IFERP

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 2, Issue 8, August 2017

Lung Cancer Detection using Log-Gabor Filter Banks

[1] Rupali Mali,[2] Dr. S. B. Bagal
[1] M.E. Student, [2] Principal
[1][2] LGNSCOE, Nashik

Abstract - **Lung cancer is the foremost reason of deaths due to cancer disease around the world. As per the report of the World Health Organization (WHO) near about 10 million patients in the world will be deceased because of lung cancer by 2030. Timely inhibition of lung tumor plays an imperative role for survival assistance enhancements. By following the notion that heavy investigation of radiographic imageries can apprise and enumerate the microenvironment and the degree of tumor level heterogeneity for personalized medicine, examination of huge numbers of image features extracted from computed tomography (CT). The focus is on high throughput that can apprehend spatial and temporal genetic heterogeneity in a without operating the patient, which preferred over intrinsic biopsy based molecular assays method. The lung cancer detection is valuable for ongoing medical research and computer-assisted diagnosis of lung cancer. In this paper, we have presented lung cancer detection algorithm that yields possible location of tumor in the lung. The algorithmic steps comprises of histogram equalization of the CT scan image followed by log- Gabor filter bank processing to enhance the CT scan image. Subsequently the image is dilated using gradient mask and after border clearing, the location of possible tumor is detected. This algorithm gives accurate results on publically shared CT scan images.**

Key words and Phrases: **Cancer Detection, Lung cancer, Log-Gabor filters, image morphology.**

1 INTRODUCTION

Lung growth which is known as disease of the bronchus is portrayed by improper cell development in tissues of lungs, much of the time because of the cancer-causing specialists. This growth is a significant supplier of enlarged tumor deprivations prompting sudden changes in human life. The succeeding cells won't develop into sound ones; they split up to shape tumors which happens to be the fundamental driver of death from growths. All circumstances this anomaly is a hereditary issue, yet nowadays it is even found in ordinary sound gathering of individuals. This is primarily because of stationary ways of life and dietary patterns, also supposed created demeanor of the general population who take precluded things for their excitement and extravagance. This class of growth turns out to be more perilous prompting aspiratory illnesses amid its course and makes the life of a patient a bad dream. Survival from lung disease is straightforwardly identified with its development and its identification time. The prior the identification is, the higher the odds of effective treatment are. When it is recognized the patient needs to experience a legitimate treatment and thorough excruciating restorative methodology by maintaining a strategic distance from the bungles they did to them with their propensities.

Screening for lung tumor has grown up when conventional ways to deal with mass screening are being tested.

Discussions about growth screening progressively happen with regards to customized pharmaceutical, individualized hazard appraisal, and shared basic leadership. Future screening techniques are joining the utilization of genomic advancements to advance early recognition of lung malignancy. In 2016, right around 13% of all disease analyze were caused by this danger, with over half happening in less industrialized regions of the world. Men keep on having the higher rate, particularly in North America, Europe, Eastern Asia, and Uruguay, with the most reduced rates in sub-Saharan Africa. Like guys, the lung malignancy rates in ladies are higher in North American and Europe, however the nations of Australia, New Zealand, North Korea, and China likewise have a higher general rate. This is accepted to reflect indoor air contamination from coal-energized stoves and cooking vapor that are unvented.

Multiple factors have been identified to cause or are associated with the development of lung cancer. Globally, smoking is identified as the primary cause of lung cancer.

A. Smoking

Dynamic smoking has the best impact on the improvement of lung disease. Inward breath of smoke from tobacco items builds the danger of lung growth, including cigarettes, funnels, stogies, and introduction to detached smoking. Indeed, even the utilization of e-cigarettes has a component

