Scientific Journal of Impact Factor (SJIF): 5.71

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 5, Issue 05, May -2018

COMPUTER-AIDED DIAGNOSIS: AN APPROACH FOR LUNG CANCER DETECTION

Mrs. Vidya Thorat[1], Pradnya Joshi[2], Pragya [3], Rupali Chaturvedi[4], Rasika Nimkarde[5]

1(Assistant Professor, Department of Computer Engineering, Pune/D.Y.Patil College Of Engineering Akurdi.
2(Department of Computer Engineering, Pune/D.Y.Patil College Of Engineering Akurdi.
3(Department of Computer Engineering, Pune/D.Y.Patil College Of Engineering Akurdi.
4(Department of Computer Engineering, Pune/D.Y.Patil College Of Engineering Akurdi.
5(Department of Computer Engineering, Pune/D.Y.Patil College Of Engineering Akurdi.

Abstract - *Lung cancer is the most powerful cancer for both men and women. It can often grow for long time before they are found. In order to treat the early lung cancer more effectively, it is important to develop an effective clinical marker or prediction model to more accurately predict cancer prognosis after cancer surgery. The main aim is to automate the classification process for the early detection of Lung Cancer. By using computed tomography (CT) images, a computer-aided detection scheme is used to segment lung tumors and computed tumor-related image features. The steps for lung cancer detection start with process of accepting CT Images.These CT images are further processed using algorithms followed by segmentation. The entire implementation is done in Python.*

Keywords— *Computer-aided diagnosis, Image features, Computed Tomography, Quantitative image feature analysis*

I. INTRODUCTION

Lung cancer is the second most prevalent cancer in both men and women. It also has the highest mortality rate, which accounts for more than a quarter of all cancer related deaths in the United States (i.e., 28% for male and 26% for female cancer patients). In lung cancer, over 75% cases are non-small-cell lung cancer (NSCLC) cases. Since lung is a relatively large human organ that may involve many other chronic lung diseases, lung tumors can often grow for a long time before they are found. Hence, the majority of lung cancers are currently detected and diagnosed at the advanced stages with a lower survival rate. Therefore, great research and promotion effort has been made to implement lung cancer screening programs using the low-dose computed tomography (CT) imaging modality for the last decade. In July 29, 2013, the U.S. Preventive Services Task Force issued a draft recommendation in favor of lung cancer screening for long-term smokers using low-dose CT tests. Meanwhile, the trend of using CT for screening and/or detecting other lung diseases (e.g., chronic obstructive pulmonary diseases) is also on the rise. As a result, more early stage lung cancers (i.e., stage I NSCLC) are detected during the regular lung cancer screening or other incident findings.

II. RELATED WORK

Paper Name: Cancer Statistics, 2015 [1]

Each year the American Cancer Society estimates the numbers of new cancer cases and deaths that will occur in the United States in the current year and compiles the most recent data on cancer incidence, mortality, and survival. Incidence data were collected by the National Cancer Institute (Surveillance, Epidemiology, and End Results [SEER] Program), the Centers for Disease Control and Prevention (National Program of Cancer Registries), and the North American Association of Central Cancer Registries. Mortality data were collected by the National Center for Health Statistics. A total of 1,658,370 new cancer cases and 589,430 cancer deaths are projected to occur in the United States in 2015.

Paper Name: Screening for lung cancer with low-dose computed tomography: a systematic review to update the U.S. Preventive Services Task Force Recommendation. [2]

Lung cancer is the leading cause of cancer-related death in the United States. Because early-stage lung cancer is associated with lower mortality than late-stage disease, early detection and treatment may be beneficial.To update the 2004 review of screening for lung cancer for the U.S. Preventive Services Task Force, focusing on screening with low-dose computed tomography (LDCT). MEDLINE (2000 to 31 May 2013), the Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews (through the fourth quarter of 2012), Scopus, and reference lists. English-language randomized, controlled trials or cohort studies that evaluated LDCT screening for lung cancer. One reviewer extracted study data about participants, design, analysis, follow-up, and results, and a second reviewer checked extractions.

