iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: VII Month of publication: July 2026

DOI: https://doi.org/10.22214/ijraset.2026.83700

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

Resi-View 360: Multimodal AI Triage System

Pallavi Marbe[1], Laxmi Sanganabasappa Soddi[2], Prajwal[3], Rohit[4], Prof. Anupama Pattanasetty[5]
Sharnbasva University, India

I. INTRODUCTION

Healthcare has undergone significant transformation due to advancements in Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL). These technologies are increasingly being utilized to assist medical professionals in disease diagnosis, patient monitoring, and clinical decision-making. By analyzing large volumes of medical data efficiently and accurately, AI-powered systems have the potential to improve healthcare accessibility, reduce diagnostic delays, and support better treatment planning.

Respiratory diseases represent one of the most common health challenges worldwide. Conditions such as pneumonia, bronchitis, asthma, chronic obstructive pulmonary disease (COPD), and other pulmonary disorders affect millions of individuals annually. Early detection of these diseases is critical because delayed diagnosis can lead to severe complications, prolonged hospitalization, and increased mortality rates.

Traditionally, respiratory disease diagnosis relies on multiple sources of information, including Chest X-Ray (CXR) imaging, respiratory sound examination, laboratory investigations, and patient clinical history. The interpretation of such information often requires experienced healthcare professionals and specialized diagnostic equipment. However, access to these resources may be limited in rural and underserved healthcare environments.

Recent developments in multimodal Artificial Intelligence have enabled the integration of different types of healthcare data into a single intelligent system. Multimodal systems can simultaneously analyze medical images, audio signals, and textual information, allowing a more comprehensive understanding of a patient's condition. Such systems have demonstrated promising results in improving diagnostic support and enhancing clinical decision-making processes.

The proposed project, **Respi-View 360: Multimodal AI-Assisted Respiratory Disease Assessment System**, leverages advanced foundation models including Google's Health Acoustic Representations (HeAR) and MedGemma to analyze respiratory audio recordings, Chest X-Ray images, and patient medical history. The system combines these diverse data sources to generate preliminary respiratory disease assessments, risk alerts, and explainable clinical reports.

The primary aim of this project is to demonstrate how modern multimodal AI technologies can be applied to healthcare decision-support systems. By integrating image analysis, acoustic analysis, and natural language reasoning into a unified platform, Respi-View 360 seeks to provide an intelligent, user-friendly, and scalable solution for respiratory disease assessment while maintaining transparency through explainable AI reporting.

A. Background Of The Project

Respiratory diseases are among the leading causes of illness and mortality worldwide. Conditions such as pneumonia, bronchitis, asthma, chronic obstructive pulmonary disease (COPD), and other pulmonary disorders affect millions of individuals every year. Early diagnosis and timely treatment are essential for reducing complications and improving patient outcomes. However, traditional diagnosis often requires specialized medical equipment, trained radiologists, and significant clinical expertise, which may not always be available in rural and resource-limited healthcare environments.

Recent advancements in Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) have transformed the healthcare industry by enabling automated disease detection and decision-support systems. AI-based medical systems are capable of analyzing large volumes of medical data, identifying hidden patterns, and assisting healthcare professionals in making informed decisions.

Medical imaging techniques such as Chest X-Rays (CXR) are commonly used to diagnose respiratory diseases. Similarly, respiratory sounds such as coughs, wheezing, crackles, and breathing patterns contain valuable information regarding a patient's pulmonary health. The combination of imaging data, respiratory audio signals, and patient clinical history provides a more comprehensive understanding of respiratory conditions.

The emergence of multimodal AI systems has made it possible to integrate multiple data sources and improve diagnostic accuracy. Instead of relying solely on a single input source, multimodal systems analyze images, audio, and textual information simultaneously to generate more informed assessments.

The proposed project, Respi-View 360, is a Multimodal AI-Assisted Respiratory Disease Assessment System that combines Chest X-Ray analysis, respiratory audio analysis, and patient history evaluation using advanced AI foundation models such as MedGemma and Google's Health Acoustic Representations (HeAR). The system aims to provide an intelligent respiratory disease assessment framework capable of assisting healthcare professionals during preliminary diagnosis and patient screening.

The feasibility of building effective multimodal triage systems has been dramatically accelerated by the development of specialized foundation models. Unlike early ML algorithms that required training from scratch on narrow datasets, modern Large Language Models (LLMs) and their multimodal variants possess a deep, pre-trained understanding of complex patterns. Google’s introduction of the MedGemma framework represents a pivotal milestone in this domain. MedGemma is an open-weights, clinically tuned model designed explicitly to handle the nuances of medical reasoning safely and accurately. It provides developers with a robust, accessible reasoning engine that can be securely integrated into clinical applications without the immense overhead of training proprietary medical models.

Contextualizing Respi-View 360. It is within this evolving technological and clinical landscape that the Respi-View 360 project is situated. Developed as a direct response to the global need for smarter, faster clinical workflows—and catalyzed by the MedGemma—this project seeks to operationalize the theory of multimodal medical AI. Respi-View 360 is designed to serve as an intelligent, preliminary triage assistant that processes both respiratory acoustics and visual clinical data through the MedGemma reasoning engine. By automating the preliminary multi-sensory synthesis, the project aims to demonstrate how modern AI can actively augment healthcare professionals, reducing their cognitive load while ensuring rapid, accurate, and scalable respiratory triage.

B. Problem Statement

Respiratory diseases are among the leading causes of morbidity and mortality across the globe. Diseases such as pneumonia, bronchitis, asthma, tuberculosis, Chronic Obstructive Pulmonary Disease (COPD), and other pulmonary disorders affect millions of people every year. Accurate and timely diagnosis of these diseases is essential to reduce complications, improve patient outcomes, and prevent the progression of severe respiratory conditions.

Traditional respiratory disease diagnosis involves multiple clinical procedures including physical examination, patient history evaluation, respiratory sound assessment, radiological investigations such as Chest X-Rays (CXR), and laboratory tests. These procedures often require the expertise of pulmonologists, radiologists, and other healthcare professionals. In many healthcare facilities, especially in rural and underdeveloped regions, access to such specialized medical resources may be limited. As a result, patients often experience delays in diagnosis and treatment.

Chest X-Ray imaging is one of the most widely used diagnostic tools for detecting respiratory abnormalities. However, interpreting X-Ray images requires experienced radiologists and is susceptible to human error due to fatigue, workload, and subjective judgment. Similarly, respiratory sounds such as coughs, wheezing, crackles, and breathing patterns contain valuable clinical information but require trained healthcare professionals for proper interpretation. Manual analysis of respiratory audio signals can be t ime-consuming and inconsistent.

In recent years, Artificial Intelligence has demonstrated significant potential in assisting medical diagnosis by automatically analyzing healthcare data. However, most existing AI-based diagnostic systems focus on a single modality such as medical images or textual patient records. These systems often fail to utilize the complete spectrum of clinical information available for a patient. As a result, diagnostic accuracy and reliability may be limited.

Another major challenge is the lack of integrated systems capable of simultaneously analyzing Chest X-Ray images, respiratory audio recordings, and patient clinical history. Since respiratory diseases often manifest through multiple symptoms and diagnostic indicators, relying on a single source of information may not provide a comprehensive assessment of a patient's condition. There is a growing need for intelligent healthcare systems capable of integrating multiple data modalities and providing a holistic understanding of respiratory health. Furthermore, healthcare professionals often require explainable and interpretable AI outputs rather than black-box predictions. Many existing AI models provide results without clearly explaining the factors influencing the decision, which reduces trust and adoption in clinical environments. Therefore, explainability and transparency are essential requirements for any AI-assisted healthcare solution. The proposed project addresses these challenges by developing **Respi-View 360**, a Multimodal AI-Assisted Respiratory Disease Assessment System that integrates Chest X-Ray analysis, respiratory audio analysis, and patient history evaluation using advanced foundation models such as MedGemma and Google's Health Acoustic Representations (HeAR). The system aims to generate explainable respiratory disease assessments, provide preliminary risk evaluation, and assist healthcare professionals during clinical decision-making.

C. Motivation

The rapid advancement of Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) technologies has created new opportunities for transforming healthcare systems across the world. These technologies have demonstrated remarkable capabilities in medical image analysis, disease prediction, clinical decision support, and healthcare automation. The ability of AI systems to process large amounts of healthcare data efficiently has encouraged researchers and healthcare organizations to explore intelligent solutions for improving diagnostic accuracy and accessibility.

Respiratory diseases continue to be one of the major public health concerns worldwide. According to global health reports, millions of people suffer from respiratory illnesses such as pneumonia, asthma, bronchitis, tuberculosis, and Chronic Obstructive Pulmonary Disease (COPD). These diseases significantly affect the quality of life of patients and often require early diagnosis for effective treatment. Unfortunately, many healthcare facilities, particularly in rural and underdeveloped regions, lack access to specialized medical professionals and advanced diagnostic infrastructure. This situation often leads to delayed diagnosis, increased healthcare costs, and poor patient outcomes.

One of the primary motivations behind the development of Respi-View 360 is the need to bridge the gap between healthcare demand and the availability of expert medical resources. Traditional respiratory disease diagnosis relies heavily on clinical expertise for interpreting Chest X-Ray images, evaluating respiratory sounds, and understanding patient symptoms. These procedures require experienced radiologists and pulmonologists who may not always be available. An AI-assisted system can support healthcare professionals by providing preliminary assessments and reducing diagnostic workload.

Another important motivation is the growing availability of multimodal healthcare data. Patients generate information in various forms, including medical images, respiratory audio recordings, laboratory reports, and textual clinical histories. Most existing healthcare systems analyze only a single data modality, which may limit diagnostic effectiveness. Human clinicians naturally combine multiple sources of information before making a diagnosis. Therefore, developing an AI system capable of integrating multiple modalities can more closely resemble real-world clinical reasoning and potentially improve assessment quality.

The emergence of foundation models has further motivated the development of this project. Recent AI models such as MedGemma and Google's Health Acoustic Representations (HeAR) have demonstrated impressive capabilities in understanding medical images and health-related acoustic signals. These models provide an opportunity to develop intelligent healthcare applications without requiring the extensive computational resources needed to train large models from scratch. By leveraging these pre-trained foundation models, the project can utilize state-of-the-art AI capabilities while maintaining practical deployment feasibility.

The COVID-19 pandemic also highlighted the importance of rapid and accessible respiratory disease screening systems. During the pandemic, healthcare facilities worldwide experienced overwhelming patient loads, creating an urgent need for automated diagnostic support tools. The experience demonstrated that AI-powered healthcare systems can play a significant role in assisting healthcare professionals during large-scale public health emergencies. This project is motivated by the vision of creating systems that can support healthcare delivery during both routine clinical practice and emergency situations.

A further motivation arises from the increasing demand for explainable Artificial Intelligence (XAI) in healthcare. Medical professionals are often hesitant to trust AI systems that provide predictions without justification. Healthcare decisions require transparency, accountability, and interpretability. Therefore, Respi-View 360 is designed not only to generate respiratory disease assessments but also to provide explainable reports that describe the evidence used in reaching those conclusions. This approach improves user trust and supports clinical decision-making.

D. Objectives Of The Project

The primary objective of the proposed project, **Respi-View 360: Multimodal AI Respiratory Disease Assessment System**, is to develop an intelligent healthcare assistance platform capable of performing preliminary respiratory disease assessment through the integration of Chest X-Ray imaging, respiratory audio analysis, and patient clinical history. The project aims to leverage advanced Artificial Intelligence and multimodal foundation models to improve the efficiency, accessibility, and explainability of respiratory disease screening.

The system is designed to combine visual findings from radiological images, acoustic biomarkers extracted from respiratory sounds, and symptom-based clinical information to generate comprehensive respiratory assessment reports. By integrating multiple sources of patient data, the proposed solution seeks to provide more informative and reliable decision-support outputs compared to traditional single-modal diagnostic systems. Another important objective of the project is to develop a multimodal reasoning framework that enables the system to analyze different healthcare data types simultaneously. The framework combines image analysis, audio interpretation, and textual symptom evaluation to generate meaningful clinical insights and preliminary disease assessments.

Main Objectives

1) To develop a multimodal AI-based respiratory disease assessment system that combines Chest X-Ray analysis, respiratory audio processing, and patient symptom evaluation.
2) To implement advanced deep learning and foundation models for extracting meaningful insights from medical images and respiratory sound recordings.
3) To design an explainable diagnostic framework that provides visual findings, symptom analysis, severity assessment, and clinical recommendations.
4) To develop a user-friendly web-based application that enables healthcare professionals and users to upload patient data and obtain assessment reports efficiently.
5) To implement an emergency risk alert mechanism for identifying potentially critical respiratory conditions and recommending immediate medical attention when required.

E. Scope Of The Project

The scope of the proposed project, **Respi-View 360: Multimodal AI Respiratory Disease Assessment System**, focuses on developing an intelligent healthcare assistance platform capable of performing preliminary respiratory disease assessment by integrating multiple sources of patient information. The system utilizes Artificial Intelligence, Deep Learning, Computer Vision, Health Acoustic Analysis, and Natural Language Processing techniques to analyze Chest X-Ray images, respiratory audio recordings, and patient clinical history simultaneously.

The project aims to provide an AI-assisted decision support system that helps healthcare professionals and medical practitioners perform initial respiratory disease screening efficiently. The system combines image-based findings, acoustic biomarkers, and symptom information to generate comprehensive respiratory assessment reports.

The primary scope of the project includes the analysis of Chest X-Ray images using advanced multimodal foundation models. The system validates uploaded medical images and identifies abnormalities such as lung opacities, consolidations, infiltrates, and other visual indicators associated with respiratory diseases. This functionality assists in detecting possible respiratory conditions while reducing the manual effort required for image interpretation. Another major component within the scope of the project is respiratory audio analysis. The system processes breathing and cough recordings using Google's Health Acoustic Representations (HeAR) model to extract meaningful acoustic biomarkers. These biomarkers provide additional information that complements radiological findings and patient symptoms. The inclusion of respiratory audio analysis enables the system to leverage non-invasive diagnostic signals for respiratory assessment. The project also incorporates patient clinical history as an important source of contextual information. Symptoms such as fever, cough, chest pain, wheezing, shortness of breath, and breathing difficulties are analyzed and integrated into the final assessment process. This multimodal approach enhances the overall reliability of the generated reports compared to systems that rely solely on image analysis. A key aspect of the project scope is the implementation of explainable Artificial Intelligence techniques. Instead of providing only a disease prediction, the system generates detailed assessment reports explaining how the conclusion was derived. The report includes visual observations from Chest X-Ray analysis, respiratory audio findings, patient symptom evaluation, severity assessment, confidence indicators, and clinical recommendations. This transparency helps users understand the reasoning behind AI-generated outputs.

