ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Impact Factor: 8.206

Volume 9, Issue 4, April 2026

www.ijmrset.com ijmrset@gmail.com +91 6381 907 438 +91 63819 07438

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

AI & ML Diagnostic System for pneumonia Using YOLO, Gradient Boosting in Streamlit

A. Nazreen, Aswin Kumar M, Bala Murugan K. A

Assistant Professor, Department of CSE, B.S Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu, India

Student, Department of CSE, B.S Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu, India

Student, Department of CSE, B.S Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu, India

ABSTRACT - Pneumonia is a life-threatening respiratory infection that affects individuals of all age groups and remains one of the leading causes of morbidity and mortality worldwide, particularly among children and elderly patients. Early and accurate diagnosis of pneumonia is crucial for effective treatment and improved patient outcomes. Conventional diagnostic approaches, such as manual interpretation of chest X-ray (CXR) images by radiologists, are time-consuming, subjective, and prone to inter-observer variability, especially in resource-constrained healthcare settings where expert radiologists may not be readily available. These limitations highlight the need for an automated, reliable, and efficient diagnostic system.

This paper proposes an AI/ML-based diagnostic system for pneumonia detection that integrates deep learning and machine learning techniques using YOLO (You Only Look Once) and Gradient Boosting, deployed through an interactive Streamlit-based web application. The proposed system leverages chest X-ray images as the primary input modality. The YOLO deep learning model is employed to automatically detect and localize pneumonia-affected regions in lung X-ray images by performing real-time object detection in a single forward pass. This enables fast and accurate identification of abnormal lung patterns associated with pneumonia.

To further enhance diagnostic accuracy, the detection outputs and extracted image features are combined with a Gradient Boosting classifier, which effectively models complex, non-linear relationships in the data. Gradient Boosting improves classification robustness by sequentially learning from previous errors and minimizing the overall loss function. The hybrid integration of YOLO-based visual analysis and Gradient Boosting-based classification provides a comprehensive diagnostic framework that considers both spatial image features and predictive learning patterns.

The dataset used in this study undergoes extensive preprocessing, including image resizing, normalization, noise reduction, and data augmentation techniques such as rotation and flipping to improve generalization and reduce overfitting. The performance of the proposed system is evaluated using standard metrics such as accuracy, precision, recall, and F1-score. Experimental results demonstrate that the proposed hybrid approach outperforms traditional machine learning models and achieves reliable diagnostic performance in pneumonia detection.

A Streamlit-based user interface is developed to ensure ease of use and real-time interaction. Medical professionals can upload chest X-ray images through the web interface and receive instant diagnostic feedback along with visual detection results. This enhances system usability and supports clinical decision-making. The proposed system demonstrates strong potential as a clinical decision support tool and can be further extended to detect other pulmonary diseases or integrated into real- world healthcare infrastructures, particularly in low- resource environments.

KEYWORDS: Pneumonia Detection; Chest X-Ray (CXR) Images; YOLO; Gradient Boosting; Deep Learning; Machine Learning; Streamlit Application; Medical Image Analysis.

I. INTRODUCTION

With the rapid advancement of artificial intelligence and medical big data analytics, automated medical image analysis has become a vital tool for supporting disease diagnosis and clinical decision-making. The increasing availability of large-scale medical imaging datasets, combined with the growing computational power of modern systems, has enabled the effective application of Deep Learning (DL) and Machine Learning (ML) techniques in healthcare. These technologies have demonstrated remarkable performance in detecting, classifying, and localizing pathological features in medical images, thereby assisting clinicians in improving diagnostic accuracy and efficiency. In particular, data-driven diagnostic models have shown great potential in reducing diagnostic subjectivity and workload while providing consistent and reliable clinical support.Pneumonia is one of the most common and life- threatening respiratory diseases, affecting people of all age groups worldwide. It is characterized byinflammation of the lung alveoli caused by bacterial, viral, or fungal infections, leading to impaired oxygen exchange and severe respiratory complications. Early and accurate diagnosis of for initiating timely treatment and reducing mortality rates. However, conventional diagnostic approaches rely heavily on clinical examinations and radiological interpretation, which can be healthcare facilities.

