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

Review of Artificial Intelligence Technique based Lung Cancer Detection

Pawan Kumar Sinha, Anand Kumar Srivastava
M.Tech Scholar, Assistant Professor
Department of Electronics and Communication Engineering
SHEAT College of Engineering, Varanasi, UP, India

Abstract : Timely diagnosis and determination to the type of lung cancer diseases has important clinical significance. Generally, it requires multiple imaging methods to complement each other to obtain a comprehensive diagnosis. Various learning methods, such as conventional clustering and classification, have been applied in diagnosing diseases to categorize samples based on their features. Artificial intelligence based machine learning techniques provides accurate prediction model of the various biomedical diseases. This paper presents review of lung diseases classification techniques using artificial intelligence techniques.

***IndexTerms* – Machine Learning, Lung, AI, Classification, Diagnosis .**

I. INTRODUCTION

Lung diseases identification based on analysis and processing of medical images is important to assist medical doctors during the diagnosis process. In this context, this work proposes a new feature extraction method based on human tissue density patterns, namely Analysis of Human Tissue Densities in Lung Diseases [1]. To counter the outbreak of corona, the accurate diagnosis of suspected cases plays a crucial role in timely quarantine, medical treatment, and preventing the spread of the pandemic. Considering the limited training cases and resources (e.g, time and budget), we propose a Multi-task Multi-slice Deep Learning System (M^3 Lung-Sys) for multi-class lung pneumonia screening from CT imaging, which only consists of two 2D CNN networks, i.e., slice- and patient-level classification networks. The former aims to seek the feature representations from abundant CT slices instead of limited CT volumes, and for the overall pneumonia screening, the latter one could recover the temporal information by feature refinement and aggregation between different slices [2].

Figure 1: Lungs Sample (google image)

Automatic screening and diagnosis of lung abnormalities from chest X-ray images has been recently drawing attention from the computer vision and medical imaging communities. Previous studies of deep neural networks have predominantly demonstrated the effectiveness of lung disease binary classification procedures. However, large numbers of medical images-which can be labeled with a variety of existing or suspected pathologies-are required to be interpreted and reported upon daily by an individual radiologist; this poses a challenge in maintaining a consistently high diagnosis accuracy. In this work, we present a competitive study of knowledge distillation (KD) in deep learning for classification of abnormalities in chest X-ray images [3].

Chest computed tomography (CT) becomes an effective tool to assist the diagnosis of coronavirus disease. Due to the outbreak of corona worldwide, using the computed-aided diagnosis technique for lungs diseases classification based on CT images could largely alleviate the burden of clinicians. In this work, we propose an Adaptive Feature Selection guided Deep Forest (AFS-DF) for lungs diseases classification based on chest CT images. Specifically, we first extract location-specific features from CT images. Then, in order to capture the high-level representation of these features with the relatively small-scale data, we leverage a deep forest model to learn high-level representation of the features. Moreover, we propose a feature selection method based on the trained deep forest model to reduce the redundancy of features, where the feature selection could be adaptively incorporated with the lungs diseases classification model [4].

