DETECTION OF LUNG CANCER AND PREDICTION OF PATIENT SURVIVAL RATE

Dissertation submitted in fulfilment of the requirements for the Degree of

MASTER OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

By

RAJAN DHARWAL

11510352

Supervisor

Loveneet Kaur

School of Computer Science and Engineering

Lovely Professional University

Phagwara, Punjab (India)

May 2017

PAC APPROVAL

LOVELY PROFESSIONAL UNIVERSITY

TOPIC APPROVAL PERFORMA

School of Computer Science and Engineering

Program : P172::M.Tech. (Computer Science and Engineering) [Full Time]

COURSE CODE : CSE546 **REGULAR/BACKLOG :** Regular **GROUP NUMBER :** CSERGD0274

Supervisor Name : Loveneet Kaur **UID :** 19341 **Designation :** Assistant Professor

Qualification : _____ **Research Experience :** _____

SR.NO.	NAME OF STUDENT	REGISTRATION NO	BATCH	SECTION	CONTACT NUMBER
1	Rajan Dharwal	11510952	2015	K1518	8195051735

SPECIALIZATION AREA : Intelligent Systems **Supervisor Signature:** _____

PROPOSED TOPIC : Detection of Lung cancer and prediction of patient survival rate

Qualitative Assessment of Proposed Topic by PAC		
Sr.No.	Parameter	Rating (out of 10)
1	Project Novelty: Potential of the project to create new knowledge	7.00
2	Project Feasibility: Project can be timely carried out in-house with low-cost and available resources in the University by the students.	6.75
3	Project Academic Inputs: Project topic is relevant and makes extensive use of academic inputs in UG program and serves as a culminating effort for core study area of the degree program.	6.75
4	Project Supervision: Project supervisor's is technically competent to guide students, resolve any issues, and impart necessary skills.	7.00
5	Social Applicability: Project work intends to solve a practical problem.	7.25
6	Future Scope: Project has potential to become basis of future research work, publication or patent.	6.25

PAC Committee Members		
PAC Member 1 Name: Prateek Agrawal	UID: 13714	Recommended (Y/N): NA
PAC Member 2 Name: Pushpendra Kumar Paterlya	UID: 14623	Recommended (Y/N): Yes
PAC Member 3 Name: Deepak Prashar	UID: 13897	Recommended (Y/N): Yes
PAC Member 4 Name: Kewal Krishan	UID: 11179	Recommended (Y/N): Yes
PAC Member 5 Name: Anupinder Singh	UID: 19385	Recommended (Y/N): NA
DAA Nominee Name: Kanwar Preet Singh	UID: 15367	Recommended (Y/N): Yes

Final Topic Approved by PAC: **Detection of Lung cancer and prediction of patient survival rate**

Overall Remarks: Approved

PAC CHAIRPERSON Name: 11024::Amandeep Nagpal **Approval Date:** 05 Mar 2017

4/24/2017 11:21:19 AM

ABSTRACT

Lung cancer is the second most common disease in India. The main cause of this disease is the unawareness in people, symptoms are only recognized in the posterior stages of the lung cancer, and at that stage, and it is very difficult to cure this disease. It is very important to detect the lung cancer at its early stage because survival rate of lung cancer patient decreases with the passage of time. In India, 90% of lung cancer is due to addiction of cigarette smoking. Passive smoking is another cause of lung cancer in which non-smokers who reside with a smoker are at increased risk of developing lung cancer as compared to other non-smokers. Exposure to certain chemicals substances like asbestos, silica and diesel exhaust may also cause lung cancer. Until there are many techniques available for detection of lung cancer but most of them are very expensive.

In this study, we propose a lung cancer detection using back propagation neural network and predict the survival rate of lung cancer patients if the result found positive. We have used the CT scan images of the lung cancer patients. The feature has extracted from these images using GLCM and then for classification purpose back, propagation neural network has used. All this implementation has done with the help of MATLAB.

After comparison, it has found that our proposed system generated the more accurate result as compared to already existing systems used for detection of lung cancer. Based on lung cancer stage the survival rate of patients has predicted.

DECLARATION STATEMENT

I hereby declare that the research work reported in the dissertation entitled "**Detection of Lung cancer and prediction of patient survival rate**" in partial fulfillment of the requirement for the award of Degree for Master of Technology in Computer Science and Engineering at Lovely Professional University, Phagwara, Punjab is an authentic work carried out under supervision of my research supervisor Ms. Research Loveneet Kaur. I have not submitted this work elsewhere for any degree or diploma.

I understand that the work presented herewith is in direct compliance with Lovely Professional University's Policy on plagiarism, intellectual property rights, and highest standards of moral and ethical conduct. Therefore, to the best of my knowledge, the content of this dissertation represents authentic and honest research effort conducted, in its entirety, by me. I am fully responsible for the contents of my dissertation work.

Signature of Candidate

Rajan Dharwal

11510352

SUPERVISOR'S CERTIFICATE

This is to certify that the work reported in the M.Tech Dissertation entitled "**Detection of Lung cancer and prediction of patient survival rate"**, submitted by **Rajan Dharwal** at **Lovely Professional University, Phagwara, India** is a bonafide record of his original work carried out under my supervision. This work has not been submitted elsewhere for any other degree.

Signature of Supervisor

(Loveneet Kaur)

Date:

Counter Signed by:

1) **Concerned HOD:**
   HoD's Signature: _____

   HoD Name: _____

   Date: _____

2) **Neutral Examiners:**

   External Examiner

   Signature: _____

   Name: _____

   Affiliation: _____

   Date: _____

   Internal Examiner

   Signature: _____

   Name: _____

   Date: _____

ACKNOWLEDGEMENT

I am using this opportunity to express my gratitude to everyone to everyone who supported me in this research work. First, I offer my sincerest gratitude to my supervisor "**Ms. Loveneet Kaur**" that has supported me throughout my Dissertation-II. Without her Dissertation-II might not write. I am very thankful for her aspiring guidelines. I invaluable constructive criticism and friendly advise during the research work. I am sincerely grateful to her for sharing their truthful and illuminating views on a number of issues related to the research. Finally, I thanks to my parents for supporting me throughout all my studies at university.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Keywords

ANN – Artificial Neural Intelligent

AI-Artificial Intelligence

BPNN-Back Propagation Neural Network

MATLAB-Matrix for Lab

GLCM-Gary Level Co-Matrix

CT-Computed Tomography

MRI-Magnetic Resonance Imaging

DC-Direct Component

FFT-Fast Fourier Transformer

ACO-Aunt Colony Optimization

CHAPTER 1
INTRODUCTION

Lung cancer is a disease of uncontrolled cell growth in tissues of cells [1]. Lung cancer is a disease of uncontrolled growth of abnormal cells that start off in one or both lungs. The abnormal cells grow rapidly and form tumors. As tumors become larger and more numerous, they underline the lung's ability to provide the bloodstream with oxygen.

1.1 Types of Lung Cancer:

a) **Small Lung Cancer:** Small lung cancer is also known oat cell cancer and this type of lung cancer tends to spread quickly. About 10 to 15% lung cancers are small lung cancer [2].
b) **Non-Small Lung Cancer:** It is a common type of lung cancer. About 85% of lung cancers are small lung cancer. Squamous cell carcinoma, adenocarcinoma are the subtypes of non-small lung cancer.

1.2 Stages of Lung Cancer:

Stages of lung cancer based on whether the lung cancer is local or it spread from the lung to the lymph nodes. The lung cancer can be categories into four stages [3]:

a) **Stage 1: Non-Small Cell Lung Cancer:** Stage 1 cancer is further categories into two sub-stages. Stage 1A and Stage 2B. In stage 1A the size of the tumor is less than 3 centimeters and in stage 2B the size of the tumor is in the range of 3 to 5 centimeters.
b) **Stage 2: Non-Small Cell Lung Cancer:** In this stage, cancer may be present underlying the lung issue but lymph nodes are unaffected. The Stage 2 can be further divided into two subcategories. Stage 2A and stage 2B. In stage 2A, the size of the tumor is greater than 5 centimeters but less than 7 centimeters and in stage 2B, the size of the tumor is more than 7 centimeters.

c) **Stage 3: Non-Small Cell Lung Cancer:** The lung cancer spread into the chest wall. Here, sometimes the tumor is difficult to remove and the tumor may be outside the structure of the lungs.
d) **Stage 4: Non-Small Cell Lung Cancer:** In this stage, the lung cancer spread to the nearby organs of lung-like heart, trachea, and esophagus. In last stage, the cancer cells present in the blood and cancer can grow anywhere inside the body part.

1.3 Cause of Lung Cancer:

Smoking habits are the major risk factor for lung cancer. Tobacco, cigarette these are the substances which damage cells and leads to lung cancer. Out of every 10, 9 people have the lung cancer due to smoking. Rests of them are suffering from lung cancer due to other environment factors as the presence of harmful chemicals in the air.

1.3.1 Symptoms of Lung Cancer:

I. Breathing problem.
II. Pain in the chest area.
III. Wheezing.
IV. Raspy, hoarse voice.
V. Drop in weight.
VI. Bone pain.
VII. A headache [4].

1.4 Image Pre-Processing:

The image pre-processing is a technique to convert the image into digital form and perform some operations on it [5]. We extract some useful features of images, which are useful for further research. Image processing system includes treats image as two-dimensional signals. In image pre-processing process, the image is passing through different techniques to improve the quality of the pictures. In image pre-processing redundancy removal in scanned images, removal of noise and blueness present in the images tasks were completed. This entire task implemented in MATLAB software. There are different preprocessing techniques that we can we used as discuss below:

1.4.1 Median Filter:

It is a nonlinear digital filter used to remove some noise in the image. To detect some edge in the image, firstly noise should be removed up to some threshold value and then edge removal is performed. Hence the median filter is placed before edge detector [6]. Its main feature is it removes noise without edge removal. Median filter is same as that of averaging filter, in which each output image pixel is set corresponding to the average value of neighboring pixel of the input image. The median filter is more sensitive to mean values and less sensitive to extreme values of pixel which helps in noise reduction.

1.4.2 Histograms Equalizers:

Histogram equalization redistributes the intensities of the image of the entire range of possible intensities [7]. Histogram equalization has been to enhance contrast. It is not necessary that contrast will always be an increase in this. There may be some cases were histogram equalization can be worse. In those cases, the contrast is decreased.

1.4.3 Image Enhancement Technique:

Image enhancement technique is of two types, spatial domain method and frequency domain method [8]. The first method deals with pixels where enhancement is achieving by changing the pixel's values and the second method deals with changing the orthogonal of the original. Image enhancement is the first step in image preprocessing process. The image enhancement process performed to improve the interpretability and better image, which used in the next steps. There are a number of enhancement techniques that we can we use in image enhancement process.

a) **Gabor Filter:** Dennis Gabor uses this process for 2-D images. A Gabor filter is a linear filter whose impose response is defined by a harmonic function multiplied by a Gaussian function [9]. Gabor filter has a significant role in image preprocessing process as it uses for fingerprint recognition and iris recognition. The enhancement percentage in this technique was 80.735%.
b) **Fast Fourier Transform:** This is another transform technique be used in image enhancement technique. It produces a complex number valued output that has

represented with real or imaginary part image with its magnitude and phase. The enhancement parentage is in this technique is 27.51% [10].

c) **Log-Gabor Filter:** Log-Gabor method had proposed by Field in 1987. This is an advanced version of Gabor filter method [9]. This method has two characteristics, first property is that it has no DC components and second log function has extended tail at high- frequency end.

