RESEARCH PAPER

Edge Biomedical Diagnostic System Integrating Signal Acquisition Circuits and Retrieval Augmented Clinical Ai

Eagala Krishnaveni Reddy[1*] and Thayyaba Khatoon Mohammed[2]

[1]*Research Scholar, Dept. of CSE, Malla Reddy University Hyderabad, India*
[2]*Supervisor, Department of CSE, Malla Reddy University, Associate Professor, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, Telangana, India.*
***Corresponding author:** ekrishnavenireddy@gmail.com*

Received: 24th April, 2026; Revised: 20th May 2026; Accepted: 30th June, 2026; Available Online: 10th July, 2026

ABSTRACT

Rapid development in the decision-making systems especially in healthcare must process heterogeneous biomedical data while maintaining reliable diagnostic reasoning. This work presents an edge enabled biomedical diagnostic architecture that integrates Biosignals acquisition circuits with retrieval augmented clinical artificial intelligence. Physiological signals captured through biomedical sensors are processed using an analog front end consisting of instrumentation amplification, noise filtering, and analog-to-digital conversion to obtain stable digital representations of patient signals. These processed signals, along with clinical text and patient history, are transformed into structured semantic features. An LSTM based contextual memory module preserves temporal dependencies in patient records to support longitudinal diagnostic analysis. The structured data are then used to query a vectorized medical knowledge repository through a retrieval augmented generation (RAG) framework. Retrieved evidence is refined using relevance re-ranking and organized into a chain-of-thought reasoning structure to guide a medical large language model for differential diagnosis and risk assessment. Experimental evaluation shows that the proposed architecture improves diagnostic consistency and interpretability by combining biomedical signal acquisition with knowledge grounded reasoning.

Keywords—Biomedical Signal Acquisition Circuits, Analog Front End, Edge Artificial Intelligence, Retrieval Augmented Generation, Clinical Decision Support, and Long Short-Term Memory.

How to cite this article: Reddy EK, Mohammed TK, Edge Biomedical Diagnostic System Integrating Signal Acquisition Circuits and Retrieval Augmented Clinical Ai. Int J Drug Deliv Technol. 2026;16(7): 206-215. DOI: 10.25258/ijddt.16.7.24

Source of support: Nil.

Conflict of interest: None

1. INTRODUCTION

The rapid advancement in healthcare information systems has led to an increasing trend of biomedical data being generated from physiological sensors, clinical records, and diagnostic imaging systems. Data collected from Physiological sensors, such as electrocardiograms, electroencephalograms, and other physiological data, play an important role in providing critical information regarding patient health status, which is commonly applied in medical diagnosis. Physiological monitoring parameters play an important role in establishing healthcare dialogues. However, these physiological monitoring parameters are normally characterized by low voltage levels and high susceptibility to environmental noise, motion artifacts, and other interferences. Biomedical signal acquisition requires specialized circuits that have the capability of amplifying weak physiological monitoring parameters before subsequent digital processing.

Physiological factors play an important role in clinical care, while biomedical signal acquisition systems apply Analog front-end (AFE) circuits that incorporate instrumentation amplifiers, filtering, and Analog-to-digital conversion modules for signal conditioning and digitization. These circuits have been presented for acquiring the monitoring parameters. It plays an important role in ensuring proper signal integrity and minimizing noise interference during physiological signal measurement. Development of low-power biomedical circuits has enabled the development of portable healthcare systems that can be applied in real-world environments. [1,2]

The analysis of the information present within the biomed domain is of prime importance, along with the advancements in biomedical hardware, artificial intelligence techniques have gained importance for the analysis of complex medical data and decision-making. Rapid advancements in artificial intelligence techniques, including Machine learning and Deep learning models, have shown promising results for

tasks such as disease classification, analysis of physiological signals, and image analysis.[3] Rapid advancements in learning paradigms such as Recurrent neural network architectures, specifically Long Short-Term Memory (LSTM) models, have shown promising results for modelling sequential data such as health records

**Author for Correspondence: ekrishnavenireddy@gmail.com* and physiological signals.[4]

Many of the above-mentioned artificial intelligence techniques led to advancements in critical care, robotic surgeries, and best healthcare decisions. Despite the above advancements, many traditional artificial intelligence-based diagnostic techniques are based on static data and do not have the ability to incorporate external knowledge during inference. This has led to predictions made by such techniques being non-grounded and lack of interpretability. Recently, retrieval augmented generation (RAG) techniques have emerged as promising techniques for addressing the above-mentioned issues faced by artificial intelligence-based diagnostic techniques. By using the above technique, it is now possible to improve the accuracy of artificial intelligence-based diagnostic techniques by combining parametric artificial intelligence models and external knowledge bases. Even though many large language models are emerging for specific domains, the above models are found to be more accurate when combined with LLMs for grounding the generated responses.[5]

