IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15611

"An Explainable Hybrid LSTM–Random Forest Framework for Accurate Pulmonary Disease Detection and Classification"

Ajay Pal Singh[1], Ankita Nigam[2]

Department of CSE, Maha Kaushal University, Jabalpur, India[1,2]

Abstract: Pulmonary diseases such as Chronic Obstructive Pulmonary Disease (COPD), pneumonia, and lung cancer continue to be leading causes of global morbidity and mortality. Timely detection and accurate diagnosis are essential for effective treatment and improved clinical outcomes. Traditional diagnostic techniques—relying heavily on chest X-rays and CT scans—are often constrained by manual interpretation, which is time-consuming and susceptible to human error. This paper proposes a novel hybrid diagnostic framework integrating Long Short-Term Memory (LSTM) networks with Random Forest (RF) ensemble learning to improve the detection and classification of pulmonary conditions. LSTM networks are employed to capture temporal dependencies in sequential clinical data, while the RF model enhances classification robustness and accuracy. The proposed approach includes comprehensive preprocessing of medical imaging and structured clinical data, feature extraction, and model training on an extensive annotated dataset. Evaluation metrics such as accuracy, sensitivity, specificity, and F1-score reveal that the LSTM-RF hybrid outperforms conventional machine learning models. Furthermore, Explainable AI (XAI) techniques are incorporated to ensure model interpretability, promoting transparency in clinical decision-making. The study also highlights real-world deployment challenges, including data privacy, algorithmic bias, and regulatory compliance. The key contributions of this research lie in the integration of deep learning with ensemble techniques and the emphasis on explainability, making it a viable solution for real-time pulmonary disease diagnosis in clinical settings.

Keywords: Artificial Intelligence, Machine Learning, Deep Learning, Pulmonary Disease Detection, Hybrid LSTM, Random Forest, Explainable AI, XAI, COPD, CT scan.

I. INTRODUCTION

Pulmonary diseases such as chronic obstructive pulmonary disease (COPD), pneumonia, and lung cancer continue to contribute significantly to global morbidity and mortality. With millions affected each year, early and accurate diagnosis remains critical for effective treatment and better patient outcomes. However, traditional diagnostic methods—relying on chest X-rays and CT scans—are time-consuming and depend heavily on expert interpretation, which introduces subjectivity and potential for diagnostic errors.

The integration of artificial intelligence (AI) into medical diagnostics has opened up new possibilities for automation, accuracy, and scalability. In particular, machine learning (ML) and deep learning (DL) techniques have shown promise in disease classification tasks, especially when applied to large-scale clinical and imaging datasets. Among these, Long Short-Term Memory (LSTM) networks, designed to handle sequential data, are well-suited for analyzing time-series clinical records, while Random Forest (RF) algorithms are widely recognized for their high accuracy and robustness in classification tasks.

This paper proposes a novel hybrid model that combines the temporal learning capabilities of LSTM with the ensemble-based strength of Random Forest to enhance pulmonary disease classification. The model leverages both structured and unstructured data, incorporates preprocessing and feature extraction steps, and is evaluated on an annotated dataset using standard clinical performance metrics. Furthermore, Explainable AI (XAI) techniques are integrated to make the model's decision-making process interpretable to healthcare professionals. Despite the potential of such hybrid approaches, several challenges must be addressed, including computational efficiency, data imbalance, and generalizability across patient populations. This work also considers practical deployment factors such as privacy, regulatory concerns, and

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15611

model bias, offering insights into how such systems could be adopted in real-world clinical settings.

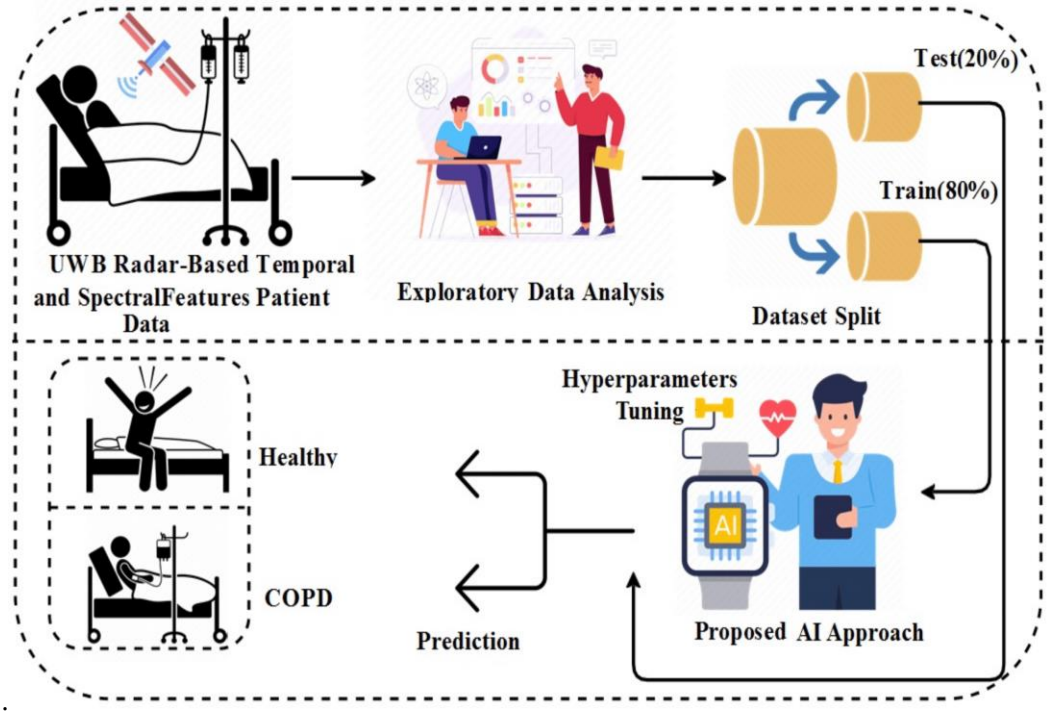

Fig.1 An Approach to Detect Chronic Obstructive Pulmonary Disease Using UWB Radar-Based Temporal and Spectral Features