Shortness of breath is a major reason that patients present to the emergency department (ED) and point-of-care ultrasound (POCUS) has been shown to aid in diagnosis, particularly through evaluation for artifacts known as B-lines. B-line identification and quantification can be a challenging skill for novice ultrasound users, and experienced users could benefit from a more objective measure of quantification. We sought to develop and test a deep learning (DL) algorithm to quantify the assessment of B-lines in lung ultrasound [5]. Lung diseases can result in acute breathing problems and prevent the human body from acquiring enough oxygen. These diseases, such as pneumonia (P), pleural effusion (Ef), lung cancer, pneumothorax (Pt), pulmonary fibrosis (F), infiltration (In) and emphysema (E), adversely affect airways, alveoli, blood vessels, pleura and other parts of the respiratory system. The death rates of P and lung cancer are higher than those of other typical lung diseases. In visualization examination, chest radiography, such as anterior-posterior or lateral image viewing, is a straightforward approach used by clinicians/radiologists to diagnose and locate possible lung abnormalities rapidly. However, a chest X-ray image of patients may show multiple abnormalities associated with coexisting conditions, such as P, E, F, Pt, atelectasis, lung cancer or surgical interventions, which further complicate diagnosis. In addition, poor-quality X-ray images and manual inspection have limitations in digital image-automated classification. Hence, this study intends to propose a multilayer machine vision classifier to automatically identify the possible class of lung diseases within a bounding region of interest (ROI) on a chest X-ray image. For digital image texture analysis, a two-dimensional (2D) fractional-order convolution (FOC) operation with a fractional-order parameter, $v=0.3-0.5$, is used to enhance the symptomatic feature and remove unwanted noises. Then, maximum pooling is performed to reduce the dimensions of feature patterns and accelerate complex computations [6]. Deep learning (DL) has proved successful in medical imaging and, in the wake of the lungs diseases pandemic, some works have started to investigate DL-based solutions for the assisted diagnosis of lung diseases. While existing works focus on CT scans, this work studies the application of DL techniques for the analysis of lung ultrasonography (LUS) images. Specifically, we present a novel fully-annotated dataset of LUS images collected from several Italian hospitals, with labels indicating the degree of disease severity at a frame-level, video-level, and pixel-level (segmentation masks).

II. LITERATURE REVIEW

S. S. Araújo et al.,[1] The proposed method uses human tissues radiological densities, in Hounsfield Units, to perform the features extraction on thorax computerized tomography images. We compared the proposed method against the Gray Level Co-occurrence Matrix and Statistical Moments to accomplish the performance evaluation alongside four machine learning classifiers. Overall, the results revealed that the proposal achieved higher accuracy ratios while it took the lowest runtime in all performed experiments. Thus, we consider our proposal as a valid alternative to be used in real-time applications.

X. Qian et al.,[2] presents lungs diseases from Healthy, H1N1, and CAP cases, our M^3 Lung-Sys also be able to locate the areas of relevant lesions, without any pixel-level annotation. To further demonstrate the effectiveness of our model, we conduct extensive experiments on a chest CT imaging dataset with a total of 734 patients (251 healthy people, 245 lungs diseases patients, 105 H1N1 patients, and 133 CAP patients). The quantitative results with plenty of metrics indicate the superiority of our proposed model on both slice- and patient-level classification tasks. More importantly, the generated lesion location maps make our system interpretable and more valuable to clinicians.

T. K. K. Ho et al.,[3] aims to either distill knowledge from cumbersome teacher models into lightweight student models or to self-train these student models, to generate weakly supervised multi-label lung disease classifications. Our approach was based on multi-task deep learning architectures that, in addition to multi-class classification, supported the visualizations utilized in saliency maps of the pathological regions where an abnormality was located. A self-training KD framework, in which the model learned from itself, was shown to outperform both the well-established baseline training procedure and the normal KD, achieving the AUC improvements of up to 6.39% and 3.89%, respectively. Through application to the publicly available ChestX-ray14 dataset, we demonstrated that our approach efficiently overcame the interdependency of 14 weakly annotated thorax diseases and facilitated the state-of-the-art classification compared with the current deep learning baselines.

L. Sun et al.,[4] presents AFS-DF on lungs diseases -19 dataset with 1495 patients of lungs diseases -19 and 1027 patients of community acquired pneumonia (CAP). The accuracy (ACC), sensitivity (SEN), specificity (SPE), AUC, precision and F1-score achieved by our method are 91.79%, 93.05%, 89.95%, 96.35%, 93.10% and 93.07%, respectively. Experimental results on the lungs diseases dataset suggest that the proposed AFS-DF achieves superior performance in lungs diseases vs. CAP classification, compared with 4 widely used machine learning methods.