IFERP

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 2, Issue 8, August 2017

of hazard. Research demonstrates that inward breath of smoke can cause transformations in oncogenes and lost heterozygosis of tumor silencer qualities which turn into the "drivers" of carcinogenesis in sporadic (substantial) lung diseases. A traditional research distinguished that as the quantity of cigarettes smoked expanded, the more noteworthy the quantity of TP53 transformations in patients with lung disease. Different analysts contemplating lung tumor patients from an assortment of ethnic populaces have discovered changes in KRAS (indicator of visualization and treatment resistance in sporadic lung growth) and different qualities on 15q25. Detached smoking presentation can be checked by testing for the essential metabolite of nicotine, cotinine, in the salivation, pee, or blood. Since 2008, introduction to uninvolved smoke (otherwise called ecological tobacco smoke or second-hand smoke) has declined. The 2007–2008 National Health and Nutrition Examination Survey recognized that non-Hispanic blacks, youngsters, in addition to people living beneath the government destitution level or tenants were at most noteworthy hazard for introduction to ecological tobacco smoke.8 Second-hand smoke expands the danger of lung tumor in non-smokers by no less than 20%, with a proposal that natural tobacco smoke is expanding not diminishing. Lung disease chance increments by the measure of smoking. On the off chance that dynamic or uninvolved smoking is ceased, the danger of growth diminishes.

B. Genetics

A family history of lung growth places people at expanded hazard for the future advancement of lung malignancy. The hazard raises if a relative is determined to have lung growth at a youthful age or if numerous relatives have a background marked by this illness. In 2008, an inclusive affiliation contemplate emphatically connected the danger of acquired lung growth helplessness to a solitary nucleotide polymorphism on chromosome 15q25. Later named as CHRNA5 this single-nucleotide polymorphism was associated with the before advancement of lung malignancy by 4 years. A little report recognized the uncommon germ line change of EGFR, T790M, is related with acquired powerlessness to lung growth, particularly in nonsmoking females. Truth be told, ladies that convey this genotype and have never smoked had a 31% danger of creating lung disease; higher than substantial smokers with or without the transformation.

C. Lung Diseases and Infections

Regardless of whether an individual is a smoker or nonsmoker, endless obstructive pneumonic malady (COPD) builds the danger of lung tumor. Another COPD lung tumor screening apparatus, the COPDLUCSS, was created to encourage expectation of lung malignancy hazard in people who take an interest in lung disease screening programs. This apparatus has four markers: a body mass list < 25 kg/square meter, pack-years history > 60, age more seasoned than 60, and a finding of emphysema. Focuses are granted to every marker met with 1 point for the body mass record, 2 focuses for the expanded pack-year history, 3 focuses for being more seasoned than 60, and 4 focuses for a radiologic conclusion of emphysema. Every one of "no" responses to the markers get 0 focuses. The aggregate focuses ascertained is the COPD lung tumor screening score (COPD-LUCSS). The okay classification is 0 to 6 focuses and high hazard is 7 to 10 focuses. The long haul objective for utilizing this scale is simple distinguishing proof of patients with the most noteworthy hazard for lung tumor, to advance early location and affect the high death rate of lung malignancy in patients with COPD.

D. Carcinogens

There are numerous cancer-causing agents that have an immediate relationship with lung malignancy. Among individuals who have contact with these operators, the danger of lung disease is higher for the individuals who have smoked than the individuals who have never smoked. More than 7000 chemicals are in tobacco smoke, with more than 70 named as cancer-causing agents. Some incorporate cyanide, benzene, formaldehyde, acetylene (fuel for wood lights), smelling salts, arsenic, vinyl chloride, cadmium, and nitrosamines. Contingent upon nature in which the tobacco leaves are developed, it is conceivable that radioactive materials, for example, radon, can be invested in the lungs through inward breath. After some time, this could cause a huge measurement of radiation.

DETECTION OF LUNG CANCER

The early recognition of lung growth has been a troublesome objective to accomplish for quite a long time. Different techniques have been utilized to recognize lung malignancy with constrained effect on related passings after analysis. In 2002, the National Lung Screening Trial was propelled to think about low-dosage helical processed tomography (LDCT) with trunk x-beam (CXR). Members were haphazardly doled out to routinely get LDCT or CXR screenings over a 3-year time frame. Results from the National Lung Screening Trial uncovered a 20% relative lessening in mortality from lung malignancy with three rounds of LDCT screening (rounds T0, T1, and T2) as contrasted and CXR.

