International Journal of Future Generation Communication and Networking
Vol. 13, No.3, (2020), pp. 2273–2279

Detection of Various Lung Diseases using Machine Learning

Prof. P.S. Sadaphule [1], Pushkar Ghogare [2], Harshal Divekar [3], Sarthak Ighe [4], Hrishikesh Lokhande [5]

AISSMS IOIT, SPPU, Pune

[1]*p.s.sadaphule@gmail.com*, [2]*pushkarghogare@gmail.com*,
[3]*harshaldivekar7@gmail.com*,
[4]*sarthakighe93@gmail.com*,
[5]*hrishikeshlokhande23@gmail.com*

Abstract

There are many applications of machine learning in various fields among which diagnosis of diseases in early stages proves to be one of the important application in the medical field. Image processing techniques plays important role in various disease detection. Among the various diseases, lung diseases are most harmful diseases with more than 225,000 peoples getting affected with it and causing death of almost 150,000 peoples. Detection of lung diseases at a very early stage can be useful specially to improve the survival rate of a patient. The aim of this research is to detect bronchiectasis, lung cancer, hanta virus and pneumonia using CT scan image. In this research a quick and well balanced system of lung disease detection is proposed. In contradiction to previous studies, fully connected network requires more number of neurons in first hidden layer of artificial neural network while CNN algorithm requires less number of neurons.

Keywords: *Machine Learning, Image Processing, Feature Extraction, CNN.*

1. Introduction

Nowadays, one of the mostly used techniques in medical fields for optimized and earlier detection of diseases is image processing technique, so that the treatment can be started at a very earlier stage. Lung diseases are the most fatal type among all the diseases national wide. Lung diseases are caused due to unrestrained cell growth in the tissue of lungs. Lung diseases are categorized mostly as small cell lung malignancy (SCLC) and non-small cell lung tumors (NSCLC). Diagnosis of diseases is mostly based on Computed Tomography (CT) Scan images. By the use of image processing techniques on CT scan images of lungs different types of lung disease can be identified. Before processing an image, it must be broken down a pixel. There are 4 components present in a pixel
i.e blue, green, red and alpha. Transparency of a pixel is determined by alpha while RGB determines color of the pixel. There is a value assigned between 0 to 255 for each and every component. If a component is missing the value assigned is 0 and if the component is fully present then the value is assigned as 255. The values 0 to 255 can be represented using 8 bits. Therefore we need 32 bits to starting pixel of image is at coordinate (0,0).

The proposed algorithm solves the problem of fully connected network. CNN algorithm can identify the infected cell is much faster time. If no affected cell is found in the test image then the report of a patient generated as normal.

ISSN: 2233-7857 IJFGCN

2. Literature Survey

[1].Recently, CNN has become superior approach in field of medical science for image analysis. Regarding detection of lung diseases, methods proposed so far are based mainly on radiology. The image based features which are related to the survival of patient are taken out from CT Scans. To study various distinct features of lung nodules training is given to the computers via analysing CT scans lung nodules.

[2]. To remove diagnostic errors by manual inspection of the specimens tends to be difficult, inaccurate and requires a skilled person. From the research, the local binary pattern provides much better performance rather than basic textural patterns.

[3]. Among these researches Tiantian Fang proposed a precise, quick, and well balanced system based on deep learning techniques which can detect lung cancer. A CNN structure similar to that of GoogLeNet was developed using a transfer learning approach.

[4]. In 1960, Lusted L.B. stated that analyzing and detecting the normal and abnormal images of the lungs could be carried out automatically. Hence analysis of image in medical areas with the help of computers was initiated.

Latter, Dr. K Manjunath, Avinash.S, Dr.J.Senthilkumar introduced a system based on gabor filter and watershed segmentation techniques to detect lung cancer. [5]. The system proposes, detection and prediction of the lung cancer and gives precise outcome. SVM is used to classify a set of textural features taken out from the differentiated ROIs.

[6]. The system focuses on the technique of identification of lung cancer by breath. The sole motive of the breath detection system was that it helped the physicians in quickly screening of lung cancer. KNN and SVM was used to analyze.

[7]. To predict malignant an EDM machine learning algorithm with vectorized histogram features was used to detect SCLC.

[8]. By analysis of different human lung images at different levels, cancer can be detected early by the system proposed by development of an automatic CAD system.

[9]. Histograms with vectors and an algorithm of EDM machine learning along with it are used to detect SCLC for early prediction of malicious cancer.

[10]. A methodology was proposed to build an automatic system to detect disease by using morphological operators to extract ROI.

3. Proposed Methodology

The figure below depicts the proposed system. The goal of image processing is to optimize the quality of image and find patterns. CNN algorithm is used for disease classification. Figure 2 shows the original CT scan image.

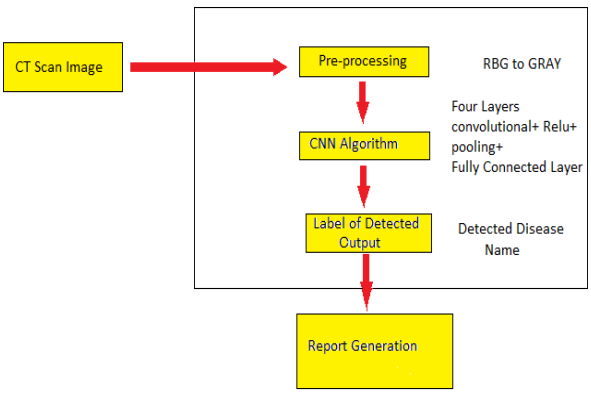


Figure 1: System Design

Pre-processing

This is the first step in which image de-noising algorithms are considered such as grayscale, thresholding, edge detection.

Converting color image into grayscale:

1. Take RGB value of the required pixel.
2. Find the mean RGB value of the pixel.

   Average= (R+G+B)/3

3. Replace the R, G and B value of pixel with mean.

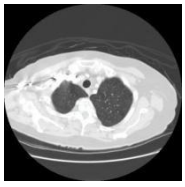
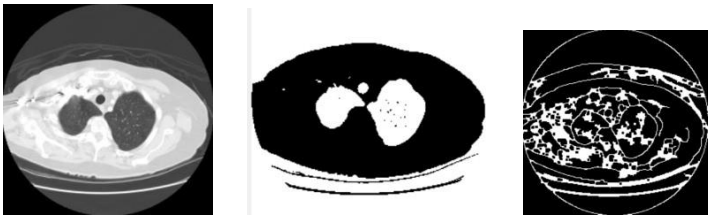


Fig 2 : Original Image



Gray Scale Threshold Edge Detection Figure 3: Pre-processed Image

The next step includes feature extraction in which we focus only on region of interest i.e. tissues of lungs which are needed for detection of any abnormal behavioral.

Convolutional Neural Network (CNN)

In the CNN, neuron in a particular layer will be connected to a small region of layer before it, instead of all neurons before it, as in fully connected network. The CNN algorithm takes the pre-processed image as an input and it passes this image through four layers of CNN which are explained below.

-1	-1	-1	-1	-1	-1
-1	1	1	1	1	-1
1	-1	1	1	-1	1
1	-1	1	1	-1	1
-1	1	1	1	1	-1
-1	-1	-1	-1	-1	-1

Fig 4: Input Image Matrix

i. Convolutional layer

CNN uses small piece of blocks usually known as patterns and compare it with test image. We are comparing pattern on input image whether it is matching or not. We are multiplying pixel values of pattern and test image block to get resultant matrix.

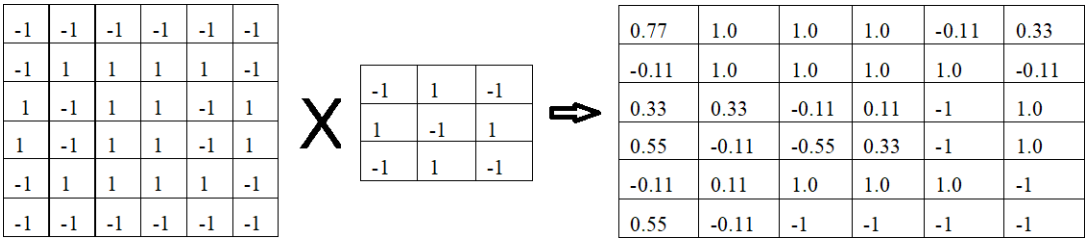


Figure 5: Convolutional layer

ii. ReLU layer

Our primary objective of using this layer is to remove negative values from output which we got from convolutional layer. Here we are replacing negative values with 0. A node is activated in a Rectified Linear Unit (ReLU) function if the input is above a certain quantity, if the input is less than zero the output is zero. When there is a rises in input above a certain threshold, their exists a linear relationship with dependent variable

.

x	**f(x)=X**	**F(x)**
-1	f(-1)=0	0
1	f(1)=1	1

Figure 6: ReLU layer

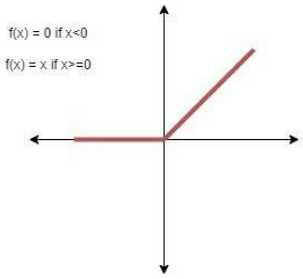


Figure 7: ReLU function

iii. Pooling layer

This is the third step of CNN algorithm in which we are reducing size of the image.

0.77	1.0	1.0	1.0	-0.11	0.33
-0.11	1.0	1.0	1.0	1.0	-0.11
0.33	0.33	-0.11	0.11	-1	1.0
0.55	-0.11	-0.55	0.33	-1	1.0
-0.11	0.11	1.0	1.0	1.0	-1
0.55	-0.11	-1	-1	-1	-1

1.0	0.33	0.55
0.33	1.0	0.33
0.55	0.55	1.0

Figure 8: Pooling layer

iv. Fully connected layer

Actual classification of diseases is carried out in this layer. Here the output of pooling layer is converted into a single list.

1.0	0.33	0.55
0.33	1.0	0.33
0.55	0.55	1.0

1.0
0.33
0.55
0.33
1.0
0.33
0.55
0.55
1.0

Figure 9: Fully connected layer

4. Expected Outcome

Based on the examination of lung nodules from input CT scan image, the image

is classified as normal or abnormal. If the image is detected as abnormal the type of disease the patient is infected with will be determined.

5. Conclusion

The manual analysis of the sample is tedious and requires intensive trained person to eliminate diagnostic error. Hence, to overcome this, a system is proposed to avoid the manual errors and increase the accuracy. It can be concluded that the system proposes the detection of lung diseases with the help of CT scan images which are being segregated using image pre-processing techniques such as gray scale imaging, thresholding and edge detection. Further these pre-processed images were trained and tested using the CNN algorithm to detect the lung diseases based on the lung nodules.

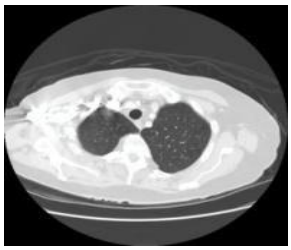


Fig 10: Lung Cancer Image.

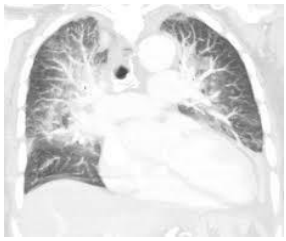


Fig 11: Lungs Infected with Hanta Virus.

Fig 12: Lungs Infected with Pneumonia.

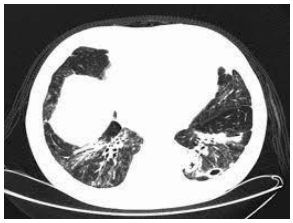


Fig 13:Lungs Infected with Bronchiectasis.

6. References

[1] Matko Sari, Mladen Russo, Maja Stella, Marjan Sikora, Marjan Sikora , “CNN-based Method for Lung Cancer Detection in Whole Slide Histopathology

Images", in *4th International*
Conference on Smart and Sustainable Technologies (SpliTech), 2019.

[2] Sanjukta Rani Jena , Thomas George , Narain Ponraj, "Texture Analysis Based Feature Extraction and Classification of lung cancer", *in [IEEE International Conference on Electrical,*
Computer and Communication Technologies (ICECCT),2019.

[3] Tiantian Fang , "A Novel Computer-Aided Lung Cancer Detection Method Based on Transfer Learning from GoogLeNet and Median Intensity Projections" , in *IEEE International*
Conference on Computer and Communication Engineering Technology (CCET), 2018.

[4] Ms. Twinkal Patel, Asst. Professor Mr. Vimal Nayak, "Hybrid Approach For Feature Extraction of Lung Cancer Detection", in *Second International Conference on Inventive*
Communication and Computational Technologies (ICICCT), 2018.

[5] Janee Alam , Sabrina Alam , Alamgir Hossan, "Multi-Stage Lung Cancer Detection and Prediction Using Multi-class SVM Classifie", *in International Conference on Computer,*
Communication, Chemical, Material and Electronic Engineering (IC4ME2) , 2018.

[6] De-Ming Wong , Chen-Yu Fang, Li-Ying Chen , Chen-I Chiu , Ting-I Chou , Cheng-Chun Wu , Shih-Wen Chiu , Kea-Tiong Tang , "Development of a breath detection method based Enose
system for lung cancer identification", *in IEEE International Conference on Applied System Invention (ICASI)*, 2018.

[7] Moffy Vas1 , Amita Dessai, "Lung cancer detection system using lung CT image processing" in *International Conference on Computing, Communication, Control and Automation*
(ICCUBEA), 2017.

[8] Rachid Sammouda, "Segmentation and Analysis of CT Chest Images for Early Lung Cancer Detection" in *Global Summit on Computer & Information Technology (GSCIT)*, 2016.

[9] Qing Wu and Wenbing Zhao, "Small-Cell Lung Cancer Detection Using a Supervised Machine Learning Algorithm" , in *International Symposium on Computer Science and Intelligent Controls (ISCSIC)* ,2017.

[10] K.Punithavathy, M.M.Ramya, Sumathi Poobal , "Analysis of Statistical Texture Features for Automatic Lung Cancer Detection in PET/CT Images", *in International Conference on Robotics, Automation, Control and Embedded Systems (RACE)* , 2015.