IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Machine Learning System to Detect Lung Infection Using CT Scan

[1]Er Dipika N. Gumble, [2]Dr R. D. Ghongade
[1]M.E.Student, [2]H.O.D.of Electronics Department
[1]Department of Electronics and Telecommunications, Amravati University
[1]P.R.Pote (Patil) College of Engineering, Amravati (M.S), India

ABSTRACT

Lung diseases and its infections are the disorders that affect the lungs, the organs that allow us to breathe and it is the most common medical conditions worldwide especially in India. Identification and grading of disease in a human organ demand expert's opinion and the patients' medications are completely dependent on the results given by that expert. However, there might be situations where experts may not be available or too busy. To tackle the emergencies, which arise due to lack of experts, in this work, we will design an algorithm that inputs an image (MRI or PET or CT one among these) to classify whether the lungs is infectious or non- infectious. If a lung is infectious, we try to identify which disease based on the symptoms and figure out the amount of degradation that the lung has gone through due to that disease. For this purpose, we rely on the image and pattern recognition with machine learning system algorithms. The purpose of the work is to review the work carried out by various researchers for detecting the lung infection in CT scan images.

Keywords: Lung infection detection, CT scan, Machine Learning, Respiration

INTRODUCTION

The term lung disease refers to many disorders affecting the lungs such as asthma, chronic obstructive pulmonary (COPD) disease, infections such as tuberculosis, influenza, lung cancer, pneumonia and other breathing problems. Lung diseases signs and symptoms can differ by the type of the affected disease. Common signs are trouble in breathing, shortness of breath, feeling like you're not getting enough air, decreased ability to exercise, a cough that won't go away, coughing up blood or mucus, pain or discomfort when breathing in or out. Medical image analysis and process requires an environment for data access, data analysis, processing, revelation and algorithm development. Medical imaging is the technique and process used to create images of the human body for clinical purposes for diagnosis and analysis or medical science (including the study of disease of normal anatomy and physiology) [2][3][5].

In the recent years, medical CT Images have been applied in clinical diagnosis widely. It assists physicians to detect and locate pathological changes with more accuracy. Computed tomography images can be distinguished for different tissues according to their different grey levels.

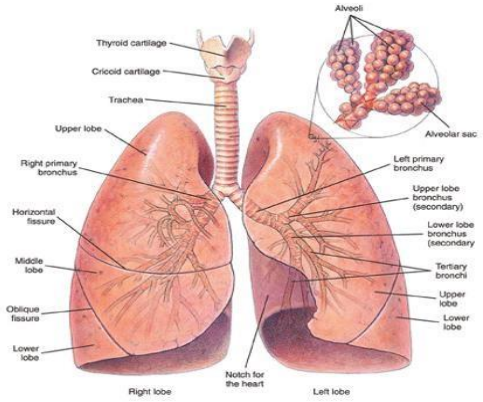


Fig 1: Healthy Lungs of Human

Lung diseases can be caused by infection, an exposure at the workplace, medications and various disorders. X-ray chest radiography and computer tomography (CT) are two common anatomic imaging modalities that are routinely used in the detection and diagnosis of a variety of lung diseases. Medical images play a vital role in patient diagnosis, therapy, surgical, medical reference, and training. The Digital Imaging and Communications in Medicine (DICOM) standard allows storing textual descriptions, known as metadata, along with the images. It was the most important breakthrough since the discovery of the X-rays, and CT has remained a cornerstone of diagnostic radiology throughout the years [10][12][14]. The respiratory tract infection due to the Coronavirus Disease (COVID-19) is emerged as one of the major threats globally due to its acuteness and the infection rate. It is one of the major communicable infectious diseases caused by Severe Acute Respiratory Syndrome-Corona Virus2 (SARS-CoV-2) and according to a recent report [7][8], it affected a larger human community, irrespective of their race and gender. The infection caused by COVID-19 severely affects the respiratory system by causing the severe pneumonia. Due to its harshness and the spreading rate, the World Health Organization (WHO) recently announced it as pandemic [9]. Even though various controlling and treatment procedures are implemented from December 2019 to till date, the mortality due to COVID-19 infection is rapidly increasing.