Before the generation of the responses from large language models, the contextual prompts are created by utilizing the retrieval augmentation methods, and many of them are in comparison with the base models healthcare analytics, the integration of biomedical signal acquisition circuits with intelligent diagnostic reasoning frameworks remains limited. Most existing clinical AI systems focus primarily on software-based analysis of pre-collected medical datasets without considering the role of signal acquisition hardware in the overall diagnostic pipeline. In this work a novel unified architecture is proposed that combines biomedical circuit design, edge processing, and knowledge driven diagnostic reasoning with RAG and LLMs could significantly enhance the reliability and applicability of intelligent clinical decision support systems.

The rapid developments in the addressed fields and Motivated by these challenges, this paper proposes an edge biomedical diagnostic system that integrates Biosignals acquisition circuits with retrieval-augmented clinical artificial intelligence. The proposed framework combines an analog biomedical signal acquisition stage with an edge processing module capable of performing contextual feature extraction and intelligent diagnostic reasoning. Physiological signals captured through biomedical sensors are first processed through an analog front-end circuit for amplification, filtering, and digitization. The processed signals are then analyzed through a contextual memory model based on LSTM networks to capture temporal dependencies in patient data. A retrieval-augmented reasoning module subsequently queries a medical knowledge repository to support evidence-based diagnostic inference. This work presents the following novel contributions for specific purposes. Another significant innovation

 Design of a biomedical signal acquisition framework

in the contemporary health care systems is the utilization of edge computing for the processing of

incorporating sensor interfaces and Analog front-end circuits for reliable physiological signal processing.

biomedical data. Edge computing allows the execution

 Development of an edge-based diagnostic architecture

of computational analyses at the data source, thus reducing the latency and communication overhead associated with the execution of such analyses in the

that combines Biosignals analysis with contextual memory modelling using LSTM networks.

cloud environment. Such a computing paradigm is Integration of retrieval-augmented reasoning with

highly advantageous for health care applications that involve the real-time analysis of physiological signals and the monitoring of the conditions of patients. The medical knowledge repositories to enable evidence-driven diagnostic support. edge computing paradigm is based on the utilization of

 Demonstration of a unified architecture that bridges

the edge devices, which act as the data sources, for the execution of the computational analyses. Such a paradigm is advantageous for health care applications owing to the reduction in latency and communication costs. The edge computing paradigm has the potential to revolutionize the health The data collected from sensors processed at the edge and decision making or alerting can be quick. Integrating biomedical sensing devices with edge AI processing modules allows efficient analysis of Biosignals while maintaining low system latency and improved privacy.[6]

Many advancements in the edge devices leverages the decision making effective in healthcare. Although considerable progress has been made in both biomedical sensing technologies and AI driven

biomedical circuit design with intelligent clinical decision support systems.

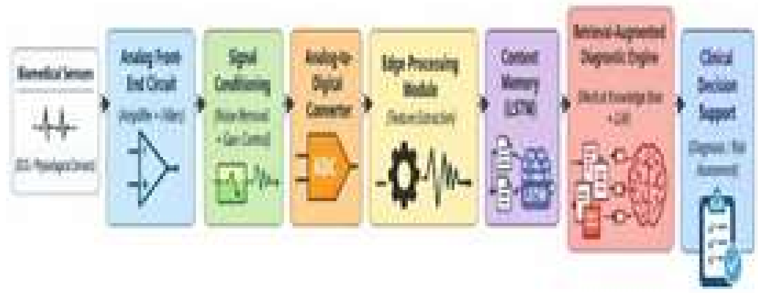

Fig-1 : Biomedical Diagnostic System

The proposed architecture of biomedical diagnostic system is represented in Fig. 1. This architecture illustrates the analysis of system on biomedical signal

data. The system consists of a biomedical sensing layer, an analog front end signal acquisition stage, and an edge based diagnostic processing module. Physiological signals captured from biomedical sensors are first processed using an analog front-end circuit that performs amplification, filtering, and signal conditioning. The conditioned signals are then converted into digital form using an analog-to-digital converter before being processed by an edge computing module. The edge processing stage performs feature extraction and contextual memory modeling using an LSTM network. The extracted features represented as vector embeddings and the Retrieved medical knowledge is incorporated through a retrieval-augmented reasoning module to support evidence-based diagnostic inference. For efficient retrieval many works proposed different database representations and multimodality fusion with RAG frameworks.