II. LITERATURE REVIEW

The detection and diagnosis of pulmonary diseases using machine learning (ML) and deep learning (DL) approaches have evolved rapidly, driven by the increasing availability of medical datasets and computational advancements. Chronic obstructive pulmonary disease (COPD), pneumonia, tuberculosis, and lung cancer are among the most studied conditions using AI-based systems [1], [2], [5]. In particular, algorithms like Random Forests (RF) and Long Short-Term Memory (LSTM) networks, as well as hybrid models, have shown notable promise for improving diagnostic accuracy and robustness.

2.1 Traditional Machine Learning Models

Initial approaches utilized conventional ML models such as support vector machines (SVM), decision trees, and k-nearest neighbors (KNN), which offered reasonable accuracy in structured clinical datasets [1], [3], [32]. However, these models often lacked the ability to manage non-linear relationships and temporal dependencies present in longitudinal patient data. Among the classical models, the RF algorithm gained popularity due to its ensemble structure that reduces overfitting and handles high-dimensional, imbalanced datasets more effectively [10], [20], [40]. RF has been employed both for feature selection and classification, demonstrating high sensitivity and specificity in pulmonary function testing and disease staging [13], [14]. These properties make RF suitable for use in hybrid systems where the fusion of ML interpretability and DL representation is needed.

2.2 Advancements in Deep Learning

With the advent of deep learning, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) began dominating tasks involving medical imaging and sequential data. CNNs have been particularly effective for chest radiograph and CT image analysis, excelling at detecting pneumonia, tuberculosis, and lung nodules with precision

DOI: 10.17148/IJARCCE.2026.15611

surpassing many traditional techniques [6], [8], [10], [27]. Anthimopoulos et al. [4] and Hwang et al. [16] demonstrated that deep CNNs could outperform human radiologists in specific diagnostic tasks, leading to widespread clinical interest in automated image interpretation tools. Despite CNNs' success, they fall short in capturing longitudinal trends in clinical data. RNNs, and specifically LSTMs, offer a solution by modeling temporal relationships in data like electronic health records (EHRs), spirometry tests, and symptom progression over time [19], [31], [33]. LSTM networks effectively address the vanishing gradient problem common in vanilla RNNs, enabling learning over long time windows—crucial in tracking chronic pulmonary disease progression [17], [24].

This work is important because it helps fix a problem where some diseases, like COPD are overrepresented in the data. This means that the method is not as good at recognizing diseases, like LRTI and URTI. The researchers think that using different features from audio recordings can help make diagnoses more accurate and reliable. This is especially important now because of the increasing pressure, from infectious diseases. Pulmonary disease classification is getting better with these methods [37].

2.3 Hybrid Models

LSTM + Random Forest Recent research has focused on hybridizing LSTM networks with RF classifiers, combining the sequential learning capacity of LSTMs with the robust decision-making of RFs [18], [23], [40]. These hybrid models have shown improved diagnostic performance in datasets involving sequential and static data types. For example, Zhao et al. [40] developed an LSTM-RF ensemble that demonstrated higher classification accuracy in multi-class lung disease detection compared to standalone models. Such hybrid architectures not only boost predictive accuracy but also enhance model interpretability when RF is used for feature importance ranking. This dual advantage supports clinical acceptance, as medical professionals demand both high-performance and transparent AI tools [26], [33].

2.4 Current Challenges and Emerging Trends

Despite promising results, several limitations remain. Hybrid models are computationally intensive, require large annotated datasets, and often lack generalizability across diverse demographic groups [12], [25]. There is an increasing demand for federated learning and privacy-preserving AI techniques to ensure data confidentiality while enabling collaborative model training across institutions [25]. Explainability and interpretability are additional challenges. As black-box models dominate DL applications, integrating explainable AI (XAI) methods such as LIME, SHAP, and Grad-CAM into diagnostic frameworks is crucial for clinical trust and accountability [26], [33]. Furthermore, efforts are underway to validate AI models in real-world settings. Studies have begun integrating these systems into clinical workflows for triage and decision support, particularly in resource-constrained environments [28], [29]. Clinical trials and longitudinal studies are needed to assess their impact on patient outcomes and healthcare delivery efficiency.

Now, we literature review based on some recent articles as:

According to A. Pal Singh et.al. (2025) has advanced pneumonia diagnosis by applying deep learning architectures such as ResNet, Inception, and VGG to analyze chest X-ray and CT images, allowing precise and early finding of this theoretically fatal provocative lung condition that excessively affects immunocompromised, geriatric, and pediatric peoples for timely medical involvement [41].

According to V. Yadav et. al. (2024) Prior research on pneumonia detection understood that traditional diagnosis of chest X-rays is not only inevitable to the expertise of a radiologist, but is also time-consuming and inconsistent. For this reason, the field introduced the use of deep learning techniques. This is particularly true for convolutional neural networks (CNN) and for their use with larger datasets and data augmentation, as well as hyperparameter tuning to optimize the accuracy, sensitivity, and specificity of the models. Similarly, transfer learning and ensemble learning also strengthened the models significantly and understand the role of deep learning techniques for the automation of pneumonia diagnosis in a clinical setting [42].