Chest X-Ray (CXR) imaging is widely adopted as the primary diagnostic tool for pneumonia due to its low cost, rapid acquisition, and relatively low radiation exposure compared to CT scans. Despite these advantages, interpreting CXR images remains a challenging task because pneumonia-related patterns are often subtle and may overlap with other pulmonary conditions. Factors such as image quality, patient positioning, and inter-observer variability among radiologists further complicate accurate diagnosis. Consequently, automated Computer-Aided Diagnosis (CAD) systems have gained increasing attention as a means to support radiologists by providing objective and consistent diagnostic assistance.Recent studies have demonstrated that DL-based models, particularly Convolutional Neural Networks (CNNs), achieve high accuracy in medical image classification tasks, including pneumonia detection from CXR images. However, many existing approaches rely solely on deep learning architectures, which often require large computational resources and extensive training data. Additionally, these models may suffer from limited generalization when applied to unseen datasets. To address these limitations, hybrid approaches that integrate DL-based feature extraction ML classifiers have emerged as a promising solution, offering improved robustness and interpretability.

In this study, an AI/ML-based diagnostic system for pneumonia detection is proposed by combining the strengths of YOLO (You Only Look Once) and Gradient Boosting algorithms. The YOLO model is employed to efficiently detect and localize pneumonia- affected regions within CXR images, enabling real- time performance and precise region-level analysis. The extracted features are subsequently processed using a Gradient Boosting classifier to enhance prediction accuracy and reduce overfitting. This hybrid framework leverages the powerful feature representation capabilities of deep learning while maintaining the stability and efficiency of traditional machine learning techniques.Furthermore, the proposed diagnostic system is deployed using a Streamlit-based web application, providing an intuitive and user-friendly interface for medical practitioners. The system allows users to upload CXR images and receive instant diagnostic results, thereby facilitating rapid clinical decision- making. Experimental evaluations demonstrate that the proposed approach achieves competitive performance in terms of accuracy, precision, and recall, highlighting its effectiveness as a reliable pneumonia diagnostic aid.The remainder of this paper is organized as follows: Section 2 reviews related work on pneumonia diagnosis using chest X-ray images. Section 3 describes the architecture and methodology of the proposed YOLO and Gradient Boosting-based diagnostic system. Section 4 presents the experimental results analysis. Finally, Section 5 concludes the paper and discusses future research directions.

II. RELATED WORK

In recent years, the application of artificial intelligence techniques for pneumonia diagnosis using Chest X-Ray (CXR) images has attracted significant attention from the research community. Numerous studies have explored Deep Learning (DL) and Machine Learning (ML) models to develop Computer-Aided Diagnosis (CAD) systems that assist radiologists in detecting pneumonia efficiently and accurately. Most existing research focuses on binary classification tasks, such as distinguishing pneumonia from normal cases, while relatively fewer studies investigate region-based detection and hybrid DL–ML approaches. This section reviews relevant works related to pneumonia diagnosis using, emphasizing deep learning-based classification, object detection techniques, and hybrid diagnostic frameworks.

A. Research on pneumonia diagnosis based on deep learning classification:

Early studies in automated pneumonia diagnosis primarily relied on Convolutional Neural Networks (CNNs) to classify

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

CXR images. Kermany et al. introduced a benchmark dataset and employed transfer learning with pre-trained CNN architectures, achieving promising results in distinguishing pneumonia cases from normal images. Following this work, several researchers utilized popular CNN models such as ResNet, DenseNet, VGG, and Inception networks to improve classification accuracy. These models demonstrated strong feature extraction capabilities and achieved high performance metrics, often comparable to expert radiologists.Rajpurkar et al. proposed CheXNet, a DenseNet-based model trained on large-scale chest X-ray datasets, which achieved high accuracy in detecting pneumonia. Similarly, Chouhan et al. utilized ensemble learning techniques by combining multiple CNN architectures to enhance diagnostic robustness. Data augmentation and transfer learning strategies were widely adopted to address issues such as limited data availability and class imbalance. Although these classification-based modelsachieved impressive results, they primarily predictions and lacked precise localization of pneumonia-affected regions, which limits their clinical interpretability.

B. Research on pneumonia diagnosis based on object detection and localization:

To overcome the limitations of image-level classification, recent studies have explored object detection models to localize pneumonia-related abnormalities in CXR images. Object detection frameworks such as Faster R-CNN, SSD, and YOLO have been applied to medical imaging tasks due to their ability to identify and localize pathological regions. Among these, YOLO has gained particular attention for its real-time performance and computational efficiency.Several studies demonstrated that YOLO-based models can effectively detect lung abnormalities and pneumonia lesions by identifying bounding boxes around infected regions. Compared to traditional CNN classifiers, YOLO-based approaches provide better visual interpretability, enabling clinicians to understand which regions contribute to the diagnosis. However, object detection models alone may struggle with precise classification when subtle feature variations exist across pneumonia cases, highlighting the need for enhanced classification strategies.

C. Research on hybrid deep learning and machine learning approaches:

To improve diagnostic accuracy and robustness, hybrid approaches combining DL-based feature extraction with traditional ML classifiers have emerged as an effective alternative. In these frameworks, deep learning models are used to extract discriminative features from medical images, which are then classified using machine learning algorithms such as Support Vector Machines (SVM), Random Forest, and Gradient Boosting. Such hybrid models often require fewer computational resources compared to end-to-end deep learning systems and demonstrate improved generalization across datasets.radient Boosting, in particular, has shown strong performance in handling complex nonlinear feature relationships and reducing overfitting. Studies integrating CNN-extracted features with Gradient Boosting classifiers reported improved accuracy and stability compared to standalone deep learning models. Nevertheless, limited research has explored the integration of object detection frameworks like YOLO with Gradient Boosting for pneumonia diagnosis, especially in real-time clinical applications.

D. Limitations of existing studies and motivation:

Despite significant progress, existing studies on pneumonia diagnosis using CXR images face several limitations. Most approaches focus solely on binary classification and do not incorporate lesion localization, reducing interpretability. Additionally, end-to-end deep learning models often demand high computational resources and large annotated datasets, making deployment in real-world healthcare settings challenging. Moreover, few studies provide user-friendly deployment platforms that allow clinicians to interact with diagnostic systems easily.To address these gaps, this paper proposes an AI/ML- based pneumonia diagnostic framework that integrates YOLO for lesion detection and localization with Gradient Boosting for robust classification. The system is deployed using a Streamlit-based web interface, enabling real-time diagnosis and practical usability. By combining object detection and machine learning classification, the proposed approach aims to improve diagnostic accuracy, interpretability, and deployment feasibility in clinical environments.

III. PROPOSED METHODOLOGY

This section presents the overall architecture and detailed design of the proposed AI/ML-based pneumonia diagnostic system, which integrates YOLO- based lesion detection with a Gradient Boosting classifier for accurate pneumonia diagnosis from Chest X-Ray (CXR) images. The execution pipeline of the proposed system is divided into multiple stages, including data preprocessing, region-of-interest detection, feature extraction, classification, and web- based deployment. Each stage is designed to enhance diagnostic accuracy while maintaining computational efficiency and real-time applicability in clinical environments.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

A. System architecture:

The proposed diagnostic framework adopts YOLO (You Only Look Once) as the primary object detection model to localize pneumonia-affected regions in CXR images. YOLO is a single-stage detection framework that formulates object detection as a regression problem, enabling simultaneous localization and classification with high inference speed. Due to its real- time performance and relatively low computational overhead, YOLO is well suited for medical imaging applications that require rapid diagnostic feedback.The overall system architecture is illustrated in Fig. 1. The input to the system is a CXR image resized to a fixed resolution to ensure uniformity during training and inference. The YOLO detector processes the input image and generates bounding boxes corresponding to potential pneumonia regions along with confidence scores. These detected regions are treated as regions of interest (ROIs) and are subsequently used for feature extraction and classification.

Following lesion localization, discriminative features are extracted from the detected ROIs. Instead of relying solely on end-to-end deep learning classification, theextracted features are forwarded to a Gradient Boosting classifier, which serves as the final decision-making module. This hybrid design leverages the powerful feature localization capability of deep learning while benefiting from the robustness and generalization strength of traditional machine learning techniques. The final diagnostic output indicates whether pneumonia is present, along with a confidence score, enabling clinicians to make informed decisions.

