A PROJECT REPORT ON

Lung Cancer Detection Using Deep Learning

SUBMITTED TO THE SAVITRIBAI PHULE PUNE UNIVERSITY, PUNE
IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE

BACHELOR OF ENGINEERING (COMPUTER ENGINEERING)

SUBMITTED BY

SWAPNIL BANDAL	**78029438L**
AKASH DESHMUKH	**72029404F**
SHREYAS PADMAWAR	**72029708H**

Under The Guidance of

Dr. Smita Chaudhari

'येथे बहुतांचे हित'

DEPARTMENT OF COMPUTER ENGINEERING
Marathwada Mitra Mandal's College of Engineering
Karvenagar, Pune-411052
Savitribai Phule Pune University

2022-23

'येथे बहुतांचे हित'

CERTIFICATE

This is to certify that the project report entitled

"Lung Cancer Detection Using Deep Learning"

Submitted by

SWAPNIL BANDAL	72029438L
AKASH DESHMUKH	72029404F
SHREYAS PADMAWAR	72029708H

are bonafide students of this institute and the work has been carried out by them under the supervision of **Dr. Smita Chaudhari** ma'am and it is approved for the partial fulfillment of the requirement of Savitribai Phule Pune University, for the award of the degree of **Bachelor of Engineering**(Computer Engineering).

Dr. Smita Chaudhari
Internal Guide,
Dept. of Computer Engg.

Dr. K. S. Thakre
Head,
Dept. of Computer Engg.

Dr.V. N. Gohokar
I/C Principal,
Marathwada Mitra Mandal's College of Engineering, Karvenagar, Pune – 411052

Date: 22-11-2022
Place: Pune

Acknowledgment

We take this to express our deep sense of gratitude towards our esteemed guide Dr. Smita Chaudhari ma'am for giving us this splendid opportunity to select and present this project and also providing facilities for successful completion.

I thank Dr. Kalpana Thakre ma'am, Head, Department of Computer Engineering, for opening the doors of the department towards the realization of the project, all the staff members, for their indispensable support, priceless suggestion and for most valuable time lent as and when required. With respect and gratitude, we would like to thank all the people, who have helped us directly or indirectly.

SWAPNIL BANDAL
AAKASH DESHMUKH
SHREYAS PADMAWAR

Abstract

Lung cancer is consistently identified as one of the deadliest diseases in human history. It is one of the most common types of cancer and one of the leading causes of death. According to the World Health Organization (WHO), more than 7.6 million people die from lung cancer worldwide each year.

Cancer is a disease caused by uncontrolled division of eccentric cells found in all parts of the human body. It is one of the most deadliest disease. This work aims to present an effective and efficient mode of computer-assisted lung cancer detection method. In this, we used a series of computed tomography scan lung images as input, retrieved from a lung image archive, and applied image processing techniques such as feature extraction.We use a well-known Convolutional Neural Network(CNN) algorithm for lung cancer classification.

Technical Keywords

(1) Pulmonary nodules

(2) lung neoplasms

(3) feature extraction

(4) Lung Cancer Nodule

(5) Convolutional Neural Network

(6) Pattern matching

(7) Image processing

Contents

List of Figures

Chapter 1

Introduction

1.1 Overview

According to projections, the number of cancer patients will continue to rise, with approximately 17 million people dying from lung cancer by 2030. Lung cancer in nonsmokers can be caused by radon, secondhand smoke, air pollution, or other causes. Occupational exposure to asbestos, diesel exhaust, or other pollutants can also lead to lung cancer in nonsmokers. Typical symptoms include coughing (often bloody), chest discomfort, wheezing, and weight loss. However, these symptoms usually do not appear until the malignancy is advanced. Cancers are caused by a variety of factors, from behavioral factors such as high BMI, tobacco and alcohol. Pain, fatigue, nausea, chronic cough, difficulty breathing, weight loss, muscle pain, bleeding, and bruising are common symptoms in cancer patients.

CT scans, which look for nodules in the lungs, are often used to detect lung cancer. By examining the nodules, doctors can determine if this scan is cancerous. physicians can evaluate nodules greater than 7 mm in diameter. Doctors also often tell patients to wait until nodules develop and see what happens. If not developed, the knot is harmless. It is more likely that the knot will not be detected. Therefore, cancer diagnosis is difficult without comprehensive diagnostic methods such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), ultrasound, or biopsy. is. In most cases, victims show little to no symptoms in the early stages and by the time symptoms appear it is usually too late. The purpose of this study is to look for alternative methods to detect and detect lung cancer at an early stage. As a result, 4,444 lung cancers can be treated before they progress to the point of being untreatable. The focus of this study is solely on the detection and prediction of lung cancer and its stages.

1.2 Motivation

1) The mortality rate is even higher than breast cancer and prostate cancer combination.

2) Lung cancer is one of the key causes of death amongst humans globally, with a mortality rate of approximately five million cases annually.

3) However, early detection and diagnosis can improve the survival rate.

1.3 Problem Definition

To develop an application that helps in Lung Cancer Detection by using convolutional neural network (CNN) and also predict stages of cancer.

1.4 Project Scope and Limitations

The project is not restricted by any one particular field and can be used by any person with the use of web-app. Hence the scope of this project is very broad and long term.

This process reduces the time complexity and increases the diagnosis confidence. The collected data contain noise, the noises are removed. And then segmentation of the lung images and after that the image is separated. The output image is trained by using the CNN model and the diagnosis is made from the output. We are trying to develop the automated medical image processing tools in which it detects the cancer cells in advance.

Chapter 2

Literature Survey

2.1 Literature Review

In paper, Muthazhagan B, Ravi T, Rajinigirinath D[1] proposed research paper on A Machine Learning Approach to Diagnosing Lung and Colon Cancer
This paper proposed an Enhanced Computer-assisted Lung Cancer Detection Method Using Content-Based Image Retrieval and Data Mining Techniques methodology used is Support Vector Machine image classification algorithm. limitations are The malignancy is classified as 'Normal' and 'Abnormal', not as Stages 1-4.

Masud M, Sikder N, et al. [2] Machine Learning Approach to Diagnosing Lung cancer Using a Deep Learning
proposed research paper on A Machine Learning Approach to Diagnosing Lung and Colon Cancer Using a Deep Learning Based Classification Framework in which methodology used is Digital Image Processing techniques with CNN. The limitations are Dataset uses microscopic cells images rather than CT/MRI scans.

Tripathi P, Tyagi S, Nath M [3] proposed research paper on Comparative Analysis of Segmentation Techniques for Lung Cancer
This paper proposed research paper on Comparative Analysis of Segmentation Techniques for Lung Cancer Detection in which methodology used is comparative analysis- image segmentation techniques. The limitations are marker-controlled watershed segmentation provides less accurate results.

Ibrahim M. Nasser et al. [4] an Artificial Neural Network (ANN) to identify the presence or absence of lung cancer

In this study in 2019, Ibrahim M. Nasser et al. created an Artificial Neural Network (ANN) to identify the presence or absence of lung cancer in the human body. Symptoms such as yellow fingers, anxiety, chronic disease, and others were utilized to identify lung cancer. They were employed as input variables to their ANN, along with additional information of the patient. The ANN model was shown to be 93.67% accurate in detecting the absence or presence of lung cancer.

Siddharth Bhatia [5] proposed use of Deep residual learning to detect lung cancer.

Deep residual learning has been used by Siddharth Bhatia to detect lung cancer. They provide a set of pre-processing algorithms for obtaining cancer-vulnerable lung characteristics from images using UNet and ResNet models. They compare the efficiency of classifiers such as Random forest and XGBoost in predicting carcinogenic CT images. The authors achieve the highest accuracy of 84% when they mix the two classifiers.

