PREDICTION OF BENIGN AND MALIGNANT TUMOURS IN LUNG CANCER USING CNN AND DEEP LEARNING

1Kollati Nandini, 2Komali Yasudha, 3V Jhansy Archana
1 PG Student, 2,3 Assistant Professor
1Department of Computer Science,
1GITAM (Deemed to be University), Visakhapatnam, India

Abstract: Lung cancer is currently the most frequently diagnosed major cancer in the world. This is largely due to the carcinogenic effects of cigarette smoke. Over the coming decades, changes in smoking habits will greatly affect primarily those influence lung cancer incidence and mortality as well as the prevalence of various histologic types of lung cancer. Since many years, more women have died each year of lung cancer than of breast cancer. Imaging techniques such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and X-ray are used for capturing images of the lungs for analysis. Among the approaches described, the CT image technique is the most prevalent. It is difficult for doctors to interpret and detect cancer. Lung cancer can be diagnosed with high accuracy using CT scans. Early detection can save you a significant amount of time. We predict the type of tumour using CT scan images of benign, malignant, and normal cases using deep learning CNN architectures such as VGG16, resNet50, and a two-layer convolutional model in the early stages.

Index Terms - Benign, Malignant, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Convolutional Neural Network

I. INTRODUCTION

Lung cancer is the most common primary malignant tumour in men and accounts for almost 30% of all cancer deaths in both sexes in developing countries. Cancer of the lung is a disease of middle and peak incidence in 55-65 years of age. The High incidence of cancer is associated with the number of etiologic factors, notably cigarette smoking. The average active smoker has 13 times higher risk, while a passive smoker has 1.5 times risk of lung cancer than a non-smoker, more so in women probably related to hormonal factors. A variety of benign and malignant tumours may arise in the lung, but 90% to 95% are malignant, about 5% to 7% are benign.

A Benign tumour is a growth of a tissue but non-cancerous but still can be dangerous in the current scenario. The symptoms of benign lung tumour are persistent cough, respiratory infections, blood vomiting (haemoptysis), wheezing. In most cases, benign lung tumours don't require treatment. Consulting a healthcare provider will want to take a series of CT scans a period of months to years to examine the growth of the tumour.

Whereas, malignant lung tumours are dangerous and cancerous and must be treated otherwise it can be dangerous to the other tissues. The symptoms are cough that does not go away and gets worse, unexplained weight loss, coughing up rust-coloured sputum, feeling of tiredness or weakness. Around 40 out of 100 every people survive their cancer for 1 year or more after diagnosis.

In this paper, by using Deep Learning CNN architectures like VGG16, Resnet50 and 2-layer convolutional model diagnosis of early stage lung cancer can be detected by using CT scans.
II. DATASET
The lung cancer dataset IQ–OTH/NCCD is an image dataset from Kaggle. In that directory, CT scan images of lung cancer is identified with three categories: benign, malignant, and normal. The dataset contains a total of 1094 photos in jpg format.

III. RESEARCH METHODOLOGY

Libraries employed Tensorflow and Keras, which are high-level and low level APIs and whereas, keras only provides high level for machine learning and current deep learning challenges. Image data generator is used to increase the performance of the model with the support of Keras. Matplotlib is a graph representation and data visualisation library. In this paper, 3 Convolutional neural network models are used:
1. VGG16
2. resNet50
3. Custom 2-layer Convolutional model.
VGG16 (Visual Geometry Group) is a 16-layer convolutional neural network, Resnet50 (Residual Network) is a 50-layer deep convolutional neural network, and a custom 2-layer convolutional model is used in this implementation. To begin, import Sequential from keras models, which transforms our neural network into a Sequential network with sequential layers. After that, import Conv2D, which will work with images. All of the images are essentially two-dimensional. Import MaxPool2D from keras.layers for Average Pooling, Sum Pooling, and Max Pooling which are to be flattened. This involves transforming all of the 2D arrays into a single long continuous Linear vector. The final phase is dense. This Dense is used to link the neural network in its whole. Rescale, rotate, zoom, and flip images with Image Data Generator. ReLu and soft max are the activation layers employed in this model. Matplotlib is used to obtain model accuracy and to illustrate model test and train accuracy. Finally, we may put the models to the test by detecting if the image is benign, malignant, or normal.
IV. RESULTS
In this paper, the models of CNN like VGG, ResNet50 are used with a two-layer customised model. VGG16 has the highest accuracy when compared to ResNet50 and the two-layer convolutional model.

The below Figure 4 is the accuracy depicted using VGG16 model with 97.81% train accuracy and 98.45% test accuracy taking into epoch value for consideration. The below Figure 5 is the output predicts CT scan image has Malignant tumour.

The below Figure 6 is the accuracy depicted using resNet50 model with 75.41% train accuracy and 68.17% test accuracy taking into epoch value for consideration. The Figure 7 is the output predicts CT scan image has Benign tumour.

The below Figure 8 is the accuracy depicted using 2-layer convolutional model with 90.59% train accuracy and 96.39% test accuracy taking into epoch value for consideration. The Figure 9 is the output Predicts CT scan image has Normal case.
4.1 Contrasting Models used:

<table>
<thead>
<tr>
<th>S.no</th>
<th>Models</th>
<th>Training Accuracy</th>
<th>Testing Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>VGG16</td>
<td>97.81%</td>
<td>98.45%</td>
</tr>
<tr>
<td>2.</td>
<td>resNet50</td>
<td>75.41%</td>
<td>68.17%</td>
</tr>
<tr>
<td>3.</td>
<td>2-layer convolutional</td>
<td>90.59%</td>
<td>96.39%</td>
</tr>
</tbody>
</table>

V. Conclusions:

Prevention is better than cure. Lung cancer can be prevented by stop smoking, avoid exposure to toxic chemicals, healthy diet and physical exercise. Since the mortality rate of lung cancer is rising every day, it is critical that it be diagnosed as early as possible because it has a tendency to spread and is incurable if it spreads too far. Lung cancer is difficult to diagnose since symptoms only appear in the latter stages, and it is nearly hard to save a person's life at this stage. Doctors have a difficult time interpreting and diagnosing cancer. Lung cancer can be accurately diagnosed using CT images whether it is Benign or Malignant. The disease can be detected as early as feasible utilizing machine learning and deep learning approaches that employ a variety of algorithms. To achieve better outcomes, three alternative models are used in this work. In this case, VGG16 gives superior outcomes and precision. As a result, deep learning approaches assist us in achieving better results.

REFERENCES