1.4.4 Image Segmentation:

Marker-driven watershed segmentation technique extracts seeds that indicate the presence of objects or background at specific image locations. Marker locations are then set to be regional minima within the topological surface and the watershed algorithm is applied [11]. Separating touching objects in an image is one of the most difficult image processing operations, where the watershed transform is often applied to such problem.

1.4.5 Features Extractions:

Feature extraction technique is performing to detect and isolates various shapes or desired portion [6]. Features have extracted according to the requirements. There are a number of features extracted techniques that can used on images. The features are extracted to check the normally and abnormally of the pictures. The features, which generally are extracted, are area, perimeter, and average intensity

a) **Binarization Technique:** This technique work on quantity of white pixels presents in the lung cancer images [12]. In the first step, the numbers of white pixels are counted. To find lung cancer stage i.e. normal or abnormal the numbers of white pixels are counted and if the number of white pixels greater that then the threshold it means that lung cancer stage is abnormal else it is in normal stage. The threshold value is 255.

b) **Gray Level Co-Occurrence:** This technique is very useful in features extraction technique. The radius and angle are the crucial input parameters of GLCM [13]. The radius can obtained from the autocorrelation function of the image and no definite result can draw regarding the value of angle. This approach has two steps. In the first step we create gray level co-occurrence matrix from the image in

MATLAB and after that, we normalize the GLCM using some formula. The fractures extracted in this method are entropy, energy, homogeneity, maximum probability, and contrast.

1.5 MATLAB

MATLAB is a short name for matrix laboratory. In the current scenario for the implementation of proposed methodology Matlab version (R2013a) is used. MATLAB provides an environment for the development of algorithms, analysis of data, visualization and numerical computation. It performs many computation-intensive tasks with considerable high speed. MATLAB is used for image processing, noise removal method, and filter method. We can train and test our code by using Matlab.

In MATLAB, variables are present in a "workspace" that correlates variables names and their values. A global workspace has defined global variables. MATLAB provides for two types of reusable code units i.e. scripts and function. Scripts take no particular input or parameters, operating directly on the caller's workspace. The caller can be either a function or the global workspace. On the other hand, functions have several input/output parameters. These parameters remain bound to the function's workspace. A symbol unbound in a function evaluated but the global workspace. It has a rich toolbox, which is the collection of various functions the above features are the main reason for opting MATLAB.

Requirements of MATLAB in systems are:

Space: I GB for Matlab.

Version: Window 8.

RAM: 1024MB.

Processor: Any Intel or AMD x86 processor supporting SSE2 instruction set (Here, Intel Core™ i5 Processor).

1.6 Image Processing Toolbox

Image Processing Toolbox in Matlab provides a wide range of functions, algorithms, and applications for image analysis, processing, and algorithm development. Image processing toolbox makes easy to carry out image analysis, image segmentation, noise reduction and performs a geometric transformation on images.

The key features of this toolbox are as follows

a) Image analysis, including image segmentation, image restoration, image properties.
b) Image enhancement, noise reduction, filtration.
c) Image transformation, including Fast Fourier Transform (FFT), DCT and Radon.
d) Geometric transformation and Visualization applications.
e) Multicore and GPU enable functions and C-code generation support.

1.7 Artificial Neural Network

Artificial intelligence is a branch of computer science that leads to the creation of the intelligent system, which can work and reacts like humans. It is a non-linear computational system inspired by the structure, behavior and learning abilities of a biological brain.

The term “neural network” resulted from the research in artificial intelligence, which attempts to understand and model human brain functionality [14]. Artificial intelligence is a part of computer science, which focuses to the creation of intelligent machines, that is, systems that work and behave like humans. From recent few years, artificial neural networks have proved themselves as a better alternative for solving complex problems in many areas. Input layer, hidden layers, and output layer are the layers of Artificial Neural Network. Hidden layers perform an intermediate computation to produce required output from the various inputs received. For pattern recognition applications, the efficiency of the neural network depends on the learning algorithm adopted. In supervised learning, the correct answer is provided for every input to the network and Un-supervised learning in which result is derived from prior assumptions and inferences; however, system and hybrid learning do not know the correct result, which contains both supervised and unsupervised learning.

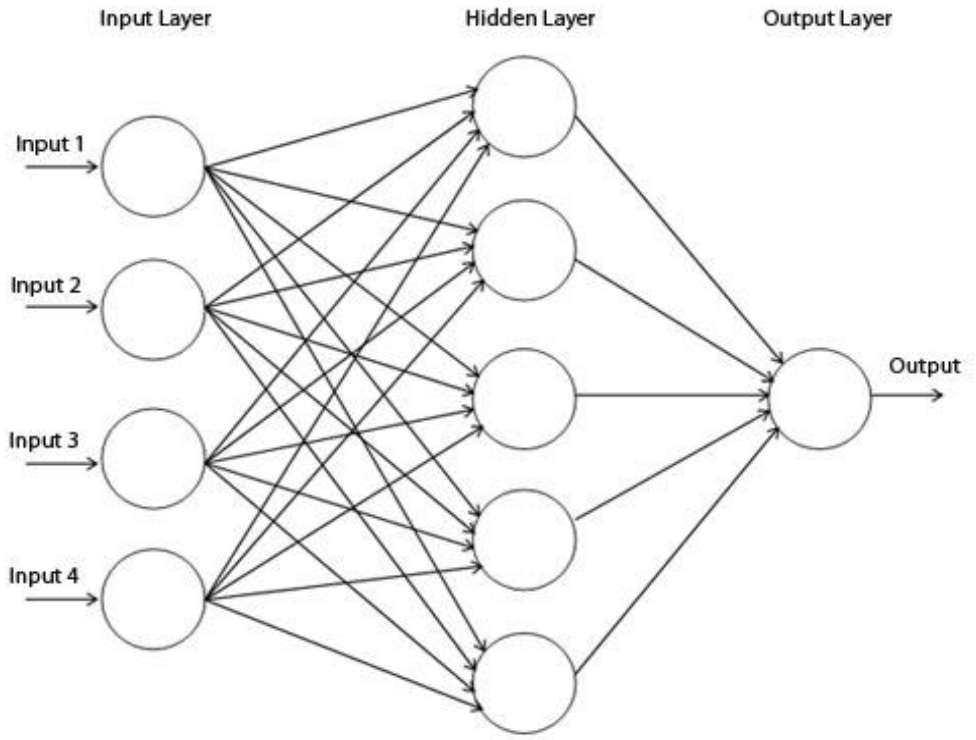


Figure 2.1: Neural network

1.8 Biological Neural Networks:

Human brain contains a billion of cells, which interconnected to receive process and transmit information. Each cell works as a simple processing unit and are core components of nervous system. Interaction between these simple processing units, the brain manages to perform extremely difficult tasks. Biological neural networks made up of these real biological neurons, one-neuron makes connections with other neurons and signals are propagated between them. These signals control the activity of brain in central nervous system. Artificial neural networks are composed of interconnected artificial neurons, modeled on the interconnecting biological neurons in the human nervous system. Artificial neural networks used to understand the behavior of the human brain or to solve complex problems. Artificial neural networks used to understand the behavior of the human brain or to solve complex problems.

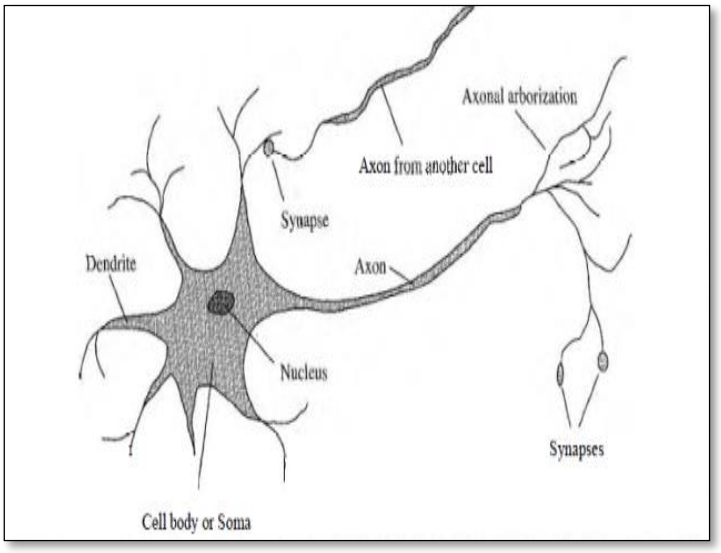


Figure 2.2: A Biological neuron

1.8.1 Elements of Artificial Neural Networks:

The fundamental building block of artificial neural network is neuron. A simple neural network consists of three layers, activation function, learning technique, and weights. All these layers consist of neurons, which were interconnected to form a network.

a) **Input Layer:** Input layer is feature vector of the problem. The extracted features are passing to the input layer of the neural network. It has the value in numeric form. Feature vector should select very carefully. It should be able to describe the information properly and it should not contain any redundant information or faulty data, otherwise, it will decrease the overall performance of the network.

b) **Output Layer:** Output layer takes processed data from hidden layers, give results to the external user. Here an output layer depends on the outcome of the problem. For example, if the problem is of classification then the output layer will give a class number.

c) **Activation Function*:*** Activation function exhibits a great variety and has a strong effect on the behavior of the neural network. The sum of weighted inputs was applied with an activation function to obtain the output response. The activation function is an abstraction representing the rate of firing in the cell. For neurons present in the same layer, same activation function is used.

Most commonly used activation functions are:

- **Linear Activation Function:**

These functions are most commonly used in single layer networks.

The expression for identity function is given by Equation

a. $F(x) = x$ for all x

The output is continuously varying value as shown in Figure.

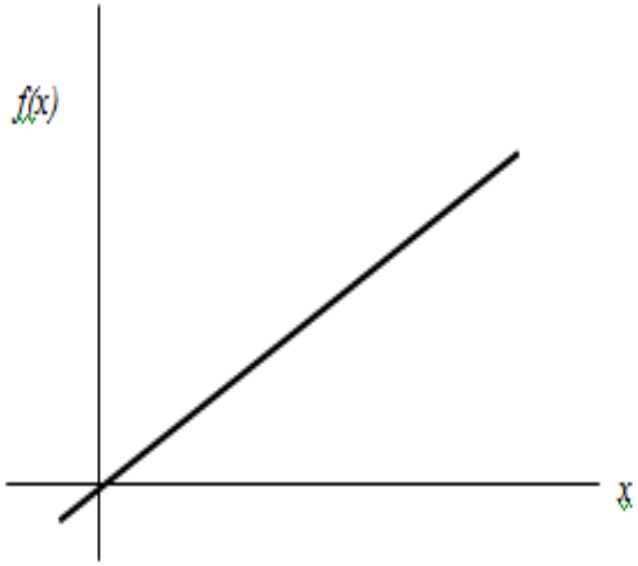


Figure 2.3: Identity activation function

a) Binary Step Activation Function:

The mathematical representation of binary step activation function is given by Equation.

b. $F(x) = 1$ if $x \geq \theta$
$= 0$ if $x < \theta$

Where, θ indicates the threshold value.

Figure shows its variation of output of function with input values.

