RESEARCH PAPER

Chronic Obstructive Pulmonary Disease

Kausar Shameem[1] and Shilpa.B.Kodli[2]

[1]*Student, Department of Computer Science and Engineering, VTU's CPGS Kalaburagi, Karnataka, India*
[2]*Assistant Professor, Department of Computer Science and Engineering, VTU's CPGS Kalaburagi, Karnataka, India*

Received: 28th Feb, 2026; Revised: 6th March 2026; Accepted: 7th April, 2026; Available Online: 20th April, 2026

ABSTRACT

Chronic Obstructive Pulmonary Disease (COPD) is a life-threatening lung disorder characterized by persistent respiratory symptoms and airflow limitation. Early detection and accurate classification of COPD are critical for timely intervention and effective treatment. In recent years, Machine Learning (ML) and Deep Learning (DL) models have emerged as powerful tools for enhancing COPD diagnosis and prognosis. This study explores the application of various ML algorithms, including Support Vector Machines (SVM), Random Forest (RF), Gradient Boosting (XGBoost), and k-Nearest Neighbors (k-NN), for COPD prediction based on patient clinical data and spirometry results. Additionally, Convolutional Neural Networks (CNN) are employed for image-based classification using chest X-rays and CT scans, enabling automated detection of COPD-related abnormalities. The comparative analysis demonstrates the potential of these models in improving diagnostic accuracy, reducing human error, and facilitating early-stage COPD identification. The integration of ML and DL techniques holds promise for advancing clinical decision-making and personalized healthcare for COPD patients.

Index Terms: *Chronic Obstructive Pulmonary Disease (COPD, Support Vector Machines (SVM), Random Forest (RF), Gradient Boosting (XGBoost), and k-Nearest Neighbors (k-NN)*

How to cite this article: Shameem K, Kodli SB. Chronic Obstructive Pulmonary Disease. Int J Drug Deliv Technol. 2026;16(61s):1631-1635. DOI: 10.25258/ijddt.16.61s.185

Source of support: Nil.

Conflict of interest: None

I. INTRODUCTION

Chronic Obstructive Pulmonary Disease (COPD) is a progressive respiratory disorder characterized by airflow limitation, chronic inflammation, and irreversible lung tissue damage. Globally, COPD ranks among the leading causes of morbidity and mortality, placing a significant burden on healthcare systems and affecting patients' quality of life. Early diagnosis and timely treatment are crucial to managing disease progression and reducing associated complications.

Traditional diagnostic methods, such as spirometry, chest radiography, and clinical assessment, though effective, are often time-consuming and subject to variability in interpretation. With the advent of Artificial Intelligence (AI), particularly Machine Learning (ML) and Deep Learning (DL) techniques, there has been a paradigm shift in medical diagnostics and predictive analytics.

ML algorithms like Support Vector Machines (SVM), Random Forest (RF), Gradient Boosting (XGBoost), and k-Nearest Neighbors (k-NN) have shown promising results in analyzing clinical and spirometric data for COPD risk prediction and classification. These models excel at recognizing patterns in high-dimensional datasets, enabling more accurate and automated diagnostic support.

On the other hand, Deep Learning models, particularly Convolutional Neural Networks (CNN), have demonstrated exceptional performance in medical image analysis. CNN-based systems can automatically extract features from chest X-rays and CT scans, facilitating the detection of COPD-induced structural changes in the lungs without manual intervention.

This study aims to explore the application of these ML and DL models in COPD detection, comparing their performance, and highlighting their potential in enhancing early diagnosis, reducing diagnostic errors, and supporting clinical decision-making in pulmonary care...

II. LITERATURE SURVEY

[1] Fischer et al. (2020) evaluated AI-based automatic chest CT emphysema quantification compared to pulmonary function testing, noting that while promising, their study was limited to emphysema quantification and may not generalize to all COPD phenotypes.

[2] Makimoto et al. (2023) compared various feature selection methods and ML classifiers for predicting COPD using texture-based CT lung radiomic features, highlighting the risk of overfitting due to high-dimensional data and the necessity for external validation.

[3] Wang et al. (2023) developed a DL model for COPD diagnosis using chest X-rays across multiple sites and modalities, although their study faced challenges due to variability in imaging protocols, underlining the need for standardization.