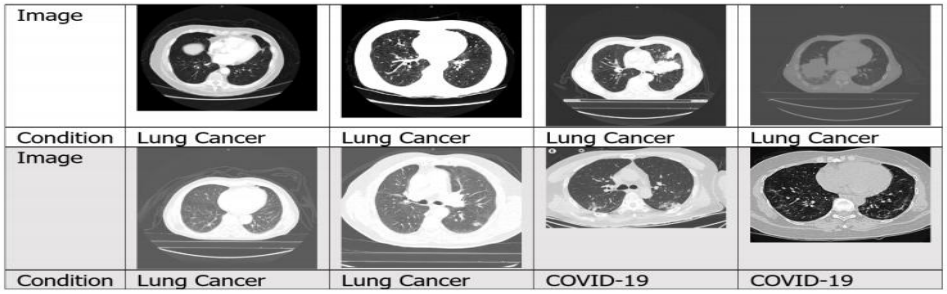


Fig 2: Examples of CT Scan Images

In this article, an automated approach for classification of the lung infections using CT images is reviewed. This review helps researchers as the assistant in detecting the lung diseases in the medical field.

LITRATURE REVIEW

Zhao Wang et.al. [1] proposes a novel joint learning framework to perform accurate COVID- 19 identification by effectively learning with heterogeneous datasets with distribution discrepancy. They build a powerful backbone by redesigning the recently proposed COVID-Net in aspects of network architecture and learning strategy to improve the prediction accuracy and learning efficiency. Muskan Lawania et.al. [2] proposes a productive strategy to identify the lung malignancy and its stages effectively and furthermore means to have progressively precise outcomes by utilizing KNN and Image Processing systems. Seifedine Kadry et.al. [3] propose a Machine-Learning-System (MLS) to detect the COVID-19 infection using the CT scan Slices (CTS). This MLS implements a sequence of methods, such as multi-thresholding, image separation using threshold filter, feature-extraction, feature-selection, feature-fusion and classification. Arnab Kumar Mishra et.al. [4] explored various Deep CNN based approaches are explored for detecting the presence of COVID19 from chest CT images. A decision fusion-based approach is also proposed, which combines predictions from multiple individual models, to produce a final prediction. Ilker Ozsahin et.al. [5] review the diagnosis of COVID-19 by using chest CT toward AI. We searched ArXiv, MedRxiv, and Google Scholar using the terms “deep learning”, “neural networks”, “COVID-19”, and “chest CT”.

Andreas Christe et.al. [6] proposed CAD system for dichotomous classification into the group needing further intervention and the group without the need for further workups was as good as that of the radiologists. Preeti Katiyar et.al. [7] proposed survey various segmentation, feature extraction and classification techniques are considered such as Artificial Neural Network, Convolutional Neural Network, SVM, Gray level co-occurrence matrix, Discrete wavelet transform and many more. Tulin Ozturk et.al. [8] presented new model for automatic COVID-19 detection using raw chest X-ray images is presented. Varalakshmi Perumal et.al. [9] the transfer learning technique has been applied to clinical images of different types of pulmonary diseases, including COVID-19. It is found that COVID-19 is very much similar to pneumonia lung disease. Li et.al. [10] designed and evaluated a three-dimensional deep learning model for detecting coronavirus disease 2019 (COVID-19) from chest CT scans. a 3D deep learning framework was proposed for the detection of COVID-19.

PROPOSED METHODOLOGY

The entire working process of the presented method is shown in Fig. 3. As shown in figure, the presented model consists a series of processes which are discussed in the figure.

Lung CT Scan Images: We are applying our methodology to standard benchmark image database is applied and observed under several aspects. The dataset has a collection of lung cancer screening CT images that can be employed for the design. The input CT scan lung images are provided as input to the presented model.

Pre-processing and Segmentation: In the initial stage, the pre-processing takes place by the use of median filter and Gaussian filter. Once the input image is pre-processed, the segmentation of images will take place by morphological segmentation algorithm which produces the output as segmented image in the binary form.

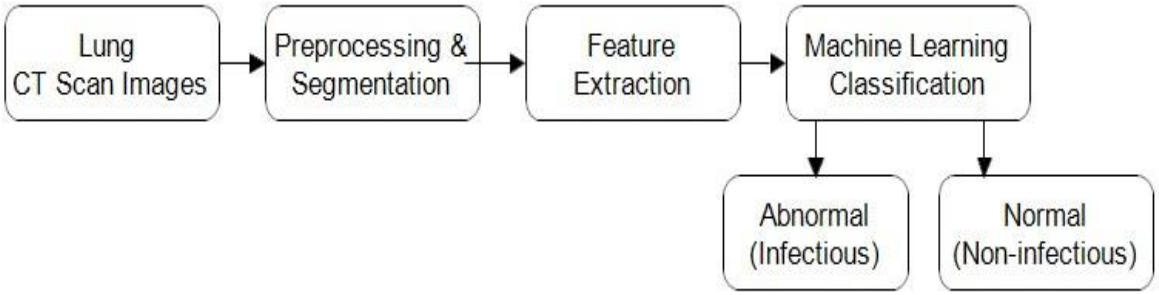


Fig 3: Proposed methodology of system

Feature Extraction: At the next stage, a collection of important features gets extracted from the segmented image namely image texture using Haralick feature and image HU moment. centroid, area, perimeter, eccentricity, pixel mean intensity and diameter.