2. BIOMEDICAL SIGNAL ACQUISITION CIRCUIT DESIGN

The biomedical signals data acquisition of physiological signals is essential for real-time biomedical diagnostics and reasoning. The proposed system integrates low power Analog front-end circuits designed to capture and condition biomedical signals such as electrocardiography (ECG) and photoplethysmography (PPG). The raw extracted signals possess low signal strength and much noise and which required signal amplification.

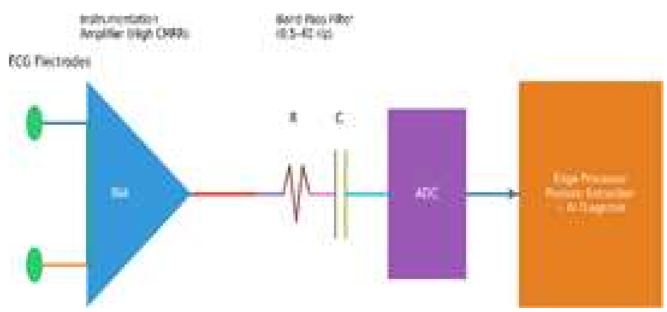

Fig-2: Biomedical signal acquisition and edge processing.

The architecture of the circuit diagram for extracting the biomedical signals, the overall biomedical signal acquisition architecture is illustrated in Fig. 2, where ECG electrodes capture bio-potential signals that are amplified using an instrumentation amplifier with high common-mode rejection ratio (CMRR). The amplified signal is then processed through a band-pass filter to isolate the clinically relevant frequency band before digitization using an Analog-to-digital converter (ADC). The data collected from these circuits and after pre-processed and is transferred to the LSTM-RAG model for further process.

2.1 *ECG Signal Acquisition Circuit*

The extracted signals are converted to digital form and the Electrocardiography signals generally range from 0.5 mV to 5 mV, making them highly sensitive to interference from muscle activity, motion artifacts, and power line noise. To address these challenges, an instrumentation amplifier-based Analog front-end is used to amplify the differential ECG signal obtained from skin electrodes. The instrumentation amplifiers used in this work for signal amplification.

These amplified signals are then used for further processing to obtain accurate diagnostic results. The instrumentation amplifiers have high input impedance, a high Common Mode Rejection Ratio (CMRR), and low noise characteristics, which are essential for biomedical signal acquisition.[3,5] The output of the amplifier is then passed through a band pass filter, which generally works within a frequency range of 0.5 to 40 Hz, corresponding to the major frequency components of ECG signals. The filtered output is then passed through an Analog to Digital Converter (ADC) before being sent to the edge processing unit for feature extraction and diagnostic inference.

2.2 *Photoplethysmography (PPG) Interface Circuit*

Photoplethysmography is widely used in wearable healthcare devices to monitor blood oxygen saturation and heart rate. The PPG circuit consists of a light emitting diode (LED) and a photodiode sensor that measures variations in reflected light caused by blood volume changes in tissue. Photoplethysmography is widely used in wearable healthcare devices for cardiovascular monitoring.

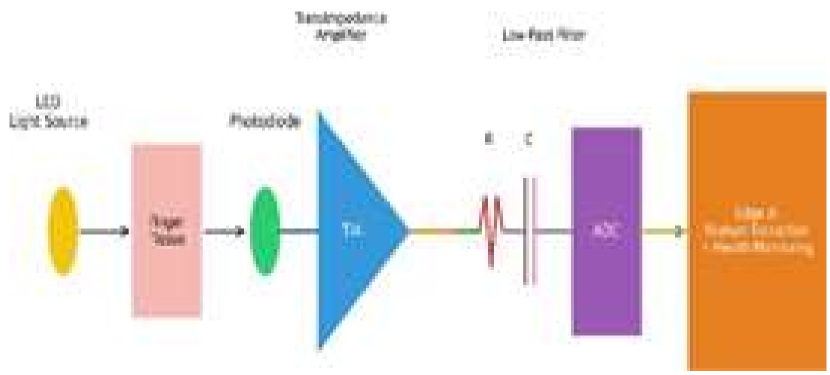

Fig-3: PPG sensing interface using photodiode, a transimpedance amplifier, and edge AI processing.