**Author for Correspondence: Kausar Shameem*

[4] Altan et al. (2019) applied DL techniques for computerized analysis of COPD using medical imaging, but their findings were constrained by a limited dataset size, potentially affecting generalizability.

[5] Haider et al. (2019) utilized ML algorithms to classify COPD severity based on respiratory sounds, yet variability in sound recording quality necessitated standardized protocols for broader application.

[6] Amaral et al. (2012) implemented ML algorithms combined with forced oscillation measurements for COPD identification; however, the study's age suggests it may not reflect the latest ML advancements.

[7] Kanwade and Bairagi (2019) classified COPD and normal lung airways using electromyographic signal features but noted limitations in sample size and called for broader validation.

[8] Siddiqui et al. (2022) incorporated ultra-wideband radar with ML algorithms for respiration-based COPD detection, requiring further validation in clinical environments. [9]Cosentino et al. (2023) employed DL on raw spirogram data for COPD inference and genetic loci identification, though integrating genetic data remains complex and demands replication studies.

[10] Wu et al. (2023) provided a comprehensive review of ML applications in COPD, including predictive models for exacerbations, though specific model performance metrics were not detailed, limiting direct applicability to model selection.

III. PROPOSED SYSTEM

The proposed system focuses on the early detection and classification of Chronic Obstructive Pulmonary Disease (COPD) by integrating Machine Learning (ML) and Deep Learning (DL) techniques within a unified diagnostic framework. The system utilizes structured clinical data such as patient demographics, spirometry readings, smoking history, and medical history, alongside medical imaging data like chest X-rays and CT scans. Clinical data undergo preprocessing steps including handling missing values, normalization, and feature selection to ensure data quality and relevance. For the analysis of clinical data, several ML algorithms—Support Vector Machines (SVM), Random Forest (RF), Gradient Boosting (XGBoost), and k-Nearest Neighbors (k-NN)—are employed to predict COPD presence and severity based on patient-specific features.

Parallelly, a Convolutional Neural Network (CNN) is designed for the classification of image data, focusing on detecting COPD-specific patterns such as emphysematous changes, airway obstructions, and structural lung abnormalities. The CNN model processes preprocessed and augmented images, learning hierarchical features critical for accurate classification. Both ML and DL models are evaluated using metrics like accuracy, precision, recall, F1-score, and ROC-AUC to validate their performance. Finally, the system integrates the predictions from both ML and DL models into a combined interface, offering a comprehensive diagnostic suggestion to assist clinicians in decision-making. This hybrid approach aims to enhance diagnostic accuracy, support early risk assessment, and ultimately improve patient outcomes in COPD management.

IV. METHODOLOGY

Step 1 – Data Collection
Collect clinical data (age, gender, spirometry results, smoking history, symptoms).

Acquire medical images (chest X-rays, CT scans) from clinical databases or hospitals.

Step 2 – Data Preprocessing

Clinical Data:
Handle missing values (imputation/removal).

Normalize numerical attributes.

Encode categorical variables.

Apply feature selection methods (e.g., correlation analysis, PCA).

Image Data:
Resize images to standard input size.

Apply contrast enhancement and noise reduction.

Perform data augmentation (rotation, flipping) to increase dataset diversity.

Step 3 – Model Development
Machine Learning Models (for clinical data):

Train classifiers:
Support Vector Machines (SVM)

Random Forest (RF)

Gradient Boosting (XGBoost)

k-Nearest Neighbors (k-NN)

Use cross-validation for reliable evaluation.

Optimize hyperparameters using Grid Search or Random Search.

Deep Learning Model (for image data):
Design and build a Convolutional Neural Network (CNN).

Train CNN on preprocessed images with labeled outcomes.

Apply dropout and batch normalization to prevent overfitting.

Step 4 – Model Evaluation
Evaluate all models on testing datasets.

Use metrics such as:
Accuracy

Precision

Recall

F1-Score

ROC-AUC

Step 5 – System Integration
Integrate ML and DL models into a unified diagnostic tool.

Combine clinical prediction and image-based classification outputs.

Design a user-friendly interface for clinicians to input data and view predictions.

Step 6 – Validation & Deployment
Validate system performance on unseen clinical datasets.

Deploy as a clinical decision support tool or web application.

Monitor model performance and update periodically with new data.