Machine Learning Classification: Then, the classification of images will be carried out using machine learning classification model which finally provides the output as classified image into 'normal' or 'abnormal'. Later if abnormal image is found then only it will show and detect the lung nodule abnormal section. Training and testing of computer aided diagnosis models for detecting and diagnosing lung cancer will be done in proposed approach. The process of extracting features takes place using image processing and classifier operation is carried out utilizing machine learning which helps to develop the trained prediction approaches from the filtered features in an easier way and rapid way.

In the training phase of the input images taken from a given location to extract input features and the known output will be found by naming the images from the type of haemorrhage. Then the net file can be generated using a train tool for the first time after going through few testing iterations by providing the saved input and output files. Once the input features are calculated and the vector is created, to add the image to train, the output will be defined according to the value that has been received as the output result. Once the input and output files are saved, system can be trained with them. This logic can be used to train the tested images as well. We have using machine learning algorithm and ensemble learning algorithm

RESULT AND DISCUSSION

In our proposed approach, CT scan image will be given to system and our designed algorithm will identify the lung infection in image and classify the infection in the lung to covid 19, tuberculosis and normal. The results will be justified using graphs and suitable parameters.

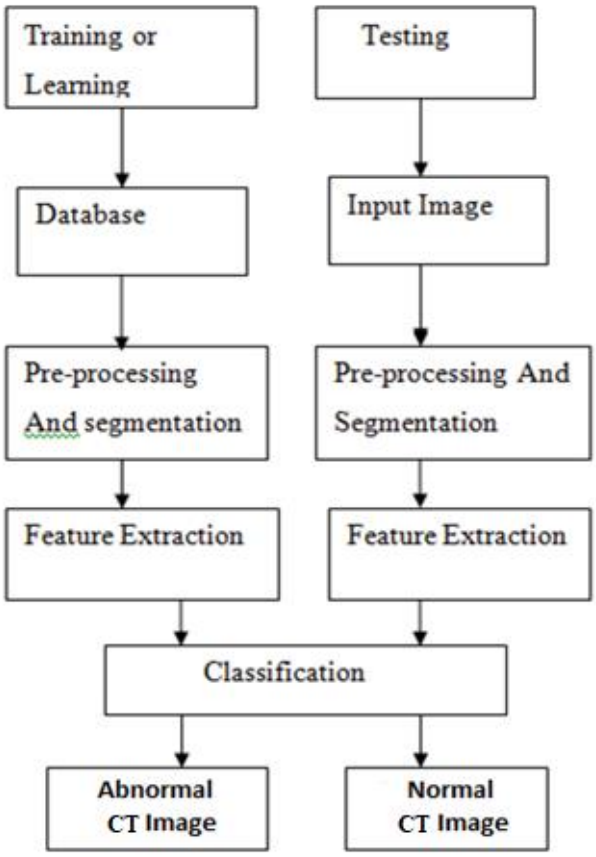


Fig 4: Flowchart

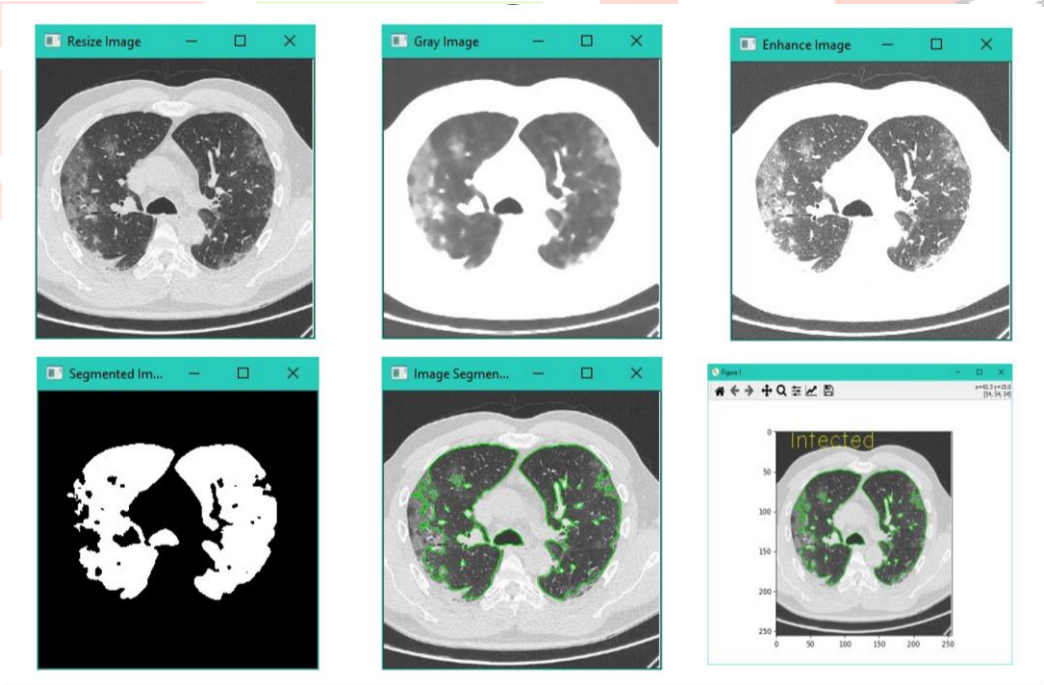


Fig 5: Testing Phase

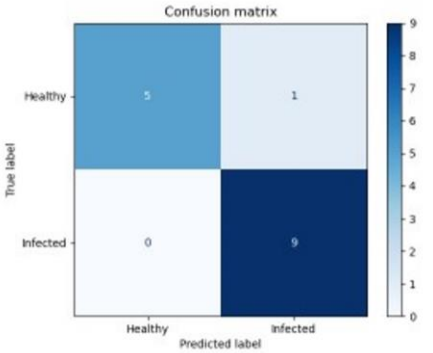
Testing Phase (SVM)

```
Testing features: (26,)
Support Vector Machine Classifier