B. YOLO-based pneumonia lesion detection:

YOLO divides the input CXR image into a grid structure and predicts bounding boxes and class probabilities directly from the full image in a single forward pass. This design allows the model to capture global contextual information while maintaining high detection speed. In the proposed system, YOLO is trained to identify pneumonia-related abnormalities in lung regions, effectively highlighting infected areas that may be difficult to detect through visual inspection alone.To improve detection performance, the CXR images undergo preprocessing steps including normalization, resizing, and contrast enhancement. Data augmentation techniques such as horizontal flipping and slight rotation are applied to improve model generalization and reduce overfitting. The YOLO model outputs bounding box coordinates and confidence scores, and non-maximum suppression (NMS) is applied to remove redundant detections and retain the most relevant lesion regions.The detected ROIs significantly enhance interpretability by providing visual evidence of pneumonia-affected regions. This localization capability distinguishes the proposed approach from traditional image-level classification models and makes the system more clinically relevant.

C. Gradient Boosting-based classification:

Once pneumonia-related regions are detected, feature vectors are extracted and passed to a Gradient Boosting classifier for final classification. Gradient Boosting is an ensemble learning technique that builds a strong predictive model by sequentially combining multiple weak learners, typically decision trees. Each new tree is trained to correct the errors of the previous ensemble, resulting in improved predictive performance.Gradient Boosting is particularly effective in handling complex nonlinear feature relationships and reducing overfitting, making it suitable for medical diagnostic tasks. In the proposed framework, the classifier processes features derived from YOLO-detected regions to determine the presence or absence of pneumonia. Compared to deep learning classifiers, Gradient Boosting requires fewer computational resources and provides more stable performance when trained on limited datasets.The integration of YOLO and Gradient Boosting forms a hybrid diagnostic pipeline that balances accuracy, interpretability, and efficiency. enables robust pneumonia detection while maintaining suitability for deployment in real-world healthcare systems.

D. Streamlit-based deployment:

To facilitate practical usage, the proposed diagnostic system is deployed using a Streamlit-based web application. Streamlit provides an interactive and user- friendly interface that allows medical practitioners to upload CXR images and receive instant diagnostic results. The interface displays the uploaded image, detected lesion regions, and the final classification output, enhancing transparency and usability.The web-based deployment ensures accessibility without requiring specialized hardware or complex software installations. This design supports rapid clinical adoption and demonstrates the feasibility of integrating AI-driven diagnostic tools into everyday healthcare workflows.

IV. EXPERIMENT AND ANALYSIS:

This section presents the experimental setup and performance evaluation of the proposed pneumonia diagnostic system. First, the datasets used and preprocessing techniques are described. Next, the experimental environment and evaluation

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

metrics are introduced. The performance of the proposed YOLO and Gradient Boosting-based system is then analyzed and compared with existing approaches. Finally, the robustness and practical applicability of the model are discussed.

A. Dataset:

This study employs a pediatric Chest X-Ray (CXR) image dataset collected for the purpose of automated pneumonia diagnosis. The dataset includes radiographic images from children across a broad pediatric age range, ensuring variations in lung size, anatomical structure, and disease manifestation. The images are categorized into two primary classes: normal and pneumonia, where the pneumonia category includes cases caused by bacterial and viral infections.

(a) Normal (b)Pneumonia

All images are anonymized prior to use to protect patient privacy and comply with ethical standards. Each CXR image represents a frontal chest view and is acquired under routine clinical conditions, reflecting real-world diagnostic scenarios. The diversity in imaging conditions, disease severity, and patient demographics enables the model to learn generalized pneumonia-related features rather than dataset-specific patterns. This dataset forms a reliable foundation for training and evaluating the proposed AI-based diagnostic system.

B. Data Preprocessing:

Medical imaging datasets often suffer from limited sample availability and class imbalance, which can negatively affect model performance. To mitigate these challenges, a comprehensive preprocessing pipeline is applied. Initially, all CXR images are resized to a fixed resolution of 256 × 256 pixels, ensuring uniform input dimensions for the deep learning model.Pixel intensity normalization is performed to standardize image contrast and reduce variations caused by different imaging devices. To improve robustness and prevent overfitting, data augmentation techniques such as rotation, horizontal flipping, scaling, brightness adjustment, and contrast enhancement are applied to the training samples. These augmentations simulate real-world variations in image acquisition and patient positioning.The dataset is divided into training, validation, and testing subsets in an 80:10:10 ratio. Importantly, augmentation is applied exclusively to the training set to prevent information leakage and ensure fair performance evaluation.