The project also includes the development of a user-friendly web-based interface using Gradio. The interface allows users to upload Chest X-Ray images, respiratory audio recordings, and patient information through an intuitive dashboard. The generated assessment reports can be viewed directly on the dashboard and downloaded as professionally formatted PDF documents for future reference.

An emergency risk assessment module is also included within the project scope. The system identifies potentially critical symptoms such as severe chest pain, breathing difficulties, oxygen-related complications, and loss of consciousness. When such symptoms are detected, high-risk alerts are generated to encourage immediate medical consultation and intervention.

The scope of the project is limited to preliminary respiratory disease assessment and clinical decision support. The system is not intended to replace professional medical diagnosis, laboratory investigations, radiological evaluation, or healthcare consultation. All generated results should be considered supportive information that assists healthcare professionals during the screening process.

Future enhancements may include integration with Electronic Health Record (EHR) systems, cloud-based deployment, mobile application support, IoT-enabled medical devices, real-time telemedicine platforms, and the incorporation of specialized medical foundation models trained specifically for respiratory disease diagnosis.

Overall, the project contributes toward improving accessibility, efficiency, and transparency in respiratory healthcare assessment by providing an intelligent multimodal AI-powered respiratory disease screening platform.

F. Project Planning And Timeline

Project planning is a critical phase in the successful development of the proposed Respi-View 360: Multimodal AI Respiratory Disease Assessment System. A well-structured project plan helps in organizing tasks, allocating resources effectively, monitoring development progress, and ensuring timely completion of project objectives. The project follows a systematic development methodology consisting of requirement analysis, literature review, model selection, data preprocessing, multimodal integration, user interface development, testing, and deployment phases.

The project begins with an extensive study of respiratory diseases, chest radiography analysis, health acoustic representations, multimodal artificial intelligence systems, and existing healthcare diagnostic solutions. Various research papers, journals, and industry reports are reviewed to identify limitations in current respiratory disease screening methods and explore opportunities for AI-based improvements.

Following the research phase, suitable foundation models and technologies are selected for implementation. Google HeAR is chosen for respiratory audio representation learning, while MedGemma/Gemma-based multimodal models are utilized for Chest X-Ray interpretation and clinical reasoning. The system architecture is designed to integrate image analysis, acoustic biomarker extraction, symptom evaluation, and report generation within a single platform.

During the preprocessing stage, respiratory audio recordings are standardized and converted into Mel-Spectrogram representations suitable for acoustic feature extraction. Similarly, Chest X-Ray images are validated, normalized, and prepared for visual analysis. Patient symptom data is structured to enable seamless integration into the multimodal reasoning pipeline.

The next phase focuses on implementing the two-pass reasoning framework. In the first pass, the Chest X-Ray image is analyzed to identify visible abnormalities and generate visual findings. In the second pass, visual findings, respiratory audio information, and patient symptoms are combined to produce a comprehensive respiratory disease assessment report. Additional modules such as confidence estimation, severity analysis, explainable AI reporting, and emergency risk detection are incorporated into the workflow. Multiple respiratory disease scenarios are tested to ensure accurate report generation, proper alert functionality, and stable multimodal integration. The system is further optimized to improve response time and resource utilization.

Finally, comprehensive documentation is prepared detailing project methodology, implementation, experimental observations, limitations, and future enhancement possibilities. The project concludes with deployment, demonstration, and presentation activities.

1) Project Timeline

Activity	Duration
Problem Identification and Requirement Analysis	Week 1
Literature Survey and Research Gap Analysis	Week 2
Technology Selection and Architecture Design	Week 3
Dataset Collection and Study	Week 4
Audio Processing Pipeline Development	Week 5
Chest X-Ray Analysis Module Development	Week 6
Multimodal Integration and Reasoning Module	Week 7
Explainable AI and Risk Assessment Module	Week 8
Gradio Dashboard Development	Week 9
PDF Report Generation and History Management	Week 10
System Testing and Performance Evaluation	Week 11
Performance Evaluation and Optimization	Week 12

Table 1.1: Project Timeline

2) Milestones

Milestone 1: Project Initiation
- Problem statement finalized.
- Project objectives identified.
- Requirement specifications documented.

Milestone 2: Literature Review and Technology Selection
- Research papers analyzed.
- Existing solutions studied.
- Suitable AI models selected.

Milestone 3: Data Processing Development
- Audio preprocessing pipeline implemented.
- Image preprocessing workflow completed.
- Data validation mechanisms integrated.

Milestone 4: AI Model Integration
- HeAR model integrated for respiratory audio analysis.
- MedGemma model integrated for Chest X-Ray interpretation.
- Multimodal reasoning framework developed.

Milestone 5: Assessment and Reporting System
- Severity assessment module implemented.
- Confidence estimation module developed.
- Explainable AI reporting integrated.

Milestone 6: Dashboard Development
- User interface designed.
- Report visualization implemented.
- PDF generation functionality completed.

Milestone 7: Testing and Deployment
- Functional testing completed.
- User acceptance testing conducted.
- Final deployment and project demonstration performed.

The planned timeline ensures systematic execution of all development activities while providing sufficient time for testing, optimization, documentation, and successful project completion.

G. Organization Of Report

The project report is organized into multiple chapters, each focusing on a specific aspect of the proposed Respi-View 360: Multimodal AI Respiratory Disease Assessment System.

The structure of the report is designed to provide a clear and systematic understanding of the problem domain, methodology, implementation, and outcomes of the project.

Chapter 1: Introduction

This chapter introduces the project and provides an overview of respiratory disease assessment challenges, project background, problem statement, motivation, objectives, project scope, planning timeline, and report organization. The chapter establishes the need for intelligent multimodal healthcare systems capable of assisting in respiratory disease screening and assessment.

Chapter 2: Literature Survey

This chapter reviews existing research works, methodologies, technologies, and healthcare systems related to respiratory disease diagnosis, chest radiography analysis, acoustic biomarker extraction, machine learning applications in healthcare, explainable AI, and multimodal artificial intelligence systems. The strengths and limitations of previous approaches are analyzed to identify research gaps addressed by the proposed system.

Chapter 3: Requirement Analysis
This chapter discusses the requirements of the proposed system including functional requirements, non-functional requirements, feasibility analysis, software requirements, hardware requirements, system architecture, use case diagrams, and workflow analysis. The chapter explains the design considerations necessary for successful implementation.

Chapter 4: System Design and Methodology
This chapter describes the methodology adopted for developing the project. It explains multimodal data processing techniques, audio feature extraction using HeAR, Chest X-Ray analysis using MedGemma, image validation mechanisms, explainable AI generation, confidence estimation, severity assessment, and emergency risk detection strategies. The complete workflow of the proposed system is illustrated and explained.

Chapter 5: Implementation
This chapter presents the implementation details of the project. It discusses the development environment, programming languages, libraries, frameworks, AI models, data processing techniques, dashboard development, PDF generation mechanisms, and report history management modules. The integration of various system components is also described in detail.

Chapter 6: Testing and Validation
This chapter presents the testing methodology and validation procedures used to evaluate the system. Functional testing, integration testing, usability testing, and performance testing results are discussed. Different respiratory disease scenarios are evaluated to verify system reliability and consistency.

Chapter 7: Results and Discussion
This chapter presents the final outputs generated by the proposed system and discusses the effectiveness of multimodal respiratory disease assessment. The chapter analyzes generated reports, risk alerts, confidence indicators, explainability features, and overall system performance. Advantages and limitations of the developed system are also discussed.

Chapter 8: Conclusion and Future Scope
This chapter summarizes the overall achievements and outcomes of the project. It highlights how the proposed system contributes toward intelligent healthcare assistance and respiratory disease screening. The chapter also discusses limitations of the current implementation and suggests future enhancements such as specialized medical foundation models, cloud deployment, mobile healthcare applications, IoT integration, Electronic Health Record (EHR) integration, and advanced multimodal fusion techniques.
Overall, the report provides a comprehensive description of the design, development, implementation, evaluation, and future prospects of the Respi-View 360 system while demonstrating the application of modern artificial intelligence technologies in healthcare.

II. LITERATURE SURVEY

A. Introduction

A literature survey is an important phase of any research and development project as it provides a comprehensive understanding of existing technologies, methodologies, and research contributions related to the proposed work. It helps identify the strengths and limitations of previously developed systems, reveals research gaps, and provides a strong foundation for designing improved solutions. In the context of healthcare and artificial intelligence, literature review plays a crucial role in understanding how advanced machine learning techniques, computer vision models, acoustic signal processing, and multimodal learning approaches can be integrated to support disease diagnosis and clinical decision-making.
Respiratory diseases remain one of the leading causes of illness and mortality worldwide. Conditions such as pneumonia, bronchitis, asthma, tuberculosis, chronic obstructive pulmonary disease (COPD), and other lung-related disorders affect millions of individuals every year. Early identification and diagnosis of these diseases are essential for timely treatment and improved patient outcomes. However, traditional diagnostic procedures often require specialized medical equipment, trained radiologists, laboratory investigations, and significant processing time, which may not always be available in rural and resource-constrained healthcare environments.

Recent advancements in Artificial Intelligence (AI), Deep Learning, Computer Vision, Natural Language Processing (NLP), and Foundation Models have significantly transformed healthcare systems. Researchers have explored various deep learning models for analyzing medical images such as Chest X-Rays, CT scans, and MRI images to assist healthcare professionals in disease diagnosis. Similarly, health-acoustic analysis techniques have emerged as promising approaches for extracting clinically relevant information from respiratory sounds, cough recordings, and breathing patterns. Several studies have focused on Chest X-Ray classification using Convolutional Neural Networks (CNNs), Vision Transformers (ViTs), and multimodal large language models. Other research works have investigated respiratory sound analysis using machine learning and acoustic representation learning techniques. However, most existing systems focus on only a single modality such as image analysis or audio analysis independently, with limited integration between multiple clinical data sources. Furthermore, many existing healthcare AI systems provide disease classification without offering explainable reasoning, confidence estimation, severity assessment, or integrated clinical reporting. The absence of intelligent multimodal frameworks capable of simultaneously analyzing Chest X-Ray images, respiratory audio recordings, and patient symptom history represents a significant research challenge. There is also limited work on combining foundation models, explainable AI techniques, and healthcare-specific acoustic representations within a unified respiratory disease assessment platform.

The findings obtained from this literature survey provide valuable insights for selecting suitable AI models, designing multimodal diagnostic pipelines, implementing explainable assessment mechanisms, and developing user-friendly healthcare applications. The knowledge gathered from previous studies serves as the foundation for developing the proposed Respi-View 360 system, which integrates Chest X-Ray analysis, respiratory audio processing, symptom evaluation, explainable AI reporting, and emergency risk assessment within a single intelligent healthcare platform.

B. Review Of Existing Systems

The diagnosis and assessment of respiratory diseases have traditionally relied on clinical examination, radiological investigations, laboratory testing, and expert medical interpretation. With the rapid advancement of Artificial Intelligence, several intelligent healthcare systems have been developed to support respiratory disease detection and diagnosis. These systems utilize medical imaging, respiratory sound analysis, machine learning algorithms, and deep learning techniques to assist healthcare professionals in identifying respiratory abnormalities.

1) Traditional Radiologist-Based Diagnosis Systems

Traditional respiratory disease diagnosis primarily depends on radiologists and physicians who manually analyze Chest X-Ray images and evaluate patient symptoms. Chest radiography is one of the most commonly used diagnostic tools for detecting pneumonia, tuberculosis, bronchitis, lung infections, and other pulmonary disorders. Radiologists interpret visual abnormalities such as opacities, consolidations, infiltrates, and lung lesions before making clinical decisions.

Although these methods provide reliable diagnostic outcomes, they require highly trained medical professionals and considerable time for interpretation. In many healthcare facilities, especially in rural and underserved regions, there is a shortage of experienced radiologists, resulting in delayed diagnosis and treatment. Manual interpretation may also be affected by fatigue, workload, and inter-observer variability, which can influence diagnostic consistency.

2) CNN-Based Chest X-Ray Classification Systems

Recent developments in deep learning have enabled automated analysis of Chest X-Ray images using Convolutional Neural Networks (CNNs). Several researchers have developed AI models capable of detecting respiratory diseases directly from radiographic images. Popular architectures such as ResNet, DenseNet, VGGNet, EfficientNet, and InceptionNet have demonstrated promising performance in medical image classification tasks.

These systems automatically extract visual features from Chest X-Ray images and classify diseases such as pneumonia, tuberculosis, COVID-19, and lung infections. Deep learning-based image analysis significantly reduces diagnostic workload and improves processing speed. However, most CNN-based systems operate as black-box models and provide limited explanation regarding their predictions. Additionally, these systems rely solely on image data and often ignore important clinical information such as symptoms and respiratory sounds.

3) Respiratory Sound Analysis Systems

Researchers have also explored respiratory audio analysis as a non-invasive approach for respiratory disease detection. Respiratory sounds such as coughing, wheezing, crackles, and breathing patterns contain valuable clinical information that can assist in identifying pulmonary abnormalities.

Machine learning techniques including Support Vector Machines (SVM), Random Forests, Artificial Neural Networks (ANN), and Deep Learning models have been applied to respiratory sound classification. More recently, Google Health introduced Health Acoustic Representations (HeAR), a foundation model trained on large-scale health acoustic datasets. HeAR provides generalized audio embeddings that can be utilized for various healthcare-related audio analysis tasks.

Although respiratory sound analysis offers promising diagnostic potential, many existing systems focus only on audio signals and do not combine them with radiological findings or patient clinical history. As a result, their diagnostic capability remains limited when compared to comprehensive clinical evaluation.

4) AI-Based Clinical Decision Support Systems

Clinical Decision Support Systems (CDSS) have been developed to assist healthcare professionals in diagnosis, treatment planning, and patient monitoring. These systems analyze patient symptoms, laboratory reports, and clinical data to generate recommendations.

Many AI-powered CDSS utilize rule-based reasoning and machine learning algorithms to support healthcare decision-making. While such systems improve efficiency and reduce manual effort, they often depend on structured clinical records and lack integration with medical imaging and acoustic analysis. Consequently, they provide only partial support for respiratory disease assessment.

5) Large Language Models in Healthcare

Large Language Models (LLMs) such as GPT, Med-PaLM, Gemma, and MedGemma have recently gained attention in healthcare applications. These models can process clinical information, generate medical summaries, answer healthcare-related questions, and assist in diagnostic reasoning.

MedGemma combines language understanding with visual analysis capabilities, enabling multimodal healthcare applications. These models demonstrate strong reasoning capabilities and can generate explainable outputs. However, many current implementations focus primarily on text-based healthcare tasks and have limited integration with respiratory audio analysis.

6) Multimodal Medical AI Systems

Multimodal AI systems represent a significant advancement in healthcare technology. These systems combine multiple sources of information such as medical images, patient history, laboratory reports, and clinical observations to improve diagnostic performance.