C. Baloescu et al.,[5] We utilized ultrasound clips (n = 400) from an existing database of ED patients to provide training and test sets to develop and test the DL algorithm based on deep convolutional neural networks. Interpretations of the images by algorithm were compared to expert human interpretations on binary and severity (a scale of 0- 4) classifications. Our model yielded a sensitivity of 93% (95% confidence interval (CI) 81%-98%) and a specificity of 96% (95% CI 84%-99%) for the presence or absence of B-lines compared to expert read, with a kappa of 0.88 (95% CI 0.79-0.97). Model to expert agreement for severity classification yielded a weighted kappa of 0.65(95% CI 0.56- 074). Overall, the DL algorithm performed well and could be integrated into an ultrasound system in order to help diagnose and track B-line severity. The algorithm is better at distinguishing the presence from the absence of B-lines but can also be successfully used to distinguish between B-line severity. Such methods could decrease variability and provide a standardized method for improved diagnosis and outcome.

J. -X. Wu et al.,[6] A multilayer machine vision classifier with radial Bayesian network and gray relational analysis is used to screen subjects with typical lung diseases. Anterior-posterior chest X-ray images from the NIH chest X-ray database (NIH Clinical Center) are enrolled. For digital chest X-ray images, with K -fold cross-validation, the proposed multilayer machine vision classifier is applied to facilitate the diagnosis of typical lung diseases on specific bounding ROIs, as promising results with mean recall (%), mean precision (%), mean accuracy (%) and mean F1 score of 98.68%, 82.42%, 83.57% and 0.8981, respectively, for assessing the performance of proposed multilayer classifier for rapidly screening lung lesions on digital chest X-ray images.

S. Roy et al.,[7] introduce a new method based on uninorms for effective frame score aggregation at a video-level. Finally, we benchmark state of the art deep models for estimating pixel-level segmentations of lungs diseases imaging biomarkers. Experiments on the proposed dataset demonstrate satisfactory results on all the considered tasks, paving the way to future research on DL for the assisted diagnosis of lungs diseases from LUS data.

S. Pang et al.,[8] propose a deep learning model to identify lung cancer type from CT images for patients in Shandong Provincial Hospital. It has a two-fold challenge: artificial intelligent models trained by public datasets cannot meet such practical requires, and the amount of collected patients' data is quite few. To solve the two-fold problem, we use image rotation, translation and transformation methods to expand and balance our training data, and then densely connected convolutional networks (DenseNet) is used to classify malignant tumor from images collected from, and finally adaptive boosting (adaboost) algorithm is used to aggregate multiple classification results to improve classification performance. Experimental results show that our method can achieve identifying accuracy 89.85%, which performs better than DenseNet without adaboost, ResNet, VGG16 and AlexNet. This provides an efficient, non-invasive detection tool for pathological diagnosis to lung cancer type.

H. Yazdani et al.,[9] presents method evaluates samples for their movement from one cluster to another. This technique allows us to find critical samples in advance those with the potential ability to belong to other clusters in the near future. BFPM was applied on metabolomics of individuals in a lung cancer case-control study. Metabolomics as proximal molecular signals to the actual disease processes may serve as strong biomarkers of current disease process. The goal is to know whether serum metabolites of a healthy human can be differentiated from those with lung cancer. Using BFPM, some differences were observed, the pathology data were evaluated, and critical samples were recognized.

F. Yan et al.,[10] The chest X-ray is a simple and economical medical aid for auxiliary diagnosis and therefore has become a routine item for residents' physical examinations. Based on 40167 images of chest radiographs and corresponding reports, we explore the abnormality classification problem of chest X-rays by taking advantage of deep learning techniques. First of all, since the radiology reports are generally templatized by the aberrant physical regions, we propose an annotation method according to the abnormal part in the images. Second, building on a small number of reports that are manually annotated by professional radiologists, we employ the long short-term memory (LSTM) model to automatically annotate the remaining unlabeled data. The result shows that the precision value reaches 0.88 in accurately annotating images, the recall value reaches 0.85, and the F1-score reaches 0.86. Finally, we classify the abnormality in the chest X-rays by training convolutional neural networks, and the results show that the average AUC value reaches 0.835.