2.2 Literature Survey

Sr.No.	Title of paper	Year	Authors	Findings	Technologies
1.	A lung cancer detection method using image retrieval and data mining techniques	2020	Muthazhagan B, Ravi T, Rajinigirinath D	This paper proposed an Lung Cancer Detection Method Using Content-Based Image Retrieval and Data Mining Techniques methodology used is SVM algorithm.	Support Vector Machine

Sr.No.	Title of paper	Year	Authors	Findings	Technologies
2.	A machine learning approach to diagnosing lung cancer using a deep learning.	2021	Masud M, Sikder N, et al	proposed research paper on A Machine Learning Approach to Diagnosing Lung cancer Using a Deep Learning Based Classification Framework	Digital Image Processing techniques with CNN.
3.	A comparative analysis of segmentation techniques for lung cancer detection.	2019	Tripathi P, Tyagi S, Nath M	This paper proposed research paper on Comparative Analysis of Segmentation Techniques for Lung Cancer Detection	Image Segmentation Techniques
4.	an Artificial Neural Network (ANN) to identify lung cancer.	2019	Yibing Xiong Takashi, Yamada Takao, Terano	This paper proposed an Artificial Neural Network (ANN) to identify the presence or absence of lung cancer in the human body. Symptoms such as yellow fingers, anxiety, chronic disease, and others were utilized to identify lung cancer.	Artificial Neural Network (ANN)
5.	Use of Deep residual learning to detect lung cancer	2019	Siddharth Bhatia	They provide a set of preprocessing algorithms for obtaining cancer-vulnerable lung characteristics from images using UNet and ResNet models.	Random forest and XGBoost

