

# **LUNGS CANCER DETECTION USING CLASSIFICATION ALGORITHM**

Submitted in partial fulfillment of the

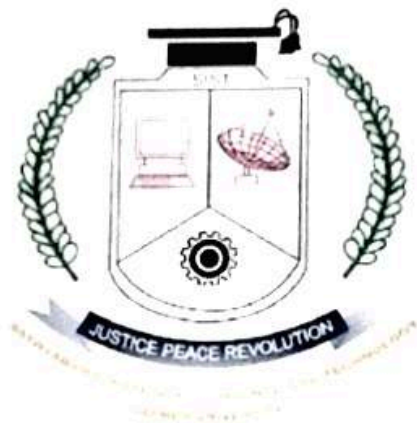
requirements for the award of

Bachelor of Engineering Degree in Electronics and Communication Engineering

by

**VIJAYA K (Reg No-39130514)**

**SNEHA S (Reg No-39130434)**



**DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING  
SCHOOL OF ELECTRICAL AND ELECTRONICS**

## **SATHYABAMA**

**INSTITUTE OF SCIENCE AND TECHNOLOGY  
(DEEMED TO BE UNIVERSITY)**

**Accredited with grade "A" by NAAC**

**JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI – 600119**

**APRIL – 2023**



# **SATHYABAMA**

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited with —All grade by NAAC

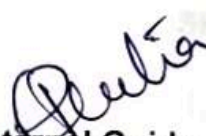
Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai – 600 119

[www.sathyabama.ac.in](http://www.sathyabama.ac.in)

## DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

### **BONAFIDE CERTIFICATE**

This is to certify that this Final Year Project Report is the bonafide work of **VIJAYA K (REG NO: 39130514)** and **SNEHA S (REG NO: 39130434)** who have done the project work as an individual who carried out the project entitled "**LUNG CANCER DETECTION USING CLASSIFICATION ALGORITHM**" under our supervision from Jan 2023 to April 2023.

  
Internal Guide

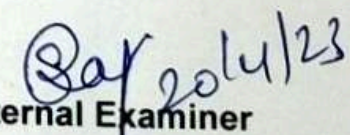
Dr. P. CHITRA M.E., Ph.D.

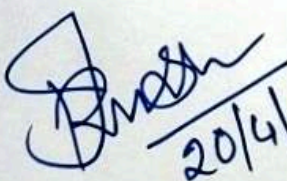
  
Head of the Department

Dr. T. RAVI, M.E., Ph.D.

Submitted for Viva voce Examination held on 20/4/23

Internal Examiner

  
External Examiner

  
20/4/23

## DECLARATION

We, VIJAYA K (REG NO:39130514) and SNEHA S (REG NO:39130434) hereby declare that the Project Report on "LUNGS CANCER DETECTION USING CLASSIFICATION ALGORITHM" done by us under the guidance of Dr. P. CHITRA at Sathyabama Institute of Science and Technology is submitted in partial fulfillment of the requirements for the award of Bachelor of Engineering degree in Electronics and Communication Engineering.

DATE: 20-04-2023

PLACE: Chennai

SIGNATURE OF CANDIDATES(S)

1. *K. Vijaya*
2. *S. Sneha*

## ACKNOWLEDGEMENT

We are pleased to acknowledge our sincere thanks to **Board of Management of SATHYABAMA** for their kind encouragement in doing this project and for completing it successfully. We are grateful to them.

We convey our thanks to **Dr. N. M. NANDHITHA, M.E., Ph.D.,** Dean, School of Electronics, **Dr. T. RAVI, M.E., Ph.D.,** Head of the Department of Electronics and Communication Engineering for providing us necessary support and details at the right time during the progressive reviews.

we would like to express our sincere and deep sense of gratitude to our Project Guide **Dr.P. CHITRA, M.E., Ph.D.,** for her valuable guidance, suggestions, and constant encouragement for the successful completion of our phase-2 project work.

we wish to express our thanks to all Teaching and Non- teaching staff members of the **Department of Electronics and Communication Engineering** who were helpful in many ways for the completion of the project.

## ABSTRACT

The most extreme and lethal illnesses on this world is mostly lung cancers. On the opposite hand, timely analysis and aid can keep lives. Detection of malignant lungs. Recently, imaged processing techniques were broadly used in many scientific fields to improve images in early detection and remedy. Various segmentation and magnification techniques were used to extract the normal and uncommon functions of the images. In this system Neuro-fuzzy classifier is used to detect cancer. For extracting features the image processing techniques applied are Pre-processing, Image Segmentation, filtering, Dilation and Image filling—to get more accurate findings. Once the lung CT image has been input, the pre-processing technique is used with the intention of improving the lung image data by suppressing unintentional distortions or enhancing lung image features crucial for future processing. The image has been segmented by thresholding after pre-processing. The threshold segmentation process can be regarded as the process of separated foreground background. By using a filtering process, the segmented image is enhanced or modified. Filtered images dilated using the dilation technique. Dilation adds images to the boundaries of objects in an image. The dilation and image filling stages will find the image is affected by cancer or not with neuro-fuzzy classification.

## 4 METHOUR

### 4.1 WORKING

11

### 4.2 METHODOLOGY

12

### 4.3 MORPHOLOGICAL IMAGE PROCESSING

14

### 4.4 GUASSIAN IMAGE PROCESSING

16

### 4.5 FOURIER TRANSFORM

17

## **CONTENTS**

| <b>S.NO</b> | <b>TITLE</b>                                      | <b>PAGE NO</b> |
|-------------|---------------------------------------------------|----------------|
|             | <b>ABSTRACT</b>                                   | <b>iv</b>      |
|             | <b>LIST OF CONTENTS</b>                           | <b>v</b>       |
|             | <b>LIST OF FIGURES</b>                            | <b>vii</b>     |
|             | <b>LIST OF TABLES</b>                             | <b>viii</b>    |
| <b>1</b>    | <b>INTRODUCTION</b>                               | <b>1</b>       |
| <b>2</b>    | <b>LITERATURE SURVEY</b>                          | <b>3</b>       |
| <b>3</b>    | <b>AIM AND SCOPE OF THE PRESENT INVESTIGATION</b> |                |
|             | 3.1 OBJECTIVE                                     | 7              |
|             | 3.2 SCOPE OF THE PROJECT                          | 7              |
|             | 3.3 EXISTING SYSTEM                               | 7              |
|             | 3.4 PROPOSED SYSTEM                               | 8              |
|             | 3.5 SOFTWARE USED                                 | 8              |
| <b>4</b>    | <b>METHODOLOGY OF WORKING SYSTEM</b>              |                |
|             | 4.1 WORKING                                       | 11             |
|             | 4.2 METHODOLOGY                                   | 12             |
|             | 4.3 MORPHOLOGICAL IMAGE PROCESSING                | 14             |
|             | 4.4 GUASSIAN IMAGE PROCESSING                     | 16             |
|             | 4.5 FOURIER TRANSFORM                             | 17             |

|          |                                        |           |
|----------|----------------------------------------|-----------|
|          | 4.6 EDGE DETECTION IN IMAGE PROCESSING | 18        |
|          | 4.7 WAVELET IMAGE SOURCEING            | 19        |
|          | 4.8 NEURO FUZZY CLASSIFIER             | 24        |
| <b>5</b> | <b>RESULTS AND DISCUSSION</b>          |           |
|          | 5.1 IMPLEMENTATION                     | 27        |
|          | 5.2 TABULATION OF CLASSIFICATION       | 30        |
|          | 5.3 ANALYSIS OF RESULTS                | 32        |
| <b>6</b> | <b>CONCLUSION</b>                      | <b>36</b> |
| <b>7</b> | <b>REFERENCES</b>                      | <b>38</b> |

## LIST OF FIGURES

| FIGURE NO | FIGURE NAME                                                  | PAGE NO |
|-----------|--------------------------------------------------------------|---------|
| 4.1       | WORK FLOW                                                    | 11      |
| 4.2       | SOURCE                                                       | 13      |
| 4.3       | SOURCE STRUCTURE                                             | 14      |
| 4.4       | ORIGIN SOURCE                                                | 15      |
| 4.5       | DILATION SOURCE                                              | 15      |
| 4.6       | ORIGINAL FILTER RESULT SOURCE                                | 17      |
| 4.7       | GRADIENT SOURCE                                              | 19      |
| 4.8       | PREPROCESSING SOURCE                                         | 20      |
| 4.9       | BASIC NEURAL NETWORK SOURCE                                  | 20      |
| 4.10      | OPERATIONS IN A SINGLE NEURON SOURCE                         | 22      |
| 4.11      | 4x4x3 RGB IMAGE                                              | 23      |
| 4.12      | BLOCK DIAGRAM FOR CLASSIFIER                                 | 24      |
| 4.13      | LAYERS OF CLASSIFIER                                         | 25      |
| 5.1       | DETECTION AND CLASSIFYING OF LUNG<br>CANCER                  | 28      |
| 5.2       | ACCURACY ANALYSIS                                            | 33      |
| 5.3       | PERFORMANCE ANALYSIS WITH RESPECT TO<br>MEAN SQUARE ERROR    | 34      |
| 5.4       | TRAINING STATE ANALYSIS AT 3 EPOCHS                          | 35      |
| 5.5       | REGRESSION ANALYSIS WITH RESPECT TO<br>THE OUTPUT OF TARGETS | 35      |

