

COVID-19 Classification and Lesion

Localization from Chest CT

Submitted in partial fulfillment of the requirements for

the award of

Bachelor of Engineering degree in Computer Science and Engineering

by

Vulli Bhaskar Venkata Satish

(Reg no:37110848)



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SCHOOL OF COMPUTING

SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited with Grade "A" by NAA

JEPPIAAR NAGAR, RAJIV GANDHI SALAI,

CHENNAI - 600 119

April - 2021





SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited with "A" grade by NAAC

Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai – 600119

www.sathyabama.ac.in



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the bonafide work of **Vulli Bhaskar Venkata Satish (37110848)** who carried out the project entitled as **COVID-19 Classification and Lesion Localization from Chest CT** under my supervision from November 2019 to April 2020.

Internal Guide

Mr. VIGNESH. R, M.E., (Ph.D.,)

Head of the Department

Dr. S. VIGNESHWARI, M.E., Ph.D.,

Dr. L. LAKSHMANAN, M.E., Ph.D.,

Submitted for Viva voce Examination held on _____



Internal Examiner

External Examiner

DECLARATION

I **Vulli Bhaskar Venkata Satish (37110848)** hereby declare that the Project Report entitled “**COVID-19 Classification and Lesion Localization from Chest CT**” done by me under the guidance of **Mr. Vignesh. R, M.E., (Ph.D.)** is submitted in partial fulfillment of the requirements for the award of Bachelor of Engineering degree in Computer Science and Engineering.

DATE:

SIGNATURE OF THE CANDIDATE

PLACE:



ACKNOWLEDGEMENT

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

I convey my thanks to **Dr. T. Sasikala, M.E., Ph.D.**, Dean, School of Computing, **Dr. S. Vigneshwari, M.E., Ph.D.**, and **Dr. L. Lakshmanan, M.E., Ph.D.**, Heads of the Department of Computer Science and Engineering for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide **Mr. Vignesh. R, M.E., (Ph.D.,)** for his valuable guidance, suggestions and constant encouragement paved way for the successful completion of our project work.

I wish to express my thanks to all Teaching and Non-teaching staff members of the **Department of Computer Science and Engineering** who were helpful in many ways for the completion of the project.



ABSTRACT

Accurate and rapid diagnosis of COVID-19 suspected cases plays a crucial role in timely quarantine and medical treatment. Developing a deep learning-based model for automatic COVID-19 diagnosis on chest CT is helpful to counter the outbreak of SARS-CoV-2. A weakly supervised deep learning framework was developed using 3D CT volumes for COVID-19 classification and lesion localization. For each patient, the lung region was segmented using a pre-trained UNet; then the segmented 3D lung region was fed into a 3D deep neural network to predict the probability of COVID-19 infectious; the COVID-19 lesions are localized by combining the activation regions in the classification network and the unsupervised connected components. 499 CT volumes were used for training and 131 CT volumes were used for testing. Our algorithm obtained 0.959 ROC AUC and 0.976 PR AUC. When using a probability threshold of 0.5 to classify COVID-positive and COVID-negative, the algorithm obtained an accuracy of 0.901, a positive predictive value of 0.840 and a very high negative predictive value of 0.982. The algorithm took only 1.93 seconds to process a single patient's CT volume using a dedicated GPU. Our weakly-supervised deep learning model can accurately predict the COVID-19 infectious probability and discover lesion regions in chest CT without the need for annotating the lesions for training. The easily-trained and high-performance deep learning algorithm provides a fast way to identify COVID-19 patients, which is beneficial to control the outbreak of SARS-CoV-2.



TABLE OF CONTENT

CHAPTER NO	TITLE	PAGE NO
	ABSTRACT	V
	LIST OF FIGURES	VIII
1.	INTRODUCTION	1
	1.1 INTRODUCTION	1
2.	LITERATURE SURVEY	5
		4
3.	DESIGN METHDOLOGY	8
	3.1 EXISTING SYSTEM	8
	3.2 PROPOSED SYSTEM	8
	3.3 MODULE DESCRIPTION	8
	3.3.1 PREPROCESSING	8
	3.3.2 SEGMENTATION	9
	3.3.3 CLASSIFICATION	9
	3.4 FUZZY C MEANS CLUSTERING	9
	3.5 NEURAL NETWORK	10
	3.5.1 THE BASICS OF NEURAL NETWORKS	11
	3.5.1.1 INPUT LAYER	11
	3.5.1.2 HIDDEN LAYER	11
	3.5.1.3 PATTERN LAYER	11
	3.5.1.4 DECISION LAYER	11
	3.5.2 USE OF NEURAL NETWORKS	11



	3.5.3 NETWORK LAYERS	13
4	SYSTEM REQUIREMENTS	14
	4.1 HARDWARE REQUIREMENTS	14
	4.2 SOFTWARE REQUIREMENTS	14
	4.3 INTRODUCTION TO MATLAB R2018a	15
	4.3.1 THE MATLAB SYSTEM	15
	4.4 IMAGE PROCESSING TOOLBOX	16
	4.5 COMPUTER VISION TOOLBOX	17
	4.6 IMAGE ACQUISITION TOOLBOX	18
	4.7 DESIGN OF PROPOSED SYSTEM	18
	4.8 UML DIAGRAM	19
	4.9 <i>ACTIVITY DIAGRAM</i>	21
	4.10 SEQUENCE DIAGRAM	22
	4.11 USE CASE DIAGRAM	22
	4.12 DATAFLOW DIAGRAM	23
5	RESULT & DISCUSSION	28
6	Summary and Conclusions	32
	6.1 CONCLUSION	32
	6.2 FUTURE ENHANCEMENT	32



LIST OF FIGURES

FIGURE NO	TITLE	PAGE NO
1.1	CHEST CT SHOWING THE EFFECT IN THE LUNG	9
4.1	BLOCK DIAGRAM	
4.2	ACTIVITY DIAGRAM	14
4.3	SEQUENCE DIAGRAM	16
4.4	USE CASE DIAGRAM	17
4.5	DATA FLOW DIAGRAM 1	18
4.6	DATA FLOW DIAGRAM 2	19
4.7	DATA FLOW DIAGRAM 3	20
4.8	DATA FLOW DIAGRAM 4	21
5.1	INPUT IMAGE	22
5.2	GRAYSCALE IMAGE	31
5.3	RESIZE IMAGE	32
5.4	NOISE IMAGE	33
5.5	FILTER IMAGE	34
5.6	RECOLOR IMAGE	35
5.7	NEURAL NETWORK	36
5.8	OUTPUT	51



CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

SINCE Dec 2019, increasing of a new corona virus was first found in Wuhan located in china [1], [2] which might cause acute lung disease and even causing death lung distress syndrome (ARDS) [3]. The International Committee of Taxonomy of Viruses announced as new coronavirus as SARS-COV-2 [4] By the world health organization coronavirus was renames as novel corona virus [5]. The COVID - 19 has been transfer from one human to another human transmission [6], [7], and because of the huge crowd and expansive population before the Chinese spring fest, this new coronavirus has spread fast to other places in china and all other countries. Data given by the National Health Commission of China [8], up to 29/03/2020, China has reported more than 80000 identified cases with novel corona virus, along with “3,311” death cases. Effected persons are more than 73% of the cases in Hubei province. Nearly 200 countries effected from corona virus after the rapid growth in COVID-19 in china. The higher metabolic demands for glucose and amino acids in malignant tumors than normal tissues have been proven by both in vitro and in vivo experiments. Although the image quality and standard deviation of the region of interest (ROI) data were deteriorated in the small injection dose. In European Countries up to “6,00,000” people was infected due to corona virus and 30,000 death cases in Europe. Countries like America (“1,03,321” cases), Italy (“92,472” cases), china (“82,000” cases), Spain (“72,248” cases), Germany (“52,547” cases), Iran (“38,309” cases), France (“37,145” cases) with huge amount of effected people, spreading the awareness of covid-19 all over the world [11]. Due to limited number of kits and social distancing it's getting difficult to check all the people, with the help of Chest CT it is easier to check the patients whether he is affected or not. And this process can be implemented in every hospital by checking the Chest CT rays. And

people who are in quarantine are rapidly increasing, this is the fast and easier process to check Covid-19.

