JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

LUNG DISEASE DETECTION METHODOLOGY ADOPTED (During Covid-19)

ABOUT THE AUTHORS

RITU GILLis a student of Computer Science and Engineering at DPG Institute of Technology and Management. She is currently doing her master's from this institute and Degree is about to complete in July 2021.Her research area is in artificial intelligence and machine learning .Before her master's she is having a bachelor's degree in Computer Science and Engineering.

Ms. ARCHANA ROHILLA is working as Assistant professor in CSE department of DPG Institute of Technology and Management, Gurugram. Her research area is Artificial Intelligence. She has more than 10 years of teaching experience. Her interest is in Cyber Security and Artificial Intelligence. She completed her master's and bachelor's degree in Computer Science and Engineering from MD University.

Mr. YASH DHANKAR is an Assistant professor in CSE department of DPG Institute of Technology and Management, Gurugram. His research work includes data mining and text summarization etc. He has published papers in UGC recognized journals. His research interest areas include Recommender systems sentiment analysis and Internet of Things. He has done B.Tech and M.Tech in Computer Science and Technology. He has an excellent academic record.

ABSTRACT

The platform will enable its users and professional diagnostic centres to upload their chest radiographs (x-rays) and get accurate predictions based on those. Chest radiography hasimportant clinical value in the diagnosis of diseases."Chest-X ray (CXR) radiography can be used as a first-line triage process for non-COVID-19 patients with pneumonia." However, the similarity between features of CXR images of COVID-19 and pneumonia caused by other infections makes the differential diagnosis by radiologists challenging. We hypothesized that machine learning-based classifiers can reliably distinguish the CXR images of COVID-19 patients from other forms of pneumonia. We useda dimensionality reduction method to generatea set of optimal features of CXR images to build an efficient machine learning classifier that can distinguish COVID-19 cases from non-COVID-19 caseswith high accuracy and sensitivity. By using global features ofthe whole CXR images, we successfullyimplemented our classifier using a relatively small dataset of CXR images. We proposethat our COVID-Classifier can be used in conjunction with other tests for optimal allocation of hospital resources by rapid triage of non-COVID-19 cases.The automatic detection of chest disease based on chest radiography has become a hot topic in medical imaging research. This project has overall two parts:

1. Backend Server - The server will cater the request from the medical diagnosis labs and individual users who do not have access professional consultation for medical diagnosis. The server should be capable of generating report whenever a chest radiograph is uploaded and provide accurate results.

Apart from the above said components, there will be a machine learning algorithm that will learn from the new x-rays being uploaded thus continuously improving its accuracy.

Back Ground : Coronavirus disease 2019 (COVID-19) has widely spread all over the world since the beginning of 2020. It is desirable to develop automatic and accurate detection of COVID-19 using chest CT

Purpose: To develop a fully automatic framework to detect COVID-19 using chest CT and evaluate its performance.

Key Words: LUNG DISEASE DETECTION METHODOLOGY, MACHINE LEARNING, ARTIFICIAL INTELLIGENCE .

Introduction:

Traditionally, lung disease can be detected via **skin test, blood test, sputum sample test** [7], chest X-ray examination and computed tomography (CT) scan examination [8]. Recently, deep learning has shown great potential when applied on medical images for disease detection, including lung disease.

A large number of diseases that affect the worldwide population are lung-related. Therefore, research in the field of Pulmonology has great importance in public health studies and focuses mainly on Infiltration, Atelectasis, Cardiomegaly, Effusion, Mass, Nodule, Pneumonia, Pneumothorax.

The World Health Organisation (WHO) estimates that there are 300 million people who suffer from asthma, and that this disease causes around 250 thousand deaths per year worldwide (Campos and Lemos, 2009). In addition, WHO estimates that 210 million people have Cardiomegaly. The disease caused the death of over 300 thousand people in 2005 (Gold Cardiomegaly, 2008). Recent studies reveal that CARDIOMEGALY is present in the 20 to 45 year-old age bracket, although it is characterised as an over-50-year-old disease. Accordingly, WHO estimates that the number of deaths due to CARDIOMEGALY will increase 30% by 2015, and by 2030 CARDIOMEGALY will be the third cause of mortalities worldwide (World…, 2014).

For the public health system, the early and correct diagnosis of any pulmonary disease is mandatory for timely treatment and prevents further death. From a clinical standpoint, diagnosis aid tools and systems are of great importance for the specialist and hence for the people's health.