## LIST OF TABLES

| TABLE NO | TABLE NAME                                                           | PAGE NO |
|----------|----------------------------------------------------------------------|---------|
| 5.1      | CLASSIFICATION RESULTS OF LUNG IMAGES (IMAGE 1 AND IMAGE 2)          | 31      |
| 5.2      | CLASSIFICATION RESULTS OF LUNG IMAGES (IMAGE 1 AND IMAGE 2)          | 31      |
| 5.3      | CLASSIFICATION RESULTS OF LUNG IMAGES (IMAGE 1 AND IMAGE 2)          | 31      |
| 5.4      | COMPARITIVE ANALYSIS OF ACCURACY WITH RESPECT TO THE PROCESSING TIME | 32      |

## **CHAPTER - 2**

### **INTRODUCTION**

The lung cancer, when it occurs and shows symptoms of a terminal condition, it's miles difficult to discover. However, early detection of the sickness and treatment can reduce the charge and likelihood of mortality. The best imaging technique Since Computer Tomography (CT) scans can hit upon each suspicious and hidden nodules of lung most cancers, they are effective in diagnosing lung most cancers. However, differences in CT scan image depth and misinterpretation of anatomical structures by way of scientific specialists which includes radiologists could make it tough to become aware of malignant cells. Recently, computer diagnostics has become a treasured and promising method to assist radiologists and clinicians efficiently diagnose most cancers. Many systems had been established and research maintains to locate lung cancer. However, a few structures are nevertheless to be developed to be able to be capable of hit upon it with the best accuracy. Lung cancer detection and classification has been carried out using processing and machine gaining knowledge of strategies. In order to pick the maximum current best techniques and to research them and reviewed these days evolved structures for most cancer detection based totally on lung CT scan images. Gindi et al (2014) used the excellent imaging approach. CT is dependable for the diagnosis of lung cancer due to the fact it could discover all suspicious and hidden nodules of lung most cancers [7]. However, Suzuki et al (2006) found that variations in CT scan image intensity and discomfort of anatomical structures by using health experts could make it hard to discover malignant cells [25]. Xiuhua et al (2011) applied models for the prediction of cancerous lung nodules primarily based on CT image [26] texture capabilities. Recently, a promising computer diagnostic approach has been evolved to assist radiologists and medical doctors to diagnose cancer successfully. Many structures had been www established and research maintains to stumble on lung cancer. However only a few structures have been created to be able to find it with a 100 percentage of accuracy. Lung cancers detection and classification has been applied the usage of image www processing and neuro fuzzy classification techniques. In order to pick the maximum superior and powerful systems. Presently advanced systems for

the diagnosis of most cancers for CT scan image of the lung was reviewed. Therefore, these elements are analyzed and proposed a brand-new model. In this system Neuro-fuzzy classifier is used to detect cancer for extracting features the image processing techniques applied are Pre-processing, Image Segmentation, filtering, Dilation and Image filling to get more accurate findings.

## **CHAPTER – 2**

### **LITERATURE SURVEY**

Muthazhagan et al (2021) states ~~ta~~with the aid of current lung cancer prediction technologies, predicting and detecting lung cancer at an early stage is a difficult challenge. An early lung tumor prediction might extend a person's life by one to five years. They created a Support Vector Machine based classification model which provided about 98% prediction accuracy in a small amount of time. However, the images were merely classified into 'abnormal' or 'normal' and did not take into account the various stages [Stage 0 – stage IV] which is what this project aims to improve on [13].

Masud et al. (2021) applied a CNN based model for classifying the image into one of five kinds: colon adenocarcinomas, benign colonic tissues, lung adenocarcinomas, lung squamous cell carcinomas and benign lung tissues. While a peak accuracy of 96.33% has been achieved in the classification, the authors state that two out of five classes can have much improved performance with further experimentation. The dataset used is Histopathological and Histopathology is the microscopic examination of a biopsy which is an invasive process. Our approach prefers to work on CT scans which is a non-invasive mechanism to detect cancer [12].

Saija et al. (2019) published a paper which worked on detecting lung cancer using the pre-trained CNN model called Google-Net. The deployed 60% of all neurons in the drop out layers to prevent overfitting and achieved a simplified and sparse network for classifying the CT images into benign or malignant. The model still requires testing on various dropout ratios to check for better performance accuracy. Our approach aims to construct a simplified CNN model to classify cancer along with providing 7 medical information costs. The dataset used is Histopathological and Histopathology is the microscopic examination of a biopsy which is an invasive process. Our approach prefers to work on CT scans which is a non-invasive mechanism to detect cancer [20].

Bhatia et al. (2019) presents a method for detecting lung cancer using deep residual learning. They offer a series of preprocessing strategies for extracting cancer-vulnerable lung features using UNet and ResNet models. They examine the likelihood of predicting carcinogenic CT scans by comparing the effectiveness of classifiers such as Random Forest and XGBoost. When the authors combine the two classifiers, they get the greatest accuracy of 84%. The constraint in this case is that the best achievable accuracy may have been higher [23].

Makaju et al. (2018) made a comparison of many probable cancer detection approaches and ranked them in order of effectiveness. They decide to upgrade that model to achieve even higher accuracy by selecting the current best approach from their survey of articles. The Median and Gaussian filters were used in the preprocessing stage, and the data was then segmented using the Watershed algorithm. They went on to use support vector machines to identify diagnosed cancerous nodules as benign or malignant. This upgraded model outperformed the previous best model by 5.4%, with an accuracy rate of 92 %. The model's sole flaw is that it does not differentiate between cancer stages (I to IV) [24]

AlphaGo et al. (2018) developed a deep learning algorithm that takes a CT image and perceives it as a collection of states, producing a classification of whether a malignant nodule is present. They employ a Reinforcement Learning algorithm that improves with time and with more data. Their research shows that the model's training data has a high accuracy of 99.1%, however the validation data has allowed accuracy of 64.4 %. The model appears to be overfitted as a result of this. The authors suggest that because this is the only flaw, the constraint can be solved with more data [2]

Aggarwal et al (2015) to distinguish nodules from a normal anatomical form. With this method, features are taken from statistics, geometry, and grey degrees. LDA is then utilized as a segmentation classifier and has the best thresholds. The accuracy of the machine is 84%, the sensitivity is 97.14%, and the specificity is 53.33%. The era detects a most cancers-like node. but its accuracy is simply too low. Instead of device mastering strategies, simple strategies are placed on the facts. For this motive, any boom it provides to the new version does no longer growth the probability of

improvement [1].

Jin et al (2016) used a convolutional neural network as a classifier in the cadmium device for lung most cancers detection. The accuracy, sensitivity, and specificity of the device are 84.6%, 82.5%, and 86.7%, respectively. The benefit of this model is the usage of a round clear out within the region of interest extraction, which decreases cost and understanding degrees. Although the fee of implementation has reduced, the accuracy remains beneath parity [10].

Sangamithraa et al (2016) created a device that uses K-means nested studying technique for clustering or segmentation. A set of pixel records is organized by way of sure traits. This example makes use of a grid community for categorization. The Gray Level Coincidence Matrix (GLCM) approach is used to extract functions consisting of entropy correlation uniformity, PSNR, and SSIM. The accuracy of the device is ready 90.7%. A median filter used in image preprocessing to take away noise can help our new version lessen noise and improve accuracy [21].

Roy et al (2015) though used a block device and a lively model to broaden a most cancers nodule detection machine version. This approach will increase visible assessment with grayscale conversion. Before segmentation, the image undergoes a binary manner, and the resulting image is segmented the usage of an active shape, thresholding technique is used to diagnose most cancers to teach the classifier, functions together with place, correlation, mean, entropy, fundamental axis duration, and minor axis duration are obtained. The overall machine's accuracy is 94.12%. The suggested version application has some restrictions, including the inability to distinguish between benign and malignant tumors [19].

Ignatious et al (2015) created a mechanism using segmentation therapy. A Gabor clear out is utilized in pre-processing to enhance image exceptional. This contrasts the accuracy of the growing region approach with the bushy neural version. Despite the rising sample and application of the approach, the segmentation accuracy of the suggested approach is 90.1%, which is better than the segmentation accuracy of

utilizing the neural model. The advantage of this approach is that it just uses a small number of water segments to solve the problem. Its downside is that, in spite of its terrific accuracy, it does not suggest benign or www malignant cancer. This model can also undergo some adjustments or additions that growth the accuracy to a first-class stage [8].

Gonzalez et al. (2016) proposed a framework that separates lung cancer into benign and malignant types. The area of interest is defined by using the system using the House field Unit (HU) and prior knowledge (ROI). To set up a vector machine and insert assist to determine whether a node is benign or malignant, form characteristics together with location, eccentricity, roundness, and fracture measurement are decided on. The following textual properties are also extracted: implication, tension, electricity, entropy, skewness, variance, and smoothness. This model's benefit is that it distinguishes between benign and malignant cancer. But it has the drawback that it calls for to know in advance the amount of hobby. Our new model can enjoy the class of benign or malignant tumors the use of vectorial mechanisms. The machine proposed by means of Ignatius and Joseph is the best option, in step with the evaluation of literature evaluations which have been completed in terms of accuracy and convenience of the tactics. In pre-processing, the image is better with a filter and cancerous nodes are detected using the threshold segmentation method. Also, the method best extracts characteristics from most cancer nodes, including proximity, perimeter, and eccentricity. It suggests the correctness of the proposed version. that is 90.1%, in opposition to different models that have been proposed before [18].