V. RESULTS

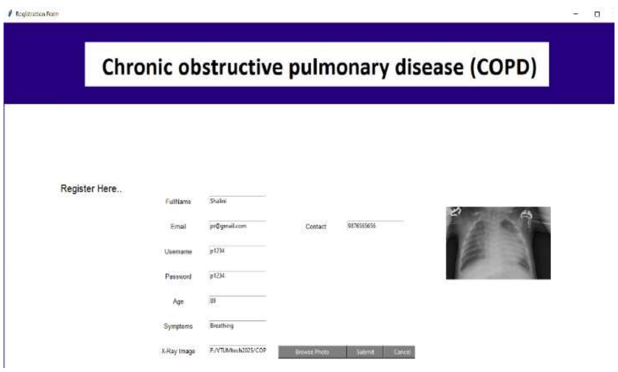


Figure 1: Patient Registration

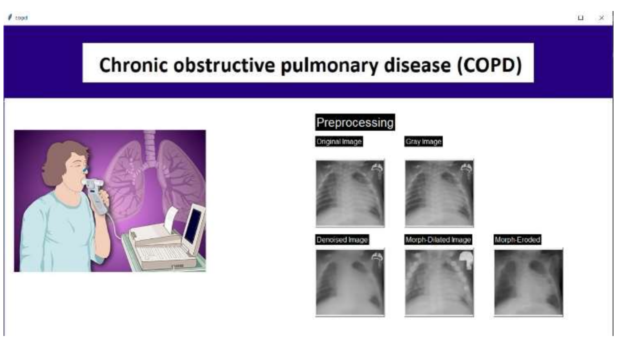


Figure 2: Preprocessing

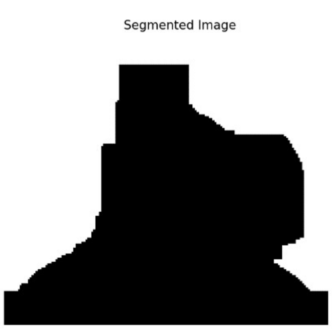


Figure 3: Segmented Image

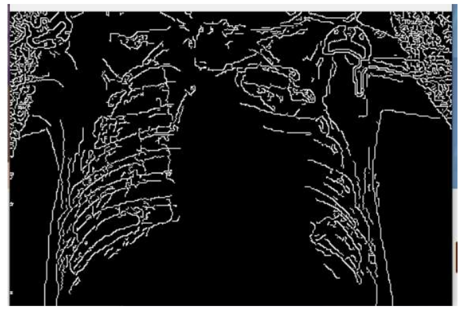


Figure 4: Feature Extraction

The model is 92.0% accurate

	precision	recall	f1-score	support
copd	0.94	0.94	0.94	78
NORMAL	0.89	0.89	0.89	47
accuracy			0.92	125
macro avg	0.91	0.91	0.91	125
weighted avg	0.92	0.92	0.92	125

Medicinal

The predicted image is : copd

OK

Figure 5: Predicted Report

Model Performance Summary:

Model	Accuracy (%)
Support Vector Machine	100.0
Random Forest	100.0
XGBoost	100.0
k-Nearest Neighbors	100.0