A. Rao et al.,[11] Accumulation of excess air and water in the lungs leads to breakdown of respiratory function and is a common cause of patient hospitalization. Compact and non-invasive methods to detect the changes in lung fluid accumulation can allow physicians to assess patients' respiratory conditions. In this work, an acoustic transducer and a digital stethoscope system are proposed as a targeted solution for this clinical need. Alterations in the structure of the lungs lead to measurable changes which can be used to assess lung pathology. We standardize this procedure by sending a controlled signal through the lungs of six healthy subjects and six patients with lung disease. We extract mel-frequency cepstral coefficients and spectroid audio features, commonly used in classification for music retrieval, to characterize subjects as healthy or diseased. Using the K-nearest neighbors algorithm, we demonstrate 91.7% accuracy in distinguishing between healthy subjects and patients with lung pathology.

O. P. Singh et al.,[12] Currently, carbon dioxide (CO_2 waveforms measured by capnography are used to estimate respiratory rate and end-tidal CO_2 ($EtCO_2$) in the clinic. However, the shape of the CO_2 signal carries significant diagnostic information about the asthmatic condition. Previous studies have shown a strong correlation between various features that quantitatively characterize the shape of CO_2 signal and are used to discriminate asthma from non-asthma using pulmonary function tests, but no reliable progress was made, and no translation into clinical practice has been achieved. Therefore, this work reports a relatively simple signal processing algorithm for automatic differentiation of asthma and non-asthma. CO_2 signals were recorded from 30 non-asthmatic and 43 asthmatic patients. Each breath cycle was decomposed into subcycles, and features were computationally extracted. Thereafter, feature selection was performed using the area under the receiver operating characteristics curve analysis.

III. CHALLENGES

The similarities in pneumonia secondary to lung cancer (such as fatigue, cough and difficulty in breathing) make it difficult to differentiate them clinically and can result in the spread of the viral infection among contacts and health staff. All lung cancer patients scheduled for anticancer treatment must be tested for lungs diseases, irrespective of symptoms or contact history, to determine the status before compromising their immune system. Unfortunately, testing kits are not readily available [4].

It is advisable to use image-guided biopsy to establish the diagnosis and avoid aerosol-generating procedures like bronchoscopic biopsy, bronchial lavage cytology and mediastinal staging with endobronchial ultrasound.

The Royal Australian and New Zealand College of Radiologists (Sydney, NSW, Australia) recommends deferring or cancelling the no urgent procedures except procedures to save life and permanent disability [5]. For lung cancer treatment, biopsy is the first step and it should be decided by the doctor, on a case-by-case basis to plan further interventions.

If a lung cancer patient on targeted therapy develops lungs diseases, the treating team would need to look at the potential drug interactions between the targeted drug(s) that the patient was taking and the medication(s) needed to treat lungs diseases.

Increased risk of hepatic and/or renal dysfunction, are distinct possibilities; if so, may mandate either a reduction in dose/frequency or temporary discontinuation of the ongoing-targeted drug(s). Theoretically, a short duration of discontinuation/dose modification is unlikely to have any adverse impact on the disease status (especially if the patient is in clinical remission) but the risk of tumor progression may increase if this interruption is sustained for several days. Targeted therapies with known cardiovascular toxicities (especially VEGF inhibitors) may need to be temporarily withheld until the patient has recovered fully from lungs diseases, especially since there is a concern that myocardial dysfunction may be an important contributing factor to mortality from lungs diseases.

The primary role of palliative and supportive care is symptom management by addressing the physical, emotional, social and spiritual needs of patients in a life-threatening illness. Advanced stages of lung cancer are rarely curative in nature, hence there is an increased need for palliative and supportive care. Symptoms that require palliation include pain, breathlessness, anxiety and depression. The management of these symptoms will not only increase survival but also improves the quality of life of lung cancer patients.

Elderly lung cancer patients are vulnerable group for lungs diseases. Acute onset breathlessness in lung cancer patients at this time of lungs diseases pandemic will always lead to a diagnostic dilemma whether it is due to disease progression or due to lungs diseases infection. The social and psychological stigma associated with lungs diseases is the major cause of suffering in this group of patients. This is associated with disproportionate anxiety and depression. So, anxiety and depression needs to be addressed simultaneously with other symptoms like pain and cough.