Table 2.1: Literature survey

Chapter 3

System Design

3.1 System Architecture

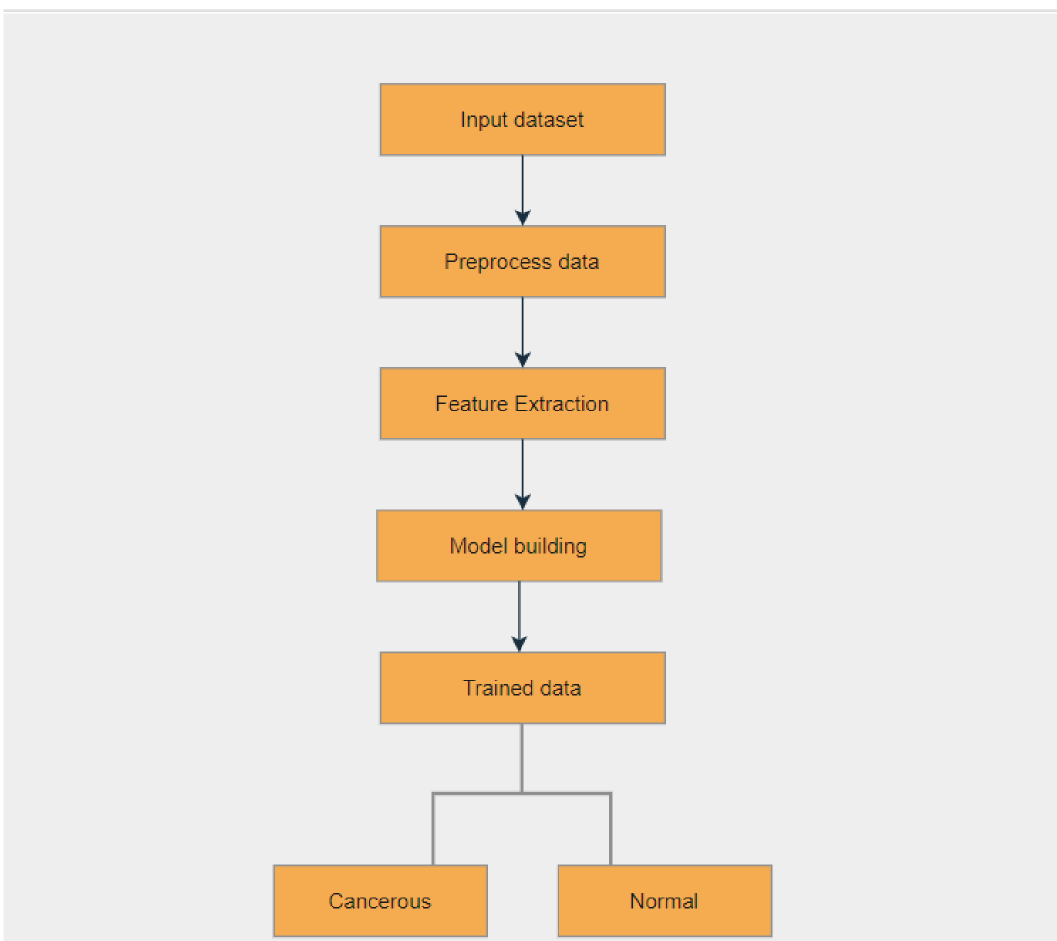

Figure 3.1: System Architecture

3.2 Class Diagram

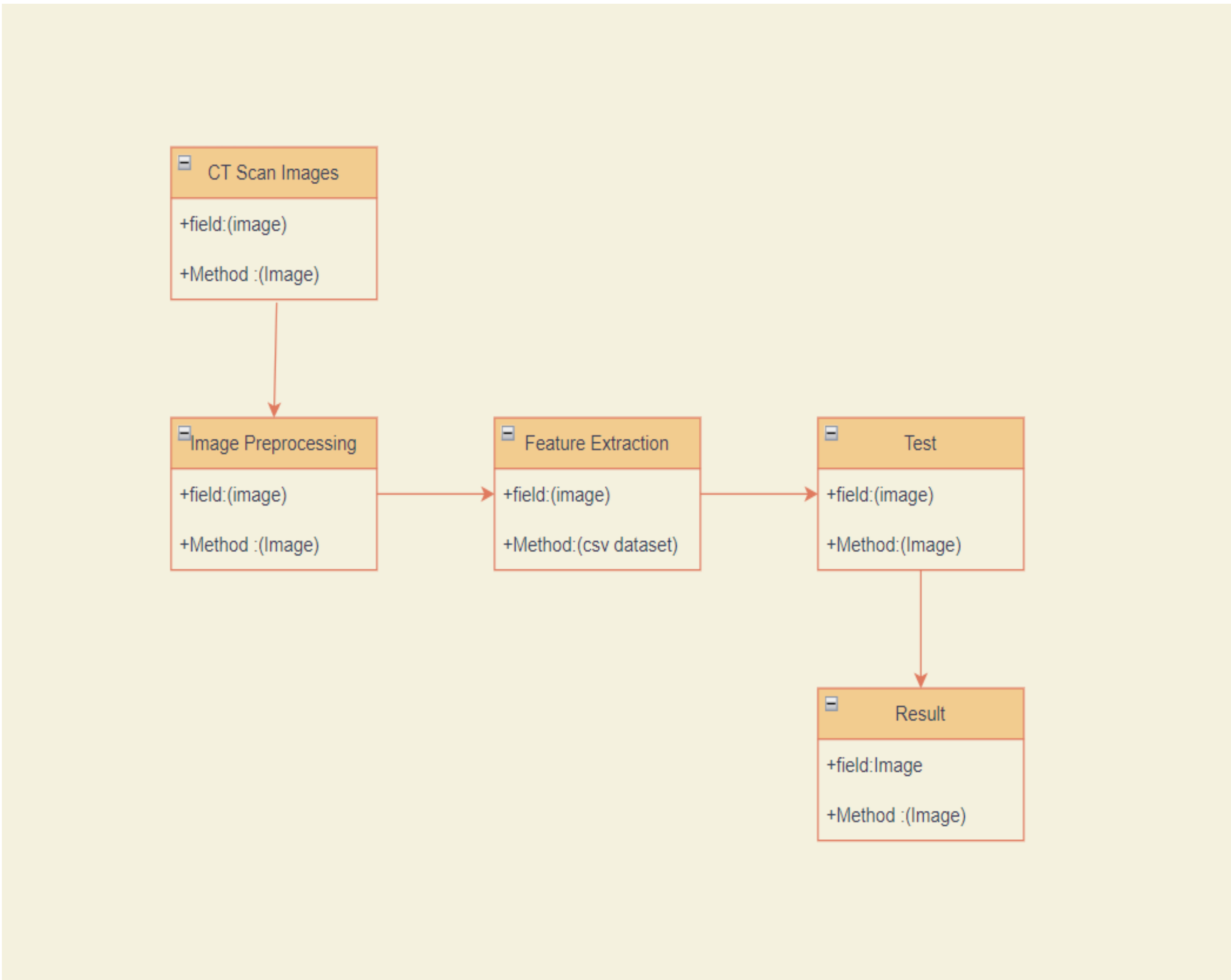

Figure 3.2: Class Diagram

3.3 Use Case Diagram

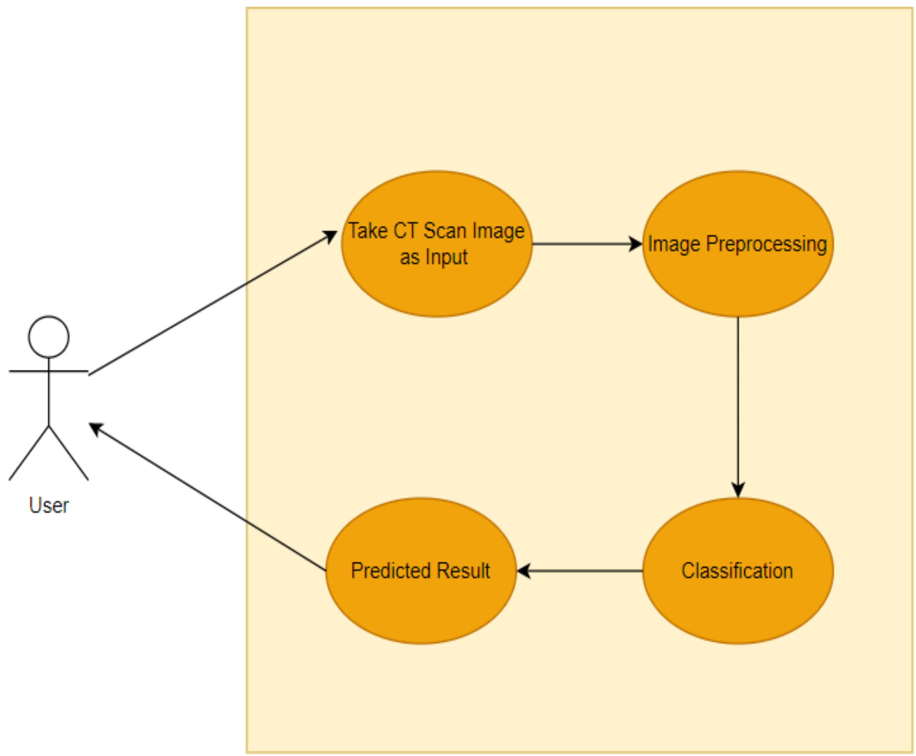

Figure 3.3: Use Case Diagram

Fig. 3.3 shown above is a use case of how the proposed system shall be used by Radiologist or any other user. User comes and Loads some CT image in which nodule detection is required. System do some processing and shows results back to user. In processing, a chain of sub processes starts with preprocessing of CT image, followed by 3D patch extraction which are needed for detection. Then results are computed, stored returned to the request initiator.

3.4 Sequence Diagram

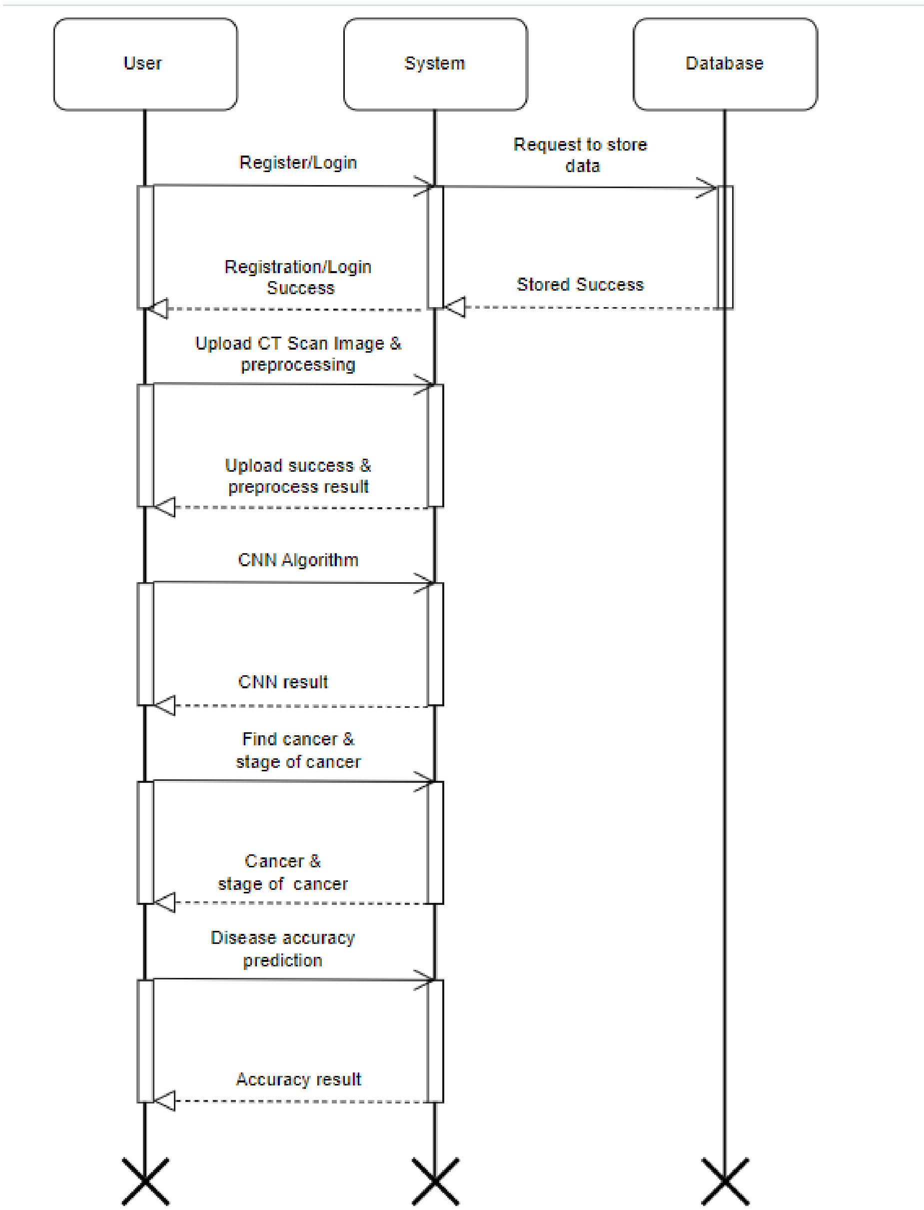

Figure 3.4: Sequence Diagram

Fig. 3.4 is a sequence diagram which is demonstrating the methods/functions along with the parameters with their calling sequence. First method named as “upload CT scan images” is the request initiator method which takes only a CT scan as a parameter and forwards it for further processing. The method named as “CNN Result” gives back the results of detected nodules to initiator for the requested CT scan

