International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 03 | Mar 2020 www.irjet.net p-ISSN: 2395-0072

Lung Cancer Detection using GLCM and Convolutional Neural Network

Suresh Babu.P[1], Jennifer.J[2], Hesima.C[3] , Preethi.S[4]

[1]*Assistant Professor-III, Department of Information Technology, Velammal College of Engineering and Technology, Madurai, Tamil Nadu, India.*
[2][3][4]*Department of Information Technology, Velammal College of Engineering and Technology, Madurai, Tamil Nadu, India*

Abstract - *Cancer is one of the deadliest diseases leading to innumerable deaths worldwide. According to WHO (World Health Organization) lung cancer contributes about 14 per cent among all the cancers. Therefore, early detection and treatment is very much required. Computed Tomography (CT) scan can provide valuable information in the diagnosis of lung diseases. The main objective of this work is to classify the tumors found in lung as malignant or benign by means of Convolutional Neural Network(CNN). The accuracy obtained by means of CNN is 96%, which is more efficient when compared to accuracy obtained by the traditional neural network systems.*

***Key Words*- Computed Tomography(CT), Gray Level Co-occurence Matrix(GLCM), Convolution Neural Network (CNN).**

1. INTRODUCTION

Cancer is the most prevalent terminal disease globally, accounting for an estimated 9.6 million deaths in 2018. Of many types of cancers, lung cancer is one of the frequently occurring diseases that causes death and is identified in both genders over an estimated death of 1.76 million in 2018. Although surgery, radiation therapy, and chemotherapy have been used in the treatment of lung cancer, the five year survival rate for all stages combined is only 14%. This has not changed in the past three decades [1]. It is difficult to detect because it arises and shows symptoms in final stage. However, mortality rate and probability can be reduced by early detection and treatment of the disease. Best imaging technique CT imaging are reliable for lung cancer diagnosis because it can disclose every suspected and unsuspected lung cancer modules.

The proposed methodology has five phases for the classification of lung cancer. In phase one, the required data is collected from the database https://www.kaggle.com/datasets. In phase two removal of noise is done by using the Median filter. In phase three, the taken input images are segmented using K Means clustering. After segmentation, in phase four, features are extracted using GLCM (Gray Level Co-occurrence Matrix). These extracted features are used in phase five for classification purpose which is carried by CNN (Convolutional Neural Network)

2. METHODOLGY

To improve the detection of lung cancer in the CT images there are four main steps involved. At each step, various techniques are applied which resulted in different accuracies in detecting the lung cancer. Firstly, the lung CT image is pre-processed to remove any noises that exists in the image. Secondly, the image is segmented to get Region of Interest (ROI). Thirdly, feature extraction is applied to extract features like energy, entropy, variance. Finally, different classification algorithm is applied on the extracted features of the lung CT image.

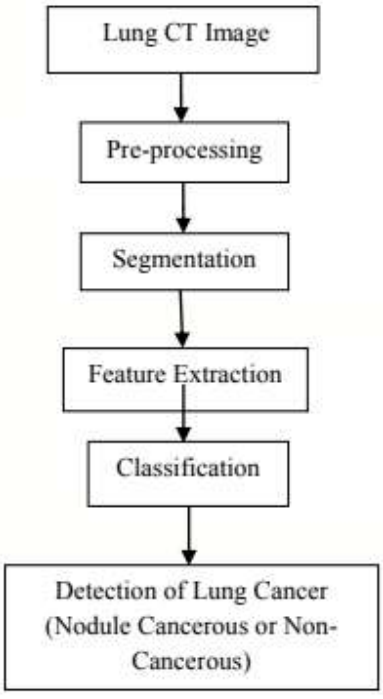

Fig 1: Steps involved to detect Lung Cancer

2.1. Input CT lung Images

The lung CT images having low noise when compared to scan image and MRI image. So we can take the CT images for detecting the lungs. The main advantage of the computer tomography image having better clarity, low noise and distortion. The mean and Variance can be easily calculated. The calculated value is very closer to the original value.

Fig.2 The Lung CT Image

2.2. Preprocessing

Image pre-processing is the first and important technique involved in lung cancer detection. Pre-processing technique is needed to improve the detection accuracy and to eliminate some regions of CT Image such as background and surrounding tissues or vessels. In preprocessing stage, the median filter is used to restore the image under test by minimizing the effects of the degradations during acquisition..The median filter simply replaces each pixel value with the median value of its neighbors including itself. Hence, the pixel values which are very different from their neighbors will be eliminated[2].