ADENOCARCINOMA:

Aden carcinoma is the most common type of lung cancer, accounting for more than 40% of lung cancers, 60% of the NSCC, and more than 70% of surgically resected cases.^{3,4} The incidence of adenocarcinoma has risen steadily over the past few decades. Lung adenocarcinoma commonly forms a peripherally located mass with central fibrosis and pleural puckering. It can also have a variety of other gross appearances, including centrally located mass, diffuse lobar consolidation, bilateral multinodular distribution, and pleural thickening. By definition, lung adenocarcinoma is a malignant epithelial neoplasm with glandular differentiation or mucin production. When such morphologic features are recognized, the tumor can be designated as adenocarcinoma, even in small biopsy specimens. Lung adenocarcinoma cells usually express pneumocytic markers. Thyroid transcription factor (TTF-1) and NapsinA are expressed in more than 85% of the lung adenocarcinoma cases and thus can serve as markers of adenocarcinoma or adenocarcinoma differentiation in poorly differentiated tumor and in limited biopsy sampling material ^{5–7} Tumor classification based on ancillary tests such as immunohistochemistry (IHC) is designated as “NSCC, favor adeno carcinoma” in a small biopsy specimen. Resection specimens allow a more detailed sub classification. There has been significant refinement in adeno carcinoma classification in recent years based on close pathologic and clinical correlation.^{3,8} The major histological types have been validated to bear prognostic significance delineated by the tumor grade.^{9–14} Multiple gene alterations can occur in adenocarcinomas, with approved molecular targeted therapy available to improve patient survival.

Differential Diagnosis of Lung Tumor with Positron Emission Tomography:

The higher metabolic demands for glucose and amino acids in malignant tumors than normal tissues have been proven by both vitro and in vivo experiments. Metabolic evaluation of cancer with positron emission tomogram.

phy (PET) has been clinically demonstrated for brain tumor (8) and lung cancer using L-' ^{14}C -methionine (MET) and " ^{18}F 2-fluoro 2-deoxy-D-glucose" (FDG) (10). However, limited resolution of 'PET' in some patients with small lesions and a known diagnosis of cancer makes it difficult to appraise the differential ability of 'PET'. To determine the utility of 'PET' in the differential diagnosis of chest diseases, the differential diagnosis of chest disease using 'MET' and 'FDG'.

Radiopharmaceuticals:

MET and FDG were synthesized using mechanical synthesis systems as described previously (9, 11,12). Radiochemical integrities were over 99%. Quality-assurance trials for clinical practice have been carried out in accordance with the safety instructions of our university. Tracers were randomly assigned to each patient. Twenty-four patients whose mean injection dose Twenty-two patients whose mean injection dose was 7.3 (3.5 - 13.0) 'MCI' (270 MBq) were studied with FDG. The effect of variance on the injection doses has been evaluated previously (9); injection dose as well as the tumor/muscle radioactivity ratio (T/ M ratio), although the image quality and standard deviation of the regionof interest (ROI) data were deteriorated in the small injection dose.

Acinar Adenocarcinoma :

Acinar adenocarcinoma is a common type of adenocarcinoma with tumor cells arranged in classic glandular structure on a fibroelastic stroma. It is important to separate demosplastic stroma in this pattern from preexisting alveolar structures with thickened fibroelastic alveolar septa sometimes seen in a lepidic pattern. It is of note that the tumor cells displaying more complex growth patterns, such as a cribriform pattern, likely represent a poor prognostic subtype conveying a significant risk of recurrence.

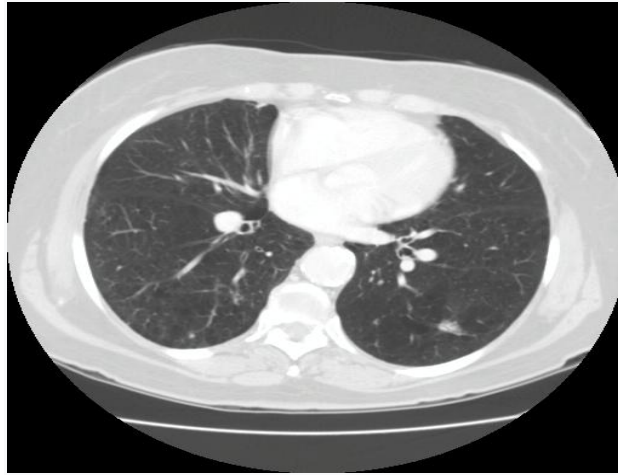


Fig 1.1 Chest CT showing the effect in the lung

Lung cancer, also known as lung carcinoma, is a malignant lung tumor characterized by uncontrolled cell growth in tissues of the lung. This growth can spread beyond the lung by the process of metastasis into nearby tissue or other parts of the body. Most cancers that start in the lung, known as primary lung cancers, are carcinomas. The two main types are small-cell lung carcinoma (SCLC) and non-small-cell lung carcinoma (NSCLC). The most common symptoms are coughing (including coughing up blood), weight loss, shortness of breath, and chest pains. Avoidance of risk factors, including smoking and air pollution, is the primary method of prevention. Treatment and long-term outcomes depend on the type of cancer, the stage (degree of spread), and the person's overall health. Most cases are not curable. Common treatments include surgery, chemotherapy, and radiotherapy. NSCLC is sometimes treated with surgery, whereas SCLC usually responds better to chemotherapy and radiotherapy.

Biomineralogy of lung tumors :

Adenocarcinomas can be located on different levels of the bronchial tree. In the main or lobar bronchi, glandular carcinomas with a very low degree of malignancy, i.e. adenomas, are generally located. They originate from glands of the bronchial wall. The main mass grows in the form of a polypoid to the bronchial lumen. This form is basically nonmetastatic and can be classified as a locally malignant tumor. In the

large bronchi, types of tumors typical for different organs may appear: cylindromatous carcinoma typical for salivary glands and mucinous epithelial tumor usually seen in salivary glands and the mucosa of the nose and paranasal sinuses. In the distal bronchioles, adenocarcinomas with a high tendency to generalization are frequently located. They occur much more often than all the other forms in this group. They are located among the richly vascularized pulmonary parenchyma and metastasize quickly through the blood vessels. Two subtypes of bronchioalveolar carcinoma may develop on the border of the respiratory part of the lung: mucinous and non-mucinous. These tumors are characterized by growth displacing pneumocytes on the surface of the alveoli and lining the alveolar spaces of the lungs. Broncho-alveolar carcinoma develops multifocally in a given pulmonary area.

CHAPTER 2

LITERATURE SURVEY

2.1 Project Title : “Convolutional STN for Weakly Supervised Object”
Author Name : ‘Akhil Meethal’, ‘Marco Pedersoli’, ‘Soufiane Belharbi’ and ‘Eric’
Year of Publish : ‘2020’
Abstract :

Weakly-supervised object localization is a daunting task during which the object of interest shall be concentrated upon learning what they look like. State of the art techniques reusing the structural design of standard set CNN by utilizing the stimulation maps of the final layer to localize the item. Although this approach is straightforward and works pretty well, object the regionalization depends on the different characteristics than categorization, therefore, a specialized location-based mechanism that is necessary during training in order to improve performance. Throughout this document, we recommend a convolutions, multi-scale geographic localization network which provides precise location for the item relevant. Laboratory findings on ImageNet datasets demonstrate that the proposed method delivers the competitiveness poorly overseen localize.



2.2 Project Title : “Development and Validation of Deep Learning Algorithms for Detection of Critical Findings in Head CT Scans”

Author Name : ‘Sasank Chilamkurthy¹’, ‘Rohit Ghosh¹’, ‘Swetha Tanamala¹’, ‘Mustafa Biviji²’.

Year of Publish : ‘2018’

Abstract :

Significance Non-contrast head CT scan is the existing standard for preliminary imaging of patients diagnosed with traumatic brain injury and stroke manifestations. Design and Environments Review accumulated a dataset that contains the ‘313,318’ head CT scans in conjunction with their medical records from many different regions. A portion of this data set has been used to authenticate and the remainder to create systems. Furthermore, dataset that has been compiled from individual centers through two consignments B1 & B2 to be clinically verify the systems. Territory under the managers characteristics curve for each and every conclusion was mostly used to assess algorithms used.

2.3 Project Title : “NormGrad: Finding the Pixels that Matter for Training”

Author Name : ‘Sylvestre Alvisé Rebuffi¹ Ruth Fong¹ Xu Ji’

Year of Publish : ‘2019’

Abstract :

Several different families the remarkable methods, based upon the comparison between the indicators, closed form recipes mixing inclines with introductions or on agitation disguises, all attention which areas of an image are solely responsible for the model’s inference. Throughout this document, we are quite curious about the sites of the picture that have contributed to the model’s training. Earliest, we recommend that the essential standard method that we have taken from the formula for calculating which will be used to calculate the steepness of the scales for a 1 -1 folding layer. In this manner, we shall not objective the parts of the picture that are contributing to the pattern’s production in more precisely the places which are

responsible for the excellent training of the prototype on this picture. On The Other Hand, we also demonstrate that the same meta-learning technique can be used to isolate the antagonistic places that could lead to the deterioration of the template.

2.4 Project Title : “Deep Residual Learning”
Author Name : ‘Asifullah Khan’
Year of Publish : ‘2016’
Abstract :

Deep characteristics are important for optical recognition assignments, but profoundly mesh have suffered from disappearing/explosion inclines. And adding multiple-layer results in a higher preparation error. The recommended ResNet: learn the remaining function rather of your unused purposes.