Several recent studies have shown that multimodal learning improves disease detection accuracy compared to single-modality systems. However, existing multimodal healthcare systems are often complex, computationally expensive, and focused on specific clinical applications. Very few systems integrate Chest X-Ray analysis, respiratory sound analysis, patient symptoms, explainable AI reporting, emergency risk assessment, and downloadable medical reports within a unified platform.

7) Disadvantages of Existing Systems

- Depend heavily on expert radiologists and clinicians.
- Diagnostic delays due to manual interpretation.
- Most systems analyze only one data modality.
- Limited integration between image, audio, and symptom data.
- Lack explainable reasoning mechanisms.
- High computational requirements for advanced models.
- Limited accessibility in remote healthcare settings.
- Poor support for real-time respiratory disease assessment.
- Insufficient emergency risk assessment capabilities.
- Absence of integrated report generation and historical record management.

C. Comparative Analysis Of Existing Systems

A comparative analysis is essential for understanding the strengths and limitations of various respiratory disease assessment approaches. Different researchers have proposed systems based on radiological analysis, respiratory audio processing, machine learning, deep learning, and clinical decision support mechanisms. While these systems contribute significantly to healthcare technology, many of them lack comprehensive multimodal integration and explainable diagnostic capabilities.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

Parallel to acoustic research, advancements in computer vision (CV) have been leveraged to observe patients passively. These systems utilize standard camera feeds combined with object detection and pose estimation algorithms (such as YOLO or MediaPipe) to track physiological markers without physical contact.

Visual triage systems are trained to continuously monitor a patient's chest excursions to calculate an accurate respiratory rate. More advanced CV models attempt to identify visual signs of respiratory fatigue, such as the use of accessory neck and abdominal muscles, specific postural distress (like the "tripod" position), or changes in skin pallor and cyanosis (bluish discoloration indicating oxygen deprivation).

The following table presents a comparative analysis of major existing approaches and highlights the research gap addressed by the proposed Respi-View 360

Existing System	Methodology Used	Advantages	Limitations
Traditional Radiologist-Based Diagnosis	Manual Chest X-Ray interpretation	High clinical reliability	Time-consuming and dependent on experts
CNN-Based X-Ray Classification Systems	Deep learning image classification	Fast disease detection	Limited explainability
Respiratory Sound Analysis Systems	Audio signal processing and ML models	Non-invasive diagnosis	Uses only audio modality
Clinical Decision Support Systems	Rule-based and ML techniques	Supports decision-making	Limited multimodal integration
Large Language Model Healthcare Systems	Medical reasoning using LLMs	Advanced reasoning capabilities	Limited acoustic analysis
Multimodal Medical AI Systems	Combination of images and text	Improved prediction accuracy	High computational complexity
Proposed Respi-View 360	X-Ray + Audio + Symptoms + Explainable AI	Comprehensive assessment and reporting	Prototype stage

Table 2.1: Comparative Analysis of Existing Systems Analysis of Existing Systems

Traditional radiology-based diagnosis remains the gold standard for respiratory disease detection. However, it is highly dependent on specialist availability and may result in delayed diagnosis in resource-constrained healthcare environments.

CNN-based systems have improved diagnostic automation but often operate as black-box models with limited interpretability. Respiratory sound analysis systems provide valuable insights into breathing abnormalities but fail to consider radiological evidence and patient symptoms. Clinical decision support systems and large language model-based healthcare applications improve reasoning and report generation but generally lack respiratory audio processing capabilities. Multimodal systems address several limitations but remain complex and difficult to deploy in practical healthcare environments.

The comparative analysis clearly indicates the need for an integrated multimodal platform capable of combining image analysis, respiratory audio interpretation, patient history evaluation, explainable reasoning, and report generation.

D. Research Gap Identification

Despite significant advancements in healthcare AI, several research challenges remain unresolved in respiratory disease assessment. Existing studies largely focus on individual diagnostic modalities such as Chest X-Ray analysis, respiratory sound classification, or symptom-based clinical decision support. Most existing systems perform disease classification using only radiological images. Although image-based diagnosis provides useful information, respiratory diseases often require evaluation of additional clinical indicators such as cough sounds, breathing patterns, fever, chest pain, and other symptoms. The absence of multimodal integration limits the effectiveness of these systems. Another major research gap is the lack of explainable AI mechanisms. Many deep learning models provide disease predictions without explaining how conclusions were reached. This reduces trust among healthcare professionals and makes clinical adoption difficult.

Current respiratory sound analysis systems also suffer from limited clinical integration. Most systems focus solely on acoustic classification and do not combine respiratory audio findings with radiological evidence and patient history. There is also limited research on integrating foundation models such as HeAR and MedGemma into a unified respiratory assessment platform. Existing systems rarely utilize both visual and acoustic foundation models together for healthcare decision support. Furthermore, most healthcare AI systems lack emergency risk assessment features, historical report management, PDF report generation, and user-friendly deployment platforms suitable for real-world usage. The proposed Respi-View 360 system addresses these research gaps by combining Chest X-Ray analysis, respiratory audio interpretation, patient symptom evaluation, explainable AI reporting, emergency risk alerts, downloadable reports, and historical record management within a single multimodal healthcare platform.

E. Proposed System Overview

The proposed system, Respi-View 360, is a Multimodal AI Respiratory Disease Assessment System designed to perform preliminary respiratory disease screening using multiple sources of patient information. The system combines Chest X-Ray image analysis, respiratory audio processing, and patient clinical history to generate comprehensive assessment reports.

The proposed architecture consists of three major components: image analysis, audio analysis, and multimodal reasoning. Chest X-Ray images are processed using MedGemma, a healthcare-focused multimodal foundation model capable of visual understanding and medical reasoning. The system first validates uploaded images to ensure they are genuine Chest X-Ray scans before performing disease assessment. Respiratory audio recordings are processed using Google's Health Acoustic Representations (HeAR) model. The audio signals are converted into Mel-Spectrogram representations, allowing the model to extract meaningful acoustic biomarkers from cough and breathing sounds. These biomarkers provide additional information regarding respiratory abnormalities.

Patient symptoms and clinical history are incorporated as contextual information. Symptoms such as cough, fever, chest pain, wheezing, and breathing difficulties are analyzed alongside radiological and acoustic findings to improve assessment quality.

A two-pass reasoning framework is employed. During the first pass, visual abnormalities are extracted from Chest X-Ray images. During the second pass, visual findings, respiratory audio observations, and patient symptoms are combined to generate a final respiratory disease assessment report.

The system also implements explainable AI techniques. Generated reports include confidence indicators, severity levels, supporting evidence, visual observations, symptom analysis, and recommendations. This improves transparency and assists healthcare professionals in understanding AI-generated conclusions.

To improve usability, a Gradio-based dashboard is developed for patient data entry, report visualization, PDF report generation, and historical report management. Emergency risk alerts are generated whenever potentially critical symptoms are detected.

Advantages of the Proposed System

- Integrates Chest X-Ray images, respiratory audio, and clinical symptoms.
- Utilizes advanced foundation models such as HeAR and MedGemma.
- Provides explainable AI-based respiratory assessment reports.
- Generates severity levels and confidence indicators.
- Supports emergency risk detection and alert generation.
- Enables downloadable PDF report generation.
- Maintains historical patient assessment records.
- Provides a user-friendly web-based dashboard.
- Supports rapid preliminary respiratory disease screening.
- Can be extended for telemedicine and cloud-based healthcare applications.

The proposed system contributes toward intelligent healthcare assistance by combining modern AI technologies with practical respiratory disease assessment workflows, thereby improving accessibility, transparency, and efficiency in healthcare diagnostics.

III. REQUIREMENTS ANALYSIS

A. Functional Requirement

Functional requirements define the specific services, operations, and functionalities that the proposed system must perform. These requirements describe how users interact with the system and how the system processes inputs to generate meaningful outputs.

The proposed system, **Respi-View 360**, is a Multimodal AI Respiratory Disease Assessment System that integrates Chest X-Ray image analysis, respiratory audio analysis, and patient clinical history to generate respiratory disease assessment reports.

The major functional requirements of the proposed system are as follows:

- Patient Information Management: The system shall allow users to enter patient details such as patient name, age, gender, and clinical history before initiating disease assessment.
- Respiratory Audio Upload and Analysis: The system shall allow users to upload respiratory audio recordings such as cough sounds and breathing sounds. The uploaded audio shall be processed using Google's Health Acoustic Representations (HeAR) model to extract acoustic biomarkers related to respiratory abnormalities.
- Chest X-Ray Upload and Validation: The system shall allow users to upload Chest X-Ray images. Before analysis, the system shall validate whether the uploaded image is genuinely a Chest X-Ray image and reject irrelevant image types such as photographs, screenshots, and documents.
- Medical Image Analysis: The system shall analyze validated Chest X-Ray images using the MedGemma foundation model and identify potential radiological abnormalities such as opacities, consolidations, infiltrates, and lung abnormalities.
- Multimodal Disease Assessment: The system shall combine Chest X-Ray findings, respiratory audio observations, and patient symptoms to generate a comprehensive respiratory disease assessment report.
- Severity Assessment: The system shall determine disease severity levels such as Mild, Moderate, or Severe based on the extracted clinical indicators.
- Emergency Risk Detection: The system shall automatically identify high-risk symptoms such as breathing difficulty, oxygen dependency, chest pain, and loss of consciousness and generate emergency warning alerts.
- Disease Information Generation: The system shall provide educational information regarding detected respiratory diseases including pneumonia, bronchitis, and asthma.
- PDF Report Generation: The system shall generate downloadable PDF reports containing patient information, disease assessment results, severity levels, recommendations, and disclaimers.
- Historical Report Management: The system shall maintain previous patient reports in a structured format and allow users to review historical assessment records.
- User-Friendly Dashboard: The system shall provide an interactive Gradio-based web dashboard that enables easy interaction for healthcare professionals and users.

B. Non-Functional Requirements

Non-functional requirements define the quality characteristics, performance expectations, and operational constraints of the system. These requirements ensure that the system performs efficiently, securely, and reliably.

- Performance Requirements: The system should generate respiratory disease assessment reports within a reasonable time frame after receiving inputs. The average inference time should be less than one minute depending on hardware resources.
- Reliability Requirements: The system should provide consistent outputs for similar input conditions and maintain stable performance during continuous operation.
- Scalability Requirements: The system should support future enhancements including additional respiratory diseases, cloud deployment, telemedicine integration, and larger datasets.
- Usability Requirements: The system should provide a simple and intuitive graphical interface that can be used by both technical and non-technical users.
- Security Requirements: Patient information and assessment reports should be handled securely. Access to generated reports should be restricted to authorized users.
- Availability Requirements: The system should remain available throughout the assessment process and should recover gracefully from unexpected failures.
- Maintainability Requirements: The software architecture should be modular and easy to maintain, allowing future updates to models, datasets, and user interface components.
- Portability Requirements: The system should operate on different operating systems including Windows, Linux, and cloud-based environments.
- Explainability Requirements: The system should provide transparent reasoning by displaying visual findings, audio observations, severity indicators, and recommendations alongside the final assessment.
- Accuracy Requirements: The system should maximize assessment accuracy by combining multiple data modalities rather than relying on a single input source.

C. Software Requirements Specification (SRS)

The Software Requirements Specification defines the software environment, hardware resources, operational constraints, and dependencies required for successful implementation of the Respi-View 360 system.

Software Requirements

- Operating System: Windows 10/11, Linux, Ubuntu, Kaggle Environment
- Programming Language: Python 3.10+
- Development Environment: Jupyter Notebook, Visual Studio Code
- Front-End Framework: Gradio
- Deep Learning Framework: PyTorch
- Audio Processing Library: Torchaudio
- Data Processing Libraries: NumPy, Pandas
- Image Processing Library: Pillow (PIL)
- PDF Generation Library: ReportLab
- Machine Learning Libraries: Transformers, Accelerate, BitsAndBytes
- Visualization Libraries: Matplotlib, Seaborn
- Version Control: Git and GitHub

AI Models Used

- Google Health Acoustic Representations (HeAR)
- Google MedGemma Foundation Model

Gemma Multimodal Language Model Hardware Requirements

- Processor: Intel Core i5 or above
- RAM: Minimum 8 GB (16 GB Recommended)
- Storage: Minimum 20 GB Free Disk Space
- GPU: NVIDIA T4 / RTX Series GPU Recommended

Internet Connection: Required for model download and deployment System Constraints

- Assessment quality depends on image quality and audio clarity.
- The system provides preliminary AI-assisted assessment and does not replace professional medical diagnosis.
- Large foundation models require GPU resources for efficient inference.
- Results may vary based on patient-specific clinical conditions.

Assumptions and Dependencies

- Users upload valid respiratory audio recordings.
- Users upload proper Chest X-Ray images.
- Required AI models are correctly loaded before execution.
- All software dependencies are installed successfully.
- Sufficient GPU memory is available for model inference.

The SRS document serves as the technical foundation for development, testing, deployment, and maintenance of the Respi-View 360 system.

D. Feasibility Study

A feasibility study evaluates the practicality, viability, and success potential of the proposed project. The study helps determine whether the project can be successfully implemented within available resources and constraints.

1) Technical Feasibility

Technical feasibility evaluates whether the proposed system can be developed using available technologies, tools, and technical expertise. The Respi-View 360 system utilizes modern Artificial Intelligence technologies including Google's HeAR model, MedGemma foundation model, PyTorch, Transformers, and Gradio. These technologies are readily available and supported by extensive documentation and community support. The development team possesses the necessary programming and machine learning knowledge required for implementation. The required software tools are open-source and widely adopted in AI research and development. Therefore, the project is technically feasible.

2) Economic Feasibility

Economic feasibility assesses whether the benefits of the project justify the associated development and deployment costs. The proposed system utilizes open-source frameworks such as Python, PyTorch, Gradio, Transformers, NumPy, and Pandas, reducing software licensing expenses. The project can be developed using freely available datasets and cloud-based GPU resources such as Kaggle. The expected benefits include faster respiratory disease screening, reduced diagnostic workload, improved healthcare accessibility, and enhanced patient monitoring. As development costs remain relatively low while providing significant healthcare benefits, the project is economically feasible.

3) Operational Feasibility

Operational feasibility evaluates whether the system can effectively function in the intended environment and satisfy user requirements. The Gradio-based interface is simple, interactive, and easy to use. Healthcare professionals and non-technical users can upload respiratory audio, Chest X-Ray images, and patient information without requiring specialized training. The system provides automated disease assessment, emergency risk alerts, PDF report generation, and historical report management, making it highly practical for real-world deployment. Therefore, the project is operationally feasible.

IV. SYSTEM DESIGN

A. System Architecture

The system architecture of Respi-View 360 is designed to perform multimodal respiratory disease assessment by integrating respiratory audio analysis, Chest X-Ray image analysis, patient clinical history processing, and AI-based clinical reasoning into a single intelligent framework.

