

LUNG AND PANCREATIC TUMOR IDENTIFICATION USING DEEP LEARNING

Submitted in partial fulfillment of the requirements for the award of
Bachelor of Engineering degree in
Electronics and Communication Engineering

By

KORIVI JAGADEESH (38130112)

DUDE ROHIT (38130058)



**DEPARTMENT OF ELECTRONICS AND COMMUNICATION
ENGINEERING**

SCHOOL OF ELECTRICAL AND ELECTRONICS ENGINEERING

SATHYABAMA

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

Accredited with Grade "A" by NAAC

JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI-600119

MAY-2022



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE

www.sathyabama.ac.in

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

BONAFIDECERTIFICATE

This is to certify that this Project Report is the bonafide work of K.Jagadeesh (38130112), D.Rohit (38130058) who carried out the project entitled **"LUNG AND PANCREATIC TUMOR IDENTIFICATION USING DEEP LEARNING"** under our supervision August,2021 to May,2022.

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_____

Internal Examiner

External Examiner

DECLARATION

I K.Jagadeesh(38130112) and D.Rohit(38130058) hereby declare that the Project Report entitled “LUNG AND PANCREATIC TUMOR IDENTIFICATION USING DEEP LEARNING” done by us under the guidance of Dr.P.CHITRA, M.E., Ph.D., at SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY is submitted in partial fulfillment of the requirements for the award of Bachelor of Technology degree in Electronics and communication Engineering.

DATE:

1) K.Jagadeesh
2) D.Rohit

PLACE:

SIGNATURE OF THE CANDIDATES

ACKNOWLEDGEMENT

We are pleased to acknowledge my 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.**, Prof & Dean of school of Electrical and Electronics Engineering and **Dr .T.RAVI, M.E., Ph.D.**, Heads of the Department, 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 their valuable guidance, suggestions and constant encouragement paved way for the successful completion of our project work.

We wish to express our thanks to all Teaching and Non-teaching staff and also panel members of the Department of Electronics and communication engineering who were helpful in many ways for the completion of the project.

ABSTRACT

In light of recent technological advancements and current trends, we have chosen to do biological research on lung cancer detection. An increasing number of medical specialties are relying increasingly on image processing tools for early diagnosis and treatment. Many different forms of cancer may be found in the body. Examples include lung cancer, breast cancer, blood cancer, esophageal cancer, brain cancer, throat cancer, tongue cancer, and mouth cancer.

Tumors in the lungs form as a result of abnormal cells developing and proliferating. In the blood or lymph liquid that encompasses lung tissue, disease cells might be shipped away from the lungs. MATLAB was used in this study to obtain cancer images from several hospitals where current work is taking place, and this colour picture was one of those images. This study focuses on picture quality and accuracy, which is reliant upon the improvement stage where unassuming pre-handling approaches are applied in light of Gabor separating inside Gaussian guidelines. The feature extraction from a normal and pathological picture is accomplished via the segmentation and enhancement technique. A comparison of normalcy is conducted based on the characteristics of the population. During this study, the most important variables for comparing photos accurately were pixels percentage and mask-labeling.

In the globe, pancreatic malignant growth is the greatest reason for disease related mortality . In order to diagnose pancreatic cancer before it has spread to other parts of the body, a reliable screening tool is needed, however such a tool does not yet exist. Early detection is key to successfully treating pancreatic cancer. To identify pancreatic cancer using CT pictures, this experiment is being undertaken. A CNN Model Architecture and image processing methods are used to identify the tumour. To locate the tumour in an image, CNN Model Architecture is used after the picture has been preprocessed. Tumor detection training accuracy has been found to be approximately 98%.

TABLE OF CONTENTS

CHAPTER NO	TITLE	PAGE NO
	ABSTRACT	vi
	LIST OF FIGURES	x
1.	INTRODUCTION	1
	1.1. INTRODUCTION TO DEEP LEARNING	1
	1.1.1. CONVOLUTIONAL NEURAL NETWORK	2
2.	LITERATURE SURVEY	4
3.	EXISTING SYSTEM	
	3.1. AIM	7
	3.2. SCOPE OF THE PROJECT	7
	3.3. EXISTING SYSTEM	7
	3.3.1. ADVANTAGES OF EXISTING SYSTEM	8
	3.4. PROPOSED SYSTEM	8
	3.4.1. ADVANTAGES OF PROPOSED SYSTEM	9
4.	PROPOSED SYSTEM	
	SYSTEM ARCHITECTURE	10
	4.1. PROPOSED ARCHITECTURE	10
	4.2. METHODOLOGY	11
	4.3. INPUT DESIGN	12
	4.4. OUTPUT DESIGN	12
	4.5. SOFTWARE REQUIREMENTS	12

4.6. IMAGE PROCESSING	14
4.6.1. PURPOSE OF IMAGE PROCESSING	15
4.6.2. TYPES OF IMAGE PROCESSING	16
4.7. PREPROCESSING	17
4.7.1. DIFFERENT TECHNIQUES	17
4.8. HISTOGRAM EQUALIZATION	18
4.9. FEATURE EXTRACTION	19
4.10. SEGMENTATION	20
4.10.1. CLASSIFICATION	21
4.10.2. APPLYING ALGORITHM FOR CLASSIFICATION	21
4.11. MATLAB	23
4.11.1. DATA TYPES IN MATLAB	24
4.11.2. IMAGE TYPES IN MATLAB	24
4.11.3. IMAGE TYPE CONVERSION	25
4.11.4. KEY FEATURES	25
4.11.5. IMAGE PROCESSING WITH MATLAB GUI	25
4.11.5.1. CREATING MATLAB GRAPHICAL USER INTERFACE FOR IMAGE PROCESSING	28
4.11.6. MATLAB GUI CODE FOR IMAGE PROCESSING	31
4.12. PARAMETERS	31
4.12.1. ACCURACY	32
4.12.2. MEAN	32
4.12.3. STANDARD DEVIATION	32
4.12.4. ENTROPY	32
4.12.5. RMS	32
4.12.6. VARIANCE	32
4.12.7. SMOOTHNESS	33
4.12.8. KURTOSIS	33
4.12.9. SKEWNESS	33
4.12.10. IDM	33
4.12.11. CONTRAST	33

4.12.12. CORRELATION	34
4.12.13. ENERGY	34
4.12.14. HOMOGENITY	
5. RESULT AND DATA COLLECTED	35
5.1.DATASET COLLECTED	
6. CONCLUSIONS AND DISCUSSIONS	
6.1. OUTPUT	37
6.2. COMPARISION	42
6.3. CONCLUSION	42
REFERENCES	44

LIST OF FIGURES

Figure No	Title	Page No
4.1	BLOCK DIAGRAM	10
4.2	METHODOLOGY	11
4.3	METHODS USING IMAGE PROCESSING	16
4.4	PREPROCESSING IMAGE	17
4.5	HISTOGRAM EQUALIZATION	18
4.6	FEATURE EXTRACTION IMAGE	19
4.7	SEGMENTATION IMAGE	20
4.8	CLASSIFICATION IMAGE	20
4.9	APPLYING IMAGE FOR CLASSIFICATION	21
5.1	DATASET COLLECTED CT IMAGES	35
5.2	DATASET COLLECTED FOR TUMOR	36
6.1	OUTPUT SCREENSHOTS	41
6.2	COMPARISION	42

CHAPTER- 1

INTRODUCTION

Lung cancer is the fourth most common cause of cancer death in men and women in the United States, yet it is the only major illness with a five-year relative survival rate in the single digits for both men and women. Cancer-related deaths from lung cancer are expected to exceed those from colon cancer by 2020, according to a new report from the Lung Cancer Action Network.

Because there are no reliable procedures for early identification of lung tumours, this often results in the tumour being discovered too late. In addition, the symptoms of a lung tumour are often nonspecific and may be attributed to a wide variety of other illnesses in the abdomen. It is also tough to cure if the disease has spread to other organs. As a result, radiologists urgently require a procedure that would aid them in the early detection of lung tumours. Lung tumour detection has received little attention. According to a literature review, lung tumour identification is done utilising symptoms and patient history, but not image processing.

The goal of this study is to identify lung tumours using CT scan pictures. Imaging and deep learning are utilised to pre-process these pictures before a simple classifier is employed to determine if the tumour region is normal or whether it has a tumour in the lung, based on its location.

1.1. INTRODUCTION TO DEEP LEARNING

Deep learning is a form of machine learning that employs artificial neural networks to learn how to represent the world as a stacked hierarchy of ideas, with each notion defined in relation to simpler concepts and more abstract representations calculated in a nested hierarchy.

When it comes to deep learning algorithms, self-learning representations play an important role. But these self-learning representations are only possible because of artificial neural networks (ANNs). When training, algorithms employ unknown components in the input distribution to extract features, organise objects and uncover relevant data patterns. To develop models, algorithms are used, much as when teaching

robots to self-learn via repetition. One algorithm is used in a number of deep learning models. While no network may be regarded flawless, certain algorithms are better adapted to do particular tasks.. Gaining a thorough grasp of all of the major algorithms is essential for making informed decisions. The following is a list of the most common deep learning algorithms:

- Convolutional Neural Networks (CNN)
- Long Short Term Memory Networks (LSTM)
- Recurrent Neural Networks (RNN)
- Generative Adversarial Networks (GAN)
- Radial Basis Function Networks (RBFN)
- Deep Belief Networks (DBN)
- Multi-layer Perceptron (MLP)
- Self Organizing Maps (SOM)
- Restricted Boltzmann Machines(RBM)

1.1.1. CONVOLUTIONAL NEURAL NETWORK

Profound learning utilizes convolutional brain organizations (CNNs) to decipher visual information, the most pervasive utilization of this sort of brain organization. Presently, when we consider a brain organization, we consider grid increases, however that isn't true with CNN.. It utilizes a technique known as Convolution to do this. A numerical procedure on two capacities that yields a third capacity that makes sense of how one is modified by the other is known as a convolution in math.

Various layers of fake neurons make up convolutional brain organizations. Counterfeit neurons are numerical capacities that process the weighted amount of many data sources and result an initiation esteem, an unrefined impersonation of

their organic partners. Input images into a CNN create numerous activation functions, which are then passed on to the next layer. Each layer generates several activation functions.

Horizontal and diagonal edges are typically extracted from the first layer. Following this output, a more complicated characteristic such as corners or arrangement edges is detected by the following layer. Objects, faces, and other complex traits may be identified as we advance further into the network.

Final convolution layer provides a collection of confidence ratings (numbers between 0 and 1) that indicate how likely a picture is to be classified as belonging to a certain "class." While utilizing a CNN to distinguish growths, the last layer's result is the likelihood that the information picture incorporates any of those tumors.

CHAPTER- 2

LITERATURE SURVEY

According to RAHIB L, et al., cancer incidence and death rates were predicted for the most frequent cancer types in the United States in 2040 based on demographic changes and average yearly percentage increases in incidence and death rates. These cancer diagnoses will continue to rise, but by 2030 thyroid and melanoma will be the fifth and sixth most prevalent diagnoses, respectively. This means that melanoma will overtake uterine cancer as the sixth most common disease. It is predicted that lung cancer will continue to be the leading cause of cancer deaths over this timeframe. For the first time ever, pancreatic and liver cancers will be the leading causes of mortality from all malignancies by 2030. In order to have a significant impact on cancer incidence and mortality rates in the future via advancements in screening, prevention, and treatment, the research and healthcare communities will need to work together today more than they have before.

Pancreatic ductal adenocarcinoma has been described as an aggressive cancer that has a high death rate by Al-HAWARY et al. One of the most critical aspects of patient care is determining the amount of illness on imaging investigations at the time of staging. A standard layout for radiology announcing, involving generally acknowledged and settled on wording for strong pancreatic neoplasm's, is required on account of the wide variety in ability and meaning of sickness degree among various experts and the regular absence of complete revealing of pertinent imaging discoveries at radiologic assessments. It was laid out by a multi-institutional gathering of experts in pancreatic ductal carcinoma, including gastroenterologists and hepato-pancreato-biliary specialists, under the help of the Society of Abdominal Radiologists. Assuming this layout is utilized, it will upgrade the consideration of patients with pancreatic ductal adenocarcinoma by giving an exhaustive revealing of illness organizing to expand helpful proposals that might be introduced to patients. As a further benefit of standardisation, it makes it possible to compare the outcomes of studies conducted at various institutions by utilising suitable and uniform staging based on respectable statuses.

YASAKA K., et al., concluded that radiological imaging diagnostics play an essential role in clinical patient care. Recently, deep learning based on convolutional neural networks (CNNs) has attracted a great deal of interest because of its superior performance in identifying pictures. Patients will benefit from better treatment if CNNs deliver on their promise in the field of radiology, according to current expectations. Here, we address late leap forwards nearby, incorporating research distributed in the flow PLOS Medicine Special Issue on Machine Learning in Health and Biomedicine, with remarks on assumptions and making arrangements for man-made consciousness (AI) in radiology facilities.