## **CHAPTER – 3**

### **AIM AND SCOPE OF THE PRESENT INVESTIGATION**

#### **3.1 OBJECTIVE:**

The overall classification of accuracy using this method is found to be 98% which is significantly good. To reduce the false positive and false negative rate and cancer patients will be diagnosed in more accurate way.

#### **3.2 SCOPE OF THE PROJECT:**

From this, it is possible to identify patients quickly and cheaply with lung cancer. The motivation of this paper is to discover the cancer cells in the lungs at earlier stage. The lung cancer detection is the extension of the image processing that produces the results of feature extraction and feature selection after segmentation. The system accepts Lung CT (Computed Tomography) images as input.

#### **3.3 EXISTING SYSTEM:**

In an existing system proposed an image processing-based method that gives improved accuracy rate of the lung cancer detection and its location. It also provides with the information that helps to determine whether the image is malignant or not. The approach you describe in this research combines a histogram-based method with an integrated thresholding procedure to detect lung cancer in MRI images and provides a comprehensive analysis.

### **3.4 PROPOSED SYSTEM:**

An image processing technique is built to detect diseases at early stage of cancer so the patient can take the treatment at early stages. The time factor is major factor to discover the abnormal tissue in target x-ray images. The accuracy and the quality of image is one the major core factor of this research. Image quality as well as image enhancement stage were adopted as low pre-processing techniques based on histogram equalization. This technique is efficient for segmentation stage so the region of interest for feature extraction obtaining. The segmentation technique is based on thresholding. On the basis of general features, a normality and abnormality comparison is made. The main feature for detection of accurate image comparison are pixel percentage and mask labeling which gives us the indication that the process of detection this disease plays a very important and essential role to avoid serious stages and to reduce its percentage distribution in the world. Three stages applied to obtain more accurate results: Image Enhancement stage, Image Segmentation stage and Features Extraction stage.

### **3.5 SOFTWARE USED:**

Here the software used, MATLAB 2013R and Python for developing an algorithm. Images define the world, each image has its own story, it contains a lot of crucial information that can be useful in many ways. This information can be obtained with the help of the technique known as Image Processing.

It is the core part of computer vision which plays a crucial role in many real-world examples like robotics, self-driving cars, and object detection. Image processing allows us to transform and manipulate thousands of images at a time and extract useful insights from them. It has a wide range of applications in almost every field.

Python is one of the widely used programming languages for this purpose. Its amazing libraries and tools help in achieving the task of image processing very efficiently. MATLAB® is a high-level technical computing language and interactive environment for algorithm development, data visualization, data analysis, and numerical computation. Using matlab, you can solve technical computing problems

faster than with traditional programming languages, such as C, C++, and FORTRAN. Matlab is a data analysis and visualization tool which has been designed with powerful support for matrices and matrix operations. As well as this, Matlab has excellent graphics capabilities, and its own powerful programming language. One of the reasons that Matlab has become such an important tool is using sets of MATLAB programs designed to support a particular task. These sets of programs are called toolboxes, and the toolbox of interest to us is the image processing toolbox. Rather than give a description of all of Matlab's capabilities, Let will limit our discussion to only those elements that relate to managing images. Functions, instructions, and procedures will be introduced as needed. A Matlab function is a keyword that accepts different inputs and outputs, such as a matrix, a text, or a graph. Sin, misunderstand, and enclose are a few examples of such functions. In Matlab, there are several functions, and as are about to discover, it is quite simple (and occasionally important) to develop our own. Matlab's standard data type is the matrix all data are matrices of some sort. Images, of course, are matrices whose elements are the grey values (or possibly the RGB values) of its pixels. Single values are considered by Matlab to be matrices, while a string is merely a matrix of characters; being the string's length. It will examine the more general Matlab commands in this chapter, and proceed to talk about images in later chapters.

When you start up Matlab, you have a blank window called the Command Window\_ in which you enter commands. Given the vast number of Matlab's functions, and the different parameters they can take, a command line style interface is in fact much more efficient than a complex sequence of pull-down menus.

Signal and image processing, communications, control design, test and measurement, financial modelling, and analysis are just a few of the many applications you can utilize MATLAB to use. Add-on toolboxes (collections of special-purpose Matlab functions) extend the Matlab environment to solve classes of problems in these application areas.

MATLAB provides a number of features for documenting and sharing your work. You can integrate your Matlab code with other languages and applications, and distribute your MATLAB algorithms and applications. When working with images in Matlab, there are many things to keep in mind such as loading an image, using the right

format, saving the data as different data types, how to display an image, conversion between different image formats.

Image Processing Toolbox provides a comprehensive set of reference-standard algorithms and graphical tools for image processing, analysis, visualization, and algorithm development. You can perform image enhancement, image deblurring, feature detection, noise reduction, image segmentation, spatial transformations, and image registration. Many functions in the toolbox are multithreaded to take advantage of multicore and multiprocessor computers.

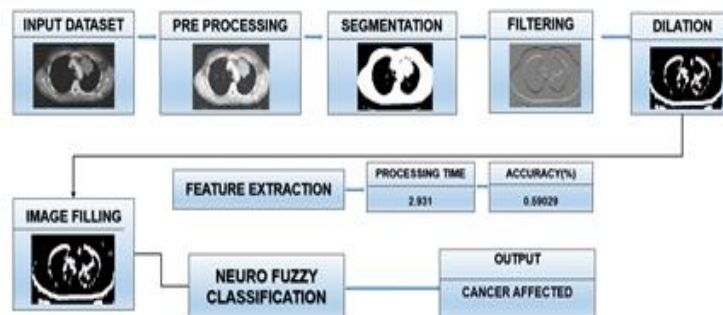
## CHAPTER 4

### METHODOLOGY OF WORKING SYSTEM

#### 4.1 WORKING:

The gadget that was described in this newsletter makes use of a combination of Gaussian thresholding and a morphology- based totally imaging segmentation approach to stumble on lung most cancers from Magnetic Resonance Imaging (MRI) and CT images and affords a comprehensive evaluation.

In order to come across diseases inside the first level of cancer, in order that the patient can get hold of remedy at an early degree, an image processing system has been evolved. In order to find aberrant tissues on target x- ray images, timing is fundamental. One of the main factors of this study is the accuracy and first-rate of the photos, Low-level processing strategies primarily based on the Gabor filter had been utilized in each the picture best and image enhancement stage. In this system neuro-fuzzy classification was used for cancer detection. It is typically feasible to discover a appropriate classifier even though by way of searching on the information. It has 3 layers and connections among the input, hidden and output layers. For extracting features, the image processing techniques applied are Pre-processing, Image Segmentation, filtering, Dilation and Image filling—to get more accurate findings.



**Fig 4.1: Work flow**

## 4.2 METHODOLOGY:

STEP-1: Collect the lung cancer images from respective cancer hospital

STEP-2: Access one image into MATLAB with the help of command

STEP-3: Enhancement process Image enhancement is to improve the interpretability or perception of information included in the image for human viewers, or to provide better input for other automated image processing techniques. For the purpose of image enhancement, people use the Gabor filter.

STEP-4: Segmentation process Segmentation divides the image into its constituents' regions or objects. It has many useful applications for the medical professional such as visualization and volume estimation of object of interest, detection of abnormalities, tissue qualification and classification, and more. For segmentation, one can be employing the Marker-Controlled Watershed Segmentation Approach.

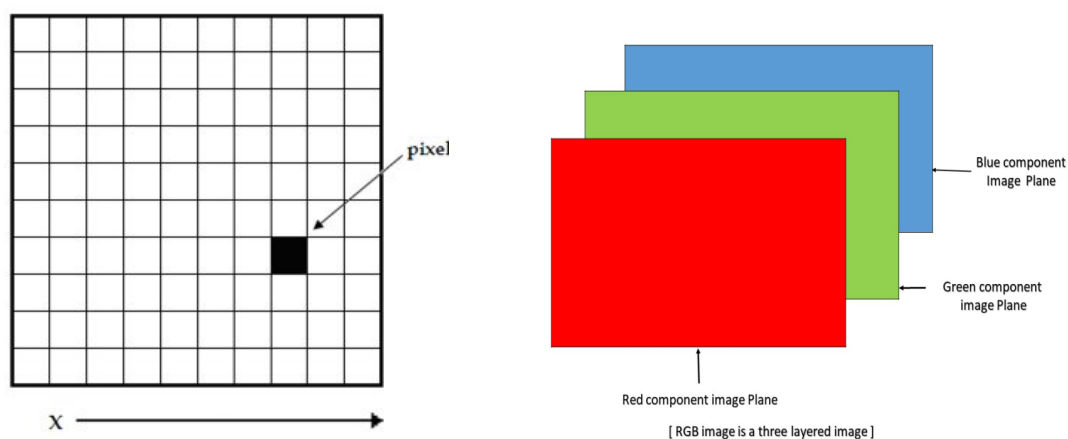
STEP-5: Extracting features This procedure is used to determine whether an image is normal or aberrant. For feature extraction, binarization and masking are used.