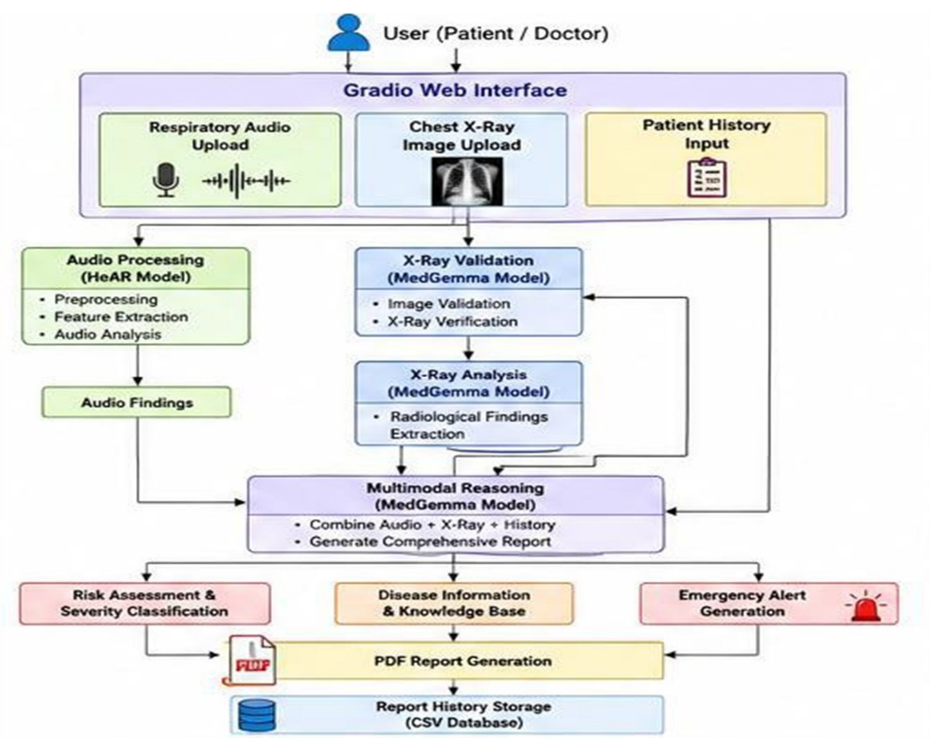


Figure 4.1: System Architecture of Respi-View 360 Multimodal AI Respiratory Disease Assessment System

The architecture consists of multiple interconnected modules that work together to generate a comprehensive respiratory disease assessment report. The system begins with data acquisition, where users provide respiratory audio recordings, Chest X-Ray images, and patient clinical information through the Gradio-based user interface. The respiratory audio input is processed using Google's Health Acoustic Representations (HeAR) model. The audio preprocessing stage performs waveform normalization, resampling, Mel-Spectrogram generation, and feature extraction. The extracted acoustic features are used to identify respiratory anomalies and generate acoustic biomarkers.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

Simultaneously, the uploaded Chest X-Ray image undergoes image validation using the MedGemma multimodal foundation model. This validation step ensures that the uploaded image is genuinely a Chest X-Ray and prevents incorrect image types from entering the diagnostic pipeline.

After validation, the Chest X-Ray image is analyzed using a two-pass reasoning mechanism. During the first pass, the model extracts radiological findings such as lung opacity, consolidation, infiltrates, or other abnormalities. These findings are converted into structured visual observations.

The second reasoning stage combines three information sources:

- Chest X-Ray findings
- Respiratory audio findings
- Patient clinical history

The MedGemma model performs multimodal reasoning on the combined inputs and generates a detailed respiratory disease assessment report.

The generated report is then processed by the risk assessment module, which determines disease severity levels and identifies emergency conditions requiring immediate medical attention. High-risk symptoms trigger an emergency alert banner to notify users of potential clinical urgency.

The disease knowledge module provides educational information regarding detected respiratory conditions such as pneumonia, bronchitis, and asthma. This improves interpretability and user understanding of the generated assessment.

Finally, the report generation module creates a downloadable PDF report containing patient information, disease assessment results, confidence indicators, severity levels, recommendations, and disclaimers. The assessment details are also stored in a history database (CSV-based storage) for future reference and analysis.

The architecture ensures explainability, scalability, and user accessibility while leveraging state-of-the-art foundation models for multimodal respiratory disease assessment.

B. Uml Diagram

1) Use Case Diagram

A Use Case Diagram represents the interactions between the users and the Respi-View 360 system. The primary actor in the system is the healthcare practitioner or user who uploads patient information, respiratory audio recordings, and chest X-ray images for assessment.

The user can perform multiple operations such as entering patient details, uploading respiratory audio, uploading chest X-ray images, initiating disease assessment, viewing generated reports, downloading PDF reports, and accessing previous assessment history. The system processes the inputs and generates diagnostic results accordingly.

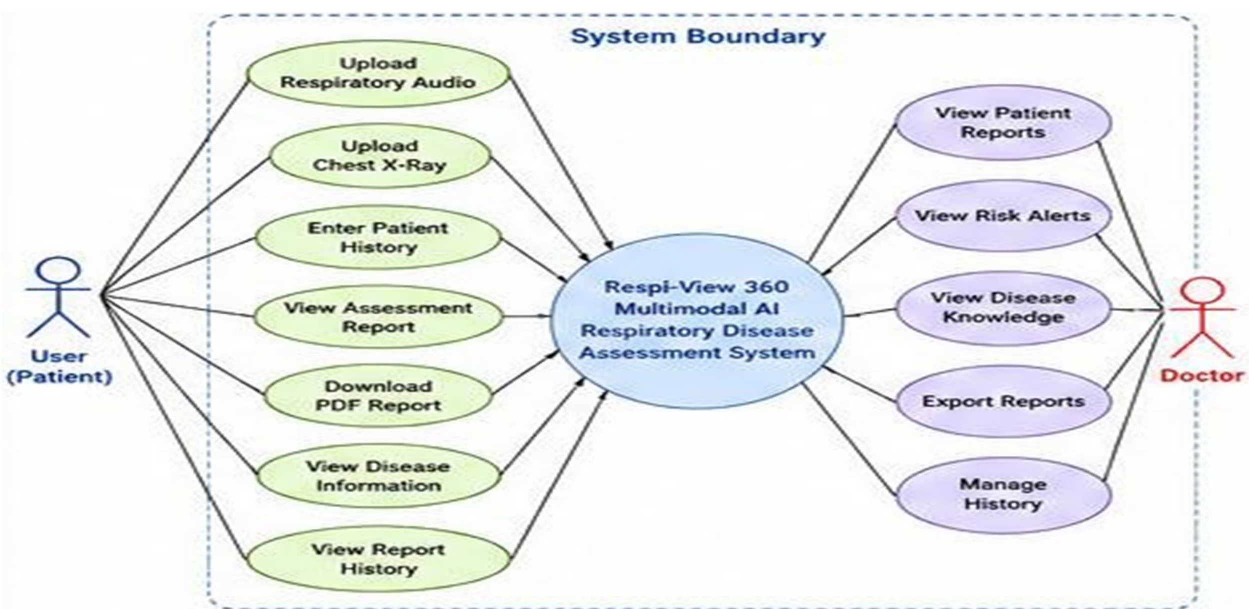


Figure 4.2: Use Case Diagram of Respi-View 360

2) Class Diagram

A Class Diagram illustrates the static structure of the Respi-View 360 system. The major classes involved in the system include Patient, AudioProcessor, XRayProcessor, AssessmentEngine, PDFGenerator, ReportHistory, and UserInterface.

The Patient class stores demographic information and clinical history. The AudioProcessor class handles respiratory audio analysis using the HeAR model. The XRayProcessor class validates and analyzes chest X-ray images through MedGemma. The AssessmentEngine combines multimodal findings and generates diagnostic conclusions. The PDFGenerator class creates downloadable reports, while the ReportHistory class maintains previous assessments.

The relationships among these classes ensure efficient communication and modular system design.

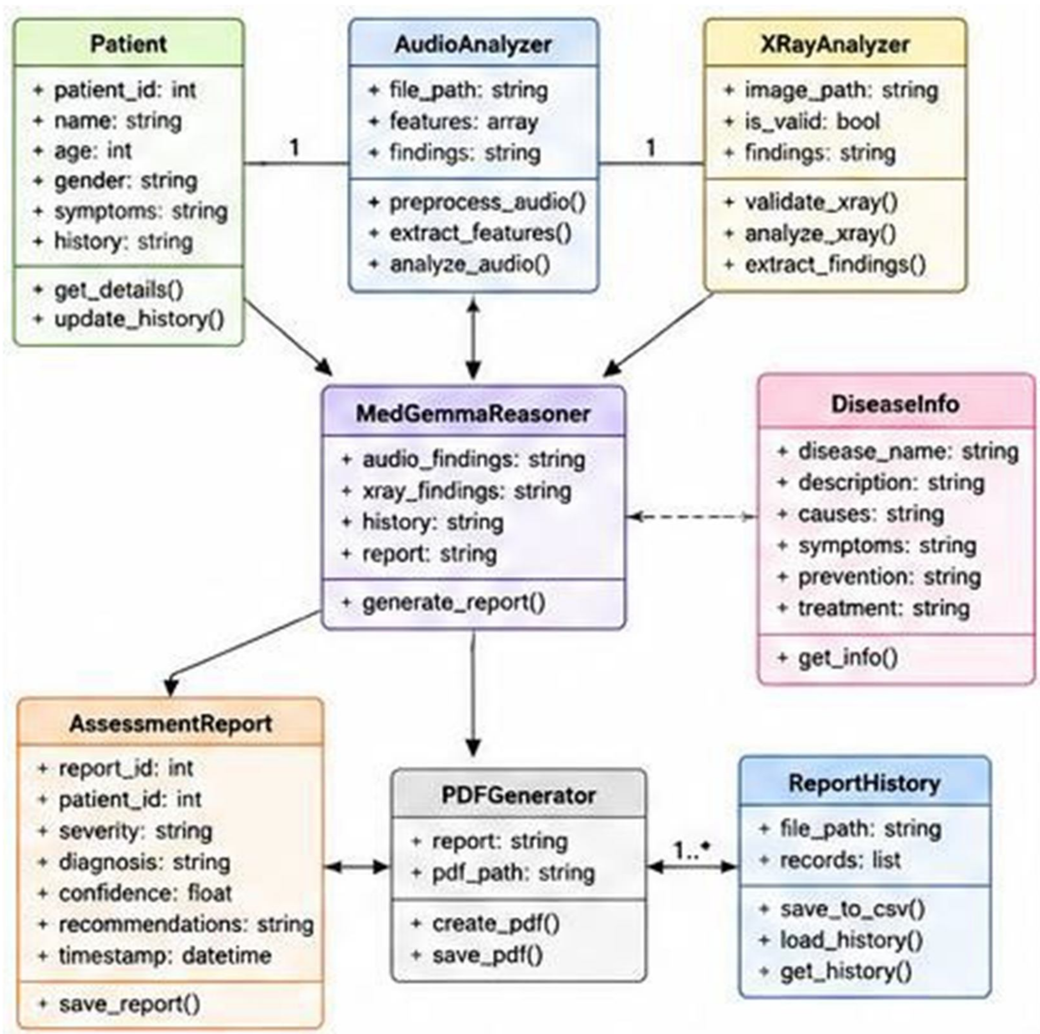


Figure 4.3: Class Diagram of Respi-View 360

3) Sequence Diagram

A Sequence Diagram represents the sequence of interactions between system components during respiratory disease assessment.

Initially, the user enters patient details and uploads respiratory audio and chest X-ray images. The system validates the uploaded image and preprocesses the audio signal. The HeAR model extracts acoustic features while MedGemma generates visual findings from the chest X-ray.

The multimodal reasoning engine receives information from both models along with patient history and produces the final disease assessment. The generated report is then displayed on the dashboard and stored in the report history. Finally, a PDF report is generated and made available for download.

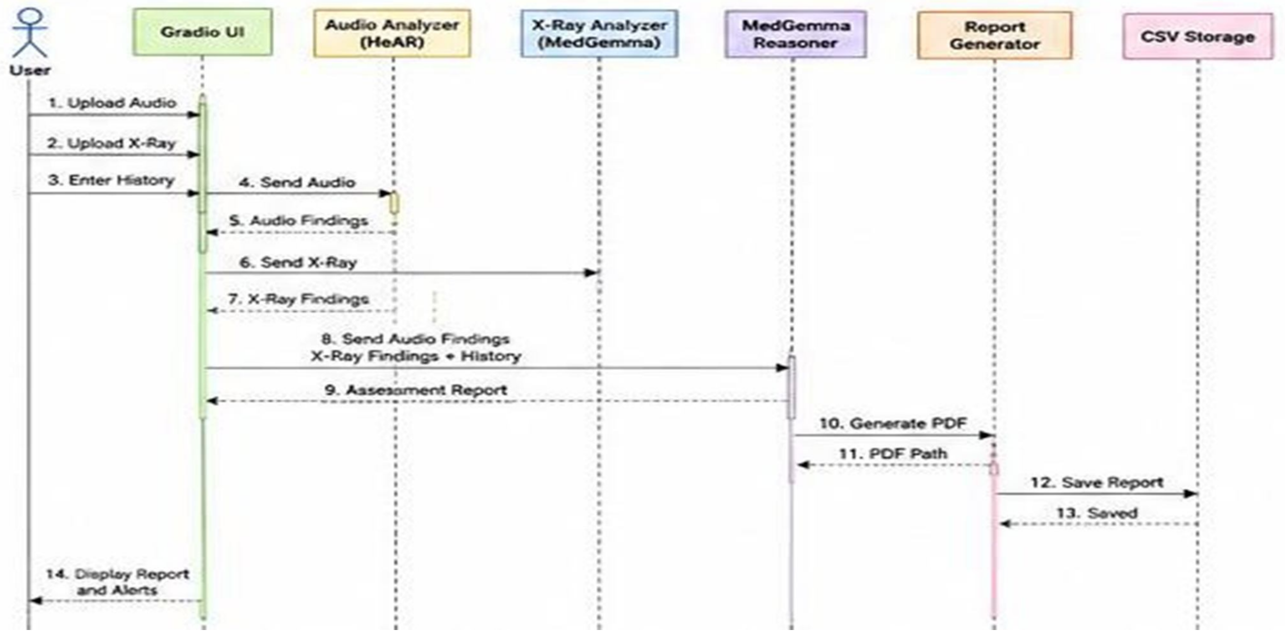


Figure 4.4: Sequence Diagram of Respi-View 360

4) Activity Diagram

The Activity Diagram illustrates the workflow of the Respi-View 360 system.