2.5 Project Title : “A Brief Introduction to Weakly Supervised Learning”
Author Name : ‘Zhi-Hua Zhou’
Year of Publish : ‘2018’
Abstract :

Supervised learning techniques create the forecasting model to learn from an enormous number of training instances, where every training example contains a stamp showing the ground-truth productivity. Even Though existing technologies have been able to achieve enormous success, it is remarkable that in various assignments is hard to get strong regulatory oversight information such as the fully ground-truth labels for at the high cost of data branding process. Therefore, it would be desirable for computer learning methods to work with weak supervision. This article reviews some research progress of weakly supervised learning, focusing on three typical types of weak supervision: incomplete supervision where only a subset of training data are given with labels; inexact supervision where the training data are given with only coarse-grained labels; inaccurate supervision where the given labels are not always ground-truth.

2.6 Project Title : U-Net: Convolutional Networks for Biomedical
Image Segmentation
Author Name : Olaf Ronneberger, Philipp Fischer, and Thomas Bro
Year of Publish : 2015
Abstract :

There is large consent that successful training of deep networks requires many thousand annotated training samples. In this paper, we present a network and training strategy that relies on the strong use of data augmentation to use the available annotated samples more efficiently. The architecture consists of a contracting path to capture context and a symmetric expanding path that enables precise localization. We show that such a network can be trained end-to-end from very few images and outperforms the prior best method (a sliding window convolutional network) on the ISBI challenge for segmentation of neuronal structures in electron microscopic stacks.

CHAPTER 3

DESIGN METHDOLOGY

3.1 Existing system:

For image specific learning Interactive segmentation often requires and to assist huge difference among the images but neural network isn't adaptive to different test images. From training the images at the testing stage this model will learn without image specific. Extract the foreground for the given region of interest was done with the help of bounding-box-based segmentation algorithm.

3.2 Proposed System:



The Proposed system concentrate on interactive tumor segmentation of medical image sequences using back propagation neural network. Using neural network function, they proposed work pattern-based classification. Fuzzy c means clustering is intended within the proposed area. For segmented area and position threshold is required. Finally, the classifier is employed to test whether the input image is COVID or Non-COVID. In this project, lung images of a patient are taken as input. The images are pre-processed and segmented and then classified for the required feature.

3.3 Module Description:

3.3.1 Preprocessing

If the input images are color images means we convert color image into gray scale image. In binary image ones will replace with zeros and zeros will replace with ones; black and white will replace as zeros and ones. And in the result image, lighter areas become darker and darker areas become lighter. In image it has some disturbance like noise. In this process, the aim is to give priority to certain image features for sequential analysis or for display image. Examples include edge enhancement and contrast, sharpening, pseudo-coloring, noise filter, and with magnifying. Image process is useful in image analysis, feature extraction of image and a display image. Removal of noise from input image is done by median filter. A median filter is effective when the goal is to decrease noise and conserve edges.

3.3.2 Segmentation

Before the classifier, the image features like color, weight, and depth and pixel information should be apply. Fuzzy C clustering algorithm is used for the segment portion at the defected areas. To find the objects and edges in images we use segmentation process. In this process labels should be assigning to every pixel in the image such that pixels with the same label will share few characteristics. In this process the output should be collectively fill the hole image with some set of segments, or with a set of extracted images.

3.3.3 Classification

In this module image and testing the image is done with the help of neural network and weight image classifier. There are three layers which are interconnected to each other like, input layer is connected to hidden layer within the output layer. In input layer the raw data is change into the network signals. Hidden process is resolved by the activities of the input process. Weights between the input units and hidden units are done by the hidden process.

And the output process is defined with the help of hidden units and the weights between the output process and hidden process. Hidden layer can build their own input representation. The hidden unit can represent after modifying the weights when each hidden unit is active. The final image will compare with every image in dataset whether it is negative or positive. The final image will compare with every image in dataset whether it is negative or positive.

3.4 Fuzzy C means Clustering:

Clustering methods have constructive tools that will examine the structure the information and emerged as well-liked methods for the unsupervised pattern acceptance and are used in numerous application fields such as image identification, data mining, automatic learning, etc. Organization of a set of models in cluster can create a low or high degree of both similarity and dissimilarity depending on whether the attachment of patterns to a well-known or unknown cluster. Generally, the grouping can be split into two groups like tiered and splitting the clustering. Tiered clustering outcomes and contributes to the appropriate output within the nested hierarchical structure pattern, where like dividing the grouping delivers the output with the target function that make up the partition of the contribution into a set of fixed quantity of bundles. In accordance with the target function concept, FCM is the best full technique and is the cause for being used extensively.

Clustering is either be a tough or fuzzy type. In the former category, the models are renowned in a well-defined group boundary area. However due to the intersecting nature of the constellation limits, certain types of patterns can be specified through a single group in the cluster or disparate collective. This property restricting the use of



hard clustering in real-world submissions. To mitigate these restrictions to the fuzzy type clustering has come into the image and helping to provide additional information about the associations of the models. Following the fuzzy logic introduced through 'Zadeh', scientists have set the fuzzy logic into the grouping. The Fuzzy clustering issues have been expansively studied and the establishment of fuzzy clustering had been suggested by 'Bellman' and 'Ruspini'. Fuzzy clustering issues could be divided into three branches: (a) Based upon the fuzzy relative; (b) Based upon the fuzzy ruling knowledge; and (c) That is based upon increasing the efficiency of the target function. The vague clustering in accordance with the goal function is somewhat universally famous for being Fuzzy c-means crowding (FCM). In FCM method, a template may belong to everybody cluster's lessons with the certain fuzzy affiliation the diploma.

3.5 Neural Network

Complex Neural Network (NN) is the information-processing model that draws inspiration from the way the biological in the nervous system, for example in the brain, processes information. The most important component of this pattern is the innovative structure of information that is contained in the handling system. It is made up of many interwoven handling elements employed in harmony solve a particular problem. NNs, as humans, learn through examples. An NN have been configured for a particular purpose, like pattern acknowledgement or data categorization, through the learning process. Instruction in organic systems requires the modifications to the synaptic pathways that exist among the nerve cells.

3.5.1 The Basics of Neural Networks

Neural network is usually within the layers. Layers have been made up of a quantity of interrelated 'nodes' that include an 'activation function'. Models are introduced to the networks through the 'input layer', communicating with one or more 'hidden layers' if the real processing is made with a system of weighted

'connections'. The hidden layers then connect an 'output layer' where the answer exported as shown in the following picture.

All NN networking has four layers:

3.5.1.1 Input layer — There's only one neuron into the input layer for every forecasting variable. In this specific case of categorical variables, $N-1$ nerve cells are used in cases where n represents the number of categories. Input to the nerve cells (and processed prior to the input layer) normalizes the range of the values through subtracting from the average and by splitting the quartiles range. Input to the neuronal nourish the values for the nerve cells in the hidden layer.

3.5.1.2 Hidden layer — This layer is the one neuron for each case within the necessary training data collection. Neuron's holds the value of the predictive variables in connection with the case in conjunction with the destination value. When you are presented with the x vector of input values from an input layer, a concealed nerve cells that will calculate the Euclidean distance of the test scenario dating back to the neurons in the central point and then applying the RBF core functionality when you use sigma's value(s). The resultant value is passed through to the neurons within the pattern layer.

3.5.1.3 Pattern layer / Summation layer — Another layer in your network is various designed for NN networks and to GRNN networks. For NN networks there is yet another pattern neuron for everyone and every classification of the destination variable. Real target category of each training case will be stored per hidden neuron; weighted average value coming out of a hidden neuron is being fed solely to the pattern neuron which is equivalent to the hidden neuron's classification. Pattern on the neurons add the values for the course they constitute (hence, it is a weighted vote for that category).

3.5.1.4 Decision layer — The decision layer is distinct from the NN and GRNN systems. For NN networks, the decision layer will compare the weighted votes for

every target category accumulated in the pattern layer and utilizes the largest vote to anticipate the target category.

For GRNN networks, the decision layer divides the value accumulated during the numerator summation unit with the value in the denominator summation unit and uses outcome with the expected target value.

3.5.2 Use of neural networks:

Neural nets, including their amazing ability to derive meaning from sophisticated or inaccurate, may be used to extract patterns and identify tendencies that are just too complex to get noticed either through humans or some other computer equipment. Trained professional neural circuits can be considered as an "expert" under the category of information it has been assigned to carry out an analysis. This specialist can also be used to estimates for new situations relevant and respond to the.

Additional benefits include:

1. Adaptive learning: A skill to find out how to perform the tasks depending on the data specified for the training or early experience.
2. Self-Organization: An NN can create its own enterprise or depiction of the information that it receives for the duration of the training time.
3. Real Time Operation: NN calculations could be carried out at the same time, and specific hardware devices are specifically designed and produced that will exploit this capability.
4. Culpability Easiness via Redundant Data Programming: In a portion of a network is leading to the appropriate deterioration in the quality of execution. Nevertheless, specific network features might be preserved even with substantial network causing harm.