[2]
Elapsed time for testing: 19.106734 seconds.

Abnormal or Infected
Area Infected -  61.29032258064516  %
```

Testing Phase (RDF)

```
Testing features: (26,)
Random Decision Forest Classifier
[2]
Elapsed time for testing: 12.251541 seconds.

Abnormal or Infected
Area Infected -  61.29032258064516  %
```

```
Testing features: (15, 26)
Testing labels: (15,)

Accuracy Score of Project Testing = 0.9333333333333333

Elapsed time for testing: 15.330322 seconds.


Confusion Matrix
[[5 1]
 [0 9]]

              precision    recall  f1-score   support

     Healthy       1.00      0.83      0.91         6
    Infected       0.90      1.00      0.95         9

    accuracy                           0.93        15
   macro avg       0.95      0.92      0.93        15
weighted avg       0.94      0.93      0.93        15
```



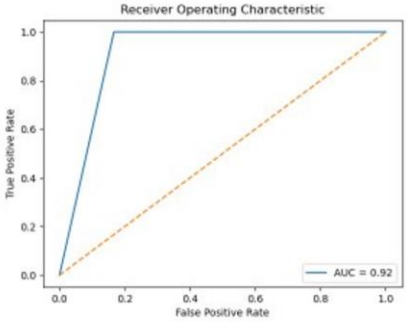
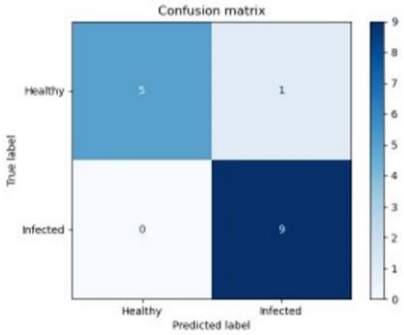


Figure 6: Result of Testing Performance (SVM)

```
Testing features: (15, 26)
Testing labels: (15,)

Accuracy Score of Project Testing = 0.8666666666666667

Elapsed time for testing: 6.979851 seconds.


Confusion Matrix
[[5 1]
 [1 8]]

              precision    recall  f1-score   support

     Healthy       0.83      0.83      0.83         6
    Infected       0.89      0.89      0.89         9

    accuracy                           0.87        15
   macro avg       0.86      0.86      0.86        15
weighted avg       0.87      0.87      0.87        15
```



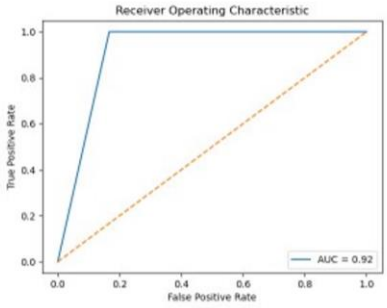


Figure 7: Result of Testing Performance (RDF)

ADVANTAGES AND APPLICATIONS

Quick and versatile recognition of respiratory nodules in chest CT checks. Proposed morphology-based nodule recognition approach is quick, versatile, and easy to comprehend and furthermore section the malignancy zone alone which prompts give right yield. Our methodology indicated a practically identical discovery ability however a lower computational expense.