For extracting features, the image processing techniques applied are Pre-processing, Image Segmentation, filtering, Dilation and Image filling—to get more accurate findings. The pre-processing approach is used after the lung CT image has been input, and its goal is to improve the lung image data by suppressing distortions or enhancing lung image features crucial for subsequent processing. Behind pre-processing, the preprocessed image has been segmented by thresholding. The threshold segmentation process can be regarded as the process of separated foreground background. By using a filtering process, the segmented image is enhanced or modified. A lung image has to be filtered to highlight or remove specific features. Filtered images dilated using the dilation technique. Dilation adds images to the boundaries of objects in an image. The dilation and Image filling stages will find the image is affected by cancer or not with neuro-fuzzy classification. Variable MRI and CT scan images were identified and recognized as malignant using this classifier approach.

Through this article, you will learn about classical algorithms, techniques, and tools to process the image and get the desired output. The final output can be either in the form of an image or a corresponding feature of that image. This can be used for further analysis and decision making.

An image can be represented as a 2D function  $F(x, y)$  where  $x$  and  $y$  are spatial coordinates. The amplitude of  $F$  at a particular value of  $x, y$  is known as the intensity of an image at that point. Inevitably refer to it as a digital image if  $x, y$ , and the amplitude value are all finite. Pixels are grouped in columns and rows in an array.

Pixels are the elements of an image that contain information about intensity and color. An image can also be represented in 3D where  $x, y$ , and  $z$  become spatial coordinates. Pixels are arranged in the form of a matrix. This is known as an **RGB image**.



**Fig 4.2 : Source**

There are various types of images:

**RGB image:** It contains three layers of 2D image, these layers are red, green, and blue channels.

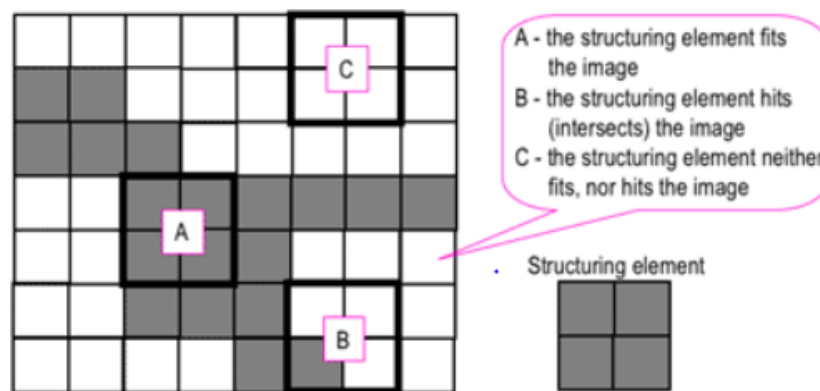
**Grayscale image:** These images contain shades of black and white and contain only a single channel.

### 4.3 MORPHOLOGICAL IMAGE PROCESSING

Morphological image processing tries to remove the imperfections from the binary images because binary regions produced by simple thresholding can be distorted by noise. It also helps in smoothing the image using opening and closing operations.

Morphological operations can be extended to grayscale images. It consists of non-linear operations related to the structure of features of an image. It depends on the related ordering of pixels but on their numerical values. This technique analyzes an image using a small template known as **structuring element** which is placed on different possible locations in the image and is compared with the corresponding neighborhood pixels. A structuring element is a small matrix with 0 and 1 values.

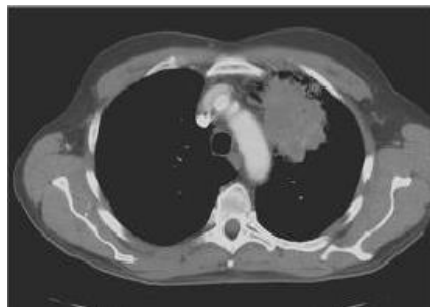
**Dilation** operation adds pixels to the boundaries of the object in an image. **Erosion** operation removes the pixels from the object boundaries. The number of pixels removed or added to the original image depends on the size of the structuring element. Structuring element is a matrix consisting of only 0's and 1's that can have any arbitrary shape and size. It is positioned at all possible locations in the image and it is compared with the corresponding neighborhood of pixels.



**Fig 4.3: Source Structure**

The rectangular structural element 'A' fits inside the item that wish to choose, 'B' intersects the object, and 'C' is outside the object. The zero-one pattern defines the configuration of the structuring element depends on the shape of the thing that everybody wants to choose. The center of the structuring element identifies the pixel being processed.

Once pre-processing, the image could be divided through the threshold value. The threshold segmentation technique may be concept of as a foreground extraction technique. Threshold in particular extracts foreground primarily based on grey price facts Therefore, segmenting images with strong contrast between focus and background objects is especially advantageous. Once segmentation, the image is applied for image filtering, then the filtered image is applied for dilation. The dilated image displayed the cancered parts.



***Fig 4.4 Original Source***

Once segmentation, the image is applied for image filtering, then the filtered image is applied for dilation. The dilated image displayed the cancered parts.



***Fig 4.5 Dilation Source***

## 4.4 GAUSSIAN IMAGE PROCESSING

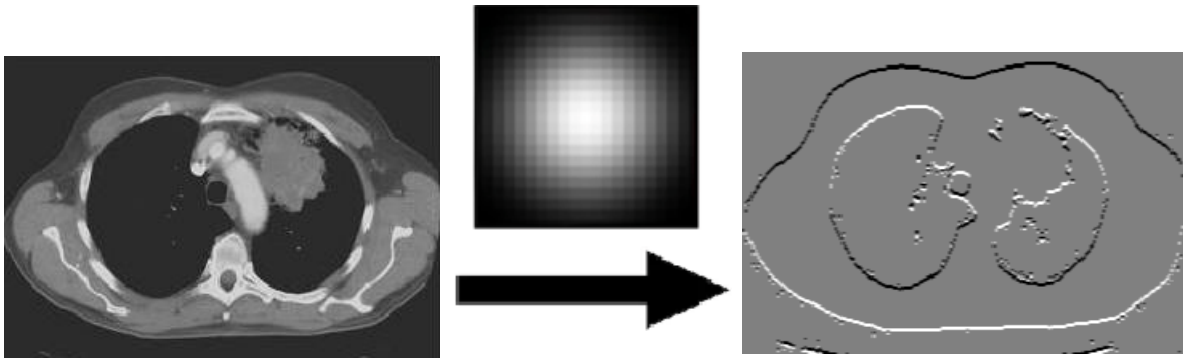
Gaussian blur which is also known as gaussian smoothing, is the result of blurring an image by a Gaussian function. It is used to reduce image noise and reduce details. The visual effect of this blurring technique is similar to looking at an image through the translucent screen. It is sometimes used in computer vision for image enhancement at different scales or as a data augmentation technique in deep learning.

The basic gaussian function looks like:

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \text{-----}(4.1)$$

In practice, it is best to take advantage of the Gaussian blur's separable property by dividing the process into two passes. In the first pass, a one-dimensional kernel is used to blur the image in only the horizontal or vertical direction. In the second pass, the same one-dimensional kernel is used to blur in the remaining direction. The resulting effect is the same as convolving with a two-dimensional kernel in a single pass. Let's see an example to understand what gaussian filters do to an image.

When a filter with a normal distribution is used on an image, the outcomes are as follows:



***Fig 4.6 : Original Filter Result Source***

It is evident that some of the margins are a little less detailed. The filter is giving more weight to the pixels at the center than the pixels away from the center. Gaussian filters are low-pass filters i.e. weakens the high frequencies. It is commonly used in edge detection.

## **4.5 FOURIER TRANSFORM**

Fourier transform breaks down an image into sine and cosine components.

It has multiple applications like image reconstruction, image compression, or image filtering. inevitably shall consider the discrete Fourier transform while are discussing images. Let's consider a sinusoid, it comprises of three things:

Magnitude – related to contrast

Spatial frequency – related to brightness Phase – related to color information

The formula for 2D discrete Fourier transform is:

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})} \text{-----}(4.2)$$

In the above formula,  $f(x, y)$  denotes the image.

The inverse fourier transform converts the transform back to image. The formula for 2D inverse discrete fourier transform is:

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(\frac{ux}{M} + \frac{vy}{N})} \text{-----}(4.3)$$

#### 4.6 EDGE DETECTION IN IMAGE PROCESSING

Edge detection is an image processing technique for finding the boundaries of objects within images. It works by detecting discontinuities in brightness.

This could be very beneficial in extracting useful information from the image because most of the shape information is enclosed in the edges. Classic edge detection methods work by detecting discontinuities in the brightness.

It can rapidly react if some noise is detected in the image while detecting the variations of grey levels. Edges are defined as the local maxima of the gradient.

The most common edge detection algorithm is sobel edge detection algorithm. Sobel detection operator is made up of 3\*3 convolutional kernels. A simple kernel  $G_x$  and a 90-degree rotated kernel  $G_y$ . Separate measurements are made by applying both the kernel separately to the image.

$$Gx = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * Image\ matrix \quad \text{-----}(4.4)$$

And,

$$Gy = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * Image\ matrix \quad \text{-----}(4.5)$$

denotes the 2D signal processing convolution operation. Resulting gradient can be calculated as:

$$G = sqrt(Gx^2 + Gy^2) \quad \text{-----}(4.6)$$



**Fig 4.7: Gradient Source**

## 4.7 WAVELET IMAGE PROCESSING

The Fourier transform that everyone saw is only applicable to frequency. Wavelets consider both time and frequency. The non-stationary signals are a good fit for this transform.