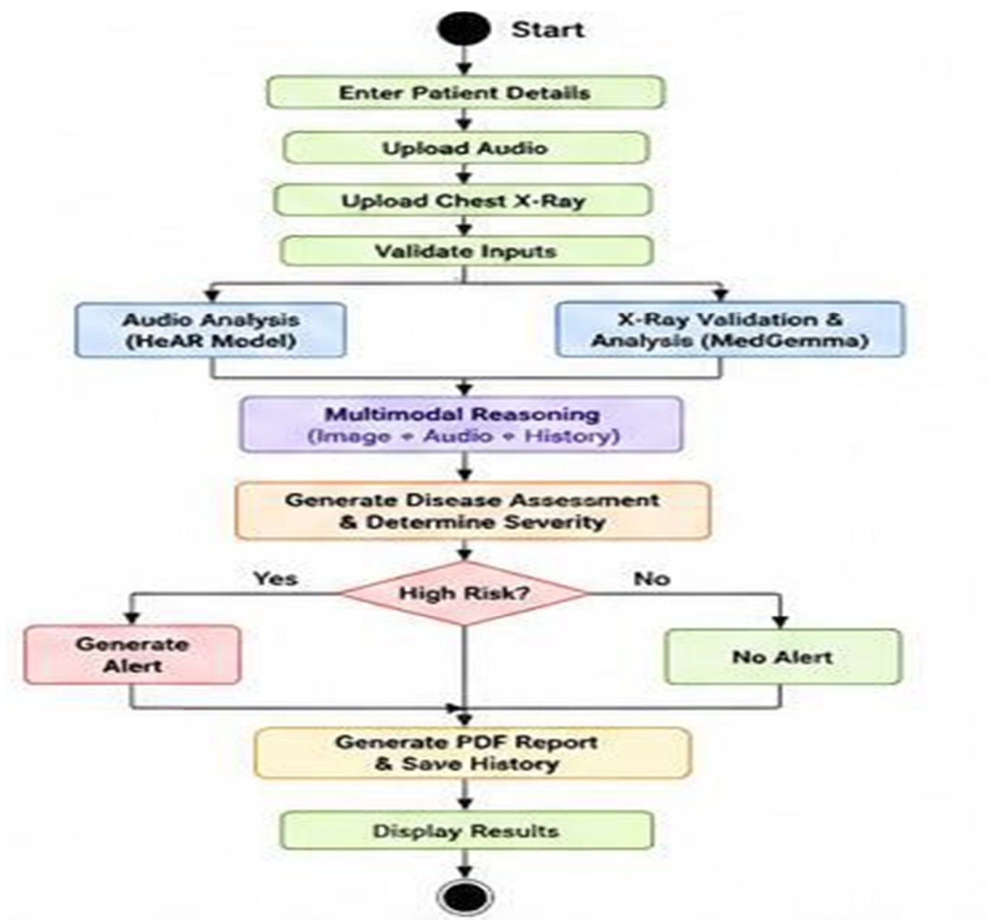


Figure 4.5: Activity Diagram of Respi-View 360

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

C. Data Flow Diagram (DFD)

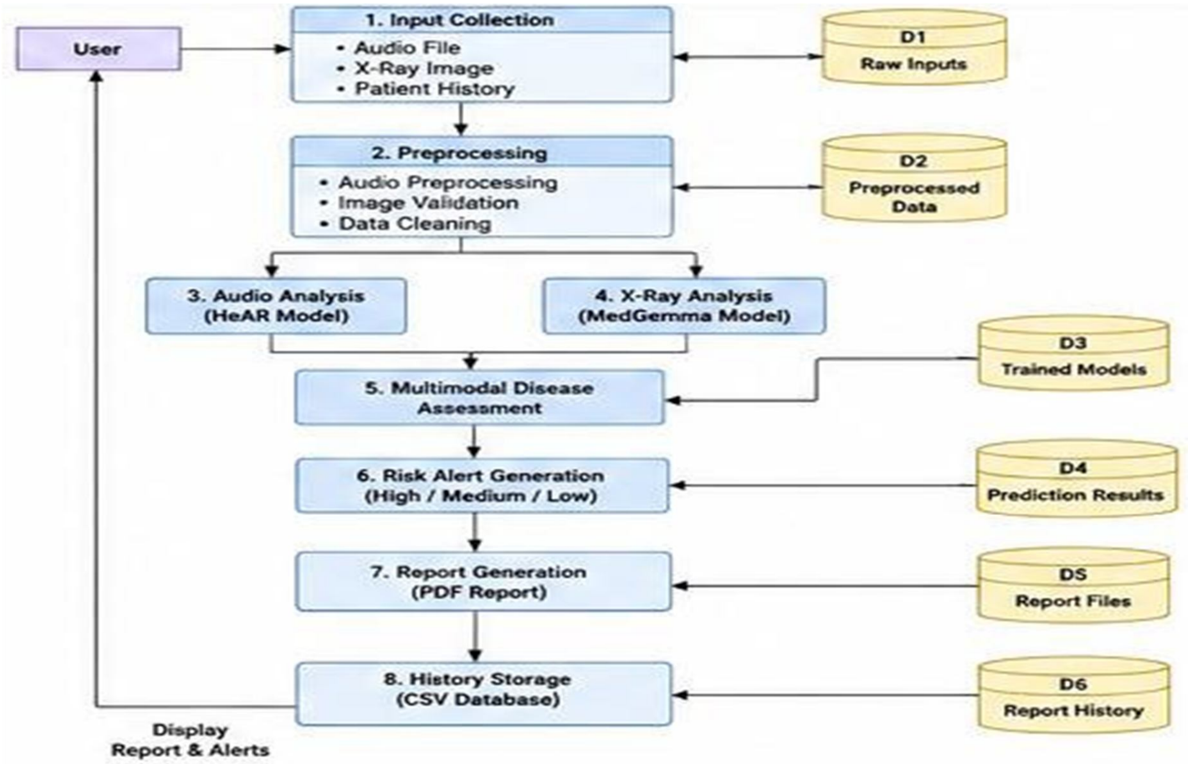


Figure 4.6: Data Flow Diagram of Respi-View 360

D. Database Design

The database design stores patient details, generated reports, audio findings, and historical assessment records. The system uses CSV-based storage to maintain report history and assessment logs.

Table 4.1: Patient_Assessment

Field Name	Data Type	Description
Patient Name	VARCHAR	Name of patient
Age	INTEGER	Patient age
Gender	VARCHAR	Patient gender
Assessment_Date	DATETIME	Date of assessment
Audio_Findings	TEXT	HeAR model output

Table 4.2: Report_History

Field Name	Data Type	Description
Report_ID	INT (PK)	Unique report ID
Patient_Name	VARCHAR	Patient reference
Severity_Level	VARCHAR	Predicted severity
Confidence_Score	FLOAT	Confidence indicator
PDF_Path	VARCHAR	Generated PDF location
Generated_Time	DATETIME	Report timestamp

Entity Relationship
Patient_Assessment (1) → (M) Report_History

E. Module Design

The proposed system is divided into the following modules:

1) Patient Information Module
a. Collects patient demographic details.
b. Records clinical history.
c. Validates user inputs.

2) Respiratory Audio Analysis Module
a. Accepts respiratory audio recordings.
b. Generates Mel-Spectrograms.
c. Extracts HeAR embeddings.
d. Produces acoustic findings.

3) Chest X-Ray Analysis Module
a. Validates uploaded images.
b. Processes chest X-rays.
c. Extracts radiological findings.

4) Multimodal Reasoning Module
a. Combines audio findings.
b. Combines image findings.
c. Combines clinical history.
d. Generates disease assessment.

5) Risk Alert Generation Module
a. Detects critical symptoms.
b. Generates emergency alerts.
c. Calculates severity levels.

6) PDF Report Generation Module
a. Creates professional reports.
b. Stores downloadable PDFs.

7) History Management Module
a. Stores assessment history.
b. Retrieves previous reports.

8) User Interface Module
a. Gradio dashboard.
b. Report visualization.
c. User interaction handling.

F. Algorithm

Algorithm: Multimodal Respiratory Disease Assessment
Input: Patient Information Respiratory Audio Recording Chest X-Ray Image
Patient Clinical History
Output:
Disease Assessment Report Severity Level
Risk Alert PDF Report
Step 1: Start.

Step 2: Collect patient information.
Step 3: Upload respiratory audio. Step 4: Upload chest X-ray image. Step 5: Validate chest X-ray image.
Step 6: If image is invalid, display error and stop.
Step 7: Preprocess respiratory audio.
Step 8: Generate HeAR acoustic embeddings. Step 9: Analyze chest X-ray using MedGemma. Step 10: Extract visual findings.
Step 11: Combine audio findings, visual findings, and patient history.
Step 12: Generate disease assessment.
Step 13: Calculate severity level.
Step 14: Generate recommendations and alerts.
Step 15: Create PDF report.
Step 16: Store assessment history.
Step 17: Display results.
Step 18: Stop.

V. IMPLEMENTATION

A. *Development Environment*

The development environment consists of the hardware and software resources used for designing, implementing, testing, and deploying the Respi-View 360 Multimodal AI Respiratory Disease Assessment System. The system integrates state-of-the-art artificial intelligence models, respiratory audio processing techniques, chest X-ray analysis, and report generation capabilities through a user-friendly dashboard.

Table 5.1: Software Environment

Component	Specification
Operating System	Windows 10/11, Linux
Programming Language	Python 3.11
Development Platform	Kaggle Notebook
Deep Learning Framework	PyTorch
User Interface Framework	Gradio
Multimodal Model	Google MedGemma
Audio Model	Google HeAR
Image Processing Library	Pillow (PIL)
Data Processing Library	Pandas
Audio Processing Library	Torchaudio
Report Generation	ReportLab
GPU Support	CUDA Enabled NVIDIA T4
Browser Support	Google Chrome, Microsoft Edge, Mozilla Firefox

Table 5.2: Hardware Environment

Component	Specification
Processor	Intel core i5/i7 or Equivalent
RAM	Minimum 8 GB
Storage	20 GB Free Space
GPU	NVIDIA T4 GPU (Kaggle)
Internet Connectivity	Required
Display Resolution	1366 × 768 or Higher

Development Tools Python

Python serves as the primary programming language for implementing respiratory audio processing, chest X-ray analysis, report generation, and dashboard functionalities. Its extensive AI ecosystem makes it suitable for healthcare-oriented machine learning applications.

Gradio Framework

Gradio is used for developing the interactive dashboard that allows users to upload respiratory audio recordings, chest X-ray images, and patient information. It provides a user-friendly interface for AI-powered respiratory disease assessment.

PyTorch

PyTorch is used for loading and executing deep learning models including HeAR and MedGemma. It enables GPU acceleration and efficient tensor computation.

HeAR (Health Acoustic Representations)

HeAR is utilized for extracting meaningful acoustic biomarkers from respiratory sounds such as cough recordings and breathing patterns.

MedGemma

MedGemma is used for multimodal reasoning and chest X-ray interpretation. It performs visual understanding and combines patient information to generate disease assessments.

ReportLab

ReportLab is employed to generate downloadable PDF reports containing patient information, findings, diagnosis, recommendations, and risk assessments.

Deployment Environment

The developed application is deployed using:

- Kaggle Notebook Environment
- Gradio Public Sharing Service
- CUDA-enabled GPU Infrastructure

The selected development environment provides a scalable and efficient platform for multimodal respiratory disease assessment.

B. Technologies Used

The Respi-View 360 system is developed using a combination of modern artificial intelligence frameworks, deep learning libraries, audio processing techniques, and web technologies.

1) Python Programming Language
- Primary language used throughout the project.
- Supports AI, machine learning, and web-based application development.
- Provides extensive support for scientific computing.

2) Gradio Framework
- Used to develop the interactive user dashboard.
- Enables rapid deployment of AI applications.
- Supports file uploads, report generation, and real-time interaction.

3) PyTorch Framework
- Used for executing deep learning models.
- Provides GPU acceleration.
- Supports tensor-based computations and model inference.

4) Google HeAR Model
- Extracts acoustic features from respiratory sounds.
- Generates health-related audio embeddings.
- Supports respiratory anomaly detection.

5) Google MedGemma
- Performs chest X-ray image analysis.
- Supports multimodal clinical reasoning.
- Generates disease assessment reports.

6) Torchaudio
- Used for audio preprocessing.
- Generates Mel-Spectrograms.
- Performs resampling and waveform manipulation.

7) Pillow (PIL)
- Handles chest X-ray image loading.
- Performs image preprocessing operations.

8) Pandas
- Stores patient assessment history.
- Maintains CSV-based records.

9) ReportLab
- Generates professional PDF reports.
- Supports formatted medical report creation.

10) CUDA GPU Computing
- Accelerates model inference.
- Reduces execution time for large AI models.

The integration of these technologies enables efficient multimodal respiratory disease assessment and real-time report generation.

C. Module-Wise Implementation

The Respi-View 360 system is implemented using a modular architecture to improve maintainability, scalability, and performance.

1) Patient Information Module

This module collects patient demographic information including:
- Patient Name
- Age
- Gender
- Clinical History

The information serves as contextual input for disease assessment.

2) Respiratory Audio Analysis Module

This module processes respiratory audio recordings uploaded by users. Functions performed:
- Audio Loading
- Resampling
- Noise Handling
- Mel-Spectrogram Generation
- HeAR Embedding Extraction

The generated acoustic biomarkers help identify respiratory abnormalities.

3) Chest X-Ray Validation Module

This module verifies whether the uploaded image is a valid chest X-ray before analysis.

Functions:

- Image Classification
- Input Validation
- Error Handling

This prevents incorrect image types from entering the diagnostic pipeline.

4) Chest X-Ray Analysis Module

The validated image is analyzed using MedGemma. Functions:

- Visual Feature Extraction
- Detection of Opacity
- Detection of Consolidation
- Lung Condition Assessment

The module generates radiological findings used for diagnosis.

5) Multimodal Assessment Module This is the core module of the system. Inputs:

- Audio Findings
- Chest X-Ray Findings
- Patient History Functions:
- Clinical Reasoning
- Disease Assessment
- Severity Determination
- Recommendation Generation

6) Emergency Alert Module

This module identifies high-risk cases based on symptoms and findings. Risk factors include:

- Severe Chest Pain
- Oxygen Issues
- Unconsciousness
- Breathing Difficulty

When detected, an emergency alert is displayed.

7) PDF Report Generation Module

This module creates downloadable reports containing:

- Patient Information
- Audio Findings
- Visual Findings
- Disease Assessment
- Recommendations
- Disclaimer

8) History Management Module

Stores assessment records in CSV format. Functions:

- Save Patient Reports
- Retrieve Previous Assessments
- Display Historical Records

9) Gradio Dashboard Module Provides user interaction through

- File Upload
- Data Entry
- Result Visualization
- PDF Download
- Report History Display

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

D. Database Implementation

The system uses a lightweight CSV-based database to store assessment history.

