International Journal of Contemporary Research in Computer Science and Technology (IJCRCST) ISSN: 2395-5325
Volume 9, Special Issue 1 (April '2026)
Special Issue on International Conference on Innovation, Engineering & Management (ICIEM-2026)
Conference held at Garden City University, Bangalore, India.

ARTIFICIAL INTELLIGENCE IN EARLY DISEASE DETECTION: ADVANCES, TECHNIQUES, AND APPLICATIONS ACROSS MAJOR HEALTH DOMAINS

Punith R
Post-Graduate (Reg. No: 24MSIC120),
Department of Computer Science and Information Technology,
Garden City University, Bengaluru, India

Abstract: Artificial Intelligence (AI) has emerged as a transformative technology in healthcare, particularly in the early detection, diagnosis, and management of diseases. Leveraging advanced machine learning, deep learning, explainable AI (XAI), ensemble learning, and edge computing techniques, AI enables the analysis of vast and complex biomedical datasets. This paper presents a comprehensive review of recent developments and applications of AI across multiple disease domains, including cardiovascular diseases, chronic kidney disease, retinal disorders, COVID-19, and plant diseases used as proxy models for XAI research. Key methodologies examined include convolutional neural networks (CNNs), transfer learning, generative adversarial networks (GANs), federated learning, Long Short-Term Memory (LSTM) networks, and ensemble models (Random Forest, XGBoost, CatBoost, LightGBM). The paper further examines emerging trends such as the integration of AI with the Internet of Medical Things (IoMT), blockchain for secure data sharing, and augmented reality for enhanced clinical decision support. Research gaps including limited explainability, dataset scarcity, clinical workflow integration, and multi-modal data fusion challenges are identified and discussed. This review provides a holistic perspective on the capabilities, challenges, and future directions of AI-driven disease detection, underscoring its potential to revolutionize healthcare delivery.

Keywords: *Artificial Intelligence, Machine Learning, Deep Learning, Disease Detection, Explainable AI (XAI), Convolutional Neural Networks, Ensemble Learning, Edge AI, Federated Learning, Cardiovascular Disease, Chronic Kidney Disease, Retinal Disorders, COVID-19, Transfer Learning, Generative Adversarial Networks*

I. INTRODUCTION

Early detection and accurate diagnosis of diseases are among the most critical determinants of effective treatment, improved patient outcomes, and reduced long-term healthcare costs. Globally, diseases such as cardiovascular conditions, chronic kidney disease, diabetic retinopathy, and infectious diseases like COVID-19 continue to impose enormous burdens on healthcare systems. Traditional diagnostic approaches — including radiological imaging, blood panel analysis, and histopathological examination — while effective, are often constrained by high operational costs, invasiveness, delayed turnaround times, geographic inaccessibility, and heavy reliance on specialist expertise. These limitations create systemic bottlenecks that delay treatment and worsen prognoses, particularly in low- and middle-income settings.

The rapid advancement of Artificial Intelligence (AI) over the past decade has introduced transformative solutions to these longstanding challenges. Modern AI techniques — encompassing machine learning (ML), deep learning (DL), ensemble methods, and explainable AI (XAI) — are now capable of processing large-scale, high-dimensional biomedical datasets with remarkable speed and accuracy. Applications range from the automated analysis of medical images using convolutional neural networks (CNNs) to the prediction of disease risk from routine blood biomarkers using gradient-boosted classifiers. These capabilities are increasingly complemented by technologies such as Edge AI, federated learning, and the Internet of Medical Things (IoMT), which enable distributed, real-time, and privacy-preserving health monitoring.

Despite these promising developments, the translation of AI research into clinical practice remains uneven. Barriers to adoption include the opacity of complex models (the 'black box' problem), limited availability of annotated clinical datasets, difficulties in integrating AI systems within existing hospital infrastructure, and unresolved questions around regulatory approval, liability, and algorithmic bias. Addressing these challenges requires not only technical innovation but also a multidisciplinary approach involving clinicians, data scientists, ethicists, and policymakers.