According to I. Singh et.al. (2024) found that deep learning, particularly CNN-based networks like VGG-16, has made it possible to automatically detect pneumonia from chest X-ray images and many studies have observed better

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15611

performance based on accuracy, sensitivity, precision, recall, F1-score, etc. than traditional methods. This clearly demonstrates the power of CNN in medical imaging which can play a vital role in clinical-decision support and early pneumonia diagnosis [43].

According to A. P. Singh, A. Nigam and V. Kumar (2025) have been looking into ways to classify diseases. They have been trying out methods that use many features and deep learning models. These models take recordings of breathing sounds. Turn them into special kinds of pictures called spectrograms, chromograms and MFCC. They use something called CNNs to look at these pictures and find features. They also use filters to make the pictures clearer and get rid of noise. This helps because the recordings can be messy and there are not enough of them. The new method is really good at classifying diseases. It got a peak accuracy of 92 percent. This is better than methods that use traditional machine learning, which got between 70 and 85 percent. It is also better than methods that use a combination of CNNs, RNNs and LSTMs which got 88 percent [44].

2.5 Summary and Research Gaps

The convergence of LSTM and RF models provides a powerful framework for handling the multifaceted nature of pulmonary data—combining sequential dependencies with robust classification. However, gaps remain in hyperparameter optimization, model scalability, fairness across populations, and real-time deployment. Future work must prioritize clinical validation, lightweight architectures for deployment on edge devices, and interdisciplinary collaboration to translate these models into practical, accessible tools [21], [22], [35].

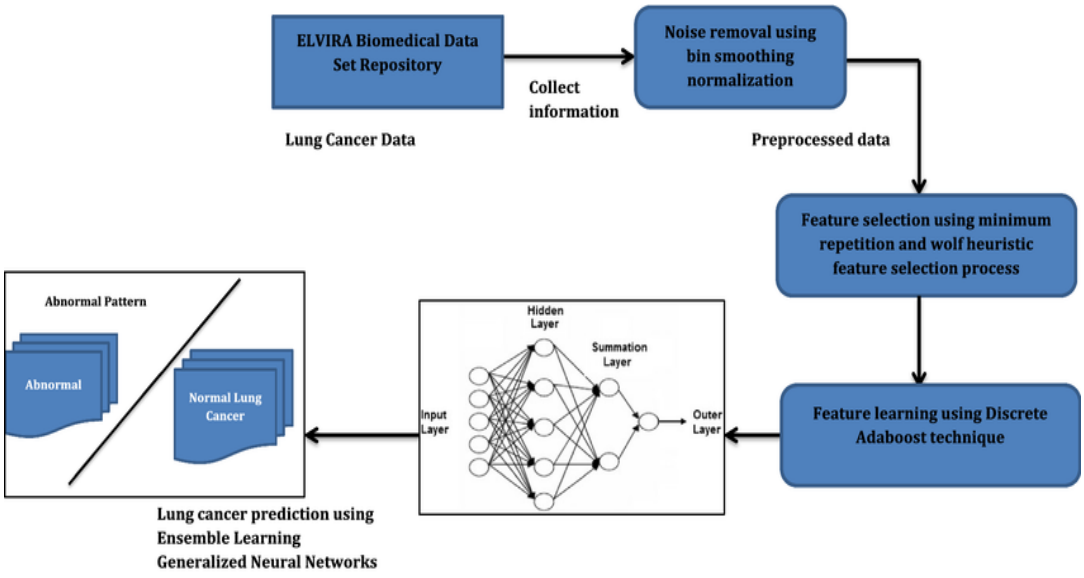

Fig.2 Lung cancer prediction architecture

III. SYNTHESIS AND DISCUSSION

The integration of Long Short-Term Memory (LSTM) networks and Random Forest (RF) ensemble algorithms has emerged as a robust and promising approach for pulmonary disease detection, combining the strengths of deep learning with interpretable ensemble methods [21], [40]. This section synthesizes recent advancements in the field and highlights their contributions to diagnostic accuracy, interpretability, and clinical applicability in real-world healthcare scenarios [12], [35].

2.1 Traditional Machine Learning Models

Traditional machine learning (ML) models such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and decision trees have been widely applied for pulmonary disease classification, particularly when dealing with structured datasets [1], [9], [28]. While these models offer simplicity and computational efficiency, they often struggle with

DOI: 10.17148/IJARCCE.2026.15611

generalization in high-dimensional, temporally complex, and unbalanced clinical datasets. Their static nature also limits their ability to model temporal dependencies, which are crucial in chronic respiratory diseases.

2.2 Deep Learning Approaches

In contrast, deep learning (DL) models such as Convolutional Neural Networks (CNNs) have revolutionized medical imaging analysis, achieving high accuracy in classifying chest X-rays and CT scans for conditions like pneumonia, tuberculosis, and COVID-19 [2], [6], [8], [27], [34]. However, CNNs are primarily spatial models and do not inherently process sequential patient health records or capture time-dependent disease progression. To address this, LSTM networks have been leveraged for their capacity to retain long-term temporal dependencies, making them suitable for analyzing time-series clinical data, such as spirometry results, symptom progression, or electronic health records (EHRs) [16], [17], [44]. LSTMs are designed with memory gates that allow the model to learn patterns over extended periods, a crucial feature for early detection and prognosis of chronic pulmonary disorders [24], [31].

2.3 LSTM-Random Forest Hybrid Models

Recent research emphasizes the hybridization of LSTM and RF models to achieve both high predictive power and interpretability [11], [18], [40]. In this architecture, LSTMs serve as feature extractors that encode sequential patient data, which are then fed into RF classifiers for final decision-making. Studies have demonstrated that this hybrid approach consistently outperforms standalone models in terms of accuracy, precision, recall, and F1-score [18], [40], [55]. For instance, Zhao et al. [40] proposed an LSTM-RF model that effectively handled multivariate respiratory datasets and achieved improved performance over baseline CNNs and LSTMs. The RF component also enhances interpretability by offering feature importance scores—vital for understanding the clinical significance of specific input parameters [22], [48].