X RAY images of lungs represent a slice of the ribcage, where a large number of structures are located, such as blood vessels, arteries, respiratory vessels, pulmonary pleura and parenchyma, each with its own specific information. Thus, for pulmonary disease analysis and diagnosis, it is necessary to segment lung structures. It is worth noting that segmentation is an essential step in image systems for the accurate lung disease diagnosis, since it delimits lung structures in X RAY images. Indeed, image processing techniques can help computer diagnosis if lung region is accurately obtained.

Following the segmentation process, an automatic procedure is applied to detect possible diseases in lung X RAY images in order to guide the radiologist diagnosis. Some studies have yielded promising disease detection results as reported by Trindade (2009) that uses texture descriptors extracted from the gray level concurrence matrix (GLCM) (Haralick et al., 1973) to describe three disease patterns (nodule, emphysema and frosted glass) and a normal one. Shimo et al. (2010) also employ GLCM texture descriptors to determine if the lungs are healthy or not. Furthermore, some papers address the detection of certain specific diseases, such as nodules (Ayres et al., 2010; Silva and Oliveira, 2010), and emphysema (Felix et al., 2007, 2011).

Machine learning (ML) based methods have shown unprecedented success in the reliable analysis of medical images. ML-based approaches are scalable, automatable, and easy to implement in clinical settings. A common application of ML-based image analysis is the classification of images with highly similar features. This approach relies on the segmentation of image region of interest, identification of effective image features extracted from the segmented area in the spatial or frequency domain, and development of an optimal machine learning-based classification method to accurately assign image samples into target classes. Recently, several ML-based methods for the diagnosis of COVID-19 medical images has been proposed. Wang et al. applied a pre-trained deep learning model called DenseNet 121 to CT images aiming to classify COVID-19 imaging tests into positive and negative categories leading to 81.24% accuracy. Also, Roy et al. studied the application of deep learning models to analyze COVID-19 infections in a small dataset of lung ultrasonography(LUS) images (only 11 patients). Zhang et al. proposed the application of the lung-lesion segmentation in CT images a ResNet-18 classifier model for three classes of COVID-19, pneumonia, and normal, generating an accuracy of 92.49%.

Here, we hypothesized that CXR images of COVID-19 patients can be reliably distinguished from other forms of pneumonia using an ML-based classifier. "We used a dimensionality reduction approach to generate a model with an optimized set of synthetic features that can distinguish COVID-19 images with an accuracy of 94% from non-COVID-19 cases."The detection and extraction of characteristics from the whole CXR picture without any segmentation procedure on chest lesions is a unique feature of our approach.

Not only does this new quantitative marker help us prevent segmentation mistakes, but it also helps us decrease the computing cost of our final model. Our findings show that basic machine learning-based categorization may be used in conjunction with other tests to aid in the differential diagnosis of COVID-19 patients' CXR pictures. More generally, we believe that our method for quickly classifying CXR pictures may be readily applied in any future viral epidemic.

METHODOLOGY ADOPTED:

In this section, the methodology used to conduct the survey of recent lung disease detection using deep learning is described. Figure 1 shows the flowchart of the methodology used. First, a suitable database, as a main source of reference, of articles was identified. The Scopus database was selected as it is one of the largest databases of scientific peer-reviewed articles. However, several significant articles, indexed by Google Scholar but not Scopus, are also included based on the number of citations that they have received. Some preprint articles on COVID-19 are also included as the disease has just recently emerged. To ensure that this survey only covers the state-of-the-art works, only articles published recently (2016–2020) are considered. However, several older but significant articles are included too. To search for all possible deep learning aided lung disease detection articles, relevant keywords were used to search for the articles. The keywords used were “deep learning”, “detection”, “classification”, “CNN”, “lung disease”, “Tuberculosis”, “pneumonia”, “lung cancer”, “COVID-19” and “Coronavirus”. Studies were limited to articles written in English only. At the end of this phase, we identified 366 articles. Second, to select only the relevant works, screening was performed. During the screening, only the title and abstract were assessed. The main selection criteria were this survey is only interested in work, whereby deep learning algorithms were applied to detect the relevant diseases. Articles considered not relevant were excluded. Based on the screening performed, only 98 articles were shortlisted.

Blood tests, pulmonary function tests (spirometry), pulse oximetry, chest x-ray, chest CT, bronchoscopy with biopsy, or surgical biopsy may be performed to help diagnose your condition. Treatment may depend on the underlying cause of the disease and your health status. Lung disease is a leading cause of death in the U.S. The three main categories of lung disease each encompass different diseases, including asthma, pulmonary fibrosis, and pulmonary edema. According to the American Lung Association, **COPD** is the third leading cause of death in the U.S. Dr. Meyer identifies COPD as one of the most serious and dangerous respiratory illnesses, and COPD is the number one problem seen in most pulmonology offices. “It's a very serious disease. Once you get COPD, you've got it.