Individuals are aware that edges are among the most significant aspects of an image, and using conventional filters has been shown to blur the image while removing noise. The wavelet transform is constructed in such a way that low frequency components have strong frequency resolution.

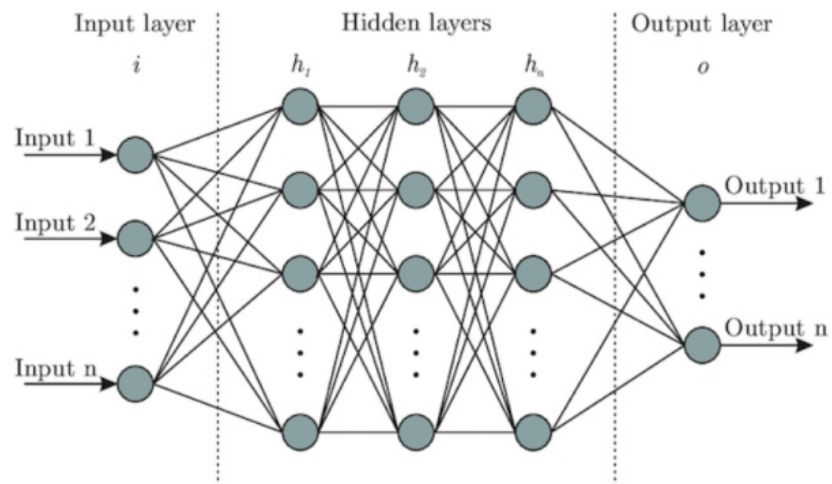


**Fig 4.8: Preprocessing Source**

Neural Networks are multi-layered networks consisting of neurons or nodes. These neurons are the core processing units of the neural network. They are designed to act like human brains. They take in data, train themselves to recognize the patterns in the data and then predict the output.

A basic neural network has three layers:

Input layer, Hidden layer and Output layer



**Fig 4.9 :Basic neural network Source**

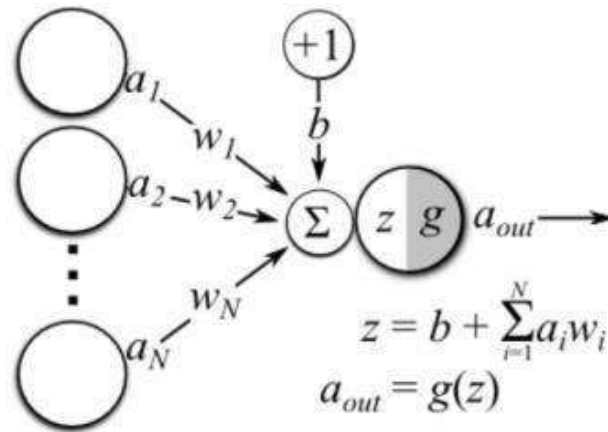
The input layers receive the input, the output layer predicts the output and the hidden layers do most of the calculations. The number of hidden layers can be modified according to the requirements. There should be at least one hidden layer in a neural network. The basic working of the neural network is as follows:

Let us consider an image, each pixel is fed as input to each neuron of the first layer, neurons of one layer are connected to neurons of the next layer through channels. Each of these channels is assigned a numerical value known as weight. The inputs are multiplied by the corresponding weights and this weighted sum is then fed as input to the hidden layers.

The output from the hidden layers is passed through an activation function which will determine whether the neuron will be activated or not. The activated neurons transmit data to the next hidden layers. In this manner, data is propagated through the network, this is known as Forward Propagation. In the output layer, the neuron with the highest value predicts the output. These outputs are the probability values. The predicted output is compared with the actual output to obtain the error. This information is then transferred back through the network, the process is known as Backpropagation.

Based on this information, the weights are adjusted. This cycle of forward and backward propagation is done several times on multiple inputs until the network predicts the output correctly in most of the cases. This ends the training process of the neural network. The time taken to train the neural network may get high in some cases.

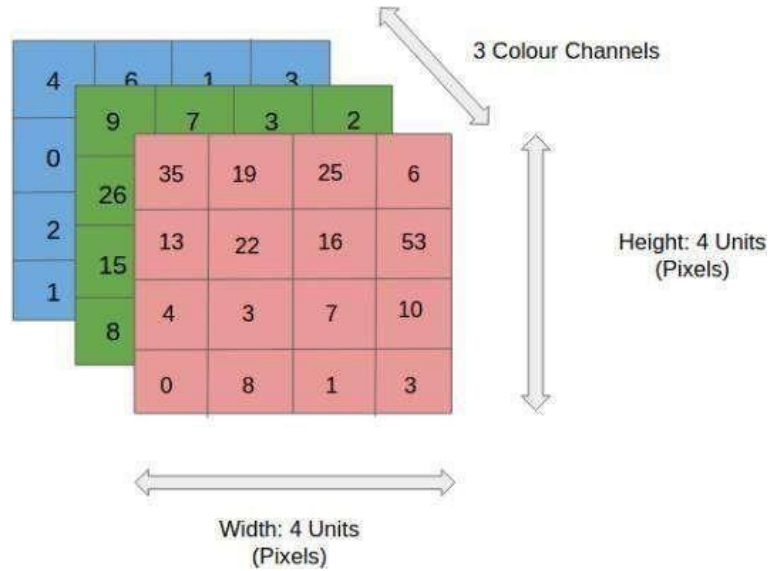
In the below image,  **$a_i$ 's** is the set of inputs,  **$w_i$ 's** are the weights,  **$z$**  is the output and  **$g$**  is any activation function.



**Fig 4.10: Operations in a single neuron Source**

Here are some guidelines to prepare data for image processing. More data needs to be fed to the model to get the better results. Image dataset should be of high quality to get more clear information, but to process them you may require deeper neural networks.

In many cases RGB images are converted to grayscale before feeding them into a neural network.



**Fig 4.11 :4x4x3 RGB Image**

In the figure, we have an RGB image which has been separated by its three-color planes — Red, Green, and Blue. There are a number of such color spaces in which images exist — Grayscale, RGB, HSV, CMYK, etc.

Once the graphics approach dimensions, then begin to understand how computationally intensive things would get. say 8K (7680×4320).

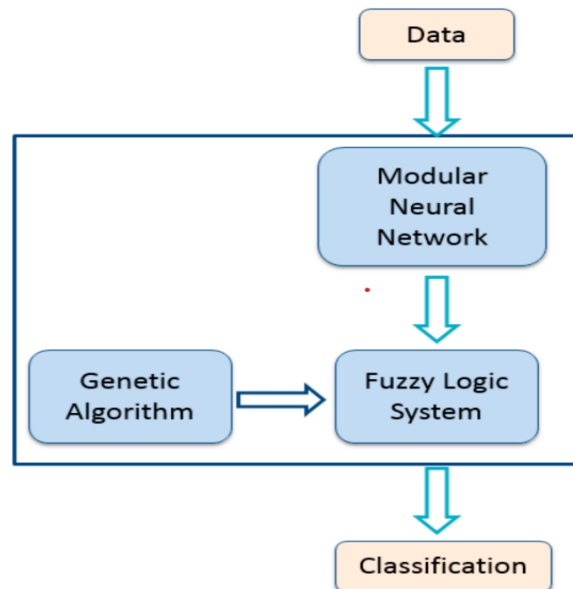
Here neuro fuzzy classifier was applied for developing our proposed system for extracting features and detecting lung cancer.

#### 4.8 NEURO FUZZY CLASSIFIER:

Neuro-fuzzy classification systems offer means to obtain fuzzy classification rules by a learning algorithm. It is usually possible to find a suitable fuzzy classifier by learning from data, but it can be hard to obtain a classifier that can be interpreted conveniently. A neuro-fuzzy system is a fuzzy system that uses a learning algorithm derived from or inspired by neural network theory to determine its parameters (fuzzy sets and fuzzy rules) by processing data samples.

A heterogeneous neuro A heterogeneous neuro-fuzzy system is fuzzy system is hybrid system that consists of a neural network and a fuzzy system working as independent components. As an example of the application of such a system, we will consider a problem of diagnosing myocardial perfusion from cardiac images.

Fuzzy classification is the process of grouping elements into fuzzy sets whose membership functions are defined by the truth value of a fuzzy propositional function.