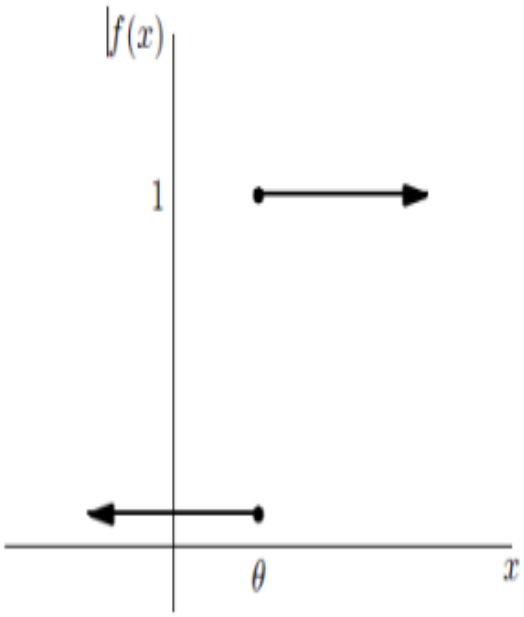


Figure 2.4: Binary step activation function

- **Log Sigmoid Transfer Function:**

This function yields output between the range 0 and 1. The output is calculated according to Equation.

c. Logs (n) $=\frac{1}{1+\exp(-n)}$

- **Hyperbolic Tangent Sigmoid Activation Function:**

This transfer function yields result scaled between -1 and 1. The hyperbolic tangent function used with gradient descent based training methods as it has a derivative. The output is calculated according to Equation.

d. n $= \frac{2}{(1+\mathrm{EXP}(-2*n))} - 1$

1.9 Learning in Artificial Neural Networks:

A learning system changes itself in order to adapt to various changes, e.g. environmental change. A processing unit is can change its input/output behavior in accordance with the changes in environment. Learning method also required so that during the training phase, weights has modified in response to input/output changes. In artificial neural networks, learning can be supervised, unsupervised and reinforced.

a) **Supervised Learning:** Supervised learning is also known as learning with the teacher as an external teacher controls learning and incorporate global information. In supervised learning, the neural network is trained by providing both, the input vectors and the expected outputs. The difference between actual output and desired output serves as error measure. Figure 2.5 shows the block diagram of supervised learning. At each instant of time, the external teacher provides input X is applied to network and desired response of the system. The distance between actual network output O and desired response deserves as error measure. The main objective of supervised learning is to reduce the difference between actual output and desired output.

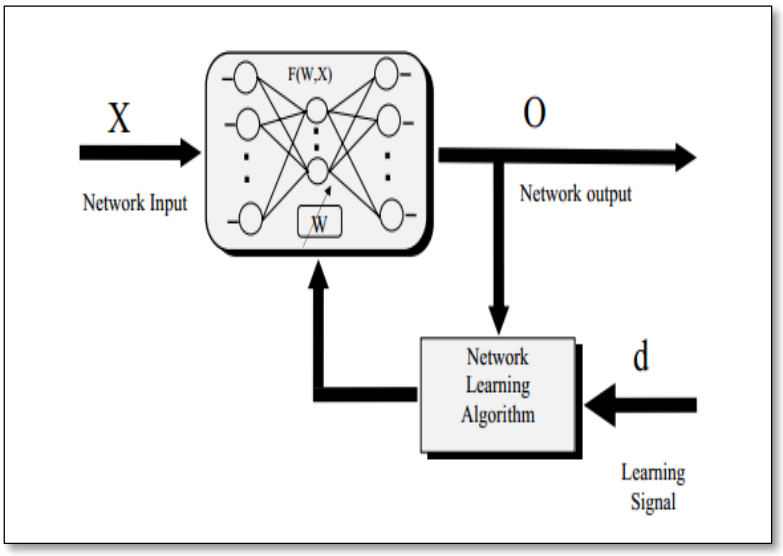


Figure 2.5: Supervised learning

This is achieving by adjusting connection weights so that the network is more likely to produce output close to the desired value. The weights adjusted continuously until the performance is satisfactory. Supervised learning is majorly using in prediction and classification problems.

b) **Unsupervised Learning:** In this type of learning, only inputs are given and no desired output of the neural network. It also called as learning without a teacher because there is no external teacher. Here learning, the network must discover any

regularity, separating properties, etc. and network changes its parameters accordingly to produce optimal results. Unsupervised learning is often used to perform clustering, i.e. unsupervised classification of input without any prior knowledge about actual classes. Figure 2.6 shows the block diagram of unsupervised learning.

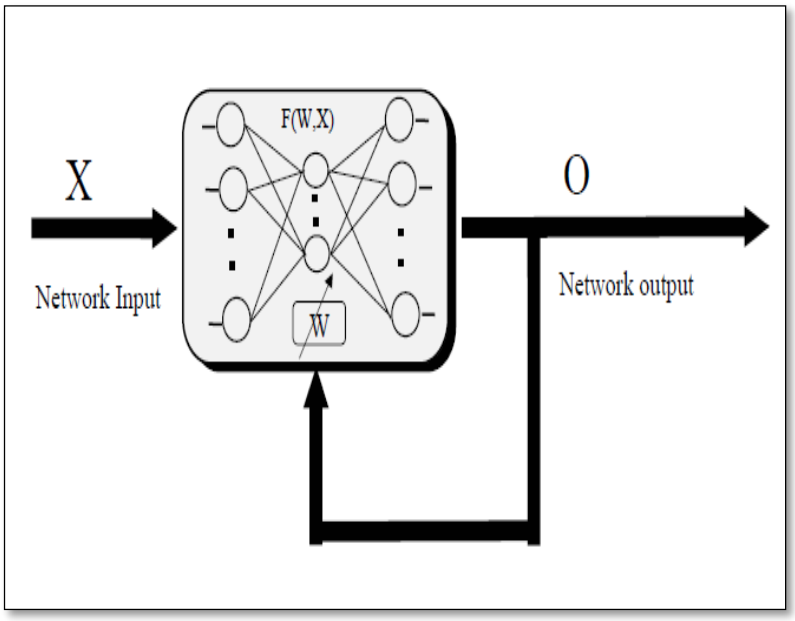


Figure 2.6: Unsupervised learning

c) **Reinforcement Learning:** In many situations, unlike unsupervised learning, less detailed information is available and reinforcement learning is method is use to deal with such situations. In Reinforcement learning where the network gets some feedback from the environment. The feedback signal is in the form of only evaluative (yes/no) not instructive, i.e. the feedback signals only specify whether the output is right or not. If the feedback signal says that output is wrong, then it does not give any hint as what the right output should be.

1.9.1 Neural Network Architecture:

There exist various neural network architectures. They differ from each other in many factors such as a number of hidden layers, learning techniques, activation functions, etc. The two mostly used neural network architectures are:

a) **Feed Forward Neural Networks**: These networks allow only one direction signal flow, i.e. from input nodes, through hidden nodes to output nodes. There is no feedback or loops in feed-forward neural networks. In these networks, like any other neural network perceptron are organized into layers. The input and hidden layer must be connected and hidden layer can be connected to another hidden layer or output layer. The information is "feed forward" from one layer to another hence, called feed forward neural network.

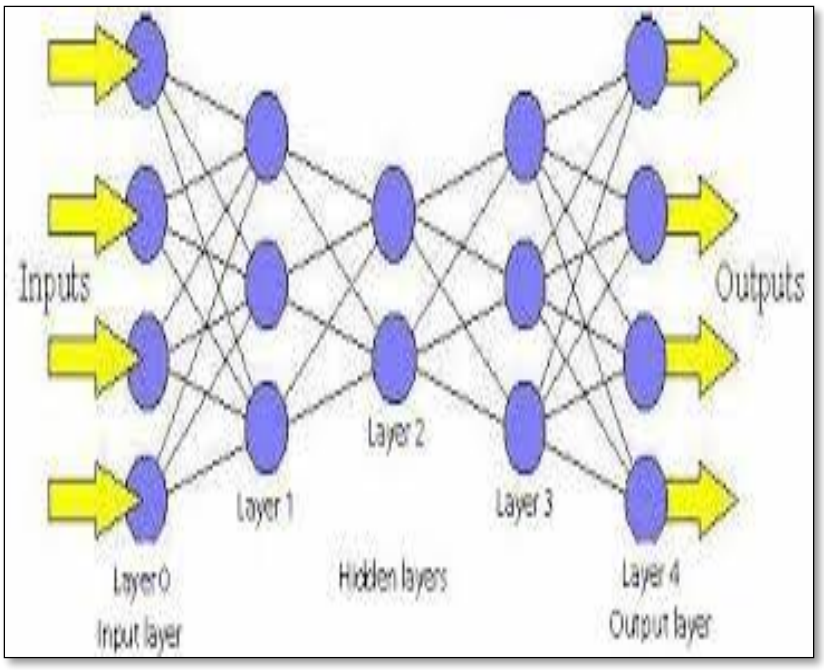


Figure 2.7: Feedback / recurrent neural network

Feedback Neural Networks: Feedback neural networks allow signals traveling in both directions by introducing loops in the network. There are feedback signals traveling from one layer to another. Feedback or recurrent networks are dynamic in nature, i.e. they change their state continuously until they get a satisfactory response.

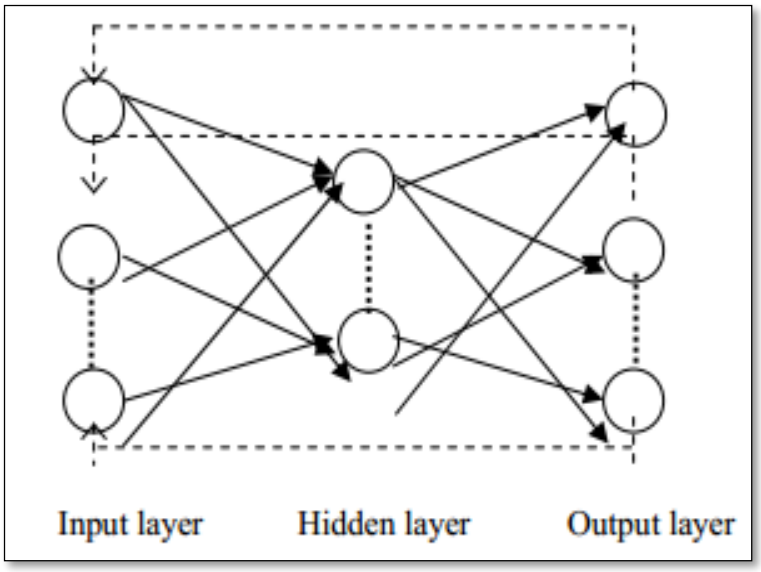


Figure 2.8: Recurrent/ feedback neural network

b) Feed Forward Back Propagation Neural Network: This network is the combination of two neural network algorithms. In "feed forward" refers where the networks will be recognizing a pattern and the term "back propagation" describes where the networks will be trained. In other words, "feed forward," describes how neural network processes and recalls patterns.