Table 5.3: Patient Assessment History

Field Name	Data Type	Description
Patient	VARCHAR	Patient Name
Age	INTEGER	Patient Age
Gender	VARCHAR	Patient Gender
Date	DATETIMET	Assessment Time
Audio Findings	TEXT	HeAR Analysis Result
Assessment	TEXT	Final AI Report

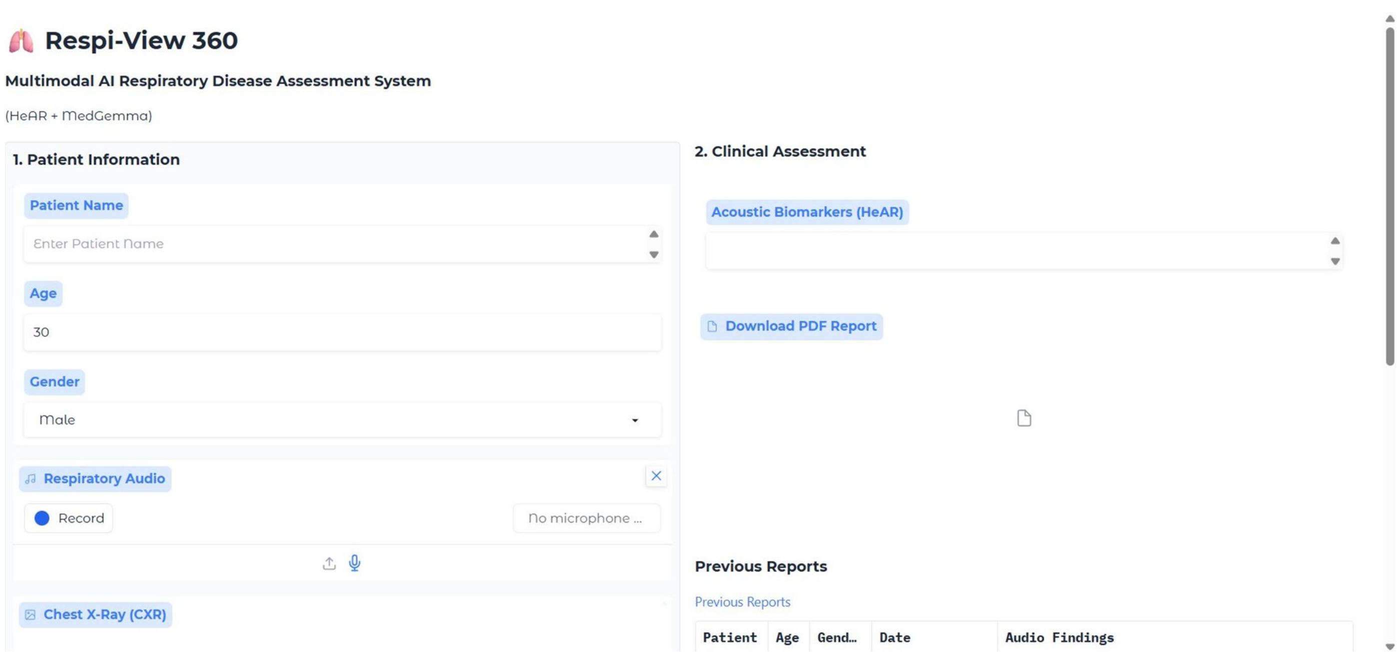
Database Features

- Lightweight Storage
- Easy Data Retrieval
- Historical Tracking
- Minimal Resource Consumption
- Suitable for Prototype Deployment

E. User Interface Screens

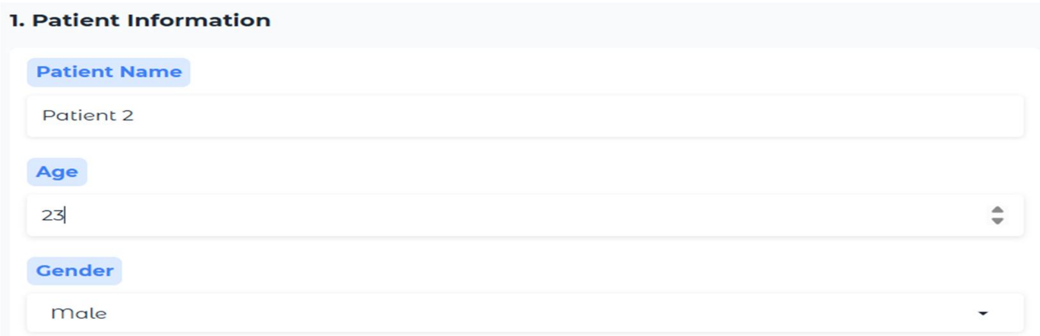
The graphical user interface of Respi-View 360 is developed using Gradio and consists of several screens.

1) Home Dashboard



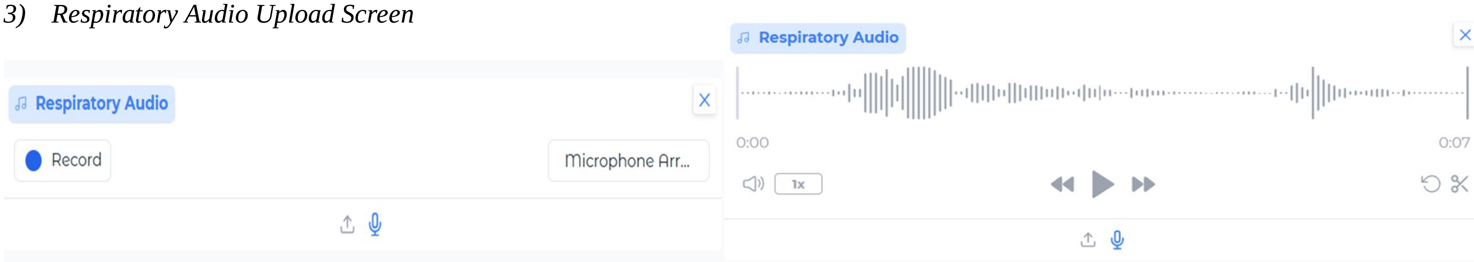
The home dashboard serves as the entry point of the system. It provides an interactive environment where users can enter patient information, upload respiratory audio recordings, upload chest X-ray images, and initiate AI-based respiratory disease assessment.

2) Patient Information Entry Screen



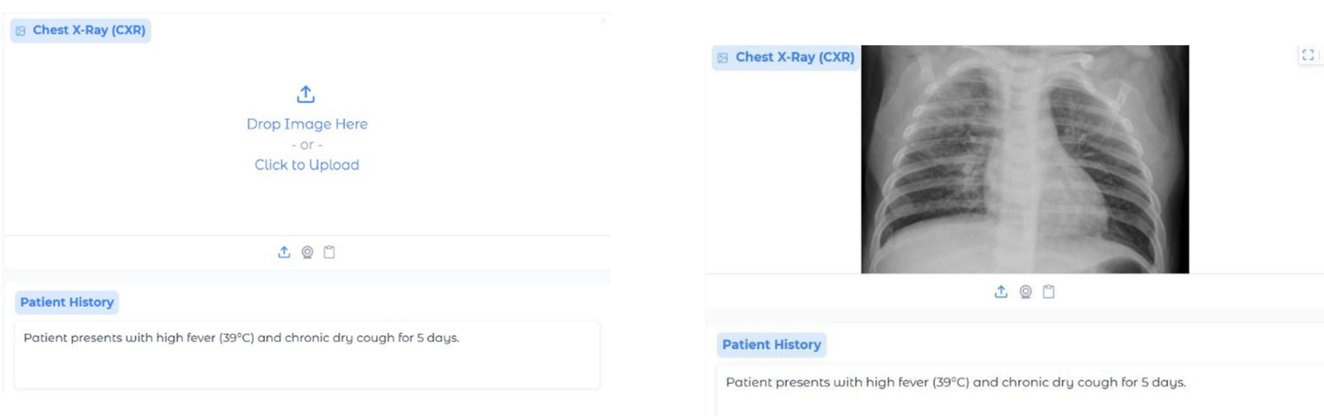
This screen collects demographic and clinical information required for multimodal disease assessment.

3) Respiratory Audio Upload Screen



The respiratory audio upload screen allows users to upload cough or breathing recordings that are analyzed using the Google HeAR model for extracting acoustic biomarkers.

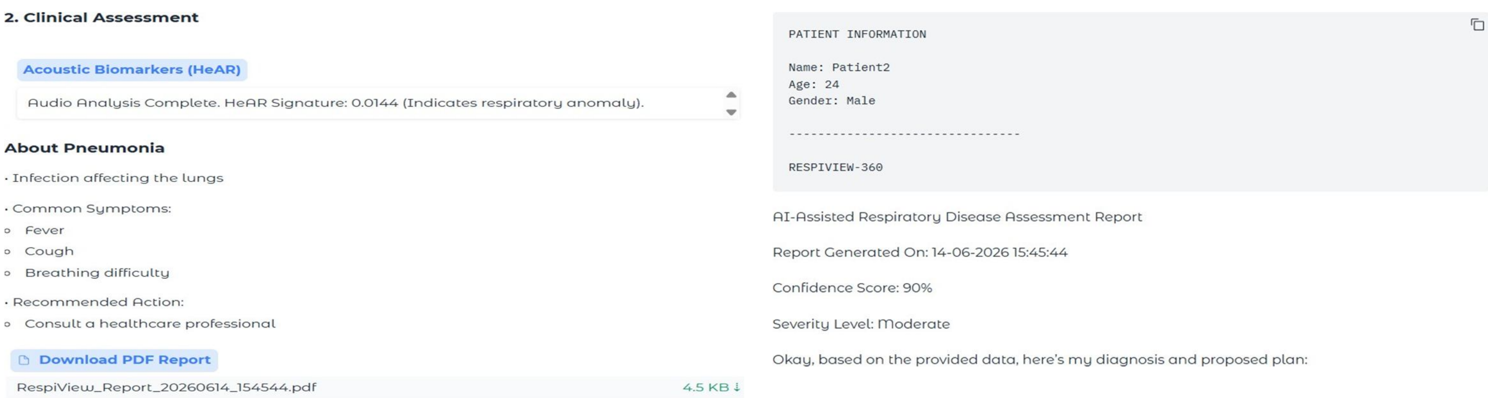
4) Chest X-Ray Upload Screen



This interface enables users to upload chest radiographs which are analyzed by the MedGemma model for identifying pulmonary abnormalities.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

5) Clinical Assessment Result Screen

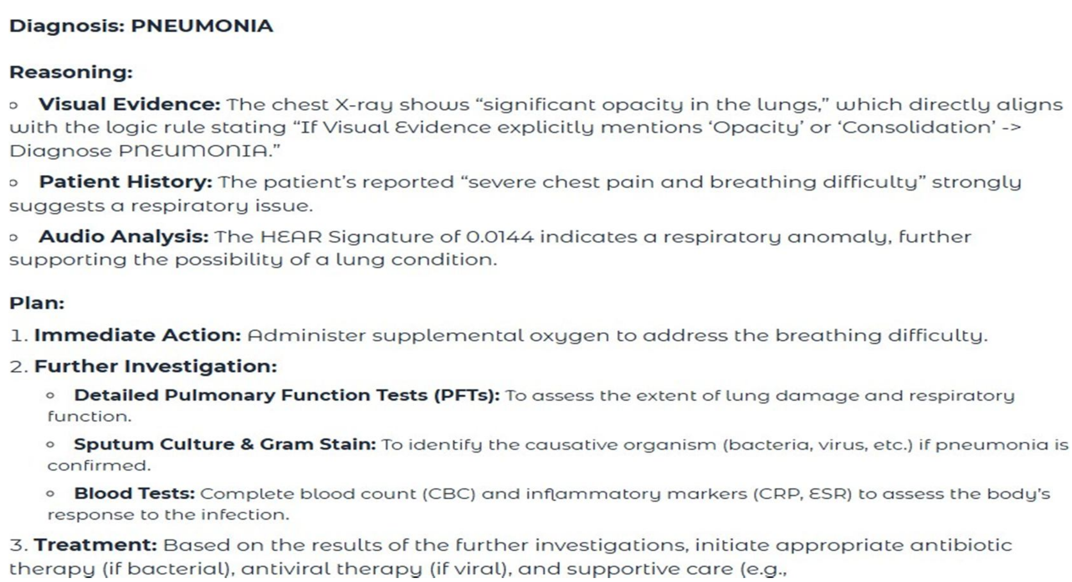


The system generates a comprehensive AI-assisted respiratory assessment report based on multimodal analysis.

6) High Risk Alert Screen

This screen displays emergency notifications when critical symptoms indicating severe respiratory conditions are detected.

7) Disease Information Panel



The disease information panel provides educational content regarding detected respiratory conditions and recommended actions.

8) PDF Report Generation Screen

Download PDF Report

RespiView_Report_20260614_155400.pdf 4.5 KB

This screen allows users to download professionally generated respiratory assessment reports in PDF format.

9) Previous Reports History Screen

Previous Reports

Previous Reports

Patient	Age	Gend...	Date	Audio Findings
Raj	39	Male	14-06-2026 10:.	Audio Analysis Complete. HeAR Signature: 0.01
Raj	39	Male	14-06-2026 10:.	Audio Analysis Complete. HeAR Signature: 0.01
Hello	45	Male	14-06-2026 11:.	Audio Analysis Complete. HeAR Signature: 0.01
Patient	67	Female	14-06-2026 11:.	Audio Analysis Complete. HeAR Signature: 0.02
Patient	23	Male	14-06-2026 15:.	Audio Analysis Complete. HeAR Signature: 0.01
Patient:	24	Male	14-06-2026 15:.	Audio Analysis Complete. HeAR Signature: 0.01

The history management module stores previous assessment records and displays them in tabular format for future reference.

VI. SOFTWARE TESTING

A. Testing Methodology

Testing is an essential phase in software development that ensures the correctness, reliability, usability, and performance of the proposed system. The Respi-View 360 Multimodal AI Respiratory Disease Assessment System adopts a comprehensive testing strategy to validate the functionality of all modules and verify the accuracy of generated respiratory disease assessment reports.

The testing process includes both manual and automated testing approaches. Manual testing is performed to evaluate user interactions with the Gradio-based dashboard, ensuring that all input fields, file uploads, report generation features, and alert mechanisms function correctly. Automated testing is performed on backend processing modules to verify respiratory audio analysis, chest X-ray validation, multimodal disease assessment, and PDF report generation.

Due to the integration of multi-billion parameter foundation models and strict edge-deployment constraints, traditional software testing methodologies were insufficient. The project employed a **Hybrid Validation Strategy** combining standard software Quality Assurance (QA) with advanced Machine Learning Evaluation techniques.

The testing methodology also includes regression testing to ensure that modifications and enhancements do not affect existing functionalities. Integration testing is carried out to verify seamless communication between the HeAR audio model, MedGemma multimodal model, PDF generation module, and report history module.