3.5 Activity Diagram

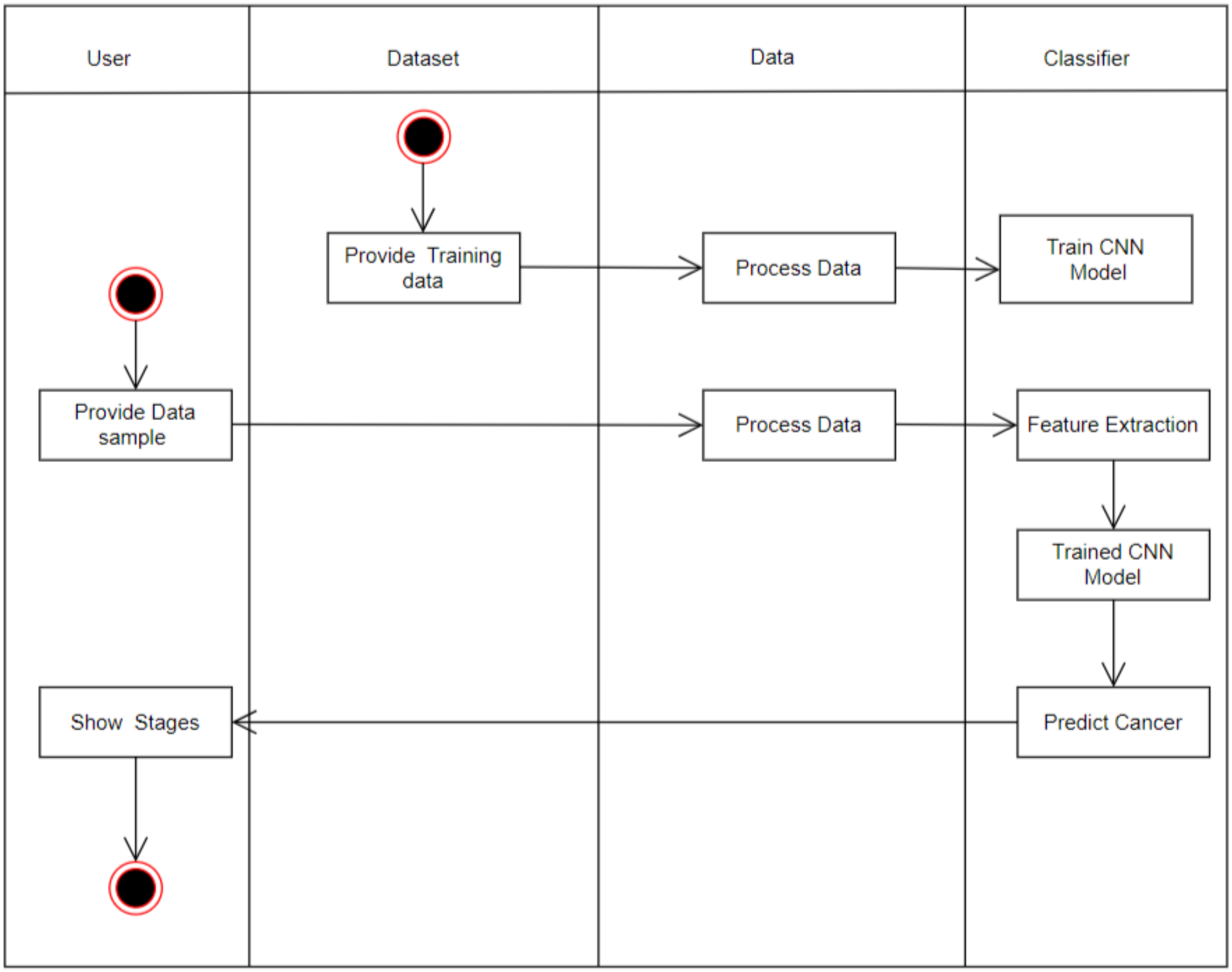

Figure 3.5: Activity Diagram

3.6 Deployment Diagram

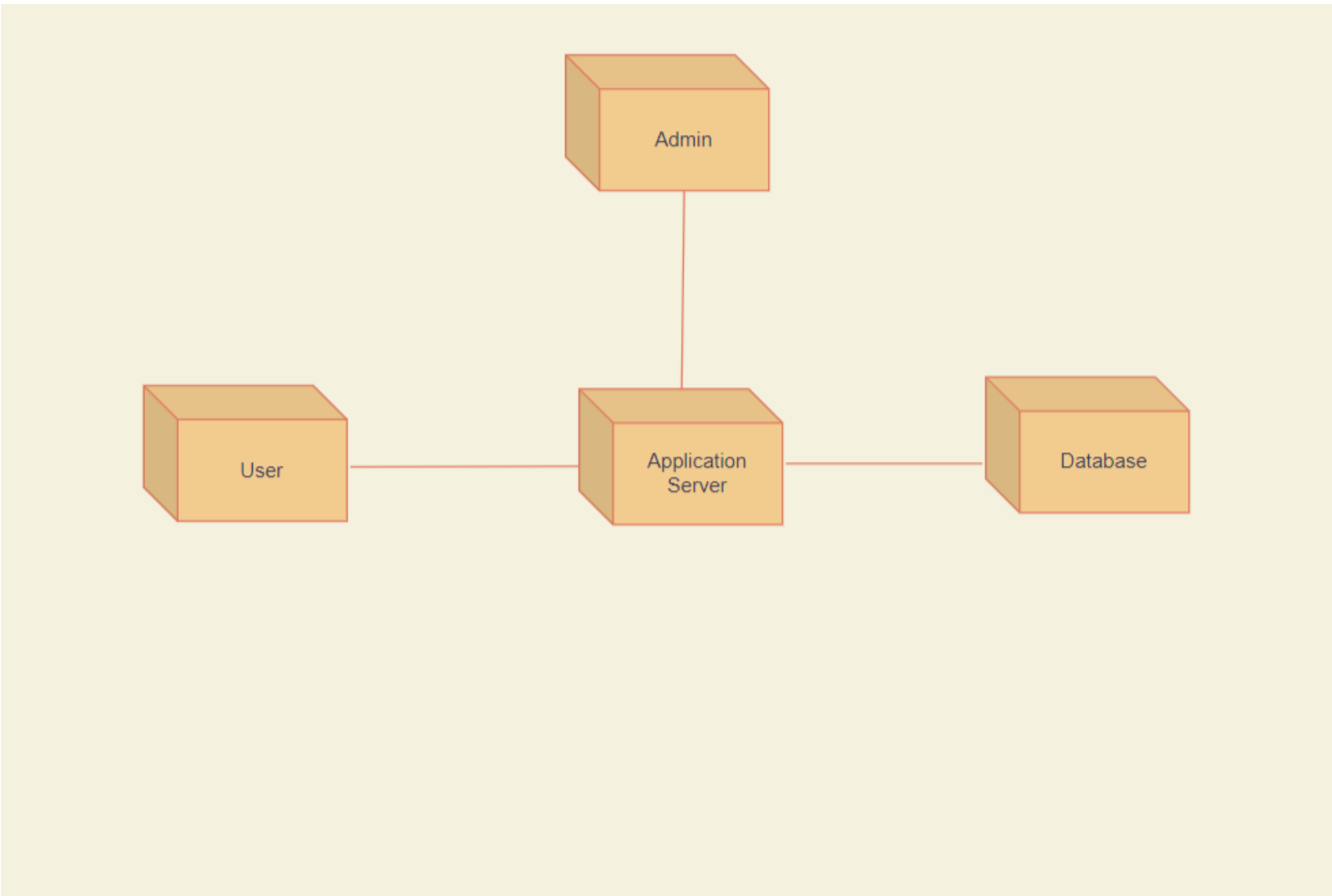

Figure 3.6: Deployment Diagram

Chapter 4

Project Plan

4.1 Project Estimate

Stake holders: 3 Team Members, 1 Internal Guide, 1 External Guide

4.1.1 Software Requirments

- IDE: Visual-studio-code
- Language: Python 3.8
- Framework: Django

4.1.2 Hardware Requirments

- Greater than 500 GB
- Greater than 4 GB
- I3 and Above

4.1.3 Dataset

A dataset containing CT scan images is required to train the model.

4.2 Risk Management

Various risks identified for our project are :

4.2.1 Technical Risks or Technical Feasibility

1. Image might be unclear or distorted.
2. Neural network might over fit the data.
3. Neural network might under fit the data.

Back up plan to Avoid:

1. Image processing.
2. Use of regularizers.
3. Get more amount of data.

4.2.2 Cost Risks or Cost Feasibility:

We are using all open source technologies for our project development like python, CNN algorithm and Django framework for better interface. So there is no cost risk.

4.3 Project Schedule

4.3.1 Project Task Set

1. Research:-
   Searching for IEEE and other publication papers related to Algorithmic Trading.
2. Software Requirement Specification:-
   For understanding customer requirements and analyzing objectives to be achieved.
3. Synopsis Formation
4. Designing of system architecture, gathering data for designing of models.
5. Implementation of system.