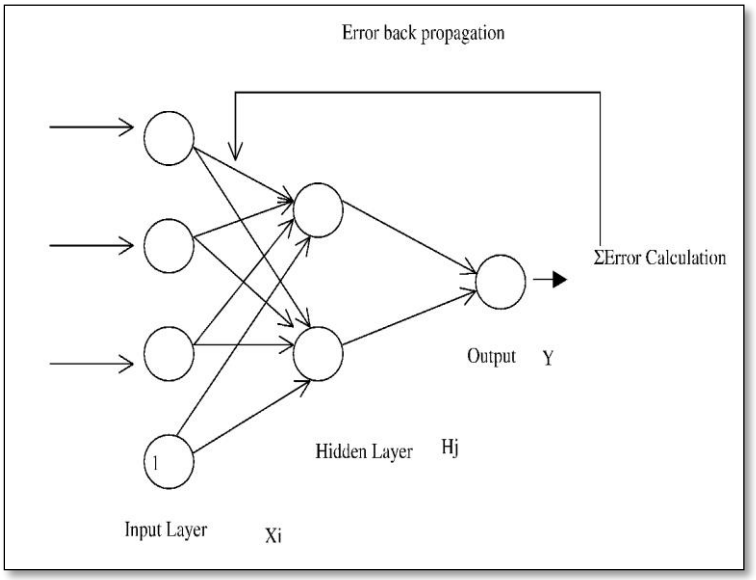


Figure 2.9: Feed forward back propagation neural network

Back propagation is a form of supervised training, i.e. network must be provided with input as well as desired output. The desired outputs are comparing with actual outputs to compute errors. Back-propagation is a method, which takes calculated error and altered in a way that causes the error to be reduced.

1.10 Applications of Artificial Neural Networks:

I. **Neural Network in Medical Area:** The Artificial neural network is very popular in medical science. It can be used for patient diagnostics. [10] In medical for cost estimation of a patient's treatment neural network can be used. For forecasting the body, temperature, and working of body neural network used.

II. **Recognition of Characters:** Character recognition has become very important in these days. It is used in devices, which used hand touch for processing like the optical character recognition. These devices are now becoming popular. Neural networks can be used to recognize handwritten characters. It is also used for checking OMR sheets and papers, which is a popular method.

III. **Neural Network in Image Compression:** Neural networks can accept and use huge information at a particular period and making that data valuable in image compression. People can use web exploration and other sites for adding further images on their sites, using neural networks for image compression.

IV. **Stock Market Prediction:** The regular business of the stock market is very complex. There are many for checking whether a given stock will rise or down on any given time period. Then neural networks can scrutinize a portion of statistics fast and sort it all out, by prediction of stock market prices.

V. **Neural Networks for Solving Business Problems:** The Neural network is also used for solving the sales man problem like checking trends etc. However, the result of this is approximate not 100% correct.

VI. **Neural Network in Energy:** Neural network can use in various areas of energy. It can be used to forecast electrical load, the demand of energy. In Hydro dams and power control systems neural network can use. For estimation of coal and load, neural network can use.

CHAPTER 2
REVIEW OF LITERATURE

MR Zali *et al.* (2011) studied about gastric cancer i.e. was the malignant growth of stomach which can develop any part of the stomach [15]. It is the fourth common type of cancer that occurs in worldwide. He developed Artificial Neural Network system by analyzed the data of 436 patients with stomach cancer from a private organization to find survival rate of gastric lung cancer. The survival rate of lung cancer was formed by using COX propagation hazard. The author concluded that the survival rate of gastric cancer patients between years 2002 to 2007 was respectively 77.9%, 53.1%, 40.8%, 32% and 17.4%.

F. Feng *et al*. (2012) evaluated the diagnosis potential of Artificial Neural Network model that combined with six tumor markers in auxiliary diagnoses of lung cancer [16]. The Artificial Neural Network has trained with 242 samples including 93 lung cancer patients, 66 lung benign disease patients and 83 healthy people from a private institute. 19 parameters were used there like smoking, dust exposure, cooking fuel, chemical exposure etc. to differentiate the lung cancer from lung benign cancer and normal controls and sensitivity, specificity and accuracy were found 98.3%, 99.5%, and 96.9% respectively. Another model with six tumors markers used to distinguish between lung cancer from gastric cancer by using three Artificial Neural Networks and sensitivity, specificity and accuracy has found 100%, 83.5%, and 93.5% respectively.

Ismail Saritas *et al*. (2012) used back-propagation algorithms in two phases feed forward and back propagation procedure [17]. It is used to predict whether the patient has breast cancer or not if yes determines its types. The author first performed CT scan of the images of patients then features are extracted from these CT scan images and final classification was made using ANN. The classification was made in term of the image had cancerous or not. During feed forward, the information was given through input layers to output layer and in back propagation, the value of the output layers and the output value was compared and the result is calculated. The type determines by the use of TIRADES SHAPE, mass

density, and mass order. The disease prediction, in this case, was 90.5% and health ratio was 80.9%.

S. Neghibi *et at.* (2012) purposed that early detection of lung cancer could reduce the significant number of breast cancer [18]. The author used Wisconsin breast cancer database, downloaded from UCI Machine repository. 699 original samples were introduced. Two classes derived from the dataset, 458 benign samples and 241 were malignant samples. The hierarchical fuzzy neural network and fuzzy Gaussian potential neural network used there. The nine features have defined in breast cancer used as an input. Here author concluded that result forms here sensitivity, accuracy and specificity were more accurate than the fuzzy neural networks and the number of rules was reduced that used in hierarchical neural networks and fuzzy Gaussian potential networks.

Rajneet Kaur *et al*. (2013) used the CT scan images of the lung cancer patients from a private hospital [19]. 900 CT scan images of lung cancer images had taken from a private hospital. The author-performed experiment was in mat lab. Pre-processing of images was done and then the classification of the lung cancer patients in the normal and abnormal state was determined. GLCM and BINARIZATION methods had used for features extraction purpose. The supervised feed-forward back propagation network was used as a tool for classification purpose and the author concluded that after 5 years of lung cancer disease the survival rate of the patient was 40%.

Ranjeet Kaur *et al.* (2013) determined the survival rate of the lung cancer [20]. The author improved the survival rate of the lung cancer previous experiment from 14% to 49% of five years of lung cancer patient. First pre-processing of images of lung cancer patients have done for image extraction purpose, 160 features have extracted, and lung field segmentation method was performed on the hidden part of the lung area. The classification has made with the help of neural networks and SVM, for detection and classification of the lung cancer.

K.Ahmed *et al.* (2013) purposed the environment factors had a significant effect on the lung cancer [1]. The author initially took the data of 400 cancer patients from different centers. The pre-processed and clustered of data was done by the used of k mean algorithm.

400 patients of that which contained 200 lung cancer patients and 200 non-lung cancer patients have taken. There were 200 male-female patients whose age between 20 to 80 years. The author developed lung cancer prediction tool based on data mining that represented the stage of lung cancer patients out of stages 3 and conclude that diagnosed of the patient at the last stage mostly impossible so, early prediction of lung cancer should play a vital role in the diagnosis of lung cancer.

M.Assefa *et al.* (2013) In this paper the author concluded that the main reason for lung cancer is smoking habit and survival rate of the patient 15% if the patients got this disease [21]. The author focused on directly Lancaster traction system rising City scan of lung cancer patient for this purpose author used a template-matching algorithm with multi-resolution analysis. The authors took CT scan of 165 images and out of those 134 images has classified correctly and accuracy of the system found 81%.

R. Manickavasagam *et al.* (2014) evaluated that one out of three people has affected by lung cancer [22]. Lung cancer was the leading cause of death related to cancer in many countries. Only 14% people can survive from lung cancer. The author used Imagining techniques, which combined both positrons emission tomography scanner and computer topology scanner in a single system, so the images gathered taken as consecutively and combine into a single superpose image. The noise has removed using a median filter. The extracted features were stored and different classifications of cancer were determined and those classifications used by the physician to give some therapy suggestions.

Prashant Naresh *et al.* (2014) focused on detecting cancer at the early stage [13]. First, the author removed the Gaussian noise from the CT Scan images of lung cancer patients with the help of none- local mean filter and to segment the lung Otsu ‘s threshing. To form feature vector, the textural and structural features has formed from processed images.3 classifiers SVM, ANN, KNN were used for detection of the lung cancer to finding the different stages of the lung cancer and comparison was made between SVM and KNN with respect to accuracy, sensitivity, and accuracy. The author concluded that SVM, ANNN, and KNN had an accuracy of 95.12%, 92.68%and 85.37% respectively for tested images.

Mr.vijay *et.al* (2014) according to author the lung cancer disease is the most common disease in the world. Around 85% males and 75% females have cancer to cigarette smoking [23]. The overall 5 yearly survival rate of lung cancer patients can be increased from 14% to 49% if the disease is found in time. The author used MATLAB for this implementation where image-preprocessing processes have been the applied on the CT scan of the lung cancer. The author proposed the detection of a lung nodule. The first author performs noise filter method on CT scan of lung cancer to remove the noise from the images. He used the watershed algorithm, which is commonly used for the unsupervised setting of segmenting an image into non-overlapping regions. The major step was thresholding where author represent the results of images is in normal form or in moral form. Here he used the rule-based approach for the thresholding, which includes extraction process, edge detection, functions or dynamic programming. Author extracted three features for identifying the nodule size, which is an area, perimeter, and eccentricity, and classification has been made by using the SVM model. The author described the stages of lung cancer into four stages.

Sukhjinder Kaur *et al*. (2015) evaluated among various diseases, cancer becomes the major thread to India [24]. In India, 90% of the people have cancer who consume tobacco and only 10% of the people have cancer disease who did not smoke cancer. Early detection of lung cancer is a key challenge. The author analyzed the two techniques neural network and C-mean clustering algorithm. The author concluded that out of these two techniques the neural networks was best to use because the clustering algorithm was not good at low-intensity variations.

S. Avinash et*.al* (2015) In this paper the author gave more intention on early detection of lung cancer using different techniques [25]. In this paper to overcome the various limitation of the Phoenix technique, the author used to give a method and watershed algorithm reproduction of lung cancer cable method is used as ethnic and for segmentation of images naked controlled watershed is used result will calculate you protect email is normal or not.

Parmjit Kaur *et al*. (2015) focused on the early stage of the lung cancer. The author took 909 real CT scan of images of lung cancer images [26]. The classification was made using neural networks. Classification made that lung cancer is in normal or abnormal form. After the CT scan of the images, pre-processing of images was done and training and testing

were performed along with features extraction. Then classification was done with back propagation neural network and GA algorithm. The author concluded the back propagation neural had the best accuracy to find the early stages of lung cancer after it is passed through GA algorithm. The result was in form of the cancer is the normal or abnormal form.

M. Hussian *et al.* (2015) the lung cancer disease causes a lot of death in the world. The cure of the diseases depends upon its initial stages [27]. Lungs have usually large in size, so tumors can grow before it can be recognized. Therefore, early stages of lung cancer stage 1 and stage 2 are difficult to find. The main objective was to found detection of lung cancer before it can grow a significant level. Images were pre- processed, features have extracted, and final classification of cancer has made by providing the images as the input to the system with neural networks. The classification has made to identify the different stages of the lung cancer.

A. Dharmarajan *et al.* (2015) The author identified the normal and abnormal stage of lung cancer by used of CT scan of the image [28]. The author used watershed segmentation and thresholding for images segmentation process and watershed technique showed more efficient than the other techniques. The author concluded that lesion size of lung cancer cell is 20mm for normal lung cancer and more that 20mm as abnormal lung cancer.

P. Eskandarian *et.al* (2015) In this paper the author generated the system that diagnoses the lung cancer disease by detecting the nodule size present in CT scan of lung cancer. For this purpose, author used CAD system of lung cancer system [29]. MATLAB used for implementation, features were extracted like area, perimeter, eccentricity, classification was made by using SVM, and author concluded the accuracy with this system 81.212%

S. Logesh *et al.* (2015) focused on the clustering algorithm that can help to identify the different stages of lung cancer and improvement in medical care [11]. K-mean and farthest first clustering algorithms used by the author and analyzed their performance on the different dataset values. The implementation done with WEKA software and author concluded that k-mean algorithm was more efficient for lung cancer dataset with arff format.