A. Sputum

Cytologic evaluation of sputum to detect lung cancer is now rarely used. Although LDCT is currently the preferred method, researchers have identified that it may be ideal to screen patients' sputum via real time-polymerase chain reaction before LDCT. With this method, miRNA biomarkers can be used to identify circulating or expired malignant cells, metabolites associated with specific types of lung cancer, and diagnose solitary pulmonary nodules. This will improve the issue with LDCT falsely diagnosing patients with lung cancer.

B. CXR

Research studies in the 1990s compared LDCT and CXR methods for feasibility of detecting lung cancer and found that CXR did not provide the detailed images available with computed tomography (CT). Images are taken from the anterior posterior, lateral, and posterior anterior positions. CXR does emit radiation, though not at a high level compared with other types of radiological procedures. According to the American Nuclear Society, the average annual dose of radiation from all sources is 620 millirems (mrems) per year. CXR is associated with 10 mrems, a mammogram emits 40 mrems, and half pack of cigarettes every day for 1 year is 18 mrems. A chest CT emits 700 mrems.

C. CT

CT scans continue to be used for detecting both acute and chronic changes in the lung parenchyma. With this modality, multiple x-ray images are taken from different angles to obtain rotating views of the tissue. A computer is used to create slices of image to view the lung tissue in this example. It is particularly relevant because normal 2-dimensional x-rays do not show lung cancer changes unless they are associated with a more advanced stage of disease.

D. LDCT

Currently, the US Preventive Services Task Force recommends annual screening for lung cancer with LDCT for asymptomatic individuals between the ages of 55 and 80 years who have a 30-pack year smoking history and currently smoke or have quit smoking within the past 15 years, extending the upper limit of the 55 to 74 years age based on criteria applied in the National Lung Screening Trial. Screening can be stopped for a person who has not smoked for 15 years, acquired a health problem causing life expectancy limitations, or the inability to have curative lung surgery. LDCT screening is covered as a preventive services benefit for Medicare beneficiaries between the ages of 55 and 77 years who meet the US Preventive Services Task Force criteria for tobacco smoking history. Screening should target those who are at highest risk to develop lung cancer. See Figure 1 for a comparison of CXR versus LDCT in the early detection of lung cancer.

E. PET/CT

The PET (positron emission tomography) is a nuclear medicine, functional imaging technique used to observe active metabolic processes in the body. The modern PET/CT scanner takes 3-dimensional images of tracer concentration within the body for analysis by computer. If the biological molecule used for PET imaging is an analogue for glucose, the concentration of tracer imaged will show enhanced tissue metabolic activity related to the regional uptake of glucose. The PET and CT scans are commonly used together to discriminate between tissue and abnormal metabolic activity. While these are important tools in the detection and monitoring of cancer, together they emit a significant amount of radiation. Research has demonstrated there is a risk with the development of cancer.

LITERATURE SURVEY

T. Sowmiya, M. Gopi, M. New Begin, L.Thomas Robinson [1] - In this paper they described Cancer as the most dangerous diseases in the world. Lung cancer is one of the most dangerous cancer types in the world. These diseases can spread worldwide by uncontrolled cell growth in the tissues of the lung. Early detection of the cancer can save the life and survivability of the patients who affected by this diseases. In this paper we survey several aspects of data mining procedures which are used for lung cancer prediction for the patients. Data mining concepts is useful in lung cancer classification. We also reviewed the aspects of ant colony optimization (ACO) technique in data mining. Ant colony optimization helps in increasing or decreasing the disease prediction value of the diseases. This case study assorted data mining and ant colony optimization techniques for appropriate rule generation and classifications on diseases, which pilot to exact Lung cancer classifications. In additionally to, it provides basic framework for further improvement in medical diagnosis on lung cancer.

Ada, Rajneet Kaur [2] - In this paper uses a computational procedure that sort the images into groups according to their similarities. In this paper Histogram Equalization is used for preprocessing of the images and feature extraction process and neural network classifier to check the state of a patient in its early stage whether it is normal or abnormal. After that we predict the survival rate of a patient by extracted features. Experimental analysis is made with dataset to evaluate the performance of the different classifiers. The performance is based on the correct and incorrect classification of the classifier. In this paper Neural Network Algorithm is implemented using open source and its performance is compared to other classification algorithms. It shows the best results with highest TP Rate and lowest FP Rate and in case of correctly classification, it gives the 96.04% result as compare to other classifiers. The second paper of this same author is based on Feature Extraction and Principal Component Analysis for Lung Cancer Detection in CT scan Images. In this paper uses a hybrid technique based on feature extraction and Principal Component Analysis (PCA).