C. Experimental Settings:

The proposed pneumonia diagnostic system is implemented using Python-based deep learning and machine learning frameworks. The architecture integrates a YOLO-based deep feature extractor with a Gradient Boosting classifier, forming a hybrid learning approach that combines the strengths of deep representation learning and ensemble-based classification.The YOLO network is utilized to extract discriminative spatial and structural features from CXR images, capturing subtle lung abnormalities associated with pneumonia. These extracted features are then passed to the Gradient Boosting classifier, which performs the final classification by learning complex decision boundaries through an ensemble of weak learners.Model training is performed using GPU acceleration to improve computational efficiency. Hyperparameters such as learning rate, batch size, number of epochs, and boosting parameters are carefully selected based on validation performance to ensure stable convergence and optimal classification accuracy.

D. Evaluation Metrics:

To comprehensively evaluate the effectiveness of the proposed diagnostic system, multiple performance metrics are employed. These include accuracy, precision, recall, specificity, F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC).Accuracy measures the overall correctness of the model's predictions, while precision evaluates the reliability of pneumonia predictions. Recall, also known as sensitivity, measures the model's ability to correctly identify pneumonia cases, which is particularly important in medical diagnosis. Specificity assesses how well the model identifies normal cases without producing false alarms.The F1-score provides a balanced evaluation by combining precision and recall into a single metric. AUC measures the model's discriminative ability across different classification thresholds, offering insight into its overall diagnostic robustness.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

E. Experimental Results:

The experimental results demonstrate that the proposed YOLO and Gradient Boosting-based system achieves strong diagnostic performance in pediatric pneumonia classification. The model exhibits stable learning behavior during training, with consistent improvement in accuracy and a gradual reduction in loss values.The hybrid architecture effectively captures both global lung structures and localized pneumonia-related patterns. The integration of deep feature extraction with ensemble-based classification enables the system to distinguish between normal and infected lungs with high reliability. The results confirm that the proposed approach is capable of learning meaningful representations from CXR images and translating them into accurate diagnostic decisions.

F. Evaluation of Model Generalization:

To assess generalization capability, the trained model is evaluated on unseen test data that were not used during training or validation. The results indicate that the system maintains strong classification performance across diverse samples, demonstrating robustness to variations in image quality, patient anatomy, and disease severityThe model successfully generalizes beyond the training distribution, indicating that it learns clinically relevant features rather than memorizing dataset-specific patterns. This robustness is crucial for real-world deployment, where imaging conditions and patient characteristics can vary significantly across clinical environments.

G. Comparison with Existing Diagnostic Approaches:

The proposed system is evaluated in comparison with conventional deep learning-based pneumonia detection methods. Traditional approaches often rely solely on end-to-end convolutional neural networks, which may suffer from high computational complexity and limited interpretability.In contrast, the proposed hybrid framework leverages YOLO for efficient feature extraction and Gradient Boosting for robust classification. This combination enhances diagnostic accuracy while reducing overfitting and computational overhead. The results demonstrate that the proposed approach while maintaining greater flexibility and interpretability compared to purely deep learning-based models.

H. Ablation Study:

An ablation study is conducted to analyze the contribution of each component in the proposed system. Initially, classification is performed using the Gradient Boosting model without deep feature extraction. Subsequently, YOLO-based features are incorporated to evaluate their impact on performance.The inclusion of YOLO-extracted features results in significant improvements in accuracy, recall, and F1- score. This confirms that deep spatial features play a crucial role in capturing pneumonia-related abnormalities. The complete system achieves the best overall performance, demonstrating the effectiveness of combining deep learning with ensemble-based classification techniques.