MR. Kishore *et al.* (2015) In this papwe the author focus on better feature extraction technique which is used to give better result of lung cancer detection. Implementation was performed by-using MATLAB [30]. According to author, the early detection of lung cancer disease of the patient can increase the survival rate from 14% to 49%. In this research paper, author focus on the selection of right features from lung cancer because with selection of accurate features he images can classify in better Way. Aunt colony algorithm was used for this purpose; here the author proposed the five step for the detection purpose. The segmentation algorithms have used to detect the cancer nodule present in the CT scan of the lung cancer and rule based learning technique is used for classification of cancer nodule and according to extracted features, the diagnosis rules are generated. The author gave more attention on CT scan of lung cancer rather than the MRI of the patients because CT scan of patients contains low noise and distortion. Here features selection technique was used to reduce the irrelevant portion and redundant portion from the image which has been not required. This technique was used to remove the missing values and filtering the most required feature from other features extracted and used to assign the ranking order to the features respectively.

Md. Badrul Alam Miah *et.al* (2015) In this paper the author proposed the detection of lung cancer in its prior stat, the author took the X ray of images of lung cancer and image processing was performed to remove unrequited parts from the images [31]. The binarization and segmentation process has performed on the images, in thresholding method was applied on the images in which grayscale image is converted into a binary image. After author used that MFFNN with SVM, from this system the author found 96.67% accuracy get this system.

S. Ignatious *et.al* (2015) In this paper the author used CT scan of images and classified the CT scan of images according to their TNM state, the author identified the T state and classified according to their state [11]. Segmentation process performed used of watershed algorithm. The author used to 200 CT scan of lung cancer images from regional Cancer Centre Trivandrum and the accuracy of a force system found 94.4% near the author use features like area perimeter Area, the eccentricity of lung cancer nodule. The random tree used for classified.

S. Singh1 *et.al* (2016) In this paper the author used Matlab software for detection of lung cancer disease [32]. The author took data from a private organization and classified the images of lung cancer in terms of normal or abnormal CT scan of the images the pre-processing of images performed to remove the unwanted parts from the mages. Histogram equalization method has used for modification of the images. Features have extracted from the images like area, eccentricity, perimeter, complexity etc. At last, the author classified the lung cancer images in terms o normal or abnormal images.

MR.Sangamithaa *et.al* (2016) In this paper the author focused on image segmentation process of lung cancer images because lung cancer images have complex properties. The author used image processing technique as remove non-required parts of the images and segmentation process was carried out using k-means clustering method and after features are extracted using by using GLCMS method for the classification purpose the author used the back-propagation neural network. This helps to determine whether the images have cancer or normal the accuracy found by using this neural network is 90.87% [6].

E.Ponomaryov *et.al* (2016) in this paper the author gave more emphasis on early detection of a lung cancer disease [33]. The author used CT scan of the lung cancer images for this purposed. Detection of a lung cancer process has 4 steps image processing, parenchyma, segmentation nodule detection and reduction of false positive. The classified has made by using SVM. The sensitivity, specificity, and accuracy of this proposed system have found 84.93%, 80.92 and 78.08% respectively.

C. Yang *et al.* (2000) studied the applications of Artificial Neural Networks in image recognition and classification of crops and weeds [34]. The back propagation ANN model was used to differentiate between corn plants from weeds. 80 images including 40 images each of corn plants and weeds were used to provide for training purpose. Kodak DC50 camera was acquired for digitals images of corn plants and weed from McDonald Campus Farm of McGill University.

Aman Tayagi *et al.* (2011) highlighted the development of low-cost temperature and soil moisture sensor, which was placed in the fields and if the soil temperature went high and soil moisture, fell low a described limit the water was sprinkled on to the fields [35]. LM-

35DZ sensor was used as a water sensor and Probes that was made up of two metal rods tied together using an insulating tape was act as a moisture sensor. The sensor modeling to compensate nonlinearity and temperature the Artificial Neural Network to line arise voltage output was used. The ANN consist one input layer, one hidden layer, and one output layer. Hidden layer consisted of 4 neurons and output layer consists of 2 nodes. The error rate was concluded less than 1% that indicated the high accuracy of this system.

P.S. Chaudhari*et al.* (2012) evaluated that three layers Back propagation neural network along with LM algorithm to construct a mathematical model that was used in the laser cutting parameters [36]. A genetic algorithm was used to evaluate the best prefers a combination of laser cutting parameters for optimized a good cutting quality.

Karishna *et al.* (2013) explained the advantages of Artificial Intelligence in the business area [37]. Artificial intelligence system used to make an automatic decision, which is helpful for companies to using their limited resources to gain more benefits. Artificial Intelligent System helps companies to identify threats and opportunities and make defensive strategies. Artificial Intelligence has a significant role in expert system, natural language processing, case-based reasoning, decision making and also dangerous tasks for human-like mining, firefighting, bomb disarming can be done with special Artificial Systems.

P. Kambam *et al.* (2015) studied that Artificial intelligence can be used in robots path planning for their independent movement from starting point to target point without any obstacles [38]. For this purpose, the ACO (Aunt Colony Optimization) algorithm was used that is based on the phenomenon that aunts follow shortest paths moving from one point to another point. Mitigating stagnation approach was used to overcome the limitations of the Ant Colony Optimization Algorithm, which include evaporation, aging and pheromone smoothing approaches. Using ACO algorithm the money can be saved and reliability of the system can be increased.

F. Machinery *et al.* (2015) purposed that Artificial Intelligence used for selection of dismantling stations in a vehicle recycling system [39]. The simulation technique using

genetic algorithm was used for locating the objects in the networks within the limited area. A special expert system was designed for this purpose.

CHAPTER 3

SCOPE OF STUDY

3.1 Scope for Patient:

a) Detection of lung cancer in its prior stage can be beneficial for increase the survival rate of lung cancer patient.
b) By detecting the stage of lung cancer, the doctors can give the theory to patients more efficiently.

3.2 Scope for Doctors:

a) Lung cancer detection using Artificial Neural Networks makes the result more accurate, errorless and measurable.
b) Artificial Neural Network can be a powerful tool to help physicians to perform diagnosis and other enforcements to cure lung cancer disease.

CHAPTER 4

PRESENT WORK

4.1 Problem Formulation

Lung cancer is the second most common disease in India. The main cause of this disease is the unawareness of the people, symptoms are only recognized in the high stages of the lung cancer, and at that stage and it is very difficult to cure this disease. In this study, we propose lung cancer detection technique artificial neural network. Detection of accurate stage of lung cancer has big challenge, because there the number of steps that are performed on lung cancer images in which noise and other unnecessary parts are removes from the lung cancer images and on the bases of feature extraction, we classified the lung cancer stage and find the survival rate of the lung cancer disease. Previous work in this domain is the detection of lung cancer with the help of back propagation neural NETWORK, our focus is on to performed better processes on the CT scan of lung cancer images, increases the accuracy of back propagation neural network, and finds the stage and survival rate of lung cancer patients.

4.2 Objective of the Study

a) Detection of Lung cancer disease using the back-propagation neural network.
b) Find the survival rate of the patient by identifying the stage of the lung cancer.
c) Comparison of the proposed system with already existing systems.

The main objective of my study is to find whether the patient is lung cancer positive or not. If found positive, then the proposed system will find the lung cancer stage from which we can predict the survival rate of the patient.

4.3 Research Methodology

In research methodology, image pre-processing process will apply on CT scan of lung cancer images to remove the redundancy in data, noise present in data. Then, features extraction technique will be applied on these images to extract the parameters from the images and for the lung cancer detection and classification of lung cancer data is made with

Back Propagation Neural Network and stages of lung cancer patients will also be determined along with a prediction of survival rate survival of patient rate.

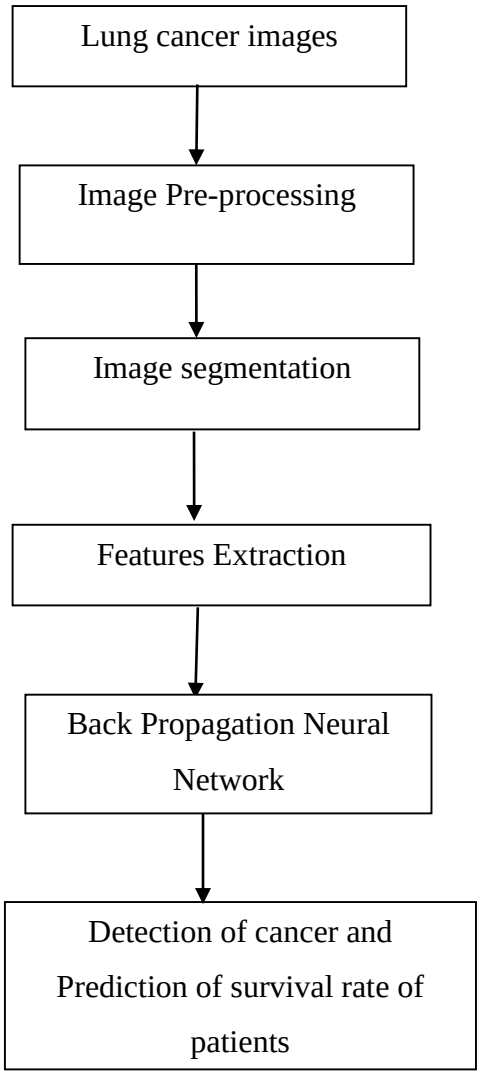


Figure 4.1: Research methodology

4.3.1 Lung Cancer image

The CT scan images have been receive from a private organization where data related to different stages of cancer is collected, after removal the noise from these images, lung cancer images will pass through from a number of processes for detection and prediction the survival rate of the patient.

Figure 4.2: CT Scan of lung cancer patient

4.3.2 Image Pre-processing:

In image preprocessing the processing 2D image converted into gray scale image. In this technique, we have used two techniques image enhancement technique and segmentation. In image processing process noise, redundancy and blueness present in images have been removing. All these steps have been implemented in MATLAB software. There are different preprocessing techniques that we implement are below.

a) Median Filter: Median Filter process has been use to remove the noise from the images [6]. In median filtering, the neighboring pixels are rank according to brightness and the median value becomes the new value ofs the central pixel. The noise reduction uses to improve the result. Median filter works by moving through the image pixel by pixel, replacing each value with the median value of neighboring pixels

```
%% Median Filter
  if(f1==3)
     med_filter = medfilt2(rgb2gray(read_img));
     else
```

```
        med_filter = medfilt2((read_img));
    end
    subplot(4, 3, 2);imshow(med_filter)
    title('Median Filtered Image')
%
%     figure, imhist(med_filter)
```

Figure 4.3: Image filter

b) Image Enhancement: Log-Gabor method has used for image enhancement purpose. Here in log-Gabor method, we extracted the actual lung from the image [9].

