Report On

Lung cancer detection using CT scan data set

Submitted in partial fulfillment of the requirements of the Course project in Semester VII of Final Year Computer Engineering

by

Prerna Gawali (Roll No. 22)
Onkar Suryavanshi (Roll No. 36)
Raj Sutar (Roll No. 37)

Supervisor
Dr. Megha Trivedi

University of Mumbai

Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

(2023-24)

Vidyavardhini's College of Engineering & Technology
Department of Computer Engineering

CERTIFICATE

This is to certify that the project entitled "Lung cancer detection using ct scan data set" is a bonafide work of "Prerna Gawali (Roll No. 22)Onkar Suryavanshi (Roll No. 36) Raj Sutar (Roll No. 37)"submitted to the University of Mumbai in partial fulfillment of the requirement for the Course project in semester VII of Final Year Computer Engineering.

Supervisor

Dr. Megha Trivedi

Dr. Megha Trivedi
Head of Department

Contents

Abstract

Deep learning technologies, particularly Convolutional Neural Networks (CNNs), have advanced rapidly, ushering in a new era of precision and efficiency in medical image analysis. In this project, we present a web application that uses the power of CNNs to make accurate and timely disease diagnoses using medical imaging data. This ground-breaking technology not only streamlines healthcare professionals' workflows, but it also plays an important role in early disease detection, ultimately leading to improved patient outcomes through prompt treatment interventions.The project incorporates cutting-edge deep learning models, data preprocessing techniques, and real-time analysis capabilities, all of which are seamlessly accessible via a user-friendly web interface. The processing and interpretation of various medical imaging modalities, such as computed tomography (CT) scans, are among the system's core functions

Acknowledgments

We would like to express our sincere gratitude to everyone who helped in completion of this project. We would like to thank Dr. Megha Trivedi for her valuable guidance and support that helped us work our way through this project. We are thankful to our principal Dr. Harish Vankudre for giving us this great opportunity to do a project on Lung cancer detection using ct scan data set. We would also like to extend to all our teachers, without their support we would not have been abl1e to complete this project

Section-1

Introduction:

1.1Introduction:

Cancer is the leading cause of death in the world, with lung cancer having the highest mortality rate. Radiologists use computer tomography scans to detect and track cancer in the body.Visual database interpretation can lead to cancer detection at later stages, resulting in late cancer treatment, which only increases cancer mortality rates.

As a result, image processing tools can be used to detect cancer at an early stage. In this project, a lung cancer detection algorithm is proposed that uses mathematical morphological operations to segment the lung region of interest, from which Haralick features are extracted and used for cancer classification by neural networks.

1.2 Problem Statement:

Using Convolutional Neural Networks (CNNs) to Create a Lung Cancer Detection System.

A lung cancer detection system is a computer-aided system that aids radiologists in the detection of lung cancer in CT scans.

1.3 Scope:

The potential benefits of developing a lung cancer detection system based on CT scan data include early cancer detection, improved diagnostic accuracy, personalized treatment, telemedicine support, cost reduction, and contributions to public health initiatives. This project also provides opportunities for learning, addresses ethical concerns, and has the potential to lead to valuable collaborations with medical professionals. It has the potential for global impact, continuous improvement, and technology transfer to healthcare institutions if it is successful. However, in healthcare, ethical and regulatory compliance must be a top priority.

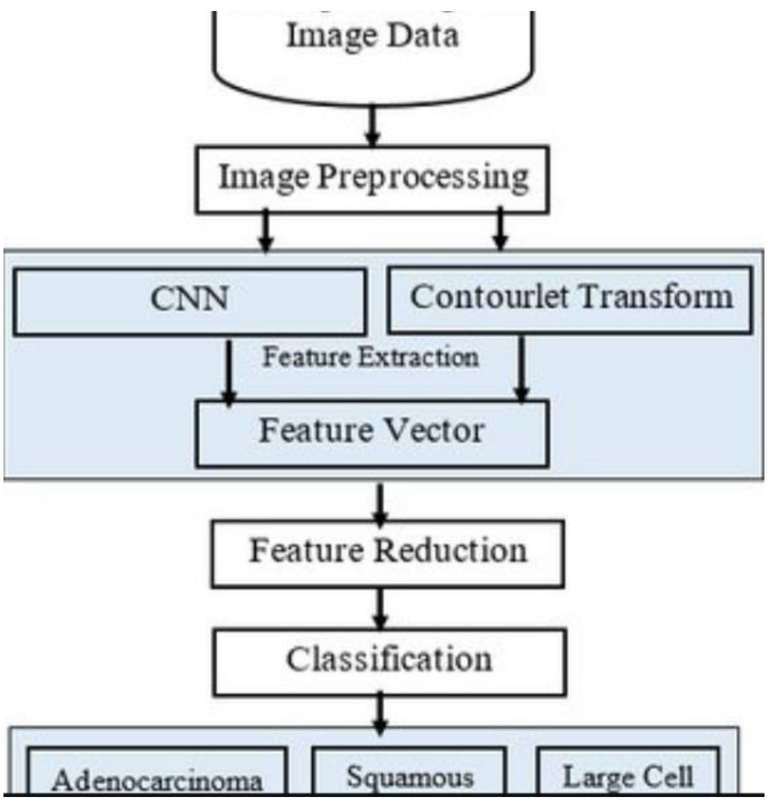
Section-2

Proposed System:

2.1 Introduction:

Cancer is the leading cause of death for a large number of people worldwide, with lung cancer having the highest mortality rates. Radiologists use computer tomography scans to detect and monitor cancer in the body. Visual database interpretation can lead to cancer detection at later stages, resulting in late cancer treatment, which only increases cancer death rates. As a result, image processing tools can be used to detect cancer early.In this project, a lung cancer detection algorithm is proposed that uses mathematical morphological operations to segment the lung region of interest, from which Haralick features are extracted and used by neural networks to classify cancer.

2.2 Architecture:



2.3 Module Description:

This module describes a lung cancer detection system using a convolutional neural network (CNN) algorithm. The system is trained on a large dataset of lung cancer CT images, and it can be used to identify lung cancer lesions in new CT images with high accuracy.

The system consists of the following steps:

Preprocessing: The input CT image is preprocessed to remove noise and normalize the pixel intensities.
Feature extraction: The CNN algorithm extracts features from the preprocessed image. These features represent the texture, shape, size, and location of the different objects in the image.
Classification: The extracted features are used to classify the image as either cancerous or non-cancerous.
The CNN algorithm is trained using a supervised learning approach. This means that the algorithm is given a set of training images that have already been labeled as cancerous or non-cancerous. The algorithm learns to identify the features in the training images that are associated with cancer. Once the algorithm is trained, it can be used to classify new images with high accuracy.
The lung cancer detection system can be used to improve the accuracy and efficiency of lung cancer screening. It can also be used to assist radiologists in making more accurate diagnoses of lung cancer.

2.4 Details of Hardware Software:

Software Requirements :
Operating System: Any modern OS (Linux, Windows, macOS).
Python: Python 3.6+.
Deep Learning Framework: TensorFlow or PyTorch.
Web Framework: Flask.
Libraries: scikit-learn, Keras, OpenCV, PIL.:

Hardware Requirements :
CPU: Multi-core processor.
GPU: Optional but beneficial (e.g., NVIDIA GeForce GTX 1080).
RAM: 8GB+ (16GB recommended).
Storage: SSD preferred.

2.5 Code:

```
import os
import numpy as np
import tensorflow

# Keras

from tensorflow.keras.models import load_model
from tensorflow.keras.preprocessing import image

# Flask utils
from flask import Flask, request, render_template
from werkzeug.utils import secure_filename
import tensorflow_hub as hub

# Define a flask app
app = Flask(_name_)

# Model saved with Keras model.save()
model = load_model(('ml_test2/ct_cnn_best_model.hdf5'),
custom_objects={'KerasLayer': hub.KerasLayer})
# print(model)
```

```
def model_predict(img_path, model):
    classes_dir = ["Adenocarcinoma","Large cell carcinoma","Normal","Squamous cell
carcinoma"]
    # classes_dir = ["Adenocarcinoma","Large cell carcinoma","Normal","Squamous cell
carcinoma"]
    # Loading Image
    img = image.load_img(img_path, target_size=(64,64))
    # Normalizing Image
    norm_img = image.img_to_array(img)/255
    # Converting Image to Numpy Array
    input_arr_img = np.array([norm_img])
    # Getting Predictions
    result = np.argmax(model.predict(input_arr_img))
    print(result)

    return (classes_dir[result])

@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
        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)

        # Make prediction
        preds = model_predict(file_path, model)
        result = preds
        return result
    return None

if _name_ == '_main_':
    app.run(debug=False
```

Section-3

3.1 Results

3.2 Conclusion:

The code demonstrates a training and deploying a CNN model for medical image classification. The Flask web application allows users to interact with the trained model for real-time predictions, which could be useful in a Lung Cancer diagnosis. Our methodology successfully developed automatic lung cancer detection for CT images using texture analysis and CNN. Pre-processing techniques enhance the accuracy of the cancer detection. Results of the proposed methodology are promising with an overall accuracy of 53.96%.

Section-4

References:

1. https://ieeexplore.ieee.org/document/8463851

2. https://ieeexplore.ieee.org/document/7097244