According to Anita chaudhary, SonitSukhraj Singh, and others, lung cancer is one of the most deadly and life-threatening diseases in the world. In spite of this, early diagnosis and treatment may save a person's life. Cancer cannot be easily detected by physicians using CT scans, in spite of the fact that they are the greatest imaging tool currently available in the medical profession. As a result, clinicians may use computer-aided diagnostics to more correctly detect malignant cells.

When it comes to reducing high death rates, FatmTaher, et.al., suggest that early detection of lung cancer is critical to success. The worldwide lung screening programme recommends PET and CT scans for the most at-risk elderly populations in order to increase the likelihood of early diagnosis.. As the illness progresses and symptoms don't present until the disease is advanced, radiologists have a tough time locating the lesions. Lung cancer patients with advanced stages of the disease had a five-year survival rate of 17.8 percent and just 4 percent for distant tumours, respectively. The foundation of disease control efforts is based on accurate information.

Lung cancer identification via image processing was shown by Nihad Mesanovic et al. in their paper published in JCI Insights. The use of image processing methods in the early identification phase of many medical disorders is frequently employed to aid in the early treatment of the disease. We used picture segmentation to develop a technique for detecting lung cancer. Image segmentation is a more advanced kind of image processing.

Images from CT scans may be segmented using a combination of watershed management watersheds and region-growing techniques. This is followed by picture enhancement using the Gabor filter, image segmentation, and feature extraction once the detection steps have been completed.

Researchers Nikita Pandey and colleagues have found that: According to the scientists, lung cancer seems to be the most prevalent cause of death for individuals throughout the globe. Patients with lung cancer who are diagnosed at an early stage have a better chance of surviving the disease. Early detection boosts the five-year survival percentage of lung cancer patients from 14 to 49 percent. Despite the fact that CT scans are more efficient than X-rays in certain situations. As a result, the time constraints for identifying lung cancer seemed to have merged into one issue.

CHAPTER -3

EXISTING SYSTEM

3.1. AIM

- Overall categorization accuracy is determined to be 98%, which is a major improvement.
- to lower the percentage of false positive and false negative results, cancer patients will be diagnosed more accurately.

3.2. SCOPE OF THE PROJECT

From this we can easily identified the lung cancer patient with low cost.

- The motivation behind this article is to track down malignant growth cells in the lungs at a prior stage. To do highlight extraction and element choice after division, the picture handling is reached out to recognize cellular breakdown in the lungs. CT (Computerized Tomography) scans of the lungs may be used as input into the system.

3.3. EXISTING SYSTEM

- In order to deal with noisy label categorization in natural photos, a slew of approaches have been presented. A approach proposed by Ren et al. uses a clean validation set to weight training data. To enhance the gradient update, they believe it is best to assign lower weights to noisy samples and to raise the weights of clean training samples. Noisy label classification approaches in medical imaging data have received little attention in the last several years.
- According to a study by Dgani et al., model label noise is included into a deep learning network for the purpose of identifying breast micro-calcifications in multi-view mammograms.

3.3.1. DISADVANTAGES OF EXISTING SYSTEM

- In terms of detecting pancreatic tumours, there has been little research.
- According to a literature review, pancreatic tumour identification is based on symptoms and patient history, but not on image processing.
- So, here we can't predict the disease in the early stage, it will become difficult when it goes to further states,
- Detecting the disease is very slow.
- It did not use any image processing techniques to predict disease at the early stage.

3.4. PROPOSED SYSTEM

The goal of this initiative is to use CT scan pictures to identify pancreatic tumours. These images are preprocessed using image processing techniques and then CNN model architecture is used to classify the tumor area in the image. The dataset is taken from the cancer imaging archive repository. These photos of a pancreatic tumour from a CT scan are sent into the system as data inputs. These images are different format, so we have converted into jpg and also the image size is too huge, so we have used few of them only. Then classification is done using CNN model architecture. In this, the system is made learn according to the classes defined in the image. Once the system learns the classification done based on the features given to it, it can then classify the test data into one of those classes.

A CNN was taught to discriminate between cancerous and healthy pancreases using contrast-enhanced CT scans of patients. In independent test sets, CNN demonstrated high accuracy and enhanced sensitivity compared to radiologist interpretation, with adequate execution in a test set got from an assortment of patients. To help and increase radiologists in the ID and determination of pancreatic malignant growth, our discoveries offer the principal significant proof of idea that CNN might catch troublesome CT qualities.

3.4.1. ADVANTAGES OF PROPOSED SYSTEM:

- Using convolutional neural networks (CNNs) for medical image processing has shown considerable potential.

- Neuronal organizations are constructed utilizing a heap of neurons with actuation capacities and boundaries to extricate and coordinate attributes from the photos and make a model that addresses the muddled connection among pictures and determination. According to some reports, CNN is capable of making accurate imaging diagnoses.
- CNN can consistently detect pancreatic cancer from non-cancerous pancreas, and with upgrades may accommodate variations in patient race and ethnicity and imaging features that are inescapable in real-world clinical practise. Pancreatic cancer CAD approaches are being studied at the CNN Institute to help radiologists in their evaluation of the results of their work.

CHAPTER- 4

PROPOSED SYSTEM

SYSTEM ARCHITECTURE

4.1. PROPOSED ARCHITECTURE:

- If you want to see the system's incoming data, the processing that takes place on it and what it generates, you may use this basic graphical formalism.
- Demonstrating the framework's components is utilized. The framework cycle, the information used simultaneously, an outer element that collaborates with the framework, and the data streams in the framework are largely parts.
- Data streams and is changed by a grouping of changes in this outline. It's a visual portrayal of information stream and the progressions that happen when information moves between different places.

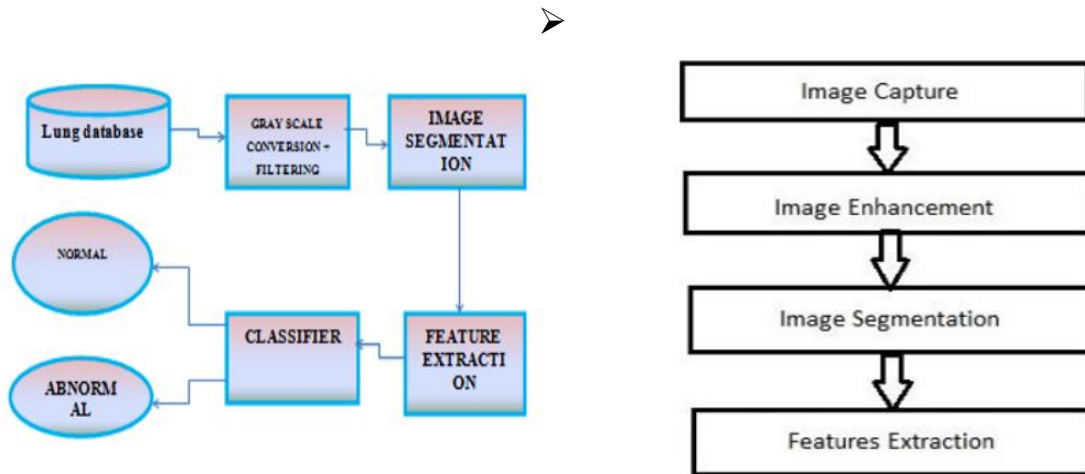


FIG-4.1.BLOCK DIAGRAM

4.2. Methodology:

- STEP-1: Take photos of the lung cancer patients at your local cancer centre.

- STEP 2: Using a command, open a certain picture in MATLAB.
- STEP 3: Process of Improvement In image enhancement, the goal is to make the information contained in a picture easier for humans to understand or better serve as input for other types of automated image processing. Gabor filters are used to improve images in our post-production workflow.
- STEP 4: The Process of Segmentation Slicing up a picture into its basic parts or areas is called segmentation. Medical professionals may utilise it for a variety of things, including seeing and estimating the volume of various objects of interest, detecting anomalies, and classifying tissues. For segmentation, we're use the Marker-Controlled Watershed Segmentation Approach.
- STEP-5: Extracting features This procedure is used to determine whether or not a picture is normal or aberrant. For feature extraction, we are using binarization and masking.

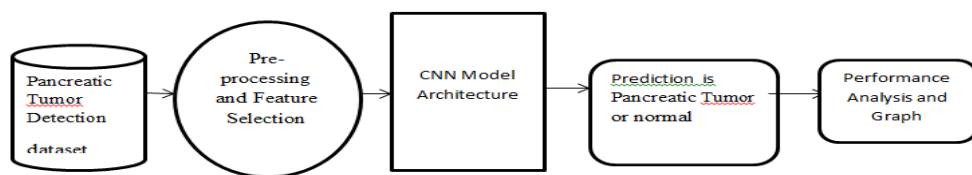


FIG: 4.2: METHODOLOGY

4.3 INPUT DESIGN

User and information system are connected through the input design. In order to handle transaction data, it is required to define specifications and processes for data preparation. These stages may be completed by checking the computer to read data from a written or printed document, or by having users type the data directly into the system. The goal of input design is to maintain the process as easy as possible while also reducing the quantity of input necessary, reducing mistakes, and preventing delays. User security and convenience are maintained while preserving privacy with the input's design.

4.4 OUTPUT DESIGN

An excellent result is one that fulfills the requirements of the end client and actually conveys the data. Outputs are used in any system to convey the outcomes of processing to the user and to other systems. For immediate use and hard copy output, output design determines how the information is to be dispensed. It is the user's primary point of contact with the most relevant and up-to-date information. The system's connection to the user is improved via efficient and intelligent output design, which aids in decision-making.

One or more of the following goals should be met by an information system's output form.

- Deliver data about the present or previous actions, status, or future prospects of the
- Future.
- Inform about significant occasions, chances, difficulties and cautions.
- Trigger an action.
- Confirm an action.

4.5. SOFTWARE REQUIREMENTS:

- Operating system : Windows 10.
- Software used : Matlab 2013R
- Filter used : Gabor filter
- Domain : Image processing.
- Algorithm used : convolutional neural network

4.6. IMAGE PROCESSING

Using image processing, a picture is converted into a digital format and processed in order to produce a better image or extract relevant data. Automated picture transformations are mainly based on algorithms that have been meticulously crafted.

The fields of mathematics, physics, optical and electrical engineering, and software engineering all assume a part in picture handling. It additionally has a wide scope of utilizations, including design acknowledgment and AI, as well as human vision studies. An optical scanner or an advanced camera might be utilized to import the image, investigate and control the picture (information pressure, picture upgrade, and separating), and create the ideal result picture in picture handling.

The development of image processing has been driven by the need to extract information from pictures and to analyse their meaning. Many fields, including as medicine, industry, the armed forces, and consumer electronics, make extensive use of image processing technology of this kind. For diagnostic imaging modalities, such as digital radiography, PET imaging, automated axial tomography (CAT), MRI, and functional magnetic resonance imaging (fMRI), it is employed (fMRI). A few examples of systems used in manufacturing are safety measures, quality assurance, and autonomous vehicle guidance.

From the identification of troops or vehicles to the guidance of missiles and the identifying and reconnaissance of objects, complex picture it are used to handle techniques. In policing security, biometric techniques like fingerprinting, face, iris, and hand acknowledgment are generally utilized. Superior quality TVs, DVD players and individual video recorders are a portion of the most popular consumer electronics products that use image processing.

Picture handling and computerized picture handling, picture catch, or handling to separates the vital data from it. Picture handling is typically done consequently and in view of a significant level calculation. Visual and field handling with commitments from different trains like arithmetic, physical science, optics, and electrical designing.

Furthermore, it coordinates with different fields, for example, plan mindfulness, AI, craftsmanship, and human examination. Different picture handling steps incorporate bringing pictures into optical optics or computerized cameras, picture examination and picture the executives (data pressure, picture improvement, sifting), and picture handling. Separating data from images and figuring out what's in them has become a significant part of the evolution of image processing. Imaging is used in a variety of industries such as pharmaceuticals, industrial products, military applications, and electrical devices.

In medication, it is utilized as X-beams, positron radiation, processed tomography, imaging, and attractive reverberation imaging. Modern applications incorporate creation frameworks like wellbeing frameworks, quality control, and self-driving vehicles.

A very much evolved calculation is utilized to look for warriors or vehicles, rocket control, data recognizable proof, and search. Biometric methods like fingerprinting, looks, eyebrows, and finger impression acknowledgment are broadly utilized in law requirement and security.

The most common electronic video-handling gear includes advanced cameras, cameras, high-quality televisions, DVD players, video recorders, and mobile phones.

There are mainly three processes to image processing.

Digital photography and optical scanning may be used to import the picture.