2.3. Segmentation

The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image Segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics [3]. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (edge detection). Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s).

2.4. Feature Extraction

The Image features Extraction stage is very important in our working in image processing techniques which using algorithms and techniques to detect and isolate various desired portions or shapes (features) of an image. After the segmentation is performed on lung region, the features can be obtained from it and the diagnosis rule can be designed to exactly detect the cancer nodules in the lungs. This diagnosis rules can eliminate the false detection of cancer nodules resulted in segmentation and provides better diagnosis. In the literature we found among the features used in the diagnostic indicators:

- Area of the interest
- Calcification
- Shape and Size of nodule
- Contrast Enhancement

2.5. Classification

Feature extraction techniques in image processing are used for extracting desired features from image like portions, shapes of an image. Normality and abnormality of an image can be determined in this stage. The detected features provide a basis for process of classification. Various features of an image can be area, perimeter, eccentricity, intensity, etc. Various feature extraction techniques which can be used are histogram of oriented gradients, local binary patterns, Gray-Level Co-Occurrence Matrix. The lung cancer can be detected in the classification stage by applying back-propagation, and in the pre-processing stage

3. RELATED WORKS

In this paper [4], input color images were first converted into grey scale images as processing of grey scale image is easier than that of the color images. histogram equalization was then applied to the images for obtaining the sharp borders which can help in further analysis as it highlights the borders thereby increasing the contrast of the images. thresholding was applied and features were extracted based on the pixel value. By applying Baack Propagation Network, an overall efficiency of 78 percent was achieved.

In this paper [5], detection of cancer was carried out with the help of ANN (Artificial Neural Network back- propagation). From the lung cancer database, a total of 50 images were considered and these images were divided into two groups-cancer and non-cancer images. The features were extracted

with the help of GLCM (Gray Level Co-occurrence Matrix) and these were given to ANN. The cancer was detected and 80% of accuracy was obtained.

In this paper [6], proposed methodology used CT (Computerized Tomography) scan images. The methodology has several steps such as pre-processing where the noise of the images is removed. The second step is image segmentation where over segmentation is removed using Marker Controlled Watershed Segmentation. The segmented image was then applied to Binarization and masking technique. Watershed achieved more accuracy when compared with Threshold which is 85.27 percent.

In this paper [7], two segmentation methods were used for early detection of lung cancer. The first method was Hopfield Neural Network (HNN) and the second method was Fuzzy C-Mean (FCM) algorithm. With the help of HNN, nuclei and cytoplasm regions were extracted successfully. HNN was preferred as FCM failed in extracting these features with accuracy.

This paper [8], proposed an automatic detection of lung cancer by suing CT scan images. While diagnosing lung cancer, lung nodule detection plays an important role. Therefore, FCN (Fully Convolutional Network) was selected for gaining more accuracy. First, FCN was used for segmentation and then lung nodules were detected. The proposed method was able to detect the lung nodules with 100 percent accuracy.

4. RESULT AND DISCUSSION

Phase 1: Here figure 3a and 3b represents the normal and abnormal images taken from the dataset for classification purpose. All the collected cancerous and non-cancerous images are CT Scan images.

Figure 3a: Normal Lung image

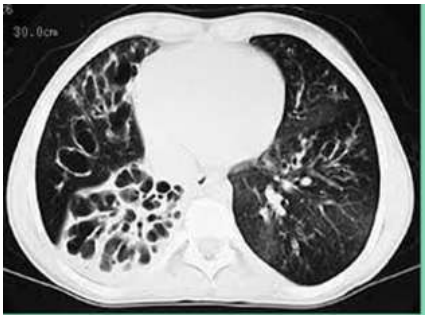

Figure 3b: Abnormal Lung image

Phase 2: Noise is removed successfully with the help of median filter while preserving all the edges carefully for classification purpose.
Phase 3: After filtering of image, segmentation of the image was done using K Means clustering. It was successfully applied and better results were shown.
Phase 4: The features of the input image are extracted by using GLCM. These are given to the classifier for the further classification.

The above image (as depicted in Figure 2b) when processed, provides the following output

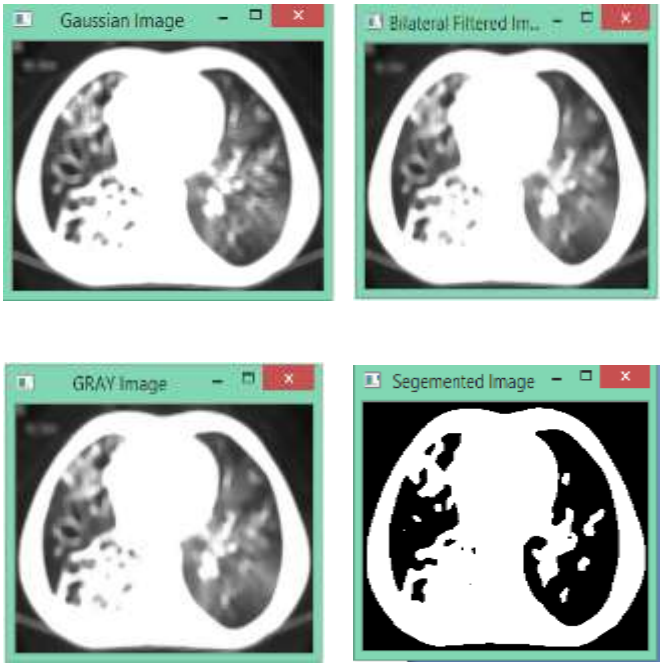

Figure 4a: Processes Involved in Lung Cancer Detection

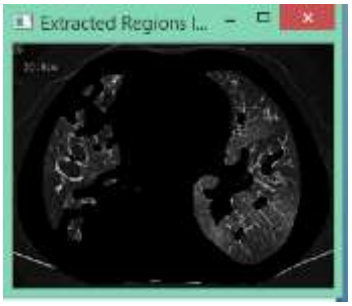

Figure 4b: Extracted Regions

Figure 4c: Accuracy of the Abnormal Lung Image

5. CONCLUSION AND FUTURE WORK

A convolutional neural network based system was implemented to detect the malignancy 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 96%.

In addition to deep convolutional network, the same dataset was classified by multilayer perceptron network Back propagation algorithm with using GLCM features. The results show only 93% accuracy [9].

In this proposed work, the specificity obtained is 100% which shows that that there is no false positive detection. Also, the accuracy, sensitivity and specificity of the proposed system is high when compared to previously available conventional neural network based systems.

In the near future, the system will be trained with large datasets to diagnose the type of cancer with its size and shape. The overall accuracy of the system can be improved using 3D Convolutional Neural Network and also by improving the hidden neurons with deep network.

REFERENCES

[1] Disha Sharma, Gagandeep Jindal, "Identifying Lung Cancer Using Image Processing Techniques", International Conference on Computational Techniques and Artificial Intelligence (ICCTAI'2011)

[2] M.S. Al-Tarawneh, "Lung cancer detection using image processing techniques," Leonardo Electronic Journal of Practices and Technologies, vol. 20, pp. 147– 58, May 2012

[3] Linda G. Shapiro and G.C. Stockman., Computer Vision: Theory and Applications. 2001: Prentice Hall.

[4] Kalaivani, S., Chatterjee, P., Juyal, S., & Gupta, R. (2017, April). Lung cancer detection using digital image processing and artificial neural networks. In 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA) (Vol. 2, pp. 100-103). IEEE. 9

[5] Anifah, L..., Haryanto, Harimurti, R.P., Permatasari, Z., Rusimamto, P.W., & Muhamad, A.R. (2017). Cancer lungs detection on CT scan image using artificial neural network backpropagation based gray level coocurrence matrices feature. 2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS).

[6] Patil, B. G., & Jain, S. N. (2014). Cancer cells detection using digital image processing methods. International Journal of Latest Trends in Engineering and Technology, 3(4), 45-49.

[7] Taher, F., & Sammouda, R. (2011, February). Lung cancer detection by using artificial neural network and fuzzy clustering methods. In 2011 IEEE GCC Conference and Exhibition (GCC) (pp. 295-298). IEEE.

[8] Chunran, Y., Yuanvuan, W., & Yi, G.X. (2018). Automatic Detection and Segmentation of Lung Nodule on CT Images. 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 1-6.

[9] A.Kavitha and P.Senthil," Design Model of Retiming Multiplier For FIR Filter &its Verification", International Journal of Pure and Applied Mathematics, Vol116 No12, 2017, pp. 239-247