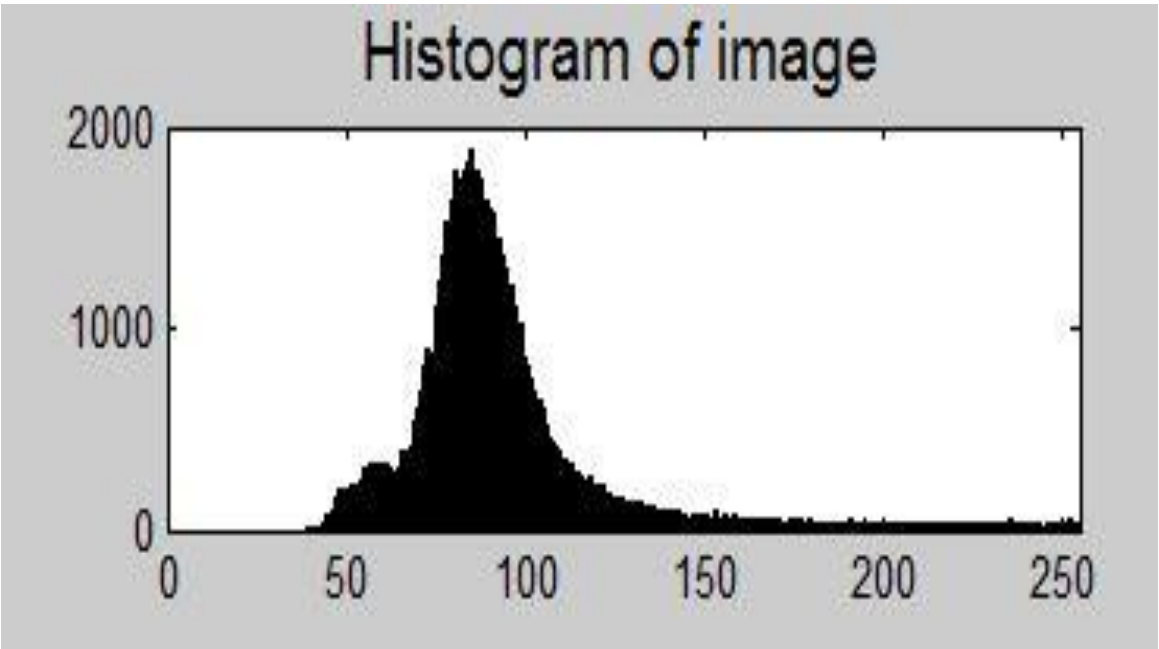
```
%% Log-Gabor
disc('Apply Filtering on Image ...')
inimg_h = med_filter;

size           =25;
nscale         =6;
norient        = 8;
minWaveLength   = 3;
mult           = 1.7;
sigmaOnf       = 0.65;

[gx]=loggabor(size,nscale,norient,minWaveLength,mult,sigmaOnf);
img_out_disp=imfilter(inimg_h,gx,'circular');
% figure;imshow(img_out_disp);title('gabor filterd enhanced image')
```

Figure: 4.4 Log Gabor method

(c) Histogram Equalizers: In Histogram Equalizers we draw a graph of the number of pixels with respect to their values of the pixel [7].



Figure; 4.5 Histogram equalizers

(d) Lung Extracted and Tumor Detection: The binary image has been converting into pixels i.e black and white pixels. After we clear white border that from boundaries. We

provided the filter on images. Area threshold and filling filter. Area threshold we defined the minimum value of the pixel and in fill filter**.** After the filters have applied, the present image mapped with the original image of lung cancer so that now we got the image without any non-required parts. Here we extracted the tumor part from the lungs. We subtracted the lung with no tumor and image segmentation process has applied on that image.

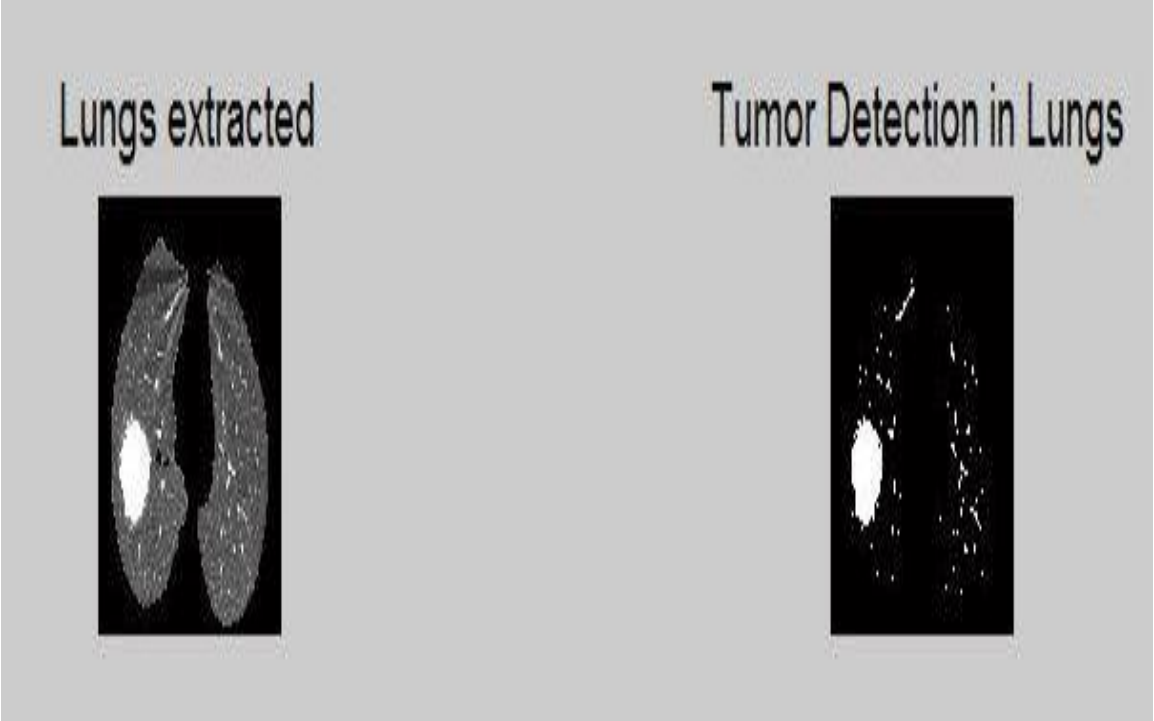


Figure 4.6: Lung extracted and tumor detection

4.3.3 Image Segmentation:

In image segmentation, the process has been performed by using Marker-Controlled Watershed segmentation [28]. The main advantage of using Marker-controlled watershed segmentation is it removes the over segmentation within the image. In image segmentation, we segment the image into multiple parts [11].

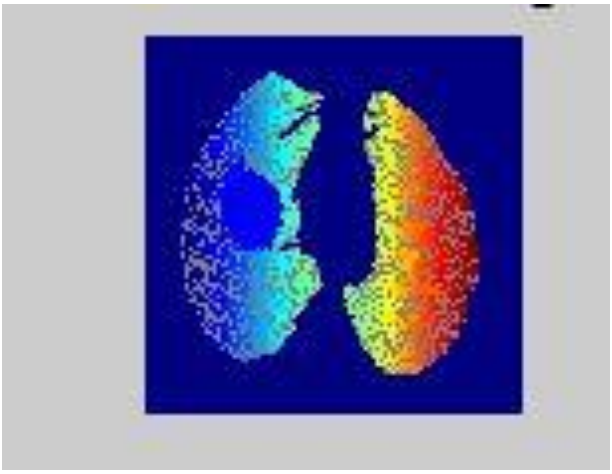


Figure 4.7: Marker-controlled watershed segmentation

4.3.4 Tumor Count:

In the below diagram after the lungs are extracted we detected the tumor part from the lungs. The actual tumor present in the lungs has been detected and GLCM method has applies on the images and various features have been extracted and on the bases of these features the training and testing of the images has been done and classified has made weather the patients has cancer or not.

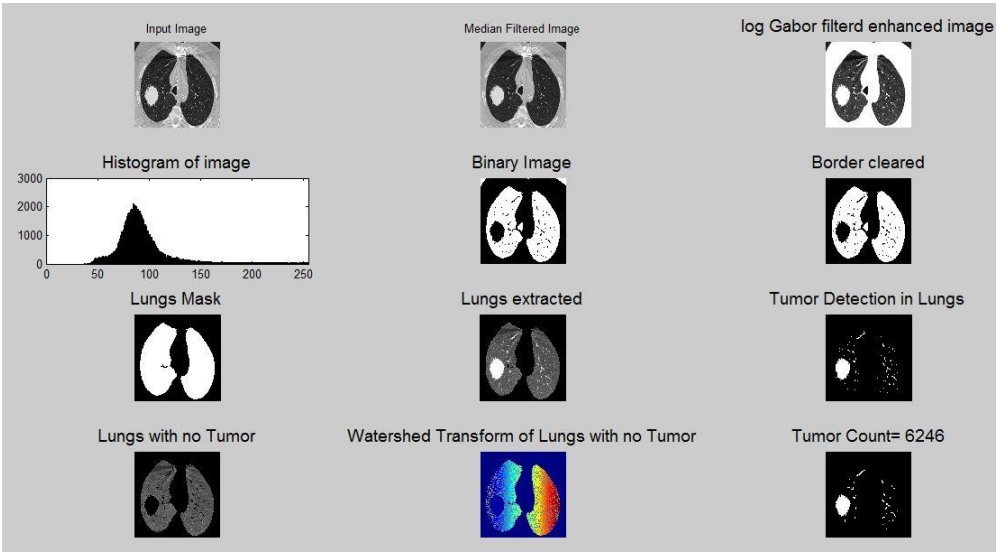


Figure: 4.8 Tumor count

4.3.5 Features Extraction Technique:

Features extraction technique has been performing to detect and isolates various shapes or desired portion. Features have extracted according to the requirements [6]. GLCM method has been use where the features like area, perimeter, and eccentricity have extracted. Feature extraction is the process, which involves for clarifying a number of resources required from a large set of data accurately. Once the features are selection its needs to be extracting for accurate classification of lung cancer disease.

```
Apply Gray Co-MAtrix

glcm = graycomatrix(Tumor_detect,'Offset',[2 0])

%% Extract Detected Tumor Properties

stats = regionprops(Tumor_detect,'all')'
```

```
Area = stats.Area
Centroid  = stats.Centroid
BoundingBox  = stats.BoundingBox
SubarrayIdx = stats.SubarrayIdx
MajorAxisLength = stats.MajorAxisLength
MinorAxisLength = stats.MinorAxisLength
Eccentricity = stats.Eccentricity
Orientation = stats.Orientation
ConvexHull = stats.ConvexHull
ConvexImage = stats.ConvexImage
ConvexArea = stats.ConvexArea
Image = stats.Image
FilledImage = stats.FilledImage
FilledArea = stats.FilledArea
EulerNumber = stats.EulerNumber
Extrema = stats.Extrema
EquivDiameter = stats.EquivDiameter
Solidity = stats.Solidity
Extent = stats.Extent
PixelIdxList = stats.PixelIdxList
PixelList = stats.PixelList
Perimeter = stats.Perimeter
```

4.3.6 Detection of Lung Cancer:

Back propagation neural network was having been used for the classification purpose. We can divide the neural network in two parts training and testing [36]. In the training part, we

provided the training and after that, we test on different images. First, we provide the training with 3 images of each category to our system. When the system executes in the training part all, those features have been extracted and when we testing has been performed different images then according to parameters matching with category, stage of lung cancer has been classified. Features extracted are area, perimeter and eccentricity, tumor count.