6. Testing of system by using different test cases.

7. Report.

Schedule of the project:
Start Date: August 2022
End Date: May 2023
Duration: Approx 10 Months

4.3.2 Task Network

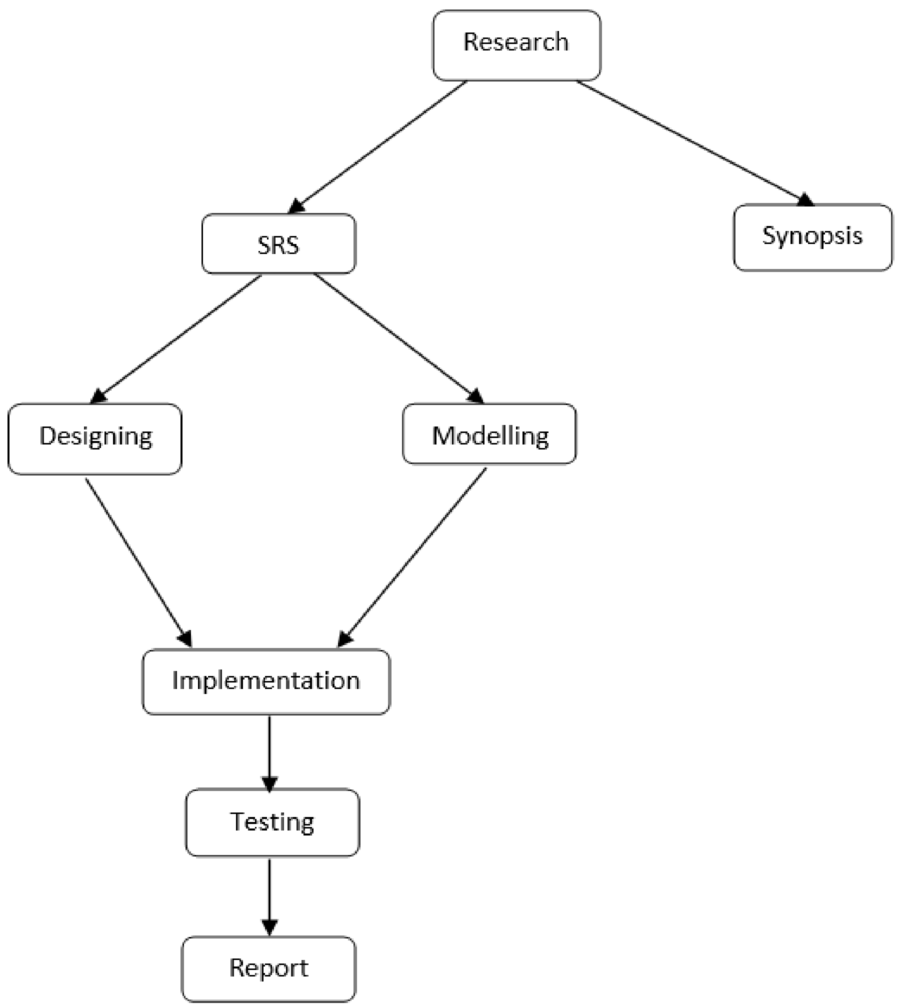

Figure 4.1: Task Network

4.3.3 Timeline Chart

Activity	I week	II week	III week	IV week	V Week	VI week	VII week	VIII week	IX week	X week	XI week	XII week
	Aug 4	Aug 11	Aug 18	Sept 2	Sept 14	Sept 21	Oct 06	Oct 15	Oct 29	Nov 07	Nov 15	Nov 23
Initiate the project												
Communication												
Literature survey												
Define scope												
Develop SRS												
Plan the project												
Design mathematical model												

Figure 4.2: Time Line

4.4 Team Organization

4.4.1 Team structure

1. Swapnil Bandal : Front-end, Data analysis and Documentation.
2. Aakash Deshmukh: Django Back-end, Data analysis and Documentation.
3. Shreyas Padmawar: Front-end, Data analysis and Documentation.

4.4.2 Management reporting and Communication

1. Weekly reporting to internal guide.
2. Continuous updating and reviewing of SRS and Development Processes.
3. Monthly reporting to external guide.
4. Expert guidance every month.

Chapter 5

Project Implementation

5.1 Overview of Project Modules

There are 4 major modules in this project namely,

1. Django Application
   Django is a Python-based free and open-source web framework that follows the model-template-views architectural pattern. A Django application is basically a Python package that is specifically intended for use in a Django projects. An application may use common Django conventions, such as having models ,tests ,urls ,and views submodules. With Django, you can take Web applications from concept to launch in a matter of hours. Django takes care of much of the hassle of web development, so you can focus on writing your app without needing to reinvent the wheel.

2. Authentication
   Authentication is the act of proving an assertion, such as the identity of a computer system user. Here, Django comes with a user authentication system. It handles user accounts, Communities based platforms , and permission sessions. Briefly, authentication verifies whether a user is who they claim to be. Requests for protected resources by unauthenticated users always result in an authentication challenge.

3. Dahsboard
   A dashboard is a type of graphical user interface which presents various graphs and charts in the form of a report. This is the part of application where user can see the information which they have uploaded at the time of the registration like mobile number, uploaded image, previous result etc

4. Pre-processing
   Pre-processing involves altering the raw CT image's quality. This involves doing specific procedures on the image in order to improve particular visual details and data.

5. Feature Extraction
   The process of turning raw data into processable numerical features while keeping the original data set's content intact. Compared to using machine learning on the raw data directly, it produces better outcomes

5.2 Tools and Technologies Used

- OS: Windows
- Tools: Visual Studio Code, Django application Server
- Programming Language Used: Python
- Scripting Language Used: Hypertext Markup Language, JavaScript, Cascading Style Sheets, Bootstrap-4

5.3 Algorithm Details

5.3.1 Convolutional Neural Network

A convolutional neural network (CNN or ConvNet) is one of the most popular algorithms for deep learning, a type of machine learning in which a model learns to perform classification tasks directly from images, video, text, or sound.

CNNs are particularly useful for finding patterns in images to recognize objects, faces, and scenes. They learn directly from image data, using patterns to classify images and eliminating the need for manual feature extraction. Like other neural networks, a CNN is composed of an input layer, an output layer, and many hidden layers in between. These layers perform operations that alter the data with the intent of learning features specific to the data. Three of the most common layers are: convolution, activation or ReLU, and pooling. Convolution puts the input images through a set of convolutional filters,each of which activates certain features from the images. Rectified linear unit (ReLU) allows for faster and more effective training by mapping negative values to zero and maintaining positive values.

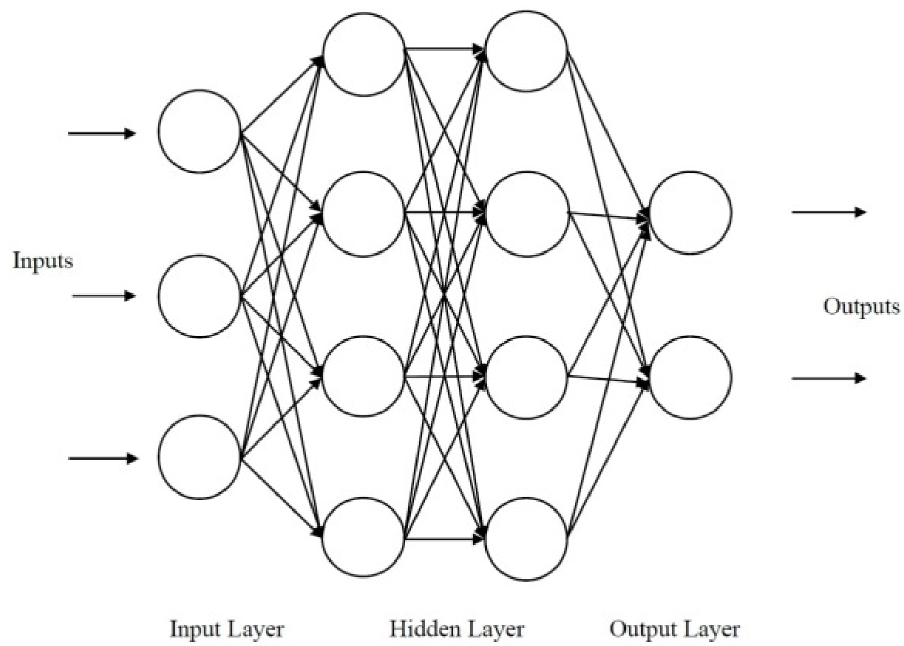

Figure 5.1: Neural Network

This is sometimes referred to as activation, because only the activated features are carried forward into the next layer. Pooling simplifies the output by performing nonlinear downsampling, reducing the number of parameters that the network needs to learn. These operations are repeated over tens or hundreds of layers, with each layer learning to identify different features.

