**Доклади на Българската академия на науките**
**Comptes rendus de l'Académie bulgare des Sciences**
*Tome 76, No 10, 2023*

*BIOLOGY*
*Bioinformatics*

# DEEP LEARNING-BASED PRECISION DIAGNOSIS OF LUNG DISEASES ON THE INTERNET OF MEDICAL THINGS (IoMT)

**Lekshmy Sushama[#], Kuttaiyur Palaniswamy Sridhar, Michaelraj Kingston Roberts[*]**

*Received on June 6, 2023*
*Presented by I. Ivanov, Member of BAS, on July 31, 2023*

**Abstract**

Lung disease is one of the common and severe pathological conditions that affect the respiratory system, causing respiratory illness and potential mortality. In recent times, deep learning paradigm based on the Internet of Medical Things (IoMT) platform has been adopted as a viable solution to address the challenges encountered in detection of lung diseases which are characterized by their diverse nature and the complexities associated with their diagnosis. In this work, we have proposed an approach that aims to achieve accurate prediction and analysis of lung diseases. The proposed research methodology presents a Deep Learning-based Accurate Lung Disease Prediction (DL-ALDP) model based on deep learning algorithms to enhance its predictive capabilities. The DL-ALDP framework integrates several preprocessing techniques, including Wiener filtering, optimized region growing method (ORGM)-based feature extraction, and Contrast limited AHE (CLAHE)-based segmentation. The accurate prediction of lung diseases is achieved by utilizing a Deep Neural Network (DNN) for classification purposes. The DL-ALDP technique, as suggested, attained a precision of 86.77%, sensitivity of 82.47%, specificity of 92.87%, accuracy of 92.08%, and F1 score of 89.42%. The findings of this research underscore the prospective utility of deep learning techniques in forecasting and analyzing lung ailments within the context of the IoMT platform. Through IoMT capabilities, healthcare practitioners can avail themselves of enhanced prognostic accuracy and timeliness, resulting in superior patient care and outcomes.

**Key words:** lung disease, deep learning, Internet of Medical Things, prediction

[#]Corresponding author.
DOI:10.7546/CRABS.2023.10.07

**1. Introduction.** Lung diseases comprise diverse afflictions that impact the respiratory system, such as Chronic Obstructive Pulmonary Disease (COPD), asthma, lung cancer, and pneumonia [1]. The illnesses present significant health hazards and constitute a primary contributor to both disease and death on a global scale. The timely and accurate detecting of lung diseases is imperative for efficient management and treatment. Diagnosing these diseases poses significant challenges for healthcare providers due to various complexities, including subtle symptoms, overlapping clinical presentations, and specialized testing requirements.

The emergence of the Internet of Medical Things (IoMT) has presented a powerful platform for transforming healthcare delivery [2]. The principal findings of this investigation are:

- The utilization of the Wiener filtering technique for preprocessing has been observed to enhance the quality of Computer Tomography (CT) [3] images by reducing noise and improving the accuracy of subsequent analysis.
- The Optimized Region Growing Method (ORGM) is a technique that facilitates the extraction of lung regions and infected regions from CT images [4].
- The utilization of Contrast Limited Adaptive Histogram Equalization (CLAHE) in segmentation is a technique aimed at improving the visibility of lung abnormalities and enabling accurate classification during image segmentation [5].

The subsequent sections are organized in the prescribed order: Section 2 presents a comprehensive review of the existing literature, Section 3 outlines the proposed methodology, which contains several steps. Section 4 presents the simulation analysis. Section 5 of the study provides a summary and analysis.

**2. Literature survey and findings.** The literature review about the detection and prediction of lung diseases encompasses a range of subjects, such as deep learning methodologies, models for risk prediction, retrieval systems for medical images, and clinical trials. The research investigates diverse techniques to accurately identify pulmonary ailments, potentially enhancing patient prognoses and therapeutic alternatives.

The deep learning methodology proposed exhibits favourable levels of precision and responsiveness in forecasting the probability of pulmonary ailments [6].

The utilization of risk score is deemed significant in the prompt identification of subclinical interstitial lung disease among patients with rheumatoid arthritis [7].

The system under consideration employs deep CNNs to efficiently retrieve pertinent medical images pertaining to lung ailments, thereby facilitating the process of diagnosis and devising treatment strategies [8].

The therapeutic efficacy of Nintedanib has been observed in patients suffering from progressive fibrosing interstitial lung diseases that are linked with autoimmune diseases [9].