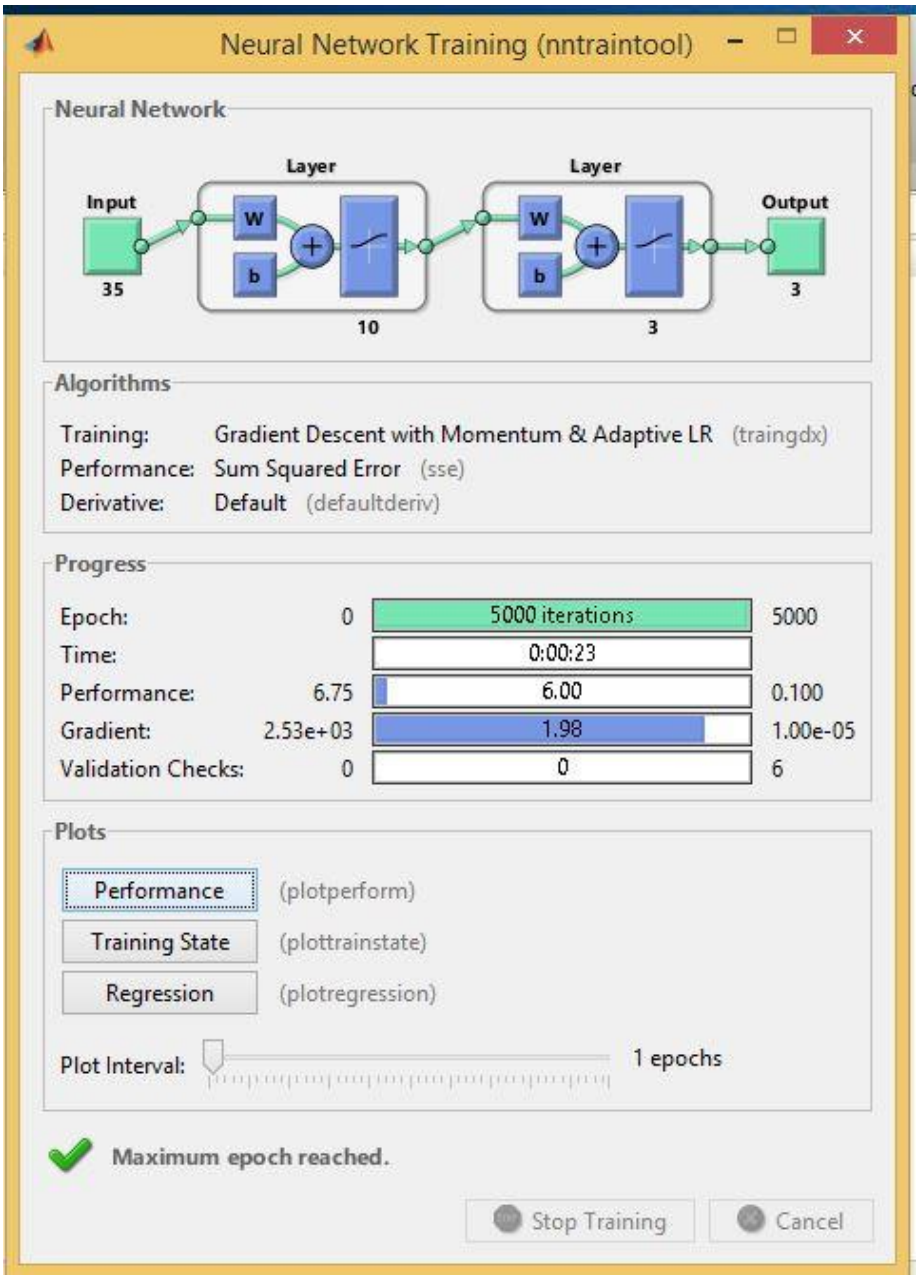


Figure 4.9: Training to system

5.3.7 Prediction of Survival Rate of Lung Cancer Patients:

Based on different stages of lung cancer, we predict the 5-yearly survival rate of patients For no cancer disease predection rate is 80%,stage 1 the predection rate is 50% and for stage is 30%.

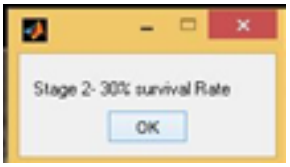


Figure 4.10: Prediction of survival rate

In the below table we discuss the survival rate of patient in terms of five-yearly.

Stages	Survival rate
Stage 0	80%
Stage 1	50%
Stage 2	30%

Table 4.1 Survival rate [40].

CHAPTER 5

RESULT AND DISCUSSION

In this chapter, we conclude our result and discuss different parameters on which result has concluded. We used 15 images of CT scan of lung cancer, 5 images of the lung of each category has used.10 images having the disease of lung cancer, rest of 5 the CT scan images have no disease.

Accuracy is the percentage of correctly identified lung cancer disease. It can be measured as the number of correctly identified images to the total number of images. Our purposed system gave best result with the previous implemented techniques using backpropagation neural network.

Now we create confusion matrix for accuracy.

Patient	Test=Positive	Test=Negative
Cancer	Tp = 9	Fn = 1
No Cancer	Fp = 0	Tn = 5

Correct Test=(Tp+Tn) / (Tp+Tn+Fp+Tn)= 14/ (15

0.9333=93.33%

False test=1/ (15=0.066

=6.66%

Correct Accuracy of our System implemented=93.33%

5.1 Result

Detection of lung cancer by using back propagation neural network gives an accuracy of purposed system is 93.33%, persistence and efficiency, which is better than the previously calculated result with the help of back propagation neural network.

Lung Cancer Detection System	Accuracy
Lung Cancer Detection using Curvelet Transform and Neural Network	90%
Gray Coefficient Mass Estimation Based Image Segmentation Technique For Lung Cancer Detection Using Gabor Filters	83%
Identifying Lung Cancer Using Image Processing Techniques	80%%
Lung Tumor Detection and Classification using EK-Mean Clustering	90.85%
Proposed system	**93.33%**

Table 5.1 Comparasion between Techiniques

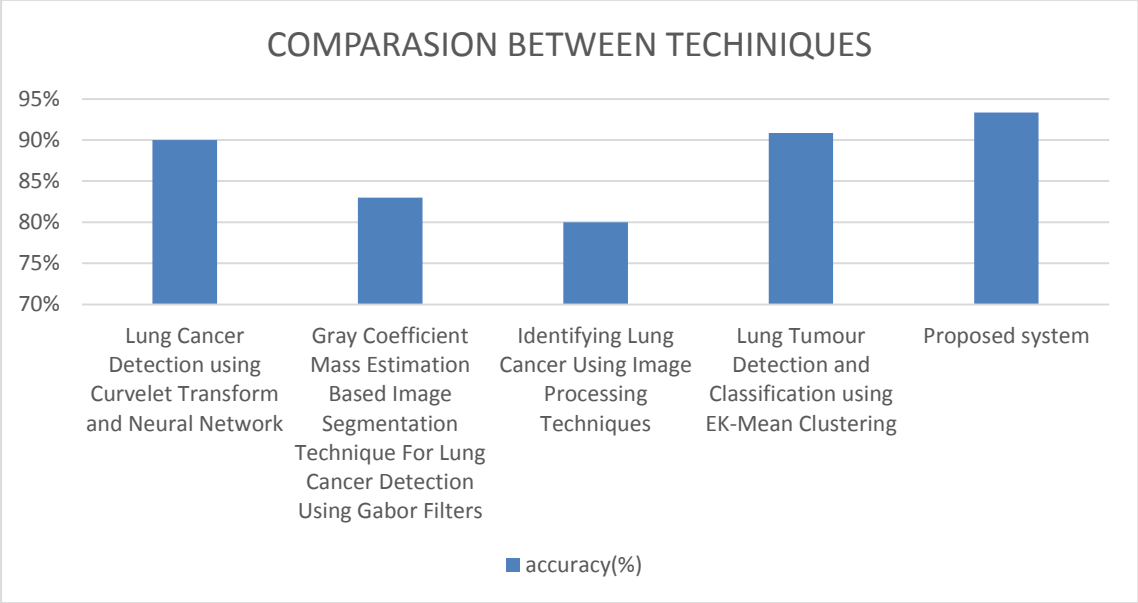


Figure 5.1: Comparison between techniques

From the above discussion, it clear that our purposed implementation gave better result than of other implemented design.

5.2 Classification of Lung Cancer Images:

The classification of the lung cancer is categorized into 3 categories.

No cancer detection= The patient do not have the lung cancer disease.

Stage 1= The patient found the lung cancer disease and stage of lung cancer is stage 1

Stage 2=The patient have lung cancer disease and the stage of this disease is stage 2.

5.3 Prediction of survival Rate of Patient

The Prediction of the survival rate of patient concluded by detection the stage of the patient. The survival rate of lung cancer patients concluded in 5 years of survival rate.

a) No cancer detected= 80% (5-year survival rate)
b) Stage 1= 50% (5-year survival rate)
c) Stage 2= 30% (5-year survival rate

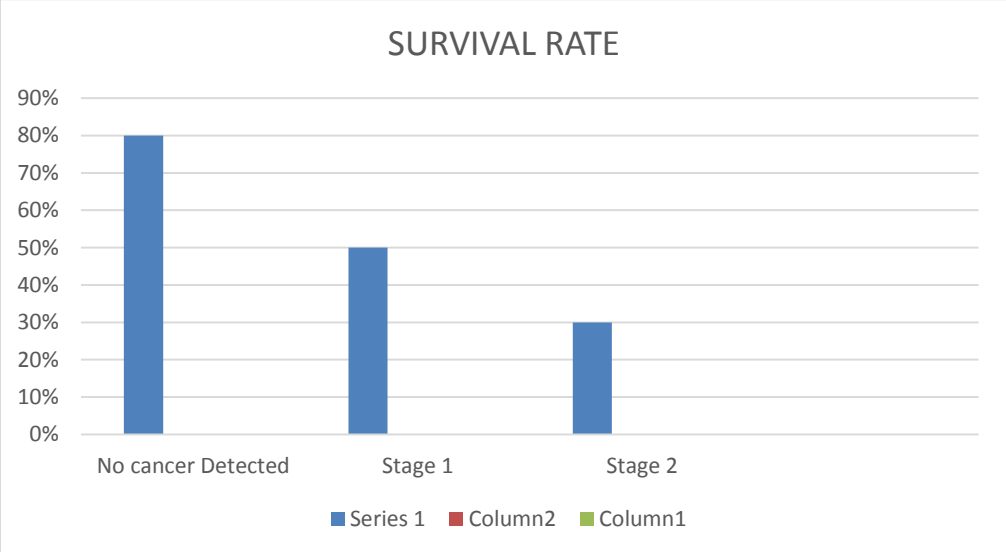


Figure 5.2: survivals rate

From the above Figure it is clear that survival rate of patient is decrease as the increase of the lung cancer stages.

CHAPTER 6

CONCLUSION AND FUTURE SCOPE

Lung cancer disease is the leading cause of the deaths in worldwide. It causes 13% deaths in 2008 and this rate is increasing rapidly every year. Therefore, the identification of this disease becomes very important. The theremajor cause is a lack of awareness about lung cancer. Specially, in India, the death rate of this cancer is very high. In India, it is the second largest disease which causes 0.3 million deaths in every year. Detection of lung cancer in the early stage is the key to the cure of lung cancer disease and it is a very challenging problem because the structure of cells that overlap with each other so, it is not easy to detect this disease. Lung Cancer Detection disease is concluded by using Back Propagation Neural Network. Early detection of this disease can increase the survival rate of the lung cancer patients because it is necessary to detect this disease in its early stage later on this disease becomes cureless Back propagation neural network is a useful tool for classification purpose. Future work Aunt Colony Optimization Algorithm can be implemented with Back propagation neural Network for classification purpose.