5.3.2 Architecture of CNN

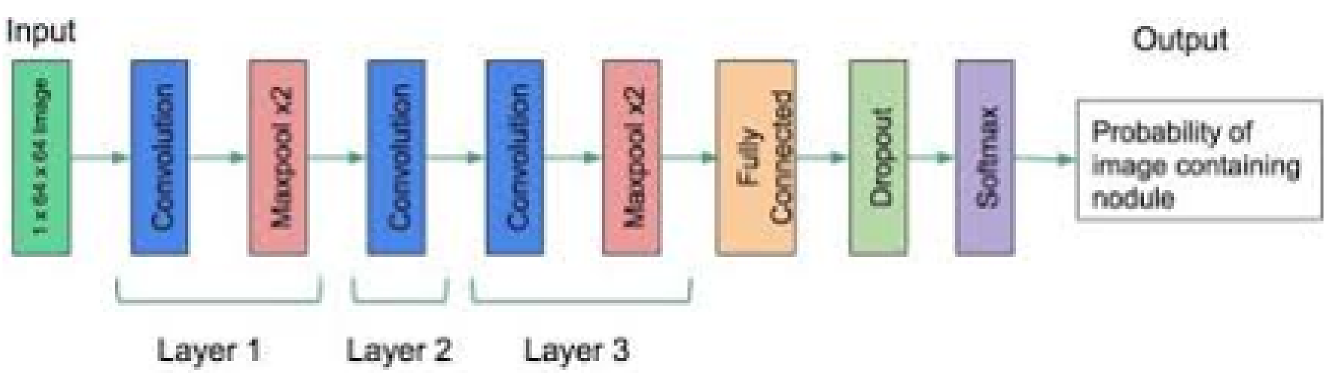

Figure 5.2: CNN Architecture

CNNs or convolutional neural nets are a type of deep learning algorithm that does really well at learning images.Consider any image, as every image represent some pixels in simple terms. We analyze the influence of nearby pixels in an image by using something called a filter (can be called as weights, kernels or features)

Filters are tensor which keeps track of spatial information and learns to extract features like edge detection, smooth curve, etc of objects in something called a convolutional layer. The major part is to detect edges in the images and these are detected by the filters. It helps to filter out unwanted information to amplify images. There are high-pass filters where the changes occur in intensity very quickly like from black to white pixel and vice-versa. The tricky part is that the CNN expects an image: a graycale image(1 channel) or a color image with three channels (RGB). Choose either 1D for a grayscale image (one feature) or 3D for a color image (to represent multiple features).

The number of parameters in a given layer is the count of "learnable" elements for a filter aka parameters for the filter for that layer. Computations in Each layer are as follows: Input layer: Input layer has nothing to learn, at it's core, what it does is just provide the input image's shape. So no learnable parameters here. Thus number of parameters = 0.

CONV layer: This is where CNN learns, so certainly we'll have weight matrices. To calculate the learnable parameters here, all we have to do is just multiply the by the shape of width m, height n, previous layer's filters d and account for all such filters k in the current layer. Don't forget the bias term for each of the filter. Number of parameters in a **CONV layer would be :** ((m * n * d)+1)* k), added 1 because of the bias term for each filter. The same expression can be written as follows: ((shape of width of the filter * shape of height of the filter * number of filters in the previous layer+1)*number of filters).

POOL layer: This has got no learnable parameters because all it does is calculate a specific number, no backprop learning involved! Thus number of parameters = 0.

Fully Connected Layer (FC): This certainly has learnable parameters, matter of fact, in comparison to the other layers, this category of layers has the highest number of parameters. It is the product of the number of neurons in the current layer c and the number of neurons on the previous layer p and as always, do not forget the bias term. Thus number of parameters here are: ((current layer neurons c * previous layer neurons p)+1*c).