Paper Name: The importance of lung cancer screening with low-dose computed tomography for Medicare beneficiaries. [3]
The National Lung Screening Trial has provided convincing evidence of a substantial mortality benefit of lung cancer screening with low-dose computed tomography (CT) for current and former smokers at high risk. The United States Preventive Services Task Force has recommended screening, triggering coverage of low-dose CT by private health insurers under provisions of the Adorable Care Act. The Centers for Medicare Medicaid Services (CMS) are currently evaluating coverage of lung cancer screening for Medicare beneficiaries. Since 70 percent of lung cancer occurs in patients 65 years or older, CMS should cover low-dose CT, thus avoiding the situation of at-risk patients being screened up to age 64 through private insurers and then abruptly ceasing screening at exactly the ages when their risk for developing lung cancer is increasing.
Paper Name: Lung Cancer Screening. [4]
Advances in imaging technology have ushered in a new era for lung cancer screening in high-risk individuals using computed topographic (CT) scans. Although most published studies are nonrandomized observational cohorts of volunteers, the ability of CT scans to detect early stage lung cancer is undisputable. What is unresolved is the ability of spiral CT screening to affect lung cancer-related mortality. A large randomized trial sponsored by the National Cancer Institute to address this question is currently under way. Genomic and proteomic approaches promise to complement the ability of spiral CT to detect early lung cancer in the next few years.
Paper Name: ERCC1 and RRM1 in the international adjuvant lung trial by automated quantitative in situ analysis. [5]
The excision repair cross completing group 1 gene product (ERCC1) and the regulatory subunit of ribonucleotide reductase (RRM1) have been reported as being prognostic of outcome and predictive of therapeutic in patients with non–small cell lung cancer. Routinely processed surgical specimens from 784 patients from the International Adjuvant Lung Trial were arrayed as tissue microarrays. In situ protein levels were scored with an automated, quantitative analysis system, dichotomized into high and low marker categories, and analyzed for associations with patients' characteristics, survival, and benefit from adjuvant chemotherapy. Scores for both markers were significantly associated with contributing center (P ¡ 0.001) and skewed, with the bulk of scores being low. High scores were more frequent in women for ERCC1 and RRM1 and in older patients and those with adenocarcinoma for RRM1.

III. PROPOSED SYSTEM

Medical data processing is one in all the main problems during this contemporary world. Medical issues square measure usually in every and each creature. Cancer is one in all the foremost dangerous diseases a person's will ever had. Carcinoma is one in all them. Carcinoma may be a un wellness that happens thanks to the uncontrolled cell growth in tissues of the respiratory organ. It's terribly troublesome to discover it in its early stages as its symptoms seem solely within the advanced stages. We'll use some techniques are essential to the task of medical image mining, respiratory organ Field Segmentation, processing, Feature Extraction, Classification mistreatment neural network. The strategies employed in this paper work states to classify digital X-ray chest films into two categories: traditional and abnormal. Totally different learning experiments were performed on two different information sets, created by suggests that of feature choice trained with completely different parameters; the results square measure compared and according. Mistreatment CT pictures, we have a tendency to developed a computer-aided detection theme to section respiratory organ tumors and computed tumor-related image options.

A. Advantages Of Proposed System

1) Instant detection of lung cancer.
2) Required time is reduced
3) Reduction of paper work

IV. ARCHITECTURE OVERVIEW

A system architecture or systems architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system.