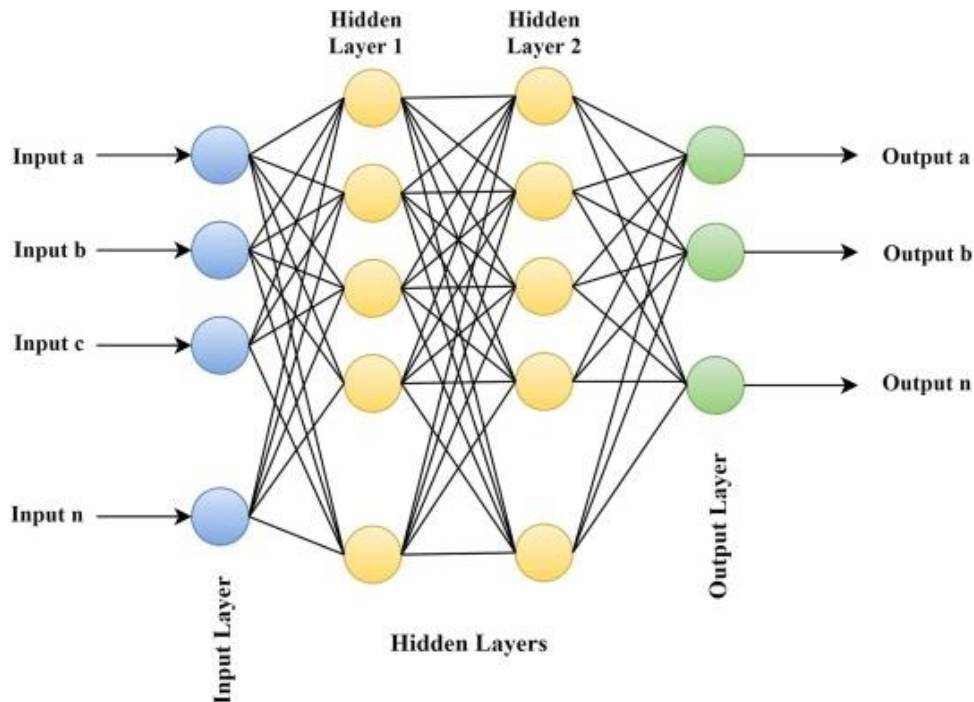
**Fig 4.12: Block Diagram of classifier**

The four most common types of neural network layers are Fully connected, Convolution, Deconvolution, and Recurrent, and below you will find what they are and how they can be used. A neural network can implement tasks that a linear program cannot.

When an item of the neural network declines, it can continue without some issues

by its parallel features. A neural network determines and does not require to be reprogrammed. It can be executed in any application.

A fuzzy neural network can be viewed as a three-layer feedforward network, with a fuzzy input layer (fuzzification), a hidden layer containing the fuzzy rules, and a final fuzzy output layer (defuzzification).



**Fig 4.13: layers of classifier**

The input layer of a neural network is composed of artificial input neurons, and brings the initial data into the system for further processing by subsequent layers of artificial neurons. The input layer is the very beginning of the workflow for the artificial neural network. A hidden layer in an artificial neural network is a layer in between input layers and output layers, where artificial neurons take in a set of weighted inputs and produce an output through an activation function. Hidden layer(s) are the secret sauce of your network. They allow you to model complex data thanks to their nodes/neurons. They are “hidden” because the true values of their nodes are unknown in the training dataset. In fact, we only know the input and output.

For example, a hidden layer functions that are used to identify human eyes and ears

may be used in conjunction by subsequent layers to identify faces in images. While the functions to identify eyes alone are not enough to independently recognize objects, they can function jointly within a neural network. Hiding a layer means that the contents of that layer are not visible and locking a layer means that the contents of that layer cannot be edited. The size of the hidden layer is normally between the size of the input and output-. It should be should be  $\frac{2}{3}$  the size of the input layer plus the size of the o/p layer The number of hidden neurons should be less than twice the size of the input layer. The output layer in an artificial neural network is the last layer of neurons that produces given outputs for the program. The output layer is the final layer in the neural network where desired predictions are obtained. There is one output layer in a neural network that produces the desired final prediction. It has its own set of weights and biases that are applied before the final output is derived. The fourth layer is the output membership layer (y). Neurons in this layer represent fuzzy sets used in the consequent of fuzzy rules. The output of this layer is the product of the normalized firing strength of the rule obtained in previous layer and first order polynomial. Size of the output layer equals to number of classes in the dataset. Further, if dataset has two classes only just one output unit is enough for discriminating these two classes. The number of neurons in the output layer equals the number of outputs associated with each input. The axon (tree roots) is the output structure of the neuron; when a neuron wants to talk to another neuron, it sends an electrical message called an action potential throughout the entire axon.

## **CHAPTER 5**

### **RESULTS AND DISCUSSION**

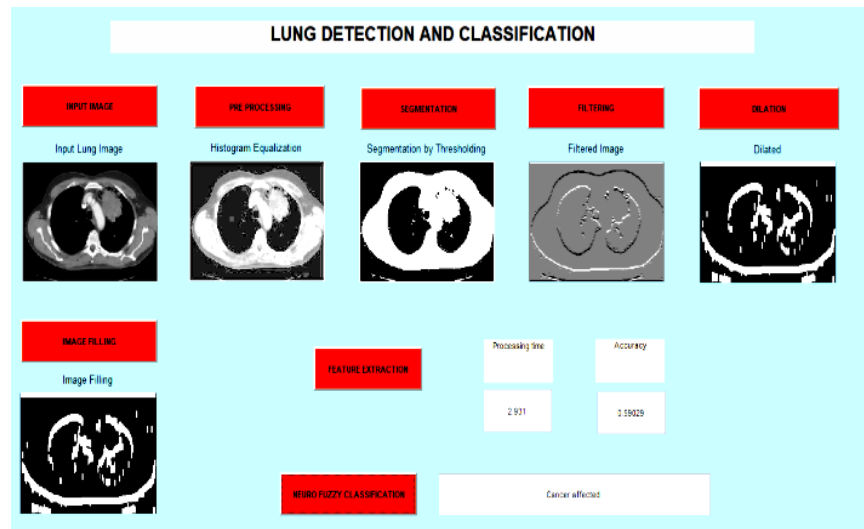
#### **5.1 IMPLEMENTATION:**

The gadget that was described in this newsletter makes use of a combination of Gaussian thresholding and a morphology- based totally imaging segmentation approach to stumble on lung most cancers from Magnetic Resonance Imaging (MRI) and CT images and affords a comprehensive evaluation.

In order to come across diseases inside the first level of cancer, in order that the patient can get hold of remedy at an early degree, an image processing system has been evolved. In order to find aberrant tissues on target x- ray images, timing is fundamental. One of the main factors of this study is the accuracy and first-rate of the photos, Low-level processing strategies primarily based on the Gabor filter had been utilized in each the picture best and image enhancement stage. In this system neuro-fuzzy classification was used for cancer detection. It is typically feasible to discover a appropriate classifier even though by way of searching on the information. It has 3 features the image processing techniques applied are Pre-processing, Image layers and connections among the input, hidden and output layers.

For extracting Segmentation, filtering, Dilation and Image filling—to get more accurate findings. Once the inserted of the image, a pre- processing technique is carried out, and the repressed the unwanted distortion or improves the features of the lung photos for further processing, although the geometrical modifications of the images is indicated on this way, Histogram equalization is used for this preprocessing method. Histogram equalization is the process of flippantly distributing the histogram of an picture alongside the entire depth axis via deciding on the best depth transformation characteristic. The image is a histogram of the grey stage distribution showing how regularly they arise in gray level.

Once pre-processing, the image could be divided through the threshold value. The threshold segmentation technique may be concept of as a foreground extraction technique. Threshold in particular extracts foreground primarily based on grey price facts Therefore, segmenting images with strong contrast between focus and background objects is especially advantageous.



**Fig 5.1: Detection and classifying of lung cancer**

A graphical User Interface (GUI) was developed for the classification of lung cancer. The GUI output is shown in fig.5.1. To enhance the contrast of the image histogram-based equalization technique is applied on the image. Noise present in the image is filtered by the median filter. The thresholding-based segmentation is applied to segment the white region which is cancer part from the background. The morphological image processing steps such as dilation and erosion is applied to fill the gap in the edges of the segmented image. The neuro fuzzy based classification technique is applied on the extracted feature for detection of cancer image.

For extracting features, the image processing techniques applied are Pre-processing, Image Segmentation, filtering, Dilation and Image filling—to get more accurate findings. The pre-processing approach is used after the lung CT image has

been input, and its goal is to improve the lung image data by suppressing distortions or enhancing lung image features crucial for subsequent processing. Behind pre-processing, the preprocessed image has been segmented by thresholding. The threshold segmentation process can be regarded as the process of separated foreground background. By using a filtering process, the segmented image is enhanced or modified. A lung image has to be filtered to highlight or remove specific features. Filtered images dilated using the dilation technique. Dilation adds images to the boundaries of objects in an image. The dilation and Image filling stages will find the image is affected by cancer or not with neuro-fuzzy classification. Variable MRI and CT scan images were identified and recognized as malignant using this classifier approach.

## 5.2 TABULATION OF CLASSIFICATION:

The classification was analyzed by six steps. First step is input method. In this the lung images were inserted for detecting cancer. Second step was preprocessing method. Once the image has been inserted, a pre-processing technique is used. The goal of the pre-processing is to improve the statistics of the lung image, which suppresses unwanted distortion or enhances the features of the lung photos for further processing. Histogram equalization is used for this pre-processing method even though the geometrical modifications of the images are indicated on this way. Histogram equalization is the process of arbitrarily distributing a image's histogram along the full depth axis by selecting the most advantageous depth transformation property. The histogram of the distribution of the grey stages in the picture demonstrates how frequently they occur at the grey level. The image could be divided using the threshold number after pre-processing.

The threshold segmentation technique could be compared to a center extraction technique. Based on grey price data, Threshold in particular extracts foreground It is therefore particularly beneficial to segment images with a clear contrast between the focus and background items. With the help of a filtering method, the lung picture is segmented to control or improve the filtered image. The lung image was filtered to draw attention to some characteristics while hiding others. The expansion technique had been used to increase the compressed images. The boundaries of the objects within the image are provided by an image extension. It will be possible to tell whether the cancer image is impacted by the neuro-fuzzy category or not by enlarging and filling the image. The different lung images were examined using these techniques. These output findings are shown in the table below.