3.5.3 Network layers:

Perhaps the most common type of neural network system consists of three clusters, or strata, of units: a single layer of **"input"** units are linked to one layer of **"hidden"** parts, that is attached to one layer of **"output"** components.

The relevant activity of input to the units is representing raw data which is supplied into the grid.

Activities carried out of every secret device shall be determined by the activities of the input apparatus and the weight to the links between input and the concealed units.

That the conduct of the production units varies depending on the activity of the concealed units and weight between both the concealed and the production units.

This straightforward type of grid is intriguing because the concealed troops have the freedom to create their own representations of the information entered. Weights are between input and concealed units to establish when every secret device has been activated, and thus by adjusting these substances, the concealed device will be able to select what it stands for.

We would also like to differentiate between 'single-layer' and 'multi-layer' styles. One single-layer association, under which each unit is linked to each other, represents the most common case and that it is of more prospective computing capacity than the hierarchical organized multi-layer corporations. Across multiple-layer grids, devices are frequently numbered on coating, as opposed to in accordance with the worldwide numbered list.



Chapter 4

System Requirements

4.1 HARDWARE REQUIREMENTS

Hard-drive	:	40GB and More
RAM	:	512MB and More
Cpu	:	Pentium III and More

4.2 SOFTWARE REQUIREMENTS

MATLAB R2018a
Operating System Is Windows XP and More

MATLAB (matrix laboratory) has multiple-paradigm numerical computing environment and fourth-generation programming language. A proprietary programming language developed by Math Works, MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages, including C, C++, Java, Fortran and Python.

Although MATLAB is intended primarily for numerical computing, an optional toolbox uses the Mu PAD symbolic engine, allowing access to symbolic computing abilities. An additional package, Simulink, adds graphical multi-domain simulation and model-based design for dynamic and embedded systems.

The basic data structure in MATLAB is the array, an ordered set of real or complex elements. This object is naturally suited to the representation of images, real-valued ordered sets of color or intensity data.



MATLAB stores most images as two-dimensional arrays (i.e., matrices), in which each element of the matrix corresponds to a single pixel in the displayed image. (Pixel is derived from picture element and usually denotes a single dot on a computer display.) For example, an image composed of 200 rows and 300 columns of different coloured dots would be stored in MATLAB as a 200-by-300 matrix. Some images, such as RGB, require a three-dimensional array, where the first plane in the third dimension represents the red pixel intensities, the second plane represents the green pixel intensities, and the third plane represents the blue pixel intensities.

This convention makes working with images in MATLAB similar to working with any other type of matrix data, it makes the full power of MATLAB available for image processing applications. For example, you can select a single pixel from an image matrix using normal matrix subscripting.

4.3 INTRODUCTION TO MATLAB R2018a:

4.3.1 The MATLAB System

The MATLAB system consists of these main parts:

Desktop Tools and Development Environment

This is the set of tools and facilities that help you use and become more productive with MATLAB functions and files. Many of these tools are graphical user interfaces. It includes the MATLAB desktop and Command Window, a command history, an editor and debugger, a code analyzer and other reports, and browsers for viewing help, the workspace, files, and the search path.

Mathematical Function Library

This is a vast collection of computational algorithms ranging from elementary functions, like sum, sine, cosine, and complex arithmetic, to more sophisticated functions like matrix inverse, matrix Eigen values, Bessel functions, and fast Fourier transforms.



The Language

This is a high-level matrix/array language with control flow statements, functions, data structures, input/output, and object-oriented programming features. It allows both "programming in the small" to rapidly create quick and dirty throw-away programs, and "programming in the large" to create large and complex application programs.

Graphics

MATLAB has extensive facilities for displaying vectors and matrices as graphs, as well as annotating and printing these graphs. It includes high-level functions for two-dimensional and three-dimensional data visualization, image processing, animation, and presentation graphics. It also includes low-level functions that allow you to fully customize the appearance of graphics as well as to build complete graphical user interfaces on your MATLAB applications.

External Interfaces

This is a library that allows you to write C and FORTRAN programs that interact with MATLAB. It includes facilities for calling routines from MATLAB (dynamic linking), for calling MATLAB as a computational engine, and for reading and writing MAT-files.

4.4 Image Processing Toolbox

The Image Processing Toolbox is a collection of functions that extend the capability of the MATLAB numeric computing environment. The toolbox supports a wide range of image processing operations, including:

- Spatial image transformations
- Morphological operations
- Neighborhood and block operations
- Linear filtering and filter design
- Transforms



- Image analysis and enhancement
- Image registration
- De blurring
- Region of interest operations

Many of the toolbox functions are MATLAB M-files, a series of MATLAB statements that implement specialized image processing algorithms. You can extend the capabilities of the Image Processing Toolbox by writing your own M-files, or by using the toolbox in combination with other toolboxes, such as the Signal Processing Toolbox and the Wavelet Toolbox.

4.5 Computer Vision Toolbox

The field is of importance for such various applications as autonomous vehicles, navigating with the help of images, captured by a mounted camera, and high precision measurements using images, taken by calibrated cameras. In this paper, we will present a number of numerical routines, implemented in MATLAB, that are useful in a variety of computer vision applications. The collection of routines will be called the Computer Vision Toolbox. One of the main problems in Computer Vision is to calculate the 3D-structure of the scene and the motion of the camera from measurements in the images taken from different view-points in the scene. This problem is called structure and motion, referring to the fact that both the structure of the scene and the motion of the camera are calculated from image measurements only. A number of different sub problems, arising from different knowledge of the intrinsic properties of the camera appear. Other important problems are to calculate the structure of the 3D-scene given the motion of the camera and to calculate the motion of the camera given the structure of the scene. These problems are somewhat simpler than the general structure and motion problem, are nevertheless important for navigation and obstacle avoidance. A related problem is to calibrate a camera, i.e. calculate the focal distance, the principal point etc., from images of a known scene. This calibration may be used to make a more precise reconstruction.

The problem of estimating structure and motion from image sequences are closely related to the field of closed range photogrammetric. In this field images are used to make precise measurements of three-dimensional objects from a number of images, often taken by calibrated cameras. The main difference is that in computer vision, the focus is not on accuracy, but instead on reliability and speed of calculation. In all these problems different kind of is called features can be used. The simplest feature is points that are easily detected in the sequence, such as corners or painted marks. Other types of features are lines and curves, which are easier to detect and track but more difficult to use in structure and motion algorithms. In order to use our computer vision routines these features must be detected in advance and the correspondence between different images must be established. The main emphasis of our computer vision toolbox is to use these detected features to solve for structure and motion.

4.6 Image Acquisition Toolbox

The Image Acquisition Toolbox's is a collection of functions that extend the capability of the MATLAB numeric computing environment. The toolbox supports a wide range of image acquisition operations including.

- Acquiring images through many types of image acquisition devices, from professional grade frame grabbers to USB-based Webcams.
- Viewing a preview of the live video stream.
- Triggering acquisitions (includes external hardware triggers).
- Configuring call back functions that execute when certain events occur.
- Bringing the image data into the MATLAB workspace.

Many of the toolbox functions are MATLAB M-files. You can extend the capabilities of the Image Acquisition Toolbox by writing your own M-files, or by using the toolbox in combination with other toolboxes, such as the Image Processing Toolbox and the



Data Acquisition Toolbox. The toolbox also includes a Simulink interface called the Image Acquisition Block set. This block set extends Simulink with a block that lets you bring live video data into a model.

4.7 Design of Proposed System:

Within the project, the respiratory images of the sufferer are considered be entered. Pictures that are pre-processed segmentation and subsequently categorized when it comes to the functionality required. Appear To Be Extracted has been done for the pictures by GLCM distinguishing characteristics. Fuzzy C clustering algorithm algorithm that is applied in sequence to define portion of tumor. In this instance the necessary threshold for segmentation is adapting depending on the partitioned region and the position. Ultimately, categorization applied by using a neural network then the results in the picture to be compared the data set pictures and it will be shown if a covid have been affected or not affected.