As illustrated in Fig. 3, the PPG sensing module consists of an LED illumination source and a photodiode sensor that captures variations in reflected light caused by blood volume changes. The photodiode output current is converted into a voltage signal using a transimpedance amplifier (TIA). Ambient light suppression and low-pass filtering are then applied to improve signal fidelity before digitization and transmission to the edge AI diagnostic module.

2.3 *Edge AI Framework for Biomedical Diagnostic Intelligence*

Following Analog signal acquisition and conditioning, the digitized biomedical signals are processed using an edge-based artificial intelligence framework designed for real-time physiological analysis and clinical decision support. In this work the system integrates signal feature extraction, temporal learning, and retrieval augmented reasoning to enhance diagnostic reliability while maintaining low computational latency. This system adopted a well-structured retrieval strategy for accurate dialogues synthesis.

Here the cloud based processing is replaced with edge processing module and it receives digitized ECG and PPG signals from the biomedical signal acquisition circuits described in Section II. These signals are first pre-processed to remove residual noise and normalize signal amplitude before feature extraction. The extracted features are then analysed using a temporal deep learning model deployed on the edge device.

The LSTM model is used to analyse the extracted features and before that the photodiode current is converted to voltage using a transimpedance amplifier (TIA). The signal is subsequently processed through a low pass filtering stage to remove high frequency noise components generated by motion artifacts and ambient light interference.[6,7] The high frequency components are interfered by different factors of the physiological aspects.

The conditioned PPG signal provides crucial physiological information, and it can be analyzed using edge AI techniques for the diagnosis of cardiovascular abnormalities and monitoring the health of patients in real time. Besides heart-related issues, many other diseases can be analyzed.

2.4 *Analog Filtering and Noise Suppression*

Physiological signals extracted are prone to environmental noise and baseline drift. To improve the quality of the signals, the proposed system incorporates both high and low pass filters at the Analog front-end stage.

High pass filters are used for removing baseline drift caused by electrode motion and respiration, whereas low pass filters are used for removing high-frequency noise, which is often caused by electronic circuits and other external factors such as signals and performs AI-based analysis. Edge computing minimizes latency and cloud connectivity requirements, allowing for real-time diagnostics in wearable devices and remote health monitoring systems.[1, 2, 9]

The integration of signal acquisition circuits with edge computing provides a compact system that can support continuous physiological monitoring and intelligent clinical decision support.

3. EDGE AI FRAMEWORK FOR BIOMEDICAL DIAGNOSTIC INTELLIGENCE

Following analog signal acquisition and conditioning, the digitized biomedical signals are processed by an edge-based artificial intelligence framework for real-time physiological signal analysis and clinical decision support. The proposed biomedical signal processing system combines signal feature extraction, temporal learning, and retrieval-based reasoning for improving the accuracy of the diagnosis with minimal computational latency.

The edge processing module receives the digitized ECG and PPG signals from the biomedical signal acquisition circuits discussed in Section II. The signals are pre-processed to remove any possible noise and normalize the signal strength before feature extraction. The signals are then analyzed by a temporal deep learning model running at the edge device.

3.1 Signal Preprocessing and Feature Extraction

The physiological signals obtained from the ECG and PPG sensor are preprocessed to improve the signal quality and ensure the consistency of the feature representation. Preprocessing includes signal normalization, segmentation, and peak detection.

Let the acquired physiological signal sequence be represented as

$$X = \{x1, x2, x3, \dots, xt\} \quad (1)$$

where x_t denotes the signal amplitude at time step

t. From this signal sequence, a set of feature vectors F

is extracted, including:

electromagnetic interference. Notch filtering can also

- Heart rate variability (HRV) be incorporated for removing power-line noise, which is often present in biomedical signals at frequencies of
- R-peak intervals 50/60 Hz which commonly affects biomedical recordings.[4,8] These filtering stages improve signal
- Pulse amplitude fidelity before digital conversion.

2.5 Analog-to-Digital Conversion and Edge Interface

Once the signals are conditioned using analog conditioning, they are digitized by a high-resolution analog-to-digital converter, which is part of the edge

- Signal morphology features

The acquired signal sequence represents the Time-based and Structural features and of the physiological signal, which are crucial for cardiovascular disease diagnosis. Fig.4 represents the acquired signal sequence and the feature vectors obtained from the computing module. The digitized signals are sent to the embedded processor, which extracts features from the signal sequence.