1 INVESTIGATIVE TECHNIQUES

The investigation project technique is of experiment type. "There are many similar models that are detecting lung diseases available but their number of diseases is limited to just one." We are trying to get an accuracy above 90% for more than 5 lung diseases.

2 PROPOSED SOLUTION

We will ask user to upload digital Jpeg format images of X rays . The uploaded image will be pre-processed i.e its contrast will be increased , ribs will be removed and images will be highlighted. "After that inference will be performed which , with the help of depth-wise convolution will detect the diseases." The detected disease will be displayed on the website with its %.

3 WORK BREAKDOWN STRUCTURE

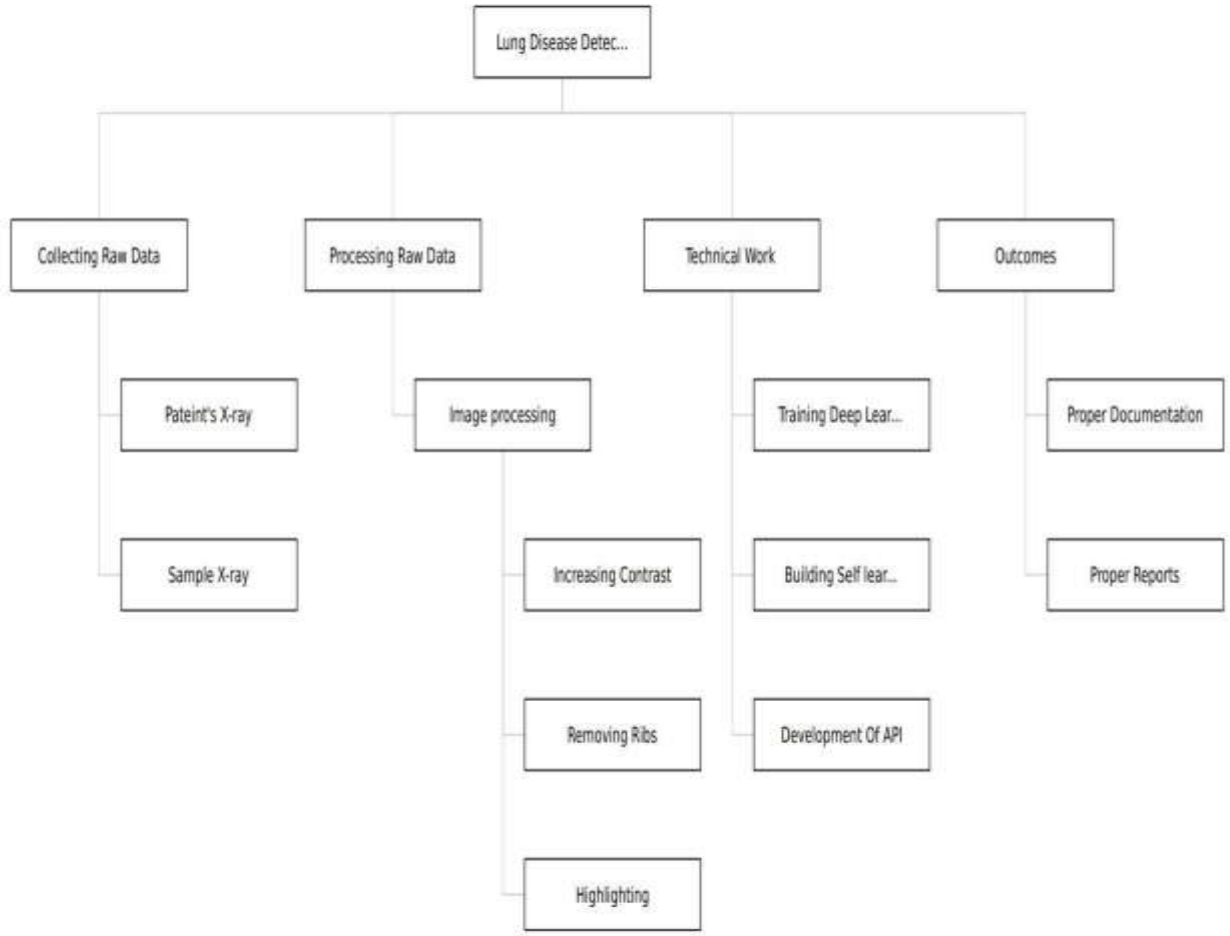

Figure - WORK BREAKDOWN STRUCTURE

4 TOOLS AND TECHNOLOGIES USED Image Processing (Segmentation)

- Machine Learning
- Deep Learning
- Web development
- Powerful Computer with GPU
- Google Colab
- Kaggle
- Python