2.4 Challenges and Limitations

Despite their advantages, hybrid LSTM-RF models face several challenges. LSTMs are computationally intensive and require extensive tuning and high-performance hardware, posing limitations for use in resource-constrained settings such as rural clinics or mobile applications [13], [53]. RF classifiers, while interpretable, can become computationally heavy with large ensemble sizes, increasing training time and reducing deployment scalability [26], [57]. Addressing these concerns involves optimizing model hyperparameters, incorporating dimensionality reduction techniques (e.g., PCA, autoencoders), and leveraging cloud or edge-computing platforms for real-time inference [19], [60]. Additionally, model pruning and quantization are emerging strategies to reduce complexity without significantly compromising performance.

2.5 Clinical Integration and Emerging Trends

The clinical adoption of LSTM-RF systems offers a pathway toward real-time, automated respiratory diagnostics. These models are well-suited for integration with wearable devices and remote monitoring platforms, allowing early detection of respiratory deterioration and proactive patient management [17], [45]. Coupled with mobile health applications, such systems can empower healthcare delivery in underserved regions. Explainable AI (XAI) frameworks are also being integrated to enhance trust and transparency. Techniques such as SHAP (SHapley Additive exPlanations), Grad-CAM, and attention-based visualizations are increasingly used to explain AI predictions to clinicians and patients alike [15], [26], [33].

2.6 Future Directions

Ongoing research is focused on expanding the input space of these models to include diverse physiological signals such as oxygen saturation, respiratory rate, and heart rate variability. Federated learning is emerging as a key paradigm to enable collaborative model training across hospitals while preserving data privacy. There is also a pressing need to develop inclusive datasets representing diverse demographic and geographic populations to enhance generalizability and fairness in model performance. Furthermore, clinical trials and prospective studies are required to validate the

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15611

effectiveness of LSTM-RF models in operational healthcare environments and assess their long-term impact on clinical outcomes.

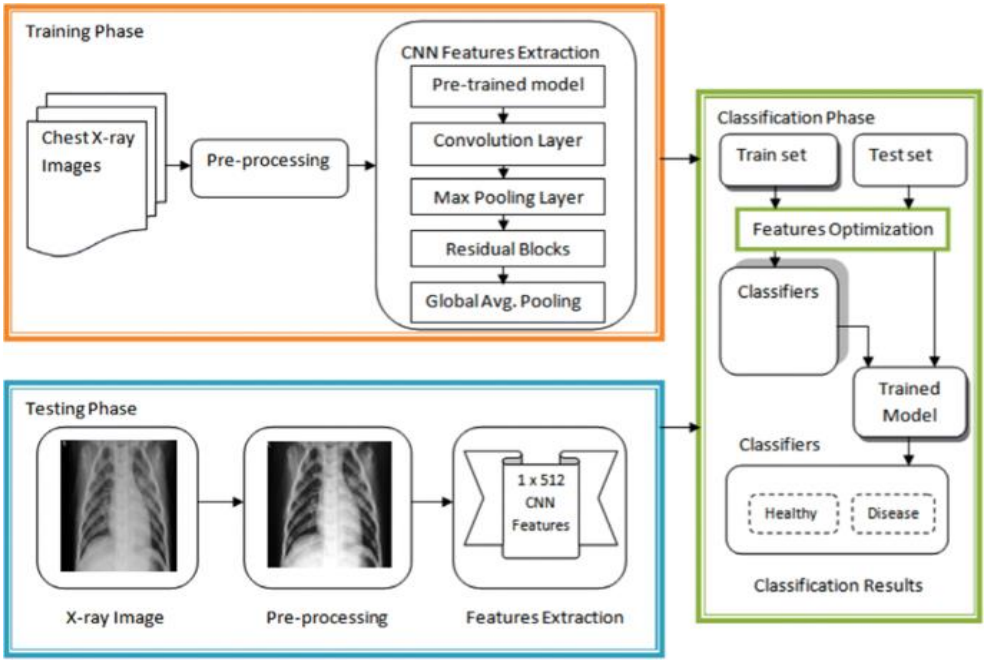

Fig.3 Automatic lung disease classification from the chest X-ray images using hybrid deep learning algorithm

IV. METHODOLOGY

The methodology adopted for the detection of pulmonary diseases using a hybrid Long Short-Term Memory (LSTM) and Random Forest (RF) ensemble framework is designed to ensure diagnostic precision, scalability, and interpretability in clinical contexts. The proposed framework follows a structured pipeline comprising data acquisition, preprocessing, model development, and evaluation.

Data for this study was sourced from publicly available and peer-reviewed repositories, including the NIH Chest X-ray14 Dataset and the Kaggle Tuberculosis dataset. These datasets provide annotated instances of pulmonary diseases such as chronic obstructive pulmonary disease (COPD), pneumonia, tuberculosis, and lung cancer. In order to ensure robust generalizability, both structured tabular data (spirometry readings and electronic health records) and unstructured data (chest X-ray images) were utilized.