IV. CONCLUSION

Lung diseases or cancer remains the main source of disease related mortality for both men and women and its frequency is expanding around the world. Lung disease is the uncontrolled development of irregular cells that begin off in one or both Lungs. The earlier detection of cancer is not easier process but if it is detected, it is curable. In this paper a study was made to analyze the lung cancer prediction using various classification algorithms. In future we develop efficient deep learning technique to more accurate prediction model to detection of lung diseases.

REFERENCES

1. S. S. Araújo Alves, E. de Souza Rebouças, S. A. Freitas de Oliveira, A. Magalhães Braga and P. P. Rebouças Filho, "Lung Diseases Classification by Analysis of Lung Tissue Densities," in IEEE Latin America Transactions, vol. 18, no. 09, pp. 1329-1336, September 2020, doi: 10.1109/TLA.2020.9381790.
2. X. Qian et al., "M^{3}Lung-Sys: A Deep Learning System for Multi-Class Lung Pneumonia Screening From CT Imaging," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 12, pp. 3539-3550, Dec. 2020, doi: 10.1109/JBHI.2020.3030853.
3. T. K. K. Ho and J. Gwak, "Utilizing Knowledge Distillation in Deep Learning for Classification of Chest X-Ray Abnormalities," in IEEE Access, vol. 8, pp. 160749-160761, 2020, doi: 10.1109/ACCESS.2020.3020802.
4. L. Sun et al., "Adaptive Feature Selection Guided Deep Forest for lungs diseases Classification With Chest CT," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 10, pp. 2798-2805, Oct. 2020, doi: 10.1109/JBHI.2020.3019505.
5. C. Baloescu et al., "Automated Lung Ultrasound B-Line Assessment Using a Deep Learning Algorithm," in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 67, no. 11, pp. 2312-2320, Nov. 2020, doi: 10.1109/TUFFC.2020.3002249.
6. J. -X. Wu, P. -Y. Chen, C. -M. Li, Y. -C. Kuo, N. -S. Pai and C. -H. Lin, "Multilayer Fractional-Order Machine Vision Classifier for Rapid Typical Lung Diseases Screening on Digital Chest X-Ray Images," in IEEE Access, vol. 8, pp. 105886-105902, 2020, doi: 10.1109/ACCESS.2020.3000186.
7. S. Roy et al., "Deep Learning for Classification and Localization of lungs diseases -19 Markers in Point-of-Care Lung Ultrasound," in IEEE Transactions on Medical Imaging, vol. 39, no. 8, pp. 2676-2687, Aug. 2020, doi: 10.1109/TMI.2020.2994459.
8. S. Pang, Y. Zhang, M. Ding, X. Wang and X. Xie, "A Deep Model for Lung Cancer Type Identification by Densely Connected Convolutional Networks and Adaptive Boosting," in IEEE Access, vol. 8, pp. 4799-4805, 2020, doi: 10.1109/ACCESS.2019.2962862.
9. H. Yazdani, L. L. Cheng, D. C. Christiani and A. Yazdani, "Bounded Fuzzy Possibilistic Method Reveals Information about Lung Cancer through Analysis of Metabolomics," in IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 17, no. 2, pp. 526-535, 1 March-April 2020, doi: 10.1109/TCBB.2018.2869757.
10. F. Yan, X. Huang, Y. Yao, M. Lu and M. Li, "Combining LSTM and DenseNet for Automatic Annotation and Classification of Chest X-Ray Images," in IEEE Access, vol. 7, pp. 74181-74189, 2019, doi: 10.1109/ACCESS.2019.2920397.

11. A. Rao, S. Chu, N. Batlivala, S. Zetumer and S. Roy, "Improved Detection of Lung Fluid With Standardized Acoustic Stimulation of the Chest," in IEEE Journal of Translational Engineering in Health and Medicine, vol. 6, pp. 1-7, 2018, Art no. 3200107, doi: 10.1109/JTEHM.2018.2863366.
12. O. P. Singh, R. Palaniappan and M. Malarvili, "Automatic Quantitative Analysis of Human Respired Carbon Dioxide Waveform for Asthma and Non-Asthma Classification Using Support Vector Machine," in IEEE Access, vol. 6, pp. 55245-55256, 2018, doi: 10.1109/ACCESS.2018.2871091.