- Keras/Tensorflow and Terminal

Deep learning is a subfield of machine learning relating to algorithms inspired by the function and structure of the brain. Recent developments in machine learning, particularly deep learning, support the identification, quantification and classification of patterns in medical images [9]. These developments were made possible due to the ability of deep learning to learned features merely from data, instead of hand-designed features based on domain-specific knowledge. Deep learning is quickly becoming state of the art, leading to improved performance in numerous medical applications. Consequently, these advancements assist clinicians in detecting and classifying certain medical conditions efficiently [10].
Numerous works on the detection of lung disease using deep learning can be found in the literature. To the best of our knowledge, however, only one survey paper has been published in the last five years to analyse the state-of-the-art work on this topic [11]. In that paper, the history of deep learning and its applications in pulmonary imaging are presented. Major applications of deep learning techniques on several lung diseases, namely pulmonary nodule diseases, pulmonary embolism, pneumonia, and interstitial lung disease, are also described. In addition, the analysis of several common deep learning network structures used in medical image processing is presented. However, their survey is lacking in the presentation of taxonomy and analysis of the trend of recent work. A taxonomy shows relationships between previous work and categorises them based on the identified attributes that could improve reader understanding of the topic. Analysis of trend, on the other hand, provides an overview of the research direction of the topic of interest identified from the previous work. In this paper, a taxonomy of deep learning applications on lung diseases and a trend analysis on the topic are presented. The remaining issues and possible future direction are also described. The aims of this paper are as follows: (1) produce a taxonomy of the state-of-the-art deep learning based lung disease detection systems; (2) visualise the trends of recent work on the domain; and (3) identify the remaining issues and describes potential future directions in this domain. This paper is organised as follows. Section 2 presents the methodology of conducting this survey. Section 3 describes the general processes of using deep learning to detect lung disease in medical images. Section 4 presents the taxonomy, with detailed explanations of each subtopic within the taxonomy. The analysis of trend, research gap and future directions of lung disease detection using deep learning are presented
in Section 5. Section 6 describes the limitation of the survey. Section 7 concludes this paper.

REFERENCES

1. Jaeger S, Karargyris A, Candemir S, Folio L, Siegelman J, Callaghan FM, Xue Z, Palaniappan K, Singh RK, Antani S, Thoma G, Wang Y, Lu P, McDonald C. Automatic tuberculosis screening using chest radiographs. IEEE Trans. Med. Imaging, 33(2):233–45, 2014.
2. Daley C, Gotway M, Jasmer R. Radiographic manifestations of tuberculosis. A Primer for Clinicians. San Francisco, CA: Curry International Tuberculosis Center, 2009.
3. Folio LR. Robochest Web Teaching Tool. Available online: www.robochest.com (last accessed 10 April 2013)
4. Lee KS, Song KS, Lim TH, et al. Adult-onset pulmonary tuberculosis: findings on chest radiographs and CT scans. AJR Am J Roentgenol. 160:753-8, 1993.
5. vanGinneken B, Hogeweg L, Prokop M. Computer-aided diagnosis in chest radiography: Beyond nodules. Eur. J. Radiol., 72(2):226–30, 2009.
6. Lodwick GS. Computer-aided diagnosis in radiology: A research plan. Investigative Radiology, 1(1):72, 1966.
7. Lodwick G, Keats T, Dorst J. The coding of Roentgen images for computer analysis as applied to lung cancer. Radiology, 81(2):185, 1963.
8. Sakai S, Soeda H, Takahashi N, Okafuji T, Yoshitake T, Yabuuchi H, Yoshino I, Yamamoto K, Honda H, Doi K. Computer-aided nodule detection on digital chest radiography: Validation test on consecutive T1 cases of resectable lung cancer. J. Digital Imaging, 19(4):376–82, 2006.
9. Shiraishi J, Abe H, Li F, Engelmann R, MacMahon H, Doi K. Computer-aided diagnosis for the detection and classification of lung cancers on chest radiographs: ROC analysis of radiologists' performance. Acad. Radiol., 13(8):995–1003, 2006.
10. Kakeda S, Moriya J, Sato H, Aoki T, Watanabe H, Nakata H, Oda N, Katsuragawa S, Yamamoto K, Doi K. Improved detection of lung nodules on chest radiographs using a commercial computer- aided diagnosis system. Am. J. Roentgenol., 182(2):505-10, 2004.
11. Doi K. Current status and future potential of computer-aided diagnosis inmedical imaging. Br. J. Radiol. 78(1):3-19, 2005.
12. vanGinneken B, terHaarRomeny B, Viergever M. Computeraided diagnosis in chest radiography: A survey. IEEE Trans. Med. Imag., 20(12):1228–41, 2001.
13. vanGinneken B, StegmannM,Loog M. Segmentation of anatomical structures in chest radiographs using supervised methods: A comparative study on a public database. Med. Image Anal., 10(1):19–40, 2006.
14. vanGinneken B, terHaarRomeny B. Automatic segmentation of lung fields in chest radiographs. Med. Phys., 27(10):2445-55, 2000.
15. Dawoud A. Fusing shape information in lung segmentation in chest radiographs. Image Anal. Recognit., pp. 70–8, 2010.
16. vanGinneken B, Katsuragawa S, terHaarRomeny B, Doi K, Viergever M. Automatic detection of abnormalities in chest radiographs using local texture analysis. IEEE Trans. Med. Imag., 21(2):139–49, 2002.
17. Hogeweg L, Mol C, de Jong P, Dawson R, Ayles H, van Ginneken B. Fusion of local and global detection systems to detect tuberculosis in chest radiographs. Proc. MICCAI, pp. 650–57, 2010.