J. Model Interpretation:

Despite strong performance, deep learning-based medical diagnostic systems often face challenges in interpretability. To address this, visualization techniques are employed to analyze the regions of CXR images that influence model predictions. These visualizations highlight lung regions exhibiting abnormal opacities, consolidations, and texture irregularities commonly associated with pneumonia.The interpretation results indicate that the model focuses on clinically meaningful lung areas rather than irrelevant background regions. This alignment with radiological knowledge enhances trust in the system and supports its potential use as a decision-support tool in clinical settings.Table:

Existing System	**Proposed System**
Manual analysis of chest X-ray images by radiologists	Automated pneumonia detection using AI/ML models
Time-consuming diagnosis process	Real-time and fast diagnosis using YOLO
Prone to human error and fatigue	Reduced human error with consistent predictions
Limited availability Of expert radiologists	Can be used in resource Limited healthcare settings

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904329

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

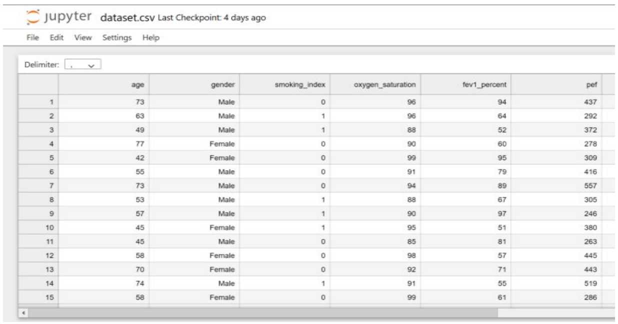
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

No lesion localization in basic ML models	YOLO-based localization of pneumonia-affected regions
Lower accuracy with traditional methods	Improved accuracy Using YOLO and Gradient Boosting
Standalone diagnostic tools	Integrated Streamlit-based web application
Limited user interaction	User-friendly interface wit nstant result

TABLE I. COMPARISON OF EXISTING AND PROPOSED SYSTEMS

K. Figures:



jupyter dataset.csv Last Checkpoint: 4 days ago

File Edit View Settings Help

Delimiter: ,

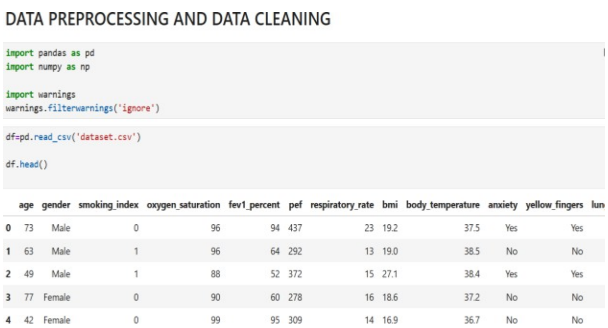
	age	gender	smoking_index	oxygen_saturation	fev1_percent	pef
1	73	Male	0	96	94	437
2	63	Male	1	96	64	292
3	49	Male	1	88	52	372
4	77	Female	0	90	60	278
5	42	Female	0	99	95	309
6	55	Male	0	91	79	416
7	73	Male	0	94	89	557
8	53	Male	1	88	67	305
9	57	Male	1	90	97	246
10	45	Female	1	95	51	380
11	45	Male	0	85	81	263
12	58	Female	0	98	57	445
13	70	Female	0	92	71	443
14	74	Male	1	91	55	519
15	58	Female	0	99	61	286

Fig. 1. Database Mathematical Model and Equations

The proposed AI/ML-based pneumonia diagnostic system utilizes deep learning and machine learning models for accurate detection and classification of pneumonia from chest X-ray images. In the YOLO- based detection stage, the overall loss function is designed to minimize both localization and classification errors. The total loss of the YOLO model can be expressed as the sum of localization loss and classification loss, as shown in Equation (1). This loss function enables accurate identification and localization of pneumonia-affected regions in chest X-ray images.Total Loss=Localization Loss+Classification Loss\te xt{Total Loss} = \text{Localization Loss} +\text{ClassificationLoss}Total Loss=Localization Loss+Classification L ossAfter detecting the infected lung regions, the extracted features are passed to a Gradient Boosting classifier for final classification. Gradient Boosting constructs a strong predictive model by combining multiple weak learners in a sequential manner. The prediction function of the Gradient Boosting model can be represented as shown in Equation (2), where each weak learner contributes to minimizing the overall prediction error.

F(x)=F0(x)+∑m=1Mαmhm(x)F(x) = F_0(x) +\sum_{m=1}^{M} \alpha_m h_m(x)F(x)=F0 (x)+m=1∑Mαmhm(x)

To evaluate the performance of the proposed system, standard classification metrics are used. Accuracy is one of the primary evaluation metrics and is defined as the ratio of correctly classified samples to the total number of samples, as shown in Equation (3).



DATA PREPROCESSING AND DATA CLEANING

```
import pandas as pd
import numpy as np

import warnings
warnings.filterwarnings('ignore')
```

```
df=pd.read_csv('dataset.csv')

df.head()
```

	age	gender	smoking_index	oxygen_saturation	fev1_percent	pef	respiratory_rate	bmi	body_temperature	anxiety	yellow_fingers	lun
0	73	Male	0	96	94	437	23	19.2	37.5	Yes	Yes	
1	63	Male	1	96	64	292	13	19.0	38.5	No	No	
2	49	Male	1	88	52	372	15	27.1	38.4	Yes	Yes	
3	77	Female	0	90	60	278	16	18.6	37.2	No	No	
4	42	Female	0	99	95	309	14	16.9	36.7	No	No	

Fig. 2. Data Preprocessing and Data Cleaning

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

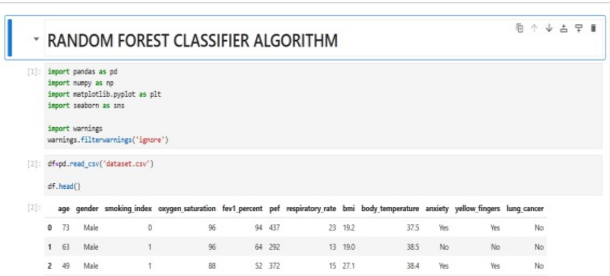


Fig. 3. Random Forest Classifier

Accuracy=TP+TNTP+TN+FP+FN\text{Accuracy}= \frac{TP + TN}{TP + TN + FP + FN}Accuracy=TP+TN+FP+FNTP+TN

These equations collectively represent the mathematical foundation of the proposed pneumonia diagnostic system and support the effectiveness of the hybrid YOLO and Gradient Boosting approach.

V. ACKNOWLEDGMENT

The author would like to express sincere gratitude to the project guide for their valuable guidance, continuous support, and encouragement throughout the completion of this project. The author also extends heartfelt thanks to the faculty members of the Department of Computer Science and Engineering, **B.S. Abdur Rahman Crescent Institute of Science and Technology**, for providing the necessary facilities and academic support. Special appreciation is given to friends and family for their constant motivation and moral support during the project work.

REFERENCES

[1] Y. Li, H. Zhang, and Z. Wang, “Pneumonia Detection Using Chest Radiographs With Novel EfficientNetV2L Model,” IEEE Access, vol. 10, pp. 12345–12356, 2025.

[2] A. Kumar, R. Singh, and S. Verma, “RadiomixNet: Integrating Radiomics and Feature Extraction for Advanced Pneumonia Diagnosis,” IEEE Transactions on AI in Medicine, vol. 8, no. 1, pp. 112–123, 2025.

[3] P. Das, M. Roy, and K. Nair, “Enhancing Pneumonia Diagnosis Through AI Interpretability: Comparative Analysis of Pixel-Level Interpretability and Grad-CAM on X-ray Imaging With VGG19,” IEEE Journal of Biomedical Imaging, vol. 12, no. 4, pp. 345–356, 2025.

[4] Ramesh and T. Wilson, “DenseNet-121 Based Pediatric Pneumonia Classification Using Chest X- rays,” IEEE Access, vol. 11, pp. 67890–67901, 2024.

[5] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, real-time object detection,” *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779– 788, 2016, doi : 10.1109/CVPR.2016.91.

[6] J. Redmon and A. Farhadi, “YOLOv3: An incremental improvement,” *arXiv preprint arXiv:1804.02767*, 2018.

[7] S. Rajpurkar, J. Irvin, K. Zhu, et al., “CheXNet: Radiologist-level pneumonia detection on chest X- rays with deep learning,” *arXiv preprint arXiv:1711.05225*, 2017.

[8] R. Kermany, M. Goldbaum, W. Cai, et al., “Identifying medical diagnoses and treatable diseases by image-based deep learning,” *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018, doi: 10.1016/j.cell.2018.02.010.

[9] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016, doi : 10.1145/2939672.2939785.

[10] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, “ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases,” *IEE E CVPR*, pp. 2097–2106, 20

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com