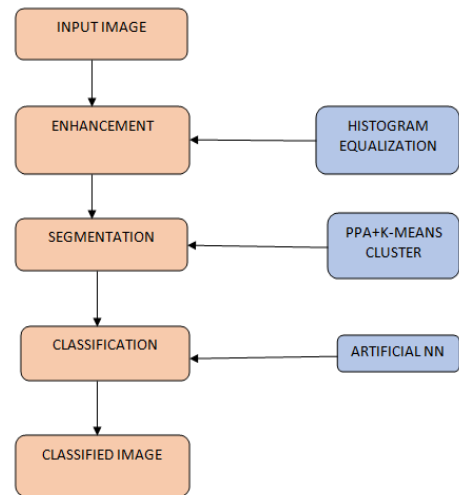


Figure 1. Process Block Diagram

- The work partially implements foggy k-means algorithm and then employs the principal pattern analysis algorithm, consequently evaluating the feature patterns.
- Neural network is used for Classification.

C. Histogram Equalization

- Histogram effort is employed to boost distinction. It's not necessary that distinction can perpetually be increase during this.
- There could also be some cases were bar chart effort will be worse. In those cases the distinction is shrunken.
- Bar chart technique sometimes will increase the world distinction of the many pictures, particularly once the usable information of the image is depicted by shut distinction values.
- Through this adjustment, the intensities will be higher distributed on the bar chart.
- This permits for areas of lower native distinction to realize a better distinction. Bar chart effort accomplishes this by effectively spreading out the foremost frequent intensity values.
- The technique is beneficial in pictures with backgrounds and foregrounds that square measure each bright or each dark.
- In explicit, the strategy will cause higher views of bone structure in x-ray pictures, and to higher detail in images that square measure over or under-exposed.
- A key advantage of the strategy is that it's a reasonably easy technique associated an invertible operator.

D. Segmentation with PPA And Foggy K-Means

- Segmentation is that the most vital half in image process.
- Fence off a whole image into many components that are some things additional meaning and easier for any method.
- These many components that area unit re-joined can cowl the complete image.
- Segmentation may additionally rely on numerous options that area unit contained within the image.
- It could also be either colour or texture.
- The most expression of segmentation is to cut back the knowledge for simple analysis.
- Segmentation is additionally helpful in Image Analysis and compression.

E. Foggy K Means Algorithm

- In medical data processing, it's helpful to urge domain information from domain subject specialists.
- Here, additionally in Foggy k means that cluster of carcinoma dataset has been mentioned with domain specialists and sure attributes with the outstanding impact issue has been known.
- Number of the clusters has been selected the premise of the worth of those attributes.
- For examples within the carcinoma dataset if the tumour size is larger than three then there's an occasion that the patient is stricken by carcinoma.
- On the premise of those a forethought points2 clusters square measure shaped and take 2 points as centre of mass.
- The result of latest attributes on the cluster is represented as- suppose next attribute is smoking, then the cluster can move either left or right to a selected distance in step with the impact of that attribute on the cluster.

F. Classifier Artificial Neural Network (ANN)

- ANN Classification is the process of studying to separate samples into different classes by finding common features between samples of known classes.
- Artificial neural networks are almost crude electronic networks of neurons based on the neural structure of the brain.

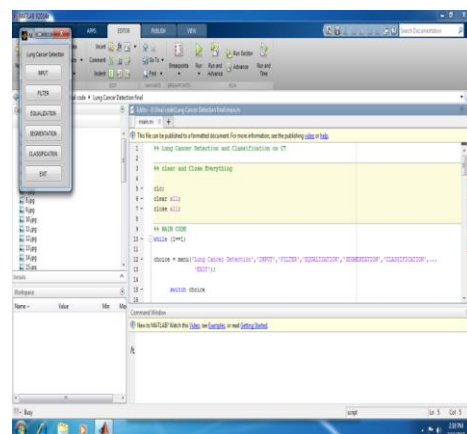
- They process records one at a time, and learn by comparing their classification of the record (i.e., largely arbitrary) with the known actual classification of the record.
- Neural networks are typically organized in layers. Layers are made up of a number of interconnected 'nodes' which contain an 'activation function'.
- Patterns are conferred to the network via the 'input layer', which communicates to one or more 'hidden layers' where the actual processing is done via a system of weighted 'connections'.
- The hidden layers then bond to an 'output layer'.