Chapter 6

Results

6.1 Outcomes

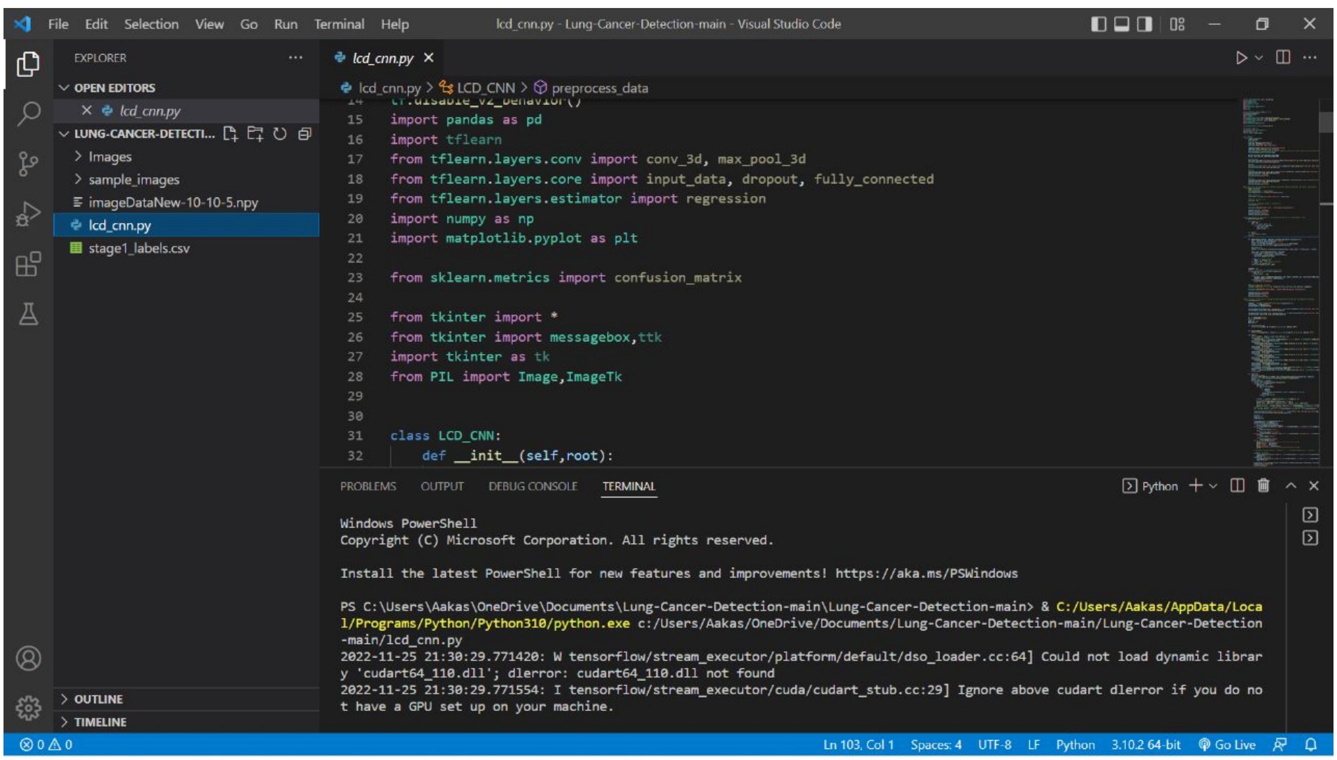

Figure 6.1: Output 1

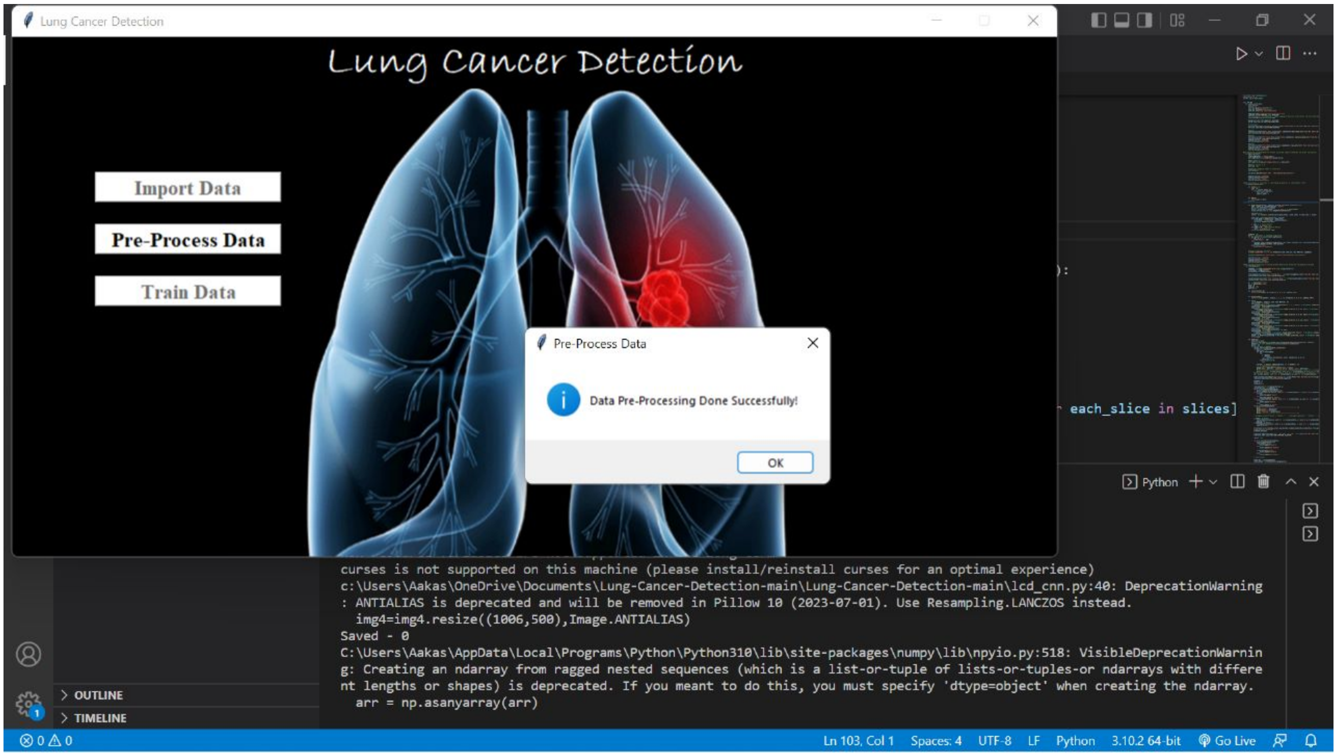

Figure 6.2: Output 2

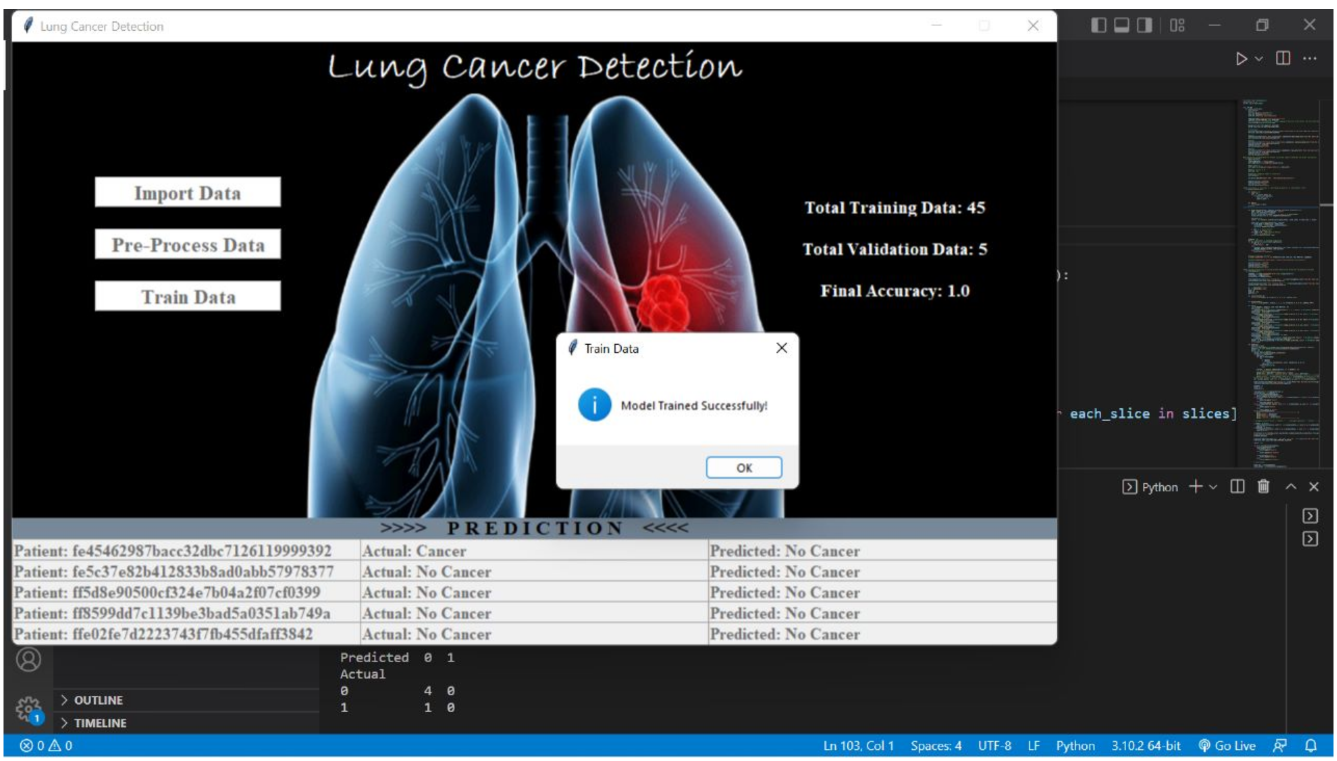

Figure 6.3: Output 3

Chapter 7

Conclusion and Future Work

7.1 Conclusion

This procedure simplifies the time commitment and boosts the accuracy of the diagnosis. Noise is removed from the data after it has been collected. Then the lung images are segmented, and finally the image is separated. The diagnosis is made from the output image, which was trained using the CNN model. We are working to create automated tools for processing medical images that can recognise cancer cells early on.

7.2 Future Work

In summary, we believe that the trend in the development of more reliable tests for early diagnosis of lung cancer should be focused on discovery that will alleviate the discomfort of the patients, as well as the burden for the health authorities, as the techniques and methodologies currently in use are not reliable.

7.3 Applications

- Hospital
- Using these system doctors are able to easily visualize and locate the particular portion or area where the disease is being affected and finally to detect them.
- This would help to make a decision about the appropriate treatment, surgery for radiologist and following-up for a series of disease control measures.

Chapter 8

APPENDIX A

8.1 Feasibility Study

8.1.1 Operational Feasibility

The system is easy to use, portable, and has more features related to the other systems that are already available in the market.

8.1.2 Technical Feasibility

1) The system will run efficiently if provided with good internet connection.
2) No issues regarding legality will be faced as the libraries and softwares used are open source and free to use.

8.2 Problem Analysis Using Algebra

8.2.1 Relevant mathematics associated with the Project

- A1: Query provided by the user. Eg: CT Scan image of Lungs.
- A2: Query provided by user. Eg: CT Scan image of Lungs.
- R1: Result provided by Lung Cancer Detection
- A3: Image missing or Other than CT Scan image provided.
- R2: Error in Lung Cancer detection.

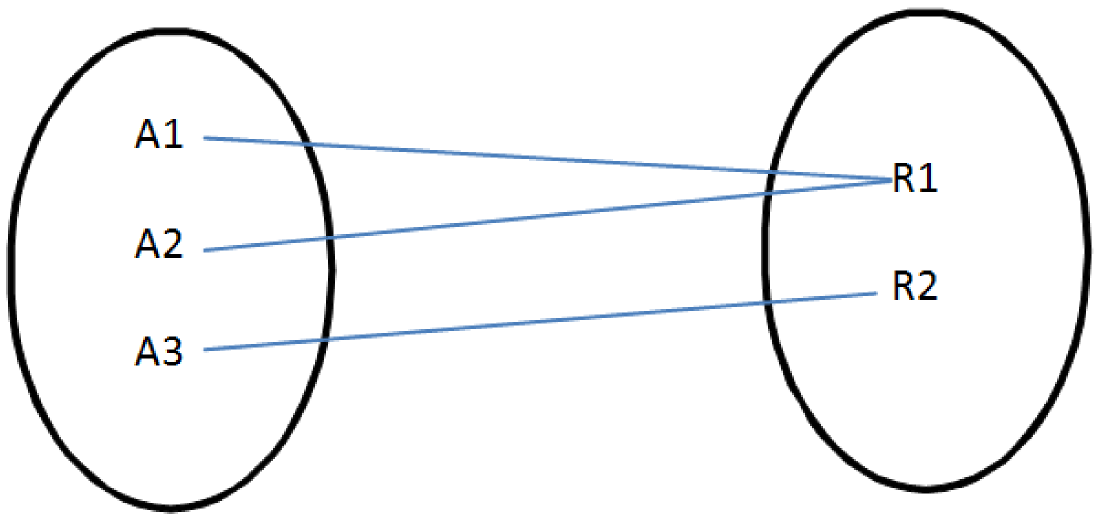

Figure 8.1: Venn Diagram

Set Theory:

S= {s, e, X, Y}

Where,

s = Start of the program.