The utilization of weakly supervised deep learning methodology presents a potentially effective resolution for precise identification and classification of chronic obstructive pulmonary disease through the analysis of CT images [10].

Prediction models offer significant insights into the likelihood of lung cancer and can assist in clinical decision-making for pulmonary nodules that are detected incidentally [11].

**3. Materials and methods.** ***Proposed deep learning-based accurate lung disease prediction.*** The accurate lung disease prediction through deep learning on the IoMT platform comprises multiple stages: collecting data from IoMT devices, preprocessing, feature extraction, segmentation, wavelet processing, and classification using DNN. The proposed system is shown in Fig. 1.

***Data collection from IoMT.*** IoMT devices, comprising wearable sensors and remote monitoring systems, gather diverse data about lung health, encompassing physiological measurements, respiratory sounds, and medical imaging scans [12].

***Preprocessing.*** The data is subjected to a preprocessing stage to eliminate extraneous noise, artefacts, or irrelevant information. Equation (1) is utilized for

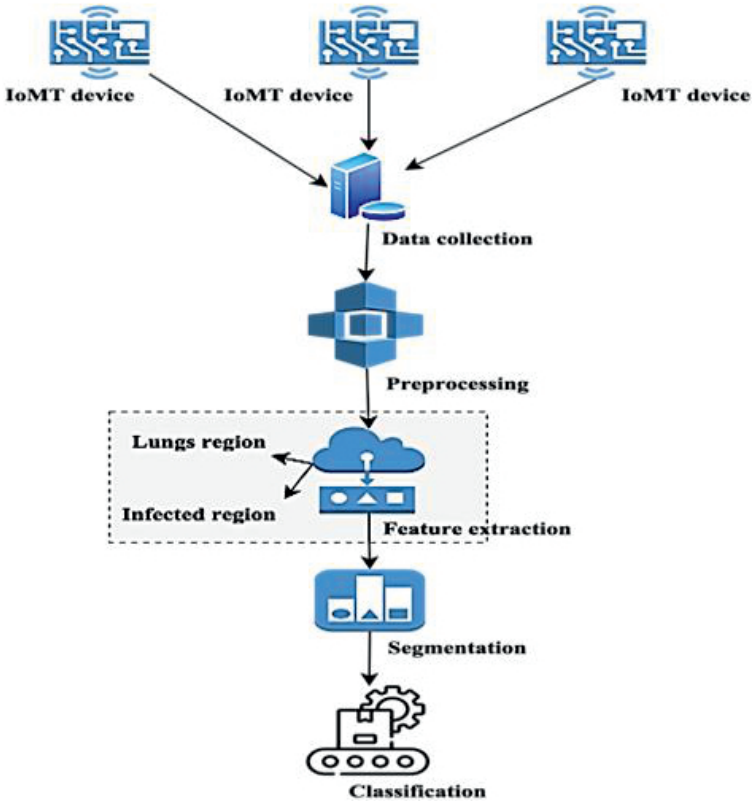



Fig. 1. Workflow of the proposed system

performing pixel-wise linear filtering.

$$Z(i,j) = n + \frac{u^2 - m^2}{u^2} * (J(i,j) - n), \tag{1}$$

where $Z(i,j)$ denotes the filtered image, while the original image is represented by $J(i,j)$. The local neighbourhood's mean and variance are denoted by $n$ and $u$, and the noise-related variance is represented by $m$.

***Feature extraction.*** For deep learning algorithms to achieve optimal performance in pattern recognition and prediction tasks, they must be provided with significant features with meaningful information [13].

***Extraction of lung regions.*** Preprocessing is a technique that enables the extraction of pertinent regions, specifically lung regions while preserving their quality [14]. The pulmonary regions of a CT scan depicting the lungs are overlaid onto the image's backdrop. Consider a CT image of the lung, denoted by $Z(i,j)$, comprising two distinct lung regions, wherein a dark background surrounds the infected areas. The process of segregating the lung region from its location can be achieved using a threshold $L$, as denoted by Equation (2).

$$Z(i,j) = \begin{cases} y & \text{if } J(i,j) > L \\ 0 & \text{else} \end{cases}. \tag{2}$$

Any coordinate pair $(i,j)$ within the image that satisfies the condition $J(i,j) > L$ is designated as the pulmonary region. The remaining pixel points are regarded as background areas.