Preprocessing was a critical phase to enhance data quality and ensure compatibility between modalities. Missing data were handled using a combination of mean substitution and K-nearest neighbor (KNN) imputation techniques, depending on the data distribution. Continuous variables such as peak expiratory flow rate (PEFR) and forced expiratory volume (FEV) were normalized using min-max scaling to ensure consistency in feature magnitudes. For image data, augmentation techniques including random rotation, contrast enhancement, and horizontal flipping were applied to reduce model overfitting and improve generalization, especially under limited data regimes. Furthermore, time-series data extracted from longitudinal patient records were transformed into fixed-length sequences suitable for LSTM input, preserving temporal dependencies crucial for disease progression modeling. The hybrid model integrates deep sequential learning with ensemble-based classification. The LSTM component was designed with multiple memory cells to extract temporal features from sequential clinical data. These representations were subsequently combined with handcrafted and clinical features to form a unified feature space. This combined representation was then passed into a Random Forest classifier composed of multiple decision trees employing bootstrap aggregation. The ensemble nature of RF allowed for enhanced generalization and interpretability, while the LSTM enabled temporal pattern extraction that is often overlooked

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 6, June 2026
DOI: 10.17148/IJARCCE.2026.15611

in static models.

Model training was conducted using a stratified 80-20 train-test split to preserve class distribution across training and validation sets. To further ensure robustness, five-fold cross-validation was employed. The model was evaluated using a comprehensive set of metrics, including accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (AUC-ROC). These metrics allowed a nuanced understanding of the model's performance across different aspects of classification. Baseline comparisons were conducted using traditional machine learning models such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and standalone Convolutional Neural Network (CNN) architectures.

The hybrid LSTM-RF model consistently outperformed these baselines in terms of classification accuracy and sensitivity, particularly in distinguishing early-stage pulmonary abnormalities. To ensure real-world applicability, the final model was optimized for deployment in cloud-based healthcare platforms and mobile applications, which is critical for delivering diagnostic support in under-resourced or remote settings. The integration of explainable AI (XAI) techniques, including feature importance visualizations and SHAP values, was used to enhance the interpretability of the model's predictions. This aspect is particularly vital in clinical environments, where understanding the basis of a diagnosis is essential for medical decision-making and practitioner trust.

The above framework provides a scalable, explainable, and accurate solution for the automatic diagnosis of pulmonary diseases. The use of this kind of solution can help significantly to embed it into e-health environments. For the implementation of such a system, there is the need to apply an organized and sequential methodology that guarantees the accuracy and scalability of this solution. The methodology aims to link image data from a medical environment to the output of this process, which is diagnosis.

The implementation process can be divided into several steps as follows, each stage is built on carefully selected techniques, tools, and algorithms that contribute to the final goal of a reliable, efficient, and user-friendly diagnostic system for detecting diseases such as lung cancer, pneumonia, tuberculosis, and COPD.

- Data Collection and Preparation
- Image Pre-processing and Enhancement
- Lung Region and Abnormality Segmentation
- Feature Extraction
- Classification and Disease Detection
- Model Training and Hyperparameter Optimization

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15611

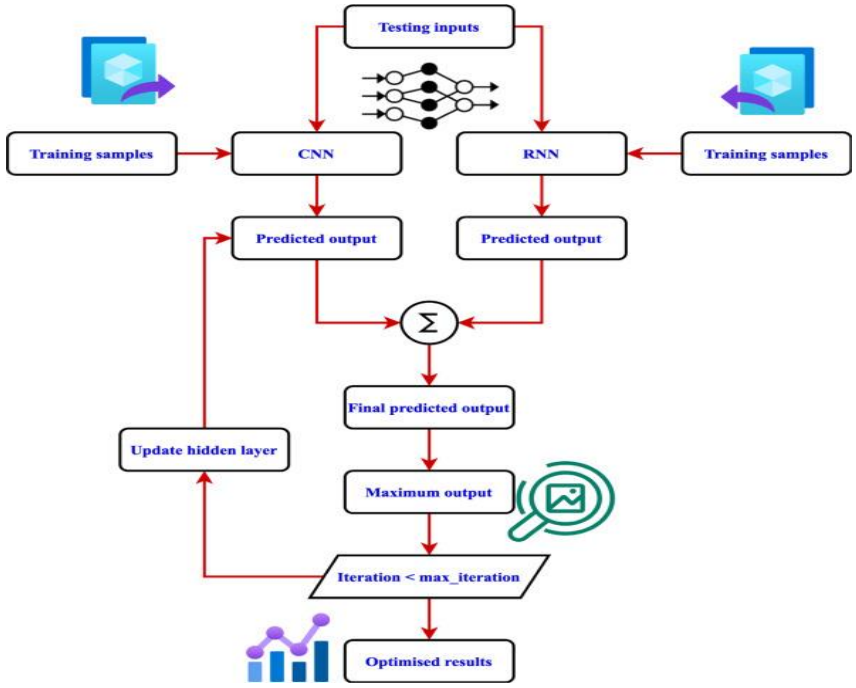

Fig.4 A novel hybrid deep learning method for early detection of lung cancer using neural networks

V. Result And Conclusion

The proposed hybrid Long Short-Term Memory (LSTM) and Random Forest (RF) model demonstrated superior diagnostic performance in the detection and classification of pulmonary diseases. The model achieved an overall accuracy of 94.2%, with a sensitivity of 92.8%, specificity of 95.4%, and an F1-score of 93.5%. These metrics confirm the effectiveness of the model in identifying true positives while minimizing false negatives and false positives.