REFERENCES

[1] K. Ahmed, T. Jesmin, R. Fatima, Z. Rahman, and F. Ahmed, "Early Detection of Lung Cancer Risk Using Data Mining," vol. 14, pp. 595–598, 2013.

[2] http://m.cancer.org/cancer/lungcancers-smallcell/

[3] http://www.cancer.net/cancer-types/lung-cancer-non-small-cell/stages

[4] http://www.healthline.com/health/lung-cancer/early-signs

[5] http://www.engineersgarage.com/articles/images-processing-tutorial-applications

[6] P. B. Sangamithraa and S. Govindaraju, "Lung Tumour Detection and Classification using EK-Mean clustering," Proc. 2016 IEEE Int. Conf. Wirel. Commun. Signal Process. Networking, WiSPNET 2016, pp. 2201–2206, 2016.

[7] S. S. Bagade, "Use of Histogram Equalization in Image Processing for Image Enhancement," vol. 1, no. 2, pp. 6–10, 2011.

[8] M. Computing, N. Panpaliya, N. Tadas, S. Bobade, R. Aglawe, and A. Gudadhe, "A Survey On Early Detection And Prediction Of Lung Cancer," vol. 4, no. 1, pp. 175–184, 2015.

[9] C. Science and M. Studies, "Lung Cancer Detection from CT Image using Image Processing Techniques," vol. 7782, pp. 249–254, 2015.

[10] M. S. Al-tarawneh, "Lung Cancer Detection Using Image Processin Techniques," no. 20, pp. 147–158, 2012.

[11] S. Logesh, M. Swathy, S. Sathish, J. Sivaraman, and M. Rajasekar, "Identification of Lung Cancer Cell using Watershed Segmentation on CT Images," vol. 9, no. January, pp. 1–4, 2016.

[12] P. Naresh and R. Shettar, "Early Detection of Lung Cancer Using Neural Network Techniques," vol. 4, no. 8, pp. 78–83, 2014.

[13] P. Naresh and R. Shettar, “Early Detection of Lung Cancer Using Neural Network Techniques,” vol. 4, no. 8, pp. 78–83, 2014

[14] R. Dharwal and L. Kaur, “Applications of Artificial Neural Networks : A Review,” vol.no ,pp . December, 2016.

[15] O. Article, “Application of Artificial Neural Network in Predicting the Survival Rate of Gastric Cancer Patients,” vol. 40, no. 2, pp. 80–86, 2011.

[16] F. Feng, Y. Wu, Y. Wu, G. Nie, and R. Ni, “The effect of artificial neural network model combined with six tumor markers in auxiliary diagnosis of lung cancer,” J. Med. Syst., vol. 36, no. 5, pp. 2973–2980, 2012.

[17] I. Saritas, “Prediction of Breast Cancer Using Artificial Neural Networks,” pp. 2901–2907, 2012.

[18] S. Naghibi and M. Teshnehlab, “Breast Cancer Classification Based on Advanced Multi Dimensional Fuzzy Neural Network,” pp. 2713–2720, 2012.

[19] R. Kaur and W. S. Email, “Early Detection and Prediction of Lung Cancer Survival using Neural Network Classifier,” vol. 2, no. 6, pp. 375–383, 2013.

[20] M. T. Cse, “A Study of Detection of Lung Cancer Using Data Mining Classification Techniques,” vol. 3, no. 3, pp. 131–134, 2013.

[21] M. Assefa, I. Faye, A. S. Malik, and M. Shoaib, “Lung nodule detection using multi-resolution analysis,” 2013 ICME Int. Conf. Complex Med. Eng., pp. 457–461, 2013.

[22] C. Engineering, “Lung Cancer Detection at Early Stage Using PET / CT Imaging Technique,” pp. 3358–3363, 2014.

[23] M. AGajdhane, “Detection of Lung Cancer Stages on CT scan Images by Using Various Image Processing Techniques,” IOSR J. Comput. Eng. Ver. III, vol. 16, no. 5, pp. 2278–661, 2014.

[24] S. Kaur, “Comparative Study Review on Lung Cancer Detection Using Neural

Network and Clustering Algorithm," vol. 4, no. 2, pp. 169–174, 2015.

[25] S. Avinash, "An Improved Image Processing Analysis for the Detection of Lung Cancer using Gabor Filters and Watershed Segmentation Technique."

[26] P. Kaur, "Identify Lung Cancer in its Early Stage Using Neural Network and GA Algorithm," vol. 4, no. 2, pp. 341–344, 2015.

[27] M. A. Hussain, T. M. Ansari, P. S. Gawas, and N. N. Chowdhury, "Lung Cancer Detection Using Artificial Neural Network & Fuzzy Clustering," vol. 4, no. 3, pp. 360–363, 2015.

[28] A. Dharmarajan and T. Velmurugan, "Lung Cancer Data Analysis by k-means and Farthest First Clustering Algorithms," vol. 8, no. July, pp. 1–8, 2015.

[29] P. Eskandarian, "Computer-Aided Detection of Pulmonary Nodules based on SVM in Thoracic CT Images," pp. 0–5, 2015.

[30] D. Etection, "An Effective and Efficient Feature s election m ethod for l ung cancer," vol. 7, no. 4, pp. 135–141, 2015.

[31] B. A. Miah, "Detection of Lung Cancer from CT Image Using Image Processing and Neural Network," no. May, pp. 21–23, 2015.

[32] S. Singh, Y. Singh, and R. Vijay, "An Evaluation of Features Extraction from Lung CT Images for the Classification Stage of Malignancy," pp. 78–83.

[33] E. Rendon-gonzalez and V. Ponomaryov, "Automatic Lung Nodule Segmentation and Classification in CT Images B ased on SVM," no. June, pp. 1–4, 2016.

[34] C. Yang, S. O. Prasher, J. Landry, H. S. Ramaswamy, and A. Ditommaso, "Application of artificial neural networks in image recognition and classification of crop and weeds," no. September, pp. 147–152, 2000.

[35] A. Tyagi, A. A. Reddy, J. Singh, and S. R. Chowdhury, "A low cost portable temperature-moisture sensing unit with artificial neural etwork based signals conditioning for smart irrigation applications," vol. 4, no. 1, pp. 94–111, 2011.

[36] P. S. Chaudhari, P. D. M. Patel, and J. L. Juneja, “Artificial Intelligence apply for prediction of Laser Cutting Process – A review,” vol. 2, no. 4, pp. 1025–1028, 2012.

[37] T. G. Krishna, “Expert Systems in Real world Business,” vol. 1, no. 7, pp. 21–27, 2013.

[38] P. R. Kambam, R. Brungi, and P. G. G, “Artificial Intelligence in Robot Path Planning,” vol. 17, no. 3, pp. 115–119, 2015.

[39] F. Machinery and S. Symposium, “System supporting location of service works in agriculture on example of vehicle recycling network,” vol. 7, pp. 87–93, 2015.

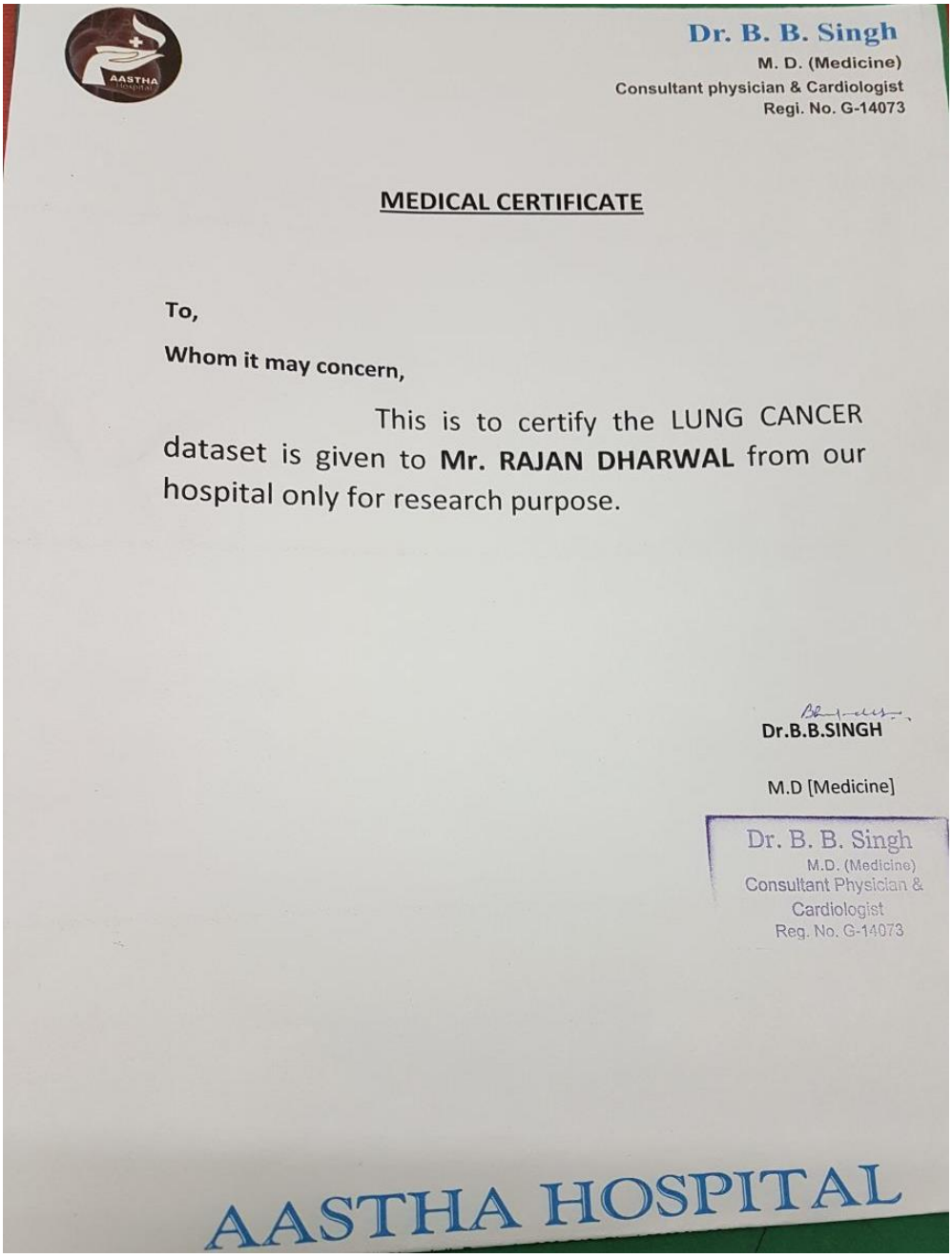
[40] http://www.lung-cancer.com/lung-cancer-survival-rate.html

PUBLICATION

[1] R. Dharwal and L. Kaur, "*Applications of Artificial Neural Networks : A Review*," vol.no 9, December, 2016

.

CERTIFICATE



Dr. B. B. Singh
M. D. (Medicine)
Consultant physician & Cardiologist
Regi. No. G-14073

MEDICAL CERTIFICATE

To,

Whom it may concern,

This is to certify the LUNG CANCER dataset is given to **Mr. RAJAN DHARWAL** from our hospital only for research purpose.

Dr.B.B.SINGH

M.D [Medicine]

Dr. B. B. Singh
M.D. (Medicine)
Consultant Physician & Cardiologist
Reg. No. G-14073

AASTHA HOSPITAL

FIGURE: CERTIFICATE