1. Log in.
2. Upload the CT scan image.
3. Preprocessing.
4. Segmentation.
5. Feature extraction.
6. Calculate area of nodule.
7. Identify the stage of cancer.

e = End of the program. Resultant output provided by the input CT scan image.

X = Input of the program.

Input should be Image file.

Y = Output of the program.

Image will be uploading. Then the further processing will be done and finally appropriate result will provided.

X, Y U

Let U be the Set of System.
U= Client, I, K, F
Where, Client, I, K, F are the elements of the set.
Client= User
I= CT scan image of lungs.
K= CNN model.
F= Output.

8.2.2 Space Complexity

The space complexity of an algorithm is the maximum amount of space used at any one time, ignoring the space used by the input to the algorithm.

8.2.3 Time Complexity

- Check No. of image file available in the datasets= n
- If (n¿1) then retrieving of information can be time consuming.
- So the time complexity of this algorithm is O(n^2).

$\Phi = Failures and Success conditions.$

Failures:

1. Huge database can lead to more time consumption to get the information.
2. Hardware failure.
3. Software failure.

Success:

1. Search the required information from available in Datasets.
2. User gets result very fast according to their needs.

So the above mathematical model is NP-Complete.

8.3 Problem Type

Our project comes into the NP Complete, because in particular time it will not give the result. For the decision problem, so that it will give the solution for the problem within polynomial time. The set of all decision problems whose solution can be provided into polynomial time by using the attribute enhanced index.

Chapter 9

APPENDIX B

9.1 Paper Publication Details

- Title: Lung Cancer Detection Using Deep Learning
- Journal :Institute of Electrical and Electronics Engineers(IEEE) (UGC Approved)
- Year : 2022
- Status: Going to submit.

Chapter 10

APPENDIX C

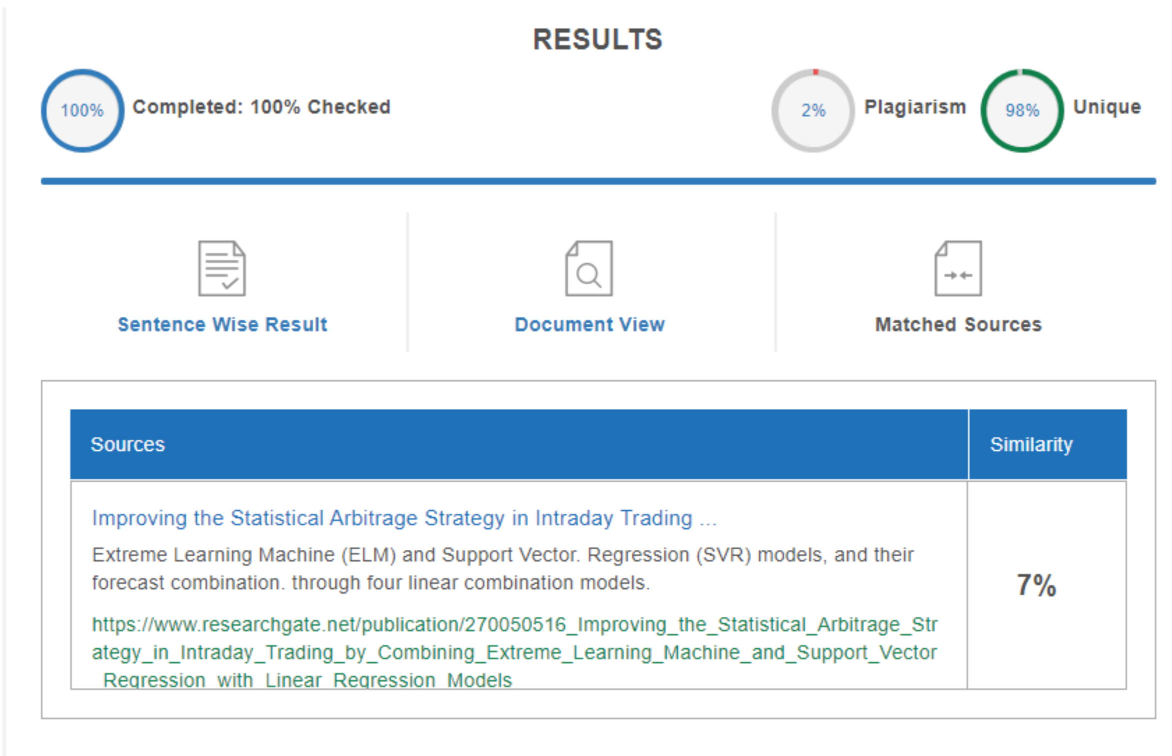

Figure 10.1: Plagarism Check

Bibliography

[1] Muthazhagan B, Ravi T, Rajinigirinath D. An enhanced computer-assisted lung cancer detection method using content-based image retrieval and data mining techniques. Journal of Ambient Intelligence and Humanized Computing. 2020 Jun 2:1-9.

[2] Sajja T, Devarapalli R, Kalluri H. Lung Cancer Detection Based on CT Scan Images by Using Deep Transfer Learning. Traitement du Signal. 2019 Oct;36(4):339-44.

[3] Masud M, Sikder N, Nahid AA, Bairagi AK, AlZain MA. A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework. Sensors. 2021 Jan;21(3):748.

[4] Tripathi P, Tyagi S, Nath M. A comparative analysis of segmentation techniques for lung cancer detection. Pattern Recognition and Image Analysis. 2019 Jan;29(1):167-73.

[5] P. Monkam, S. Qi, H.Ma,W.Gao,Y.YaoandW.Qian," Detection and Classification of Pulmonary Nodules Using Convolutional Neural Networks: A Survey," in IEEE Access, vol. 7, pp. 78075-78091, 2019,doi:10.1109/ACCESS.2019.2920980

[6] H. Guo , U. Kruger, G. Wang, M. K. Karla and P. Yan, "Knowledge-Based Analysis for Mortality Prediction From CT Images," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 2, pp. 457- 464,Feb.2020,doi:10.1109/JBHI.2019.2946066.

[7] D. Kumar et al., "Automatic Detection of White Blood Cancer From Bone Marrow Microscopic Images Using Convolutional Neural Net-works,"inIEEEAccess,vol.8,pp.142521-142531,2020,doi:10.1109/ACCESS.2020.3012292

[8] Eng.OnLine,vol.17,p.113,Aug.2018 R. Yamashita, M. Nishio , R. K. G. Do, and K. Togashi, "Convolutional neural networks: An overview and application in radiology," InsightsImag.vol.9,pp.611–629,Aug.2018.

[9] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, s Mougiakakou, lungpattern classification for ınterstitial lung diseases using a deep convolution neural network. IEEE Tran. Med. Imag. 35, 1207–1216 (2016)

[10] H. Mahersia, M. Zaroug, L. Gabralla, Lung cancer detection on CT scan ımages: a review on the analysis techniques. Int. J. Adv. Res. Artif. Intel. (IJARAI), 4(4) (2015)

[11] K. Dımılıler, B. Ugur, Y.K. Ever, Tumor detection on CT lung ımage using ımage enhancement. Online J. Sci. Technol. (TOJSAT), 7(1) (2017)

[12] J. Kim, H. Lee and T. Yoon, Automated Diagnosis of Lung Cancer with the Use of Deep Convolutional Neural Networks on Chest CT, in Proceedings of the 2017 4th International Conference on Biomedical and Bioinformatics Engineering. 2017, Association for Computing Machinery: Seoul, Republic of Korea. p. 126–132.

[13] Zhang C, Sun X, Dang K, Li K, Guo XW, Chang J, Yu ZQ, Huang FY, Wu YS, Liang Z, Liu ZY, Zhang XG, Gao XL, Huang SH, Qin J, Feng WN, Zhou T, Zhang YB, Fang WJ, Zhao MF, Yang XN, Zhou Q, Wu YL, Zhong WZ. Toward an Expert Level of Lung Cancer Detection and Classification Using a Deep Convolutional Neural Network. Oncologist. 2019 Sep;24(9):1159-1165. doi: 10.1634/theoncologist.2018-0908. Epub 2019 Apr 17. PMID: 30996009; PMCID: PMC6738288.

[14] A. Shaffie, A. Soliman, H. A. Khalifeh, M. G., F. Taher, R. Keynton, et al., "On the integration of ct-derived features for accurate detection of lung cancer", 2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 435-440, 2018.