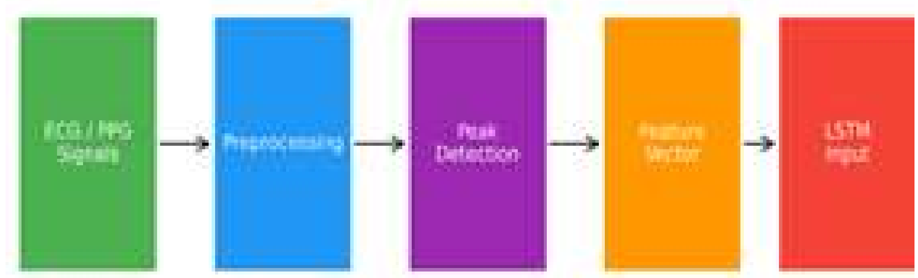

Fig-4: Feature Extraction Pipeline

3.2 Temporal Modeling Using LSTM Networks

Physiological signals have strong temporal dependencies, and the modeling of such signals is best done through sequential modeling techniques. To achieve this, the proposed system makes use of a Long Short-Term Memory (LSTM) neural network for time-series analysis.

The hidden state of the LSTM network at time t is calculated as

$$h_t = (W_h h_{t-1} + W_x x_t + b) \quad (2)$$

where

- h_t = hidden state at time t
- W_h = recurrent weight matrix
- W_x = input weight matrix
- b = bias vector
- F (·) = nonlinear activation function

The final diagnostic class probability is obtained using the SoftMax function

D= predicted diagnostic class

The predicted diagnostic output is subsequently fed into the retrieval augmented reasoning module, which is discussed in the following sections, to enhance interpretability and clinical reliability.

3.3 Edge Biomedical Diagnostic Pipeline

The overall workflow of the proposed system is illustrated in Algorithm 1, where biomedical signal acquisition, feature extraction, temporal learning, and clinical knowledge retrieval are incorporated for diagnostic predictions.

Algorithm 1: Edge Biomedical Diagnostic Pipeline Input: Physiological signals S (ECG, PPG) Output: Diagnostic prediction D

1: Acquire physiological signals from biomedical sensors

2: Amplify signals using instrumentation amplifier

3: Apply analog filtering (high-pass, low-pass, notch) 4: Convert signals to digital form using ADC

5: Perform signal preprocessing and segmentation 6: Extract physiological feature vector F

7: Input F into LSTM-based temporal model 8: Generate feature embedding E

$$p\left({}^{y}\!/_{x}\right) = e^{z}\, y$$

$$K \qquad (3)$$

9: Query clinical knowledge database using E

$$\sum_{k=1} e \quad k$$

where K represents the number of diagnostic categories.

LSTM networks are suited for biomedical time-series analysis because they can retain long-term temporal information and detect subtle physiological abnormalities.

a. Edge-Based Diagnostic Inference

The trained model of the LSTM network is implemented and embedded in the edge computing platform for real-time diagnostic inference. The edge computing device carries out the computation for the physiological signals locally without the need to send the signals to the cloud servers for processing. This approach minimizes latency and improves data privacy for the healthcare system.

The prediction of the diagnostic inference by the LSTM network is represented as

$$D = f_{LSTM}(F) \qquad (4)$$

where

- F = extracted feature vector
- f_{LSTM} = trained temporal learning model

10: Retrieve top-k relevant medical knowledge entries

11: Combine retrieved knowledge with model prediction

12: Generate final diagnostic output D

13: Return diagnostic result to clinical interface

The algorithm represents an example of how biomedical signal acquisition circuitry, edge computing, and retrieval-based reasoning can be combined for providing intelligent diagnostic capabilities in clinical environments.

3.4 Computational Complexity Analysis

The computational complexity of the proposed edge diagnostic framework is mainly influenced by the temporal learning model and the knowledge retrieval process.

Let n denote the physiological signal length and h denote the number of hidden units in the LSTM network. The complexity of the LSTM processing stage can be approximated as

$$(nh^2) \qquad (5)$$

since each time step involves matrix multiplications between input vectors and hidden states.

The complexity of the retrieval module is based on the comparison of the query vector with the mmm medical knowledge vectors embedded in the knowledge repository.