This paper presents a structured and critical review of ten significant studies published between 2021 and 2025, spanning AI applications in heart disease, blood cell analysis, COVID-19 prediction, retinal disorders, cardiovascular imaging, and chronic kidney disease. Through this synthesis, the paper identifies methodological trends, domain-specific achievements, common limitations, and unresolved research gaps. It further contextualizes these findings within broader discussions on ethical AI, clinical deployment strategies, and future research priorities.

A. Motivation and Scope

The convergence of big data, cloud computing, and sophisticated machine learning algorithms has created an unprecedented opportunity to reimagine disease diagnostics. Existing reviews in this space tend to focus on individual disease domains or specific algorithmic families. This paper takes a cross-domain perspective, deliberately spanning multiple conditions and AI paradigms to identify patterns and transferable insights. The ten reviewed studies were selected to represent a diversity of modalities, target diseases, and AI approaches, enabling a more holistic evaluation of the field.

B. Organization of the Paper

The remainder of this paper is organized as follows: Section 2 provides background on AI techniques relevant to disease detection. Section 3 presents the detailed literature review structured by study. Section 4 offers a comparative analysis and synthesis across the reviewed works. Section 5 identifies key research gaps. Section 6 discusses findings in context. Section 7 concludes the paper, and Section 8 outlines the future research scope.

II. BACKGROUND: AI TECHNIQUES IN HEALTHCARE

Before examining the reviewed literature, it is useful to briefly situate the key AI techniques that recur across the studies. This section provides a concise technical overview of the primary methods employed in AI-driven disease detection.

A. Machine Learning Classifiers

Classical machine learning classifiers — including Random Forests, Support Vector Machines (SVM), and logistic regression — operate on manually engineered feature sets derived from structured clinical data such as blood test results, patient demographics, and physiological measurements. These models are computationally efficient, relatively interpretable, and effective on moderate-sized datasets. Random forests build an ensemble of decision trees using bootstrap aggregation (bagging) and feature randomization, producing robust predictions with built-in feature importance scores.

B. Deep Learning and Convolutional Neural Networks

Deep learning models — particularly Convolutional Neural Networks (CNNs) — have revolutionized medical image analysis. CNNs automatically learn hierarchical spatial features from raw image data through successive layers of convolution, pooling, and non-linear activation. This eliminates the need for hand-crafted feature engineering and enables the detection of subtle, clinically significant patterns in radiographs, fundus photographs, histological slides, and other imaging modalities. Transfer learning further amplifies CNN utility by initializing models with weights pre-trained on large general-purpose datasets.

C. Ensemble Learning

Ensemble methods combine the predictions of multiple base learners to produce more accurate and stable outputs than any single model. Boosting algorithms — including XGBoost, LightGBM, and CatBoost — sequentially train weak learners, each correcting the residual errors of its predecessors, resulting in powerful gradient-boosted decision tree ensembles. These methods excel on tabular clinical data and have demonstrated state-of-the-art performance across numerous medical prediction benchmarks.

D. Explainable AI (XAI)

As AI models become increasingly complex, the demand for interpretability has intensified — particularly in healthcare, where clinical accountability is paramount. Explainable AI (XAI) refers to a suite of techniques designed to make model predictions transparent and understandable to human users. Post-hoc explanation methods such as SHAP (SHapley Additive exPlanations) quantify the contribution of each input feature to a given prediction. Gradient-weighted Class Activation Mapping (Grad-CAM and its extensions) produces visual saliency maps highlighting image regions that most influenced a CNN's classification decision.

E. Federated Learning and Edge AI

Federated learning is a distributed machine learning paradigm in which model training occurs locally on client devices without raw data ever leaving the device. Only model gradients or parameters are shared with a central aggregator, preserving data privacy and enabling training on geographically dispersed datasets. Edge AI refers to the deployment of AI inference at or near the data source — on IoT devices, wearables, or local hospital hardware — minimizing latency, bandwidth consumption, and privacy risks.

F. Recurrent Neural Networks and LSTMs

Recurrent Neural Networks (RNNs) and their advanced variant, Long Short-Term Memory (LSTM) networks, are designed to process sequential and time-series data. LSTMs incorporate memory cells with gating mechanisms that enable the model to capture long-range temporal dependencies — a critical capability for modeling disease progression over time, predicting epidemic trajectories, or analyzing continuous physiological signals from wearable devices.