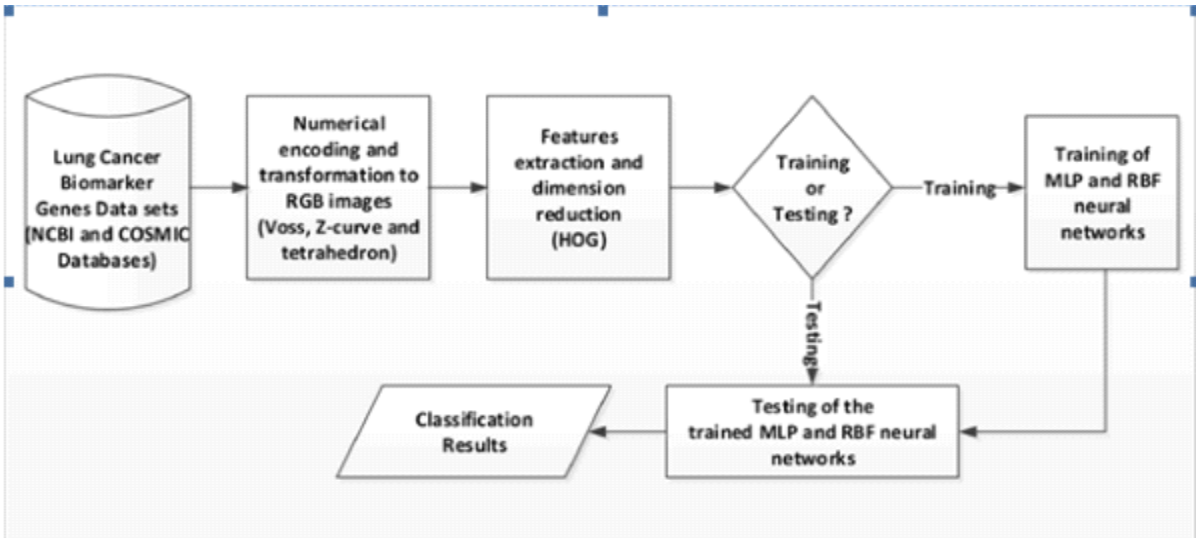


Figure: System Architecture

V. IMPLEMENTATION DETAILS

A. INTRODUCTION:

Recently, image processing techniques has been widely used in several medical areas, clinic and Hospitals for image improvement in earlier detection and treatment stages, where the time factor is very important to discover the abnormality issues and level of cancer in target images, especially in various cancer tumors such as lung cancer, breast cancer, etc. Image quality and accuracy and time taken to do the process of finding are 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. The segmentation principle is an enhanced region of the object of interest that can be used as a basic foundation of feature extraction. Relying on general features, a normality comparison is made between the images. The main detected features for accurate images comparison are pixels percentage and mask-labeling. The Lung cancer often spreads toward the centre of the chest because of the natural flow of lymph is out of the lungs and toward the centre of the chest. Metastasis occurs when a cancer cell leaves the site where **it** began to grow and moves into a lymph node or to another part of the body through the blood stream. Cancer that begins in the lung is called primary lung cancer. There are several different types of lung cancer, and these are divided into two main groups: Small cell lung cancer and non-small cell lung cancer which has three subtypes: Carcinoma, Adenocarcinoma and Squamous cell carcinomas.

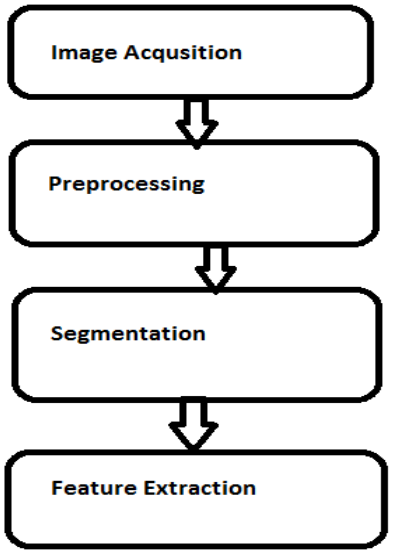


Figure : Lung Cancer Image Processing Stages

B. LIBRARIES USED

a. NumPy

NumPy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays. The ancestor of NumPy, Numeric, was originally created by Jim Hugunin with contributions from several other developers. In 2005, Travis Oliphant created NumPy by incorporating features of the competing Numarray into Numeric, with extensive modifications. NumPy is open-source software and has many contributors.

b. OpenCV

OpenCV-Python is a library of Python bindings designed to solve computer vision problems.
Python is a general purpose programming language started by Guido van Rossum that became very popular very quickly, mainly because of its simplicity and code readability. It enables the programmer to express ideas in fewer lines of code without reducing readability.
Compared to languages like C/C++, Python is slower. That said, Python can be easily extended with C/C++, which allows us to write computationally intensive code in C/C++ and create Python wrappers that can be used as Python modules. This gives us two advantages: first, the code is as fast as the original C/C++ code (since it is the actual C++ code working in background) and second, it easier to code in Python than C/C++. OpenCV-Python is a Python wrapper for the original OpenCV C++ implementation.
OpenCV-Python makes use of Numpy, which is a highly optimized library for numerical operations with a MATLAB-style syntax. All the OpenCV array structures are converted to and from Numpy arrays. This also makes it easier to integrate with other libraries that use Numpy such as SciPy and Matplotlib.

C. ALGORITHMS USED

a. MEDIAN FILTER

- Median filtering is a nonlinear method used to remove noise from images. It is widely used as it is very effective at removing noise while preserving edges. It is particularly effective at removing ‘salt and pepper’ type noise.
- The median filter works by moving through the image pixel by pixel, replacing each value with the median value of neighboring pixels. The pattern of neighbors is called the ”window”, which slides, pixel by pixel over the entire image 2 pixel, over the entire image.
- The median is calculated by first sorting all the pixel values from the window into numerical order, and then replacing the pixel being considered with the middle (median) pixel value.