***Extraction of the affected region.*** Segmenting the infected area involves the separation of the affected regions from the surrounding lung tissue in the CT scan image [15]. The methods for the proposed ORGM-optimized region-growing algorithm are outlined as follows:

1. Provide the CT image section that displays abnormal lung segmentation.
2. The optimized threshold symbolized by $\alpha$ can be generated using the ORGM algorithm.
3. The point $\alpha$ is the initial seed for the region-growing algorithm.
4. Incorporate the four adjacent pixels within the neighbourhood.
5. The task involves assessing the distance ($\beta$) between the mean intensity of newly acquired neighbouring pixels and the mean power of the region in question.
6. The region growing process should be executed when the value of $\alpha$ is less than or equal to $\beta$. The closest neighbouring pixel to the average intensity of the region incorporated if it has not already been included, and the location of the new pixel should be saved.
7. Record the mean of the recently established region and proceed to Step 2.
8. Conduct a region-growing process until all pixels that share similarities are grouped.

***Segmentation.*** Segmentation techniques are utilized to predict lung disease to isolate the region of interest in medical images, such as X-rays or CT scans, specifically about the lungs [16]. The clip limit, determined by the neighbourhood region's size, is the value at which histograms are truncated through histogram normalization. Equation (3) is utilized to assess the clip point.

$$\beta = \frac{P}{R}\left(1 + \frac{\alpha}{100} c_{\max}\right). \tag{3}$$

The variables in question are denoted as follows: $M$ represents the pixel count within each block, $N$ represents the block ranges, $c_{\max}$ represents the maximum slope, and $\alpha$ represents the clip factor. When $\alpha$ approaches zero, the Clipping Point (CP) is determined to be $P/R$, resulting in a uniform pixel distribution within the block [17]. As $\alpha$ approaches 100, there is a significant enhancement in contrast. CP is a crucial determinant in regulating advancements in comparison. The Mapping Function (MF) is obtained to reassign the gray levels of a block image as traces, as described by Equations (4) and (5).

$$MF(d) = \sum_{i=0}^{D} pdf(i), \tag{4}$$

$$T(d) = cdf(d) d_{\max}. \tag{5}$$

Equation (5) involves a remapping function denoted as $T(d)$, where $d$ represents the gray level and $d_{\max}$ refers to the maximum pixel value within the block. The process of interpolation serves to remove any artefacts present in the data. The proposed method exhibits reduced computational complexity for enhancement due to block processing.

***DNN classification.*** DNNs, a class of deep learning algorithms, are utilized for classification tasks [18]. The Restricted Boltzmann Machine (RBM) is a neural network architecture consisting of two layers. It connects stochastic binary inputs to stochastic binary outputs through symmetrically-weighted associations. The vector undergoes a reverse course through RBM, resulting in confabulation or reprocessing of the essential data, shown in Equation (6).

$$F(w, b) = -\sum_{x=0}^{M}\sum_{y=0}^{N} P_{xy} w_x b_y - \sum_{x=0}^{M} \alpha_x w_x - \sum_{y=0}^{N} \beta_y b_y. \tag{6}$$

Here $P_{xy}$ denotes the symmetrical interaction between the visible unit $w_x$ and the hidden unit $b_y$. The variables $\alpha$ and $\beta$ indicate the bias terms, while $x$ and $y$ correspond to the number of visible and remote units, respectively. The predictive model demonstrated important accuracy, sensitivity, specificity, precision, and F1 score levels in identifying lung pathologies on CT scans.

**4. Simulation analysis and findings.** The CT lung image classification models were implemented using MATLAB 2016 on an operating platform with system configurations that included an i5 processor and 4 GB RAM. The proposed model was subjected to comparison with established classifiers such as Support Vector Machine (SVM), CNN, Random Forest (RF), Naïve Bayes (NB), Linear Discriminant Analysis (LDA), and Principal Component Analysis (PCA), based on various measures of the classification model [19]. The NIH Chest X-rays dataset is a comprehensive collection of 112 120 frontal-view chest X-ray images from 30 805 distinct patients [20]. It is valuable for training and assessing deep learning models for detecting and classifying lung diseases.

The DL-ALDP technique exhibits superior performance in terms of specificity and precision across various iterations compared to alternative algorithms such as SVM, CNN, RF, NB, LDA, and PCA in Fig. 2(a) and (b). The results presented in Fig. 3 demonstrate a noteworthy enhancement in specificity, with a range of 5.8% to 9.8%, compared to the most effective baseline algorithm. The DL-ALDP technique exhibits enhanced precision in forecasting lung ailments in contrast to alternative algorithms in Fig. 3. The results demonstrate superior precision values compared to the best-performing baseline algorithm, with improvements ranging from 8.6% to 12.2%.