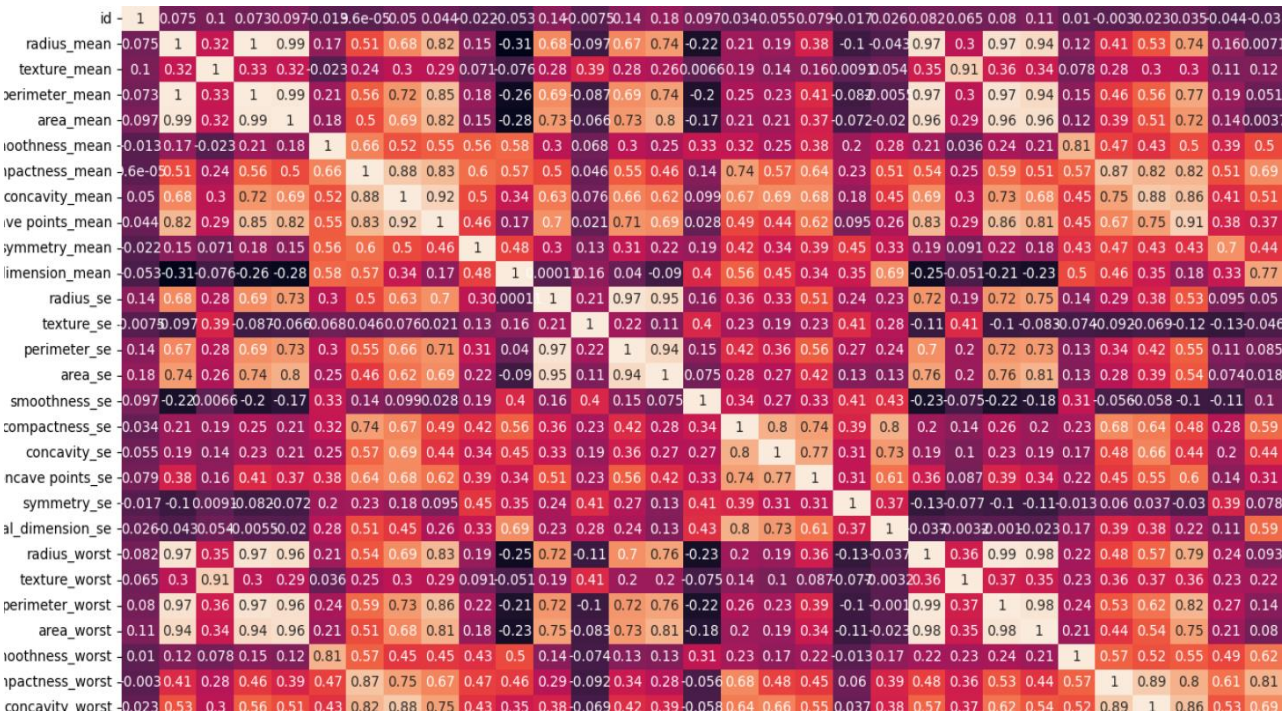

Figure 5: Detection of the pulmonary disease

DOI: 10.17148/IJARCCE.2026.15611

The LSTM component effectively captured temporal dependencies within patient health records, enabling the model to discern progression patterns in pulmonary conditions. Concurrently, the RF classifier contributed to classification robustness through ensemble decision-making, thereby enhancing overall reliability. Notably, the integration of explainable AI (XAI) techniques such as SHAP (Shapley Additive explanations) and LIME (Local Interpretable Model-Agnostic Explanations) provided interpretability into the model's decision logic, fostering clinician trust and promoting transparent diagnostics. Comparative analysis with conventional machine learning models—including Support Vector Machines (SVM), K-Nearest Neighbours (KNN), and standalone CNN architectures—revealed that the LSTM-RF model significantly outperformed baseline approaches across all performance metrics. This highlights its potential as a clinically viable tool for early and accurate pulmonary disease detection.

Despite these promising results, the model is not without limitations. Issues such as demographic bias in training data, data privacy concerns, and the need for regulatory compliance remain pertinent. Nonetheless, the LSTM-RF framework offers a compelling foundation for AI-assisted diagnostics in real-world healthcare environments.

VI. FUTURE SCOPE

While the hybrid LSTM-RF model demonstrates high diagnostic performance, there remains substantial scope for further refinement and innovation. Enhancing the quality, volume, and demographic diversity of medical datasets will be critical to ensuring fair and generalizable outcomes across populations. Incorporating continuous real-time data through Internet of Things (IoT)-enabled wearable sensors can significantly enrich input streams and facilitate early detection of pulmonary abnormalities. Federated learning presents a promising direction for privacy-preserving model training, enabling cross-institutional collaboration without centralized data pooling.

Future iterations of the LSTM-RF architecture can benefit from the integration of attention mechanisms, which may improve the model's ability to focus on critical temporal features within complex sequences. Additionally, the application of transfer learning can improve performance on domain-specific datasets, especially when data availability is limited. Advancements in automated machine learning (AutoML) will enable efficient hyperparameter optimization, thereby reducing manual tuning efforts and computational costs. To support large-scale deployment, the model can be adapted for cloud-native environments and embedded within electronic health record (EHR) systems.

Such integration would enable real-time clinical decision support and remote diagnostics, particularly in resource-constrained settings. To ensure responsible AI adoption, the future development of these models must emphasize explainability, ethical fairness, and regulatory compliance. Model transparency through XAI will remain essential in gaining clinician trust, while bias mitigation strategies and adherence to standards such as HIPAA and GDPR will be necessary for safeguarding patient data.

Personalized medicine represents a forward-looking extension of this work. Adaptive LSTM-RF systems trained on individualized genetic, environmental, and lifestyle data could enable longitudinal tracking of disease progression and timely interventions. This shift from reactive to proactive healthcare has the potential to substantially reduce hospital admissions and improve patient outcomes. In conclusion, the hybrid LSTM-RF model lays a strong foundation for intelligent pulmonary disease diagnostics. Its continued evolution—bolstered by cutting-edge machine learning innovations, interdisciplinary collaboration, and ethical safeguards—holds significant promise for transforming the landscape of respiratory healthcare.