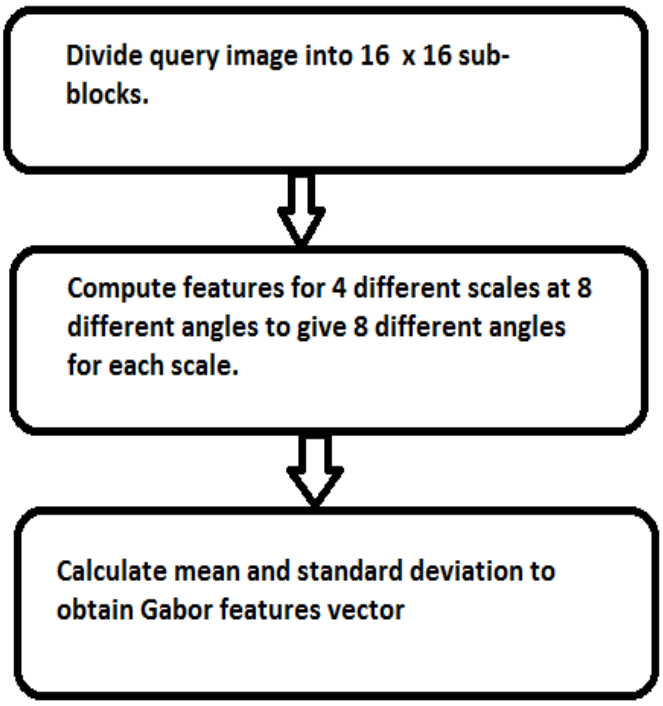
b. OTSU FILTER

- Thresholding is an important technique in image segmentation applications.
- The basic idea of thresholding is to select an optimal gray-level threshold value for separating objects of interest in an image from the background based on their gray-level distribution. While humans can easily differentiable an object from complex background and image thresholding is a difficult task to separate them.
- The gray-level histogram of an image is usually considered as efficient tools for development of image thresholding algorithms. Thresholding creates binary images from grey-level ones by turning all pixels below some threshold to zero and all pixels about that threshold to one.
- If g(x, y) is a threshold version of f(x, y) at some global threshold T, it can be defined as [1],
- g(x, y) = 1 if f(x, y) T = 0 otherwise Thresholding operation is defined.
- T = M [x, y, p(x, y), f (x, y)] In this equation, T stands for the threshold.
- f (x, y) is the gray value of point (x, y) and p(x, y) denotes some local property of the point such as the average gray value of the neighborhood centered on point (x, y) Based on this, there are two types of thresholding methods.
- Global thresholding: When T depends only on f (x, y) (in other words, only on gray-level values) and the value of T solely relates to the character of pixels, this thresholding technique is called global thresholding.

- Local thresholding: If threshold T depends on f (x, y) and p(x, y), this thresholding is called local thresholding. This method divides an original image into several sub regions, and chooses various thresholds T for each sub region reasonably.
- Otsu method is type of global thresholding in which it depend only gray value of the image. Otsu method was proposed by Scholar Otsu in 1979. Otsu method is global thresholding selection method, which is widely used because it is simple and effective.

c. GABOR FILTER

Gabor Filter is a linear filter used for texture analysis, which means that it basically analyses whether there are any specific frequency content in the image in specific directions in a localized region around the point or region of analysis. Frequency and orientation representations of Gabor filters are claimed by many contemporary vision scientists to be similar to those of the human visual system, though there is no empirical evidence and no functional rationale to support the idea. They have been found to be particularly appropriate for texture representation and discrimination. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave.



VI. CONCLUSION AND FUTURE SCOPE

A. Conclusion

- A quantitative image feature-based classifier yielded significantly higher discriminatory power than a genomic biomarker- based classifier in predicting cancer recurrence risk. Fusion of prediction scores generated by two classifiers further improved prediction performance.
- Developing precision medicine or a more effective personalized strategy for treating and managing Stage 1 NSCLC patients require a more accurate clinical marker and/or assessment tool to predict cancer prognosis.

B. Future Scope

- NSCLC is the lung cancer which has no remedies so it is important to detect it in its early stages. Doctors sometimes may go wrong on prediction about lung cancer. So, it is necessary that we design a system that can help to detect the cancer.
- In future, this project can help us in detetection of NSCLC lung Cancer and in addition with detection of its stages.

VII. REFERENCES

1. R. Siegel, et al., ”Cancer statistics, 2015,” CA Cancer J Clin., vol. 65, pp.5-29, Jan/Feb., 2015.
2. L. Humphrey, et al.,Screening for lung cancer with low-dose computed tomography: a systematic review to update the U.S. Preventive Services Task Force Recommendation”, Ann. Intern. Med., vol. 159, no. 6, pp. 411-420, Sep. 2013.
3. American Cancer Society, (2014, Nov.).Non-small cell lung cancer survival rates by stage.http://www.cancer.org/cancer/lungcancer-nonsmallcell/detailedguide/non small-cell-lung-cancer-survival-rates.
4. G. Bepler, et al.,”ERCC1 and RRM1 in the international adjuvant lung trial by automated quantitiative in situ analysis,” Am. J. Pathol., vol. 178, no. 1,pp. 69-78, Jan. 2011.
5. G. Loannidis, et al.,”How close are we to customizing chemotheraphy in early non-small-cell lung cancer,” Ther. Adv. Med. Oncol., vol. 3, no. 4, pp.185-205, Jul. 2011