III. LITERATURE REVIEW

This section presents a detailed, critical analysis of ten peer-reviewed studies on AI-based disease detection, published between 2021 and 2025. Each study is examined with respect to its objectives, methodology, key findings, and critical limitations.

A. AI Model for Heart Disease Detection (Chang et al., 2022)

1) Objective and Context

Chang et al. (2022) sought to design and implement an AI-based heart disease detection system using machine learning, with an emphasis on improving early diagnosis and predictive accuracy for a condition that remains the leading cause of global mortality [1]. Cardiovascular disease (CVD) accounts for approximately 17.9 million deaths annually according to the WHO.

2) Methodology

The authors developed a Python-based application using standard ML libraries (Pandas, NumPy, Matplotlib, SciPy). The core classifier employed was a Random Forest trained on the Cleveland Heart Disease dataset — a widely used benchmark in cardiac ML research.

3) Key Findings

The system achieved approximately 83% accuracy in classifying heart disease presence, with preprocessing steps shown to significantly boost model performance over raw data baselines.

4) Critical Analysis

The study's primary limitations are its reliance on a relatively small benchmark dataset and the absence of comparative evaluation against alternative algorithms such as SVM, XGBoost, or neural networks. Future work should incorporate larger, multi-site datasets and rigorous cross-validation schemes.

B. Microscopic Blood Cell Analysis with AI (Deshpande et al., 2021)

1) Objective and Context

Deshpande et al. (2021) conducted a systematic review of AI-enhanced microscopic blood cell analysis, focusing on automation of segmentation, classification, and anomaly detection in peripheral blood smear images [2]. Blood cell morphology analysis is fundamental to diagnosing hematological conditions including anemia, leukemia, malaria, and thrombocytopenia.

2) Key Findings

Automated blood cell analysis via AI was found to substantially reduce diagnostic errors compared to manual pathological examination. The review identified three major challenge areas: (1) cell overlap and clumping during staining, (2) detection of intracellular parasitic components such as Plasmodium, and (3) lack of explainability in current AI classification models.

3) Critical Analysis

The study provides a broad and useful survey but lacks rigorous head-to-head benchmarking of algorithms on standardized datasets. The absence of a standard public benchmark for blood cell classification further limits cross-study comparability.

C. AI-Based COVID-19 Detection: D-espy (Tanwar et al., 2022)

1) Objective and Context

Tanwar et al. (2022) proposed D-espy, a cloud-oriented, AI-powered framework for predicting COVID-19 case trajectories and optimizing medical resource distribution across high-risk regions in India [3]. The COVID-19 pandemic exposed critical weaknesses in healthcare resource allocation systems worldwide.

2) Methodology

Three time-series models — ARIMA, Vanilla LSTM, and Stacked LSTM — were trained on JHU COVID-19 dataset spanning January–May 2020 and evaluated against three-week rolling prediction windows.

3) Key Findings

Stacked LSTM outperformed ARIMA and Vanilla LSTM across all metrics, achieving 96.2% prediction accuracy for case counts, recoveries, and fatalities. The MRD mechanism demonstrated effective distribution of limited healthcare resources to high-risk geographic zones.

4) Critical Analysis

Results are dataset-specific (India, early pandemic lockdown period), and the model has not been validated against post-lockdown data or other national contexts. The 3-week forecast horizon is short for long-range resource planning.

D. Emerging Trends in AI for Disease Detection (Yarman & Rathore, 2025)

Yarman and Rathore (2025) authored a forward-looking chapter exploring next-generation AI paradigms for disease detection, arguing that ensemble models augmented with XAI techniques consistently outperform standalone models in both predictive accuracy and robustness [4]. XAI integration was identified as central to overcoming clinician resistance and enabling regulatory compliance. The chapter remains conceptual, drawing on secondary literature without presenting original experimental results.

E. Explainable AI for Deep Learning-Based Disease Detection (Kinger & Kulkarni, 2021)

Kinger and Kulkarni (2021) integrated Grad-CAM++ into a plant disease classification CNN, generating visual saliency maps highlighting regions most influential to classification [5]. The approach preserved accuracy while making predictions interpretable. Its primary limitation is domain scope — results are not validated in human clinical contexts, and Grad-CAM++ does not address feature-level explanations for structured clinical data.