The DL-ALDP approach performs better than alternative algorithms such as SVM, CNN, RF, NB, LDA, and PCA, as evidenced by its consistently higher F1 score across multiple iterations as shown in Fig. 3. The results indicate a noteworthy enhancement in the F1 score, with a range of 6.2% to 12.2% compared to the algorithm with the highest performance among the baseline models. The DL-ALDP method, as proposed, attained a precision of 86.77%, sensitivity of 82.47%, specificity of 92.87%, accuracy of 92.08%, and F1 score of 89.42%.

**5. Conclusion.** The present study introduces DL-ALDP, a deep learning approach, to predict lung disease. The simulation outcomes exhibited good performance, achieving an accuracy rate of 92.08%, sensitivity rate of 82.47%, specificity rate of 92.87%, precision rate of 86.77%, and F1 score of 89.42%. The

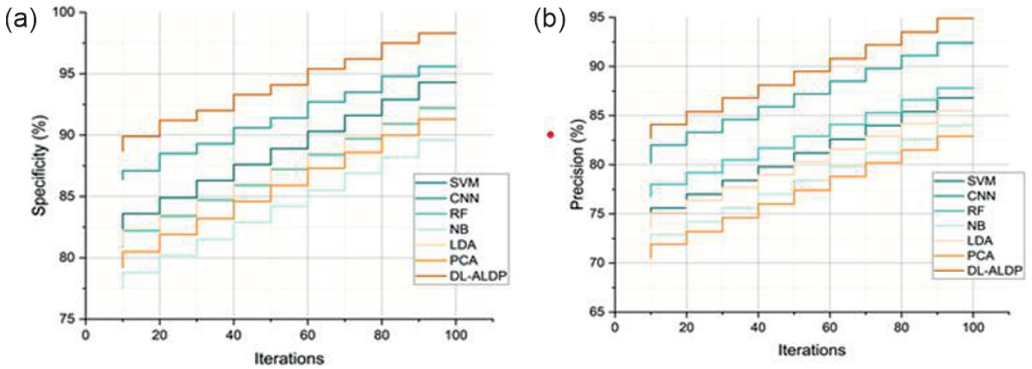



Fig. 2. (a) Specificity analysis; (b) Precision analysis

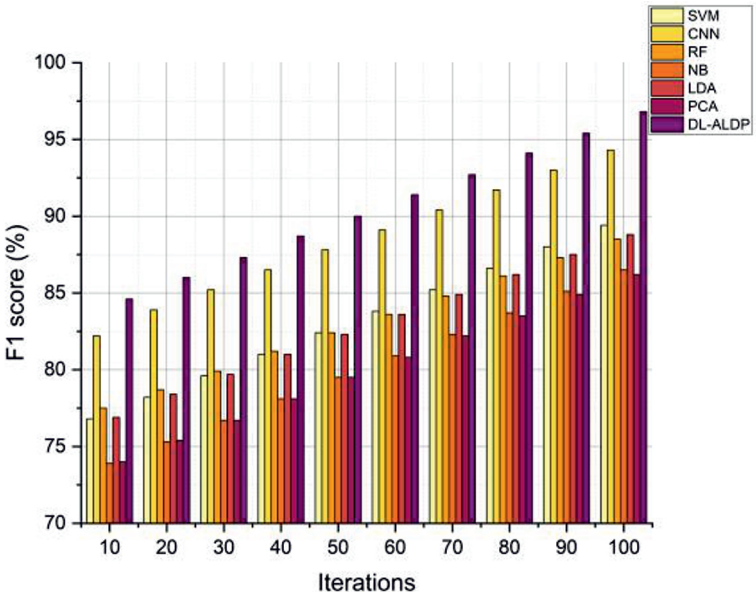


Fig. 3. F1 score analysis

findings of this study underscore the capacity of deep learning algorithms to forecast pulmonary ailments effectively. The efficacy of DL-ALDP in early detection and intervention is evidenced by its superior accuracy and performance metrics, resulting in improved patient outcomes and decreased healthcare expenditures.

## REFERENCES

[1] Hage Chehade A., N. Abdallah, J.-M. Marion, M. Oueidat, P. Chauvet (2022) Lung and colon cancer classification using medical imaging: a feature engineering approach, Phys. Eng. Sci. Med., **45**(3), 729–746.

[2] Dwivedi R., D. Mehrotra, S. Chandra (2022) Potential of the Internet of Medical Things (IoMT) applications in building an intelligent healthcare system: A systematic review, J. Oral Biol. Craniofac. Res., **12**(2), 302–318.

