**LUNG CANCER PREDICTION**

Using Deep Learning

**Developed by: Sheetal P, Meghana H D, Vidyashree M B, Smruthi S Kiran**

**Smart Bridge – Remote Summer Internship Program**

1. **INTRODUCTION**

Lung cancer is one of the most dreadful diseases in the developing countries and its mortality rate is 19.4%. Early detection of lung tumor is done by using many imaging techniques such as Computed Tomography (CT), Sputum Cytology, Chest X-ray and Magnetic Resonance Imaging (MRI). Detection means classifying image into two classes non-cancerous tumor and cancerous tumor . The chance of survival at the advanced stage is less when compared to the treatment and lifestyle to survive cancer therapy when diagnosed at the early stage of the cancer. Manual analysis and diagnosis system can be greatly improved with the implementation of image processing techniques. A number of researches on the image processing techniques to detect the early stage cancer detection are available in the literature. But the hit ratio of early stage detection of cancer is not greatly improved. With the advancement in the machine learning techniques, the early diagnosis of the cancer is attempted by lot of researchers. Neural network plays a key role in the recognition of the cancer cells among the normal tissues, which in turn provides an effective tool for building an assistive AI based cancer detection. The cancer treatment will be effective only when the tumor cells are accurately separated from the normal cells Classification of the tumor cells and training of the neural network forms the basis for the machine learning based cancer diagnosis. This paper presents a Convolutional Neural Network (CNN) based technique to classify the lung cancerous or not.

* 1. **Overview**

A risk factor is anything that increases chances of getting a disease, such as cancer. But having a risk factor, or even many, does not confirm to get the disease. While we cannot change some cancer risk factors like family history and aging, etc. While there are some factors that can be controlled. Thus diagnosing the cancer at the earliest stage is very important. Nowadays any cancer affects the human and may lead to death and lung cancer is one of its kind. To decrease the mortality rate and give a good treatment for the affected ones we need a better technique to diagnosis the lung cancer in initial stage itself. We can also analysis the risk of lung that it has been affected with cancer or not. So we have used the convolutional neural network (CNN) which is a class of deep neural network. By using the algorithm a Flask model has been implemented and tested. The results has been discussed and a full comparison between algorithms was conducted. CNN was selected as best algorithm based on accuracy.

* 1. **Purpose**

Tumors are continuously evolving biological systems, and medical imaging is uniquely positioned to monitor changes throughout treatment. Smoking causes the majority of lung cancers both in smokers and in people exposed to secondhand smoke. But lung cancer also occurs in people who never smoked and in those who never had prolonged exposure to secondhand smoke. In these cases, there may be no clear cause of lung cancer Early prediction of Lung cancer will help with the survival of cancer patients. Machine learning and Deep Learning have been widely used in the diagnosis of Lung cancer and on the early detection. The main aim of this research is to review the role of deep learning techniques in Lung cancer detection and diagnosis

# 2. LITERATURE SURVEY

A literature review showed that there have been several studies on the survival prediction problem using statistical approaches and Convolutional Neural Networks. However, we could only find a few studies related to medical diagnosis and recurrence using data mining approaches, artificial neural networks, convolutional neural networks to develop prediction models for breast cancer survival by analyzing a large dataset. There are a wide range of prognostic factors which determine survival duration of lung cancer patients. Among the most important factors are patient demographics, disease factors and health care related factors. Lung cancer remains one of the leading causes of the overall cancer burden worldwide.

## Existing Problem

Nowadays any cancer affects the human and may lead to death and lung cancer is one of its kind. To decrease the mortality rate and give a good treatment for the affected ones we need a better technique to diagnosis the lung cancer in initial stage itself. The existing lung cancer diagnosing systems using CNN systems for the symptoms based study are suffering with lack of sensitivity, specificity and accuracy. The existing systems were designed with fuzzy rule based systems, neural systems and supervised learning systems. However, in future hybrid techniques using more than two systems as a combination may improve sensitivity, specificity and accuracy.

## Proposed Solution

**Deep Learning (Convolutional Neural Networks):**

This proposed solution presents lung cancer detection based on chest CT images using CNN. In the first stage, lung regions are extracted from CT image and in that region each slices are segmented to get tumors. The segmented tumor regions are used to train CNN architecture. Then, CNN is used to test the patient images. The main objective of this study is to detect whether the person has cancer or not by analysing the given input image.

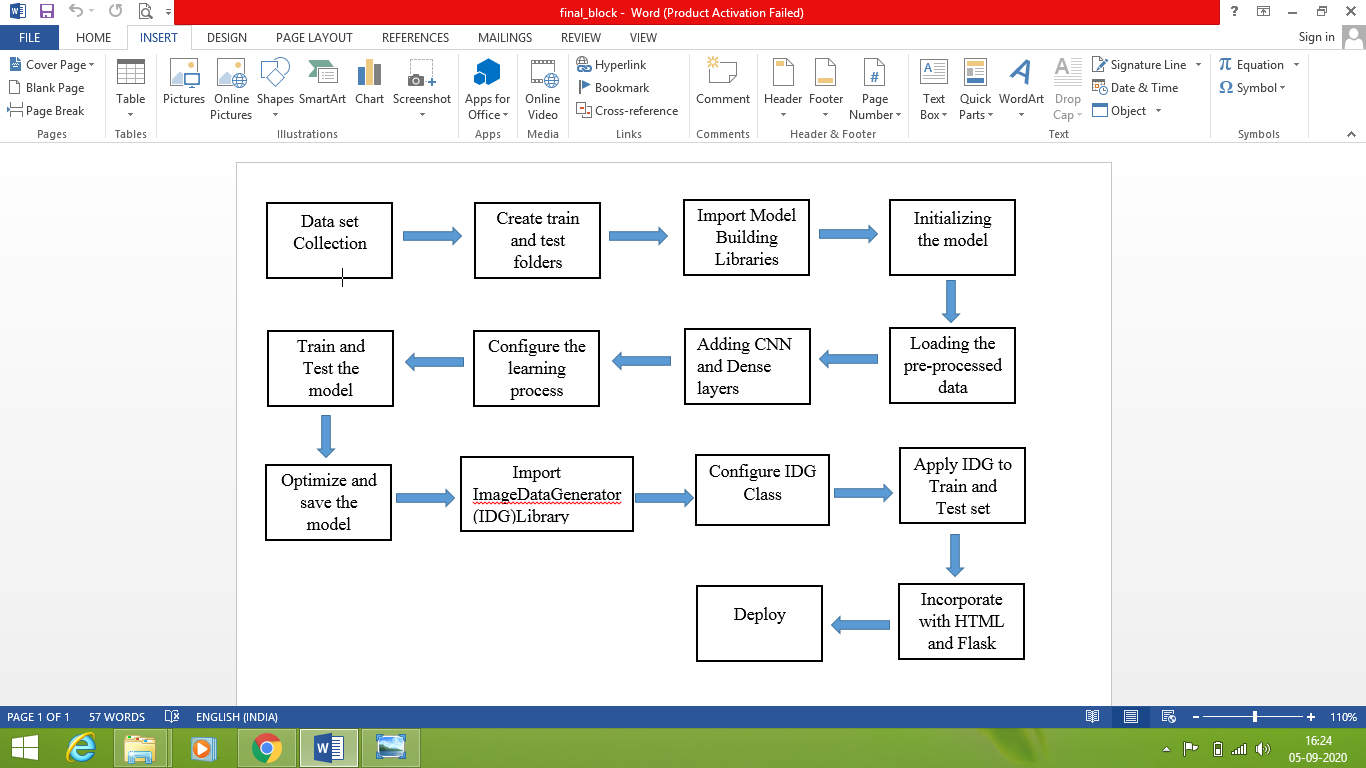
1. **THEORETICAL ANALYSIS**

Convolutional Neural Networks is all about using Deep Learning with Computer Vision. A good way to gain intuition about this is to think about a Neural Network Architecture and how it is applied to visual tasks i.e. Images and Video. Convolutional Neural Networks have allowed us to create Facial Recognition, Object Recognition.

Like a Neural Network, a typical Convolutional Neural Network consists of a multiple hidden layers called a Convolutional Layer where the linear function computes the strided convolutions over an image to extract features. It also consists of a pooling layer that computes another function such as Max Pool or Average Pool to reduce the size of the image in the neuron to speed up the computation. It does it by extracting the features of the neuron image and ignoring the rest, this makes the network more robust. There is also fully connected layer which is like a hidden layer in a neural network where the sum of the outputs of each layer are flattened and where each value is an input to the next layer followed by an activation function and an output. In a Convolutional Neural Network the linear function that is used is called a convolutional layer. Each node in the hidden layer extracts different features by using image processing feature detectors. For example, in the first layer, the first node may extract the horizontal edges of an image, the second node may extract vertical edges and etc. These features are extracted using a kernel.

The pooling layer happens tends to be computed after the convolutional layer. The reason why pooling is done is to further reduce the dimensions of the convolutional layer and just extract out the features to make the model more robust. There are two types of pooling done: max pooling and average pooling. Max pooling extracts out the highest pixel value out of a feature while average pooling calculates the average pixel value that has to be extracted.

* 1. **Block Diagram**

****

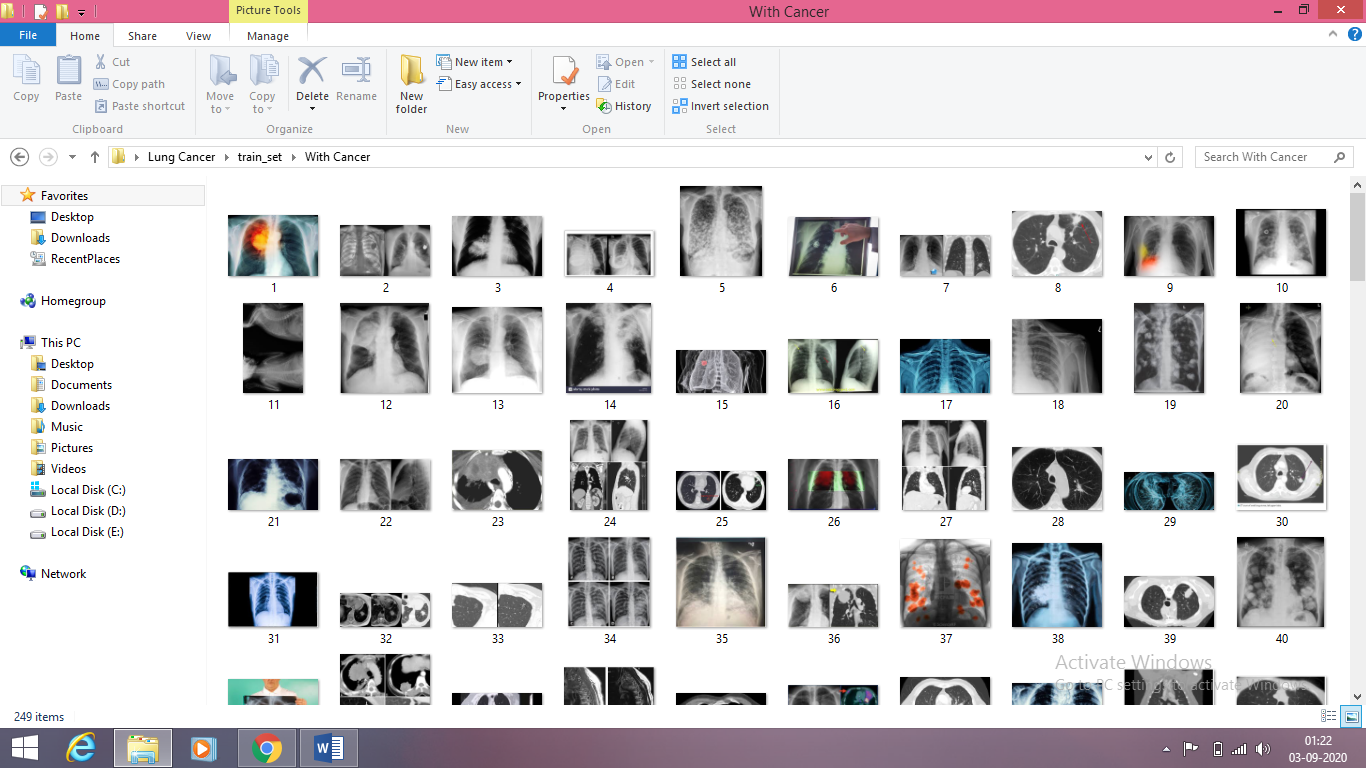
**3.2 Software Designing**

* + - Jupyter Notebook Environment
    - Spyder
    - Deep Learning Algorithms
    - Python(Sequential, Dense,Conv2D,MaxPool2D,Flatten)
    - HTML

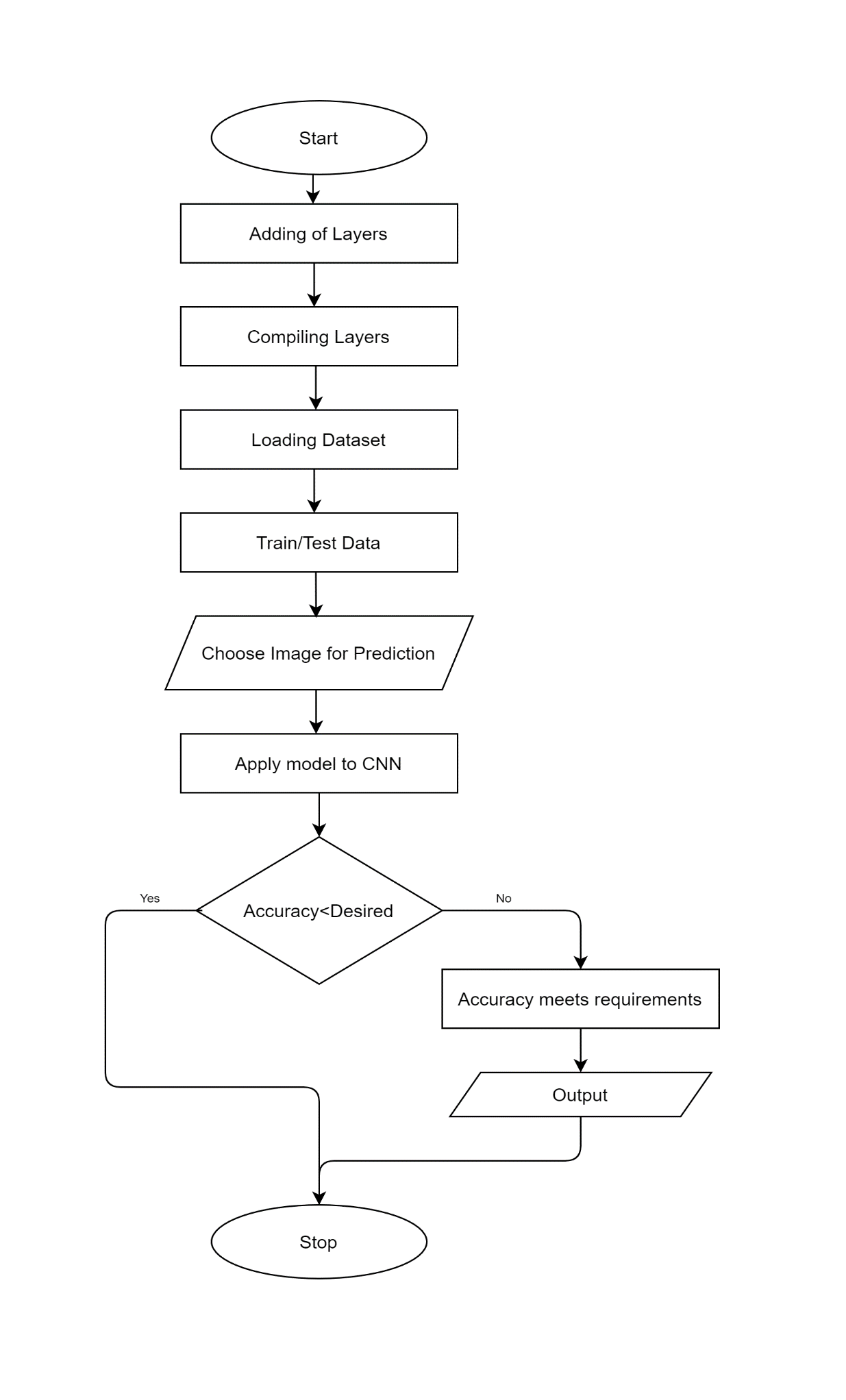
We developed this lung cancer prediction by using the Python language which is a interpreted and high level programming language and using the Deep Learning algorithms. For coding we used the Jupyter Notebook environment of the Anaconda distributions and the Spyder, it is an integrated scientific programming in the python language. For creating an user interface for the prediction we used the Flask. It is a micro web framework written in Python. It is classified as a microframework because it does not require particular tools or libraries. It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions, and a scripting language to create a webpage is HTML by creating the templates to use in th functions of the Flask and HTML.

1. **EXPERIMENTAL INVESTIGATION**

In this paper, the dataset we used is derived from <https://www.kaggle.com> and our dataset contains more than 500 images. This dataset contains images of lungs which is used for the prediction of Cancer.

****

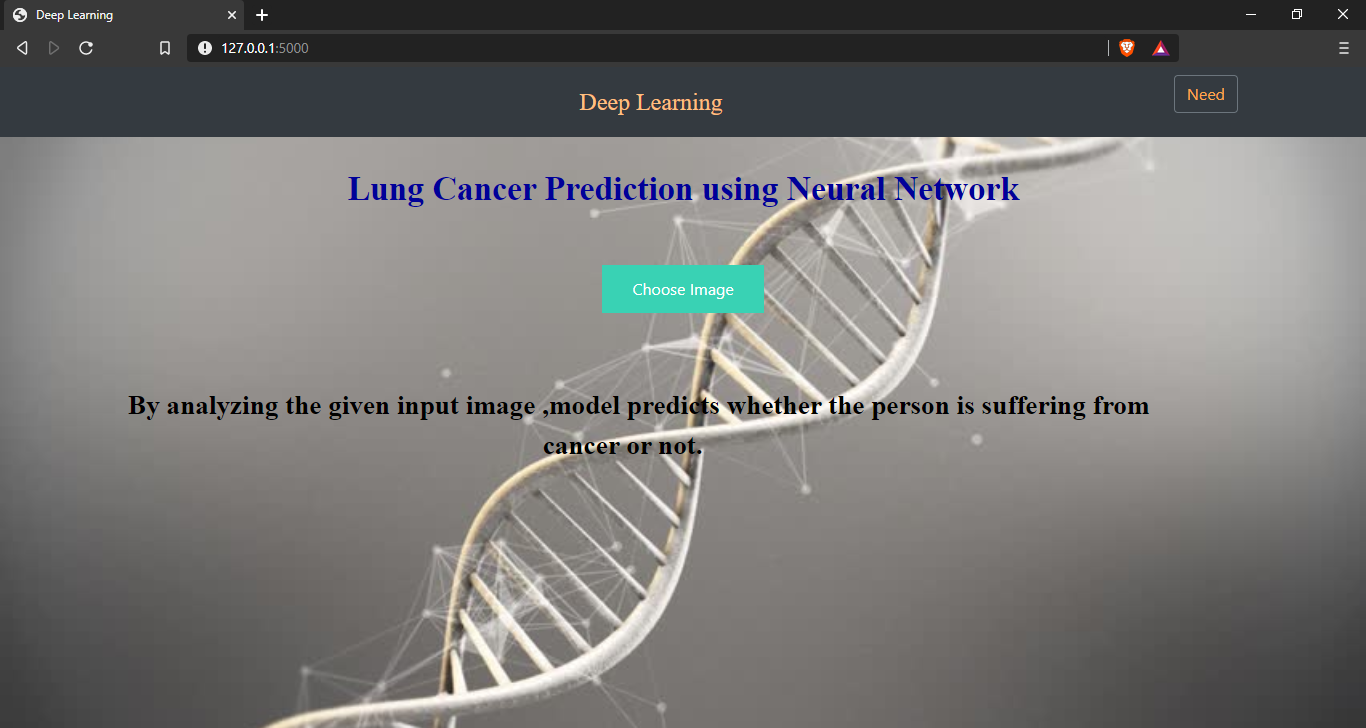
**5.FLOWCHART**

****

1. **RESULT**

A convolutional neural network based system was implemented to detect the Cancer tissues present in the input lung CT image. Lung image with different shape, size of the cancerous tissues has been fed at the input for training the system. The proposed system is able to detect the presence and absence of cancerous cells with accuracy of about 80% after running it for about 50 epoch.

**Snapshots:**

 Fig 1: Home Page

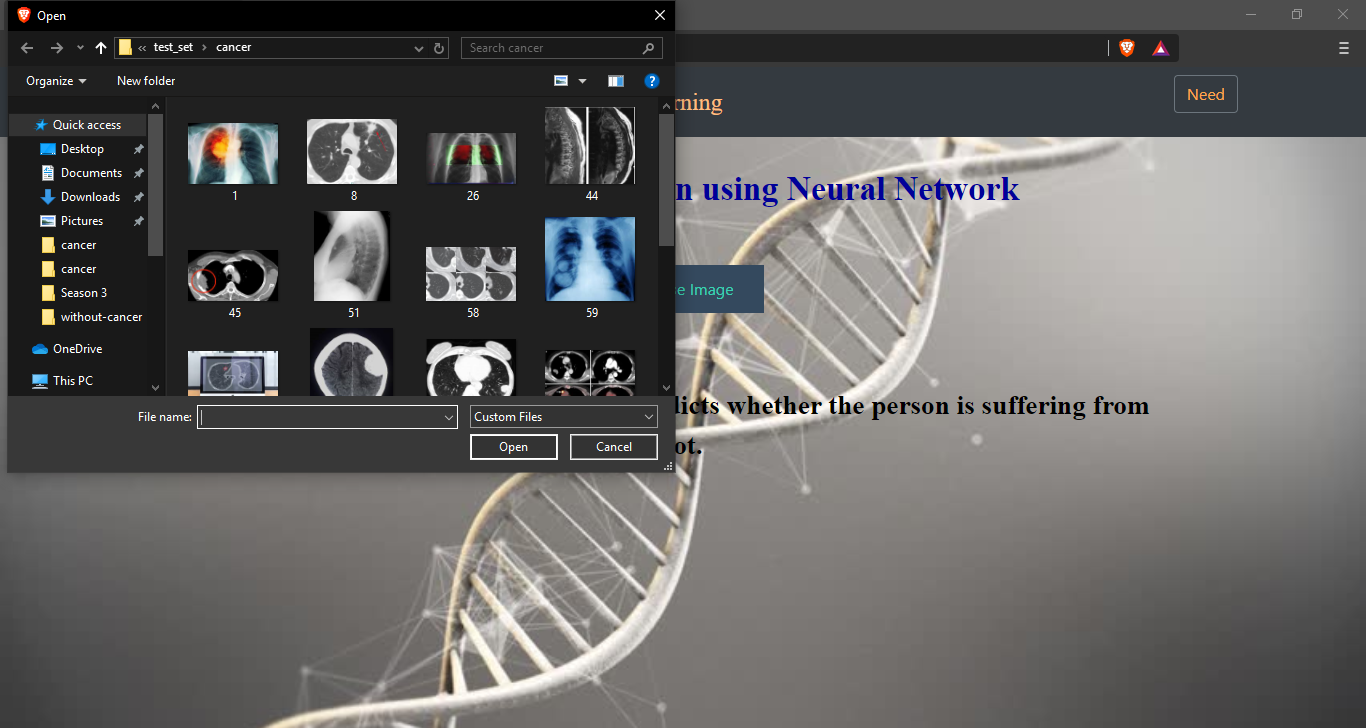


Fig 2: When you click on choose image button, a small dialog box is opened to choose image

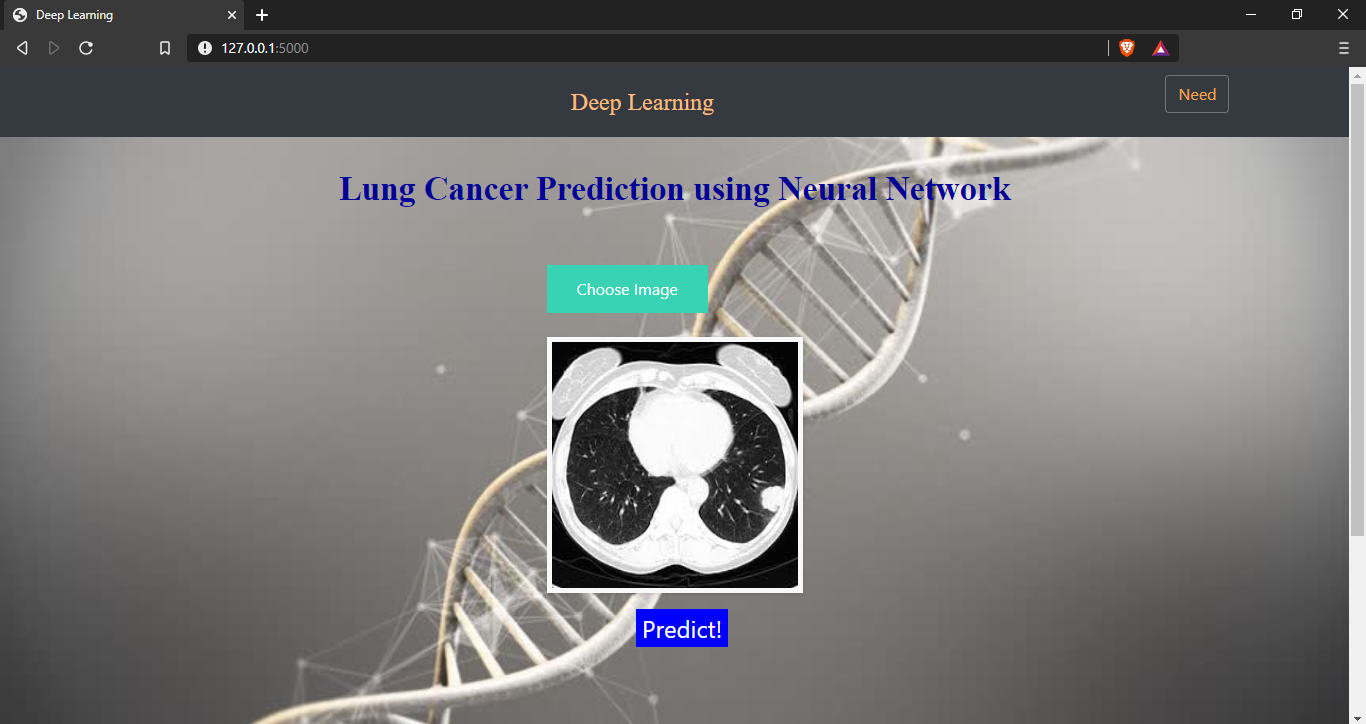


Fig 3: After selecting the image, it will displayed on screen

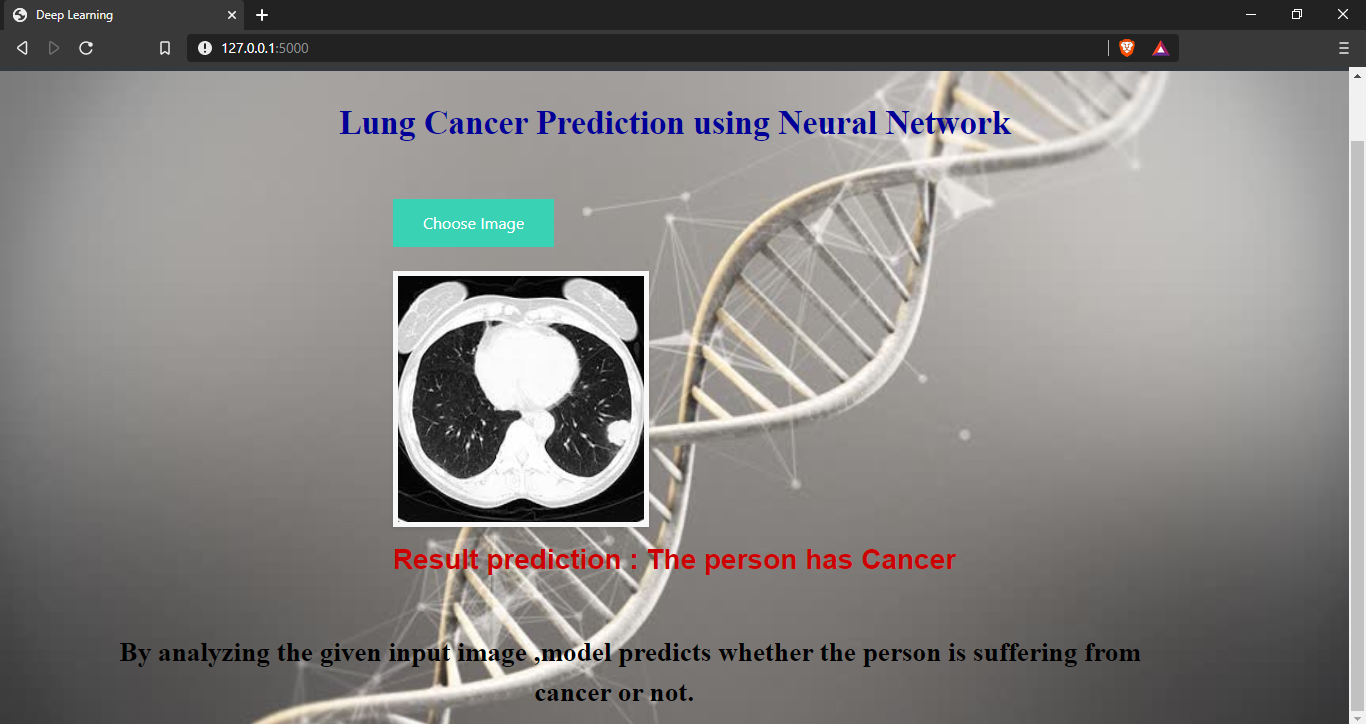


fig 4: Output screen 1

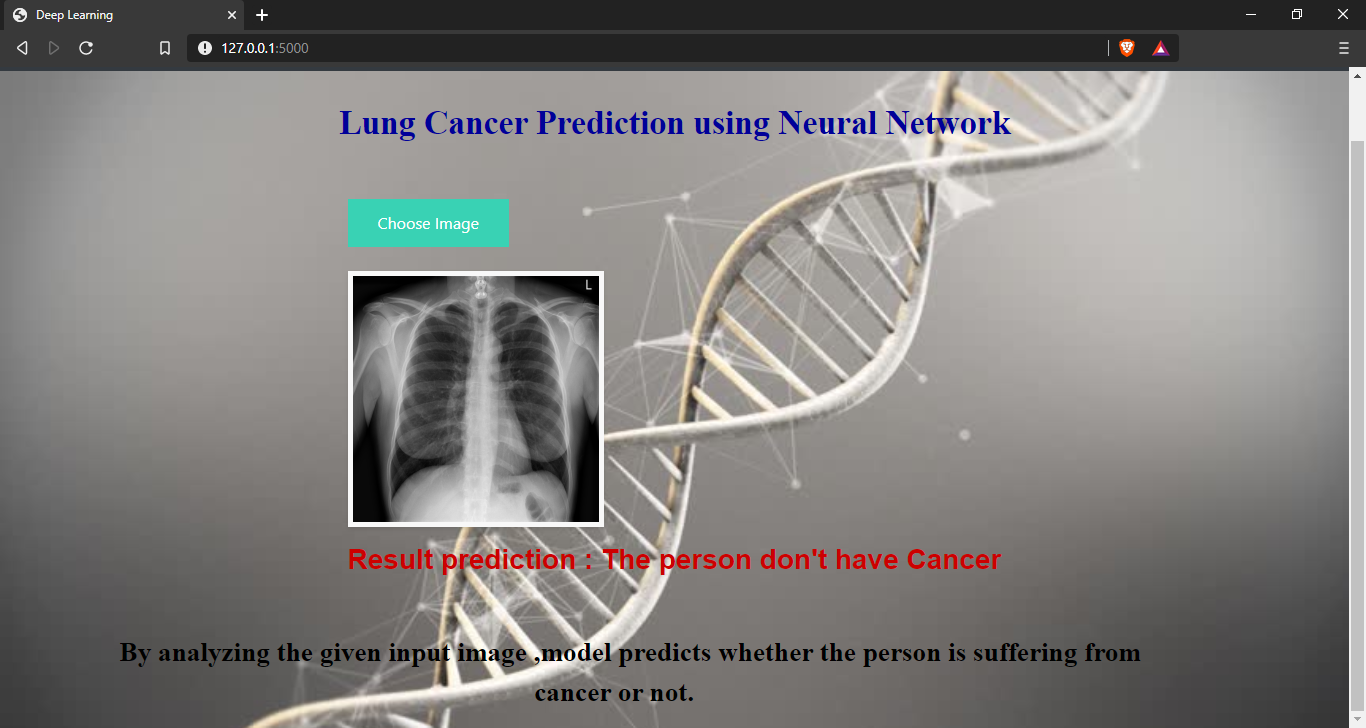


fig 5: Output screen 2



fig 6: When you click on need menu present at top right corner, you will be redirected to the above page displayed.

# ADVANTAGES AND DISADVANTAGES

**Advantages:**

* Easy and simple User Interface for detection of lung cance at early stage using radiology images
* The known benefits of lung cancer screening include reduced lung cancer mortality and using it as a “teachable moment for smoking cessation, which is the single best way to prevent lung cancer,”
* It is widely used for managing risks in medical domine.
* It is composed using the HTML and Python for the web usage in real time.
* It can work in real time and predict as soon as the necessary details for prediction are given to the model.

**Disadvantages:**

* There are also potential harms to doing these studies that are usually related to “detection, management of abnormalities, and treatments,” Screening patients who are low risk or who will not benefit from screening results in unnecessary diagnostic tests, potentially unnecessary invasive procedures.
* It could not work anywhere like an web-application, if one is using other should be quiet.
* Needs more than a single value for the prediction.

# APPLICATIONS

* It is widely used in many medical areas for early detection of cancer so the the proper treatment will be provided to the patient.
* This image processing technique can also be used to detect other cancers also such as breast cancer and tumor in our body part.
* To have an idea of patients health data.
* So we use deep learning CNN Algorithms to detect lung cancer by radiology images

# CONCLUSION

An image processing technique is built to detect diseases at early stage of cancer so the patient can take the treatment at early stages. The time factor is major factor to discover the abnormal tissue in target x-ray images. The accuracy and the quality of image is one the major core factor of this research. The empirical results demonstrate the advantage of the proposed CNN system for detecting lung cancer. The main feature for detection of accurate image comparison are pixel percentage and mask labeling which gives us the indication that the process of detection this disease plays a very important and essential role to avoid serious stages and to reduce its percentage distribution in the world. Machine learning algorithms can be used for medical oriented research, it advances the system, reduces human errors and lowers manual mistakes

# FUTURE SCOPE

This project mainly deals with CNN for early detection of lung cancer nodules from the Chest Computer Tomography (CT) image. This technique is highly successful in detecting lung cancer. For improving the performance of the proposed approaches, some future enhancements would be essential to improve upon the current research work. The main aim of the future enhancements would be to increase the sensitivity and specificity of the system.

The future scope for this research is to use Extreme Learning Machine in the CNN system. The use of ELM would provide even better results. Moreover the training time and the accuracy of the system would also be bettered by using ELM. The second enhancement in this research would be to use new clustering algorithms which can provide better results than the proposed clustering algorithms. Further future enhancements would be incorporating cutting edge technologies like nanotechnology, genetic algorithm etc., into the CNN system for better performance of the system.

# BIBILOGRAPHY

* Hastie, Friedman, and Tibshirani, *The Elements of Statistical Learning*, 2001
* Bishop, *Pattern Recognition and Machine Learning*, 2006
* Ripley, *Pattern Recognition and Neural Networks*, 1996
* Duda, Hart, and Stork, *Pattern Classification*, 2nd Ed., 2002
* Tan, Steinbach, and Kumar, Deep Learning, Addison-Wesley, 2005.
* Scholkopf and Smola, *Learning with Kernels*, 2002
* Mardia, Kent, and Bibby, *Multivariate Analysis*, 1979
* Computational Statistics (online book)
* Sutton and Barto, *Reinforcement Learning: An Introduction*, MIT Press, 1998.
* Bertsekas and Tsitsiklis, *Neuro-Dynamic Programming*, Athena Scientific, 1996.

# APPENDIX

**index.html**

{% extends "base.html" %} {% block content %}

<h2 style="text-align: center; color: #000099; font-family: Times New Roman;font-size:2.1em";><b>Lung Cancer Prediction using Neural Network</b></h2>

<br>

<br>

<div class="d-flex justify-content-center">

<div class="flex-column-center">

<form id="upload-file" method="post" enctype="multipart/form-data">

<label for="imageUpload" class="upload-label">

Choose Image

</label>

<input type="file" name="file" id="imageUpload" accept=".png, .jpg, .jpeg">

</form>

<div class="image-section" style="display:none;">

<div class="img-preview">

<div id="imagePreview">

</div>

</div>

<div>

<button type="button" style="background-color: blue;border: none; color: white; cursor: pointer; font-size: 1.5em;margin-left:35%;margin-right:auto;display:inline-block;margin-top:0%;margin-bottom:75%; class="btn btn-primary btn-lg " id="btn-predict">Predict!</button>

</div>

</div>

<div class="loader" style="display:none;"></div>

<h3 id="result" style="color:cc0000; text-align: center;font-weight:bold; font-family: sans-serif";>

<span> </span>

</h3>

<br>

<br>

</div>

</div>

<p style="color:#000000;font-family:Times New Roman;font-size:1.7em"><b>By analyzing the given input image ,model predicts whether the person is suffering from &emsp;&emsp;&emsp;&emsp;&emsp;&emsp;&emsp;&emsp;&emsp;&emsp;&emsp;&emsp;&emsp;&emsp;&emsp; cancer or not.</b></p>

{% endblock %}

**base.html**

<html lang="en">

<head>

<meta charset="UTF-8">

<meta name="viewport" content="width=device-width, initial-scale=1.0">

<meta http-equiv="X-UA-Compatible" content="ie=edge">

<title style="color:#ffbf80; text-align: center; font-family: Serif; font-size: 1.5em; margin-left:auto;margin-right:auto">Deep Learning</title>

<link href="https://cdn.bootcss.com/bootstrap/4.0.0/css/bootstrap.min.css" rel="stylesheet">

<script src="https://cdn.bootcss.com/popper.js/1.12.9/umd/popper.min.js"></script>

<script src="https://cdn.bootcss.com/jquery/3.3.1/jquery.min.js"></script>

<script src="https://cdn.bootcss.com/bootstrap/4.0.0/js/bootstrap.min.js"></script>

<link href="{{ url\_for('static', filename='css/main.css') }}" rel="stylesheet">

</head>

<body style="background-color: #ffc34d;background-image:url('https://ak.picdn.net/shutterstock/videos/13222175/thumb/1.jpg');background-repeat:no-repeat;background-attachment:fixed;background-size:cover">

<nav class="navbar navbar-dark bg-dark">

<div class="container">

<a class="navbar-brand" style="color:#ffbf80; text-align: center; font-family: Serif; font-size: 1.5em; margin-left:auto;margin-right:auto" href="#" >Deep Learning</a>

<form action="https://www.rsna.org/en/news/2019/November-December/Deep-Learning-For-Detecting-Lung-Cancers" method="get" target="\_blank">

<button class="btn btn-outline-secondary my-2 my-sm-0" type="submit" style="color:#ffa64d">Need</button>

</form>

</nav>

<div class="container">

<div id="content" style="margin-top:2em">{% block content %}{% endblock %}</div>

</div>

</body>

<footer>

<script src="{{ url\_for('static', filename='js/main.js') }}" type="text/javascript"></script>

</footer>

</html>

**app.py**

from \_\_future\_\_ import division, print\_function

# coding=utf-8

import sys

import os

import glob

import numpy as np

from keras.preprocessing import image

from keras.applications.imagenet\_utils import preprocess\_input, decode\_predictions

from keras.models import load\_model

from keras import backend

import tensorflow as tf

global graph

tf.compat.v1.disable\_eager\_execution()

graph=tf.compat.v1.get\_default\_graph()

#global graph

#graph = tf.get\_default\_graph()

from skimage.transform import resize

# Flask utils

from flask import Flask, redirect, url\_for, request, render\_template

from werkzeug.utils import secure\_filename

from gevent.pywsgi import WSGIServer

# Define a flask app

app = Flask(\_\_name\_\_)

# Model saved with Keras model.save()

model = tf.keras.models.load\_model("models/testmodel10.h5")

# Necessary

# print('Model loaded. Start serving...')

# You can also use pretrained model from Keras

# Check https://keras.io/applications/

#from keras.applications.resnet50 import ResNet50

#model = ResNet50(weights='imagenet')

#model.save('')

print('Model loaded. Check http://127.0.0.1:5000/')

@app.route('/', methods=['GET'])

def index():

# Main page

return render\_template('index.html')

@app.route('/predict', methods=['GET', 'POST'])

def upload():

if request.method == 'POST':

# Get the file from post request

f = request.files['file']

# Save the file to ./uploads

basepath = os.path.dirname(\_\_file\_\_)

file\_path = os.path.join(

basepath, 'uploads', secure\_filename(f.filename))

f.save(file\_path)

img = image.load\_img(file\_path, target\_size=(64, 64))

x = image.img\_to\_array(img)

x = np.expand\_dims(x, axis=0)

with graph.as\_default():

preds = model.predict\_classes(x)

#index = ['No Tumor','Yes Tumor']

#text = "prediction : "+[preds[0]]

if preds[0][0]==0:

prediction="The person has Cancer"

else:

prediction="The person don't have Cancer"

text = "prediction : "+prediction

# ImageNet Decode

return text

if \_\_name\_\_ == '\_\_main\_\_':

app.run(debug=False,threaded = False)