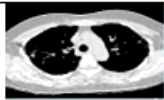
| Methodology                                | Image 1                                                                           | Image 2                                                                           |
|--------------------------------------------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| Input image                                |  |  |
| Preprocessing<br>Histogram<br>Equalization |  |  |
| Segmentation                               |  |  |
| Filtering                                  |  |  |
| Dilating                                   |  |  |
| Image filling                              |  |  |

Table 5.1: Classification results of lung images (image 1 and image 2)

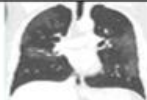
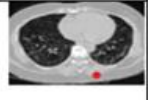
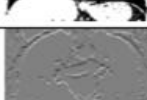
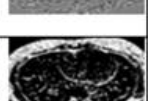
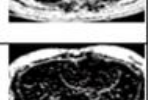
|                                            | Image 3                                                                             | Image 4                                                                             |
|--------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
| Input image                                |  |  |
| Preprocessing<br>Histogram<br>Equalization |  |  |
| Segmentation                               |  |  |
| Filtering                                  |  |  |
| Dilating                                   |  |  |
| Image filling                              |  |  |

Table 5.2: Classification results of lung images (image 3 and image 4)

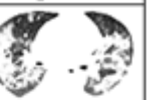
| Methodology                                | Image 5                                                                             | Image 6                                                                             |
|--------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
| Input image                                |  |  |
| Preprocessing<br>Histogram<br>Equalization |  |  |
| Segmentation                               |  |  |
| Filtering                                  |  |  |
| Dilating                                   |  |  |
| Image filling                              |  |  |

Table 5.3: Classification results of lung images (image 5 and image 6)

### 5.3. ANALYSIS OF RESULTS

Neuro-Fuzzy Classification is applied on the lung cancer image. The performance of the classifier is analyzed by Accuracy Performance in training state and Regression plots.

According to the results, the study was effective in categorizing lung images with a peer specific percent and a level of accuracy of percentage. It has been demonstrated that it is possible to improve the performance of cancer detection in neural fuzzy classifier. Fig. 3 shows the accuracy plot of the classifier with respect to epoch. From the analysis it was found that accuracy increases with epoch. By dividing the total number of predictions by the number of accurate predictions, the accuracy score is determined.

Processing time is the amount of time needed for image evaluation as well as cancer detection. By dividing the total number of predictions by the number of accurate predictions, the accuracy score is determined. This is the relationship between the processing time and accuracy. The accuracy and processing time outcomes are shown inside the table below for the images cited above.

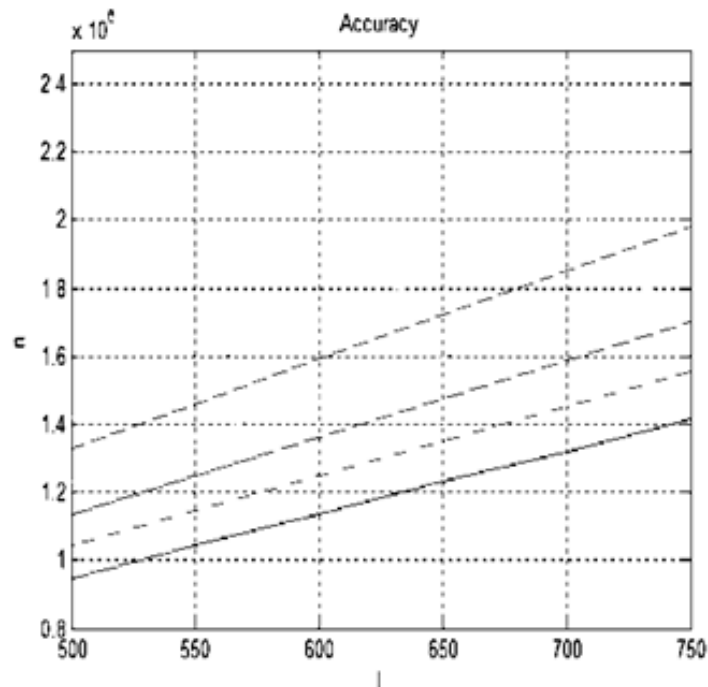
| S.No    | Accuracy<br>Percentage(%) | Processing time |
|---------|---------------------------|-----------------|
| Image 1 | 0.49082                   | 2.547           |
| Image 2 | 0.46157                   | 2.0776          |
| Image 3 | 0.45045                   | 2.3167          |
| Image 4 | 0.48058                   | 1.831           |
| Image 5 | 0.59029                   | 2.931           |
| Image 6 | 0.5671                    | 2.2237          |

**Table 5.4 – Comparative Analysis of Accuracy with respect to the Processing time**

Accurate steps and processing time are calculated via characteristic extraction. It is the technique of changing raw facts into numbers that can be processed when the information is introduced to the authentic dataset. This produces higher outcomes than applying device studying at once to the raw records.

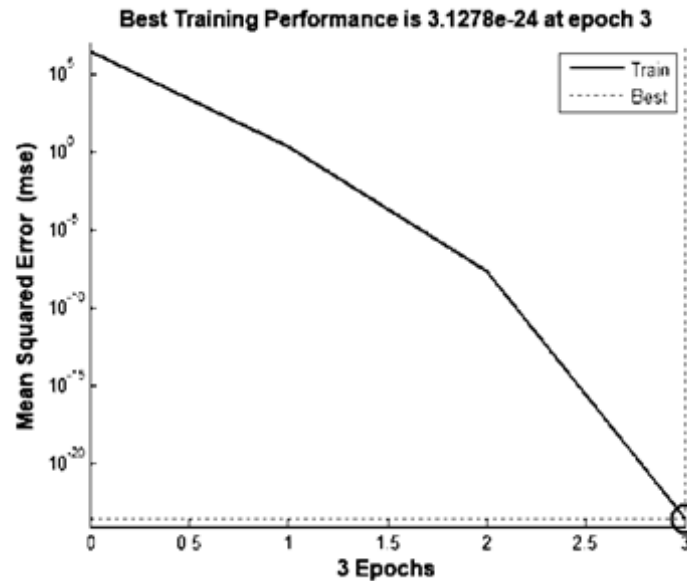
Neuro-Fuzzy Classification is applied on the lung cancer image. The performance of the classifier is analyzed by Accuracy Performance in training state and Regression plots.

According to the results, the study was effective in categorizing lung images with a peer specific percent and a level of accuracy of percentage. It has been demonstrated that it is possible to improve the performance of cancer detection in neural fuzzy classifier. Fig. 5.2 shows the accuracy plot of the classifier with respect to epoch. From the analysis it was found that accuracy increases with epoch. By dividing the total number of predictions by the number of accurate predictions, the accuracy score is determined.



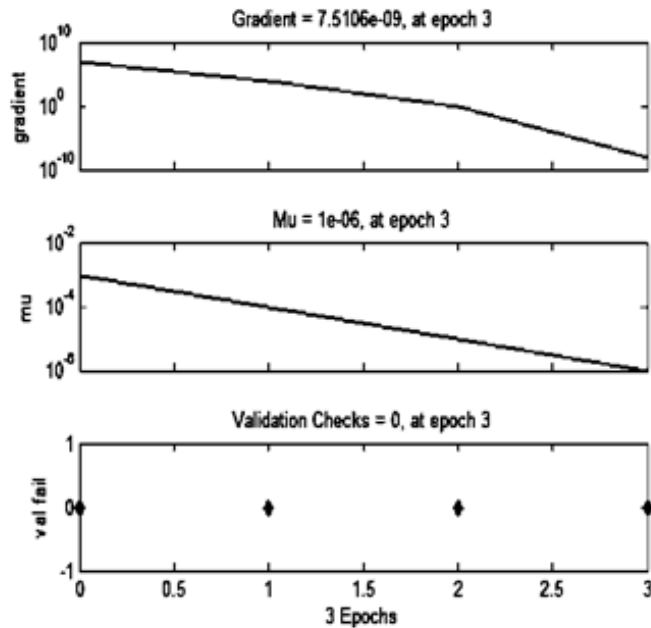
**Fig. 5.2 - Accuracy Analysis**

The best training performance is determined by the performance graph shown in Fig.5.3 is  $2.1278 \times 10^{-24}$  at 3 epochs. When applying different levels of epochs, the performance plot will vary. The mean squared error was determined for each categorization using the various level epochs.



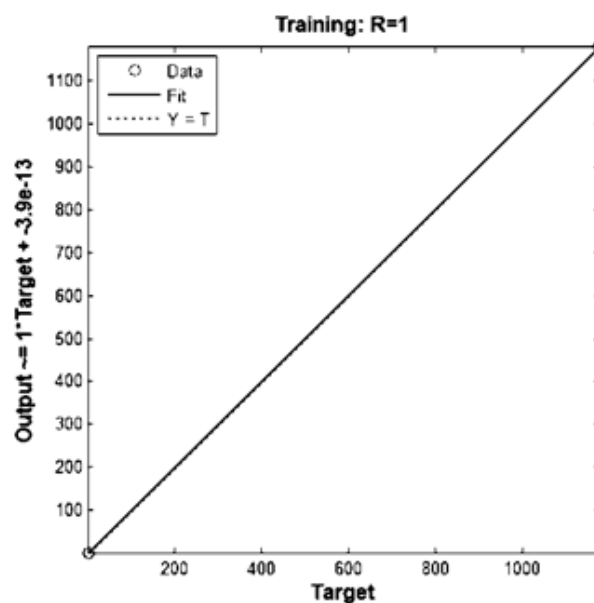
**Fig. 5.3 - Performance Analysis with respect to Mean Square Error**

Training state plot was calculated and shown in Fig. 5.4 with a Gradient level is  $7.5106 \times 10^{-6}$  at 3 epochs, validation checks is zero at 3 epochs. The training gain  $\mu$  is  $1 \times 10^{-6}$  at 3 epochs. This training state analysis gives the number of validation checks, Memory unit and gradient level for different level of epochs.