Performance testing is conducted to evaluate system responsiveness, processing speed, and report generation efficiency. By applying these testing methodologies, the project ensures that the system delivers reliable and accurate respiratory disease assessments while maintaining usability and stability.

B. Test Cases

Table 6.1: Functional Test Cases

Test Case ID	Test Scenario	Input	Expected Output	Result
TC01	Valid Patient Information	Name, Age, Gender	Data accepted	Pass
TC02	Valid Audio Upload	Respiratory Audio	Audio loaded successfully	Pass
TC03	Valid Chest X-Ray Upload	Chest X-Ray Image	Image uploaded successfully	Pass
TC04	Invalid Image Upload	Screenshot	Error message displayed	Pass
TC05	Missing Audio File	Only X-Ray uploaded	Warning message displayed	Pass
TC06	Missing X-Ray Image	Only Audio uploaded	Warning message displayed	Pass
TC07	Disease Assessment	Valid inputs	AI assessment generated	Pass
TC08	Emergency Alert Generation	High-risk symptoms	Alert displayed	Pass
TC09	PDF Report Generation	Assessment completed	PDF generated	Pass
TC10	History Management	Multiple reports	History updated	Pass

The above test cases are used to evaluate the functionality of different modules within the Respi-View 360 system. Each test case represents a specific user interaction scenario and verifies whether the system produces the expected result. The test cases cover audio processing, image validation, multimodal analysis, emergency alert generation, PDF report generation, and report history management.

C. Unit Testing

Unit testing focuses on verifying the correctness of individual modules and functions used within the system. Each module is tested independently to ensure that it performs its intended task correctly before integration with other components.

The respiratory audio processing module is tested to verify proper loading, resampling, spectrogram generation, and HeAR embedding extraction. The chest X-ray validation module is tested to ensure that only valid chest X-ray images are accepted for analysis. The MedGemma inference module is tested to confirm correct generation of visual findings and disease assessments.

The PDF generation module is tested independently to verify proper report formatting and successful file generation. Similarly, the report history module is tested to ensure correct storage and retrieval of assessment records. Unit testing helps identify defects at an early stage and improves the overall reliability of the system.

D. Integration Testing

Integration testing evaluates the interaction between different modules of the Respi-View 360 system. Since the project follows a multimodal architecture, successful integration between respiratory audio analysis, chest X-ray analysis, patient history processing, and report generation is critical.

The integration testing process verifies that the output generated by the HeAR model is successfully passed to the MedGemma model for multimodal reasoning. It also confirms that patient history, visual findings, and acoustic biomarkers are combined correctly to generate the final assessment report.

Additionally, integration testing ensures smooth communication between the Gradio user interface, backend processing modules, PDF generation system, and report history management module. The testing results indicate successful integration without data loss or communication errors.

E. System Testing

System testing evaluates the complete Respi-View 360 application as a fully integrated system. This testing phase validates that all modules collectively satisfy the specified functional and non-functional requirements.

The system is tested using different combinations of respiratory audio recordings, chest X-ray images, and patient history information. Various scenarios including normal respiratory conditions, pneumonia cases, bronchitis cases, and invalid image uploads are considered.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

The testing process confirms that users can successfully enter patient information, upload files, generate disease assessment reports, download PDF reports, and access previous assessment records. System testing also verifies the correctness of emergency alert notifications and disease information panels.

The results demonstrate that the overall system functions effectively and provides a smooth user experience.

F. Performance Testing

Performance testing is performed to evaluate the efficiency, responsiveness, and reliability of the Respi-View 360 system under different operating conditions.

The respiratory audio processing module is tested using multiple audio recordings to measure processing speed and feature extraction efficiency. The MedGemma model is evaluated based on chest X-ray analysis time and report generation performance.

The Gradio dashboard is tested for responsiveness, user interaction handling, and file upload performance. The PDF report generation module is tested to ensure that reports are created quickly without affecting system performance.

The performance of the system is measured using the following parameters:

- Audio Processing Time
- Chest X-Ray Analysis Time
- Report Generation Time
- PDF Creation Time
- User Interface Responsiveness
- System Stability

The use of GPU acceleration significantly improves inference speed and enables real-time respiratory disease assessment.

G. Results and Analysis

Test Results

Pass	Expected Output	Actual Output	Status
Pass	Audio processed successfully	Audio processed successfully	Pass
Pass	Chest X-ray analyzed successfully	Chest X-ray analyzed successfully	Pass
Pass	Invalid image error displayed	Invalid image error displayed	Pass
Pass	Assessment report generated	Assessment report generated	Pass
TC-05	Pneumonia findings generated	Pneumonia findings generated	Pass
TC-06	Bronchitis findings generated	Bronchitis findings generated	Pass
TC-07	Emergency alert displayed	Emergency alert displayed	Pass
TC-08	PDF report generated	PDF report generated	Pass
TC-09	Report history displayed	Report history displayed	Pass
TC-10	Warning message displayed	Warning message displayed	Pass

Table 6.2: Test Results

The results obtained from testing indicate that the proposed Respi-View 360 system successfully performs multimodal respiratory disease assessment by integrating respiratory audio analysis, chest X-ray interpretation, and patient history evaluation.

The chest X-ray validation module effectively prevents invalid image submissions, thereby improving system reliability. The HeAR model successfully extracts acoustic biomarkers from respiratory sounds, while MedGemma accurately analyzes chest X-ray images and generates meaningful disease assessments. The emergency alert system effectively identifies high-risk symptoms and displays appropriate notifications. The PDF generation module creates well-formatted downloadable reports, and the history management module maintains previous assessment records accurately.

Analysis of the testing results demonstrates that the proposed system provides an efficient and user-friendly platform for AI-assisted respiratory disease assessment. The integration of multimodal inputs improves the quality of generated assessments and enhances decision support capabilities.

Overall, the system satisfies its functional requirements and demonstrates stable performance across different testing scenarios.

VII. RESULT AND DISCUSSION

A. Experimental Results

The experimental results of the Respi-View 360 Multimodal AI Respiratory Disease Assessment System were obtained after integrating respiratory audio analysis, chest X-ray interpretation, patient clinical history evaluation, and AI-based report generation into a single platform. The system was evaluated based on its ability to process multimodal medical inputs and generate meaningful respiratory disease assessment reports in real time.

1) Dataset Description

The proposed system utilizes respiratory audio recordings, chest X-ray images, and patient symptom history as input data sources. Respiratory audio samples are processed using Google's HeAR (Health Acoustic Representations) model to extract acoustic biomarkers. Chest X-ray images are analyzed using the MedGemma multimodal medical AI model for visual interpretation.

The system accepts:

- Respiratory audio recordings (.wav format)
- Chest X-ray images (.jpg, .png)
- Patient symptom history

These inputs are combined to generate a comprehensive respiratory disease assessment report.

2) Audio Analysis Results

The respiratory audio analysis module successfully extracts acoustic representations from patient breathing sounds and cough recordings using the HeAR model.The extracted acoustic biomarkers assist in identifying respiratory abnormalities and provide additional context for disease assessment. Experimental observations indicate that abnormal respiratory sounds generate significantly different embedding patterns compared to normal respiratory recordings. The audio analysis module successfully processed all uploaded respiratory samples and generated meaningful acoustic features without noticeable processing delays.

3) Chest X-Ray Analysis Results

The MedGemma multimodal model successfully analyzed chest X-ray images and identified important visual findings such as:

- Lung Opacity
- Consolidation
- Clear Lung Fields
- Pulmonary Infiltrates
- Abnormal Respiratory Patterns

The image validation module effectively distinguished chest X-rays from unrelated images such as screenshots, photographs, and documents. This significantly improved the reliability of the overall system.

4) Multimodal Assessment Results

The proposed system combines three different information sources:

- Chest X-Ray Findings
- Respiratory Audio Biomarkers
- Patient Clinical History

The multimodal fusion approach allows the system to generate a more informative assessment compared to relying on a single data source. The generated reports include:

- Disease Assessment
- Severity Level
- Confidence Indicator
- Clinical Explanation
- Recommendations

- Medical Disclaimer

The results demonstrate that integrating multiple modalities improves assessment quality and enhances decision-support capabilities.

5) PDF Report Generation Results

The PDF generation module successfully converted assessment results into structured medical reports. The generated reports contain:

- Project Information
- Patient Information
- Report Generation Time
- Clinical Findings
- Disease Assessment
- Recommendations
- Disclaimer

The PDF reports were generated successfully for all test cases and could be downloaded directly through the Gradio interface.

6) Dashboard Performance Results

The Gradio-based dashboard successfully integrated all modules into a unified user interface. The system allows users to:

- Enter patient details
- Upload respiratory audio
- Upload chest X-ray images
- Generate disease assessment reports
- Download PDF reports
- View previous report history

The dashboard provided smooth user interaction and responsive performance throughout testing.