CONCLUSION AND FUTURE SCOPE

Our proposed system efficiently categorizes between infected or healthy CT scan images. From experimental discussion by comparing performance among SVM and RDF, SVM is superior. Future implementation – deep learning (required very large dataset and Higher GPU), can add different types of Infected Lung Diseases or can classify among various disease categories

REFERENCES

[1] Wang, Zhao & Liu, Quande & Dou, Qi. (2020). Contrastive Cross-site Learning with Redesigned Net for COVID-19 CT Classification.

[2] R. Yamini, Muskan Lawania, Aashna Bansal. MORPHOLOGY BASED LUNG NODULE DETECTION USING MACHINE LEARNING. JCR. 2020; 7(4): 237-242. doi:10.31838/jcr.07.04.44

[3] Kadry, Seifedine & Rajinikanth, Venkatesan & Rho, Seungmin & Raja, N. & Rao, Vaddi & Krishnan, Palani. (2020). Development of a Machine-Learning System to Classify Lung CT Scan Images into Normal/COVID-19 Class.

[4] Mishra, Arnab & Das, Sujit & Roy, Pinki & Bandyopadhyay, Sivaji. (2020). Identifying COVID19 from Chest CT Images: A Deep Convolutional Neural Networks Based Approach. Journal of Healthcare Engineering. 2020. 1-7. 10.1155/2020/8843664.

[5] Ilker Ozsahin, Boran Sekeroglu, Musa Sani Musa, Mubarak Taiwo Mustapha, Dilber Uzun Ozsahin, "Review on Diagnosis of COVID-19 from Chest CT Images Using Artificial Intelligence", Computational and Mathematical Methods in Medicine, vol. 2020, Article ID 9756518, 10 pages, 2020. https://doi.org/10.1155/2020/9756518

[6] Christe, Andreas & Peters, etc. Computer-Aided Diagnosis of Pulmonary Fibrosis Using Deep Learning and CT Images. Investigative Radiology. 54. 1. 10.1097/RLI.0000000000000574.

[7] P. Katiyar and K. Singh, "A Comparative study of Lung Cancer Detection and Classification approaches in CT images," 2020 7th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2020, pp. 135-142, doi: 10.1109/SPIN48934.2020.9071240.

[8] Aboughazala, Laila & Mohammed, Kamel K.. (2020). Automated Detection of Covid-19 Coronavirus Cases Using Deep Neural Networks with X-ray Images. Al-Azhar University Journal of Virus Researches and Studies. 10.21608/aujv.2020.106722.

[9] Perumal, Varalakshmi & Narayanan, Vasumathi & Rajasekar, Sakthi. (2020). Detection of COVID-19 using CXR and CT images using Transfer Learning and Haralick features. Applied Intelligence. 10.1007/s10489-020-01831-z.

[10] Li L, Qin L, etc. Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT Aug;296(2): E65-E71. doi: 10.1148/radiol.2020200905. Epub 2020 Mar 19. PMID: 32191588; PMCID: PMC7233473.

[11] Ramalho, Geraldo & Filho, Pedro Pedrosa & Medeiros, F.N.s & Cortez, Paulo. (2014). Lung Disease Detection Using Feature Extraction and Extreme Learning Machine. RBEB. 30. 207. 10.1590/rbeb.2014.019.

[12] Hussain, Lal & Nguyen, etc. Machine-learning classification of texture features of portable chest X-ray accurately classifies COVID-19 lung infection. BioMedical Engineering OnLine. 19. 10.1186/s12938-020-00831-x.

[13] X. Gao, R. Comley and M. H. -M. Khan, "An Enhanced Deep Learning Architecture for Classification of Tuberculosis Types from CT Lung Images," 2486-2490, doi: 10.1109/ICIP40778.2020.9190815.

[14] Harmon SA, Sanford TH, Xu S, Turkbey EB, etc. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. Nat Commun. 2020 Aug 14;11(1):4080. doi: 10.1038/s41467-02017971-2. PMID: 32796848; PMCID: PMC7429815.