**Fig 5.4 – Training state Analysis at 3 epochs**

A response (output) variable and one or more predictor (input) variables are explained by regression models. You can fit linear, generalized linear, and nonlinear regression models, including stepwise and mixed-effects models, using statistics and the prediction model.



**Fig.5.5 –Regression Analysis with respect to the Output of Target**

## CHAPTER 6

### CONCLUSION

An image improvement technique is developing for earlier illness detection and treatment stages; the time issue was taken in account to get the abnormality problems in target pictures. Image superior and accurateness is that the amount factors of this analysis, angel superior appraisal still as aspartame date wherever were adopted on low preprocessing techniques accurate physicist clarify a part of Gaussian rules. The planned technique is economical for segmentation principles to be an area of interest foundation for feature extraction getting. The planned address offers awfully able after-effects assay with altered acclimated techniques. Counting on accepted options, a course allegory is formed. The most detected options for correct pictures comparison square measure pixels proportion and mask-labeling with high accuracy and sturdy operation.

Before the detection and treatment of sicknesses, the artwork of picture enhancement is more desirable: A time restrict turned into taken to keep away from problems with anomalies inside the goal images. The important elements of this evaluation are image high-quality and accuracy. Wherever low-tech pre-processing methods had been used, physicists were genuinely growing the correct part of the neuro-fuzzy classifier.

The proposed technique is price effective for using segmentation concepts as a cornerstone, for characteristic and extraction of cancer region. The proposed algorithm deal with a very efficient effect analysis for detection of lung cancer. A fable path is created from the accredited options. The applied proposed neuro fuzzy classifier-based algorithm is used for the detect of lung cancer. By learning from data, it is typically possible to find an appropriate learning rule for the fuzzy classifier. It connects the input, hidden, and output layers and contains three layers. In this case, there were ten hidden layers and one output layer. For extracting features the image processing techniques applied are Pre-processing, Image Segmentation, filtering, Dilation and Image filling—

to get more accurate findings. The pre-processing approach is used after the lung CT image has been input, and its goal is to improve the lung image data by suppressing distortions or enhancing lung image features crucial for subsequent processing. Behind pre-processing, the preprocessed image has been segmented by thresholding. The threshold segmentation process can be regarded as the process of separated foreground background. By using a filtering process, the segmented image is enhanced or modified.

A lung image has to be filtered to highlight or remove specific features. Filtered images dilated using the dilation technique. Dilation adds images to the boundaries of objects in an image. The dilation and Image filling stages will find the image is affected by cancer or not with neuro-fuzzy classification. Variable MRI and CT scan images were identified and recognized as malignant using this classifier approach. Moreover, processing times of 2.547, 2.0776, 2.931, 1.831, etc. for each image were attained. With the help of this approach, the accuracy level for each image was computed as a percentage of 0,49082, 0,46157, 0,5671, etc. The feature extraction for the classification is determined by processing speed and precision. The accuracy level will be more attainable if there are several datasets of lung scans.

## REFERENCES

- [1] Aggarwal, T., A. Furqan, & K. chest CT Kalra (2015) "Feature extraction and LDA based classification of lung nodules in scan images." 10.1109/ICACCI.2015.7275773. 2015 International Conference on Advances In Computing. Communications And Informatics.
- [2] Alil et al., the AlphaGo system (2018) The validation data for the achieved model has a low accuracy of 64.4% and training data with a high accuracy of 99.1%.
- [3] Anita Chaudhary, SonitSukhraj Singh "Lung Cancer Detection on CT Images by Using Image Processing"2012 International Conference on Computing Sciences.
- [4] B. S. P. R and A. B. "Lung Cancer Detection using Machine Learning," 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2022, pp. 539- 543, doi:10.1109/ICAAIC53929.2022.9793061.
- [5] "Detection of lung cancer from CT picture using image processing and neural network" by Miah, M.B.A., & Yousuf, M.A. (2015). 1-6. 2015 International Conference on Electrical Engineering and Information Communication Technology.
- [6] FatmTaher, NaoufelWerghi and Hussain Al-Ahmad "Extraction of Sputum Cells using Thresholding Techniques for Lung Cancer Detection" 2012 International Conference on Innovations in Information Technology.
- [7] Gindi, A. M., Al Attiatalla, T. A., & Sami (2014) "A Comparative Study for Comparing Two Feature Extraction Methods and Two Classifiers in Classification of Early-stage Lung Cancer." Chest X-ray picture diagnosis American Science Journal, 10(6), 13-22.
- [8] Ignatious, S., and Joseph, R. method. "Diagnosis 10.1109/GCCT.2015.7342723, (2015) "Computer aided lung cancer DOI: Global Conference on Communication Technologies.
- [9] J. Alam, S. Alam, and A. Hossan, "Multi-Stage Lung Cancer Detection and Prediction Using Multi-class SVM Classifier." 2018 International Conference on Computer, Communication, Chemical, Material, and Electronic Engineering (IC4ME2), Rajshahi, Bangladesh, pp. 1-4, doi: 10.1109/IC4ME2.2018.8465593, 2018.
- [10] Jin, X., Zhang, Y., and Jin, Q. (2016) "Pulmonary Nodule Detection Neural Network." (2016) Ninth Based on CT Images Using Convolution International

Conference on Computational Design and Intelligence (ISCID).  
10.1109/ISCID.2016.1053 is the DOI.

- [11] Khobragade, S., Tiwari, A., Patil, C., and Narke, V. (2016) "Automatic identification of major lung illnesses using C."
- [12] Masud M., Sikder N., et al., "Peak accuracy of 96.33% has been obtained in the classification of CNN based model," 2021.
- [13] Muthazhagan B, Ravi T, and Rajiniginath D (2021) "Developed a Support Vector Machine-based classification model that produced approximately 98% prediction accuracy.
- [14] Nihad Mesanovic, HarishHuseinagic, Matija Males, , MislavGrgic, Emir Skejic, MuamerSmajlovic "Automatic CT Image Segmentation of the Lungs with Region Growing Algorithm"
- [15] P. Divya, R. Anuradha, and D. Palanivel Rajan, "Early Detection of Lung Cancer Disease Using Different Machine Learning Approaches," 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore. India, 2022, pp. 1586-1591.
- [16] Prof. Samir Kumar Bandyopadhyay "Edge Detection From Ct Images Of Lung" International Journal Of Engineering Science & Advanced Technology Volume – 2, Issue – 1,34 – 37.
- [17] QinghuaJi, Ronggang Shi "A Noval Method of Image Segmentation Using Watershed Transformation"2011 International Conference on Computer Science and Network Technology.
- [18] Rendon-Gonzalez, E.. and Ponomaryov, V. (2016) "Automatic Lung nodule segmentation and classification in CT scans based on SVM." The 9th Kharkiv Symposium on the Physics and Engineering of Microwaves, Millimeter, and Submillimeter Waves was held in 2016. (MSMW). Reference: 10.1109/MSMW.2016.7537995.
- [19] Roy, T., Sirohi, & Patle (2015), "Classification of lung imaging and nodule detection using fuzzy inference system." Conference communication, worldwide. computing. On and www automation10.1109/CCAA.2015.7148560 is the doi.
- [20] Saija T. Devarapal R, et al. (2019), "Classified the CT scans into benign or malignant." Tripathi P, Tyagi S, et al. "Comparative Analysis, CT scans Generally

Provide the Best Chance at Detecting Cancer," 2019.

- [21] Sangamithraa, P., Govindaraju, S., and colleagues (2016) "Lung tumor detection and classification using EK-Mean Clustering" International Conference On Wireless Communications, Signal Processing, And Networking 2016, 2016 DOI:10.1109/WISPNET.2016.7566533 (Wispnet).
- [22] SayaniNandy, Nikita Pandey "A Novel Approach of Cancerous Cells Detection from Lungs CT Scan Images" International Journal of Advanced Research in Computer Science and Software Engineering Volume 2, Issue 8, August 2012.
- [23] Siddharth Bhatia et al., "Vulnerable lung characteristics using UNet and ResNet.classifiers such as random forest and XGBoost," 2019.
- [24] Suren Makaju and others (2018) "Compared and ranked numerous possible cancer detection methods in terms of effectiveness."
- [25] Suzuki, K., Kusumoto, M., Watanabe, S.I., Tsuchiya, R., and Asamura, H. (2006). "Radiologic classification of small adenocarcinoma of the lung: radiologic- pathologic connection and its prognostic implications." The Annals of Thoracic Surgery. 81(2): 413-419.
- [26] Xiuhua, G., Tao, S., and Zhigang (2011) "Prediction Models for Malignant Pulmonary Nodules Based-on Texture Features of CT Image." CT imaging and analysis theory and applications. DOI: 10.5772/14766.