International Journal of Research in Engineering, Science and Management
Volume 5, Issue 4, April 2022
https://www.ijresm.com | ISSN (Online): 2581-5792

Detection of Lung Cancer Using Deep Learning

Y. Nagateja[1*], G. Renukesh[2], L. Suhas Gowda[3], K. M. Shreyas[4], J. T. Thirurkrihshna[5]

[1,2,3,4]*B.E. Student, Department of Information Science and Engineering, Dayananda Sagar Academy of Technology and Management, Bangalore, India*
[5]*Associate Professor, Department of Information Science and Engineering, Dayananda Sagar Academy of Technology and Management, Bangalore, India*

***Abstract*: In recent years, many computer-aided diagnostic (CAD) systems have been developed to diagnose various diseases. Early detection of lung cancer is becoming increasingly important, made possible through image processing and deep learning technologies. Nodule detectors and feature-based classifiers make up the majority of CAD systems. Computed tomography can help doctors detect lung cancer early. In many cases, the diagnosis of lung cancer is based on the doctor's experience, which can lead to: some patients being overlooked and causing issues. In several fields of medical imaging diagnostics, deep learning has shown to be a popular and effective strategy. Two types of deep neural networks are used to classify lung cancer, such as CNN and DNN.**

***Keywords*: lung nodules, deep learning, CNN, computed tomography, deep neural network.**

1. Introduction

Lung cancer is a cancer type that begins in the lung. Lung Cancer is the most influential cause of cancer deaths around the world. According to the WHO, 96 million people died from cancer itself in 2018. Lung cancer results from the uncontrolled mass growth of abnormal cells that start in one or both lungs. According to a report published by the American Cancer Society, the highest mortality rate is only 18%. Lung cancer can be divided into small cell lung cancer and non-cell lung cancer. The number of patients with non-cell lung cancer is greater than the number of patients with small-cell lung cancer. Early detection of lung cancer is essential to prevent death and improve survival. A lung nodule is a small mass of tissue that can be cancerous or benign, also called malignant or benign tissue, benign tissues are most common non-cancerous and does not have much growth where malignant tissues grows very fast and can affect to the other body parts and are dangerous to health.

Fig. 1. Three different categories of lung nodules

Benign nodules are usually smooth, triangular in periphery, filled with fat and calcium, whereas 10 out of 10 malignant nodules have marginal variations in lobular shape, vascular convergence, cystic air morphology, pleural depression, vesicular transparency, and sub solid morphology. I see it. It is related to the size and growth of the node.

Three different categories of pulmonary nodules: benign, primary and metastatic malignancy.

There are several methods of diagnosing lung cancer. Currently, computed tomography (CT) is widely used to diagnose lung nodules, and MRI is being used to diagnose lung disease with the development of magnetic resonance imaging (MRI).

Deep learning-based systems with pattern recognition and computer vision play an important role in the current trend of high-precision nodule detection. A possible malicious method for this requirement is called Convolution Neural Network (CNN), a class of deep neural networks that can learn textures from available embedding images.

2. Related Work and Method

A. Dataset

The dataset used in this study was taken from a TCIA repository called the Lung Imaging Database Consortium and Imaging Database Resource Initiative (LIDCIDRI). These data included 1010 patient cases and 1018 chest CT scans obtained in dicom format. Four radiologists have annotated lung lesions according its size as nodules>=3mm, nodules< 3mm and non-nodules>=3mm. This also contains the labels of malignancy level of lung nodules. This dataset contains 4 levels of malignancy: 0 = unknown, 1= benign or non-malignant disease, 2 = malignant, primary lung cancer, 3 = malignant metastasis. Benign are the lung tissues which grow gradually and this growth stop at certain point. These tissues are commonly non-cancerous and does not affect seriously to health. Malignant tissue is cancerous and grows very quickly. Its main application is to detect the location of nodes in the lungs based on computed tomography and type classification.

B. Deep Learning Architectures

Many different types of Deep Learning Architectures are proposed in recent researches for image segmentation and object classification in images. Some architecture are also proposed for medical imaging disease diagnosis. Two

*Corresponding author: nagateja.2k@gmail.com

architectures among them are adapted and modified in this study. Convolution Neural Network is also known as CNN and ConvNet. CNN is useful for feature extraction and classification of objects in the image. This CNN is nothing but a stack of different layers. Convolution layer and pooling layer has the responsibility of feature extraction of objects provided to CNN where convolutional layer extract the features and pooling layer selects the important features from them also known as sub sampling of convolved features. There are two types of joins: max join and average join, but max join is widely used in so many researches where maximum value among the values in pooling window is selected as sampled feature from convolved features. We also come across segmentation which is used in the detecting the lung nodules.

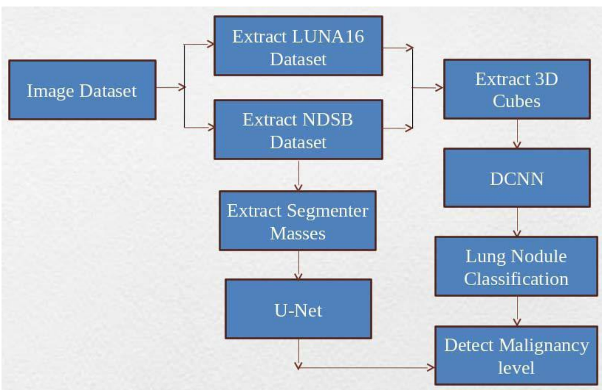

Fig. 2. Proposed approach

C. Lung Nodule Detection

FASTER R_CNN- Consists of 2 modules. The first module is the area proposal network, which can generate a suggested area for each image. The second module is the Fat RCNN detector, which classifies proposals in regions. In RPN, the convolution layer of the pretrained network is a 3x3 layer. Then two 1x1 layers are added for classification and regression. Anchors are introduced in RPN to handle different scales and aspect ratios of objects.

D. Feature Extraction

MODEL: Different models can be used. But here we use 3 CNN models. VGG16 and other two are

RESNET50 and RESNET100. RESNET makes it easier to train deep networks using residual learning frameworks.

Direction Using Faster R-CNN: When we use the MR images for training if the feature extraction model parameters are over fitting. First, we use a pre-trained model from natural images and then the fine tuning for initial and final parameters respectively. This is done by using the three-channel image, including the R, G, B channels. One way to do this is by using the gray level images in each channel and each channels are the same.

E. Nodule Detection Using DNN

Advanced medical images from the National Cancer Institute and University of Washington image databases are used for CAD analysis. This method is divided into two parts: nodule detection and feature-based classification. Nodule detection is performed in the following steps.

Preprocessing: The image used for this retrieval process is 512x512 pixels. Image intensities vary with the CT scan, and in addition to varying image intensities, CT images contain distortion and noise. This method is used to accurately identify nodules and is used to remove unnecessary blood vessels and tissues.

Image adjustment: For normal histogram equalization, every pixel is taken into account for the transformation, and it won't work if the pixel values are unevenly distributed. Adaptive histogram equalization is used to transform pixels according to a local cumulative distribution function. Used to restore the brightness value of an image.

Image segmentation: The Otsu threshold is used in the image segmentation process. Split up the image into two classes.

Edge detection: Image filtering is used to make edges sharper and smoother. The slope of the image intensity function is approximated using the Sobel operator. The Sobel operator uses two 3x3 kernels to detect edges.

Region of interest: The image has a lot of unnecessary information that does not contribute to the diagnostic procedure. Hansel et al. noted that nodules larger than 3.00 cm are likely to be malignant. Iwano et al also noted that the various forms must be malignant.

F. Traditional Methods for Nodule Classification

A brief overview of the subsections shows that various methods have been used to detect malignant neoplasms and classify lung nodules using traditional methods. Different detection methods using LUNA 16 data sets at different slice thicknesses. Node classification using linear SVM, logistic regression, kNearest Neighbors kNN and AdaBoost classifier.

The huge amount of data can effectively improve the accuracy of CNN training and testing by reducing the loss function and ultimately improving the reliability of the network. Image segmentation is a very good technique to improve network performance with little training data. Artificially generate training images using a variety of image processing operations, including translation, scaling, random rotation, flipping, and skewing. The size of the data set increases with transformations. Generate artificial training images for deep CNNs with random rotation and image processing operations.

G. Data Augmentation

The diameters of the pulmonary nodules are well known to vary. The size of the lung nodules was measured in order to determine their textural and size properties. The number of pulmonary nodules is uniformly fixed at 28. To begin with, the Binary imaging was used to obtain an image of the pulmonary nodules. Fully training a neural network requires a huge number of positive and negative samples. The image processing operations of translation, rotation, and scaling are discussed in this study. Before the image was submitted into the programme, it was given a rotation and a flip. The neural network, which increased the number of samples in the dataset image to be input The use of a large amount of sample data can be quite useful enhance the accuracy of training and testing the neural network reduces the loss function and consequently increases the reliability of the whole-body neural network.

H. Deep Neural Network

In major neural networks, DNN is an increase in the number of hidden nodes. Since each hidden layer may be a nonlinear conversion of the output layer, the neural network can be used to perform a more complex input calculation, and the deep neural network is higher than the "shallow" network. If the activation function is linear, if the depth of the hidden layer on the network does not improve the functions represented in one hidden layer, you must use nonlinear FX on each hidden layer. DNN disassembles sections of waste node processing so that you can use multiple network layers to obtain the characteristics of a closed cord of a variety of sizes. DNN has difficulty in spreading additional local extremities and gradients.

I. 3D CNN Architecture

A series of highly successful CNN models with very large training datasets of millions of different images have been used for the visual identification of 2D natural images. This model influenced the design of the 3D CNN architecture. We find it relatively simple compared to conventional systems, 3D CNN architectures, but still deep, due to the nature of the task, the dimensions of the cube, the size of the training data and the available GPU processing power. The input layer is 32x32x32. Three convolution layers are used with 32, 16 and 16 small 3x3x3 kernels respectively. A max pooling layer with a 2x2x2 overlapping window follows each convolution layer. Three fully connected layers are used, each with 64, 64, 2 neurons.