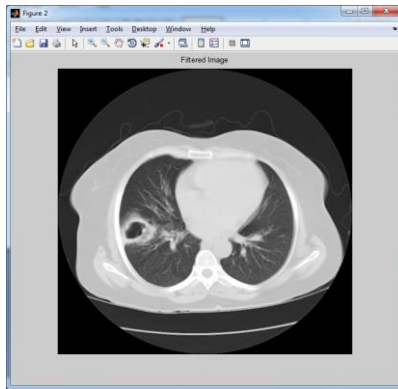
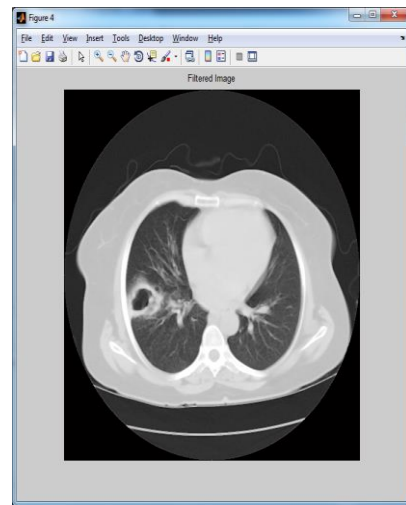
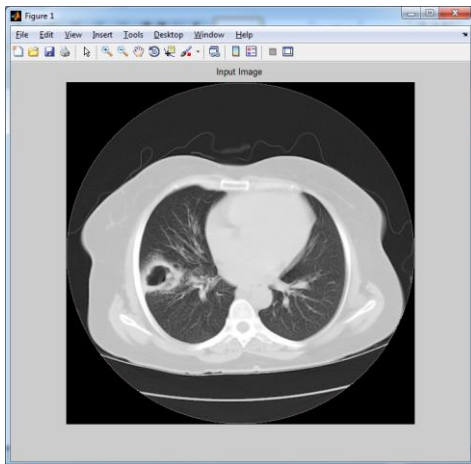
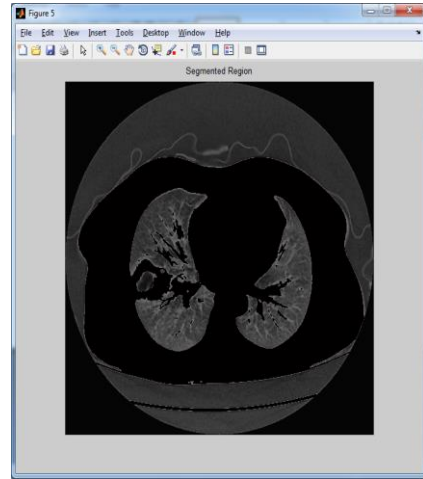
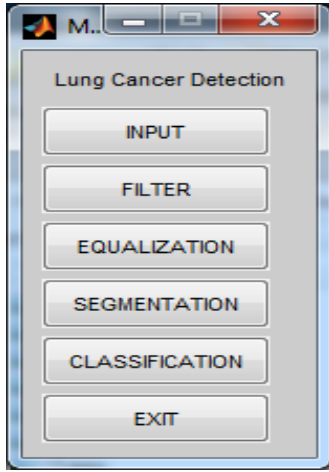
III. RESULTS AND DISCUSSION

A. Principal Pattern Analysis (PPA) Algorithm

- Still there is an ever spreading needs for techniques related to the dimensionality reduction and classification.
- A novel algorithm called Principal Pattern Analysis algorithm (PPA) is presented in our proposed work.
- The work partially implements k-means algorithm and then employs the principal pattern analysis algorithm, consequently evaluating the feature patterns.
- The weight with which each principal-pattern commit to the intensity-pattern can be represented on a set of orthogonal axes that span a previously introduced pattern space.