F. Medical Imaging and AI in Retinal Disease Detection (Saleh et al., 2022)

Saleh et al. (2022) provided a comprehensive survey of AI-driven methodologies for early detection and grading of retinal diseases, focusing on diabetic retinopathy (DR) and age-related macular degeneration (AMD) [6]. AI systems based on CNNs and transfer learning demonstrated strong performance in automated grading of DR severity, approaching specialist ophthalmologist-level accuracy in controlled trials. Standardized benchmarks and multi-center validation studies are needed to enable reliable cross-study comparison.

G. Edge AI for Chronic and Infectious Disease Detection (Badidi, 2023)

Badidi (2023) examined the role of Edge AI in enabling distributed, real-time, and privacy-preserving disease monitoring, with applications in chronic disease prediction (diabetes, CVD, cancer) and infectious disease outbreak detection [7]. Federated learning allows collaborative model improvement without centralizing patient data, directly addressing privacy regulations such as GDPR and HIPAA. Significant barriers to large-scale deployment remain, including heterogeneity of edge hardware and the challenge of maintaining model accuracy under non-IID federated data.

H. AI Applications in Cardiovascular Disease Detection (Srinivasan & Sharma, 2025)

Srinivasan and Sharma (2025) reviewed state-of-the-art deep learning innovations applied to CVD detection through medical imaging, covering multiscale convolutional architectures, attention mechanisms, transfer learning, GANs for synthetic data augmentation, and self-supervised contrastive learning [8]. The chapter is technically rigorous but lacks discussion of model validation in prospective clinical trials and regulatory pathways.

I. CVD Diagnosis via Blood Biomarkers and AI (Pezoulas et al., 2024)

1) Methodology

Pezoulas et al. (2024) employed anonymized data from over 500,000 UK Biobank participants [9]. A hybrid XGBoost classifier with a custom scalable loss function was applied to 701 input features spanning demographics, routine blood tests, blood pressure measurements, and medical history.

2) Key Findings

The model achieved accuracy of 0.83, sensitivity of 0.82, and specificity of 0.84 on holdout test data — performance competitive with imaging-based diagnostics at a fraction of the cost. The most informative predictors were blood pressure, BMI, and age.

3) Critical Analysis

The UK Biobank's demographic composition (predominantly White British participants) raises concerns about cross-ethnic generalizability. External validation on independent cohorts representing diverse ethnic, socioeconomic, and geographic backgrounds is essential before clinical deployment.

J. Ensemble Learning for CKD Detection During Sleep (Jang et al., 2024)

1) Methodology

Jang et al. (2024) analyzed data from 730 participants of the Cleveland Family Study [10]. PSG signals were processed to extract 1,210 phenotypic parameters. Four ensemble classifiers — CatBoost, Random Forest, XGBoost, and LightGBM — were trained to classify CKD into four severity categories.

2) Key Findings

All four ensemble models achieved AUC values exceeding 89% for severe CKD classification, demonstrating that PSG — acquired during routine sleep studies — can simultaneously serve as a non-invasive CKD pre-screening tool at no additional cost or patient burden.

3) Critical Analysis

The reliance on a single research cohort raises generalizability concerns. PSG is not universally available in primary care settings, limiting the approach's scalability as a population-level screening tool.

IV. COMPARATIVE ANALYSIS AND SYNTHESIS

Having reviewed the ten studies individually, this section synthesizes their findings thematically, drawing cross-cutting observations about methodological trends, domain-specific achievements, and recurring limitations.

A. Dominance of Ensemble and Deep Learning Methods

Across the reviewed studies, ensemble learning methods and deep learning architectures emerged as the dominant AI paradigms. Ensemble approaches demonstrated particular strength on structured tabular clinical data, consistently achieving AUC values above 0.85. Deep learning models — especially CNNs — demonstrated superiority on image-based tasks. The complementarity of these two paradigms suggests that hybrid architectures represent a promising direction for future AI diagnostics.

B. The Centrality of Explainability

A consistent theme across the reviewed literature is the tension between model performance and interpretability. The most accurate models — deep CNNs and complex ensembles — are also the least transparent. This opacity is a practical barrier to clinical adoption, as clinicians require not only accurate predictions but also actionable explanations to guide treatment decisions.

C. Dataset Scale and Diversity

The study by Pezoulas et al. (2024), leveraging over 500,000 UK Biobank participants, stands out as a methodological exemplar in dataset scale. However, demographic diversity is equally critical. Most reviewed studies draw on datasets that are geographically and ethnically homogeneous, limiting the generalizability of their findings to global populations.

D. Privacy, Regulation, and Clinical Integration

Badidi's review of Edge AI and federated learning highlights a critical infrastructural dimension often overlooked in algorithm-focused publications. Regulatory frameworks for AI medical devices impose rigorous evidence requirements that go beyond retrospective accuracy benchmarking. Prospective clinical trials, real-world performance monitoring, bias audits, and post-market surveillance are all mandatory components of responsible clinical AI deployment.

V. RESEARCH GAPS

A. Limited Explainability and Clinical Trust

Despite the advancement of deep learning models, many AI systems continue to function as opaque 'black boxes,' limiting clinicians' ability to understand, critique, and trust AI-based predictions. A universal, modality-agnostic XAI framework that can explain predictions from multi-modal AI systems remains a critical unsolved research problem.

B. Data Availability, Quality, and Diversity

Many AI diagnostic models require large volumes of accurately labeled training data — a resource that is scarce, expensive to produce, and ethically complex to acquire. High-quality, diverse, and representative datasets remain particularly scarce for rare diseases, pediatric populations, and patients from low-income countries.

C. Clinical Workflow Integration

The predominant focus of reviewed studies on algorithmic innovation leaves largely unaddressed the practical challenge of integrating AI systems into existing clinical workflows. EHR interoperability, real-time inference latency, user interface design for clinical end-users, and institutional liability frameworks are all critical determinants of successful clinical AI deployment.

D. Edge AI Deployment Constraints

Although Edge AI and federated learning address fundamental privacy and latency concerns, their large-scale deployment remains constrained by heterogeneous hardware capabilities, battery limitations in wearables, inconsistent regulatory frameworks, and the technical challenge of maintaining model accuracy under non-IID federated data distributions.

E. Multi-Modal Data Fusion

Real disease processes are inherently multi-modal — combining imaging findings, laboratory biomarkers, genomic data, physiological signals, and patient-reported outcomes.

However, the reviewed studies predominantly exploit single data modalities. Standardized architectures for fusing heterogeneous biomedical data types represent an underdeveloped but high-impact research frontier.

F. Longitudinal and Prospective Validation

Nearly all reviewed studies employ retrospective study designs on historical datasets. Prospective validation — where AI predictions are tested in real-time clinical settings over extended follow-up periods — is largely absent from the literature.

VI. DISCUSSION

The reviewed studies collectively paint a picture of a field in rapid and productive evolution. Machine learning and deep learning have demonstrably advanced the state of automated disease detection across cardiovascular disease, hematological disorders, retinal conditions, chronic kidney disease, and infectious diseases.

At the same time, the field faces significant maturity gaps. The overwhelming majority of published studies report retrospective performance on curated research datasets under idealized conditions, with limited engagement with the complexities of real-world clinical deployment. This gap between research performance and clinical utility — sometimes termed the 'AI chasm' — is not merely a technical challenge but also a regulatory, institutional, and cultural one.

The theme of explainability is particularly prominent in the reviewed literature. Clinicians are not passive recipients of algorithmic outputs — they are professionals who bear legal and ethical responsibility for diagnostic and treatment decisions. The development of XAI tools that operate at the clinician's level of understanding is therefore as much a communication design challenge as a technical one.

A final consideration is the ethical dimension of AI in healthcare. Algorithmic bias — where models perform differentially across demographic groups due to underrepresentation in training data — has been documented across multiple medical AI applications. Ensuring equitable AI performance across age, sex, ethnicity, and socioeconomic groups is both an ethical imperative and a regulatory requirement.