24. Shen R, Cheng I, Basu A. A hybrid knowledge-guided detection technique for screening of infectious pulmonary tuberculosis from chest radiographs. IEEE Trans. Biomed. Eng. 57(11): 2646–56, 2010.
25. Xu T, Cheng I, Mandal M. Automated cavity detection of infectious pulmonary tuberculosis in chest radiographs. Int. IEEE Eng. Med. Biol. Soc., pp. 5178–81, 2011.
26. Hogeweg L, Sánchez CI, de Jong PA, Maduskar P, van Ginneken B. Clavicle segmentation in chest radiographs. Med. Image Anal. 16(8):1490–1502, 2012.
27. Freedman M, Lo S, Seibel J, Bromley C. Lung nodules: Improved detection with software that suppresses the rib and clavicle on chest radiographs. Radiology, 260(1): 265–73, 2011.
28. Arzhaeva Y, Tax D, van Ginneken B, Dissimilarity-based classification in the absence of local ground truth: Application to the diagnostic interpretation of chest radiographs. Pattern Recognition, 42(9):1768–76, 2009.
29. Jaeger S, Karargyris A, Candemir S, Siegelman J, Folio L, Antani S, Thoma G. Automatic screening for tuberculosis in chest radiographs: a survey. Quant Imaging Med Surg. 2013; 3(2):89-99.
30. Shiraishi J, Katsuragawa S, Ikezoe J, Matsumoto T, Kobayashi T, Komatsu K, Matsui M, Fujita H, Kodera Y, Doi K. Development of a digital image database for chest radiographs with and without a lung nodule: Receiver operating characteristic analysis of radiologists detection of pulmonary nodules. Am. J. Roentgenol., 174:71–4, 2000.
31. vanGinneken B, Stegmann M, Loog M. Segmentation of anatomical structures in chest radiographs using supervised methods: A comparative study on a public database. Med. Image Anal., 10(1):19-40, 2006.
32. Jaeger S, Candemir S, Antani S, Wáng YX, Lu PX, Thoma G. Two public chest X-ray datasets for computer- aided screening of pulmonary diseases. Quant Imaging Med Surg 2014;4(6):475-477.
33. Jain R, Kasturi R, Schunck BG. Machine Vision. McGraw-Hill, Inc., ISBN 0-07-032018-7, 1995
34. Bhattacharyya A. On a measure of divergence between two statistical populations defined by their probability distributions. Bull. Calcutta Math. Soc. 35:99–109, 1943.
35. Liu C, Yuen J, Torralba A. SIFT flow: Dense correspondence across different scenes and its applications. IEEE Trans. Pattern Anal. Mach. Intell., 33(5): 978-94, 2011.
36. Lowe D. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis., 60(2):91-110, 2004.
37. Boykov Y, Funka-Lea G. Graph cuts and efficient n-d image segmentation. Int. J. Comput. Vis., 70:109–131, 2006.
38. Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts. IEEE Trans. Pattern Anal. Mach. Intell., 23(11):1222–39, 2001.
39. Gururajan A, Sari-Sarraf H, Hequet E. Interactive texture segmentation via IT-SNAPS. IEEE Southwest Symp. Image Anal. Interpret., 2010.