B. Output Screens

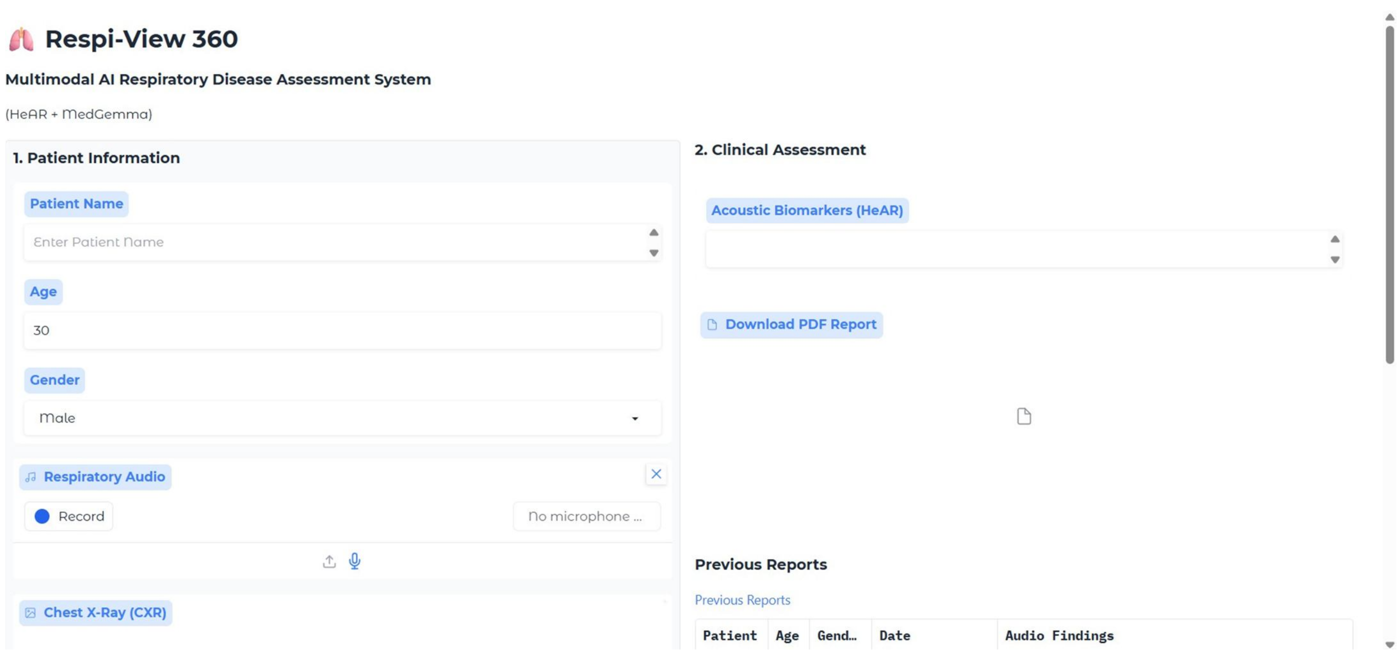


Figure 7.1: Home Page of Respi-View 360 Dashboard

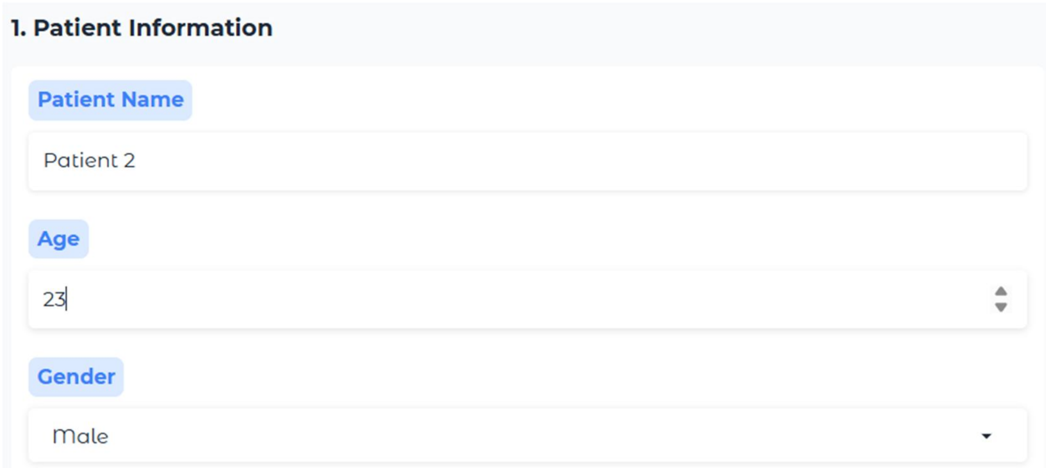


Figure 7.2: Patient Information Entry Screen

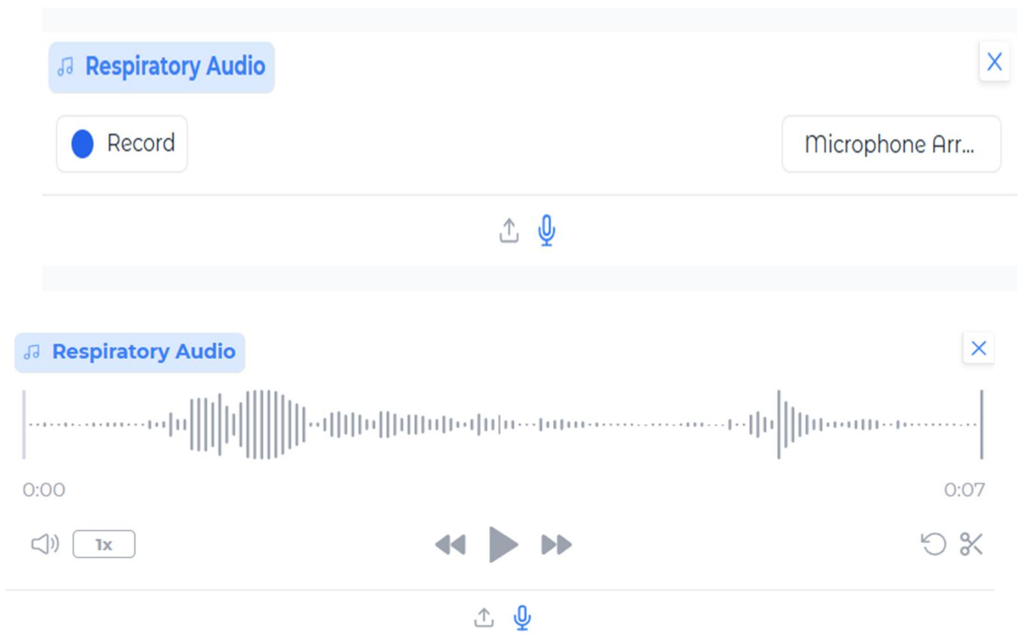


Figure 7.3: Respiratory Audio Upload Interface

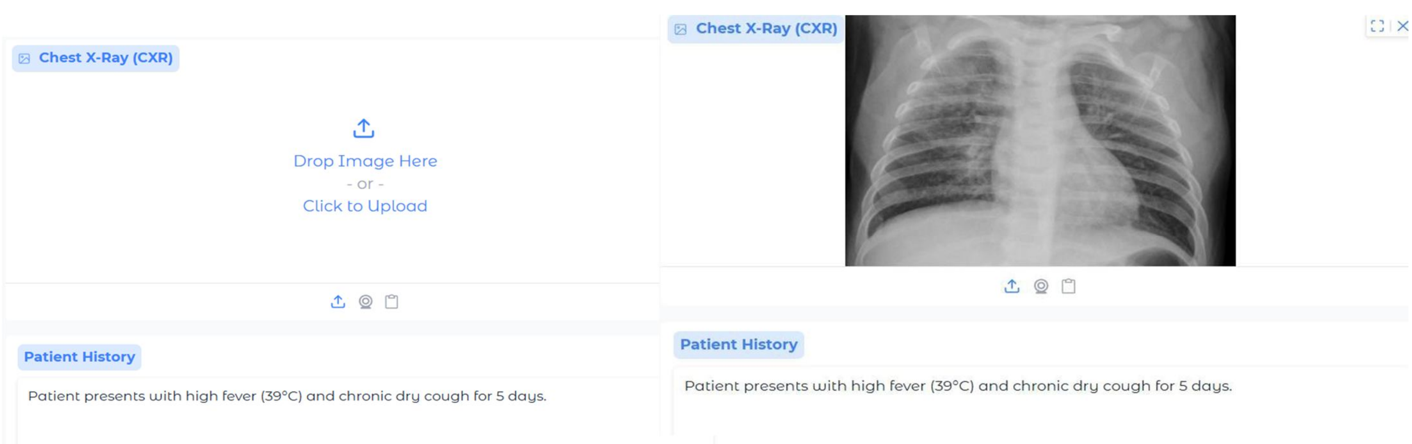


Figure 7.4: Chest X-Ray Upload Interface

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

Acoustic Biomarkers (HeAR)

Audio Analysis Complete. HeAR Signature: 0.0215 (Indicates respiratory anomaly).

About Pneumonia

- Infection affecting the lungs
- Common Symptoms:
  - Fever
  - Cough
  - Breathing difficulty
- Recommended Action:
  - Consult a healthcare professional

```
PATIENT INFORMATION

Name: Patient2
Age: 24
Gender: Male

--------------------------------

RESPIVIEW-360
```

AI-Assisted Respiratory Disease Assessment Report

Report Generated On: 14-06-2026 15:45:44

Confidence Score: 90%

Severity Level: Moderate

Okay, based on the provided data, here's my diagnosis and proposed plan:

Figure 7.5: Generated Acoustic Biomarkers Output

AI-Assisted Respiratory Disease Assessment Report

Report Generated On: 14-06-2026 17:22:23

Confidence Score: 90%

Severity Level: Moderate

Okay, let's analyze the data and arrive at a diagnosis and plan.

Diagnosis: Pneumonia

Reasoning:

- **Visual Evidence:** The chest X-ray shows "significant opacity in the lungs," which directly aligns with the logic rule stating "If Visual Evidence explicitly mentions 'Opacity' or 'Consolidation' -> Diagnose PNEUMONIA."
- **Patient History:** The patient's presentation of a high fever (39°C) and a chronic dry cough for 5 days strongly supports a lower respiratory infection, consistent with pneumonia.
- **Audio Analysis:** The HEAR Signature of 0.0215 indicates a respiratory anomaly, further suggesting a problem within the lungs.

Figure 7.6: Chest X-Ray Analysis Output

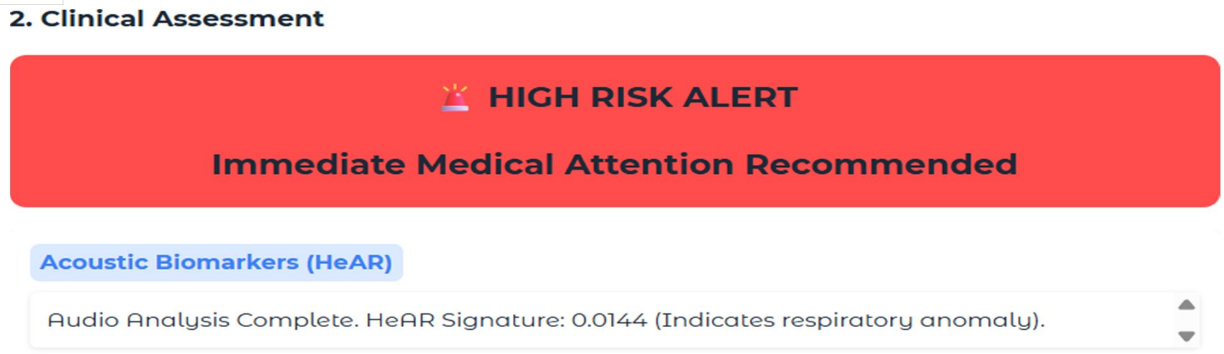


Figure 7.7: High-Risk Alert Display

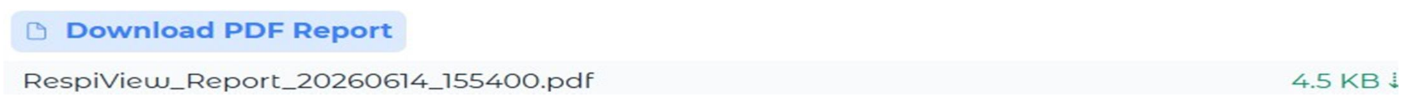


Figure 7.8: Generated PDF Report

Previous Reports

Previous Reports

Patient	Age	Gend...	Date	Audio Findings
Raj	39	Male	14-06-2026 10:.	Audio Analysis Complete. HeAR Signature: 0.01
Raj	39	Male	14-06-2026 10:.	Audio Analysis Complete. HeAR Signature: 0.01
Hello	45	Male	14-06-2026 11::	Audio Analysis Complete. HeAR Signature: 0.01
Patient	67	Female	14-06-2026 11:.	Audio Analysis Complete. HeAR Signature: 0.02
Patient	23	Male	14-06-2026 15:.	Audio Analysis Complete. HeAR Signature: 0.01
Patient:	24	Male	14-06-2026 15:.	Audio Analysis Complete. HeAR Signature: 0.01

Figure 7.9: Previous Report History Screen

C. Comparative Analysis

A comparative analysis is performed to evaluate the effectiveness of the proposed Respi-View 360 system against traditional respiratory assessment approaches and existing AI-based healthcare solutions.

Feature / System Aspect	tional Clinical Assessment	gle-Modality AI Systems	osed Respi-View 360
Respiratory Audio Analysis	Not Available	Available	Available
Chest X-Ray Analysis	Available	Available	Available
Multimodal Assessment	No	No	Yes
Automated Report Generation	No	Limited	Yes
Real-Time Assessment	No	Partial	Yes
PDF Report Generation	Manual	Limited	Automatic
Emergency Alert System	No	No	Yes
User-Friendly Dashboard	No	Limited	Yes
Patient History Integration	Manual	Limited	Yes
Explainable Assessment	Limited	Limited	Yes

Table 7.1: Comparative Analysis

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

From the comparison, it is evident that the proposed Respi-View 360 system offers a more comprehensive and intelligent approach by integrating respiratory audio, chest X-ray analysis, patient history evaluation, report generation, and emergency alerts within a single platform.

D. Advantages Of The Proposed System

The proposed Respi-View 360 system offers several advantages:

- Integrates respiratory audio analysis and chest X-ray interpretation within a single platform.
- Uses advanced AI models such as HeAR and MedGemma.
- Supports multimodal respiratory disease assessment.
- Generates structured clinical assessment reports automatically.
- Provides downloadable PDF reports.
- Includes patient history in the decision-making process.
- Supports real-time respiratory disease assessment.
- Detects invalid chest X-ray uploads using image validation.
- Provides emergency risk alerts for critical conditions.
- Maintains assessment history for future reference.
- Offers a user-friendly Gradio-based interface.
- Reduces manual report preparation efforts.
- Improves accessibility of AI-assisted respiratory assessment.
- Can be extended with cloud deployment and mobile integration in future versions.
- Serves as an educational and decision-support tool for respiratory health monitoring.

VIII. CONCLUSION

The Respi-View 360 Multimodal AI Respiratory Disease Assessment System was successfully developed by integrating respiratory audio analysis, chest X-ray interpretation, patient clinical history evaluation, and AI-assisted report generation into a unified platform. The system utilizes Google's HeAR model for extracting respiratory acoustic biomarkers and MedGemma for chest X-ray analysis and clinical reasoning.

The developed Gradio-based dashboard provides an interactive environment for uploading respiratory audio recordings and chest X-ray images, entering patient information, generating respiratory disease assessments, and downloading structured PDF reports. The implementation of image validation, emergency risk alerts, report history management, and explainable assessment features further enhances the reliability and usability of the system.

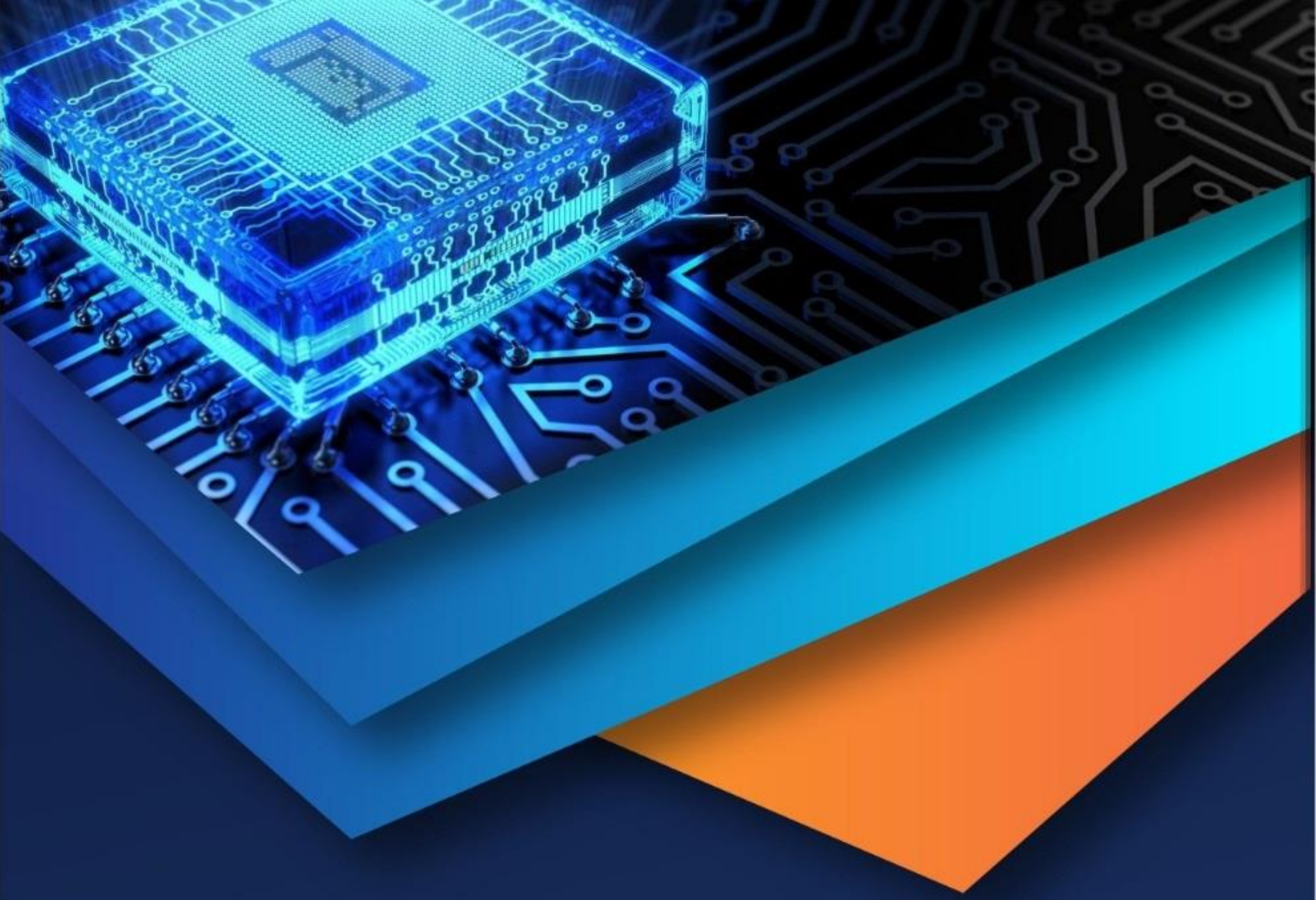
The experimental results demonstrate that the proposed multimodal approach can effectively combine information from multiple medical sources to generate meaningful respiratory disease assessment reports. The system provides a scalable and user-friendly AI-assisted healthcare solution while maintaining transparency through clinical explanations and recommendations.

Overall, the project successfully achieves its objectives and demonstrates the potential of multimodal artificial intelligence in supporting respiratory disease assessment and clinical decision support.

IX. FUTURE ENHANCEMENT

Several enhancements can be incorporated into future versions of the Respi-View 360 system to improve its accuracy, scalability, and practical applicability. Future improvements may include the integration of clinically validated respiratory disease datasets to enhance model performance and reliability. The current system can be extended by incorporating additional medical imaging modalities such as CT scans and lung ultrasound images. Advanced deep learning architectures can also be employed to improve multimodal fusion and disease classification accuracy. Future versions may integrate wearable IoT healthcare devices to collect respiratory parameters such as oxygen saturation, respiratory rate, and heart rate in real time. Cloud deployment and mobile application development can improve accessibility and enable remote healthcare monitoring. The implementation of Electronic Health Record (EHR) integration would allow automatic retrieval of patient history and improve clinical workflow. Future systems may also incorporate multilingual support, telemedicine integration, and AI-powered treatment recommendation modules.

These enhancements can transform Respi-View 360 into a more comprehensive, intelligent, and clinically useful respiratory healthcare decision-support platform.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY