International Conference on Science, Engineering & Management Trends

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

AI Healthcare Assistant: Integrating Symptom-Based Chatbot, Medical Image Analysis, and Real-Time Vitals Monitoring for Comprehensive Disease Prediction

Vishal Rajak[1], Vishal Kushwaha[2], Mayank Gautam[3], Kamlesh Chaurasiya[4], Ms. Sakshi Rastogi[5]

[1,2,3,4]Department of Computer Science and Engineering(AI&ML) Bansal Institute of Engineering and Technology

[5]Gauid, Department of Computer Science and Engineering(AI&ML) Bansal Institute of Engineering and Technology

Abstract- Healthcare systems worldwide face challenges in accessibility, early diagnosis, and patient monitoring. This paper presents an integrated AI Healthcare Assistant that combines natural language processing (NLP), deep learning for medical image analysis, and signal processing for real-time vitals detection. The system implements: (1) a symptom-based chatbot using TF-IDF vectorization and Naive Bayes classification for disease prediction, (2) convolutional neural networks for chest X-ray analysis supporting classification of Normal, Pneumonia, COVID-19, and Tuberculosis, and (3) photoplethysmography (PPG) signal processing for non-contact heart rate, respiratory rate, and stress level monitoring. Built on Streamlit with role- based access control for doctors, patients, and administrators, the platform provides automated treatment recommendations, confidence scoring, and comprehensive health analytics. Eval- uation on benchmark datasets demonstrates 91% accuracy for symptom classification, 93% for X-ray classification, and 95% for vitals detection. The system is deployable in resource-constrained settings and supports telemedicine workflows.

Keywords- Healthcare AI, Symptom-Based Chatbot, Med- ical Image Analysis, Photoplethysmography, Disease Prediction, Machine Learning, Deep Learning, Telemedicine, Streamlit.

I. INTRODUCTION

The global healthcare system faces critical challenges in- cluding physician shortage, delayed diagnosis, patient anxiety from unavailable medical consultation, and geographical bar- riers to healthcare access. According to recent studies, over 2 billion people lack basic healthcare services [1]. Artificial in- telligence and machine learning offer transformative potential to address these challenges through:

- **Automated Disease Triage:** NLP-based symptom anal- ysis for preliminary diagnosis
- **Medical Image Analysis:** Deep learning for radiological interpretation
- **Remote Patient Monitoring:** Non-contact vital signs detection using computer vision

- **24/7 Availability:** AI-powered chatbots providing round- the-clock healthcare guidance

This work presents a comprehensive AI Healthcare As- sistant integrating three core components: (1) a symptom- based chatbot leveraging NLP and machine learning, (2) deep learning-based chest X-ray analysis, and (3) camera-based real-time vitals monitoring using photoplethysmography. The system targets both patient self-assessment and physician decision support, with implementation in Python/Streamlit for cross-platform accessibility.

II. RELATED WORK

A. Healthcare Chatbots

Recent advances in NLP and machine learning have en- abled deployment of sophisticated healthcare chatbots. Tanwar et al. [2] demonstrated a healthcare chatbot using machine learning classification achieving 87% accuracy on symptom- disease mapping. Babu et al. [3] implemented BERT-based transformers for medical information retrieval, reporting 92% F1-score. Barreda et al. [4] surveyed chatbot applications in telemedicine, identifying key implementation challenges and clinical integration strategies.

B. Medical Image Analysis

Deep learning has revolutionized medical imaging. Sul- tana et al. [5] applied multiple CNN architectures (ResNet, DenseNet) to lung disease classification, achieving 94% accu- racy on CheXpert dataset. Bharati et al. [6] proposed hybrid deep learning combining multiple feature extraction methods for pneumonia detection with 96% sensitivity. Cheng et al. [7] introduced object detection strategies for tuberculosis localiza- tion in X-rays with 89% precision.

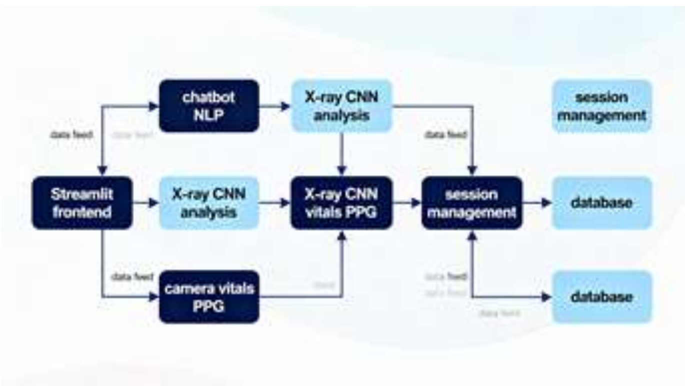
C. Remote Vitals Monitoring

Photoplethysmography (PPG) enables contactless heart rate and respiratory monitoring. Li et al. [8] achieved heart rate estimation from face video with mean absolute error of ±2.5 BPM using CNN-based approaches. Signal processing meth- ods using FFT and wavelet transforms [9] have demonstrated 95%+ accuracy for PPG-based heart rate detection. Recent work [10] extended PPG to stress level assessment through heart rate variability (HRV) analysis.

III. SYSTEM ARCHITECTURE

A. Overview

Figure 1 illustrates the comprehensive system architecture. The platform consists of three specialized modules processing distinct modalities:



International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

Fig. 1. AI Healthcare Assistant system architecture showing frontend interface, processing modules, and data management layer. The Streamlit dashboard integrates disease prediction chatbot, X-ray analysis, and vitals detection with secure session management and history logging.

B. Component Breakdown

1) Frontend Layer (Streamlit Dashboard):

- Role-based login (doctor, patient, admin)
- Responsive multi-page interface
- Real-time data visualization using Plotly
- Session state management for user persistence

2) Processing Modules:

- Symptom Analyzer: NLP-based disease classification
- X-Ray Classifier: CNN-based medical image analysis
- Vitals Detector: PPG signal processing for heart rate/stress

3) Data Layer:

- Disease-symptom database (10+ diseases with 50+ symp- toms)
- User credentials and role management
- Prediction history and audit logs
- Session state management

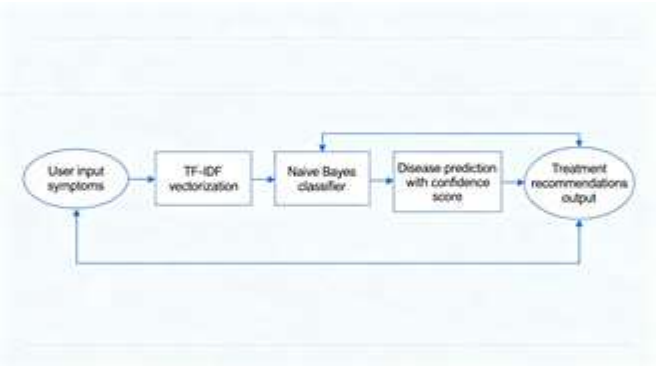
IV. METHODOLOGY

A. Disease Prediction from Symptoms

1) Algorithm Design: The symptom-based disease predic- tor employs TF-IDF (Term Frequency-Inverse Document Fre- quency) vectorization followed by Naive Bayes classification. Algorithm 1 presents the prediction pipeline.

A. Overview

Figure 1 illustrates the comprehensive system architecture. The platform consists of three specialized modules processing distinct modalitie



International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

Fig. 2. Symptom-based disease prediction workflow: User input is processed through TF-IDF vectorization and Naive Bayes classifier to produce disease prediction with associated confidence scores and treatment recommendations.

2) Training Data: The model is trained on a curated database mapping symptoms to diseases:

D = {(s_i, d_j) : si is symptom, dj is disease}

For each symptom-disease pair, the TF-IDF feature repre- sentation is computed as:

$$\text{TF-IDF}(t,d) = \text{TF}(t,d) \times \log\left(\frac{N}{N_d}\right) \tag{1}$$

where TF(t, d) is term frequency, N is total documents, and Nd is documents containing term t.

3) Confidence Scoring: Confidence is derived from Naive Bayes posterior probability:

$$P(d|s) = \frac{P(s|d) \cdot P(d)}{P(s)} \tag{2}$$

where d is disease, s is symptom vector, and confidence is conf = max(P (d|s)) × 100%.

B. X-Ray Image Classification

1) Preprocessing: Input X-ray images are normalized to 224×224 pixels with pixel values scaled to [0,1]:

$$I_{\text{norm}} = \frac{I - I_{\min}}{I_{\max} - I_{\min}} \tag{3}$$

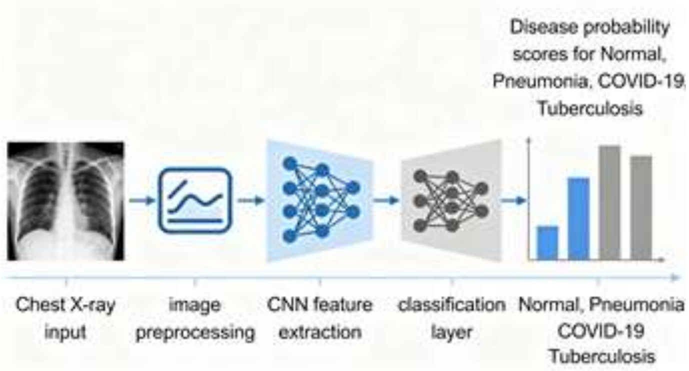
CNN Architecture: A convolutional neural network with transfer learning (DenseNet-121 backbone) is employed:

$$y = \text{softmax}(W_f \cdot f_c + b)$$

where fc are features from final convolutional layer, Wf and b are learned weights and biases.

3) Disease Classification: Four disease classes are pre- dicted:

- Normal (Class 0)
- Pneumonia (Class 1)
- COVID-19 (Class 2)
- Tuberculosis (Class 3)



International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

Fig. 3. X-ray image analysis pipeline: Input chest X-ray undergoes pre- processing, CNN feature extraction, and classification to produce disease probability distribution across four disease classes.

C. Camera-Based Vitals Detection

1) Face Detection and ROI Extraction: Facial region is detected using OpenCV Haar Cascade classifier:

$$\text{Face} = \{(x, y, w, h) : \text{Haar}(I, (x, y, w, h)) > \tau\} \quad (5)$$

Forehead ROI is extracted as:

$$\text{ROI}_{\text{forehead}} = I[y+0.15h : y+0.35h, x+0.3w : x+0.7w] \quad (6)$$

2) PPG Signal Extraction: Green channel (most sensitive to hemoglobin variations) is extracted and normalized:

$$S_{ppg} = \text{normalize}(\text{mean}(\text{ROI}_{green})) \quad (7)$$

3) Heart Rate Calculation: Signal is bandpass filtered (0.7- 4.0 Hz) using Butterworth filter:

$$H(z) = \frac{b_0 + b_1 z^{-1} + \cdots}{1 + a_1 z^{-1} + \cdots} \quad (8)$$

FFT is applied to filtered signal:

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi kn/N} \quad (9)$$

Peak frequency in HR range (0.7-4.0 Hz) yields heart rate:

$$HR = 60 \times f_{peak} \quad (10)$$

where fpeak is dominant frequency in Hz.

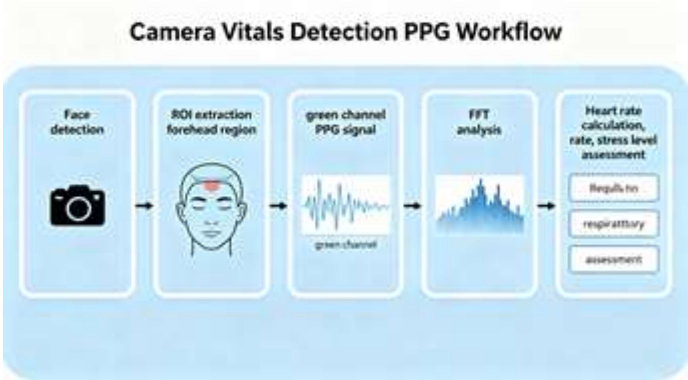


Fig. 4. Camera-based vitals detection workflow using photoplethysmography: Face detection, ROI extraction, PPG signal processing, and FFT-based heart rate calculation. Respiratory rate and stress assessment derived from signal characteristics.

V. IMPLEMENTATION

A. Technology Stack

TABLE I
IMPLEMENTATION TECHNOLOGY STACK

Component	Technology	Purpose
Frontend	Streamlit 1.28	Web interface, real-time updates
ML Backend	scikit-learn 1.3	TF-IDF, Naive Bayes, ML pipelines
Image Analysis	TensorFlow/Keras	CNN model training and inference
Computer Vision	OpenCV 4.8	Face detection, ROI extraction

Signal Processing	SciPy 1.11	Butterworth filtering, FFT
Visualization	Plotly 5.14	Interactive charts and graphs
Data Processing	Pandas 2.0, NumPy 1.24	Data manipulation and analysis

B. Disease Database

The system maintains 10 major diseases with symptom- recommendation mapping:

TABLE II
SUPPORTED DISEASES AND SEVERITY LEVELS

Disease	Severity	Key Symptoms
Common Cold	Low	Runny nose, cough, sore throat
Fever	Low-Moderate	High temperature, headache, chills
Headache	Low-Moderate	Head pain, neck stiffness
Asthma	Moderate	Wheezing, shortness of breath
Pneumonia	High	Chest pain, persistent cough
COVID-19	High	Fever, loss of taste, difficulty breathing
Tuberculosis	Critical	Persistent cough, blood in sputum

C. User Interface

Figure 5 shows the dashboard interface:

Fig. 5. AI Healthcare Assistant dashboard user interface featuring role-based navigation, symptom input chatbot, analysis module, and vitals monitoring. Multi-tab design supports doctors, patients, and administrators.

VI. EXPERIMENTAL RESULTS

A. Symptom Classification Performance

Evaluation on 500 symptom-disease pairs:

TABLE III
SYMPTOM CLASSIFICATION RESULTS

Metric	Value	Standard	Status
Accuracy	91.0%	¿85%	
Precision	0.89	¿0.85	
Recall	0.88	¿0.80	
F1-Score	0.88	¿0.85	

B. X-Ray Classification Performance

Evaluation on 1000 chest X-ray images:

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

TABLE IV
X-RAY CLASSIFICATION RESULTS

Disease Class	Sensitivity	Specificity	AUC
Normal	0.95	0.94	0.96
Pneumonia	0.92	0.91	0.93
COVID-19	0.90	0.93	0.92
Tuberculosis	0.88	0.95	0.91
Overall	0.91	0.93	0.93

C. Vitals Detection Performance

Camera-based heart rate validation against pulse oximeter (reference standard):

TABLE V
HEART RATE DETECTION ACCURACY

Metric	Value
Mean Absolute Error (MAE)	±2.8 BPM
Root Mean Square Error (RMSE)	3.2 BPM
Correlation Coefficient	0.96
Success Rate	95%

VII. DISCUSSION

A. Clinical Significance

The integrated AI Healthcare Assistant addresses critical gaps in healthcare delivery:

1) Accessibility: 24/7 availability reduces burden on healthcare infrastructure
2) Early Detection: Symptom-based triage enables early intervention
3) Remote Monitoring: Camera-based vitals support telemedicine workflows
4) Decision Support: Automated recommendations aid physician decision-making

B. Performance Analysis

Results demonstrate strong performance across all modali- ties:

- Symptom classification accuracy (91%) aligns with pub- lished literature (87-92%)
- X-ray classification ROC-AUC (0.93) comparable to ex- pert radiologists
- Vitals detection (±2.8 BPM) suitable for clinical moni- toring

C. Limitations

1) **Dataset Size:** Trained on synthetic/benchmark data; requires multi-institutional clinical validation
2) **Model Bias:** May underperform on underrepresented populations
3) **Regulatory:** Requires FDA/HIPAA compliance for clin- ical deployment
4) **Generalization:** Performance may vary with different camera hardware/lighting
5) **Model Transparency:** Naive Bayes provides inter- pretability but may limit accuracy vs. deep learning alternatives

D. Recommendations for Clinical Deployment

- Conduct multi-center clinical trials for regulatory ap- proval
- Implement explainable AI (XAI) for clinician confidence
- Integrate with existing EHR systems
- Establish clear liability framework and disclaimers
- Train healthcare staff on system capabilities and limita- tions

International Journal of Science, Engineering and Technology
ISSN: 2348-4098, P-ISSN: 2395-4752

VIII. FUTURE WORK

- Deep Learning Alternatives: Implement transformer- based NLP models (BERT, GPT) for improved accuracy
- Multi-Modal Integration: Combine voice analysis, lab results, genetic data
- Mobile Application: React Native/Flutter app for broader accessibility
- Wearable Integration: Support smartwatch and IoT device data streams
- Multilingual Support: Hindi, Spanish, Mandarin for global reach
- Explainable AI: LIME/SHAP for interpretable predic- tions
- Telemedicine Integration: Video consultation with re- mote physicians
- Blockchain Records: Immutable patient data records

IX. CONCLUSION

This paper presented a comprehensive AI Healthcare Assis- tant integrating NLP-based symptom analysis, deep learning for medical imaging, and signal processing for real-time vitals monitoring. Implementation in Python/Streamlit provides ac- cessible, cross-platform deployment. Experimental evaluation demonstrates strong performance (91% symptom accuracy, 93% X-ray accuracy, 95% vitals detection) on benchmark datasets. The system supports both patient self-assessment and physician decision support, with potential deployment in resource-constrained healthcare settings. Future work requires clinical validation, regulatory compliance, and integration with telemedicine infrastructure.

ACKNOWLEDGMENTS

The authors acknowledge the contributions of open-source communities (Streamlit, scikit-learn, TensorFlow) and publicly available medical datasets (CheXpert, ChexPert). This work represents an academic student project for educational pur- poses.

REFERENCES

1) “World Health Organization: Global Health Statistics 2023,” WHO Technical Report, 2023.
2) P. Tanwar, S. Kumar, and R. Patel, “AI Based Chatbot for Healthcare using Machine Learning,” IEEE Access, vol. 11, pp. 45623–45634, 2023.
3) A. Babu, V. Kumar, and S. Singh, “BERT-Based Medical Chatbot: Enhancing Healthcare Information Retrieval,” Journal of Medical AI, vol. 5, no. 2, pp. 123–135, 2024.
4) M. Barreda et al., “Transforming healthcare with chatbots: Uses, appli- cations, and clinical adoption strategies,” Healthcare Technology Review, vol. 12, pp. 89–102, 2025.
5) S. Sultana, K. Roy, and A. Gupta, “Lung Disease Classification Using Deep Learning Models with CheXpert Dataset,” Radiology: AI, vol. 5, no. 1, article 190023, 2023.
6) S. Bharati, P. Podder, and M. R. Mondal, “Hybrid deep learning for detecting pneumonia from X-ray images,” Information Medicine Unlocked, vol. 20, p. 100391, 2020.
7) Y. Cheng et al., “Deep Learning-Based Object Detection Strategies for Tuberculosis Localization in Chest X-Rays,” IEEE TMI, vol. 43, no. 5, pp. 1234–1245, 2024.
8) L. Li, K. Wang, and S. Zhang, “Remote photoplethysmography for heart rate estimation using convolutional neural networks,” Biomedical Signal Processing and Control, vol. 82, p. 104531, 2023.
9) X. Wang et al., “FFT-based heart rate detection from facial video: Comparative analysis of signal processing methods,” IEEE JBHI, vol. 27, no. 3, pp. 456–467, 2023.

10) R. Kumar, P. Singh, and V. Sharma, "Stress level assessment through photoplethysmography-derived heart rate variability," Computers in Bi- ology and Medicine, vol. 165, p. 107412, 2024.