J. Vessel Filter

The lung parenchyma contains a vast number of tissue, polyps, veins, and other pollutants. Vascular morphology has recently been discovered in the lungs, which may affect lung polyp diagnosis. Cells and polyps appear in the same black pixels as bright tissues in CT scans of the lungs. The arteries, on the other hand, have a different structure and composition than the lung nodules. As ellipses, tiny circles, or cloth-like structures developed from lung nodules, tubular forms resembled tubes. The goal of this stage is to replace the venous networks in the lungs in such a way that the nodule formations may be examined more easily. Several advanced vascular algorithms, various filters and vascular filters have been proposed in recent years for Sato, Vascular Enhancement Diffusion (VED) and Vessel filters. The main application of these vascular filters is in the retina of the eye and they work great.

3. Literature Survey

Table 1
Comparison of table

Work	Database (samples)	Accuracy (%)	Sensitivity (%)	Specificity (%)
Nascimento et al. [21]	LIDC (73)	92.78	85.64	97.89
Orozco and Villegas [22]	NBIA-ELCAP (113)	N/A	96.15	52.17
Krewer et al. [7]	LIDC-IDRI (33)	90.91	85.71	94.74
Dandil et al. [23]	Private (128)	90.63	92.30	89.47
Parveen and Kavitha [24]	Private (3278)	N/A	91.38	89.56
Kuruvilla and Gunavathi, 2014 [6]	LIDC (110)	93.30	91.40	100
Gupta and Tiwari [25]	Private (120)	90	86.66	93.33
Hua et al. [10]	LIDC (2545)	N/A	73.30	78.70
Kumar et al. [8]	LIDC (4323)	75.01	83.35	N/A
da Silva [26]	LIDC-IDRI (8296)	82.3	79.4	83.8
CNN (this paper)	LIDC-IDRI (5024)	84.15%	83.96%	84.32%
DNN (this paper)	LIDC-IDRI (5024)	82.37%	80.66%	83.9%
SAE (this paper)	LIDC-IDRI (5024)	82.59%	83.96%	81.35%

4. Conclusion

Faster RCNN is designed for lung nodule detection with optimized parameters, three-channel spatial input, and transfer learning. This detection method avoids candidate selection and is less scale dependent. Imaging techniques and deep neural networks are used to detect nodules in the lungs on CT scans. First, AHE was applied to clarify latent image information. Second, we convert the image to a binary image using the Otsu threshold method. Transferable Texture CNN architecture for lung cancer classification problem. We have introduced EL to remove general form information and explore texture capabilities. The experiment shows that the proposed method is successful for positive and malignant categories that do not require complicated pretreatment or complex pretreatment. Neural networks have been used to detect the closing of a neural network to consider the current trend and the difficulties of the future. As a result, the optimal approach to the whole of lung cancer is still that this review training paper is available for researchers and professionals. Patients with lung cancer are evaluated.

References

[1] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay & A. Jemal, Global Cancer Statistics American Cancer Society. Cancer Facts & Figures 2019, Amer. Cancer Soc., Atlanta, GA, USA, 2019.

[2] A. M. Santos, A. O. de Carvalho Filho, A. C. Silva, A. C. de Paiva, R. A. Nunes, and M. Gattass, “Automatic detection of small lung nodules in 3D CT data using Gaussian mixture models, Tsallis entropy and SVM,” Eng. Appl. Artif. Intell., vol. 36, pp. 27-39, Nov. 2014

[3] G. A. P. Singh and P. Gupta, ‘‘Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans,’’ Neural Comput. Appl., vol. 31, no. 10, pp. 6863–6877, 2018.

[4] N. Nasrullah, J. Sang, M. S. Alam, M. Mateen, B. Cai, and H. Hu, “Automated lung nodule detection and classification using deep learning combined with multiple strategies,” Sensors, vol. 19, no. 17, pp. 3722, Aug. 2019.

[5] R. Shojaii, J. Alirezaie and P. Babyn, "Automatic lung segmentation in CT images using watershed transform," IEEE International Conference on Image Processing 2005, Genova, 2005, pp. II-1270.

[6] D. R. Aberle, A.M. Adams, et al., “Reduced lung-cancer mortality with low-dose computed tomographic screening.,” The New England journal of medicine, vol. 365, no. 5, pp. 395–409, 2011.

[7] K. Murphy, B. van Ginneken, et al., “A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification.,” Medical image analysis, vol. 13, no. 5, pp. 757–70, Oct. 2009.

[8] T.W. Way, B. Sahiner, et al., “Computer-aided diagnosis of pulmonary nodules on CT scans: improvement of classification performance with nodule surface features,” Medical physics, vol. 36, no. 7, pp. 3086–3098, 2009.