VII. CONCLUSION

This paper has presented a comprehensive review of ten significant studies on AI-driven disease detection, spanning cardiovascular disease, blood disorders, COVID-19, retinal conditions, chronic kidney disease, and XAI methodology. The reviewed evidence confirms that AI techniques — particularly deep learning, ensemble learning, LSTM-based sequential modeling, and explainable AI — have achieved clinically significant performance levels across multiple disease domains.

Among the most significant individual contributions reviewed, Pezoulas et al.'s application of XGBoost to over 500,000 UK Biobank participants demonstrates the power of large-scale data for CVD biomarker classification; Tanwar et al.'s D-espy framework showcases the unique value of combining disease prediction with resource allocation; and Jang et al.'s PSG-based CKD staging illustrates the potential for repurposing existing clinical data streams for novel diagnostic applications.

The paper has also identified critical research gaps — including limited explainability, dataset diversity deficits, absence of prospective validation, clinical workflow integration challenges, and underdeveloped multi-modal fusion architectures — that must be systematically addressed for AI to fulfill its promise as a transformative healthcare technology. The ultimate goal is not to replace clinical judgment but to augment it — delivering AI tools that are accurate, interpretable, equitable, and seamlessly integrated into the fabric of compassionate patient care.

VIII. FUTURE SCOPE

The future of AI in disease detection is both expansive and deeply consequential. In the near term, the integration of large language models (LLMs) with structured clinical data and medical imaging represents an emerging and potentially transformative direction. Multimodal

foundation models pre-trained on vast corpora of biomedical text, images, and genomic data may enable unprecedented generalization across disease domains.

The maturation of federated learning frameworks will be essential for enabling multi-institutional collaborative AI development at scale. Multi-modal AI systems that seamlessly fuse imaging, genomics, proteomics, wearable physiological signals, and clinical notes into unified diagnostic and prognostic models represent a critical frontier.

The integration of AI with IoMT ecosystems will enable continuous, longitudinal health monitoring at population scale — shifting the diagnostic paradigm from episodic clinical encounters to persistent physiological surveillance. Finally, the field must invest in AI governance infrastructure: international standards for AI device evaluation, mandatory bias auditing, algorithmic transparency requirements, and post-market surveillance frameworks.

IX. ACKNOWLEDGMENT

The author gratefully acknowledges Professor Vaishnavi, Department of Computer Science, for her supervision and guidance throughout this research. This work was submitted in partial fulfillment of the requirements for the subject Deep Learning, M.Sc. Information Technology programme.

X. REFERENCES

[1]. V. Chang, V. R. Bhavani, A. Q. Xu, and M. A. Hossain, "An Artificial Intelligence Model for Heart Disease Detection Using Machine Learning Algorithms," Healthcare Analytics, vol. 2, 100016, 2022. doi: 10.1016/j.health.2022.100016.

[2]. N. M. Deshpande, S. Gite, and R. Aluvalu, "A Review of Microscopic Analysis of Blood Cells for Disease Detection with an AI Perspective," Microscopy Research and Technique, vol. 84, no. 9, pp. 2087-2106, 2021. doi: 10.1002/jemt.23782.

[3]. S. Tanwar, A. Kumari, D. Vekaria, N. Kumar, and R. Sharma, "An AI-based Disease Detection and Prevention Scheme for

COVID-19 (D-espy)," IEEE Transactions on Industrial Informatics, vol. 18, no. 8, pp. 5513-5522, 2022. doi: 10.1109/TII.2021.3120533.

[4]. B. S. Yarman and S. P. S. Rathore, "The Future of AI in Disease Detection - A Look at Emerging Trends and Future Directions," in AI in Disease Detection: Advancements and Applications, 1st ed. Wiley-IEEE Press, 2025, ch. 12. doi: 10.1002/9781394278695.ch12.

[5]. S. Kinger and V. Kulkarni, "Explainable AI for Deep Learning Based Disease Detection," in Proc. 2021 Thirteenth International Conference on Contemporary Computing (IC3), pp. 209-216, 2021. doi: 10.1145/3474124.3474154.