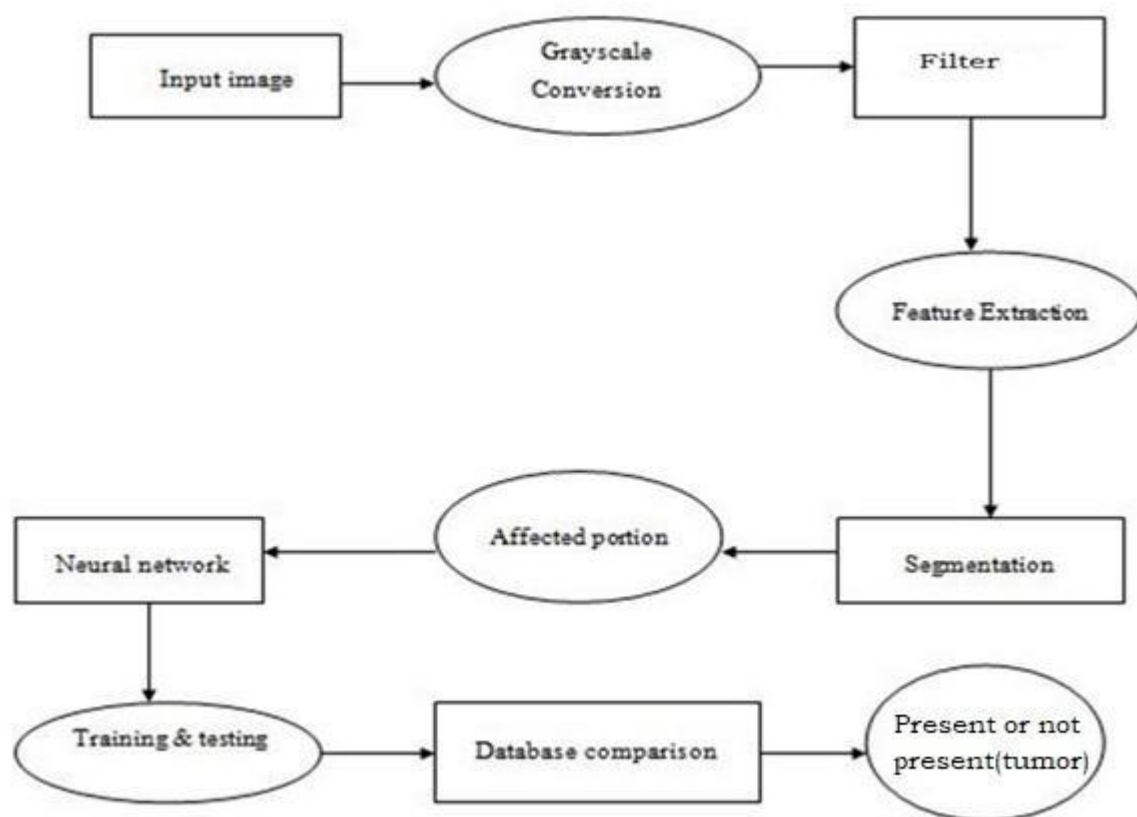


Fig 4.1 Block Diagram

4.8 UML DIAGRAM:

UML is synonymous with the Unified Modelling Language. UML is the standard general-purpose simulation language available in the sphere of object-oriented software technology. Standard specification is administered, and was established by, an Object That The objective is for UML to develop into a shared language that can be used to create models of object-oriented software applications. In its present form UML is comprised of two major components: a meta-model and notation system. The UML represents a set of best manufacturing techniques that have been proven successful through the modeling of large and sophisticated methods. The UML is an extremely important part of developing the pieces software as well as your software development process. The UML primarily used graphical representations articulate the design of software development projects.

UML can be applied to diverse application domains (e.g., banking, finance, internet, aerospace, healthcare, etc.) It can be used with all major object and component software development methods and for various implementation platforms (e.g., J2EE, .NET).

UML is not a complete and it is not completely visual. Given some UML diagram, we can't be sure to understand depicted part or behavior of the system from the diagram alone. Some information could be intentionally omitted from the diagram, some information represented on the diagram could have different interpretations, and some concepts of UML have no graphical notation at all, so there is no way to depict those on diagrams. For example, semantics of multiplicity of actors and multiplicity of use cases on use case diagrams is not defined precisely in the UML specification and could mean either concurrent or successive usage of use cases.

UML is a standard modeling language, not a software development process. UML1.4.2 Specification explained that process:



- Provides guidance as to the order of a team's activities.
- Specifies what artifacts should be developed.
- Directs the tasks of individual developers and the team as whole.
- Offers criteria for monitoring and measuring a project's products and activities.

The Primary goals in the design of the UML are as follows:

1. Provide users a ready-to-use, expressive visual modeling Language so that they can develop and exchange meaningful models.
2. Provide extendibility and specialization mechanisms to extend the core concepts.

UML is intentionally process independent and could be applied in the context of different processes. Still, it is most suitable for use case driven, iterative and incremental development processes. An example of such process is Rational Unified Process (RUP).

4.9 Activity Diagram:

The activity diagram is the graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. The activity diagrams can be used to describe the business and operational step-by- step workflows of components in a system. Activity diagram consist of Initial node, activity final node and activities in between.

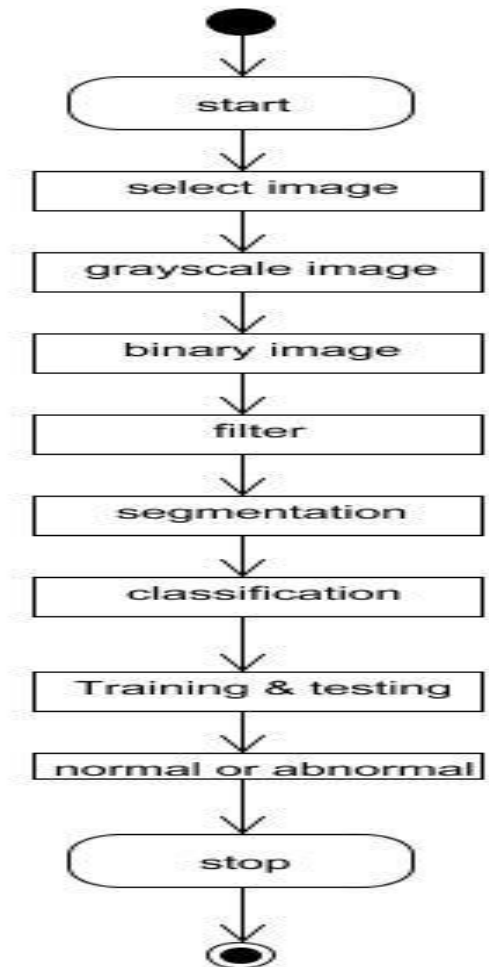


Fig 4.2 Activity Diagram

4.10 Sequence Diagram:

Sequence diagram model the flow of logic within the system in a visual manner, enabling both the document and validate the logic, and commonly used for both analysis and design purpose. Sequence diagram are the most popular UML artifact for dynamic modeling, which focuses on identifying the behavior within the system.

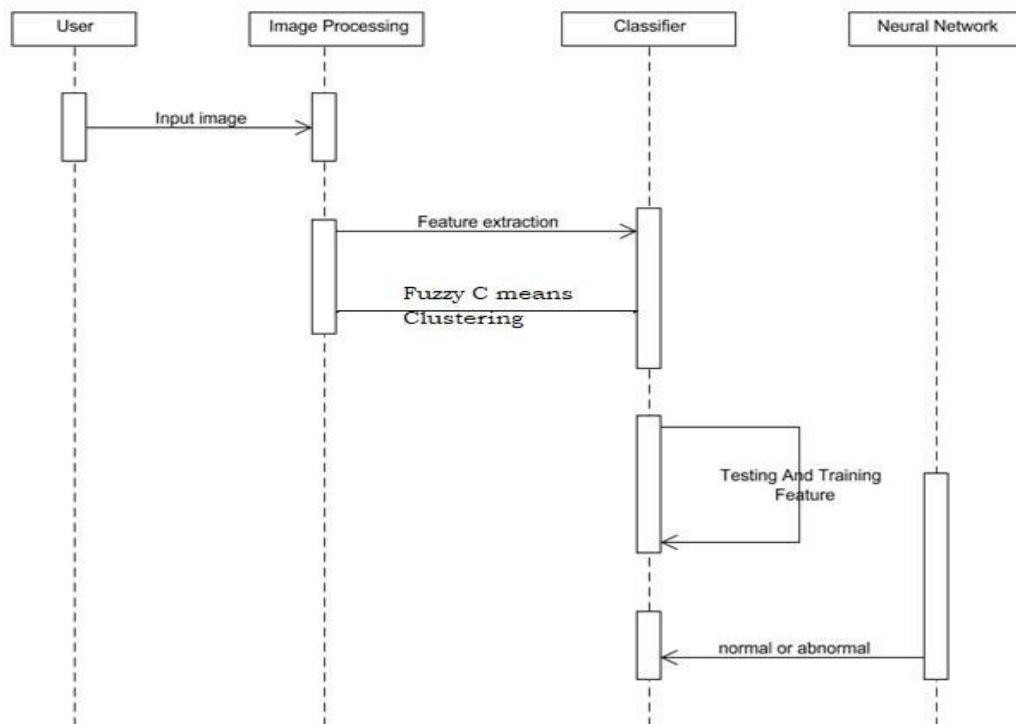


Fig 4.3 Sequence Diagram

4.11 Use case diagram:

As the most known diagram type of the behavioral UML diagrams, Use case diagrams give a graphic overview of the actors involved in a system, different functions needed by those actors and how these different functions interact. It's a great starting point for any project discussion because user can easily identify the main factors involved and the main processes of the system. The user can create use case diagrams using our tool and/or get started instantly using our use case templates.