The complexity of the retrieval module can therefore be expressed as

$$(m \cdot d) \qquad (6)$$

m = number of documents in the medical knowledge database

d = embedding dimension.

$$D = \{d_1, d_2, d_3, \dots, d_m\} \qquad (8)$$

where m denotes the number of clinical documents stored in the knowledge repository.

The query embedding q is then used to retrieve the most relevant clinical information associated with the detected physiological signal patterns.

4.2 Similarity-Based Knowledge Retrieval

The retrieval module identifies relevant medical knowledge by computing the similarity between the query embedding and document embeddings stored in the knowledge base.

The similarity score is computed using cosine similarity:

Thus, the overall computational complexity of the

$$Sim(q, d_i) = \frac{q \cdot d_i}{\|q\|\|d_i\|}$$

□□□

proposed framework can be approximated as

$$(nh^2 + m \cdot d) \qquad (7)$$

Despite the inclusion of knowledge retrieval, the system remains suitable for edge deployment due to the relatively small size of feature representations and optimized indexing techniques.

4. RETRIEVAL-AUGMENTED CLINICAL INTELLIGENCE

While it is possible for machine learning models to recognize patterns within physiological signals, it is often necessary for accurate medical diagnoses to be informed by contextual knowledge, which is typically acquired through clinical guidelines and biomedical literature. To improve the accuracy of the medical diagnosis, the proposed system incorporates a retrieval

Where, q = query embedding derived from physiological features

d_i= embedding of clinical document iii.

The system retrieves the top-k documents with the highest similarity scores. These documents contain clinically relevant information related to the physiological abnormalities detected by the edge AI model.

4.3 Knowledge-Augmented Diagnostic Inference

The retrieved knowledge is then used in conjunction with the predictions of the LSTM-based diagnostic model. This leads to enhanced interpretability and enhanced prediction confidence.

The final diagnostic decision can be expressed as

augmented clinical intelligence component, which augments the predictions of the edge AI model using

$$D_{final} = f(D_{AI}, K_r) \quad (10)$$

medical knowledge.

This retrieval component works by mapping the extracted physiological features into a semantic embedding space and searching for relevant clinical documents within a structured medical knowledge repository. This process enables the diagnostic system to combine data-driven learning with validated clinical evidence.

4.1 Knowledge Repository and Embedding Representation

The retrieval-augmented module maintains a structured clinical knowledge repository containing medical guidelines, biomedical publications, and annotated clinical cases related to cardiovascular diagnostics.

Each document d_i in the repository is converted into a semantic embedding vector using a biomedical text embedding model. Similarly, the physiological feature representation generated by the LSTM model is

converted into a query embedding q.

Let the document embeddings be represented as

Where, D_{AI} = prediction generated by the edge AI model

K_r = retrieved clinical knowledge entries

(·) = knowledge integration function

This reasoning mechanism allows the system to combine physiological signal analysis with clinical context, thereby improving the reliability of diagnostic predictions.

4.4 Retrieval-Augmented Diagnostic Pipeline

The overall retrieval-enhanced diagnostic architecture is illustrated in Fig. 5.

5.2 Dataset and Signal Processing

In the proposed framework, physiological signal datasets with ECG and PPG signals were used for testing. These datasets contain annotated cardiovascular signals that include both normal and abnormal physiological states, enabling reliable testing of diagnostic algorithms.

The acquired signals undergo preprocessing steps including:

- noise suppression

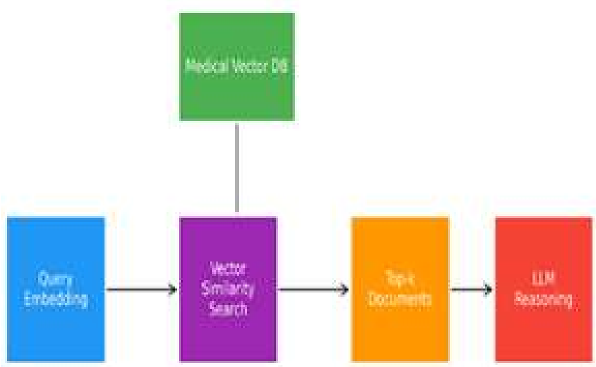

Fig-5: Retrieval-augmented clinical AI pipeline

- baseline drift removal integrating physiological signal features with medical
- signal normalization

knowledge to generate clinically informed diagnostic predictions.