[6]. G. A. Saleh et al., "The Role of Medical Image Modalities and AI in the Early Detection, Diagnosis and Grading of Retinal Diseases: A Survey," Bioengineering, vol. 9, no. 8, p. 366, Aug. 2022. doi: 10.3390/bioengineering9080366.

[7]. E. Badidi, "Edge AI for Early Detection of Chronic Diseases and the Spread of Infectious Diseases: Opportunities, Challenges, and Future Directions," Future Internet, vol. 15, no. 11, p. 370, Nov. 2023. doi: 10.3390/fi15110370.

[8]. S. M. Srinivasan and V. Sharma, "Applications of AI in Cardiovascular Disease Detection - A Review," in AI in Disease Detection: Advancements and Applications, 1st ed. Wiley-IEEE Press, 2025, pp. 123-146. doi: 10.1002/9781394278695.ch6.

[9]. V. Pezoulas, G. Ehret, K. Dobretz, D. I. Fotiadis, and A. I. Sakellarios, "The Diagnosis of Cardiovascular Disease Using Simple Blood Biomarkers Through AI and Big Data," in Proc. 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 15-19 Jul. 2024. doi: 10.1109/EMBC53108.2024.10782688.

[10]. T. G. Jang, S. Y. Park, J. A. Lee, H. Y. Park, T. Dorjmaa, and D. Kang, "Ensemble Learning Approaches for Automatic Detection of Chronic Kidney Disease Stages during Sleep," in Proc. 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 15-19 Jul. 2024. doi: 10.1109/EMBC53108.2024.10782972.

AUTHOR PROFILE

PUNITH R is a postgraduate student enrolled in the Master of Science — Information Technology programme at the School of Computational Science and IT, Department of Computer Science. Registration No: 24MSIC120. This paper was submitted under the supervision of Professor Vaishnavi as partial fulfillment of the subject Deep Learning (September 2025). Research interests include artificial intelligence, machine learning, deep learning, and healthcare informatics.

Table 1: Summary of Key AI Techniques and Their Healthcare Applications

AI Technique	Strengths	Typical Application in Healthcare
Random Forest / SVM	Interpretable, handles tabular data well	Heart disease risk scoring, CKD staging
CNN / Deep Learning	Automatic feature extraction from images	Retinal disease, histopathology, radiology
XGBoost / LightGBM / CatBoost	High accuracy on structured data, fast training	CVD biomarker classification, CKD detection
Explainable AI (XAI / Grad-CAM)	Transparency, clinical trust, regulatory compliance	Plant/human disease DL interpretation
LSTM / RNN	Temporal sequence modeling	Epidemic forecasting, physiological monitoring
Federated Learning	Privacy-preserving distributed training	Multi-hospital AI model training
GANs	Synthetic data generation	Augmenting scarce medical imaging datasets
Transfer Learning	Effective with small labeled datasets	Medical image classification

(Table 1 above summarizes the key AI techniques and their typical healthcare applications discussed in Section II.)

Table 2: Performance Summary of Reviewed AI Models

Study	Disease Domain	Primary AI Method	Key Performance Metric
Chang et al. (2022)	Cardiovascular	Random Forest	~83% Accuracy
Deshpande et al. (2021)	Hematological	CNN / ML (Review)	Qualitative (Review)
Tanwar et al. (2022)	COVID-19	Stacked LSTM	96.2% Accuracy
Yarman & Rathore (2025)	Multi-domain	Ensemble + XAI (Conceptual)	N/A (Conceptual)
Kinger & Kulkarni (2021)	Plant Disease (XAI)	CNN + Grad-CAM++	Qualitative Interpretability
Saleh et al. (2022)	Retinal (DR, AMD)	CNN / Transfer Learning	Qualitative (Survey)
Badidi (2023)	Chronic/Infectious	Federated Learning	Qualitative (Survey)
Srinivasan & Sharma (2025)	Cardiovascular	DL (Attention, GAN)	Qualitative (Review)
Pezoulas et al. (2024)	Cardiovascular	XGBoost (Hybrid)	0.83 Acc / 0.82 Sens / 0.84 Spec
Jang et al. (2024)	Chronic Kidney Disease	CatBoost / XGBoost / LightGBM	AUC > 89%