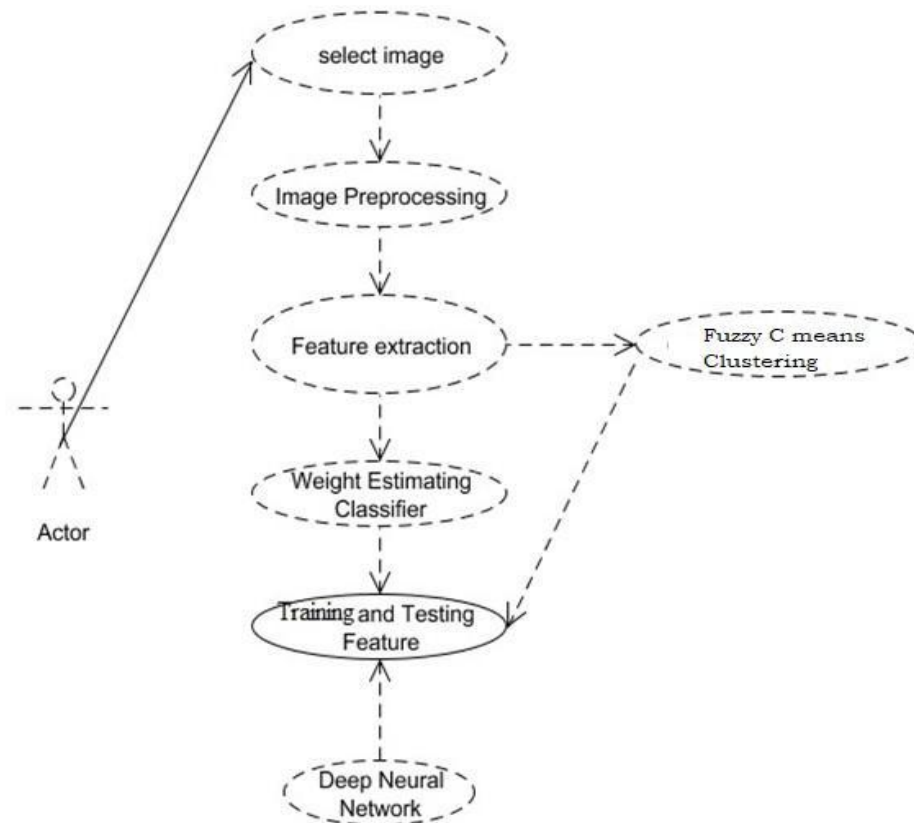


Fig 4.4 Use case diagram

4.12 Dataflow Diagram:

A data flow diagram (DFD) is a graphical representation of the “flow” of data through an information system, modelling its *process* aspects. A DFD is often used as a preliminary step to create an overview of the system without going into great detail, which can later be elaborated. DFDs can also be used for the visualization of data processing (structured design).

A DFD shows what kind of the information will be input to and output from the system, how the data will advance through the system, and where the data will be stored. It does not show information about the timing of process or information about whether processes will operate in a sequence or in parallel unlike a flowchart which also shows this information.

Data flow diagrams are also known as bubble charts. DFD is a designing tool used in the top-down approach to Systems Design. This context-level DFD is next "exploded", to produce a Level 1 DFD that shows some of the detail of the system being modeled. The Level 1 DFD shows how the system is divided into sub-systems (processes), each of which deals with one or more of the data flows to or from an external agent, and which together provide all of the functionality of the system as a whole. It also identifies internal data stores that must be present in order for the system to do its job, and shows the flow of data between the various parts of the system.

Data flow diagrams are one of the three essential perspectives of the structured-systems analysis and design method SSADM. The sponsor of a project and the end users will need to be briefed and consulted throughout all stages of a system's evolution. With a data flow diagram, users are able to visualize how the system will operate, what the system will accomplish, and how the system will be implemented.

Step 1:

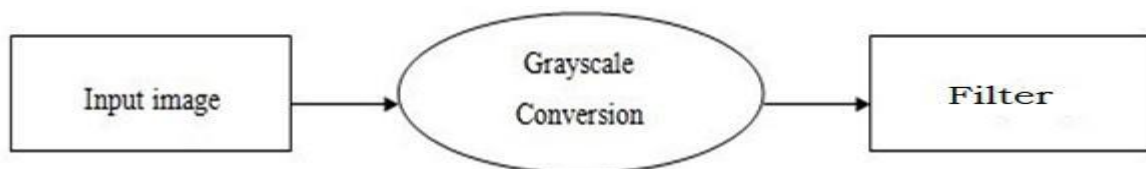


Fig 4.5 Data flow diagram

If the input images are color images means we are convert to gray scale from that color images. A median filter is more effective than convolution when the goal is to simultaneously reduce noise and preserve edges.

Step 2:

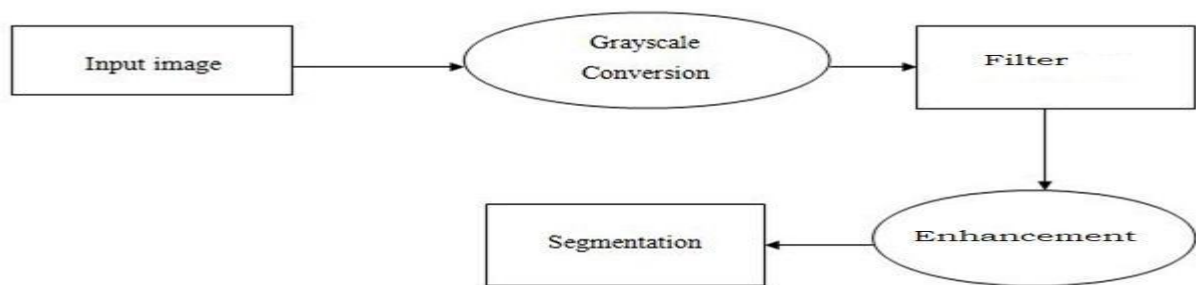


Fig 4.6 Data flow diagram

If the input images are color images means we are convert to gray scale from that color images. Image enhancement refers to sharpening of image features such as edges, boundaries, or contrast to make a graphic display more useful for display and analysis.

Step 3:

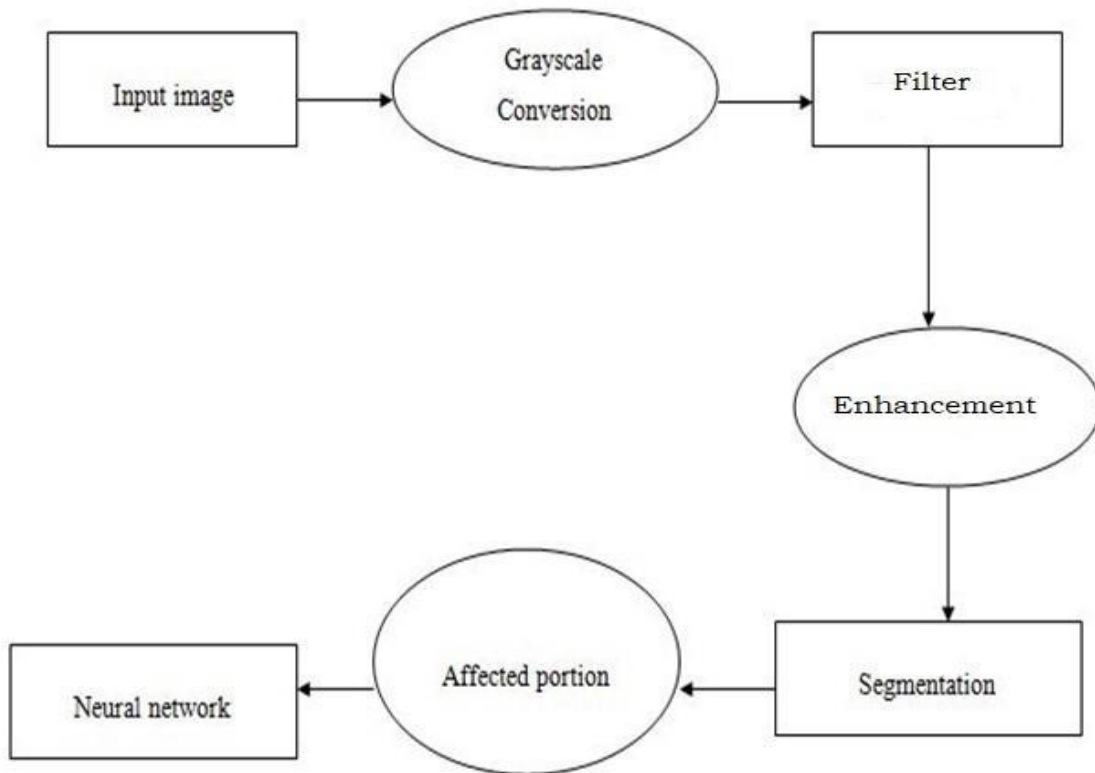


Fig 4.7 Data flow diagram

The RGB panel is used to view the red, green and blue components of the image separately. It has already been mentioned that an RGB image is overlap of three two-dimensional matrix. If the input images are color images means we are convert to gray scale from that color images. So the RGB image is converted into Grayscale image. Image enhancement refers to sharpening of image features such as edges, boundaries, or contrast to make a graphic display more useful for display and analysis. Here we used the Fuzzy C means clustering algorithm is used in order to segment the portion of defected areas. Image segmentation is typically used to locate objects and boundaries in images.

Step 4

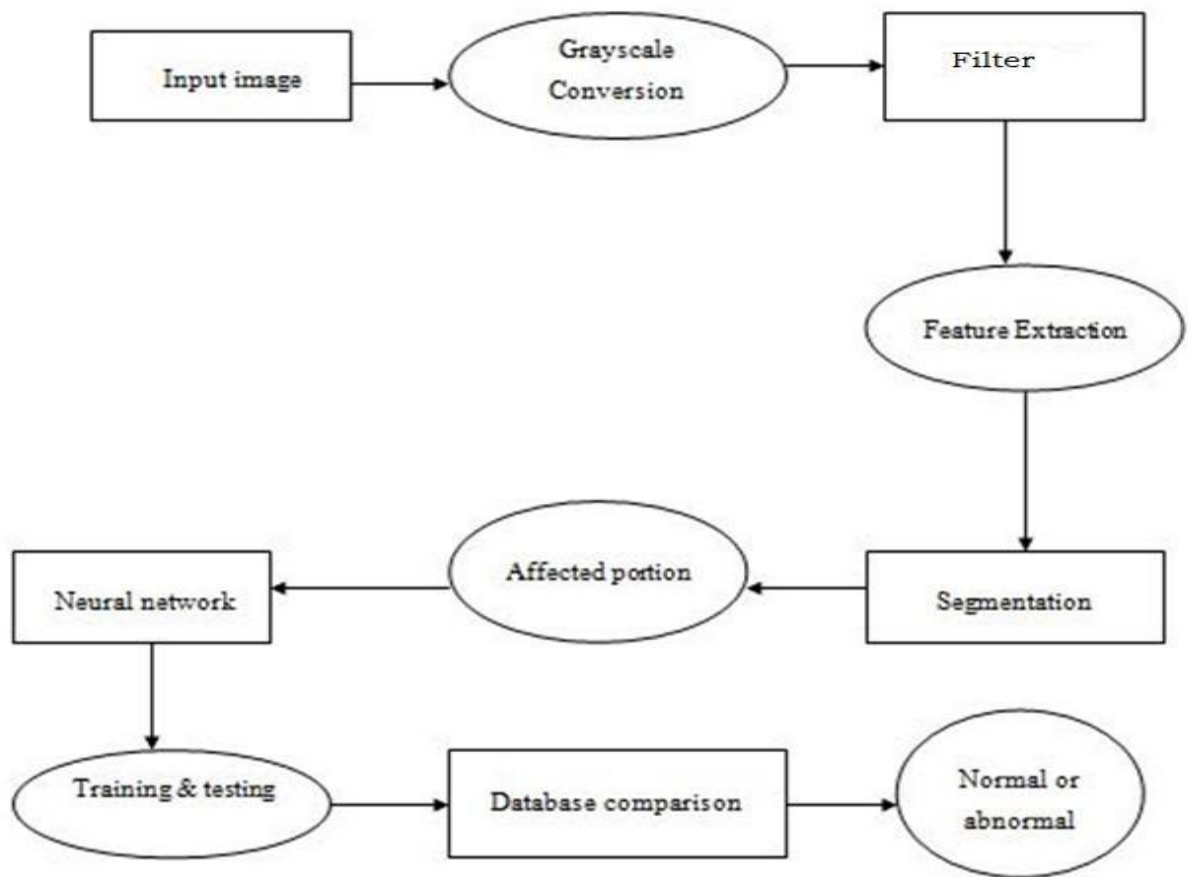


Fig 4.8 Data flow diagram

If the input images are color images means we are convert to gray scale from that color images. A median filter is more effective than convolution when the goal is to simultaneously reduce noise and preserve edges. The back propagation neural network concept for training the image and testing the image with the help of weight estimating classifier. The result image will compared with the dataset images and it will display on the LCD whether it is present or not present.

CHAPTER 5

Result & Discussion



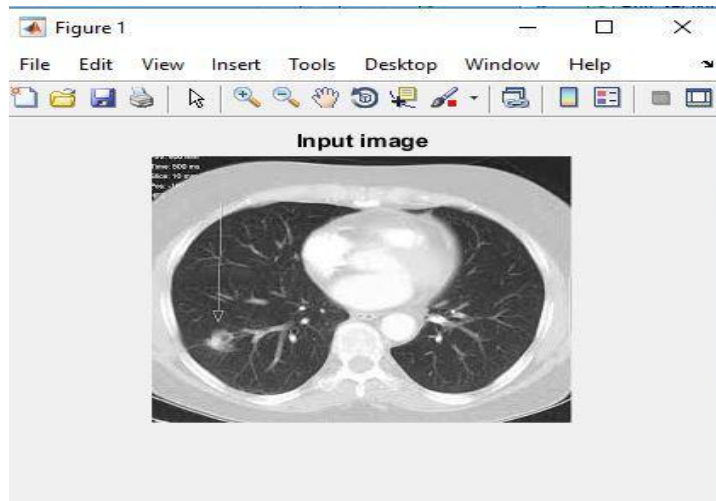


Fig 5.1 Input image

The above Fig 5.1 is the input image of the chest CT. Using this input image we check whether the given chest CT is Positive or Negative.

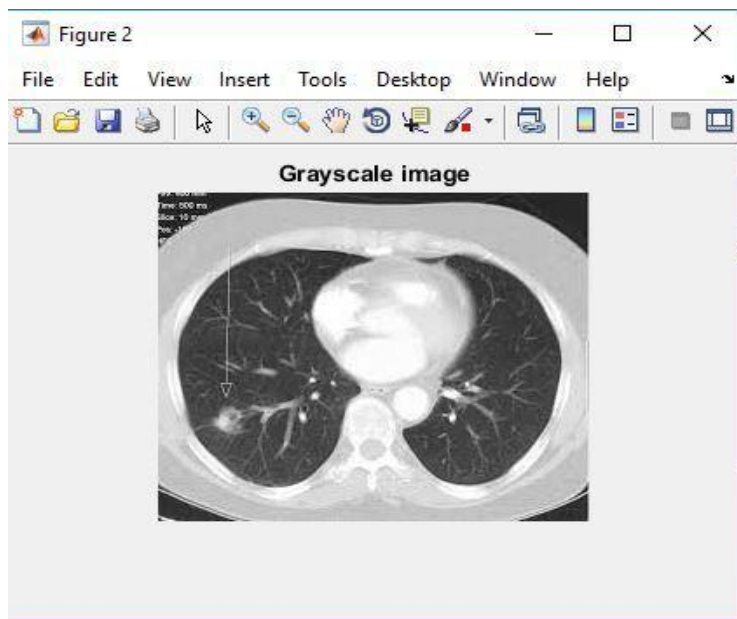


Fig 5.2 Grayscale image

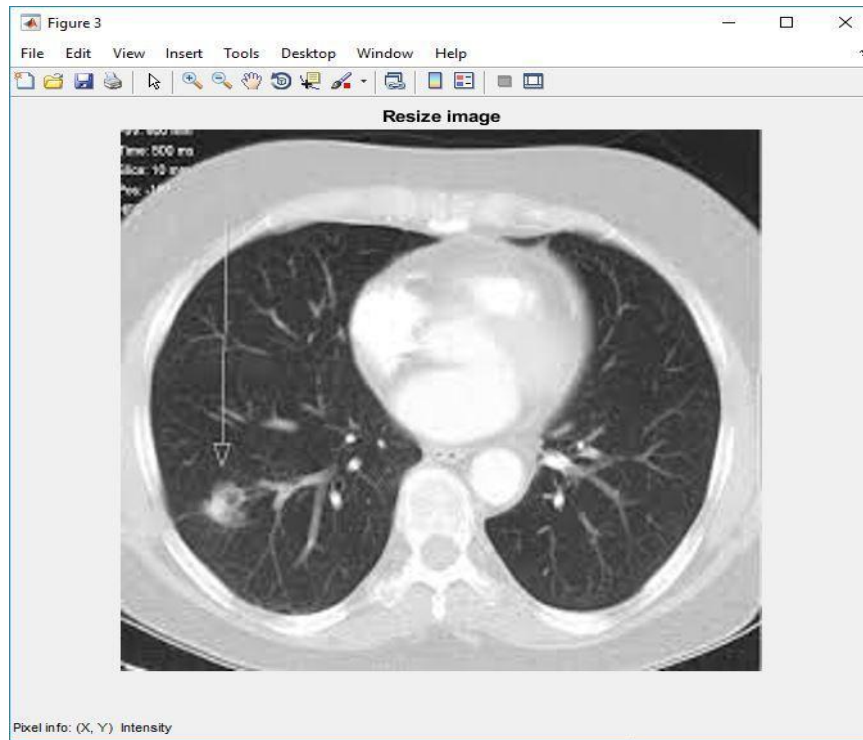


Fig 5.3 Resize image

After that the input image was converted into RGB to Gray scale image if it is necessary. As shown in fig 5.2 the input image is converted into gray scale image. And also the input image will resize into its necessary size. If the image size is less than or greater than fixed size then it will convert into its required size as shown in fig 5.3.