- **temporal segmentation**

As illustrated in Fig. 4, the physiological signals obtained from the biomedical sensors are fed into the signal acquisition circuits and the edge AI model discussed in the previous sections. The feature embeddings obtained from the model are then used to query the clinical knowledge repository, where relevant medical information is obtained to support the diagnostic inference.

This approach to AI improves the transparency of the diagnostic process while minimizing the probability of incorrect predictions.

5. EXPERIMENTAL SETUP AND SYSTEM EVALUATION

In order to verify the effectiveness of the proposed Edge Biomedical Diagnostic System, experiments were carried out to assess the accuracy of signal processing, reliability in diagnosis, and the latency in making inferences. The evaluation focuses on the performance of the proposed architecture that integrates biomedical signal acquisition circuits with retrieval-augmented edge artificial intelligence.

5.1 Hardware Platform

The biomedical signal acquisition module comprises a low-power analog front-end architecture that consists of an instrumentation amplifier, a cascade of active filters, and a high-resolution analog-to-digital converterADC). These circuits capture physiological signals such as electrocardiography (ECG) and photoplethysmography (PPG), as described in Section II.

Following preprocessing, relevant temporal and statistical features are extracted and provided as inputs to the LSTM-based diagnostic model described in Section III.

5.3 Evaluation Metrics

The performance of the proposed biomedical diagnostic system was evaluated using widely adopted classification metrics commonly used in healthcare AI systems. The evaluation metrics include:

Accuracy – overall classification correctness **Precision** – proportion of correctly predicted positive cases

Recall – the ability of the system to identify true abnormal conditions

F1-Score – harmonic mean of precision and recall

In addition to classification metrics, inference latency was measured to evaluate the efficiency of real-time diagnostic processing in the edge computing environment.

5.4 Diagnostic Performance Results

Table -I summarizes the diagnostic performance comparison between conventional diagnostic models and the proposed edge-enabled retrieval-augmented system.

Table-I: Diagnostic Performance Comparison

Method	Accuracy (%)	Precision	Recall	F1 Score
CNN-based Cloud Model	88.2	0.87	0.86	0.86
Edge LSTM Model	91.5	0.90	0.91	0.90
Proposed Edge AI + RAG	**95.1**	**0.94**	**0.95**	**0.94**

The physiological signals are then processed using the edge computing platform, which can be used to execute machine learning algorithms. The edge computing platform can be used to process biomedical signals in real time without incurring significant communication overhead and compromising the privacy of the patient.[1, 2]

The results indicate that integrating retrieval-augmented clinical reasoning improves diagnostic accuracy compared with standalone AI models.

5.5 System Latency Analysis

Real-time healthcare monitoring systems require minimal diagnostic delay. Table II compares the inference latency of cloud-based and edge-based diagnostic frameworks.

Table II: System Latency Comparison

System Architecture	Processing Location	Latency (ms)
Cloud-based AI	Remote Cloud	320
Edge AI Model	Edge Device	140
Proposed Edge AI + RAG	Edge Device	85

The proposed architecture achieves significantly lower latency because signal processing and diagnostic inference are performed directly at the edge device without continuous cloud communication.

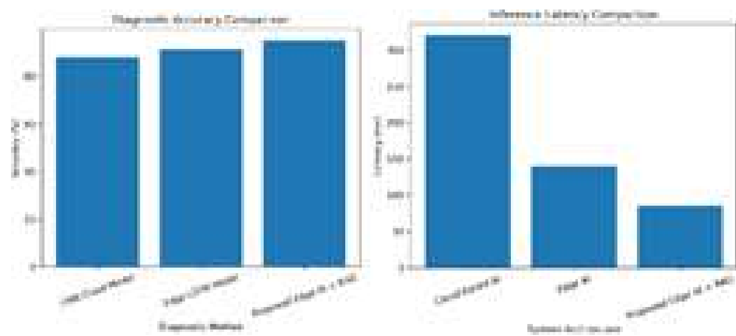

Fig-6: Diagnostic and inference accuracy

Table III: Ablation Study of System Components

System Configuration	Accuracy (%)	Precision	Recall	F1 Score
Edge AI (LSTM only)	91.5	0.90	0.91	0.90
Edge AI + Feature Enhancement	93.0	0.92	0.92	0.92
Proposed Edge AI + RAG	**95.1**	**0.94**	**0.95**	**0.94**

The experimental results are summarized in Table

III. The ablation results show that the diagnostic accuracy and classification reliability are improved with the integration of the clinical knowledge retrieval. Although the baseline LSTM model performs reasonably well by learning the temporal patterns from the physiological signals, the integration of the clinical knowledge retrieval provides more contextual information for improving the diagnostic accuracy.