Dasu Vaman Ravi Prasad [3] - In this paper image quality and accuracy is the core factors of this research, image quality assessment as well as improvement are depending on the enhancement stage where low pre-processing techniques is used based on Gabor filter within Gaussian rules. Following the segmentation principles, an enhanced region of the object of interest that is used as a basic foundation of feature extraction is obtained. Relying on general features, a normality comparison is made. In this research, the main detected features for accurate images comparison are pixels percentage and masklabeling.

S Vishukumar K. Patela and Pavan Shrivastavab [4] - In this paper authors mostly focus on significant improvement in contrast of masses along with the suppression of background tissues is obtained by tuning the parameters of the proposed transformation function in the specified range. The manual analysis of the sputum samples is time consuming, inaccurate and requires intensive trained person to avoid diagnostic errors. The segmentation results will be used as a base for a Computer Aided Diagnosis (CAD) system for early detection of cancer, which improves the chances of survival for the patient. In this paper, authors proposed gabor filter for enhancement of medical images. It is a very good enhancement tool for medical images.

Fatma Taher et al. [5] - This paper presents two segmentation methods, Hopfield Neural Network (HNN) and a Fuzzy C-Mean (FCM) clustering algorithm, for segmenting sputum color images to detect the lung cancer in its early stages. The manual analysis of the sputum samples is time consuming, inaccurate and requires intensive trained person to avoid diagnostic errors. The segmentation results will be used as a base for a Computer Aided Diagnosis (CAD) system for early detection of lung cancer which will improve the chances of survival for the patient. However, the extreme variation in the gray level and the relative contrast among the images make the segmentation results less accurate, thus we applied a thresholding technique as a pre-processing step in all images to extract the nuclei and cytoplasm regions, because most of the quantitative procedures are based on the nuclear feature. The thresholding algorithm succeeded in extracting the nuclei and cytoplasm regions. Moreover, it succeeded in determining the best range of thresholding values. The HNN and FCM methods are designed to classify the image of N pixels among M classes. In this study, we used 1000 sputum color images to test both methods, and HNN has shown a better classification result than FCM, the HNN succeeded in extracting the nuclei and cy-toplasm regions. In this paper authors uses a rule based thresholding classifier as a pre-processing step. The thresholding classifier is succeeded in solving the problem of intensity variation and in detecting the nuclei and cytoplasm regions, it has the ability to mask all the debris cells and to determine the best rang of threshold values. Overall, the thresholding classifier has achieved a good accuracy of 98% with high value of sensitivity and specificity of 83% and 99% respectively.

MATERIALS AND METHODS

Log-Gabor filter

CT scan images typically have the property of frequency localization. Gabor filters are a traditional choice for obtaining localized frequency information. They offer the best simultaneous localization of spatial and frequency information. However, one cannot construct Gabor functions of arbitrarily wide bandwidth and still maintain a reasonably small DC component in the even-symmetric filter. Alternatively, we turn to choose the Log-Gabor filter as the filtering kernel. There are two important characteristics to note. Firstly, Log-Gabor functions, by definition, always have no DC component, which contributes to improve the contrast between the ridges and valleys of digital images. Secondly, the transfer function of the Log-Gabor function has an extended tail at the high frequency end, which enables us to obtain wide spectral information with localized spatial extent and consequently helps to preserve true ridge structures of images. Gabor functions have Gaussian transfer functions when viewed on the linear frequency scale. Similarly, Log-Gabor functions have Gaussian transfer functions when viewed on the logarithmic frequency scale. Due to the singularity in the log function at the origin, the 2D Log-Gabor filter is constructed in the frequency domain. In polar coordinates system, it can be divided into two components: the radial filter and the angular filter. The radial filter has a frequency response described by (1).

International Journal of Science, Engineering and Management (IJSEM)
Vol 2, Issue 8, August 2017

$$G_r(r) = \exp\left(-\frac{[\log(r/f_o)]^2}{2\cdot\sigma_r^2}\right) \qquad (1)$$

And the angular filter has a frequency response described by (2).

$$G_\theta(\theta) = \exp\left(-\frac{(\theta-\theta_o)^2}{2\cdot\sigma_\theta^2}\right) \qquad (2)$$

The two components are multiplied together to construct the overall Log-Gabor filter which has the transfer function as in (3).

$$G(r,\theta) = G_r(r)\cdot G_\theta(\theta) \qquad (3)$$

Here, (r,θ) represents the polar coordinates, f_0 is the center frequency of the filter, θ_0 is the orientation angle of the filter, σ_r determines the scale bandwidth and σ_θ determines the angular bandwidth. From the definition of formulas, we can see that the Log-Gabor filter is determined by four parameters: f_0, θ_0, σ_r and σ_θ. Where, f_0 and θ_0 respectively correspond to the local frequency and orientation of CT scan images, σ_r and σ_θ are specified according to the empirical data for CT scan images.

Image Histogram Equalization

The luminance histogram of a typical natural scene that has been linearly quantized is usually highly skewed toward the darker levels; a majority of the pixels possess a luminance less than the average. In such images, detail in the darker regions is often not perceptible. One means of enhancing these types of images is a technique called histogram equalization, in which the original image is rescaled so that the histogram of the enhanced image follows some desired form. The histogram of a monochrome image is a graphical representation of the frequency of occurrence of each gray level in the image. The data structure that stores the frequency values is a 1D array of numerical values, h, whose individual elements store the number (or percentage) of image pixels that correspond to each possible gray level. Each individual histogram entry can be expressed mathematically as in (4).

$$h(k) = n_k = card\{(x,y) \mid f(x,y) = k\} \qquad (4)$$

Here, $k = 0,1,\ldots,(L-1)$ where L is the number of gray levels of the digitized image, and card $\{\cdot\ \cdot\ \cdot\}$ denotes the cardinality of a set, that is, the number of elements in that set (n_k). A normalized histogram can be mathematically defined as in (5).

$$p(r_k) = \frac{n_k}{n}$$

(5)

Where, n is the total number of pixels in the image and $p(r_k)$ is the probability (percentage) of the k^{th} gray level (r_k). Histograms are normally represented using a bar chart, with one bar per gray level, in which the height of the bar is proportional to the number (or percentage) of pixels that correspond to that particular gray level.

***Mathematical Morphology**: Image Dilation*

Mathematical morphology is a branch of image processing that has been successfully used to provide tools for representing, describing, and analyzing shapes in images. It was initially developed in the early 1980s and because of its emphasis on studying the geometrical structure of the components of an image named after the branch of biology that deals with the form and structure of animals and plants. In addition to providing useful tools for extracting image components, morphological algorithms have been used for pre- or post-processing the images containing shapes of interest. The basic principle of mathematical morphology is the extraction of geometrical and topological information from an unknown set (an image) through transformations using another, well-defined, set known as structuring element. In morphological image processing, the design of SEs, their shape and size, is crucial to the success of the morphological operations that use them. The structuring element is the basic neighborhood structure associated with morphological image operations. It is usually represented as a small matrix, whose shape and size impact the results of applying a certain morphological operator to an image.

Dilation is a morphological operation whose effect is to "grow" or "thicken" objects in a binary image. The extent and direction of this thickening are controlled by the size and shape of the structuring element.

Mathematically, the dilation of a set A by B, denoted A ⊕ B, as defined in (6).

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \phi\}$$

(6)

PROPOSED APPROACH

The proposed approach is depicted in Fig. 1.