Spotting patterns that aren't visible to the human eye, as those in satellite photos, may be accomplished by picture analysis and manipulation that involves data reduction and image augmentation.

The last step before an image or report based on image analysis may be adjusted is called output.

4.6.1. Purpose of Image processing

There are five main purposes for image processing. In order, they are as follows:

- Use your imagination to see what isn't seen.

- To improve the quality of a picture, you may use image sharpening and restoration techniques.
- Seek for the picture you're looking for in this step.
- Measurement of pattern - Measures different elements in a picture.
- Distinguish between the various elements of a picture.

4.6.2. Types of images processing

Images may be processed using either Analog or Digital Image Processing techniques. Images on hard copies, such as prints and pictures, may be processed visually or through analogue methods of image processing. When interpreting images with the use of these visual methodologies, picture examiners utilize an assortment of reasonable establishments. The picture handling isn't simply restricted to the locale that must be broke down, yet additionally to the investigator's aptitude. Image processing using visual approaches also makes use of association as a valuable tool for processing images. So, analysts use a mix of their own expertise and data from other sources to do picture processing.

In order to manipulate digital photos using computers, digital processing methods are used. There are inherent flaws in the raw data from satellite imaging sensors, As a result, the material must go through a series of processing steps to ensure that it is original. Preprocessing, Enhancement and Display, and Information Extraction are the three fundamental cycles that a wide range of information should encounter while utilizing computerized strategies.

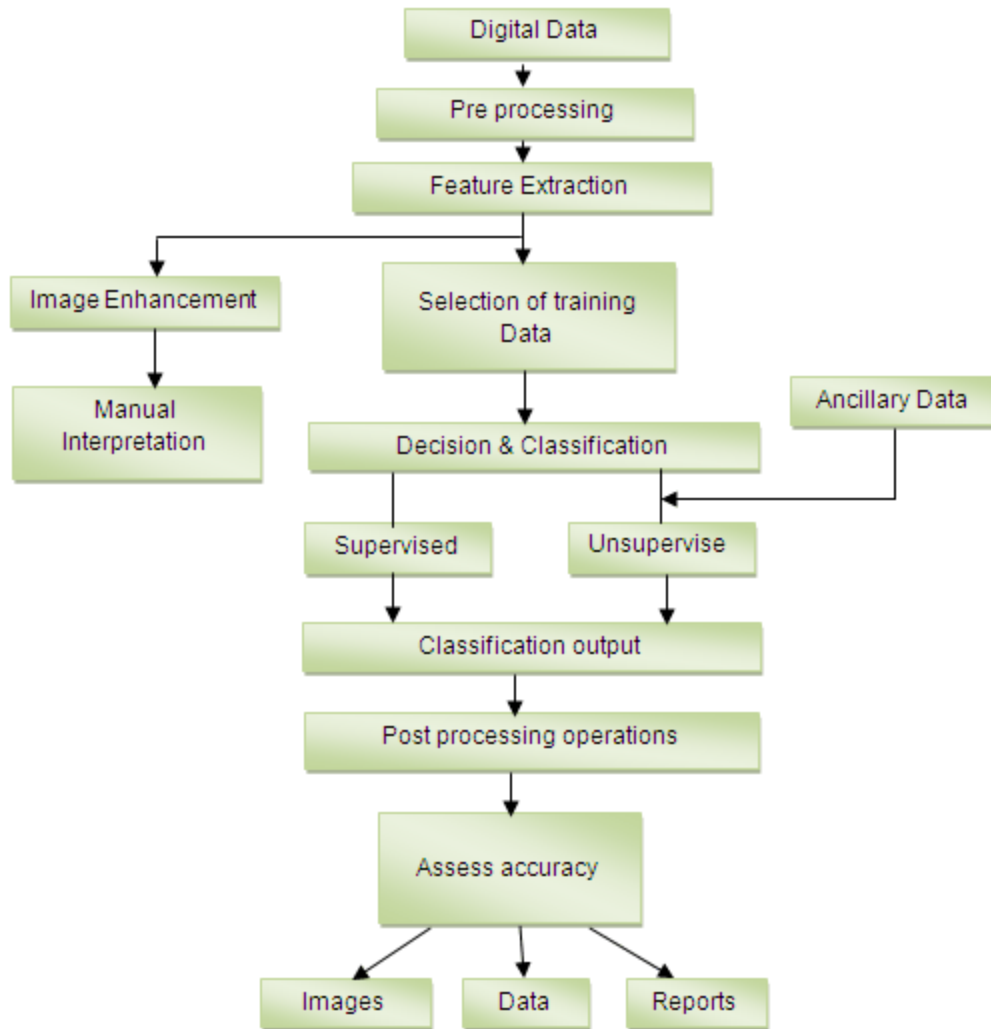


FIG: 4.3: METHODS USING IMAGE PROCESSING

4.7. Preprocessing:

Image preprocessing is a frequent term used to describe processes on photographs when the input and output are intensity images. It is the goal of preprocessing to enhance the picture data in order to reduce or eliminate any undesirable distortion.

Point handling strategies incorporate difference extending, worldwide edge, histogram adjustment, log changes, and power regulation changes. Different veil handling strategies incorporate averaging channels, honing channels, neighborhood thresholding, etc.

4.7.1. Different techniques:

Information preprocessing is an information mining method that includes changing crude information into a justifiable configuration. ... Information preprocessing is a demonstrated technique for settling such issues. Information preprocessing plans crude information for additional handling or improves some picture highlights significant for additional handling.

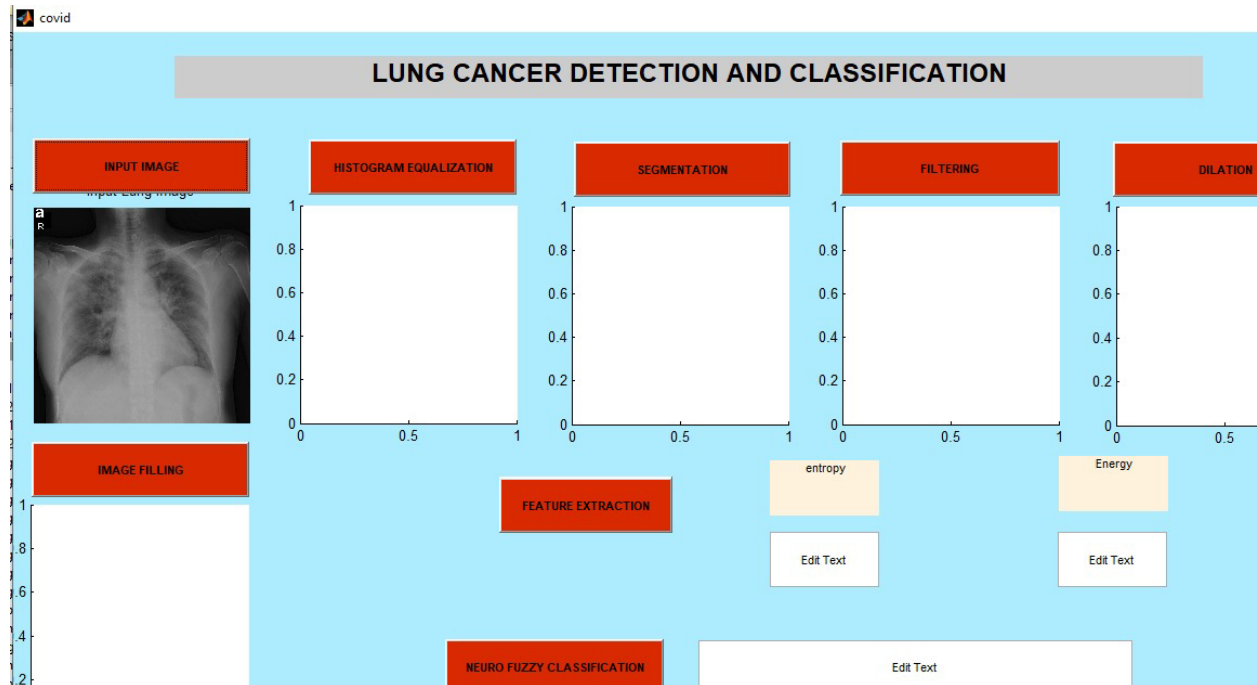


FIG.4.4: PREPROCESSING IMAGE

4.8. HISTOGRAM EQUALIZATION

A grayscale picture's histogram shows how often different shades of grey appear in the image. The graph shows how many times each grey level appears in the picture. Gray levels (and hence the histogram) are clogged towards the bottom end of a dark picture. Gray levels (and hence the histogram) are congested towards the top end of a consistently bright picture. Gray levels (and hence the histogram) would be evenly distributed throughout a wide range in a well-contrasted picture.

Fig: 4.5. HISTOGRAM EQUALIZATION

4.9. FEATURE EXTRACTION:

To begin with, all raw data sets are separated into manageable subsets and decreased in dimensionality by feature extraction. Even though these qualities are simple to implement, they accurately and uniquely characterise the underlying data collection.

An object (also known as segment) in Feature Extraction is a set of pixels with comparable spectral, spatial, and/or texture qualities that is used to identify pictures. As it stands, the most common approach to image classification relies on the usage of color-coded pixels to categorise a picture.

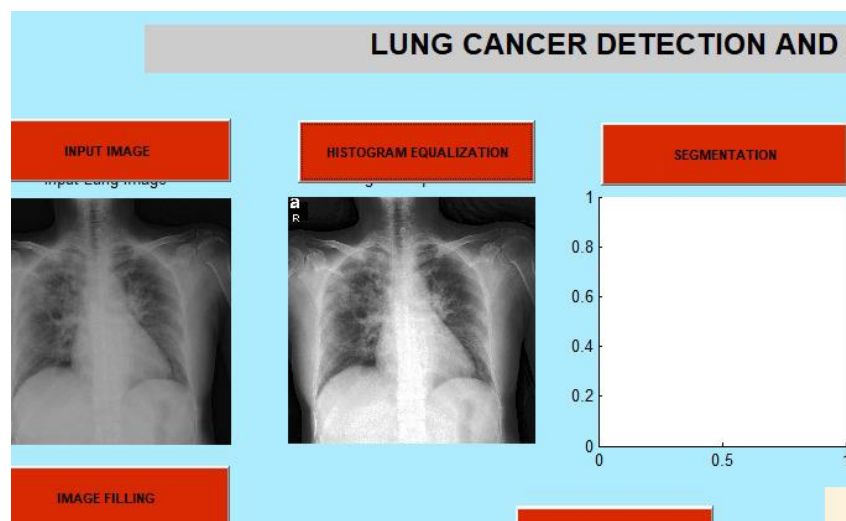


FIG: 4.6: FEATURE EXTRACTION IMAGE

4.10. SEGMENTATION

Image Segmentation is the process of dividing up a digital image into smaller groupings of pixels called Image Objects, which makes it easier to evaluate the picture.

The pixels in a picture are separated and organised using a variety of image segmentation methods. When we give labels to pixels, we're really creating a category for those pixels with the same name that have a similar characteristic.

We may establish borders, draw lines, and distinguish the most significant parts of a picture from the less important parts by using these labels. Here, we've tried to isolate the important elements, such as a chair and a table, from a larger picture on the left, and as a result, all the chairs have the same shade of grey. Since each chair has a distinct colour, we've found them in the next tab and may use them to identify them.

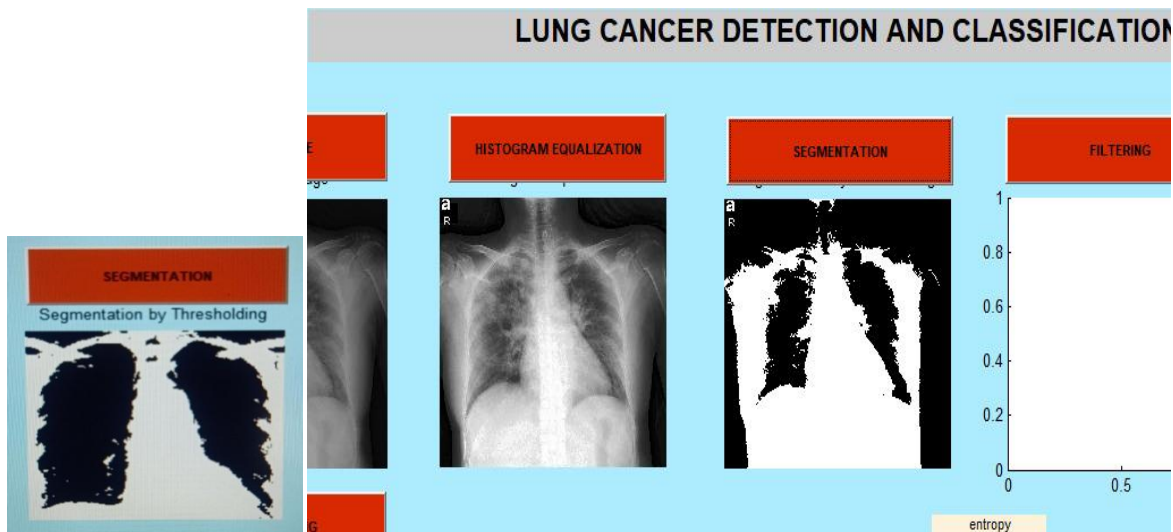


FIG: 4.7: SEGMENTATION IMAGE

4.10.1. CLASSIFICATION

In the classification technique, it is used to compare the filtered or segmented image and dataset. Then it will give entropy and energy level. For classification we are using CNN algorithm since it gives high accuracy.

