Detection of Lung Cancer Using Image Processing

Pagidi Shravani

Student
Department of InformationTechnology
B.V. Raju Institute of TechnologyAffiliated to JNTUH
Vishnupur, Narsapur, Medak,Telangana State, India 19211a1287@bvrit.ac.in

P Om Sri Venkata Sai

Student
Department of InformationTechnology
B.V. Raju Institute of TechnologyAffiliated to JNTUH
Vishnupur, Narsapur, Medak,Telangana State, India 19211a1290@bvrit.ac.in

Mr. L Sandeep Assistant professor
Department of InformationTechnology
B.V. Raju Institute of TechnologyAffiliated to JNTUH
Vishnupur, Narsapur, Medak,Telangana State, India sandeep.l@bvrit.ac.in

Abstract

Cancer is one of the most serious and widespread disease that is responsible for large number of deaths every year. Among all different types of cancers, lung cancer is the most prevalent cancer having the highest mortality rate. X Rays are used for identification of lung cancer. Image processing techniques are used widely in medical fields for early stage detection of lung cancer. This work presents an automated approach for detection of lung cancer in X Ray images. Geometrical features are computed from the extracted region of interest and used to classify X Ray images into normal and abnormal by using support vector machine.

Keywords: *Image Pre-processing,X rays.*

I.INTRODUCTION

Machine learning is programming computers to optimize a performance criterion using example data or experience. We have a model defined up to some parameters, and learning is the execution of a computer program to optimize the parameters of the model using the training data or experience. The model may be predictive to make predictions in the future,or descriptive to gain knowledge from data.

The field of study known as machine learning is concerned with the question of how to construct computer programs that automatically improve with experience.

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks T, as measured by P, improves with experience.

Cancer is a noteworthy general heath issue worldwide with mortality rates increasing day by day. Lung cancer, among all other cancer types is the most common and deadly that occur both in men and women. Lung cancer, additionally known carcinoma is formation of malignant lung tumors due to uncontrolled growth of cells in lung tissues. Eating tobacco and smoking are the leading risk factors for causing cancerous lung nodules. The survival rate of lung cancer patients combining all stages is very less roughly 14% with a time span of about 5-6 years. The main problem with lung cancer is that most of these cancer cases are diagnosed in later stages of cancer making treatments more problematic and significantly reducing the survival chances. Hence detection of lung cancer in its earlier stages can increase the survival chances up to 60-70% by providing the patients necessary fast treatment and thus it curbs the mortality rate.

Lung cancer prediction method employs deep learning is significantly more effective than our Proposed model. This will improve with the inclusion of further functionalities for future upgrades.

II.IMPLEMENTATION

Image processing is the process of transforming an image into a digital form and performing certain operations to get some useful information from it. The image processing system usually treats all images as 2D signals when applying certain predetermined signal processing methods.

Types of Image Processing:

There are five main types of image processing:

- Visualization - Find objects that are not visible in the image
- Recognition - Distinguish or detect objects in the image
- Sharpening and restoration - Create an enhanced image from the original image
- Pattern recognition - Measure the various patterns around the objects in the image
- Retrieval - Browse and search images from a large database of digital images that similar to the original image.

Why Image Processing?

Image Processing is only an aspect of Computer Vision, and they are not the same. Image Processing systems focus on transforming images from one form to another, and Computer Vision systems help the computer to understand and get meaning from an image.

Many Computer Vision systems employ Image Processing algorithms.For example, a face enhancement app may use computer vision algorithms to detect faces in a photo, and then apply Image Processing techniques like smoothing or Gray scale filters to it.

Many advanced Image Processing methods leverage Machine Learning Models like Deep Neural Networks to transform images on a variety of tasks,like applying artistic filters, tuning an image for optimal quality, or enhancing specific image details to maximize quality for computer vision tasks.

Convolutional Neural Networks (CNN) take in an input image and use filters on it, in a way that it learns to do things like object detection, image segmentation and classification.

Convolutional Neural Networks:

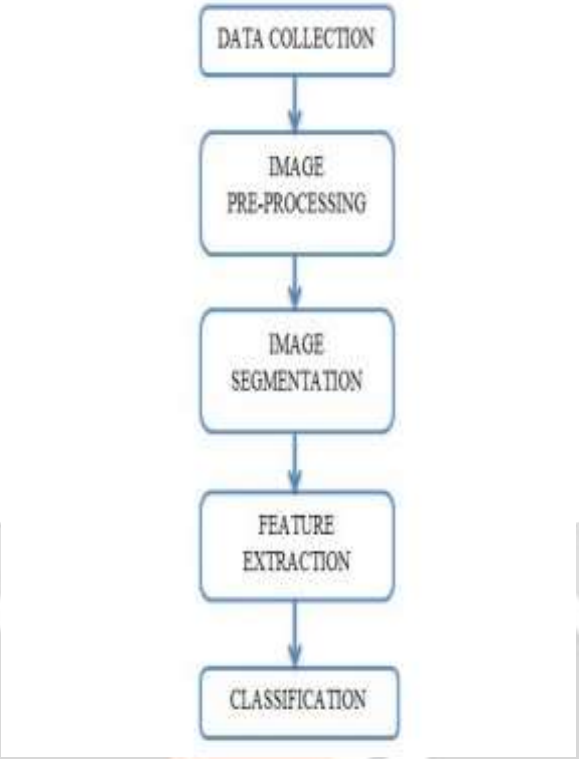
A convolutional neural network (CNN) is a subset of machine learning. It is one of the various types of artificial neural networks which are used for different applications and data types. A CNN is a kind of network architecture for deep learning algorithms and is specifically used for image recognition and tasks that involve the processing of pixel data.

There are other types of neural networks in deep learning, but for identifying and recognizing objects, CNNs are the network architecture of choice. This makes them highly suitable for computer vision (CV) tasks and for applications where object recognition is vital, such as self-driving cars and facial recognition.

Proposed System:

- The proposed system for lung cancer detection in X Ray images is shown with the help of a flowchart.

- This methodology is carried out in five main steps.
    a. Data Collection
    b. Image Pre-Processing
    c. Image Segmentation
    d. Feature Extraction
    e. Classification



Advantages of proposed system:-

- Increase in accuracy of cancer nodule detection than the best current model.
- Classifies the detected lung cancer as malignant or benign.
- Remove salt-pepper noises and speckle noise that creates false detection of cancer.

Existing System:

- Gonzalez and Ponomaryvo proposed a system that classifies lung cancer as benign or malignant. The system uses the priori information and House field Unit(HU) to calculate the Region of Interest(ROI).
- This model classifies cancer as benign, however the limitation of it is that prior information is required about region of interest.
- In image pre- processing it uses Gabor filter to enhance the image and uses marker controlled watershed method for segmentation and detects the cancer nodule.

Disadvantages of Existing system :

- Even the system is current best solution, it has some limitations.
- Only few features have been extracted for cancer nodules.
- No preprocessing like noise removal, image smoothing which can probably assists in increasing the detection of nodules accurately has been implemented.
- No classification as benign or malignant of extracted cancer has been performed.

System Design:

Software Requirement Specification (SRS) is the starting point of the software developing activity. As system grew more complex it became evident that the goal of the entire system cannot be easily comprehended. Hence the need for the requirement phase arose. The software project is initiated by the client needs. The SRS is the means of translating the ideas of the minds of clients (the input) into a formal document (the output of the requirement phase.)

System-Wide Requirements: The system will process exceptions in a consistent fashion. If the exception is a user error that can be corrected the system displays an error message with an explanation. The system allows the user to correct the mistake without starting over.

MODELS USED

Data Collection:

- The first step is to obtain lung X ray images of cancer patients. For research work,the images have been downloaded from the Cancer Imaging Archive database.
- The images are stored in DICOM format. The image database contains X Ray images of patients with and without lung cancer.

Image Pre-Processing:

- The objective of image pre-processing stage is to suppress unwanted distortions present in the image and to enhance some features useful for further processing. It includes two main steps such as image smoothing and image enhancement.
- X Ray images are prone to salt and pepper noise, hence median filtering is found to be quite effective technique in eliminating this impulse noise while preserving the edges. Median filtering gives the best results for image smoothing as it removes noise without blurring the image.

Image Segmentation:

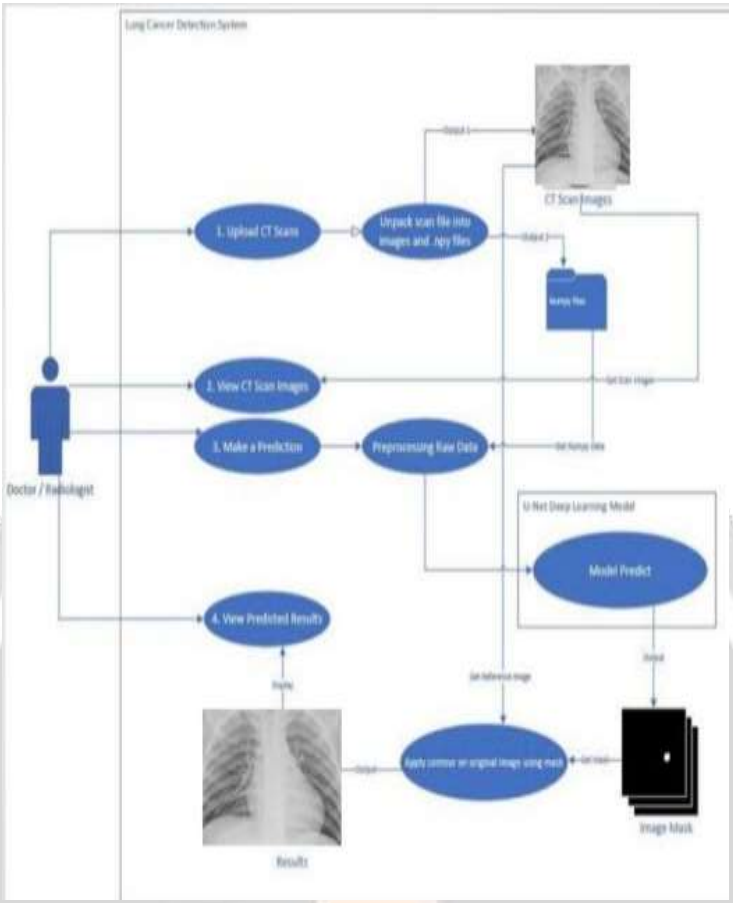
- The process of separating out required region of interest from the image is known as segmentation. Mathematical morphological operations are powerful tools in acquiring lung region from binary images.
- In our methodology, first the preprocessed gray scale images were converted to binary images. Morphological opening operation was performed to the binary image with disk structuring element for removal of unwanted components from the image.
- The opened image was then complemented and clear border operation was performed to it. The lung masks were obtained by filling the holes and gaps present in the lungs.
- Finally exclusive OR operation was performed to lung mask output and clear border output to give us the segmented tumor region.

UML Diagrams:-

TheUMLis alanguagefor

- Visualizing
- Specifying
- Constructing
- Documenting

Use Case Diagram:



III. CODE

```
From keras.layers import Input, Lambda, Dense, Flatten
fromkeras.models import Model
from keras.applications.vgg16 import VGG16
from keras.applications.vgg16 import pre-process_input
fromkeras.preprocessing import image
fromkeras.preprocessing.image import Image Data Generator
fromkeras.models import Sequential
importnumpy as np
from glob import glob
importmatplotlib.pyplot as plt

IMAGE_SIZE = [224, 224]

train_path = 'Datasets/train'
valid_path = 'Datasets/test'


vgg = VGG16(input_shape=IMAGE_SIZE + [3], weights='imagenet', include_top=False)

for layer in vgg.layers:
layer.trainable = False

folders = glob('Datasets/train/*')
```

```
x = Flatten()(vgg.output)

prediction = Dense(2, activation='softmax')(x)
model = Model(inputs=vgg.input, outputs=prediction)

model.summary()

model.compile(
loss='categorical_crossentropy',
optimizer='adam',
metrics=['accuracy']
)

fromkeras.preprocessing.image import ImageDataGenerator

train_datagen = ImageDataGenerator(rescale = 1./255,
shear_range = 0.2,
zoom_range = 0.2,
horizontal_flip = True)

test_datagen = ImageDataGenerator(rescale = 1./255)

importos
os.getcwd()

training_set = train_datagen.flow_from_directory('Datasets/train',
target_size = (224, 224),
batch_size = 32,
class_mode = 'categorical')

test_set = test_datagen.flow_from_directory('Datasets/test',
target_size = (224, 224),
batch_size = 32,
class_mode = 'categorical')

r = model.fit_generator(
generator=training_set,
validation_data=test_set,
epochs=5,
steps_per_epoch=len(training_set),
validation_steps=len(test_set)
)

importtensorflow as tf
fromkeras.models import load_model
model.save('model.h5')
```

IV. RESULTS

Input:

Fig: input photo

Fig: output screen shot

V. CONCLUSION

1. A system for automatic detection of lung cancer in X Rays images was successfully developed using image processing technique. The adopted methodology performs well in enhancing, segmenting and extracting features from X Ray images.
2. Hand detection is done using OpenCV and TensorFlow object detector. And further it is enhanced for interpretation of gestures by the computer to perform actions.
3. We have used CNN and achieved an accuracy of around 89% and also by using the decision tree algorithm, for training data we got 94.57 and for testing we got 96.72

Future Work Enhancements:-

The lung cancer prediction method that employs deep learning is significantly more effective, offering excellent results to physician and assisting them. This will improve with the inclusion of further functionalities for future upgrades. This processing technology will assist the radiologist in accurately detecting the afflicted patients.

Machine learning is critical to AI's success, and data science is the way of the future in healthcare. The machine learning algorithms to assist patients obtain accessibility to better medicine has resulted in advancements in AI displays, customized medication, forecasting medicine, and diagnosis

VI. REFERENCES

1. Sannasi Chakravarthy SR & Harikumar Rajaguru.2019. Lung Cancer Detection using Probabilistic Neural Network with modified Crow-Search Algorithm. Asian Pac J Cancer Prev.2019;20(7): 2159-2166.

2. K.A.G. Udeshani, R.G.N. Meegama , T.G.I. Fernando. 2011. Statistical Feature-based Neural Network Approach for the Detection of Lung Cancer in Chest X-Ray Images. International Journal of Image Processing(IJIP).

3. Afrianis Lubis. 2018. Identifikasi Penyakit Kanker Paru-Paru Pada Citra Medis Chest X-Ray (CXR) Menggunakan Metode Backpropagation Neural Network. Skripsi. Universitas