B. Output





Neural Network Training (nntraintool)

Neural Network

Input: 1, Hidden: 12, Output: 1

Algorithms

Data Division: Random (divisorand)
 Training: Scaled Conjugate Gradient (trainscg)
 Performance: Cross-Entropy (crossentropy)
 Calculations: MEX

Progress

Epoch:	0	10 iterations	1000
Time:		0:00:02	
Performance:	0.956	0.504	0.00
Gradient:	0.764	0.0851	1.00e-06
Validation Checks:	0	6	6

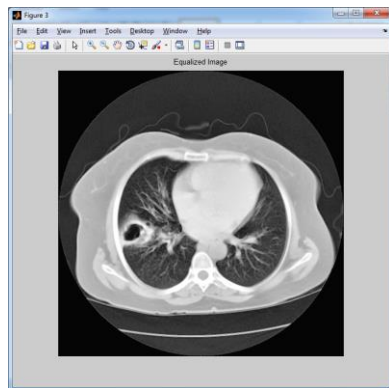
Plots

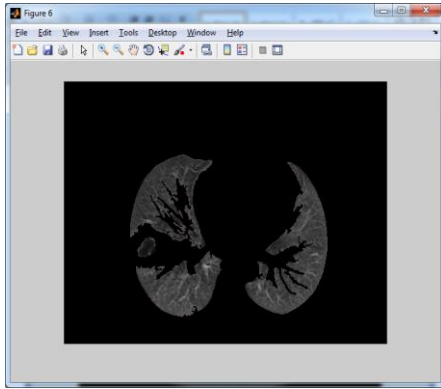
Performance (plotperform)
 Training State (plottrainstate)
 Error Histogram (ploterrhist)
 Confusion (plotconfusion)
 Receiver Operating Characteristic (plotroc)

Plot Interval: 1 epochs

Validation stop.

Stop Training Cancel





These are the screenshots obtained during the execution of the modules of our proposed system.

IV. CONCLUSION

Hence, a lung cancer detection system using image processing is used to classify the existence of lung cancer in CT- images. In this work, MATLAB has been used through every procedure. This involves image pre-processing, segmentation based on PPA, Foggy K-Means clustering and feature extraction by neural network. This paper concludes to provide the detection of lung cancer cells using Foggy k-means clustering algorithm on the reduced images of Principal Pattern Analysis algorithm. This has shown to provide optimum results. The objective of the proposal is to materialize a means to fasten the process as well as the accuracy of detecting the damaged cells to a valuable extent it helps in saving human lives.

V. REFERENCES

- [1] "Feature Extraction and Principal Component Analysis for Lung Cancer Detection in CT scan Images" by Ada, RajneetKaur, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 3, March 2013.
- [2] "A Survey on Early Detection and Prediction of Lung Cancer" by NehaPanpaliya, NehaTadas, SurabhiBobade, RewtiAglawe, AkshayGudadhe, International Journal of Computer Science and Mobile Computing, IJCSMC, Vol. 4, Issue. 1, January 2015, pg.175 – 184.
- [3] "Prediction of Lung Cancer Using Image Processing Techniques: A Review" by ArvindKumar Tiwari, Advanced Computational Intelligence: An International Journal (ACIJ), Vol.3, No.1, January 2016.
- [4] "Clustering of Lung Cancer data using Foggy Kmeans" by DivyaTomar, SonaliAgarwal, 2013 International Conference on Recent Trends in Information Technology (ICRTIT).
- [5] "Clustering of Lung Cancer Data Using Foggy K-Means" by Akhilesh Kumar Yada, DivyaToma, SonaliAgarwal, 2013 International Conference on Recent Trends in Information Technology (ICRTIT).
- [6] "Lung Cancer Detection by Using Artificial Neural Network and Fuzzy Clustering Methods "In 2011 by Rashmita Sehgal, Saurabh Gupta
- [7] "Clustering of Lung Cancer Data Using Foggy K-means" In 2013 By Akhilesh Kumar Yadav, DivyaTomar, Sonali Agarwal.
- [8] "Feature Extraction and Principal Component Analysis for Lung Cancer Detection in Ct Scan Images"-march 2013 By Ada, RajneetKaur
- [9] "Lung Cancer Data Analysis By K-means And Farthest First Clustering Algorithms" In 2015 By A. Dharmarajan and T. Velmurugan.
- [10] "K-means Clustering Using Fuzzy C-means Based Image Segmentation for Lung Cancer" In 2016 By K. Kaviarasu , V. Sakthivel.
- [11] "A Novel Approach for Lung Cancer Detection Using Morphological Operations" In 2016 by Gangireddy Hemanth Kumar, Dr. G. Umamaheswara Reddy.