This improvement highlights the advantage of combining data The benefits of the integration are evident in the application of the diagnostic system for biomedical diagnosis, where the clinical context plays a vital role.

comparison

Fig 6 Illustrates the difference between

6. CONCLUSION AND FUTURE WORK

The paper has discussed an edge-enabled

conventional AI models and the proposed retrieval-augmented edge diagnostic system. And the efficiency of edge-based biomedical diagnostic processing.

5.6 Ablation Study of the Retrieval-Augmented Module

To analyse the contribution of different components in the proposed framework, an ablation study was conducted. The objective of this study is to evaluate how the integration of edge-based processing and retrieval-augmented clinical reasoning affects diagnostic performance. Three system configurations were evaluated:

- Baseline Edge AI Model – Only the LSTM-based signal classification model without external knowledge retrieval.
- Edge AI + Clinical Feature Enhancement – Signal-based classification with additional contextual signal features.
- Proposed Edge AI + Retrieval-Augmented Reasoning

– Full architecture integrating physiological signal analysis with clinical knowledge retrieval.

biomedical diagnostic system, which integrates biomedical signal acquisition circuits with artificial intelligence and retrieval-augmented clinical reasoning. The proposed architecture integrates analog front-end circuits for biomedical signal acquisition with an edge computing framework for real-time diagnostic analysis.

The integration of the Long Short-Term Memory(LSTM)-based signal processing model with retrieval-augmented clinical knowledge improves the accuracy of the diagnostic system by utilizing data-driven learning and contextual medical knowledge. The proposed architecture has been experimentally evaluated to show the superior classification accuracy and lower latency compared to the conventional cloud-based diagnostic models.

The edge computing framework reduces communication overhead and facilitates faster diagnostic decision-making, which makes the framework suitable for various wearable healthcare and remote patient monitoring applications.

Future research directions include the integration of various physiological sensing modalities, improving the robustness of the system to motion artifacts, and developing energy-efficient hardware for large-scale biomedical monitoring applications.

REFERENCES

[1] J. G. Webster, *Medical Instrumentation: Application and Design*, 4th ed. New York, NY, USA: Wiley, 2010.

[2] R. Sarpeshkar, *Ultra Low Power Bioelectronics: Fundamentals, Biomedical Applications, and Bio-Inspired Systems*. Cambridge, U.K.: Cambridge University Press, 2010.

[3] T.-H. Tsai, J.-H. Hong, L.-H. Wang, and S.-Y. Lee,

"Low-power analog integrated circuits for wireless ECG acquisition systems," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 5, pp. 907–917, Sep. 2012.

[4] D.-L. Shen and Y.-Y. Hsu, "A low-mismatched instrumentation amplifier for ECG front-end readout circuit," *Analog Integrated Circuits and Signal Processing*, vol. 111, pp. 215–221, 2022.

[5] Kim, J. *"Simultaneous Voltage and Current Measurement Instrumentation Amplifier for ECG and PPG Monitoring. Electronics* 2021, 10, 679." 2021,.

[6] Bhat, Ganesh V., and Bhimasen G. Tasaganva. *"ECG Acquisition System."* International Journal of Engineering Research & Technology (IJERT) *3.3 (2014).*

[7] W. Shi and S. Dustdar, "The promise of edge computing," *Computer*, vol. 49, no. 5, pp. 78–81, May 2016.

[8] A. M. Rahmani *et al.*, "Exploiting smart e-health gateways at the edge of healthcare internet-of-things," *Future Generation Computer Systems*, vol. 78, pp. 641–658, Jan. 2018.

[9] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[10] A. Esteva *et al.*, "A guide to deep learning in healthcare," *Nature Medicine*, vol. 25, no. 1, pp. 24–29, Jan. 2019.

[11] P. Lewis *et al.*, "Retrieval-augmented generation for knowledge-intensive NLP tasks," in *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459–9474, 2020.

[12] T. Kojima, S. Gu, M. Reid, Y. Matsuo, and Y. Iwasawa, "Large language models are zero-shot reasoners," in *Advances in Neural Information Processing Systems*, vol. 35, 2022.