Figure 6: Model Accuracy Summary

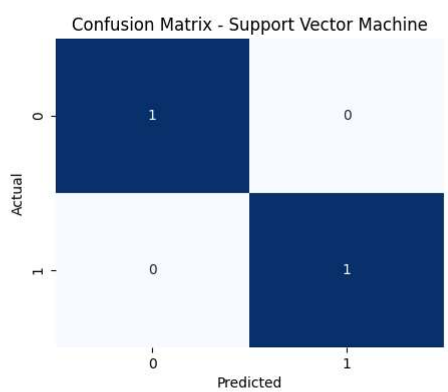


Figure 7: SVM Confusion Matrix

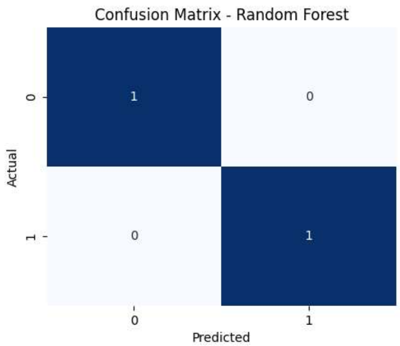


Figure 8: RF Confusion Matrix

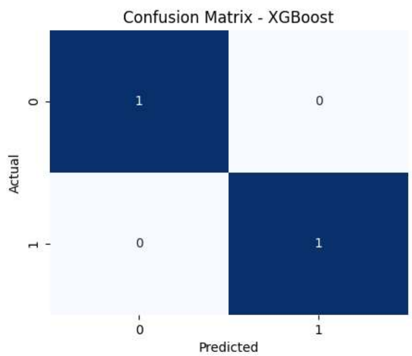


Figure 9: XG Boost Confusion Matrix

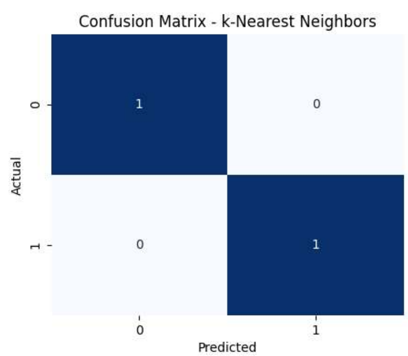
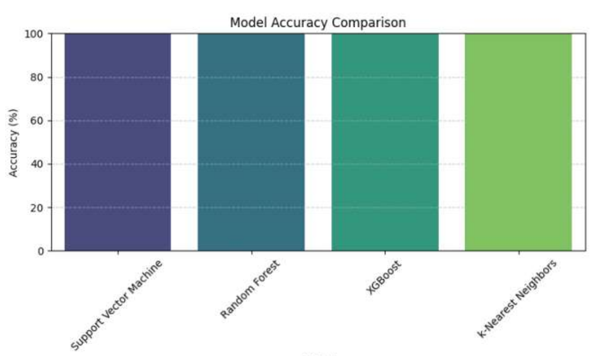


Figure 10: KNN Confusion Matrix



Graph 2: Model Accuracy Graph

VI. CONCLUSION AND FUTURE WORKS

The increasing global burden of Chronic Obstructive Pulmonary Disease (COPD) necessitates advanced diagnostic approaches that enable early detection and accurate classification. In this study, a hybrid system combining Machine Learning and Deep Learning models was proposed to enhance COPD diagnosis. Machine Learning algorithms such as Support Vector Machines, Random Forest, Gradient Boosting, and k-Nearest Neighbors were effectively applied to clinical and spirometry data, while Convolutional Neural Networks were utilized for image-based diagnosis from chest X-rays and CT scans.

The integration of both approaches demonstrated the potential to improve diagnostic accuracy, reduce dependence on manual interpretation, and assist clinicians in making informed decisions. By leveraging structured clinical data and medical imaging, the system offers a comprehensive solution for COPD detection and risk assessment. The promising results highlight the importance of AI-driven tools in healthcare and suggest that such hybrid models could play a significant role in early intervention, personalized treatment planning, and better patient outcomes in COPD management.

REFERENCES

[1] Fischer et al., “Comparison of AI-Based Fully Automatic Chest CT Emphysema Quantification to Pulmonary Function Testing,” 2020.

[2] Makimoto et al., “Comparison of Feature Selection Methods and Machine Learning Classifiers for Predicting COPD Using Texture-Based CT Lung Radiomic Features,” 2023.

[3] Wang et al., “Enabling COPD Diagnosis through Chest X-Rays: A Multi-Site and Multi-Modality Study,” 2023.

[4] Altan et al., “Deep Learning on Computerized Analysis of COPD,” 2019.

[5] Haider et al., “Respiratory Sound-Based Classification of COPD: A Risk Stratification Approach in Machine Learning Paradigm,” 2019.

[6] Amaral et al., “Machine Learning Algorithms and Forced Oscillation Measurements Applied to the Automatic Identification of COPD,” 2012.

[7] Kanwade and Bairagi, “Classification of COPD and Normal Lung Airways Using Feature Extraction of Electromyographic Signals,” 2019.

[8] Siddiqui et al., “Respiration-Based COPD Detection Using UWB Radar Incorporation with Machine Learning,” 2022.

[9] Cosentino et al., “Inference of COPD with Deep Learning on Raw Spirograms Identifies New Genetic Loci and Improves Risk Models,” 2023.

[10] Wu et al., “Machine Learning in COPD,” 2023.

*Author for Correspondence: Kausar Shameem