International Journal of Science, Engineering and Management (IJSEM)
Vol 2, Issue 8, August 2017

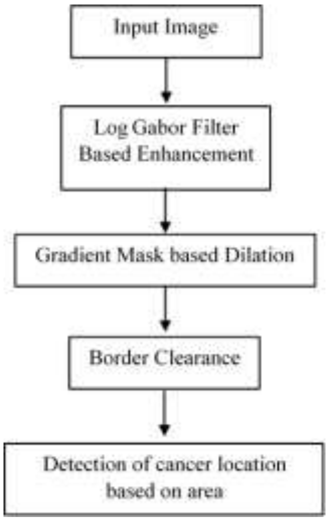

Fig.1. Proposed Approach

Initially the CT scan image of lung part of human body is given as input to lung cancer detection system. As a part of image preprocessing, image is converted into gray scale if it is in RGB color format. After converting into floating integers, histogram equalization is performed. For applying the log Gabor filter the following parameters were selected: size=25, nscale=6, norient=6, minWaveLength=3, mult=1.7, sigmaOnf= 0.65. The output image from Log Gabor filter was sifted with circular filter. Subsequently the image is converted into binary image. Two structural elements viz. 1. Line at 90 degree and 2. Line at 0 degree were defined. The combination of these two structural elements forms gradient mask with which the binary image is dilated. The redundant image border is removed and resultant image is monitored for closed area regions without any holes. The prominent area regions are marked as possible cancer regions. These steps are indicated in following figures.

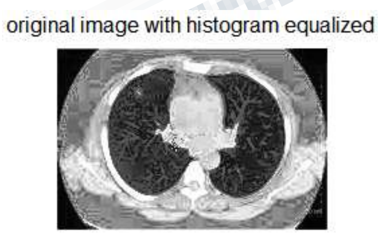

Fig.2. Original image with histogram equalization

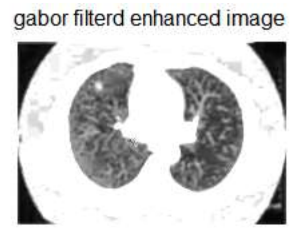

Fig.3. Image after Log-Gabor Filtering

Fig.4. Image Dilation with Gradient Mask

Fig.5. Border cleared image

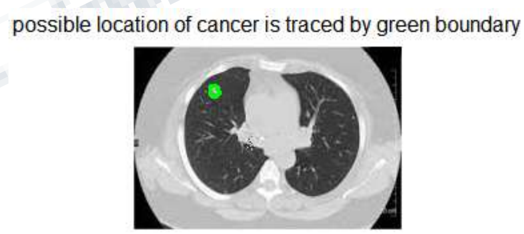

Fig.6. Image indicating detected cancer cell

CONCLUSION

Timely inhibition of lung tumor plays an imperative role for survival assistance enhancements. By following the notion that heavy investigation of radiographic imageries can apprise and enumerate the microenvironment and the degree of tumor level heterogeneity for personalized medicine, examination of huge numbers of image features extracted from computed tomography (CT). In this paper, we have presented lung cancer detection algorithm that yields possible location of tumor in the lung. The algorithmic steps comprises of histogram equalization of the CT scan image followed by log- Gabor filter bank processing to enhance the CT scan image. Subsequently the image is dilated using gradient mask and after border clearing, the location of possible tumor is detected. This algorithm gives accurate results on publically shared CT scan images.

REFERENCES

[1] Sowmiya, T., Gopi, M., New, B. M., Thomas, R. L. (2014). Optimization of lung cancer using modern data mining techniques. Int J Eng Res, 3(5), 309-14.

[2] Kaur, A. R. (2013). Feature extraction and principal component analysis for lung cancer detection in CT scan images. International Journal of Advanced Research in Computer Science and Software Engineering, 3(3).

[3] Dasu Vaman Ravi Prasad, "Lung cancer detection using image processing techniques", http://www.ijltet.org/journal_details.php?id=894&j_id=2523, Volume 3 Issue 1 - September 2013

[4] S Vishukumar K. Patela and Pavan Shrivastavab, “Lung A Cancer Classification Using Image Processing”, International Journal of Engineering and Innovative Technology Volume 2, Issue 3, September 2012.

[5] Taher, Fatma, and Rachid Sammouda. "Lung cancer detection by using artificial neural network and fuzzy clustering methods." GCC Conference and Exhibition (GCC), 2011.