[3] Lasek J. (2022) The impact of data preprocessing on the accuracy of CNN-based heart segmentation. In: Progress in Image Processing, Pattern Recognition and Communication Systems: Proc. Conf. (CORES, IP&C, ACS) June 28–30 2021, 12, 173–180, Springer Int. Publ.

[4] Bangare S. L. (2022) Classification of optimal brain tissue using dynamic region growing and fuzzy min-max neural network in brain magnetic resonance images, Neurosci. Inf., **2**(3), 100019.

[5] Jebadass J. R., P. Balasubramaniam (2022) Low contrast enhancement technique for color images using interval-valued intuitionistic fuzzy sets with contrast limited adaptive histogram equalization, Soft Comp., **26**(10), 4949–4960.

[6] Harinath Reddy C., B.-V. Koushik Kumar, N. SaiTejaVarma, S. Vidya, P. Nagaraj et al. (2022) Risk prediction of lung disease using deep learning ap-

proach. In: 2nd Int. Conf. on Image Processing and Capsule Networks, ICIPCN 2021, 2, 462–471.
[7] Juge P.-A., B. Granger, M.-P. Debray, E. Ebstein, F. Louis Sidney et al. (2022) A risk score to detect subclinical rheumatoid arthritis-associated interstitial lung disease, Arthritis Rheumatol., **74**(11), 1755–1765.
[8] Agrawal S., A. Chowdhary, S. Agarwala, V. Mayya, S.-S. Kamath (2022) Content-based medical image retrieval system for lung diseases using deep CNNs, Int. J. Inf. Technol., **14**, 3619–3627.
[9] Matteson E. L., C. Kelly, J. H. W. Distler, A.-M. Hoffmann-Vold, J.-R. Seibold et al. (2022) Nintedanib in patients with autoimmune disease-related progressive fibrosing interstitial lung diseases: Subgroup analysis of the inbuild trial, Arthritis Rheumatol., **74**(6), 1039–1047.
[10] Sun J., X. Liao, Y. Yan, X. Zhang, J. Sun et al. (2022) Detection and staging of chronic obstructive pulmonary disease using a computed tomography-based weakly supervised deep learning approach, Eur. Radiol., **32**(8), 5319–5329.
[11] Vachani A., C. Zheng, I.-L. A. Liu, B.-Z. Huang, T.-A. Osuji et al. (2022) The probability of lung cancer in patients with incidentally detected pulmonary nodules: clinical characteristics and accuracy of prediction models, Chest, **161**(2), 562–571.
[12] Razdan S., S. Sharma (2022) Internet of medical things (IoMT): overview, emerging technologies, and case studies, IETE Tech. Rev., **39**(4), 775–788.
[13] Benyahia S., B. Meftah, O. Lézoray (2022) Multi-features extraction based on deep learning for skin lesion classification, Tissue and Cell, **74**, 101701.
[14] Punitha S., T. Stephan, R. Kannan, M. Mahmud, M. S. Kaiser et al. (2023) Detecting COVID-19 from lung computed tomography images: A swarm optimized artificial neural network approach, IEEE Access, **11**, 12378–12393.
[15] Yu X., X. Ye, S. Zhang (2022) Floating pollutant image target extraction algorithm based on immune extremum region, Digit. Signal Process., **123**, 103442.
[16] Gu W., S. Bai, L. Kong (2022) A review on 2D instance segmentation based on deep neural networks, Image Vis. Comput., 104401.
[17] Ermshaus A., P. Schäfer, U. Leser (2023) ClaSP: parameter-free time series segmentation, Data Min. Knowl. Discov., **37**(3), 1262–1300.
[18] Nallasivan G., V. S. Akshaya, C. Padmavathy (2023) A computational model for medical image retrieval using orthogonal moment, J. Intell. Fuzzy Syst., **44**(2), 3261–3269.
[19] Christina Magneta S., C. Sundar, M. S. Thanabal (2023) Lung lobe segmentation and feature extraction-based hierarchical attention network for COVID-19 prediction from chest X-Ray images, Comp. J., **66**(2), 508–522.
[20] Int. J. Syst. Assur. Eng. Manag., 1–11, `https://www.kaggle.com/datasets/nih-chest-xrays/data`.

*Karpagam Academy of Higher Education*
*Coimbatore, India-6412021*
*e-mails*: `lekshmyajayan@gmail.com`
`capsridhar@gmail.com`

*[*]Sri Eshwar College of Engineering*
*ECE Department*
*Centre for Advances in Signal Processing and Sensor Networks*
*Coimbatore, India-641202*
*e-mail*: `king.pane@gmail.com`