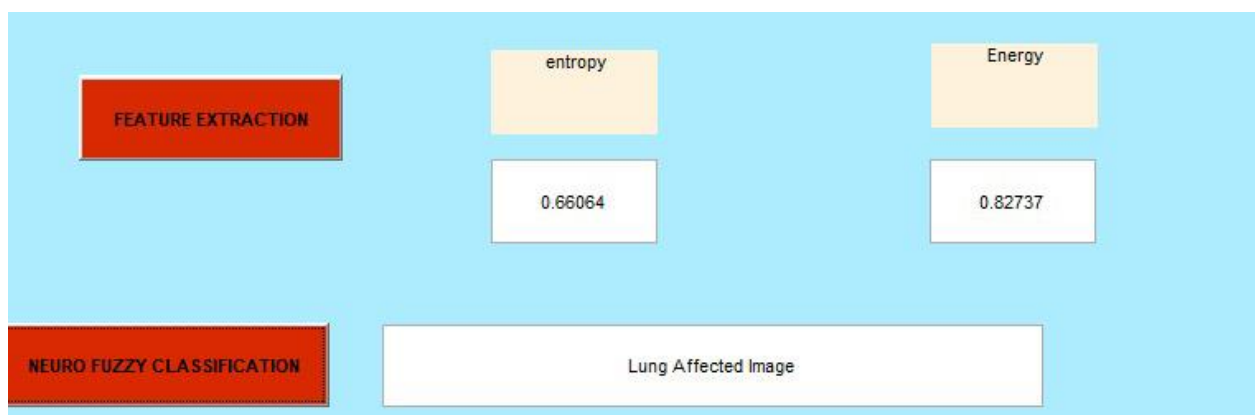


FIG: 4.8: CLASSIFICATION IMAGE

4.10.2. Applying algorithm for classification:

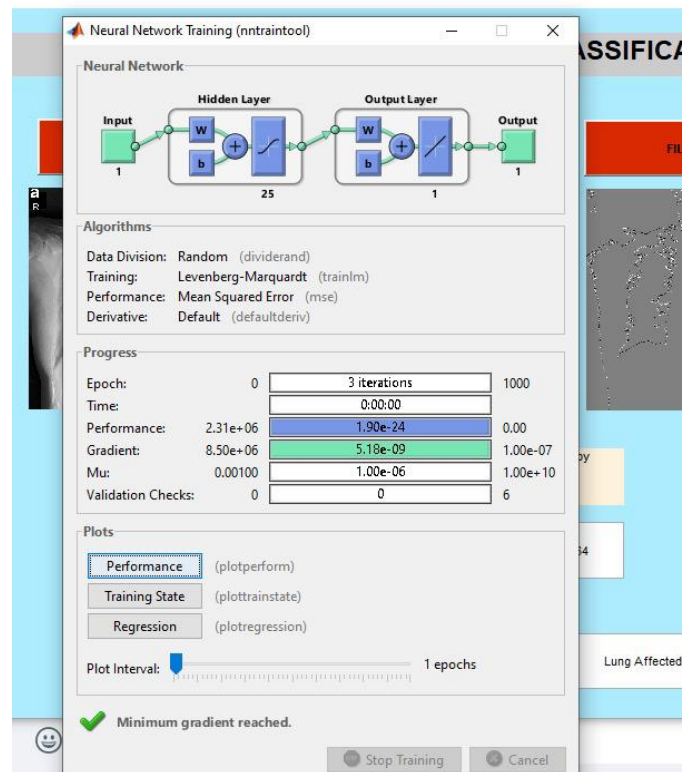


FIG: 4.9: APPLYING ALGORITHM FOR CLASSIFICATION

4.11. MATLAB

MATLAB® is a significant level specialized registering language and intelligent climate for calculation creation, information perception, information examination and mathematical estimation. Conventional programming dialects like C, C++, and FORTRAN are more slow than MATLAB in tackling specialized PC issues. Matlab is the fastest way to do it.

This data analysis and visualisation software is built on Matlab's robust support for matrices and matrix operations. In addition, Matlab offers a robust programming language and outstanding graphical capabilities of its own. As a result, Matlab has become one of the most widely used tools in industry. For our purposes, we'll focus on . the picture handling tool kit, which contains an assortment of related instruments. Rather than depicting Matlab's elements in general, we'll zero in on just those parts connected with

handling pictures. We will add capacities, directions, and approaches depending on the situation. A Matlab function is a term that takes one or more arguments and outputs one or more results, such as a matrix, text, or graph. As an example, consider the function `sin`. Many functions are available in Matlab, and as we'll see, writing your own is both simple and occasionally required.

As a general rule, all data in Matlab are matrices, and the usual data type is matrix. Images, of course, are matrices whose members are the grey values (or maybe RGB values) of its pixels. A single value is considered a matrix by Matlab, but the length of a string is just a matrix of letters. A general overview of Matlab operations will be presented in this chapter, followed by a discussion of pictures in subsequent sections.

A command window called the Command Window_ appears when Matlab is launched. An order line-style point of interaction is more proficient than a progression of pull-down menus in view of Matlab's many capacities and the numerous contentions they might acknowledge.

Sign and picture handling, interchanges, control plan, test and measurement and financial modelling may all benefit from MATLAB's extensive range of capabilities.. A set of MATLAB functions called a toolbox may be used to handle certain types of issues in various application areas by extending the MATLAB environment with additional functions.

Documenting and sharing your work is easy using MATLAB's built-in tools. Matlab may be used to create other languages and programmes, or it can be used to create and share your own programmes.

For example, loading a picture, using the correct format, storing the data as various data types and how to display and convert images are just a few of the things to bear in mind while dealing with images in Matlab.

A broad assortment of reference-standard calculations and graphical devices for picture handling, investigation, perception, and technique creation might be found in Image Processing Toolbox (ITB). Highlight acknowledgment, sound decrease, picture division, spatial changes, and picture enrollment are only a portion of the picture handling capacities you might utilize. To take utilization of multicore and multiprocessor

frameworks, a few capacities in the tool stash are multithreaded.

- MATLAB and images
- MATLAB treats a picture as a matrix
- Matrix elements are found in every pixel.
- MATLAB's operations are all defined in terms of
- If you are working with photos, you may utilise matrices such as "+", "-", "*", "/", and others like sqrt, sin, and cos to manipulate the results.
- Importing and exporting many picture formats is supported by the MATLAB software
- The BMP (Microsoft Windows Bitmap)
- Animated Graphics File (Graphics Interchange Files)
- The HDF (Hierarchical Data Format)
- JPEG is a popular image format (Joint Photographic Experts Group)
- In addition to PCX, there are a number of other (Paintbrush)
- Papua-New-Guinea (Portable Network Graphics)
- JPEG, GIF, and PNG (Tagged Image File Format)
- XWD (X Window Dump)
- It is possible to input raw data or other picture data into MATLAB

4.11.1. Data types in MATLAB

- Increase (64-bit double-precision floating point)
- Unmarried (32-bit single-precision floating point)
- Int32; (32-bit signed integer)
- Int16 (16-bit signed integer)
- Int8 (8-bit signed integer)
- The 32-bit Uint32 type (32-bit unsigned integer)
- At least 16-bit integers are supported (16-bit unsigned integer)
- 8-bit unsigned integer (8-bit unsigned integer)
- In MATLAB, you can create images.
- A binary image is one that has a value of zero or one.
- Intensity images: [0,1] or uint8, double, and so forth. • Uint8:

- m-by-n-by-3 RGB pictures
- m-by-3 colour map index pictures
- Images that are three-dimensional The m-by-n-by-p (p is the number of layers)

4.11.2. IMAGE TYPES IN MATLAB

• Outside of Matlab, pictures may be black-and-white, gray-scale, or coloured in any combination. Matlab, on the other hand, has four different sorts of pictures. It is known as a binary picture because it has a value of 1 for white and 0 for black. The integers 0 to 255 or 0 to 1 used to represent grayscale pictures are referred to as intensity images. It is possible to represent coloured pictures in RGB or Indexed Image format.

These three photos may be found by searching for "RGB Images." These three images are arranged in a way that each includes a distinct section of a picture. As a result, the matrix for a 640 by 480 picture is 640 x 480 x 3. Listed Image is an elective method for showing hued pictures. The picture lattice and the guide grid are the main two frameworks that exist. There is an interesting list number allocated to each tone in the image, and this file number shows up in the picture framework. In the guide network, each tone has its own exceptional file number.

4.11.3. IMAGE TYPE CONVERSION

- RGB Image to Intensity Image (rgb2gray)
- RGB Image to Indexed Image (rgb2ind)
- RGB Image to Binary Image (im2bw)
- Indexed Image to RGB Image (ind2rgb)
- Indexed Image to Intensity Image (ind2gray)
- Indexed Image to Binary Image (im2bw)
- Intensity Image to Indexed Image (gray2ind)
- Intensity Image to Binary Image (im2bw)
- Intensity Image to RGB Image (gray2ind, ind2rgb)

4.11.4. Key Features

- A significant level programming language for logical and designing applications.
- Advancement climate for dealing with code, records, and information. —
- Iterative investigation, plan and issue goal utilizing intelligent instruments
- Mathematical mix and sifting are among the numerous numerical capacities
- available for use in linear algebra and statistics.
- Data visualisation using 2-D and 3-dimensional graphics
- Graphical user interface (GUI) design tools
- Integrating MATLAB-based algorithms with external programmes is made easier by these functions and languages, such as C, C++, FORTRAN, Java, COM, and Microsoft Excel.

In this four-part series, we cover both the basics and more advanced aspects of image processing in MATLAB. MATLAB image processing toolbox (IPT) functions and their effects on various pictures are covered in these articles, which range from basic to sophisticated. An introduction to digital pictures and MATLAB is given in the first part of this series before MATLAB activities such as image reading and display are shown.

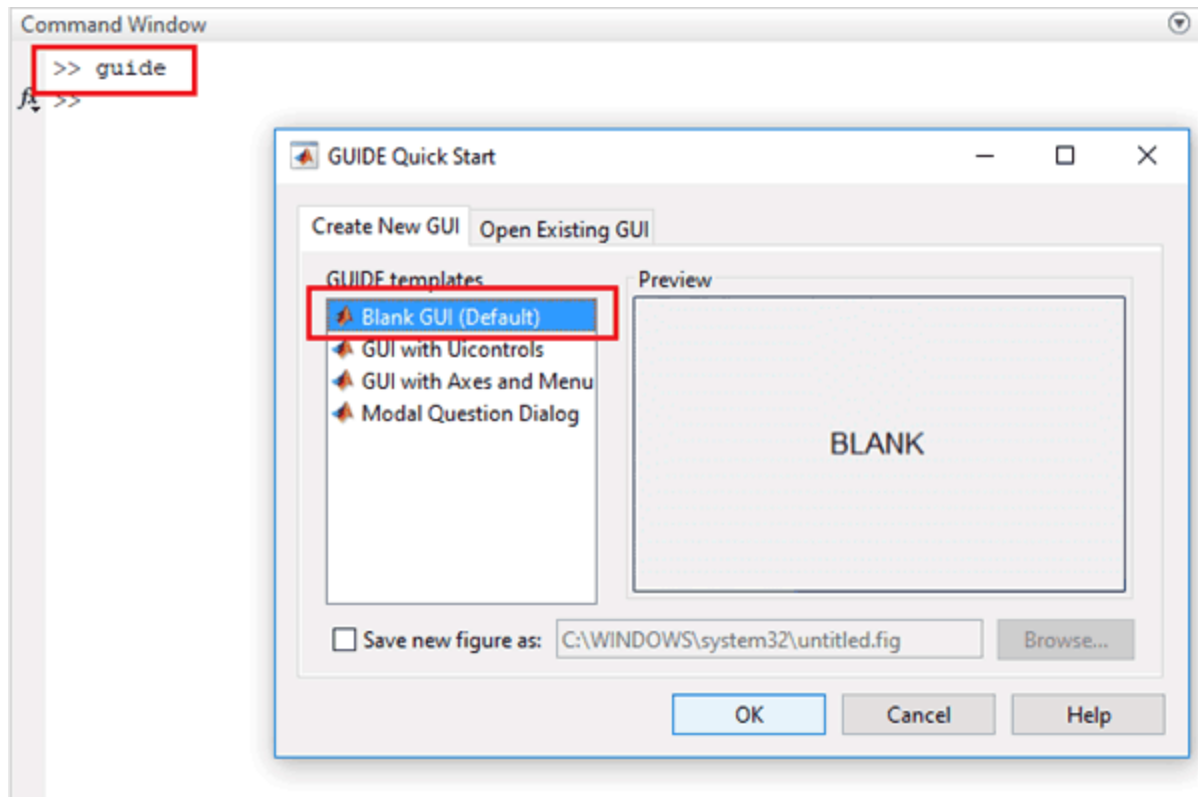
4.11.5. Image Processing with MATLAB GUI

4.11.5.1. Creating MATLAB Graphical User Interface for Image Processing

Launch the Image Processing GUI by entering the following command in the command window to begin developing a GUI (Graphical User Interface).