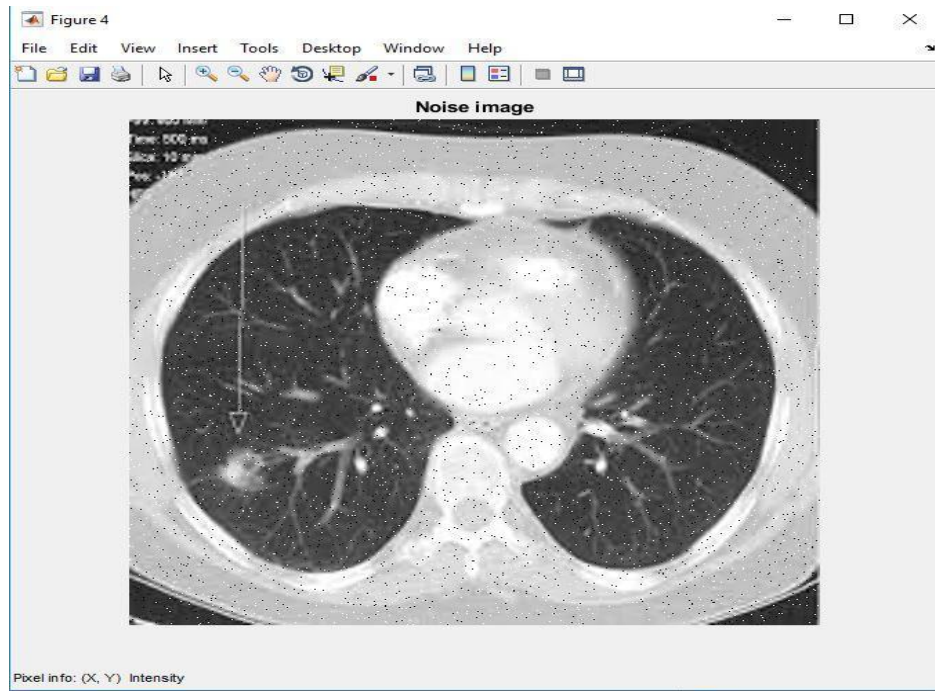


Fig 5.4 Noise image

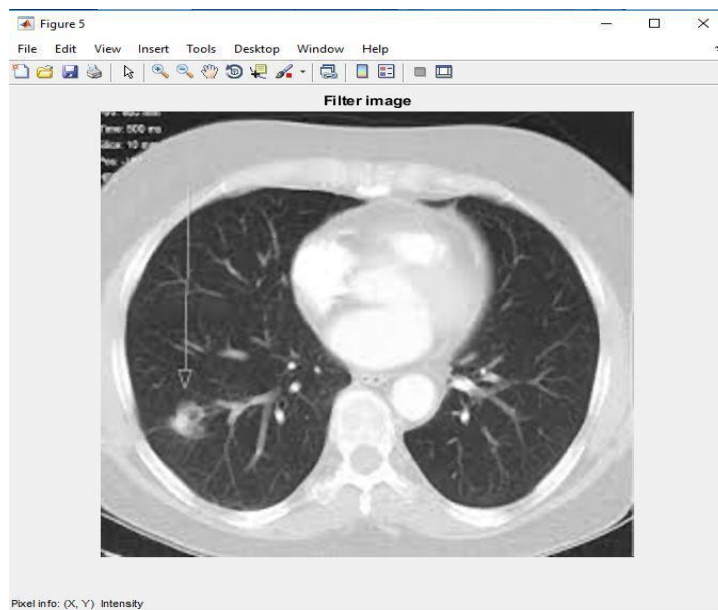


Fig 5.5 Filter image

As shown in fig 5.4 the noise created during this process is removed using salt and pepper process. Then the input image will converted into filter image as shown in fig 5.5.

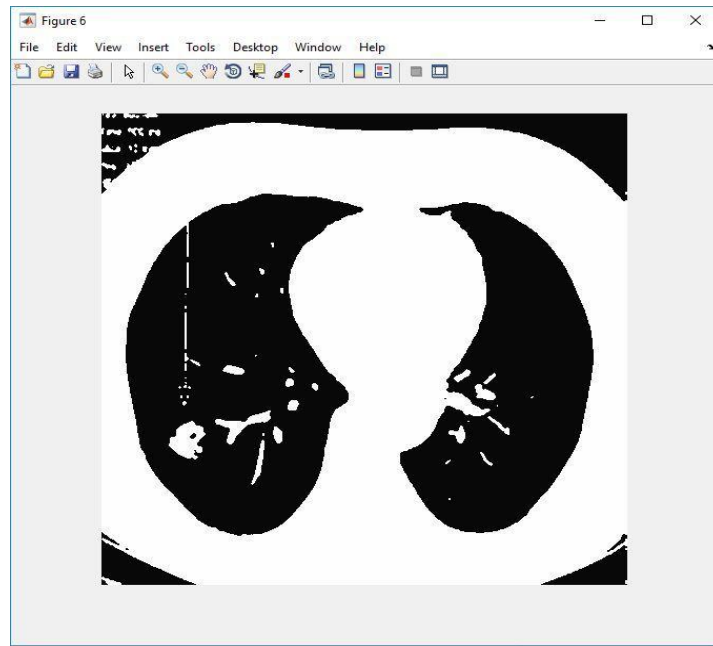


Fig 5.6 Recolor image

From the fig 5.6 and 5.7 the area which is in black was transferred into white and the white area will transfer into black to detect the effected area clearly.

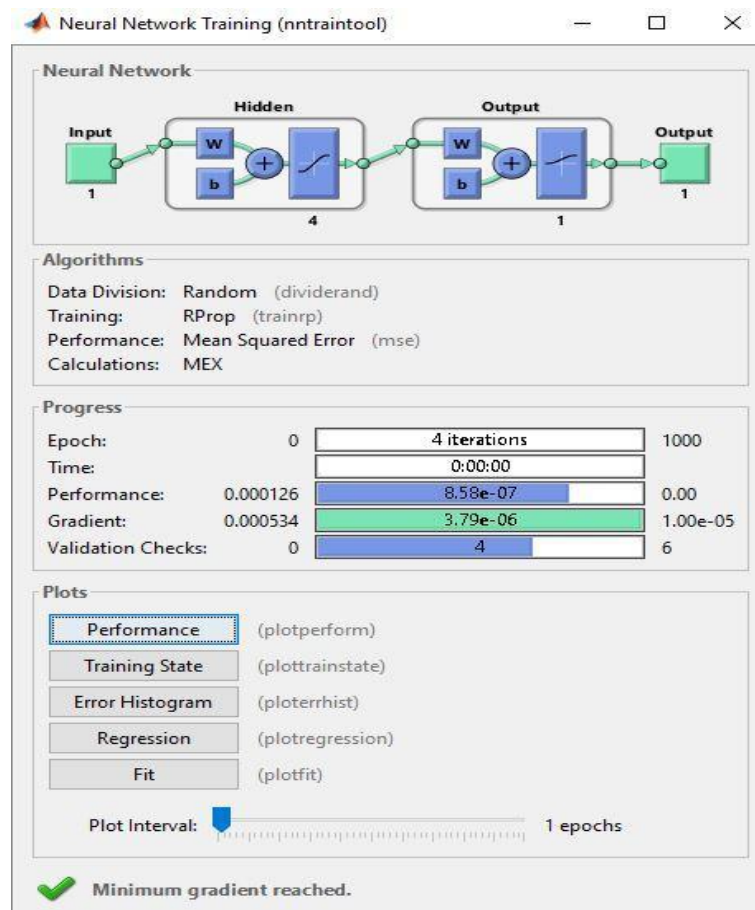


Fig 5.8 Neural network



Fig 5.9 Output

At last after comparing the image with the dataset images to check whether the given Chest CT is effected or not. As shown in fig 5.9 the output will display as COVID or NON COVID.

CHAPTER 6

6.1 CONCLUSION

In this paper, we study a general inference framework for extracting lung tumors from medical image sequences. A collaborative formulation of tumor segmentation is discussed by jointly integrating region and boundary information. Here we used the Fuzzy c means clustering algorithm is used in order to segment the portion of defected areas. We used to establish the neural network concept for training the image and testing the image with the help of weight estimating classifier. The result image will compared with the dataset images and it will display whether it is normal or abnormal.

6.2 Future Enhancement:

For future work, we can also know the size of the tumor in the lungs such as we can find the staging of the cancer and also find the eccentricity, intensity, Average intensity, Perimeter of the tum