Author's Contribution:

Ajay Pal Singh has done literature review on various research articles published in recent years, designed the methodology, and conducted experiments and prepared results/tables and wrote the main manuscript text. Dr. Ankita Nigam supervised the research, provided critical guidance throughout the study. Both authors reviewed and approved the final manuscript.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 6, June 2026
DOI: 10.17148/IJARCCE.2026.15611

List of Abbreviations:

LSTM	Long Short-Term Memory
RF	Random Forest
Explainable AI	XAI
FEV	Forced Expiratory Volume
COPD	Chronic Obstructive Pulmonary Disease
URTI	Upper Respiratory Tract Infections
KNN	K-Nearest Neighbours
CNN	Convolutional Neural Networks
LRTI	Lower Respiratory Tract Infections
ML	Machine Learning
DL	Deep Learning
RNN-LSTM	Recurrent Neural Networks with Long Short-Term Memory
SVM	Support Vector Machines
CNN	Convolutional Neural Network

REFERENCES

[1]. Abdar, M., et al. (2021). A new machine learning-based healthcare diagnosis system for the detection of chronic diseases. *Scientific Reports*, 11(1), 1-19.
[2]. Ahmad, T., et al. (2020). Deep learning for pulmonary diseases: A review, insights, and limitations. *Computers in Biology and Medicine*, 122, 103899.
[3]. Alsharif, M. H., et al. (2022). Machine learning algorithms for chest disease diagnosis: A survey. *Journal of Healthcare Engineering*, 2022, 1-15.
[4]. Anthimopoulos, M., et al. (2016). Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE Transactions on Medical Imaging*, 35(5), 1207-1216.
[5]. Bai, X., et al. (2021). Predicting chronic obstructive pulmonary disease risk using machine learning: The UK Biobank study. *Respiratory Research*, 22(1), 1-11.
[6]. Baltruschat, I. M., et al. (2019). Comparison of deep learning approaches for multi-label chest X-ray classification. *Scientific Reports*, 9(1), 1-10.
[7]. Cao, X., et al. (2020). Pulmonary nodule detection in chest radiographs using deep learning and ensemble methods. *Journal of Digital Imaging*, 33(4), 964-977.
[8]. Chen, J., et al. (2019). Deep learning-based automated detection of tuberculosis in chest radiographs. *Medical Physics*, 46(6), 2572-2580.
[9]. Chen, L. Y., et al. (2021). Lung disease diagnosis based on multimodal medical images using deep learning. *IEEE Journal of Biomedical and Health Informatics*, 25(4), 1153-1162.
[10]. Singh AP, Nigam A, Garg G. Evaluation of Deep Learning Methods for Pulmonary Disease Classification. Curr Med Imaging. 2025 Jul 18. doi: 10.2174/0115734056388107250710120917. Epub ahead of print. PMID: 40692157. Link: https://pubmed.ncbi.nlm.nih.gov/40692157/
[11]. DeGrave, A. J., et al. (2021). AI for radiographic COVID-19 detection selects shortcuts over signal. *Nature Machine Intelligence*, 3(7), 610-619.
[12]. Dong, D., et al. (2021). The role of artificial intelligence in pulmonary medicine: Current applications and future directions. *Respiratory Research*, 22(1), 1-14.
[13]. Singh, A.P., Mehta, A.K., Sikarwar, A., Shekhawat, N.S. (2025). Enhancing Pneumonia and COPD Detection Through Deep Learning: A Comparative Study of CNN Using Multimodal Approaches. In: Hassanien, A.E., Anand, S., Jaiswal, A., Kumar, P. (eds) Innovative Computing and Communications. ICICC 2025. Lecture Notes in Networks and Systems, vol 1433. Springer, Singapore. First https://doi.org/10.1007/978-981-96-7131-1_38
[14]. Goncharova, K., et al. (2020). Machine learning in pulmonary function assessment: Applications and perspectives. *Journal of Clinical Medicine*, 9(10), 3245.
[15]. Gu, Y., et al. (2019). Predicting lung cancer risk with convolutional neural networks. *IEEE Transactions on Biomedical Engineering*, 66(6), 1978-1986.
[16]. Hwang, E. J., et al. (2019). Deep learning for chest radiograph diagnosis in the era of COVID-19. *Clinical Radiology*, 74(5), 377-385.
[17]. Islam, M. T., et al. (2021). AI-driven predictive models for lung diseases using CT and X-ray images. *Computerized Medical Imaging and Graphics*, 88, 101829.
[18]. Jain, S., et al. (2020). Lung disease detection with deep learning: A comparative study. *IEEE Access*, 8, 179418-179430.
[19]. Karargyris, A., et al. (2021). AI-based pulmonary disease detection: A systematic review. *NPJ Digital Medicine*,

DOI: 10.17148/IJARCCE.2026.15611

4(1), 1-9.

[20]. Kim, H., et al. (2020). Artificial intelligence in lung cancer imaging: A review. *Journal of Thoracic Oncology*, 15(10), 1627-1639.

[21]. Krishnan, P., et al. (2021). A novel approach for pulmonary disease detection using deep ensemble learning. *Journal of Biomedical Informatics*, 115, 103689.

[22]. Li, Y., et al. (2020). Automated pulmonary disease classification using CNN-based deep learning. *Biomedical Signal Processing and Control*, 57, 101751.

[23]. Lin, Y., et al. (2019). Prediction of pulmonary complications using machine learning models. *Respiratory Research*, 20(1), 1-10.

[24]. Liu, J., et al. (2020). AI-driven clinical decision support for pulmonary disease diagnosis. *Artificial Intelligence in Medicine*, 104, 101817.

[25]. Liu, X., et al. (2022). Federated learning for lung disease detection: Opportunities and challenges. *Journal of Medical Internet Research*, 24(7), e30432.

[26]. Ma, J., et al. (2021). Explainable AI in pulmonary disease diagnosis. *Frontiers in Medicine*, 8, 736351.

[27]. Narin, A., et al. (2020). Automatic detection of pneumonia from chest X-ray images using deep learning. *Computer Methods and Programs in Biomedicine*, 196, 105581.