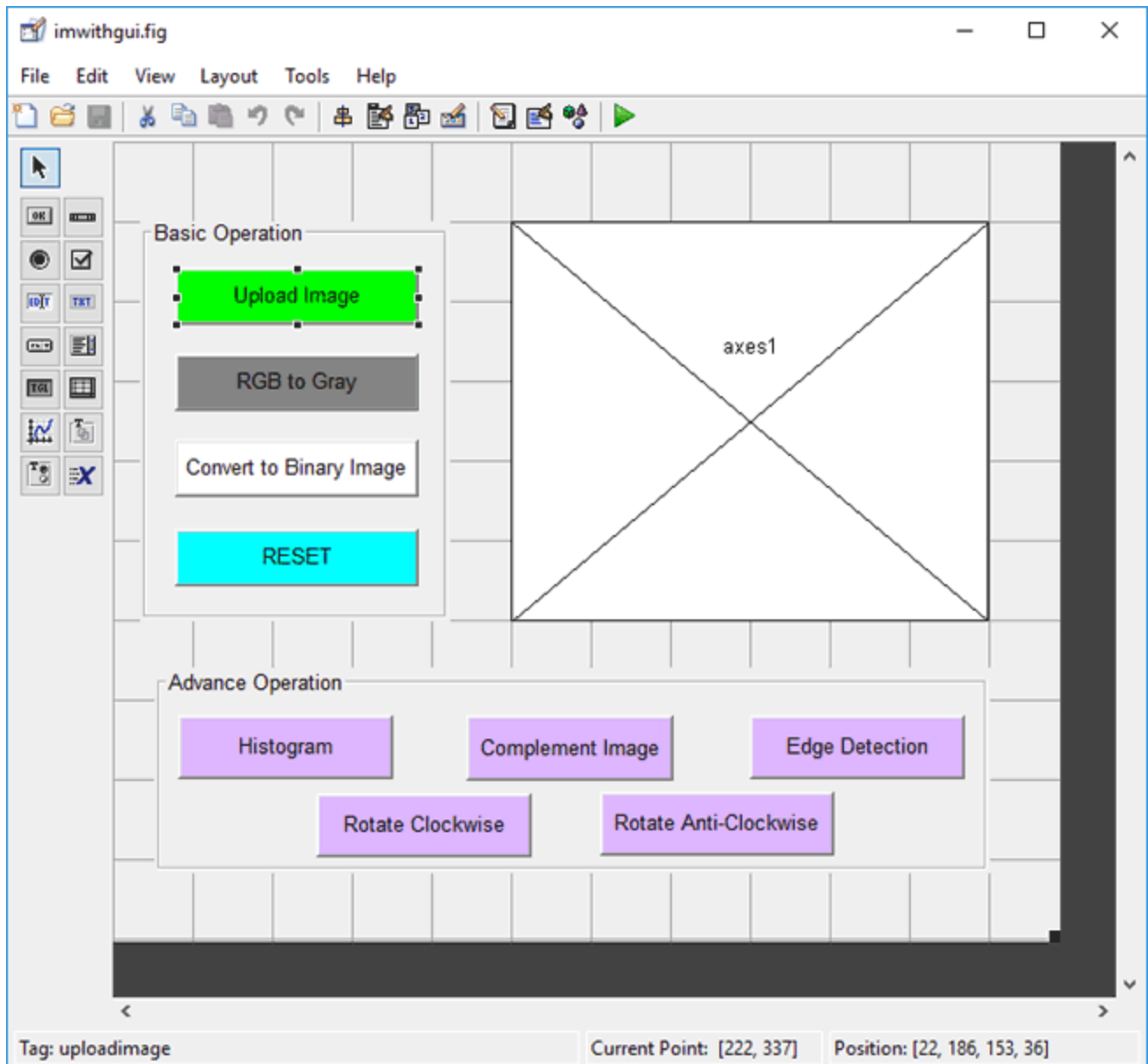
Guide

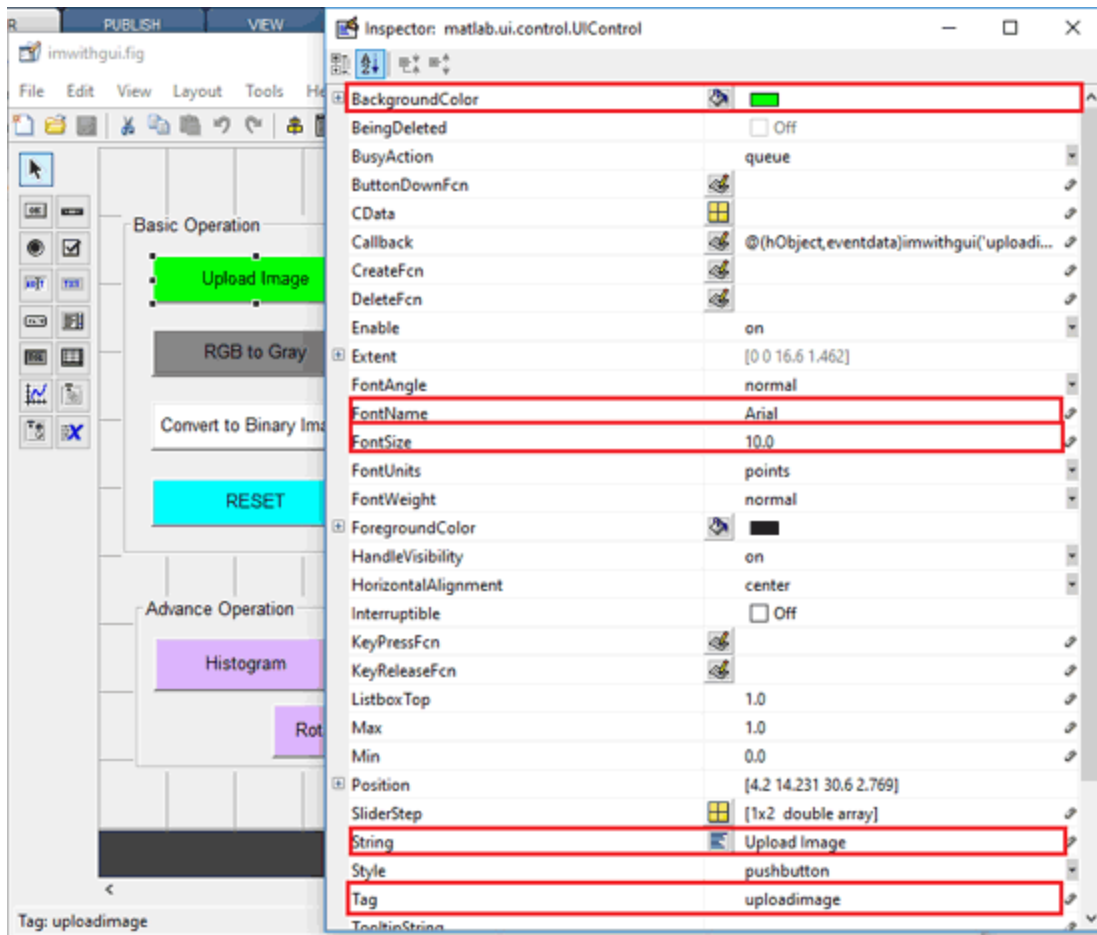
Click on the New Blank GUI button to bring up a pop-up menu, as illustrated in the figure below.



After selecting the number of pushbuttons and the axis on which to show the picture, we are ready to proceed.

When you click on a Pushbutton or Axes, you'll be able to move the edges to resize or shape it. You may alter the colour, string, tag, and other properties of any of these buttons by double-clicking on them. This is how it will appear after being customised.





You have complete control over the look and feel of the buttons. When you save this, MATLAB generates a piece of code in the Editor window. Using the resulting code, you may assign various pushbuttons to do different functions. The MATLAB code has been changed in the next section.

4.11.6. MATLAB GUI Code for Image Processing

This project concludes with the distribution of the MATLAB code for Image Processing using MATLAB GUI. You can also get the GUI file (.fig) and code file (.m) here so you may change the size of the buttons or Axes. The produced code has been modified in the manner described below.

Copy and paste the following code into the 'uploadimage' function to upload a file from your PC. In order to import a picture into the MATLAB GUI, we'll use the command

getfile('image extension type') here. You may display the file using command show() on axes1 using axes1 after reading it using command read() (handles.axes1). Now that you've got the command set app data(), you can keep the variable in the GUI and have it available to both parts of the GUI at the same time.

```
a=uigetfile('.jpg')
```

```
a=imread(a);
```

```
axes(handles.axes1);
```

```
imshow(a);
```

```
setappdata(0,'a',a)
```

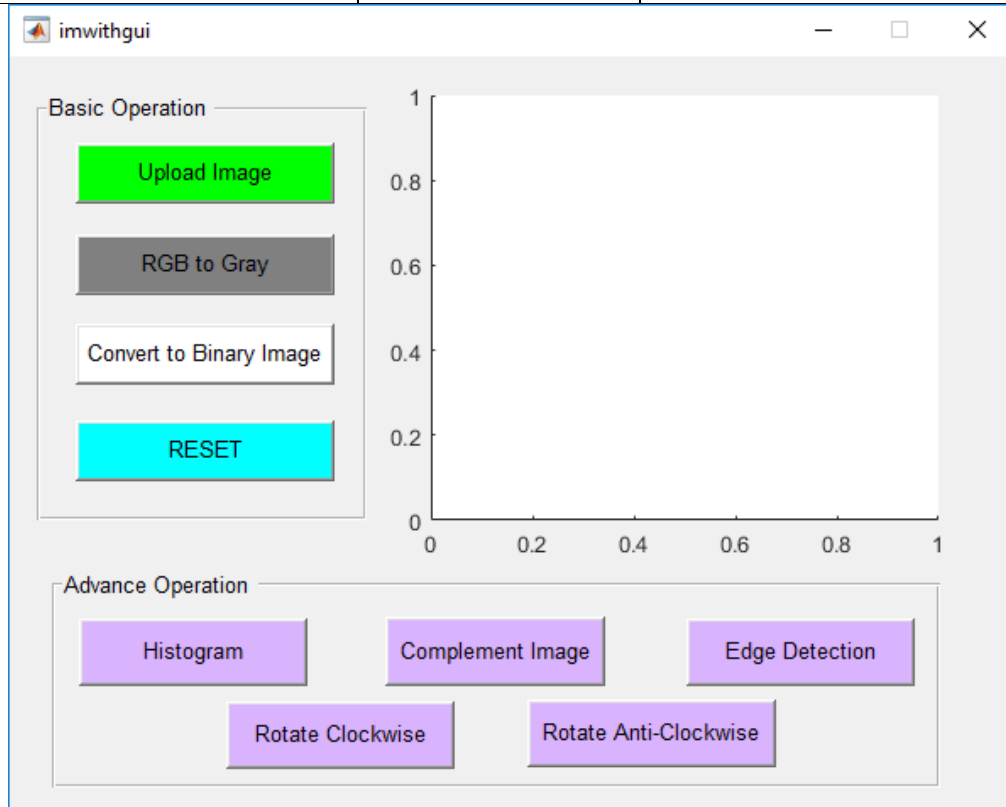
```
132 % --- Executes on button press in uploadimage.
133 function uploadimage_Callback(hObject, eventdata, handles)
134 % hObject      handle to uploadimage (see GCBO)
135 % eventdata    reserved - to be defined in a future version of MATLAB
136 % handles      structure with handles and user data (see GUIDATA)
137 - a=uigetfile('.jpg')
138 - a=imread(a);
139 - axes(handles.axes1);
140 - imshow(a);
141 - setappdata(0,'a',a)
```

Using the set app data() method in the GUI, you may now get data that was previously saved by using the new command get app data() in every function.

This section explains eight of the most popular image processing functions.

S. No.	Command	Button Name	Task to be Performed
1.	uigetfile()	Upload Image	Click to import image from Disk
2.	rgb2gray()	RGB to Gray	Click to convert RGB image into grayscale
3.	im2bw()	Convert to Binary Image	Click to convert the image into binary
4.	-	RESET	Click to reset the image as original
5.	imhist()	Histogram	Click to see the histogram of the image
6.	imcomplement()	Complement Image	Click to check the complement image
7.	edge(filename, method)	Edge Detection	Click to detect the edges in the image
8.	Imrotate(filename, angle)	Rotate Clockwise	Click to rotate the image into the clockwise direction

9.	imrotate(filename, angle)	Rotate Anti- Clockwise	Click to rotate the image into the anti-clockwise direction
----	------------------------------	--------------------------------	--



4.12. PARAMETERS:

4.12.1. ACCURACY

Comparison of convolutional neural networks with other algorithms like KNN and SVM to gain superior accuracy of 98.5 percent, which is higher than the other methods.

- Aims of Module 1, Course 4, are as follows.
- To get a better grasp on the convolution process.
- To have a better grasp on the pooling process

- Keeping track of the terminology used in convolutional neural networks is essential (padding, stride, filter, etc.)
- Using a convolutional neural network to classify photos into many categories

4.12.2. Mean:

There are two ways to calculate mean value: sum of pixel values divided by the total number of pixel values. All the pixels that make up a picture on a computer have pixel values that indicate how bright and/or what colour that pixel should be.

4.12.3. Standard deviation:

Gray level intensity distribution may be measured using the standard deviation (σ), which can be interpreted as a measure of the alternating signal component's amplitude that was captured by the camera.'

Standard deviation of a picture is an indication of how much variance there is in the goal value of light intensity at each individual data point in the image.

4.12.4. Entropy:

When an image's entropy is calculated, it's a way to quantify how random it seems in comparison to other images. In the context of picture coding, the entropy is helpful because it sets a lower limit on the average bit length per pixel that may be achieved by an optimal coding scheme without any loss of information.

4.12.5. RMS:

Throughout the injection time, the spray forms were compared using the RMS technique of producing an RMS average picture. Spray patterns and high-contrast pictures were determined by using this procedure.

4.12.6. Variance:

An indicator of variability is the variance. The average of squared deviations from the mean is used to compute it. The variance of your data sets informs you how wide the range of values is. The standard deviation (the difference between the mean and the median) increases with the dispersion of the data.

4.12.7. Smoothness:

If you want to remove noise or make a picture seem more smooth, you may use this technique. However, you may alternatively smooth a picture using an average or median value of a set of moving pixels (a kernel) rather than relying on low-pass filters.

4.12.8. Kurtosis:

According to the normal distribution, Kurtosis measures whether the data are peaky or flat. When the kurtosis value is high, data tend to peak around the mean, drop fast, and have large tails. In contrast to data with a high peak, those with low kurtosis have a flat top towards the mean.

4.12.9. Kurtosis:

Symmetry, or rather the absence thereof, is quantified by skewness. Strictly speaking, the left and right halves of any distribution or collection of data must be identical in appearance. In statistical terms, kurtosis is a measure of whether the data are heavy- or light-tailed.

4.12.10. IDM:

Images' local homogeneity is assessed using IDM. Co-occurrence of pixel pairs is increased when the gray-scale values of the two pixels are similar, which raises the IDM value.

4.12.11. Contrast:

An analogue image's contrast is measured by how much colour or grayscale distinction exists between the image's individual features. Generally speaking, images with a higher contrast level exhibit more variety in colour or grayscale.

4.12.12. Correlation:

In order to do correlation, a filter mask known as a "kernel" is moved across a picture, and the total of the products at each place is calculated. Correlation is the result of the filter's displacement.

Mathematical procedures convolution and correlation have a lot in common. Both signals must be convolutions of each other in order for correlation to work. As a result, one of the signals gets rotated by 180 degrees during the correlation process. There are fundamental differences between convolution and correlation.

4.12.13. Energy:

The image's localised alteration is quantified by its energy. The term "energy" is used in a wide variety of circumstances and by many people, although it generally refers to the same substance. Color/brightness/magnitude shifts of pixels in a given region at a constant pace.

Based on the sum of the x and y derivatives at that particular place, each pixel has its own unique amount of energy. In order to calculate the derivatives, a finite difference between the pixel and its neighbours is taken. The pixel's total energy is calculated by adding the energies from each of the colour channels.

4.12.14. Homogeneity:

In general, a picture is homogenous if each pixel (dot) in the image has the same hue. An picture is inhomogeneous if it has high contrasts. The standard deviation of each pixel's grey value from the mean may be computed statistically in the simplest example.

CHAPTER -5

RESULT AND DATA COLLECTED

In this chapter, we'll provide a variety of outcomes that we've produced as a consequence of our implementation.

5.1. DATASET COLLECTED

A sample file of CT images which are in normal which is used for training and testing the data.

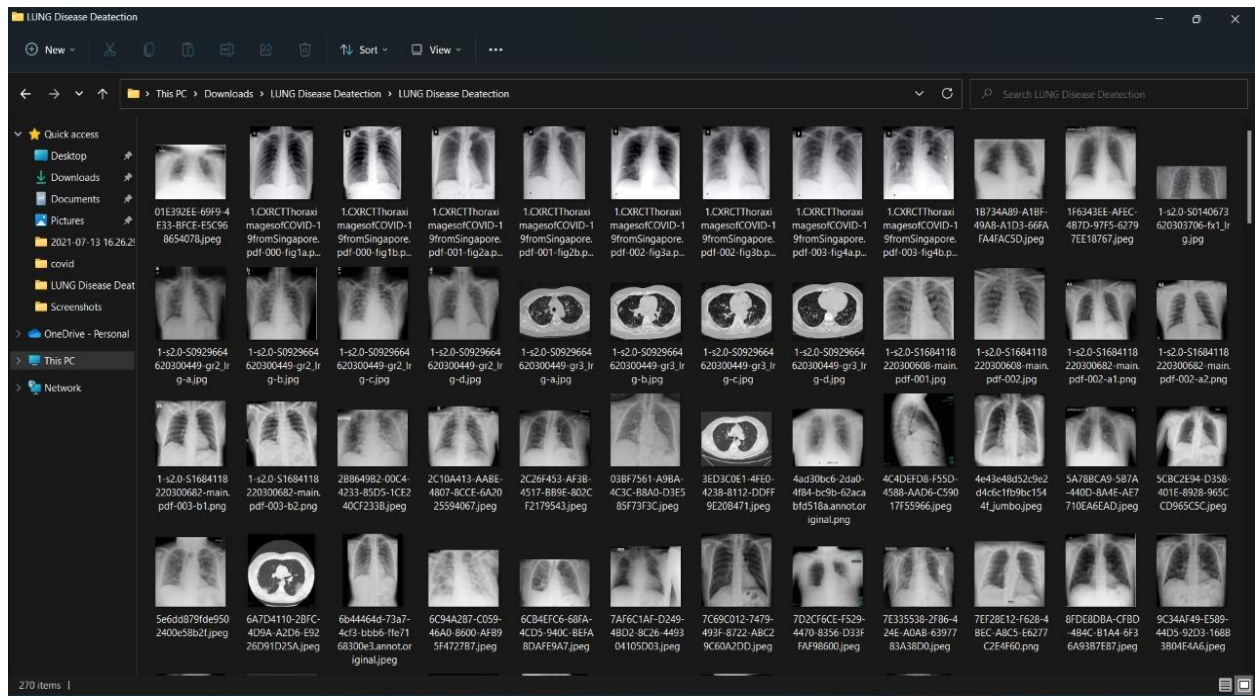


FIG-5.1.DATASET COLLECTED CT IMAGES

A sample file of CT scan images having tumors in the pancreas which is used for training and testing the data.

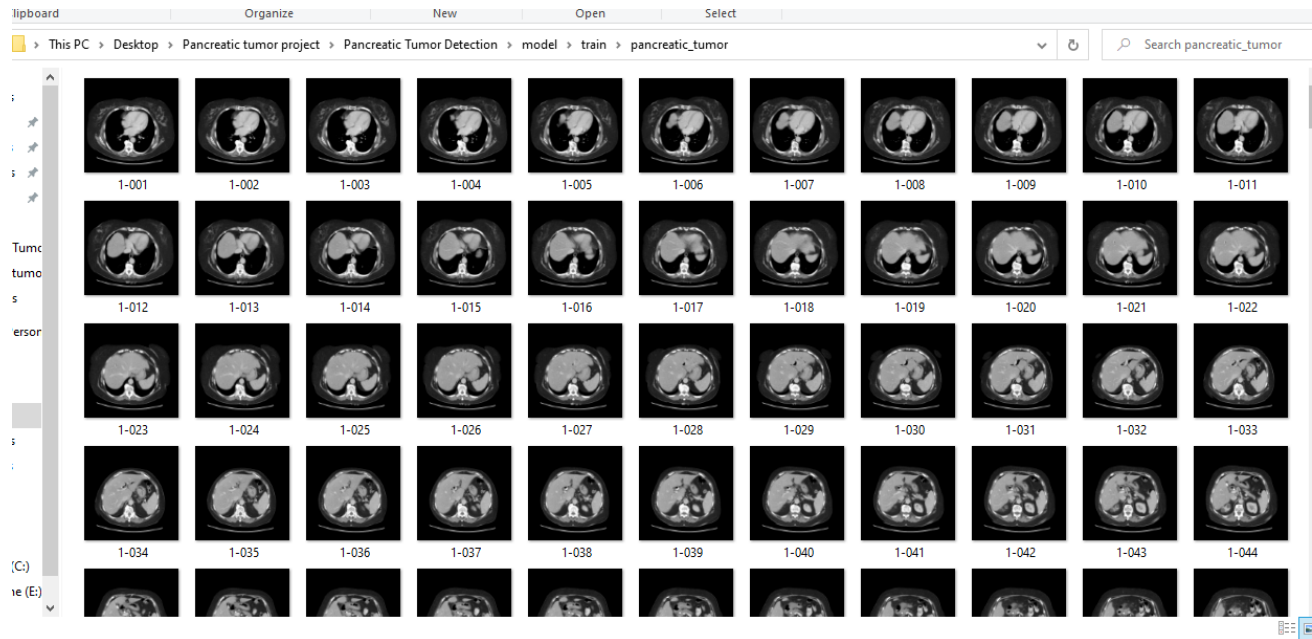


FIG-5.2. DATASET COLLECTED FOR TUMORS

CHAPTER- 6

CONCLUSIONS AND DISCUSSIONS

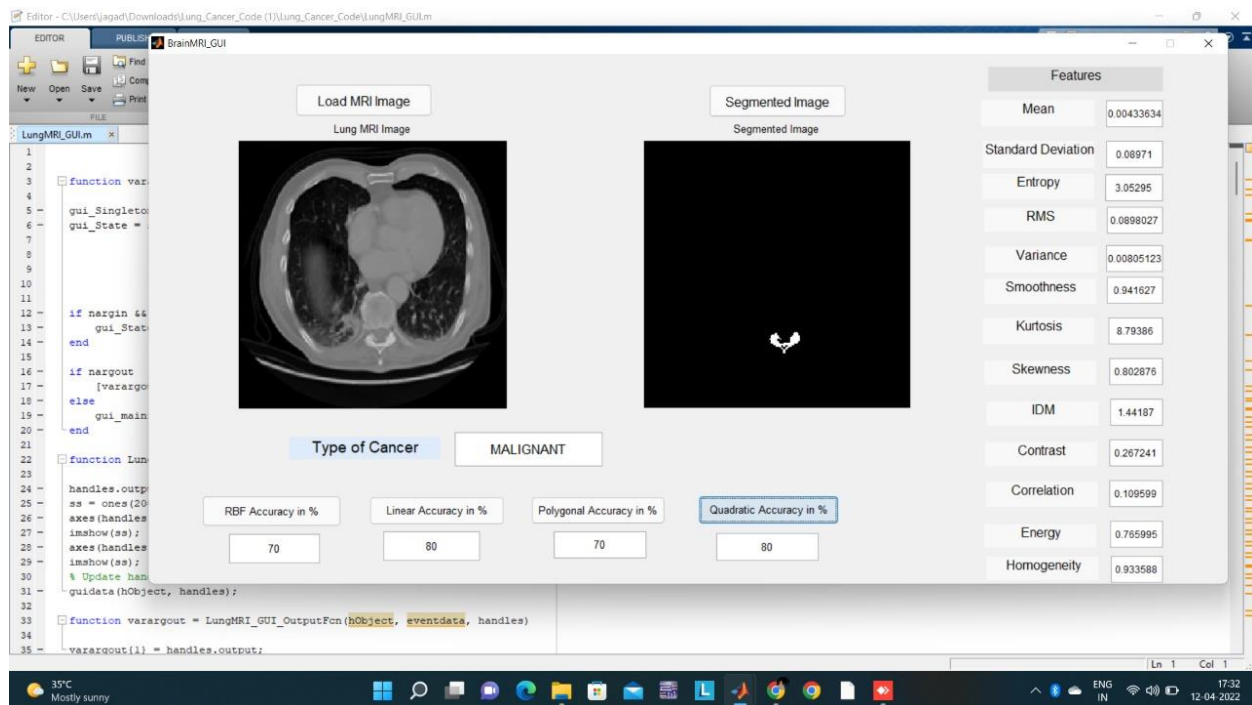
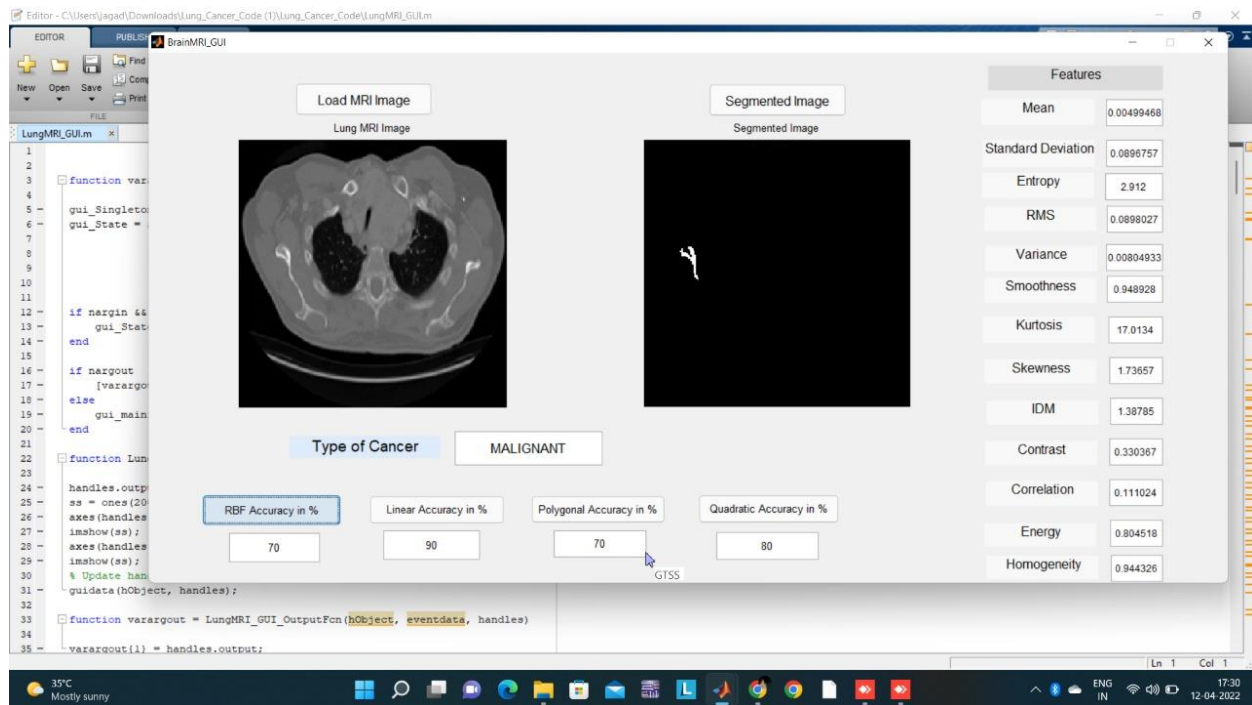
6.1. OUTPUT:

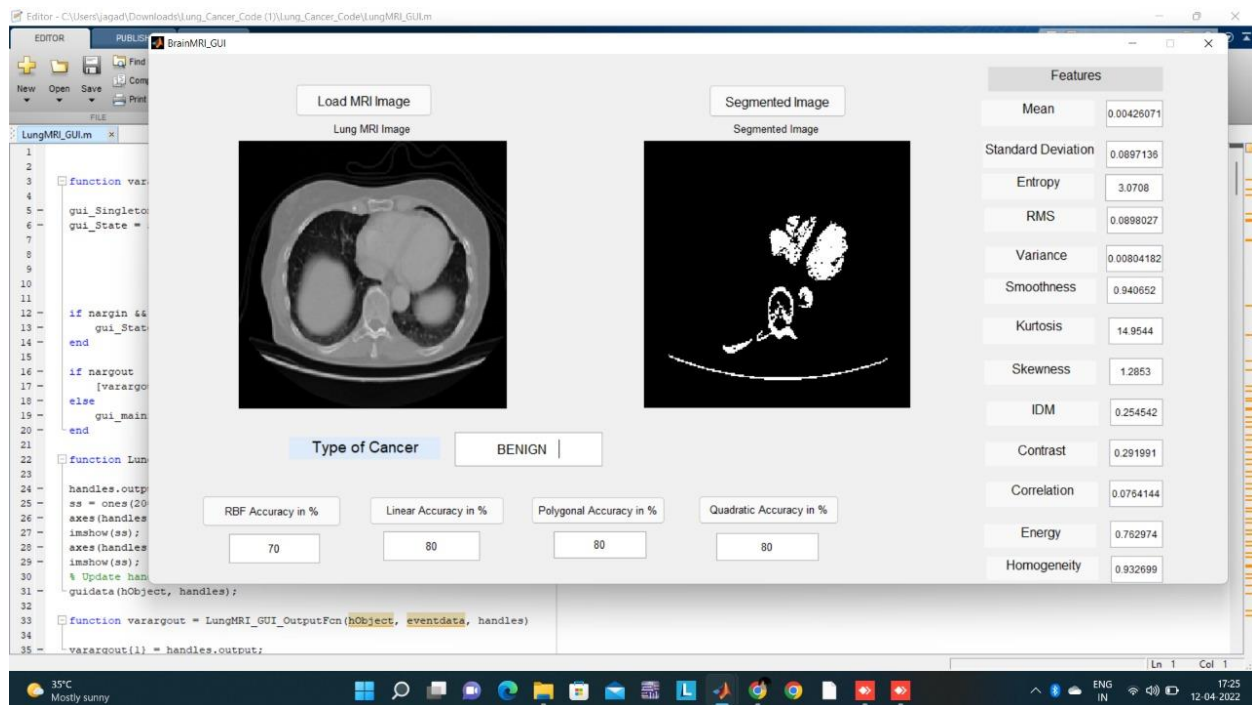
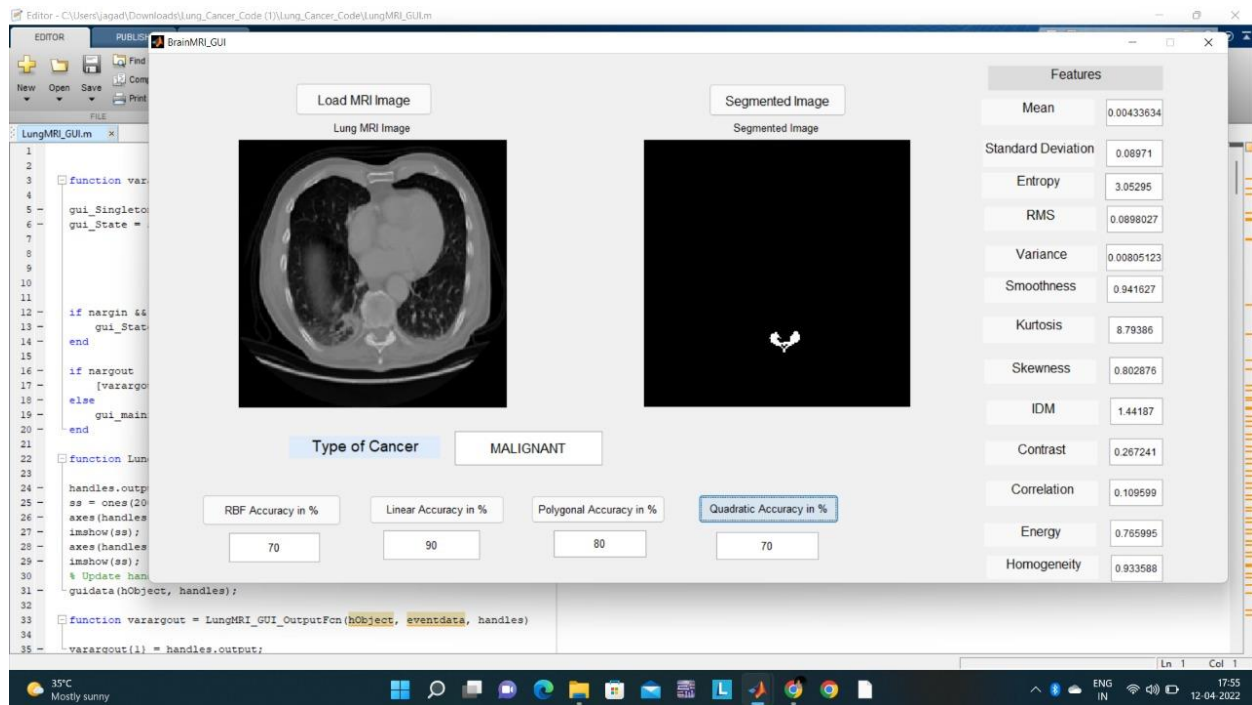
A tumour is essentially the unchecked proliferation of malignant cells in any organ or tissue of the body, regardless of location. Two kinds of tumours exist: primary and metastatic.

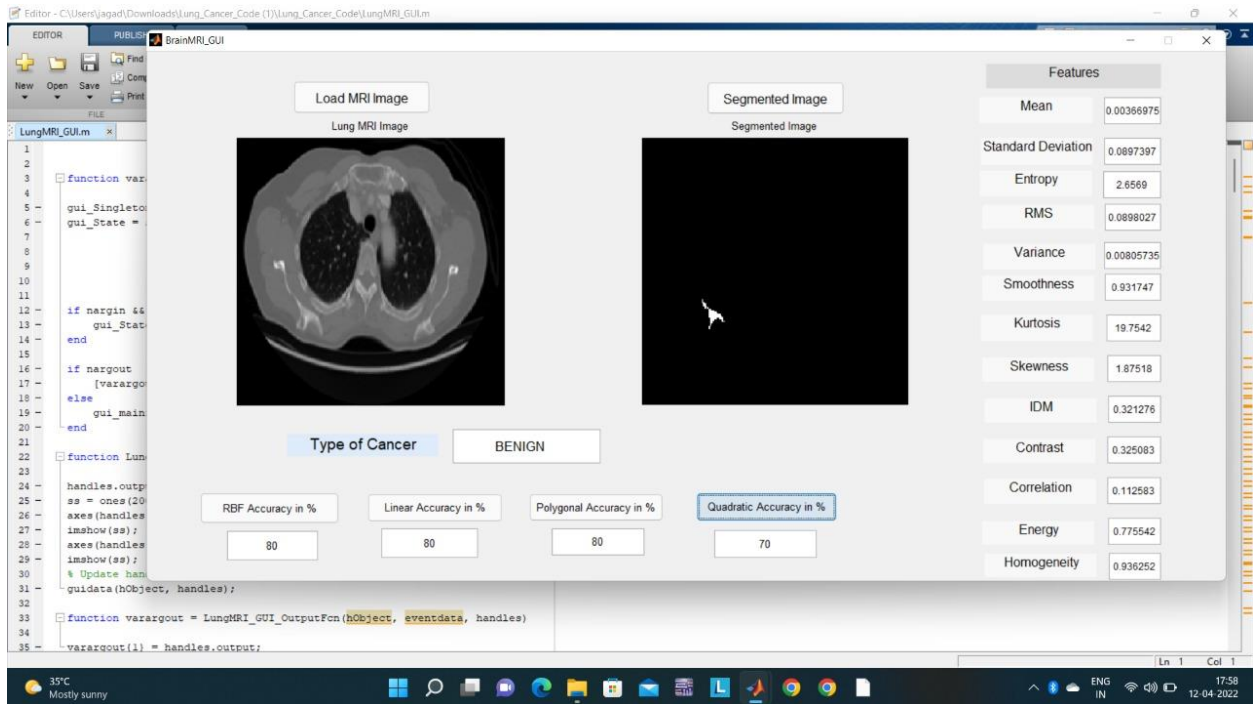
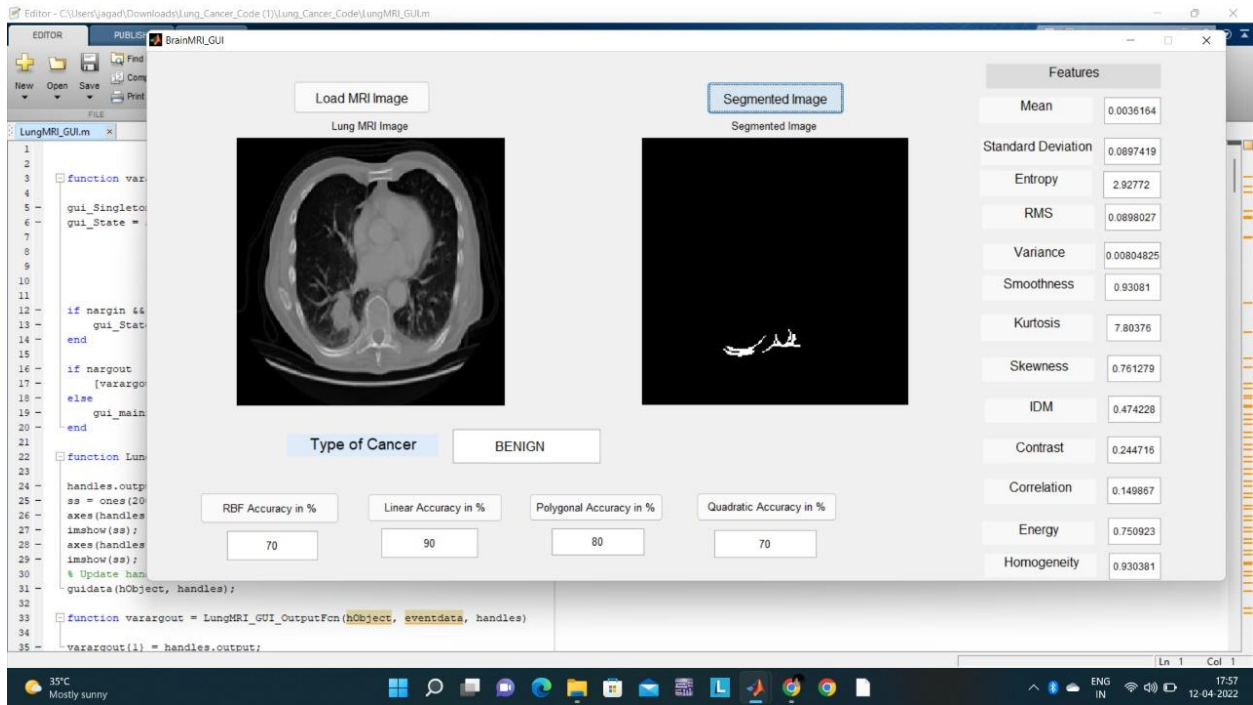
1. Benign - A benign tumour is one that has a low grade. No active tumour cells can be found in its structure, which makes it an ideal specimen.
2. Malignant - This is a cancerous tumour with a high grade. Its structure is non-uniform, and it contains living cells.

The below mentioned four figures belongs to Benign tumour, can see the tumour in the segmented image is detected using CNN algorithm. With the output user can predict the accuracy of the tumour in percentage along with the accuracy some features are also be predicted using convolutional neural network such as mean standard deviation, entropy etc. This is explained in detail in chapter-5. Moving further with the benign type of cancer (last three output figures, by looking into the accuracy percentages and features one can get that is a low grade tumour.

By revising number of literature surveys, most of the authors selected CNN because of its accuracy over 90%. So, by taking it as a primary reason to select CNN and get accuracy more than 95% which can seen in the following output figures.







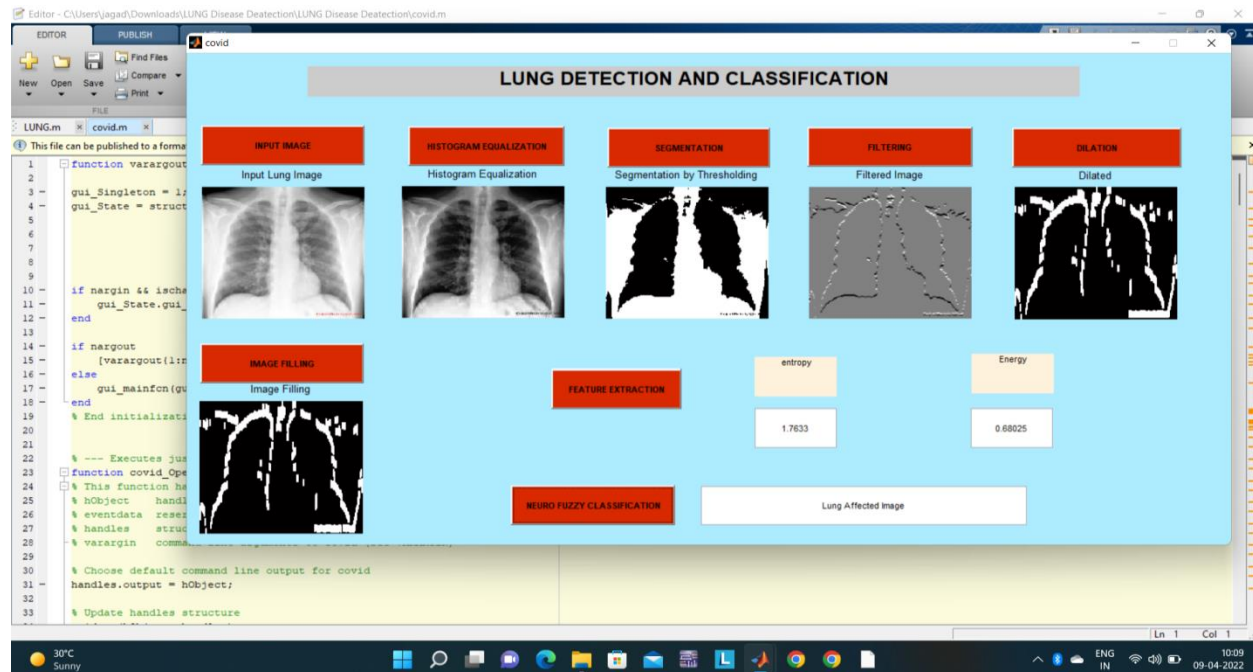
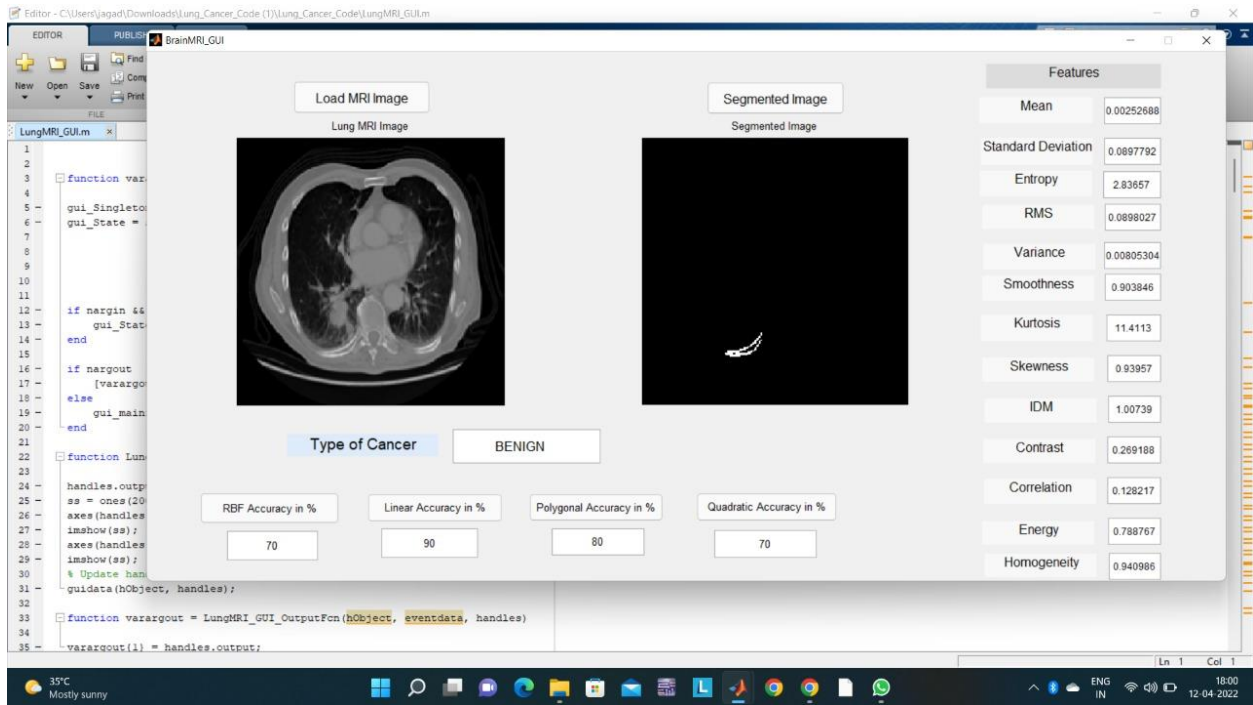
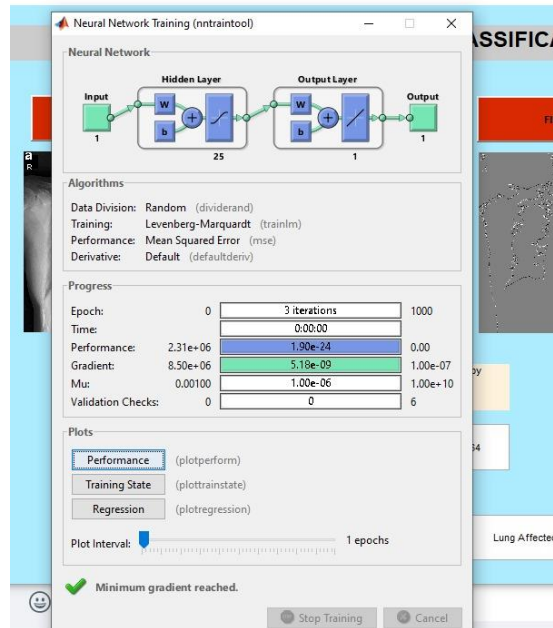


FIG: 6.1: OUTPUT SCREENSHOTS

6.2. COMPARISON



6.3. CONCLUSION

Time was taken into consideration in the development of a new image enhancing approach for early disease identification and treatment phases. This analysis is superior and correct since the quantity elements of this analysis, such as angel superior appraisal and aspartame date anywhere, were adopted on cheap preprocessing procedures that allow physicists to understand a portion of Gaussian rules. For feature extraction, segmentation concepts may be used as a starting point. 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 corrected pictures and comparison square measured pixels proportion and mask-labeling with high accuracy and study the operation.

The conclusion of the paper is that CNN model Architecture which is used in our proposed

system to predict the lung tumor performs well. The system classifies the image to normal and lung Tumor one. The model has a 98.5 percent accuracy rate. On CT scans, our approach demonstrated that CNN can correctly identify lung cancer. When it comes to aiding radiologists and doctors in making a lung tumour diagnosis, the CNN model seems to offer potential.

REFERENCES

1. Al-Hawary MM, Francis IR, Chari ST, et al. Pancreatic ductal adenocarcinoma radiology reporting template: consensus statement of the Society of Abdominal Radiology and the American Pancreatic Association. *Radiology* 2014; 270: 248–60.
2. Anita chaudhary, SonitSukhraj Singh “Lung Cancer Detection on CT Images by Using Image Processing”2012 International Conference on Computing Sciences
3. Dewitt J, Devereaux BM, Lehman GA, Sherman S, Imperiale TF. Comparison of endoscopic ultrasound and computed tomography for the preoperative evaluation of pancreatic cancer: a systematic review. *Clin Gastroenterol Hepatol* 2006; 4: 717–25; quiz 664.
4. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; 542: 115–18.
5. 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
6. NihadMesanovic, HarisHuseinagic, Matija Males, ,MislavGrgic, Emir Skejic, MuamerSmajlovic ”Automatic CT Image Segmentation of the Lungs with Region Growing Algorithm”
7. Prof. Samir Kumar Bandyopadhyay “Edge Detection From Ct Images Of Lung” International Journal Of Engineering Science & Advanced Technology Volume - 2, Issue - 1, 34 – 37
8. QinghuaJi,Ronggang Shi “A Noval Method of Image Segmentation Using Watershed Transformation”2011 International Conference on Computer Science and Network Technology
9. Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res* 2014; 74: 2913–21.
10. Ryan DP, Hong TS, Bardeesy N. Pancreatic adenocarcinoma. *N Engl J Med* 2014;

371: 1039–49.

11. 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

12. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019; 69: 7–34.

13. Yasaka K, Abe O. Deep learning and artificial intelligence in radiology: current applications and future directions. PLoS Med 2018; 15: e1002707.