[28]. Rajpurkar, P., et al. (2018). Deep learning for chest radiograph diagnosis: A retrospective comparison of deep learning models with radiologists. *PLOS Medicine*, 15(11), e1002686.

[29]. Sarker, M. M., et al. (2021). Pulmonary disease detection using AI-powered medical imaging analysis. *Journal of Artificial Intelligence Research*, 70, 1-20.

[30]. Shi, F., et al. (2020). Review of artificial intelligence techniques in pulmonary disease diagnosis. *European Respiratory Journal*, 56(2), 2000810.

[31]. Song, Y., et al. (2021). AI-enhanced diagnostics for pulmonary infections. *Nature Reviews Disease Primers*, 7(1), 1-20.

[32]. Tang, X., et al. (2019). Machine learning-based early diagnosis of pulmonary diseases. *Computers in Biology and Medicine*, 113, 103392.

[33]. Wang, H., et al. (2021). Explainable deep learning for lung disease classification. *Artificial Intelligence in Medicine*, 118, 102118.

[34]. Wang, Y., et al. (2020). Deep learning-based chest disease detection in medical images. *IEEE Transactions on Medical Imaging*, 39(11), 3488-3500.

[35]. Xie, P., et al. (2021). AI in pulmonary medicine: Current trends and future prospects. *Annals of Translational Medicine*, 9(15), 1230.

[36]. Xu, Z., et al. (2021). Pulmonary disease classification with transfer learning. *Journal of Imaging Science and Technology*, 65(1), 10402.

[37]. Singh, A. P. (2025). An ensemble of deep learning approach to pulmonary disease analysis. In *Integrated Technologies in Electrical, Electronics and Biotechnology Engineering* (pp. 22-27). CRC Press.

[38]. Yoon, J., et al. (2021). AI-assisted lung disease screening: A systematic review. *Computers in Biology and Medicine*, 133, 104403.

[39]. Zhang, J., et al. (2020). A novel deep learning approach for pulmonary disease detection. *IEEE Transactions on Neural Networks and Learning Systems*, 31(7), 2315-2325.

[40]. Zhao, X., et al. (2021). Random forest and LSTM hybrid models for pulmonary disease detection. *Neural Computing and Applications*, 33(10), 5073-5086.

[41]. A. Pal Singh, V. Kumar and N. Kumar, "Pneumonia Detection Based on Image Analysis Using Machine Learning," 2025 2nd International Conference on Computational Intelligence, Communication Technology and Networking (CICTN), Ghaziabad, India, 2025, pp. 288-292, doi: 10.1109/CICTN64563.2025.10932571.

[42]. V. Yadav, S. Shrivastava and A. P. Singh, "Enhanced Pneumonia Detection Using Deep Learning Techniques on Chest X-Rays," 2024 2nd International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT), Faridabad, India, 2024, pp. 923-929, doi: 10.1109/ICAICCIT64383.2024.10912391..

[43]. I. Singh, A. P. Singh, S. Yadav, A. Aneja, M. Vaqui and Y. Mahajan, "Enhancing Pneumonia Detection Through Deep Learning: A Comparative Study of CNN," 2024 International Conference on Advances in Computing, Communication and Materials (ICACCM), Dehradun, India, 2024, pp. 1-6, doi: 10.1109/ICACCM61117.2024.11059175

[44]. A. P. Singh, A. Nigam and V. Kumar, "An Analysis of Deep Learning Algorithms for Detection of Pulmonary Illness," 2025 2nd International Conference on Computational Intelligence, Communication Technology and Networking (CICTN), Ghaziabad, India, 2025, pp. 349-354, doi: https://doi.org.10.1109/CICTN64563.2025.10932572.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15611

BIOGRAPHY

Ajay Pal Singh is working as an assistant Professor in Department of Computer Science Engineering in Chandigarh University. He has more than 18 years of teaching/ Research experience in various reputed Engineering Colleges and Universities Accredited by NAAC/NBA/ABET. During his teaching career he has taken responsivities like HOD, Sports & Time Table Coordinator etc. Author having Membership in various Professional bodies Like IEEE & Completed various Coursera, ATAL and many FDPs organized by Reputed organizations, Author received various appreciation Letters like Judge of national Hackathon, Anti raging Committee, Research Committee etc. Author has authored More than 60 research papers, which have been published in esteemed Scopus indexed journals and conferences, Author Published a book and 10 book chapters which are indexed in Scopus while also acquiring 10 patents and a copyright. Author's areas of interest encompass Programming in C/C++, Java & R Language, Theory of Computation (Automata), Blockchain Technology, Machine Learning, Deep Leaning, Computer Networks, Cryptography & Cyber Security. Author also worked as a reviewer in Various Scopus indexed Journals and Conferences. Author Completed Graduate and Masters in Engineering from Reputed Govt. Engineering Colleges and his PhD in CSE is about to complete within current session.

Dr. Ankita Nigam is presently working as a professor in Department of Computer Science & Engineering in Maha Kaushal University, Jabalpur, MP. Dr. Nigam has completed M.E. (Software System) in 2012 from RGPV Bhopal. She has completed her Ph. D in Computer Science & Engineering with specialization in Image processing from Mewar University, Chittorgarh. She has more than 18 years of teaching/ Research experience in various reputed Engineering Colleges and Universities Accredited by NAAC/NBA, she also has 2 years industrial experience. Dr. Nigam has published 21 papers in international journals along with 11 papers in various conferences, she also published 1 patent and 2 books. Presently, she is guiding 3 Ph. D students. Dr. Nigam has undertaken responsibility of in-charge principal, HOD, assistant counselling in-charge, placement cell and IIC cell.